ON THE DYNAMICS AND RECURSIVE PROPERTIES OF MULTIDIMENSIONAL SYMBOLIC SYSTEMS

MICHAEL HOCHMAN

Abstract. We study the (sub)dynamics of multidimensional shifts of finite type and sofic shifts, and the action of cellular automata on their limit sets. Such a subaction is always an *effective dynamical system*: i.e. it is isomorphic to a subshift over the Cantor set the complement of which is a recursive sequence of basic sets.

Our main result is that, to varying degrees, this recursive-theoretic condition is also sufficient. We show that the class of expansive subactions of multidimensional sofic shifts is the same as the class of expansive effective systems, and that a general effective system can be realized, modulo a small extension, as the subaction of a shift of finite type or as the action of a cellular automaton on its limit set (after removing a dynamically trivial set).

As applications, we characterize, in terms of their computational properties, the numbers which can occur as the entropy of cellular automata, and construct SFTs and CAs with various interesting properties.

1. Introduction

1.1. **Background.** Let Σ be a finite set of symbolic and let $\Sigma^{\mathbb{Z}^d}$ be the compact space of configurations, i.e. of Σ-colorings of $\mathbb{Z}^d$. This space is the *full shift*, and $\mathbb{Z}^d$ acts on it by the translations$\{T_u\}_{u\in\mathbb{Z}^d}$ given by $(T_u(x))(v) = x(v+u)$, $v \in \mathbb{Z}^d$.

By a d-dimensional pattern over Σ we mean a Σ-coloring of a finite subset of $\mathbb{Z}^d$. If L is a finite set of patterns then they define a *shift of finite type* (SFT) by

$$S_L = \{x \in \Sigma^{\mathbb{Z}^d} : \text{no element of } L \text{ appears in } x\}$$

here, a pattern $a \in \Sigma^F$ is said to appear in a configuration x if $(T_u x)|_F = a$ for some $u \in \mathbb{Z}^d$. The set S_L is easily seen to be closed and invariant under the shift action, so may be regarded as a $\mathbb{Z}^d$-dynamical system. Background on topological dynamics can be found in section 2.

If Δ is some other finite alphabet and $Y \subseteq \Delta^{\mathbb{Z}^d}$ is a subshift which is the factor of an SFT, then Y is called a *sofic shift*.

SFTs and sofic shifts have been studied in topological dynamics, physics and computer science as models for interacting systems, and in dimensions $d \geq 2$ they are capable of very complex behavior. Indeed, not only are these systems hard to analyze but most questions about them are formally intractable, in the sense that, for most non-trivial questions, there is no algorithm which

Date: August 20, 2008.
2000 *Mathematics Subject Classification.* 37B15, 37B40, 37B50, 94A17, 03D45.

decides them given a finite presentation of the system as input. [Ber66, Rob71, Hur87, HKC92, Kar94]. The recursive structure of SFTs is also known to be complex [Mye74, Hur90, Sim07].

One can nonetheless ask meaningful questions about the dynamics of SFTs and sofic shifts. Research in this direction has been motivated by questions from particle physics and crystallography, and also by the mathematical theory of one-dimensional SFTs and sofic shifts, which are in many respects quite well understood. Many striking examples have emerged, e.g. [Moz89, BS94], and also the feeling that almost anything can happen for higher dimensional SFTs. Below we show that in a certain precise sense this is true.

1.2. **Effective systems and subactions of SFTs and sofic shifts.** In this work we study SFTs and sofic shifts through their subactions, by which we mean the restriction of the $\mathbb{Z}^d$-action to a subgroup $H < \mathbb{Z}^d$. We are interested in understanding what dynamics can arise in this way.[1]

It turns out that to a large extent the subdynamics can be characterized, and the characterization is recursion-theoretic. This is another example of a general philosophy that has emerged recently: for many characteristics of these systems there is a trivial restriction of a recursive nature, and this turns out to be the only restriction. An example of this is our work in [HM07], where the entropies of SFTs and sofic shifts are characterized. Another example is the Medvedev degree invariant introduced by Simpson [Sim07].

To state our results we require some definitions. A sequence (a_n) of integers is *recursive* (R) if there is an algorithm A (formally a Turing machine) that, upon input $n \in \mathbb{N}$, outputs a_n. A set of integers is *recursively enumerable* (RE) if it is the set of elements of some recursive sequence.

By identifying the integers with other sets we can speak of recursive sequences of other elements. For example, since $\mathbb{N} \cong \mathbb{N}^2$ (and the bijection can be made effective), we can speak of recursive sequences of pairs of integers; and in the same way of sequences of finite sequences of integers.

Let $\{0,1\}^{\mathbb{N}}$ denote the Cantor set and for a finite $I \subseteq \mathbb{N}$ and $a \in \{0,1\}^I$ let

$$[a] = \{x \in \{0,1\}^{\mathbb{N}} \,:\, x|_I = a\}$$

denote the cylinder set determined by a. As a ranges over all finite patterns of this sort, $[a]$ provides a basis of closed and open sets for the topology. Notice that the set of such a's, which parametrizes the cylinder sets, can be put into effective bijection with $\mathbb{N}$.

Definition 1.1. A subset $X \subseteq \{0,1\}^{\mathbb{N}}$ is *effectively closed* if its complement is the union of a recursive sequence of cylinder sets.[2]

[1] Since passing to finite index subgroup action on an SFT is still an SFT, and similarly for sofic shifts, it is not important which subgroup we consider, and in formulating our results we will take the canonical $\mathbb{Z}^k$-subgroups of $\mathbb{Z}^d$ generated by the first k of the standard generators $e_1, \ldots, e_d$ of $\mathbb{Z}^d$.

[2] An effectively closed set is also the complement of the union of a recursive *set* $\mathcal{C}$ of cylinder sets, i.e. there is an algorithm that decides in finite time whether a given cylinder set $[a]$ is in $\mathcal{C}$. This condition is a-priori stronger and we thank S. Simpson for pointing this equivalence out to us.

Since there are only countably many recursive sequences, there are only countably many effectively closed sets.

We may similarly define effectively closed subsets of the Cantor set when parametrized as $\Sigma^{\mathbb{Z}^d}$ for arbitrary finite Σ, or closed subset of $(\{0,1\}^{\mathbb{N}})^{\mathbb{Z}^d}$. Let T_u denote the shift actions on these spaces.

Definition 1.2. A subset $X \subseteq \Sigma^{\mathbb{Z}^d}$ (Σ finite) is an *effectively symbolic system* (ESS) if it is effectively closed and invariant under the shift.

A subset $X \subseteq (\{0,1\}^{\mathbb{N}})^{\mathbb{Z}^d}$ is an *effectively dynamical system* (EDS) if it is effectively closed and invariant under the shift.[3]

We will also say that a dynamical system is effective if it is isomorphic to an ESS or EDS.

Once again, both these classes are countable; but we know of no "natural" type of dynamics of subshifts or totally disconnected systems which cannot arise as ESSs or EDSs, respectively. Indeed, all combinatorial constructions which appear in the literature give systems of this type, assuming they are defined by recursive parameters (for example, Sturmian sequences will be ESSs if the rotation and partition used are computable). We note that an expansive EDS is (isomorphic to) an ESS, and every ESS is (isomorphic to) an EDS (though an ESS can also be embedded in $(\{0,1\}^{\mathbb{N}})^{\mathbb{Z}^d}$ in a non-effective way).

If L is a finite set of patterns then the set of all translates of patterns from L is an RE set, and thus the SFT S_L is an ESS. A symbolic factor of an ESS is an ESS, so sofic shifts are ESSs; and the subaction of an EDS is an EDS (we prove both these statements in section 3). This provides a restriction on the subdynamics of SFTs:

Theorem 1.3. *The subaction of an SFT or sofic shift is an EDS.*

This necessary condition turns out to be (almost) sufficient. We have the following characterization of the expansive subdynamics of sofic shifts:

Theorem 1.4. *A symbolic system is isomorphic to the subaction of a sofic shift if and only if it is effective.*

Specifically, we can realize a $\mathbb{Z}^k$-ESS X as the $\mathbb{Z}^k$-subaction of a $\mathbb{Z}^{k+2}$ sofic shift; in fact, the $\mathbb{Z}^{k+2}$ subshift obtained by extending each configuration in X identically in the directions e_{k+1}, e_{k+2} complementary to the subgroup $\mathbb{Z}^k$, is a sofic shift. We do not know whether either of these statements holds when $k+2$ is replaced with $k+1$.

The analog of theorem 1.4 is false for SFTs. Indeed, there are ESSs which cannot be realized as subactions of SFTs, such as the Chacon system (see proposition 6.2 below). It is an interesting open problem whether one can characterize the expansive subactions of SFTs.

[3]Equivalently, a $\mathbb{Z}^d$-system is an EDS if it is the inverse limit of a recursive sequence of ESSs. For a system X, this means that X is the inverse limit of a diagram

$$\ldots \to X_m \xrightarrow{\pi_m} X_{m-1} \to \ldots \to X_0$$

where $X_m \subseteq (\{0,1\}^m)^{\mathbb{Z}^d}$ and (a) there is a recursive array $a_{m,n}$ of patterns with $a_{m.n}$ having symbols in $\{0,1\}^m$, such that $X_m = S_{\{a_{m,n}\,:\,n\in\mathbb{N}\}}$, and (b) the sequence π_m is recursive (note that they are block codes).

Let us now turn to the non-expansive case of EDS. Note that in order for an EDS to occur as the subaction of an SFT or sofic shift, it must first of all arise as the subaction of some ESS. This is a non-trivial restriction, since not all EDS have this property; for example, there exist effective odometers, and these do not arise as the subaction of any symbolic system, effective or not (see section 6.2). There are also other obstructions; for example, the topological Chacon system is not the subaction of an SFT, though it is an effective symbolic system and therefore is the subaction of a sofic shift. See section 6.2.

These problems disappear if one is willing to allow a small extension. We shall say that a factor map $\pi : Y \to X$ is *almost-1-1* if the set of points in X with unique pre-image has full measure with respect to every invariant Borel probability measure on X.[4] Although weaker than isomorphism, this relation implies that the statistical behavior of the systems Y and X are identical in a strong sense: T induces a bijection of the invariant probability measures on X and Y, and for every invariant measure ν on Y the factor map is a continuous isomorphism of (Y, ν) and $(X, \pi\nu)$.

Another type of extension $Y \to X$ which may be considered trivial occurs when Y extends X by a direct product with a well-understood system, that is, $Y = X \times W$, and the factor is projection onto the first coordinate. We will deal with the particularly simple case where W is an isometric action on a totally disconnected space.

Definition 1.5. An extension $Z \to Y$ of $\mathbb{Z}^k$-dynamical systems is an *almost trivial isometric extension* (ATIE) if we can interpolate a factor

$$Z \to Y \times W \to Y$$

where W is an isometric action on a totally disconnected space, $Z \to Y \times W$ an almost-1-1 extension, and $Y \times W \to Y$ is projection onto the first coordinate.

The composition of ATIEs is an ATIE, and ATIEs do not increase topological entropy. The invariant measures of a system and an ATIE extension of it differ by at most the addition of some pure-point rational spectrum, and for this reason and those explained above ATIEs can be considered small from the point of view of the ergodic behavior of orbits. We remark, however, that from a purely topological point of view many properties are not preserved by ATIEs, such as transitivity, expansiveness and equicontinuity.

Our main result for SFTs is:

Theorem 1.6. *The subaction of an SFT is an EDS. Conversely, if Y is an effective $\mathbb{Z}^k$-system, then there is an SFT X and a $\mathbb{Z}^k$-subaction of X which is an ATIE of Y.*

As before, we can prove this with X a $k+2$-dimensional SFT; we do not know if the dimension can be reduced to $k+1$. We have quite good control of the isometric part of the ATIE, and can make it an odometer. On the other hand, the almost-1-1 part of the extension partly comes

[4] Note that the set of points in X with unique pre-image is a G_δ-set. Some authors define almost-1-1 extensions by the condition that this set is dense. This notion is distinct from ours, though in the presence of a globally supported invariant measure, and in particular when X is minimal, our definition implies the other.

from the dynamics of a certain Turing machine associated to the EDS, and we have little control over it. We mention that the SFT X which we construct in the proof has entropy 0 with respect to the full $\mathbb{Z}^{k+2}$-action.

It is not clear how far one can reduce the size of the extension in theorem 1.6. We show in proposition 6.2 below that if the subaction of an SFT factors onto the Chacon system then, with respect to the unique invariant probability measure on the factor, almost every fiber contains more than one point, and in particular the extension cannot be almost-1-1. It remains an interesting question whether every effective system has a finite-to-1 extension that is the subaction of an SFT.

1.3. **Cellular automata.** A *cellular automaton* (CA) is a continuous transformation $f : \Sigma^{\mathbb{Z}^d} \to \Sigma^{\mathbb{Z}^d}$ of the full shift (with Σ finite), which commutes with the shift action in the sense that $T_u f = f T_u$ for $u \in \mathbb{Z}^d$. The Curtis-Hedlund-Lyndon [Hed69]theorem characterizes such maps as those which are defined locally: the site u in $f(x)$ is determined by the coloring of a neighborhood of u in x, and the size of the neighborhood is independent of x. This makes CA an attractive discrete model for interacting systems. CAs were introduced in the 1940's by von Neumann [vN66]. They were popularized by J. Conway in the 1970's with the introduction of his "Game of Life", and in the 1980's by the work of Wolfram [Wol84]. The recursive properties of CA have been studied by several authors [Hur87, CHY90, Hur90, Kar94, Sut05], and as for SFTs, most properties are undecidable. See [Kar05] for a recent survey.

The limit set of a CA f is $\Lambda = \cap_{n=1}^{\infty} f^n(\Sigma^{\mathbb{Z}^d})$; this is the largest set on which f acts surjectively. In order to get an action which is also injective, we pass to the natural extension (Λ^*, f^*) of (Λ, f), i.e. the inverse limit of the diagram $\ldots \xrightarrow{f} \Lambda \xrightarrow{f} \Lambda \xrightarrow{f} \ldots \xrightarrow{f} \Lambda$. This is the smallest invertible system extending (Λ, f); we call (Λ^*, f) the *limit $\mathbb{Z}$-action of f*.

The classes of limit $\mathbb{Z}$-actions of CA and of $\mathbb{Z}$-subactions of SFTs are closely related; they are essentially the same after removing the unavoidable periodic point from the limit sets of CAs. See section 3 for details. Using this, theorem 1.6 gives

Theorem 1.7. *The limit $\mathbb{Z}$-action of a CA is an EDS. Conversely, if Y is a $\mathbb{Z}$-EDS, then there is a 3-dimensional CA f such that, after removing from its limit $\mathbb{Z}$-action a fixed point and its basin of attraction, the remaining system is an ATIE of Y.*

We do not know to what extent theorem 1.7 holds in dimension 1 and 2, or what analog there may be for one-dimensional CA and for injective or surjective CAs in any dimension.

1.4. **Applications.**

Entropy of CAs. Entropy is perhaps the most important invariant of a dynamical system (see section 7.1 for definitions). It has been known for some time that, in general, one cannot compute the entropy of an SFT or CA from its combinatorial description. For SFTs this follows from Berger's theorem [Ber66], and for CA was proved by Hurd, Kari and Culik [HKC92].

We are interested in a somewhat different question, namely, what are the possible numbers that can arise as the entropy of SFT subactions and CAs; a-priori there are only countably many such numbers, and for SFTs in dimension 1 they have a simple algebraic characterization [Lin84]. In [HM07], we recently proved the following recursive-theoretic characterization of entropies of higher dimensional SFTs:

Theorem. [HM07] *Fix $d \geq 2$. Then a real number $h \geq 0$ is the entropy of a $\mathbb{Z}^d$-SFT if and only if it is the infimum of a recursive sequence of rational numbers.*

This is actually the same class of numbers which arises as the entropies of ESSs. In a similar vein, we can now prove the following:

Theorem 1.8. *For each $d \geq 1$, a real number $h \geq 0$ is the entropy of a $\mathbb{Z}^d$-EDS if and only if it is the* $\liminf$ *of a recursive sequence of rational numbers.*

Applying theorem 1.8, and the fact that ATIEs do not increase entropy, we get:

Corollary 1.9. *For $d \geq 3$, the entropies of d-dimensional CA are precisely the non-negative numbers that are the* $\liminf$ *of a recursive sequence of rational numbers.*

Real numbers with various recursive properties have been studied in [ZW01], where a countably infinite hierarchy of number types is described. Let us mention here one interesting fact; the class of number which are entropies of EDS (and hence of CA) contains numbers which are not the limit of any recursive sequence (in contrast, the entropies of SFTs always are). Thus the entropy of some CAs is truly out of reach. This sharpens a theorem of Hurd, Kari and Culik [HKC92], who showed that the function $f \mapsto h(f)$, which assigns to a CA f its entropy, cannot be approximated. What we now know is the stronger fact that not only can the function not be approximated, but there are individual values which in a very strong sense cannot be approximated.

It is an interesting open problem to characterize the entropies of CA in dimensions 1 and 2.

Measures of maximal entropy on SFTs. Interest in measures of maximal entropy on SFTs is motivated by the study of phase transitions in particle physics, and there are by now several examples of SFTs with multiple measures of maximal entropy [BS94, QŞ03].

Theorem 1.6 gives a general mechanism for producing SFTs whose measures of maximal entropy behave in various ways. Given a zero-entropy SFT $Y \subseteq \Sigma^{\mathbb{Z}^d}$, $\Sigma_0 \subseteq \Sigma$ and $k \in \mathbb{N}$, define an SFT W by superimposing one of k new symbols over each symbol in Σ_0; formally, take the SFT $W \subseteq Y \times \{0, 1, \ldots, k-1\}^{\mathbb{Z}^d}$ defined by the condition that $(y, y') \in W$ if and only if $y'(u) = 0$ whenever $y(u) \notin \Sigma_0$. We call W the k-extension (with respect to Σ_0). Then

$$h(W) = \max\{\mu(\bigcup_{\sigma \in \Sigma_0} [\sigma]) \cdot \log k \,:\, \mu \text{ an invariant probability measure on } W\}$$

and the measures of maximal entropy are in 1-1 correspondence with the measures maximizing the quantity above (this approach was used in [HM07], but the control there over Y was poorer). One should note however that this technique cannot produce irreducible (or even mixing) examples, since in the procedure described the system we get factors onto Y, which, using our present techniques, always has some discrete spectrum.

Another interesting consequence for CA is that, since there are $\mathbb{Z}$-EDS without measures of maximal entropy, there must be 3-dimensional CA with this property as well. This is not otherwise obvious.

Factoring relations. A problem which has received some attention recently is that of determining the factoring relations between SFTs, and particularly the question whether every SFT with entropy $\geq \log N$ factors onto a full shift on N symbols [JM05]. Using the technique above and results from [HM07], Boyle and Schraudner recently showed that this is false. As another application, we answer question 2.10 of [BS07], albeit in three rather than two dimensions:

Proposition 1.10. *There is an SFT $Y \subseteq \Sigma^{\mathbb{Z}^3}$ with entropy $\log 2$ which does not factor onto the full shift $\{0,1\}^{\mathbb{Z}^3}$, and it can be obtained as a 4-extension of a uniquely ergodic SFT with respect to a set of symbols Σ_0 having density $1/2$. On the other hand, there is an infinite, uniquely ergodic SFT and a set of symbols of density $1/2$ whose 4-extension does factor onto the full shift on N symbols.*

1.5. **Organization.** The rest of this paper is organized as follows. In the next section we give some background in topological dynamics. In section 3 we discuss some general properties of EDS and prove theorem 1.3. In section 4 we describe some auxiliary constructions, and in section 5 we construct sofic shifts and SFTs with specified dynamics, proving theorem 1.4 and 1.6. In section 6 we discuss the relation between CAs and subactions of SFTs, proving theorem 1.7, and give some (counter-)examples. In sections 7 we discuss entropy of EDS, and in section 8 we prove theorem 1.10 about the factoring of SFTs onto full shifts.

Acknowledgement. I am grateful to Benjamin Weiss for many interesting and stimulating discussions. I also thank the referee for a careful reading and helpful suggestions.

2. Topological dynamics

We collect here some standard definitions from topological and symbolic dynamics.

A topological $\mathbb{Z}^d$-dynamical system (X,T) (sometimes written $(X,\mathbb{Z}^d)$) is an action of $\mathbb{Z}^d$ by homeomorphisms $\{T_u\}_{u\in\mathbb{Z}^d}$ on a compact metric space X; in this paper we assume that X is totally disconnected. Two dynamical $\mathbb{Z}^d$-systems (X,T) and (Y,S) are isomorphic, or conjugate, if there is a homeomorphism $\pi : X \to Y$ satisfying $S_u\pi = \pi T_u$ for $u \in \mathbb{Z}^d$. If π is merely onto then it is called a *factor map* from X to Y, and X is called an *extension* of Y.

Let Σ be a finite set of symbols. The space $\Sigma^{\mathbb{Z}^d}$ of colorings of $\mathbb{Z}^d$ by Σ is called the *full d-dimensional shift over* Σ, or just the full shift, and its points are called *configurations*. Topologically the full shift is a Cantor set, and it comes equipped with a natural $\mathbb{Z}^d$ action, called the *shift action*, in which $u \in \mathbb{Z}^d$ acts via the translation $T_u : \Sigma^{\mathbb{Z}^d} \to \Sigma^{\mathbb{Z}^d}$ defined by

$$(T_u x)(v) = x(u+v)$$

We let $e_1,\ldots,e_d$ denote the standard generators of $\mathbb{Z}^d$, and write $T_1,\ldots,T_d$ for the corresponding shift elements.

A subset $X \subseteq \Sigma^{\mathbb{Z}^d}$ which is closed and invariant to the shift (i.e. $T_u X = X$ for $u \in \mathbb{Z}^d$) is called a *subshift*, or a *symbolic system*.

By the Curtis-Hedlund-Lyndon theorem [Hed69], factor maps between subshifts of the same dimension (but possibly distinct alphabets) are given by a *block code*: if $Y \subseteq \Delta^{\mathbb{Z}^d}$, $X \subseteq \Sigma^{\mathbb{Z}^d}$ and $\pi : Y \to X$ is a factor map, then there is a finite $F \subseteq \mathbb{Z}^d$ and a function $\pi_0 : \Delta^F \to \Sigma$, so that π acts on each site of $x \in \Delta^{\mathbb{Z}^d}$ by applying π_0 to the local neighborhood of the site: $(\pi x)(u) = \pi_0((T_u x)|_F)$. The diameter of F is called the *window size* of π. Conversely, any such map $\pi_0 : \Delta^F \to \Sigma$ gives rise to a factor map π in this way (the image is automatically a subshift).

The property of (X, T) being isomorphic to a subshift can be characterized intrinsically: it is equivalent to being totally disconnected and *expansive*, i.e. there is some $\varepsilon > 0$ such that, for any $x \neq y$, there is some $u \in \mathbb{Z}^d$ such that $\delta(T_u x, T_u y) > \varepsilon$, where δ is some fixed metric (but the condition does not depend on the metric). See [Wal82].

3. Basic properties of ESSs and EDSs

In this section we develop some general properties of EDS, in the course of which we will prove theorem 1.3.

Theorem 3.1. *A subaction of an effective system is effective.*

Proof. Let $X \subseteq (\{0,1\}^{\mathbb{N}})^{\mathbb{Z}^d}$ be an effectively closed subset, invariant under the $\mathbb{Z}^d$-shift action, and let $H < \mathbb{Z}^d$ be a subgroup. Let $K \subseteq \mathbb{Z}^d$ be a recursive cross-section of the projection $\mathbb{Z}^d \to \mathbb{Z}^d/H$. Then

$$(\{0,1\}^{\mathbb{N}})^{\mathbb{Z}^d} \cong (\{0,1\}^{\mathbb{N}\times K})^H$$

and the homeomorphism is effective (it is induced by a computable identification of $(\mathbb{N}\times K)\times H$ with $\mathbb{N}\times\mathbb{Z}^d$). Thus the recursive set of cylinder sets which together constitute the complement of X in $(\{0,1\}^{\mathbb{N}})^{\mathbb{Z}^d}$ is a recursive set of cylinder sets with respect to the new parametrization, and shows that X is an effective (and clearly shift-invariant) subset of $(\{0,1\}^{\mathbb{N}\times K})^H$, as required. □

Since an SFT is an ESS, theorem 1.3 follows immediately for SFTs. To obtain the same for sofic shifts we first need a classical fact from recursion theory. Recall that a set $A \subseteq \mathbb{N}$ is recursive if there is an algorithm that, given $n \in \mathbb{N}$, decides whether $n \in A$.

Lemma 3.2. *Suppose $L \subseteq U$ is an RE set and $R \subseteq U \times V$ is a recursive set, and let*

$$M = \{b \in V \,:\, (a,b) \in R \text{ for some } a \in L\}$$

Then M is RE.

Proof. Let A be an algorithm that on input $a \in U$ halts if $a \in L$ and runs forever otherwise. Let B be the algorithm which, upon input $b \in V$, iterates over all pairs $(n, a) \in \mathbb{N} \times U$, and for each pair runs the algorithm A for n steps (or until it halts) on the input a. If A halts before n steps are up, it checks whether $(a, b) \in R$, and if so it halts; otherwise it continues to the next pair (n', a'). It is easily seen that this algorithm halts on input b if and only if $b \in M$, so M is RE. □

Proposition 3.3. *A symbolic factor of an EDS is an ESS.*

Proof. Let $Y \subseteq (\{0,1\}^{\mathbb{N}})^{\mathbb{Z}^d}$ be an EDS and $U_1, U_2, \ldots$ a computable sequence of cylinder sets whose union are the complement of Y. Let $Z \subseteq \Sigma^{\mathbb{Z}^d}$ be a subshift and $\pi : Y \to Z$ a factor map. We must show that Z is effective.

For each $\sigma \in \Sigma$ let $C_\sigma = \{x \in \Sigma^{\mathbb{Z}^d} : x(0) = \sigma\}$ and let $V_\sigma = \pi^{-1}(C_\sigma)$. Each V_σ is closed and open, so can be written as a finite union of cylinder sets. Thus we have a concrete, finite representation of the factor map.

A cylinder set $C \subseteq \Sigma^{\mathbb{Z}^d}$ intersects Z non-trivially if and only if for there is some $y \in Y$ with $\pi(y) \in Z$. By compactness, this occurs if and only if for each n there is a point $y \in Y$ so that $T^u y \in V_{x(u)}$ for $u \in [-n; n]^d$. Therefore, C is disjoint from Z if and only if for some n,

$$Y \cap \bigcap_{u \in [-n;n]^d} T^{-u} V_{x(u)} = \emptyset$$

which occurs if and only if for some n, k,

$$\bigcap_{u \in [-n;n]^d} T^{-u} V_{x(u)} \subseteq \bigcup_{i=1}^{k} U_i$$

In summary,

$$C \cap Z = \emptyset \iff \exists n, k \bigcap_{u \in [-n;n]^d} T^{-u} V_{x(u)} \subseteq \bigcup_{i=1}^{k} U_i$$

Now, the relation on the right hand side is computable, since the Boolean operations between cylinder sets are computable; hence by the preceding lemma, we see that the collection of cylinder sets disjoint from Z is RE, as required. □

Corollary 3.4. *An expansive effective system is isomorphic to an ESS.*

4. Auxiliary Constructions

In this section we describe some constructions which we use later on in the proof of theorems 1.4 and 1.6.

4.1. **Superposition.** Given an SFT $X = S_L$ defined by a finite set of patterns L, *superposition* is a combinatorial construction which gives an SFT X' that factors into (generally not onto) X. Informally, this is done by adding data to each symbol of X and enriching L with rules relating to this new layer of data.

More precisely, suppose X is an SFT defined by a set $L \subseteq \Sigma^F$. A system Y is superimposed over X if it is obtained by the following process. (a) Fix a finite set Δ, and replace each symbol $\sigma \in \Sigma$ with one or more symbols of the form $(\sigma, \delta) \in \Sigma \times \Delta$. Let Σ' be the set of these pairs. For the new symbol $(\sigma, \delta) \in \Sigma'$, we say that δ is superimposed over σ; we also frequently refer to this pair as the symbol σ marked with δ. (b) We extend each pattern $a \in L \subseteq \Sigma^F$ in every possible way to a pattern in $a' \in (\Sigma')^F$ by superimposing new symbols over each symbol of a. (c) Let L' be the extended patterns from (b), together with possibly other patterns. Then the SFT X' defined by L' is superimposed over X, and has the property that every pattern appearing in X' consists of a Δ-configuration superimposed Σ-configurations from X.

Note that $S_{L'}$ may be empty, but if it is not then the map $\pi : X' \to \Sigma^{Z^2}$ which erases the superimposed layer of data maps X' to a subsystem of X. We say that $x \in X$ is represented in X' if one can turn x into a point of X' by superimposing a suitable Δ-pattern over x; i.e., if $x = \pi(x')$ for some $x' \in X'$.

4.2. **Subshifts Defined by Substitution.** One of the building blocks of our construction will be certain SFTs whose configurations possess a simple hierarchical structure. We will not need anything more complicated than Robinson's classical aperiodic SFT [Rob71], but rather than describe that system and the modifications we would require of it, we will instead rely on a general construction due to Mozes [Moz89], which allows a shorter and more transparent exposition.

For the remainder of this section we fix the dimension $d = 2$. Given a finite alphabet Σ, a *substitution rule* is a map $s : \Sigma \to \Sigma^{F_k}$ for some integer $k > 1$, where $F_k = \{1, \ldots, k\}^2$ (in the terminology of [Moz89], this is a deterministic $k \times k$ substitution system with property A). The map s extends naturally to a map $s_n : \Sigma^{F_n} \to \Sigma^{F_{n \cdot k}}$ by identifying $\Sigma^{F_{n \cdot k}}$ with $(\Sigma^{F_k})^{F_n}$.

Starting from a single symbol located at $(1, 1) \in \mathbb{Z}^2$ and iterating the substitution map, we obtain a sequence of colorings of F_{k^n} for $n = 0, 1, 2 \ldots$. Such patterns are called s-blocks. A point $x \in \Sigma^{\mathbb{Z}^2}$ is admissible for s if every finite subpattern of x appears in some s-block. The subshift $W \subseteq \Sigma^{\mathbb{Z}^2}$ associated with s is the set of admissible patterns; this is seen to be closed and shift invariant.

For each configuration $x \in W$ one can find a *derivation tree* of x. This is an infinite tree whose node set V is a disjoint union $V = \cup_{n=0}^{\infty} V_n$. Each V_n is identified with a k^n-periodic subset $\widehat{V}_n \subseteq \mathbb{Z}^2$ (that is, a coset of $k^n\mathbb{Z}^2$) in such a way that $\widehat{V}_0 = \mathbb{Z}^2$ and $\widehat{V}_n \subseteq \widehat{V}_{n-1}$, and nodes $v \in V_n$, $v' \in V_{n-1}$ are connected if $\hat{v}' \in \hat{v} + \{0, \ldots, k^n - 1\}^2$, where $\hat{v}, \hat{v}' \in \mathbb{Z}^2$ correspond to v, v' respectively. Each node in the tree also carries a label from Σ. A derivation tree for $x \in W$ must satisfy the condition that the labeling of $\widehat{V}_0$ agrees with x, and for $n > 0$, every $a \in \Sigma$ and each $u \in \widehat{V}_n$ labeled a, the labeling of the $k \times k$ square of elements of $\widehat{V}_{n-1}$ of which u is the lower left corner are labeled according to the block $s(a)$. In other words, the labeling of $\widehat{V}_1$ corresponds to a decomposition of x into $k \times k$ blocks arranged in a grid, and then replacing each block with the symbol from which it is derived; this gives a pattern on a coset of $k\mathbb{Z} \times k\mathbb{Z}$, which we may identify with a point in $\Sigma^{\mathbb{Z}^2}$, decompose it into $k \times k$ blocks and repeat this procedure to get the labeling of $\widehat{V}_2$, and so on. One can prove by induction that a finite version of this procedure can be carried out n times for each block $s_n(a)$, $a \in \Sigma$; since each sub-pattern of x is contained a block of this form, a compactness argument now shows that for any $x \in W$ these finite derivations can be pasted together consistently, giving a derivation tree for x.

A substitution rule s has *unique derivation* if each $x \in X$ has a unique derivation tree. The derivation tree need not be connected; a tree is connected if and only if every pair of vertices have a common parent. However, it is easy to see that if C is a connected component of the tree then $C \cap \widehat{V}_0$ is either the whole plane, a half-plane or a quarter-plane, and hence there are at most four connected components which meet along horizontal and/or vertical lines.

Theorem 4.1. *(Theorem 4.5 of* [Moz89]*) Let $s : \Sigma \to \Sigma^{F_k}$ be a substitution rule with unique derivation and let W be the associated dynamical system. Then there exists an alphabet Δ, an SFT $\widetilde{W} \subseteq \Delta^{\mathbb{Z}^2}$, and a one-block factor map $\varphi : \widetilde{W} \to W$ such that $x \in W$ has a unique pre-image under φ whenever the derivation tree of x is connected.*

Note that theorem 4.1 is false in dimension $d = 1$.

Corollary 4.2. *Let s be a substitution rule with unique derivation and $\varphi : \widehat{W} \to W$ as in theorem 4.1. Then φ is an almost-1-1 extension with respect to the action of T_u for any $u \notin \mathbb{Z}e_1 \cup \mathbb{Z}e_2$.*

Proof. To any $x \in W$ let D_x denote its derivation tree, and to x associate the collection

$$C_x = \{C \cap \widehat{V}_0 \,:\, C \text{ is a connected component of } D_x\}$$

The function $C : x \mapsto C_x$ is measurable and takes on countable many values. Note that according to our identification of tree nodes with points in the plane, the tree is acted on by the shifts in a natural way, and in particular $C_{T_u x} = T_u(C_x)$, where T_u acts on C_x by shifting each element of C_x. If C_x contains a half- or quarter-space then $T_u^n C_x \cap C_x = \emptyset$ for $n \in \mathbb{Z}$, because $u \notin \mathbb{Z}e_1 \cup \mathbb{Z}e_2$; hence any T_u-invariant measure gives mass 0 to those x with C_x non-trivial. Hence by the Poincare recurrence principle, the set of x's with disconnected derivation tree has measure zero for every T_u-invariant probability measure on W, and the claim follows from theorem 4.1. $\square$

4.3. **Almost Odometers.** Fix $2 \leq p \in \mathbb{N}$ for the remainder of this section, and consider the substitution $s : \{\circ, \bullet\} \to \{\circ, \bullet\}^{\{1,\ldots,p\}^2}$ defined by the rule that maps $\bullet$ to a $p \times p$ block with $\bullet$'s on the diagonal in direction $\nearrow$ and $\circ$'s everywhere else, and maps $\circ$ to the same block except that the upper right corner is a $\circ$ instead of a $\bullet$. For $p = 5$ this gives the rule

$$\bullet \;\mapsto\; \begin{matrix} \circ & \circ & \circ & \circ & \bullet \\ \circ & \circ & \circ & \bullet & \circ \\ \circ & \circ & \bullet & \circ & \circ \\ \circ & \bullet & \circ & \circ & \circ \\ \bullet & \circ & \circ & \circ & \circ \end{matrix} \qquad\qquad \circ \;\mapsto\; \begin{matrix} \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \bullet & \circ \\ \circ & \circ & \bullet & \circ & \circ \\ \circ & \bullet & \circ & \circ & \circ \\ \bullet & \circ & \circ & \circ & \circ \end{matrix}$$

Let $\Omega = \Omega_p$ denote the dynamical system defined by these rules. See figure 4.1 for three iterations of the rule with $p = 2$.

Let us say that a row or column of an s-block is of type r if the $\bullet$'s in it appear in an infinite arithmetic progression with gap p^r, and of type $\geq r$ if it is of type r' for some $r' \geq r$; for finite sequences we adopt the convention that a sequence of length n containing only one occurrence

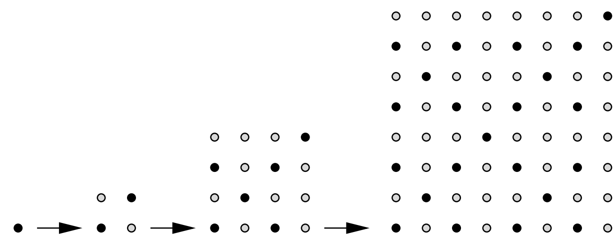

FIGURE 4.1. Three iterations of the substitution for $p = 2$.

of $\bullet$ has period n, and if it has no occurrences of $\bullet$ its period is ∞. One readily verifies by induction that for each $r \geq 1$ and all sufficiently large blocks $a \in \{s_n(\bullet), s_n(\circ)\}$, the set

$$I_r(a) = \{i \,:\, 1 \leq i \leq p^n \,:\, \text{the } i\text{-th column of } a \text{ is of type } \geq r\}$$

is the intersection of $[1; p^n]$ with a coset of $p^{r-1}\mathbb{Z}$. Similarly the set

$$J_r(a) = \{i \,:\, 1 \leq j \leq p^n \,:\, \text{the } j\text{-th row of } a \text{ is of type } \geq r\}$$

is the intersection of $[1; p^n]$ with a coset of $p^{r-1}\mathbb{Z}$.

We define $I_r(\omega), J_r(\omega)$ similarly for infinite configurations $\omega \in \Omega_p$, except that now a row or column with a unique $\bullet$ is also considered type ∞. It follows from the finite case that in the infinite case I_r, J_r are cosets of $p^{r-1}\mathbb{Z}$.

Denote $\mathbb{Z}_k = \mathbb{Z}/k\mathbb{Z}$, and let $U_t = \mathbb{Z}_{p^t} \times \mathbb{Z}_{p^t}$, and let $T_1, T_2 : U_t \to U_t$ be the maps $T_1(m, n) = (m - 1, n) \bmod p^t$ and $T_2(m, n) = (m, n - 1) \bmod p^t$. We consider T_1, T_2 as the generators of a free abelian group $\mathbb{Z}^2$ acting on U_t, giving a $\mathbb{Z}^2$-dynamical system (the action of course is not faithful). Let π_t denote reduction modulo p^t (we allow π_t to act in the obvious manner on elements of $\mathbb{Z}$, $\mathbb{Z}_{p^{t'}}$ and $U_{t'}$ for $t' \geq t$). We obtain the following compatible sequence of factor maps of $\mathbb{Z}^2$ dynamical systems:

$$\ldots \xrightarrow{\pi_t} U_t \xrightarrow{\pi_{t-1}} \ldots \xrightarrow{\pi_1} U_1 \tag{4.1}$$

Define $\widehat{\pi}_r : \Omega \to U_r$ by

$$\widehat{\pi}_r(\omega) = (\pi_r(I_{r+1}(w)), \pi_r(J_{r+1}(w))$$

note that $\pi_r(I_{r+1})$ is well-defined because I_{r+1}, J_{r+1} are p^r-periodic, and that $\pi_{r-1} \circ \widehat{\pi}_r = \widehat{\pi}_{r-1}$ because the I_r's and J_r's are decreasing sequences. Note also that by definition

$$\begin{aligned} I_r(T_{ae_1+be_2}w) &= I_r - a \\ J_r(T_{ae_1+be_2}w) &= J_r - b \end{aligned}$$

so $\widehat{\pi}(T_u\omega) = T_u\widehat{\pi}(\omega)$, where $u \in \mathbb{Z}^2$ and T_u is the element of the appropriate $\mathbb{Z}^2$-action generated by T_1, T_2. We see that the system of factors $\widehat{\pi}_r : \Omega \to U_r$ is compatible with the maps $\pi_r : U_{r+1} \to U_r$, so Ω factors into the inverse limit of the sequence (4.1), which we denote by U. Since the $\mathbb{Z}^2$-action on each U_r is transitive (in the strict sense that every orbit is the entire space), we see that Ω actually maps *onto* U. Denote the factor map thus defined by $\varphi : \Omega \to U$. Note that $\varphi(\omega)$ completely determines the sets $I_r(\omega), J_r(\omega)$

With a little more work, we can show that φ is almost-1-1 with respect to the action of T_u for any $u \notin \mathbb{Z}e_1 \cup \mathbb{Z}e_2$. In fact, another induction shows that given large n and a block $a \in \{s_n(\bullet), s_n(\circ)\}$, the sets $I'_r = I_{r-1} \setminus I_r$ and $J'_r = J_{r-1} \setminus J_r$ uniquely determine the pattern $a|_{I'_r \times J'_r}$. Now, for $\omega \in \Omega$, the point $\varphi(\omega)$ determines the sets $I_r(\omega), J_r(\omega)$, and these in turn determine sets I'_r, J'_r defined as above, and $a|_{I' \times J'}$ is determined by $\varphi(\omega)$. It is easy to check that $\mathbb{Z} \setminus \cup_{r=1}^{\infty} I'_r$ contains of at most one point i_*, and similarly $\mathbb{Z} \setminus \cup_{r=1}^{\infty} J'_r$ contains at most one point j_*, so w is determined by $\varphi(\omega)$ except possibly on the column $\{i_*\} \times \mathbb{Z}$ and the row $\mathbb{Z} \times \{j_*\}$. We call these the *exceptional rows and columns*, if they exist. The set of configurations containing exceptional rows or columns is wandering with respect to the action of any T_u with $u \notin \mathbb{Z}e_1 \cup \mathbb{Z}e_2$ (the argument is similar to corollary 4.2), so for any T_u-invariant measure on Ω almost every point is the unique pre-image of its image under φ. We have shown that for such u, (Ω, T_u) is an almost-1-1 extension of (U, T_u).

Let $\widehat{\Omega} = \widehat{\Omega}_p$ be the SFT cover of Ω_p promised by theorem 4.1. The division of $\omega \in \Omega$ into columns of type $r+1$ but not r uniquely determines the locations of blocks of the type $s_r(\bullet)$ and $s_r(\circ)$ in ω, because the right column/top row of these blocks can only be located on columns/rows which are type $r+1$ but not r. This shows that s has unique derivation, and it follows from corollary 4.2 that $(\widehat{\Omega}, T_u)$ is an almost-1-1 extension of (Ω, T_u), and hence of (U, T_u), for every $u \notin \mathbb{Z}e_1 \cup \mathbb{Z}e_2$.

Finally, fix $\omega \in \Omega$ and a segment $I \subseteq \mathbb{Z}$, and consider the segments $I'_n = I \times \{n\}$ obtained by embedding I in $\mathbb{Z}^2$ and translating it vertically a distance of n. Each I'_n lies in a row of type r for some $r = r(n)$, and every r occurs as $r(n)$ for some n. Now, it is easily seen that if a row of type r contains a $\bullet$ in the i-th column then a row of a different type cannot contain a $\bullet$ in this column. Hence I'_n can contain a $\bullet$ in column i only when it intersects a row of type r but not $r-1$. It follows that for any $|I|+1$ distinct values of r, at least one of them is such that when I'_n is contained in a row of type r then it does not contain any $\bullet$'s; in particular, one of the segments $I'_0, I'_1, \ldots, I'_{|I|}$ contains no $\bullet$'s. A similar statement holds for translates of $\{0\} \times I$ in the e_1 direction.

4.4. **Rectangular partitions of $\mathbb{Z}^3$.** An important role in our constructions will be played by $\mathbb{Z}^3$-SFTs whose configurations partition $\mathbb{Z}^3$ into rectangular regions in a special way. By a rectangle we mean a set of the form $\{i\} \times I \times J \subseteq \mathbb{Z}^3$ where $I, J \subseteq \mathbb{Z}$ are segments of integers, possibly infinite on one or both sides. Write $H_i = \{i\} \times \mathbb{Z} \times \mathbb{Z}$, and identify configurations on H_i with configurations in $\mathbb{Z}^2$; thus for a rectangle $\{i\} \times I \times J \subseteq H_i$ we will say that $|I|$ is the width and $|J|$ the height, and refer to e_2 as the horizontal direction and e_3 as the vertical one.

We next construct a $\mathbb{Z}^3$-SFT W and a factor map ρ from W into a subshift defined over the alphabet $\{\circ, \bullet\}$, such that for $w \in W$ the configuration $\rho(w)|_{H_i}$ consists of rows and columns of $\bullet$'s, and all other symbols are equal to $\circ$. Such a configuration can be naturally interpreted as inducing a partition of H_i into rectangles, e.g. with the convention that the bottom and left borders of a rectangle belong to the rectangle, and the other borders do not. We call such a configuration a *rectangular partition*.

We begin with the 2-dimensional SFTs $\widehat{\Omega}_3$ and $\widehat{\Omega}_5$ of the previous section, and extend each of them to a 3-dimensional SFT as follows. We identify configurations of $\widehat{\Omega}_3$ with configurations in the plane $\mathbb{Z}\times\mathbb{Z}\times\{0\}$, with $e_1\in\mathbb{Z}^2$ identified with $e_1\in\mathbb{Z}^3$ and $e_2\in\mathbb{Z}^2$ with $e_2\in\mathbb{Z}^3$, and extend the symbols in the e_3 direction; we obtain the system

$$\widehat{W}_3=\{x\in\Sigma^{\mathbb{Z}^3}\;:\;\exists\omega\in\widehat{\Omega}_3\text{ with }x(i,j,k)=\omega(i,j)\text{ for }i,j,k\in\mathbb{Z}\}$$

Similarly, we identify configurations of $\widehat{\Omega}_5$ with configurations in the plane $\mathbb{Z}\times\{0\}\times\mathbb{Z}$, with $e_1\in\mathbb{Z}^2$ identified with e_1 and $e_2\in\mathbb{Z}^2$ with e_3. We obtain the system

$$\widehat{W}_5=\{x\in\Sigma^{\mathbb{Z}^3}\;:\;\exists\omega\in\widehat{\Omega}_5\text{ with }x(i,j,k)=\omega(i,k)\text{ for }i,j,k\in\mathbb{Z}\}$$

We define W_3, W_5 similarly starting with Ω_3,Ω_5; there are natural factor maps $\widehat{W}_3\to W_3$ and $\widehat{W}_5\to W_5$ induced from the factor maps $\widehat{\Omega}_3\to\Omega_3$ and $\widehat{\Omega}_5\to\Omega_5$, respectively. Both $\widehat{W}_3$ and $\widehat{W}_5$ are SFTs, and that the action of T_u on each is an almost-1-1 extension of an isometric system as long as $u\notin\cup_{i=1}^3\mathbb{Z}e_i$.

Let $W=\widehat{W}_3\times\widehat{W}_5$, and define the factor map ρ on W so that $\rho(w',w'')(u)=\bullet$ if and only if one of the projections of w',w'' onto W_3,W_5, respectively, contains a $\bullet$ at u. Thus each $w=(w',w'')\in W$ induces, via $\rho(w)$, a partition of $\mathbb{Z}^3$ into rectangles. See figure 4.2.

Proposition 4.3. *Let $y=\rho(w)$ for some $w\in W$.*

(1) *There are at most finitely many planes H_i containing infinite rectangles in y.*
(2) *For each finite horizontal segment $I\subseteq\{0\}\times\mathbb{Z}\times\{0\}$, each $M>|I|$ and each $N\in\mathbb{N}$, there is a translate of I in the direction e_1 which is contained in some rectangle R induced from y, with width between M and $M\cdot 3^{|I|+2}$, and height $>N$.*

Proof. The first statement follows from the fact that points in $\widehat{\Omega}_p$ contain at most one exceptional row or column.

Next we verify the second statement. Fix a point $w\in W$ induced by a $w'\in\widehat{\Omega}_3$ and $w''\in\widehat{\Omega}_5$, fix a segment $I\subseteq\mathbb{Z}$, $M>|I|$ and $N\in\mathbb{N}$, and let $I'=\{0\}\times I\times\{0\}$. A translate $I'+ne_1$ is located between two vertical lines in H_n at distance ℓ from each other if and only if $\{0\}\times I+ne_1$ is in a column in w' between two $\bullet$'s at distance ℓ apart. The fact that this holds for some translate and $M\leq\ell\leq M\cdot 3^{|I|+2}$ follows from the remark at the end of section 4.3, and this occurs for a set of n's which has period 3^k for some k. Similarly, a translate $I'+ne_1$ is located in H_n between horizontal lines at distance $>N$ from each other if and only if ne_1 is on a column in w'' between $\bullet$'s at least $>N$ apart. This occurs for a set of n's with period 5^m for some m. Since $3,5$ are relatively prime, there is an n satisfying both simultaneously. □

From the first part of the proposition we deduce the following, as in corollary 4.2:

Corollary 4.4. *For any $u\in\mathbb{Z}^3\setminus H_0$ and for any T_u-invariant probability measure μ on W, the set of w such that $\rho(w)$ contains infinite rectangles has μ-probability 0.*

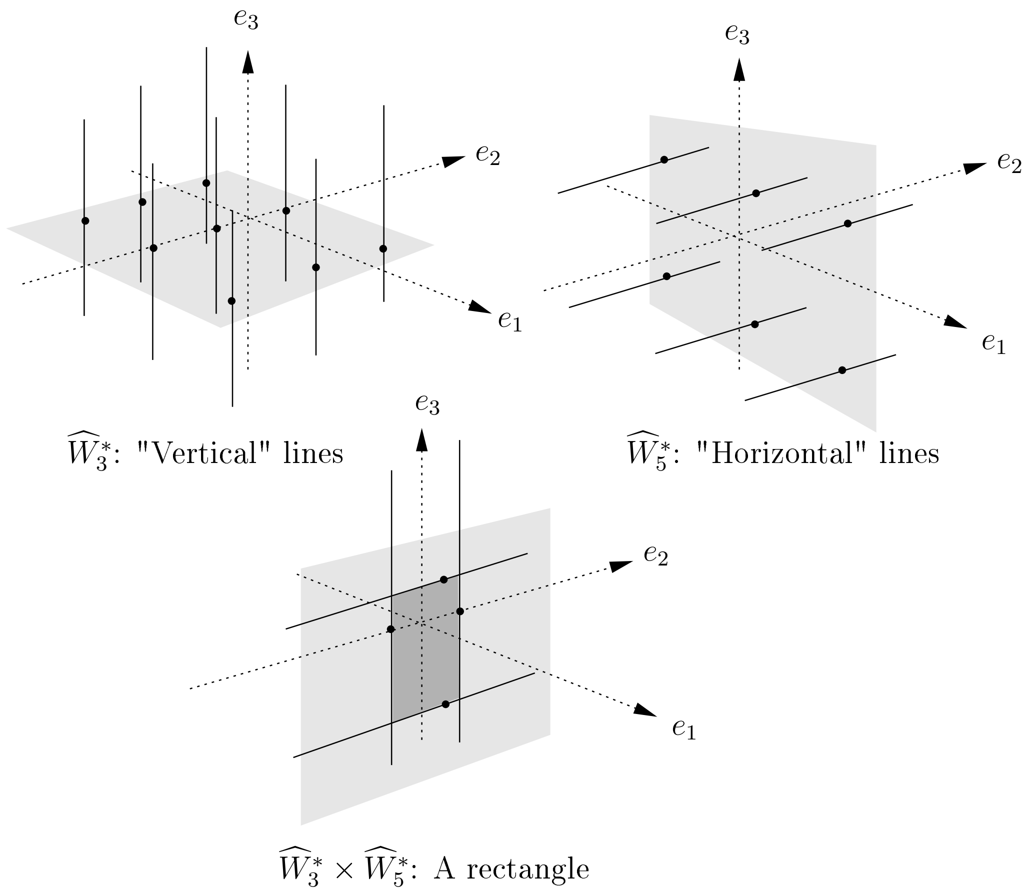

FIGURE 4.2. Two components of a point in W and an induced rectangle.

4.5. **Turing Machines in Rectangles.** The classical *Turing machine* is an automaton with a finite number of internal states which reads and writes data on a one-sided infinite array of *cells* indexed by $\mathbb{N}$, called the *tape*. Each cell contains one symbol from a fixed alphabet. The computation begins with the machine located at the 0-th (leftmost) cell and in a special initial state, and the tape is initially blank, or contains some data which is the input to the computation. The state of the tape along with the location and internal state of the machine are called a *configuration*; a configuration uniquely determines all future configurations. The computation proceeds in discrete time steps. At each iteration the machine is located at some cell. The machine reads the symbol at its current positions and, based on this data and on its internal state (and only on these parameters), it performs three actions: (a) it replaces the current data symbol with a new one, (b) it moves one cell to the left or to the right, and (c) it updates its internal state. The computation may halt after a finite number of steps if the machine either moves off the tape (steps left at cell 0) or enters a designated state, called the halting state. Barring these occurrences, the computation continues forever.

Although a very simple model, any algorithm written in a modern computer programming language can be implemented as a Turing machine, and it is generally accepted that any effective computation can be performed by a Turing machine; this is Church's thesis. For background and basic facts on this subject, see [HU79].

4.6. **Representing Turing machines in SFTs.** It is well known that one can use SFTs to represent the runs of Turing machines. Given a machine T, we first fix an encoding of the configurations of the machine as bi-infinite sequences of symbols (mark the cells at positions < 0 as "off limits" by using some special symbol), with each cell containing a symbol from the tape alphabet and possibly the state of the machine, indicating that the machine is located at that position. Now consider two-dimensional configurations in which each row represents a configuration of the machine, and is obtained by iterating by one step the computation in the row below it. The transition from row to row is determined locally, and can be encoded in the rules of an SFT, giving a system whose configurations describe infinite computations of the machine, assuming their rows represents states of the machine. To make this work one must allow the initial configuration of the machine to remain unchanged so that it can extend infinitely downward; then every infinite (non-halting) run of the machine can be represented as a two-dimensional array in which the initial configuration occurs for all negative times, say, and the computation starts at row 0; and no halting run can be represented. Of course, in addition to arrays representing runs the SFT will also contain "pathological" configurations which do not correspond to computations, such as configurations containing only data and no machine, tapes with multiple machines, or computations which extend back in time indefinitely and do not begin in an initial state.

If one wants better control of these matters, one can start with an "infrastructure" layer which partitions space into rectangles, such as the SFT W constructed in the previous section. Then we can superimpose a layer in which each rectangle represents a finite portion of a machine run on the rectangle, with the lower border of each rectangle initialized to a blank tape and the machine starting in the lower left corner in its initial state (this can be enforced by requiring that the data superimposed over a point in one of the rectangular partition's horizontal lines is the "blank" symbol, and over the intersection of a horizontal and vertical line there must be superimposed the machine in its initial state. Both these conditions are local). Note that we still have no control over initialization of infinite rectangles.

A crucial observation is the following. If the machine halts on blank input, then sufficiently large rectangles cannot be completed by this tiling scheme; since each configuration of W contains arbitrarily large rectangles this means the system we have defined is empty. On the other hand if the machine does not halt, then the system is not empty. One technical point to note is that there are configurations containing rectangles much higher than they are wide. In this case it can occur that the machine tries to move past the right border of the rectangle in the course of the computation. If this occurs then no further changes occur in the configuration, since once it is gone it cannot come back, but a pattern can still be superimposed over this rectangle. In this case the configuration on the rectangle does not coincide with the true computation viewed through this rectangle, but this does not affect our previous analysis.

Following Robinson, this construction can be used to prove that it is undecidable whether an SFT is empty. Indeed, it follows that if there were an algorithm for deciding emptiness of an SFT then we could use it to decide if a given Turing machine halts by constructing from T the SFT above and checking if it is empty.

4.7. **Real-time Turing machines.** It will be convenient for us to work with a slightly modified model of a Turing machine, in which the machine also receives data in "real time" (that is, in the course of the computation) from some external source. To model this we use machines which have two tapes, the *input tape*, which the machine can read but cannot write to, and whose state at each moment is determined externally and can change independently at each iteration of the computation; and the *memory tape*, which is initially blank, and which behaves like the ordinary tape: the machine can both read and write to it, and other than changes made by the machine, it retains its data unaltered from step to step. There is still only one machine head, which at each step reads a symbol from each of the tapes.

This model will occur in our construction as follows. We will start with the product $W \times X$, where X is some SFT, and superimpose a layer so that rectangles induced by W contain representations of runs of a certain Turing machine of the modified type. Note that the rows of each rectangle contain a row of symbols coming from the X component, and these will serve as the data on the input tape, so successive rows in X provide successive rows of input. The memory tape is simulated in the usual manner as part of the machine's configuration, represented by symbols from the superimposed layer.

In our application we will want the machine to be able to read (and store for later use) k symbols from the i-th cell of the input tape in the course of k consecutive time steps, where i, k are determined in the course of the computation. This cannot be implemented in the machine model we have described, because after reading a symbol the machine must take time to store it out of the way, and by the time it gets back to the i-th cell it will have missed one or more input symbols, which are constantly changing.

However, we can implement this functionality in the SFT representation with the following trick. Assume for simplicity that the input language is $\{0, 1\}$. We assume that the memory language contains the symbols $0, 1$, and in addition a pair of special symbols $\widehat{0}, \widehat{1}$. These special symbols are not merely symbols, but are simple automata in their own right: when $\widehat{0}$ or $\widehat{1}$ appears in the memory tape, it tends to move one step to the right with each time step, overwriting whatever was there before (the cell it previously occupied becomes blank, unless written to by the machine or moved into by another special symbol). The only thing that can stop a moving special symbol is an ordinary symbol 0 or 1, or a special stop symbol '|'. When a moving symbol $\widehat{0}$ or $\widehat{1}$ comes up against a $0, 1$ or '|' from the left it turns into an ordinary 0 or 1, respectively. Furthermore this transformation is instantaneous: a special symbol cannot be the neighbor of $0, 1$ or '|'. Thus a change from special to ordinary symbol propagates instantly to the left. For example, if we have the sequence

$$x\widehat{0}\widehat{1}\widehat{1}y|$$

on the memory tape, and $x, y \neq 0, 1, \widehat{0}, \widehat{1}, |$, then after one more time step we will have

$$xb011|$$

where b is the blank symbol. Breaking this down, what happened is that each of the special symbols moved one step to the right (the rightmost one overwriting the y), came up against a '|', turned into a 1. Therefore its neighbor on the left turned into a 1 and the next symbol to the left into a 0. This instantaneous transformation cannot be implemented in a one-dimensional automaton on rows because it requires transmission of information across long distances, but is easily implemented as part of the SFT rules. It is similar to the instantaneous machines of Robinson [Rob71], and we omit the details. Finally, this instantaneous change will also affect the machine's state as follows: there is a special state s of the machine so that if the machine was located at a cell containing a special symbol $\widehat{0}$ or $\widehat{1}$, and in the next step its neighbor on the right is $0, 1$ or '|', then the machine enters state s.

Returning now to our objective, if the machine wants to read k input symbols from cell i at consecutive times, it first erases the memory cells between i and $i + k$, and writes a $'|'$ at cell $i+k$. It then returns to cell i, and enters a special state t. While in this state all it does is read a symbol 0 or 1 from the input tape, and print $\widehat{0}$ or $\widehat{1}$, respectively, on the memory tape. Things have now been arranged so that the sequence of symbols printed moves one step to the right with each time step, making room for the new symbol, and this continues until k symbols have been read. At that point the segment $[i; i + k - 1]$ is filled with $\widehat{0}$'s and $\widehat{1}$'s, and the rightmost has come up against the "stop" symbol; this transmutes the symbols to ordinary symbols and forces the machine out of the state t and into the state s, at which point it resumes its usual operation, but has at its disposal the k symbols of input recorded to its right on the memory tape.

We introduce one more modification: the machine may also run on finite tapes, i.e. tapes which extend only a finite distance to the right. We enable the machine to detect when it is near the right side of the tape, and use this in its decision procedure. We note that we have used memory to the right of a cell to store the input data captured at that cell, and this won't work near the right edge of the tape, but one easily introduces a similar procedure, whose details we omit, which uses memory to the left of the cell. We remark that this feature will be used when we run machines that need to read data from the entire width of their finite tape several times. If not for this, we could just have defined that the machine halts when it tries to store data off the right end of the tape.

We call machines of the type above, which are Turing machines which can run on finite tapes and are capable of capturing k bits of input in real time for arbitrary k, *real-time Turing machines*. Together with the previous discussion, we have proved the following theorem:

Theorem 4.5. *Let T be a real-time Turing machine with input language Σ, and let $X \subseteq \Sigma^{\mathbb{Z}^3}$ be an SFT. Then there is an SFT Z superimposed over $W \times X$ so that $(w,x) \in W \times X$ is represented in Z if and only if, for each rectangle R of dimensions $m \times n$ induced by w, when the machine T is run and input data defined by the array $x|_R$, it runs at least n steps or leaves the region $[1;m]$ in less than n steps.*

We remark that any ordinary Turing machine can be implemented trivially as a real-time Turing machine by simply not using any of the added functionality.

5. Realizing EDS as subactions of SFTs and sofic shifts

In this section we prove theorems 1.4 and 1.6. We begin with theorem 1.4, whose proof is slightly easier.

5.1. **Realizing ESSs as subactions of sofic shifts.** We prove theorem 1.4 for $\mathbb{Z}$-systems, that is, we show that every one-dimensional ESS (= expansive EDS) is the subaction of a 3-dimensional sofic shift. The proof of the general case is very similar, requiring one to define higher-dimensional analogues of the rectangular partitions described in the previous section, and a definition of Turing machines with multidimensional tapes. These modifications are straightforward, and in the interest making the presentation readable we omit them.

Let $L \subseteq \Sigma^*$ be a RE set of finite sequences over Σ; we are out to realize the system $S_L \subseteq \Sigma^{\mathbb{Z}}$ as the subaction of a 3-dimensional sofic shift. We will do so for the subaction generated by the transformation $\tilde{T} = T_{e_1+e_2+e_3}$.

Let $Y \subseteq \Sigma^{\mathbb{Z}^3}$ be the shift of finite type defined by the conditions $y(u) = y(u+e_1) = y(u+e_3)$ for $u \in \mathbb{Z}^3$, so symbols are constant in directions e_1 and e_3. Let $\pi : Y \to \Sigma^{\mathbb{Z}}$ be the map $(\pi y)(n) = y(ne_2)$; a moment's reflection shows that π conjugates $(Y, \tilde{T})$ to the full one-dimensional shift over Σ. Let $Y_L = \pi^{-1}(S_L)$.

Let W be the system defined in section 4.4, and set $Z_0 = Y \times W$. For a rectangle $R = \{k\} \times I \times J$ and $y \in Y$ we define $\pi'_R(y) \in \Sigma^*$ to be the word of length $|I|$ induced by y on the bottom row of R, that is: $(\pi'_R y)(i) = y(ke_1 + ie_2)$ for $i \in I$. See figure 5.1

We now superimpose a layer over Z_0 whose object is to "kill" points $(y,w) \in Z_0$ with $y \notin Y_L$. This is done using theorem 4.5, utilizing the rectangles R induced by w to represent runs of Turing machines which use as input the pattern $\pi'_R(y)$ induced by y on the bottom edge of the rectangle, and use the vertical (e_3) direction to represent time (since symbols in Y are constant in the e_3 direction, the input does not change in real-time, and we may use "traditional" Turing machines).

The machine we run on the rectangles performs the following computation: it generates the elements of L one after the other (this can be done by the assumption that L is RE), and for each word it checks if the word appears in the input. We denote the resulting SFT by Z.

We claim that the effect of this is that (y,w) is represented in Z if and only if $y \in Y_L$. By theorem 4.5, we need to check that if $y \in Y_L$ then in every rectangle the machine does not halt in fewer steps than the height of the rectangle, and that if $y \notin Y_L$ then there is some rectangle

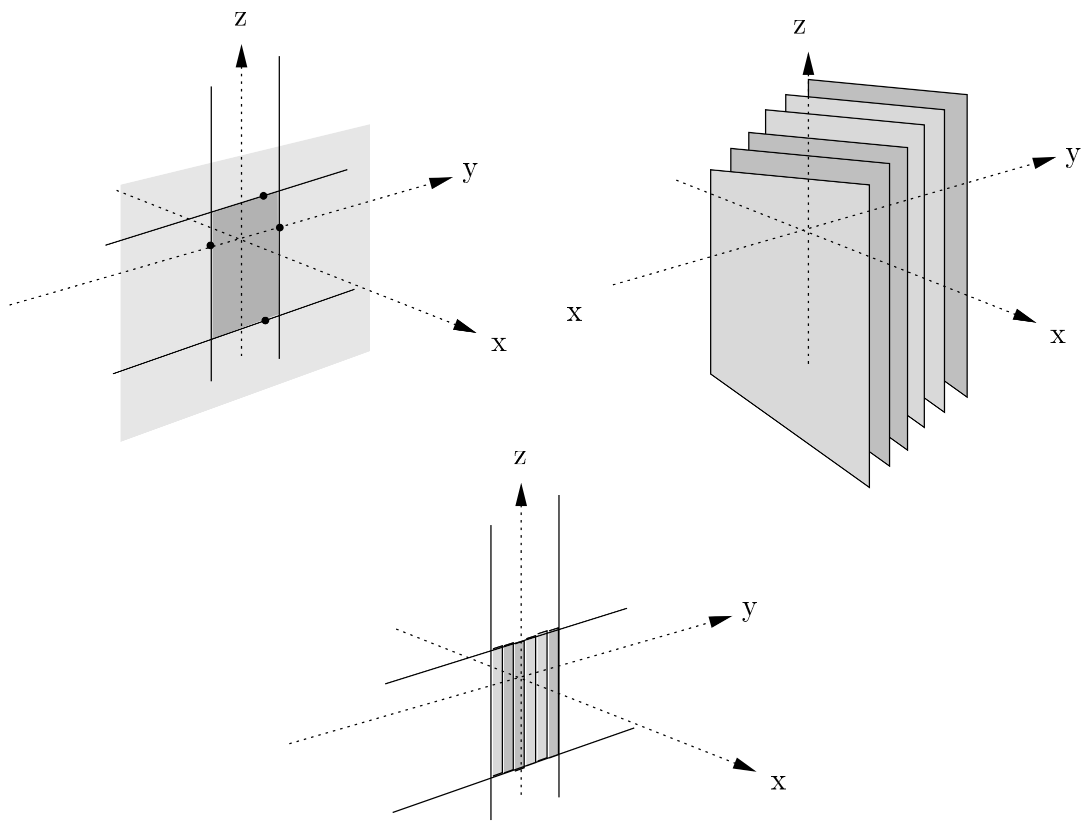

Figure 5.1. The configuration of stripes induced by Y on one of the rectangles induced from W. The shades of gray represent symbols on planes perpendicular to e_2 (or densities of symbols in the construction of section 5.3

for which the machine will halt in fewer steps. The first statement is obvious. As for the second, for any subword of $\pi'(y)$, corresponding to the word $y|_{\{0\}\times I\times\{0\}}$ for a finite segment $I \subseteq \mathbb{Z}$, there are translates of I in the direction e_1 contained in rectangles of arbitrarily large width and of height arbitrarily large as a function of the width; this follows from the second part of proposition 4.3. The word induced by y on this translate is the same as the original word. If the word is in L, then when the width and height are large enough the machine will halt prematurely.

One must be a little cautious regarding the analysis above, since it depends on the implementation of the algorithm the machine is running. It is important that the machine first calculate the n-th disallowed word, and only then check it against its "input". Given n, this first stage (where we compute the n-th word of L) uses some amount of memory and time which depends only on n and not on either the dimensions of the rectangle or the input, and therefore can be performed on any rectangle which is wide and high enough. On the other hand, the second stage (where we check if the word appears in the input) requires an amount of memory which depends only on n, and time which depends only on n and on the width of the rectangle, so it can be performed on any sufficiently wide rectangle as long as the height is sufficiently large compared to its width. These observations imply that if $\pi'(y)|_I$ contains the n-th disallowed word, then there is a rectangle wide enough and high enough for this to be discovered.

This completes the proof of theorem 1.4, since Y_L is a sofic shift via the $\mathbb{Z}^3$-factor map $Z \to Y_L$ which forgets the second coordinate (the W-coordinate) and the machine symbols (we need only re-parametrize the action to transform $\widetilde{T}$ into T^{e_1}). More is true: as we show next, this map, which is automatically a factor map also with respect to the action of $\tilde{T}$, is an ATIE with respect to the $\tilde{T}$-action. This proves a special case of theorem 1.6 for ESSs.

To see that the factor map is an ATIE with respect to $\tilde{T}$, note that we have the sequence of factors

$$Z \to Y_L \times W \to Y_L$$

(the first map forgets the Turing machine symbols, and the second factors onto the first coordinate). Since $(W, \tilde{T})$ is an almost-1-1 extension of an isometric system, in order to deduce that $Z \to Y_L$ is ATIE we need to show that $Z \to Y_L \times W$ is an ATIE with respect to $\tilde{T}$. To see this, merely note that given $(y, w) \in Y_L \times W$, the superimposed layer representing the machine run is completely determined on finite rectangles; it is undetermined only on infinite rectangles (actually, only on rectangles which are infinite in the $-e_2$ or $-e_3$ directions). However, this occurs with probability 0 for any $\tilde{T}$-invariant measure by corollary 4.4. Hence $Z \to Y_L \times W$ is almost-1-1 with respect to $\tilde{T}$, and we are done.

5.2. **The Striped System.** We need one more auxiliary construction. Let $\widehat{\Omega}_2$ be the system defined in section 4.3. In [HM07] it was shown how to superimpose a layer over $\widehat{\Omega}_2$ in such a way that each row is colored 0 or 1, the rows whose coordinates are in $J_{r+1} \setminus J_r$ all have the same color, and any combination of colors occurs subject to these restrictions; in particular if there is a row not of the above type it may have any color; there is at most one such row (in [HM07] this was done with columns in place of rows, but the modification is trivial). We denote this system S and call it the *striped system*. The main property of this system that we will use is that, to each $s \in S$, there is associated the density $\delta(s)$ of 1's, which is well defined, and if s has a unique binary representation then the n-th digit in its binary expansion is a 1 if and only if the rows with coordinates $J_{n+1} \setminus J_n$ are marked 1. It was shown in [HM07] that for any $s \in S$ and any $\varepsilon > 0$ one can estimate $\delta(s)$ with error $\varepsilon \in (0, 1)$ by observing the $0,1$-pattern induced by s on any vertical segment $\{i\} \times J$, as long as $|J| > 10/\varepsilon$. We also note that if we fix $0 < \delta < 1$ with unique binary representation, then the set $S_\delta = \{s \in S : \delta(s) = \delta\}$ is a closed subset of S, and the projection from it to $\widehat{\Omega}_2$ is 1-1 except when the image contains an exceptional row, in which case the projection is 2-to-1 (because the only thing not determined is the color of that row).

Finally, we note that in [HM07] the coloring on rows in S was performed in such a way that, given the coloring of rows, all other auxiliary symbols were determined.

5.3. **Realizing EDSs.** For simplicity of notation we prove theorem 1.6 for $d = 1$. Let $X \subseteq (\{0,1\}^{\mathbb{N}})^{\mathbb{Z}}$ be an effectively closed subsystem and let $U_1, U_2, \ldots$ be a recursive sequence of cylinder sets whose union is the complement of X. More precisely, we are given an algorithm which, on input k, outputs a function $c : I \to \mathcal{C}$, where $I \subseteq \mathbb{Z}$ is finite and $\mathcal{C}$ is the parametrized family of cylinder sets in $\{0,1\}^{\mathbb{N}}$, such that $x \in U_k$ if and only if $x(i) \in c(i)$ for $i \in I$.

The method of realizing X as an ATIE of the subaction of an SFT is similar to that in section 5.1. We will construct a 3-dimensional SFT Y so that the configurations induced by $y \in Y$ on each translated plane

$$F_n = \mathrm{span}_{\mathbb{Z}}\{e_1, e_3\} + ne_2$$

encodes an infinite sequence $s_n = (s_n(1), s_n(2), \ldots) \in \{0,1\}^{\mathbb{N}}$ (rather than a single symbol as in the expansive case), and we will want this to be done in a manner which is invariant to the shifts T_{e_1} and T_{e_3}. Assuming we have such a representation, we can define $\pi : Y \to (\{0,1\}^{\mathbb{N}})^{\mathbb{Z}}$ by $\pi(y) = (s_n)_{n\in\mathbb{Z}}$; this is a factor map from $(Y, \tilde{T})$ to $((\Sigma^{\mathbb{N}})^{\mathbb{Z}}, T_1)$, where $\tilde{T} = T_{e_1+e_2+e_3}$. Then, as in the expansive case, we will superimpose another layer which kills all points $y \in Y$ which don't map to X under π.

There are many ways to extract a sequence $(s_n(i))_{i=1}^{\infty}$ from a two-dimensional configuration on F_n: the trivial one is to fix an enumeration of the elements of $\mathbb{Z}^2$ and use this to identify each configuration in F_n with an element of $\Sigma^{\mathbb{N}}$. This solution will not work for us for two reasons. First it is not invariant to T_{e_1}, T_{e_3}. Furthermore, in order to perform the selection stage of our strategy we will need the information encoding the sequence in each plane F_n to be "spread all over"; this is because each run of a Turing machine has access to only a small region of space, and from this sample it must extract sufficient information about the encoded sequences to make its decision. To accomplish both invariance and redundancy, we encode a sequence in configurations by using densities of symbols. By making the distribution of symbols in F_n's sufficiently uniform we can guarantee that the configuration on the intersection of each large enough rectangle with F_n gives sufficient information to decode increasingly long portions of the encoded sequence.

Here are the details. We start with the 2-dimensional system $\widehat{\Omega}_2$ of section 5.2, and extend it to a 3-dimensional system by sending each 2-dimensional configuration $\omega \in \widehat{\Omega}_2$ to the three-dimensional configuration w which is constant in the direction e_2, and on F_0 the pattern is obtained from ω by identifying the direction $e_1 \in \mathbb{Z}^2$ with $e_1 \in \mathbb{Z}^3$, and $e_2 \in \mathbb{Z}^2$ with $e_3 \in \mathbb{Z}^3$; that is, for each $\omega \in \widehat{\Omega}_2$ a point y is defined by

$$y(ie_1 + je_2 + ke_3) = \omega(i,k)$$

This defines a three-dimensional SFT which we denote Y_0. In particular, columns in $\widehat{\Omega}_2$ correspond to lines in direction e_3 in Y_0. Also, the $\tilde{T}$-action on Y_0 is an almost-1-1 extension of an isometric action.

Next, in each plane F_n we extend the configuration to a striped system as in section 5.2; we allow the stripes of each translate of the xz-plane to be colored independently. Due to the embedding of $\mathbb{Z}^2$ in $\mathbb{Z}^3$ which we have chosen, stripes now form lines in the direction e_1. Denote the resulting system by Y.

For $y \in Y$ we denote by $\delta_n(y)$ the density of 1's occurring in the plane F_n, so $\delta_n(\tilde{T}y) = \delta_{n+1}(x)$. The pattern induced by x on F_n allows us to recover $\delta_n(y)$; this information can also be obtained from the pattern induced on any vertical line (i.e. in direction e_3) contained in F_n, and furthermore in order to recover $\delta_n(y)$ up to an error of 2^{-k} it suffices to examine the pattern on any vertical segment of length e.g. 2^{k^2} in F_n.

We will use $\delta_n(y)$ to associate a sequence s_n of 0's and 1's to $y|_{F_n}$. The most straightforward way to do this would be to set s_n to the digits of the binary representation of $\delta_n(y)$, but then we run into ambiguities related to the non-uniqueness of binary expansions of dyadic numbers. To avoid this, we do not define this association $y|_{F_n} \mapsto s_n$ for every $y \in Y$, but rather only for those y's such that for every n, the binary expansion of $\delta_n(y)$ is of the form $\delta_n(y) = 0.b_1 0 b_2 0 b_3 0 \ldots$ for some sequence $b_i \in \{0,1\}$. The set $Y' \subseteq Y$ so that $\delta_n(y)$ is of this form for all n is a closed subset of Y, though not an SFT. For convenience, to each $y \in Y \setminus Y'$ we define $\delta_n(y) = 0,0,0,\ldots$ for all n.

Let $\pi : Y \to (\Sigma^{\mathbb{N}})^{\mathbb{Z}}$ denote continuous shift-commuting map $y \mapsto (\delta_n(y))_{n\in\mathbb{Z}}$ from $(Y, \hat{T})$ to $(\mathbb{Z}^{\mathbb{N}})^{\mathbb{Z}}$. We now superimpose another layer over Y which kills points outside of Y' and also point from Y' which do not project into X under π. As in the expansive case, we first take the product of Y with W, and over this product we superimpose another layer which represents the run of a Turing machine over rectangles induced by W. We use real-time machines as described in section 4.6, which allows the machine to read arbitrarily long vertical segments of data from the Y-layer. See figure 5.1.

The machine implements the following algorithm. We denote by y_R the two-dimensional array of input symbols on a rectangle R, which in our setting comes from the restriction of a point $y \in Y$ to a rectangle R induced by some $w \in W$. The rows of y_R represent the input at a given time and the vertical one represents passage of time. The algorithm iterates over integers $k \in \mathbb{N}^2$ in some order, and for each k it applies the given algorithm which calculates U_k. Suppose U_k is the basis element specified by a finite subset $I \subseteq \mathbb{Z}$ and $c : I \to \mathcal{C}$, where $\mathcal{C}$ is the collection of cylinder sets in $\{0,1\}^{\mathbb{N}}$. Assume that each cylinder set $c(i)$, $i \in I$ is defined by indices in a set $J \subseteq [1;m] \subseteq \mathbb{N}$. Next, the algorithm iterates over the tape, and at the i-th cell it captures 2^{m^2} vertical bits from its input, and does this also for cells $i+1, i+2, \ldots, i+|I|-1$. This data suffices to determine, up to m bits, the densities $\delta_j(y)$ of the planes F_j to which the cells $i, i+1, \ldots, i+|I|-1$ belong. If these finite expansions are not of the form $0.b_1 0 b_1 0 \ldots$ the algorithm halts. Finally, with the information at hand the algorithm can determine whether some translate of U_k intersects the projection under π of the current point, and if the answer is affirmative the algorithm halts.

As before, to make this work we assume that the machine first calculates U_k and only then checks it against its "input". This first stage uses some amount of memory and time which depend only on k, and therefore can be performed on any wide enough rectangle. Again, the second stage requires an amount of memory depending only on $k, |I|, |J|$ and time which is additionally a function of the width of the rectangle, so can be performed on any rectangle whose height is sufficiently large compared to its width. These observations, and the fact that

W induces admissible partitions, imply that every subword encoded by a point in Y will be checked by some machine run in some rectangle.

The above implies that if we let Z denote the resulting SFT then it consists of extensions of pairs (y,w) with $y \in Y'$ and such that y projects to X via π.

To complete the proof we must verify this projection is an ATIE with respect to T_1. As at the end of section 5.1 we have map $\tau : Z \to Y$ which is an ATIE with respect to its image, but instead of the image of τ being the sought-after system, as in section 5.1, we now must project one more time, using the map $\pi : \tau(Z) \to Y$. But this map is also an ATIE, because we can break it into

$$\tau(Z) \cong X \times Y_0 \to X \times U \to X$$

where the first map acts coordinate-wise as the identity on the first component and as the ATIE $Y_0 \to U$ on the second, where U is an isometric system; and the second factor map is projection onto the first component. Finally, since the composition of ATIEs is an ATIE, we are done.

5.4. **Variations.** With a little more effort the equicontinuous system that appears in the ATIE of theorem 1.6 can be made to be transitive, i.e. an odometer. This can be achieved if, instead of the systems Ω_p which we constructed in section 4.3, we work with systems in which the spectrum of T_1 and T_2 come from distinct primes. In this case the action of T_u for $u \notin \mathbb{Z}e_1 \cup \mathbb{Z}e_2$ will be transitive, i.e. an odometer. We omit the details.

We also note that other types of substitution systems may be used instead of odometers. For the dynamical possibilities this provides see [Moz89].

6. Constructions and counterexamples

6.1. **The Relation Between subactions of SFTs and Limit Sets of CAs.** There is a close connection between the action of a CA on the natural extension of its limit set, and one-dimensional subactions of SFTs and sofic shifts: given a system belonging to one of these classes one can construct a member of the other class which captures most of the dynamics of the original system.

In order to go from a CA f acting on $\Sigma^{\mathbb{Z}^d}$ to an SFT, one considers the subshift $X \subseteq \Sigma^{\mathbb{Z}^{d+1}}$ defined by the property that $x \in X$ if and only if $x|_{\{i+1\}\times\mathbb{Z}^d} = f(x|_{\{i\}\times\mathbb{Z}^d})$ for every $i \in \mathbb{N}$, where we identify configurations on $\{i\} \times \mathbb{Z}^d$ with configurations on $\mathbb{Z}^d$ in the obvious way. Since f acts by a local rule this system is an SFT, and since for $x \in X$ the sequence of configurations $(x|_{\{n\}\times\mathbb{Z}^d})_{n\in\mathbb{Z}}$ constitutes a two-sided f-orbit, the subaction (X, T_1) is isomorphic to the natural extension the action of f on its limit set.

To go the other way, suppose that $X = S_L \subseteq \Sigma^{\mathbb{Z}^d}$ is an SFT defined by a finite set $L \subseteq \Sigma^E$ of disallowed patterns, $E \subseteq \mathbb{Z}^d$ finite (a standard argument shows that every SFT is of this form). We construct a d-dimensional CA which has similar dynamics to (X, T_1). We do this by introducing a "destructive" symbol which comes into being at sites where the SFTs rules are broken, and "spreads"; and on the other hand the CA acts like T_1 on legal configurations. To be precise, let $*$ be a symbol not appearing in Σ and let $\Delta = \Sigma \cup \{*\}$. Define a CA $f : \Delta^{\mathbb{Z}^d} \to \Delta^{\mathbb{Z}^d}$ acting on $x \in \Delta^{\mathbb{Z}^d}$ according to the rules:

- If $x(u) = *$, or $x(u \pm e_i) = *$ for some $i = 1, \ldots, d$, or $(T_u x)|_E \in L$, then $(fx)(u) = *$;
- otherwise $(fx)(u) = x(u + e_1)$.

Note that $X \subseteq \Sigma^{\mathbb{Z}^d} \subseteq \Delta^{\mathbb{Z}^d}$, and the restriction of f to X acts like the shift T_1. On the other hand, if $x \in \Delta^{\mathbb{Z}^d} \setminus X$ then fx contains a $*$ and this symbol will spread: for any $u \in \mathbb{Z}^d$ we will have $(f^n x)(u) = *$ for all large enough n. Hence the configuration consisting of $*$'s is the unique attracting point of $\Delta^{\mathbb{Z}^d} \setminus X$ under f. It follows that the only nontrivial dynamics of f can occur in X, where f acts like T_1.

6.2. **Systems which cannot be realized as subactions.** In this section we present some examples of systems which cannot be realized as subactions of SFTs or sofic shifts. The first class of examples are the odometers, defined as equicontinuous, transitive actions on infinite spaces. There exist EDS of this type: it is easy to check that the map $x \mapsto x + 1$ on the group of 2-adic integers is an EDS. As is well known, every automorphism of this system is also given by a translation. Thus if this system could be realized as a subaction of a $\mathbb{Z}^d$-action it would follow that the $\mathbb{Z}^d$ action is itself an equicontinuous action; and it is well known the any subshift on which the shift acts isometrically is finite, a contradiction.

Since a direct proof is not long we include it.

Proposition 6.1. *Odometers cannot be realized as the subaction of a symbolic system.*

Proof. Suppose to the contrary that $U \subseteq \Sigma^{\mathbb{Z}^d}$ for some d and (U, T) is an odometer for $T = T_1$ (the proof for $T = T_u$is the same). Fix a compatible metric on U and choose $\varepsilon > 0$ so that $d(u, u') < \varepsilon$ implies $u(0) = u'(0)$. It is well known (and not hard to check) that for every $\varepsilon' > 0$ there is an $n > 0$ so that $d(T_1^n u, u) < \varepsilon'$ for every $u \in U$, and in particular this holds for $\varepsilon' = \varepsilon$. Now for any $u \in U$ we see that for $v \in \mathbb{Z}^d$ we have

$$d(T_v u, T_1^n (T_v u)) < \varepsilon$$

so $u(v) = u(v + ne_1)$. This holds for all $v \in \mathbb{Z}^d$, so as a configuration in $\Sigma^{\mathbb{Z}^d}$ we have that u has period n in direction e_1, hence the action of T_1 on it is periodic. But this is impossible because a transitive isometric action is minimal and therefore, if it's infinite, cannot have periodic points. □

Next, we exhibit an ESS which is not the subaction of an SFT, although by theorem 1.4 it surely is a subaction of some sofic shift. Recall that the (topological) Chacon system is obtained by the following process. Define words $a_n \in \{0, 1\}^*$ by setting $a_1 = 0$, and given a_n define $a_{n+1} = a_n a_n 1 a_n$, so

$$a_2 = 0010 \quad , \quad a_3 = 0010001010010 \quad \ldots$$

The Chacon system $X \subseteq \{0, 1\}^{\mathbb{Z}}$ is the subshift such that a finite word appears in X if and only if it appears in some a_n.

The condition that a word appear as a subword of some a_n is decidable. Indeed, it is easy to show by induction that for each k, every a_n for $n \geq k$ is a concatenation of the words a_k and $a_k 1$. Thus if b is a word whose length does not exceed that of a_k, then b is a subword of a_n if and only if it is a subword of $a_k a_k$ or of $a_k 1 a_k$. It follows that the set of subwords of X is recursive, so X is an ESS.

We remark that this argument can be applied to show that many other constructions in symbolic dynamics give ESSs; it works for any explicit construction by block concatenation.

Proposition 6.2. *Let Z be an SFT such that the subaction (Z, T_1) factors onto the Chacon system (X, T) via $\pi : Z \to X$. Then $|\pi^{-1}(x)| > 1$ for almost every $x \in X$ with respect to the unique invariant probability measure on X. In particular, (X, T) is not the subaction of an SFT.*

Proof. Let $X_0 \subseteq X$ be the G_δ subset of points with a unique pre-image, and let μ be the unique invariant measure on X, which is ergodic. Note that X_0 is invariant under T, so $\mu(X_0) = 0$ or 1.

Assume that $\mu(X_0) = 1$. Let $\widetilde{\mu}$ denote the lift of μ to $\pi^{-1}(X_0)$, so $\widetilde{\mu}$ is an invariant measure on Z, and it is the only one since any other invariant measure would have to be supported on $Z \setminus \pi^{-1}(X_0)$, hence would project under π to an invariant measure on X supported on $X \setminus X_0$, hence is different from μ, a contradiction.

Each of the shifts T_i maps $\widetilde{\mu}$ to a T_1-invariant measure, so by uniqueness of $\widetilde{\mu}$ we see that $T_2, \ldots, T_d$ act as automorphisms on the measure preserving system $(Z, \widetilde{\mu}, T_1)$. By a theorem of del Junco [dJ78], μ (and hence $\widetilde{\mu}$) has minimal self joinings, and in particular has no non-trivial automorphisms. Thus for $\widetilde{\mu}$-almost all $z \in Z$ the shifts T_i act as powers of T_1. Hence for a fixed typical z_0, we have $T_i z_0 = T_1^{n(i)} z_0$ for $i = 2, 3, \ldots, d$ and some integers $n(2), \ldots, n(d)$.

The group $\mathbb{Z}^d$ is generated by $T_1' = T_1, T_2' = T_1^{-n(2)} T_2, \ldots, T_d' = T_1^{-n(d)} T_d$, and the $\mathbb{Z}^d$-action they generate on X is also an SFT, so by this re-parametrization of the action we may assume that $n(2) = \ldots = n(d) = 0$. Now the action on z_0 is trivial for $T_2, \ldots, T_d$. Let $R > 0$ be the maximum diameter of a pattern defining Z (with respect to this new parametrization). Since the alphabet is finite we can find $m < n$ so that $z_0((m+i)e_1) = z_0((n+i)e_1)$ for $1 \leq i \leq R$, and hence $z_0((m+i)e_1 + u) = z_0((n+i)e_1 + u)$ for any $u \in \operatorname{span}_{\mathbb{Z}}\{e_2, \ldots e_d\}$.

Write $k = n - m$; it follows that the point z_1 defined by $z_1(\sum_{i=1}^d s_i e_i) = z_0(s_1' e_1)$ for $s_1' = s_1 \bmod k$ belongs X, since we are merely "gluing" together patterns on strips of the form $[m; n + R - 1] \times \mathbb{Z}^{d-1}$, the gluing taking place along the boundary of depth R, on which the patterns agree.

For the point z' we now have that $z'(u) = z'(u + ke_1)$ for any $u \in \mathbb{Z}^d$, i.e. it is periodic in the e_1-direction. Hence its image in X is as well. But the Chacon system does not contain any periodic points, being infinite and minimal; a contradiction. □

This proof works generally for any $\mathbb{Z}^k$-system X with trivial centralizer.

It remains an interesting open question whether the Chacon system can occur as a finite-to-one factor of the subaction of an SFT. Another is whether a uniquely ergodic subaction of an SFT can be measure-theoretically isomorphic to the Chacon system.

7. Entropy

7.1. **Entropy.** The entropy of a dynamical system is a non-negative number measuring the asymptotic rate of growth of the number of distinct orbits at smaller and smaller scales. A definition for the general setting may be found in [Wal82]. For our purposes the entropy of a subshift $X \subseteq \Sigma^{Z^d}$ may be defined by

$$h(X) = \lim_{n\to\infty} \frac{1}{|F_n|} \log \#\{a \in \Sigma^{F_n} \,:\, a \text{ appears in } x\}$$

where $F_n = \{1, 2, \ldots, n\}^d$ is the d-dimensional cube of side n. This limit can be shown to exist and converges to its limit from above, and is decreasing along the sequence $n = 2^k$, $k = 1, 2, \ldots$.

To define entropy for a totally disconnected system X, let $\mathcal{P}_1, \mathcal{P}_2, \ldots$ be a refining sequence of closed and open partitions of X whose union, together with all shifts of atoms by the action, separates points in X. Let X_n be the symbolic factor defined by $\mathcal{P}_n$ (that is, if $\mathcal{P}_n = \{A_{n,1}, \ldots, A_{n,k(n)}\}$, then $X_n \subseteq \{1, \ldots, k(n)\}^{\mathbb{Z}^d}$ such that $y \in X_n$ if and only if there is some $x \in X$ with $T^u x \in A_{n,y(u)}$ for $u \in \mathbb{Z}^d$). Then $h(X) = \lim h(X_n)$.

Entropy is an isomorphism invariant. If $Y \to X$ is a factor then $h(Y) \geq h(X)$, and $h(X \times Y) = h(X) + h(Y)$. Also if $X \subseteq Y$ is a subsystem than $h(Y) \geq h(X)$. If $X_1 \supseteq X_2 \supseteq \ldots$ are symbolic systems then $h(X_n) \searrow h(\cap X_n)$. The entropy of a non-invertible system is the entropy of the natural extension of the original system.

There is a notion of entropy $h(\mu)$ for invariant measure μ on X, and the variational principle gives the relation $h(X) = \sup_\mu h(\mu)$, where μ ranges over all invariant probability measures on X (see [Wal82]). This is true also for non-invertible systems.

Applying the above we see that if X is an SFT and f is the CA associated to X in section 6.1, then $h(X, T_1) = h(\Lambda, f)$, where Λ is the limit set of f. To see this, note that except for the Dirac measure on the fixed point, there is a 1-1 correspondence between f-invariant measures on Λ and T_1-invariant measures on X, and corresponding measures are isomorphic, so have the same entropy. Since the measure on the fixed point has entropy 0, the claim follows.

7.2. **A characterization of the entropy of EDS.** In this section we prove theorem 1.8: $h \geq 0$ is the entropy of an EDS if and only if it is the $\liminf$ of a recursive sequence, or equivalently, there is a recursive array $(m, n) \mapsto h_m(n) \in \mathbb{Q}$ with $h_m(n) \searrow h_m \nearrow h$. To see that $\liminf a_n$ has this form when $(a_n)_{n=1}^\infty$ is recursive, define the recursive array $h_m(n) = \min\{a_m, a_{m+1}, \ldots, a_{m+n}\}$. The other direction is slightly more involved, and we refer the reader to [ZW01].

To simplify notation we give the proof for $\mathbb{Z}$-systems; the multidimensional case is similar.

Let us first show necessity, which is the easier direction. Denote by Σ^* the set of finite words over Σ. Recall that $h \in \mathbb{R}$ is *right-recursively enumerable* if there is a recursive sequence $a_n \in \mathbb{Q}$ with $a_n \searrow h$. Define S_L for infinite sets L in the same way as in section 1.1.

Lemma 7.1. *If $L \subseteq \Sigma^*$ is RE, then $h(S_L)$ is right-recursively enumerable*

Proof. Let $a_1, a_2, \dots$ be a recursive sequence with $L = \{a_n : n \in \mathbb{N}\}$. Let $N_{n,m}$ denote the number of patterns $a \in \Sigma^{2^n}$ which do not contain any of the words $a_1, a_2, \dots, a_m$. Then $h_{n,m} = \frac{1}{2^n} \log N_{n,m}$ is a recursive array, and by enumerating the pairs (m,n) and taking the minimum of $h_{m,n}$ for initial segments of this enumeration, we see that $h = \inf_{m,n} h_{m,n}$ is right recursively enumerable. We claim that $h = h(S_L)$. Clearly $\geq$ holds. On the other hand, $h_{n,k} \to h(S_{\{a_1,\dots,a_n\}})$ as $k \to \infty$ by [Fri97] or [HM07]. Therefore $h \leq \inf_n h(S_{\{a_1,\dots,a_n\}} = h(S_L)$ as desired. □

Returning to the general (non-expansive) case, let X be an EDS. Let $\mathcal{P}_n$ be a recursive sequence of refining partitions separating points in X; such a sequence clearly exists. Let X_n be the factor defined by $\mathcal{P}_n$. By proposition 3.3 and the previous lemma $h(X_n)$ is right recursively enumerable, but an inspection of the proof of the lemma shows that the recursive sequence of numbers descending to $h(X_n)$ can be computed from $\mathcal{P}_n$ and the effective data defining X (see proposition 3.3). In other words, we can compute a recursive array $(m,n) \mapsto h_m(n)$ with $h_m(n) \searrow h(X_m) \nearrow h(X)$. This proves one direction of theorem 1.8.

Before proving the opposite direction, we demonstrate the technique in a simpler case.

Lemma 7.2. *Let h be a right recursively enumerable number; then there exists an ESS X with $h(X) = h$.*

Proof. We may assume that $0 < h < 1$, since we can increase entropy by integer increments by taking products with full shifts.

For a word $a \in \{0,1\}^\ell$, define $N_k(a)$ to be the number of distinct subwords of length k in a. Given a decreasing recursive sequence of numbers $0 < h(n) < 1$ with $h = \lim h(n) > 0$, we define a sequence of numbers $\ell(n)$ recursively by

$$\begin{aligned} \ell(1) &= 2 \\ \ell(n+1) &= \ell(n)^{\ell(n)} \end{aligned}$$

and for $n \geq 2$ define sets $L_n \subseteq \{0,1\}^{\ell(n)}$ by

$$L_n = \{a \in \{0,1\}^{\ell(n)} \,:\, N_{\ell(k)}(a) > \ell(k) \cdot \left\lceil 2^{h(k)\cdot\ell(k)} \right\rceil \text{ for some } k < n\}$$

Clearly if $a \notin L_n$ and $k < n$ then $a|_I \notin L_k$ for any segment I of length $\ell(k)$. It follows that S_{L_n} is a decreasing family of subshifts. Set $L = \cup L_n$; this is clearly an RE set. Let $X = S_L = \cap S_{L_n}$. We claim that $h(X) = \lim h(n)$.

Indeed, the inequality $h(X) \geq h$ follows from the fact that we can construct a subshift $X_0 \subseteq X$ with entropy h. To do this select $\lceil 2^{h(1)\cdot\ell(1)} \rceil$ blocks of length $\ell(1)$. Form all possible concatenations of $\ell(2)/\ell(1)$ of these blocks; this gives a collection of $\lceil 2^{h(1)\cdot\ell(1)} \rceil^{\ell(2)/\ell(1)}$ of blocks of length $\ell(2)$, none belonging to L_2. Choose a subset of size $\lceil 2^{h(2)\cdot\ell(2)} \rceil$ of these blocks – the fact that the $h(i)$ decreases means that there are enough blocks to do this – and again form all concatenations of length $\ell(3)/\ell(2)$ of them, arriving at a collection of $\lceil 2^{h(2)\cdot\ell(2)} \rceil^{\ell(3)/\ell(2)}$ blocks of length $\ell(3)$ none of which belong to L_3; etc. Taking the limit of these collections of blocks gives a subshift $X_0 \subseteq X$ with entropy h.

For the other direction we rely on an empirical version of the Shannon-MacMillan-Breiman, theorem 2.2 of [OW90], which states that if $(\xi_n)_{n=1}^{\infty}$ is a typical sample from a finite valued process with entropy t, then for all $\varepsilon > 0$, all M sufficiently large and all $N \geq M^M$, one cannot cover $1-\varepsilon$ of the word $\xi_1\xi_2 \ldots \xi_N$ with a collection of less than $2^{M(t-\varepsilon)}$ words of length M; choosing a large enough n we can take $M = \ell(n)$ and $N = \ell(n+1)$, and we get that $\xi_1\xi_2 \ldots \xi_N \in L_{n+1}$. Now suppose by way of contradiction that $h(X) > h$. By the variational principle there is an invariant measure μ on X with entropy $> h$ and the support of μ contains points containing subwords belonging to some L_N, and this is impossible, in contradiction to the Ornstein-Weiss result. Hence $h(X) \leq h$. □

We now turn to the proof of sufficiency in theorem 1.8. Suppose that $(n,k) \mapsto h_n(k)$ is recursive, that $h_n(k) \searrow h_n$ and that $h_n \nearrow h$. We may assume without loss of generality that $h_{n+1}(k) \geq h_n(k)$, since we can always replace $h_n(k)$ with $\max_{m<n} h_m(k)$. As before, we may also assume that $0 < h < 1$. We can further assume that $0 < h_n(k) < 1$. We are out to construct an EDS $X = X(h)$ with entropy h.

We describe an effectively closed subset of $(\{0,1\}^{\mathbb{N}})^{\mathbb{Z}}$, which we think of as the space of colorings of $\mathbb{Z} \times \mathbb{N}$, by specifying a sequence of disallowed 2-dimensional patterns of rectangular shape. For a rectangular pattern $a \in \{0,1\}^{[1;m]\times[1;n]}$ we think of it as a word of length m over the alphabet $\{0,1\}^n$ and define $N_k(a)$ as above; i.e. $N_k(a)$ is the number of distinct sub-patterns of a of the form $a|_{[i,i+k]\times[1;n]}$ (for the purpose of counting we identify patterns which differ only up to a translation).

Define $\ell(m)$ as in the proof of lemma 7.2, i.e. $\ell(1) = 2$ and $\ell(m+1) = \ell(m)^{\ell(m)}$. Define languages $L_{n,m} \subseteq (\{0,1\}^n)^{\ell(m)}$ by

$$L_{n,m} = \{a \in (\{0,1\}^n)^{\ell(m)} \ : \ N_{\ell(k)}(a|_{[1;\ell(m)]\times[1;i]}) > \ell(k) \cdot \left\lceil 2^{h_i(k)\cdot \ell(k)} \right\rceil \text{ for some } k < m \text{ and } i \leq n\}$$

Let L be the union of the $L_{n,m}$ together with all translates of patterns from this union in the $\mathbb{Z}$-direction (that is if a pattern $a \in \{0,1\}^{[1,m]\times[1,n]}$ is in the union then so are all translates of it on rectangles $[i, i+m] \times [1;n]$). This is clearly an RE set. Let X be the complement of the corresponding cylinder sets in $(\{0,1\}^{\mathbb{N}})^{\mathbb{Z}}$; this is an EDS.

If we consider the partition $\mathcal{P}_n$ of X according to the pattern induced on $\{0\} \times [1;n]$ then the resulting symbolic system X_n has entropy h_n, as can be seen from the lemma above. Since the sequence $\mathcal{P}_n$ is recursive, and together with all translates generates the topology of X, we see that $h(X) = \lim h(X_n) = h$, as desired.

8. Factoring SFTs onto Full Shifts

We next prove proposition 1.10, which answers question 2.10 of [BS07]. We refer to [BS07] for background. Our construction is based on the following lemma:

Lemma 8.1. *There is a uniquely ergodic ESS $X \subseteq \{0,1\}^{\mathbb{Z}}$ such that the frequency of 1's in each point $x \in S_L$ is $1/2$, and such that for every integer $r > 0$ there is an n and a block a of length n appearing in X, such that*

$$\#\{i : a(i) = 1\} < \frac{(n-2r)^3}{2n^3}$$

Proof. We briefly sketch the construction, which is by block concatenation. We define pairs of words $a_r, b_r \in \{0,1\}^{k(r)}$ recursively, satisfying the following properties:

(1) a_r, b_r are concatenations of a_{r-1}, b_{r-1}, and contains both as subwords.
(2) The frequency of 1's in a_r is $< (k(r) - 2r)^3/2k(r)^3$ and in b_r is $> (k(r) + 2r)^3/2k(r)^3$, and the sum of these frequencies is $1/2$.
(3) for $r \geq 3$, for every word w of length $< k(r-2)$, and any two subwords c', c'' of a_r, b_r of length $k(r-1)$, the frequencies of the occurrence of w in c' and in c'' differ by at most $1/r$.

Here w^i means the concatenation of w with itself i times. We begin with $k(1) = 3$ and $a_1 = 000$, $b_1 = 111$, and it is easy to check that given a_{r-1}, b_{r-1}, the choice $a_r = (a_{r-1}^m b_{r-1}^m)a_{r-1}^n$ and $b_r = (b_{r-1}a_{r-1})^m b_{r-1}^n$ will satisfy these requirements for suitably chosen large integers $m > n$ (which may depend on r).

A standard argument now shows that one-sided infinite sequences a, b which are the limits of a_r and of b_r, respectively, have the same forward orbit closure, which is uniquely ergodic with 1's appearing with frequency $1/2$. Now take the natural extension. The fact that this is an ESS is clear, and given r, the desired block is a_r (with $n = k(r)$). □

Let $Y \subseteq \{0,1\}^{\mathbb{Z}^3}$ be the subshift obtained by extending each point on X in directions e_2 and e_3 (that is: $y \in Y$ if and only if for some $x \in X$ we have $y(i,j,k) = x(i)$ for all (i,j,k). Thus $(Y, T_1) \cong (X, T)$. By theorem 1.4 and its proof, Y is a sofic shift and there is an SFT Z and $\mathbb{Z}^3$-factor map $\pi : Z \to Y$ so that the extension $Z \to Y$ is ATIE with respect to T_1. In particular, $h(Z) = 0$ with respect to the $\mathbb{Z}^3$-action.

We may assume that π is a 1-block code, so each symbol in Z contains a component from Y's alphabet and π simply forgets all other information. Starting with Z, superimpose 4 symbols over the occurrence of 1's in Z (these are the 1's coming from Y), with no restrictions on configurations, and call the resulting system W. Since the density of 1's in Y is $1/2$, we have $h(\widehat{Y}) = \log 2$. Write $\pi : W \to Y$ as well. As in [BS07], we observe that there is a unique invariant measure μ on W with entropy $\log 2$, and if $\mu = \int \mu_y d\nu(y)$ is the disintegration of μ over Y then for ν-a.e. $y \in Y$, the measure μ_y is obtained by uniformly and independently choosing the symbol over each 1 in y.

We shall show that no factor map exists from W to the full shift. Indeed, suppose $f : W \to \{0,1\}^{\mathbb{Z}^3}$ were a factor map given by a sliding block code with radius r. As in [BS07], for a typical $y \in Y$ the measure μ_y must map under f to the uniform Bernoulli measure on $\{0,1\}^{\mathbb{Z}^3}$. Now since y is typical we can, using the properties of X, choose a large cube $Q \subseteq \mathbb{Z}^3$of dimensions $n \times n \times n$, so that the density of 1's in $y|_Q$ is $< (n-2r)^3/2n^3$. Thus the entropy of the measure

μ_y with respect to the partition according to symbols in Q is $< \log 4 \cdot (n-2r)^3/2$. But because the window width of f is r, for any $w \in W$ the values of $w|_Q$ completely determine the pattern $f(w)|_{\widehat{Q}}$, where $\widehat{Q}$ is the cube obtained by deleting from Q every point within r of the complement of Q, in the ℓ_∞ metric. But the uniform measure on $\{0,1\}^{\mathbb{Z}^3}$, when restricted to the coordinates $\widehat{Q}$, has entropy $|\widehat{Q}| \cdot \log 2 \geq \log 4 \cdot (n-2r)^3/2$; a contradiction. This completes the proof of the first part of the proposition.

On the other hand, it is not hard to explicitly construct a non-trivial, uniquely ergodic subshift $X \subseteq \{0,1\}$ with 1's appearing with density $\frac{1}{2}$ and such that if we extend the one-dimensional subshift by adding one of four random colors over each 1, the resulting system factors onto the full shift $\{0,1\}^{\mathbb{Z}}$. Starting now from a uniquely ergodic zero entropy 3-dimensional SFT whose subaction factors onto X, the same process gives a subshift that factors onto $\{0,1\}^{\mathbb{Z}^3}$.

To construct such a system X, pick any uniquely ergodic ESS whose points are concatenations of the words $a = 111000$ and $b = 110100$ with equal frequencies (for example, repeat the construction of lemma 8.1, starting with $a_1 = a$ and $b_1 = b$). Let $A \subseteq \{a\} \times \{0,1,2,3\}^6$ be those pairs where $(a,a') \in A$ if and only if $(a(i)=0) \Rightarrow (a'(i)=0)$, and define $B \subseteq \{b\} \times \{0,1,2,3\}^6$ similarly. Note that $|A| = |B| = 2^6$. If $Y \subseteq X \times \{0,1,2,3\}^{\mathbb{Z}}$ is the subshift obtained by coloring the 1's in X arbitrarily with $0,1,2,3$ and coloring the 0's with 0, then each point in Y is the concatenation of words from A and B. We can define a factor map $\pi : Y \to \{0,1\}^{\mathbb{Z}}$ which, for $(x,x') \in X \times \{0,1,2,3\}^{\mathbb{Z}}$, first identifies the intervals I so the $a|_I = a$ or b (these are determined uniquely and locally), and then act so that $\pi(x.x')|_I = \pi_0(y|_I)$, where π_0 is an arbitrary function so that $\pi_0|_A$ and $\pi_0|_B$ are bijections to $\{0,1\}^6$.

9. Discussion and Problems

In this section we collect some comments and questions regarding this work. We have seen that the class of subactions of SFTs are very rich; almost as rich as the category of general effective dynamics. This reflects the richness of the full dynamics. Another indication of this richness is Simpson's work [Sim07], where the complications are of a more recursive-theoretic nature.

A major challenge is to understand the full dynamics. One approach is to try to control the full dynamics via subactions; this is the approach taken in [HM07, BS07]. Some information can be obtained directly from the fact that the system is effective; theorem 4.1 of [HM07] is a step in this direction, but this gives rather poor information.

A reasonable intermediate step towards the full dynamics might be to complete the picture of the $\mathbb{Z}^{d-1}$-subactions $\mathbb{Z}^d$-SFTs; these are effective but our constructions do not work for them.

Problem 9.1. Characterize the $\mathbb{Z}^{d-1}$ subactions of $\mathbb{Z}^d$-SFTs and sofic shifts.

Even with regard to the $\mathbb{Z}^{d-2}$-subactions there are some interesting questions of a topological nature. It would be desirable, for example, to get extensions which are smaller then ATIEs.

Problem 9.2. Can every EDS be realized as a finite-to-1 factor of the subaction of some SFT?

Another interesting question is the following:

Problem 9.3. What are the expansive subactions which can occur for SFTs, particularly in dimension 2?

Such systems are closely related (though more complicated than) expansive cellular automata, on which some progress has been made in the one-dimensional case [BM97, BM00, Nas08].

One can also use the recursive-theoretic approach to differentiate between potentially tractable systems and intractable ones (e.g. systems with nontrivial Medvedev degree). Two important classes of systems at opposite ends of the dynamical spectrum are the strongly irreducible SFTs and the minimal SFTs (minimal means every orbit is dense). For X in these classes the globally admissible patterns can be decided, i.e. the extension problem can be solved for them. For strongly irreducible systems this was demonstrated in [HM07]; since the proof for minimal SFTs is short we include one here (We note that related results have been proved independently in [DKB06]):

Proposition 9.4. *There is an algorithm which, given a finite set L of patterns defining a non-empty minimal SFT and a pattern $b \in \Sigma^F$, decides whether b appears in S_L or not.*

Proof. The algorithm is as follows: For each n enumerate all $[-n;n]^d$-patterns $a_1, \ldots a_{k(n)}$ which do not contain patterns from L. If $a_i|_F \neq b$ for all $i = 1, \ldots k(i)$, output that b does not appear in X, and halt. If, on the other hand, b occurs in all a_i's, output that b appears in S_L, and halt.

To see that this algorithm halts, note that if b does not appear in S_L then by compactness the first alternative will eventually hold; otherwise b appears in S_L, so the second alternative will eventually occur, since if it does not then again by compactness there is a point in x not containing b, contradicting minimality of X. □

It is important to note that we assume S_L is nonempty; in general, it cannot be decided whether S_L is empty or not.

This recursive-theoretic property of minimal and strongly irreducible SFTs severely limits the applicability to them of the methods presented here. In fact, it seems that any scheme which tries to use sufficiently strong computation (e.g. Turing machines) to introduce dynamical features into SFTs must fail to produce SFTs in these classes, since such a scheme would probably allow us to leave the recursive universe.

The following basic problem underscores the contrast between what we can construct in general and in the minimal case. Recall that the universal $\mathbb{Z}^d$-odometer is a minimal $\mathbb{Z}^d$-action on the Cantor set which factors onto every $\mathbb{Z}^d$-action on a finite abelian group, and such factors separate points. This system is unique up to isomorphism. Using our results one can construct an SFT that factors onto the universal odometer but this SFT will be far from minimal; it will not even be transitive.

Problem 9.5. Is there a minimal SFT extending the universal odometer?

For cellular automata, we have characterized entropy and obtained a fairly good understanding of the possible dynamics on the limit set in dimension ≥ 3, a least with regard the invariant measures. Once again, topologically we are far from a good understanding.

Problem 9.6. Can one describe the dynamics and entropies of surjective or injective CA?

Problem 9.7. Can anything be said about the dynamics of 1- and 2-dimensional CA on their limit sets, analogous to theorem 1.7?

References

[Ber66] Robert Berger. The undecidability of the domino problem. *Mem. Amer. Math. Soc. No.*, 66:72, 1966.

[BM97] F. Blanchard and A. Maass. Dynamical properties of expansive one-sided cellular automata. *Israel J. Math.*, 99:149–174, 1997.

[BM00] Mike Boyle and Alejandro Maass. Expansive invertible onesided cellular automata. *J. Math. Soc. Japan*, 52(4):725–740, 2000.

[BS94] Robert Burton and Jeffrey E. Steif. Non-uniqueness of measures of maximal entropy for subshifts of finite type. *Ergodic Theory Dynam. Systems*, 14(2):213–235, 1994.

[BS07] Mike Boyle and Michael Schraudner. Z^d shifts of finite type without equal entropy full shift factors. *preprint*, 2007.

[CHY90] K. Culik, II, L. P. Hurd, and S. Yu. Computation theoretic aspects of cellular automata. *Phys. D*, 45(1-3):357–378, 1990. Cellular automata: theory and experiment (Los Alamos, NM, 1989).

[dJ78] Andrés del Junco. A simple measure-preserving transformation with trivial centralizer. *Pacific J. Math.*, 79(2):357–362, 1978.

[DKB06] Jean-Charles Delvenne, Petr Kurka, and Vincent Blondel. Decidability and universality in symbolic dynamical systems. *Fund. Inform.*, 74(4):463–490, 2006.

[Fri97] Shmuel Friedland. On the entropy of $\mathbf{Z}^d$ subshifts of finite type. *Linear Algebra Appl.*, 252:199–220, 1997.

[Hed69] G. A. Hedlund. Endormorphisms and automorphisms of the shift dynamical system. *Math. Systems Theory*, 3:320–375, 1969.

[HKC92] Lyman P. Hurd, Jarkko Kari, and Karel Culik. The topological entropy of cellular automata is uncomputable. *Ergodic Theory Dynam. Systems*, 12(2):255–265, 1992.

[HM07] Michael Hochman and Tom Meyerovitch. A characterization of the entropies of multidimensional shifts of finite type. *preprint, www.arxiv.org/math.DS/0207206*, 2007.

[HU79] John E. Hopcroft and Jeffrey D. Ullman. *Introduction to automata theory, languages, and computation.* Addison-Wesley Publishing Co., Reading, Mass., 1979. Addison-Wesley Series in Computer Science.

[Hur87] Lyman P. Hurd. Formal language characterizations of cellular automaton limit sets. *Complex Systems*, 1(1):69–80, 1987.

[Hur90] Lyman P. Hurd. Nonrecursive cellular automata invariant sets. *Complex Systems*, 4(2):131–138, 1990.

[JM05] Aimee Johnson and Kathleen Madden. Factoring higher-dimensional shifts of finite type onto the full shift. *Ergodic Theory Dynam. Systems*, 25(3):811–822, 2005.

[Kar94] Jarkko Kari. Rice's theorem for the limit sets of cellular automata. *Theoret. Comput. Sci.*, 127(2):229–254, 1994.

[Kar05] Jarkko Kari. Theory of cellular automata: a survey. *Theoret. Comput. Sci.*, 334(1-3):3–33, 2005.

[Lin84] D. A. Lind. The entropies of topological Markov shifts and a related class of algebraic integers. *Ergodic Theory Dynam. Systems*, 4(2):283–300, 1984.

[Moz89] Shahar Mozes. Tilings, substitution systems and dynamical systems generated by them. *J. Analyse Math.*, 53:139–186, 1989.

[Mye74] Dale Myers. Nonrecursive tilings of the plane. II. *J. Symbolic Logic*, 39:286–294, 1974.

[Nas08] M. Nasu. Textile systems and onesided resolving automorphisms and endomorphisms of the shift. *Ergodic Theory Dynam. Systems*, 28(1):167–209, 2008.

[OW90] Donald S. Ornstein and Benjamin Weiss. How sampling reveals a process. *Ann. Probab.*, 18(3):905–930, 1990.

[QŞ03] Anthony Quas and Ayşe A. Şahin. Entropy gaps and locally maximal entropy in $\mathbb{Z}^d$ subshifts. *Ergodic Theory Dynam. Systems*, 23(4):1227–1245, 2003.

[Rob71] Raphael M. Robinson. Undecidability and nonperiodicity for tilings of the plane. *Invent. Math.*, 12:177–209, 1971.

[Sim07] Stephen G. Simpson. Medvedev degrees of 2-dimensional subshifts of finite type. *preprint, www.math.psu.edu/simpson/papers/2dim.pdf*, 2007.

[Sut05] K. Sutner. Universality and cellular automata. In *Machines, computations, and universality*, volume 3354 of *Lecture Notes in Comput. Sci.*, pages 50–59. Springer, Berlin, 2005.

[vN66] Jon von Neumann. Theory of self-reproducing automata. 1966.

[Wal82] Peter Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1982.

[Wol84] Stephen Wolfram. Universality and complexity in cellular automata. *Phys. D*, 10(1-2):1–35, 1984. Cellular automata (Los Alamos, N.M., 1983).

[ZW01] Xizhong Zheng and Klaus Weihrauch. The arithmetical hierarchy of real numbers. *MLQ Math. Log. Q.*, 47(1):51–65, 2001.

Current address: Fine Hall, Washington Road, Princeton University, Princeton, NJ 08544

E-mail address: hochman@math.princeton.edu