UNIQUELY ERGODIC MINIMAL TILING SPACES WITH POSITIVE ENTROPY

IAN PALMER AND JEAN BELLISSARD

Abstract. Strictly ergodic spaces of tilings with positive entropy are constructed using tools from information and probability theory. Statistical estimates are made to create a one-dimensional subshift with these dynamical properties, yielding a space of repetitive tilings of $\mathbb{R}^D$ with finite local complexity that is also equivalent to a symbolic dynamical system with a $\mathbb{Z}^D$ action.

1. Introduction

A method for constructing uniquely ergodic minimal (also known as strictly ergodic) tiling spaces with positive entropy is demonstrated:

Theorem 1. *For every positive integer D there are repetitive aperiodic tilings T of $\mathbb{R}^D$ with finite local complexity such that, together with an $\mathbb{R}^D$-transation action, the hull Ω of T is a uniquely ergodic minimal dynamical system with positive configurational entropy.*

1.1. **Tiling Spaces.** This result was motivated by recent developments concerning the topology of tiling spaces. For the basic properties of tiling spaces, good references are [28, 29, 34]. Inspired by the description of aperiodic solids through the formalism of Noncommutative Geometry [2, 3], Kellendonk [21] realized that a similar approach applies to tilings in $\mathbb{R}^D$. Soon after, Anderson and Putnam [1] showed that substitution tiling spaces can be seen as an inverse limit of CW-complexes. Independently, Gambaudo and Martens [9] developed a similar method for dynamical systems leading to the construction of a large class of uniquely ergodic dynamical systems with positive entropy. The present paper was motivated by this latter result. Indeed,

Date: June 2009.

Work supported by NSF grant DMS-0600956 and CRC 701 of the DFG (German Research Council).

the Anderson-Putnam construction was extended to arbitrary tiling spaces which are aperiodic, repetitive and have finite local complexity [4], so the construction of tilings in this class with unique ergodic measure and positive configurational entropy was a natural question to address. Tiling space theory provides a convenient conceptual bridge between the theory of symbolic dynamics and more general dynamical systems; the present work considers tilings of $\mathbb{R}^D$ by cubic tiles, colored by a label taken from a finite alphabet. The resulting tiling space supports a minimal, uniquely ergodic $\mathbb{Z}^D$-action with positive entropy. Furthermore, the construction is sufficiently general to allow the following extension of Krieger's Theorem[25] to tiling spaces:

Theorem 2. *Let the tilings consist of cubes colored by labels from a finite alphabet $\mathcal{A}$. For a positive probability distribution μ on the alphabet $\mathcal{A}$, the tiling space can be constructed such that i) the occurrence probability of each letter a is arbitrarily close to $\mu(a)$, and ii) its configurational entropy is arbitrarily close to the Shannon entropy of μ.*

1.2. **Sub-Shifts.** The method of construction reduces the problem to constructing a sub-shift. Given a finite alphabet $\mathcal{A}$, the set $\mathcal{A}^{\mathbb{Z}}$ of doubly infinite sequences admits a dynamics given by the shift S. Sub-shifts are produced from S-invariant subsets, either measurable, in the context of ergodic theory, or closed, if the topology is emphasized, with an invariant ergodic probability measure on it. The same line of reasoning has been followed in the past to obtain seminal results on unique ergodicity. The original reference, inspired by the Birkhoff theorem and the von Neumann approach to ergodicity, is an early work of Bogolioubov and Kryloff [26] in which ergodicity was systematically studied. In particular, this work gives the first proof that a topological dynamical system on a compact space is uniquely ergodic if and only if the Birkhoff sums converge uniformly. A classical review can be found in [31]. In the late sixties several strictly ergodic sub-shifts were constructed by Kakutani [19], Keane [20], and Jacobs and Keane [17]. More recent results in substitution dynamical systems are found in the book [32], and some criteria for strict ergodicity of tiling spaces are found in [4]. The first example of a weakly mixing strictly ergodic system was produced by Jacobs [15]. Hahn and Katznelson [13] constructed strictly ergodic systems with arbitrarily high entropy, and Grillenberger [12] constructed strictly ergodic sub-shifts with entropy arbitrarily close to a prescribed value. Jewett [18] and Krieger [25] then showed that any ergodic measure preserving transformation T of a Lebesgue measure space with finite entropy $h(T)$ is isomorphic, by bi-measurable isomorphisms, to a uniquely ergodic homeomorphism of the Cantor set (see also the more recent work [33]). Similar results where obtained for flows [16]. Finally, Cortez has constructed a uniquely ergodic $\mathbb{Z}^2$-Toeplitz system with positive entropy [6].

1.3. **Probability and Information.** The present method of construction derives from the information theoretic formalism introduced by Shannon [35]. Given a probability measure μ on the finite alphabet $\mathcal{A}$, the set of words of length n with empirical distribution close to μ can be chosen so that (i) it has a large probability with respect to the product measure $\mu^{\otimes n}$, and (ii) the subset of probability measures given μ for the occurrence probability of the various letters is dramatically reduced in size as $n \to \infty$. Thanks to (i) the entropy per letter of such measure is not too far from the entropy of the original measure μ. This argument, together with an iteration process, leads to a sub-shift with configurational entropy close to the entropy of μ. It is remarkable that this line of argument can be found in Krieger's work, with the restriction that μ be the uniform distribution on $\mathcal{A}$ and without any reference to Information Theory. The present work, however, is closer in spirit to the Central Limit Theorem on Bernoulli processes. This line of ideas is almost as old as probability theory itself. The posthumous book by Bernoulli [5], published in 1713, defined the framework for probabilities on finite sets and proved the Law of Large Numbers. Twenty years later, De Moivre [8] gave the first proof of the Central Limit Theorem. This theorem was later proved in a more general setting in the book by Laplace [30], using an argument developed by Gauss [10, 11]. Shannon followed the same line of argument to define the information content of a finite probability space leading to his explicit formula for the entropy [35]. The Russian school led by Kolmogoroff was also inspired by information theory, in particular for the definition of the Kolmogoroff-Sinai entropy [23, 24, 36]. The information theoretic point of view is that the shift operator, interpreted as a time evolution in ergodic theory, is rather a sampling process: a statistical definition of the probability distribution μ is given by a limit of an infinite number of trials, without reference to time.

Acknowledgements. Both authors thank the CRC 701 in Bielefeld for financial support during the completion of this work. I.P. wishes to thank A. Grigo for several helpful conversations. Both authors are indebted to P. Arnoux, J. Kellendonk and L. Sadun for comments and useful new references.

2. Estimates from Probability and Information Theory

An excellent reference for this section is the book by Cover and Thomas [7], especially Chapter 11. Given a finite set $\mathcal{A}$ of cardinality $|\mathcal{A}|$, let $\mathcal{M}_1(A)$ denote the space of probabilities on $\mathcal{A}$. Fix an integer word length $l > 1$ and let $\mathcal{A}^l$ denote the set of words w of length l with letters $w_i \in \mathcal{A}, i = 1, \ldots, l$. For $w \in \mathcal{A}^l$, the *empirical distribution* $\pi^w \in \mathcal{M}_1(\mathcal{A})$ of w gives the occurrence frequencies of the letters in $\mathcal{A}$:

$$\pi^w(a) := \frac{1}{l}\sum_{i=1}^{l} \delta_{aw_i}$$

There are three related entropies to be considered for measures on $\mathcal{A}$:

(Shannon) Entropy	$S(\mu) := -\sum_{a\in\mathcal{A}} \mu(a)\,\mathrm{Log}\,(\mu(a))$
Relative Entropy	$D_{KL}(\mu,\nu) := \sum_{a\in\mathcal{A}} \mu(a)\,\mathrm{Log}\left(\frac{\mu(a)}{\nu(a)}\right)$
Cross Entropy	$S(\mu,\nu) := -\sum_{a\in\mathcal{A}} \mu(a)\,\mathrm{Log}\,(\nu(a))$

It follows from the definitions that $S(\mu,\nu) = S(\mu) + D_{KL}(\mu,\nu)$. The distance between two measures $\nu, \nu' \in \mathcal{M}_1(X)$ is

$$d(\nu,\nu') := \sum_{x\in X} |\nu(x) - \nu'(x)|$$

Lemma 1. *Let $\epsilon \leq \min_{a\in\mathcal{A}} \mu(a)$. For $\mu,\nu \in \mathcal{M}_1(\mathcal{A})$, if $d(\mu,\nu) < \epsilon$, then $|S(\mu,\nu) - S(\nu)| < \epsilon\, Log\left(\frac{1}{\epsilon}\right)$.*

Proof.

$$\begin{aligned} |S(\mu,\nu) - S(\nu)| &= \left| -\sum_{a\in\mathcal{A}} [\mu(a) - \nu(a)]\,\mathrm{Log}\,(\nu(a)) \right| \\ &\leq \max_{a\in\mathcal{A}} \{-\mathrm{Log}\,(\nu(a))\} \sum_{a\in\mathcal{A}} |\mu(a) - \nu(a)| \leq \epsilon\,\mathrm{Log}\left(\frac{1}{\epsilon}\right) \end{aligned}$$

□

The property of having the same empirical distribution is an equivalence relation on the words in $\mathcal{A}^l$. As is well known (see [7]), the number of equivalence classes is bounded by a polynomial in the word length l (of degree $|\mathcal{A}|$), while the size of a class is exponential in l:

$$|[w]| = \mu^l([w])e^{lS(\pi^w,\mu)} \quad \forall \mu \in \mathcal{M}_1(\mathcal{A}) \tag{1}$$

Let W be the set of words whose empirical distributions are ϵ-close to a given probability $\mu \in \mathcal{M}_1(\mathcal{A})$:

$$W := \{w \mid d(\pi^w, \mu) < \epsilon\}$$

With (1) and Lemma 1 it is now possible to give an estimate on the number of words in W:

Lemma 2.

$$e^{lS(\mu)-\epsilon\, Log\left(\frac{1}{\epsilon}\right)} \leq |W| \leq e^{lS(\mu)+\epsilon\, Log\left(\frac{1}{\epsilon}\right)}$$

Proof.

$$\left|\frac{1}{l}\operatorname{Log}(|W|) - S(\mu)\right| = \left|\frac{1}{l}\operatorname{Log}\left(\sum_{[w]\in\mathcal{N}(\mu,\epsilon)} \mu^l([w])e^{l[S(\pi^w,\mu)-S(\mu)]}\right)\right|$$
$$\leq \epsilon \operatorname{Log}\left(\frac{1}{\epsilon}\right) + \frac{1}{l}\operatorname{Log}\left(\mu^l(W)\right) < \epsilon \operatorname{Log}\left(\frac{1}{\epsilon}\right)$$

□

Let W^C be the complement of W in $\mathcal{A}^l$. The following concentration inequality is a direct result of the Chernoff-Hoeffding Theorem[14]:

Lemma 3.

$$\mu^l(W) > 1 - 2\,|\mathcal{A}|\, e^{-2l\left(\frac{\epsilon}{|\mathcal{A}|}\right)^2}$$

Proof. For $a \in \mathcal{A}$ and $\epsilon > 0$, because the expectation with respect to μ of each δ_{aw_i} is $\mu(a)$, the Chernoff-Hoeffding Theorem gives the bound

$$\mu^l\left\{w \in \mathcal{A}^l \middle| |\pi^w(a) - \mu(a)| \geq \epsilon\right\} < 2e^{-2l\epsilon^2}$$

The result follows, since

$$\left\{w \in \mathcal{A}^l \middle| d(\pi^w, \mu) \geq \epsilon\right\} \subset \bigcup_{a\in\mathcal{A}} \left\{w \in \mathcal{A}^l \middle| |\pi^w(a) - \mu(a)| \geq \frac{\epsilon}{|\mathcal{A}|}\right\}$$

□

For ease of notation, let $\gamma := 2l\left(\frac{\epsilon}{|\mathcal{A}|}\right)^2$.

Lemma 4. *Let $\hat{\mu}$ be the normalized restriction of μ^l to $W \subset \mathcal{A}^l$. As a consequence of the previous lemmas,*

$$d\left(\mu^l, \hat{\mu}\right) < 4e^{-\gamma} \qquad \textit{and} \qquad 0 < lS(\mu) - S(\hat{\mu}) < 4e^{-\gamma}\, Log\left(\frac{1}{\epsilon}\right)$$

Proof. The first inequality is a consequence of Lemma 3:

$$d\left(\mu^l, \hat{\mu}\right) = \sum_{w \in W} \left|\mu^l(w) - \hat{\mu}(w)\right| + \sum_{w \in W^C} \mu^l(w) = 2\left(1 - \mu^l(W)\right) < 4e^{-\gamma}$$

The second inequality comes from combining the technique of Lemma 1 with the previous inequality and the fact that $S(\mu^l) = lS(\mu)$. The difference $S(\mu^l) - S(\hat{\mu})$ is positive because $\hat{\mu}(w)$ is the conditional probability of $\mu^l(w)$ with the condition $d\left(\pi^w, \mu\right) < \epsilon$. Since

$$\left|S(\mu^l, \hat{\mu}) - S(\hat{\mu})\right| \leq \sum_{w \in W} \left|\mu^l(w) - \hat{\mu}(w)\right| \operatorname{Log}\left(\frac{1}{\hat{\mu}(w)}\right) \leq 4e^{-\gamma} \operatorname{Log}\left(\frac{1}{\epsilon}\right)$$

and $S(\hat{\mu}) \leq S(\mu^l) \leq S(\mu^l, \hat{\mu})$ (because $S(\mu) + D_{KL}\left(\mu, \nu\right) = S(\mu, \nu)$, with all quantities positive), the result follows. □

3. Construction of Anderson-Putnam Complex

3.1. **Alphabets and Measures.** The first goal is to construct a sequence of alphabets $\{\mathcal{A}_n\}_{n=0}^{\infty}$ so that each $\mathcal{A}_n$ is a family of words of length l_n with letters in $\mathcal{A}_{n-1}$. Each letter of $\mathcal{A}_{n-1}$ is forced to occur in each word of $\mathcal{A}_n$ by controlling the distributions of letters. Also, only the distributions of letters are controlled; the families of words are invariant under permutations of the letters in a word.

Let $\mathcal{A}_0$ be an initial finite alphabet with $|\mathcal{A}_0| \geq 2$ symbols, and fix a measure $\mu_0 \in \mathcal{M}_1\left(\mathcal{A}_0\right)$ positive on all $a \in \mathcal{A}_0$. For $\beta \in \left(0, \frac{1}{2}\right)$ and $n = 0, \ldots, \infty$ choose an integer

$$l_{n+1} > \max_{a \in \mathcal{A}_n} \left(\mu_n(a)\right)^{-\frac{1}{\beta}} \qquad \text{(Constraint 1)}$$

and $\epsilon_n = l_{n+1}^{-\beta}$. Then

$$\mathcal{A}_{n+1} := \left\{w \in \mathcal{A}_n^{l_{n+1}} \mid d\left(\pi^w, \mu_n\right) < \epsilon_n\right\} \qquad (2)$$

is a family of words in $\mathcal{A}_n^{l_{n+1}}$ whose empirical distributions are close to μ_n; let μ_{n+1} be the normalized product measure $\mu_n^{l_{n+1}}$ after restricting to $\mathcal{A}_{n+1} \subset \mathcal{A}_n^{l_{n+1}}$. With these conditions, $\epsilon_n < \min_{a \in \mathcal{A}_n} \mu_n(a)$, which forces each letter to occur at least once in each word of $\mathcal{A}_{n+1}$ (because $|\pi^w(a) - \mu_n(a)| < \epsilon_n$), and thus $l_{n+1} > |\mathcal{A}_n|$. Let $\Phi_n^m := \Phi_n^{n+1} \circ \cdots \circ \Phi_{m-1}^m$, where

$$\Phi_n^{n+1}\nu := \sum_{w \in \mathcal{A}_n^{l_{n+1}}} \nu(w)\pi^w$$

is a projection $\mathcal{M}_1\left(\mathcal{A}_n^{l_{n+1}}\right) \to \mathcal{M}_1\left(\mathcal{A}_n\right)$

Proposition 1. *For each $n \in \mathbb{N}$, the set $\mathcal{A}_n$ is nonempty.*

Proof. Given a collection of $|\mathcal{A}|$ point measures $\{\delta_a\}_{a\in\mathcal{A}}$, the measure with the greatest distance from the δ_a is the uniform distribution $\frac{1}{|\mathcal{A}|}\sum_{a\in\mathcal{A}}\delta_a$; this maximum distance is

$$d\left(\delta_{a'}, \frac{1}{|\mathcal{A}|}\sum_{a\in\mathcal{A}}\delta_a\right) = \sum_{b\in\mathcal{A}}\left|\delta_{a'b} - \frac{1}{|\mathcal{A}|}\sum_{a\in\mathcal{A}}\delta_{ab}\right| = 2\left(1-\frac{1}{|\mathcal{A}|}\right)$$

The family $\{\pi^w \mid w \in \mathcal{A}^l\}$ of empirical distributions over the words in $\mathcal{A}^l$ is a lattice in $\mathcal{M}_1\left(\mathcal{A}^l\right)$, and the distance between neighboring points is $\frac{1}{l}$, so the distance to the center of one of the cells is $\frac{2}{l}\left(1-\frac{1}{|\mathcal{A}|}\right)$. With

$$\epsilon = l^{-\beta} > \frac{2}{l}\left(1-\frac{1}{|\mathcal{A}|}\right) \tag{3}$$

there is guaranteed to be a word $w \in \mathcal{A}^l$ satisfying $d(\pi^w, \mu) < \epsilon$ for any $\mu \in \mathcal{M}_1(\mathcal{A})$. □

Let $L_n := l_n \cdots l_1$ be the number of letters of $\mathcal{A}_0$ in a word of $\mathcal{A}_n$, and, for ease of notation, let $\gamma_n := 2l_{n+1}\left(\frac{\epsilon_n}{|\mathcal{A}_n|}\right)^2$.

Lemma 5. *For each $n \in \mathbb{N}$,*

$$\left|\frac{1}{L_n}\,\mathit{Log}\left(|\mathcal{A}_n|\right) - S(\mu_0)\right| \le \frac{\epsilon_{n-1}}{L_{n-1}}\,\mathit{Log}\left(\frac{1}{\epsilon_{n-1}}\right) + \sum_{k=1}^{n-1}\frac{4e^{-\gamma_{k-1}}}{L_k}\,\mathit{Log}\left(\frac{1}{\epsilon_{k-1}}\right)$$

Proof.

$$\left|\frac{1}{L_n}\operatorname{Log}\left(|\mathcal{A}_n|\right) - S(\mu_0)\right| \le$$
$$\frac{1}{L_{n-1}}\left|\frac{1}{l_n}\operatorname{Log}\left(|\mathcal{A}_n|\right) - S(\mu_{n-1})\right| + \sum_{k=1}^{n-1}\frac{1}{L_k}\left|S(\mu_k) - l_k S(\mu_{k-1})\right|$$

so the result follows by Lemmas 2 and 4. □

3.1.1. *Unique Ergodicity.* An essential ingredient of the present construction is the contraction property of the projections Φ_n^{n+1} (see [9, 4]), demonstrated in the following lemma. The contractions of the Φ_n^{n+1} will cause the space of probability measures to collapse to a point in the inverse limit.

Lemma 6. *Given arbitrary choices of $\nu_m \in \mathcal{M}_1(\mathcal{A}_m)$ for each m, the sequence of measures $\{\Phi_n^m \nu_m\}_{m=n+1}^{\infty}$ in $\mathcal{M}_1(\mathcal{A}_n)$ is Cauchy for any n. Thus there is a unique $\tilde{\nu}_n$ in each $\mathcal{M}_1(\mathcal{A}_n)$ such that*

$$\lim_{m\to\infty} \Phi_n^m \nu_m = \tilde{\mu}_n$$

and, furthermore,

$$\Phi_n^{n+1} \tilde{\mu}_{n+1} = \tilde{\mu}_n$$

Proof.

$$\begin{aligned} d\left(\Phi_n^{n+1}\nu, \Phi_n^{n+1}\nu'\right) &= \sum_{a\in\mathcal{A}_n} \left| \sum_{w\in\mathcal{A}_{n+1}} \left[\pi^w(a) - \mu(a)\right]\left[\nu(w) - \nu'(w)\right] \right| \\ &\leq \sum_{w\in\mathcal{A}_{n+1}} d\left(\pi^w, \mu\right) \left|\nu(w) - \nu'(w)\right| \leq \epsilon_n d\left(\nu, \nu'\right) \end{aligned}$$

and by induction, given $m' > m > n$,

$$d\left(\Phi_n^m \nu_m, \Phi_n^{m'} \nu_{m'}\right) \leq \prod_{k=n}^{m-1} \epsilon_k \; d\left(\nu_m, \Phi_m^{m'} \nu_{m'}\right) \leq 2 \prod_{k=n}^{m-1} \epsilon_k$$

Since $\epsilon_k = l_{k+1}^{-\beta}$ decays exponentially, $\{\Phi_n^m \nu_m\}_{m>n}$ is Cauchy, converging to a measure $\tilde{\mu}_n$.

Furthermore,

$$\begin{aligned} d\left(\Phi_n^{n+1}\tilde{\mu}_{n+1}, \tilde{\mu}_n\right) &\leq d\left(\Phi_n^{n+1}\tilde{\mu}_{n+1}, \Phi_n^m \nu_m\right) + d\left(\Phi_n^m \nu_m, \tilde{\mu}_n\right) \\ &\leq \epsilon_n d\left(\tilde{\mu}_{n+1}, \Phi_{n+1}^m \nu_m\right) + d\left(\Phi_n^m \nu_m, \tilde{\mu}_n\right) \end{aligned}$$

for all $m > n$ so $\Phi_n^{n+1}\tilde{\mu}_{n+1} = \tilde{\mu}_n$. The choice of the original sequence $\{\nu_m\}_{m=0}^{\infty}$ was arbitrary, so the sequence $\{\tilde{\mu}_n\}_{n=0}^{\infty}$ is unique. □

3.2. **Construction of Tilings.** The system of words $\{\mathcal{A}_n\}_{n=0}^{\infty}$ can be used to label a tiling of $\mathbb{R}^D$ by D-cubic tiles. By fixing an ordering on D-cubic configurations of tiles, a one-to-one correspondence between words of a given length and configurations of a corresponding size is established. The constraints on the word lengths l_n are all lower bounds, so it is possible to impose the additional constraint that for each $n \in \mathbb{N}$,

$$l_n = (2r_n + 1)^D \qquad \text{for some } r_n \in \mathbb{N} \qquad \text{(Constraint 2)}$$

In this way it will be possible to map each word in $\mathcal{A}_n$ to a D-cube of volume $L_n := l_n \cdots l_1$, the number of letters of the initial alphabet $\mathcal{A}_0$ in a word in $\mathcal{A}_n$.

For $n = 0, \dots, \infty$, an *n-patch* is the interior of a D-cube of volume $L_n := l_n \cdots l_1$; each n-patch is the interior of the union of the closures of l_n $(n-1)$-patches, arranged $(2r_n + 1)$ on a side. A 0-patch is also called a *tile*. The location of an n-patch is determined by its center, assumed to be the origin unless stated otherwise, and always an element of the integer lattice $\mathbb{Z}^D$ seen as a subset of $\mathbb{R}^D$. For an n-patch P, the point set $Ł_P$ consists of the centers of the l_n $(n-1)$-patches that comprise P.

An *order* on $Ł_P$ is induced by a bijection $\{1, \dots, l_n\} \to Ł_P$. Fix an order $\sigma_{n,\mathbf{0}}$ on $Ł_P$ for $n \in \mathbb{N}$, and for an n-patch P centered at $\mathbf{x}$, let $\sigma_{n,\mathbf{x}}(i) = \sigma_{n,\mathbf{0}}(i) + \mathbf{x}$ so that the ordering is translation invariant in each generation n.

A *labeled*, or *colored*, n-patch is a pair (P, w), where $w \in \mathcal{A}_n$. If $w = w_1 \dots w_{l_n}$, the $(n-1)$-patch at $\sigma_{n,\mathbf{0}}(i)$ is labeled by w_i. By recursion, this assigns a label in $\mathcal{A}_0$ to each of the L_n tiles in P. Let $\mathcal{P}_n$ be the collection of all labeled n-patches, which is in one-to-one correspondence with $\mathcal{A}_n$. For $n = 0, \dots, \infty$, since $\mathcal{A}_n$ is invariant under permutations of the letters in a word, $\mathcal{P}_n$ is invariant under permutations of the $(n-1)$-patches in an n-patch.

To ensure that a tiling formed from the system of patches $\{\mathcal{P}_n\}_{n=0}^{\infty}$ forces its border [21], it is sufficient to collar the tiles. The *geometric base* of an n-patch P is the interior of the union of the closures of the tiles in P, and the *geometrical boundary* of a patch is the boundary of its geometric center. For any patch $P \in \mathcal{P}_n$, the geometric base is restricted to that of the $2r_n - 1$ $(n-1)$-patches at the center of P; the label w on P now also determines the $(n-1)$-patches adjacent to P. Thus the labeled patch P now represents a collared patch—a patch whose nearest neighbors are known.

The *boundary* ∂P of an n-patch P can be decomposed into a disjoint union of codimension-k boundary components, each of which is labeled by the 2^k n-patches adjacent to it. For $\mathbf{x} \in \partial P$, $\mathbf{y} \in \partial P'$, let $\mathbf{x} \sim_n \mathbf{y}$ when $\mathbf{x}$ and $\mathbf{y}$ occupy the same position in a boundary component with the same 2^k labels.

Following the philosophy of the Anderson-Putnam complex as presented in [4], for $n \in \mathbb{N}$, the branched manifolds $\mathcal{B}_n$ are quotients of the disjoint union of the collared, colored n-patches in $\mathcal{P}_n$ by the relation $\sim_n$. Because $\mathcal{P}_n$ is invariant under permutations of the $(n-1)$-patches in an n-patch, each $(n-1)$-patch shares each of its boundary components with every other $(n-1)$ patch in $\mathcal{P}_{n-1}$ in some n-patch in $\mathcal{P}_n$.

The projections $\varphi^n_{n-1}\colon \mathcal{B}_n \to \mathcal{B}_{n-1}$ of the Anderson-Putnam complex are determined by the labeling of n-patches by $(n-1)$-patches. The singular locus of $\mathcal{B}_n$ is the set of points belonging to more than one n-patch; the *regions* of $\mathcal{B}_n$ are the connected components of $\mathcal{B}_n$ after removing the singular locus, one for each collared, colored n-patch. Each region is tiled by $(n-1)$-patches, and the image under φ^n_{n-1} of a point $\mathbf{x}$ in a region is determined by its location in the $(n-1)$-patch that contains it. Points in the boundaries of multiple $(n-1)$-patches are sent to the corresponding points in the singular locus of $\mathcal{B}_{n-1}$. The tiling space Ω is the inverse limit $\varprojlim(\mathcal{B}_n, \varphi^n_{n-1})$.

4. Properties of the Tiling Space

The proof of the theorem is completed by establishing the properties of Ω. For any $T \in \Omega$, because there are finitely many tiles of a given size and they always meet full-face to full-face, T has FLC. For all $n \in \mathbb{N}$, every $(n-1)$-patch occurs in every n-patch, so T is also repetitive, implying minimality of the dynamical system $(\Omega, \mathbb{R}^D)$. It follows that Ω is the closure of the orbit of a repetitive FLC tiling under the $\mathbb{R}^D$-action, so every tiling in Ω is aperiodic if any one is [22]. To construct an aperiodic tiling in Ω it is sufficient to choose a sequence of words $\{w^{(n)}\}_{n=0}^{\infty}$, with $w^{(n)} \in \mathcal{A}_n$ so that $w^{(n)}_i = w^{(n-1)}$ for i such that $\sigma_{n,\mathbf{0}}(i) = \mathbf{0}$, choosing each $w^{(n)}$ so that the n-patches do not allow translation symmetry.

4.1. **Unique Ergodicity.** In [4] it was shown that the top dimension singular homology $H_D(\mathcal{B}_n, \mathbb{R})$ has a canonical orientation, and that the positive weights on the $\mathcal{B}_n$ are the elements of the positive cone in $H_D(\mathcal{B}_n, \mathbb{R})$. Furthermore, with maps $(\varphi^{n+1}_n)_* : H_D(\mathcal{B}_n, \mathbb{R}) \to H_D(\mathcal{B}_n, \mathbb{R})$ induced by the projections $\varphi^{n+1}_n : \mathcal{B}_{n+1} \to \mathcal{B}_n$, the invariant measures on the transversal Ξ of Ω are determined by the elements of the inverse limit $\varprojlim\,(H_D(\mathcal{B}_n, \mathbb{R}), (\varphi^{n+1}_n)_*)$.

A positive weight on $\mathcal{B}_n$ is a distribution on the regions of $\mathcal{B}_n$ satisfying Kirchhoff's Law. This is equivalent to the condition that, given a face of codimension-1, the sum of the weights on the region on one side of the face be equal to the sum of the weights on the region on the opposite side. Due to the invariance of the $\mathcal{P}_n$ under permutations of the $(n-1)$-patches in an n-patch, for any measure in $\mathcal{M}_1(\mathcal{A}_n) = \mathcal{M}_1(\mathcal{P}_n)$, either sum is equal to a conditional probability given by the product of the measure on the two regions.

Thus the weights of mass one on $\mathcal{B}_n$ are given by the probabilities on $\mathcal{P}_n$, and the projections Φ_n^{n+1} defined on the $\mathcal{A}_n$ are the same as the induced maps $(\varphi_n^{n+1})_*$. By lemma 6 there is a unique sequence of measures $\tilde{\mu}_n$ on the $\mathcal{A}_n$ such that $\Phi_n^{n+1}\tilde{\mu}_{n+1} = \tilde{\mu}_n$ for $n = 0, \ldots \infty$, determining the unique invariant measure on Ξ.

4.2. **Entropy.**

Proposition 2. *The configurational entropy of $T \in \Omega$ is positive and can be made arbitrarily close to $S(\mu_0)$.*

Proof. Let $h_n := \frac{1}{L_n}\operatorname{Log}(|\mathcal{A}_n|)$. Since $\mathcal{A}_n$ is a proper subset of $\mathcal{A}_{n-1}^{l_n}$, $\{h_n\}_{n\in\mathbb{N}}$ is a decreasing sequence of positive numbers. By Lemma 5, if $h_\infty := \lim_{n\to\infty} h_n$, then

$$|h_\infty - S(\mu_0)| \leq \sum_{n=1}^{\infty} \frac{4e^{-\gamma_{k-1}}}{L_k} \operatorname{Log}\left(\frac{1}{\epsilon_{k-1}}\right)$$

so $|h_\infty - S(\mu_0)| < \sum_{n=1}^{\infty} \frac{1}{L_n}$. Since

$$S(\mu_0) = \operatorname{Log}\left(e^{S(\mu_0)}\right) = \sum_{n=1}^{\infty} \frac{1}{n}\left(1 - e^{-S(\mu_0)}\right)^n$$

$\sum_{n=1}^{\infty} \frac{1}{L_n} < S(\mu_0)$ provided $l_n > \frac{n}{(n-1)(1-e^{-S(\mu_0)})}$, which is trivial since l_n grows exponentially in n. Thus $h_\infty > 0$, and by choosing the sequence of word lengths l_n to be larger, the difference between $S(\mu_0)$ and the configurational entropy can be made arbitrarily small. □

Remark: For $n = 0, \ldots, \infty$, the distance between μ_n and the product measure $\mu_{n-1}^{l_n}$ is small and is controlled by the parameter β. By Lemma 4, $d(\tilde{\mu}_n, \mu_n) \leq \sum_{m>n} 4e^{-\gamma_m}$, with $\gamma_m = \frac{2}{|\mathcal{A}_m|^2} l_m^{1-2\beta}$; just as the parameter δ can be adjusted to control the growth rate of the words, β can be adjusted to control the deviation of $\tilde{\mu}_n$ from μ_n. Since Ω is uniquely ergodic, the metric entropy of $\tilde{\mu}$ is equal to the configurational entropy and hence the difference between these and the entropy of the initial choice of measure μ_0 can be made arbitrarily small.

References

[1] Anderson, J.E., Putnam, I.F.: Topological invariants for substitution tilings and their associated C^*-algebras. *Ergod. Th. & Dynam. Sys.*, **18** 1998, 509–537.

[2] Bellissard, J.: K-Theory of C^*-algebras in Solid State Physics. in *Statistical Mechanics and Field Theory, Mathematical Aspects*, T.C. Dorlas, M.N. Hugenholtz & M. Winnink, *Lecture Notes in Physics*, **257** 1986, pp. 99–156.

[3] Bellissard, J.: Gap Labelling Theorems for Schrödinger Operators. in *From Number Theory to Physics*, J.M. Luck, P. Moussa & M. Waldschmidt Eds., Les Houches March 89, Springer, 1993, pp. 538–630.

[4] Bellissard, J., Benedetti, R., Gambaudo, J.-M.: Spaces of tilings, finite telescopic approximations and gap-labeling. *Communications in Mathematical Physics*, **261**(3) 2006, pp. 1–41.

[5] Bernoulli, J.: *Ars Conjectandi.*

[6] Cortez, M.I.: $\mathbb{Z}^d$-Toeplitz arrays. *Discrete Contin. Dyn. Syst.* **15**(3) 2006, pp. 859–881.

[7] Cover, T.M., Thomas, J.A.: *Elements of Information Theory*, 2nd ed., Wiley, 2006.

[8] De Moivre, A.: *The Doctrine of Chance*, 3rd ed., 1756. (This result was first written in latin in 1733.)

[9] Gambaudo, J.-M., Martens, M.: Algebraic topology for minimal cantor sets. *Annales Henri Poincare*, **7**(3) 2006, pp. 423–446. (*A preprint was available in 1999.*)

[10] Gauss, C.F.: Disquisitio de elementis ellipticis Palladis. *Commentationes Societatis Regiae Scientarium Gottingensis Recentiores*, **1**, 1810. Reprinted in [11] Vol.6, pp. 1–64.

[11] Gauss, C.F.: *Werke*, Königlichen Gesellschaft der Wissenschaften zu Göttingen, 1870–1928.

[12] Grillenberger C.: Constructions of strictly ergodic systems, I: Given entropy. *Z.Wahrscheinlichkeitstheorie verw. Geb.*, **25** 1973, pp. 323–334.

[13] Hahn, F., Katznelson, Y.: On the entropy of uniquely ergodic transformations. *Transactions of the American Mathematical Society*, **126**(2) Feb. 1967, pp. 335–360.

[14] Hoeffding, W.: Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, **58**(301) Mar. 1963, pp. 13–30.

[15] Jacobs, K.: Systèmes dynamiques Riemanniens. (French) *Czechoslovak Math. J.*, **20**(95) 1970, pp. 628–631.

[16] Jacobs, K.: Lipschitz functions and the prevalence of strict ergodicity for continuous-time flows. *Contributions to Ergodic Theory and Probability (Proc. Conf., Ohio State Univ., Columbus, Ohio, 1970), Springer, Berlin* 1970, pp. 87–124.

[17] Jacobs, K., Keane, M.: $0-1$-sequences of Toeplitz type. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **13** 1969, pp. 123–131.

[18] Jewett, R.I.: The prevalence of uniquely ergodic systems. *J. Math. Mech.*, **19** 1969/1970, pp. 717–729.

[19] Kakutani, S.: Ergodic theory of shift transformations. *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. II:*

Contributions to Probability Theory, Part 2 Univ. California Press, Berkeley, Calif. 1967, pp. 405–414.

[20] Keane, M.: Generalized Morse sequences. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **10** 1968, pp. 335–353.

[21] Kellendonk, J.: Non-commutative geometry of tilings and gap labelling. *Reviews in Mathematical Physics*, **7**(07) 1995, pp. 1133–1180.

[22] Kellendonk, J., Putnam, I.F.: Tilings, C^*-algebras and K-theory. in *Directions in Mathematical Quasicrystals*, CRM Monograph Series, Volume **13**, M.P. Baake, R.V. Moody, Eds., AMS, Providence, RI: 2000, pp. 177–206.

[23] Kolmogorov, A.N.: A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces. (Russian) *Dokl. Akad. Nauk SSSR (N.S.)*, **119** 1958, pp. 861–864.

[24] Kolmogorov, A.N.: Entropy per unit time as a metric invariant of automorphisms. (Russian) *Dokl. Akad. Nauk SSSR*, **124** 1959, pp. 754–755.

[25] Krieger, W.: On unique ergodicity. *Proc. Sixth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1970/1971), Vol. II: Probability theory,* Univ. California Press, Berkeley, Calif., 1972, pp. 327–346.

[26] Kryloff, N., Bogoliouboff, N.: La théorie générale de la mesure dans son application à l'étude des systèmes dynamiques de la mécanique non linéaire. (French) *Ann. of Math.* (2), **38**(1) 1937, pp. 65–113.

[27] Kullback, S., Leibler, R.A.: On information and sufficiency. *Ann. Math. Statist.* **22**(1) 1951, pp. 79–86.

[28] Lagarias, J.C.: Geometric models for quasicrystals I. Delone sets of finite type. *Discrete Comput. Geom.*, **21** 1999, pp. 161–191.

[29] Lagarias, J.C., Pleasants, P.A.B.: Repetitive Delone sets and quasicrystals. *Ergod. Th. & Dynam. Sys.*, **23** 2003, pp. 831–867.

[30] Laplace, P.S.: *Théorie analytique des probabilités*, Courcier Editeur, Paris, 1812. (*Accessible through Google Book Search.* The Central Limit Theorem is proved in Chap. IV.)

[31] Oxtoby, J. C.: Ergodic sets. *Bull. Amer. Math. Soc.*, **58** 1952, pp. 116–136.

[32] Queffelec, M.: *Substitution Dynamical Systems - Spectral Analysis.* Lecture Notes in Mathematics, 1294. Springer 1987.

[33] Rosenthal, A.: Strictly ergodic models for non-invertible transformations. *Israel J. Math.*, **64**(1) 1988, pp. 57–72.

[34] Sadun, L.: *Topology of Tiling Spaces.* University Lecture Series, 46. American Mathematical Society, Providence, RI, 2008.

[35] Shannon, C. E.: A mathematical theory of communication. *The Bell Syst. Tech. J.*, **27** 1948, pp. 379–423 & 623–656.

[36] Sinaĭ, Ja.: On the concept of entropy for a dynamic system. (Russian) *Dokl. Akad. Nauk SSSR*, **124** 1959, pp. 768–771.

Georgia Institute of Technology, School of Mathematics, Atlanta GA 30332-0160