Faithful zero-dimensional principal extensions

Tomasz Downarowicz* Dawid Huczek*

Abstract

We prove that every topological dynamical system (X,T) has a faithful zero-dimensional principal extension, i.e. a zero-dimensional extension (Y,S) such that for every S-invariant measure ν on Y the conditional entropy $h(\nu|X)$ is zero, and, in addition, every invariant measure on X has exactly one preimage on Y. This is a strengthening of the result in [D-H] (in which the extension was principal, but not necessarily faithful).

1 Introduction

All the terms used in the introduction are standard, nonetheless they are explained in the Preliminaries. Given a topological dynamical system (X,T) in the form of the action of the iterates of a continuous transformation T on a compact metric space X, it is of interest to replace it by another, more familiar system, easier to describe and handle. One of the possibilities is to lift (i.e. extend) the system to some (Y,S), whose phase space Y is zero-dimensional. Any zero-dimensional system admits a pleasant symbolic array form in which every point is represented as an infinite array filled with discrete symbols. This representation allows one to apply many methods of symbolic dynamics, for instance to calculate the topological entropy or the entropies of the invariant measures, and, in fact, many other invariants. In case we find such a system (Y,S), we say, that we have found a *zero-dimensional extension*. It is an elementary exercise to show that every system has a zero-dimensional extension. On the other hand, we are naturally interested in minimizing the "distance" between the original system (X,T) and its extension (Y,S). There are at least two ways in which this minimization can be performed:

*Research of both authors is supported from resources for science in years 2009-2012 as research project (grant MENII N N201 394537, Poland).

1. We may be interested in minimizing, for each invariant measure μ on X, the increase of entropy as we lift that measure to an invariant measure ν on Y.
2. We may want that for each μ on X there exist a unique lift ν on Y.

The increase of entropy specified in 1. is measured by the conditional dynamical entropy $h(\nu|X)$. In case μ has finite dynamical entropy, this conditional entropy simply equals the difference $h(\nu)-h(\mu)$, but the conditional notation is universal. The best one can get in the category 1. is a *principal extension*, i.e., such that $h(\nu|X) = 0$ for every invariant measure ν on Y. An extension which satisfies the postulate 2. will be called here a *faithful extension*.

One may also ask whether there exists an extension that satisfies a stronger version of postulate 2., namely that for each μ on X there exist a unique lift ν on Y and that (Y, ν, S) is measure-theoretically isomorphic to (X, μ, T). Such an extension is called *isomorphic*. Clearly, an isomorphic extension is automatically principal and faithful.

For invertible systems (i.e., such that T is a homeomorphism) with finite topological entropy and satisfying an additional assumption that it has a nonperiodic minimal factor, the existence of zero-dimensional isomorphic extensions has already been established as a consequence of the deep results in mean dimension theory developed by E. Lindenstrauss and B. Weiss ([L-W] and [Li]). Every such system satisfies the so-called "small boundary property", which allows to rather easily construct its zero-dimensional isomorphic extension. If the assumption about the existence of a minimal factor is dropped, the above theory still allows to easily build a zero-dimensional principal (but not necessarily faithful) extension: It is elementary to see that the direct product with any system of zero topological entropy is a principal extension and that the composition of two principal extensions is a principal extension. Thus, an arbitrary (invertible) system of finite topological entropy is first extended to its direct product with some infinite minimal system of zero topological entropy (for example an irrational rotation of the circle, or an odometer) and since such a product already has a minimal nonperiodic factor, it can be isomorphically extended to a zero-dimensional system.

Lindenstrauss provided examples showing that both assumptions (finite entropy and minimal factor) are essential for the small boundary property ([Li]). Thus, for systems with infinite entropy, even those which admit a minimal nonperiodic factor, we cannot hope to prove the existence of an isomorphic zero-dimensional extension.

In [D-H] we have shown that a principal zero-dimensional extension can be constructed for any topological dynamical system. However, this extension was almost certainly not faithful, let alone isomorphic. In the present paper we strengthen this result, showing that any topological dynamical system has a faithful, principal zero-dimensional extension. The construction is in fact very similar to the one in [D-H] and the constructed extension is essentially the limit of a sequence of block codes.

Historically, the term *principal* was probably first used by F. Ledrappier in [Le]. The construction of a principal extension for systems of finite topological entropy via the mean dimension theory was heavily exploited in [B-D] in the theory of symbolic extensions. The detailed description of the passage from the small boundary property to the principal extension can be found in [D1] (in earlier papers it is considered more or less obvious and left to the reader).

2 Preliminaries

2.1 Basic notions

Dynamical systems. Throughout this work a *dynamical system* will be a triple $(X, T, \mathbb{S})$, where X is a compact metric space with metric d, T is a continuous map on X (invertible or not) and $\mathbb{S} \in \{\mathbb{Z}, \mathbb{Z}_+ \cup \{0\}\}$ is the index set, depending on whether we consider both the negative and positive iterates of T (i.e. the action on X of $\mathbb{Z}$) or only positive ones (i.e. the action on X of $\mathbb{Z}_+ \cup \{0\}$) — our final result applies in both cases, but a few details of the proof differ, so we need to be able to make the distinction. For brevity, where the index set or transformation are obvious or irrelevant, we will omit them.

Factors, extensions and conjugacies. A dynamical system $(X, T, \mathbb{S})$ is a *factor* of the dynamical system $(Y, S, \mathbb{S})$ if there exists a continuous map π from Y onto X such that $T \circ \pi = \pi \circ S$. In this situation we also say that Y is an *extension* of X. If the map π is a homeomorphism, we say that X and Y are *conjugate*.

2.2 Zero-dimensional dynamical systems

A dynamical system $(Y, S, \mathbb{S})$ is called *zero-dimensional* if Y is a zero-dimensional space, i.e. if it has a base consisting of sets which are both closed and open. A particularly important class of such systems are symbolic systems over an uncountable alphabet which we will call *array systems* and which are constructed as follows: Let Λ_k be a finite set with discrete topology and let $Y_c = \prod_{k=1}^{\infty} \Lambda_k^{\mathbb{S}}$. The points of Y_c can be thought of as arrays $\{y_{k,n}\}_{k \geq 1, n \in \mathbb{S}}$ where $y_{k,n} \in \Lambda_k$. With the action of the horizontal shift S (i.e. $(Sy)_{k,n} = y_{k,n+1}$) Y_c becomes a zero-dimensional dynamical system. An array system is any closed, shift-invariant subset Y of such a Y_c.

A *j by m rectangle* will mean a rectangle of height j and width m, that is a finite matrix $C = (C_{k,n})$, $k = 1, \ldots, j$; $n = 0, \ldots, m-1$, with $C_{k,n} \in \Lambda_k$. Any such rectangle C can be identified with a cylinder in Y in the standard way: $y \in C$ iff $y_{k,n} = C_{k,n}$, $k = 1, \ldots, j$; $n = 0, \ldots, m-1$.

2.3 The marker lemma

The following lemma is a standard tool in zero-dimensional dynamics, the proof and several generalizations can be found e.g. in [D2].

Lemma 2.1. *Let $(X, T, \mathbb{Z})$ be zero-dimensional topological dynamical system without any periodic points. For every $n \geq 1$ there exists a clopen set F such that:*

1. *$T^{-i}(F)$ are pairwise disjoint for $i = 0, 1, \ldots, n-1$.*
2. *$X = \bigcup_{i=-n+1}^{n-1} T^{-i}(F)$.*

2.4 Invariant measures

Denote the set of all probability measures on the sigma-algebra of Borel subsets of X by $\mathcal{M}(X)$. It is a compact, metrizable, convex set. The map T induces on $\mathcal{M}(X)$ a continuous map (which we also denote by T) by the formula $(T\mu)(B) = \mu(T^{-1}(B))$ (for Borel sets B). Let $\mathcal{M}_T(X)$ be the set of all T-invariant probability measures on X, i.e. such that $T\mu = \mu$. This is a nonempty, closed and convex subset of $\mathcal{M}(X)$. The extreme points of $\mathcal{M}_T(X)$ are ergodic measures, i.e. the ones for which any invariant set has measure either 0 or 1. We will denote the set of ergodic measures by $\mathcal{M}_T^e(X)$.

For a measure $\mu \in \mathcal{M}(XC)$ we define

$$\mathbf{A}_n^T(\mu) = \frac{1}{n}\left(\mu + T\mu + \ldots + T^{n-1}\mu\right).$$

For $x \in X$ the symbol δ_x denotes the point mass at x. We will later need the following two facts, both of which are fairly obvious:

Fact 2.2. *For any n the set $\mathcal{M}_T(X)$ is within the closure of the convex hull of the set $\left\{\mathbf{A}_n^T(\delta_x);\, x \in X\right\}$.*

Fact 2.3. *Let U be an open subset of $\mathcal{M}(X)$ containing $\mathcal{M}_T(X)$. There exists an N such that for any $n > N$ and any $x \in X$ the measure $\mathbf{A}_n^T(\delta_x)$ is in U.*

Let (Y, S) be an array system and let d be a metric consistent with the weak-star topology on $\mathcal{M}_S(Y)$. We will make use of the following standard fact:

Fact 2.4. *Let $\mu, \nu \in \mathcal{M}_S(Y)$. For any $\varepsilon > 0$ there exist $\delta > 0$ and $j, k > 0$ such that if*

$$|\mu(C) - \nu(C)| < \delta$$

for all j by k-rectangles C, then $d(\mu, \nu) < \varepsilon$.

2.5 Entropy

We recall the basic definitions and facts of the entropy theory of dynamical systems. Let (X, T) be a dynamical system and let $\mu \in \mathcal{M}_T(X)$. For any finite partition $\mathcal{A}$ of X into measurable sets we define the entropy of a partition as

$$H(\mu, \mathcal{A}) = -\sum_{A \in \mathcal{A}} \mu(A) \ln \mu(A).$$

$$H_n(\mu, \mathcal{A}) = \frac{1}{n} H(\mu, \mathcal{A}^n),$$

where $\mathcal{A}^n = \bigvee_{j=0}^{n-1} T^{-j}(\mathcal{A})$. The sequence H_n is known to converge to its infimum, which allows one to define

$$h(\mu, \mathcal{A}) = \lim H_n(\mu, \mathcal{A}).$$

Finally the *entropy of a measure* is given as

$$h(\mu) = \sup_{\mathcal{A}} h(\mu, \mathcal{A}),$$

where the supremum is taken over all finite partitions of X.

If $\mathcal{A}$ and $\mathcal{B}$ are two finite partitions of X, then we can define the conditional entropy of $\mathcal{A}$ with respect to $\mathcal{B}$ as

$$H(\mu, \mathcal{A}|\mathcal{B}) = \sum_{B \in \mathcal{B}} \mu(B) H(\mu_B, \mathcal{A}),$$

where $\mu_B(A) \in \mathcal{M}(X)$ is defined by $\mu_B(A) = \mu(A|B)$ for all Borel sets. Then we can proceed to define

$$H_n(\mu, \mathcal{A}|\mathcal{B}) = \frac{1}{n} H(\mu, \mathcal{A}^n|\mathcal{B}^n).$$

Again, the sequence H_n is known to converge to its infimum, which allows one to define

$$h(\mu, \mathcal{A}|\mathcal{B}) = \lim H_n(\mu, \mathcal{A}|\mathcal{B}).$$

Suppose now that we have a system (Y, S) that is an extension of (X, T) by a map π. Let ν be an invariant measure on Y. Any partition $\mathcal{B}$ of X can be lifted to a partition $\pi^{-1}\mathcal{B}$ of Y. Define

$$h(\nu, \mathcal{A}|X) = \inf_{\mathcal{B}} h(\nu, \mathcal{A}|\pi^{-1}\mathcal{B}),$$

where the infimum is taken over all finite partitions of X. Finally, define

$$h(\nu|X) = \sup_{\mathcal{A}} h(\nu, \mathcal{A}|X),$$

where once again the supremum is taken over all finite partitions of Y. If the image $\pi\nu$ of ν by the factor map π has finite entropy, then it is not difficult to see that $h(\nu|X) = h(\nu) - h(\pi\nu)$.

We will make use of the following fact.

Fact 2.5. *Let Y be an array system and let X be a factor of Y. Let $\mathcal{R}_k$ be the partition defined by cylinders of height k and length 1. Then for any measure ν on Y we have $h(\nu|X) = \lim_k h(\nu, \mathcal{R}_k|X)$.*

To see that it is so, it suffices to observe two facts. Firstly, that the family $\{\mathcal{R}_k\}$ together with its images under iterates of S generates the Borel σ-algebra on Y. Secondly, if $j < k$ then $\mathcal{R}_j \prec \mathcal{R}_k$, and therefore $h(\nu, \mathcal{R}_j|X) < h(\nu, \mathcal{R}_k|X)$.

We recall the key definition from the Introduction:

Definition 2.6. *Suppose a dynamical system (Y, S) is an extension of the system (X, T) via the map π. (Y, S) is a* principal extension *if $h(\nu|X) = 0$ for every $\nu \in \mathcal{M}_S(Y)$.*

If (X, T) has finite topological entropy, then by the variational principle $\pi\nu$ has finite entropy for each ν in $\mathcal{M}_S(Y)$ so the extension is principal if and only if $h(\nu) = h(\pi\nu)$ for each ν. In particular (Y, S) has the same topological entropy as (X, T) (this holds also in case of infinite entropy).

2.6 Continuity of the entropy functions

In the main proof we will consider entropy as a function of the measure, and we will need several basic facts about the continuity of this function which we state without proof. First of all:

Fact 2.7. *The function $\mu \mapsto \mu(A)$ on $\mathcal{M}(X)$ is upper semicontinuous if A is closed and lower semicontinuous if A is open.*

Since $\mu(\mathrm{Int}(A)) \leq \mu(A) \leq \mu(\overline{A})$ and the three are equal if the boundary of A has measure 0, we have the following:

Fact 2.8. *The function $\mu \mapsto \mu(A)$ on $\mathcal{M}(X)$ is continuous at every μ such that $\mu(\partial A) = 0$.*

Using the fact that the limit defining $h(\mu, \mathcal{A}|\mathcal{B})$ is also the infimum, we easily arrive at the following:

Fact 2.9. *For any finite partitions $\mathcal{A}, \mathcal{B}$ of X the function $\mu \mapsto h(\mu, \mathcal{A}|\mathcal{B})$ on $\mathcal{M}_T(X)$ is upper semicontinuous at every μ such that $\mu(\partial A) = 0$ for every $A \in \mathcal{A}$ and $\mu(\partial B) = 0$ for every $B \in \mathcal{B}$.*

Finally:

Fact 2.10. *If (Y, S) is an extension of (X, T) and $\mathcal{A}$ is a finite partition of Y, then the function $\mu \mapsto h(\mu, \mathcal{A}|X)$ on $\mathcal{M}_S(Y)$ is upper semicontinuous at every μ such that $\mu(\partial A) = 0$ for every $A \in \mathcal{A}$.*

To observe that, note that $h(\mu, \mathcal{A}|X)$ is the infimum of any sequence $h(\mu, \mathcal{A}|\mathcal{B}_n)$, provided that the diameter of the largest set in $\mathcal{B}_n$ tends to 0. Since we can construct partitions into sets of arbitrarily small diameter that all have boundaries whose measure μ is 0, Fact 2.10 now follows.

3 The main result

Theorem 3.1. *Any topological dynamical system has a faithful zero-dimensional principal extension.*

Remark 3.2. *Moreover, the extension we construct has no periodic points.*

The proof of theorem 3.1 will occupy the remainder of this section.

Let $(X, T, \mathbb{S})$ denote the dynamical system for which we will be constructing the desired extension. Without loss of generality we can assume that T is invertible. Indeed, if T is surjective, we can simply replace the system by its natural extension (which is principal and faithful). If T is not surjective, we can replace X by the set

$$X' = (X \times \mathbb{Z}_+) \cup \{\infty\}$$

and define a metric d' on it as follows (with the original metric on X denoted by d):

$$\begin{aligned} d'((x_1, n_1), (x_2, n_2)) &= \left| \frac{1}{n_1} - \frac{1}{n_2} \right|, \text{ if } n_1 \neq n_2, \\ d'((x_1, n), (x_2, n)) &= \frac{1}{n} d(x_1, x_2), \\ d'((x, n), \infty) &= \frac{1}{n}. \end{aligned}$$

In other words, X' can be seen as infinitely many copies of X arranged in a sequence and shrinking to a single point. Now, we define the action T' on X' as follows:

$$\begin{aligned} T'(\infty) &= \infty, \\ T'(x,n) &= (x,n-1), \text{ if } n>1, \\ T'(x,1) &= (Tx,1). \end{aligned}$$

Now T' is surjective on X', and, with the exception of the measure concentrated on the fixed point ∞, all T'-invariant measures are supported by the set $X\times\{1\}$, so they are the same as the original measures on X. The system $(X',T',\mathbb{S})$ has a natural extension. If we now construct a faithful zero-dimensional principal extension of $(X',T',\mathbb{S})$ (via the natural extension), then the system on the preimage of $X\times\{1\}$ will be a faithful zero-dimensional principal extension of $(X,T,\mathbb{S})$. Therefore, from now on we will assume T to be invertible.

Let I denote the one-dimensional torus, i.e. the interval $[0,1]$ with the endpoints identified, and let λ be the Lebesgue measure on I. Any function $f : X \to [0,1]$ induces a partition $\mathcal{A}_f$ of $X\times I$ into two sets: $\{(x,t) : 0\le t< f(x)\}$ and $\{(x,t) : f(x)\le t<1\}$ (i.e. the sets of points below and above the graph of f). For a family $\mathcal{F}$ of functions we denote by $\mathcal{A}_{\mathcal{F}}$ the partition $\bigvee_{f\in\mathcal{F}}\mathcal{A}_f$. Two useful observations are that $\mathcal{A}_{\mathcal{F}\cup\mathcal{G}} = \mathcal{A}_{\mathcal{F}}\vee\mathcal{A}_{\mathcal{G}}$ and that $\mathcal{F}\subset\mathcal{G}$ implies $\mathcal{A}_{\mathcal{F}}\prec\mathcal{A}_{\mathcal{G}}$.

Let R_0 be some irrational rotation of I, chosen completely arbitrarily but fixed throughout this paper. Let $\mathcal{A}_j$ be a sequence of partitions of $X\times I$, each of which is induced by a finite family of continuous functions $\mathcal{F}_j$ such that $\mathcal{F}_j\subset\mathcal{F}_{j+1}$. Let η_j be the diameter of the largest set in $\mathcal{A}_j$ (in the product metric on $X\times I$). We will require that $2\eta_{j+1}<\eta_j$ (which obviously implies that the η_j tend to 0). We will also request that $(T\times R_0)^{-1}(\mathcal{A}_j)\vee\mathcal{A}_j\vee(T\times R_0)\mathcal{A}_j\prec\mathcal{A}_{j+1}$.

Let $\pi^{(1)}$ denote the projection of $X\times I$ onto X. Consider the space of all formal arrays $y = y_{j,n}$ ($j\ge 1, n\in\mathbb{S}$), such that $y_{j,n}\in\mathcal{A}_j$ (we treat each finite partition $\mathcal{A}_j$ as the alphabet in row j). For an array y define the sets $K_{j,n}(y) = \left\{x\in X : d(x,\pi^{(1)}(y_{j,n}))\le\eta_j\right\}$. ($K_{j,n}(y)$ is the η_j-neighborhood of the projection onto X of the cell of $\mathcal{A}_{\mathcal{F}_j}$ appearing as a symbol in y at the position (j,n)). An array y will be said to satisfy the *column condition* if for each n the sequence $K_{j,n}(y)$ is descending (as j increases). Since the diameter of $K_{j,n}$ tends to 0 with j, the column condition implies that the intersection $\bigcap_{j=0}^{\infty}K_{j,n}(y)$ is a single point in X which we will denote by $x_n(y)$. Note that $x_n(y)$ is within η_j of each set $\pi^{(1)}(y_{j,n})$ – a fact that will be useful later.

Now, let Y_C be the space of all arrays y satisfying the column condition with the additional requirement that $x_{n+1}(y) = T(x_n(y))$. It is easy to see that with the action of the horizontal shift S, Y_C forms a continuous extension of (X, T), where the factor map is $\pi_X(y) = x_0(y)$ (by which we mean $x_n(y)$ for $n = 0$). We will construct the desired faithful zero-dimensional principal extension, Y, as a subsystem of Y_C which will in some sense be the limit of an auxiliary sequence of (disjoint) mutually conjugate subsystems $Y_k \subset Y_C$. We will define Y_k inductively, by constructing the maps $\Phi_k : Y_{k-1} \to Y_k$ (these maps will in fact be block codes defined on rectangles of some order). The main goal will be to ensure that for any k and all $k' > k$ the set $\mathcal{M}_S(Y_{k'})$ is contained within an open set $\mathcal{U}_k \subset \mathcal{M}_S(Y_C)$, where the sequence $\{\mathcal{U}_k\}$ (which we will also define inductively) satisfies the following properties:

- U1. $\mathcal{U}_{k+1} \subset \mathcal{U}_k$.
- U2. For any $k > 0$ and any measure $\nu \in \mathcal{U}_k$ we have $h(\nu, \mathcal{R}_k | X) < \varepsilon_k$, (where, recall, $\mathcal{R}_k$ is the partition defined by cylinders of height k and length 1).
- U3. For any $k > 0$ and any two measures $\nu_1, \nu_2 \in \mathcal{U}_k$ the condition $\pi_X(\nu_1) = \pi_X(\nu_2)$ implies that $d^*(\nu_1, \nu_2) < \varepsilon_k$, where d^* is a chosen metric on $\mathcal{M}_S(Y_C)$ consistent with the weak-star topology.
- U4. For any $k > 0$, $\mathcal{U}_k$ does not contain any periodic measures of period less than k.

To begin with, let Y_0 be the closure of the set of array-names of points in $X \times I$ under the action of $T \times R_0$ with respect to the partitions $\mathcal{A}_j$. In other words, Y_0 is the closure of the set of all points $y \in Y_C$ such that for some pair $(x, t) \in X \times I$ and for any j and n we have $(T^n x, R_0^n t) \in y_{j,n}$. By a standard argument, Y_0 is an extension of $X \times I$ (we will denote the corresponding map by π_0) as well as of X itself and the following diagram commutes:

$$\begin{tikzcd}
& Y_0 \arrow[dl, "\pi_0"'] \arrow[dd, "\pi_X"] \\
X \times I \arrow[dr, "\pi^{(1)}"] & \\
& X
\end{tikzcd}$$

Let the set $\mathcal{U}_0$ be all of $\mathcal{M}_S(Y_C)$ (all our requirements on the properties of $\mathcal{U}_k$ only apply to the case $k > 0$).

There are two important observations to be made here: Firstly, the only points in $X \times I$ that have multiple preimages under π_0 are the ones whose orbits enter the graph of a function from some $\mathcal{F}_j$ (we are using the fact that the graphs of continuous functions are closed). The product measure $\mu \times \lambda$ of the graph of any function is 0 (recall that λ

denotes the Lebesgue measure on the circle). Therefore whenever ν is a measure on Y_0 that factors onto a measure $\mu \times \lambda$ on $X \times I$, then the set of points in $X \times I$ with multiple preimages by π_0 has zero measure $\mu \times \lambda$. This implies that the measure-theoretic systems (Y_0, S, ν) and $(X \times I, T \times R_0, \mu \times \lambda)$ are isomorphic and ν is a unique preimage of $\mu \times \lambda$. Secondly, any j by n rectangle in Y_0 is associated with a unique cell of $\mathcal{A}_j^n$, the closure of which is the image (by π_0) of this rectangle.

We will now proceed to create the systems Y_k, requiring them to have the following properties:

Y1. For each k, $\mathcal{M}_S(Y_k) \subset \mathcal{U}_k$.

Y2. For each k, $Y_k = \Phi_k(Y_{k-1})$, where Φ_k is a conjugacy, and there exists an increasing sequence j_k such that Φ_k leaves the rows with indices greater than or equal to j_{k+1} unchanged.

Observe that the property (Y2) ensures that the diagram

$$\begin{tikzcd}
 & Y_0 \arrow[r, leftrightarrow, "\Phi_0"] \arrow[dl, "\pi_0"'] \arrow[dd, "\pi_X"] & Y_1 \arrow[r, leftrightarrow, "\Phi_1"] \arrow[ddl, "\pi_X"] & Y_2 \arrow[r, leftrightarrow, "\Phi_2"] \arrow[ddll, "\pi_X"] & \cdots \\
X \times I \arrow[dr, "\pi^{(1)}"] & & & & \\
 & X & & &
\end{tikzcd}$$

commutes and that for $j \geq j_k$ we still have the one-to-one correspondence between rectangles of size j by n in Y_k and the cells of $\mathcal{A}_j^n$, since this correspondence depends only on the contents of row j. Throughout, π_k will denote the factor map of Y_k onto $X \times I$ defined by composing the factorization π_0 of Y_0 with the conjugacy between Y_0 and Y_k.

To facilitate describing the steps of the induction (and demonstrating the properties of the obtained systems), we will endow Y_C (and thus every Y_k) with two additional rows, each of them over the alphabet $\{0, 1, \ldots, \infty\}$. We will call them the -1st row (the marker row) and the -2nd row (the jump-point row). (To avoid accumulation of references to the index zero, we do not define the 0th row.) These new rows are purely auxiliary; they are determined by other rows, and whenever we calculate entropies or the distance between measures, we ignore them (that is, we technically calculate entropies and distances on the factor of Y_C obtained by discarding these two rows). Furthermore, for any Y_k the marker and jump-point rows contain only symbols not exceeding k, in particular in Y_0 both rows consist entirely of zeroes. We will say that a point $y \in Y_c$ has a *marker of order l at position k* if $y_{-1,n} \geq k$. Let F_k^* be the set of points that have a marker of order k

at position 0 and let N_k^* be the smallest gap between markers of order k for all $y \in Y_C$.

We proceed with the induction. Y_0 (and $\mathcal{U}_0$) has already been defined. Suppose we have defined the system Y_{k-1} and the set $\mathcal{U}_{k-1}$ (and the numbers j_1 through j_{k-1}). Our task is to create the set $\mathcal{U}_k$ satisfying the requirements (U1)-(U4) and a system Y_k such that $\mathcal{M}_S(Y_k) \subset \mathcal{U}_k$. Let $\mathcal{P}_{k-1}$ be the set of all measures on Y_{k-1} that factor by π_{k-1} onto measures of the form $\mu \times \lambda$ on $X \times I$. As stated above, if $\nu \in \mathcal{P}_{k-1}$ and ν factors onto $\mu \times \lambda$, then (Y, S, ν) and $(X \times I, T \times R_0, \mu \times \lambda)$ are measure-theoretically isomorphic. It follows that for any ν in $\mathcal{P}_{k-1}$ we have

$$h(\nu|X) = h(\mu \times \lambda|X) = 0,$$

so in particular $h_\nu(\mathcal{R}_k|X) = 0$. As we have noted earlier, $h(\nu, \mathcal{R}_k|X)$ is upper semicontinuous at ν, provided $\nu(\partial R) = 0$ for every $R \in \mathcal{R}_k$. This is the case for any ν since cylinders (being clopen) have empty boundaries. Therefore every measure in $\mathcal{P}_{k-1}$ has a neighborhood where $h(\nu, \mathcal{R}_k|X) < \varepsilon_k$. $\mathcal{P}_{k-1}$ is compact, so by choosing a finite number of such neighborhoods covering $\mathcal{P}_{k-1}$ we can simply assume that there exists some neighborhood $\mathcal{V}_k$ of $\mathcal{P}_{k-1}$ such that for any measure $\nu \in \mathcal{V}_k$ we have $h(\nu, \mathcal{R}_k|X) < \varepsilon_k$. Since Y_{k-1} is conjugate to Y_0, each $\mu \times \lambda$ has a unique preimage on Y_k, which is to say that every measure μ on X has exactly one preimage in $\mathcal{P}_{k-1}$. This implies that there exists a neighborhood $\mathcal{V}'_k$ of $\mathcal{P}_{k-1}$ satisfying the condition (U3). Finally, since the sets $\mathcal{P}_{k-1}$ and the set of all periodic measures with period up to $k-1$ are closed and disjoint ($\mathcal{P}_{k-1}$ contains no periodic measures), there exists an open neighborhood $\mathcal{V}''_k$ of $\mathcal{P}_{k-1}$ that does not contain any periodic measures with period up to $k-1$. It is clear that the set $\mathcal{U}_k = \mathcal{U}_{k-1} \cap \mathcal{V}_k \cap \mathcal{V}'_k \cap \mathcal{V}''_k$ has the properties (U1)-(U4). We must now construct the system Y_k whose invariant measures satisfy the condition (Y1), i.e. are all in $\mathcal{U}_k$ (also making sure to satisfy the requirement (Y2)). In other words, we must ensure that every invariant measure on Y_k is close (in the space $\mathcal{M}_S(Y_C)$) to some measure on Y_{k-1} that factors onto a product measure of the form $\mu \times \lambda$ on $X \times I$. To this end we will employ the following lemma (which we copy from [D-H] with the proof):

Lemma 3.3. *For any measure μ on X and any neighborhood $U_{\mu\times\lambda}$ of $\mu \times \lambda$ in $\mathcal{M}(X \times I)$ there exists a neighborhood U_μ of μ in $\mathcal{M}(X)$, an irrational rotation R_μ of the one-dimensional torus and a number N_μ such that for any $(x, t) \in X \times I$ and any $n > N_\mu$ the condition $\mathbf{A}_n^T(\delta_x) \in U_\mu$ implies that $\mathbf{A}_n^{T\times R_\mu}(\delta_{(x,t)}) \in U_{\mu\times\lambda}$.*

Proof: First note that for any measure μ in $\mathcal{M}_T(X)$ there exists an irrational rotation R_μ disjoint from μ (i.e. the only $(T \times R_\mu)$-invariant measure on $X \times I$ with marginals μ and λ is $\mu \times \lambda$). Indeed, if μ is an ergodic measure and $e^{\alpha 2\pi i}$ is rationally independent from all its eigenvalues, then the rotation of the circle by α is disjoint from μ. Since an ergodic measure has at most countably many eigenvalues, for any ergodic μ there exist at most countably many rotations that are not disjoint from μ. If μ is not ergodic, denote its ergodic decomposition by ξ (ξ is a measure on the set $\mathcal{M}_T^e(X)$ of ergodic measures on X) and consider the product $\mathcal{M}_T^e(X) \times I$ with the measure $\xi \times \lambda$. The set $\{(\nu, \alpha) : e^{\alpha 2\pi i}$ is an eigenvalue of $\nu\}$ is a measurable subset of the product and has measure 0 (because all its vertical sections are countable), so almost every horizontal section of this set has measure 0. Therefore there exists an α such that the measures for which $e^{\alpha 2\pi i}$ is an eigenvalue have zero mass in the ergodic decomposition of μ. Setting R_μ to be the rotation by α we obtain a rotation disjoint from μ.

Suppose the statement of the lemma is not true. Then there exists a sequence of measures $\mathbf{A}_n^{T\times R_\mu}(\delta_{(x_n,t_n)})$ such that $\mathbf{A}_n^T(\delta_{x_n})$ converge to μ yet the $\mathbf{A}_n^{T\times R_\mu}(\delta_{(x_n,t_n)})$ all lie outside $U_{\mu\times\lambda}$ (remember that the averaging in $X \times I$ is with respect to $T \times R_\mu$). Choose the limit ν of some subsequence of $\mathbf{A}_n^{T\times R_\mu}(\delta_{(x_n,t_n)})$. It is a $T \times R_\mu$-invariant measure which is outside $U_{\mu\times\lambda}$ and whose marginals are μ (being the limit of $\mathbf{A}_n^T(\delta_{x_n})$) and λ (being the only R_μ-invariant measure on I). But the only $T \times R_\mu$-invariant measure with marginals μ and λ is $\mu \times \lambda$, which is in $U_{\mu\times\lambda}$ – a contradiction. ■

Note that Y_0 has the property that each of its elements is entirely determined by its rows with indices larger than some j (this is true for every $j \geq 1$). Since Y_{k-1} is topologically conjugate to Y_0 and has the same rows from j_{k-1} onwards, every element of Y_{k-1} is entirely determined by the rows from j_{k-1} onwards. Recall that the system Y_{k-1} is equipped with two extra rows: the marker row labeled -1 and the jump-point row labeled -2 (which we have not discussed yet, and which are entirely zeros in Y_0), which are determined by the positive-indexed rows. From what was said in the preceding sentence, the rows $-1, -2$ are determined by the rows from j_{k-1} onwards. It follows that for any $y \in Y_{k-1}$ we can determine $y_{j,0}$ for $j \leq j_{k-1}$ (including $j = -1, -2$) by looking at a large enough rectangle (symmetric about column zero) in y, of size independent of y, contained in the rows from j_{k-1} onwards.

Moreover, for $j \geq j_{k-1}$, the inequality $(T \times R_0)^{-1}(\mathcal{A}_j) \vee \mathcal{A}_j \vee (T \times R_0)\mathcal{A}_j \prec \mathcal{A}_{j+1}$ means that each symbol determines the three consecutive symbols in the preceding row (i.e. knowing $y_{j,m}$ we know $y_{j-1,m-1}, y_{j-1,m}$ and $y_{j-1,m+1}$). Therefore if we choose j_k large enough, the symbol $y_{j_k-1,0}$ alone suffices to determine the aforementioned rectangle in y, and thus to determine $y_{j,0}$ for all $j \leq j_{k-1}$ (including $j = -1$ and $j = -2$). Clearly, the symbols $y_{j,0}$ for $j_{k-1} < j \leq j_k - 1$ are also determined by $y_{j_k-1,0}$. Concluding, for every $y \in Y_{k-1}$ and every n, the symbol $y_{j_k-1,n}$ determines all the symbols $y_{j,n}$ for $j = -1, -2, 1, 2, \ldots, j_k - 1$.

As the open set $\mathcal{U}_k$ contains the compact set $\mathcal{P}_{k-1}$, there exists some number ε such that the ε-neighborhood of $\mathcal{P}_{k-1}$ is contained in $\mathcal{U}_k$. By increasing j_k (if necessary) we can find δ and d_0 such that if two measures differ by less than δ on all rectangles of height $j_k - 1$ (in rows 1 through $j_k - 1$) and length d_0 (call the family of these rectangles $\mathcal{D}_k$), then the distance between these measures is less than ε.

Let d be an integer so large that $\frac{d_0}{d} < \frac{\delta}{5}$. For any T^d-invariant measure μ on X the partition $\mathcal{A}^{d_0}_{j_k-1}$ (where the "exponent" d_0 refes to the action of T rather than T^d, i.e., $\mathcal{A}^{d_0}_{j_k-1} = \bigvee_{n=0}^{d_0-1}(T \times R_0)^{-n}(\mathcal{A}_{j_k-1})$) has boundaries of measure $\mu \times \lambda$ equal to 0. Therefore there exists a neighborhood $U_{\mu\times\lambda}$ of $\mu \times \lambda$ in $\mathcal{M}(X \times I)$ such that if $\nu \in U_{\mu\times\lambda}$ then $|\nu(A) - (\mu \times \lambda)(A)| < \frac{\delta}{5}$ for every $A \in \mathcal{A}^{d_0}_{j_k-1}$. Moreover, by making, in necessary, $U_{\mu\times\lambda}$ even smaller, we can also have $\big|(T^j \times R_0)(\nu)(A) - (T^j \times R_0)(\mu \times \lambda)(A)\big| < \frac{\delta}{5}$ for $j = 0, \ldots, d-1$.

Applying Lemma 3.3 to $U_{\mu\times\lambda}$ and T^d, we obtain for every $\mu \in \mathcal{M}_{T^d}(X)$ an open set U_μ around μ in $\mathcal{M}(X)$. Out of these we select a finite family $\mathcal{W}$ of measures such that the union of U_μ, with μ ranging over $\mathcal{W}$, covers $\mathcal{M}_{T^d}(X)$ in $\mathcal{M}(X)$. The union of this cover is an open set in $\mathcal{M}(X)$. There exists a number N such that every measure of the form $\mathbf{A}^{T^d}_N(\delta_x)$ is in U_μ for some $\mu \in \mathcal{W}$ (we are using Fact 2.3). We can also assume that N is larger than the numbers N_μ of Lemma 3.3 for all $\mu \in \mathcal{W}$.

The system Y_{k-1} is aperiodic, since it is an extension of the aperiodic system $(X \times I, T \times R_0)$. By applying the marker lemma (Lemma 2.1) to Y_{k-1} with the constant $N_k = Nd$, we obtain a set F_k (which we can assume to be contained in the set F^*_{k-1}) such that every point of Y_{k-1} visits F_k with gaps between N_k and $2N_k$. The set F_k decomposes into disjoint, clopen sets $F^0_k, F^1_k, \ldots, F^{N_k}_k$, such that F^i_k consists of points returning to F_k after exactly $N_k + i$ steps. Each F^i_k can in turn be decomposed into disjoint sets of the form $F^i_k \cap R$, where $R \in \mathcal{R}^{N_k+i}_{j_k}$ (i.e. all points in the same $F^i_k \cap R$ have the same symbols in rows 1 through j_k between the coordinate 0 and the next occurrence of the marker). Therefore we can decompose F_k into a family $\mathcal{C}_k$ of disjoint clopen sets (of the form $C = F^i_k \cap R$) such that all points in a chosen $C \in \mathcal{C}_k$ return to F_k after the same time and they all have the same content of rows 1 through j_k until that time. We will call such sets *k-rectangles.* Notice that, by the above described rules of determining the symbols in rows $-1, -2$, the k-rectangle also determines the symbols in these two rows at the corresponding horizontal coordinates.

To summarize, all points in the same k-rectangle share the following properties:

- They have a marker of order k at position 0.
- They all have the same symbols in rows 1 through j_k between coordinate zero and the next marker of order k.
- They all have the same content of the marker and jump-point rows at the above coordinates.
- All the above is entirely determined by the last row j_k between coordinates zero and the next marker of order k.

Now, let C be a k-rectangle in Y_{k-1} and let N_C be its length (i.e. the shared return time to F_k for all points in C; a number between Nd and $2Nd$). Set $N'_C = \lceil N_C/d \rceil$ (the number of non-overlapping subblocks of length d in C). Since row j_k has so far remained unaltered (and it determines the other rows in C), $\pi_{k-1}(C)$ is the closure of a single cell of $\mathcal{A}_k^{N_C}$. We choose a point (x_C, t_C) from the (easily seen to be nonempty) interior of this cell, and we choose $\mu_C \in \mathcal{W}$ such that $\mathbf{A}_{N'_C}^{T^d}(\delta_{x_C})$ belongs to U_{μ_C} (this is possible since $N'_C \geq N$). We can assume that $(x_C, R_{\mu_C}^i R_0^{-i \cdot d} t_C)$ (for all $i = 0, 1, 2, \ldots, N'_C$) does not enter the boundary of any set from $\mathcal{A}_{j_k}$ over $2N_k$ iterations of $T \times R_0$.

This implies that, for each i as above, all preimages by π_{k-1} of the point $(x_C, R_{\mu_C}^i R_0^{-i \cdot d} t_C)$ have a common k-rectangle (the content of the rows 1 through j_k at the coorinates 0 through N_C). We denote this k-rectangle by C_i (in particular, $C_0 = C$).

We now establish the jump points of order k for C as follows: For $i = 1, \ldots, N'_C$, let n_i be the closest number to $i \cdot d$ such that:

- For each $l < k$, if we set $L_l(i)$ to be the last marker of order l before position n_i in C_{i-1}, and $R_l(i)$ to be the first marker of order l after position n_i in C_i, then both $R_l(i) - n_i$ and $n_i - R_l(i)$ are greater than $\frac{N_l}{3}$.
- With the notation as above, there are no jump points of order greater than l between the positions $L_l(i)$ and n_i in C_{i-1}, and between n_i and $R_l(i)$ in C_i.

In informal terms, n_i is a position close to $i \cdot d$ that is "reasonably" distant from all markers and does not fall between such pairs of consecutive markers of order l, which are already separated by a jump point of order greater than l (and smaller than k). It is not hard to see that if N_k grow fast enough, such numbers n_i exist, moreover, we can assume that each n_i differs from $i \cdot d$ by no more than N^*_{k-1} (recall that N^*_{k-1} is the smallest possible gap between markers of order $k-1$). To make the notation in the following formula homogeneous, we also let $n_0 = 0$.

Now we define the image $\Phi_k(C)$ of C as follows: For each $i = 1, \dots, N'_C$, in rows from 1 to $j_k - 1$, we replace in C the content of the columns n_i up to (not including) n_{i+1} with the content of the same columns from C_i, including markers and jump points. (Note that we make no changes in row j_k, the final row of C). Also, for every i (as above) we put the symbol k at the position n_i of the jump point row. The mapping Φ_k (formally defined on k-rectangles) induces a map on Y_{k-1} (which we will denote by the same symbol Φ_k) in a natural way, as a code replacing k-rectangles by their images. Let Y_k be the image of Y_{k-1} by this block code. Note that the column condition is preserved (hence $Y_k \subset Y_C$): Since the code replaces the columns up to row $j_k - 1$ by other columns existing in Y_{k-1}, it can only introduce a violation of the column condition between rows $j_k - 1$ and j_k. However, observe that the symbols in an image k-rectangle $\Phi_k(C)$ appearing in rows $j_k - 1$ and j_k at the same position n both correspond to cells (of the partitions $\mathcal{A}_{j_k-1}$ and $\mathcal{A}_{j_k}$, respectively) whose projections on the first coordinate both contain the point $T^n x_C$, which easily implies the column condition.

We will now show that any invariant measure on Y_k is in $\mathcal{U}_k$. In order to do this, it suffices to show that if C is a k-rectangle in Y_k, then any rectangle $D \in \mathcal{D}_k$ (of length d_0) occurs in C with frequency close to some $\nu_C(D)$, where $\nu_C \in \mathcal{P}_k$. For ease of calculation, we first assume that the jump points in C are placed exactly every d positions, i.e. that $n_i = i \cdot d$. Since $\frac{d_0}{d} < \frac{\delta}{5}$, the frequency of occurrences of D in C (say, $F_C(D)$) differs by at most $\frac{\delta}{5}$ from the average frequency of its occurrences in consecutive segments of length d, i.e. from the expression

$$\frac{1}{N'_C} \sum_{i=0}^{N'_C - 1} F_{C[i \cdot d, (i+1) \cdot d - 1]}(D).$$

However, due to the aforementioned correspondence between rectangles and cells in the product,

$$F_{C[i \cdot d, (i+1) \cdot d - 1]}(D) \approx \mathbf{A}_d^{T \times R_0} \delta_{(T^{i \cdot d} x_C, R^i_{\mu_C} t_C)}(\pi_{k-1}(D)),$$

because the bottom row of $C[i \cdot d, (i+1) \cdot d - 1]$ is a fragment of the orbit-name of the point $(x_C, R^i_{\mu_C} R_0^{-i \cdot d} t_C)$ under $T \times R_0$ with respect to $\mathcal{A}_{j_k}$. The error of this approximation comes only from the "end effect" and is at most $\frac{d_0}{d} < \frac{\delta}{5}$. Therefore,

$$\begin{aligned}
F_C(D) &\overset{\frac{2\delta}{5}}{\approx} \frac{1}{N'_C}\sum_{i=0}^{N'_C-1} \mathbf{A}_d^{T\times R_0}\delta_{(T^{i\cdot d}x_C, R^i_{\mu_C}t_C)}(\pi_{k-1}(D)) = \\
&= \frac{1}{N'_C}\sum_{i=0}^{N'_C-1}\left(\frac{1}{d}\sum_{j=0}^{d-1}\mathbf{1}_{\pi_{k-1}(D)}(T^{i\cdot d+j}x_C, R^i_{\mu_C}R_0^j t_C)\right) = \\
&= \frac{1}{dN'_C}\sum_{i=0}^{N'_C-1}\sum_{j=0}^{d-1}\mathbf{1}_{\pi_{k-1}(D)}(T^{i\cdot d+j}x_C, R^i_{\mu_C}R_0^j t_C) = \\
&= \frac{1}{d}\sum_{j=0}^{d-1}\left(\frac{1}{N'_C}\sum_{i=0}^{N'_C-1}\mathbf{1}_{\pi_{k-1}(D)}(T^{i\cdot d+j}x_C, R^i_{\mu_C}R_0^j t_C)\right) = \\
&= \frac{1}{d}\sum_{j=0}^{d-1}\mathbf{A}_{N'_C}^{T^d\times R_{\mu_C}}\delta_{(T^j x_C, R_0^j t_C)}(\pi_{k-1}(D)).
\end{aligned}$$

By lemma 3.3, $\mathbf{A}_{N'_C}^{T^d\times R_{\mu_C}}\delta_{(x_C,t_C)} \in U_{\mu_C\times\lambda}$, which implies

$$\mathbf{A}_{N'_C}^{T^d\times R_{\mu_C}}\delta_{(T^j x_C, R_0^j t_C)}(\pi_{k-1}(D)) \overset{\frac{\delta}{5}}{\approx} (T^j\mu_C\times\lambda)(\pi_{k-1}(D)).$$

Therefore,

$$\begin{aligned}
F_C(D) &\overset{\frac{3\delta}{5}}{\approx} \frac{1}{d}\sum_{j=0}^{d-1}(T^j\mu_C\times\lambda)(\pi_{k-1}(D)) = \\
&= \left(\left(\frac{1}{d}\sum_{j=0}^{d-1}T^j\mu_C\right)\times\lambda\right)(\pi_{k-1}(D)).
\end{aligned}$$

Since $\frac{1}{d}\sum_{j=0}^{d-1}T^j\mu_C$ is a T-invariant measure (as μ_C was T^d-invariant), we conclude that for every rectangle $D\in\mathcal{D}_k$ we have (provided that d and then N were chosen large enough)

$$|F_C(D)-\nu_C(D)| < \frac{3\delta}{5},$$

where $\nu_C\in\mathcal{P}_k$ is the unique measure on Y_{k-1} which projects onto $\left(\frac{1}{d}\sum_{j=0}^{d-1}T^j\mu_C\right)\times\lambda$. Recall that the above calculation was based on the assumption that the jump points were exact multiples of d. Even if that is not the case, each n_i differs from $i\cdot d$ by less than N^*_{k-1}.

Therefore, if $\frac{N^*_{k-1}}{d} < \frac{\delta}{5}$ (which we can assume), the relative error in $n_i \approx i \cdot d$ is also less than $\frac{\delta}{5}$, and thus in the general case

$$|F_C(D) - \nu_C(D)| < \frac{4\delta}{5}.$$

If N_k is large enough, then for any invariant measure ν on Y_k, $\nu(D)$ differs by no more than $\frac{\delta}{5}$ from some convex combination of the form $\sum_{m=1}^M \alpha_m F_{C_m}(D)$ (with the α_m and C_m depending only on ν, and not on D). Therefore $\left|\nu(D) - \sum_{m=1}^M \alpha_m \nu_{C_m}(D)\right| < \delta$ for all D. But $\sum_{m=1}^M \alpha_m \nu_{C_m} \in \mathcal{P}_k$, and thus $\nu \in \mathcal{U}_k$, as requested.

We have now ensured that in Y_k any rectangle D from $\mathcal{D}_k$ occurs between two consecutive markers of order k with frequency close to some product measure in $X \times I$ of the cell corresponding to D. However, in subsequent steps of the induction we will "cut" rectangles at the jump points, which fall between markers. The way we choose the jump points ensures that any rectangle between two markers of order k will be cut at most once throughout the remaining steps and that the cut will fall between 1/3 and 2/3 of its length. Therefore if we now increase the length of $k-$rectangles so much that any $D \in \mathcal{D}_k$ will occur with controlled frequency over 1/3 of the new length, this property will be preserved throughout the subsequent steps of the induction.

Given the above, let $N^*_k = \frac{3N_k}{\delta}$ and apply the marker lemma to Y_k with the constant N^*_k. We obtain a set F^*_k which we can assume to be a subset of F_k. For any $y \in Y_k$, if $S^n(y) \in F^*_k$, we put the symbol k in the marker row. As a result, if C is any rectangle (in rows 1 through j_k) appearing in Y_k, starting or ending at a marker k and extending at least $\frac{N^*_k}{3}$ to the right (resp. left) of it but not beyond the next marker k, then for any $D \in \mathcal{D}_k$ $F_C(D) \approx \nu(D)$ (up to δ) for some $\nu \in \mathcal{P}_k$. The rule of choosing the jump-points ensures that this last property (involving the markers k) passes to elements of all systems $Y_{k'}$ for all $k' \geq k$.

We are now ready to define the desired faithful zero-dimensional principal extension. First replace each Y_k by a system obtained by deleting the rows -1 and -2. Now set

$$Y = \bigcap_{m=1}^{\infty} \overline{\bigcup_{k=m}^{\infty} Y_k}.$$

In other words, Y is the set of all points y such that $y = \lim_k y_k$, $y_k \in Y_k$ (without regard to markers and jump points).

Note that if $k > k'$, then any k'-rectangle in Y_k is either exactly some k'-rectangle from $Y_{k'}$ (if it contains no jump points of order greater than k') or a concatenation of fragments of two such k'-rectangles, each being at least $\frac{1}{3}$ (in length) of the original. Let ν be any invariant measure on Y. For each large enough k, ν is well-approximated by the periodic measure carried by some rectangle C of height $j_{k'}$ and length $N > 2N_{k'}$. Since C must occur in some Y_k, it is a concatenation of segments of k'-rectangles from $Y_{k'}$ (each such segment of length at least $\frac{N^*_{k'}}{3}$. This, however, means that any $D \in \mathcal{D}_k$ occurs in C with frequency close to $\nu_{k'}(D)$ for some $\nu_{k'} \in \mathcal{P}_{k'}$, and thus $\nu \in \mathcal{U}_{k'}$.

Since any invariant measure on Y is in $\mathcal{U}_k$ for every k, Y obviously has no periodic invariant measures due to property (U4).

To show that Y is a principal extension of X we need to show that the conditional entropy of Y with respect to X is 0 for every measure $\nu \in \mathcal{M}_S(Y)$. For any $k > 0$ and for any $k' > k$ we have $h(\nu, \mathcal{R}_k|X) \leq h(\nu, \mathcal{R}_{k'}|X)$, since $\mathcal{R}_{k'} \succ \mathcal{R}_k$. On the other hand, since ν is in the set $\mathcal{U}_{k'}$, using the property U2, we know that $h(\nu, \mathcal{R}_{k'}|X) < \varepsilon_{k'}$. It follows that for any $k' > k$ $h(\nu, \mathcal{R}_k|X) < \varepsilon_{k'}$, and thus $h(\nu, \mathcal{R}_k|X) = 0$. Thus we conclude that $h(\nu|X) = 0$.

Similarly, since $\mathcal{M}_S(Y) \subset \mathcal{U}_k$ for every k, using the property U3, if two invariant measures on Y factor onto the same measure on X, then they must be closer to each other than ε_k for all k, and thus every invariant measure on X has exactly one preimage on Y. ■

4 Final remarks

The result presented in this paper can be combined with the following result of J. Serafin in [S], which uses the notion of extension entropy: For a factor map $\phi : Y \to X$, we define

$$h^{\phi}_{\text{ext}}(\mu) = \sup \{h(S, \nu) : \nu \in \mathcal{M}_S(Y) \text{ and } \phi(\nu) = \mu\}.$$

The result of [S] can be phrased in two ways, the latter of which is in terms of entropy structures and affine superenvelopes discussed in [B-D]:

Theorem 4.1. *Let (X, T) be a finite entropy zero-dimensional dynamical system without periodic points, and let $\psi : (Y, S) \to (X, T)$ be a symbolic extension. Then there exists another symbolic extension $\phi : (W, S) \to (X, T)$ such that*

1. $h^{\phi}_{ext} \equiv h^{\psi}_{ext}$ *on $\mathcal{M}_T(X)$;*
2. *Every measure in $\mathcal{M}_T(X)$ has exactly one preimage in $\mathcal{M}_S(W)$.*

Theorem 4.2. *Let (X,T) be a finite entropy zero-dimensional dynamical system without periodic points, with a given entropy structure $\mathcal{H} = (h_k)$. Suppose that E_A is a bounded affine superenvelope of $\mathcal{H}$. Then there exists a symbolic extension $\phi : (W,S) \to (X,T)$ such that*

1. *$h^{\phi}_{ext} = E_A$;*
2. *Every measure in $\mathcal{M}_T(X)$ has exactly one preimage in $\mathcal{M}_S(W)$.*

Combined with our Theorem 3.1, the above results yield the following corollaries:

Corollary 4.3. *Let (X,T) be a finite entropy dynamical system and let $\psi : (Y,S) \to (X,T)$ be a symbolic extension. Then there exists another symbolic extension $\phi : (W,S) \to (X,T)$ such that*

1. *$h^{\phi}_{ext} \equiv h^{\psi}_{ext}$ on $\mathcal{M}_T(X)$;*
2. *Every measure in $\mathcal{M}_T(X)$ has exactly one preimage in $\mathcal{M}_S(W)$.*

Corollary 4.4. *Let (X,T) be a finite entropy dynamical system with a given entropy structure $\mathcal{H} = (h_k)$. Suppose that E_A is a bounded affine superenvelope of $\mathcal{H}$. Then there exists a symbolic extension $\phi : (W,S) \to (X,T)$ such that*

1. *$h^{\phi}_{ext} = E_A$;*
2. *Every measure in $\mathcal{M}_T(X)$ has exactly one preimage in $\mathcal{M}_S(W)$.*

In particular, note that we do not have to assume the lack of periodic points in (X,T), since the (intermediate) zero-dimensional system we obtain in Theorem 3.1 has no periodic points, and thus we can immediately apply the theorems of [S].

References

[B-D] M. Boyle and T. Downarowicz, *The entropy theory of symbolic extensions*, Inventiones Math. **156** (2004), pp 119-161.

[D1] T. Downarowicz, *Entropy Structure*, Journal d'Analyse **96** (2005), pp. 57-116

[D2] T. Downarowicz, *Faces of simplices of invariant measures*, Israel J. Math., *to appear*

[D-H] T. Downarowicz and D. Huczek, *Zero-dimensional principal extensions*, Acta Applicandae Mathematica, *to appear*

[Le] F. Ledrappier *A variational principle for the topological conditional entropy*, Springer Lec. Notes in Math. **729**, Springer-Verlag (1979), pp 78-88

[Li] E. Lindenstrauss, *Mean dimension, small entropy factors and an imbedding theorem*, Publ. Math. I.H.E.S. **89** (1999), pp 227–262

[L-W] E. Lindenstrauss and B. Weiss, *Mean Topological Dimension*, Israel J. Math **115** (2000), pp 1-24

[S] J. Serafin, *A faithful symbolic extension*, to be published.