ERGODIC UNIVERSALITY OF SOME TOPOLOGICAL DYNAMICAL SYSTEMS

ANTHONY QUAS AND TERRY SOO

ABSTRACT. The Krieger generator theorem says that every invertible ergodic measure-preserving system with finite measure-theoretic entropy can be embedded into a full shift with strictly greater topological entropy. We extend Krieger's theorem to include toral automorphisms, and more generally, any topological dynamical system on a compact metric space that satisfies almost weak specification, asymptotic entropy expansiveness, and the small boundary property. As a corollary, one obtains a complete solution to a natural generalization of an open problem in Halmos's 1956 book regarding an isomorphism invariant that he proposed.

1. INTRODUCTION

Let S be a self-homeomorphism of a compact metric space (Y, d). Let T be an invertible ergodic measure-preserving transformation on a non-atomic probability space (Ω, μ). An ***embedding*** of (Ω, μ, T) into (Y, S) is a measurable mapping $\Psi : \Omega \to Y$ such that the restriction of Ψ to a set of full measure Ω' is an injection, and $\Psi(T(\omega)) = S(\Psi(\omega))$ for all $\omega \in \Omega'$. We say that the topological dynamical system (Y, S) is ***universal*** if for every invertible ergodic measure-preserving system (X, μ, T) with measure-theoretic entropy strictly less than the topological entropy of S there exists an embedding of (Ω, μ, T) into (Y, S), and we say that (Y, S) is ***fully universal*** if the embedding can be chosen so that the push-forward of the measure on Ω is fully supported on Y. The Krieger finite generator theorem [30, 31] says that the full shift on a finite number of symbols is universal. We prove the following extension of Krieger's theorem.

Theorem 1. *Toral automorphisms are universal.*

Following Lind [36], we say that a toral automorphism is **quasi-hyperbolic** if its associated matrix A has no roots of unity as eigenvalues, and **hyperbolic** if A does not have an eigenvalue of modulus 1. Lind and Thouvenot [37, Section 5] proved that hyperbolic (two-dimensional) toral automorphisms are fully universal and asked whether the same is true in the quasi-hyperbolic case, and in the more general case of an automorphism of a compact group. Lind and Thouvenot made use of the fact that hyperbolic toral automorphisms can be represented as irreducible shifts of finite type; this is not true in the non-hyperbolic case [35, Section 6], [36, Theorem 4]. We prove an affirmative answer to their question, which easily implies Theorem 1.

Theorem 2. *A quasi-hyperbolic toral automorphism is fully universal.*

Theorem 2 will be proved as part of a more general result which we state in Section 2. The result has three conditions, one of which is a form of specification.

Let S be a self-homeomorphism of a compact metric space (Y, d). Suppose that for every $\varepsilon > 0$ there exists a function $L_\varepsilon : \mathbb{Z}^+ \to \mathbb{Z}^+$ such that given a finite number of points $y_1, \ldots, y_n \in Y$ and finite sequence of integers $a_1 \le b_1 < a_2 \le b_2 \cdots a_n \le b_n$ with $a_i - b_{i-1} \ge L_\varepsilon(b_i - a_i)$ for all $i \in [2, n]$, there is a $y \in Y$ with $\mathrm{d}(S^k y, S^k y_i) \le \varepsilon$ for all $k \in [a_i, b_i]$ and $i \in [1, n]$. We call L_ε a **gap** function. If the gap function L_ε satisfies $L_\varepsilon(m)/m \to 0$ as $m \to \infty$, then we say that S satisfies **almost weak specification**; if the function L_ε is a constant function, then S satisfies **weak specification**, and in addition, if y can be chosen to be a periodic point, then S satisfies **specification** (for background see [4, 13, 51, 57]).

Marcus [39] proved that quasi-hyperbolic toral automorphisms satisfy almost weak specification. Let us remark that a toral automorphism is ergodic with respect to Haar measure if and only if it is quasi-hyperbolic [24].

Conjecture 1. *A self-homeomorphism with almost weak specification on a compact space is (fully) universal.*

Recently, we also proved that the time-one map of a geodesic flow on a compact surface of negative curvature is universal [45]. With the help of the symbolic dynamics for geodesic flows developed by Bowen [6] and Ratner [46], it was sufficient to show that the time-one map of a topologically weak-mixing suspension flow over an irreducible subshift of finite type is universal; in our proof of this result we were aided by the symbolic nature of the suspension flow and the fact that the time-one map satisfies (weak) specification [45, Proposition 5].

Krieger also proved that mixing subshifts of finite type are universal [33] (see also the proof given by Denker [14, Theorem 28.1]). We also prove the following generalization of Krieger's theorem, and Lind and Thouvenot's result [37, Theorem 2] that every mixing subshift of finite type is fully universal.

Theorem 3. *A subshift with almost weak specification on a finite number of symbols is fully universal.*

We will prove a weaker version of Conjecture 1, Theorem 7, from which Theorems 2 and 3 will follow. We will require some additional conditions that are satisfied in Theorems 2 and 3. We will give the precise statement of these additional conditions in the next section.

In our proof of Theorem 7, we make use of an idea of Burton and Rothstein [11], further developed in work of Burton, Keane and Serafin [10], and Downarowicz and Serafin [18] that involves producing the required injections using the Baire Category theorem.

In Section 3 we will prove Theorems 2 and 3 using our more general result, and apply Theorem 3 to show that β-shifts are universal. We will also generalize Theorem 2 to include any automorphism of a compact metric abelian group that is ergodic with respect to Haar measure. In Section 4 we set the stage for the proof of our more general result; the proof will be carried out in the remaining sections.

We would like to thank Jean-Paul Thouvenot for introducing us to the problem.

2. Non-expansive homeomorphisms and the small boundary condition

We recall in this section some basic tools to deal with non-expansive homeomorphisms. Let (X, d) denote a compact metric space. For $r > 0$, we let $B(x, r)$ denote the open ball about a point $x \in X$, and for a subset $A \subset X$, we let $B(A, r) := \bigcup_{x \in A} B(x, r)$. If $\mathcal{P}$ is a finite (Borel-)measurable partition of X its ***diameter*** is defined to be the maximum of the diameters of the elements of the partition. The set $\partial \mathcal{P}$ is the union of the topological boundaries of the elements of $\mathcal{P}$ and $\partial_r \mathcal{P}$ denotes $\overline{B(\partial \mathcal{P}, r)}$. We recall that a partition is said to be ***generating*** if for any distinct pair of points, x and y, there exists an $n \in \mathbb{Z}$ such that $T^n x$ and $T^n y$ lie in different elements of $\mathcal{P}$.

A self-homeomorphism T of a metric space (X, d) is ***expansive*** if there exists a $\delta > 0$ (the ***expansiveness constant***) such that for all $x, y \in X$ if $\mathrm{d}(T^n x, T^n y) < \delta$ for all $n \in \mathbb{Z}$, then $x = y$. Expansive homeomorphisms have many generating partitions: Indeed any partition of diameter less than the expansiveness constant is generating. We

recall that any subshift (that is the restriction of the shift map to a non-empty closed shift-invariant subset of the full shift) is expansive.

We write $M_T(X)$ for the collection of T-invariant Borel probability measures on X. Another desirable feature of expansive homeomorphisms is that the entropy functional sending an invariant measure to its measure-theoretic entropy is upper semi-continuous with respect to the weak* topology (see Lemma 14 for an explicit metric which generates the weak* topology). In order to consider non-symbolic spaces we will make use of two successive weakenings due to Bowen [5] and Misiurewicz [41] which allow us to recover the upper semi-continuity of the entropy functional.

For each $\delta > 0$, define the set

$$\Gamma_\delta(x) := \{y \in X \colon \mathrm{d}(T^j x, T^j y) < \delta \text{ for all } j \geq 0\}.$$

If there is a constant $\delta > 0$ such that $h_{\mathrm{top}}(T, \Gamma_\delta(x)) = 0$ for all $x \in X$, then T is **entropy expansive**, and if $\sup_{x \in X} h_{\mathrm{top}}(T, \Gamma_\delta(x)) \to 0$ as $\delta \to 0$, then T is **asymptotically entropy expansive**.

That these are strict weakenings of expansiveness is seen by simple examples such as the identity map, twist maps such as $T(x, y) = (x, y + x) \bmod 1$. More substantial examples are given by the following lemma.

Lemma 4 (Bowen [5] Example 1.2). *Toral automorphisms are entropy expansive.*

Thus while hyperbolic toral automorphisms are expansive, in general toral automorphisms are only entropy expansive.

Lemma 5 (Misiurewicz [41] Corollary 4.1). *Let T be an asymptotically entropy expansive self-homeomorphism of a compact metric space X. Then the entropy functional $\mu \mapsto h_\mu(T)$ is upper semi-continuous with respect to the weak* topology on $M_T(X)$.*

Bowen proved Lemma 5 in the case where T is entropy expansive. For an example of a map that satisfies specification, but is not asymptotically entropy expansive and does not have an upper semi-continuous entropy functional see [44, Page 952].

Another property that we need is sometimes called the small boundary property or the existence of an essential partition. A metric space is **d-dimensional** (i.e. it has topological dimension d) if there exists an open cover with arbitrarily fine diameter with the property that the intersection of any $d + 2$ distinct elements is empty. A self-homeomorphism, T, of a compact metric space X is said to have the **small boundary property** if for each $\delta > 0$, there exists a partition $\mathcal{P}$ of X such that $\mathrm{diam}(\mathcal{P}) < \delta$ and $\mu(\partial \mathcal{P}) = 0$ for each T-invariant

measure μ. Notice that if $\mu(\partial\mathcal{P}) = 0$, then $\mu(\partial(T^{-1}\mathcal{P})) = 0$. Also if $\mu(\partial\mathcal{P}) = 0$ and $\mu(\partial\mathcal{Q}) = 0$, then $\mu(\partial(\mathcal{P} \vee \mathcal{Q})) = 0$.

Both the small boundary property and asymptotic entropy expansiveness are important properties in the theory of symbolic extensions and entropy structure; for more information see the recent book of Downarowicz [17] and the recent articles of Boyle and Downarowicz [7], Burguet [9], Downarowicz [16, 15], and Lindenstrauss [38]. We make use of the following result of Kulesza.

Lemma 6 (Kulesza [34] Lemma 3.7). *A self-homeomorphism of a compact finite-dimensional metric space with the property that the periodic points form a zero-dimensional set has the small boundary property.*

Krieger [33] raised the question of when a homeomorphism of a compact metric space is universal. We have the following partial answer.

Theorem 7. *A self-homeomorphism of a compact metric space is fully universal whenever if satisfies:*

  (1) almost weak specification;
  (2) asymptotic entropy expansiveness; and
  (3) the small boundary property.

Notice that Theorem 3 follows immediately from Theorem 7.

3. EXAMPLES

3.1. **Subshifts.** We call a subshift ***non-trivial*** if it does not consist of a finite set of points. Subshifts have a natural generating partition given by $\mathcal{P} := \{P_1, \ldots P_n\}$, where $P_i = \{x \in X : x_0 = i\}$ for all $1 \leq i \leq n$.

Familiar examples of subshifts that satisfy specification are mixing subshifts of finite type, and mixing sofic subshifts. However, there are many subshifts that do not satisfy almost weak specification, but are still universal.

Corollary 8. *If a subshift contains subshifts that satisfy almost weak specification and have topological entropy arbitrarily close to the original topological entropy, then it is universal.*

Proof. Immediate from Theorem 3 and the definition of universality. $\square$

The β-shifts were introduced by Rényi [47] and can be defined as follows. For $\beta > 1$, let T be the full-shift on $\Sigma_\beta = \Sigma := \{0, 1, \ldots, \lfloor\beta\rfloor\}$ and let $\preceq$ denote the lexicographic order on $\Sigma^{\mathbb{Z}}$ and $\Sigma^{\mathbb{Z}^+}$. Let $\omega \in \Sigma^{\mathbb{Z}^+}$

be the locally greedy β-expansion of 1; that is if $a \in \Sigma^{\mathbb{Z}^+}$ is a solution to the equation

$$1 = \sum_{j=1}^{\infty} \frac{a_j}{\beta^j},$$

then $a \preceq \omega$. The β-shift is the subshift of the full-shift such that $x \in \Sigma^{\mathbb{Z}}$ belongs to the subshift if and only if $(T^k x)_{i=1}^{\infty} \preceq \omega$ for all $k \geq 1$.

Schmeling [50, Theorem B, E] proved that the set of all of $\beta \in (1, \infty)$ such that β-shift does *not* satisfy specification is residual and its symmetric difference with $(1, \infty)$ is a set of measure zero; see [1, 2] for background.

Proposition 9. *For each $\beta > 1$ the β-shift is universal.*

Let $\mathcal{L}$ be the set of all allowed (finite) words in the subshift X, so that $\mathcal{L}(X) = \mathcal{L} := \bigcup_n \mathcal{L}_n$, where for $n \geq 1$ we define

$$\mathcal{L}_n := \{x_1 \cdots x_n : x \in X\}.$$

A subshift generated by free concatenations of a *finite* set of words is often called a *renewal system*. Renewal systems are sofic [56].

Proof. Climenhaga and Thompson [12, Section 3.1] show that for β-shifts, by taking

$$\mathcal{G}_n := \{g_1 \cdots g_n \in \mathcal{L}_n : g_n = 0 \text{ and } \omega_n \neq 0\},$$

we have that

$$\limsup_{n \to \infty} \frac{1}{n} \log |\mathcal{G}_n| = \log(\beta) = h_{\text{top}}(T);$$

and furthermore, given any k words in $\mathcal{G}_n$ can be concatenated to produce an element of $\mathcal{G}_{kn}$. Thus the subshift X_n generated by $\mathcal{G}_n$ is a mixing sofic renewal system. Clearly, the sequence of subshifts X_n has topological entropy arbitrary close to $\log(\beta)$. Thus β-shifts are universal by Corollary 8. $\square$

3.2. **Toral automorphisms.** Lind [35, Section 3] classifies hyperbolic and quasi-hyperbolic toral automorphisms into three disjoint classes according to the spectral properties of the associated matrix. These classes correspond precisely to the various forms of specification that were defined in Section 1. Recall that associated matrix for a hyperbolic toral automorphism has no roots of unity, and hence all quasi-hyperbolic toral automorphisms in dimensions two or three are hyperbolic. Bowen [4] proved that hyperbolic toral automorphisms satisfy specification. The associated matrix for a **central spin** automorphism has some eigenvalues on the unit circle, and the Jordan blocks for these

eigenvalues have no off-diagonal 1's, and all other quasi-hyperbolic toral automorphisms are **central skew** automorphisms. Lind [35, Theorem (ii)] proved that central spin automorphisms (which only occur in dimensions 4 or higher) satisfy weak specification, but never satisfy specification. Lind [35, Theorem (iii)] also proved that central skew automorphisms (which only occur in dimensions 8 or higher) never satisfy weak specification, but nevertheless Marcus [39] proved that they still must satisfy almost weak specification. It is easy to give explicit examples of all the different types of quasi-hyperbolic toral automorphisms (see for example [36, Section 3]).

Proof of Theorem 2. By Lemma 12, we know that toral automorphisms are entropy expansive. Since we assume quasi-hyperbolicity, almost weak specification is satisfied [39], and furthermore an easy linear algebra argument shows the periodic points are exactly the points on the torus with rational coordinates; it is an elementary fact [21, Proposition 1.2.4] that countable sets have zero-dimension. Thus Theorem 2 follows from Theorem 7 and Lemma 6. □

We will use the following elementary factorization lemma to prove Theorem 1.

Lemma 10. *Any toral automorphism is the direct product of a quasi-hyperbolic toral automorphism and a toral automorphism of zero topological entropy.*

Proof. The lemma is a consequence of the following facts. The topological entropy of a toral automorphism is given by the sum of the logarithms of the moduli of the associated eigenvalues of modulus at least one (see for example [3, Corollary 16]). A toral automorphism with a characteristic polynomial f that can be expressed as a product $f = gh$ such that g and h are polynomials over $\mathbb{Z}$ with $\gcd(g, h) = 1$ is the direct product of toral automorphisms with characteristic polynomials g and h. A polynomial over $\mathbb{Z}$ with a root of unity is the product of a polynomial over $\mathbb{Z}$ with only roots of unity and a polynomial over $\mathbb{Z}$ without roots of unity. □

Proof of Theorem 1. Let S be an automorphism of the d-dimensional torus $\mathbb{T}^d$. By Lemma 10, let $S = S^q \times S^o$, where S^q is a quasi-hyperbolic toral automorphism of a d_1-dimensional torus $\mathbb{T}^{d_1}$ with the same topological entropy as S, and S^o is a zero entropy toral automorphism of a d_2-dimensional torus $\mathbb{T}^{d_2}$. Note that if T is an invertible ergodic-measure preserving transformation of a probability space (Ω, μ) and Ψ is an embedding of (Ω, μ, T) into $(\mathbb{T}^d, S^q)$, then the map $\Psi'(y, z) :=$

$(\Psi(y), 0)$ gives an embedding of (X, μ, T) into $(\mathbb{T}^{d_1} \times \mathbb{T}^{d_2}, S^q \times S^o)$, since 0 is a fixed point of any toral automorphism. Thus the universality of $(\mathbb{T}^d, S)$ follows from the universality of $(\mathbb{T}^{d_1}, S^q)$ and Theorem 2. $\square$

We give the following easy application of Theorem 1. Let T be a $C^{1+\delta}$ $(\delta > 0)$ diffeomorphism of a compact smooth manifold. Let

$$E(T) = \{h_\mu(T) : \mu \text{ is an ergodic invariant measure for } T\}.$$

Katok conjectures $E(T) \supset [0, h_{\text{top}}(T))$ and proved this to be true in the case where T is a diffeomorphism of a two-dimensional surface [55, 27]; the proof follows immediately from [27, Theorem 4]. Recently, Sun [55] showed that $E(T)$ is dense in $[0, h_{\text{top}}(T)]$ for a toral automorphism T; Theorem 1 extends the result of Sun and implies that Katok's conjecture is true for toral automorphisms.

3.3. **Automorphisms of a compact metric abelian group.** Toral automorphisms form an important class of examples of automorphisms of compact metric abelian groups. Note that an automorphism always preserves the Haar measure. Generalizing the earlier result of Marcus [39], Dateyama [13] proved that an automorphism of a compact metric abelian group satisfies almost weak specification if and only if it is ergodic with respect to the Haar measure.

Ergodicity in the setting of an automorphism S of a compact metric abelian group Y still has an algebraic characterization. Recall that the character group of Y is the set of all continuous group homomorphisms of Y into the complex unit circle with multiplication and denoted by $\hat{Y}$. Let $U_S : L^2 \to L^2$ be the Koopman representation given by $U_S(\chi) = \chi \circ S$, where L^2 is the set of all square integrable functions on Y with respect to Haar measure. Since S is also a group homomorphism, we have that $U_S : \hat{Y} \to \hat{Y}$. Halmos [24] proved that S is ergodic with respect to Haar measure if and only if U_S has no finite orbits other than the trivial character; that is for all $\chi \in \hat{Y}$ and all $n \in \mathbb{Z}^+$, we have $U_S^n(\chi) = \chi$ if and only if $\chi = 1$.

Theorem 11. *An automorphism of a compact finite-dimensional metric abelian group with finite topological entropy whose Koopman representation has no finite orbits on the character group, other than the trivial character, satisfies almost weak specification and is universal.*

The proof of Theorem 11 is similar to the proof of Theorem 2 making use of slightly more technical tools to verify the conditions of Theorem 7. The following lemma substitutes entropy expansiveness with asymptotic entropy expansiveness.

Lemma 12 (Misiurewicz [41] Example 7.1). *A self-homomorphism of a compact group with finite topological entropy is asymptotically entropy expansive.*

Finally, we also need to argue that the periodic points are countable. We do not have an elementary argument for this, and instead will refer to a recent paper of Miles [40] who gives a formula for the number of periodic points of each finite order of a ergodic finite entropy automorphism of a finite-dimensional compact abelian group; we will not need to use the full force of this formula.

Proof of Theorem 11. By Lemma 12, we have asymptotic entropy expansiveness. Since the Koopman representation has no non-trivial finite orbits, we have that the automorphism is ergodic with respect to Haar measure [24] and thus satisfies almost weak specification [13]. Since the automorphism is ergodic and has finite entropy, by [40, Lemma 4.3], the number periodic points of each finite order is finite thus the number of periodic points is countable and has zero-dimension [21, Proposition 1.2.4]. Theorem 11 now follows from Theorem 7 and Lemma 6. □

In his book [25], Halmos proposed an invariant of two-dimensional toral automorphisms. He noted that if $f\colon \mathbb{T}^2 \to \mathbb{T}$ is one of the coordinate functions then $p(T)f = 0$, where p is the characteristic polynomial of the automorphism, even though the function f is a non-constant function mapping into $\mathbb{T}$. Here, by $p(T)$, if $p(t) = \sum a_k t^k \in \mathbb{Z}[t]$, we mean $p(T)f = \sum_k a_k f \circ T^k$. Clearly the existence of a non-constant $\mathbb{T}$-valued f such that $p(T)f = 0$ almost everywhere is an isomorphism invariant. Halmos asked whether this invariant (with the polynomials p taken to be of the form $p(t) = t^2 - at + 1$) distinguishes two-dimensional toral automorphisms with distinct characteristic polynomials. A natural generalization is to set $\mathbb{P}(X, T, \mu) := \{p \in \mathbb{Z}[t]\colon p(t)$ is irreducible over $\mathbb{Z}$ and there exists a non-constant $\mathbb{T}$-valued f such that $p(T)f = 0$, μ-a.e.$\}$. Notice that for a reducible polynomial $q(t) \in \mathbb{Z}[t]$, there is a non-constant $\mathbb{T}$-valued f such that $q(T)f = 0$ μ-a.e. if and only if one of its irreducible factors belongs to $\mathbb{P}(X, T, \mu)$.

The following corollary establishes that $\mathbb{P}(X, T, \mu)$ is no stronger an invariant than the rational spectrum (in the sense of its ability to distinguish transformations). We would like to thank Benjy Weiss for pointing this corollary out to us, and for suggesting we include it in this paper.

Corollary 13 (Weiss). *Let (X, T, μ) and (X', T', μ') be ergodic invertible measure-preserving transformations of non-atomic probability spaces. If T and T' have the same rational spectrum, then $\mathbb{P}(X, T, \mu) = \mathbb{P}(X', T', \mu')$.*

Proof. Let NCI denote the collection of non-cyclotomic irreducible polynomials in $\mathbb{Z}[t]$. If (X, T, μ) is non-trivial and $p(t) \in NCI \setminus \{\pm 1, \pm t\}$, then let $\mathbb{T}^{\mathbb{Z}}$ be equipped with the shift map σ. Let $I_p = \{z \in \mathbb{T}^{\mathbb{Z}} : p(\sigma)z = 0\}$. This satisfies the conditions of Theorem 11, so that provided (X, T, μ) has entropy smaller than $h_{\text{top}}(I_p, \sigma)$, there is a non-constant equivariant map $\pi \colon (X, \mu) \to (I_p, \sigma)$. If $h_\mu(T) \geq h_{\text{top}}(I_p, \sigma)$, then we first take a two-element partition $\mathcal{P}$ of X such that $H(\mathcal{P}) < h_{\text{top}}(I_p, \sigma)$ and consider the factor of X induced by $\mathcal{P}$. This has strictly smaller entropy than $h_{\text{top}}(I_p, \sigma)$, which is then embedded as before. In either case, we end up with a non-constant equivariant map π from (X, T, μ) to (I_p, σ). Taking $f(z) = z_0$, we see $p(T)f \circ \pi = 0$ for μ-a.e. x, so that $p \in \mathbb{P}(X, T, \mu)$.

For a positive integer m, let S denote the transformation $S(t) = t + 1 \bmod m$ of $[m] = \{0, \ldots, m-1\}$ and c denote normalized counting measure on $[m]$. We denote the nth cyclotomic polynomial by Φ_n. If $\ell > 1$ is a factor of m, then one can check that $\Phi_\ell \in \mathbb{P}([m], S, c)$ (as witnessed by the function $g(x) = \cos(2\pi(x - \alpha)/\ell)$ for α irrational – notice that the α guarantees that g is not constant for $\ell > 1$). However, $\mathbb{P}([m], S, c)$ may contain Φ_n's for integers n that are not factors of m also. For example the function $g(0) = \frac{1}{3}$, $g(1) = \frac{2}{3}$ shows that $\Phi_6 \in \mathbb{P}([2], S, c)$.

For a non-trivial invertible ergodic measure-preserving transformation (X, T, μ), let $R(T) := \{m \colon T \text{ has an } m\text{-point factor}\}$. We claim the following, from which the statement of the corollary holds.

$$\mathbb{P}(X, T, \mu) = \big(NCI \setminus \{\pm 1, \pm t\}\big) \cup \bigcup_{m \in R(T)} \mathbb{P}([m], S, c).$$

$NCI \setminus \{\pm 1, \pm t\}$ is contained in both sides by the above, so that it suffices to show that the two sides contain the same cyclotomic polynomials. That $\mathbb{P}([m], S, c)$ is contained in $\mathbb{P}(X, T, \mu)$ for $m \in R(T)$ is clear by lifting a function defined on $[m]$ to X through the factor map. For the converse, suppose that $\Phi_n \in \mathbb{P}(X, T, \mu)$. Let $f \colon X \to \mathbb{T}$ be non-constant but be such that $\Phi_n(T)f = 0$. Since $\Phi_n(T)f = 0$ and $\Phi_n(x) | x^n - 1$, we see that $(T^n - 1)f = 0$, so that $f \circ T^n(x) = f(x)$ μ-a.e. Let m be the period of the sequence $(f(T^j x))_j$, so that $m \in R(T)$. Then (X, T, μ) factors onto $[m]$ and transferring the function f to $[m]$ shows that $\Phi_n \in \mathbb{P}([m], S, c)$ as required. $\qquad \square$

3.4. **Geodesic flows and suspension flows.** The small boundary property is a technical condition that is convenient for our method of proof, but is not necessary for universality. The time-one map of a geodesic flow on a compact surface of negative curvature is entropy expansive [5, Example 1.6 *] and time-map of a suspension flow can also be shown to be entropy expansive by [5, Example 1.6]. The time-one map of a geodesic flow of a compact surface of negative curvature satisfies specification [51, (Example) F] and the time-one map of a topologically weak-mixing suspension flow satisfies (weak) specification [45, Proposition 5]. However, clearly the time-one maps of these flows may have uncountably many periodic *points*, and moreover, it is not difficult to construct examples of topologically weak-mixing suspensions flows that do not have the small boundary property. In spite of this, we proved universality for these flows [45, Theorems 1 and 2].

3.5. **Non-examples.** We give an example to show that it is too much to ask that the strict entropy difference in the definition of universality can be relaxed. The same example also shows that universality can not always be satisfied; this fact also easily follows from the Krieger-Jewett theorem [32].

Let S be the full-shift on $\Omega := \{0,1\}^{\mathbb{Z}}$ endowed with the Bernoulli measure ζ that is the unique measure of maximal entropy–$\log(2)$. Let S' be an irrational rotation of the circle $\mathbb{T}^1$. The invertible measure-preserving transformation $S \times S'$ is ergodic and has entropy $\log(2)$. The invertible measure-preserving system $(\Omega \times \mathbb{T}^1, S \times S')$ can not be embedded within (Ω, S), since the embedding would yield a measure-theoretic isomorphism of $S \times S'$ and S; this implies that S has a non-trivial zero entropy factor which contradicts that fact that S is a K-automorphism [53, Part 1, Paper 2]. Similarly, if the topological dynamical system $(\Omega \times \mathbb{T}^1, S \times S')$ were universal, then every measure-preserving automorphism with entropy strictly less than $\log(2)$ would have a non-trivial zero entropy factor.

4. Soft methods

4.1. **Basic idea.** Roughly, the idea of Burton and Rothstein [11], is that one defines an notion of ϵ-approximate embedding (see Proposition 18) with the property that a point lying in the intersection of the ϵ-approximate embeddings is a true embedding. One then introduces a Polish space (a separable completely metrizable topological space) consisting of potential injections (these are in fact joinings, see Lemma 17) and shows that the ϵ-approximate embeddings form a dense open subset of the set of potential embeddings. Baire's theorem gives the

desired result. The essential (and surprising) feature of this idea is that in a topological sense, almost any candidate works. Burton, Keane and Serafin reproved the Krieger generator theorem [30], the Sinai factor theorem [52, 53], and the Ornstein isomorphism theorem [42].

Keane and Smorodinsky [28, 29] gave an explicit proof of the Ornstein isomorphism theorem. Similarly, there is an alternative approach to producing embeddings that we used in our earlier paper [45]. In that approach, one carefully produces a sequence of approximate embeddings, takes a pointwise limit, and proves that this limit has the desired properties. (However, our construction does not yield a finitary map, unlike those of Keane and Smorodinsky.)

4.2. **Joinings and Baire category.** In this subsection, we make precise the ideas outlined above. We will make use of the following explicit metric giving rise to the weak*-topology on measures. Let Z be a compact metric space. We let $\mathrm{Lip}_1(Z)$ denote the space of all real-valued Lipschitz continuous functions on Z taking values in $[0, 1]$ with Lipschitz constant no greater than 1. Let $\mathcal{M}(Z)$ denote the collection of all Borel probability measures on Z. Write $\mu(f) := \int f d\mu$ for $\mu \in \mathcal{M}(Z)$ and an integrable function $f : Z \to \mathbb{R}$.

Lemma 14 (Distance giving weak*-topology on measures)**.** *Let Z be a compact metric space. Define a metric d^* on $\mathcal{M}(Z)$ as follows:*

$$\mathrm{d}^*(\mu, \nu) = \sup_{f \in \mathrm{Lip}_1} |\mu(f) - \nu(f)|.$$

The metric d^* gives rise to the weak topology on $\mathcal{M}(Z)$.*

The proof is a simple adaptation of the proof in Dudley's book ([20], Theorem 11.3.3) to take account of the fact that functions in Lip_1 are required to take values in $[0, 1]$.

In general, a measure-preserving transformation is defined on a measure space. In order to use the Baire category machinery, we will need the space to be embedded to be a metric space. The following lemma allows us to assume that the transformation to be embedded lives on a metric space.

Lemma 15. *Let S be a self-homeomorphism of a compact metric space Y. Then (Y, S) is (fully) universal if and only if there exists an embedding of every non-trivial ergodic subshift (X, T, μ) with $h_\mu(T) < h_{\mathrm{top}}(S)$ into (Y, S) (such that the push-forward of μ is fully supported on Y).*

Proof. The only if part is clear from the definition of universality. For the converse by the Krieger generator theorem, for an arbitrary invertible measure-preserving transformation T_0 of the probability space

(X_0, μ_0), there is a measure-theoretic isomorphism between (X_0, T_0, μ_0) and a subshift (X, T, μ) (which, of course, preserves the entropy). By the assumption of the lemma, this subshift may be embedded into (Y, S). Composing the isomorphism and the embedding gives the result. □

Thus by Lemma 15, we may always assume that the space to be embedded is a subshift on a finite number of symbols.

Let $(X, \mathcal{F}, \mu, T)$ be a non-trivial ergodic subshift with invariant measure μ and Borel σ-algebra $\mathcal{F}$. Let S be a self-homeomorphism of a compact metric space Y with Borel σ-algebra $\mathcal{B}$. By a **μ-joining**, we mean a $(T \times S)$-invariant measure on the product metric space $X \times Y$ (with the product σ-algebra $\mathcal{F} \otimes \mathcal{B}$)) whose X-marginal is μ. (Note that unlike the standard definition of a joining, we do not make any requirement on the Y-marginal; for more background on joinings see [49, 22, 48].) We let $J_\mu(T, S)$ denote the space of μ-joinings. It is well known that $J_\mu(T, S)$ is a (non-empty) compact metric (hence complete) space with the weak* topology. We let π_1 and π_2 be the co-ordinate projections from $X \times Y$ to X and Y, respectively, and denote the push-forward maps by π_1^* and π_2^*, so that for $\xi \in J_\mu(T, S)$, we have $\pi_1^*(\xi) = \mu$. Before we prove topological properties regarding subsets of the space of joinings, we first state an elementary but useful fact which motivates how joinings are related to embeddings.

Let $(Z, \mathcal{G}, \xi, U)$ be a measure-preserving system. Let $\mathcal{F}, \mathcal{B} \subset \mathcal{G}$ be sub-σ-algebras. We write $\mathcal{F} \subset \mathcal{B} \mod \xi$ if for each $F \in \mathcal{F}$, there is a $B \in \mathcal{B}$ such that the ξ-measure of the symmetric difference is zero. Similarly, for each $\varepsilon > 0$, we write $\mathcal{F} \overset{\epsilon}{\subset} \mathcal{B} \mod \xi$, if for each $F \in \mathcal{F}$, there is a $B \in \mathcal{B}$ such that the ξ-measure of the symmetric difference is strictly less than ϵ. For a partition $\mathcal{P}$, we denote by $\sigma(\mathcal{P})$ the finite σ-algebra that it generates. Note that $\mathcal{P}$ is a generating partition for $Z \pmod \xi$ if and only if $\mathcal{G} \subset \bigvee_{i \in \mathbb{Z}} T^{-i} \mathcal{P} \mod \xi$. Also let $\mathcal{T}_Z = \{\emptyset, Z\}$ denote the trivial σ-algebra on Z.

Proposition 16. *Let $(X, \mathcal{F}, \mu, T)$ be a non-trivial ergodic subshift with invariant measure μ and Borel σ-algebra $\mathcal{F}$. Let S be a self-homeomorphism of a compact metric space Y with Borel σ-algebra $\mathcal{B}$. Let ξ be a μ-joining satisfying the following conditions*

(i) $\mathcal{T}_X \otimes \mathcal{B} \subset \mathcal{F} \otimes \mathcal{T}_Y \mod \xi$;
(ii) $\mathcal{F} \otimes \mathcal{T}_Y \subset \mathcal{T}_X \otimes \mathcal{B} \mod \xi$.

Then there exists an embedding $\Psi : X \to Y$ of (X, T, μ) into (Y, S) such that

$$\xi(\mathcal{F} \times B) = \mu(\mathcal{F} \cap \Psi^{-1}(B)) \text{ for all } (F, B) \in \mathcal{F} \times \mathcal{B}. \qquad (1)$$

Similarly, if Ψ is an embedding of (X, T, μ) into (Y, S), then ξ defined by (1) satisfies the first two conditions.

Proof. Proposition 16 follows easily from [49, Theorem 2.8]. $\qquad\square$

Thus if ξ satisfies the two containment conditions in Proposition 16, we also say that ξ is an ***embedding***. The Burton–Rothstein argument, that we follow, defines approximate embeddings and works by showing that there are a large collection of approximate embeddings at each scale.

Lemma 17 (The Baire Space). *Let (X, μ, T) be a non-trivial ergodic subshift with invariant measure μ, and let S be a self-homeomorphism of a compact metric space Y that is asymptotically entropy expansive. Let $h_\mu(T) < h_{\mathrm{top}}(S)$. The space defined by*

$$\mathcal{M}_0 := \{\xi \in J_\mu(T, S) \colon \xi \text{ is ergodic and } h_{\pi_2^*(\xi)}(S) \geq h_\mu(T)\}$$

is a Baire space.

If U is an invertible measure-preserving transformation on a space Z, and $f : Z \to \mathbb{R}$ is any real-valued function, we let $\mathbb{A}_m^n(f)$ denote the Césaro average given by

$$\mathbb{A}_m^n(f)(x) := \frac{1}{n-m} \sum_{k=m}^{n-1} f(U^k x).$$

Let μ be a U-invariant measure on Z. By the Birkhoff ergodic theorem, μ is ergodic if for any $f \in L^1$, $\mathbb{A}_0^n(f)$ converges in measure to a constant. In the case that Z is a compact metric space, then we say that $x \in Z$ is ***generic*** (for μ) if for every continuous function $f : Z \to \mathbb{R}$, we have that $\mathbb{A}_0^n(f)(x) \to \mu(f)$ as $n \to \infty$. An elementary argument gives that if μ is ergodic if and only if μ-a.e. $x \in Z$ is generic for μ.

Proof of Lemma 17. Since S is assumed to be asymptotically entropy expansive, by Lemma 5 the entropy functional is upper-semicontinuous. Thus the subset of joinings of $J_\mu(T, S)$ satisfying the entropy inequality is closed, and hence complete. We next show that the subset of these joinings that are ergodic forms a G_δ subset of this set.

Let $D \subset C(X \times Y)$ be a countable dense collection of continuous functions. The collection of ξ in $\mathcal{M}_0$ satisfying the condition: for all $f \in D$ and for all j, there exists an n such that $\xi(\mathbb{A}_0^n(f)^2) - \xi(f)^2 < 1/j$ is a G_δ set; clearly, the condition is that the limit inferior of the variances of the n-step Césaro averages is 0. This condition is satisfied if and only if $\mathbb{A}_0^n(f)$ converges in measure (with respect to ξ) to a constant for each $f \in D$, which holds if and only if ξ is ergodic.

A G_δ subset of a complete metric space is a Polish space by a theorem of Alexandrov [54, Theorem 2.2.1]; and the Baire category theorem tells us that every Polish space is a Baire space [54, Theorem 2.5.5]. □

Proposition 18. *Let (X, μ, T) be a non-trivial ergodic subshift with invariant measure μ, with its natural generating partition $\mathcal{P}$. Let S be a self-homeomorphism with almost weak specification on a compact metric space Y that satisfies the small boundary condition witnessed by a sequence of refining partitions $(\mathcal{Q}_\ell)$, where $\mathrm{diam}(\mathcal{Q}_\ell) < 1/\ell$. Let $h_\mu(T) < h_{\mathrm{top}}(S)$. For each $\ell, n \geq 1$, let $E_\mu^{n,\ell}$ be the set of elements ξ of $\mathcal{M}_0$ (defined in Lemma 17) satisfying the following two conditions:*

(i) $\mathcal{T}_X \otimes \sigma(\mathcal{Q}_\ell) \overset{1/n}{\subset} \left(\bigvee_{i \in \mathbb{Z}} T^{-i} \mathcal{P} \right) \otimes \mathcal{T}_Y \mod \xi$;

(ii) $\sigma(\mathcal{P}) \otimes \mathcal{T}_Y \overset{1/n}{\subset} \mathcal{T}_X \otimes \bigvee_{i \in \mathbb{Z}} T^{-i} \mathcal{Q}_\ell \mod \xi$.

Let $E_\mu^n = \bigcup_{\ell \geq n} E_\mu^{n,\ell}$.

(I) The set $E_\mu^{n,\ell}$ is a relatively open subset of $\mathcal{M}_0$.
(II) If $\xi \in \bigcap_{n \geq 1} E_\mu^n$, then ξ is an embedding.
(III) The set E_μ^n is a dense subset of $\mathcal{M}_0$.

Let us make a few remarks. First, by (II), we may call E_μ^n the collection of **$1/n$-approximate embeddings** of (X, T, μ) into (Y, S). Second, the small boundary condition is used to prove (I), and almost weak specification is used to prove the density condition (III). Finally, most of the hard work will done in verifying (III).

Before, we prove the easier parts of Proposition 18, let us put together a proof of Theorem 7. We need one more lemma in order to obtain fully supported measures on Y.

Lemma 19. *Let (X, μ, T) be a non-trivial ergodic subshift with invariant measure μ, and let S be a self-homeomorphism of a compact metric space Y with almost weak specification. Let $h_\mu(T) < h_{\mathrm{top}}(S)$. Let $\mathcal{M}_0'$ be the set of all $\xi \in \mathcal{M}_0$ such that $\pi_2^*(\xi)$ is fully supported on Y. Then $\mathcal{M}_0'$ is an intersection of countably many relatively open dense subsets of $\mathcal{M}_0$.*

We give the proof of Lemma 19 in Section 5.

Proof of Theorem 7. By Lemma 15, we may assume that (X, T, μ) is a non-trivial ergodic subshift. By Lemma 17, $\mathcal{M}_0$ is a Baire space, and by Proposition 18 (I) and (III), we have that $\mathcal{E} := \bigcap_{n \geq 1} E_\mu^n$ is a (non-empty) dense subset of $\mathcal{M}_0$; furthermore, by Lemma 19, $\mathcal{M}_0' \cap \mathcal{E}$ is also a (non-empty) dense subset of $\mathcal{M}_0$. It follows from Proposition 18 (II) that the joinings of $\mathcal{E}$ are also embeddings. □

Proof of Proposition 18 (I). If $\xi \in E_\mu^n$, then each element, $X \times A$ with $A \in \sigma(\mathcal{Q}_\ell)$ agrees with an element $B \times Y$ with $B \in \bigvee_{i \in \mathbb{Z}} T^{-i}\mathcal{P}$ up to a symmetric difference of measure strictly less than $1/n$; moreover, since $\bigvee_{i \in \mathbb{Z}} T^{-i}\mathcal{P}$ is the limit of $\bigvee_{|i| \leq m} T^{-i}\mathcal{P}$, there exists $B \in \bigvee_{i=-m}^{m} T^{-i}\mathcal{P}$ for some finite m with

$$\xi((X \times A) \triangle (B \times Y)) < 1/n. \tag{2}$$

Let m_0 be the maximum of these m's over the finite set of A's in $\sigma(\mathcal{Q}_\ell)$.

For all $\xi' \in \mathcal{M}_0$ we have $\xi'(X \times \partial A) = 0$, since $\pi_2^*(\xi')$ is an S-invariant measure and by the small boundary property, we have $\pi_2^*(\xi')(\partial A) = 0$. Since the map $\nu \mapsto \nu(A)$ is continuous at all points where $\nu(\partial A) = 0$, inequality (2) persists for all sufficiently small perturbations of ξ. The openness of the second condition is proved similarly (using the fact that each element of $\bigvee_{|i| \leq k} S^{-i}\mathcal{Q}_\ell$ has boundary of measure 0 for each k and ℓ and every invariant measure). $\qquad\square$

Proof of Proposition 18 (II). The proof follows from Proposition 16. $\qquad\square$

The remainder of this paper is dedicated to proving Proposition 18 (III) and Lemma 19. We choose a $\xi \in \mathcal{M}_0$ and a weak*-neighbourhood of ξ. We show the denseness in several stages. Firstly it will be convenient to assume that $h_{\pi_2^*(\xi)}(S)$ strictly exceeds $h_\mu(T)$ (it is already at least $h_\mu(T)$ by assumption). In the next section, we show the strict entropy gap may be assumed without loss of generality. We show that we can make an arbitrarily small perturbation to ξ (in the weak* topology) so that the entropy condition is satisfied. (Similar techniques will also be used to prove Lemma 19.) Secondly, using the fact that $h_{\pi_2^*(\xi)}(S) > h_\mu(T)$, we build a mapping from blocks of X into separated orbit segments in Y. In a third stage, we put this together with some marker blocks in Y, verify that the weak*-closeness is satisfied, the entropy condition still holds and show that the almost embedding property is satisfied.

5. UPPING THE ENTROPY

Lemma 20 (Perturbation to increase marginal entropy). *Let (X, μ, T) be a non-trivial ergodic subshift with invariant measure μ. Let S be a self-homeomorphism with almost weak specification of a compact metric space Y. Let $\xi \in \mathcal{M}_0$ have the property that $h_\mu(T) < h_{\text{top}}(S)$ and let V be a weak*-neighbourhood of ξ in $\mathcal{M}_0$. Then there exists a $\xi' \in V$ such that $h_{\pi_2^*(\xi')}(S) > h_\mu(T)$.*

To prove Lemma 20, we alter the measure ξ by the following procedure. Given an ergodic measure λ on Y with $h_\lambda(S) > h_\mu(T)$, a ξ-generic point (x, y^1), and a λ-generic point y^2 that is chosen independently of (x, y^1), we produce another point (x, y^3), where y^3 is obtained from y^1 by splicing segments of y^1 with segments of y^2. This is made possible by almost weak specification of S, and a careful choice of parameters which tune a perturbation of ξ so that we have an entropy increase, but remain within a prescribed weak* neighborhood. The proof may appear somewhat technical because an arbitrary generating partition for Y may not have any relation to the specification properties of Y, and we also need to construct the perturbation so that we can verify the required properties.

We define the d-bar distance between two processes in the following way. By a **joining** of two probability measure spaces each endowed with a transformation, we mean a probability measure on the product space that is invariant with respect to the product transformation and has coordinate projections equal to the original measures. Let (X, T) be the full-shift on a finite number of symbols with invariant measures μ_1 and μ_2. Let

$$\bar{\mathrm{d}}(\mu_1, \mu_2) := \inf_{\xi \in J(\mu_1, \mu_2)} \int \mathbf{1}[x_0 \neq x_0'] d\xi(x, x'), \qquad (3)$$

where $J(\mu_1, \mu_2)$ denotes the set joinings of (X, μ_1, T) and (X, μ_2, T).

More generally, let T be a measure-preserving automorphism (X, μ) with a measurable partition $\mathcal{P} = \{B_1, \ldots, B_n\}$. For each $x \in X$, let $\mathcal{P}(x) \in \mathcal{P}$ denotes the part that contains x. The $\mathcal{P}$-**name** is the function $\bar{\mathcal{P}} : X \to \{1, \ldots, n\}^{\mathbb{Z}}$ given by $\bar{\mathcal{P}}(x)_i := \mathcal{P}(T^{-i}x)$. Thus $\mu_\mathcal{P} := \mu \circ \bar{\mathcal{P}}^{-1}$ is a shift-invariant measure on the full-shift of n symbols. Thus we may define the d-bar distance between two invertible measure-preserving transformations (X, μ, T) and (Z, ν, U) endowed with finite measurable partitions $\mathcal{P}$ and $\mathcal{Q}$ with n and m elements, respectively. We consider $\mu_\mathcal{P}$ and $\nu_\mathcal{Q}$ to be shift-invariant measures on the full-shift of $k = \max(m, n)$ symbols. Set

$$\bar{\mathrm{d}}(\mu, \mathcal{P}; \nu, \mathcal{Q}) := \bar{\mathrm{d}}(\mu_\mathcal{P}, \nu_\mathcal{Q}). \qquad (4)$$

We define $\mathcal{H} : (0, 1) \to \mathbb{R}$ via

$$\mathcal{H}(\eta) := -\eta \log \eta - (1 - \eta) \log(1 - \eta). \qquad (5)$$

Lemma 21. *Let (X, μ, T) and (Z, ν, U) be two invertible ergodic measure preserving transformations endowed with measurable partitions $\mathcal{P}$ and $\mathcal{Q}$. Let $\eta > 0$. If $\bar{\mathrm{d}}(\mu, \mathcal{P}; \nu, \mathcal{Q}) < \eta$ and $|\mathcal{P}| \geq |\mathcal{Q}|$, then*

$$|h_\mu(T, \mathcal{P}) - h_\nu(U, \mathcal{Q})| \leq \mathcal{H}(\eta) + \eta \log |\mathcal{P}|.$$

Proof. See Rudolph's book [48, Theorem 7.9]   □

Proof of Lemma 20. If $h_{\pi_2^*(\xi)}(S) > h_\mu(T)$, then we are done, so we assume that $h_{\pi_2^*(\xi)}(S) = h_\mu(T)$ in what follows.

Choose an ergodic measure λ on Y satisfying $h_\lambda(S) > h_{\pi_2^*(\xi)}(S) = h_\mu(T)$. Choose $\epsilon \in (0,1)$ so that V contains the ball $B(\xi, 3\epsilon) \cap \mathcal{M}_0$. Let $\delta = \frac{1}{4}(h_\lambda(S) - h_{\pi_2^*(\xi)}(S))$.

By a variation of a standard argument (see for example [5, Lemma 3.2]), we may choose a partition $\mathcal{P} = \{B_1, \ldots, B_n\}$ of Y such that $\pi_2^*(\xi)(\partial\mathcal{P}) = \lambda(\partial\mathcal{P}) = 0$ and $h_\lambda(S, \mathcal{P}) > h_\lambda(S) - \epsilon\delta$ and $h_{\pi_2^*(\xi)}(S, \mathcal{P}) > h_{\pi_2^*(\xi)}(S) - \epsilon\delta$. Let $\mathcal{H}$ be as in (5) and choose $\gamma < \epsilon$ so that $\mathcal{H}(2\gamma) + 2\gamma \log(|\mathcal{P}| + 1) < \epsilon\delta$ and $2\gamma h_{\text{top}}(S) < \epsilon\delta$. Let $r < \epsilon$ be chosen so that $\pi_2^*(\xi)(\partial_r \mathcal{P}) < \gamma$ and $\lambda(\partial_r \mathcal{P}) < \gamma$.

We now build a mixing process taking values in $R = \{0, 1, 2\}^{\mathbb{Z}}$. Choose N sufficiently large so that the gap function satisfies

$$L_r(N) < k_0 := \lfloor (\gamma/2)N \rfloor;$$

in addition, we require that

$$\frac{k_1}{k_0 + k_1 + k_2} > 1 - \varepsilon - \gamma \text{ and } \frac{k_2}{k_0 + k_1 + k_2} > \varepsilon - \gamma,$$

where $k_1 = \lfloor (1 - \epsilon - \gamma/2)N \rfloor$ and $k_2 = \lfloor (\epsilon - \gamma/2)N \rfloor$. The process consists of concatenations of the blocks $1^{k_1} 0^{k_0} 2^{k_2} 0^{k_0}$ and $1^{k_1+1} 0^{k_0} 2^{k_2} 0^{k_0}$, where the blocks are placed independently with equal probabilities of using the longer and the shorter block. (We need to use values k_1 and $k_1 + 1$ in order to make ζ mixing.) Write ζ for the measure on R.

It will be convenient to have notation that will allow us to distinguish various copies of Y in certain product spaces. Let $Y^{(i)} = Y$ for $i = 1, 2, 3$. Let σ be the shift on R and consider the transformation $\tau = \sigma \times (T \times S) \times S$ acting on $\Omega = R \times (X \times Y^{(1)}) \times Y^{(2)}$, preserving the (not necessarily ergodic) measure $\zeta \times \xi \times \lambda$. Define a new partition $\mathcal{Q}$ on the product by $D_i = \{(z, y^{(1)}, y^{(2)}) : z_0 \in \{1, 2\} \text{ and } y_0^{(z_0)} \in B_i\}$ and $D_0 = \{(z, y^{(1)}, y^{(2)}) : z_0 = 0\}$. Note that $|\mathcal{P}| + 1 = |\mathcal{Q}|$. The idea is knowing which element of $\mathcal{Q}$ a point $(z, y^{(1)}, y^{(2)})$ lies in tells us which element of $\mathcal{P}$ the z_0'th copy of y lives in if z_0 is 1 or 2; or else tells us that $z_0 = 0$. Let $\mathcal{R}$ denote the σ-algebra of Ω giving information about the R coordinates.

Using the fact that $y^{(1)}$ and $y^{(2)}$ are chosen independently from the measures $\pi_2^*(\xi)$ and λ implies that

$$\lim_{M\to\infty} \frac{1}{M} H_{\zeta\times\xi\times\lambda}\left(\bigvee_{i=0}^{M-1} \tau^{-i}\mathcal{Q}\right) \geq \lim_{M\to\infty} \frac{1}{M} H_{\zeta\times\xi\times\lambda}\left(\bigvee_{i=0}^{M-1} \tau^{-i}\mathcal{Q}\bigg|\mathcal{R}\right)$$
$$\geq (1-\epsilon-\gamma)h_{\pi_2^*(\xi)}(S,\mathcal{P}) + (\epsilon-\gamma)h_\lambda(S,\mathcal{P}).$$

By affineness of entropy and the ergodic decomposition theorem, we can take an ergodic component, ι, of $\zeta\times\xi\times\lambda$ such that

$$h_\iota(\tau,\mathcal{Q}) \geq (1-\epsilon-\gamma)h_{\pi_2^*(\xi)}(S,\mathcal{P}) + (\epsilon-\gamma)h_\lambda(S,\mathcal{P}). \tag{6}$$

Since ζ is mixing, both $\zeta\times\pi_2^*(\xi)$ and $\zeta\times\lambda$ are ergodic. Thus the projections of ι on $R\times Y^{(1)}$ and $R\times Y^{(2)}$ are (still) $\zeta\times\pi_2^*(\xi)$ and $\zeta\times\lambda$, respectively.

By induction and the definition of almost weak specification, for an ι-generic point $\boldsymbol{\omega} = (\mathbf{z}, (\mathbf{x}, \mathbf{y}^{(1)}), \mathbf{y}^{(2)}) \in \Omega$, we can build a $\tilde{\mathbf{y}}$ such that

$$\mathrm{d}(S^n\tilde{\mathbf{y}}, S^n\mathbf{y}^{(\mathbf{z}_n)}) < r \text{ for all } n \text{ such that } \mathbf{z}_n \in \{1,2\}. \tag{7}$$

Thus by standard arguments (see the proof of the Krylov-Bogolyubov theorem [22, Page 97]) and (7), there exists a $(\tau\times S)$-invariant measure $\tilde{\iota}$ on $\Omega' := \Omega\times Y^{(3)}$ such that for $\tilde{\iota}$-a.e. point $(z, (x, y^{(1)}), y^{(2)}, y^{(3)})$, we have

$$\mathrm{d}(S^n y^{(3)}, S^n y^{(z_n)}) < r \text{ for all } n \text{ such that } z_n \in \{1,2\}. \tag{8}$$

Furthermore, by taking an ergodic component ι' of $\tilde{\iota}$ we may obtain an ergodic measure such that (8) holds for ι'-a.e. point. Since ι is ergodic and the original point $\boldsymbol{\omega}$ was chosen to be generic, the projections of ι' and $\tilde{\iota}$ on Ω are ι.

We now define $\tilde{\pi}(z, (x, y^{(1)}), y^{(2)}, y^{(3)}) = (x, y^{(3)})$. Note that by construction, $y^{(3)}$ is within r of $y^{(1)}$ whenever $z_0 = 1$. Let $\xi' = \tilde{\pi}^*(\iota')$. Clearly, ξ' is ergodic.

We now give estimates on the weak* distance between ξ and ξ' and on the entropy $h_{\pi_2^*(\xi')}(S)$. Let $f \in \mathrm{Lip}_1(X\times Y)$ (recalling that f takes values in $[0,1]$ by definition). Extend f to functions f^1, f^3 on Ω' by setting $f^1(z, (x, y^{(1)}), y^{(2)}, y^{(3)}) = f(x, y^{(1)})$ and $f^3(z, (x, y^{(1)}), y^{(2)}, y^{(3)}) = f(x, y^{(3)})$. Also let

$$D := \left\{(z, (x, y^{(1)}), y^{(2)}, y^{(3)}) \in \Omega' : z_0 = 1\right\}.$$

We then have

$$|\xi(f) - \xi'(f)| = |\iota'(f^1) - \iota'(f^3)|$$
$$\leq \iota'(\mathbf{1}_D|f^1 - f^3|) + \iota'(\mathbf{1}_{D^c}|f^1 - f^3|).$$

Thus by (8), we have

$$\mathrm{d}^*(\xi, \xi') \le r + (\epsilon + \gamma) < 3\varepsilon. \tag{9}$$

As for entropy, we estimate it by using (6) and the $\bar{\mathrm{d}}$ distance between $\mathcal{P}$-name of $y^{(3)}$ and the $\mathcal{Q}$-name of $(z, (x, y^{(1)}), y^{(2)}, y^{(3)})$.

For $\omega' \in \Omega'$, if $z_0 = 1$ and $y^{(1)} \notin \partial_r \mathcal{P}$, then $y^{(1)}$ and $y^{(3)}$ lie in the same element of $\mathcal{P}$. Similarly if $z_0 = 2$ and $y^{(2)} \notin \partial_r \mathcal{P}$, then $y^{(2)}$ and $y^{(3)}$ lie in the same element of $\mathcal{P}$. In particular, if $z_0 = 1$ and $y^{(1)} \notin \partial_r \mathcal{P}$ or $z_0 = 2$ and $y^{(2)} \notin \partial_r \mathcal{P}$, then $y^{(3)}$ lies in B_j if and only if $(z, (x, y^{(1)}), y^{(2)}, y^{(3)})$ lies in D_j.

On the other hand, since the projection of ι' on Ω is ι and the projections of ι on $R \times Y^{(1)}$ and $R \times Y^{(2)}$ are $\zeta \times \pi_2^*(\xi)$ and $\zeta \times \lambda$, we have

$$\iota'\{\omega' \in \Omega' : y^{(1)} \in \partial_r \mathcal{P} \text{ and } z_0 = 1\} = \pi_2^*(\xi)(\partial_r \mathcal{P}) \cdot \zeta\{z : z_0 = 1\}; \text{ and}$$

$$\iota'\{\omega' \in \Omega' : y^{(2)} \in \partial_r \mathcal{P} \text{ and } z_0 = 2\} = \lambda(\partial_r \mathcal{P}) \cdot \zeta\{z : z_0 = 2\}.$$

Note that $\pi_2^*(\xi')$ is the projection of ι' on $Y^{(3)}$. Let $\mathcal{P}'$ be the partition of Ω' whose parts are given by

$$\{\omega' \in \Omega' : y^{(3)} \in B_i\} \text{ for } 1 \le i \le n.$$

Let $\mathcal{Q}'$ be the partition of Ω' whose parts are given by

$$\{\omega' \in \Omega' : (z, (x, y^{(1)}), y^{(2)}) \in D_i\} \text{ for } 0 \le i \le n.$$

It follows that the d-bar distance $\bar{\mathrm{d}}(\pi_2^*(\xi'), \mathcal{P}'; \iota', \mathcal{Q}')$ is bounded above by

$$\pi_2^*(\xi)(\partial_r \mathcal{P}) \cdot \zeta\{z : z_0 = 1\} + \lambda(\partial_r \mathcal{P}) \cdot \zeta\{z : z_0 = 2\} + \zeta\{z : z_0 = 0\} < 2\gamma.$$

By Lemma 21, we have

$$|h_{\pi_2^*(\xi')}(S, \mathcal{P}) - h_{\iota'}(\tau \times S, \mathcal{Q})| \le \mathcal{H}(2\gamma) + 2\gamma \log(|\mathcal{P}| + 1) < \epsilon\delta.$$

Thus by (6) and the fact that $h_\iota(\tau, \mathcal{Q}) = h_{\iota'}(\tau \times S, \mathcal{Q}')$, we have

$$\begin{aligned}
&h_{\pi_2^*(\xi')}(S) \\
&\ge h_{\pi_2^*(\xi')}(S, \mathcal{P}) \\
&\ge (1 - \epsilon - \gamma)h_{\pi_2^*(\xi)}(S, \mathcal{P}) + (\epsilon - \gamma)h_\lambda(S, \mathcal{P}) - \epsilon\delta \\
&\ge (1 - \epsilon)h_{\pi_2^*(\xi)}(S) - \epsilon\delta + \epsilon(h_{\pi_2^*(\xi)}(S) + 4\delta) - \epsilon\delta - 2\gamma h_{\mathrm{top}}(S) - \epsilon\delta \\
&> h_{\pi_2^*(\xi)}(S). \qquad \qquad \qquad \square
\end{aligned}$$

The proof of Lemma 19 uses similar ideas, but is made easier by the fact that we can use Lemma 20.

Proof of Lemma 19. Let $(V_n)_{n\in\mathbb{N}}$ be a countable collection of open sets forming a neighbourhood basis for Y. Let $C_n = \{\xi \in \mathcal{M}_0 \colon \pi_2^*(\xi)(V_n) > 0\}$. By standard properties of the weak*-distance, C_n is a relatively open subset of $\mathcal{M}_0$ in the weak* topology. It suffices to show that C_n is dense in $\mathcal{M}_0$ for each n.

Fix $n \in \mathbb{N}$ and $\epsilon > 0$. Let $\xi_0 \in \mathcal{M}_0$. By Lemma 20, there exists $\xi_1 \in \mathcal{M}_0$ such that $h_{\pi_2^*(\xi_1)}(S) > h_\mu(T)$ and $\mathrm{d}^*(\xi_0, \xi_1) < \epsilon/2$. Let $\mathcal{Q}$ be a partition such that $\pi_2^*(\xi_1)(\partial Q) = 0$ and $h_{\pi_2^*(\xi_1)}(S, \mathcal{Q}) > h_\mu(T)$. Let $\delta < \epsilon/4$ be such that the following inequality holds:

$$\mathcal{H}(2\delta) + 2\delta \log |\mathcal{Q}| < h_{\pi_2^*(\xi_1)}(S, \mathcal{Q}) - h_\mu(T),$$

where $\mathcal{H}$ is defined in (5).

Let $r < \epsilon/4$ be chosen such that $\pi_2^*(\xi_1)(\partial_r \mathcal{Q}) < \delta$ and $B(\mathbf{y}, r) \subset V_n$ for some fixed $\mathbf{y} \in Y$. We then choose N such that $L_r(N) < M := \lfloor \delta N/2 \rfloor$. Let ζ be a mixing measure on $Z = \{0, 1, 2\}^{\mathbb{Z}}$ supported on concatenations of the blocks $1^{N-2M} 0^M 2 0^M$ and $1^{N+1-2M} 0^M 2 0^M$. As in the proof of Lemma 20, we start with the (ergodic) product measure $\xi_1 \times \zeta$ on $(X \times Y) \times Z$ and obtain an ergodic measure ι on $(X \times Y) \times Z \times Y$ whose marginal on the first pair of coordinates is ξ_1 and which satisfies $\mathrm{d}(y', y) < r$ for ι-almost every $((x, y), z, y')$ such that $z_0 = 1$ and $\mathrm{d}(y', y_0) < r$ for ι-almost every $((x, y), z, y')$ such that $z_0 = 2$. Projecting ι onto the initial X coordinate and the final Y coordinate gives a measure ξ' that can be checked using the same arguments as in Lemma 20 to satisfy $\mathrm{d}^*(\xi_1, \xi') < \epsilon/2$ and $\bar{\mathrm{d}}(\pi_2^*(\xi_1), \mathcal{Q}; \pi_2^*(\xi'), \mathcal{Q}) < 2\delta$. This is sufficient to ensure that $h_{\pi_2^*(\xi')}(S, \mathcal{Q}) > h_\mu(T)$, so that $\xi' \in \mathcal{M}_0$. Finally, $\pi_2^*(\xi')(B(\mathbf{y}, r)) > 1/(N+1)$, so that $\xi' \in C_n$. $\qquad\square$

6. Marriage via Brin–Katok

Let (X, μ, T) be a non-trivial ergodic subshift with invariant measure μ. Let S be a self-homeomorphism on a compact metric space Y. Given a μ-joining, ξ, with $h_{\pi_2^*(\xi)}(S) > h_\mu(T)$, we will define an injective map from blocks of X to orbit segments of Y, with certain properties that will be useful in constructing perturbations of ξ that are approximate embeddings.

This section, while in spirit closely following the proof of the Krieger generator theorem given by Burton, Keane and Serafin [10], differs in the details because we embed into a general compact metric space, rather than into a shift space. Burton, Keane and Serafin make use of the Shannon-Macmillan-Breiman theorem and Hall's marriage theorem, which are also important ingredients in the proofs of the Sinai factor theorem and Ornstein isomorphism theorem given by Keane and

Smorodinsky [28, 29]. We substitute the Shannon-Macmillan-Breiman theorem with a topological analogue due to Brin and Katok [8]. In order to define *markers* in a general setting that is not necessarily symbolic, we will also make use of a generalization due to Downarowicz and Weiss [19] of the Ornstein and Weiss [43] return time formula for entropy.

Let S be a self-homeomorphism on a compact metric space (Y, d). For integers $m < n$ let

$$\mathrm{d}_m^n(y, z) = \max_{m \leq j < n} \mathrm{d}(S^j y, S^j z) \quad \text{for all} \quad y, z \in Y.$$

We define the $(\mathbf{d}_m^n, \boldsymbol{\delta})$-**Bowen ball** about a point $y \in Y$ by

$$B_m^n(y, \delta) := \{z \in Y : \mathrm{d}_m^n(y, z) < \delta\}.$$

For an S-invariant measure λ on Y we define for $(z, \eta) \in Y \times \mathbb{R}^+$

$$\underline{h}_\lambda^{\mathrm{BK}}(z, \eta) = \liminf_{n \to \infty} -(1/n) \log \lambda(B_0^n(z, \eta))$$

and

$$\underline{h}_\lambda^{\mathrm{DW}}(z, \eta) = \liminf_{n \to \infty} (1/n) \log \min\{i > 0 \colon \mathrm{d}_0^n(z, S^i z) < \eta\}.$$

It can be easily checked that $\underline{h}_\lambda^{\mathrm{BK}}(z, \eta)$ and $\underline{h}_\lambda^{\mathrm{DW}}(z, \eta)$ are S-invariant functions of z, so that if λ is an ergodic measure, there are monotone functions $\underline{h}_\lambda^{\mathrm{BK}}(\eta)$ and $\underline{h}_\lambda^{\mathrm{DW}}(\eta)$ such that $\underline{h}_\lambda^{\mathrm{BK}}(z, \eta) = \underline{h}_\lambda^{\mathrm{BK}}(\eta)$ and $\underline{h}_\lambda^{\mathrm{DW}}(z, \eta) = \underline{h}_\lambda^{\mathrm{DW}}(\eta)$ for λ-almost every $z \in Y$.

Theorem 22 (Brin and Katok [8]). *Let S be a self-homeomorphism of a compact metric space Y preserving an ergodic measure λ and suppose that $h_\lambda(S) < \infty$. Then for λ-almost every $z \in Y$, we have*

$$\lim_{\eta \to 0} \underline{h}_\lambda^{\mathrm{BK}}(z, \eta) = h_\lambda(S).$$

Theorem 23 (Downarowicz and Weiss [19]). *Suppose that S is a self-homeomorphism of a compact metric space Y preserving an ergodic measure λ and suppose that $h_\lambda(S) < \infty$. Then for λ-a.e. $z \in Y$, we have*

$$\lim_{\eta \to 0} \underline{h}^{\mathrm{DW}}(z, \eta) = h_\lambda(S).$$

Lemma 24 (Marker Lemma). *Let S be a self-homeomorphism of a compact metric space Y preserving an ergodic measure ν with $0 < h_\nu(S) < \infty$. Let $\alpha \in (0, 1)$ and let r be such that $\underline{h}_\nu^{\mathrm{BK}}(2r) > 0$ and $\underline{h}_\nu^{\mathrm{DW}}(2r) > 0$.*

Then for all sufficiently large integers M, there is a point $y_{\mathrm{mark}} \in Y$ with the following properties:

(1) $\mathrm{d}_0^{2M}(S^i y_{\mathrm{mark}}, S^{6M} y_{\mathrm{mark}}) \geq 2r$ for $i \in \{0, \ldots, 6M - 1\}$;

(2) $\nu(H_1) < \alpha/M$ where $H_1 := B_0^{2M}(y_{\mathrm{mark}}, 2r)$.
(3) $\nu(H_2) < \alpha/M$ where $H_2 := B_0^{2M}(S^{6M}y_{\mathrm{mark}}, 2r)$.

Proof. Property (1) comes from applying Theorem 23 to S^{-1}. Properties (2) and (3) follow from Theorem 22. $\square$

Corollary 25 (Marker decipherability). *Let Y, S, M, r, H_1 and H_2 be as in the statement of Lemma 24. Let $z \in Y$ be a point such that*

(i) $S^i z \notin H_1$ for $0 \le i < L$;
(ii) $S^i z \notin H_2$ for $0 \le i < L$.

Also suppose that S satisfies almost weak specification. Let $L > 0$ and assume that the gap function $L_r(L) < M$. Let $\tilde{z} \in Y$ be a point satisfying the specification conditions:

(a) $d_M^{L-10M}(\tilde{z}, z) < r$
(b) $d_0^{8M}(S^{L-9M}\tilde{z}, y_{\mathrm{mark}}) < r$.

such a point exists by almost weak specification. Then $\tilde{z}$ satisfies

$$d_0^{8M}(S^i\tilde{z}, y_{\mathrm{mark}}) \ge r \text{ for } 0 \le i < L - 9M.$$

In particular, if $0 \le j \le L - 9M$, then j may be recovered from $S^j\tilde{z}$ by $j = (L - 9M) - \min\{i \ge 0 \colon d_0^{8M}(S^i\tilde{z}, y_{mark}) < r\}$

Proof. We will check three cases.

If $0 \le i < L - 18M$, then we have

$$\begin{aligned}
d_0^{8M}(S^i\tilde{z}, y_{\mathrm{mark}}) &\ge d_0^{2M}(S^{i+6M}\tilde{z}, S^{6M}y_{\mathrm{mark}}) \\
&\ge d_0^{2M}(S^{i+6M}z, S^{6M}y_{\mathrm{mark}}) - r \text{ (by (a))} \\
&\ge 2r - r = r \text{ (by (ii))}.
\end{aligned}$$

If $M \le i < L - 11M$, then we have

$$\begin{aligned}
d_0^{8M}(S^i\tilde{z}, y_{\mathrm{mark}}) &\ge d_0^{2M}(S^i\tilde{z}, y_{\mathrm{mark}}) \\
&\ge d_0^{2M}(S^i z, y_{\mathrm{mark}}) - r \text{ (by (a))} \\
&\ge 2r - r = r \text{ (by (i))}.
\end{aligned}$$

If $L - 11M \le i < L - 9M$, then we have

$$d_0^{8M}(S^i\tilde{z}, y_{\mathrm{mark}}) \ge d_0^{2M}(S^{i+6M}\tilde{z}, S^{6M}y_{\mathrm{mark}}) \ge$$
$$d_0^{2M}(S^{i+15M-L}y_{\mathrm{mark}}, S^{6M}y_{\mathrm{mark}}) - d_0^{2M}(S^{i+15M-L}y_{\mathrm{mark}}, S^{i+6M}\tilde{z})$$
$$\ge 2r - r = r \text{ (by Lemma 24 (1)) and (b).} \qquad \square$$

The following lemma involves the choice of many constants (which are necessary for our proof of Proposition 18 (III)), but it follows easily from Theorem 22 and the ergodic theorem. We will refer to Bowen balls in the product space $X \times Y$, where X is a non-trivial subshift, and Y

is a compact metric space endowed with a self-homeomorphism. For the definition of these, we will take the metric to be

$$\mathrm{d}_{X \times Y}((x, y), (x', y')) = \max(\mathrm{d}_X(x, x'), \mathrm{d}_Y(y, y')).$$

Lemma 26. *Let (X, μ, T) be a non-trivial ergodic subshift with invariant measure μ and natural generating partition $\mathcal{P}$. Let S be a self-homeomorphism with almost weak specification of a compact metric space Y that satisfies the small boundary condition witnessed by a sequence of refining partitions $(\mathcal{Q}_\ell)$, where $\mathrm{diam}(\mathcal{Q}_\ell) < 1/\ell$. Let ξ be a μ-joining with $h_{\pi_2^*(\xi)}(S) > h_\mu(T)$. Let $\varepsilon > 0$.*
  Set

$$\Delta := (h_{\pi_2^*(\xi)}(S) - h_\mu(T))/10. \tag{10}$$

Let $0 < \delta < \epsilon/80$ satisfy the inequalities:

$$\delta\big(1 + h_\xi(T \times S)\big) < \Delta; \tag{11}$$

$$16\delta(1 + \log|\mathcal{P}|) < \epsilon\Delta; \tag{12}$$

$$4(1 - \delta)(1 - 15\delta) \geq 3. \tag{13}$$

By Theorem 22 let $\eta < \epsilon/12$ be chosen so that

$$\underline{h}_{\pi_2^*(\xi)}^{BK}(2\eta, S) > h_{\pi_2^*(\xi)}(S) - \delta \text{ and} \tag{14}$$

$$\underline{h}_\xi^{BK}(2\eta, T \times S) > h_\xi(T \times S) - \delta. \tag{15}$$

Let $\ell > 10/\eta$ so that $\mathrm{diam}(\mathcal{Q}_\ell) < \eta/10$. Let $r < \eta/10$ satisfy the condition of Lemma 24 and be such that

$$\pi_2^*(\xi)(\partial_r \mathcal{Q}_\ell) < \delta. \tag{16}$$

Set

$$M := \lfloor \delta N/11 \rfloor. \tag{17}$$

Then then following conditions hold for all sufficiently large N:
(a) $e^{N\Delta} > 2$;
(b) $2^{-M} < r$;
(c) $\frac{1}{N} < \delta$;
(d) the gap function satisfies $L_r(N) < M$;
(e) Lemma 24 holds with $\alpha = \delta^2/22$ and the value of M as defined in (17);
(1) $\mu(S_{1,N}) > 1 - \delta$ where

$$S_{1,N} = \{x \in X : \mu(B_0^N(x, \eta)) > e^{-(h_\mu(T) + \Delta)N}\};$$

(2) $\pi_2^*(\xi)(S_{2,N}) > 1 - \delta$ where

$$S_{2,N} = \{y \in Y : \pi_2^*(\xi)(B_M^{N-10M}(y, 2\eta)) < e^{-(h_{\pi_2^*(\xi)}(S)-\Delta)N}\};$$

(3) $\xi(S_{3,N}) > 1 - \delta$ where

$$S_{3,N} = \{(x,y) \in X \times Y : \xi(B_M^{N-10M}(x,y), 2\eta) < e^{-(h_\xi(T \times S)-\Delta)N}\};$$

(4) $\xi(S_{4,N}) > 1 - \delta$ where

$$S_{4,N} = \{(x,y) \in X \times Y : \xi(B_0^N((x,y),\eta)) > e^{-(h_\xi(T \times S)+\Delta)N}\};$$

(5) $\xi(S_{5,N}) > 1 - \delta$ where

$$S_{5,N} = \left\{(x,y) : |\mathbb{A}_M^{N-10M}f(x,y) - \xi(f)| < \tfrac{\epsilon}{12} \text{ for all } f \in \mathrm{Lip}_1(X \times Y)\right\};$$

(6) $\pi_2^*(\xi)(S_{6,N}) > 1 - \delta$ where

$$S_{6,N} = \{y \in Y : \mathbb{A}_M^{N-10M}\mathbf{1}_{\partial_r \mathcal{Q}_\ell}(y) < \delta\}.$$

Note that in Lemma 26, our choices of constants give:

$$120r, \ 80\delta, \ 12\eta \ < \ \varepsilon. \tag{18}$$

Proof of Lemma 26. Conditions (a), (b) and (c) are trivial. By Lemma 24, for all sufficiently large M (hence for sufficiently large N), the conditions are satisfied establishing (e). Condition (d) holds for large N as a consequence of the definition of almost weak specification.

Conditions (1) and (4) follow for large N from the upper estimate in Theorem 22.

For (2), we note that

$$\pi_2^*(\xi)(B_M^{N-10M}(y, 2\eta)) = \pi_2^*(\xi)(B_0^{N-11M}(y, 2\eta)).$$

Hence for $\pi_2^*(\xi)$-almost every y, we have

$$\liminf_{N \to \infty} -(1/N)\log \pi_2^*(\xi)(B_M^{N-10M}(y, 2\eta)) = (1-\delta)\underline{h}_{\pi_2^*(\xi)}^{\mathrm{BK}}(2\eta)$$
$$> (1-\delta)(h_{\pi_2^*(\xi)}(S) - \delta) \ \text{(by (14))}$$
$$> h_{\pi_2^*(\xi)}(S) - \delta(1 + h_{\pi_2^*(\xi)}(S))$$
$$> h_{\pi_2^*(\xi)}(S) - \Delta \ \text{(by (11))}.$$

A similar argument shows (3) (using (15)) holds for large N.

That condition (6) holds for large N, follows from the Birkhoff ergodic theorem and (16). That condition (5) holds for large N, would follow immediately from the Birkhoff ergodic theorem, if Lip_1 were a

finite set. Recall that by Lip_1, we mean those functions with Lipschitz constant 1 taking values in $[0, 1]$. By the Arzelà-Ascoli theorem, Lip_1 is totally bounded with respect to the uniform norm on continuous functions. Hence there exists a finite collection $F \subset \mathrm{Lip}_1$ that is $\epsilon/24$-dense in Lip_1 with respect to the uniform norm. Now set

$$S_{5',N} = \{(x, y) \colon |\mathbb{A}_M^{N-10M} f(x, y)| < \epsilon/24 \text{ for all } f \in F\}.$$

An application of the triangle inequality shows that $S_{5,N} \supset S_{5',N}$. Notice that since $M = cN$, $\mathbb{A}_M^N f = N/(N-M) \cdot (1/N)(N\mathbb{A}_0^N f - cN\mathbb{A}_0^{cN} f)$, so that for a fixed $f \in \mathrm{Lip}_1(X \times Y)$, we have $\mathbb{A}_M^N f(x) \to \xi(f)$ ξ-a.e. Since F is finite, we have that $\xi(S_{5',N}) \to 1$ as $N \to \infty$ by the Birkhoff ergodic theorem. $\qquad\square$

Lemma 26 and the following variant of Hall's marriage theorem will be used to define the injective map from blocks of X to orbit segments of Y. Given a relation $R \subset \mathsf{B} \times \mathsf{G}$, we let $R(b, \cdot) := \{g \in \mathsf{G} : (b, g) \in R\}$ and $R(\cdot, g) := \{b \in \mathsf{B} : (b, g) \in R\}$.

Theorem 27 (Hall's marriage theorem [23])**.** *Let B and G be finite sets. Let $R \subset \mathsf{B} \times \mathsf{G}$ be a relation with the property that there exists $K > 0$ such that for all $b \in \mathsf{B}$ we have $|R(b, \cdot)| \geq K$, and for all $g \in \mathsf{G}$ we have $|R(\cdot, g)| \leq K$. Then there exists an injection $\phi : \mathsf{B} \to \mathsf{G}$ such that as a relation $\phi \subset R \subset \mathsf{B} \times \mathsf{G}$.*

A proof of Theorem 27 can be found in Downarowicz's book [17, Appendix A]. In Theorem 27, the sets B and G are often referred to as boys and girls, respectively; we call the integer K a ***marriage bound*** for R, and the map ϕ a ***dictionary***.

We say that a subset $A \subset Y$ is $(\mathbf{d}_m^n, \boldsymbol{\eta})$-***separated*** if $\mathrm{d}_m^n(a, a') \geq \eta$ for each distinct pair a, a' in A; and is said to be a $(\mathbf{d}_m^n, \boldsymbol{\eta})$-***spanning*** subset of $B \subset Z$ if $A \subset B$ and for each $b \in B$, there exists an $a \in A$ with $\mathrm{d}_m^n(a, b) \leq \eta$. Notice that a maximal (d_m^n, η)-separated subset of B is necessarily (d_m^n, η)-spanning.

Given a finite word, $B = b_0 \ldots b_{N-1}$, with symbols in the alphabet of X, we write $[B]$ for the cylinder set $\{x \in X \colon x_0 = b_0, \ldots, x_{N-1} = b_{N-1}\}$.

Corollary 28 (Corollary to Lemma 26)**.** *Fix N so that conditions of Lemma 26 hold, with a point $y_{\mathrm{mark}} \in Y$ satisfying* (e)*. Let H_1 and H_2 be defined as in Lemma 24. Let*

$$A := \{y \colon T^j y \notin H_1 \cup H_2 \text{ for } 0 \leq j < N\}.$$

Let G be a maximal $(\mathrm{d}_M^{N-10M}, \eta)$-separated subset of $A \cap S_{2,N} \cap S_{6,N}$. Let $U \subset X \times Y$ be given by

$$U = (X \times (S_{2,N} \cap S_{6,N} \cap A)) \cap (S_{1,N} \times Y) \cap S_{3,N} \cap S_{4,N} \cap S_{5,N}.$$

Let B be the set of elements $B \in \mathcal{P}^N$ such that $\mu([B]) \geq e^{-N(h_\mu(T)+\Delta)}$ and $\xi(([B] \times Y) \cap U) \geq \frac{1}{2}\mu([B])$.

Define

$$R = \left\{ (B, y) \in \mathsf{B} \times \mathsf{G} \colon \exists (u, v) \in U \text{ with } u \in [B], \ \mathrm{d}_M^{N-10M}(y, v) < \eta \right\}.$$
$$K = \tfrac{1}{2} e^{N(h_\xi(T \times S) - h_\mu(T) - 2\Delta)}.$$

Then

(A) $\xi(U) > 1 - 7\delta$;

(B) $|\mathsf{G}| > \frac{1}{2} e^{N(h_{\pi_2^*(\xi)}(S) - \Delta)}$;

(C) $|\mathsf{B}| \leq e^{N(h_\mu(T)+\Delta)}$ and $\mu\left(\bigcup_{B \in \mathsf{B}} [B]\right) > 1 - 15\delta$;

(D) K is a marriage bound for R;

(E) there exists an injection $\phi \colon \mathsf{B} \to \mathsf{G}$ such that $\phi \subset R$;

(F) if $B \in \mathsf{B}$, then for all $x \in [B]$, $y \in B_M^{N-10M}(\phi(B), r)$, and for all $f \in \mathrm{Lip}_1(X \times Y)$, we have

$$\left| \mathbb{A}_M^{N-10M} f(x, y) - \xi(f) \right| < \tfrac{\epsilon}{12} + \eta + r.$$

Proof. (A) Since ξ is a joining, from Lemma 26 (2), (6), and (e), we see that $\xi(X \times (S_{2,N} \cap S_{6,N} \cap A)) > 1 - 3\delta$ and from Lemma 26 (1) $\xi(S_{1,N} \times Y) > 1 - \delta$ so that Condition (A) holds.

(B) By the definition of G we have,

$$\bigcup_{y \in \mathsf{G}} B_M^{N-10M}(y, \eta) \supset A \cap S_{2,N} \cap S_{6,N}.$$

By Lemma 26 (2) we have for each $y \in \mathsf{G}$, $\pi_2^*(\xi)(B_M^{N-10M}(y, \eta)) < e^{-N(h_{\pi_2^*(\xi)}(S) - \Delta)}$. We deduce $|\mathsf{G}| > \frac{1}{2} e^{N(h_{\pi_2^*(\xi)}(Y) - \Delta)}$.

(C) The first claim follows from the definition of B. For the second, let $\mathsf{B}_1 = \{B \in \mathcal{P}^N \colon \mu([B]) > e^{-(h_\mu(T)+\Delta)N}\}$. For $x \in [B]$, we have $[B] = B(x, N, \frac{1}{2}) \supseteq B(x, N, \eta)$. By Lemma 26 (1), we have $\mu(\bigcup_{B \in \mathsf{B}_1} [B]) > 1 - \delta$. We then have

$$7\delta > \xi(U^c) \geq \xi\left(\bigcup_{B \in \mathsf{B}_1 \setminus \mathsf{B}} ([B] \times Y) \cap U^c \right) \geq \sum_{B \in \mathsf{B}_1 \setminus \mathsf{B}} \tfrac{1}{2}\mu([B]).$$

This yields $\mu\left(\bigcup_{B \in \mathsf{B}} [B] \right) > 1 - 15\delta$ as required.

(D) Let $B \in \mathsf{B}$ and let $R(B, \cdot) = \{y_1, \ldots, y_n\}$. We will show that $n \geq K$. Since $(B, y_j) \in R$, it follows there exist $(u_j, v_j) \in U$ such that $u_j \in [B]$ and $\mathrm{d}_M^{N-10M}(v_j, y_j) < \eta$. Let $S_j = B_M^{N-10M}((u_j, v_j), 2\eta)$.

Since $(u_j, v_j) \in U$, by Lemma 26 (3), $\xi(S_j) < e^{-(h_\xi(T \times S) - \Delta)N}$. We claim that the S_j cover $([B] \times Y) \cap U$. To see this, let $(u, v) \in U$ satisfy $u \in [B]$. Since $(u, v) \in U$, we have $v \in A \cap S_{2,N} \cap S_{6,N}$. Hence, by the definition of G, there exists $y \in \mathsf{G}$ with $\mathrm{d}_M^{N-10M}(y, v) < \eta$, so that $(B, y) \in R$; hence $y = y_j$ for some j. It follows that $\mathrm{d}_M^{N-10M}(v, v_j) < 2\eta$. Since $u, u_j \in [B]$, we have $\mathrm{d}_M^{N-10M}(u, u_j) < 2^{-M} < \eta$ so that $\mathrm{d}_M^{N-10M}((u, v), (u_j, v_j)) < 2\eta$ and $(u, v) \in S_j$ as required. Since $B \in \mathsf{B}$, we have $\xi(([B] \times Y) \cap U) \geq \frac{1}{2}\mu([B]) \geq \frac{1}{2}e^{-(h_\mu(T) + \Delta)N}$. Since $([B] \times Y) \cap U$ is covered by the S_j's, we see that $n \geq \frac{1}{2}e^{N(h_\xi(T \times S) - h_\mu(T) - 2\Delta)} = K$ as required.

For the other half of the argument, let $y \in \mathsf{G}$ and let $R(\cdot, y) = \{B_1, \ldots, B_m\}$. Pick witnesses $(u_i, v_i) \in U$ so that $u_i \in [B_i]$ and $d_M^{N-10M}(v_i, y) < \eta$. Let $D_i = B_0^N((u_i, v_i), \eta)$. These sets are all contained in $X \times B_M^{N-10M}(y, 2\eta)$, which, since $y \in \mathsf{G}$, has measure at most $e^{-(h_{\pi_2^*(\xi)}(S) - \Delta)N}$ by Lemma 26 (2). Since $(u_i, v_i) \in U$, by Lemma 26 (4), each D_i has measure at least $e^{-(h_\xi(T \times S) + \Delta)N}$. Finally, if $i \neq i'$, then $\mathrm{d}_0^N(u_i, u_{i'}) = 1$ so that the D_i are disjoint. In particular, we deduce from (10) and Lemma 26 (a) that $m \leq e^{N(h_\xi(T \times S) - h_{\pi_2^*(\xi)}(S) + 2\Delta)} < K$.

(E) Property (E) follows immediately from Property (D) and Theorem 27.

(F) Let $B \in \mathsf{B}$ and $x \in [B]$. Let $f \in \mathrm{Lip}_1$. Since $(B, \phi(B)) \in R$, there exists $(x_0, y_0) \in U$ such that $x_0 \in [B]$ and $y_0 \in B_M^{N-10M}(\phi(B), \eta)$. Thus from the definition of U and Lemma 26 (5), we have

$$|\mathbb{A}_M^{N-10M} f(x_0, y_0) - \xi(f)| < \epsilon/12.$$

If $x \in [B]$ and y is any point in $B_M^{N-10M}(\phi(B), r)$, then we have $\mathrm{d}((T^i x, S^i y), (T^i x_0, S^i y_0)) < r + \eta$ for each $M \leq i < N - 10M$, so that by the Lipschitz property,

$$\left|\mathbb{A}_M^{N-10M} f(x, y) - \xi(f)\right| < \tfrac{\epsilon}{12} + \eta + r,$$

as required. $\qquad\square$

7. PROOF OF PROPOSITION 18 (III)

In this section, we use Corollary 25, Lemma 26 and Corollary 28 to build a joining. Using Lemma 20, we then prove the following restatement of Proposition 18 (III), where we assume a strict entropy gap.

Lemma 29. *Consider the setup of Lemma 26. Let (X, μ, T) be a nontrivial ergodic subshift with invariant measure μ and natural generating*

partition $\mathcal{P}$. Let S be a self-homeomorphism with almost weak specification on a compact metric space Y that satisfies the small boundary condition witnessed by a sequence of refining partitions $(\mathcal{Q}_\ell)$, each having zero measure boundary, where $\mathrm{diam}(\mathcal{Q}_\ell) < 1/\ell$. Let ξ be a μ-joining with $h_{\pi_2^*(\xi)}(S) > h_\mu(T)$. Let $\varepsilon > 0$. There exist ℓ and a μ-joining, $\tilde{\xi}$, satisfying the following properties:

- *weak*-closeness: $\mathrm{d}^*(\xi, \tilde{\xi}) < \varepsilon$;*
- *approximate embedding properties: $\mathcal{Q}_\ell \overset{\varepsilon}{\subset} \bigvee_{i \in \mathbb{Z}} \mathcal{T}^{-i} P \mod \tilde{\xi}$; and $\mathcal{P} \overset{\varepsilon}{\subset} \bigvee_{i \in \mathbb{Z}} S^{-i} \mathcal{Q}_\ell \mod \tilde{\xi}$;*
- *entropy preservation: $h_{\pi_2^*(\tilde{\xi})}(S) \geq h_\mu(T)$.*

Proof of Proposition 18 (III). Let $\xi \in \mathcal{M}_0$. Let $n > 0$. Without loss of generality, let $0 < \varepsilon < 1/n$. We need to find a $\tilde{\xi} \in E_\mu^n$ such that $\mathrm{d}^*(\xi, \tilde{\xi}) < \varepsilon$. By Lemma 20, we may assume without loss of generality that $h_{\pi_2^*(\xi)}(S) > h_\mu(T)$. So the result follows from Lemma 29. $\qquad\square$

We need one more tool before we can define the joining in the proof of Lemma 29. We will make use of the following variation of the Rokhlin tower theorem.

Lemma 30 (Rokhlin tower theorem: independent base version). *Let $(\Omega, \mathcal{F}, \mu, T)$ be a non-periodic measure-preserving system. Let N be a positive integer and $\delta > 0$. For any finite measurable partition $\mathcal{P}$, there exists $F \in \mathcal{F}$ (the* base*) with the following properties.*

(1) *The sets $F, TF, \ldots, T^{N-1}F$ are pairwise disjoint.*
(2) *The complement of their union $E := \Omega \setminus \bigcup_{i=0}^{N-1} T^i F$ (the* error set*) has measure exactly δ.*
(3) *The σ-algebras generated by F and $\mathcal{P}$ are independent, so that $\mu(F \cap P) = \mu(F)\mu(P)$ for all $P \in \mathcal{P}$.*

For a proof, see the book of McCutcheon and Kalikow [26, Theorem 184]. After defining the joining in the proof of Lemma 29, we motivate why it satisfies the desired properties before we proceed with technical calculations.

Let (Y, d) be a compact metric space. Given two sequences of points $u = (y_i)$, $v = (y_i')$ of Y, and $r > 0$, we say that ***u r-shadows v in I*** if $\mathrm{d}(y_i, y_i') < \varepsilon$ for all $i \in I$.

Proof of Lemma 29: definition of $\tilde{\xi}$. Choose $\Delta, \delta, \eta, \ell, r, M$, and N as in Lemma 26. By Lemma 30, let F be the base of a Rokhlin tower in X with height N and error set E_0 of measure δ with the property that F is independent of $\bigvee_{i=0}^{N-1} T^{-i} \mathcal{P}$. Given a point x that we assume to

be generic for μ, we let $n_0(x) = \sup\{k \le 0 : T^k x \in F \cup E_0\}$. We then
let $(n_j(x))_{j \in \mathbb{Z}}$ be the enumeration of $\{k \in \mathbb{Z} : T^k x \in F \cup E_0\}$ satisfying

$$\ldots < n_{-2}(x) < n_{-1}(x) < n_0(x) < n_1(x) < \ldots.$$

The jth block of x is then the block

$$B_j(x) = x_{n_j(x)} x_{n_j(x)+1} \ldots x_{n_{j+1}(x)-1}.$$

The blocks are of length N (for those j such that $T^{n_j(x)}(x) \in F$) and
1 for those points on the orbit that land in the error set E_0.

Let y_{mark}, B, G, and ϕ be given by Corollary 28. We introduce special
symbols $\mathbf{D}, \mathbf{V} \notin Y$ (with $\mathbf{D}$ standing for 'dictionary' and $\mathbf{V}$ standing
for 'vacuous'). Let $\mathsf{G_D} = \mathsf{G} \cup \{\mathbf{D}\}$ and $Y_{\mathbf{V}} = Y \cup \{\mathbf{V}\}$. Equip $\mathsf{G_D^{\mathbb{Z}}}$
with the usual shift map σ, and the Bernoulli measure ζ giving each
coordinate mass $\epsilon/(2|\mathsf{G}|)$ to each member of G and $1 - \epsilon/2$ to $\mathbf{D}$. If
$n_{j+1}(x) - n_j(x) = N$, let $m_j(x) := n_j(x) + N - 9M$ and set

$$R_j^{\text{info}}(x) = [n_j(x) + M, n_j(x) + N - 10M); \text{ and}$$
$$R_j^{\text{mark}}(x) = [m_j(x), m_j(x) + 8M);$$

otherwise, let $R_j^{\text{info}}(x) = R_j^{\text{mark}}(x) = \emptyset$. We now define a map $\psi :$
$X \times \mathsf{G_D^{\mathbb{Z}}} \to Y_{\mathbf{V}}^{\mathbb{Z}}$ by

$$\psi(x, z)_n = \begin{cases} S^{n - n_j(x)} z_{n_j(x)} & n \in R_j^{\text{info}}(x), \, B_j(x) \in \mathsf{B}, \, z_{n_j(x)} \in \mathsf{G}; \\ S^{n - n_j(x)} \phi(B_j(x)) & n \in R_j^{\text{info}}(x), \, B_j(x) \in \mathsf{B}, \, z_{n_j(x)} = \mathbf{D}; \\ S^{n - m_j(x)} y_{\text{mark}} & n \in R_j^{\text{mark}}(x), \, B_j(x) \in \mathsf{B}, \, z_{n_j(x)} = \mathbf{D}; \\ \mathbf{V} & \text{otherwise.} \end{cases}$$

The sequence $\psi(x, z)$ of points in $Y_{\mathbf{V}}$ is the sequence of points that
we wish to r-shadow, while $\mathbf{V}$ denotes the null constraint. By condition
(d) of Lemma 26, the orbit segments that we are attempting to shadow
are of length less than N and are separated by at least $M > L_r(N)$, so
that given $(\mathbf{x}, \mathbf{z})$, a $\mu \times \zeta$-generic point of $X \times \mathsf{G_D^{\mathbb{Z}}}$, there exists $\mathbf{y} \in Y$
satisfying the specification constraint given by

$$\mathrm{d}(S^n \mathbf{y}, \psi(\mathbf{x}, \mathbf{z})_n) < r \text{ whenever } \psi(\mathbf{x}, \mathbf{z})_n \neq \mathbf{V}, \text{ for all } n \in \mathbb{Z}.$$

Define a product space Ω and an invertible measure-preserving trans-
formation τ via:

$$\Omega := X \times \mathsf{G_D^{\mathbb{Z}}} \times Y \text{ and } \tau := T \times \sigma \times S.$$

As in the proof of Lemma 20, there exists an ergodic invariant measure
ι on Ω such that for ι-almost all points $\omega = (x, z, y)$

$$\mathrm{d}(S^n y, \psi(x, z)_n) < r \text{ whenever } \psi(x, y) \neq \mathbf{V}, \text{ for all } n \in \mathbb{Z}. \qquad (19)$$

Define $\tilde{\xi}$ to be the ergodic measure obtained by projecting ι onto the first and last coordinates of the tuple. By the genericity assumption above, $\tilde{\xi}$ is a μ-joining. $\qquad\square$

It remains to verify that $\tilde{\xi}$ satisfies the required properties. Using Corollary 28 (F) on the blocks of length $N-11M$ where we are applying the dictionary ϕ on B, we will obtain the weak*-closeness by Corollary 28 (C) and (F).

The fact that we are placing markers in between the coded boys will enable us to recover most of x from y, giving us the latter half of the approximate embedding property. To guess which element of $\mathcal{Q}_\ell$ that a point y belongs to from x, we have no chance if $B_0(x)$ does not belong to B. If it does, and the dictionary ϕ is applied, instead of using a random element of G, then we know that $(S^k y)$ shadows $(S^{k-n_0(x)}\phi(B_j(x)))$ in $R_0^{\text{info}}(x)$. Unless $S^{n_0(x)}\phi(B_j(x))$ lies close to the boundary of $\mathcal{Q}_\ell$, we can deduce which element y lies in, giving us the other approximate embedding property.

Finally, the entropy of S with respect to $\pi_2^*(\tilde{\xi})$ has a lower bound that is the sum of two contributions: one term is the entropy obtained by encoding most of the blocks of X, and another is given by sometimes using random elements of G. We do not encode blocks that are not members of B and because of the error set in the Rokhlin's lemma the contribution from encoding blocks of X is strictly less than $h_\mu(T)$; however, this entropy loss is at most

$$\left(\mu\Big(\bigcup_{B\notin\mathsf{B}}[B] \Big) + \mu(E_0) \right) \cdot \log|\mathcal{P}|.$$

For an $\epsilon/2$ proportion of blocks that do belong to B, we do not apply the dictionary ϕ, but instead use a random element of G; we will see that this results in a net gain of entropy that is large enough to cover the losses incurred by encoding only members of B and by not being able to encode anything on the set E_0.

Proof of Lemma 29: weak-closeness.* Let $P \in \bigvee_{i=0}^{N-1} T^{-i}\mathcal{P}$, where $\mathcal{P}$ is the natural generating partition for X. By Lemma 30, $\mu(P\cap F) = \mu(P)\mu(F)$. Since T preserves the measure μ and for all $0 \le k \le N-1$ $\{x \in X : T^{-k}x \in F\} = \{x \in X : x \notin E_0, n_0(x) = -k\}$, we have

$$\mu\big\{x \in X : T^{n_0(x)}x \in P, x \notin E_0, n_0(x) = -k\big\} = \mu(P)\mu(F) \qquad (20)$$

for all $0 \le k \le N-1$ and hence

$$\mu\big\{x \in X : T^{n_0(x)}x \in P : x \notin E_0\big\} = \mu(P)(1-\delta). \qquad (21)$$

Define the 'bad set' by

$$\mathsf{BS}_1 := \{(x, z, y) \in \Omega \colon x \in E_0\}$$
$$\mathsf{BS}_2 := \{(x, z, y) \in \Omega \colon x \notin E_0, x_{n_0(x)} \cdots x_{n_0(x)+N-1} \notin \mathsf{B}\};$$
$$\mathsf{BS}_3 := \{(x, z, y) \in \Omega \colon x \notin E_0, z_{n_0(x)} \neq \mathbf{D}\};$$
$$\mathsf{BS}_4 := \{(x, z, y) \in \Omega \colon x \notin E_0, 0 \notin R_0^{\mathrm{info}}(x)\};$$
$$\mathsf{BS} := \mathsf{BS}_1 \cup \mathsf{BS}_2 \cup \mathsf{BS}_3 \cup \mathsf{BS}_4.$$

We have that $\iota(\mathsf{BS}_1) = \mu(E_0) = \delta$. By (21) and Corollary 28 (C), we get

$$\iota(\mathsf{BS}_2) = (1 - \delta)\mu\left(\bigcup_{B \in \mathcal{P}^N \setminus \mathsf{B}} [B]\right) \leq (1 - \delta)15\delta.$$

By independence of ζ and μ, we have $\iota(\mathsf{BS}_3) = (1 - \delta)(\varepsilon/2)$. Using the fact that the length of the interval $R_0^{\mathrm{info}}(x)$ is $N - 11M$ when $x \notin E_0$, by our choice of $M = \lfloor \delta N/11 \rfloor$ in (17), we have $\iota(\mathsf{BS}_4) \leq \delta$. Thus by (18), we have

$$\iota(\mathsf{BS}) \leq 17\delta + \varepsilon/2 < 3\epsilon/4. \tag{22}$$

Let $f \in \mathrm{Lip}_1(X \times Y)$. For $0 \leq k < N$, let $A_k = \{(x, z, y) \colon x \in \mathsf{BS}_1^c \cap \mathsf{BS}_2^c \cap \mathsf{BS}_3^c, n_0(x) = -k\}$. Notice that

$$\mathsf{BS}^c = \mathsf{BS}_1^c \cap \mathsf{BS}_2^c \cap \mathsf{BS}_3^c \cap \mathsf{BS}_4^c$$
$$= \mathsf{BS}_1^c \cap \mathsf{BS}_2^c \cap \mathsf{BS}_3^c \cap \bigcup_{k=M}^{N-10M-1} \{(x, z, y) \in \Omega : n_0(x) = -k\}$$
$$= \bigcup_{k=M}^{N-10M-1} A_k.$$

So we have

$$\int f(x, y)\, d\tilde{\xi}(x, y) = \int f(x, y)\, d\iota(x, z, y)$$
$$= \int_{\mathsf{BS}} f(x, y)\, d\iota(x, z, y) + \sum_{k=M}^{N-10M-1} \int_{A_k} f(x, y)\, d\iota(x, z, y)$$
$$= \int_{\mathsf{BS}} f(x, y)\, d\iota(x, z, y) + (N - 11M) \int_{A_0} \mathbb{A}_M^{N-10M} f(x, y)\, d\iota(x, z, y).$$

By Corollary 28 (F), for $(x, z, y) \in A_0$, we have $|\mathbb{A}_M^{N-10M} f(x, y) - \xi(f)| < \varepsilon/12 + \eta + r$. We also have $(N - 11M)\mu(A_0) = \mu(\mathsf{BS}^c)$. Hence we by (22) and (18), we have

$$|\tilde{\xi}(f) - \xi(f)| < \mu(\mathsf{BS}) + \epsilon/12 + \eta + r < \epsilon. \qquad \square$$

Proof of Lemma 29: approximate embedding properties. We first show that $\mathcal{Q}_\ell \overset{\epsilon}{\subset} \bigvee_{i\in\mathbb{Z}} T^i\mathcal{P} \mod \tilde{\xi}$. Enumerate $\mathcal{Q}_\ell$ as $\{D_1, \ldots, D_n\}$ and regard $\mathcal{Q}_\ell$ as a map from Y to $\{1, \ldots, n\}$ where $\mathcal{Q}_\ell(y) = j$ if $y \in D_j$. Define

$$\chi(x) = \begin{cases} S_0^{-n_0(x)}\phi(B_0(x)) & \text{if it's defined;} \\ y_{\text{mark}} & \text{otherwise.} \end{cases}$$

We define $\mathcal{Q}'$ from X to $\{1, \ldots, n\}$ by

$$\mathcal{Q}'(x) = \mathcal{Q}_\ell(\chi(x)).$$

Since $\mathcal{Q}'$ is $\sigma(\bigvee_{i\in\mathbb{Z}} T^i\mathcal{P})$-measurable, it suffices to show that

$$\iota(\{(x, z, y)\colon \mathcal{Q}_\ell(y) \neq \mathcal{Q}'(x)\}) < \epsilon.$$

Let

$$\mathsf{BS}_5 := \{(x, z, y) \notin \mathsf{BS}\colon \chi(x) \in \partial_r\mathcal{Q}_\ell\}.$$

Notice that by definition of ι, we have

$$\{(x, z, y)\colon \mathcal{Q}_\ell(y) \neq \mathcal{Q}'(x)\} \subset \mathsf{BS} \cup \mathsf{BS}_5 \mod \iota,$$

By the definition of the set G and ϕ, we have that $\phi(B) \in S_{6,N}$ for all $B \in \mathsf{B}$. Thus by Lemma 26 (6) we have $\mathbb{A}_M^{N-10M}\mathbf{1}_{\mathsf{BS}_5}(x, z, y) < \delta$ on A_0. We compute as in the weak*-closeness section.

$$\iota(\mathsf{BS}_5) = \sum_{k=M}^{N-10M-1} \iota(\mathsf{BS}_5 \cap A_k)$$

$$= (N - 11M)\int_{A_0} \mathbb{A}_M^{N-10M}\mathbf{1}_{\mathsf{BS}_5}(x, z, y)\, d\iota(x, z, y) < \delta.$$

By (22) and (18), we have

$$\tilde{\xi}\{(x, y)\colon \mathcal{Q}_\ell(y) \neq \mathcal{Q}'(x)\} \leq \iota(\mathsf{BS} \cup \mathsf{BS}_5) \leq \tfrac{3}{4}\varepsilon + \delta < \epsilon.$$

For the approximate embedding in the opposite direction, we use the markers and the invertibility of ϕ to argue similarly. We let

$$\mathsf{BS}_6 = \{(x, z, y) \in \Omega\colon n_0(x) < -N + 9M\}.$$

We see $\iota(\mathsf{BS}_6) < 9M/N < \delta$. Define

$$\tilde{n}_0(y) = \min\left(\min\{k \geq 0\colon S^k(y) \in B_0^{8M-1}(y_{\text{mark}}, r)\} - (N - 9M), 0\right).$$

Equip B with an arbitrary total order and define a map $b\colon Y \to \mathsf{B}$ by

$$b(y) = \operatorname*{argmin}_{B\in\mathsf{B}} \mathrm{d}_M^{N-10M}(y, \phi(B)),$$

breaking ties lexicographically if necessary. Finally, set

$$\tilde{\mathcal{P}}(y) = b(S^{\tilde{n}_0(y)}(y))_{-\tilde{n}_0(y)}.$$

For ι-a.e. $(x, z, y) \in \Omega \setminus (\mathsf{BS} \cup \mathsf{BS}_6)$, we have $\tilde{n}_0(y) = n_0(x)$ by the definition of ι and Corollary 25. Since G is a (d_m^n, η)-separated set, by (19) and choice of parameters in Lemma 26, $\mathrm{d}_M^{N-10M}(y, \phi(B_0(x)) < r < \eta/10$. Since ϕ is one-to-one, there is only one girl, namely $\phi(B_0(x))$ within $\eta/5$ of the corresponding block of y; thus, $b(y) = B_0(x)$.

Hence we see that

$$\tilde{\xi}(\{(x, y) \colon \mathcal{P}(x) \neq \tilde{\mathcal{P}}(y)\}) \leq \iota(\mathsf{BS} \cup \mathsf{BS}_6) \leq 19\delta < \epsilon. \qquad \square$$

Proof of Lemma 29: entropy preservation. We define a number of partitions of Ω that we shall need in order to do calculations. Regard the partition $\mathcal{Q}_\ell$ of Y as a partition of Ω. We will express a typical point of $\omega \in \Omega$ as $\omega = (x, z, y)$. Define

$$C_1 := \{\omega \in \Omega \colon x \notin E_0, B_0(x) \in \mathsf{B} \text{ and } z_{n_0(x)} = \mathbf{D}\},$$

and

$$C_2 := \{\omega \in \Omega \colon x \notin E_0, B_0(x) \in \mathsf{B} \text{ and } z_{n_0(x)} \neq \mathbf{D}\}.$$

Let $C_0 := \Omega \setminus (C_1 \cup C_2)$, and $\mathcal{C} := \{C_0, C_1, C_2\}$. Set

$$\mathcal{R} := \sigma\left(\bigvee_{i \in \mathbb{Z}} \tau^{-i}\mathcal{C}\right).$$

Let

$$\mathcal{Q}_\ell^{1,2} := \{Q \cap C_0^c \colon Q \in \mathcal{Q}_\ell\} \cup \{C_0\}.$$

For each $j \in \{1, 2\}$, let

$$\mathcal{Q}_\ell^j := \{Q \cap C_j \colon Q \in \mathcal{Q}_\ell\} \cup \{\Omega \setminus C_j\}.$$

Let

$$\tilde{\mathcal{Q}}_\ell^{1,2} := \{\{\omega \in \Omega \colon \psi(x, z)_0 \in Q\} \colon Q \in \mathcal{Q}_\ell\} \cup \{C_0\}.$$

For each $j \in \{1, 2\}$, let

$$\tilde{\mathcal{Q}}_\ell^j := \{\{\omega \in \Omega \colon \psi(x, z)_0 \in Q\} \cap C_j \colon Q \in \mathcal{Q}_\ell\} \cup \{\Omega \setminus C_j\}.$$

Let $\bigvee_{i=M}^{N-10M-1} S^{-i}\mathcal{Q}_\ell = \{A_1, \ldots, A_L\}$. Let $\mathcal{Q}_\ell^{[\mathrm{BL}]1,2}$ be the partition given by $\{A_0^{[\mathrm{BL}]1,2}, A_1^{[\mathrm{BL}]1,2}, \ldots, A_L^{[\mathrm{BL}]1,2}\}$, where

$$A_i^{[\mathrm{BL}]1,2} := \{\omega \in C_1 \cup C_2 \colon n_0(x) = 0 \text{ and } y \in A_i\}$$

for $1 \leq i \leq L$ and $A_0^{[\mathrm{BL}]1,2}$ is the complementary set. (Here the 'BL' stands for 'block.') Let $\tilde{\mathcal{Q}}_\ell^{[\mathrm{BL}]1,2}$ be the partition given by

$$\{\tilde{A}_0^{[\mathrm{BL}]1,2}, \tilde{A}_1^{[\mathrm{BL}]1,2}, \ldots, \tilde{A}_L^{[\mathrm{BL}]1,2}\},$$

where

$$\tilde{A}_i^{[\mathrm{BL}]1,2} := \{\omega \in C_1 \cup C_2 \colon n_0(x) = 0 \text{ and } \psi(x, z)_M^{N-10M-1} \in A_i\}$$

for $1 \leq i \leq L$ and $\tilde{A}_0^{[\text{BL}]1,2}$ is the complementary set. Similarly for $j \in \{1,2\}$, let $\tilde{\mathcal{Q}}^{[\text{BL}]j}$ be the partition $\{\tilde{A}_0^{[\text{BL}]j}, \tilde{A}_1^{[\text{BL}]j}, \ldots, \tilde{A}_L^{[\text{BL}]j}\}$ where

$$\tilde{A}_i^{[\text{BL}]j} := \{\omega \in C_j : n_0(x) = 0 \text{ and } \psi(x,z)_0 \in A_i\}$$

and $\tilde{A}_0^{[\text{BL}]j}$ is the complementary set.

Finally let $\mathcal{P}^{[\text{BL}]1}$ be the partition with elements

$$\{\omega \in C_1 : x \in P \text{ and } n_0(x) = 0\},$$

where $P \in \bigvee_{i=0}^{N-1} T^{-i}\mathcal{P}$, together with the complement of the union of this collection; $\mathcal{P}^{[\text{BL}]2}$ be the partition with elements

$$\{\omega \in C_2 : x \in P \text{ and } n_0(x) = 0\},$$

where $P \in \bigvee_{i=0}^{N-1} T^{-i}\mathcal{P}$, together with the complement; and $\mathcal{P}^0$ be the partition with elements $P \cap \{\omega \colon B_0(x) \notin \mathsf{B} \text{ or } x \in E_0\}$, where $P \in \bigvee_{i=0}^{N-1} T^{-i}\mathcal{P}$, again along with the complement.

Let us pause to explain the above notations. The partition $\mathcal{C}$ tells you whether you are attempting to shadow an element of G determined by the dictionary ϕ (C_1), a random element of G (C_2) or if there is no constraint (C_0). For the $\mathcal{Q}_\ell$ partitions, the superscript 1 indicates that one is looking at those times when one is shadowing an element of G determined by ϕ; 2 indicates that one is shadowing an random element of G; and $1,2$ indicates that one is shadowing either of these two. The tildes indicate the partition element that one is aiming for (i.e. the partition element that $\psi(x,z)$ lies in) rather than the partition element that y actually ends up lying in. Also the superscript 'BL' indicates that one is getting a whole block's worth of information at once, whereas otherwise one gets the information a symbol at a time. The partition $\mathcal{Q}_\ell^{1,2}$ tells you which element of $\mathcal{Q}_\ell$ the point y ends up in for the parts that are constrained by $\psi(x,z)$. The partition $\mathcal{Q}_\ell^{[\text{BL}]1,2}$ tells you which element of $\bigvee_{i=M}^{N-10M-1} S^{-i}\mathcal{Q}_\ell$ the point y ends up in if x is at the base of the Rokhlin tower and is at the start of a B block; $\tilde{\mathcal{Q}}_\ell^{1,2}$ tells you which element of $\bigvee_{i=M}^{N-10M-1} S^{-i}\mathcal{Q}_\ell$ the orbit segment you are aiming for $((\psi(x,z)_i)_{M \leq i < N-10M})$ belongs to when x is at the base of the Rokhlin tower and is at the start of a B block. The partition $\tilde{\mathcal{Q}}_\ell^1$ tells you which element of $\bigvee_{i=M}^{N-10M-1} S^{-i}\mathcal{Q}_\ell$ the orbit segment $(\psi(x,z)_i)_{M \leq i < N-10M}$ belongs to when you are shadowing an element of G determined by ϕ; and $\tilde{\mathcal{Q}}_\ell^2$ tells you which element of $\bigvee_{i=M}^{N-10M-1} S^{-i}\mathcal{Q}_\ell$ the orbit segment $(\psi(x,z)_i)_{M \leq i < N-10M}$ belongs to when you are shadowing an random element of G. The partitions $\mathcal{P}^{[\text{BL}]j}$ (for $j = 1, 2$) tell you the block x_0^{N-1} when x is at the base of the tower and $B_0(x) \in \mathsf{B}$ if the dictionary is being used ($j = 1$) or if the word

is being randomized ($j = 2$). Note that the partition $\mathcal{P}^0$ tells you the symbol x_0 when $B_0(x) \notin \mathsf{B}$. or $x \in E_0$. The partition $\mathcal{P}^0 \vee \mathcal{P}^{[\mathrm{BL}]1} \vee \mathcal{P}^{[\mathrm{BL}]2}$ is a generating partition for μ

Let

$$\mathsf{E} := \{\omega \in \Omega : x \notin E_0, n_0(x) = 0, B_0(x) \in \mathsf{B}, z_0 \neq \mathbf{D}\}.$$

Then by (20) and Corollary 28 (C), and the independence of μ and ζ, we have $\iota(\mathsf{E}) \geq (1/N)(1 - \delta)(1 - 15\delta)(\epsilon/2)$, so that by (13), we have

$$8N\Delta \cdot \iota(\mathsf{E}) > 3\epsilon\Delta. \tag{23}$$

We also note that C_0 is the set of points in Ω whose first coordinate belongs to $E_0 \cup \bigcup_{i=0}^{N-1} \tau^{-i}\left(F \cap \bigcup_{B \notin \mathsf{B}}[B]\right)$. We therefore calculate

$$\iota(C_0) = \mu(E_0) + (1 - \mu(E_0))/N \cdot N \cdot \mu\left(\bigcup_{B \notin \mathsf{B}}[B]\right) < 16\delta. \text{ Using (12)}$$

we obtain

$$\iota(C_0) \log |\mathcal{P}| < \epsilon\Delta \tag{24}$$

The following facts will be used to complete the calculation.

(a) $h_\iota(\tau, \mathcal{Q}_\ell | \mathcal{R}) \geq h_\iota(\tau, \mathcal{Q}_\ell^{1,2} | \mathcal{R})$;

(b) $h_\iota(\tau, \mathcal{Q}_\ell^{1,2} | \mathcal{R}) \geq h_\iota(\tau, \tilde{\mathcal{Q}}_\ell^{1,2} | \mathcal{R})$;

(c) $h_\iota(\tau, \tilde{\mathcal{Q}}_\ell^{1,2} | \mathcal{R}) \geq h_\iota(\tau, \tilde{\mathcal{Q}}_\ell^{[\mathrm{BL}]1,2} | \mathcal{R})$;

(d) $h_\iota(\tau, \tilde{\mathcal{Q}}_\ell^{[\mathrm{BL}]1,2} | \mathcal{R}) = h_\iota(\tau, \tilde{\mathcal{Q}}_\ell^{[\mathrm{BL}]1} | \mathcal{R}) + h_\iota(\tau, \tilde{\mathcal{Q}}_\ell^{[\mathrm{BL}]2} | \mathcal{R})$;

(e) $h_\iota(\tau, \tilde{\mathcal{Q}}_\ell^{[\mathrm{BL}]1} | \mathcal{R}) = h_\iota(\tau, \mathcal{P}^{[\mathrm{BL}]1} | \mathcal{R})$;

(f) $h_\iota(\tau, \tilde{\mathcal{Q}}_\ell^{[\mathrm{BL}]2} | \mathcal{R}) = \mu(\mathsf{E}) \log |\mathsf{G}|$;

(g) $h_\iota(\tau, \mathcal{P} | \mathcal{R}) \leq h_\iota(\tau, \mathcal{P}^{[\mathrm{BL}]1} | \mathcal{R}) + h_\iota(\tau, \mathcal{P}^{[\mathrm{BL}]2} | \mathcal{R}) + h_\iota(\tau, \mathcal{P}^0 | \mathcal{R})$;

(h) $h_\iota(\tau, \mathcal{P}^0 | \mathcal{R}) \leq \iota(C_0) \log |\mathcal{P}| < \epsilon\Delta$;

(i) $h_\iota(\tau, \mathcal{P}^{[\mathrm{BL}]2} | \mathcal{R}) \leq \mu(\mathsf{E}) \log |\mathsf{B}|$;

(j) $h_\iota(\tau, \mathcal{C}) \leq (\delta + \frac{1}{N}) \log 3 \leq \epsilon\Delta$.

Notice that (a), (c) and (e) follow from the fact that if $\mathcal{P}_1$ and $\mathcal{P}_2$ are partitions such that $\sigma(\bigvee_{i \in \mathbb{Z}} T^{-i}\mathcal{P}_1) \supseteq \sigma(\bigvee_{i \in \mathbb{Z}} T^{-i}\mathcal{P}_2)$, then $h(\tau, \mathcal{P}_1) \geq h(\tau, \mathcal{P}_2)$. Facts (g) and (h) follow from standard entropy results, and (24). To see that (b) holds, notice that observing the elements of $\mathcal{R}$ and $\mathcal{Q}_\ell^{1,2}$ the point ω lies in is sufficient to determine which element of G was being targeted; this follows from the same reasoning used in the second half of the proof of the approximate embedding property. That $n_0(x) = 0$ and $\omega \in C_1$ is a $\mathcal{R}$-measurable event since ω points stay in C_1 for an exact multiple of N steps. Hence one can deduce the partition element of $\tilde{\mathcal{Q}}_\ell^{1,2}$ that ω belongs to given the sequence of $\mathcal{R}$ and $\mathcal{Q}_{1,2}$ partition elements. For (f) and (i), we work with the information functions $I_\iota(\tilde{\mathcal{Q}}_\ell^{B2} | \mathcal{R})$ and $I_\iota(\mathcal{P}^{B2} | \mathcal{R})$; these are 0 if $\omega \notin \mathsf{E}$, whereas if $\omega \notin \mathsf{E}$, then they are $\log |\mathsf{G}|$ (since each girl appears independently with

equal likelihood) and at most $\log |\mathsf{B}|$ respectively. To establish (d), it suffices to show that $h_\iota(\tilde{\mathcal{Q}}_\ell^{[\mathrm{BL}]2}|\tilde{\mathcal{Q}}_\ell^{[\mathrm{BL}]1} \vee \mathcal{R}) = h_\iota(\tilde{\mathcal{Q}}_\ell^{[\mathrm{BL}]2}|\mathcal{R})$. This follows since the second coordinate of Ω is independent of the first. Finally, to see (j), we use Abramov's formula with the induced transformation of τ to the set

$$A := \{\omega \in \Omega : x \in E_0 \cup F\}.$$

Notice that between visits to A, the system stays entirely in a single element of $\mathcal{C}$. Hence we see that $h_\iota(\tau, \mathcal{C}) \leq \iota(A) \log 3 \leq (\delta + \frac{1}{N}) \log 3$; this is bounded above by $\epsilon\Delta$ using (12) and Lemma 26 (c).

We then have the following calculation. By Corollary 28 (B) and (C),

$$\log(|\mathsf{G}|/|\mathsf{B}|) \geq N(h_{\pi_2^*(\xi)}(S) - h_\mu(T) - 2\Delta). \tag{25}$$

Let $h = h_{\pi_2^*(\tilde{\xi})}(S)$. We have that

$$h \geq h_{\pi_2^*(\tilde{\xi})}(S, \mathcal{Q}_\ell)$$

$$\geq h_\iota(\tau, \mathcal{Q}_\ell|\mathcal{R}) \geq h_\iota(\tau, \mathcal{Q}_\ell^{1,2}|\mathcal{R}) \geq h_\iota(\tau, \tilde{\mathcal{Q}}_\ell^{1,2}|\mathcal{R}) \text{ (by (a) and (b))}$$

$$\geq h_\iota(\tau, \tilde{\mathcal{Q}}_\ell^{[\mathrm{BL}]1,2}|\mathcal{R}) = h_\iota(\tau, \tilde{\mathcal{Q}}_\ell^{[\mathrm{BL}]1}|\mathcal{R}) + h_\iota(\tau, \tilde{\mathcal{Q}}_\ell^{[\mathrm{BL}]2}|\mathcal{R}) \text{ (by (c), (d))}$$

$$= h_\iota(\tau, \mathcal{P}^{[\mathrm{BL}]1}|\mathcal{R}) + \iota(\mathsf{E}) \log |\mathsf{G}| \text{ (by (e) and (f)).}$$

Hence

$$h \geq h_\iota(\tau, \mathcal{P}|\mathcal{R}) - h_\iota(\tau, \mathcal{P}^{[\mathrm{BL}]2}|\mathcal{R}) - h_\iota(\tau, \mathcal{P}^0|\mathcal{R}) + \iota(\mathsf{E}) \log |\mathsf{G}| \text{ (by (g))}$$

$$\geq h_\iota(\tau, \mathcal{P}) - h_\iota(\tau, \mathcal{C}) - \epsilon\Delta + \iota(\mathsf{E}) \log(|\mathsf{G}|/|\mathsf{B}|) \text{ (by (h) and (i))}$$

$$\geq h_\mu(T) - 2\epsilon\Delta + 8N\Delta\iota(\mathsf{E}) \text{ (by (j), (25) and (10))}$$

$$\geq h_\mu(T) \text{ (by (23)).} \qquad \square$$

References

[1] A. Bertrand-Mathis. Développement en base θ; répartition modulo un de la suite $(x\theta^n)_{n \geq 0}$; langages codés et θ-shift. *Bull. Soc. Math. France*, 114:271–323, 1986.

[2] F. Blanchard. β-expansions and symbolic dynamics. *Theoret. Comput. Sci.*, 65:131–141, 1989.

[3] R. Bowen. Entropy for group endomorphisms and homogeneous spaces. *Transactions of the American Mathematical Society*, 153:401–414, 1971.

[4] R. Bowen. Periodic points and measures for Axiom A diffeomorphisms. *Trans. Amer. Math. Soc.*, 154:377–397, 1971.

[5] R. Bowen. Entropy-expansive maps. *Trans. Amer. Math. Soc.*, 164:323–331, 1972.

[6] R. Bowen. Symbolic dynamics for hyperbolic flows. *Amer. J. Math.*, 95:429–460, 1973.

[7] M. Boyle and T. Downarowicz. The entropy theory of symbolic extensions. *Invent. Math.*, 156:119–161, 2004.

[8] M. Brin and A. Katok. On local entropy. In *Geometric dynamics (Rio de Janeiro, 1981)*, volume 1007 of *Lecture Notes in Math.*, pages 30–38. Springer, Berlin, 1983.

[9] D. Burguet. A direct proof of the tail variational principle and its extension to maps. *Ergodic Theory Dynam. Sys.*, 29:357–369, 2009.

[10] R. Burton, M. Keane, and J. Serafin. Residuality of dynamical morphisms. *Colloq. Math.*, 85:307–317, 2000.

[11] R. Burton and A. Rothstein. Isomorphism theorems in ergodic theory. Technical report, Oregon State University, 1977.

[12] V. Climenhaga and D. J. Thompson. Intrinsic ergodicity beyond specification: beta-shifts, S-gap shifts, and their factors. *Israel Journal of Mathematics, (to appear)*, 2010.

[13] M. Dateyama. The almost weak specification property for ergodic group automorphisms of abelian groups. *J. Math. Soc. Japan*, 42:341–351, 1990.

[14] M. Denker, C. Grillenberger, and K. Sigmund. *Ergodic theory on compact spaces*. Lecture Notes in Mathematics, Vol. 527. Springer-Verlag, Berlin, 1976.

[15] T. Downarowicz. Entropy structure. *J. Anal. Math.*, 96:57–116, 2005.

[16] T. Downarowicz. Minimal models for noninvertible and not uniquely ergodic systems. *Israel J. Math.*, 156:93–110, 2006.

[17] T. Downarowicz. *Entropy in dynamical systems*, volume 18 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2011.

[18] T. Downarowicz and J. Serafin. A short proof of the Ornstein theorem. *Ergodic Theory and Dynamical Systems*, 32:587–597, 2012.

[19] T. Downarowicz and B. Weiss. Entropy theorems along times when x visits a set. *Illinois J. Math.*, 48:59–6, 2004.

[20] R. Dudley. *Real analysis and probability*. Wadsworth, Pacific Grove, CA, 1989.

[21] R. Engelking. *Dimension theory*. North-Holland Publishing Co., Amsterdam, 1978.

[22] E. Glasner. *Ergodic theory via joinings*, volume 101 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2003.

[23] P. Hall. On representatives of subsets. *J. London Math. Soc.(1)*, 10:26–30, 1935.

[24] P. R. Halmos. On automorphisms of compact groups. *Bull. Amer. Math. Soc.*, 49:619–624, 1943.

[25] P. R. Halmos. *Lectures on ergodic theory*. Chelsea, 1956.

[26] S. Kalikow and R. McCutcheon. *An outline of ergodic theory*, volume 122 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[27] A. Katok. Nonuniform hyperbolicity and structure of smooth dynamical systems. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Warsaw, 1983)*, pages 1245–1253, Warsaw, 1984. PWN.

[28] M. Keane and M. Smorodinsky. A class of finitary codes. *Israel J. Math.*, 26:352–371, 1977.

[29] M. Keane and M. Smorodinsky. Bernoulli schemes of the same entropy are finitarily isomorphic. *Ann. of Math. (2)*, 109:397–406, 1979.

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

[31] W. Krieger. Erratum to: "On entropy and generators of measure-preserving transformations". *Trans. Amer. Math. Soc.*, 168:519, 1972.

[32] W. Krieger. On unique ergodicity. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, pages 327–346, Berkeley, Calif., 1972. Univ. California Press.

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

[34] J. Kulesza. Zero-dimensional covers of finite dimensional dynamical systems. *Ergodic Theory Dynam. Systems*, 15:939–950, 1995.

[35] D. Lind. Ergodic group automorphisms and specification. In M. Denker and K. Jacobs, editors, *Ergodic Theory*, volume 729 of *Lecture Notes in Mathematics*, pages 93–104. Springer Berlin / Heidelberg, 1979.

[36] D. Lind. Dynamical properties of quasihyperbolic toral automorphisms. *Ergodic Theory Dynam. Systems*, 2:49–68, 1982.

[37] D. A. Lind and J.-P. Thouvenot. Measure-preserving homeomorphisms of the torus represent all finite entropy ergodic transformations. *Math. Systems Theory*, 11:275–282, 1977.

[38] E. Lindenstrauss. Mean dimension, small entropy factors and an embedding theorem. *Pub. Math. I. H. E. S.*, 89:227–262, 1999.

[39] B. Marcus. A note on periodic points for ergodic toral automorphisms. *Monatsh. Math.*, 89:121–129, 1980.

[40] R. Miles. Periodic points of endomorphisms on solenoids and related groups. *Bull. Lond. Math. Soc.*, 40:696–704, 2008.

[41] M. Misiurewicz. Topological conditional entropy. *Studia Math.*, 55:175–200, 1976.

[42] D. Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Advances in Math.*, 4:337–352 (1970), 1970.

[43] D. S. Ornstein and B. Weiss. Entropy and data compression schemes. *IEEE Trans. Inform. Theory*, 39:78–83, 1993.

[44] C.-E. Pfister and W. G. Sullivan. On the topological entropy of saturated sets. *Ergodic Theory Dynam. Systems*, 27:929–956, 2007.

[45] A. Quas and T. Soo. Weak mixing suspension flows over shifts of finite type are universal. *Journal of Modern Dynamics, to appear*, 2011.

[46] M. Ratner. Markov partitions for Anosov flows on n-dimensional manifolds. *Israel J. Math.*, 15:92–114, 1973.

[47] A. Rényi. Representations for real numbers and their ergodic properties. *Acta Math. Acad. Sci. Hungar*, 8:477–493, 1957.

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

[49] T. de la Rue. An introduction to joinings in ergodic theory. *Discrete Contin. Dyn. Syst.*, 15:121–142, 2006.

[50] J. Schmeling. Symbolic dynamics for β-shifts and self-normal numbers. *Ergodic Theory Dynam. Systems*, 17:675–694, 1997.

[51] K. Sigmund. On dynamical systems with the specification property. *Trans. Amer. Math. Soc.*, 190:285–299, 1974.

[52] J. G. Sinaĭ. On a weak isomorphism of transformations with invariant measure. *Mat. Sb. (N.S.)*, 63 (105):23–42, 1964.

[53] Y. G. Sinai. *Selecta. Volume I. Ergodic theory and dynamical systems.* Springer, New York, 2010.

[54] S. M. Srivastava. *A course on Borel sets*, volume 180 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.

[55] P. Sun. Density of metric entropies for linear toral automorphisms. *Dynamical Systems*, 27:197–204, 2012.

[56] S. Williams. Notes on renewal systems. *Proc. Amer. Math. Soc.*, 110:851–853, 1990.

[57] K. Yamamoto. On the weaker forms of the specification property and their applications. *Proc. Amer. Math. Soc.*, 137:3807–3814, 2009.

(A. Quas and T. Soo) DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF VICTORIA, PO BOX 3060 STN CSC, VICTORIA, BC V8W 3R4, CANADA

E-mail address: `aquas(at)uvic.ca`

URL: `http://www.math.uvic.ca/∼aquas/`

E-mail address: `tsoo(at)uvic.ca`

URL: `http://www.math.uvic.ca/∼tsoo/`