PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 127, Number 11, Pages 3309–3318
S 0002-9939(99)05215-6
Article electronically published on May 13, 1999

ON THE ABSENCE OF INVARIANT MEASURES WITH LOCALLY MAXIMAL ENTROPY FOR A CLASS OF $\mathbb{Z}^d$ SHIFTS OF FINITE TYPE

E. ARTHUR ROBINSON, JR. AND AYŞE A. ŞAHİN

(Communicated by Michael Handel)

ABSTRACT. We prove that for a class of $\mathbb{Z}^d$ shifts of finite type, $d > 1$, any invariant measure which is not a measure of maximal entropy can be perturbed a small amount in the weak* topology to an invariant measure of higher entropy. Namely, there are no invariant measures which are strictly local maxima for the entropy function.

1. INTRODUCTION

In this paper we consider invariant measures for a class of $\mathbb{Z}^d$ subshifts of finite type with positive topological entropy. In many cases–although not always–a $\mathbb{Z}^d$ shift of finite type, $d > 1$, has a unique measure of maximal entropy (cf. [1], [2]). In this paper we identify a natural class of $\mathbb{Z}^d$ subshifts that satisfy a strong mixing condition, and we show that for shifts in this class, there are no strictly local maxima for the entropy function. In particular, we show that any invariant measure of submaximal entropy can be perturbed a small amount in the weak or the $\bar{d}$ topology, into a measure with higher entropy.

The organization of the paper is as follows. In Section 2 we introduce the shifts of finite type which we will be studying, and we establish the basic definitions necessary to state our main result. Section 3 contains an entropy lemma necessary for the main proof, and has some independent interest. Finally, Section 4 contains the proof of our main result.

2. DEFINITIONS AND RESULTS

2.1. **The uniform filling property.** Let A be a finite set (the *alphabet*) and let $Y_A = A^{\mathbb{Z}^d}$. We denote the $\vec{m}$th entry of $y \in Y_A$ by $y[\vec{m}]$. Let S be the $\mathbb{Z}^d$ *shift*

$$(S^{\vec{n}}y)[\vec{m}] = y[(\vec{m} + \vec{n})]$$

Received by the editors February 6, 1998.
1991 *Mathematics Subject Classification.* Primary 28D15; Secondary 28D20.
Key words and phrases. Ergodic theory, $\mathbb{Z}^d$ actions, entropy, symbolic dynamics.
The research of the first author was partially supported by the NSF under grant number DMS 9303498.
The research of the second author was partially supported by the NSF under grant number DMS 9501103.

on Y_A. In the product topology, Y_A is compact and metrizable, and the shift S is a continuous $\mathbb{Z}^d$ action. A $\mathbb{Z}^d$ *subshift* (Y, S) is the restriction of S to a closed S-invariant subspace $Y \subset Y_A$.

A *shift of finite type* (abbreviated as SFT) is a subshift (Y, S) consisting of those elements of (Y_A, S) that omit a given finite collection of finite blocks. To make this precise, we need a little terminology. If $R \subseteq \mathbb{Z}^d$, we call $b \in A^R$ a *block* with *shape* R. A block is *finite* if R is finite. The block obtained by restricting $y \in Y_A$ to R is denoted $y[R]$. If $\mathcal{F} = \{f_1, \ldots, f_n\}$ is a finite collection of finite blocks, with shapes $R_1, \ldots, R_n$, we define the *shift of finite type* (SFT) $Y_{\mathcal{F}} = \{y \in Y_A : y[R_j - \vec{n}] \neq f_j$ for any $f_j \in \mathcal{F}\}$. We refer to $\mathcal{F}$ as the set of *forbidden blocks.*

Given $\mathcal{F}$, let $m = \max_j\{\operatorname{diam}(R_j)\}$ (with respect to the box norm on $\mathbb{Z}^d$). Let $s = (m-1)/2$ if m is odd and $s = m/2$ if m is even. We call s the *step-size* of $Y = Y_{\mathcal{F}}$. Without loss of generality, we can assume that every $f \in \mathcal{F}$ satisfies $f \in A^{B_s}$ where $B_s = \{\vec{n} \in \mathbb{Z}^d : ||\vec{n}|| \leq s\}$.

A symbol $\sigma \in A$ is called a *safe symbol* for a SFT (Y, S) (cf. [3]) if $y' \in Y$ whenever y' is obtained from some $y \in Y$ by replacing arbitrary entries with σ. For example the *golden mean shift* $Y_{\mathcal{F}} \subseteq \{0, 1\}^{\mathbb{Z}^2}$ is defined by the rule: *two* 1*s may not be vertically or horizontally adjacent.* Then $Y_{\mathcal{F}}$ has 0 as a safe symbol.

Let us fix a SFT (Y, S). A *configuration* $R = \bigcup_j R_j$ is a finite or countable set of disjoint shapes in $\mathbb{Z}^d$. An *R–specification* is an element $y' \in A^R$ such that, for each j, there is a $y_j \in Y$ with $y'[R_j] = y_j[R_j]$, i.e., each block $y'[R_j]$ occurs in a word of (Y, S). A specification $y' \in A^R$ is said to be *extendable* if there is $y \in Y$ such that $y[R] = y'$. We say that a SFT (Y, S) has *R–specification* if every R–specification is extendable. Note that if (Y, S) has a safe symbol, then it has R–specification for any R with $dist(R_i, R_j) > 0$ for $i \neq j$.

A SFT (Y, S) is said to have the *uniform filling property* with *filling length* ℓ if it has R-specification for all configurations of the form $R = R_1 \cup R_2$, where

$$R_1 = B_m + \vec{n}, \qquad R_2 = B^c_{m+2l} + \vec{n}, \tag{2.1}$$

for $m \in \mathbb{N}$ and $\vec{n} \in \mathbb{Z}^2$. It is easy to see that the uniform filling property implies topological mixing (cf. [1]). If (Y, S) has a safe symbol, then it has the uniform filling property with $\ell = 1$. The existence of a safe symbol is not, however, necessary for the uniform filling property. Consider, for example, the *iceberg model* (cf. [1]):

$$Y_{\mathcal{F}} \subset \{-M, -M+1, \ldots, -1, 1, \ldots, M-1, M\}^{\mathbb{Z}^d}, \quad M \in \mathbb{N},$$

where $\mathcal{F}$ is defined by the rule: *the only numbers of opposite sign allowed to be vertically or horizontally adjacent are* ± 1. It is easy to verify that this SFT has the uniform filling property with $\ell = 2$.

2.2. **Processes, process topologies.** A *process* is an ergodic $\mathbb{Z}^d$ action (X, μ, T) together with a finite labeled measurable partition Q on X. Elements of Q will be labeled by the symbols from an alphabet A. For $a \in A$ we write $Q(x) = a$ if $x \in Q_a \in Q$. We let $Q^n = \bigvee_{\vec{v} \in B_n} T^{\vec{v}} Q$, writing Q^n_a, $a \in A^{B_n}$, for the elements of Q^n.

The (T, Q)-*name* of $x \in X$, denoted $\phi_Q(x)$, is the element $y \in Y_A$ such that $y[\vec{n}] = Q(T^{\vec{n}} x)$. Given a SFT $(Y, S) \subseteq (Y_A, S)$, we say a partition Q is *type* (Y, S) if for μ a.e. $x \in X$, the (T, Q)-name of x satisfies $\phi_Q(x) \in Y$. We also define a map $\phi_Q : \mathcal{M}(X, T) \to \mathcal{M}(Y, S)$, (where $\mathcal{M}(Y, S)$ denotes the S invariant Borel measures on (Y, S)), by $\phi_Q(\mu)(E) = \mu(\{x \in X : Q(x) \in E\})$. Letting P denote the time

$\vec{0}$ partition of (Y, S), if $\nu = \phi_Q(\mu)$, the processes (X, μ, T, Q) and (Y, ν, S, P) are *identical* (cf [6]).

We denote the set of *ergodic joinings* of two ergodic $\mathbb{Z}^d$ actions (X, μ, T) and (Z, γ, U) by $J_e((X, \mu, T), (Z, \gamma, U))$ (cf. [6]). Given partitions Q and R on X and Z, define the partitions $\overline{Q} = Q \times Z$ and $\overline{R} = X \times R$ on $X \times Z$. The $\overline{d}$*-distance* between processes (cf. [6]) is defined by

$$\overline{d}((X, \mu, T, Q), (Z, \gamma, U, R)) = \min \left\{ \lambda(\overline{Q} \triangle \overline{R}) : \lambda \in J_e((X, \mu, T), (Z, \gamma, U)) \right\}$$

where we set $\overline{Q} \triangle \overline{R} = \{(x, z) : \overline{Q}(x, z) \neq \overline{R}(x, z)\} = \{(x, z) : Q(x) \neq R(z)\}$. For $n \geq 1$ the $\overline{d}^n$ metric is defined by:

$$\overline{d}^n((X, \mu, T, Q), (Z, \gamma, U, R)) = \overline{d}((X, \mu, T, Q^n), (Y, \gamma, U, R^n)).$$

For $\nu, \nu' \in \mathcal{M}(Y, S)$ we write $\overline{d}^n(\nu, \nu') = \overline{d}^n\left((Y, \nu, S, P), (Y, \nu', S, P)\right)$. All of these metrics are equivalent (cf. [6]) and they generate the $\overline{d}$ *topology* on $\mathcal{M}(Y, S)$. The classical *weak* topology* on $\mathcal{M}(Y, S)$, as a subset of $C(Y)^*$, is given by the metric

$$d(\nu, \nu') = \sum_{n=1}^{\infty} \frac{1}{2^{n+1}} \sum_{a \in A^{B_n}} |\nu(P_a^n) - \nu'(P_a^n)|.$$

The weak* topology is strictly weaker than $\overline{d}^n$ [6].

2.3. **Entropy.** The *entropy* of a partition Q on (X, μ) is defined

$$H(Q) = \sum_{a \in A} -\mu(Q_a) \log(\mu(Q_a)).$$

For (X, μ, T, Q), the *process entropy* is defined

$$h(X, \mu, T, Q) = \lim_{N \to \infty} \frac{1}{N^d} H\Big(\bigvee_{\vec{v} \in B_N} T^{\vec{v}} Q\Big)$$

(cf. [7]). Process entropy can be computed in terms of the *conditional information* $\mathbf{I}$ of Q given its *past* (cf. [7]) $\bigvee_{\vec{n} \prec 0} T^{\vec{n}} Q$ via the integral

$$h(X, \mu, T, Q) = \int \mathbf{I}\Big(Q | \bigvee_{\vec{n} \prec 0} T^{\vec{n}} Q\Big)\, d\mu \tag{2.2}$$

(cf. [7] or [6]). Here, $\prec$ denotes lexacographic order on $\mathbb{Z}^d$. The *metric entropy* of the action (X, μ, T) is defined

$$h(X, \mu, T) = \sup \left\{ h(X, \mu, T, Q) : Q \text{ a measurable partition} \right\}.$$

For a SFT (Y, S) and $\nu \in \mathcal{M}(Y, S)$, the time $\vec{0}$ partition P is a *generating partition*, and as in the one dimensional case $h(Y, \nu, S) = h(Y, \nu, S, P)$. Let $W_N(Y, S) = \operatorname{card}\{y[B_N] : y \in Y\}$ be the number of distinct blocks of shape B_N that appear in $y \in Y$. The *topological entropy* of (Y, S) is given by

$$h(Y, S) = \lim_{N \to \infty} \frac{1}{N^2} \log(W_N(Y, S)).$$

This limit always exists, but it can be exceedingly difficult to compute, even for seemingly simple $\mathbb{Z}^d$ SFT's with $d > 1$. For example, the exact value is unknown for the golden mean shift (cf. [3]).

Lemma 2.1. *If (Y, S) is a SFT with the uniform filling property, then $h(Y, S) > 0$.*

Proof. The uniform filling property implies positive entropy using the same argument as in [1]. □

2.4. **The entropy function.** We fix a $\mathbb{Z}^d$ SFT (Y,S) and let $h : \mathcal{M}(Y,S) \to \mathbb{R}$, $h(\nu) = h(Y,\nu,S)$. The *Variational Principle* says that $h(Y,S) = \sup\{h(\nu) : \nu \in \mathcal{M}(Y,S)\}$ (cf. [3]). For a SFT h is weak* upper semicontinuous and the maximum is always achieved (cf. [4]) by a measure ν_0 of *maximal entropy.* If ν_0 is unique, then (Y,S) is *intrinsically ergodic.* For $d > 1$ this may not be the case, even if the uniform filling property holds (cf. [3], [1]). The next theorem, our main result, shows that for a broad class of shifts there can be no local maxima for the entropy function other than absolute maxima: measures of maximal entropy.

Theorem 2.2. *Let (Y,S) be a $\mathbb{Z}^d$ shift of finite type with the uniform filling property. Let $\nu \in \mathcal{M}(Y,S)$ be an ergodic measure and suppose $\nu^* \in \mathcal{M}(Y,S)$ is a weakly mixing measure with $h(\nu^*) > h(\nu)$. Given $\delta > 0$ and $n \geq 1$, there exists an ergodic measure $\nu' \in \mathcal{M}(Y,S)$ so that*

$$\overline{d}^n(\nu,\nu') < 2\delta \tag{2.3}$$

and

$$h(\nu') \geq h(\nu) + \delta\big(h(\nu^*) - h(\nu)\big). \tag{2.4}$$

The theorem is stated in terms of the $\overline{d}$ topology, but can also be reformulated in the weak topology.

Corollary 2.3. *Under the same hypotheses as Theorem 2.2, there exists an ergodic measure $\nu' \in \mathcal{M}(Y,S)$ so that $d(\nu,\nu') \leq 2\delta$ and $h(\nu') \geq h(\nu) + \delta\big(h(\nu^*) - h(\nu)\big)$.*

This follows by applying exercise 7.7 from [6].

Note that $\overline{d}(\nu,\nu') < \delta$ implies $|h(\nu) - h(\nu')| \leq -\delta\log(\delta) + \delta\log(|A|)$ [6], so h is $\overline{d}$ continuous, but (2.4) is not contradicted since $h(\nu*) \leq h(Y,S) \leq \log(|A|)$.

The "mixing" hypotheses on (Y,S) in Theorem 2.2 are now discussed. The first hypothesis is that (Y,S) have the uniform filling property. We do not know if the theorem holds for SFTs that do not satisfy this condition. The second hypothesis is the existence of a weakly mixing measure ν^* of relatively high entropy.

If (Y,ν_0,S) is weakly mixing for every maximal entropy measure ν_0, then we say (Y,S) is *intrinsically weakly mixing.* Intrinsic weak mixing is a fairly common property because in many cases measures of maximal entropy are Bernoulli. For example, Burton and Steif [2] state that the golden mean shift has a unique Bernoulli measure of maximal entropy. When (Y,S) is intrinsically weakly mixing, one can substitute $h(Y,S)$ for $h(\nu^*)$ in the statements of Theorem 2.2 and Corollary 2.3. However, ν_0 need not be weak mixing, even if it is the unique measure of maximal entropy for (Y,S), a SFT satisfying the uniform filling condition [2].

Sufficient conditions for intrinsic weak mixing can sometimes be obtained from conditions for intrinsic ergodicity. Markley and Paul [3] consider $\mathbb{Z}^d$ SFTs (Y,S) defined by matrices satisfying certain commutation conditions. They show (Y,S) is intrinsically ergodic if

$$r/|A| > 2d/(1+\sqrt{4d^2+1}), \tag{2.5}$$

where r is the number of safe symbols (a more general version is also given). It is easy to see that if (Y,S) and $(Y \times Y, S \times S)$ are both intrinsically ergodic, then

(Y, S) is intrinsically weak mixing. Applying (2.5), we have that if

$$r/|A| > \sqrt{2d/(1+\sqrt{4d^2+1})},$$

then (Y, S) is intrinsically weak mixing.

3. Proofs

3.1. **An "Abramov Lemma" for partitions.** In this section we prove a lemma that estimates the entropy of a partition constructed as a "skew product" of other partitions.

Lemma 3.1. *Let (X_0, μ_0, T_0) and (Y_i, ν_i, S_i), $i = 1, \dots, k$, be measure preserving $\mathbb{Z}^d$ actions such that*

$$(X, \mu, T) = (X_0, \mu_0, T_0) \times \prod_{i=1}^{k} (Y_i, \nu_i, S_i)$$

is ergodic. For $i = 1, \dots, k$, let P^i be a finite partition of (Y_i, ν_i) with labels in A. Let $Q^0 = \{Q_1^0, \dots, Q_k^0\}$ be a partition of (X_0, μ_0) with labels $\{1, \dots, k\}$. We define a skew product *partition Q of (X, μ) with labels in A by*

$$Q(x, y_1, \dots, y_k) = P^{Q^0(x)}(y_{Q^0(x)}). \tag{3.1}$$

Then

$$\sum_{i=1}^{k} h(Y_i, \nu_i, S_i, P^i)\, \mu_0(Q_i^0) \leq h(X, \mu, T, Q) \tag{3.2}$$

$$\leq \sum_{i=1}^{k} h(Y_i, \nu_i, S_i, P^i) + h(X_0, \mu_0, T_0, Q^0). \tag{3.3}$$

Proof. To prove the first inequality, we let $Y = \prod Y_i$ and $\nu = \prod \nu_i$.

$$\overline{P}^i = X_0 \times (\prod_{j<i} Y_j) \times P^i \times (\prod_{j>i} Y_j),$$

$i = 1, \dots, k$, be partitions of X. Then by (2.2)

$$\begin{aligned} h(X, \mu, T, Q) &= \int \mathbf{I}(Q | \bigvee_{\vec{n} \prec 0} T^{\vec{n}} Q)\, d\mu \tag{3.4} \\ &\geq \int \mathbf{I}(Q | \bigvee_{\vec{n} \prec 0} T_0^{\vec{n}} Q^0 \times \prod_{i=1}^{k} \bigvee_{\vec{n} \prec 0} S_i^{\vec{n}} P^i)\, d(\mu_0 \times \nu) \\ &= \sum_{i=1}^{k} \int_{Q_i^0 \times Y} \mathbf{I}(\overline{P}^i | \bigvee_{\vec{n} \prec 0} T_0^{\vec{n}} Q^0 \times \prod_{i=1}^{k} \bigvee_{\vec{n} \prec 0} S_i^{\vec{n}} P^i)\, d(\mu_0 \times \nu) \\ &= \sum_{i=1}^{k} \int_{Q_i^0 \times Y_i} \mathbf{I}(P^i | \bigvee_{\vec{n} \prec 0} S_i^{\vec{n}} P^i)\, d(\mu_0 \times \nu_i) \\ &= \sum_{i=1}^{k} h(Y_i, \nu_i, S_i, P^i)\, \mu_0(Q_i^0). \end{aligned}$$

The inequalities in (3.4) and (3.3) both follow from

$$Q \subseteq Q^0 \times \prod_{i=1}^{k} P^i. \tag{3.5}$$

□

3.2. **Filling infinitely many collars.** Let (Y, S) be a SFT with step size s and filling distance ℓ. A *proper collar configuration* is a configuration $R = \bigcup_{j \geq 0} R_j$ where

$$R_j = B_{m_j} + \vec{n}_j$$

for $j > 0$, and

$$R_0 = \left(\bigcup_j B_{m_j+2l} + \vec{n}_j \right)^c,$$

where

$$\|\vec{n}_j - \vec{n}_i\| > m_j + m_i + 4l + s \tag{3.6}$$

for $i \neq j$. The *collars* are the sets $C_j = (B_{m_j+2l} + \vec{n}_j) \backslash (B_{m_j} + \vec{n}_j)$.

Lemma 3.2. *Let (Y, S) be a SFT with step size s. Then (Y, S) has the uniform filling property with filling length ℓ if and only if it has R-specification for any proper collar configuration R.*

Proof. Let $R = \bigcup_{j \geq 0} R_j$ be a proper collar configuration and let y' be an R-specification. Then for $j \geq 0$, there exist $y_j \in Y$ such that $y_j[R_j] = y'[R_j]$. For $j > 0$ note that by the uniform filling property there exists $y_j'' \in Y$ such that

$$y_j''[R_j] = y'[R_j] \qquad \text{and} \qquad y_j''[(B_{m_j+2l} + \vec{n}_j)^c] = y_j[(B_{m_j+2l} + \vec{n}_j)^c].$$

Note that $y_j''[R_0] = y_i''[R_0]$ for all $i, j \geq 0$.

Let C_j be the collar centered at $\vec{n}_j$. Then

$$\mathbb{Z}^d = R_0 \cup \bigcup_{j>0} (R_j \cup C_j),$$

where the union is disjoint. Thus we can define $y \in Y_A$ by $y[R_0] = y'[R_0]$, and $y[R_j \cup C_j] = y_j[R_j \cup C_j]$ for $j \geq 1$.

To show $y \in Y$, it suffices to show for any $\vec{v} \in \mathbb{Z}^d$ that $y[B_s + \vec{v}]$ is not forbidden, i.e. is not a translate of a block in $\mathcal{F}$. If $B_s + \vec{v} \subset R_0$, then $y[B_s + \vec{v}] = y_0[B_s + \vec{v}]$, which is not forbidden since $y_0 \in Y$. It follows from (3.6) that for any $\vec{v}$, the shape $B_s + \vec{v}$ can intersect the set $R_j \cup C_j = B_{m_j+2\ell} + \vec{n}_j$ for at most one j. In this case, $y[B_s + \vec{v}] = y_j''[B_s + \vec{v}]$, which is not forbidden since $y_j'' \in Y$. □

3.3. **The proof of Theorem 2.2.** Suppose (Y, S) has stepsize s and filling length ℓ.

Without loss of generality we will assume that there are no safe symbols and that

$$0 < \delta < \frac{1}{2}. \tag{3.7}$$

Step 1: Perturbing entropy into ν. Suppose $\sigma \notin A$. We will create a "faux" safe symbol by constructing a new alphabet $A_\sigma = A \cup \{\sigma\}$. Let (Y_σ, S) be a new SFT with alphabet A_σ that has the same forbidden blocks as (Y, S). This new SFT has σ as a safe symbol. Note that $\mathcal{M}(Y,S) \subset \mathcal{M}(Y_\sigma, S)$, where $\nu \in \mathcal{M}(Y_\sigma, S)$ is in $\mathcal{M}(Y,S)$ if it gives measure zero to every cylinder set with a σ in its defining block. The $\overline{d}$ metric on $\mathcal{M}(Y,S)$ is the same as it inherits as a subset of $\mathcal{M}(Y_\sigma, S)$. Let (X_0, μ_0, T_0) be a weakly mixing $\mathbb{Z}^2$ action. Define $y_\sigma \in Y_\sigma$ by $y_\sigma[\vec{n}] = \sigma$ for all $\vec{n} \in \mathbb{Z}^d$, a fixed point for S. Let δ_σ be the unit point mass at y_σ.

We are going to apply Lemma 3.1 for $k = 3$. Let $X = X_0 \times Y_\sigma \times Y_\sigma \times Y_\sigma$, $T = T_0 \times S \times S \times S$ and $\mu = \mu_0 \times \nu \times \nu^* \times \delta_\sigma$, noting that (X, μ, T) is ergodic. Let $P^i = P$ for $i = 1, 2, 3$.

The construction of Q^0 requires some care. We will use parameters $m, e \in \mathbb{N}$, and $\theta > 0$ which depend on n, δ, ℓ, s, $h(\nu)$ and $h(\nu^*)$. The values will be specified below, but for now we need the following *initial estimate* $0 < 2\ell < e < m - n - s$.

By the Rohlin Lemma (cf. [5]), there exists a *Rohlin tower* E for (X_0, μ_0, T_0) with *base* $D \subset X_0$, in the *shape* B_m, with *error* θ. By this we mean that $T^{\vec{n}}D \cap T^{\vec{m}}D = \emptyset$ for $\vec{n}, \vec{m} \in B_m$, $\vec{n} \neq \vec{m}$, and that $E = T_0^{B_m} D$ satisfies $\mu_0(E) > 1 - \theta$. Let $G = \{\ell, \dots, e-\ell\}^2 \subset B_e$ and $C = B_e \backslash G$. We define Q^0 as follows. Let $Q_2^0 = T_0^G D \subset X_0$, let $Q_3^0 = T_0^C D$ (these are both Rohlin sub-towers of E), and let $Q_1^0 = X_0 \backslash (Q_2^0 \cup Q_3^0)$. Applying Lemma 3.1, we obtain a partition Q of (X, μ, T) such that the collar C is painted with σ, G is painted with high entropy names, and the rest preserves its original labelling.

Since σ is a safe symbol, Q is of type (Y_σ, S). Let $\nu_\sigma = \phi_Q(\mu) \in \mathcal{M}(Y_\sigma, S)$. Then ν_σ is ergodic and

$$h(\nu_\sigma) = h(X, \mu, T, Q). \tag{3.8}$$

Further, by Lemma 3.1

$$\begin{aligned} h(\nu_\sigma) &\geq h(\nu)\mu_0(Q_1^0) + h(\nu^*)\mu_0(Q_2^0) + h(\theta_\sigma)\mu_0(Q_3^0) \\ &= h(\nu)\mu_0(Q_1^0) + h(\nu^*)\mu_0(Q_2^0). \end{aligned} \tag{3.9}$$

Our next goal is to prove

$$h(\nu_\sigma) \geq h(\nu) + \frac{3\delta}{2}(h(\nu^*) - h(\nu)), \tag{3.10}$$

which is a preliminary version of (ii). This will follow from the choice of parameters, which we now specify.

For $0 < r \leq 1$, let $H(r) = -r \log r$, with $H(0) = 0$. Choose real numbers $0 < \tau, \theta < \frac{\delta}{2000}$ so that

$$H(\tau) + \tau \log(|A|) < \frac{\delta}{2}\big(h(\nu*) - h(\nu)\big), \tag{3.11}$$

$$1 + \tau \leq \frac{201}{200}, \tag{3.12}$$

$$1 - 2\tau - \theta \geq \frac{99}{100}. \tag{3.13}$$

Now set $p = \frac{h(\nu*) - h(\nu)}{h(\nu)}$ and pick $e \in \mathbb{N}$ such that

$$\max\left\{\frac{2d\ell}{e}, \frac{2d\ell}{ep}, \frac{2dn}{e}\right\} < \tau. \tag{3.14}$$

Finally, we pick $m \in \mathbb{N}$ such that

$$\left(\frac{200}{201}\right)\left(\frac{999}{1000}\right)\frac{8}{5}\delta \geq \frac{e^d}{m^d} \geq \frac{3}{2}\left(\frac{100}{99}\right)\delta. \tag{3.15}$$

Notice that equations (3.15) and (3.14) imply

$$\frac{2dn}{m} < \frac{\delta}{2000}. \tag{3.16}$$

Now note that since $\mu(Q_1^0) + \mu(Q_2^0) + \mu(Q_3^0) = 1$ we can rewrite equation (3.9) as:

$$h(\nu_\sigma) = h(\nu) + \mu(Q_2^0)\big(h(\nu*) - h(\nu)\big) - \mu(Q_3^0)h(\nu). \tag{3.17}$$

Note that

$$\mu(Q_2^0) \geq (1-\theta)\frac{e^d}{m^d}\left[1 - \frac{2d\ell}{3}\right], \tag{3.18}$$

and

$$\mu(Q_3^0) \leq \frac{2de^{d-1}\ell}{m^d} = \frac{e^d}{m^d}\left[\frac{2d\ell}{e}\right] \tag{3.19}$$

so by equations (3.17) and (3.14) $h(\nu_\sigma)$ is

$$\begin{aligned} &= h(\nu) + \frac{e^d}{m^d}\left[1 - \frac{2d\ell}{e} - \frac{2d\ell}{pe} - \theta\right]\big(h(\nu*) - h(\nu)\big) \\ &\geq h(\nu) + \frac{3}{2}\left(\frac{100}{99}\delta\right)\left[1 - \frac{2d\ell}{e} - \frac{2d\ell}{pe} - \theta\right]\big(h(\nu*) - h(\nu)\big) \\ &\geq h(\nu) + \frac{3}{2}\delta\left(\frac{100}{99}\right)(1 - 2\tau - \theta)\big(h(\nu*) - h(\nu)\big) \end{aligned}$$

which by equation (3.13) is

$$\geq h(\nu) + \frac{3}{2}\delta\big(h(\nu*) - h(\nu)\big). \tag{3.20}$$

Step 2: The $\overline{d}^n$ distance between ν_σ and ν. Since $\nu_\sigma = \phi_Q(\mu)$, we have

$$\overline{d}^n(\nu, \nu_\sigma) = \overline{d}^n\big((X, \mu, T, Q), (Y_\sigma, \nu, S, P)\big).$$

We define an ergodic joining $\rho \in J_e((X, \mu, T), (Y_\sigma, \nu, S))$ by

$$\rho\big((F \times E_1 \times E_2 \times E_3) \times E\big) = \mu\big(F \times (E \cap E_1) \times E_2 \times E_3\big). \tag{3.21}$$

Let $H = \{n, \ldots, m-n\}^d \cap B_{e+n}^c$ and let $I = T^H D$. By the *initial estimate*, I is a Rohlin sub-tower of E.

If $x = (x_0, y_1, y_2, y_3) \in X$ has $x_0 \in I$, then x_0 lies inside a thickness n collar around the inside of the subtower $T_0^{B_m \setminus B_e} D \subset Q_1^0$ and it follows that

$$Q^n(x) = \phi_{Q^n}(x)[\vec{0}] = \phi_Q(x)[B_n] = \phi_P(y_1)[B_n] = P^n(y_1).$$

Thus we have

$$\begin{aligned}\overline{d}^n(\nu,\nu_\sigma) &\le \rho(\overline{Q^n}\triangle\overline{P^n})\\ &\le 1-\sum_{a\in A(P^n)}\rho(I\times P^n_a\times Y_\sigma\times Y_\sigma\times P^n_a)\\ &= 1-\sum_{a\in A(P^n)}\mu_0(I)\nu(P^n_a)\nu^*(Y_\sigma)\theta_\sigma(Y_\sigma)\\ &= 1-\mu_0(I). \end{aligned}\tag{3.22}$$

On the other hand we know that

$$\begin{aligned}\mu_0(I) &\ge (1-\theta)\left[1-\left(\frac{e^d}{m^d}+\frac{2de^{d-1}\ell}{m^d}+\frac{2dm^{d-1}n}{m^d}\right)\right]\\ &=(1-\theta)\left[1-\frac{e^d}{m^d}\left[1+\frac{2d\ell}{e}\right]-\frac{2dn}{m}\right];\end{aligned}$$

therefore equation (3.22) is

$$\le \frac{e^d}{m^d}(1+\frac{2d\ell}{e})+\frac{2dn}{m}+\theta$$

so by equations (3.14), (3.12), and (3.16) and by our choice of θ we have

$$\overline{d}^n(\nu,\nu_\sigma)\le\frac{e^d}{m^d}\left(\frac{201}{200}\right)+\frac{\delta}{1000}.$$

Putting this together with equation (3.15) we now have

$$\overline{d}^n(\nu,\nu_\sigma)<\frac{8}{5}\delta.$$

Step 3: Erasing σ. Our task now is to erase σ to obtain $\nu'\in\mathcal{M}(Y,S)$ that satisfies (i), (ii) and (iii). Recall that $C=B_e\backslash G$. Suppose $V\subset\mathbb{Z}^2$ is finite or countable and let

$$F=\bigcup_{\vec{n}\in V}C+\vec{n}\tag{3.23}$$

be the collars for a proper collar configuration F^c. Given $y\in Y_\sigma$, let $F_y=\{\vec{n}\in\mathbb{Z}^2 : y[\vec{n}]=\sigma\}$. Let $Y'_\sigma\subset Y_\sigma$ be the set of all $y\in Y_\sigma$ such that F_y is a proper collar configuration. Then for $y\in Y'_\sigma$ $y[F^c_y]$ is a proper collar specification and by Lemma 3.2, $y[F^c_y]$ has an extension to $y'\in Y$.

The extension y' may not be unique, but we pick a specific extension of $y[F^c_y]$ as follows. If we consider one of the collars $C+\vec{n}$ from F_y, the "legal" possibilities for filling $C+\vec{n}$ depend only on $y[\{-s,\dots,e+s\}^2+\vec{n}]$, called the *filling neighborhood.* We call a choice of $y'[C+\vec{n}]$ a *legal filling.* Up to translation, there are only finitely many filling neighborhoods, and for each of them, only finitely many legal fillings. For each collar, we choose the lexicographically minimal legal filling to define y'. Writing $y'=\psi(y)$, we have defined $\psi : Y'_\sigma\to Y$, satisfying $\psi(y)[\vec{n}]=y[\vec{n}]$ if and only if $y[\vec{n}]\ne\sigma$. It is not hard to see that ψ is a sliding block code.

For μ almost every $x\in X$ it follows from (3.8) that $\phi_Q(x)\in Y'_\sigma$, which is to say that for μ a.e. x, the symbols σ in $\phi_Q(x)$ lie in a proper collar configuration. Thus $\nu_\sigma=\phi_Q(\mu)\in\mathcal{M}(Y'_\sigma,S)$. Let $\nu'=\psi(\nu_\sigma)\in\mathcal{M}(Y,S)$. We claim that ν' satisifes (i), (ii) and (iii).

Since (Y, ν', S) is a factor of $(Y_\sigma, \nu_\sigma, S)$, and ν_σ is an ergodic, we have that ν' is ergodic, which proves (i).

By equation (3.14)

$$\bar{d}^n(\nu_\sigma, \nu') \le \mu_0(T^C D) < \frac{2de^{d-1}\ell}{m^d} = \frac{e^d}{m^d}\left\lceil\frac{2dl}{e}\right\rceil < \frac{2dl}{e} < \tau \tag{3.24}$$

so we have (iii),

$$\bar{d}^n(\nu, \nu') \le \bar{d}^n(\nu, \nu_\sigma) + \bar{d}^n(\nu_\sigma, \nu') < \frac{8}{5}\delta + \frac{\delta}{2000} < 2\delta.$$

By the continuity of h in $\bar{d}^n$, (3.24) implies

$$h(\nu') \ge h(\nu_\sigma) - (H(\tau) + \tau \log_2(|A|)). \tag{3.25}$$

Putting together (3.20),and (3.11), we have

$$h(\nu') \ge h(\nu) + \delta\big(h(\nu^*) - h(\nu)\big)$$

which is (ii). □

References

[1] 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. MR **95f:**28023

[2] R. M. Burton and J. E. Steif. Some 2-d symbolic dynamical systems: entropy and mixing. In *Ergodic theory of $\mathbb{Z}^d$ actions (Warwick, 1993–1994)*, volume 228 of *London Math. Soc. Lecture Note Ser.*, pages 297–305. Cambridge Univ. Press, Cambridge, 1996. MR **97e:**58066

[3] N. G. Markley and M. E. Paul. Maximal measures and entropy for $\mathbb{Z}^\nu$ subshifts of finite type. In *Classical mechanics and dynamical systems (Medford, Mass., 1979)*, volume 70 of *Lecture Notes in Pure and Appl. Math., 70*, pages 135–157. Dekker, New York, 1981. MR **83c:**54059

[4] M. Misiurewicz. A short proof of the variational principle for a F_+^N action on a compact space. In *International Conference on Dynamical Systems in Mathematical Physics (Rennes, 1975)*, pages 147–157. Astérisque, No. 40. Soc. Math. France, Paris, 1976. MR **56:**3250

[5] D. S. Ornstein and B. Weiss. Entropy and isomorphism theorems for actions of amenable groups. *J. Analyse Math.*, 48:1–141, 1987. MR **88j:**28014

[6] D. J. Rudolph. *Fundamentals of measurable dynamics*. Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1990. MR **92e:**28006

[7] K. Schmidt. *Dynamical systems of algebraic origin*, volume 128 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 1995. MR **97c:**28041

Department of Mathematics, George Washington University, Washington, DC 20052
E-mail address: robinson@math.gwu.edu

Department of Mathematics, North Dakota State University, Fargo, North Dakota 58105
E-mail address: sahin@plains.nodak.edu