Higher-dimensional subshifts of finite type, factor maps and measures of maximal entropy

Ronald Meester   Jeffrey E. Steif*

University of Utrecht
Mathematical Institute
P.O. Box 80.010
3508 TA Utrecht
The Netherlands
meester@math.uu.nl

School of Mathematics
Georgia Institute of Technology
Atlanta, Georgia
30332-0160
USA
steif@math.gatech.edu

Abstract

We investigate factor maps of higher-dimensional subshifts of finite type. In particular, we are interested in how the number of ergodic measures of maximal entropy behaves under such factor maps. We show that this number is preserved under almost invertible maps, but not in general under finite to one factor maps. One of our tools, which is of independent interest, is a higher-dimensional characterization of entropy-preserving factor maps that extends the well known one-dimensional characterization result.

1 Introduction

In this paper we discuss some aspects of higher-dimensional subshifts of finite type. The book of Lind and Marcus ([4]) is an excellent introduction to the theory of one-dimensional symbolic dynamics. It turns out however, as is well known, that the higher-dimensional theory is different from the one-dimensional theory. In higher dimensions, new concepts arise that do not have one-dimensional analogues. In addition, many one-dimensional results are simply not true in higher dimensions. In this paper we address issues of both types. An example of a new question without a one-dimensional analogue is the behaviour of the number of ergodic measures of maximal entropy under various types of factor maps. For an example of the second type, we show that the one-dimensional characterization of entropy preserving factor maps (which says that under an obvious irreducibility condition, entropy preservation is equivalent to the map being finite to one) must be replaced by something different in higher dimensions. We also mention here the recent paper [6] which contains a number of interesting results concerning higher-dimensional subshifts of finite type.

Let us start with a description of the setup. Throughout, X denotes a d-dimensional **shift space** with finite alphabet $\mathcal{A}(X)$, that is, X is a closed (in the product topology) translation invariant subset of $\{\mathcal{A}(X)\}^{\mathbf{Z}^d}$. A **pattern** in X is the restriction of a

*Research supported by a grant from the Swedish National Science Foundation

configuration in X to a *finite* subset of $\mathbf{Z}^d$. The restriction of a configuration x to a finite set $A \subseteq \mathbf{Z}^d$ is denoted by $\pi_A(x)$. We write B_n for the box $\{-n, \ldots, n\}^d$ and $\pi_n(x)$ for $\pi_{B_n}(x)$. The collection of all patterns $\{\pi_n(x) : x \in X\}$ is denoted by $\mathcal{B}_n(X)$.

For a countable set of forbidden patterns $\mathcal{F} = \{F_1, F_2, \ldots\}$, we write $S_{\mathcal{F}}$ for the shift space consisting of those configurations that do not contain any of the patterns in $\mathcal{F}$. We can and often will assume that for all i, F_i is a pattern on a box. The **size** of a pattern on B_n is defined to be n. A shift space X is called a **subshift of finite type** (SOFT) if X can be written as $X = S_{\mathcal{F}}$, where $\mathcal{F}$ consists of only finitely many patterns.

Consider two shift spaces X and Y. Let, for some number α, Φ be a map $\Phi : \mathcal{B}_\alpha(X) \to \mathcal{A}(Y)$. We call Φ a **block map**. We can use this block map to define a map $\phi : X \to Y$ by putting

$$\phi(x)_z = \Phi(\pi_\alpha(T_z(x)))$$

where T_z denotes translation by the vector z. Clearly ϕ commutes with shifts. ϕ is called a **factor map** from X to Y. (Clearly ϕ is continuous in the product topology and it is well known that all shift commuting continuous maps from X to Y are of this form.) The smallest α for which a given factor map ϕ arises in this way is called the **range** of ϕ. All factor maps in this paper are assumed to be *onto*, i.e., $Y = \phi(X)$. An invertible factor map is called a **conjugacy**, and if there exists a conjugacy between X and Y, then X and Y are called **conjugate**. We will sometimes abuse notation and write $\phi(x)$ where $x \in \mathcal{B}_n(X)$ for some n, that is, we view ϕ as a mapping $\phi : \mathcal{B}_n(X) \to \mathcal{B}_{n-\alpha}(Y)$.

If X is a shift space, let $h(X)$ denote the **topological entropy** of X and if μ is a translation invariant measure on X, let $h(\mu)$ denote the **measure-theoretical entropy** of μ. See [9] for precise definitions. For a finite partition $\mathcal{P}$ of a probability space (X, μ), we let $h_\mu(\mathcal{P})$ be its **entropy** defined to be $-\sum_i \mu(P_i) \log \mu(P_i)$ where the P_i's denote the atoms of $\mathcal{P}$. The set of translation invariant probability measures on X is denoted by $\mathcal{M}(X)$. The *variational principle* (see [5]) states that

$$h(X) = \sup_{\mu \in \mathcal{M}(X)} h(\mu).$$

Moreover, the supremum is achieved at some measure: such a measure is called a **measure of maximal entropy**. Let $\mathcal{M}_{max}$ denote the set of measures of maximal entropy. In one dimension, an irreducible SOFT (precise definitions follow later) has a unique measure of maximal entropy, the so-called Parry measure. In higher dimensions this is no longer true. Burton and Steif ([1]) gave examples of strongly irreducible SOFT's with multiple ergodic measures of maximal entropy. It is well known that the set of extremal elements in $\mathcal{M}_{max}$ is exactly the set of ergodic measures of maximal entropy. This implies that $\mathcal{M}_{max}$ is a face in the simplex $\mathcal{M}$ and hence is also a simplex. Cardinality is denoted by $|\cdot|$ throughout.

This paper is organized as follows: Section 2 contains all the non-standard technical definitions, Section 3 contains our main results, and the last two sections are devoted to the proofs of the results.

2 Further definitions and preliminaries

In this section we give some technical definitions.

Definition 2.1 *A shift space X is called* **strongly irreducible** *(s.i.) if there is an $s \geq 0$ such that whenever we have two (finite) patterns η_1 on $A_1 \subset \mathbf{Z}^d$ and η_2 on $A_2 \subset \mathbf{Z}^d$ of X and the distance between A_1 and A_2 is greater than s, then there is an $\eta \in X$ that is an extension of both η_1 and η_2.*

We call the smallest such s with the above property the **separation distance** of X.

Definition 2.2 *X is called* **weakly irreducible** *if for every η_1 and η_2 as above, and every $z \in \mathbf{Z}^d$, there is a $u \in \mathbf{Z}^d$ on the halfline through z and the origin, and an η, such that η extends both η_1 and $T_u(\eta_2)$.*

Remark 1 The definition of weak irreducibility in [1] is not satisfactory: the one-dimensional SOFT with only one disallowed patterns $(1, 0)$ is weakly irreducible according to their definition. But clearly this is not desirable, since all elements of this SOFT are of the form $\ldots 0000011111 \ldots$ In our definition this SOFT is not weakly irreducible.

Remark 2 If X is s.i. and $\phi : X \to Y$ is a (onto) factor map, then it is easy to see that Y is also s.i.

Definition 2.3 *A factor map $\phi : X \to Y$ is said to be* **finite to one** *if every $y \in Y$ has only finitely many pre-images.*

For the definition of *almost invertibility* we need the concept of *transitive points*. In one dimension, a point $x \in X$ is called transitive if in x we see all patterns of X in both directions. There are several ways to extend this idea to higher dimensions. We choose for maybe the strongest possible definition, in order to make the forthcoming Theorem 3.4 as strong as possible.

Definition 2.4 *A point x in a shift space X is called* **transitive** *if for all half-lines ℓ starting at the origin with rational direction, and all $n \geq 0$, the set*

$$\left\{ \pi_n(T_z(x)) \ : z \in \mathbf{Z}^d \cap \ell \right\}$$

contains all patterns of X of size n.

Observe that this definition is stronger than requiring that the $\mathbf{Z}^d$-orbit of x is dense in X, which is one of the usual definitions.

Definition 2.5 *A factor map $\phi : X \to Y$ is called* **almost invertible** *if every transitive point in Y has only one pre-image.*

In one dimension, an almost invertible factor map from an (weakly) irreducible SOFT is necessarily finite to one (Proposition 9.2.2 in [4]). This is no longer true in higher dimensions as the following example shows.

Example Consider, in one dimension, a factor map $\phi : X \to Y$ that is almost invertible but not invertible and where X is s.i. (Such an example can be given by sending the golden mean shift to the even shift (Examples 1.2.3 and 1.2.4 in [4]) induced by the block map $00 \to 1, 01 \to 0, 10 \to 0$.) Next, define a two-dimensional shift space X' by the requirement that $x \in X'$ if for all k, $(x_{(z,k)}, z \in \mathbf{Z})$ is an element of X. Y' is defined similarly. Define a factor map $\phi' : X' \to Y'$ by just applying ϕ row by row. It is clear that ϕ' is not finite to one. On the other hand, if $y' \in Y'$ is transitive, the restriction of y' to a horizontal line is transitive in Y, and therefore has only one pre-image in X. It follows that y' has only one pre-image and we conclude that ϕ' is almost invertible.

3 Results

We first discuss two known basic facts concerning conjugacies between SOFT's. They are not related to our results but we give them for the sake of the reader. Theorem 2.1.10 in [4] shows that conjugacies preserve SOFTness in one dimension. Their proof does not extend to higher dimensions. The following extension of this result is due to Klaus Schmidt and reported in [7].

Theorem 3.1 *Let X and Y be two d-dimensional shift spaces, and suppose that X is a SOFT and that X and Y are conjugate. Then Y is also a SOFT.*

A **nearest neighbor** SOFT is a SOFT all whose forbidden patterns are between nearest neighbors, that is, all forbidden patterns are on sets of the form $\{x, x + e_i\}$ for $x \in \mathbf{Z}^d$ and where the e_i's denote the unit vectors. The following result is proved in [10].

Theorem 3.2 *Every SOFT is conjugate to a nearest neighbor SOFT.*

We now consider measures of maximal entropy. In particular, we are interested in what happens with the number of ergodic measures of maximal entropy under factor maps. Without any assumption, there is no hope for an interesting result, as indicated by the following examples.

Example Let, for positive integers N and M, $X_{N,M}$ be the following nearest neighbor SOFT. The alphabet $\mathcal{A}(X_{N,M})$ is the set $\{-N, \ldots, -1, 1, \ldots, M\}$. The forbidden patterns are all neighboring pairs of the form (a, b) with $ab \leq -2$, i.e. a positive number cannot sit next to a negative number unless they both have absolute value 1. These SOFT's were studied in [1] and [2]. In [1], it was shown that when $N = M$ and M is sufficiently large, then X has exactly two ergodic measures of maximal entropy which we call μ^+ and μ^-. If we have a configuration chosen according to μ^+, flipping all coordinates (i.e., multiplying them by minus one) leads to a choice distributed according to μ^-. In [2], it was shown that if $N = M - 1$ and M is sufficiently large, then $X_{N,M}$ has only one measure of maximal entropy. We can now define a factor map $\phi : X_{M,M+1} \to X_{M,M}$ simply by changing all the $M + 1$'s to M. When M is large enough, our previous comments imply that this is a factor map that maps a space with a unique measure of maximal entropy onto a space with two ergodic measures of maximal entropy. On the other hand, mapping $X_{M,M}$ onto the full M-shift by taking absolute values coordinatewise leads, for large M, to a factor map that maps a space with two ergodic measures of maximal entropy onto a space with a unique measure of maximal entropy.

In both of the last two mappings, the image shift space has strictly lower entropy than the domain space and so let us see what happens when entropy is preserved. We shall write $\phi\mu$ for the push-forward of the measure μ under ϕ, that is, $\phi\mu(A) = \mu(\phi^{-1}A)$, for all Borel sets A.

Theorem 3.3 *Let X and Y be shift spaces with $h(X) = h(Y)$ and let $\phi : X \to Y$ be a factor map. Then*
(i) Every $\mu \in \mathcal{M}_{max}(Y)$ is the push-forward of a measure ν in $\mathcal{M}_{max}(X)$, i.e. $\mu = \phi\nu$ for some $\nu \in \mathcal{M}_{max}(X)$;
(ii) If X is a strongly irreducible SOFT and $\mu \in \mathcal{M}(X)$, then $h(\phi\mu) = h(\mu)$. In particular, if $\mu \in \mathcal{M}_{max}(X)$, then $\phi\mu \in \mathcal{M}_{max}(Y)$ and so ϕ takes $\mathcal{M}_{max}(X)$ to $\mathcal{M}_{max}(Y)$ and is surjective by (i).

The question of whether the induced mapping from $\mathcal{M}_{max}(X)$ to $\mathcal{M}_{max}(Y)$ when $h(X) = h(Y)$ is injective will arise in Theorems 3.4 and 3.5. We note here the obvious fact that if ϕ is a conjugacy from X to Y, then the induced mapping ϕ from $\mathcal{M}_{max}(X)$ to $\mathcal{M}_{max}(Y)$ is an isomophism of simplices; in particular, the number of ergodic measures of maximal entropy are the same.

Remark 3 Note that all statements in Theorem 3.3 are false without the entropy condition. The first example preceding Theorem 3.3 shows that (i) cannot generally be true, since we can map from a space with a unique measure of maximal entropy onto a space with two ergodic measures of maximal entropy. The second example preceding Theorem 3.3 shows that (ii) is false without the entropy condition since it is not hard to see that the image of μ^+ is not i.i.d. uniform, the unique measure of maximal entropy on the full M–shift. (In fact, it is not even i.i.d.)

Next we investigate what happens when the factor map is not quite invertible. We start with almost invertible.

Theorem 3.4 *Let X be a strongly irreducible SOFT and let $\phi : X \to Y$ be almost invertible. Then $h(X) = h(Y)$ and the induced mapping $\phi : \mathcal{M}_{max}(X) \to \mathcal{M}_{max}(Y)$ is bijective. In particular, since this mapping is convex and continuous, it is an isomorphism of these simplices and hence the number of ergodic measures of maximal entropy for the two systems are the same.*

Note that we do not assume that Y is a SOFT.

Another relaxation of invertibility is to require the factor map to be *finite to one.* The following theorem tells us that as far as measures of maximal entropy are concerned, finite to one is very different from almost invertibility.

Theorem 3.5 *There exists a strongly irreducible SOFT X and a shift space Y with the same entropy, together with a two to one factor map ϕ from X onto Y such that X has two ergodic measures of maximal entropy and Y has a unique measure of maximal entropy.*

In our proofs of the above results we shall need a higher-dimensional characterization of entropy preserving factor maps. For one-dimensional irreducible SOFT's, preservation of entropy under factor maps is characterized by the requirement that the factor map is finite to one (Theorem 8.1.16 in [4]). This is no longer true in higher dimensions. For instance, one can map the full 2–dimensional 2-shift onto itself via the block map Φ of range 1 given by

$$\Phi((x_{ij})_{i,j\in\{-1,0,1\}}) = x_{00} + x_{10} (\text{mod } 2).$$

(Note that this map operates row by row.) Clearly this map preserves entropy but it is not finite to one.

Remark 4 It follows from Theorem 3.3 that a mapping between spaces of equal entropy can never be from a space with a unique measure of maximal entropy to a space with multiple ergodic measures of maximal entropy. The reader should compare this with Theorem 3.5.

For a factor map ϕ from X onto Y, a **diamond** of ϕ is a set of two elements $x \neq x'$ in X which differ in only finitely many coordinates and for which $\phi(x) = \phi(x')$. (The name "diamond" comes from the one-dimensional graphical representation of these objects.) The proper analogue of Theorem 8.1.16 in [4] in higher dimensions is given by the following theorem. We remark that the equivalence of (a) and (b) follows from Theorem 3.2 in [6] together with the fact that a strongly irreducible SOFT is automatically entropy minimal; see the forthcoming Lemma 4.1. We will need (in particular) characterization (c) below, and since the proof of the equivalence of (c) with the other statements depends on the proof of (a) $\Leftrightarrow$ (b), we include a full proof of the theorem.

Theorem 3.6 *Let X be a strongly irreducible SOFT and let $\phi : X \to Y$ be a factor map based on a block map $\Phi : \mathcal{B}_\alpha(X) \to \mathcal{A}(Y)$. Then the following six statements are equivalent:*

(a) $h(X) > h(Y)$;

(b) ϕ has a diamond;

(c) There exist $y \in Y$, a positive constant k_1 and a constant $c_1 > 1$ such that for infinitely many n we have

$$|\{x \in \mathcal{B}_n(X) \,:\, \phi(x) = \pi_{n-\alpha}(y)\}| \geq k_1 c_1^{(2n+1)^d};$$

(d) There exist $y \in Y$, a positive constant k_2 and a constant $c_2 > 1$ such that for infinitely many n we have

$$\left|\pi_n(\phi^{-1}(y))\right| \geq k_2 c_2^{(2n+1)^d};$$

(e) There exist $y \in Y$, a positive constant k_3 and a constant $c_3 > 1$ such that for all $n \geq 0$ we have

$$|\{x \in \mathcal{B}_n(X) \,:\, \phi(x) = \pi_{n-\alpha}(y)\}| \geq k_3 c_3^{(2n+1)^d};$$

(f) There exist $y \in Y$, a positive constant k_4 and a constant $c_4 > 1$ such that for all $n \geq 0$ we have

$$\left|\pi_n(\phi^{-1}(y))\right| \geq k_4 c_4^{(2n+1)^d}.$$

Remark 5 The proof readily shows that the implication from (a) to (b) above is true for any SOFT X.

4 Factor maps and entropy: Proofs

In this section, we prove Theorems 3.6. We shall need the following lemma.

Lemma 4.1 *Let X be a strongly irreducible SOFT, and consider a pattern p that occurs in X. If we add this pattern to the list of forbidden patterns, obtaining a new SOFT X', then $h(X') < h(X)$.*

Proof Our proof uses measures, which might appear a little strange since the statement is purely topological. Let μ be a measure of maximal entropy for X'. If $h(X) = h(X')$ then it would follow that μ is also a measure of maximal entropy for X. Now every measure of maximal entropy has so called uniform conditional probabilities (see [2], Proposition 1.20), that is, the conditional distribution of patterns on a finite set A given any configuration η on the outside of A is uniform over all compatible configurations on A that extend η.

Let the size of p be M. Next let $N > M + s$ (where s is the separation distance of X) and consider an allowed configuration on the external boundary ∂B_N of B_N. Since X is s.i., the extra forbidden pattern p, when placed in B_M, is compatible with this boundary condition, and by the property of uniform conditional probabilities, there is positive μ probability to see p in B_M. But this is a contradiction since μ concentrates on X' in which p does not occur. $\square$

Proof of Theorem 3.6

(a) $\Rightarrow$ (b): The factor map ϕ is based on a block map $\Phi : \mathcal{B}_\alpha(X) \to \mathcal{A}(Y)$ for some α. We choose m so large that $m > 2\alpha, 2\ell$ where ℓ is the size of the largest forbidden pattern in X. Observe that ϕ maps patterns of size $n+m$ onto patterns of size $n+m-\alpha$. From the definition of entropy we have that

$$|\mathcal{B}_{n+m}(X)| = e^{(2n+1)^d(h(X)+o(1))}, \tag{1}$$

and

$$|\mathcal{B}_{n+m-\alpha}(Y)| = e^{(2n+1)^d(h(Y)+o(1))}, \tag{2}$$

as $n \to \infty$. Denote the set of patterns of X in the annulus $B_{n+m} \backslash B_n$ by $\mathcal{E}(n,m)$. It is clear that the cardinality of $\mathcal{E}(n,m)$ is bounded above by $e^{C(2n+1)^{d-1}}$, for some suitable positive constant C (which depends on m). For $\eta \in \mathcal{E}(n,m)$, we denote by $\mathcal{P}(\eta)$ the set of all extensions of η in B_{n+m} that occur in X. Let $\epsilon > 0$ such that $h(X) > h(Y) + 2\epsilon$. From (1) it follows that for n large enough there exists a $\eta_0 \in \mathcal{E}(n,m)$ such that

$$|\mathcal{P}(\eta_0)| \geq e^{(2n+1)^d(h(X)-\epsilon)-C(2n+1)^{d-1}}.$$

From (2) we see that $\phi(\mathcal{P}(\eta_0))$ contains for large n at most $\exp\Big((2n+1)^d(h(Y)+\epsilon)\Big)$ elements, it follows (since $h(X) > h(Y) + 2\epsilon$) that for n large enough, there exists at least one element y_0 in $\mathcal{B}_{n+m-\alpha}(Y)$ which has two pre-images in $\mathcal{P}(\eta_0)$. We denote these two pre-images by $z(1)$ and $z(2)$. (Since $z(1)$ and $z(2)$ agree on $B_{n+m} \backslash B_n$ and $m > 2\ell$, we can extend $z(1)$ and $z(2)$ so that they are also equal outside of B_{n+m}. Since also $m > 2\alpha$, it follows that the two extensions have the same image under ϕ and therefore form a diamond.

(b) $\Rightarrow$ (a): Assume there is a diamond $\{x(1), x(2)\}$ and assume that $x(1)$ and $x(2)$ differ inside the box B_n only. Let m be large (we shall see later how large), and construct a new shift space $X^{(m)}$ as follows. The alphabet of $X^{(m)}$ consists of all elements of $\mathcal{B}_m(X)$. An element $\eta \in \{\mathcal{B}_m(X)\}^{\mathbf{Z}^d}$ is in $X^{(m)}$ if the configuration obtained by centering each η_z at location $(2m+1)z$ is an element of X. There is a natural one-to-one correspondence (which is not a conjugacy!) between X and $X^{(m)}$: cutting a configuration of X into appropriate disjoint translates of B_m leads to an element in $X^{(m)}$ and this process is reversible. It is easy to see that $X^{(m)}$ is a s.i. SOFT and that its entropy is equal to $(2m+1)^d h(X)$. A similar construction leads to the shift space $Y^{(m)}$, whose entropy is equal to $(2m+1)^d h(Y)$.

Next we define a second SOFT $X^{(m)*}$ which is obtained from $X^{(m)}$ by disallowing those elements of the alphabet $\mathcal{B}_m(X)$ that have a copy of $\pi_n(x(1))$ in the middle. It follows from Lemma 4.1 that $h(X^{(m)*}) < h(X^{(m)})$. The map ϕ induces a factor map $\phi^{(m)}$ from $X^{(m)}$ to $Y^{(m)}$ and a map $\phi^{(m)*}$ from $X^{(m)*}$ to $Y^{(m)}$ which satisfies all the conditions of a factor map except perhaps being surjective. We claim that this latter map is onto. To see this, just note that if $y^{(m)} \in Y^{(m)}$ has a pre-image $x^{(m)} \in X^{(m)} \backslash X^{(m)*}$, then $y^{(m)}$ also has a pre-image in $X^{(m)*}$ which is obtained by replacing all "middle copies" of $\pi_n(x(1))$ in $x^{(m)}$ by a copy of $\pi_n(x(2))$, provided m is sufficiently large. This leads to the inequality $h(X^{(m)*}) \geq h(Y^{(m)})$. Putting everything together we obtain

$$\begin{aligned} h(X) &= \frac{1}{(2m+1)^d} h(X^{(m)}) > \frac{1}{(2m+1)^d} h(X^{(m)*}) \\ &\geq \frac{1}{(2M+1)^d} h(Y^{(m)}) = h(Y), \end{aligned}$$

which is what we wanted to prove.

(b) $\Rightarrow$ (f): Assume there is a diamond $\{x(1), x(2)\}$ and assume that they differ inside the box B_n only. Let m as above be such that $m > 2\alpha, 2\ell$, where ℓ denotes the size of the largest forbidden pattern in X. Consider a regular rectangular grid of translates of B_{n+m}, where any two such boxes are separated by a distance s, the separation distance of the SOFT. We can 'fill' each of these boxes by the appropriate translate of $\pi_{n+m}(x(1))$. By strong irreducibility and compactness, we can extend this configuration to a configuration $x_0 \in X$. Define $y_0 = \phi(x_0)$. Next, we want to replace any of the patterns $\pi_{n+m}(x(1))$ in one of the boxes of the grid by $\pi_{n+m}(x(2))$. We claim that this can be done, in that the new configuration is still in X. To see this, note that $x(2) \in X$ and therefore contains no forbidden pattern. Hence, the only possibility for a forbidden pattern to be created by the replacement of $\pi_{n+m}(x(1))$ by $\pi_{n+m}(x(2))$ is that this forbidden pattern intersects both the translate of $\pi_{n+m}(x(2))$ and the complement of that translate. But since $m > 2\ell$ and $x(1)$ and $x(2)$ agree on $B_{n+m} \backslash B_n$, this is impossible. It follows by construction, using that $m > 2\alpha$, that the image under ϕ has not changed by this replacement. Since we can do this in any of the boxes in the grid, it follows straightforwardly that y_0 satisfies the requirement in (f).

(f) $\Rightarrow$ (d): This is obvious.

(d) $\Rightarrow$ (c): This is obvious.

(f) $\Rightarrow$ (e): This is obvious.

(e) $\Rightarrow$ (c): This is obvious.

(c) $\Rightarrow$ (b): (This is similar to the proof that (a) $\Rightarrow$ (b)) Choose m so large that $m > 2\alpha, 2\ell$. Since the number of configurations in $B_{n+m} \backslash B_n$ is of order $c^{(2n+1)^{d-1}}$ (as m is fixed), there exists η_0 defined on $B_{n+m} \backslash B_n$ such that

$$\{x \in B_{n+m}(X) : x = \eta_0 \text{ on } B_{n+m} \backslash B_n, \phi(\pi_n(x)) = \pi_{n-\alpha}(y)\} \geq k_1 c_1^{(2n+1)^d}.$$

In particular, there exist $x_1 \neq x_2$ in the above set. Since $m > 2\ell$, we can extend x_1 and x_2 to $\tilde{x}_1$ and $\tilde{x}_2$ in X such that they agree outside of B_{n+m} (and hence outside of B_n). Since $m > 2\alpha$, $\phi(\tilde{x}_1) = \phi(\tilde{x}_2)$ yielding a diamond. $\square$

5 Factor maps and measures of maximal entropy: Proofs

For the proof of Theorem 3.3 we need the following lemma which comes from [3]. In fact, there is a small detail missing in their proof. When applying the Hahn-Banach Theorem, one should require that the extension also has norm one so that one can then use the fact (see [8], p. 116) that an operator of norm 1 on $C(X)$ which sends 1 to 1 is a *positive* operator. It is then necessarily given by integration against some probability measure.

Lemma 5.1 *Let X and Y be shift spaces and let $\phi : X \to Y$ be a factor map. Let μ be a probability measure on Y. Then there exists at least one measure ν on X such that μ is the push-forward of ν under ϕ, i.e. $\mu = \phi\nu$. If in addition μ is stationary, then ν can also be taken to be stationary.*

Proof of Theorem 3.3
(i) This follows straightforwardly from Lemma 5.1, the fact that factor maps cannot increase measure-theoretic entropy and the variational principle.
(ii) This is more involved and we use an argument based on conditional entropy. We write $\mu' = \phi\mu$. Since $h(X) = h(Y)$ it will be enough (using the variational principle) to show that $h(\mu) = h(\mu')$. The first thing to do is to write $h(\mu')$ in terms of μ and a partition on X. Let $\mathcal{P}_n(Y)$ be the partition of Y that specifies all coordinates in B_n, and denote by $\mathcal{P}_n(X)$ the partition of X that specifies all coordinates in B_n. Finally, $\mathcal{Q}_{n+\alpha}(X)$ is the partition of X that specifies the projection on B_n of the image under ϕ. Note that $\mathcal{P}_{n+\alpha}(X)$ refines $\mathcal{Q}_{n+\alpha}(X)$. We now write

$$
\begin{aligned}
h(\mu') &= \lim_{n\to\infty} \frac{h_{\mu'}(\mathcal{P}_n(Y))}{(2n+1)^d} \\
&= \lim_{n\to\infty} \frac{-\sum_{y\in\mathcal{B}_n(Y)} \mu(\phi^{-1}(y)) \log \mu(\phi^{-1}(y))}{(2n+1)^d} \\
&= \lim_{n\to\infty} \frac{h_\mu(\mathcal{Q}_{n+\alpha}(X))}{(2n+2\alpha+1)^d}.
\end{aligned}
$$

(Here $\phi^{-1}(y)$ are the elements in $B_{n+\alpha}(X)$ which map to y.)

We can write the entropy $h(\mu)$ as follows:

$$
\begin{aligned}
h(\mu) &= \lim_{n\to\infty} \frac{h_\mu(\mathcal{P}_{n+\alpha}(X))}{(2n+2\alpha+1)^d} \\
&= \lim_{n\to\infty} \left\{ \frac{h_\mu(\mathcal{Q}_{n+\alpha}(X))}{(2n+2\alpha+1)^d} + \frac{h_\mu(\mathcal{P}_{n+\alpha}(X) \mid \mathcal{Q}_{n+\alpha}(X))}{(2n+2\alpha+1)^d} \right\} \\
&= h(\mu') + \lim_{n\to\infty} \frac{h_\mu(\mathcal{P}_{n+\alpha}(X) \mid \mathcal{Q}_{n+\alpha}(X))}{(2n+2\alpha+1)^d}.
\end{aligned}
$$

Therefore, we need to show that the limit in the last expression is equal to zero. The numerator inside the limit is by definition equal to

$$\sum_{q\in\mathcal{Q}_{n+\alpha}(X)} \mu(q)\left\{-\sum_{p\in\mathcal{P}_{n+\alpha}(X)} \frac{\mu(p\cap q)}{\mu(q)}\log\frac{\mu(p\cap q)}{\mu(q)}\right\}. \tag{3}$$

Clearly, the term between the curly brackets is itself the entropy of a partition with respect to a probability measure and therefore bounded above by the logarithm of the number of elements in this partition which have positive probability. By Theorem 3.6, we know that for all $y \in Y$, $\epsilon > 0$, for n large enough we have that

$$|\{x \in \mathcal{B}_{n+\alpha}(X) : \phi(x) = \pi_n(y)\}| \le e^{\epsilon(2(n+\alpha)+1)^d}. \tag{4}$$

Let $N(y) = N_\epsilon(y)$ be the smallest number k so that (4) holds for all $n \ge k$. There exists a number N such that

$$\mu'(\{y : N(y) \le N\}) = \mu(\{\phi^{-1}\{y : N(y) \le N\}\}) \ge 1 - \epsilon.$$

This implies that up to at most an ϵ-portion, all atoms of the partition $\mathcal{Q}_{N+\alpha}$ contain less than $\exp\left(\epsilon(2(N+\alpha)+1)^d\right)$ elements. The remaining atoms contain (by the previous remark) at most $|\mathcal{A}(X)|^{(2(N+\alpha)+1)^d}$ elements. Hence the expression in (3) is bounded above by

$$\epsilon(2(N+\alpha)+1)^d + \epsilon(2(N+\alpha)+1)^d \log|\mathcal{A}(X)|,$$

and the proof is complete. □

For the proof of Theorem 3.4 we need the following lemma.

Lemma 5.2 *Let X be a strongly irreducible SOFT, and let $\mathcal{T}(X)$ be the set of transitive points of X. For every $\mu \in \mathcal{M}_{max}(X)$, we have $\mu(\mathcal{T}(X)) = 1$.*

Proof Fix a half-line ℓ as in the definition of transitivity, and a pattern p of size n. Consider a collection of translates $B(1), B(2), \dots$ of B_n centered at vertices on ℓ and such that any two boxes in this collection are separated by at least s (the separation distance of the shift space). By strong irreducibility, the pattern p has positive μ-probability. Hence there is positive μ-probability to see p in $B(1)$. By the property of uniform conditional probabilities (explained in the proof of Lemma 4.1) and the fact that the distances between the different translates is at least s, it is easy to see that there exists $\delta = \delta(n)$ such that for all k and for any conditioning of the configuration on $\cup_{i=1}^k B(i)$, the conditional probability to see p in $B(k+1)$ is at least δ. This easily implies that μ-a.s. we see p in some box $B(k)$.

Finally, we note that there are only countably many halflines ℓ and countably many patterns to check. This shows that the μ-probability to see all patterns on every half-line with rational direction is equal to one. □

Proof of Theorem 3.4 The proof of Proposition 9.2.2 in [4] does not depend on the dimension and implies that ϕ has no diamonds, which implies by Theorem 3.6 that $h(X) = h(Y)$. In view of Theorem 3.3, it suffices to show that if $\mu, \nu \in \mathcal{M}_{max}(X)$ with $\mu \ne \nu$, then $\phi\mu \ne \phi\nu$.

According to Lemma 5.2, both μ and ν live on $\mathcal{T}(X)$. Since it is immediate to check that ϕ maps $\mathcal{T}(X)$ into $\mathcal{T}(Y)$, it follows that $\phi\mu$ and $\phi\nu$ both live on $\mathcal{T}(Y)$. If $\mu \neq \nu$, then there exists a cylinder set A such that $\mu(A) \neq \nu(A)$. Note that $\phi(A)$ is measurable, since it is the continuous image of a compact set. We claim that

$$\phi\mu(\phi(A)) \neq \phi\nu(\phi(A)), \tag{5}$$

which implies that $\phi\mu \neq \phi\nu$. To do this, we first observe that

$$A \subseteq \phi^{-1}\phi(A) \subseteq A \cup (X \backslash \mathcal{T}(X)). \tag{6}$$

The first inclusion is obvious. For the second, suppose that $x \in \phi^{-1}\phi(A)$ and $x \in \mathcal{T}(X)$. Since the image of a transitive point is automatically transitive (as is easily verified), $\phi(x) \in \mathcal{T}(Y)$ and by the almost invertibility of ϕ, $\phi(x)$ has only one pre-image. This unique pre-image must then be x, and it follows that $x \in A$. Since by definition $\mu(\phi^{-1}\phi(A)) = \phi\mu(\phi(A))$, we take the μ-measure in (6), giving

$$\mu(A) \leq \phi\mu(\phi(A)) \leq \mu(A) + \mu(X \backslash \mathcal{T}(X)) = \mu(A),$$

where the last equation follows from Lemma 5.2. Hence all inequalities in the last equation are in fact equalities. A similar statement is true when we replace μ by ν and we finally obtain (5). □

Proof of Theorem 3.5 Let X be the SOFT $X_{M,M}$ defined in Section 3, and take M so large that X has exactly two ergodic measures of maximal entropy. Denote a block of size 1 by $\{b_{i,j} \, : \, -1 \leq i,j \leq 1\}$. Define the following block map $\Phi : \mathcal{B}_1(X) \to \{\mathcal{A}(X)\}^8$:

$$\Phi(\{b_{i,j}\}_{-1\leq i,j\leq 1}) = \mathrm{sgn}(b_{0,0})\,(b_{-1,1}, b_{0,1}, b_{1,1}, b_{-1,0}, b_{1,0}, b_{-1,-1}, b_{0,-1}, b_{1,-1}),$$

and denote by ϕ the factor map on X based on Φ. Finally, we write $Y = \phi(X)$.

We first claim that ϕ is two to one. To see this fix $y \in Y$. Suppose x is a pre-image of y. Given the value of y in the origin, we know the absolute value of the values of x in the set $B_1 \backslash (0,0)$; in addition, we know which of these values have the same sign. This leaves us with exactly two possibilities. Choose one of these. If we shift the block one unit to the right, say, then we know from the value of y at $(1,0)$ what the absolute value of x is at the origin. But given our previous choice, we have no freedom in the sign anymore. Continue in this way; it is then clear that our initial choice exhausts all potential freedom, and y has at most two pre-images. As it is clear that $\phi(x) = \phi(y)$, every point has exactly two pre-images.

Now, one observes that $h(X) = h(Y)$ as easily follows from the argument in the previous paragraph. We next claim that Y has a unique measure of maximal entropy. It follows from Theorem 3.3 that every measure of maximal entropy for Y must be a push forward of a measure of maximal entropy for X. However, since the two ergodic measures of maximal entropy μ^+ and μ^- are obtained from each other by flipping coordinates, we see that $\phi\mu^+ = \phi\mu^-$. This last measure is therefore the only measure of maximal entropy for Y. □

Acknowledgement: We thank one of the referees for very useful comments which very much improved the paper.

References

[1] Burton, R.M. and Steif, J. (1994) Nonuniqueness of measures of maximal entropy for subshifts of finite type, *Erg. Th. Dyn. Syst.* **14**, 213-235.

[2] Burton, R.M. and Steif, J. (1995) New results on measures of maximal entropy, *Isr. J. Math.* **89**, 275-300.

[3] Coven, E. and Paul, M. (1974) Endomorphisms of irreducible shifts of finite type, *Math. Syst. Theory* **8**, 167-175.

[4] Lind, D. and Marcus, B. (1995) *An introduction to symbolic dynamics and coding*, Cambridge University Press.

[5] Misiurewicz, M. (1975) A short proof of the variational principle for a $\mathbf{Z}_+^N$–action on a compact space, *Asterique* **40**, 147-157.

[6] Quas, A. and Trow, P. (2000) Subshifts of multi-dimensional shifts of finite type, *Erg. Th. Dyn. Syst.* **20**, 859-874.

[7] Radin, C. and Sadun, L. (1998) *Isomorphism of hierarchical structures*, to appear in *Erg. Th. Dyn. Syst.*

[8] Rudin, W. (1974) *Real and Complex Analysis*, 2nd edition, McGraw-Hill.

[9] Schmidt, K. (1995) *Dynamical Systems of Algebraic Origin*, Birkhäuser.

[10] Schmidt, K. (1998) Tilings, fundamental cocycles and fundamental groups of symbolic Z^d-actions. *Erg. Th. Dyn. Syst.* **18**, 1473-1525.