8-1-2018

Surface Entropy of Shifts of Finite Type

Dennis Pace
University of Denver

Surface Entropy of Shifts of Finite Type

Abstract

Let χ be the class of 1-D and 2-D subshifts. This thesis defines a new function, $H_S : \chi \times \mathbb{R} \to [0,\infty]$ which we call the surface entropy of a shift. This definition is inspired by the topological entropy of a subshift and we compare and contrast several structural properties of surface entropy to entropy. We demonstrate that much like entropy, the finiteness of surface entropy is a conjugacy invariant and is a tool in the classification of subshifts. We develop a tiling algorithm related to continued fractions which allows us to prove a continuity result about surface entropy in the 2-D case, namely that while it is only upper semicontinuous with respect to eccentricity that there are bounds on how badly discontinuous it can behave.

A known result about entropy is that the class of entropies of 2-D SFTs is the class of CFA numbers. In the second part of this thesis we show that all such CFA numbers can be realized as the surface entropy of a 2-D SFT. Furthermore we construct an example of a 2-D SFT demonstrating that the class of surface entropies is a strict superset to the class of entropies.

Document Type
Dissertation

Degree Name
Ph.D.

Department
Mathematics

First Advisor
Ronnie Pavlov, Ph.D.

Second Advisor
Nicholas Ormes

Third Advisor
Alvaro Arias

Keywords
Surface entropy, Entropy

Subject Categories
Mathematics | Physical Sciences and Mathematics

Publication Statement

Surface Entropy of Shifts of Finite Type

————————

A Dissertation

Presented to

the Faculty of Natural Sciences and Mathematics

University of Denver

————————

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

————————

by

Dennis Pace

August 2018

Advisor: Dr. Ronnie Pavlov

Author: Dennis Pace
Title: Surface Entropy of Shifts of Finite Type
Advisor: Dr. Ronnie Pavlov
Degree Date: August 2018

Abstract

Let $\mathscr{X}$ be the class of 1-D and 2-D subshifts. This thesis defines a new function, $H_S : \mathscr{X} \times \mathbb{R} \to [0, \infty]$ which we call the surface entropy of a shift. This definition is inspired by the topological entropy of a subshift and we compare and contrast several structural properties of surface entropy to entropy. We demonstrate that much like entropy, the finiteness of surface entropy is a conjugacy invariant and is a tool in the classification of subshifts. We develop a tiling algorithm related to continued fractions which allows us to prove a continuity result about surface entropy in the 2-D case, namely that while it is only upper semicontinuous with respect to eccentricity that there are bounds on how badly discontinuous it can behave.

A known result about entropy is that the class of entropies of 2-D SFTs is the class of CFA numbers. In the second part of this thesis we show that all such CFA numbers can be realized as the surface entropy of a 2-D SFT. Furthermore we construct an example of a 2-D SFT demonstrating that the class of surface entropies is a strict superset to the class of entropies.

Acknowledgements

I would like to thank Dr. Ronnie Pavlov, my adviser, for his unwavering support, guidance, and patience throughout my graduate career. I would also like to thank the other members of my committee; Dr. Nic Ormes, Dr. Alvaro Arias, and Dr. Brian Majestic for helping shape this thesis into something of which I can be proud. I'm grateful to Megan Daut for offering much needed support and motivation during far too many late night writing sessions. Finally I thank my parents, Wesley and Elizabeth Pace; without them I wouldn't have become someone capable of such an achievement.

In loving memory of my father, Wesley Pace (1947-2017).

Table of Contents

List of Figures

Chapter 1

Introduction

This thesis is about properties of shifts of finite type, which are a symbolically defined subclass of dynamical systems. A **topological dynamical system** $(X, \{T_g\}_{g \in G})$ is a compact topological space X, along with a G-action by homeomorphisms $\{T_g\}_{g \in G}$. The field of dynamics concerns itself with the long term average behavior when many T_g are applied to points of X or its subsets. In the area of symbolic dynamics, we work with **subshifts**. A subshift is defined as $X \subseteq A^G$, where A is some finite set called the alphabet, endowed with the discrete topology, and G is a group (for our purposes, G will always be $\mathbb{Z}$ or $\mathbb{Z}^2$.) The homeomorphisms of the system are the G-action of translations $\{\sigma_g\}_{g \in G}$. X is a subshift if it is closed in the product topology on A^G and invariant for all shifts, i.e. $\forall g \in G, \sigma_g(X) = X$. When $G = \mathbb{Z}$, X will be a space consisting of bi-infinite sequences consisting of letters of A, and when $G = \mathbb{Z}^2$ it will be a space consisting of two-dimensional bi-infinite arrays; for these spaces we assume $T_g = \sigma_g$.

Example 1.0.1. *Let $X \subset \{0, 1\}^{\mathbb{Z}^2}$ such that $x \in X$ iff all rows in x are constant (i.e. each row must be all 0s or all 1s) and exactly one row of x contains 1s. Then X is shift invariant; any shift of a point with a single row of 1s will also have a single row*

of 1s. However X is not closed. For each $n \in \mathbb{N}$ there is a point, x_n, where there is row of all 1s n digits above the origin; then $\lim_{n \to \infty} x_n$ is a point of all 0s, x^0. Let $X' = X \cup \{x^0\}$. Then X' is a subshift.

The class of subshifts we mostly treat in this thesis are the so-called shifts of finite type (or SFTs.) Let C be a finite subset of $\mathbb{Z}^d$ and define $L_X(C) = \{x|_C : x \in X\}$. The **language** of X, L_X, is the union of all such $L_X(C)$ and any $w \in L_X(C_w) \subset L_X$ is called a **word** of X of shape C_w. When $G = \mathbb{Z}$ we will say that w has length n if $|C_w| = n$ and C_w is contiguous, similarly we will say w is a (n, m) rectangle in $\mathbb{Z}^2$ if C_w is an $n \times m$ rectangle. A **shift of finite type** is then defined thusly; let F_X be a finite list of words in $A^{\mathbb{Z}^d}$ then $X \subset A^{\mathbb{Z}^d}$ is an SFT if for all $x \in A^{\mathbb{Z}^d}$, $x \in X$ unless x contains a word in F_X. Although there are no restrictions on the size of words appearing in F_X, since it is finite for any particular SFT, there will be a maximum size forbidden word. In this sense, restricting to the class of SFTs is about restricting to systems that can be defined entirely with local information.

Example 1.0.2. *Let X' be the subshift defined in Example 1.0.1. X' is not an SFT. Assume that it were, then the set of finite forbidden words has a maximum height h. Let x be a point with all 1s at the origin and all 1s at height $2h$ above the origin. Any subword of x with height at most h would have at most one row of 1s and thus would not be forbidden, but x is not a legal point of X'.*

Example 1.0.3. *Let $\mathbb{D}_d$ be the SFT with alphabet $\{B, R, T\}$ and forbidden list*

$$\mathscr{F}_d = \{BR, BT, RT, RB, TB, TR, \tfrac{T}{B}, \tfrac{B}{T}, \tfrac{R}{R}, \tfrac{B}{R}, \tfrac{R}{T}\}.$$

$\mathbb{D}_d$ is the SFT consisting of points which are either all Bs, all Ts, or 1 distinguished row of Rs with Bs below and Ts above.

It appears that Example 1.0.2 and Example 1.0.3 have the same properties, a shift space with at most one distinguished row, and yet one is not an SFT while the other is. The important difference is whether there is local information encoding when the distinguished row has appeared in the shift. In $\mathbb{D}_d$ this is encoded when the symbols switch from B to T but in X' the 0s above and below the distinguished row are the same. This idea that we can take a non-SFT and produce a similar SFT that preserves a desired property will be used several times in constructing examples and in the main results of Chapter 4.

The isomorphism in the class of dynamical systems is called a **conjugacy** and a major question in the field is the classification of equivalent systems. One of the first major tools and the first numerical tool in answering this question was entropy. In general the entropy of a system is defined using open covers, but a key technique in the study of symbolic dynamics is noticing that we can study the behavior of the system by studying the behavior of L_X. **Topological entropy** (hereafter simply entropy) is then defined for any $\mathbb{Z}^d$ subshift as:

$$h(X) = \lim_{n \to \infty} \frac{\log L_X([1, n]^d)}{n^d}.$$

The entropy of a system is a conjugacy invariant; thus entropy is important in answering the question of classification of dynamical systems. Entropy has also found broad application in many fields related to dynamics. In statistical physics, the entropy of a system is called the free energy and determines the equilibrium behavior of the system. In information theory the encoding schema for information storage generates a dynamical system and the entropy of this system is the carrying capacity of the schema. A concrete example of this can be found in the (2,7) run-length limited scheme (RLL) which was widely used in storing data on magnetic computer

3

hard drives [7]. While we think of computers as using binary to store information there are physical constraints in the magnetic media which force a more sophisticated approach to be used to physically store the information. A short summary of these constraints are that two consecutive 1s stored in the media cause interference and the read head has trouble differentiating between a pair of 1s and a single 1, similarly if there are too many adjacent 0s the read head can have trouble accurately determining how many 0s were suppose to be present. To correct these issues the (2,7) RLL forced information to be stored so that at least two 0s but no more than seven 0s could occur between consecutive 1s. This means that not every bit on the media was data, but that some of the bits occurred simply to conform to the RLL standard. Decoding a section of media required removing these extraneous bits; which meant that a track encoding in RLL did not hold quite as much information as one encoded in pure binary. The amount of data stored in an n digit binary sequence is $2^n = e^{n \log 2}$ where $\log 2$ is the entropy of an unrestricted shift on two letters whereas the amount of data stored in an n digit (2,7) RLL sequence is about e^{nh} where $h \approx \log 1.4$ is the entropy of the shift generated by (2,7) RLL.

From the above definition we can view entropy as a first order (here we will assume $d = 2$ for purpose of discussion) estimate of the exponential growth of L_X,

$$L_X([1,n]^2) \approx e^{n^2 h(X)}.$$

However, for some constant $C \in [0, \infty)$ consider the following two estimates:

$$L_X^1([1,n]^2) \approx e^{n^2 h(X)} \qquad L_X^2([1,n]^2) \approx e^{n^2 h(X) + 2Cn}.$$

4

Entropy is not enough to distinguish these two cases. In a shift where such a C exists define this higher order term in the estimate as **surface entropy** (it will be 0 or ∞ in the case that this term is non-linear.)

This thesis has two major sections. First, we begin by proving some structural results about surface entropy in both the 1-D and 2-D settings. We develop surface entropy as a companion to entropy. The conjugacy invariance of entropy was of noteworthy importance and so we ask a similar question of surface entropy. It is not the case that surface entropy is conjugacy invariant in full generality as we construct examples of two conjugate systems with different surface entropy. However, among other results, we do show that the finiteness of surface entropy is invariant. We give examples (Example 3.2.4 and 3.2.5) that this distinction is meaningful by demonstrating two subshifts with 0 entropy that do not agree on the finiteness of their surface entropy.

Corollary 3.2.17. *If X and Y are conjugate subshifts then $H_S(X, \alpha) < \infty$ if and only if $H_S(Y, \alpha) < \infty$.*

In the 1-D setting we also show that the finiteness of the surface entropy can be calculated algorithmically.

Theorem 3.1.6. *Let X be a 1-D SFT, there exists an algorithm to determine if $H_S(X) = \infty$ or if $H_S(X) < \infty$.*

In the 2-D case we again show several structural results and then prove that surface entropy further deviates from entropy in its behavior. Note that in the definition of entropy only square words are used; it turns out that we can use this definition because any sequence of rectangular word sizes which sees the size of the words increasing to infinity will generate the same value. Interestingly, this is no longer the case when calculating surface entropy; the **eccentricity** (ratio of a rectangle's height to width)

of the words used affect the value of the calculation. Thus we define surface entropy of a particular eccentricity α to be

$$H_S(X, \alpha) = \sup_{\{\{(x_n, y_n)\} | \frac{y_n}{x_n} \to \alpha, x_n \to \infty\}} \left(\varlimsup_{n \to \infty} \frac{\log |L_X(x_n, y_n)| - x_n y_n h(X)}{x_n + y_n} \right).$$

Fixing an SFT and considering surface entropy as a function of α we produce an example demonstrating surface entropy is not continuous but it is upper semi-continuous. Moreover we produce bounds on how badly discontinuous it can be.

Theorem 3.3.2. *Let X be a subshift. Then $H_S(X, \alpha)$ is an upper semi-continuous function of α.*

Theorem 3.3.18. *Let X be a subshift. Let $\beta, \alpha \in (0, \infty)$ and assume $\frac{\beta}{\alpha} \notin \mathbb{Q}$. Let $P = \max(2 + \alpha, 2 + \frac{1}{\alpha})$ then $H_S(X, \beta) \leq PH_S(X, \alpha)$.*

Theorem 3.3.19. *Let X be a subshift. Let $\beta, \alpha \in (0, \infty)$ and assume $\frac{\beta}{\alpha} \in \mathbb{Q}$ where $\frac{p}{q}$ is the reduced form of $\frac{\beta}{\alpha}$. Let $z = \max\{p, q\}$ and let $\gamma = \max\{1 + z + \alpha, 1 + z + \frac{1}{\alpha}\}$ then $H_S(X, \beta) \leq \gamma H_S(X, \alpha)$.*

In the proof of Theorem 3.3.18 and Theorem 3.3.19 we need an upper bound on $H_S(X, \beta)$ which essentially means we need an upper bound on the word count of eccentricity β rectangles. We only have information about the word count on rectangles of eccentricity α with which to construct our upper bound. Since surface entropy is measuring the linear exponential growth rate instead of quadratic exponential growth rate these bounds must have a smaller margin of error than in similar proofs regarding entropy. We develop a tiling algorithm which accomplishes this by filling the β (eccentricity) words with many different sizes of α words. Surprisingly the algorithm that succeeds is one closely related to the Euclidean algorithm and the properties of the continued fraction expansion of $\frac{\beta}{\alpha}$ plays a role in the proof of these theorems.

6

Peculiarly, several questions in 2-D SFTs are unavoidably related to computability theory. For instance, given an alphabet and set of forbidden words it is algorithmically undecidable if the resulting SFT is non-empty [2]. We can classify elements of $\mathbb{R}$ based on how they can be approximated by Turing machines, examples of such numbers are **computable**, **computable from above**, and **computable from below** with the latter two classes strictly containing than the first. There are interesting arguments that computable numbers encapsulate the concept of how we fundamentally think about numbers and yet there are only countably many computable numbers and so they are measure 0 in $\mathbb{R}$.

The second half of the thesis moves on to the question of which numbers can be realized as surface entropy. A number which can be arbitrarily estimated from above by a Turing machine is said to be **computable from above (CFA)** and a major result in symbolic dynamics by Hochman and Meyerovitch [4] showed that the set of entropies of 2-D SFTs was equal to the the set of non-negative CFA numbers. Our results in Section 4.2 use this result to show that all such numbers can be realized as the surface entropy of a 2-D SFT.

Theorem 4.2.4. *For any CFA $\gamma \in [0, 1]$ there is a subshift X such that $H_S(X, 1) = \gamma$.*

We then demonstrate in Section 4.3 that there are non-CFA numbers which can also be realized as surface entropy, proving that the set of surface entropies is a strict superset to the set of entropies. Unlike our result in the CFA case this construction cannot directly use the results in [4]. Their construction is a highly technical result which requires using an embedded Turing machine to sample the frequencies appearing in so-called **Toeplitz** sequences. Using the methods from both their construction or ours these Toeplitz sequences always result in a final surface entropy that is infinite. Much like the previous discussion about Theorem 3.3.18 and Theorem 3.3.19

7

we require finer control of the word count in order to control surface entropy. Our construction uses the embedded Turing machine to sample from a different type of sequence, so-called **Sturmian** sequences. The upside to using these sequences is that we can better control the word counts, the downside is that these sequences have less redundancy and are harder to control with a Turing machine. We are required to use results about the Weyl equidistribution theorem to show that this limited control is still sufficient.

Theorem 4.3.20. *There exists a subshift X such that $H_S(X, 1)$ is CFB and not CFA.*

Chapter 2

Definitions

2.1 Subshifts

Definition 2.1.1. *Let A be a finite set of symbols, the **alphabet**, endowed with the discrete topology. Define $\sigma : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ such that $\forall a \in A^{\mathbb{Z}}$ and $i \in \mathbb{Z}$, $\sigma(a_i) = a_{i+1}$. Then $(A^{\mathbb{Z}}, \sigma)$ along with the product topology on $A^{\mathbb{Z}}$ is the **full shift in 1-D**.*

Definition 2.1.2. *Let A be a finite set of symbols, the alphabet, endowed with the discrete topology. For each $v \in \mathbb{Z}^2$ define $\sigma_v : A^{\mathbb{Z}^2} \to A^{\mathbb{Z}^2}$ such that $\forall a \in A^{\mathbb{Z}^2}, \sigma_v(a_w) = a_{v+w}$. Then $\left(A^{\mathbb{Z}^2}, \{\sigma_v\} \right)$ along with the product topology on $A^{\mathbb{Z}^2}$ is the **full shift in 2-D**.*

Definition 2.1.3. *If $X \subset A^{\mathbb{Z}^d}$ is closed and σ_v invariant for all $v \in \mathbb{Z}^d$ then $(X, \{\sigma_v\})$ is a **d-dimensional subshift**.*

Definition 2.1.4. *Let X be a subshift. The collection of all finite configurations appearing in any element of X is the **language of** X, denoted L_X. An element $w \in L_X$ is called a **word**.*

Definition 2.1.5. *Let X be a 1-D subshift then $L_X(n) \subseteq L_X$ is the set of words of length n.*

Definition 2.1.6. *Let X be a 2-D subshift then $L_X(n, m) \subseteq L_X$ is the set of rectangular $n \times m$ words. (Here and thereafter $n \times m$ denotes a rectangle of width n and height m.)*

Definition 2.1.7. *The **eccentricity** of an $n \times m$ rectangle is $\frac{m}{n}$.*

Definition 2.1.8. *Let X be a subshift. The **entropy** of X, denoted $h(X)$, is defined as follows:*

$$h(X) = \begin{cases} \lim_{n \to \infty} \frac{\log|L_X(n)|}{n} & \text{if } X \text{ is a 1-D subshift} \\ \lim_{n \to \infty} \frac{\log|L_X(n,n)|}{n^2} & \text{if } X \text{ is a 2-D subshift.} \end{cases}$$

Definition 2.1.9. *Let $X \subseteq A^{\mathbb{Z}^d}$ be a subshift. Then X is a **shift of finite type**, an SFT, if $\exists F$ a finite set of words such that $\forall a \in A^{\mathbb{Z}^d}$, $a \in X$ iff no word in F appears in a.*

Definition 2.1.10. *Let F be a finite list of forbidden words and $X \subseteq A^{\mathbb{Z}^d}$ be the associated SFT. X is a **nearest neighbor SFT** if F only contains words of two adjacent letters.*

Definition 2.1.11. *Let $(X, \{\sigma_v\})$ and $(Y, \{\sigma_v\})$ be subshifts. $f : X \to Y$ is a **factor map** if f is onto Y and $\forall v \in \mathbb{Z}^d$, $f \circ \sigma_v = \sigma_v \circ f$. If such an f exists Y is a **factor** of X.*

Definition 2.1.12. *Let X, Y be subshifts and $f : X \to Y$ be a factor map. f is a **conjugacy** iff f is a bijection. If such an f exists then X is said to be **conjugate** to Y.*

Definition 2.1.13. *Let G be a directed graph and A be the edges of G. Let $X \subseteq A^{\mathbb{Z}}$ such that $x \in X$ iff x is a bi-infinite walk of edges of G. Then X is a subshift and is called the **edge shift** corresponding to G.*

Definition 2.1.14. *Let $X \subseteq A^{\mathbb{Z}}$ be an edge shift and G be the corresponding graph. Let V be the vertices of G and M be a $|V| \times |V|$ matrix where $M_{(I,J)}$ is equal to the number of edges from vertex I to vertex J. Then M is the **adjacency matrix** of X.*

Definition 2.1.15. *Let G be a directed graph and A be the vertices of G. Let $X \subseteq A^{\mathbb{Z}}$ such that $x \in X$ iff x is a bi-infinite walk of vertices of G. Then X is a subshift called the **vertex shift** corresponding to G.*

We note that by Theorem 2.3.2 and Proposition 2.3.9 in Lind and Marcus [7] that any 1-D SFT is conjugate to both an edge shift and a vertex shift.

Definition 2.1.16. *Let $X \subseteq A^{\mathbb{Z}}$ be a vertex shift. Let M be an $|A| \times |A|$ matrix where each row corresponds to a unique element of A and each column corresponds to a unique element of A. Let $a, b \in A$ then if*

$$
M_{(a,b)} = \begin{cases} 1 & \text{if } ab \in L_X \\ 0 & \text{if } ab \notin L_X \end{cases}
$$

*M is the **adjacency matrix** of X.*

Definition 2.1.17. *Let $X \subseteq A^{\mathbb{Z}^d}$ be a subshift and $x \in X$. x is **periodic** with period $n > 0$ if $\exists v \in \mathbb{Z}^d$ such that $\sigma_v^n(x) = x$. The **least period** of x is the smallest $n > 0$ such that x is periodic.*

Definition 2.1.18. *Let X be a subshift. X is **irreducible** if $\forall u, v \in L_X, \exists w \in L_X$ such that $uwv \in L_X$.*

Definition 2.1.19. *Let X be an edge shift with adjacency matrix A. X is **primitive** iff $\exists N > 1$ such that $A^N > 0$.*

Definition 2.1.20. *Let G be a directed graph and X_G its associated edge shift. Let I be a maximal subgraph of G such that for any $i, j \in V(I)$ there is a path from i to j consisting of edges in I. Let X_I be the edge shift associated with I, then X_I is an **irreducible component** of X_G.*

Definition 2.1.21. *Let A be a nonnegative irreducible matrix, the eigenvalue of A with maximal modulus is its **Perron eigenvalue** and an eigenvector corresponding to a Perron eigenvalue is a **Perron eigenvector**.*

Theorem 2.1.22. *[7] Let X be a 1-D SFT and X_G be a conjugate edge shift with adjacency matrix A. If λ_A is the Perron eigenvalue of A then $h(X) = \log \lambda_A$.*

Definition 2.1.23. *Let X be a 1-D SFT. X is **mixing** if $\forall u, v \in L_X, \exists N_{(u,v)}, \forall n > N_{(u,v)} \exists w \in L_X(n)$ such that $uwv \in L(X)$.*

Definition 2.1.24. *Let $X \subseteq A^{\mathbb{Z}^2}$. X is **block gluing with gap size** g if for any pair of two solid blocks $B_1, B_2 \subsetneq \mathbb{Z}^2$ such that the distance between B_1 and B_2 is greater than g and for any two points $y, z \in X$ then $\exists x \in X$ such that $x|_{B_1} = y|_{B_1}$ and $x|_{B_2} = z|_{B_2}$.*

Definition 2.1.25. *Let X, Y be 2-D subshifts where A_X is the alphabet of X and A_Y is the alphabet of Y. Choose $k \in \mathbb{N}$ and let $B : L_X(2k+1, 2k+1) \to A_Y$. Define $B_{(i,j)} \subsetneq \mathbb{Z}^2$ as the $(2k+1, 2k+1)$ rectangle centered at (i, j). Define $f : X \to Y$ such that $y_{(i,j)} = B(X|_{B_{(i,j)}})$. Such an f is a **sliding block code of window size** k.*

Definition 2.1.26. *Let X be a subshift with alphabet A_X. Define a subshift Y such that $A_Y = L_X(M, M)$ and the identity map from $L_X(M, M) \to A_Y$ induces a sliding block code between X and Y. Then Y is the **Mth higher block presentation** of X.*

There are analogous definitions for sliding block code and Mth higher block presentation in 1-D where intervals of X, instead of squares of X, are mapped to letters of Y. (See [7] for more details.)

Definition 2.1.27. *[11] A sequence $x \in \{0, 1\}^{\mathbb{Z}}$ is **k-balanced** if for any two subwords of x with the same length, the number of 1s appearing in each word differs by no more than k, i.e. $\forall n \in \mathbb{N}, \forall i, j \in \mathbb{Z} \left| \#_1(x_{[i,i+n-1]}) - \#_1(x_{[j,j+n-1]}) \right| \leq k$ where $\#_1(x_{[i,i+n-1]})$ is the number of 1s appearing in the n letter word starting at x_i.*

Definition 2.1.28. *[3] A sequence, $(x_z) \in \mathbb{Z}^2$, is **Sturmian** if $\forall n \in \mathbb{N}, \left| L_{(x_z)}(n) \right| = n + 1$.*

It can be shown that an aperiodic 1-balanced sequence is Sturmian.

Definition 2.1.29. *Choose $\rho, \gamma, \delta \in [0, 1)$. For each $z \in \mathbb{Z}$ let $x_z = \chi_{[1-\rho,1]}(z\gamma + \delta)$ mod 1. Then the sequence (x_z) is **generated from a γ circle rotation** and the frequency of 1s appearing in (x_z) is ρ.*

Proposition 2.1.30. *[3] The set of Sturmian sequences is equivalent to the set of sequences generated from irrational circle rotations.*

Definition 2.1.31. *[5] Let $(x_z) \in \{0, 1\}^{\mathbb{Z}}$ such that $\forall n \in \mathbb{N}, \forall i, j \in \mathbb{Z}$, if $i \pmod{2^{n+1}} = j \pmod{2^{n+1}} = 2^n \pmod{2^{n+1}}$ then $x_i = x_j$. Define $X = \overline{\{\sigma_n((x_z))\}_{n \in \mathbb{Z}}}$. Then any $x \in X$ is a **(dyadic) Toeplitz** sequence.*

2.2   Computability

Definition 2.2.1. *A **Turing machine**, T, is a theoretical model of computation consisting of the following components.*

- *A finite alphabet, A, which contains a "blank" symbol.*

- *A finite list of **states**, S, which contains a distinguished initial and halting state.*

- *A finite table of **instructions**, $I : (A \times S) \to (A, S, \{-1, 0, 1\})$.*

- *A **tape**, $(t_n) \in A^{\mathbb{N}}$, which stores the input and output of T. The input and output of T is a finite sequence of elements of A proceeded by all blanks.*

- *A **head**, $H = (H_S, H_P) \in S \times \mathbb{N}$, which stores the current state of the Turing machine and the current location on the tape where the machine and read and/or write.*

We can think of T as a partial function from the set of tapes to itself where the output of T can be computed using the following algorithm.

Upon initialization, T sets its current state, H_S, to the initial state. The position of the head, H_P, is set to 0. It then performs the following loop unless it reaches the halting state, at which point it halts.

Begin Loop

(L1) Read the symbol on the current location of the tape and use this along with the current state to determine which instruction to perform. This is done by calculating $(a, s, m) = I(t_{H_P}, H_S)$.

(L2) Set the state to s, replace the symbol at the current location with a, and move

the head to $H_P + m$.

End Loop

A **step** *of the Turing machine is one iteration of this loop. When T halts, we denote the contents of the tape after halting by $T((t_n))$; else $T((t_n))$ is undefined.*

We can also think of T as acting on natural numbers by bijectively corresponding $n \in \mathbb{N}$ to its base-k expansion for some k. In this case, both the input and output have only finitely many non-blank symbols. For T we consider in this work, the output in this case will be a pair of natural numbers, which we interpret as the numerator and denominator of a rational number $T(n)$.

*The **Church-Turing Thesis** is an axiom of computation theory which states that any type of effective computation can be performed by a Turing machine.*

While a Turing machine can simulate any effective computation, it should be noted that this is by no means efficient. A simple multiplication machine which multiplied $10 * 10$ required 11500 steps[12].

Definition 2.2.2. *A number $x \in \mathbb{R}$ is **computable from above (CFA)** if there is a Turing machine T such that $T(n) \to x$ and for each input $n \in \mathbb{N}$, $T(n) \geq x$.*

Definition 2.2.3. *A number $x \in \mathbb{R}$ is **computable from below (CFB)** if there is a Turing machine T such that $T(n) \to x$ for each input $n \in \mathbb{N}$, $T(n) \leq x$.*

Definition 2.2.4. *A number $x \in \mathbb{R}$ is **computable** if there is a Turing machine T such that for each input $n \in \mathbb{N}$, $T(n) - \frac{1}{n} \leq x \leq T(n) + \frac{1}{n}$. Alternatively, $T(n)$ outputs the first n digits of x.*

It is easily checked that a number is computable iff it is both CFA and CFB.

Definition 2.2.5. *A function $f : \mathbb{N} \to \mathbb{N}$ is a **computable function** if there is a Turing machine T such that $T(n) = f(n)$.*

The set of computable functions is closed under any algebraic operation which can be computed with a finite algorithm; in particular it contains all elementary functions and is closed under sums, products, and composition. We use the term computable function to better match the literature. For our purposes we consider computable functions simply to be Turing machines and so if T is a Turing machine, f is a computable function, and T_f is the Turing machine implementing f then $f \circ T = T_f \circ T$.

Definition 2.2.6. *The **IP set** generated from an infinite set of generators, $\{g_n\}_{n \in \mathbb{N}}$, denoted $IP - (g_1, g_2, ...,)$, is the sequence of the set of all finite sums of the generators ordered by size.*

Chapter 3

Behavior of Surface Entropy

We begin by a definition for surface entropy in 1-D and Theorem 3.1.2 shows how to calculate surface entropy in the 1-D setting for certain shifts of finite type. We then classify when a SFT will have finite or infinite surface entropy. We give two equivalent definitions of surface entropy in 2-D and prove several general properties of surface entropy. We show that the surface entropy of a product is subadditive and that if the subshift is block gluing the surface entropy has an upper bound related to the entropy and block gluing gap. Examples are then presented showing calculations of surface entropy and highlighting the fact that unlike entropy, surface entropy can depend on the eccentricity of the rectangles used to count words. Several examples describing surprising behavior of surface entropy are presented; in particular surface entropy can be infinite, it can be 0 when X is not the full shift, and $\lim_{n \to \infty} S_X(x_n, y_n)$ as defined in Definition 3.2.1 may not converge along certain sequences in $\Xi(\alpha)$ (Definition 3.2.2.) We also show that surface entropy is not a conjugacy invariant, but Theorem 3.2.16 shows that whether surface entropy is infinite is a conjugacy invariant.

3.1   1-D Shifts of Finite Type

Definition 3.1.1. *Let X be a 1-D subshift and $h(X)$ be the entropy of X. The* **surface entropy** *of X is defined as:*

$$H_S(X) = \varlimsup_{n \to \infty} \log |L_X(n)| - n h(X).$$

Theorem 3.1.2. *Let X be a 1-D, mixing, edge shift with adjacency matrix A that has Perron eigenvalue λ. Let r be a right eigenvector of A and ℓ be a left eigenvector of A normalized s.t. $\ell \cdot r = 1$. Then $H_S(X) = \log \left(\frac{1}{\lambda} \sum r_i \sum \ell_i \right)$.*

Proof. By Theorem 4.5.12 in Lind and Marcus [7]

$$(A^n)_{ij} = [r_i \ell_j + p_{ij}(n)] \lambda^n \text{ where } \lim_{n \to \infty} p_{ij}(n) = 0.$$

Also recall that by Theorem 2.1.22 $h(X) = \log \lambda$, thus

$$
\begin{aligned}
H_S(X) = \varlimsup_{n \to \infty} \log |L_X(n)| - n h(X) &= \varlimsup_{n \to \infty} \log \left(\sum_{i,j} (A^{n-1})_{ij} \right) - n \log \lambda \\
&= \varlimsup_{n \to \infty} \log \left(\sum_{i,j} [r_i \ell_j + p_{ij}(n-1)] \lambda^{n-1} \right) - n \log \lambda \\
&= \varlimsup_{n \to \infty} \log \left(\sum_{i,j} [r_i \ell_j + p_{ij}(n-1)] \right) - \log \lambda \\
&= \log \left(\frac{1}{\lambda} \sum_{i,j} r_i \ell_j \right) = \log \left(\frac{1}{\lambda} \sum_i r_i \sum_i \ell_i \right).
\end{aligned}
$$

$\square$

Theorem 3.1.3. *Let X be a 1-D, irreducible, edge shift with period p. Then $H_S(X) < \infty$.*

Proof. Let A be the adjacency matrix of X and λ be the Perron eigenvalue of A. Then λ^p is the Perron eigenvalue of A^p. By Theorem 4.5.6 of [7] the vertices of A can be ordered by period class so that:

$$A = \begin{pmatrix} 0 & B_0 & 0 & \ldots & 0 \\ 0 & 0 & B_1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & B_{p-2} \\ B_{p-1} & 0 & 0 & \ldots & 0 \end{pmatrix}$$

and

$$A^p = \begin{pmatrix} A_0 & 0 & 0 & \ldots & 0 \\ 0 & A_1 & 0 & \ldots & 0 \\ 0 & 0 & A_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & A_{p-1} \end{pmatrix}$$

where each A_i is primitive and $A_i = B_i...B_{p-1}B_0...B_{i-1}$. Since A_i is primitive, $\lim_{n \to \infty} \frac{A_i^n}{\lambda^n} = C_i$ for some real valued matrix C_i. Thus $\lim_{n \to \infty} \frac{A^{pn}}{\lambda^{pn}} = C$ where C is a block diagonal matrix with C_i along the diagonal.

Instead of considering sequences of any word length, we restrict to word lengths in a single residue class mod p. For each residue class we will find a finite upper bound of the surface entropy restricted to sequences in that residue class; such an upper bound on each residue class will give an upper bound on a sequence of any word length. Let $0 < o \leq p$.

$$\varlimsup_{n \to \infty} \log |L_X(pn + o)| - (pn + o)h = \varlimsup_{n \to \infty} \log \left(\sum_{i,j} (A^{pn+o-1})_{ij} \right) - (pn + o) \log \lambda$$

$$= \overline{\lim_{n \to \infty}} \log \left(\sum_{i,j} (A^{pn} A^{o-1})_{ij} \right) - pn \log \lambda - o \log \lambda$$

$$= \overline{\lim_{n \to \infty}} \log \left(\sum_{i,j} \left(\frac{A^{pn}}{\lambda^{pn}} A^{o-1} \right)_{ij} \right) - o \log \lambda$$

$$= \lim_{n \to \infty} \log \left(\sum_{i,j} \left(\frac{A^{pn}}{\lambda^{pn}} A^{o-1} \right)_{ij} \right) - o \log \lambda$$

$$= \log \left(\sum_{i,j} \left[\lim_{n \to \infty} \left(\frac{A^{pn}}{\lambda^{pn}} \right) A^{o-1} \right]_{ij} \right) - o \log \lambda$$

$$= \log \left(\sum_{i,j} \left[C A^{o-1} \right]_{ij} \right) - o \log \lambda < \infty.$$

For any sequence of word lengths the residue classes mod p finitely partition the index set and so the bounding of the residue sequences gives a bound for any sequence.

$$\overline{\lim_{n \to \infty}} \log |L_X(n)| - (n)h = \sup_{0 \le o \le p} \left[\log \left(\sum_{i,j} \left[C A^{o-1} \right]_{ij} \right) - o \log \lambda \right] < \infty.$$

$\square$

Lemma 3.1.4. *Let X be an edge shift that is not irreducible, where $h(X) = \log \lambda$, X has two irreducible components Y and Z each with entropy $\log \lambda$ such that $\exists y \in L(Y), z \in L(Z), w \in L(X)$ such that $ywz \in L(X)$. Then $H_S(X) = \infty$.*

Proof. There is a letter $y_z \in Y$ that can transition to some $z_y \in Z$ in $|w|$ steps. Since Y is irreducible there is some $G \in \mathbb{N}$ so that any letter in Y can precede y_z by at most G number of steps. Similarly Z being irreducible means that after at most H steps z_y can transition to any letter in Z. To create a word in $L_X(n)$ we can choose $w_Y \in L(Y)$ and $w_Z \in L(Z)$ along with transition word $w_G \in L(Y)$ of length no more than G and $w_H \in L(Z)$ with length no more than H such that $|w_Y w_G w w_H w_Z| = n$.

$$|L_X(n)| \geq \sum_{i=G}^{n-|w|} |L_Y(i-G)|\,|L_Z(n-i-H)|\,.$$

Calculating surface entropy we find.

$$\varlimsup_{n\to\infty} \left[\log|L_X(n)| - n\log\lambda\right]$$

$$\geq \varlimsup_{n\to\infty} \left[\log\left(\sum_{i=G}^{n-|w|} |L_Y(i-G)|\,|L_Z(n-i-H)|\right) - n\log\lambda\right]$$

$$\geq \varlimsup_{n\to\infty} \left[\log\left(\frac{\sum_{i=G}^{n-|w|} |L_Y(i-G)|\,|L_Z(n-i-H)|}{\lambda^n}\right)\right]$$

$$\geq \varlimsup_{n\to\infty} \left[\log\left(\frac{1}{\lambda^{G+H}} \sum_{i=G}^{n-|w|} \left(\frac{|L_Y(i-G)|}{\lambda^{i-G}}\right)\left(\frac{|L_Z(n-i-H)|}{\lambda^{n-i-H}}\right)\right)\right]$$

$$\geq \varlimsup_{n\to\infty} \left[\log\left(\frac{1}{\lambda^{G+H}} \sum_{i=G}^{n-|w|} (1)(1)\right)\right] \geq \infty.$$

Thus $H_S(X) = \infty$. $\qquad\square$

Theorem 3.1.5. *Let X be a non-irreducible edge shift with $h(X) = \lambda$ such that X does not satisfy the hypotheses of Lemma 3.1.4. Then $H_S(X) < \infty$.*

Proof. For $0 \leq i \leq \psi$ let M_i be the irreducible components of X with entropy $\log\lambda$ and for $0 \leq j$ let A_j be all the other irreducible components of X. Let $\phi = \max\{h(A_j)\}$. Let A be the induced subgraph on $\bigcup A_j$ and X_A the edge shift corresponding to A; by Theorem 4.4.2-4.4.4 of [7] it follows that $h(X_A) = \phi$. By the assumption that the hypotheses of Lemma 3.1.4 do not hold $\forall w \in L(X)$, w contains subwords from at most one M_i. Let $\epsilon > 0$ such that $(\phi + \epsilon) - \lambda < -\epsilon < 0$ then there

is a C such that $\forall n \, |L_A(n)| \leq C e^{n(\phi+\epsilon)}$.

$$|L_X(n)| \leq \sum_{i \leq \psi} \sum_{k=0}^{n} \sum_{v=0}^{k} |L_A(v)| \, |L_{M_i}(n-k)| \, |L_A(k-v)|$$

$$\log|L_X(n)| - n\lambda \leq \log\left[\sum_{i \leq \psi} \sum_{k=0}^{n} \sum_{v=0}^{k} |L_A(v)| \, |L_{M_i}(n-k)| \, |L_A(k-v)| \right] - \log(e^{n\lambda})$$

$$\leq \log\left[\sum_{i \leq \psi} \sum_{k=0}^{n} \sum_{v=0}^{k} C e^{v\phi+\epsilon} \, |L_{M_i}(n-k)| \, C e^{(k-v)\phi+\epsilon} \right] - \log(e^{n\lambda})$$

$$\leq \log\left[\sum_{i \leq \psi} \sum_{k=0}^{n} \sum_{v=0}^{k} \frac{C^2 e^{k\phi+\epsilon}}{e^{k\lambda}} \frac{|L_{M_i}(n-k)|}{e^{(n-k)\lambda}} \right]$$

$$\leq \log\left[\sum_{i \leq \psi} \sum_{k=0}^{n} \sum_{v=0}^{k} \frac{C^2 e^{k\phi+\epsilon}}{e^{k\lambda}} \right]$$

$$\leq \log\left[\sum_{i \leq \psi} \sum_{k=0}^{n} \sum_{v=0}^{k} \frac{C^2}{e^{kQ}} \right]$$

$$= \log\left[\psi C^2 \sum_{k=0}^{n} \frac{k}{e^{kQ}} \right]$$

$$\varlimsup_{n \to \infty} (|L_X(n)| - n\lambda) \leq \varlimsup_{n \to \infty} \left(\log\left[\psi C^2 \sum_{k=0}^{n} \frac{k}{e^{kQ}} \right] \right) < \infty.$$

$\square$

Theorem 3.1.6. *Let X be a 1-D SFT, there exists an algorithm to determine if $H_S(X) = \infty$ or if $H_S(X) < \infty$.*

Proof. Let X be an SFT then there is an algorithm to determine Y, a nearest neighbor, edge shift, such that X is conjugate to Y. Theorem 3.1.2, Theorem 3.1.3, Lemma 3.1.4, and Theorem 3.1.5 provide an algorithm to determine if $H_S(Y) < \infty$ or if $H_S(Y) = \infty$ and by Corollary 3.2.17 $H_S(X) < \infty$ if and only if $H_S(Y) < \infty$ is. $\square$

Theorem 3.1.7. *Let X be a 1-D primitive, nearest neighbor SFT. Then $H_S(X) = 0$ if and only if X is a full shift.*

Proof. Assume X is a full shift on A letters. Then $|L_X(n)| = A^n$ and $h(X) = \log A$. Thus

$$\varlimsup_{n \to \infty} [\log |L_X(n)| - n \log A] = \varlimsup_{n \to \infty} \left[\log \frac{|L_X(n)|}{A^n} \right] = \log 1 = 0$$

By way of contradiction assume X is a primitive, nearest neighbor SFT that is not a full shift. Let A be the adjacency matrix of X and Θ be a matrix the same size of A will all entries of 1. Since X is not a full shift then $A \neq \Theta$. Let ℓ, r be defined as in Theorem 3.1.2. Since X is primitive all entries of ℓ and r are positive. Then since A is a $0 - 1$ matrix it follows that:

$$\ell A r < \ell \Theta r$$

$$\lambda \ell r < \sum_{i,j} \ell_i r_j.$$

Since ℓr are normalized,

$$\lambda < \sum_{i,j} \ell_i r_j$$

$$1 < \frac{1}{\lambda} \sum_{i,j} \ell_i r_j.$$

Then by Theorem 3.1.2,

$$0 < \log \left(\frac{1}{\lambda} \sum_{i,j} \ell_i r_j \right) = H_S(X).$$

$\square$

Lemma 3.1.8. *Let X be a 1-D SFT such that $H_S(X) = 0$ then $\varlimsup_{n\to\infty} \frac{|L_X(2n)|}{|L_X(n)|^2} = 1$.*

Proof. Since $H_S(X) = 0$ by the definition of surface entropy:

$$\varlimsup_{n\to\infty} \log |L_X(n)| - nh = 0$$
$$\varlimsup_{n\to\infty} \log \frac{|L_X(n)|}{e^{nh}} = 0$$
$$\varlimsup_{n\to\infty} \frac{|L_X(n)|}{e^{nh}} = 1.$$

By subadditivity $|L_X(n)| \geq e^{nh}$. Let $\epsilon > 0$ then $\exists N$ such that $\forall n > N$ the following holds.

$$1 \leq \frac{|L_X(n)|}{e^{nh}} \leq 1 + \epsilon \tag{3.1}$$

$$1 \leq \frac{|L_X(2n)|}{e^{2nh}} \leq 1 + \epsilon. \tag{3.2}$$

Using equations 3.1 and 3.2 we obtain lower and upper bounds on the quotient in question.

$$\frac{1}{(1+\epsilon)^2} \leq \frac{\frac{|L_X(2n)|}{e^{2nh}}}{\left(\frac{|L_X(n)|}{e^{nh}}\right)^2} \leq \frac{1+\epsilon}{1}$$

$$\frac{1}{(1+\epsilon)^2} \leq \frac{|L_X(2n)|}{|L_X(n)|^2} \leq \frac{1+\epsilon}{1}.$$

Since ϵ was arbitrary the result holds. $\qquad\square$

Theorem 3.1.9. *Let X be a 1-D, mixing, M-step SFT such that X is not a full shift, then $H_S(X) \neq 0$.*

Proof. Since X is not a full shift $\exists u, v \in L(X)$ such that $uv \notin L(X)$ and WLOG $|u| = |v| = M$. Let $Y = X^M$ be the higher block presentation of X. X is mixing so Y is a

primitive edge shift. Let A be the adjacency matrix of Y, λ be the Perron eigenvalue and r, ℓ be the normalized eigenvectors of A. Define $S_n = \frac{|\{w \in L_X(n): \text{w ends with u}\}|}{|L_X(n)|}$ and $P_n = \frac{|\{w \in L_X(n): \text{w begins with v}\}|}{|L_X(n)|}$.

$$|L_X(2n)| \leq |L_X(n)|^2 - S_n P_n |L_X(n)|^2$$
$$\frac{|L_X(2n)|}{|L_X(n)|^2} \leq 1 - S_n P_n$$

Since X is irreducible by Corollary 3.1.6 it has finite surface entropy so $\exists C$ such that $|L_X(n)| \leq C\lambda^n$. Define $S'_n = \frac{|\{w \in L_Y(n-M+1): \text{w ends with u}\}|}{C\lambda^n}$ and $P'_n = \frac{|\{w \in L_Y(n-M+1): \text{w begins with v}\}|}{C\lambda^n}$. Then $S_n \geq S'_n$ and $P_n \geq P'_n$.

Let j be the column of A representing the letter u.

$$S'_n = \frac{1}{C\lambda^n} \sum_i \left(A^n_{i,j} \right).$$

So by Theorem 4.5.12 of [7]

$$\overline{\lim} \, S'_n = \overline{\lim} \, \frac{1}{C\lambda^n} \sum_i \left(A^n_{i,j} \right) = \frac{1}{C} \sum_i [(r\ell)_{i,j}] = \frac{1}{C} \sum_i r_i \ell_j.$$

Since A was primitive, $\forall i,j$ it is the case that $r_i > 0, \ell_j > 0$ and thus $\overline{\lim} \, S_n > \overline{\lim} \, S'_n > 0$. The same argument provides that $\overline{\lim} \, P_n > \overline{\lim} \, P'_n > 0$. Hence $\overline{\lim} \, \frac{|L_X(2n)|}{|L_X(n)|^2} \leq \overline{\lim} \, (1 - S_n P_n) < 1$ so by Lemma 3.1.8 $H_S(X) > 0$.

$\square$

Theorem 3.1.10. *Let X be a 1-D, irreducible, M-step SFT with period p such that X is not a full shift, then $H_S(X) \neq 0$.*

Proof. Since X is not a full shift $\exists u, v \in L(X)$ such that $uv \notin L(X)$ and WLOG $|u| = |v| = M$. Let $Y = X^M$ be the higher block presentation of X, so Y is an edge

shift. Let A be the adjacency matrix of Y, λ be the Perron eigenvalue of A. Define

$$S_n = \frac{|\{w \in L_X(n) : \text{w ends with u}\}|}{|L_X(n)|}$$

and

$$P_n = \frac{|\{w \in L_X(n) : \text{w begins with v}\}|}{|L_X(n)|}.$$

$$|L_X(2n)| \leq |L_X(n)|^2 - S_n P_n |L_X(n)|^2$$

$$\frac{|L_X(2n)|}{|L_X(n)|^2} \leq 1 - S_n P_n$$

Since X is irreducible by Theorem 3.1.3 it has finite surface entropy so $\exists C$ such that $|L_X(n)| \leq C\lambda^n$. Define

$$S'_n = \frac{|\{w \in L_Y(n - M + 1) : \text{w ends with u}\}|}{C\lambda^n}$$

and

$$P'_n = \frac{|\{w \in L_Y(n - M + 1) : \text{w begins with v}\}|}{C\lambda^n}.$$

Then $S_n \geq S'_n$ and $P_n \geq P'_n$.

$$A^p = \begin{pmatrix} A_0 & 0 & 0 & \ldots & 0 \\ 0 & A_1 & 0 & \ldots & 0 \\ 0 & 0 & A_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & A_{p-1} \end{pmatrix},$$

where each A_i is primitive. Let j be the column of A representing the letter u and $B = A_i$ be the submatrix of A^p containing the non-zero entries of j. Let λ_B be the Perron eigenvalue and r, ℓ be the normalized eigenvectors of B. Note that $\lambda_B = \lambda^p$.

$$S'_{pn} = \frac{1}{C\lambda^{pn}} \sum_i \left(A^p_{i,j} \right) = \frac{1}{C\lambda^{pn}} \sum_i \left(B_{i,j} \right).$$

By Theorem 4.5.12 of [7]

$$\varlimsup_{n\to\infty} S'_n \geq \varlimsup_{n\to\infty} S'_{pn} = \varlimsup_{n\to\infty} \frac{1}{C\lambda^{pn}} \sum_i \left(B_{i,j} \right) = \varlimsup_{n\to\infty} \frac{\lambda^n_B}{C\lambda^{pn}} \sum_i r_i \ell_j = \frac{1}{C} \sum_i r_i \ell_j.$$

Since B was primitive, $\forall i, j$ it is the case that $r_i > 0, \ell_j > 0$ and thus $\varlimsup S_n > \varlimsup S'_n > 0$. The same argument provides that $\varlimsup P_n > \varlimsup P'_n > 0$. Hence $\varlimsup \frac{|L_X(2n)|}{|L_X(n)|^2} \leq \varlimsup (1 - S_n P_n) < 1$ so by Lemma 3.1.8 $H_S(X) > 0$. $\qquad\square$

3.2   Surface Entropy in 2-D Shifts of Finite Type

We now wish to define surface entropy in 2-D. Our goal is to isolate the linear coefficient, C, in $L_X(n,n) \approx e^{n^2 h(X) + n2C}$; S_X below formalizes how to do so. We allow surface entropy to depend on eccentricity α, which we will formalize using $\Xi(\alpha)$, instead of just using squares since as mentioned previously the eccentricity of the word size sequence will effect the value of surface entropy. We allow $\Xi(\alpha)$ to be all sequences approaching the eccentricity α instead of just using sequences of rectangles with exactly eccentricity α because if α is irrational there are no such rectangles and we want parity between a rational and irrational definition.

Definition 3.2.1.

$$S_X(x_n, y_n) = \frac{\log |L_X(x_n, y_n)| - x_n y_n h(X)}{x_n + y_n}.$$

Definition 3.2.2.

$$\Xi(\alpha) = \{\{(x_n, y_n)\} \in (\mathbb{N}^2)^{\mathbb{N}} | \frac{y_n}{x_n} \to \alpha \ and \ x_n \to \infty\}.$$

Definition 3.2.3. *If X is a subshift then the **surface entropy** of X with eccentricity α is:*

$$H_S(X, \alpha) = \sup_{\{\{(x_n, y_n)\} | \frac{y_n}{x_n} \to \alpha, x_n \to \infty\}} \left(\varlimsup_{n \to \infty} \frac{\log |L_X(x_n, y_n)| - x_n y_n h(X)}{x_n + y_n} \right)$$

$$= \sup_{\{(x_n, y_n)\} \in \Xi(\alpha)} \left(\varlimsup_{n \to \infty} S(x_n, y_n) \right).$$

By sub-additivity the surface entropy of a subshift is always nonnegative and by a standard diagonalization argument there is always a sequence of word sizes that achieves the surface entropy as a limsup.

We now present two examples of calculating surface entropy in 2-D SFTs, the first of which illustrates that surface entropy can be infinite and the second that a non-full shift can have 0 surface entropy.

Example 3.2.4. *Let $X \subset \{0, 1, 2\}^{Z^2}$ where the forbidden words of X are*

$$F_X = \{02, 10, 11, 20, 21\}$$

where no vertical configurations are forbidden. A legal configuration in X has independent rows where each row is either ...00000... or ...22222... or ...0001222... Consider $w \in L_X(n, n)$. A row of w is either all 0s, all 2s, or has n locations which could

be 1 so there are $n + 2$ choices for a row in w. There are n such rows in w thus $|L_X(n, n)| = (n + 2)^n$.

$$h(X) = \lim_{n \to \infty} \frac{\log(n + 2)^n}{n^2} = 0.$$

$$H_S(X, 1) \geq \varlimsup_{n \to \infty} \frac{\log |L_X(n, n)| - n^2 h(X)}{2n} = \varlimsup_{n \to \infty} \frac{\log(n + 2)^n}{2n} = \infty.$$

Example 3.2.5. Let $X \subset \{0, 1, 2\}^{Z^2}$ where the forbidden words of X are

$$\mathscr{F}_X = \{02, 10, 11, 20, 21, {}^0_1, {}^1_0, {}^0_2, {}^2_0, {}^1_2, {}^2_1\}.$$

A legal configuration in X has rows that are either ...00000... or ...22222... or ...0001222... and rows are copied vertically thus $|L_X(n, m)| = (n + 2)$.

$$h(X) = \lim_{n \to \infty} \frac{\log(n + 2)}{nm} = 0.$$

$$\begin{aligned}
H_S(X, 1) &= \sup_{(x_n, y_n) \in \Xi(1)} \left(\varlimsup_{n \to \infty} \frac{\log |L_X(x_n, y_n)| - x_n y_n h(X)}{x_n + y_n} \right) \\
&= \sup_{(x_n, y_n) \in \Xi(1)} \left(\varlimsup_{n \to \infty} \frac{\log(x_n + 2)}{x_n + y_n} \right) \\
&= 0.
\end{aligned}$$

Note that the subshifts in Examples 3.2.4 and 3.2.5 both have entropy 0 but will turn out to not be conjugate by Corollary 3.2.17 since Example 3.2.4 has infinite surface entropy and Example 3.2.5 does not.

Definition 3.2.6.

$$\Xi'(X,\alpha) = \{\{(x_n, y_n)\} \in (\mathbb{N}^2)^{\mathbb{N}} | \frac{y_n}{x_n} \to \alpha \ and \ x_n \to \infty \ and \ \lim_{n\to\infty} S(x_n, y_n) \ exists\}.$$

Definition 3.2.7. *Define H_s' as follows:*

$$H_s'(X,\alpha) = \sup_{\{(x_n, y_n)\} \in \Xi'(\alpha)} \left(\lim_{n\to\infty} S(x_n, y_n) \right).$$

Definitions 3.2.6 and 3.2.7 are used to generate an equivalent formulation of surface entropy using a restricted class of word size sequences where only sequences with defined limits are considered.

Proposition 3.2.8. *If X is a subshift then $H_S(X,\alpha) = H_s'(X,\alpha)$.*

Proof. Since $\Xi'(X,\alpha) \subseteq \Xi(\alpha)$ it follows that $H_S(X,\alpha) \geq H_S'(X,\alpha)$. For any $\{(x_n, y_n)\} \in \Xi(\alpha)$ let $S = \overline{\lim} S_X(x_n, y_n)$, then by compactness there is a subsequence $C = \{(x_{n_k}, y_{n_k})\}$ such that $\lim S_X(x_{n_k}, y_{n_k}) = S$ and $C \in \Xi'(X,\alpha)$ thus $H_S'(X,\alpha) \geq S$ and so $H_S'(X,\alpha) \geq H_S(X,\alpha)$. $\square$

Theorem 3.2.9. *For subshifts X, Y and $\alpha \in [0,\infty)$, $H_S(X \times Y, \alpha) \leq H_S(X,\alpha) + H_S(Y,\alpha)$.*

Proof. Let X, Y be subshifts. We note that $h(X \times Y) = h(X) + h(Y)$ [9]. Choose a sequence $O = \{(x_n, y_n)\}$ such that $\overline{\lim} S_{X \times Y}(x_n, y_n) = H_S(X \times Y)$ then $\overline{\lim} S_X(x_n, y_n) \leq H_S(X,\alpha)$ and $\overline{\lim} S_Y(x_n, y_n) \leq H_S(Y,\alpha)$. Thus $\forall (x_n, y_n) \in O$ the following holds:

$$|L_{X \times Y}(x_n, y_n)| = |L_X(x_n, y_n)| \times |L_Y(x_n, y_n)|$$

$$\log |L_{X \times Y}(x_n, y_n)| = \log |L_X(x_n, y_n)| + \log |L_Y(x_n, y_n)|$$

$$\log |L_{X\times Y}(x_n, y_n)| - x_n y_n h(X \times Y)$$

$$= \log |L_X(x_n, y_n)| + \log |L_Y(x_n, y_n)| - x_n y_n (h(X) + h(Y))$$

$$\log |L_{X\times Y}(x_n, y_n)| - x_n y_n h(X \times Y)$$

$$= \log |L_X(x_n, y_n)| - x_n y_n h(X) + \log |L_Y(x_n, y_n)| - x_n y_n h(Y)$$

$$S_{X\times Y}(x_n, y_n) = S_X(x_n, y_n) + S_Y(x_n, y_n)$$

$$\overline{\lim} \, S_{X\times Y}(x_n, y_n) \le \overline{\lim} \, S_X(x_n, y_n) + \overline{\lim} \, S_Y(x_n, y_n) \qquad (*)$$

$$H_S(X \times Y, \alpha) \le H_S(X, \alpha) + H_S(Y, \alpha).$$

$\square$

Lemma 3.2.10. *If $\exists \{(x_n, y_n)\}$ such that $\overline{\lim} \, S_X(x_n, y_n) = H_S(X)$ and $\overline{\lim} \, S_Y(x_n, y_n) = H_S(Y)$ then $H_S(X \times Y, \alpha) = H_S(X, \alpha) + H_S(Y, \alpha)$. In particular this is true if either subshift has the property that any sequence of rectangular word sizes will attain the surface entropy.*

Proof. Follows from equation (*) in the proof of Theorem 3.2.9. $\square$

Recall from Definition 2.1.24 that block gluing is a mixing property of 2-D subshifts. Informally a block gluing subshift is one where there is a universal gap g, such that subwords sufficiently far apart can be chosen independently of each other when determining a legal configuration.

Theorem 3.2.11. *If X is a block gluing subshift of gap size g then $H_S(X, \alpha) < gh$.*

Proof. Let $n, m \in \mathbb{N}$. Then $|L_X(n, m)| < \infty$. For any $k \in \mathbb{N}$ a word of size $(k(n + g), k(m + g))$ contains at least k^2 subwords of size (n, m) that each can be freely chosen. This is accomplished by placing each (n, m) subword a distance of g away from each of the other (n, m) subword as illustrated in Figure 3.1.
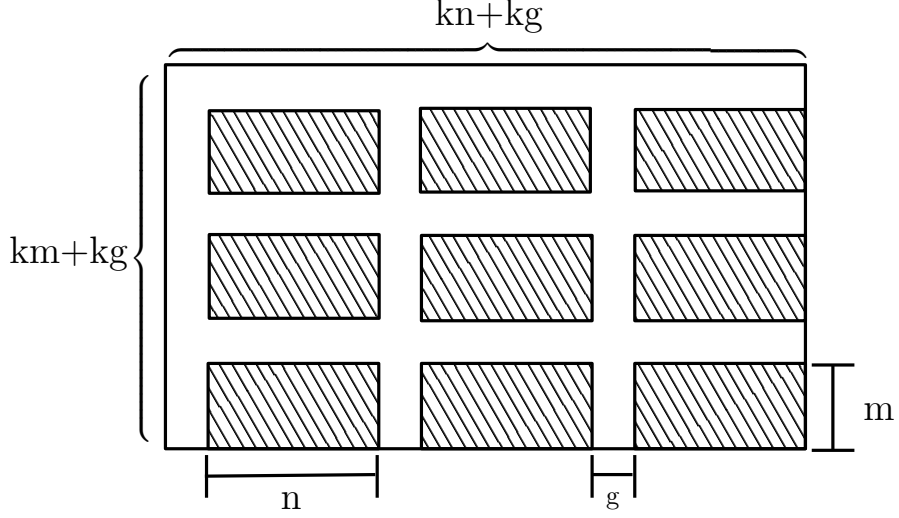
Figure 3.1: Block Gluing: Independent subwords within a large word.

$$|L_X(n,m)|^{k^2} \leq |L_X(k(n+g), k(m+g))|$$

$$k^2 \log |L_X(n,m)| \leq \log |L_X(k(n+g), k(m+g))|$$

$$\frac{k^2 \log |L_X(n,m)|}{k^2(n+g)(m+g)} \leq \frac{\log |L_X(k(n+g), k(m+g))|}{k^2(n+g)(m+g)}$$

$$\frac{\log |L_X(n,m)|}{(n+g)(m+g)} \leq \lim_{k \to \infty} \frac{\log |L_X(k(n+g), k(m+g))|}{k^2(n+g)(m+g)}$$

$$\frac{\log |L_X(n,m)|}{(n+g)(m+g)} \leq h(X)$$

$$\log |L_X(n,m)| \leq h(X)(n+g)(m+g)$$

$$\log |L_X(n,m)| - nmh(X) \leq nmh(X) + ngh(X) + mgh(X) + g^2 h(X) - nmh(X)$$

$$\frac{\log |L_X(n,m)| - nmh(X)}{n+m} \leq \frac{ngh(X) + mgh(X) + g^2 h(X)}{n+m}. \tag{*}$$

Let $\{(x_n, y_n)\} \in \Xi(\alpha)$ then by (*),

$$H_S(X, \alpha) = \lim_{n \to \infty} \frac{\log |L_X(x_n, y_n)| - x_n y_n h(X)}{x_n + y_n} \leq \lim_{n \to \infty} \frac{x_n gh(X) + y_n gh(X) + g^2 h(X)}{x_n + y_n}$$

32

$$\le \lim_{n\to\infty} \frac{gh(X) + \frac{y_n}{x_n}gh(X) + \frac{g^2h(X)}{x_n}}{1 + \frac{y_n}{x_n}}$$

$$= gh.$$

$\square$

Our next examples demonstrate several interesting properties of surface entropy.

Example 3.2.12. *Let* $X = \{0,1\}^{Z^2}$ *then* $|L_X(n,m)| = 2^{nm}$ *and* $h(X) = \log 2$.

$$H_S(X,1) = \sup_{(x_n,y_n)\in\Xi(1)} \left(\varlimsup_{n\to\infty} \frac{\log(2^{x_ny_n}) - x_ny_n\log 2}{x_n + y_n} \right) = 0.$$

Example 3.2.13. *Let* Y *be the 2nd higher block presentation of the fullshift on 2 letters as defined in Definition 2.1.26. Then from Example 3.2.12 and Corollary 3.2.18* $H_S(Y,1) = 2\log(2)$.

The SFTs in Example 3.2.12 and 3.2.13 are conjugate (the proof of this in 2-D is the same as the proof of the 1-D case in [7]) but have different surface entropies. Demonstrating that surface entropy is not a conjugacy invariant.

Example 3.2.14. *Let* $X \subset \{0,1\}^{Z^2}$ *where the forbidden words of* X *are* $\{\begin{smallmatrix}0\\1\end{smallmatrix}, \begin{smallmatrix}1\\0\end{smallmatrix}\}$. *The rows of this shift look like 1-D fullshifts and are copied vertically.* $|L_X(n,m)| = 2^n$ *thus* $h(X) = 0$.

$$H_S(X,\alpha) = \sup_{(x_n,y_n)\in\Xi(\alpha)} \left(\varlimsup_{n\to\infty} \frac{\log(2^{x_n})}{x_n + y_n} \right) = \frac{\log 2}{1 + \alpha}.$$

Example 3.2.14 demonstrates that the surface entropy of a subshift can depend on the eccentricity of the rectangles used. This is different than entropy which is independent of the eccentricity.

Example 3.2.15. *Let Y be the SFT described in Theorem 16 of [10] with $k = 10^6$. Y has an alphabet of the integers $0, 1, ..., k$ along with 6 grid symbols (Figure 3.2). A point $y \in Y$ has grid symbols partitioning the plane into rectangles and each rectangle is filled with rows of integers. These rows of integers must alternate between rows of all 0s and rows of all non-0s. If the bottom of a rectangle is closed by grid symbols, then the row of integers directly above it must be all 0s. Y is block gluing, a rather strong mixing property, and yet we will use the bounds on L_Y proved in [10] to show that $\lim_{n \to \infty} S_Y(n, n)$ does not exist. (Though those bounds are only proved for so-called locally admissible words, they in fact also apply to $L_Y(n, n)$ as needed for our calculation.) It is an immediate consequence of the bounds in [10] that $h(Y) = \frac{\log k}{2}$.*

Consider square words with odd dimensions, $(2n - 1, 2n - 1)$, then

$$k^{(2n-1)(n)} \leq L_Y(2n - 1, 2n - 1)$$

$$\lim_{n \to \infty} \frac{\log\left(k^{(2n-1)(n)}\right) - (2n-1)^2 \frac{\log k}{2}}{4n - 2} = \frac{\log k}{4} \leq \lim_{n \to \infty} S_Y(2n - 1, 2n - 1).$$

Consider square words with even dimensions, $(2n, 2n)$, then

$$L_Y(2n, 2n) \leq 2^{10n-4}(4n^2 + 1)48^{4n}k^{2n^2}$$

$$\varlimsup_{n \to \infty} S_Y(2n, 2n) \leq \varlimsup_{n \to \infty} \frac{\log\left(2^{10n-4}(4n^2 + 1)48^{4n}k^{2n^2}\right) - 4n^2 \frac{\log k}{2}}{4n} = \log 48 + \frac{5}{2} \log 2.$$

Since the sequences of even squares and the sequence of odd squares are both subsequences of the sequence of squares

$$\varliminf_{n \to \infty} S_Y(n, n) \leq \varlimsup_{n \to \infty} S_Y(2n, 2n) \leq \log 48 + \frac{5}{2} \log 2$$

$$< \frac{\log k}{4} \leq \varliminf_{n \to \infty} S_Y(2n - 1, 2n - 1) \leq \varlimsup_{n \to \infty} S_Y(n, n).$$

34

Figure 3.2: The grid symbols of Example 3.2.15

s_1 s_2 s_3 s_4 s_5 s_6
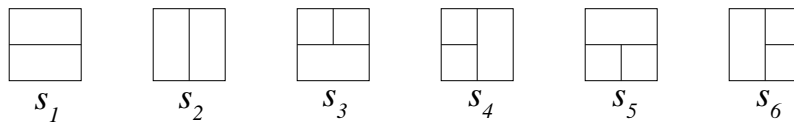
Example 3.2.15 serves to demonstrate that surface entropy is not as well behaved in 2-D as it was in 1-D. In the 1-D case an SFT having the mixing property not only guaranteed that this limit would exist, but allowed us to prove an explicit algebraic formula for its value. Demonstrated here however we see that even a property as strong as block gluing can't even guarantee the existence of this limit. This also serves to show that many of the "problems" that may arise in calculations or theorems concerning surface entropy might not be fixed by simply assuming a mixing property; there is something more subtle causing issues. This example also serves to illustrate the necessity of using a limit superior in the definition of surface entropy. Even when only considering a sequence of squares, which are not even all of the sequences in $\Xi(1)$, we cannot assume convergence of limits.

Theorem 3.2.16. *Let X and Y be conjugate SFTs where the sliding block code realizing the conjugacy has a window of size k. If $H_S(X, \alpha) < \infty$ then $H_S(Y, \alpha) < H_S(X, \alpha) + 2kh(X)$.*

Proof. Let f be the $k \times k$ sliding block code from X to Y as defined in Definition 2.1.25. Then f is a surjection from $L_X(n + 2k, m + 2k)$ to $L_Y(n, m)$ so $\forall m, n \in \mathbb{N}$ the following holds:

$$|L_Y(n, m)| \leq |L_X(n + 2k, m + 2k)|. \tag{3.3}$$

Thus

$$|L_Y(n,m)| - (n+2k)(m+2k)h(X) \le |L_X(n+2k,m+2k)| - (n+2k)(m+2k)h(X)$$

$$\log|L_Y(n,m)| - nmh(X) \le \log|L_X(n+2k,m+2k)| - (nm+2nk+2mk+4k^2)h(X)$$

$$+ (2nk+2mk+4k^2)h(X)$$

$$\frac{\log|L_Y(n,m)| - nmh(X)}{n+m}$$

$$\le \frac{\log|L_X(n+2k,m+2k)| - (nm+2nk+2mk+4k^2)h(X) + (2nk+2mk+4k^2)h(X)}{n+m}$$

$$= \frac{\log|L_X(n+2k,m+2k)| - (nm+2nk+2mk+4k^2)h(X)}{n+m} + \frac{(2nk+2mk+4k^2)h(X)}{n+m}$$

$$= \frac{(n+m+4k)}{n+m}\left[\frac{\log|L_X(n+k,m+k)| - (nm+2nk+2mk+4k^2)h(X)}{n+m+4k}\right]$$

$$+ \frac{(2nk+2mk+4k^2)h(X)}{n+m}$$

$$= \frac{(n+m+4k)}{n+m}\left[\frac{\log|L_X(n+2k,m+2k)| - (nm+2nk+2mk+4k^2)h(X)}{n+m+4k}\right] + 2kh(X).$$

Let $m_i, n_i \in \Xi(\alpha)$, by the previous inequality it follows that:

$$H_S(Y,\alpha) = \varlimsup_{i\to\infty} \frac{\log|L_Y(n_i,m_i)| - n_i m_i h(X)}{n_i+m_i} \le$$

$$\varlimsup_{i\to\infty} \frac{(n_i+m_i+4k)}{n_i+m_i}\left[\frac{\log|L_X(n_i+2k,m_i+2k)| - (n_i m_i + 2n_i k + 2m_i k + 4k^2)h(X)}{n_i+m_i+4k}\right]$$

$$+ 2kh(X)$$

$$H_S(Y,\alpha) \le H_S(X,\alpha) + 2kh(X).$$

□

Corollary 3.2.17. *If X and Y are conjugate subshifts then $H_S(X, \alpha) < \infty$ if and only if $H_S(Y, \alpha) < \infty$.*

Proof. Follows immediately from Theorem 3.2.16. □

Corollary 3.2.18. *If Y is the Mth higher block representation of X and $H_S(X) < \infty$ then $H_S(Y, \alpha) = H_S(X, \alpha) + (M - 1)h(X)$.*

Proof. This is similar to the proof of Theorem 3.2.16 except that the sliding block code for the Mth higher block presentation yields a bijection from $L_X(n+M-1, m+M-1)$ to $L_Y(n, m)$, and so the inequality (3.3),

$$|L_Y(n, m)| \leq |L_X(n + 2k, m + 2k)|,$$

becomes the following equality

$$|L_Y(n, m)| = |L_X(n + M - 1, m + M - 1)|.$$

□

3.3   2-D Surface Entropy as a function of α

To this point we've considered surface entropy by fixing an α and comparing the behaviors of the surface entropies of different shifts. A natural question is whether fixing a shift and allowing α to vary would also yield interesting results. In fact when fixing X, $H_S(X, \alpha)$ becomes a function on the positive reals and so questions we might ask about such functions become relevant. Specifically is $H_S(X, \alpha)$ a continuous function? It turns out that it is not and we present a counter example demonstrating

that as a function of α, $H_S(X, \alpha)$ need not be continuous. However, we then show that $H_S(X, \alpha)$ is upper semicontinuous. In general the discontinuities in an upper semicontinuous function can be arbitrarily large, we prove that this is not the case for $H_S(X, \alpha)$ in Theorems 3.3.18 and 3.3.19 which give a type of upper bound on how badly discontinuous the function can be.

The following example will demonstrate a subshift in which $H_S(X, \alpha)$ will be discontinuous.

Example 3.3.1. *Let X be an SFT with an alphabet of $0, 1$ along with border tiles lower left corner, upper left corner, upper right corner, lower right corner, left edge, right edge, upper edge, and bottom edge. Each border tile will also have one of 1000 colors. Points of X must have the following structure: at most one square with perimeter of border tiles, colored independently, all locations inside of such squares (by inside we mean to the right of the left edge, above the bottom edge, etc) are independently colored by $\{0, 1\}$ and all other locations are 0. (Since X is closed and shift-invariant, it must also contain "limit points" such as a point with no border tiles, and those containing 2 sides of an "infinite" square.) Though X is technically not an SFT as written (there is no way to preclude border tiles forming non-square rectangles), it can be made an SFT by adding a "diagonal signal" within the rectangle which negligibly affects word count.*

We claim that $H_S(X, \alpha)$ is not continuous wrt α. To demonstrate this we must find upper and lower approximations for the $L_X(x_n, y_n)$, we will assume WLOG that $\alpha > 1$ and thus eventually $y_n > x_n$.

The number of words on an $x_n \times y_n$ rectangle is trivially at least $2^{x_n y_n}$, since all 0-1 words are legal.

For an upper bound, first note that every word on an $x_n \times y_n$ rectangle, $y_n > x_n$, either contains an entire perimeter of a square consisting of border tiles or does not.

A word containing such an entire perimeter is determined entirely by the side length of the square (less than or equal to x_n possibilities), the location of its upper left corner (less than or equal to $x_n y_n$ possibilities), and the letters on the border and interior of the square. This gives an upper bound of

$$x_n^2 y_n 1000^{4x_n-4} 2^{(x_n-2)^2}$$

on the number of such words.

A word which does not contain an entire perimeter of a square is determined entirely by the subrectangle of 0-1 symbols it contains (which may be empty), the letters within that rectangle, and all border letters inside the word. There are less than or equal to $x_n^2 y_n^2$ choices for the subrectangle (from less than or equal to $x_n y_n$ choices for the upper-left and lower-right corners). There are less than or equal to $2x_n + y_n$ border tiles (coming from the case where three edges contain border tiles), and so the maximum number of choices for the border and 0-1 tiles is $1000^{2x_n+y_n} 2^{x_n y_n - (2x_n+y_n)}$. This gives an upper bound of

$$x_n^2 y_n^2 1000^{2x_n+y_n} 2^{x_n y_n - (2x_n+y_n)}$$

on the number of such words. Combining these bounds,

$$2^{x_n y_n} \leq |L_X(x_n, y_n)| \leq x_n^2 y_n 1000^{4x_n-4} 2^{(x_n-2)^2} + x_n^2 y_n^2 1000^{2x_n+y_n} 2^{x_n y_n - (2x_n+y_n)}$$

These bounds on $|L_X(x_n, y_n)|$ allow us to calculate bounds on $h(X)$.

$$\lim_{n\to\infty} \frac{\log\left(2^{x_n y_n}\right)}{x_n y_n} \leq \lim_{n\to\infty} \frac{\log\left(|L_X(x_n, y_n)|\right)}{x_n y_n}$$

$$\leq \lim_{n\to\infty} \frac{1}{x_n y_n} \log\left(x_n^2 y_n 1000^{4x_n-4} 2^{(x_n-2)^2} + x_n^2 y_n^2 1000^{2x_n+y_n} 2^{x_n y_n-(2x_n+y_n)}\right)$$

$$\log 2 \leq h(X) \leq \lim_{n\to\infty} \frac{1}{x_n y_n} \log\left(2^{x_n y_n}\left(\frac{2^{x_n^2}}{2^{x_n y_n}} + 1\right)\right)$$

$$= \lim_{n\to\infty} \frac{1}{x_n y_n} \log\left(2^{x_n y_n}\left(2^{x_n(x_n-y_n)} + 1\right)\right) = \log 2.$$

We now give a lower bound on $H_S(X,1)$. $(n,n) \in \Xi(1)$ and for each n there is a point in X which contains an $n \times n$ square. So every configuration of such a square is in $L_X(n,n)$.

$$1000^{4n-4} 2^{(n-2)^2} \leq |L_X(n,n)|$$

$$\varliminf_{n\to\infty} \frac{\log\left(1000^{4n-4} 2^{(n-2)^2}\right) - n^2 \log 2}{2n} \leq \varlimsup_{n\to\infty} S_X(n,n) \leq H_S(X,1)$$

$$2\log 500 \leq 2 \leq H_S(X,1).$$

We now give an upper bound on surface entropy. We first simplify the upper bound:

$$|L_X(x_n, y_n)| \leq x_n^2 y_n^2 2^{x_n y_n}(1000^{4x_n-4} 2^{-x_n(y_n-x_n)-4x_n+4} + 500^{2x_n+y_n}).$$

Clearly the first term inside the parentheses approaches 0 as $n \to \infty$, and so will be eventually dominated by the second. This yields

$$|L_X(x_n, y_n)| \leq 2x_n^2 y_n^2 2^{x_n y_n} 500^{2x_n+y_n}.$$

Taking logs, subtracting $x_n y_n h(X)$, dividing by $x_n + y_n$, and taking a limsup as $n \to \infty$ yields

$$H_S(X, \alpha) \leq \overline{\lim} \frac{\log 2 + 2\log x_n + 2\log y_n + (2x_n + y_n)\log 500}{x_n + y_n} = \frac{2 + \alpha}{1 + \alpha} \log 500.$$

Since $\lim_{\alpha \to 1+} \frac{2+\alpha}{1+\alpha} \log 500 = 1.5 \log 500$, and since $H_S(X, 1) \geq 2 \log 500$, we've shown the desired discontinuity of $H_S(X, \alpha)$ at $\alpha = 1$.

Theorem 3.3.2. *Let X be a subshift. Then $H_S(X, \alpha)$ is an upper semicontinuous function of α.*

Proof. Let $\alpha \in \mathbb{R}^+$ and $\alpha_j \to \alpha$. For each α_j there exists a sequence $\{(x_n^j, y_n^j)\} \in \Xi'(\alpha_j)$ s.t $\lim_{n \to \infty} S(x_n^j, y_n^j) = H_S(X, \alpha_j)$. We construct a new eccentricity sequence by the following. For each $k \in \mathbb{N}$ choose (x_n^j, y_n^j) so that the following hold: $n \geq k$, $j = k$ and $H_S(X, \alpha_j) - \frac{1}{2^n} \leq S(x_n^j, y_n^j) \leq H_S(X, \alpha_j) + \frac{1}{2^n}$. Let (x_k, y_k) be the element chosen. Then $\{(x_k, y_k)\} \in \Xi(\alpha)$ and $\overline{\lim}_{j \to \infty} H_S(X, \alpha_j) = \overline{\lim}_{k \to \infty} S(x_k, y_k) \leq H_S(X, \alpha)$. $\square$

In order to later prove Theorems 3.3.18 and 3.3.19 we will need to develop a few tools. These theorems will be proved by bounding the word count of a large rectangle using the word count of smaller rectangles.
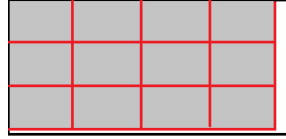
Figure 3.3: "Standard" Tiling

A standard approach to such a proof is to use smaller rectangles of a fixed size and simply fill as many copies as will fit into the larger rectangle as demonstrated in Figure 3.3. This has an error involved, the uncovered area of the larger rectangle, that is large in comparison to either dimensions of the larger rectangle but is small in comparison to the area of the larger rectangle. For calculations involving entropy this error is small enough and will vanish in the limit. When dealing with surface entropy however this error is not necessarily small in

comparison to the perimeter and thus may not vanish. The following definitions and theorems are intended to design a different sort of tiling algorithm that produces an error term that will vanish in comparison to the perimeter.

Definition 3.3.3. *Let N be a set of numbers. A **tileset**, (T, N), is a set of rectangles, T such that $\forall n \in N$ there is a rectangle in T with width n and a rectangle in T of height n.*

Definition 3.3.4. *Given a rectangle $R = (w \times h)$ and a tileset (T, N), generate an ordered pair (C, R_r) with the following non-deterministic algorithm:*

1. *Choose a tile, $(t_w \times t_h) \in T$ such that $t_h = h$ and $t_w \leq w$. If no such tile exists, terminate the algorithm.*

2. *Repeat for dimensions $(w - t_w, h)$.*

*When the algorithm terminates, C is the number of tiles used before termination and R_r the remaining rectangle (the **remainder**) that couldn't be tiled. R_r has the dimension used in step 1 when it terminated.*

$TileLEFT(R,(T,N))$ is the set of all ordered pairs that could be generated by this algorithm.

Since this algorithm is possibly not deterministic, each (C, R_r) represents a possible choice path of the algorithm. Some tilesets will produce a deterministic algorithm, in which case (C, R_r) will be unique. A later lemma shows that given the proper restrictions that the initial segments of (C, R_r) will be deterministic.

Definition 3.3.5. *Given a rectangle $R = (w \times h)$ and a tileset (T, N), generate an ordered pair (C, R_r) with the following non-deterministic algorithm:*
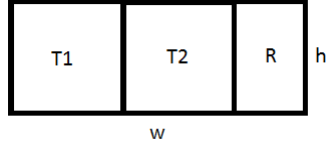
(a) Tiling from the Left
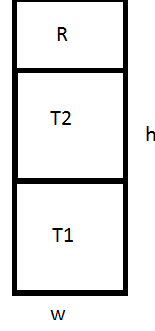
(b) Tiling from the bottom

Figure 3.4: Tiling algorithm examples.

1. *Choose a tile, $(t_w \times t_h) \in T$ such that $t_h \leq h$ and $t_w = w$. If no such tile exists, terminate the algorithm.*

2. *Repeat for dimensions $(w, h - t_h)$.*

When the algorithm terminates, C is the number of tiles used before termination and R_r is the remainder.

TileBOTTOM(R,(T,N)) *is the set of all ordered pairs that could be generated by this algorithm.*

Definition 3.3.6. *Given a rectangle, R, and a tileset, (T, N), the **tiling algorithm** will return a set of sequences of ordered pairs (C_n, R_n) where C_n is the number of tiles used at the nth step of the algorithm and R_n is the remainder uncovered by the first n steps of the algorithm.*

TILE(R,(T,N)) $= \{\{(C_n, R_n)\} | \{(C_n, R_n)\}$

was produced by a run of the following non-deterministic algorithm$\}$

43

Algorithm:

Let $C_0 = 0$ and $R_0 = R$. Let $O = 0$ if $\frac{h}{w} \leq 1$, otherwise $O = 1$. Starting with $n = 1$,
for each n do the following:

If $n + O$ is odd, choose $(C_n, R_n) \in TileLEFT(R_{n-1}, (T, N))$.

If $n + O$ is even, choose $(C_n, R_n) \in TileBOTTOM(R_{n-1}, (T, N))$.

Remark: If for any $\exists n > 1$ s.t $C_n = C_{n+1} = 0$ then $\forall i > n, C_i = 0$

Definition 3.3.7. *A **perfect tiling** of R is $TILE(R, (Squares, \mathbb{R}))$.*

Note that in the case of a perfect tiling that $[C_1; C_2, C_3...]$ will be the continued fraction expansion of $\frac{h}{w}$; since this tiling algorithm is just an application of the Euclidean algorithm.

Definition 3.3.8. *A rectangle of eccentricity α is an α-**tile**.*

Definition 3.3.9. *Define Γ to be the function such that if x is the eccentricity of a rectangle then $\Gamma(x)$ is the eccentricity of the remainder after performing 1 step of a perfect tiling.*

Based on the above definition the following can easily be checked:

$$\Gamma(x) = \begin{cases} x - \lfloor x \rfloor & x \geq 1 \\ \\ \frac{1}{\frac{1}{x} - \lfloor \frac{1}{x} \rfloor} & x < 1. \end{cases}$$

$$\Gamma\left(\frac{1}{x}\right) = \frac{1}{\Gamma(x)}.$$

Definition 3.3.10. *An $w \times h$ rectangle such that either:*

1) $|h - w\alpha| \leq z$

OR

2) $\left|w - \frac{h}{\alpha}\right| \leq z$ *is called an* (α, z)**-tile**.

Definition 3.3.11. $(\alpha, 1)$*-tiles are called are **near** α**-tiles**.*

Since we are working in an integer lattice, near α-tiles are rectangles where one side is fixed and the other is as close to being eccentricity of α as possible.

Definition 3.3.12. *For a rectangle, R, $TILE(R, ($**near** α**-tiles**, \mathbb{N}))$ is a **near** α **tiling** of R.*

$TILE(R, (\alpha\text{-tiles}, \mathbb{R}))$ is deterministic since there is only one tile of each width and height of eccentricity α. Using (α, z)-tiles is not necessarily deterministic, but the following proposition will show the needed level of determinism.

Proposition 3.3.13. *If $\frac{\beta}{\alpha} \in \mathbb{R} - \mathbb{Q}$ then there exists a γ and V such that if $R = w \times h$ is a rectangle, $\left|\frac{h}{w} - \beta\right| < \gamma$, and $w > V$ then if $(K_1, R_1) = TileBOTTOM((1 \times \beta), (\alpha\text{-tiles}, \mathbb{R}))$ and $(C_1, R_1) \in TileBOTTOM(R, ((\alpha, z)\text{-tiles}), \mathbb{N})$ it follows that $K_1 = C_1$.*
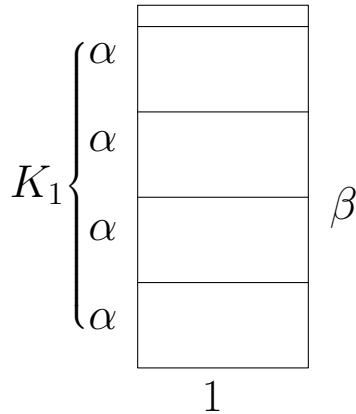

Figure 3.5: K_1 is maximized.

Proof. For use later we will need the following fact.

$$\left| \frac{h}{w} - \beta \right| < \gamma \Rightarrow \left| \frac{\frac{h}{\alpha}}{w} - \frac{\beta}{\alpha} \right| < \frac{\gamma}{\alpha} \Rightarrow \frac{\beta - \gamma}{\alpha} < \frac{\frac{h}{\alpha}}{w} < \frac{\beta + \gamma}{\alpha}.$$

(3.4)

Since the tiling algorithm is greedy, this means that K_1 will be as large as possible so that the tiles are still inside the $(1, \beta)$ rectangle (See Figure 3.5).

$$K_1 \alpha < \beta < (K_1 + 1)\alpha \implies K_1 < \frac{\beta}{\alpha} < K_1 + 1.$$

K_1 is an integer, and $\frac{\beta}{\alpha} \in \mathbb{R} - \mathbb{Q}$, so for small enough γ, $K_1 < \frac{\beta - \gamma}{\alpha}$ and $\frac{\beta + \gamma}{\alpha} < K_1 + 1$. Since these inequalities are strict, there is a V such that if $w > V$, $K_1 < \frac{\beta - \gamma}{\alpha} \frac{1}{1 + \frac{z}{w\alpha}}$, and $\frac{\beta + \gamma}{\alpha} \frac{1}{1 - \frac{z}{w\alpha}} < K_1 + 1$. Choose such a γ and V and any $w > V$. Then

$$K_1 < \frac{\beta - \gamma}{\alpha} \frac{1}{1 + \frac{z}{w\alpha}} \implies K_1(1 + \frac{z}{w\alpha}) < \frac{\beta - \gamma}{\alpha} < \frac{\frac{h}{\alpha}}{w} \text{ this follows from inequality (3.4).}$$

$$\implies K_1(w\alpha + z) < h.$$

In TileBOTTOM the width of a tile used must be the width of the original rectangle and so $(w, (w\alpha + z))$ is the tallest tile that can be used in the algorithm that produces C_1. The above shows that even if only tiles of this size are used $C_1 \geq K_1$, since at least K_1 such tiles will fit by the definition of K_1. Similarly,

$$\frac{\beta + \gamma}{\alpha} \frac{1}{1 - \frac{z}{w\alpha}} < K_1 + 1 \implies$$

$$\frac{\frac{h}{\alpha}}{w} < \frac{\beta + \gamma}{\alpha} < (K_1 + 1)\left(1 - \frac{z}{w\alpha}\right) \implies h < (K_1 + 1)(w\alpha - z).$$

46

Again since the width is fixed in TileBOTTOM $(w, (w\alpha - z))$ is the shortest tile that can be used to produce C_1. The above shows that K_1 is the maximum number of such tiles that can fit, and so $K_1 \geq C_1$. Thus $C_1 = K_1$.

$\square$

Proposition 3.3.14. *If $\frac{\beta}{\alpha} \in \mathbb{R} - \mathbb{Q}$ then there exists a γ and V such that if $R = w \times h$ is a rectangle, $\left|\frac{h}{w} - \beta\right| < \gamma$, and $h > V$, then if $(K_1, R_1) = TileLEFT((1 \times \beta), (\alpha\text{-tiles}, \mathbb{R}))$ and $(C_1, R_1) \in TileLEFT(R, ((\alpha, z)\text{-tiles}), \mathbb{N})$ it follows that $K_1 = C_1$.*

Proof. This follows a similar proof to Proposition 3.3.13. $\square$

It is easily checked that if $\{(C_i, R_i)\} = TILE((1 \times \frac{\beta}{\alpha}, (squares, \mathbb{R}))$ then $\{(C_i, \alpha R_i)\} = TILE((1 \times \beta, ((\alpha - \text{tiles}, \mathbb{R}))$.

Proposition 3.3.15. *Let $R = (w, h)$ have eccentricity β and $\{(k_i, R_i)\} \in TILE(R, ((\alpha, z)\text{-tiles}, \mathbb{N}))$. Let T_k be the tiles used by the algorithm and $w_k \times h_k$ be the dimensions of T_k. Then $\sum_{k=1}^{\infty} (2w_k + 2h_k) \leq \gamma(2w + 2h)$ where $\gamma = max\left(1 + z + \alpha, 1 + z + \frac{1}{\alpha}\right)$.*

In other words the sums of the perimeters of the tiles used by the tiling algorithm are bounded by a constant multiple of the perimeter of the rectangle they are tiling. This constant only depends on α and z.

Proof. Let Γ be the set of tiles used during TileLEFT and Γ' the tiles used during TileBOTTOM. Note $z + \alpha \geq \frac{h_k}{w_k}$ and $z + \frac{1}{\alpha} \geq \frac{w_k}{h_k}$ since these are (α, z)-tiles.

$$2w + 2h \geq \sum_{k \in \Gamma} 2w_k + \sum_{k \in \Gamma'} 2h_k$$

$$\gamma(2w + 2h) \geq \sum_{k \in \Gamma} 2w_k(1 + (\alpha + z)) + \sum_{k \in \Gamma'} 2h_k \left(1 + \left(\frac{1}{\alpha} + z\right)\right)$$

47

$$\geq \sum_{k \in \Gamma} 2w_k \left(1 + \frac{h_k}{w_k}\right) + \sum_{k \in \Gamma'} 2h_k \left(1 + \frac{w_k}{h_k}\right)$$

$$\geq \sum_{k \in \Gamma \cup \Gamma'} 2w_k + 2h_k.$$

$\square$

Lemma 3.3.16. $\forall \alpha, \beta$ such that $\frac{\beta}{\alpha} \notin \mathbb{Q}$ and $\forall M \in \mathbb{N}, \delta > 0, Z \in \mathbb{N}$ then $\exists \gamma \in \mathbb{R}, V \in \mathbb{N}$ such that for any rectangle R with dimensions $w \times h$ satisfying $\left|\frac{h}{w} - \beta\right| < \gamma, \left|\frac{w}{h} - \frac{1}{\beta}\right| < \gamma$ and $w > V, h > V$ and for any $(C_i, R_i) \in TILE(R, ((\alpha, 1), \mathbb{N}))$ the following will hold.

Let $R^s = w^s \times h^s$ where $w^s = 1$ and $h^s = \frac{\beta}{\alpha}$ and $(K_i, (R^s)_i) = TILE(R_s, (squares, \mathbb{R}))$. Define w_i, h_i as the width and height of R_i, define $(w^s)_i, (h^s)_i$ similarly for $(R^s)_i$.

(L1) $\forall i \leq M; C_i = K_i$.

(L2) $\left|\frac{h_i}{h} - \frac{(h^s)_i}{h_s}\right| < \delta$ and $\left|\frac{w_i}{w} - \frac{(w^s)_i}{w_s}\right| < \delta$.

(L3†) $\left|\frac{h_i}{w_i} - \alpha \Gamma^i(\frac{\beta}{\alpha})\right| < \delta$.

(L3‡) $\left|\frac{w_i}{h_i} - \frac{1}{\alpha}\Gamma^i(\frac{\alpha}{\beta})\right| < \delta$.

(L4) $w_i > Z$ and $h_i > Z$.

Proof. Choose $Z \in \mathbb{N}$ and $\delta > 0$. We will proceed by induction on M. First begin by considering only one step of the tiling algorithm; $M = 1$. Let $0 < \delta^\dagger \leq \min\{\delta, \frac{1}{2}\alpha\Gamma\left(\frac{\beta}{\alpha}\right), \frac{\delta}{2}\left[\alpha\Gamma\left(\frac{\beta}{\alpha}\right)\right]^2\}$. Note that $\delta^\dagger \neq 0$ since $\frac{\beta}{\alpha} \notin \mathbb{Q}$. Let $\delta' < \min\{\frac{\delta^\dagger}{2(\beta+1)}, \frac{\delta^\dagger}{2\left(\frac{1}{\beta}+1\right)}\}$. Choose $\gamma < 1$ and V as in the hypothesis of Proposition 3.3.13 and 3.3.14 such that all of the following hold: $\frac{V}{\alpha} > K_1, \frac{K_1}{V} < \delta, \lfloor\frac{V}{\alpha}\rfloor > 0, C_1\alpha(\frac{1}{V} + \gamma) < \delta', \frac{C_1}{V} + \frac{C_1}{\alpha}\gamma < \delta', \frac{\alpha}{\beta}h^s\gamma \leq \delta', V\left|\frac{(w^s)_i}{w^s} - \delta\right| > Z$ and $V\left|\frac{(h^s)_i}{h^s} - \delta\right| > Z$.

From Proposition 3.3.13 and 3.3.14 it follows that $C_1 = K_1$ and so (L1) holds. We now proceed to verify (L2) and (L3) by considering which step of the of the tiling algorithm is performed.

Case 1: Assume first step of the tiling algorithm is TileBOTTOM. We begin by verifying $(L2)$. Since this is TileBOTTOM the widths are unchanged so, $\frac{(w^s)_1}{w^s} = 1 = \frac{w_1}{w}$, verifying the first half of $(L2)$. The tiles used are $(\alpha, 1)$-tiles thus the height of each tile is in the interval $[w\alpha - \alpha, w\alpha + \alpha]$. If each tile had height exactly $w\alpha$ then the height of the remainder would be $h - C_1 w\alpha$. This is likely not the case, however each tile can only differ from $w\alpha$ by at most α and thus the actual height of the remainder can only differ from $h - C_1 w\alpha$ by at most $C_1 \alpha$, giving the following inequality,

$$(|h_1 - (h - C_1 w\alpha)|) < C_1 \alpha. \tag{3.5}$$

Recall from the hypothesis that $h^s = \frac{\beta}{\alpha}$, thus when tiling by squares with $w^s = 1$ it follows that $(h^s)_1 = h^s - C_1 w^s = \frac{\beta}{\alpha} - C_1$.

We now verify the remaining half of $(L2)$.

$$
\begin{aligned}
\left| \frac{h_1}{h} - \frac{(h^s)_1}{h^s} \right| &\leq \left| \frac{h_1}{h} - \frac{\alpha}{\beta}\left(\frac{\beta}{\alpha} - C_1 \right) \right| \\
&\leq \left| \frac{h_1}{h} - \left(1 - \frac{\alpha}{\beta} C_1 \right) \right| \\
&\leq \left| \frac{h_1}{h} - \left(1 - \frac{w}{h} C_1 \alpha + \frac{w}{h} C_1 \alpha - \frac{\alpha}{\beta} C_1 \right) \right| \\
&\leq \left| \frac{h_1}{h} - \left(1 - \frac{w}{h} C_1 \alpha \right) \right| + \left| \frac{w}{h} C_1 \alpha - \frac{\alpha}{\beta} C_1 \right| \\
&\leq \left| \frac{h_1}{h} - \left(1 - \frac{w}{h} C_1 \alpha \right) \right| + C_1 \alpha \left| \frac{w}{h} - \frac{1}{\beta} \right| \\
&\overset{(3.5)}{\leq} \frac{C_1 \alpha}{h} + C_1 \alpha \left| \frac{w}{h} - \frac{1}{\beta} \right| \\
&\leq C_1 \alpha \left(\frac{1}{h} + \gamma \right) < \delta' < \delta.
\end{aligned}
$$

The above implies (L2) for Case 1. We will need the following inequality (*) to complete the proof of this case.

$$\left| \frac{h_1}{h} - \frac{(h^s)_1}{h^s} \right| < \delta'$$

$$\frac{h}{w} \left| \frac{h_1}{h} - \frac{(h^s)_1}{h^s} \right| < \delta'(\beta + \gamma)$$

$$(*) \quad \left| \frac{h_1}{w} - \frac{\alpha}{\beta}(h^s)_1 \frac{h}{w} \right| < \delta'(\beta + \gamma).$$

Then the proof of the inequality in $(L3^\dagger)$ is as follows:

$$\left| \frac{h_1}{w_1} - \alpha\Gamma\left(\frac{\beta}{\alpha}\right) \right| = \left| \frac{h_1}{w} - \alpha(h^s)_1 \right|$$

$$\leq \left| \frac{h_1}{w} - \frac{\alpha}{\beta}(h^s)_1 \frac{h}{w} + \frac{\alpha}{\beta}(h^s)_1 \frac{h}{w} - \alpha(h^s)_1 \right|$$

$$\leq \left| \frac{h_1}{w} - \frac{\alpha}{\beta}(h^s)_1 \frac{h}{w} \right| + \left| \frac{\alpha}{\beta}(h^s)_1 \frac{h}{w} - \alpha(h^s)_1 \right|$$

$$\overset{(*)}{\leq} \delta'(\beta + \gamma) + \frac{\alpha}{\beta}(h^s)_1 \left| \frac{h}{w} - \beta \right|$$

$$\leq \delta'(\beta + \gamma) + \frac{\alpha}{\beta}(h^s)_1 \gamma$$

$$\leq \delta'(\beta + \gamma) + \frac{\alpha}{\beta}h^s\gamma \leq \delta^\dagger \leq \delta.$$

Hence $(L3^\dagger)$ holds for Case 1.

The previous also implies that $\frac{h_1}{w_1} \geq \frac{1}{2}\alpha\Gamma\left(\frac{\beta}{\alpha}\right)$. So we use the fact that

$$\left| \frac{1}{x} - \frac{1}{y} \right| = \frac{|x - y|}{|xy|} \qquad \text{and} \qquad \Gamma\left(\frac{1}{x}\right) = \frac{1}{\Gamma(x)}$$

to conclude that

50

$$\left| \frac{1}{\frac{h_1}{w_1}} - \frac{1}{\alpha\Gamma\left(\frac{\beta}{\alpha}\right)} \right| = \left| \frac{w_1}{h_1} - \frac{1}{\alpha}\Gamma\left(\frac{\alpha}{\beta}\right) \right| = \frac{\left| \frac{h_1}{w_1} - \alpha\Gamma\left(\frac{\beta}{\alpha}\right) \right|}{\left| \frac{h_1}{w_1}\alpha\Gamma\left(\frac{\beta}{\alpha}\right) \right|} \leq \frac{\left| \frac{h_1}{w_1} - \alpha\Gamma\left(\frac{\beta}{\alpha}\right) \right|}{\left| \frac{1}{2}\left[\alpha\Gamma\left(\frac{\beta}{\alpha}\right)\right]^2 \right|} \leq \delta,$$

which proves $(L3^\ddagger)$ for Case 1.

Case 2: Assume the first step of the tiling algorithm is TileLEFT. In TileLEFT the heights are unchanged so $\frac{(h^s)_1}{h^s} = 1 = \frac{h_1}{h}$. Since the tiles used are $(\alpha, 1)$-tiles the width of the tiles used are in the interval $[\frac{h}{\alpha} - 1, \frac{h}{\alpha} + 1]$ which implies $\left| \frac{w^1}{w} - \left(1 - \frac{C_1}{\alpha}\frac{h}{w}\right) \right| < C_1$. Also note that $w^s = \frac{\alpha}{\beta}$.

$$\begin{aligned}
\left| \frac{w_1}{w} - \frac{(w^s)_1}{w^s} \right| &\leq \left| \frac{w_1}{w} - \left(1 - C_1\frac{\beta}{\alpha}\right) \right| \\
&\leq \left| \frac{w_1}{w} - \left(1 - \frac{C_1}{\alpha}\frac{h}{w} + \frac{C_1}{\alpha}\frac{h}{w} - C_1\frac{\beta}{\alpha}\right) \right| \\
&\leq \left| \frac{w_1}{w} - \left(1 - \frac{C_1}{\alpha}\frac{h}{w}\right) \right| + \left| \frac{C_1}{\alpha}\frac{h}{w} - C_1\frac{\beta}{\alpha} \right| \\
&\leq \frac{C_1}{w} + \frac{C_1}{\alpha}\gamma < \delta' < \delta.
\end{aligned}$$

This implies (L2) for Case 2. As in Case 1 we need the following inequality $(**)$.

$$\left| \frac{w_1}{w} - \frac{(w^s)_1}{w^s} \right| < \delta'$$

$$\frac{w}{h}\left| \frac{w_1}{w} - \frac{(w^s)_1}{w^s} \right| < \delta'\left(\frac{1}{\beta} + \gamma \right)$$

$$(**)\quad \left| \frac{w_1}{h} - \frac{\beta}{\alpha}(w^s)_1\frac{w}{h} \right| < \delta'\left(\frac{1}{\beta} + \gamma \right)$$

51

We then prove $(L3^\ddagger)$ as follows.

$$
\begin{aligned}
\left| \frac{w_1}{h_1} - \frac{1}{\alpha}\Gamma\left(\frac{\alpha}{\beta}\right) \right| &= \left| \frac{w_1}{h_1} - \frac{1}{\alpha}(w^s)_1 \right| \\
&\leq \left| \frac{w_1}{h} - \frac{\beta}{\alpha}(w^s)_1\frac{w}{h} + \frac{\beta}{\alpha}(w^s)_1\frac{w}{h} - \frac{1}{\alpha}(w^s)_1 \right| \\
&\leq \left| \frac{w_1}{h} - \frac{\beta}{\alpha}(w^s)_1\frac{w}{h} \right| + \left| \frac{\beta}{\alpha}(w^s)_1\frac{w}{h} - \frac{1}{\alpha}(w^s)_1 \right| \\
&\overset{(**)}{\leq} \delta\left(\frac{1}{\beta} + \gamma\right) + \frac{\beta}{\alpha}(w^s)_1\left|\frac{w}{h} - \frac{1}{\beta}\right| \\
&\leq \delta\left(\frac{1}{\beta} + \gamma\right) + \frac{\beta}{\alpha}(w^s)_1\gamma \leq \delta^\dagger < \delta.
\end{aligned}
$$

This implies $(L3^\ddagger)$ for Case 2. The same reasoning that gives $(L3^\dagger)\to(L3^\ddagger)$ in Case 1 will apply to show that $(L3^\dagger)$ holds for Case 2.

In either case $h, w > V$, so from $(L2)$:

$$
\begin{aligned}
\left| \frac{h_1}{h} - \frac{(h^s)_1}{h^s} \right| &< \delta \\
-\delta < \frac{h_1}{h} - \frac{(h^s)_1}{h^s} &< \delta \\
\frac{(h^s)_1}{h^s} - \delta < \frac{h_1}{h} &< \frac{(h^s)_1}{h^s} + \delta \\
h\left(\frac{(h^s)_1}{h^s} - \delta\right) < h_1 &< h\left(\frac{(h^s)_1}{h^s} + \delta\right) \\
Z < V\left(\frac{(h^s)_1}{h^s} - \delta\right) < h\left(\frac{(h^s)_1}{h^s} - \delta\right) &< h_1.
\end{aligned}
$$

A similar argument holds for w hence $(L4)$ for both cases.

We now proceed with the inductive step. We say $R = w \times h$ satisfies $*(\gamma, V, \Phi)$ if $\left|\frac{h}{w} - \Phi\right| < \gamma, \left|\frac{w}{h} - \frac{1}{\Phi}\right| < \gamma$ and $w > V, h > V$.

Let $\delta > 0$ and $Z \in \mathbb{N}$.

By the 1 step case $\exists \gamma, V$ such that for any rectangle R satisfying $*(\gamma, V, \Phi)$ then

(1-1) $C_1 = K_1$.

(1-2) $\left|\frac{h_1}{h} - \frac{h_1^s}{h^s}\right| < \delta$ and $\left|\frac{w_1}{w} - \frac{w_1^s}{w^s}\right| < \delta$.

(1-3†) $\left|\frac{h_1}{w_1} - \alpha\Gamma^1\left(\frac{\Phi}{\alpha}\right)\right| < \delta$.

(1-3‡) $\left|\frac{w_1}{h_1} - \frac{1}{\alpha}\Gamma^1\left(\frac{\alpha}{\Phi}\right)\right| < \delta$.

(1-4) $w_1 > Z$ and $h_1 > Z$.

By the inductive hypothesis for a given m $\exists \gamma_m, V_m$ such that any rectangle R satisfying $*(\gamma_m, V_m, \beta)$ then $\forall 1 \le i \le m$.

(I1) $C_i = K_i$.

(I2) $\left|\frac{h_i}{h} - \frac{(h^s)_i}{h^s}\right| < \gamma_m$ and $\left|\frac{w_i}{w} - \frac{(w^s)_i}{w^s}\right| < \gamma$.

(I3†) $\left|\frac{h_i}{w_i} - \alpha\Gamma^i\left(\frac{\beta}{\alpha}\right)\right| < \gamma$.

(I3‡) $\left|\frac{w_i}{h_i} - \frac{1}{\alpha}\Gamma^i\left(\frac{\alpha}{\beta}\right)\right| < \gamma$.

(I4) $w_i > V$ and $h_i > V$.

Let R be a rectangle satisfying $*(\gamma_m, V_m, \beta)$ Then we can apply m steps of the tiling algorithm to R and by (I3†), (I3‡), and (I4) R_m satisfies $*\left(\gamma, V, \alpha\Gamma^m\left(\frac{\beta}{\alpha}\right)\right)$ and hence when 1 more step of the tiling algorithm is applied to R_m properties (1-1) to (1-4) will hold for the result R_{m+1}. To complete the proof we need to show that properties $(L1)$ to $(L4)$ hold for R_{m+1}.

(L1) Since $\left|\frac{h}{w} - \alpha\Gamma^m\left(\frac{\beta}{\alpha}\right)\right| < \gamma$, $(C_{m+1}, \cdot) = (C_1, \cdot) \in TILE(R_m, ((\alpha, 1), \mathbb{N})) = (K_1, \cdot) \in TILE\left(\left(1 \times \alpha\Gamma^m\left(\frac{\beta}{\alpha}\right)\right), ((\alpha, 1), \mathbb{N})\right) = (K_{m+1}, \cdot) \in TILE\left(\left(1 \times \Gamma\left(\frac{\beta}{\alpha}\right)\right), ((\alpha, 1), \mathbb{N})\right)$.

(L2) This follows from the fact that $(h_m)_1 = h_{m+1}$.

(L3†)By the definition of Γ, $\alpha\Gamma^m\left(\frac{\beta}{\alpha}\right)$ is the eccentricity of the R_m remainder, so by the application of (1-3†) to $\Phi = \alpha\Gamma^m\left(\frac{\beta}{\alpha}\right)$;

$$\left|\frac{h_{m+1}}{w_{m+1}} - \alpha\Gamma^{m+1}\left(\frac{\beta}{\alpha}\right)\right| = \left|\frac{(h_m)_1}{(w_m)_1} - \alpha\Gamma\left(\frac{\alpha\Gamma^m\left(\frac{\beta}{\alpha}\right)}{\alpha}\right)\right| = \left|\frac{(h_m)_1}{(w_m)_1} - \alpha\Gamma\left(\frac{\Phi}{\alpha}\right)\right| < \delta.$$

53

(L3‡) As above $\Phi = \alpha \Gamma^m \left(\frac{\beta}{\alpha} \right)$. Then apply (1-3‡) to (I3‡)

$$\left| \frac{w_{m+1}}{h_{m+1}} - \frac{1}{\alpha} \Gamma^{m+1} \left(\frac{\alpha}{\beta} \right) \right| = \left| \frac{(w_m)_1}{(h_m)_1} - \frac{1}{\alpha} \Gamma \left(\frac{\alpha}{\alpha \Gamma^m \left(\frac{\beta}{\alpha} \right)} \right) \right| = \left| \frac{(w_m)_1}{(h_m)_1} - \frac{1}{\alpha} \Gamma \left(\frac{\alpha}{\Phi} \right) \right| < \delta.$$

(L4) Is a direct consequence of (I4).

$\square$

Lemma 3.3.17. *Let R be a rectangle for which the tiling algorithm does not terminate for at least M steps. Then each step after the first will return a non-zero number of tiles used. After M steps of the algorithm the following will hold: $R_x \leq \frac{x_n}{2^{\frac{M}{2}}}, R_y \leq \frac{y_n}{2^{\frac{M}{2}}}$ where R_x is the width of the remainder and R_y is the height of the remainder.*

Proof. Consider a step in the tiling algorithm that applies TileLEFT and uses a non-zero number of tiles. The width of the remainder must be less than half the width of the original, otherwise more tiles of the size already used could be placed. Similarly the height of the remainder is reduced by more than half when applying TileBOTTOM.

Assume toward contradiction that a step in the tiling algorithm after the first returns 0 tiles used but the algorithm has not yet terminated. Since the algorithm is has not terminated the previous step and next step cannot return 0 tiles as this is the end condition. Assume WLOG that the previous and next steps are both TileLEFT. Since the previous TileLEFT ended this implies that no more tiles can fit into that configuration from the left, however since the current TileBOTTOM returned 0 it has not changed the configuration and simpled passed it back to TileLEFT. This is the same configuration that has already caused TileLEFT to halt and so cannot be tiled by TileLEFT. This contradiction implies that no step after the first can return 0 until the algorithm terminates in 0s.

Since every step will apply a tile and ever time a tile is applied the corresponding dimensions of the remainder is reduced by at least a factor of $\frac{1}{2}$ after M steps it follows that $R_x \leq \frac{x_n}{2^{\frac{M}{2}}}, R_y \leq \frac{y_n}{2^{\frac{M}{2}}}$. $\qquad\square$

Theorem 3.3.18. *Let X be a subshift. Let $\beta, \alpha \in (0, \infty)$ and assume $\frac{\beta}{\alpha} \notin \mathbb{Q}$. Let $P = \max(2 + \alpha, 2 + \frac{1}{\alpha})$ then $H_S(X, \beta) \leq PH_S(X, \alpha)$.*

Proof. Choose $\epsilon > 0$. Let $P = \max(2 + \alpha, 2 + \frac{1}{\alpha})$, this is the bound given by Proposition 3.3.15 to the ratio of the perimeters of the tiles to the perimeters of the rectangle they are tiling.

Let $\{(x_n, y_n)\} \in (\mathbb{N}^2)^{\mathbb{N}}$ such that $\frac{y_n}{x_n} \to \beta$ and $x_i \to \infty$. Consider the set $T = \{(n, \lfloor n\alpha \rfloor) | n \in \mathbb{N}\} \cup \{(\lfloor \frac{n}{\alpha} \rfloor, n) | n \in \mathbb{N}\} \cup \{(n, \lceil n\alpha \rceil) | n \in \mathbb{N}\} \cup \{(\lceil \frac{n}{\alpha} \rceil, n) | n \in \mathbb{N}\}$; T contains every width and every height rectangle in $\mathbb{Z}^2$, furthermore $T \subseteq (\alpha, 1)$-tiles. $\exists N_T$ such that $\forall n > N_T$; $S(n, \lfloor n \rfloor) < H_S(\alpha) + \epsilon$, $S(\lfloor \frac{n}{\alpha} \rfloor, n) < H_S(\alpha) + \epsilon$, $S(n, \lceil n \rceil) < H_S(\alpha) + \epsilon$, and $S(\lceil \frac{n}{\alpha} \rceil, n) < H_S(\alpha) + \epsilon$.

For each $(w_n, h_n) \in T$ there exists an ϵ_n such that $S(w_n, h_n) < H_S(\alpha) + \epsilon_n$. There are only finitely many $\epsilon_n > \epsilon$. Let $E = \max(\epsilon_n)$.

Choose M such that $2^{\frac{-M}{2}} PE < \epsilon$

By the previous Lemma 3.3.16 $\exists \gamma, N_M$ such that if (x_n, y_n) satisfies $*(\gamma, N_M)$ then the tiling of (x_n, y_n) will agree with the perfect tile of β by α for the first M steps and so that after M steps the remainder will have dimensions greater than N_T. There is an N_γ such that if $n > N_\gamma$ (x_n, y_n) will satisfy $*(\gamma, N_M)$.

Let $N > \max(N_T, N_\gamma)$. Let $\{K_j, R_j\} = Tile((x_n, y_n), (T, \mathbb{Z}^2))$. Let l be the number of non-zero K_j, note that $l > M$. Let C_1 be the number of $(n, \lfloor n\alpha \rfloor)$ tiles used in the K_1 step. C_2 be the number of $(n, \lceil n\alpha \rceil)$ tiles used in the K_1 step. Continue to define C_i for $1 \leq i \leq 4l$. Then $\sum_{j=1}^{l} K_j = \sum_{i=1}^{4l} C_i$. Let (w_i, h_i) be the type of tile counted by C_i.

$$\frac{\log|L_X(x_n, y_n)| - x_n y_n h}{x_n + y_n}$$

$$\leq \frac{\sum_{i=1}^{4l}[C_i \log|L_X(w_i, h_i)|]] - x_n y_n h}{x_n + y_n}$$

$$= \frac{\sum_{i=1}^{4l}[C_i \log|L_X(w_i, h_i)| - C_i w_i h_i h]}{x_n + y_n}$$

$$\leq \frac{\sum_{i=i}^{4l}[C_i(H_S(\alpha) + \epsilon_i) * (w_i + h_i)]}{x_n + y_n}$$

$$= H_S(\alpha)\frac{\sum_{i=1}^{4l}[C_i(w_i + h_i) + C_i(w_i + h_i)\frac{\epsilon_i}{H_S(\alpha)}]}{x_n + y_n}$$

$$= H_S(\alpha) * \frac{\textbf{sum of perimeter of } \alpha\textbf{-rectangles}}{\textbf{perimeter } \beta\textbf{-rectangle}} + \frac{\sum_{i=1}^{4l}[C_i(w_i + h_i)\epsilon_i]}{x_n + y_n}$$

$$= PH_S(\alpha) + \frac{\sum_{i=1}^{l}[C_i(w_i + h_i)\epsilon_i]}{x_n + y_n}$$

$$= PH_S(\alpha) + \frac{\sum_{w_i \wedge h_i > N_T}[C_i(w_i + h_i)\epsilon]}{x_n + y_n} + \frac{\sum_{w_i \vee h_i \leq N_T}[C_i(w_i + h_i)\epsilon_i]}{x_n + y_n}$$

$$\leq PH_S(\alpha) + P\epsilon + \frac{\sum_{w_i \vee h_i \leq N_T}[C_i(w_i + h_i)\epsilon_i]}{x_n + y_n}$$

$$\leq PH_S(\alpha) + P\epsilon + \frac{\sum_{w_i \vee h_i \leq N_T}[C_i(w_i + h_i)E]}{x_n + y_n}.$$

Since the dimensions of the remainder after M steps are each greater than N_T, the tiles in the following sum are all contained within the remainder after M steps. Rx and Ry be the dimensions of the remainder after M steps. Then the sum of the dimensions of these tiles are also bounded by Proposition 3.3.15 by $P(R_x + R_y)$.

$$\frac{\sum_{w_i \vee h_i \leq N_T}[C_i(w_i + h_i)E]}{x_n + y_n} \leq \frac{PE(R_x + R_y)]}{x_n + y_n}.$$

By Lemma 3.3.17 $R_x \leq \frac{x_n}{2^{\frac{M}{2}}}, R_y \leq \frac{y_n}{2^{\frac{M}{2}}}$.

$$\frac{\sum_{w_i \vee h_i \leq N_T}[C_i(w_i + h_i)E]}{x_n + y_n} \leq \frac{PE(R_x + R_y)]}{x_n + y_n} \leq PE\left(\frac{x_n}{2^{\frac{M}{2}}} + \frac{y_n}{2^{\frac{M}{2}}}\right)\frac{1}{x_n + y_n} \leq \frac{PE}{2^{\frac{M}{2}}}.$$

Thus:

$$\frac{\log|L_X(x_n, y_n)| - x_n y_n h}{x_n + y_n} \leq PH_S(\alpha) + P\epsilon + \frac{\sum_{w_i \vee h_i \leq N_T}[C_i(w_i + h_i)E]}{x_n + y_n}$$

$$\leq PH_S(\alpha) + P\epsilon + \frac{PE}{2^{\frac{M}{2}}}$$

$$\leq PH_S(\alpha) + P\epsilon + \epsilon.$$

Since ϵ was arbitrary, for any sequence in $\Xi(\beta), \overline{\lim}_{n \to \infty}\left(\frac{\log|L_X(x_n, y_n)| - x_n y_n h}{x_n + y_n}\right) \leq PH_S(\alpha)$ and thus $H_S(\beta) \leq PH_S(\alpha)$.

$\square$

Theorem 3.3.19. *Let X be a subshift. Let $\beta, \alpha \in (0, \infty)$ and assume $\frac{\beta}{\alpha} \in \mathbb{Q}$ where $\frac{p}{q}$ is the reduced form of $\frac{\beta}{\alpha}$. Let $z = \max\{p, q\}$ and let $\gamma = \max\{1 + z + \alpha, 1 + z + \frac{1}{\alpha}\}$ then $H_S(X, \beta) \leq \gamma H_S(X, \alpha)$.*

Proof. Choose $\epsilon > 0$. Let $\frac{p}{q}$ be the reduced fraction of $\frac{\beta}{\alpha}$.

Let $\{(x_i, y_i)\} \in (\mathbb{N}^2)^{\mathbb{N}}$ such that $\frac{y_i}{x_i} \to \beta$ and $x_i \to \infty$.

Let $\{(C_i, R_i)\} = TILE(1 \times \beta, (\alpha\text{-tiles}, \mathbb{R}))$. Since $\frac{\beta}{\alpha} \in \mathbb{Q} \; \exists M$ such that $\forall i > M$; $C_i = 0$ and the area of $R_i = 0$. Call this tiling the template. Let $Z = \sum_{i=1}^{M}(C_i)$.

The $(1, \beta)$ rectangle can be subdivided such that there are p equal rows and q equal columns; this forms a grid on the template. Note that the tiles will fall along grid lines.

Based on this template a tiling can be defined for each (x_i, y_i). Using only widths of $\lfloor \frac{x_i}{q} \rfloor$ or $\lceil \frac{x_i}{q} \rceil$ and heights of $\lfloor \frac{y_i}{p} \rfloor$ or $\lceil \frac{y_i}{p} \rceil$ subdivide (x_i, y_i) into p rows of integer height and q columns of integer width. There is a correspondence between the grid on the template and this grid on (x_i, y_i); use this correspondence to map the tiles from the template to tiles on (x_i, y_i). See Figure 3.6.
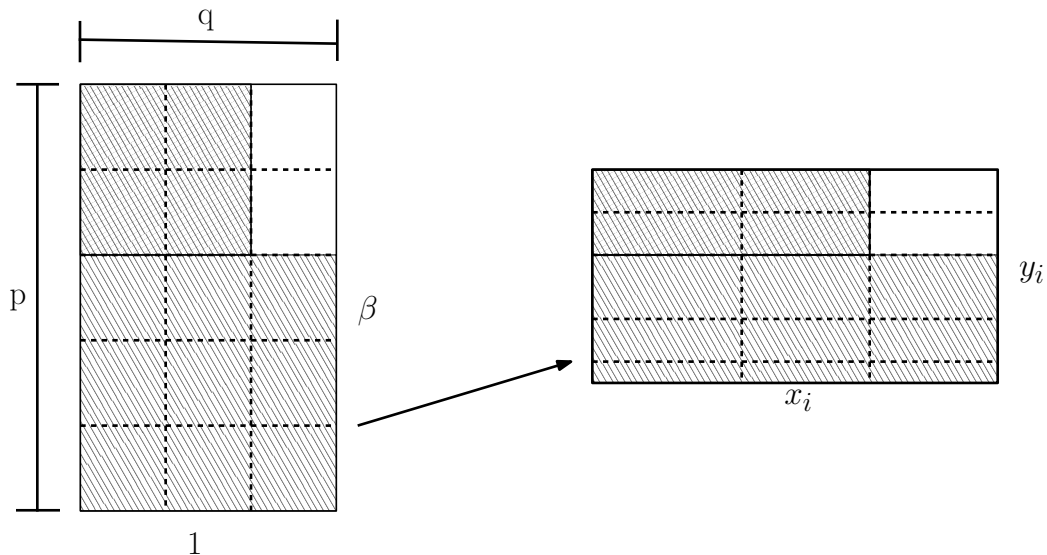
Figure 3.6: The template procedure used to prove Theorem 3.3.19.

For $1 \leq k \leq Z$ let T_k be a tile used in the template and $T_k^i = (w_k^i, h_k^i)$ be the corresponding tile used to tile (x_i, y_i) and g_k^i be the number of grid partitions used vertically or horizontally in T_k^i (Since the grid partitions are themselves α tiles the number of partitions horizontally will always equal the number of partitions vertically in an α-tile. This gives us bounds on the actual height and width of the T_k^i tiles based only on the number of grid partitions they use:

$$g_k^i \frac{y_i}{p} - g_k^i \leq h_k^i \leq g_k^i \frac{y_i}{p} + g_k^i$$
$$g_k^i \frac{x_i}{q} - g_k^i \leq w_k^i \leq g_k^i \frac{x_i}{q} + g_k^i.$$

58

Using these bounds we can bound the eccentricity of the tiles used to cover (x_i, y_i),

$$\frac{g_k^i \frac{y_i}{p} - g_k^i}{g_k^i \frac{x_i}{q} + g_k^i} \leq \frac{h_k^i}{w_k^i} \leq \frac{g_k^i \frac{y_i}{p} + g_k^i}{g_k^i \frac{x_i}{q} - g_k^i}$$

$$\frac{\frac{y_i}{p} - 1}{\frac{x_i}{q} + 1} \leq \frac{h_k^i}{w_k^i} \leq \frac{\frac{y_i}{p} + 1}{\frac{x_i}{q} - 1}$$

$$\lim_{i \to \infty} \frac{\frac{y_i}{p} - 1}{\frac{x_i}{q} + 1} \leq \lim_{i \to \infty} \frac{h_k^i}{w_k^i} \leq \lim_{i \to \infty} \frac{\frac{y_i}{p} + 1}{\frac{x_i}{q} - 1}$$

$$\beta \frac{q}{p} \leq \lim_{i \to \infty} \frac{h_k^i}{w_k^i} \leq \beta \frac{q}{p}.$$

Since $\beta \frac{q}{p} = \alpha$ this implies that $\lim_{i \to \infty} \frac{h_k^i}{w_k^i} = \alpha$. For each k the proportion of w_i^k to x_i is fixed by the template so since $x_i \to \infty$, $\lim_{i \to \infty} w_i = \infty$ it follows that $(w_k^i, h_k^i) \in \Xi(\alpha)$. Thus $\exists N_k$ such that $\forall i > N_k$, $S(w_k^i, h_k^i) \leq H_S(\alpha) + \epsilon$. When $i > \max\{N_k\}_{k=1}^Z$,

$$|L_X(x_i, y_i)| \leq \prod_{k=1}^Z |L_X(w_k^i, h_k^i)|$$

$$\log |L_X(x_i, y_i)| \leq \sum_{k=1}^Z \log |L_X(w_k^i, h_k^i)|$$

$$\log |L_X(x_i, y_i)| - x_i y_i h \leq \sum_{k=1}^Z \left[\log |L_X(w_k^i, h_k^i)| - w_k^i h_k^i h \right]$$

$$\frac{\log |L_X(x_i, y_i)| - x_i y_i h}{x_i + y_i} \leq \frac{\sum_{k=1}^Z \left[\log |L_X(w_k^i, h_k^i)| - w_k^i h_k^i h \right]}{x_i + y_i}$$

$$\frac{\log |L_X(x_i, y_i)| - x_i y_i h}{x_i + y_i} \leq \frac{\sum_{k=1}^Z \left[(H_S(\alpha) + \epsilon)(w_k^i + h_k^i) \right]}{x_i + y_i}$$

$$\frac{\log |L_X(x_i, y_i)| - x_i y_i h}{x_i + y_i} \leq (H_S(\alpha) + \epsilon) \frac{\sum_{k=1}^Z \left[(w_k^i + h_k^i) \right]}{x_i + y_i}.$$

Let $z = \max\{p, q\}$ and by Proposition 3.3.15 let $\gamma = \max\{1 + z + \alpha, 1 + z + \frac{1}{\alpha}\}$. Since ϵ was arbitrary,

$$\varlimsup_{i \to \infty} \frac{\log |L_X(x_i, y_i)| - x_i y_i h}{x_i + y_i} \leq \gamma H_S(\alpha).$$

This is true for any such sequence so $H_S(\beta) \leq \gamma H_S(\alpha)$. $\qquad\square$

Corollary 3.3.20. *Let X be a subshift. If $\exists \alpha \in (0, \infty)$ such that $H_S(X, \alpha) = 0$ then $\forall \beta \in (0, \infty), H_S(X, \beta) = 0$.*

Proof. Direct consequence of Theorem 3.3.18 and 3.3.19. $\qquad\square$

Corollary 3.3.21. *Let X be a subshift. If $\exists \alpha \in (0, \infty)$ such that $H_S(X, \alpha) = \infty$ then $\forall \beta \in (0, \infty), H_S(X, \beta) = \infty$.*

Proof. Direct consequence of Theorem 3.3.18 and 3.3.19. $\qquad\square$

Chapter 4

Realizations of Surface Entropy

4.1 Computability Properties of Entropies

An immediate question in the field of symbolic dynamics is to classify the set of numbers that are the entropy of an SFT. Since 1 dimensional SFTs are closely related to linear algebra, the property classifying the entropies of 1-D SFTs, being a Perron number, is similar to being an algebraic number. Unfortunately such an algebraic property cannot classify the entropies obtained from 2-D SFTs. Instead it turns out that the classification comes from computability theory. One framing of the computability properties of numbers is in regard to how they can be approximated by Turing machines. We will first examine these properties and then show how they relate to entropy and surface entropy.

There are only countably many possible Turing machines and so there can only be countably many computable numbers. It cannot be the case that all real numbers are computable. However, demonstrating a non-computable number is difficult since the digits of such a number cannot be effectively listed. One possible method to construct

such a number is to leverage undecidable problems in computational theory; one such problem is known as the halting problem.

Definition 4.1.1. *Let $\mathscr{T} = \{T_k\}_{k \in \mathbb{N}}$ be the set of all Turing machines. The **Halting Problem** states that there cannot exist $H \in \mathscr{T}$ with $\forall k \in \mathbb{N}, H(k) \in \{0, 1\}$ such that $\forall T_k \in \mathscr{T}, H(k) = 1$ if and only if T_k halts in finite time when run with empty input.*

We can now use this information to give an explicit construction of a non-computable number.

Example 4.1.2. *The following example demonstrates how to encode a non-computable number. Let T_i be the ith Turing machine. Define the following sequence, where $\mathbb{C}_n(i)$ is the ith digit of $\mathbb{C}_n$ in binary.*

$$\mathbb{C}_n(i) = \begin{cases} 0 & \text{if } i \leq n \text{ and } T_i \text{ has halted after } n \text{ steps} \\ 1 & \text{otherwise} \end{cases}$$

Now consider $\mathbb{C} = \lim_{n \to \infty} \mathbb{C}_n$. $\mathbb{C}(i) = 0$ if and only if T_i halted. If $\mathbb{C}$ were computable, then there would exist a Turing machine T with $T(i) = \mathbb{C}(i)$, which would solve the unsolvable halting problem.

However $\mathbb{C}$ is not completely without computability properties. It is not hard to check that there exists T with $T(n) = \mathbb{C}_n$, $\mathbb{C}_n \to \mathbb{C}$, and $\forall n, \mathbb{C}_n \geq \mathbb{C}$, and so $\mathbb{C}$ is CFA. It is this particular computability property that turns out to be important in classifying the entropies of 2-D SFTs.

Hochman and Meyerovitch[4] showed that the entropies of 2-D SFTs were exactly the CFA numbers. Our main results are the following: every CFA number can be realized as the surface entropy of a 2-D SFT, and more surprisingly, there are non-CFA numbers that can also be realized as the surface entropy of a 2-D SFT.

4.2 Realizing CFA Surface Entropy

To define an SFT we need to provide a finite alphabet and a finite list of forbidden words. For a simple system like the golden mean shift, this can often be done directly, i.e. 1s cannot be adjacent. However for more complex SFTs the size of the alphabet and forbidden list can become so enormous as to render the SFT unrecognizable. In order to combat this phenomenon, we will describe our SFT as a series of layers. The first layer of the construction will be a complete SFT with a defined alphabet and finite forbidden list. Each subsequent layer will be additions to the previous layers instead of being a complete SFT in its own right. Each new layer will add information to the existing letters and add forbidden patterns depending on both this new and previous information. Individual layers will be named so they they can be referred to separately as projections of the full SFT. We will first construct an example SFT to demonstrate this technique.

Example 4.2.1. *Let $\mathbb{G}$ be the 2-D golden mean shift on $\{0,1\}$. The alphabet is $\mathscr{A}_{\mathbb{G}} = \{0,1\}$ and the forbidden list is $\mathscr{F}_{\mathbb{G}} = \{11, \frac{1}{1}\}$.*

Let $(\mathbb{G}, \mathbb{C})$ be the SFT that colors each 0 either blue or red with the rule that a 0 can only be red if it is not adjacent to any 1s. $\mathscr{A}_{(\mathbb{G},\mathbb{C})} = \{1, 0_B, 0_R\}$ and $\mathscr{F}_{(\mathbb{G},\mathbb{C})} = \{11, \frac{1}{1}, O_R 1, 10_R, \frac{1}{0_R} \frac{0_R}{1}\}$.

We will refer to the projection of $(\mathbb{G}, \mathbb{C})$ onto the 2nd coordinate as $\mathbb{C}$.

We now wish to demonstrate that any CFA number can be the surface entropy of a 2-D SFT. For this result we will use the SFT constructed in [4] with the only modification being how many 1s are independently 2 colored. As such this SFT will only require 3 layers; $\mathbb{HM}$ will be the SFT constructed in [4], $\mathbb{D}_d$ will (possibly) distinguish a single row of $\mathbb{HM}$, and $\mathbb{D}_{\mathbb{HM}}$ will double all of the 1s in the distinguished row.

Choose $\gamma \in [0,1]$ such that γ is CFA. Let $\gamma' = \frac{2\gamma}{\log 2}$. Since γ is CFA, $\frac{2}{\log 2}$ is computable, and CFA is closed under multiplication then γ' is CFA.

4.2.1   Layer $\mathbb{HM}$

By [4] there exists an SFT $\mathbb{HM}$ with the following properties. $\mathbb{HM}$ has a sublayer consisting of 0s and 1s. For each $x \in \mathbb{HM}$ each row of x has the same $\{0,1\}$ Toeplitz sequence embedded as one of its layers. By the construction of $\mathbb{HM}$ there is an x_f, the frequency of 1s appearing in each row of this sublayer. Then $\exists x \in \mathbb{HM}$ such that $x_f = \gamma'$ and $\forall x \in \mathbb{HM}, x_f \leq \gamma'$. Moreover $h(\mathbb{HM}) = H_S(\mathbb{HM}, 1) = 0$. To achieve $H_S(\mathbb{HM}, 1) = 0$ we choose the SFT constructed in [4] at the end of Section 7 before entropy is added by doubling. Calculations similar to those used to construct the SFT in our main result show the surface entropy is as desired.

4.2.2   Layer $\mathbb{D}_d$

Let $\mathbb{D}_d$ be the SFT with alphabet $\{B, R, T\}$ and forbidden list

$$\mathscr{F}_d = \{BR, BT, RT, RB, TB, TR, \tfrac{T}{B}, \tfrac{B}{T}, \tfrac{R}{R}, \tfrac{B}{R}, \tfrac{R}{T}\}.$$

$\mathbb{D}_d$ is the SFT consisting of points which are either all Bs, all Ts, or 1 distinguished row of Rs with Bs below and Ts above.

Proposition 4.2.2. $H_S(\mathbb{D}_d, 1) = 0$.

Proof. By using the same calculation as below it is easy to check that $h(\mathbb{D}_d) = 0$. Let $\{(x_n, y_n)\} \in \Xi(1)$. In a word of size (x_n, y_n) there are $y_n + 1$ ways to choose which, if

any, row is distinguished. Thus

$$H_S(\mathbb{D}_d, 1) = \lim_{n \to \infty} \frac{\log(y_n + 1)}{x_n + y_n} = 0.$$

$\square$

4.2.3 Layer $\mathbb{D}_{\mathbb{HM}}$

Define $\mathbb{D}_{\mathbb{HM}} = (\mathbb{HM} \times \mathbb{D}_d, \mathbb{D}_1)$ as the SFT with alphabet $A_{\mathbb{D}_{\mathbb{HM}}} = A_{\mathbb{HM}} \times A_{\mathbb{D}_d} \times$ $\{Red, Black\}$ and the following rules.

For all $(i, j) \in \mathbb{Z}^2$ if $\mathbb{HM}(i, j) = 0$ then $\mathbb{D}_1(i, j) = Black$.

For all $(i, j) \in \mathbb{Z}^2$ if $\mathbb{HM}(i, j) = 1$ and $\mathbb{D}_d \neq R$ then $\mathbb{D}_1(i, j) = Black$.

For all $(i, j) \in \mathbb{Z}^2$ if $\mathbb{HM}(i, j) = 1$ and $\mathbb{D}_d = R$ then $\mathbb{D}_1(i, j) \in \{Red, Black\}$.

Proposition 4.2.3. $H_S(\mathbb{D}_{\mathbb{HM}}, 1) = \gamma$.

Proof. It is easily checked that $h(\mathbb{D}_{\mathbb{HM}}) = 0$. By Theorem 3.2.9 we need only consider the words arising from independently coloring the 1s. By [5] $\exists O_v(x_n)$ with $\frac{O_v(n)}{n} \to 0$ such that the number of 1s appearing in an x_n length subword of a Toeplitz sequence of frequency x_f is at least $x_n x_f - O_v(x_n)$ and at most $x_n x_f + O_v(x_n)$. There is an $x \in \mathbb{D}_{\mathbb{HM}}$ which contains a distinguished row such that $x_f = \gamma'$.

$$\gamma = \frac{\gamma'}{2} \log 2 = \overline{\lim_{n \to \infty}} \frac{\log 2^{x_n \gamma' - O_v(x_n)}}{x_n + y_n} \leq H_S(\mathbb{D}_{\mathbb{HM}}, 1).$$

To find an upper bound we can partition based on the number of 1s appearing in a row of the (x_n, y_n) window.

$$H_S(\mathbb{D}_{\mathbb{HM}}, 1) \leq \varlimsup_{n \to \infty} \left[\frac{1}{x_n + y_n} \log \sum_{j=0}^{x_n \gamma' + O_v(x_n)} 2^j \right] = \frac{\gamma'}{2} \log 2 = \gamma.$$

□

Theorem 4.2.4. *For any $\gamma \in [0, 1]$ with γ is CFA there is a subshift X such that $H_S(X, 1) = \gamma$.*

Proof. Choose X to be $\mathbb{HM}$ as constructed above. □

We would also like to achieve a CFB surface entropy. At first glance it may seem possible that we could simply double the 0s along a row, seemingly achieving a surface entropy of $1 - \gamma$. There are two important technical difficulties with such a construction.

First note $\sum_{j=0}^{x_n \gamma' + O_v(x_n)} 2^j$, which appears in the calculation of surface entropy in Theorem 4.2.3. This counts all the possible number of 1s that can appear in a point of the SFT. Since the frequency of 1s is bounded above by γ' the largest term in this sum is $2^{x_n \gamma' + O_v(x_n)}$. Unfortunately, since the number of 1s is only bounded from above, there is always a point of all 0s. This point of all 0s means that the corresponding sum in the doubling 0s construction would contain a 2^{x_n} term regardless of γ', and therefore the surface entropy could not depend on γ'. Our construction will alleviate this by coloring all 1s with 2^5 colors and coloring some 0s with many more colors. The words containing the most 1s will dominate the word count so even though there will be points that contain all 0s, the points that contribute to entropy and surface entropy will have the "correct" number of 0s.

The second difficulty arises from the solution to the first difficulty, needing to double all the 1s. In the construction from [4] the 1s and 0s are embedded along and controlled by Toeplitz sequences. A Toeplitz sequence of frequency x_f does not

66

have exactly $x_f n$ 1s in a given n length subword. Instead the maximum number of 1s is controlled by the overage function, $O_v(n)$. When doubling all 1s in a point this leads to a term $2^{nO_v(n)}$ in the word count. When calculating entropy, on the order of n^2, $\frac{O_v(n)}{n} \to 0 \implies \frac{\log 2^{nO_v(n)}}{n^2} \to 0$ and this overage has no effect. When calculating surface entropy, on the order of n, however, since these are dyadic Toeplitz sequences $O_v(n) \approx \log(n)$, $\frac{\log 2^{nO_v(n)}}{n} \to \infty$. Our solution to this issue is to instead use embedded Sturmian sequences which have a constant overage function.

4.3   Realizing non-computable CFB Surface Entropy

For any $x \in (0, 1)$ denote the *ith* digit of the binary expansion of x as $x(i)$. Define a function g such that for each $i \in \mathbb{N}$,

$$g(i, x) = \begin{cases} x(\sqrt{i}) + 1 & \sqrt{i} \in \mathbb{N} \\ \\ 0 & \text{otherwise.} \end{cases}$$

Choose any such x where where x is CFA but not CFB; for the remainder of the paper we consider the base 9 number $\alpha_x \in [0, 1)$ where $\alpha_x(i) = g(i, x)$. We will suppress the dependence on x for readability.

For example if $x = .011001...$ then $\alpha = .1002000020000001000000001000000000002...$

Proposition 4.3.1. α *is a non-computable CFA number.*

Proof. Assume BWOC that α is computable. Then there is a Turing machine, T, which computes α. For fixed x, $g^{-1}(i, x)$ is a computable function and thus $g^{-1} \circ T$ would compute the digits of x. However x was non-computable and thus α is non-computable. Similarly since x is CFA there is a Turing machine S such that $S(n) \searrow \alpha_k$. We note that g is order preserving thus $(g \circ S)(n) \searrow \alpha$ so α is CFA. $\square$

Since α is CFA there is a Turing machine, T_α, such that $T_\alpha(n) \searrow \alpha$. Define $\alpha_n = T_\alpha(n)$.

In the construction of our SFT we will use a Turing machine to prune $\{0, 1\}$ sequences which contain a frequency of 1s greater than α. However, since α is non-computable CFA the Turing machine we use will only have access to the upper approximations α_n. Also, as will be clarified in our construction later, when Turing machines are embedded in SFTs, it turns out that the machine only ever have access to finitely many letters of the aforementioned sequences. In order for sequences with unwanted frequencies to be pruned, but sequences with the desired frequencies to remain, we need bounds on how far the frequencies of these finite portions deviate from α. Since these bounds also need to be generated by Turing machine they must also be computable. Furthermore our Turing machine will only sample a subsequence of the sequence it is pruning, the subsequence being the IP-set (β_j). A priori there is no reason to suspect that sampling a subsequence should control the entire sequence; fortunately since α is in base 9, the IP-set is generated by 9^n, and we are working with 1 balanced sequences we can bound the behavior of the entire sequence with the subsequence. The following propositions provide all needed bounds.

Proposition 4.3.2. *Define* $g(N) = \begin{cases} 0 & N \leq 256 \\ \frac{\lfloor \sqrt{(N)} \rfloor}{13122} & otherwise. \end{cases}$

If $r \in \mathbb{N}$ *such that* $r \leq 9^{\lfloor \sqrt[3]{N} \rfloor}$, $g(N) \leq \sum_{n=1}^{N} (||r(9^n)\alpha||^2 + ||r2(9^n)\alpha||^2 + ||r5(9^n)\alpha||^2)$.

Proof. If $N \leq 256$ the statement is trivially true. Assume $N > 256$. Let $k = \lfloor \sqrt{N} \rfloor$. First consider the case when $r = 1$. If $\alpha(n) \in \{1, 2, 3, 4, 5, 6, 7, 8\}$ and $\alpha(n + 1) = 0$ then $||(9^n)\alpha||^2 > \frac{1}{6561}$. Due to the construction of α, namely that $\alpha(n) \neq 0$ only if n is a perfect square, in the first N digits of α there are at least k locations where

68

this configuration will occur. We would like a similar statement for $r > 1$. First note that since $r \leq 9^{\sqrt[3]{N}}$, the base 9 expansion of r has less than $\sqrt{N}$ digits. We refer to a run of 0s between non-zero digits as a gap. Since $N > 256$, after the first $\frac{k}{2}$ non-zero digits of α its gaps are larger than the number of digits of r. Let $s \in (\frac{k}{2}, k]$ satisfy $\alpha(s^2) \neq 0$. Since $\alpha(s^2) \cdot r$ has at most one more digit than r and the gap between $\alpha(s^2)$ and $\alpha((s+1)^2)$ was greater than the number of digits in r it follows that $(r\alpha)(s^2)$ and $(r\alpha)((s+1)^2)$ have a zero between them. Thus there are at least $\frac{k}{2}$ shifts of $r\alpha$ where $||r(9^n)\alpha||^2 > \frac{1}{6561}$. It follows that $g(N) = \frac{\lfloor \sqrt{(N)} \rfloor}{13122} = \frac{k}{2}\frac{1}{6561} \leq \sum_{n=1}^{N} ||r(9^n)\alpha||^2 \leq \sum_{n=1}^{N} (||r(9^n)\alpha||^2 + ||r2(9^n)\alpha||^2 + ||r5(9^n)\alpha||^2)$. $\qquad \square$

Definition 4.3.3. *Let (β_j) denote the sequence $IP - (9^n\alpha, 2(9^n), 5(9^n)\alpha)$.*

Definition 4.3.4. *Define $\sigma_r(N, \alpha)$ as follows:*

$$\sigma_r(N, \alpha) = \sum_{j < N} e^{2i\pi r \beta_j \alpha}.$$

Proposition 4.3.5. *$\exists \Psi(N)$ computable such that $2\sum_{1 \leq r \leq 9^{\lfloor \sqrt[3]{N} \rfloor}} \left| \frac{\sigma_r(8^N, \alpha)}{8^N} \right| + \frac{4}{\pi\sqrt{9^{\lfloor \sqrt[3]{N} \rfloor}}} \leq \Psi(N)$ and $\Psi(N) \to 0$.*

Proof. Let $g(N)$ be defined as in Proposition 4.3.2 and $r \leq 9^{\lfloor \sqrt[3]{N} \rfloor}$. Then

$$g(N) \leq \sum_{n=1}^{N} \left(||r(9^n)\alpha||^2 + ||r2(9^n)\alpha||^2 + ||r5(9^n)\alpha||^2 \right).$$

Since $0 \leq ||r9^n\alpha||^2 \leq .5$ always holds and $x \leq \sin(\pi x)$ when $x \in [0, .7]$,

$$g(N) \leq \sum_{n=1}^{N} [\sin^2(\pi r(9^n)\alpha) + \sin^2(\pi r2(9^n)\alpha) + \sin^2(\pi r5(9^n)\alpha)].$$

69

$$-g(N) \geq \sum_{n=1}^{N} [\cos^2 (\pi r(9^n)\alpha) - 1 + \cos^2 (\pi r 2(9^n)\alpha) - 1 + \cos^2 (\pi r 5(9^n)\alpha) - 1]$$

$$\geq \sum_{n=1}^{N} \left[\ln \cos^2 (\pi r(9^n)\alpha) + \ln \cos^2 (\pi r 2(9^n)\alpha) + \ln \cos^2 (\pi r 5(9^n)\alpha) \right]$$

$$= \ln \left(\prod_{n=1}^{N} \cos^2 (\pi r(9^n)\alpha) \cos^2 (\pi r 2(9^n)\alpha) \cos^2 (\pi r 5(9^n)\alpha) \right).$$

$$e^{-g(N)} \geq \prod_{n=1}^{N} \left[\cos^2 (\pi r(9^n)\alpha) \cos^2 (\pi r 2(9^n)\alpha) \cos^2 (\pi r 5(9^n)\alpha) \right]$$

$$= \left(\frac{1}{(2^3)^N} \sum_{n=1}^{(2^3)^N} e^{2i\pi r \beta_j \alpha} \right)^2 \tag{*}$$

$$= \left(\frac{1}{8^N} \sigma_r(8^N, \alpha) \right)^2$$

$$\frac{8^N}{\sqrt{e^{g(N)}}} \geq \sigma_r(8^N, \alpha).$$

(*) comes from the equality on page 578 of [1] and noting that the first N generators of (β_j) generate 8^N elements of (β_j).

And thus it follows that

$$2 \sum_{1 \leq r \leq 9\lfloor \sqrt[3]{N} \rfloor} \left| \frac{\sigma_r(8^N, \alpha)}{8^N} \right| + \frac{4}{\pi \sqrt{9 \lfloor \sqrt[3]{N} \rfloor}} \leq 2 \sum_{1 \leq r \leq 9\lfloor \sqrt[3]{N} \rfloor} \left| \frac{8^N}{8^N \sqrt{e^{g(N)}}} \right| + \frac{4}{\pi \sqrt{9 \lfloor \sqrt[3]{N} \rfloor}}$$

$$= 2 \sum_{1 \leq r \leq 9\lfloor \sqrt[3]{N} \rfloor} \left| \frac{1}{\sqrt{e^{g(N)}}} \right| + \frac{4}{\pi \sqrt{9 \lfloor \sqrt[3]{N} \rfloor}} = \Psi(N).$$

$$\lim_{N \to \infty} \Psi(N) = \lim_{N \to \infty} 2 \sum_{1 \leq r \leq 9\lfloor \sqrt[3]{N} \rfloor} \left| \frac{1}{\sqrt{e^{g(N)}}} \right| + \frac{4}{\pi \sqrt{9 \lfloor \sqrt[3]{N} \rfloor}}$$

$$\leq \varlimsup_{N\to\infty} 2 \sum_{1\leq r\leq 9\lfloor \sqrt[3]{N}\rfloor} \left| \frac{1}{e^{\frac{\sqrt{N}}{13122}}} \right| + \frac{4}{\pi\sqrt{9\lfloor \sqrt[3]{N}\rfloor}}$$

$$\leq \varlimsup_{N\to\infty} \left| \frac{2\left(9\lfloor \sqrt[3]{N}\rfloor\right)}{e^{\frac{\sqrt{N}}{13122}}} \right| + \frac{4}{\pi\sqrt{9\lfloor \sqrt[3]{N}\rfloor}} = 0.$$

$\Psi(N)$ is computable by construction and so the conclusion holds. □

Corollary 4.3.6. *For any irrational x_f, $\exists \Psi_{x_f}(N)$ such that*

$$2 \sum_{1\leq r\leq 9\lfloor \sqrt[3]{N}\rfloor} \left| \frac{\sigma_r(8^N, x_f)}{8^N} \right| + \frac{4}{\pi\sqrt{9\lfloor \sqrt[3]{N}\rfloor}} \leq \Psi_{x_f}(N)$$

and $\Psi_{x_f}(N) \to 0$. $\Psi_{x_f}(N)$ may not be computable.

Proof. The proof of Proposition 4.3.5 hinges on the existence (and computability) of $g(N) \to \infty$. In particular, if there is a $g_{x_f}(N)$ such that

$$g_{x_f}(N) \leq \sum_{n=1}^{N} ||r(9^n)x_f||^2 \leq \sum_{n=1}^{N} \left(||r(9^n)x_f||^2 + ||r2(9^n)x_f||^2 + ||r5(9^n)x_f||^2 \right)$$

and $g_{x_f}(N) \to \infty$ then the existence of $\Psi_{x_f}(N)$ would follow from the proof of Proposition 4.3.5. We use an argument similar to that in Proposition 4.3.2 to show such a $g_{x_f}(N)$ exists. For any k, if $(rx_f)(k) \neq (rx_f)(k+1)$ then $||r(9^k)x_f||^2 > \frac{1}{6561}$ and since rx_f is irrational, $(rx_f)(k) \neq (rx_f)(k+1)$ an infinite number of times. Thus $\lim_{N\to\infty} \sum_{n=1}^{N} ||r(9^n)x_f||^2 = \infty$, so $g_{x_f}(N) = \sum_{n=1}^{N} ||r(9^n)x_f||^2$ satisfies the needed criteria and we can construct $\Psi_{x_f}(N)$. □

We begin by briefly summarizing the goal and mechanics of our SFT. For any non-computable CFA number, x, it is the case that $1 - x$ is a non-computable CFB number. Due to technical limitation of the construction we need to embed a CFA

number into the SFT using a Turing machine, we will use α in our construction. We will then produce a surface entropy related to $1 - \alpha$ which will be CFB. Embedding α into the SFT is accomplished by forcing a certain letter of the alphabet to appear with frequency α, in our SFT we will be using the letter 1 for this.

The first layer of the SFT, $\mathbb{P}$, will be a modification to the construction of Pavlov in [11]. It will force each point of the SFT to contain a sequence of 0s and 1s copied along each of its rows. These sequences will be periodic, 1-balanced or 2-balanced.

The second layer of the SFT, $\mathbb{S}$, will use a substitution to place arbitrarily large non-overlapping rectangular "boards" on top of the structure of $\mathbb{P}$. These boards will provide the structure to embed a Turing machine in a later stage. The precise construction of the boards generate the IP-set (β_j) along which the Turing machine will sample the embedded sequences in $\mathbb{P}$. Sampling along this IP-set will allow the Turing machine to calculate the the frequency of 1s occurring in the entire sequence and thus control this frequency. This Turing machine embedding will be very similar to the one used by Hochman and Meyerovitch[4].

The Turing layer, $\mathbb{T}$, will embed a Turing machine into the SFT which will calculate the frequency of 1s in each point of the SFT. Any point of the SFT with a frequency of 1s greater than α will be forbidden.

At this point of the construction the surface entropy of the resulting SFT will still be 0 and all points of the SFT will have 1s appearing with frequency at most α. The final layer of the SFT will add independent colorings to all of the 1s appearing in the SFT and to some of the 0s, these colorings will increase the word count to create the desired amount of surface entropy.

4.3.1 Layer $\mathbb{P}$

Layer $\mathbb{P}$ will have 7 sublayers. The first sublayer is a technical layer, the next 3 will embed sequences of 0s and 1s along the rows of the shift; the last 3 will compare these sequences to each other, forcing the sequences that survived to be either periodic, 1-balanced, or 2-balanced and in fact there will be points containing all such sequences. If any surviving embedded sequence is 2-balanced its counterparts must be periodic or 1-balanced.

The alphabet of $\mathbb{P} = (\mathbb{P}_p, \mathbb{P}_V, \mathbb{P}_R, \mathbb{P}_L, \mathbb{P}_{VR}, \mathbb{P}_{VL}, \mathbb{P}_{RL})$ is $A = \{Purple, Green\} \times \{0,1\}^6$. where $(\mathbb{P}_p, \mathbb{P}_V, \mathbb{P}_R, \mathbb{P}_L, \mathbb{P}_{VR}, \mathbb{P}_{VL}, \mathbb{P}_{RL})$ are the respective projections of $\mathbb{P}$.

Layer $\mathbb{P}_p$ has forbidden list $\mathscr{F} = \{PG, GP, \frac{P}{P}, \frac{G}{G}\}$. This will produce a shift that is just alternating stripes of purple and green rows.

Layer $\mathbb{P}_V$ has forbidden list $\mathscr{F} = \{\frac{0}{1}, \frac{1}{0}\}$. For each point $x \in \mathbb{P}_V(\mathbb{P})$ there will be a bi-infinite sequence of 0s and 1s, s_V, along its rows that is copied vertically.

Layer $\mathbb{P}_R$ is such that $x \in A^{Z^2}$ is forbidden iff $\forall i, j$ if $\mathbb{P}_p(x_{(i,j)}) = Purple$ and $\mathbb{P}_R(x_{(i,j)}) \neq \mathbb{P}_R(x_{(i+1,j+1)})$ or if $\mathbb{P}_p(x_{(i,j)}) = Green$ and $\mathbb{P}_R(x_{(i,j)}) \neq \mathbb{P}_R(x_{(i,j+1)})$. For each point $x \in \mathbb{P}_R(\mathbb{P})$ there will be a bi-infinite sequence of 0s and 1s, s_R, along its rows and each upward vertical shift from a purple row will shift the sequence 1 digits to the right. Upward vertical shifts of a green row will remain constant.

Layer $\mathbb{P}_L$ is such that $x \in A^{Z^2}$ is forbidden iff $\forall i, j$ if $\mathbb{P}_p(x_{(i,j)}) = Green$ and $\mathbb{P}_R(x_{(i,j)}) \neq \mathbb{P}_R(x_{(i-1,j+1)})$ or if $\mathbb{P}_p(x_{(i,j)}) = Purple$ and $\mathbb{P}_R(x_{(i,j)}) \neq \mathbb{P}_R(x_{(i,j+1)})$. For each point $x \in \mathbb{P}_L(\mathbb{P})$ there will be a bi-infinite sequence of 0s and 1s, s_L, along its

rows and each upward vertical shift of a green row will shift the sequence 1 digits to the left. Upward vertical shifts of a purple row will remain constant.

Layer $\mathbb{P}_{VR}$ is such that $x \in A^{Z^2}$ is forbidden iff $\mathbb{P}_{VR}(x_{(i,j)}) \neq \mathbb{P}_{VR}(x_{(i-1,j)}) + \mathbb{P}_V(x_{(i,j)}) - \mathbb{P}_R(x_{(i,j)})$. $\mathbb{P}_{VR}$ is counting the number of digits where s_V and s_R differ using only local information. As proved in [11], $\mathbb{P}_{VR}$ will only have a legal configuration if s_V and s_R are 1-balanced with the same frequency of 1s or if exactly one of s_V and s_R is 2-balanced and the other is periodic with the same frequency of 1s.

Layers $\mathbb{P}_{VL}$ and $\mathbb{P}_{RL}$ will be defined analogously to $\mathbb{P}_{VR}$ with $\mathbb{P}_{VL}(x_{(i,j)}) = \mathbb{P}_{VL}(x_{(i-1,j)}) + \mathbb{P}_V(x_{(i,j)}) - \mathbb{P}_L(x_{(i,j)})$ and $\mathbb{P}_{RL}(x_{(i,j)}) = \mathbb{P}_{RL}(x_{(i-1,j)}) + \mathbb{P}_R(x_{(i,j)}) - \mathbb{P}_L(x_{(i,j)})$.

Each sequence is checked against the other two, with the stipulation that they're both 1-balanced or exactly one is 2-balanced and the other is periodic. The result of all checking will be that all three sequences have the same frequency of 1s and that one of the following cases hold: all three sequences are 1-balanced or exactly one of the three sequences is 2-balanced and the other two are periodic. There is no other restriction on the sequences so any triple of sequences which fulfill the above will appear in some point of $\mathbb{P}$.

4.3.2 Layer $\mathbb{S}$

Layer $\mathbb{S}$ is the substitution layer of the system. $\mathbb{S}$ will be an SFT comprised of a grid like arrangement of increasingly large rectangular boards which will allow the structure to later embed a Turing machine. Our construction is the same as page 15 of [4] with one difference. In their paper the boards are $5^n \times 5^n$ and in our SFT the boards will be $9^n \times 5^n$. To affect this change we replace columns 2 and 4 in each of

74

the substitution rules of Figure 4.1 with 3 columns that are each a copy of the column they are replacing.

First define a substitution shift $\mathbb{S}'$ with alphabet $A = \{|,-,\llcorner,\lrcorner,\ulcorner,\urcorner,\top,\bot,\vdash,\dashv$ $,+,\blacksquare,\square\}$ and substitution rules as described in Figure 4.1 with the changes noted above. $\mathbb{S}'$ fulfills all the assumptions in the hypothesis of Mozes Theorem 4.5[8], which says that there then exists an SFT, $\mathbb{S}$, that is isomorphic a.e. to $\mathbb{S}'$. We will show in the proof of Proposition 4.3.11 that the points where this isomorphism fails won't effect the construction; i.e. for the purposes of our arguments, any reasoning about $\mathbb{S}'$ will hold for $\mathbb{S}$.

Definition 4.3.7. ***Nodes*** *(defined in [4]) are tiles which contain one of the following symbols from $A_{\mathbb{S}}$: $\{\llcorner,\lrcorner,\ulcorner,\urcorner,\top,\bot,\vdash,\dashv,+\}$.*

Nodes are of note in $\mathbb{S}$ because it is only on these symbols that the embedded Turing machine will be able to sample from $\mathbb{P}$ or change its state.

Definition 4.3.8. *Let $R_N[i,j]$ be a 9^N by 5^N rectangular subword of a point in $\mathbb{S}$ whose lower left corner has coordinates (i,j) and contains the symbol $\llcorner$. An **N-board**, $B_N[i,j] \subseteq R_N[i,j]$ is the set of nodes contained in $R_N[i,j]$. (Here and thereafter, informal references to intersections/containments involving words formally refer to the coordinates they occupy in $\mathbb{Z}^2$.) These boards are where the embedded Turing machine will perform calculations as in [4]. $[i,j]$ will be omitted when the location of $B_N[i,j]$ is arbitrary. We define B_0 as either $\square$ or $\blacksquare$. Define $\Omega_S(R_N)$ to be the result of applying one step of the substitution rules to R_N. Define $\Omega_S(B_N[i,j] = B_{N+1}[\Omega_S(i),\Omega_S(j)]$ as the (N+1)-board which is a subset of $\Omega_S(R_N[i,j])$. We note that the construction of the substitution that for any B_{N+1} that there exists a B_N such that $B_{N+1} = \Omega_S(B_N)$.*

Figure 4.1: Substitution rules for $\mathbb{S}$. (Taken from [4].)
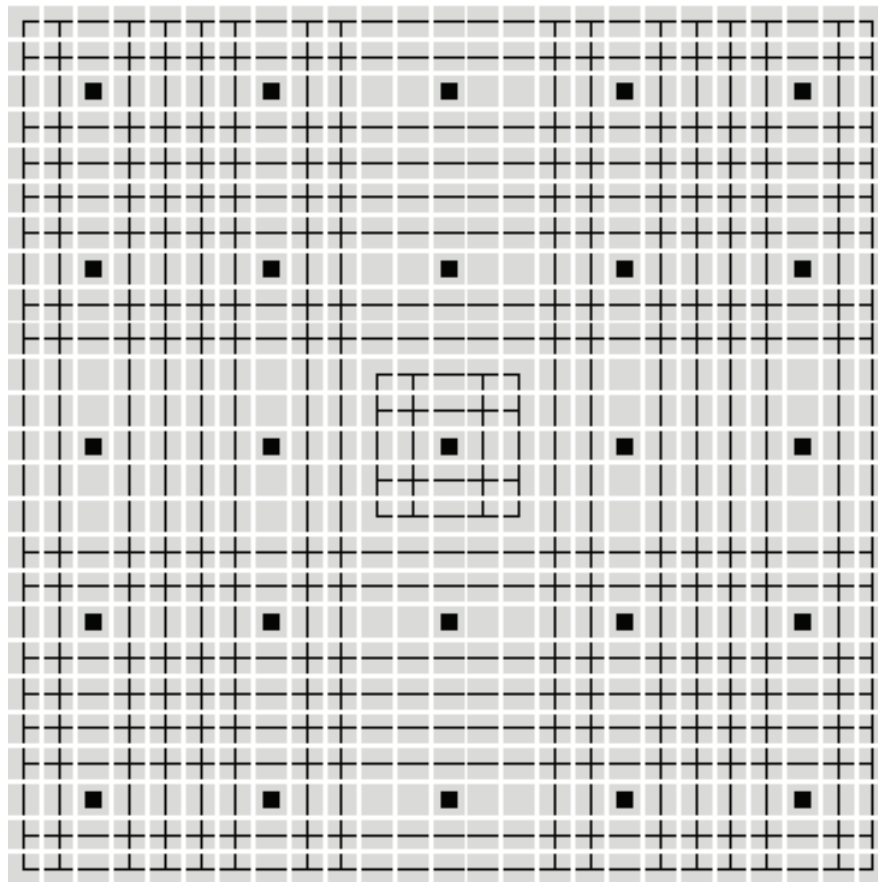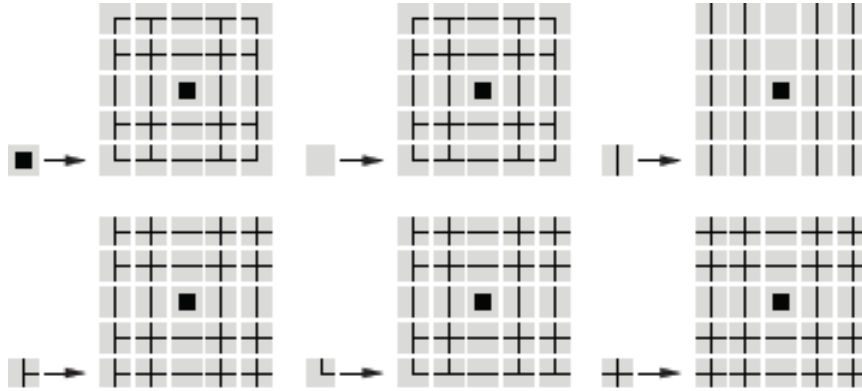
We note from Figure 4.1 that 0-boards may be adjacent horizontally and vertically, but not both. For any N, we can see that the same is true of N-boards by reversing the substitution.

We define 2 distance functions between boards, δ_h and δ_v. $\delta_h(B_n[x,y], B_m[i,j])$ is the minimum horizontal separation between the coordinates of any letter of $B_n[x,y]$ and the coordinates of any letter of $B_m[i,j]$. Likewise define δ_v as the minimum vertical separation.

Example 4.3.9. *Let* $w = \llcorner \bot - \bot$ *be a* 4×1 *subword of some point of* $\mathbb{S}$. *Then* $d_h(\llcorner, -) = 2$ *and* $\delta_h(\llcorner, \bot) = 1$.

Lemma 4.3.10. *For any point* $s \in \mathbb{S}$ *let* $B_{n_i}[a_i, b_i], B_{n_j}[a_j, b_j]$, *and* $B_{n_k}[a_k, b_k]$ *be* N-*boards in* s *such that* $1 \leq n_i < n_j < n_k$. *Then for some pair of these boards, A and B one of the following hold:*

$$\delta_h(A, B) \geq 2(9^{n_i-1}) - \sum_{k=0}^{n_i-2} 8(9^k) = 9^{n_i-1} + 1$$

$$\delta_v(A, B) \geq 2(5^{n_i-1}) - \sum_{k=0}^{n_i-2} 4(5^k) = 5^{n_i-1} + 1.$$

Proof. Let w and v be any two single letter words in s. If $\delta_h(w, v) > 1$, then there are $\delta_h(w, v) - 1$ columns between w and v. After applying the substitution, each of these columns is replaced by a strip 9 letters wide, and so there are $9(\delta_h(w, v) - 1)$ columns strictly between $\Omega_S(w)$ and $\Omega_S(v)$. This means that $\delta_h(\Omega_S(w), \Omega_S(v)) = 9(\delta_h(w, v) - 1) + 1$. Now let W and V be rectangular words in s such that $\delta_h(W, V) = d > 1$. Then $\forall (w, v) \in W \times V, \delta_h(w, v) \geq d$ thus $\delta_h(\Omega_S(w), \Omega_S(v)) \geq 9d - 8$ so $\delta_h(\Omega_S(W), \Omega_S(V)) \geq 9d - 8$. Similarly, $\delta_v(W, V) = d > 1 \implies \delta_v(\Omega_S(W), \Omega_S(V)) = 5d - 4$.

We note that in any configuration of 3 different sized boards if one of the boards is a B_1 then it cannot be the case that all pairs in the configuration are adjacent

both vertically and horizontally, i.e $\delta_v \geq 2$ or $\delta_h \geq 2$. Then $\Omega_S^{-(n_i)}(B_{n_i-1}[a_i, b_i])$, $\Omega_S^{-(n_i-1)}(B_{n_j}[a_j, b_j])$, and $\Omega_S^{-(n_i-1)}(B_{n_k}[a_k, b_k])$ is such a triple and contains a pair that are not adjacent. $\Omega_S^{n_i-1}$ applied to this pair gives the desired separation. Namely, if they are not adjacent horizontally we let $E(d) = 9d - 8$ then

$$\delta_h(\Omega_S^{-(n_i-1)}(A), \Omega_S^{-(n_i-1)}(B)) > 2 \implies \delta_h(A, B) \geq E^{n_i-1}(2) = 2(9^{n_i-1}) - \sum_{k=0}^{n_i-2} 8(9^k).$$

A similar argument holds if they are not adjacent vertically with $E(d) = 5d - 4$. $\quad\square$

Proposition 4.3.11. $H_S(\mathbb{S}, 1) = 0$.

Proof. Let $\{(x_n, y_n)\} \in \Xi(1)$. Substitution shifts are known to have polynomial word counts, and so $|L_{\mathbb{S}'}(n, m)| \leq (nm)^k$ for some $k \in \mathbb{N}$. By Theorem 4.4 of [8] and the remark following, $x \in \mathbb{S}$ if either $x \in \mathbb{S}'$ or x would be in $\mathbb{S}'$ except for at most 1 vertical and 1 horizontal separating line. Therefore,

$$|L_{\mathbb{S}}(n, m)| \leq |L_{\mathbb{S}'}(n, m)| + nm\,|L_{\mathbb{S}'}(n, m)|^4 \leq (nm)^{k+2}$$

Since $h(\mathbb{S}) = 0$ it follows that:

$$H_S(\mathbb{S}, 1) = \sup_{\{(x_n, y_n)\} \in \Xi(1)} \left(\varlimsup_{n \to \infty} \frac{\log |L_{\mathbb{S}}(x_n, y_n)| - x_n y_n h(\mathbb{S})}{x_n + y_n} \right)$$

$$\leq \sup_{\{(x_n, y_n)\} \in \Xi(1)} \left(\varlimsup_{n \to \infty} \frac{(k+2) \log(x_n y_n)}{x_n + y_n} \right) = 0.$$

$\quad\square$

4.3.3 Layer $\mathbb{T}$

Let T be the Turing machine implementing Algorithm 4.3.12 below. By the Church-Turing Thesis since Algorithm 4.3.12 is finite with well defined calculations there exists a Turing machine whose outputs will be the same as the algorithm's; it is this Turing machine we refer to when we say T implements the algorithm. Let $\mathbb{T}_R$ be an SFT superimposed over $\mathbb{P} \times \mathbb{S}$ that implements T the same way that Section 7.2 of [4] implements a Turing machine; the Turing machine takes as input the 0s and 1 of layer $\mathbb{P}_R$ which are superimposed over the nodes of the bottom row of the board in which T is running. T then performs calculations along the nodes of that board. These nodes in $\mathbb{S}$ are the locations in the SFT that correspond to locations of the tape in T, therefor these locations are the only places in the SFT where T change read, write, or update its state. The nodes of $\mathbb{S}$ fall along the IP-set (β_j) and so what T sees as a contiguous word of length N as its input is actually located at N elements of the IP-set. Let $\mathbb{T}_L$ be defined analogously with input $\mathbb{P}_L$. Finally define the SFT $\mathbb{T} = (\mathbb{P}, \mathbb{S}, \mathbb{T}_R, \mathbb{T}_L)$.

Algorithm 4.3.12. *To be implemented by Turing Machine*

Input $\{x_n\} \in \{0,1\}^{\mathbb{N}}$.

Let N=1.

Begin Loop

(AL1) Let a_N be the N-th approximation of α from above.

(AL2) Calculate $\Psi(N)$ as in Proposition 4.3.5.

(AL3) Set $S_{8^N}(x_n)$ to be the number of 1s in the first 8^N digits of $\{x_n\}$.

(AL4) If $\frac{S_{8^N}(x_n)}{8^N} > a_N + \Psi(N)$ halt.

(AL5) Increment N by 1.

End Loop

Notice that for each step of the algorithm the number of calculations required only depends on N; the runtime of the Turing machine that will implement this algorithm will not depend on the input (x_z).

Proposition 4.3.13. *Let $t \in \mathbb{T}$ and let $(x_z)_{z \in \mathbb{Z}}$ be the bi-infinite sequence embedded in the $\mathbb{P}_V$ layer of t and x_f be the frequency of 1s appearing in (x_z), then $x_f \leq \alpha$.*

Proof. The frequency of 1s in any of the sequences embedded in t, is the same for all three embedded sequences of t. It follows that any upper bound on the frequency appearing in either of $\mathbb{P}_L$ or $\mathbb{P}_R$ is automatically an upper bound on the frequency appearing in $\mathbb{P}_V$. At least one of $\mathbb{P}_R$ or $\mathbb{P}_L$ will have a sequence that is aperiodic 1-balanced or periodic, this is the layer that will bound the frequency of 1s in t. WLOG assume (x_z) is the sequence embedded in $\mathbb{P}_R$ and (x_z) is either aperiodic 1-balanced or periodic. Assume BWOC that $x_f > \alpha$.

Case 1: Assume (x_z) is aperiodic 1-balanced. Since (x_z) is 1-balanced it can be generated as follows: $\exists C_N \in [0,1]$ such that $x_z = \chi_{[1-x_f,1]}(zx_f + C_N)$. Define $S_N = \sum_{j=1}^{N} \left[\chi_{[1-x_f,1]}(\beta_j x_f + C_N) \right]$ equivalently, $S_N = \sum_{j=1}^{N} \left[\chi_{[1-x_f-C_N,1-C_N]}(\beta_j x_f) \right]$. Note that by the construction of the SFT, the Turing machine only samples (x_z) along the nodes of $\mathbb{S}$ and so the first N digits that the Turing machine counts are the N digits corresponding to the first N elements of the (β_j) IP-set. We need to show that bounding the frequency of 1s along this IP-set will be sufficient to bound the frequency of 1s in the entire sequence; results related to Weyl's equidistribution theorem from [6] will show this to be the case. S_N is the number of 1s calculated by the Turing machine for some run of N digits of (x_z). By the results on page 34 of [6] letting $u_n = \beta_n x_f$ and $(a, b) = (1 - x_f - C_n, 1 - C_n)$:

$$\forall R > 1, \left| \frac{S_{8^N}}{8^N} - x_f \right| < 2 \sum_{1 \leq r \leq R} \left| \frac{\sigma_r(8^N, x_f)}{8^N} \right| + \frac{4}{\pi \sqrt{R}}.$$

In particular when $R = 9^{\lfloor \sqrt[3]{N} \rfloor}$ by Proposition 4.3.5;

$$\left| \frac{S_{8^N}}{8^N} - x_f \right| < 2 \sum_{1 \le r \le 9^{\lfloor \sqrt[3]{N} \rfloor}} \left| \frac{\sigma_r(8^N, x_f)}{8^N} \right| + \frac{4}{\pi \sqrt{9^{\lfloor \sqrt[3]{N} \rfloor}}} \le \Psi_{x_f}(N).$$

$\Psi_{x_f}(n) \to 0$ (Corollary 4.3.6) so $\frac{S_{8^N}}{8^N} \to x_f$. Since $\alpha_n \to \alpha < x_f$ and $\Psi(n) \to 0$ (Proposition 4.3.5) there must exist n such that $\alpha_n + \Psi(n) < \frac{S_{8^N}}{8^N}$. However this value of n would cause the Turing machine to halt. Since the Turing machine is halted there is no valid infinite configuration of $\mathbb{T}$ and $t \notin \mathbb{T}$.

Case 2: Assume (x_z) is periodic with period P. Let M be such that $x_f > \alpha_M + \Psi(M)$ and let $N > M$. Recall that $\mathbb{P}_R$ embeds (x_z) along each row shifting (x_z) to the right by one digit for each two vertical shifts. Let $B_N(i, j)$ be an N-board in layer $\mathbb{S}(t)$ that has a width of $9^N > P$ and height of $5^N > P$. Then either 9^{N+1} or $9^{N+1} + 5^{N+1}$ is coprime with P. Let m be such that x_m is located at $\mathbb{P}_R(t_{(i,j)})$.

Subcase 1: 9^{N+1} is coprime with P. There is another N-board $B_N(i + 9^{N+1}, j)$ that is 9^{N+1} tiles to the right of $B_N(i, j)$ and $(i + 9^{N+1}, j)$ corresponds to $x_{m+9^{N+1}}$. Since 9^{N+1} is coprime with P, we can iterate this process to find an N-board starting with a digit of (x_z) from each residue class mod P.

Subcase 2: $9^{N+1} + 5^{N+1}$ is coprime with P. Similarly if $9^N + 5^N$ is coprime with P then there is another N-board, $B_N\left(i + 9^{N+1}, j - 2(5^{N+1})\right)$, that is 9^{N+1} tiles to the right and $2(5^{N+1})$ tiles below the original. Since (x_z) is shifted right every other vertical shift, this other N-board starts with $x_{m+9^{N+1}+5^{N+1}}$ and again there is an N-board beginning with a digit of (x_z) from each residue class mod P.

For $0 \leq j \leq P-1$, let T_j be a run of the Turing machine that begins on some x_m where $m \mod P = j$. Let $T_j(k)$ be the kth digit sampled by T_j. Since there is one Turing run for each residue class of P $\forall k, \frac{1}{P} \sum_{j=0}^{P-1} T_j(k) = x_f$. Thus

$$x_f = \frac{1}{8^N} \sum_{k=1}^{8^N} \left[\frac{1}{P} \sum_{j=0}^{P-1} T_j(k) \right] = \frac{1}{8^N} \left[\frac{1}{P} \sum_{j=0}^{P-1} \sum_{k=1}^{8^N} T_j(k) \right]$$

$$8^N x_f = \frac{1}{P} \sum_{j=0}^{P-1} \left[\sum_{k=1}^{8^N} T_j(k) \right].$$

$8^N x_f$ is the average of P elements so at least one of those elements is greater than the average. This implies that

$$\exists i, S_N(x_z) = \sum_{k=1}^{8^N} T_i(k) \geq 8^N x_f.$$

For this particular T_i,

$$\frac{S_{8^N}}{8^N} > \alpha_N + \Psi(N).$$

Then it follows that T_i halts and $t \notin \mathbb{T}$, a contradiction. Therefore our original assumption was wrong and $x_f \leq \alpha$. $\qquad\square$

Corollary 4.3.14. *There is a point $t \in \mathbb{T}$ such that the frequency of 1s appearing in t is α.*

Proof. Let (x_z) be the aperiodic 1-balanced sequence generated by α. By the proof of Proposition 4.3.13 if $x_f = \alpha$ then

$$\forall j \in \mathbb{N}, \alpha < \alpha_j \implies \forall j, n \in \mathbb{N}; \frac{S_{8^n}}{8^n} < \alpha + \Psi(n) < \alpha_j + \Psi(n).$$

Thus the Turing machine does not halt on input of (x_z) and $\exists t \in \mathbb{T}$ such that (x_z) is embedded in the $\mathbb{P}_R$ layer of t and $x_f = \alpha$.

82

$\square$

Proposition 4.3.15. *Let $n \in \mathbb{N}$. For any $w \subseteq B_i$ define $\#(w)$ to be the number of nodes appearing in w. Define $\zeta_i(n) = \max_{0 \leq x \leq 9^i - n} \# \left(B_i|_{[x,x+n-1] \times [0,0]} \right)$. Define $\zeta(n) = \sup_{\{i|9^i > n\}} \zeta_i(n)$. Then $\frac{\zeta(n)}{n} \to 0$.*

Proof. Let $n \in \mathbb{N}$ and $i \in \mathbb{N}$ such that $9^i > n$. Let $w_x = B_i|_{[x,x+n-1] \times [0,0]}$. Let $k_n \in \mathbb{N}$ such that $9^{k_n} \leq n < 9^{k_n+1}$ and define $0 \leq Q_n < 9, 0 \leq R_n < 9^{k_n}$ such that $n = Q_n 9^{k_n} + R_n$. Let $g_x = B_i|_{[x,x+9^{k_n}-1] \times [0,0]}$ then by the substitution rules of $\mathbb{S}$: $\forall x, \#(g_x) \leq 8^{k_n}$. Then w_x is the concatenation of Q_n 9^{k_n}-digit subwords and a subword of length R_n. From this it follows that $\#(w_x) \leq Q_n 8^{k_n} + 8_n^k \leq 9(8^{k_n}) + 8_n^k = 10(8^{k_n})$. Since this bound did not depend on i, $\zeta(n) \leq 10(8^{k_n})$. It follows that $\lim_{n \to \infty} \frac{\zeta(n)}{n} \leq \lim_{n \to \infty} \frac{10(8^{k_n})}{9^{k_n}} = 0$. $\square$

The following corollary is immediate.

Corollary 4.3.16. *If $\zeta(n)$ is defined as in Proposition 4.3.15 then $\frac{\zeta(3n)}{n} \to 0$.*

Proposition 4.3.17. *Let $\{(x_n, y_n)\} \in \Xi(1)$. $\exists N$ such that $\forall n > N \; \forall G \in L_{\mathbb{P} \times \mathbb{S}}(3x_n, 3y_n)$ there are at most $|A_{\mathbb{T}_R}|^{12\zeta(3x_n)}$ words, $w \in \mathbb{T}_R(x_n, y_n)$, such that $(\mathbb{P}, \mathbb{S})(w) = G|_{[x_n, 2x_n - 1] \times [y_n, 2y_n - 1]}$ where $\zeta(x_n)$ is defined as in Proposition 4.3.15.*

Proof. Let G be a $3x_n \times 3y_n$ subword of a point of $\mathbb{P} \times \mathbb{S}$. Let

$$C(G) =$$

$$\{(\mathbb{P}, \mathbb{S}, \mathbb{T}_R)|_{([i+x_n, i+2x_n-1],[j+y_n+2y_n-1])}(t) | t \in \mathbb{T}, (\mathbb{P}, \mathbb{S})|_{([i,i+3x_n-1],[j+3y_n-1])}(t) = G\}$$

and choose any $w \in C(G)$. Then w is the same as the center $x_n \times y_n$ subword of G save for the Turing machine running in the boards of w. Our goal is to count the number of possible different Turing configurations in $C(G)$.

83

The Turing machines running on $\mathbb{T}$ occupy boards B_i of various sizes; we break into various cases dependent on i.

Case 1: Let $i \in \mathbb{N}$ such that $9^i \leq x_n$ and $5^i \leq y_n$; then $B_i \cap w \neq \varnothing \implies B_i \subseteq G$. Since $B_i \subseteq G$ the bottom row of B_i along with the $\{0,1\}$ sequence which B_i samples is entirely contained in G. These two facts mean that the Turing machine running in B_i is completely determined by G.

Case 2: Let $i \in \mathbb{N}$ such that $9^i > x_n$ and $5^i > y_n$.

Sub-case 1: Let $i \in \mathbb{N}$ such that $9^{i-1} \geq x_n$ and $5^{i-1} \geq y_n$. By Lemma 4.3.10 there are at most 2 different board sizes that have non-empty intersection with w, call the boards of these 2 sizes B_X and B_Y. The dimensions of B_X are bigger than w, and we recall that same size boards are only adjacent either horizontally or vertically but not both. Therefore, there are at most 2 different B_X that have non-empty intersection with w. The same holds for B_Y by similar reasoning. Let R_X be the bottom row of B_X which intersects w. Then knowing the state of the Turing machine in all of the nodes of $R_X \cap G$ would completely determine $B_X \cap w$. By Proposition 4.3.15 and Corollary 4.3.16 there are at most $|A_{\mathbb{T}_R}|^{4\zeta(3x_n)}$ possible configurations of these nodes.

Sub-case 2: Let $i \in \mathbb{N}$ such that $9^{i-1} < x_n < 9^i$ or $5^{i-1} < y_n < 5^i$. There is at most 1 such i_x such that this is true for x_n and at most 1 such i_y for y_n. w can have non-empty intersection with at most 4 such B_{i_k} so similar to Sub-case 1 there are at most $|A_{\mathbb{T}_R}|^{8\zeta(3x_n)}$ possible configurations of these nodes.

Thus are at most $|A_{\mathbb{T}_R}|^{12\zeta(3x_n)}$ possible configurations of these boards in w.

Case 3: Let $i \in \mathbb{N}$ such that $9^i \leq x_n$ and $5^i > y_n$. $\exists N$ such that $\forall n > N$ this case is impossible since $\frac{x_n}{y_n} \to 1$.

Case 4: Let $i \in \mathbb{N}$ such that $9^i > x_n$ and $5^i \leq y_n$ and $B_i \cap w \neq \varnothing$. Choose $N \in \mathbb{N}$ such that for all $n > N, x_n > \log_9 10$ so that $i > 10$. A priori the Turing machine performs calculations in any node of an N-board. To complete the proof of Case 4 we first show that this is actually a huge over estimate of where these calculations can occur. An upward vertical shift in $\mathbb{T}$ is equivalent to a single step of the Turing machine and so the Turing head can move at most 1 location along its tape. The locations of the tape correspond to nodes in B_i and so after a single upward vertical shift the Turing head has moved to the right by at most 1 node. We note that while B_i has a width of 9^i that is only contains 8^i columns of $\mathbb{S}$, it is not a contiguous shape. Since B_i contains only 4^i rows the Turing head can only move to the right at most 4^i columns. We wish to show that the horizontal distance from the left edge of B_i to the 4^ith column of B_i is at most 5^i. By an induction using the substitution rules the maximum width of a strip containing 8^k columns of nodes is 9^k. Thus a strip containing $8^{\lceil \frac{2k}{3} \rceil}$ columns of nodes would be no wider than $9^{\lceil \frac{2k}{3} \rceil} \leq 9^{\frac{2k}{3}+2/3} \leq 5^k$ (this last inequality is for any $k > 10$). From this we conclude that a strip containing $4^i = 8^{\frac{2i}{3}} \leq 8^{\lceil \frac{2i}{3} \rceil}$ columns of nodes would have a width of no more than 5^i. This means that as the size of B_i increases that the locations in the board where there are any changes in the Turing layer is a decreasingly small proportion of the left side of the board. We also note that the width of w is wider than the width of the strip where this change can occur.

Thus if the lower left corner of B_i, $B_i(i,j) \notin G$ no actual computation happens inside of w. If $B_i(i,j) \in G$ then the entire computation is also contained within G and it is determined by G.

A bound on the total number of possible Turing configurations in $C(G)$ is the product of the bounds for the number of configurations in each of the 3 cases possible. Since Case 1 and Case 4 are uniquely determined, this essentially means that the bound comes from Case 2.

$\square$

4.3.4 Layer $\mathbb{D}$

We have now constructed an SFT $\mathbb{T}$ in which the frequency of 1s appearing in the $\mathbb{P}_V$ layer is at most α, and where there exists $t \in \mathbb{T}$ achieving frequency α. We now need to add a final layer $\mathbb{D}$ which will add colors independently to all 1s and some 0s in the $\mathbb{P}_V$ layer to create the desired surface entropy.

Let $\mathbb{D}_d$ be the SFT with alphabet $\{B, R, T\}$ and forbidden list

$$\mathscr{F}_d = \{BR, BT, RT, RB, TB, TR, \tfrac{T}{B}, \tfrac{B}{T}, \tfrac{R}{R}, \tfrac{B}{R}, \tfrac{R}{T}\}.$$

$\mathbb{D}_d$ is the SFT consisting of points which are either all Bs, all Ts, or 1 distinguished row of Rs with Bs below and Ts above. This is the same as $\mathbb{D}_d$ defined in Section 4.2.

Fix $\phi \in \mathbb{N}$ such that $\phi > \frac{40}{1-\alpha}$.

Define $\mathbb{D} = (\mathbb{T} \times \mathbb{D}_d, \mathbb{D}_D)$ as the SFT with alphabet $(A_{\mathbb{T}} \times \{B, R, T\}) \times (\{C_j\}_{j=1}^{2^\phi})$ and the following list of forbidden words.

- For all $x \in A^{\mathbb{Z}^2}$, x is forbidden in $\mathbb{D}$ if for any (i, j) $\mathbb{D}_d(x_{(i,j)}) = R$ and it is not the case that $\mathbb{P}_R(x_{(i,j)}) = \mathbb{P}_V(x_{(i,j)}) = \mathbb{P}_L(x_{(i,j)})$.

- For all $x \in A^{\mathbb{Z}^2}$, x is forbidden in $\mathbb{D}$ if $\mathbb{P}_V(x_{(i,j)}) = 1$ and $\mathbb{D}_D(x_{(i,j)}) \notin \{C_j\}_{j=1}^{2^5}$.

- For all $x \in A^{\mathbb{Z}^2}$, x is forbidden in $\mathbb{D}$ if $\mathbb{P}_V(x_{(i,j)}) = 0$ and $\mathbb{D}_d(x_{(i,j)}) \neq R$ and $\mathbb{D}_D(x_{(i,j)}) \neq C_1$.

For the remainder of this section, all references to 0s and 1s refer to those in the $\mathbb{P}_V$ layer only. $\mathbb{D}$ will color all 1s appearing in a point as one of 2^5 colors. It will also possibly have a distinguished row; which may only happen if sublayers $\mathbb{P}_L, \mathbb{P}_R, \mathbb{P}_V$ all agree, which enforces that the sequence embedded in $\mathbb{P}_V$ is 1-balanced. In this case all of the 0s in the distinguished row will be independently colored one of 2^ϕ colors.

Recall that layers $\mathbb{P}_{VR}, \mathbb{P}_{VL}$ and $\mathbb{P}_{RL}$ enforce restrictions on the sequences embedded in their respective sublayers. This enforcement is achieved by keeping a sort of running total of the difference between the number of 1s appearing in each of the two checked layers. As long as the two layers being checked involve different sequences this checksum is uniquely determined by those layers. However, in the case that both layers have identical sequences the checksum can either be all 0s or all 1s. This lack of determination inflates the word counts of such a point; the following lemma gives a bound on how badly this layer can behave in the case where the number of 1s is maximal, which we will see later to be the dominant case for the word count.

Lemma 4.3.18. *Let $\{(x_n, y_n)\} \in \Xi(1)$. Let $L_{\mathbb{D}}^{DR}(x_n, y_n, j) \subseteq L_{\mathbb{D}}(x_n, y_n)$ be the $x_n \times y_n$ subwords, w, which have a distinguished row in $\mathbb{D}_d$ and j 1s in a row of $\mathbb{P}_V(w)$. Let $R_j(x_n, y_n)$ be the maximum number of rows in any such w where $\mathbb{P}_V(w)$ and $\mathbb{P}_R(w)$ agree. Then $\lim_{n \to \infty} \frac{R_{\lceil x_n \alpha \rceil}(x_n, y_n)}{y_n} = 0$.*

Proof. Let $w \in L_{\mathbb{D}}^{DR}(x_n, y_n, \lceil x_n \alpha \rceil)$ and

$$x \in \{x_f : d \in \mathbb{D}, w = d|_{(x_n, y_n)}, x_f \text{ is the frequency of 1s in a row of } \mathbb{P}_V(d)\}.$$

It is clear that $|x_n x - \lceil x_n x \rceil| \leq 1$ and $|x_n \alpha - \lceil x_n \alpha \rceil| \leq 1$. w having a distinguished row implies that any point for which w is a subword has all 1-balanced embedded sequences, thus $|\lceil x_n \alpha \rceil - \lceil x_n x \rceil| \leq 1$. Therefore,

$$|x_n \alpha - x_n x| = |x_n \alpha - \lceil x_n \alpha \rceil + \lceil x_n \alpha \rceil - \lceil x_n x \rceil + \lceil x_n x \rceil - x_n x| \leq 3.$$

Thus $|x - \alpha| \leq \frac{3}{x_n}$. Let $q \in \mathbb{N}$ and assume BWOC that for infinitely many n, $R_{\lceil x_n \alpha \rceil}(x_n, y_n) > \frac{y_n}{2q}$. Assume WLOG $x_n > q$. This means that there are two rows in w whose vertical distance apart is more than 1, but less than $2q$ on which $\mathbb{P}_R$ and $\mathbb{P}_V$ agree. $\mathbb{P}_R$ shifts right every second vertical shift so it has shifted right at most q times; we treat only the case of exactly q shifts here, as others are similar. Choose any $d \in \mathbb{D}$ containing w. Since w contained a distinguished row this means that the sequence embedded in $\mathbb{P}_V(d)$ is the same as the one embedded in $\mathbb{P}_R(d)$. The sequence embedded in $\mathbb{P}_V(d)$ contains an $x_n + q$ letter word which contains two copies of the same x_n letter word that start at different indexes that are q shifts apart. This $x_n + q$ letter subword contains a subword, v, of length q such that vv is a subword of the sequence embedded in $\mathbb{P}_V(d)$ so vv is 1-balanced. The $x_n + q$ letter word is a subword of a 1-balanced periodic sequence with period q (the sequence $..vvvv..$); $\exists 1 \leq p \leq q$ such that the frequency of 1s in this periodic sequence is $\frac{p}{q}$. By the same argument as above $\left| x - \frac{p}{q} \right| \leq \frac{3}{x_n}$. α is irrational thus $\alpha \neq \frac{p}{q}$ and $\exists N_q$ such that $\forall n \geq N_q \; \frac{3}{x_n} \leq \left| \alpha - \frac{p}{q} \right|$. Thus $\forall n > N_q \; R_{\lceil x_n \alpha \rceil}(x_n, y_n) \leq \frac{y_n}{2q}$. Since q was arbitrary, $\frac{R_{\lceil x_n \alpha \rceil}(x_n, y_n)}{y_n}$ approaches 0.

$\square$

Proposition 4.3.19. $H_S(\mathbb{D}, 1) = \frac{(1-\alpha)}{2}\log(2^\phi)$.

Proof. Let $\{(x_n, y_n)\} \in \Xi(1)$. Choose $N \in \mathbb{N}$ such that $\forall n > N, \frac{y_n}{2} \leq x_n \leq 2y_n$, $x_n - \lceil x_n\alpha \rceil > \frac{x_n - x_n\alpha}{2}$, and $y_n > \phi$. To find a lower bound for $L_\mathbb{D}(x_n, y_n)$ recall that from Corollary 4.3.14 $\exists t \in \mathbb{T}$ such that the frequency of 1s appearing in t is α; thus $\exists d \in \mathbb{D}$ where the frequency of 1s appearing is α and it has a distinguished row. The number of 1s that can appear in a subword of length x_n in a 1-balanced sequence of frequency α is $\lceil x_n\alpha \rceil$. For each 1 there is a choice of 2^5 colors and for each 0 there is a choice of 2^ϕ colors so:

$$2^{5y_n\lceil x_n\alpha\rceil}(2^\phi)^{x_n - \lceil x_n\alpha\rceil} \leq |L_\mathbb{D}(x_n, y_n)|.$$

We claim that the following is an upper bound for $L_\mathbb{D}(x_n, y_n)$;

$$|L_\mathbb{D}(x_n, y_n)| \leq$$

$$|L_\mathbb{S}(x_n, y_n)|\,|L_{\mathbb{T}_R}(x_n, y_n)|\,|L_{\mathbb{T}_L}(x_n, y_n)|$$

$$\left[\sum_{j=0}^{\lceil x_n\alpha\rceil}(x_n+1)(y_n+1)2^{5y_nj}2^{\phi(x_n-j)}2^{3R_j(x_n,y_n)} + \sum_{j=0}^{\lceil x_n\alpha\rceil+1}2^{5y_nj}2^{3y_n}2^{3x_n}\right].$$

This sum came from partitioning $\mathbb{D}$ into points containing a distinguished row and points that do not.

Case 1: For $d \in \mathbb{D}$ such that d contains a distinguished row and x_f is the frequency of 1s in $\mathbb{P}_V(d)$ which is constant vertically; consider $(\mathbb{P}, \mathbb{D})(d)$. Recall that containing a distinguished row forces the embedded sequence to be 1-balanced. There are $y_n + 1$ choices for which row, if any, is distinguished and $x_n + 1$ words of length x_n in a 1-balanced sequence. Let j be the number of 1s appearing in a row of $\mathbb{P}_V(d)$ in the (x_n, y_n) window. Then $0 \leq j \leq \lceil x_n\alpha \rceil$. If j is the number of 1s appearing then $x_n - j$

will be the number of 0s appearing in the same row. Each 1 has 2^5 choices of color and there are $y_n j$ 1s in the (x_n, y_n) window. Each 0 has 2^ϕ choices of color. For each j there are at most $2^{R_j(x_n, y_n)}$ rows in each of $\mathbb{P}_{VR}$, $\mathbb{P}_{VL}$ and $\mathbb{P}_{RL}$ where the checksum is not determined and has a choice of 0 or 1(see Lemma 4.3.18 and preceding discussion. Note that technically as defined $R_j(x_n, y_n)$ applies to the checksum in $\mathbb{P}_{VR}$, analogous results hold for the other 2 checksum layers.) Giving an upper bound on this word count of:

$$\sum_{j=0}^{\lceil x_n \alpha \rceil} y_n 2^{5y_n j} 2^{\phi(x_n - j)} 2^{3R_j(x_n, y_n)}.$$

Case 2: Consider $d \in \mathbb{D}$ such that d does not contain a distinguished row and x_f is the frequency of 1s in $\mathbb{P}_V(d)$. Since there is no distinguished row the embedded sequence need only be 2-balanced, and for lack of a better bound we only use that there are at most 2^{x_n} words of length x_n in such a sequence. Again consider $(\mathbb{P}, \mathbb{D})(d)$; for each possible number of 1s in a row, $0 \le j \le \lceil x_n \alpha \rceil + 1$, there are 2^5 choices to color those 1s. No 0s will be colored in a point without a distinguished row. Let $R_j(x_n, y_n)$ be as defined in Lemma 4.3.18 then $R_j(x_n, y_n) \le y_n$. Giving an upper bound on this word count of:

$$\sum_{j=0}^{\lceil x_n \alpha \rceil + 1} 2^{5y_n j} 2^{3y_n} 2^{3x_n}.$$

Adding the bounds from our 2 disjoint cases and simplifying yields the following:

$$\sum_{j=0}^{\lceil x_n \alpha \rceil} (x_n + 1)(y_n + 1) 2^{5y_n j} 2^{\phi(x_n - j)} 2^{3R_j(x_n, y_n)} + \sum_{j=0}^{\lceil x_n \alpha \rceil + 1} 2^{5y_n j} 2^{3y_n} 2^{3x_n}$$

$$= \left(2^{5y_n\lceil x_n\alpha\rceil}2^{\phi(x_n-\lceil x_n\alpha\rceil)}2^{3R_{\lceil x_n\alpha\rceil}(x_n,y_n)}\right)$$

$$\left[\sum_{j=0}^{\lceil x_n\alpha\rceil}\frac{(x_n+1)(y_n+1)2^{5y_nj}2^{\phi(x_n-j)}2^{3R_j(x_n,y_n)}}{\left(2^{5y_n\lceil x_n\alpha\rceil}2^{\phi(x_n-\lceil x_n\alpha\rceil)}2^{3R_{\lceil x_n\alpha\rceil}(x_n,y_n)}\right)}\right.$$

$$\left.+\sum_{j=0}^{\lceil x_n\alpha\rceil+1}\frac{2^{5y_nj}2^{3y_n}2^{3x_n}}{\left(2^{5y_n\lceil x_n\alpha\rceil}2^{\phi(x_n-\lceil x_n\alpha\rceil)}2^{3R_{\lceil x_n\alpha\rceil}(x_n,y_n)}\right)}\right]$$

$$= \left(2^{5y_n\lceil x_n\alpha\rceil}2^{\phi(x_n-\lceil x_n\alpha\rceil)}2^{3R_{\lceil x_n\alpha\rceil}(x_n,y_n)}\right)$$

$$\left[(x_n+1)(y_n+1)\sum_{j=0}^{\lceil x_n\alpha\rceil}\left(\frac{2^{3(R_j(x_n,y_n)-R_{\lceil x_n\alpha\rceil}(x_n,y_n))}}{2^{(5y_n-\phi)(\lceil x_n\alpha\rceil-j)}}\right)\right.$$

$$\left.+\sum_{j=0}^{\lceil x_n\alpha\rceil+1}\frac{2^{5y_nj}2^{3y_n}2^{3x_n}}{\left(2^{5y_n\lceil x_n\alpha\rceil}2^{\phi(x_n-\lceil x_n\alpha\rceil)}2^{3R_{\lceil x_n\alpha\rceil}(x_n,y_n)}\right)}\right]$$

$$\leq \left(2^{5y_n\lceil x_n\alpha\rceil}2^{\phi(x_n-\lceil x_n\alpha\rceil)}2^{3R_{\lceil x_n\alpha\rceil}(x_n,y_n)}\right)$$

$$\left[(x_n+1)(y_n+1)\left(1+\sum_{j=0}^{\lceil x_n\alpha\rceil-1}\left(\frac{2^{3y_n}}{2^{(5y_n-\phi)(\lceil x_n\alpha\rceil-j)}}\right)\right)\right.$$

$$\left.+\left(\frac{2^{5y_n(\lceil x_n\alpha\rceil+1)}2^{9x_n}}{2^{5y_n\lceil x_n\alpha\rceil}2^{\phi(x_n-\lceil x_n\alpha\rceil)}}+\sum_{j=0}^{\lceil x_n\alpha\rceil}\frac{2^{9x_n}2^{5y_nj}}{2^{5y_n\lceil x_n\alpha\rceil}2^{\phi(x_n-\lceil x_n\alpha\rceil)}}\right)\right]$$

$$\leq \left(2^{5y_n\lceil x_n\alpha\rceil}2^{\phi(x_n-\lceil x_n\alpha\rceil)}2^{3R_{\lceil x_n\alpha\rceil}(x_n,y_n)}\right)$$

$$\left[(x_n+1)(y_n+1)\left(1+\sum_{j=0}^{\lceil x_n\alpha\rceil-1}\left(\frac{2^{3y_n}}{2^{(5y_n-\phi)(\lceil x_n\alpha\rceil-j)}}\right)\right)\right.$$

$$\left.+\left(\frac{2^{5y_n(\lceil x_n\alpha\rceil+1)}2^{9x_n}}{2^{5y_n\lceil x_n\alpha\rceil}2^{\phi(x_n-\lceil x_n\alpha\rceil)}}+\frac{2^{9x_n}}{2^{\phi(x_n-\lceil x_n\alpha\rceil)}}\sum_{j=0}^{\lceil x_n\alpha\rceil}\frac{1}{2^{5y_n(\lceil x_n\alpha\rceil-j)}}\right)\right]$$

$$\leq \left(2^{5y_n\lceil x_n\alpha\rceil}2^{\phi(x_n-\lceil x_n\alpha\rceil)}2^{3R_{\lceil x_n\alpha\rceil}(x_n,y_n)}\right)$$

$$\left[(x_n+1)(y_n+1)\left(1+\sum_{j=0}^{\lceil x_n\alpha\rceil-1}\left(\frac{2^{3y_n}}{2^{(5y_n-\phi)(\lceil x_n\alpha\rceil-j)}}\right)\right)\right.$$

$$\left.+\left(\frac{2^{5y_n}2^{9x_n}}{2^{\phi(x_n-\lceil x_n\alpha\rceil)}}+\frac{2^{9x_n}}{2^{\phi(x_n-\lceil x_n\alpha\rceil)}}\sum_{j=0}^{\lceil x_n\alpha\rceil}\frac{1}{2^{5y_n(\lceil x_n\alpha\rceil-j)}}\right)\right].$$

Since $x_n-\lceil x_n\alpha\rceil > \frac{x_n-x_n\alpha}{2}$; so the above is

$$\leq \left(2^{5y_n\lceil x_n\alpha\rceil}2^{\phi(x_n-\lceil x_n\alpha\rceil)}2^{3R_{\lceil x_n\alpha\rceil}(x_n,y_n)}\right)$$

$$\left[(x_n+1)(y_n+1)\left(1+\sum_{j=0}^{\lceil x_n\alpha\rceil-1}\left(\frac{2^{3y_n}}{2^{(5y_n-\phi)(\lceil x_n\alpha\rceil-j)}}\right)\right)\right.$$

$$\left.+\left(\frac{2^{19x_n}}{2^{x_n\phi\frac{(1-\alpha)}{2}}}+\frac{2^{9x_n}}{2^{x_n\phi\frac{(1-\alpha)}{2}}}\sum_{j=0}^{\lceil x_n\alpha\rceil}\frac{1}{2^{5y_n(\lceil x_n\alpha\rceil-j)}}\right)\right]$$

Recall that $\phi > \frac{40}{1-\alpha}$, so the above is

$$\leq \left(2^{5y_n\lceil x_n\alpha\rceil}2^{\phi(x_n-\lceil x_n\alpha\rceil)}2^{3R_{\lceil x_n\alpha\rceil}(x_n,y_n)}\right)$$

$$\left[(x_n+1)(y_n+1)\left(1+\sum_{j=0}^{\lceil x_n\alpha\rceil-1}\left(\frac{2^{3y_n}}{2^{(5y_n-\phi)(\lceil x_n\alpha\rceil-j)}}\right)\right)\right.$$

$$\left.+\left(\frac{2^{19x_n}}{2^{20x_n}}+\frac{2^{9x_n}}{2^{20x_n}}\left(1+\sum_{j=0}^{\lceil x_n\alpha\rceil-1}\frac{1}{2^{5y_n(\lceil x_n\alpha\rceil-j)}}\right)\right)\right]$$

$(\lceil x_n\alpha\rceil-j)>0$ since $\lceil x_n\alpha\rceil-1$ is the upper bound of the summation over j. $5y_n-\phi>0$ since ϕ is fixed with $y_n>\phi$ so $\exists U\in\mathbb{R}$ such that

$$\forall n>N,\ \sum_{j=0}^{\lceil x_n\alpha\rceil}\left(\frac{2^{3y_n}}{2^{(5y_n-\phi)(\lceil x_n\alpha\rceil-j)}}\right)<U$$

and

$$\forall n>N, 1+\sum_{j=0}^{\lceil x_n\alpha\rceil-1}\frac{1}{2^{5y_n(\lceil x_n\alpha\rceil-j)}}<U.$$

Therefor the quantity above is less than or equal to

$$\left(2^{5y_n\lceil x_n\alpha\rceil}2^{\phi(x_n-\lceil x_n\alpha\rceil)}2^{3R_{\lceil x_n\alpha\rceil}(x_n,y_n)}\right)\left[(x_n+1)(y_n+1)(1+U)+\frac{1}{2^{x_n}}+\frac{U}{2^{11x_n}}\right].$$

$\exists k\in\mathbb{N}$ such that Proposition 4.3.11 gives an upper bound on $|L_\mathbb{S}(x_n,y_n)|\leq(x_ny_n)^{k+2}$ and once there is a count on $|L_\mathbb{S}(x_n,y_n)|$ and $|L_\mathbb{P}(x_n,y_n)|$ Proposition 4.3.17 gives an upper bound on $|L_{\mathbb{T}_R}(x_n,y_n)|$ and $|L_{\mathbb{T}_L}(x_n,y_n)|$.

$$2^{5y_n\lceil x_n\alpha\rceil}(2^\phi)^{x_n-\lceil x_n\alpha\rceil}\leq|L_\mathbb{D}(x_n,y_n)|$$

$$\leq(x_ny_n)^{k+2}A_{\mathbb{T}_R}^{2\zeta(3x_n)}\left(2^{5y_n\lceil x_n\alpha\rceil}2^{\phi(x_n-\lceil x_n\alpha\rceil)}2^{3R_{\lceil x_n\alpha\rceil}(x_n,y_n)}\right)$$

$$\left[(x_n+1)(y_n+1)(1+U)+\frac{1}{2^{x_n}}+\frac{U}{2^{11x_n}}\right].$$

Recalling from Lemma 4.3.18 that $\frac{R_{\lceil x_n \alpha \rceil}(x_n, y_n)}{y_n} \to 0$ and from Corollary 4.3.16 that $\frac{\zeta(3n)}{n} \to 0$ we can calculate $h(\mathbb{D})$ as follows:

$$\lim_{n \to \infty} \frac{1}{x_n y_n} \log \left(2^{5 y_n \lceil x_n \alpha \rceil} (2^\phi)^{x_n - \lceil x_n \alpha \rceil} \right) \leq \lim_{n \to \infty} \frac{1}{x_n y_n} \log \left(|L_{\mathbb{D}}(x_n, y_n)| \right)$$

$$\leq \lim_{n \to \infty} \frac{1}{x_n y_n} \log \left((x_n y_n)^{k+2} A_{\mathbb{T}_R}^{2\zeta(3 x_n)} \left(2^{5 y_n \lceil x_n \alpha \rceil} 2^{\phi(x_n - \lceil x_n \alpha \rceil)} 2^{3 R_{\lceil x_n \alpha \rceil}(x_n, y_n)} \right) \right.$$

$$\left. \left[(x_n + 1)(y_n + 1)(1 + U) + \frac{1}{2^{x_n}} + \frac{U}{2^{11 x n}} \right] \right).$$

$$5\alpha \log(2) \leq h(\mathbb{D})$$

$$\leq \lim_{n \to \infty} \frac{1}{x_n y_n} \left((k+2) \log(x_n y_n) + 2\zeta(3 x_n) \log(A_{\mathbb{T}_R}) \right.$$

$$5 y_n \lceil x_n \alpha \rceil \log(2) + \phi(x_n - \lceil x_n \alpha \rceil) \log(2) + 3 R_{\lceil x_n \alpha \rceil}(x_n, y_n) \log(2)$$

$$\left. + \log \left[(x_n + 1)(y_n + 1)(1 + U) + \frac{1}{2^{x_n}} + \frac{U}{2^{11 x n}} \right] \right) = 5\alpha \log(2).$$

Therefore $h(\mathbb{D}) = 5\alpha \log 2$. Similarly we can calculate $H_S(\mathbb{D}, 1)$.

$$\lim_{n \to \infty} \frac{\log \left(2^{5 y_n \lceil x_n \alpha \rceil} (2^\phi)^{x_n - \lceil x_n \alpha \rceil} \right) - x_n y_n 5\alpha \log(2)}{x_n + y_n} = \frac{(1 - \alpha)}{2} \log(2^\phi) \leq H_S(\mathbb{D}, 1).$$

$\overline{\lim}_{n\to\infty} S_X(x_n, y_n) \le$

$$\lim_{n\to\infty} \frac{1}{x_n + y_n} \left[\log\left((x_n y_n)^{k+2} A_{\mathbb{T}_R}^{2\zeta(3x_n)} \left(2^{5y_n \lceil x_n \alpha \rceil} 2^{\phi(x_n - \lceil x_n \alpha \rceil)} 2^{3R_{\lceil x_n \alpha \rceil}(x_n, y_n)} \right) \right. \right.$$

$$\left. \left. \left[(x_n + 1)(y_n + 1)(1 + U) + \frac{1}{2^{x_n}} + \frac{U}{2^{11xn}} \right] \right) - x_n y_n 5\alpha \log(2) \right]$$

$$= \lim_{n\to\infty} \frac{\log\left(2^{5y_n \lceil x_n \alpha \rceil} 2^{\phi(x_n - \lceil x_n \alpha \rceil)} 2^{3R_{\lceil x_n \alpha \rceil}(x_n, y_n)} \right) - x_n y_n 5\alpha \log(2)}{x_n + y_n} = \frac{(1-\alpha)}{2} \log(2^{\phi}).$$

As $\{(x_n, y_n)\}$ was an arbitrary sequence in $\Xi(\alpha)$ it follows that as the sup over all such $\overline{\lim}_{n\to\infty} S_X(x_n, y_n)$ that $H_S(\mathbb{D}, 1) \le \frac{(1-\alpha)}{2} \log(2^{\phi})$. Thus $H_S(\mathbb{D}, 1) = \frac{(1-\alpha)}{2} \log(2^{\phi})$.

$\square$

Theorem 4.3.20. *There exists a subshift X such that $H_S(X, 1)$ is CFB and not CFA.*

Proof. Choose X to be $\mathbb{D}$ as defined above. Then $H_S(X, 1) = \frac{(1-\alpha)}{2} \log(2^{\phi})$.

Since α was not computable, $(1 - \alpha)$ is also not computable. We now show that $\frac{(1-\alpha)}{2} \log(2^{\phi})$ is CFB. Recall that since α was CFA that there is a Turing machine, T such that $T(n) \searrow \alpha$. Thus there is a Turing machine T' where for each n, $T'(n) = \frac{1-T(n)}{2} \log(2^{\phi})$ and $\frac{1-T(n)}{2} \log(2^{\phi}) \nearrow \frac{(1-\alpha)}{2} \log(2^{\phi})$. And thus $\frac{(1-\alpha)}{2} \log(2^{\phi})$ is CFB but not computable so it is not CFA.

$\square$

Our results show that we can realize any CFB number for which a computable $g(N)$ as in Proposition 4.3.2 exists; however this is not equivalent to the class of all CFB numbers. It remains an open question whether all CFB numbers can be realized as $H_S(X, 1)$ for some 2-D SFT. Furthermore a full classification of surface entropies of SFTs has yet to be found, however we now discuss without formal proof an upper bound in the arithmetical hierarchy on the set $\{H_S(X, 1) | X \text{ is an SFT}\}$.

The arithmetical hierarchy, $\mathscr{A}$, classifies the complexity of defining subsets of $\mathbb{N}$ using formal first-order logic and is closely related to determining the computability

properties of such subsets. A subset of $\mathbb{N}$ which satisfy a logical formula which uses no existential or universal quantifier (bounded quantification is allowed) is in the class Π_0^0 and Σ_0^0. The hierarchy is then built inductively by quantifying formula from the previous class; if ψ is Π_k^0 then $\forall n, \psi$ is Σ_{k+1}^0 and if ψ is Σ_k^0 then $\exists n, \psi$ is Π_{k+1}^0. A set which is both Π_k^0 and Σ_k^0 is Δ_k^0.

We can use the binary expansion of an element of $\mathbb{R}$ as the characteristic function of a subset of $\mathbb{N}$ and this natural correspondence between $2^{\mathbb{N}}$ and $\mathbb{R}$ allows us to think of elements of $\mathbb{R}$ as being in $\mathscr{A}$. It is a known result that computable numbers are Δ_1^0, CFA numbers are Π_1^0, and CFB numbers are Σ_1^0. We denote $\overline{\lim}, \underline{\lim}, \sup, \inf$ as limit operations. Since $\mathbb{Q}$ is dense and computable, in general a limit operation of computable numbers results in all of $\mathbb{R}$. We can informally denote a computable limit operation, $f(n) \to x$, as one where there is a Turing machine T such that $\forall n, T(n) = f(n)$; these sorts of computable limit operations generally increment rank in $\mathscr{A}$. Through personal communication with Pascal Vanier it seems that the limit operations in the calculation of surface entropy are of these second type and we apply two such limit operations to the set of entropies. Since entropies are CFA which are Π_1^0 it appears that surface entropies must be at most Π_3^0.

Bibliography

[1] Daniel Berend. "IP-sets on the circle". *Canad. J. Math.* 42.4 (1990), pp. 575–589. ISSN: 0008-414X. URL: https://doi.org/10.4153/CJM-1990-030-8.

[2] Robert Berger. "The undecidability of the domino problem". *Mem. Amer. Math. Soc. No.* 66 (1966), p. 72. ISSN: 0065-9266.

[3] N. Pytheas Fogg. *Substitutions in dynamics, arithmetics and combinatorics.* Vol. 1794. Lecture Notes in Mathematics. Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel. Springer-Verlag, Berlin, 2002, pp. xviii+402. ISBN: 3-540-44141-7. DOI: 10.1007/b13861. URL: https://doi-org.du.idm.oclc.org/10.1007/b13861.

[4] Michael Hochman and Tom Meyerovitch. "A characterization of the entropies of multidimensional shifts of finite type". *Ann. of Math. (2)* 171.3 (2010), pp. 2011–2038. ISSN: 0003-486X. URL: https://doi.org/10.4007/annals.2010.171.2011.

[5] Konrad Jacobs and Michael Keane. "$0 - 1$-sequences of Toeplitz type". *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 13 (1969), pp. 123–131. DOI: 10.1007/BF00537017. URL: https://doi-org.du.idm.oclc.org/10.1007/BF00537017.

[6]   Aditi Kar. "Weyl's equidistribution theorem". *Resonance* 8.5 (May 2003), pp. 30–37. ISSN: 0973-712X. DOI: `10.1007/BF02867127`. URL: `https://doi.org/10.1007/BF02867127`.

[7]   Douglas Lind and Brian Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, Cambridge, 1995, pp. xvi+495. ISBN: 0-521-55124-2; 0-521-55900-6. DOI: `10.1017/CBO9780511626302`. URL: `https://doi-org.du.idm.oclc.org/10.1017/CBO9780511626302`.

[8]   Shahar Mozes. "Tilings, substitution systems and dynamical systems generated by them". *J. Analyse Math.* 53 (1989), pp. 139–186. ISSN: 0021-7670. URL: `https://doi.org/10.1007/BF02793412`.

[9]   Donald S. Ornstein and Benjamin Weiss. "Entropy and isomorphism theorems for actions of amenable groups". *J. Analyse Math.* 48 (1987), pp. 1–141. ISSN: 0021-7670. DOI: `10.1007/BF02790325`. URL: `https://doi.org/10.1007/BF02790325`.

[10] Ronnie Pavlov. "Approximating the hard square entropy constant with probabilistic methods". *Ann. Probab.* 40.6 (2012), pp. 2362–2399. ISSN: 0091-1798. DOI: `10.1214/11-AOP681`. URL: `https://doi-org.du.idm.oclc.org/10.1214/11-AOP681`.

[11] Ronnie Pavlov. "Topologically completely positive entropy and zero-dimensional topologically completely positive entropy". *Ergodic Theory Dynam. Systems* 38.5 (2018), pp. 1894–1922. ISSN: 0143-3857. DOI: `10.1017/etds.2016.120`. URL: `https://doi-org.du.idm.oclc.org/10.1017/etds.2016.120`.

[12] Stefan Sidler and Roman Lickel. *Singletaple Turing Machine*. 2015. URL: `http://turingmaschine.klickagent.ch/einband/?&lang=en#10_*_10` (visited on 06/12/2018).