ENTROPY GAPS AND LOCALLY MAXIMAL ENTROPY IN $\mathbb{Z}^d$ SUBSHIFTS

ANTHONY QUAS AND AYŞE A. ŞAHİN

ABSTRACT. In this paper we study the behaviour of the entropy function of higher dimensional shifts of finite type. We construct a topologically mixing $\mathbb{Z}^2$ shift of finite type whose ergodic invariant measures are connected in the $\overline{d}$ topology and whose entropy function has a strictly local maximum. We also construct a topologically mixing $\mathbb{Z}^2$ shift of finite type X with the property that there is a uniform gap between the topological entropy of X and the topological entropy of any subshift of X with stronger mixing properties. Our examples illustrate the necessity of strong topological mixing hypotheses in existing higher dimensional representation and embedding theorems.

1. INTRODUCTION

It is well known that higher dimensional symbolic dynamical systems are much more complicated than their one-dimensional counterparts. Therefore it is of interest to understand which particular families of higher dimensional systems have behaviour and properties similar to the one-dimensional subshifts. Here we focus on one specific property that all one dimensional topologically mixing shifts of finite type (SFT) have: that of being a *universal model* in both the measure theoretic and topological categories.

We say that a $\mathbb{Z}^d$ SFT Y is a universal model in the measure theoretic category if given any ergodic non-atomic $\mathbb{Z}^d$ action (X, μ, T) whose measure theoretic entropy is strictly less than the topological entropy of Y, there is a shift invariant measure ν on Y such that (Y, ν, S) is measurably isomorphic to (X, μ, T). In the topological category we say Y is a universal model if given any SFT X whose topological entropy is strictly less than that of Y and whose periodic points can be embedded into the periodic points of Y is a factor of Y.

For $d = 1$ Krieger has shown that topologically mixing SFT are universal models in both categories [5],[6]. In higher dimensions [10] Robinson and Şahin show that if Y satisfies a strong topological mixing property called the *uniform filling property* (UFP), then it is a universal model in the measure theoretic category. In [7] Lightwood has shown that for $d = 2$ a SFT Y is a topological universal model if it is topologically mixing and satisfies an extension property called square filling. Together these properties imply that the SFT has the UFP.

The question we address in this paper is the following: how far can you weaken the mixing property of the SFT and still get a universal model? In particular, for the topological category, one can ask the following question (originally posed to us by M. Boyle). Let $h(Y)$ denote the topological entropy of a shift space Y.

Suppose Y is a topologically mixing SFT. Given $\epsilon > 0$, does there exist a sub-shift of finite type $Y' \subset Y$ with the UFP such that $h(Y') > h(Y) - \epsilon$?

We show that the answer is *no* and that topological mixing alone is not sufficient for a higher dimensional SFT to be a topological universal model. Let $\mathcal{M}_e(Y, S)$ denote the ergodic, shift invariant Borel probability measures on Y and for $\nu \in \mathcal{M}_e(Y, S)$ denote the measure theoretic entropy of (Y, ν, S) by $h_\nu(Y)$. We construct a SFT called *the checkerboard island shift* $\overline{X}$ which has positive entropy, is topologically mixing, does not have the UFP and:

Theorem 1.1. *There is a number $0 < h_0 < h(\overline{X})$ such that if $Y \subset \overline{X}$ is a sub-shift of finite type which has the UFP, then $h(Y) \leq h_0$.*

This theorem is not quite sufficient to provide an obstruction for a measure theoretic result. In particular, it follows from our proof of Theorem 1.1 that there are many high entropy measures in $\mathcal{M}_e(\overline{X}, S)$. On the other hand, the following result shows that those measures correspond to a particular type of dynamical behaviour and thus $\overline{X}$ cannot be a universal model in the measure theoretic category either.

Theorem 1.2. *Let $\overline{X}$ be the checkerboard island shift. Then there is a number $0 < h_0 < h(\overline{X})$ such that for any measure $\mu \in \mathcal{M}_e(\overline{X}, S)$, if $(\overline{X}, S, \mu)$ is weakly mixing, then $h_\mu(\overline{X}) \leq h_0$.*

We note that Theorem 1.2 contrasts strongly with results that one obtains for SFT with stronger mixing properties. If Y has the UFP and dense periodic points then Y satisfies the variational principle where the supremum is taken over only measures that are weak mixing, or even K [9]. If Y satisfies a stronger mixing condition, namely is *strongly irreducible*, one can drop the requirement of dense periodic points and prove a variational principle using Bernoulli measures alone [1].

Finally, in this paper we investigate the existence of strictly local maxima for the entropy function of subshifts. Our motivation is the proof in [10] which uses the Burton-Rothstein joinings machinery [2]. A crucial ingredient in that argument is a *Randomization Lemma.* This requires that the entropy function $h : \mathcal{M}(Y, S) \to \mathbb{R}$ given by $h(\nu) = h_\nu(Y)$ does not have a strictly local maximum where $\mathcal{M}_e(Y, S)$ is given the weak* topology. A natural question to ask is if this can ever happen for a sub-shift Y.

In the one dimensional case it is not possible to have a strictly local maximum of the entropy function if Y is a topologically mixing SFT. On the other hand Haydn has shown that it is possible if Y is not required to be of finite type [4]. In his examples the space of invariant ergodic measures are disconnected in the weak* topology. Here we use similar techniques to provide examples of $\mathbb{Z}^d$ sub-shifts with disconnected ergodic invariant measures and strictly local maxima for the entropy function, $d \geq 2$.

The next question then is whether a disconnected set of invariant measures is necessary for such a result, and we show that it is not. We construct examples exhibiting this behaviour in both $\mathbb{Z}$ and $\mathbb{Z}^d$. Perhaps more significant is that our higher dimensional example is a SFT.

The organization of the paper is as follows. In the first section we introduce our notation and basic definitions. Section 3 contains the checkerboard island example. In Section 4 we have the higher dimensional examples whose entropy functions have strictly local maxima, but have disconnected measure spaces. Finally in Section 5 we have the one dimensional sub-shift and higher dimensional SFT with connected measures but local maxima for their entropy functions.

2. Notation and Basic Definitions

Let A be a finite alphabet and let $S = \{S_{\vec{v}}\}_{\vec{v}\in\mathbb{Z}^d}$ on $A^{\mathbb{Z}^d}$ denote the usual shift action. If Y is a closed and shift invariant subset of $A^{\mathbb{Z}^d}$, we call (Y, S) a sub-shift or shift space.

Given a sub-shift (Y, S), a subset $R \subset \mathbb{Z}^d$, and $y \in Y$ we denote the symbols appearing in the locations determined by R in y by $y[R]$. We call $\mathbf{a} \in A^R$ a *configuration from the shift Y* if there is $y \in Y$ such that $y[R] = \mathbf{a}$. For $\vec{v} = (v_1, \ldots, v_d) \in \mathbb{Z}^d$, we let $\|\vec{v}\| = \max |v_i|$. For subsets R_1 and R_2 of $\mathbb{Z}^d$ we set $d(R_1, R_2) = \inf_{\vec{v}\in R_1, \vec{u}\in R_2} \|\vec{v} - \vec{u}\|$. Finally, we let $B_n = \{\vec{v} \in \mathbb{Z}^d : 0 \leq v_i < n\}$.

We call Y a shift of finite type (SFT) if there exist $s \in \mathbb{N}$ and a finite collection of blocks $\mathcal{F} \subset A^{B_s}$ such that $y \in Y$ if and only if $y[B_s + \vec{v}] \notin \mathcal{F}$ for all $\vec{v} \in \mathbb{Z}^d$.

There are a variety of filling conditions that have already been discussed in the literature for $\mathbb{Z}^d$ shifts of finite type. To avoid any confusion we explicitly state the definitions of the conditions we will be using here.

Definition 2.1. *A shift space (Y, S) is topologically mixing if for all finite subsets $R_1, R_2 \subset \mathbb{Z}^d$, there exists an $\ell > 0$ such that for all $\vec{v} \in \mathbb{Z}^d$ such that $d(R_1, R_2 + \vec{v}) > \ell$, for all $y_1, y_2 \in Y$ there exists a point $y \in Y$ such that*

$$y[R_1] = y_1[R_1] \qquad \textit{and} \qquad y[R_2 + \vec{v}] = y_2[R_2 + \vec{v}].$$

Definition 2.2. *A shift space (Y, S) is strongly irreducible (SI) if there exists $l > 0$ such that if $R_1, R_2 \subset \mathbb{Z}^d$ are finite subsets such that $d(R_1, R_2) > l$ then for all $y_1, y_2 \in Y$ there is a point $y \in Y$ such that*

$$y[R_1] = y_1[R_1] \qquad \textit{and} \qquad y[R_2] = y_2[R_2].$$

Definition 2.3. *A SFT (Y, S) has the uniform filling property (UFP) with filling length $l > 0$ if for all points $y_1, y_2 \in Y$ there is a point $y \in Y$ such that*

$$y[B_n] = y_1[B_n] \qquad \textit{and} \qquad y[B^c_{n+l}] = y_2[B^c_{n+l}].$$

Topological mixing is the weakest of the three properties. In particular, it does not imply that the SFT has positive topological entropy, whereas both the UFP and SI imply positive entropy. It is clear that if a SFT is SI then it also has the UFP. While all known examples of shifts with the UFP are also SI, it is not clear whether the implication holds in general.

We shall consider two different topologies on $\mathcal{M}_e(Y,S)$. One is the classical weak* topology, and the other is the $\overline{d}$ topology. To define the $\overline{d}$ metric we let $P = \{p_a : a \in A\}$ denote the time zero partition of Y and $\overline{P}_1$ and $\overline{P}_2$ the partitions on $Y \times Y$ whose atoms are of the form $\{p_a \times Y : a \in A\}$ and $\{Y \times p_a : a \in A\}$ respectively. Then for $\mu, \nu \in \mathcal{M}_e(Y,S)$ we set

$$\overline{d}(\mu,\nu) = \inf_{\gamma} \gamma(\overline{P}_1 \triangle \overline{P}_2)$$

where the infimum is taken over all joinings of (Y,μ,S) and (Y,ν,S). The $\overline{d}$ topology is stronger than the weak* topology (cf [11]). It follows that $\overline{d}$ connectedness implies weak* connectedness.

We shall say that a measure $\mu \in \mathcal{M}_e(Y,S)$ is a *strictly local maximum* of the entropy function if there exists a weak*-neighbourhood O of μ in $\mathcal{M}_e(Y,S)$ such that $h_\nu(Y) \le h_\mu(Y)$ for every $\nu \in O$ and there exists a $\mu' \in \mathcal{M}_e(Y,S)$ with $h_{\mu'}(Y) > h_\mu(Y)$.

3. The Checkerboard Island System

We start with an alphabet A consisting of the following symbols:

Figure 1. The alphabet of the sub-shift

We first obtain a $\mathbb{Z}^2$ SFT X by letting the set of allowable 2×2 configurations be exactly those contained in the configuration shown below in Figure 2.

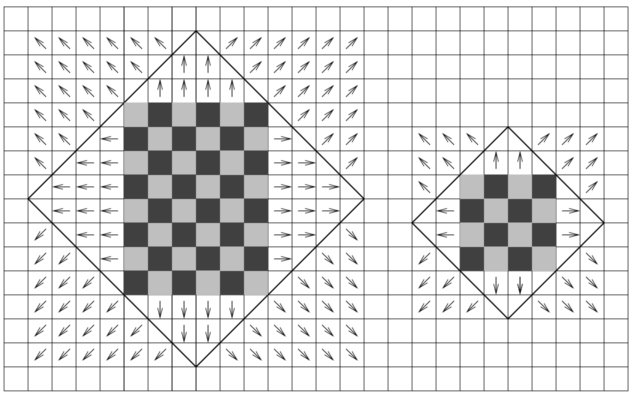

Figure 2. A typical configuration

We call the checkerboard regions *islands*. From the legal 2×2 configurations, the reader will see that if a square has a horizontal and a vertical neighbour that are shaded checkerboard symbols, then the square itself is forced to contain a checkerboard symbol. From this, it follows that any connected patch of checkerboard symbols is forced to to be rectangular (possibly with one or more infinite sides). Since the top left and bottom right corners (if they exist) are forced to be light grey, while the top right and bottom left corners are forced to be dark grey, we see that any finite length edge of an island is forced to have even length.

We claim that each checkerboard island with finite length edges is forced to have a neighbourhood consisting of arrows and lines in the particular arrangement that appears in Figure 2. As an illustration, we show in the case of a 6×6 checkerboard how the rules force the neighbourhood of the island to be as shown above.

Given that this is the full extent of the checkerboard region, the adjacency rules force a northeast diagonal line to be placed in the square above the top left corner. The square northwest of the top left corner is forced to contain a northwest arrow. This is repeated at the other corners, with the directions modified appropriately. It is then easy to check that the placement of the corner diagonal lines and arrows forces the extended configuration on an 8×8 region to be the one shown below in Figure 3:

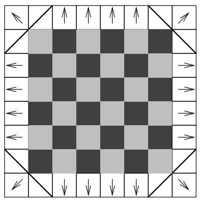

FIGURE 3. 8×8 region forced by a 6×6 checkerboard

The arrows and lines now force the placement of unique arrows and lines in their neighbouring locations, finally extending the checkerboard to the configuration on the 12×12 region to be as shown below in Figure 4.

Finally, next to the places where the diagonal lines meet, the allowed configurations force blank symbols to be placed. This then propagates around the shape forcing a collar of blank symbols. The final configuration is shown below in Figure 5.

We call the above forced neighbourhood the *shadow* of the island, and we note that even though the shift is of finite type, the adjacency rules force the shadow size to grow linearly in the dimension of the checkerboard island. A finite configuration $\mathbf{a}$ from X is called a *safe* configuration if all the islands occurring in $\mathbf{a}$ have the property that their shadows are also contained in $\mathbf{a}$.

Since we require X to be a sub-shift it also contains the limit of sequences of configurations in X with larger and larger checkerboards. These

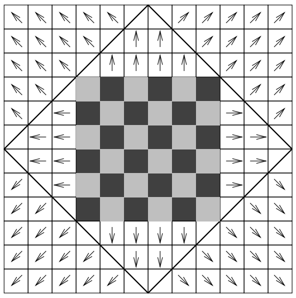

FIGURE 4. 12×12 region forced by a 6×6 checkerboard

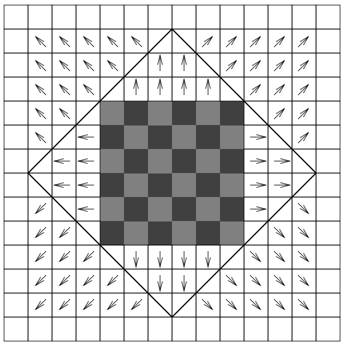

FIGURE 5. 14×14 region forced by a 6×6 checkerboard

are configurations which contain infinite checkerboard islands. There are various types of these with one or both of the dimensions being infinite (where this can mean either singly- or doubly-infinite). We note that each of these infinite checkerboards with the exception of the finite by singly-infinite checkerboard determines the entire configuration. Two such configurations are schematically illustrated in Figure 6. Note that the shadow of the first fills a half plane, whereas the shadow of the second fills the entire plane. We observe for future reference that the maximum number of infinite islands in a single point of X is two (as any greater number would necessarily have the property that their shadows overlap).

For notational convenience we let $\mathcal{I}$ denote those elements of X which contain an infinite island, and $\mathcal{F}$ denote the complement of $\mathcal{I}$.

Lemma 3.1. *X is topologically mixing.*

Proof. The result follows from two observations. If $R \subset \mathbb{Z}^d$ is finite, then there is a finite set $\tilde{R} \supset R$ such that for all $x_1 \in X$, there exists a $x_2 \in \mathcal{F}$ such that $x_1[R] = x_2[R]$ and $x_2[\tilde{R}]$ is a safe configuration.

Given two safe configurations $\mathbf{a} \in A^{R_1}$ and $\mathbf{b} \in A^{R_2}$ from X it is clear that if $\vec{v} \in \mathbb{Z}^d$ is such that $R_1 \cap (R_2 + \vec{v}) = \emptyset$, then there is $x \in X$ with $x[R_1] = \mathbf{a}$ and $x[R_2 + \vec{v}] = \mathbf{b}$.

FIGURE 6. Two configurations containing infinite checkerboards

$\square$

We next extend the SFT by replacing the light checkerboard symbol with 200 distinct light checkerboard symbols and the dark checkerboard symbol similarly. The adjacency restrictions remain as before, except now the checkerboards alternate between light and dark symbols each of which may taken from any one of the 200 symbols of the respective color. We call the new system $\bar{X}$ the *checkerboard island system.*

Let $C \subset X$ denote the sub-shift of X consisting of the two infinite checkerboard configurations covering the entire plane, and let $\overline{C} \subset \overline{X}$ be the configurations obtained by the various colorings of the two configurations in C. The proofs of both Theorem 1.1 and 1.2 rely on the following lemma.

Lemma 3.2. $h(\bar{X}) = \log 200$.

Proof. We note that $\overline{C}$ is a sub-shift of $\overline{X}$ with topological entropy $\log 200$. Thus it follows that $h(\overline{X}) \geq \log 200$.

If we can prove that $h_\nu(\overline{X}) \leq \log 200$, for every ergodic $\nu \in \mathcal{M}_e(\overline{X}, S)$, then the usual variational principle will give $h(\overline{X}) \leq \log 200$.

Given such a measure ν, note that since the set of configurations $\overline{C}$ is a shift invariant set $\nu(\overline{C})$ is either 0 or 1. If $\nu(\overline{C}) = 1$ then since $h(\overline{C}) = \log 200$, by the usual variational principle we have $h_\nu(\overline{X}) \leq \log 200$.

Suppose now that $\nu(\overline{C}) = 0$. Letting A be the event that the origin is at the corner of an infinite island, we see that this event occurs at most finitely many times in any point of $\bar{X}$. Hence, by the Poincaré Recurrence Theorem we have $\nu(A) = 0$. It follows that $\nu(\mathcal{I}) = 0$.

Then for ν a.e. configuration $\bar{x} \in \bar{X}$ the density of the checkerboard symbols is at most $\frac{1}{2}$. To see this note that any checkerboard island occurring in $\bar{x}$ is accompanied by its shadow, which occupies at least as many cells as the island itself.

Let $\mathcal{P}$ denote the time zero partition on $\overline{X}$. Define $\prec$ to be the lexicographic ordering on $\mathbb{Z}^2$ and set $\mathcal{B} = \bigvee_{\vec{v} \prec \vec{0}} S^{-\vec{v}}(\mathcal{P})$. We note that for $\overline{x} \in \overline{X}$, the symbols occurring in $\overline{x}$ at locations preceding the origin horizontally and vertically determine whether $\overline{x}[\vec{0}]$ is a checkerboard symbol or not. If it is a checkerboard symbol then $I_\nu(\mathcal{P}|\mathcal{B})(\overline{x}) \leq \log 200$. If not, then it follows from the size of the alphabet of X that $I_\nu(\mathcal{P}|\mathcal{B})(\overline{x}) \leq \log 13$. Hence integrating the information function we have

$$h_\nu(\overline{X}) = \int I_\nu(\mathcal{P}|\mathcal{B})(\overline{x})\,d\nu(\overline{x}) \le \log 13 + \frac{1}{2}\log 200 < \log 200. \tag{1}$$

□

As a corollary of the above proof we have the following key proposition.

Proposition 3.3. *There exists $h_0 < h(\overline{X})$ such that if $\nu \in \mathcal{M}_e(\overline{X}, S)$ and $\nu(\bar{C}) = 0$ then $h_\nu(\overline{X}) \le h_0$.*

Proof. Set $h_0 = \frac{1}{2}\log 200 + \log 13$, and the result follows from equation (1) in the proof of Lemma 3.2 above. □

We can now prove both Theorems 1.1 and 1.2.

Proof. (**of Theorem 1.1**) Suppose Y is a sub-shift of finite type of $\bar{X}$ that satisfies the UFP with filling length l. Suppose that there is a point $y \in Y$ which contains a checkerboard island with at least one of its side lengths $r > 2l$. Then any checkerboard configuration of shape $s = \{(0,n) : 0 \le n \le r-1\}$ is a configuration from Y. Let $\mathbf{a}$ be one such configuration and place two copies of $\mathbf{a}$, one in location $R_1 = \vec{0} + s$ and the other in location $R_2 = (2l+1, 0) + s$. Since $\mathbf{a}$ is a configuration from Y and the copies are placed at a distance $2l$ away we can use the UFP to conclude that there exists $y \in Y$ such that $y[R_1] = \mathbf{a}$ and $y[R_2] = \mathbf{a}$.

On the other hand, since R_1 and R_2 are in opposite parity locations, each checkerboard configuration must be part of a safe configuration in y. By construction, however, the shadows in the safe configuration are of width at least $\frac{r}{2} > l$, so the shadows of the two safe regions must intersect. But then y cannot be in $\bar{X}$.

Thus for any sub-shift of finite type Y of $\bar{X}$ with the UFP, if $\mu \in \mathcal{M}_e(Y, S)$ then $\mu(\overline{C}) = 0$. Then by Proposition 3.3 we have $h_\mu(Y) = h_\mu(\bar{X}) \le h_0 < h(\overline{X})$. □

Proof. (**of Theorem 1.2**).

Suppose that $\nu \in \mathcal{M}_e(\overline{X}, S)$ is a weak mixing measure. Since ν is ergodic $\nu(\overline{C})$ is either 0 or 1. If the former holds, the result follows by Proposition 3.3. If the latter holds we use the underlying checkerboard structure of the configurations in $\overline{C}$ to define the function

$$f(y) = \begin{cases} -1 & \text{if } y[\vec{0}] \text{ is a light checkerboard symbol} \\ 1 & \text{otherwise} \end{cases}$$

which is an eigenfunction for $(\overline{X}, S, \nu)$ with eigenvalue $(-1,-1)$. This contradicts our weak mixing assumption. □

We note that by changing the system so that instead of checkerboards, each of the central rectangles consists of symbols of a single shade of grey (either light or dark grey), one recovers a mixing system with exactly two measures of maximal entropy. Although it is not strongly irreducible, the example seems to us to be somewhat in the spirit of the example of Burton and Steif [3].

4. Local maxima for the entropy function I: disconnected invariant measures

Here we consider examples with alphabet

$$A = \{-n_-, -n_- + 1, \cdots, -1, 0, 1, 2, \cdots, n_+\}$$

with $n_-, n_+ \in \mathbb{N}$ and $n_- < n_+$. We denote the subset of A consisting of negative symbols only by A_- and the positive symbols only by A_+. Informally, our idea is to construct subshifts whose configurations consist of rectangles of only positive or negative symbols in a sea of 0's. We obtain a family of sub-shifts by varying the distance regions of signed integers have to be separated from one another. We use 0 as a neutral symbol to separate such regions. The details are given below.

For $R \subset \mathbb{Z}^d$ we let $\partial R = \big(\cup_{\vec{v} \in [-1,1]^d} R + \vec{v}\big) \setminus R$. For $y \in A^{\mathbb{Z}^d}$ and $R = B_k + \vec{v}$ for some $k \in \mathbb{N}$ and $\vec{v} \in \mathbb{Z}^d$ we say $y[R]$ is a *negative island* if $y[R] \subset A_-^R$ and $y[\partial R] = \{0\}^{\partial R}$. A *positive island* is defined similarly.

Given a function $f : \mathbb{N}^2 \to \mathbb{N}$, we define the subshift (Y_f, S) in the following way. A point $y \in A^{\mathbb{Z}^d}$ is in Y_f if each negative (positive) symbol belongs to a negative (positive) island and for any pair of islands $y[R_1]$, $y[R_2]$ we have

$$d(R_1, R_2) \geq f(r_1, r_2)$$

where $R_i = B_{r_i} + \vec{v}_i$ for $\vec{v}_i \in \mathbb{Z}^d$ and $i = 1, 2$.

For a sub-shift Y from this family of examples, $\mathcal{M}_e(Y, S)$ has two disjoint subsets $\mathcal{M}_e^+(Y, S)$ and $\mathcal{M}_e^-(Y, S)$, the measures concentrated only on $(A_+)^{\mathbb{Z}^d}$ and $(A_-)^{\mathbb{Z}^d}$ respectively. Since $((A_+)^{\mathbb{Z}^d}, S)$ is the full shift on n_+ symbols, the entropy function has no strictly local maxima on $\mathcal{M}_e^+(Y, S)$. Further, the measure μ_+ which weights each positive symbol exactly by $\frac{1}{n_+}$ is the measure of maximal entropy on $\mathcal{M}_e^+(Y, S)$. Similarly for $\mathcal{M}_e^-(Y, S)$ and the corresponding measure μ_-. In addition, $h_{\mu_+}(Y) = \log(n_+) > \log(n_-) = h_{\mu_-}(Y)$. We write C_+ for $\{y \in Y : y[\vec{0}] \in A_+\}$ and define C_- similarly.

We now show that whether μ_- is a strictly local maximum for the entropy function depends on the behaviour of the function f.

Proposition 4.1. *Suppose $f(m_1, m_2) = F(\max(m_1, m_2))$.*

If $\underline{\lim} \frac{F(m)}{m} = 0$ then $\mathcal{M}_e(Y_f, S)$ is connected and μ_- is not a strictly local maximum for the entropy function on $\mathcal{M}_e(Y_f, S)$ in the weak topology.*

Proof. We first argue that the entropy function has no strictly local maximum. For a fixed $m \in \mathbb{N}$ we define a measure $\nu_m \in \mathcal{M}_e(\{+, -, 0\}^{\mathbb{Z}^d}, S)$ in the following way. Divide the integer lattice into translates of B_m separated by collars of width $F(m)$ in both directions. Build a word $x \in \{+, -, 0\}^{\mathbb{Z}^d}$ by placing 0's in the collars separating the translates of B_m and by placing the symbol $-$ in every row of B_m's except every k^{th} one, starting at the origin. In every k^{th} row place the symbol $+$. We let

$$\nu_m = \frac{1}{|P|} \sum_{\vec{v} \in P} \delta_{S^{\vec{v}}x}$$

where $P \subset \mathbb{Z}^d$ is the fundamental domain of x. The measure ν_m is shift invariant and ergodic.

We then construct $\mu \in \mathcal{M}_e(Y_f, S)$ by replacing each $-$ symbol with an independently chosen element of A_-, and replacing each $+$ symbol with an independently chosen element of A_+. The symbol 0 remains as is. The measure μ will also be shift invariant and ergodic. Further we will have that

$$\begin{aligned} h_\mu(Y_f) &= \mu(C_+) \log n_+ + \mu(C_-) \log n_- \\ &= \mu(C_+) \log n_+ + (1 - \mu(C_+) - \mu(0)) \log n_- \\ &= \log n_- + \mu(C_+)(\log n_+ - \log n_-) - \mu(0) \log n_-. \end{aligned}$$

Since $\mu(0) < (\frac{F(m)}{m} + 1)^d - 1 = O(F(m)/m)$, by choosing m large enough we can guarantee that

$$h_\mu(Y_f) > \log n_-.$$

Since μ_- and μ can be joined perfectly on the regions inside the negative squares, we have

$$\overline{d}(\mu_-, \mu) < \frac{1}{k} + O(F(m)/m)$$

Again, by choosing m and k large enough we can make this distance arbitrarily small. In particular, it is clear that μ_- is not a measure of locally maximal entropy.

We now argue that $\mathcal{M}_e(Y_f, S)$ is connected in the $\overline{d}$ topology. First notice that any $\overline{d}$-neighbourhood of μ_- and μ_+ contains an ergodic measure μ for which $\mu(0) > 0$. We can construct such a measure using the technique used above, without alternating the sign of the symbols every kth row. It now suffices to show that any $\mu \in \mathcal{M}_e(Y_f, S)$ with the property that $\mu(0) > 0$ can be continuously deformed in the $\overline{d}$ topology into the measure δ_0, the point mass supported on the fixed point of all 0's.

Given such a μ, form the product measure $\tilde{\mu} = \mu \times \lambda^{\mathbb{Z}^2}$ on $\tilde{Y} = Y_f \times [0,1]^{\mathbb{Z}^2}$, where λ is the restriction of Lebesgue measure to the unit square. Define a family of measurable maps $\Phi_r : \tilde{Y} \to Y_f$ by the following rule: $\Phi_r((y,x)_{\vec{v}})$ is equal to 0 if $y_{\vec{v}}$ is in an island $y[B_m + \vec{u}]$ and $\max\{x[\vec{w}] : \vec{w} \in B_m\} \le r$, and is $y_{\vec{v}}$ otherwise. Note that this is a family of measurable maps which gradually turn the islands into blocks of 0's. Since λ is Bernoulli and hence weakly mixing, it follows that $\tilde{\mu}$ is ergodic with respect to $S \times S$ on $\tilde{Y}$. Thus $\mu_r = \tilde{\mu} \circ {\Phi_r}^{-1}$ is an ergodic measure on Y_f.

We now let $\psi(r) = \mu_r$. We can check that this function is $\bar{d}$-continuous by using the natural joining of μ_r and $\mu_{r'}$ given by $\nu(A \times B) = \tilde{\mu}({\Phi_r}^{-1}(A) \cap {\Phi_{r'}}^{-1}(B))$. It is clear that $\psi(0) = \mu$ and $\psi(1) = \delta_0$. $\square$

Proposition 4.2. *If* $\underline{\lim}\frac{F(m)}{m} > 0$ *then* $\mathcal{M}_e(Y,S)$ *is disconnected and* μ_- *is a strictly local maximum for the entropy function.*

Proof. Let $\mu \in \mathcal{M}_e(Y,S)$ be such that $\mu(0) > 0$ and $\mu(C_-) > 0$. There is a $\beta > 0$ such that for all m we have $\frac{F(m)}{m} > \beta$. Thus any island of shape B_m is surrounded by a collar of 0's of width at least βm. Using the ergodic theorem we can then conclude that

$$\mu(0) > \beta\mu(C_-). \tag{2}$$

Let $\epsilon > 0$ be given. If μ is in a sufficiently small weak* neighbourhood of μ_- then $\mu(C_-) > 1-\epsilon$. Then by (2) we must have $\mu(0) > \beta(1-\epsilon)$. Clearly for ϵ sufficiently small this gives a contradiction as we have two disjoint sets whose measures sum to more than 1.

This shows that all sufficiently small weak* neighbourhoods of μ_- in $\mathcal{M}_e(Y_f,S)$ are entirely contained in $\mathcal{M}_e^-(Y_f,S)$. Since μ_- is the unique measure of maximal entropy in $\mathcal{M}_e^-(Y,S)$, we see that it is indeed a local maximum of the entropy function on $\mathcal{M}_e(Y_f,S)$. Clearly since μ_+ has higher entropy, μ_- is a strictly local maximum. □

5. Local maxima for the entropy function II: connected invariant measures

We have seen that there exist examples of mixing higher dimensional shift spaces where the set of invariant measures is disconnected and the entropy function has a strictly local maximum. In this section we show that it is possible to have strictly local maxima of the entropy function even when the invariant measures are connected in the $\overline{d}$ metric. In two dimensions it is possible even when the shift is a mixing SFT.

5.1. **The ribbon shift: a mixing two dimensional example of finite type.** To construct the *ribbon shift* X we use our usual alphabet and four new symbols. Let $A_- = \{-n_-,\dots,-1\}, A_+ = \{1,\dots,n_+\}$ with

$$2 < n_- < n_+. \tag{3}$$

and set $\bar{A} = A_- \cup A_+ \cup \{\rightarrow, \searrow, \uparrow, \nwarrow\}$. X will be a subshift of $\bar{A}^{\mathbb{Z}^2}$.

Informally, to define X we will use the symbols $\rightarrow, \searrow$ $(\uparrow, \nwarrow)$ to construct infinite horizontal (vertical) ribbons of constant height (width) in a sea of negative symbols. The ribbons can bend using the $\searrow$ ($\nwarrow$ for vertical ribbons). Each horizontal (vertical) ribbon will have a fixed but arbitrary height (width). When a horizontal and vertical ribbon intersect they will do so in a rectangular region full of positive symbols. In particular, ribbons may not change direction in a region of positive symbols. An example of a configuration from X is shown below in Figure 7. Note that the blank spaces between ribbons are filled with negative symbols.

In spite of the fact that the ribbons are required to each keep a constant height or width (which may be arbitrarily large) X is a shift of finite type. We define X formally by specifying that all the allowable 2×2 configurations

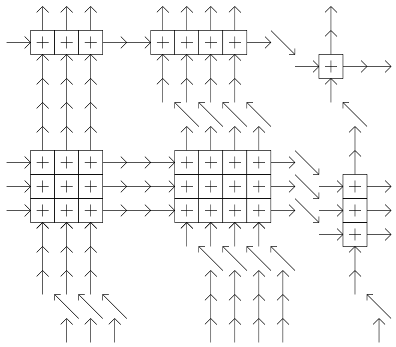

FIGURE 7. Ribbon example

are those which arise in the configuration shown in Figure 7. It is easy to check that these rules force the horizontal (vertical) ribbons to be of constant height (width) and to go in an essentially horizontal (vertical) direction. A pair of same direction ribbons are separated from each other by at least a distance of one. Each time a ribbon changes direction, it must stay in that direction for at least one more step.

As in the checkerboard example, the positive symbols are forced by the finite type rules to occur in rectangular patches. Any patch of positive symbols is forced to be surrounded by by $\uparrow$s on the top and bottom and $\rightarrow$s on the left and right.

Finally, there are exceptional configurations in which there is one or more infinite block of positive or negative symbols, or one infinite ribbon (either horizontal or vertical). These arise as limits of points with finite blocks of positive and negative symbols.

Lemma 5.1. *X is topologically mixing.*

Proof. Fix a $k > 0$. We consider the problem of placing two configurations on B_k in a single point of X. The obstruction to doing this is the positions and dimensions of the horizontal and vertical ribbons leaving the boundaries of the blocks. A pair of horizontal ribbons may not intersect unless they are of the same width and similarly for vertical ribbons. So we need enough room for the ribbons leaving the boundary of the first block to move out of the way to accommodate the ribbons leaving the second block. Note that the horizontal ribbons have slope between 0 and $-\frac{1}{2}$, thus if $\vec{v} \in \mathbb{Z}^2$ is such that $\|\vec{v}\| \geq 4k$ and $x_1, x_2 \in X$, then there is a point $x \in X$ agreeing with $\mathbf{a} = x_1[B_k]$ and $\mathbf{b} = x_2[B_k + v]$. This is illustrated below in Figure 8 $\square$

We define two subsets of the alphabet $\bar{A}$ consisting of the symbols forming horizontal and veritcal ribbons respectively: $H = A_+ \cup \{\rightarrow, \searrow\}$ and $V = A_+ \cup \{\uparrow, \nwarrow\}$. We define C_H to be $\{x \in X : x[\vec{0}] \in H\}$ and define C_V similarly.

The following lemma is a technical result that we need in our argument.

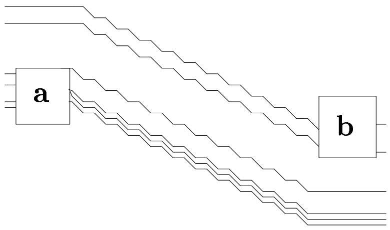

FIGURE 8. The ribbon system is topologically mixing

Lemma 5.2. *Let $\mu \in \mathcal{M}_e(X, S)$. Then*

$$\mu(C_+) \leq \tfrac{9}{4}\mu(C_H)\mu(C_V). \tag{4}$$

In fact, a more careful argument can show that the factor of $\frac{9}{4}$ is unnecessary but the weaker estimate is sufficient for our purposes.

Proof. We begin by arguing that if we let

$$A_{(0,n)}\chi_{C_H}(x) = \frac{1}{n}(\chi_{C_H}(x) + \ldots + \chi_{C_H}(S^{(0,n-1)}x))$$

then for μ-a.e. $x \in X$

$$A_*\chi_{C_H}(x) = \lim_{n\to\infty} A_{(0,n)}\chi_{C_H}(x) = \mu(C_H). \tag{5}$$

Since $S^{(0,1)}$ is a measure preserving transformation the above averages converge for μ-almost every x and clearly the limit $A_*\chi_{C_H}(x)$ is $S^{(0,1)}$-invariant. Since $|A_{[0,n]}\chi_{C_H}(x) - A_{[0,n]}\chi_{C_H}(S^{(1,0)}x)| \leq \frac{1}{n}$, we have that $A_*\chi_{C_H}$ is also $S^{(1,0)}$ invariant. It now follows by ergodicity that $A_*\chi_{C_H}$ is almost everywhere a constant. Integrating then shows that $A_*\chi_{C_H}(x) = \mu(C_H)$ almost everywhere.

Arguing similarly we have that for a.e. $x \in X$ if we let

$$A_{(n,0)}\chi_{C_V}(x) = \frac{1}{n}(\chi_{C_V}(x) + \cdots + \chi_{C_V}(S^{(n,0)}(x)))$$

then

$$\lim_{n\to\infty} A_{(n,0)}\chi_{C_V}(x) = \mu(C_V). \tag{6}$$

To obtain the estimate in (4) we will count the proportion of positive symbols that occur in $x[B_n]$ for typical points $x \in X$. Recall that these symbols must belong to the intersection of horizontal and vertical ribbons intersecting inside $x[B_n]$. It will thus suffice to count the possible number of such intersections.

We note that for μ-a.e. $x \in X$ given an $\epsilon > 0$ we can choose $m(x)$ large enough so that for all $n \geq m$ we have $A_{(n,0)}\chi_{C_V}(x) \leq (1+\epsilon)\mu(C_V)$ and $A_{(0,n)}\chi_{C_H}(x) \leq (1+\epsilon)\mu(C_H)$. Choose m_0 such that for x belonging to a set of measure $1-\epsilon$, $m(x) \leq m_0$. We fix such an x and an $n \geq m_0$ and consider $x[B_n]$.

Since the vertical ribbons can bend to the left, the vertical ribbons in $x[B_n]$ hit the x axis anywhere along $A_h = \{\vec{v} \in \mathbb{Z}^2 \colon 0 \le v_1 \le 3n/2\}$. Similarly, horizontal ribbons hit the y axis on $A_v = \{\vec{v} : v_1 = 0, 0 \le v_2 \le 3n/2\}$. By our choice of n, the number of points in $x[A_h]$ that are contained in vertical ribbons is at most $\frac{3n}{2}\mu(C_V)(1+\epsilon)$. Similarly, the number of points in $x[A_v]$ that are contained in horizontal ribbons is at most $\frac{3n}{2}\mu(C_H)(1+\epsilon)$.

Since the vertical ribbons have a constant width, and the horizontal ribbons have a constant height and must intersect in a rectangular region we see that the set of points belonging both to a horizontal ribbon in $x[B_n]$ and to a vertical ribbon in $x[B_n]$ is of cardinality at most $9\frac{n^2}{4}\mu(C_V)\mu(C_H)(1+\epsilon)^2$.

Thus the number of positive symbols in $x[B_n]$ has the same upper bound and since n is arbitrarily large and ϵ is arbitrary, we see that (4) holds, as required.

□

As before we let μ_- and μ_+ denote the maximal entropy Bernoulli measures on A_- and A_+ respectively.

Proposition 5.3. *The measure $\mu_- \in \mathcal{M}_e(X,S)$ is a strictly local maximum for the entropy function on $\mathcal{M}_e(X,S)$ with the weak* topology.*

Proof. Let $\mu \in \mathcal{M}_e(X,S)$ and let $\epsilon > 0$ be arbitrary. Using the ergodic theorem choose n such that for x belonging to a set of μ measure at least $1-\epsilon$, we have

$$(7) \qquad \begin{aligned} (1-\epsilon)\mu(C_V)n^2 < \#\{\vec{v} \in B_n : S^{\vec{v}}(x) \in C_V\} < (1+\epsilon)\mu(C_V)n^2 \\ (1-\epsilon)\mu(C_H)n^2 < \#\{\vec{v} \in B_n : S^{\vec{v}}(x) \in C_H\} < (1+\epsilon)\mu(C_H)n^2 \end{aligned}$$

Also from Lemma 5.2 we can guarantee that

$$(8) \qquad \#\{\vec{v} \in B_n : S^{\vec{v}}(x) \in C_+\} < 3\mu(C_V)\mu(C_H)n^2.$$

Together (7) and (8) imply that for such a point x

$$\#\{\vec{v} \in B_n : S^{\vec{v}}(x) \in C_-\} < \big(1-(1-\epsilon)\mu(C_V)-(1-\epsilon)\mu(C_H)+3\mu(C_V)\mu(C_H)\big)n^2.$$

We can now estimate $h_\mu(X)$ by counting the number of different configurations possible for $\cup_{\vec{v}\in B_n} S^{\vec{v}}(x)$ for the set of points x satisfying (7) and (8). We first consider the vertical ribbons. Since there are exactly two positions that a vertical ribbon can occupy in a row once its position in the previous row is known, the number of configurations of the vertical ribbons in $x[B_n]$ is bounded above by $2^{(1+\epsilon)\mu(C_V)n^2}$. Similarly, the total number of configurations of the horizontal ribbons is less than $2^{(1+\epsilon)\mu(C_H)n^2}$

Putting this together and using the fact that ϵ is arbitrary, we have

$$\begin{aligned} h_\mu(X) \le& \log n_- + 3\mu(C_V)\mu(C_H)\log n_- n_+ \\ &- (\mu(C_V)+\mu(C_H))(\log n_- - \log 2). \end{aligned}$$

Note that if

$$3\mu(C_V)\mu(C_H)\log(n_-n_+) < (\mu(C_V)+\mu(C_H))(\log n_- - \log 2)$$

then $h_\mu(X) < \log n_-$. Since n_- was assumed to be greater than 2, there exists an $\eta > 0$ such that if $\mu(C_V) < \eta$ and $\mu(C_H) < \eta$, then $h_\mu(X) \leq \log n_-$.

The above argument shows that μ_- is a local maximum for the entropy function on $\mathcal{M}_e(X,S)$. Since clearly $h_{\mu_+}(X) > h_{\mu_-}(X)$ we have that μ_- is a strictly local maximum for the entropy function. $\square$

We introduce some new terminology which will be useful in the following argument. We call a measure $\mu \in \mathcal{M}(X,S)$ *horizontally well supported* if it is supported only on configurations with infinitely many horizontal ribbons. We make a similar definition for *vertically well supported* measures. If a measure is both horizontally and vertically well supported we call it *well supported.*

Proposition 5.4. *The space $\mathcal{M}_e(X,S)$ is connected in the $\overline{d}$ topology.*

Proof. Arguing in a manner similar to Proposition 4.1 it is easy to see that any measure in $\mathcal{M}_e^-(X,S)$ can be continuously deformed into the measure δ_-, the point mass on the fixed point consisting of all -1s. Similarly, any measure $\mu \in \mathcal{M}_e^+(X,S)$ can be deformed into δ_+, the measure supported on the point consisting of all $+1$s.

In addition, any $\overline{d}$ neighbourhood of δ_+ must contain a well supported measure ν. To construct such a measure let ν_k be the measure supported on the periodic configurations in X where the horizontal (vertical) ribbons all of height (width) k are regularly spaced (with a gap of say 2 between two ribbons) and do not have any bends, and the intersection of a horizontal and a vertical ribbon is a region which containing only the symbol $+1$. As k is allowed to grow, it is clear that $\overline{d}(\nu_k, \delta_+)$ will tend to 0.

Further, an argument identical to that in Proposition 4.1 will show that any well supported measure in $\mathcal{M}_e(X,S)$ can be deformed continuously in the $\overline{d}$ topology into a well supported measure μ with the property that the only negative symbols it sees is -1.

To prove the connectedness of $\mathcal{M}_e(X,S)$ it now suffices to show that any such measure μ can be continuously deformed into the point mass δ_-. Namely, we need to find a way to continuously erase the ribbons from the support of μ. Here we cannot argue as in Proposition 4.1: if we randomly start replacing arrow symbols with the symbol -1 we will obtain configurations that are not in X. Instead, we develop a scheme for labelling the ribbons which will allow us to erase entire ribbons a few at a time. We give the details of this for the horizontal ribbons only.

We can impose a natural ordering on the horizontal ribbons in each $x \in X$ by calling the first horizontal ribbon lying above the origin the 0th ribbon. For $n \in \mathbb{Z}$ if $n > 0$ ($n < 0$) the nth ribbon is then the nth ribbon above (below) the 0th ribbon. Using this ordering on the ribbons we now construct a new subshift with alphabet

$$\{-1,\uparrow,\nwarrow\}\cup\{(1,t),\cdots,(n_+,t),(\rightarrow,t),(\searrow,t)\colon t\in[0,1]\}$$

and consisting of configurations from X whose horizontal ribbons have been labeled by a number from $[0,1]$.

We then define a map Φ from $X\times[0,1]^{\mathbb{Z}}$ into the new subshift by mapping (x,ω) to $\hat{x}\in\hat{X}$ by labelling the symbols from the nth horizontal ribbon in x by $\omega[n]$. $\Phi(X)=\hat{X}$ then consists of all configurations in the new subshift such that for all $n\in\mathbb{Z}$, the symbols in the nth ribbon all have the same label.

We set $\hat{\mu}=(\mu\times\lambda^{\mathbb{Z}})\circ\Phi^{-1}$. Assuming for the moment that $\hat{\mu}$ is ergodic and invariant we define maps $G_r\colon\hat{X}\to X$ by

$$G_r(\hat{x})_{i,j}=\begin{cases}\hat{x}[(i,j)] & \text{if } x[(i,j)] \text{ does not belong to a horizontal ribbon}\\ \pi(\hat{x}[(i,j)]) & \text{if } \hat{x}[(i,j)]\in(A_+\cup\{\rightarrow,\searrow\})\times[0,r)\\ -1 & \text{if } \hat{x}[(i,j)]\in\{\rightarrow,\searrow\}\times[r,1]\\ \uparrow & \text{if } \hat{x}[(i,j)]\in A_+\times[r,1],\end{cases}$$

where π applied to a labelled element (s,t) in the extended alphabet is defined to be s. The effect of the map is to delete all horizontal ribbons with labels that are less than r leaving -1 for symbols outside vertical ribbons and $\uparrow$ for symbols inside vertical ribbons. As before, we see that the measures μ_r given by $\mu_r=\hat{\mu}\circ G_r^{-1}$ form a $\bar{d}$-continuous family of ergodic measures with $\mu_0=\mu$ and μ_1. The measure μ_1 is an ergodic invariant vertically well supported measure. The configurations in the support of μ_1 do not contain any horizontal ribbons do not contain any negative symbols apart from -1.

Constructing a similar function for vertical ribbons, we can deform the original well supported measure μ into δ_- in a $\overline{d}$-continuous manner. It now remains to prove that $\hat{\mu}$ is ergodic and shift-invariant.

We first argue that $\hat{\mu}$ is a shift invariant measure as follows. Since we now have a shift action on three different spaces: X, $\hat{X}$, and $[0,1]^{\mathbb{Z}}$ we will distinguish between the one and two dimensional actions by writing $S^{\vec{v}}$ for shifts of X and $\hat{X}$ and S^n for shifts of $[0,1]^{\mathbb{Z}}$.

Since Φ is a bi-measurable bijection, we see that sets of the form $\hat{A}=\Phi(A\times R)$, where $A\subset X$ and $R\subset[0,1]^{\mathbb{Z}}$ are measurable, generate the σ-algebra on $\hat{X}$. It thus suffices to show that $\hat{\mu}(S^{\vec{v}}(\hat{A}))=\hat{\mu}(\hat{A})$ for such a set $\hat{A}$, and $\vec{v}\in\mathbb{Z}^2$.

Define

$$A_n=\{x\in A:\ \text{the 0th ribbon in } x \text{ becomes the } n\text{th ribbon in } S^{\vec{v}}(x)\}.$$

Observe that

$$S^{\vec{v}}(\hat{A})=\bigcup_{n\in\mathbb{Z}}\Phi(S^{\vec{v}}(A_n)\times S^{-n}R).$$

Since the sets A_n partition A and μ and $\lambda^{\mathbb{Z}}$ are invariant measures for $S^{\vec{v}}$ and S^n respectively, we see that

$$\begin{aligned}\hat{\mu}(S^{\vec{v}}(\hat{A})) &= \sum_{n\in\mathbb{Z}} \hat{\mu}\circ\Phi(S^{\vec{v}}(A_n)\times S^{-n}R) = \sum_{n\in\mathbb{Z}} \mu(S^{\vec{v}}(A_n))\lambda^{\mathbb{Z}}(S^{-n}R) \\ &= \sum_{n\in\mathbb{Z}} \mu(A_n)\lambda^{\mathbb{Z}}(R) = \mu(A)\lambda^{\mathbb{Z}}(R) = \hat{\mu}(\hat{A}).\end{aligned}$$

Thus, $\hat{\mu}$ is invariant.

To show ergodicity of the measure we show, as in the Bernoulli setting, that any invariant set has measure equal to the square of its measure. Assume that $\hat{A}$ is an invariant subset of $\hat{X}$ with $0 < \hat{\mu}(\hat{A}) = \alpha < 1$. Write $\hat{A} = \Phi(E)$ where E is a measurable subset of $X \times [0,1]^{\mathbb{Z}}$. Then of course we have $\mu \times \lambda^{\mathbb{Z}}(E) = \hat{\mu}(A)$. Write $E_x = \pi_2(E \cap (\{x\} \times [0,1]^{\mathbb{Z}}))$ for the fibre of E over x.

By Fubini's theorem, $\lambda^{\mathbb{Z}}(E_x)$ depends in a measurable way on x and $\int \lambda^{\mathbb{Z}}(E_x)\, d\mu(x) = \alpha$. By the invariance of $\hat{A}$ we have that $E_{S^{\vec{v}}x} = S^{n(x,v)}E_x$, where $n(x,\vec{v})$ measures the relative difference in the numbering of the bands between x and $S^{\vec{v}}x$. It follows that $\lambda^{\mathbb{Z}}(E_{S^{\vec{v}}x}) = \lambda^{\mathbb{Z}}(E_x)$ so that by ergodicity, $\lambda^{\mathbb{Z}}(E_x) = \alpha$ for a.e. x.

Consider cylinder sets of the form

$$C = \bigcup_{W\in\mathcal{W}_n} W \times L_W.$$

where $\mathcal{W}_n$ is the collection of cylinder sets on the central $2n+1 \times 2n+1$ block and each L_W is an open set in $[0,1]^{\mathbb{Z}}$ whose characteristic function depends only on the the central $2n+1$ coordinates. Fix $\epsilon > 0$ and choose C as above with the property that $\mu \times \lambda^{\mathbb{Z}}(C \triangle E) < \epsilon$. Since the L_W are approximations for the sets E_x which all have measure α and $[0,1]^{\mathbb{Z}}$ is a continuous measure space, we can choose the L_W such that $\lambda^{\mathbb{Z}} L_W = \alpha$ for each $W \in \mathcal{W}_n$.

Let $\hat{C} = \Phi(C)$ and $\vec{v}_t = (t,t)$. We will show that as $t \to \infty$ we have $\hat{\mu}(\hat{C} \cap S^{-\vec{v}_t}(\hat{C})) \to \alpha^2$. Since $\hat{\mu}(\hat{A} \cap \hat{S}^{\vec{v}_t}A) = \hat{\mu}(\hat{A} \cap \hat{A}) = \alpha$, this will show that $\alpha = \alpha^2$ as desired.

To this end, for a given t, partition X into subsets D_k^t on which the number of the central ribbon in the $2n+1 \times 2n+1$ square centered at $(0,0)$ differs from the number of the central ribbon in the $2n+1 \times 2n+1$ square centered at (t,t) by exactly k. Write $E^t_{k,W,W'} = D^t_k \cap W \cap S^{-\vec{v}}W'$ and note that these sets form a partition of X as k runs over $\mathbb{Z}$ and W and W' run over $\mathcal{W}_n$. Furthermore,

$$\begin{aligned}\hat{\mu}(\hat{C} \cap S^{-\vec{v}_t}\hat{C}) &= \sum_{W,W'} \hat{\mu}(\Phi(W \times L_W) \cap S^{-\vec{v}_t}(\Phi(W' \times L_{W'}))) \\ &= \sum_{k,W,W'} \hat{\mu}(\Phi(E^t_{k,W,W'} \times (L_W \times S^{-k}L_{W'}))) \\ &= \sum_{k,W,W'} \mu(E^t_{k,W,W'})\lambda^{\mathbb{Z}}(L_W \times S^{-k}L_{W'}).\end{aligned}$$

Since $\lambda^{\mathbb{Z}}$ is a Bernoulli measure, for $k > 2n$ we have $\lambda^{\mathbb{Z}}(L_W \times S^{-k}L_{W'}) = \alpha^2$. On the other hand, $\mu(\cup_{k\le 2n} D^t_k)$ is the probability that there are less than $2n$ ribbons between the lower box centered at the origin and the upper box centered at (t,t), and so for sufficiently large t this quantity is arbitrarily small. It follows that for large t, $\hat{\mu}(\hat{C} \cap S^{-\vec{v}_t}\hat{C})$ approaches α^2 as claimed, completing the proof.

□

5.2. **A one dimensional example.** Here we construct an example of a one-dimensional subshift with the property that its entropy function has a strictly local maximum. Unlike Haydn's example [4] the set of ergodic invariant measures in this example is $\bar{d}$-connected. To define the example, we introduce the function $f(x) = \sqrt{x}(1-x)$ and we use our usual alphabet $A = A_+ \cup A_-$, where as before $n_- < n_+$. The rules for the subshift X are:

(1) legal words consist of alternating blocks of negative symbols and blocks of positive symbols separated by blocks of 0s;
(2) in any contiguous substring containing both positive and negative symbols, the density of neutral symbols (ρ_0) is required to be at least $\max(f(\rho_-), f(\rho_+))$, where ρ_- and ρ_+ are respectively the densities of negative and positive symbols in the substring.

As before we will show:

Proposition 5.5. *The measure μ_- is a strictly local maximum for the entropy function on $\mathcal{M}_e(X,S)$, and $\mathcal{M}_e(X,S)$ is connected in the $\bar{d}$ topology.*

Proof. To show that μ_- is a strictly local maximum of the entropy function, fix $\delta > 0$ and let μ be any ergodic measure with the property that $1-\delta < \rho_- < 1$. Write $p = 1-\rho_-$. We will demonstrate that provided δ is sufficiently small, $h(\mu) < \log n_-$. We will estimate this entropy by covering a large part of the space with words of some length and counting the number of words needed. Let $\epsilon > 0$ be given. Using the ergodic theorem, we see that there is a length R with the property that with probability greater than $1-\epsilon$, the density of non-negative symbols is between $p/2$ and $2p$. Since we have $\rho_0 > (1-2p)\sqrt{\rho_+}$, we have $\rho_+ < (2p/(1-2p))^2 < 5p^2$ for sufficiently small p. We now are able to estimate the number of words of this form as follows:

A frame consists of alternating blocks of +s and −s with 0s between. We can estimate the number of frames by noting that the frame is determined by boundaries between the blocks. Since the total number of + elements is bounded by $5p^2R$, it follows that the number of endpoints of boundaries is at most $20p^2R$. The number of frames is therefore bounded above by

$$\binom{R}{0} + \binom{R}{1} + \ldots + \binom{R}{20p^2R}.$$

For small p, this is bounded above by $2\binom{R}{20p^2R}$.

Given a frame, we can form a word by replacing +s by any of the n_+ positive elements and −s by any of the n_- negative elements. Since there are at most $5p^2R$ positive elements and at most $(1-\frac{p}{2})R$ negative elements, we see that the total number of configurations of this type is bounded by

$$2\binom{R}{20p^2R} n_-{}^{(1-\frac{p}{2})R} n_+^{5p^2R}.$$

We therefore get the estimate that

$$h(\mu) \leq (1-p/2)\log n_- + 5p^2R\log n_+ + 20p^2(-\log p^2).$$

Since the subtracted terms are of order p, while the terms added on are of strictly lower order, we see that for small $p > 0$, there is a strict decrease in entropy. Since the positive measure has strictly greater entropy than the negative measure as before, we see that the negative measure is a strictly local maximum of the entropy function.

To show that $\mathcal{M}_e(X, S)$ is a $\bar{d}$-connected set we argue as before that there is a $\bar{d}$-continuous path connecting an arbitrary measure μ to the measure concentrated on the fixed point $\ldots 0000 \ldots$. The argument is exactly parallel to the case in the proof of Proposition 4.1 and we omit the details. □

References

[1] Bollobas B. Balister, P. and A. Quas. Convexity, random tilings and shifts of finite type. Preprint, 2000.

[2] R. M. Burton and A. Rothstein. Isomorphism theorems in ergodic theory. Oregon State University Technical Report No. 54, 1986.

[3] R. M. Burton and J. E. Steif. Non-uniqueness of measures of maximal entropy for subshifts of finite type. *Ergodic Theory Dynam. Systems*, 14(2):213–235, 1994.

[4] N. Haydn. Multiple measures of maximal entropy and equilibrium states for one-dimensional subshifts.

[5] W. Krieger. On entropy and generators of measure preserving transformations. *Trans. Amer. Math. Soc.*, 149:453–464, 1970.

[6] W. Krieger. On generators in ergodic theory. In *Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 2*, pages 303–308. Canad. Math. Congress, Montreal, Que., 1975.

[7] S. Lightwood. An aperiodic embedding theorem for square filling subshifts of finite type. Preprint, 2000.

[8] E. A. Robinson, Jr. and A. A. Şahin. On the absence of invariant measures with locally maximal entropy for a class of $\mathbb{Z}^d$ shifts of finite type. *Proc. Amer. Math. Soc.*, 127(11):3309–3318, 1999.

[9] E. Arthur Robinson, Jr. and Ayşe A. Şahin. Mixing properties of nearly maximal entropy measures for $\mathbb{Z}^d$ shifts of finite type. *Colloq. Math.*, 84/85(, part 1):43–50, 2000. Dedicated to the memory of Anzelm Iwanik.

[10] E. Arthur Robinson, Jr. and Ayşe A. Şahin. Modeling ergodic, measure preserving actions on $\mathbf{Z}^d$ shifts of finite type. *Monatsh. Math.*, 132(3):237–253, 2001.

[11] D. J. Rudolph. *Fundamentals of measurable dynamics.* Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1990.

Department of Mathematical Sciences, Dunn Hall, University of Memphis, Memphis, TN 38152-3240

E-mail address: quasa@msci.memphis.edu

Department of Mathematics, DePaul University, 2320 North Kenmore St, Chicago, IL 60614

E-mail address: asahin@condor.depaul.edu