

CHAOTIC ALMOST MINIMAL ACTIONS

VAN CYR, BRYNA KRA, AND SCOTT SCHMIEDING

ABSTRACT. Motivated by Furstenberg’s Theorem on sets in the circle invariant under multiplication by a non-lacunary semigroup, we define a general class of dynamical systems possessing similar topological dynamical properties. We call such systems chaotic almost minimal, reflecting that these systems are chaotic, but in some sense are close to minimal. We study properties of the acting group needed to admit such an action, and show the existence of a chaotic almost minimal \mathbb{Z} -action. We show there exists chaotic almost minimal \mathbb{Z}^d -actions which support multiple distinct nonatomic ergodic probability measures.

1. FURSTENBERG’S THEOREM

As celebrated result of Furstenberg [9] shows that the action of a pair of commuting maps $x \mapsto px \pmod{1}$ and $x \mapsto qx \pmod{1}$ on the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, where $p, q \geq 2$ are multiplicatively independent integers, is highly constrained. Namely, any closed set that is invariant under multiplication by such p and by q is either finite or is all of \mathbb{T} . Here, we explore what dynamical behavior is behind this type of dichotomy.

We focus on two properties: that there is a dense set of points with finite orbit, and that finite orbits are the the only proper closed invariant sets. This leads us to define a faithful and transitive action of a group (or semigroup) on a compact metric space X to be *chaotic almost minimal* if there is a dense set of points with finite orbit and every proper, closed invariant subset of X is finite (see Definition 2.1 for the precise conditions). The motivation for the name is that a transitive system with dense periodic points is chaotic, for example in the sense of [6] (see also [1, 10]), and the condition on the proper, closed invariant subsets is termed almost minimal in the context of algebraic actions by Schmidt [25].

Extending the definition, we call a group (or semigroup) *chaotic almost minimal* if it admits a chaotic almost minimal action on some compact metric space. In this terminology, Furstenberg’s Theorem shows that \mathbb{N}^2 is chaotic almost minimal, and the invertible symbolic cover of this system, such as that used by Rudolph [24], shows that \mathbb{Z}^2 is chaotic almost minimal. Similarly one can show that \mathbb{N}^d and \mathbb{Z}^d are chaotic almost minimal for any $d \geq 2$. While we focus on discrete groups, we can extend these notions to continuous actions, defining an action of a locally compact group G to be *chaotic almost minimal* if there is a dense set of points whose stabilizer is a lattice in G and every proper compact G -invariant subset is the quotient of G by a lattice. In this setting, there are also well-known examples of continuous actions which are chaotic almost minimal. For instance, Ratner’s Theorem [23] implies the action of the unipotent flow on the space $\mathrm{SL}_2(\mathbb{R})/\mathrm{SL}_2(\mathbb{Z})$

BK was partially supported by the National Science Foundation grant DMS-2348315 and SS by the National Science Foundation grant DMS-2247553.

is chaotic almost minimal. By taking the one-point compactification, we obtain a chaotic almost minimal \mathbb{R} -action on a compact space. We note however that for any time $t > 0$, the time t -map of the flow φ_t on $SL_2(\mathbb{R})/SL_2(\mathbb{Z})$ is *not* chaotic almost minimal according to our definition: since there are periodic trajectories of arbitrary length, for any time $t > 0$, there are invariant circles on which the time t map acts by irrational rotation.

Several other examples of chaotic minimal groups have been well studied, including $SL_d(\mathbb{Z})$ for $d \geq 2$ and the automorphism group of a shift of finite type, and we discuss some of these in Section 2.3. All of these examples, and the variety of other groups that we show have chaotic almost minimal actions (such as those following from the results in Section 3.1) are all either higher rank or are complicated in some other way. It is therefore natural to ask about the existence of a \mathbb{Z} -action that is chaotic almost minimal. At the same time, with an eye toward the well-known question of Furstenberg on measures invariant under multiplication by two multiplicatively independent integers, it is also natural to wonder whether being chaotic almost minimal is compatible with having multiple nonatomic, ergodic probability measures.

We show that not only do chaotic almost minimal \mathbb{Z} -systems exist, but that such systems are capable of supporting more than one nonatomic, ergodic probability measure. Specifically, we construct a symbolic chaotic almost minimal \mathbb{Z} -system with (at least) two distinct, nonatomic, ergodic probability measures.

Theorem 1.1. *There exists a topological \mathbb{Z} -system that is chaotic almost minimal. Furthermore, this system supports two distinct nonatomic ergodic probability measures.*

Our definition of CAM systems is strongly motivated by Furstenberg's Theorem on $\times p, \times q$ invariant subsets of the torus. Moreover, we are also motivated here by Furstenberg's question regarding the possible nonatomic ergodic measures for $\times p, \times q$ on the circle. A consequence of Theorem 1.1 is that if the answer to Furstenberg's question is positive, meaning there is a unique nonatomic ergodic measure invariant under a non-lacunary semigroup, then the CAM structure alone is not responsible for this uniqueness. Moreover, in higher ranks the same result holds, and we prove that in fact, there exists \mathbb{Z}^d -CAM systems possessing two distinct nonatomic ergodic measures.

Theorem 1.2. *For every $d \geq 1$, there exists a topological \mathbb{Z}^d -system that is chaotic almost minimal and supports two distinct nonatomic ergodic probability measures.*

As in the one dimensional case of Theorem 1.1, our constructed system is a subshift over the alphabet $\{0, 1\}$.

The details of the construction of the CAM \mathbb{Z} -system are in Section 4.1. We show that the resulting system is chaotic almost minimal in Proposition 4.5, and in Section 4.2 we prove that this system has more than one nonatomic, ergodic measure. In Section 5, we adapt this construction to a \mathbb{Z}^d -action, and show the existence of a d -dimensional subshift that is chaotic almost minimal and supports multiple distinct nonatomic ergodic probability measures.

A property shared by all of the groups that we show are chaotic almost minimal is that they are all residually finite, and in fact it is not hard to show that any group which admits a chaotic action must be residually finite (see Section 3.1).

This leaves open the question of whether every residually finite group gives rise to a chaotic almost minimal action, and we conjecture (see Section 6) that this holds.

We comment that we define chaotic almost minimal systems in terms of invariant subsets and not in terms of orbits of points. As a tool to showing the existence of chaotic almost minimal systems, we often proceed by showing a weaker property, which we call *weakly chaotic almost minimal*, in which we replace the condition on subsets by the condition that all orbits are either finite or dense. While every chaotic almost minimal system is weakly chaotic almost minimal, the converse does not hold and a construction showing that the two notions are not the same is given in Example 2.18. However, in Proposition 2.17 we show these notions are the same for expansive systems, which is why we are able to make use of the weakly chaotic almost minimal property in our constructions. A related question, along with some others and a conjecture, are discussed in Section 6.

2. CAM SYSTEMS

2.1. General definitions and notation. For a group G and compact metric space X , a *topological G -system* (X, T) is an action of G on X by homeomorphisms, meaning a homomorphism $T: G \rightarrow \text{Homeo}(X)$. Given such a system and $g \in G$ we write $T_g: X \rightarrow X$ for the action by the element g on X . We use the same terminology when G is a semigroup, referring to a topological G -system (X, T) , but assuming that the associated actions $T_g: X \rightarrow X$ are continuous maps which are not necessarily invertible. When $G = \mathbb{Z}$ we write simply (X, T) to mean T is a homeomorphism from X to itself.

The system (Y, S) is a (*topological*) *factor* of the system (X, T) if there exists a continuous surjective map $\pi: X \rightarrow Y$ such that $\pi \circ T = S \circ \pi$.

The G -system (X, T) is *topologically transitive* if for all open sets $U, V \subset X$, there is some $g \in G$ such that $T_g U \cap V \neq \emptyset$, and is *point transitive* if there exists some $x \in X$ such that the orbit $\{gx : g \in G\}$ is dense in X .

Some of our examples and constructions involve symbolic systems. For a finite set \mathcal{A} , called the *alphabet*, we consider the compact metric space $\mathcal{A}^{\mathbb{Z}}$, writing $x \in \mathcal{A}^{\mathbb{Z}}$ as $x = (x_n)_{n \in \mathbb{Z}}$, where $\mathcal{A}^{\mathbb{Z}}$ is endowed with the metric $d(x, y) = 2^{-\{\inf |n|: x_n \neq y_n\}}$. The shift $\sigma: \mathcal{A}^{\mathbb{Z}} \rightarrow \mathcal{A}^{\mathbb{Z}}$ is defined by $(\sigma x)_n = x_{n+1}$ for all $n \in \mathbb{Z}$. A *subshift* is a system (X, σ) , where $X \subset \mathcal{A}^{\mathbb{Z}}$ is a compact σ -invariant subset and \mathcal{A} is some finite alphabet.

For $n \geq 0$, a concatenation $w = w_0 \dots w_{n-1} \in \mathcal{A}^n$ is called a *word* w and we say that this word has *length* n . Given a subshift (X, σ) over the alphabet \mathcal{A} and word $w \in \mathcal{A}^n$, the *cylinder set* $C_k(w)$ is the collection of all $x \in X$ such that $x_k = w_0, \dots, x_{k+n-1} = w_{n-1}$. By a cylinder set in (X, σ) we mean a set $C_k(w)$ for some $k \in \mathbb{Z}$ and word w over the respective alphabet. A word w is in the *language* $\mathcal{L}(X)$ of (X, σ) if there is some nonempty cylinder set determined by w . The language $\mathcal{L}(X) = \bigcup_{n \geq 0} \mathcal{L}_n(X)$, where $\mathcal{L}_n(X)$ denotes all of the words of length n in $\mathcal{L}(X)$.

2.2. Definition of CAM systems. We now define our main object of study. All groups and semigroups here are assumed to be discrete.

Definition 2.1. A topological G -system (X, T) is *chaotic almost minimal (CAM)* if all of the following hold:

- (1) The system (X, T) is topologically transitive and the action of T on X is faithful.
- (2) The space X contains a dense set of points whose T -orbit is finite.
- (3) Every proper closed T -invariant subset of X is finite.

When (X, T) satisfies these conditions, we say that (X, T) is a *CAM system*, or shorten this further and say that it is *CAM*. When we want to emphasize the group acting, we call this a *G-CAM system*.

This definition can be rephrased: the system is CAM if it is topologically transitive, has a dense set of points with finite orbit, and the only compact minimal subsystems are finite. We note that for general groups, we can easily construct examples showing that no two of these conditions implies the third. There are various variations of these assumptions possible, and one which is useful for our purposes is a weakening of the third condition to be about points, and not invariant sets (see Section 2.7).

Definition 2.2. The group or semigroup G is (*topologically*) *CAM* if there exists an infinite compact metric space X such that the system (X, G) is a CAM system.

2.3. Classical examples of systems that are CAM. Before proceeding further, we give several examples of CAM systems.

Example 2.3 (Systems invariant under a non-lacunary (semi-) group). Our motivating example is Furstenberg's Theorem [9], which proves that $\times p, \times q$ for multiplicatively independent integers $p, q \geq 2$ acting on the torus is an \mathbb{N}^2 -CAM system. A symbolic \mathbb{Z}^2 -CAM system is given by Rudolph's coding [24] of of this system. By taking the restriction of this system to \mathbb{N}^2 , we also obtain an example of a symbolic \mathbb{N}^2 -CAM system.

Generalizations of this example for actions on higher dimensional tori that are CAM are given in Berend [3]. Another CAM action on higher dimensional tori is given in the next example.

Example 2.4 (The action of $\mathrm{SL}_d(\mathbb{Z})$ on \mathbb{T}^d). We make use of results in the literature to check that the action of $\mathrm{SL}_d(\mathbb{Z})$ on \mathbb{T}^d is CAM. Several papers prove that any infinite set has dense orbit (see for example [11, 20, 7]). It suffices then to check that periodic points are dense, but this is trivial. Namely, if $v \in \mathbb{Q}^d/\mathbb{Z}^d$, write v using a common denominator for all coordinates. Then for any $A \in \mathrm{SL}_d(\mathbb{Z})$ the denominators of the coordinates of Av are again bounded above by the original denominators, and so the orbit of v must be finite. We note that any probability measure invariant under this action is a linear combination of Haar measure and atomic measures (this is a simple case of the deep results in [2, 4]). These results cover more general settings, as shown in [11, 20, 2, 4].

A different type of example arises by considering the symmetries of symbolic systems.

Example 2.5 (The automorphism group of a mixing shift of finite type). Let (X, T) be an infinite mixing \mathbb{Z} -shift of finite type and let $\mathrm{Aut}(X, T)$ denote the group of automorphisms of (X, T) , meaning the group of self-homeomorphisms of X that commute with T . Then the action of $\mathrm{Aut}(X, T)$ on X is CAM. Indeed, the density of periodic points is immediate, as the action of $\mathrm{Aut}(X, T)$ preserves T -periodic points of a given period and T -periodic points are dense in X . If Y

is an infinite, closed $\text{Aut}(X, T)$ -invariant set, then in particular it is T -invariant, and so Y contains an aperiodic point. It follows from [5, Theorem 9.2] that the $\text{Aut}(X, T)$ -orbit of this aperiodic point is dense in X , and so the system is weakly CAM. Since it is expansive, it follows from Proposition 2.17 that the system is CAM. It is interesting to note that it is also proved in [5] that the only nonatomic measure on X invariant under $\text{Aut}(X, T)$ is the measure of maximal entropy for T .

2.4. Basic dynamical properties of CAM systems. For any group (or semi-group) G , the collection of G -CAM systems is closed under factors, orbit equivalence, and flow equivalence, but is not closed under taking products. We verify these and other simple properties of CAM systems.

Proposition 2.6. *If (X, T) is infinite, has dense periodic points, and is point transitive, then X has no isolated points. In particular, if (X, T) is CAM then X is perfect and uncountable.*

Proof. We proceed by contradiction and assume that $x \in X$ is an isolated point. Since the periodic points are dense in X , it follows that x itself is periodic. If $(U_n)_{n \in \mathbb{N}}$ is a countable base of open sets for X , then for each $n \in \mathbb{N}$, the set $V_n = \bigcup_{m > 0} T^{-m}U_n$ is open and dense. The intersection $\bigcap_{n \in \mathbb{N}} V_n$ is a residual set and by transitivity, every point in it has forward dense orbit. In particular, there is some point with a dense forward orbit and so the periodic point x itself has a dense forward orbit, contradicting the assumption that the system is infinite. Thus X has no isolated points. Since X is compact, it follows that it is perfect and uncountable. \square

Proposition 2.7. *A system with dense periodic points that is point transitive admits a Borel invariant probability measure of full support.*

Again, this applies immediately to any CAM system.

Proof. Let $(O_i)_{i \in \mathbb{N}}$ be a collection of disjoint periodic orbits whose union is dense in the space and let $(m_i)_{i \in \mathbb{N}}$ be a sequence of positive real numbers such that $\sum_{i \in \mathbb{N}} m_i = 1$. Setting $\mu = \sum_i \frac{m_i}{|O_i|} (\sum_{x \in O_i} \delta_x)$, where δ_x denotes the Dirac measure at x , we produce such a measure. \square

Proposition 2.8. *Any nonatomic invariant Borel probability measure on a CAM system has full support.*

Proof. As the support of the measure is a closed, invariant set, any nonatomic measure has some infinite closed invariant set. Invariance implies that the measure has full support. \square

Proposition 2.9. *For any factor of a CAM system, if the group action on the factor is faithful, then the factor is also a CAM system. In particular, any infinite factor of a \mathbb{Z} -CAM system is \mathbb{Z} -CAM.*

Proof. Suppose the system (X, T) is CAM and let $\varphi: X \rightarrow Y$ be a factor map onto some system (Y, S) . Clearly S -periodic points are dense in Y , since T -periodic points are dense in X , and (Y, S) is transitive since (X, T) is. If $A \subset Y$ is an infinite closed S -invariant subset, then $\pi^{-1}(A)$ is an infinite closed T -invariant subset of X . Since the system (X, T) is CAM, it follows that $\pi^{-1}(A) = X$ and so we have that $A = Y$. Thus when the action S on Y is faithful, we have that (Y, S) is CAM.

The conclusion about \mathbb{Z} -CAM factors follows from the fact that the action on an infinite factor of a \mathbb{Z} -system is automatically faithful. \square

The next result follows quickly from the definitions (see for example [22] for background).

Proposition 2.10. *Any system that is topologically orbit equivalent to a CAM system is also CAM.*

Proof. The first two properties of Definition 2.1 are immediate. Suppose (X, S) and (Y, T) are topologically orbit equivalent G -systems via an orbit equivalence $h: X \rightarrow Y$. If the system (X, S) is not CAM, then there exists a proper infinite G -invariant closed subset $Z \subset X$. But then $h(Z)$ would be a proper infinite G -invariant closed subset of Y , and so (Y, T) is not CAM. \square

2.5. Many \mathbb{Z} -CAM systems. In section 4 we prove that there exists a \mathbb{Z} -CAM subshift. We use this here to show that the existence of one such subshift implies that CAM \mathbb{Z} -subshifts are abundant in a certain sense. In fact, for any $n \geq 2$, they are dense in the space of infinite totally transitive subsystems of the full shift (X_n, σ_n) with the Hausdorff metric. They are not however a generic set, as [21, Theorem 1.4] shows there is a generic set of minimal subshifts in the space of totally transitive subshifts.

Theorem 2.11. *If (Y, σ_Y) is an infinite mixing shift of finite type, then (Y, σ_Y) contains a CAM subshift.*

Proof. Let (X, σ) be the \mathbb{Z} -CAM shift constructed in Section 4, and note that (X, σ) is contained in the full 2-shift (X_2, σ_2) . Given $m \geq 1$, let $(X^{(m)}, \sigma^{(m)})$ denote the height m discrete tower over (X, σ) , meaning that $X^{(m)} = X \times \{0, \dots, m-1\}$ and

$$\sigma^{(m)}(x, i) = \begin{cases} (x, i+1) \bmod m & \text{if } i < m-1 \\ (\sigma(x), 0) & \text{if } i = m-1. \end{cases}$$

Since (X, σ) is CAM, it is straightforward to check that $(X^{(m)}, \sigma^{(m)})$ is also CAM.

Given an infinite mixing shift of finite type (Y, σ_Y) , we show that there exists $m \geq 1$ such that $(X_2^{(m)}, \sigma_2^{(m)})$ embeds into (Y, σ_Y) . Since (X, σ) is contained in (X_2, σ_2) , once we have this result the theorem follows.

Let $q_n(Z)$ denote the points of least period n in the subshift Z . By Krieger's Embedding Theorem [17], it suffices to show there exists $m \geq 1$ such that both of the following hold:

- (1) $h(\sigma_2^{(m)}) < h(\sigma_Y)$.
- (2) $q_n(X_2^{(m)}) \leq q_n(Y)$ for all $n \geq 1$.

For the entropy condition, note that $h(\sigma_2^{(m)}) = \frac{1}{m}h(\sigma_2)$ for all $m \geq 1$, and so by choosing m sufficiently large we can guarantee that $h(\sigma_2^{(m)}) < h(\sigma_Y)$. Thus it suffices to find m satisfying the periodic point condition.

To show this, we recall some notation and results (see for example [18, Section 10.1]). Let A be a \mathbb{Z}_+ -matrix presenting the shift (Y, σ_Y) . The function q_n satisfies the formula

$$(1) \quad q_n(Y) = \text{tr}_n(A) = \sum_{d|n} \mu\left(\frac{n}{d}\right) \text{tr}(A^d),$$

where μ denotes the Möbius function and tr the trace. Letting $s^\times(A)$ denote the nonzero spectrum of A (excluding eigenvalues of modulus one) and λ_A denote the Perron-Frobenius eigenvalue for A , then for any $mn > 1$, we have

$$\text{tr}_{mn}(A) \geq \lambda_A^{mn} - \frac{\lambda_A^{mn/2+1} - \lambda_A}{\lambda_A - 1} - \sum_{\mu \in s^\times(A), \mu \neq \lambda_A} \frac{|\mu|^{mn+1} - |\mu|}{|\mu| - 1}.$$

Let $S = \max\{|\mu| : \mu \in s^\times(A), \mu \neq \lambda_A\}$ and $K_1 = \frac{\lambda_A}{\lambda_A - 1}$. It follows that

$$\text{tr}_{mn}(A) \geq \lambda_A^{mn} - K_1 \left(\lambda_A^{mn/2} - 1 \right) - K_2 (S^{mn} - 1)$$

for some constant K_2 which does not depend on either m or n . By the Perron-Frobenius Theorem, we have that $S/\lambda_A < 1$ and so

$$\frac{1}{\lambda_A^{mn}} \text{tr}_{mn}(A) \geq 1 - K_1 (\lambda_A^{-mn/2} - \lambda_A^{-mn}) - K_2 \left[\left(\frac{S}{\lambda_A} \right)^{mn} - \lambda_A^{-mn} \right].$$

Combining this with (1), there is some $m_1 \geq 1$ such that for all $m \geq m_1$ and all $n \geq 1$,

$$(2) \quad \frac{1}{\lambda_A^{m_1 n}} q_{m_1 n}(Y) \geq \frac{2}{3}.$$

Considering the full 2-shift (X_2, σ_2) , we have that

$$q_{mn}(X_2^{(m)}) = m q_n(X^{(2)}) \leq m 2^n.$$

Thus for all $m, n \geq 1$

$$\frac{1}{\lambda_A^{mn}} q_{mn}(X_2^{(m)}) \leq m \left(\frac{2}{\lambda_A^n} \right)^n.$$

Taking m_2 such that $\frac{2}{\lambda_A^m} < \frac{1}{3m}$ for all $m \geq m_2$, then for for all $m \geq m_2$ and any $n \geq 1$, we have that

$$\frac{1}{\lambda_A^{mn}} q_{mn}(X_2^{(m)}) \leq m \left(\frac{2}{\lambda_A^n} \right)^n < m \left(\frac{1}{3m} \right)^n < 2/3.$$

Taking $M = \max\{m_1, m_2\}$ and combining this with (2), we have that for all $n \geq 1$,

$$\frac{1}{\lambda_A^{Mn}} q_{Mn}(Y) \geq \frac{2}{3} > \frac{1}{\lambda_A^{Mn}} q_{Mn}(X_2^{(M)}).$$

Thus it follows that $q_{Mn}(Y) \geq q_{Mn}(X_2^{(M)})$, verifying the necessary periodic point condition for the application of Krieger's Embedding Theorem. \square

For a full shift (X_n, σ_n) , let $\mathcal{S}(X_n)$ be the space of all subshifts of X_n , meaning that $\mathcal{S}(X_n)$ is the set of all compact σ_n -invariant subsets of X_n endowed with the Hausdorff metric (more background on these spaces can be found in [21]). We use Theorem 2.11 to show that for any $n \geq 2$, the closure of the set of CAM shifts in $\mathcal{S}(X_n)$ contains all infinite totally transitive subshifts.

Theorem 2.12. *Fix $n \geq 2$ and let (W, σ_W) be an infinite totally transitive subshift in X_n . Then W is a limit of \mathbb{Z} -CAM subshifts in $\mathcal{S}(X_n)$.*

Proof. Fix $m \geq 1$. Since W is infinite and totally transitive, using standard Markov approximation there exists an infinite mixing shift of finite type (Z, σ_Z) such that $d_H(Z, W) < 2^{-m}$ where d_H denotes the Hausdorff distance. This means that Z and W have the same m -languages. Using the proof of [21, Theorem 6.4], it follows that there is an infinite mixing shift of finite type $Y \subset Z$ such that every point in Y contains every word of length m from W . In particular, the systems Y and W have the same m -languages. Applying Theorem 2.11 to the system Y , we have that Y contains a CAM subshift X such that the m -languages of X and W agree, and hence $d_H(X, W) < 2^{-m}$. \square

We note that we can not directly apply Theorem 2.11; while it shows that we can embed a CAM shift into Z , it does not a priori guarantee that the embedded CAM is close in the Hausdorff metric to Z .

2.6. Dynamical restrictions on CAM actions.

Theorem 2.13. *If X is locally connected compact metric space, then X does not support an expansive \mathbb{Z} -CAM action.*

Our argument closely follows the one used by Mañé [19].

Proof. Suppose (X, T) is CAM. Fix some periodic points $x, y \in X$ and choose $\varepsilon > 0$ such that $\bigcup_{i=1}^{\ell} T^i(B_\varepsilon(x))$ is a proper subset of X and $y \notin B_{2\varepsilon}(x)$; such points exist since (X, T) is CAM. Without loss of generality, we can assume that both x and y have period ℓ .

Consider the ε -stable set of x , defined by

$$W_\varepsilon^s(x) = \{y \in X : d(T^n x, T^n y) \leq \varepsilon \text{ for all } n \geq 0\},$$

and let $CW_\varepsilon^s(x)$ denote the connected component of x in W_ε^s . By [12, Proposition C], making $\varepsilon > 0$ smaller if necessary, there exists $\delta > 0$ such that for all $x \in X$, we have $\text{diam}(CW_\varepsilon^s(x)) \geq \delta$.

Given such δ , by [15, Corollary 2.4], there exists $M > 0$ such that for all $m \geq M$, if $\text{diam}(CW_\varepsilon^s(x)) \geq \delta$, then $\text{diam}(T^{-m}(CW_\varepsilon^s(x))) \geq 4\delta$. Choose such an m and without loss of generality, assume δ is sufficiently small such that $B_\delta(y) \cap B_\varepsilon(x) = \emptyset$ and $\bigcup_{i=1}^{\ell} T^i(B_\delta(y))$ is proper in X . Moreover we may assume that ℓ and m are relatively prime. It follows there exists two disjoint connected components A_1^1, A_2^1 of $T^{-m}(CW_\varepsilon^s(x)) \setminus B_{\delta/4}(y)$ for which $\text{diam}(A_i^1) = \delta$ for $i = 1, 2$. We now apply the same procedure to both A_1^1 and A_2^1 to obtain four connected components $\{A_i^2\}_{i=1}^4$ of $T^{-2m}(CW_\varepsilon^s(x)) \setminus B_{\delta/4}(y)$ for which $\text{diam}(A_i^2) = \delta$ for all $1 \leq i \leq 4$. Inductively we continue the procedure to generate, for every $r \geq 1$, 2^r connected components $\{A_i^r\}_{i=1}^{2^r}$ of $T^{-rm}(CW_\varepsilon^s(x)) \setminus B_{\delta/4}(y)$ such that $\text{diam}(A_i^r) = \delta$ for all $1 \leq i \leq 2^r$. Let $C = \bigcap_{r=1}^{\infty} \bigcup_{i=1}^{2^r} T^{mr}(A_i^r)$. Note that $C \subset CW_\varepsilon^s(x)$, is compact, and is uncountable (in fact C is a Cantor set). Since (X, T) is expansive, there are at most countably many T -periodic points, and so we may choose a T -aperiodic point $p \in C$. Since $C \subset CW_\varepsilon^s(x)$, $T^k(C) \subset B_\varepsilon(T^k x)$ for all $k \geq 0$, and it follows that the forward T -orbit of p is contained in $B_\varepsilon(x)$. Since $B_\varepsilon(x) \cap B_\delta(y) = \emptyset$, y is not a limit point of the forward orbit of p . We claim that the backward orbit of p does not accumulate on y . By construction we have that $T^{-km}(C) \cap B_{\delta/4}(y) = \emptyset$ for all $k \geq 1$. Continuity implies we may choose $\gamma > 0$ such that if $d(u, v) < \gamma$ then $\max\{d(T^i(u), T^i(v)) : 1 \leq i \leq \ell m\} < \delta/4$. Now suppose there exists $k \geq 0$ such that $T^{-k}(p) \in B_\gamma(y)$; without loss of generality assume that k is much larger than ℓm . Since ℓ and m are

relatively prime we may choose a, b such that $k = al + bm$ where $1 \leq a \leq m$ and $b \geq 1$. Then $al \leq \ell m$ so $T^{-k+al}(p) \in B_{\delta/4}(T^{al}(y)) = B_{\delta/4}(y)$. But $-k+al = -bm$ so this implies $T^{-bm}(p) \in B_{\delta/4}(y)$; since $b \geq 1$, this is a contradiction and the claim is proven \square

Many existing examples (see Section 2.3) of CAM systems are *algebraic*, meaning that the system consists of the action of a group by automorphisms on some compact group. However, this can not happen for a \mathbb{Z} -action.

Theorem 2.14. *If $\alpha: X \rightarrow X$ is an automorphism of a compact group X , then (X, α) is not a CAM system.*

To prove this, we use several results in the literature. Suppose β is an action of \mathbb{Z}^d by automorphisms on some compact group X . Recall that Schmidt [25, Section 29] defines an action to be *almost minimal* if every closed proper β -invariant subgroup of X is finite. Clearly if β is CAM, then it is almost minimal. (We note that we do not know if the converse of this holds.)

Proof of Theorem 2.14. Without loss of generality, we can assume that $\alpha: X \rightarrow X$ is transitive, and hence ergodic with respect to the Haar measure on X . If α is CAM, then it is almost minimal. It is a consequence of [25, Theorem 29.2] that X is either zero-dimensional, a torus, or a solenoid.

First consider the toral case. If α has no eigenvalues of modulus one, then it is hyperbolic and hence not CAM. But if α has an eigenvalue on the unit circle, then it induces a rotation on some subset, and again it is not CAM. By passing to an inverse limit, the solenoid case follows.

Thus we are left with the case that X is zero dimensional. We follow ideas in Kitchens [16]. Since X is compact and zero dimensional, we can choose an open normal subgroup K containing the identity. Then the cosets of K are disjoint and partition X ; since X is compact, finitely many cosets $\{a_i K\}$ suffice for partitioning X and the subgroup K is closed.

We claim for all $n \in \mathbb{N}$, there exists $x_n \in X$ such that $x_n^{-1}\alpha^n(x_n) \notin K$. We verify the claim by contradiction, supposing that K does not have this property. Note that K is infinite (since it is open and X is perfect) and so there exists $m \in \mathbb{N}$ such that for all $x \in X$ we have $x^{-1}\alpha^m(x) \in K$, and hence $\alpha^m(x) \in xK$. This implies that $\alpha^m(K) \subset K$, and since K is closed, proper and infinite, it follows that α is not CAM, a contradiction. This verifies the claim about the subgroup K .

Since X/K is a finite group, we can code orbits using the partition $\{a_i K\}$ to get a system $((X/K)^\mathbb{Z}, \sigma)$ and factor map $\pi: X \rightarrow X/K$. Let $Y \subset (X/K)^\mathbb{Z}$ be the image of π . Then Y is a compact totally disconnected topological group and (Y, σ) is expansive, and so by Kitchens [16, Theorem 1] the system (Y, σ) is a one-step shift of finite type. Since (X, α) is transitive, the factor (Y, σ) has a point with dense forward orbit. We claim that Y is infinite. Suppose instead that Y is finite, and choose N such that $\sigma^N = \text{Id}$ on Y . Then for all $x \in X$, we have $\sigma^N(xK) = xK$. It follows that $\alpha^N(x)K = xK$, and hence $x^{-1}\alpha^N(x) \in K$. However, we chose K such that for all $n \in \mathbb{N}$ there exists $x_n \in X$ such that $x_n^{-1}\alpha^n(x_n) \notin K$, a contradiction, proving that Y is infinite. Summarizing, we have that Y is an infinite shift of finite type that is point transitive, and hence contains a proper infinite closed σ -subset $Z \subset Y$. Let $A = \pi^{-1}(Z)$. Then A is closed, α -invariant, infinite, and proper, and so (X, α) is not CAM. \square

2.7. Weakly CAM systems.

Definition 2.15. The topological G -system (X, T) is a *weakly CAM system* if all of the following hold:

- (1) The system (X, T) is transitive and the action of T on X is faithful.
- (2) The space X contains a dense set of T -periodic points.
- (3) If $x \in X$, then the orbit of x is either finite or dense in X .

It is immediate that a CAM system is a weakly CAM system (see Proposition 2.17), and to explore the relation between weakly CAM and CAM we use a result that is well known for subshifts, adapting the argument to the general case.

Lemma 2.16. *If (X, T) is infinite and expansive, then it contains an aperiodic point.*

Proof. Choose a sequence of distinct elements in X and pass to a convergent subsequence $x_i \rightarrow x$. If x is aperiodic, we are done and so suppose that x has period p . Since the system is expansive, there are only a finite number of points of period p , and so by removing finitely many terms of the sequence x_i , we can assume that none of the x_i are of period p . Let ε_c be an expansivity constant for (X, T) . Then for each $i \in \mathbb{N}$, there exists $j_i \in \mathbb{N}$ such that $d(T^{j_i}(x_i), T^{j_i+p}(x_i)) > \varepsilon_c$ (if instead for all $k \in \mathbb{N}$ we have that $d(T^k(x_i), T^{k+p}(x_i)) \leq \varepsilon_c$, then $x_i = T^p(x_i)$ contradicting that x_i does not have period p). Passing to a subsequence if necessary, we can assume that all of the j_i are either positive or negative; let us assume they are positive, as the negative case is analogous. Choose for each i the minimal such j_i which satisfies the conditions. Then $d(T^k(x_i), T^{k+p}(x_i)) \leq \varepsilon_c$ for all $0 \leq k < j_i$. Since the sequence x_i consists of distinct element and converges to x , the j_i 's must be unbounded. Choosing a limit point y of the sequence $T^{j_i}(x_i)$, it follows that y is backward asymptotic to a point of period p . If the point y is periodic, then its period would necessarily be p , but by construction, $d(y, T^p(y)) \geq \varepsilon_c$. Thus the point y is not periodic. \square

Proposition 2.17. *If (X, T) is CAM, then it is weakly CAM. Moreover, if (X, T) is expansive and weakly CAM, then it is CAM.*

Proof. If (X, T) is CAM and $x \in X$ has infinite orbit, then the orbit closure of x is an infinite closed T -invariant subset, and hence is all of X .

For the second statement, suppose (X, T) is expansive and weakly CAM. Let $Y \subset X$ be an infinite closed T -invariant subset. Then $(Y, T|_Y)$ is an infinite expansive system, and so by Lemma 2.16, Y contains an aperiodic point. Since (X, T) is weakly CAM, that point has dense orbit under T , and hence $Y = X$. \square

This proposition motivates the following example showing that the two notions are distinct (a related issue is raised in Question 6.5).

Example 2.18. (A weakly CAM but not CAM system) Let (X, T) be a CAM \mathbb{Z} -subshift and assume that (X, T) has a fixed point p . (The existence of a \mathbb{Z} -CAM system is shown in Section 4, and this system has two fixed points.)

We start by defining the space for our new action. Choose a sequence of proper clopen subsets A_n in X such that for every $n \in \mathbb{N}$, both of the following are satisfied:

- (1) A_n contains p .
- (2) A_{n+1} is a proper subset of A_n .

For $n \in \mathbb{N}$, let $\mathcal{U}_n = \{\frac{k}{2^n} : 0 \leq k \leq 2^n - 1\}$ denote the additive group of 2^n -th roots of unity in $S^1 = \mathbb{R}/\mathbb{Z}$. Let $Z = X \times S^1$, and consider the subset Y of Z defined as the set of points $(x, t) \in Z$ such that:

- (3) If $x \notin A_1$, we require that $t = 0$.
- (4) If $x \in A_1$ and n is chosen such that $x \in A_n$ but $x \notin A_{n+1}$, we require $t \in \mathcal{U}_{n+1}$.
- (5) If $x = p$, there is no constraint on t .

Note that any point of the form (p, t) is a limit of points of the form (x, t) with $x \neq p$.

The map $f: Y \rightarrow Y$ is defined as follows.

- (6) If $x \notin A_1$, then $f(x, 0) = (T(x), 0)$.
- (7) For any $t \in S^1$, $f(p, t) = (p, t)$.
- (8) Suppose $x \in A_0$ and let n be such that $x \in A_n$ but $x \notin A_{n+1}$. Consider (x, t) where $t = \frac{k}{2^{n+1}} \in \mathcal{U}_{n+1}$. We then define

$$f(x, t) = \begin{cases} (x, t + \frac{1}{2^{n+1}}) & \text{for } t \neq \frac{2^{n+1}-1}{2^{n+1}} \\ (T(x), 0) & \text{for } t = \frac{2^{n+1}-1}{2^{n+1}}. \end{cases}$$

In other words, points of the form (x, t) where $x \in A_n$ (and n is maximal such) rotate through the 2^{n+1} -st roots of unity in the S^1 term, and at the end of that excursion, map back to X via T .

We claim that the map $f: Y \rightarrow Y$ is a homeomorphism and the system (Y, f) is weakly CAM but not CAM.

To check that f is a homeomorphism, we consider the various cases. If $(x, t) \in Y$ and $x \neq p$, then continuity of f at (x, t) is immediate. For a point (p, t) with $t \neq 0$, consider a sequence $(x_n, t_n) \rightarrow (p, t)$ as $n \rightarrow \infty$. Consider $\{m(n) : x_n \in A_{m(n)} \text{ and } m(n) \text{ is maximal among these choices}\}$ (note that by the nesting property, this set is finite). By construction, $m(n) \rightarrow \infty$ as $n \rightarrow \infty$, and since $t \neq 0$, the sequence t_n is bounded away from 0. Thus $f(x_n, t_n) = (x_n, t_n + \frac{1}{2^{m(n)+1}}) \rightarrow (p, t)$ as $n \rightarrow \infty$ and continuity of f at (p, t) follows. Continuity at $(p, 0)$ is similar.

We now consider the system (Y, f) . Since the set $\{(p, t) : t \in S^1\}$ is proper, infinite, closed, and f -invariant, this system is not CAM. To check that (Y, f) is weakly CAM, note that if x is periodic then (x, t) is also periodic. Since X is CAM, periodic points are dense in X , and so the same holds for (Y, f) . If $x \in X$ is not periodic then since (X, T) is weakly CAM, the orbit of x is dense in X . It follows immediately that the orbit of $(x, 0)$ is also dense in Y .

This example is useful in that it shows two things: that a weakly CAM \mathbb{Z} -system need not be expansive, and that it need not be totally disconnected. A slight variation of the example can be used to produce a weakly CAM \mathbb{Z} -system which is totally disconnected but not expansive.

3. GROUPS THAT ADMIT CAM ACTIONS

3.1. Residually finite. Recall that a group G is *residually finite* if for every element $g \in G$ that is not the identity, there is a group homomorphism from G to a finite group such that g is not mapped to the identity.

Proposition 3.1. *If G is a CAM group, then G is residually finite.*

Proof. Assume that the G -system (X, T) is CAM and suppose that $g \in G$ is not the identity. Since G acts faithfully and the periodic points are dense, there is a periodic point p such that $T_g p \neq p$. It follows that G acts on the finite orbit of p and g acts nontrivially on this orbit. \square

We note that this proof only relies on the faithfulness of the action and the existence of a dense set of periodic points, and not the other properties of a CAM group. However, this clarifies that many types of groups, such as the additive group of rationals or any infinite simple group, are not CAM. However, we do not know if residual finiteness is the only restriction (see Conjecture 6.4).

3.2. Finite index subactions. The main result in this section is showing that a finite index subgroup of a CAM group is also CAM. To show this, we start with a lemma on how orbits act under the action of a CAM system.

Lemma 3.2. *Suppose (X, T) is a G -CAM system and H is finite index normal subgroup of G . Let $\{g_1 H, g_2 H, \dots, g_m H\}$ denote the set of left cosets of H in G . If $Y \subseteq X$ is a compact H -invariant subspace containing an infinite H -orbit, then $\bigcup_{i=1}^m g_i(Y) = X$. Furthermore, the H -periodic points are dense in Y .*

Proof. We first prove that $\bigcup_{i=1}^m g_i(Y) = X$. Choose $y \in Y$ which has an infinite H -orbit. Then the G -orbit of y is infinite, and since the action of G on X is CAM, it follows that the orbit of y is dense in X . Let $x \in X$ and choose a sequence of elements $k_n \in G$ such that $k_n \cdot y \rightarrow x$ as $n \rightarrow \infty$. For each $n \in \mathbb{N}$, there exists $1 \leq i(n) \leq m$ and $h_n \in H$ such that $k_n = g_{i(n)} h_n$. By passing to a subsequence if necessary, we can assume that $i(n) = i$ for all $n \in \mathbb{N}$. It follows that $g_i h_n \cdot y \rightarrow x$ as $n \rightarrow \infty$, and $h_n \cdot y \rightarrow g_i^{-1}(x)$ as $n \rightarrow \infty$. Since Y is closed and H -invariant we have that $g_i^{-1}(x) \in Y$, and so $x \in g_i(Y)$ for some $i \in \{1, \dots, m\}$, proving the first statement.

We next show that the boundary ∂Y is a finite set consisting of G -periodic points (note that ∂Y is not necessarily G -invariant). To check this, let $B = \bigcup_{i=1}^m g_i(\partial Y)$. The set B is closed, and we claim it is G -invariant. Indeed, given $g \in G$, write $g = g_j h$ for some j and $h \in H$. Then

$$\begin{aligned} g(B) &= g\left(\bigcup_{i=1}^m g_i(\partial Y)\right) = \bigcup_{i=1}^m g g_i(\partial Y) = \bigcup_{i=1}^m g_j h g_i(\partial Y) \\ &= \bigcup_{i=1}^m g_j g_i h'(\partial Y) = \bigcup_{i=1}^m g_{\ell(i)} h_{\ell(i)}(\partial Y) = \bigcup_{i=1}^m g_i(\partial Y) = B, \end{aligned}$$

where the third to last equality follows from normality of H in G and the second to last from the fact that for each i there exists a unique $\ell(i)$ such that $g_j g_i = g_{\ell(i)} h_{\ell(i)}$ for some $h_{\ell(i)} \in H$. Since Y is closed, ∂Y is nowhere dense, and hence so is $g_i(\partial Y)$ for $i = 1, \dots, m$. Thus B is a proper subset of X . Therefore, B is a G -invariant subspace of X and since the action of G on X is CAM, it follows that B is finite and only consists of G -periodic points. Since $\partial Y \subset B$, it follows that the same holds for ∂Y .

Since G -periodic points are dense in X , it follows that H -periodic points are dense in the interior $\text{Int}(Y)$ of Y . Since $Y = \text{Int}(Y) \cup \partial Y$ and ∂Y consists of H -periodic points, the second statement follows. \square

Theorem 3.3. *Any finite index subgroup of a CAM group is CAM.*

Proof. Suppose (X, G) is CAM and let H be a finite index subgroup of G . We first show that we can reduce to the case that H is normal in G . Since H has finite index, there exists a subgroup $K \subseteq H$ such that K is finite index and normal in G . Suppose there exists a compact K -invariant subspace $Z \subset X$ such that the system (Z, K) is CAM. Let $\{\alpha_i\}_{i=1}^m$ be a set of left coset representatives of K in H . Setting $Z' = \bigcup_{i=1}^m \alpha_i(Z)$, we have that $Z' \subseteq X$ is H -invariant and compact. We claim that the system (Z', H) is CAM. It is easy to check that periodic points (with respect to the H -action) are dense in Z' , since K -periodic points are dense in Z and each α_i takes a K -periodic point in Z to a K -periodic point (and hence H -periodic point) in $\alpha_i(Z)$.

Now suppose A is a nonempty proper compact H -invariant subset of Z' . It follows that $A' = A \cap \alpha_j(Z) \neq \emptyset$ for some $j \in \{1, \dots, m\}$. Furthermore, A' is K -invariant, since K is normal in G and both A and $\alpha_j(Z)$ are K -invariant. Since K is normal, this implies that $\alpha_j^{-1}(A') \subset Z$ is also K -invariant. Since (Z, K) is CAM, we must have $\alpha_j^{-1}(A') = Z$, and so $\alpha_j(Z) \subset A$. But A is H -invariant and so $Z' \subset A$, and it follows that the system (Z', H) is CAM.

We can thus assume that H is a finite index, normal subgroup in G . Let \mathcal{C}_H denote the set of all compact H -invariant subsets of X which contain some point with an infinite H -orbit. The set \mathcal{C}_H is partially ordered by setting $Y_1 \preceq Y_2$ if $Y_1 \subset Y_2$. Suppose $\{Y_j\}_{j \in J}$ is a totally ordered collection of such subsets and let $\{g_1, \dots, g_m\}$ be a set of left coset representatives for H in G .

Set $Y_\infty = \bigcap_{j \in J} Y_j$. We claim that Y_∞ contains a point with infinite H -orbit. Since H has finite index in G , it suffices to show that Y_∞ contains a point with infinite G -orbit. In fact we show the stronger statement that for any $x \in X$ with infinite G -orbit, we have that Y_∞ contains at least one point from the set $\{g_1^{-1}x, \dots, g_m^{-1}x\}$; since each of these points has infinite G -orbit, the claim follows. Arguing by contradiction, if this does not hold, then for each $1 \leq i \leq m$, there exists j_i such that $g_i^{-1}(x) \notin Y_{j_i}$. We set $J = \max_{1 \leq i \leq m} \{j_i\}$, and note that J is finite since any finite subset of a totally ordered set has a maximum. Then $g_i^{-1}(x) \notin Y_J$ for all $i \in \{1, \dots, m\}$, since $Y_J \subset Y_{j_i}$ for all $i \in \{1, \dots, m\}$. This then implies that $x \notin g_i Y_J$ for all $i \in \{1, \dots, m\}$. However, Y_J belongs to \mathcal{C}_H and hence contains some point with infinite H -orbit. Thus by Lemma 3.2, it follows that $\bigcup_{i=1}^m g_i(Y_J) = X$, a contradiction. In particular, Y_∞ contains a point with infinite H -orbit.

Thus we have that Y_∞ belongs to \mathcal{C}_H and is an upper bound for the collection $\{Y_j\}_{j \in J}$. By Zorn's Lemma, there is a maximal element in $Z \in \mathcal{C}_H$, and this element Z is a compact H -invariant subset of X containing some point with an infinite H -orbit. We claim the action of H on Z is a CAM H -system. If $x \in Z$ has infinite H -orbit, then that orbit is dense, for otherwise its orbit closure would be a proper compact H -invariant subspace of Z containing an infinite H -orbit, contradicting the maximality of Z . Thus if A is a proper compact H -invariant subset of Z , then A consists only of H -periodic points. Consider $A' = \bigcup_{i=1}^m \alpha_i(A) \subset X$. Then A' is compact, nonempty G -invariant set in X , and it is proper since it consists only of G -periodic points. Since (X, G) is CAM it follows that A' is finite and so A is also finite.

Finally, by Lemma 3.2, the H -periodic points are dense in Z . □

We note the following corollary.

Corollary 3.4. *Suppose H is group acting by a CAM action on X and that H is a finite index subgroup of a group G . If there is an action of G on X extending the action of H , then the G action is also CAM.*

Example 3.5 (Free group). Since $\mathrm{SL}_2(\mathbb{Z})$ is CAM (see Example 2.4), Theorem 3.3 implies the free group on two generators is CAM. More generally, the free group on n generators for any $n \geq 3$ is a finite index subgroup of the free group on 2 generators, and so is also CAM.

Another corollary of Theorem 3.3 gives a sharpening of Corollary 3.4.

Corollary 3.6. *Suppose G acts on X by a CAM action and let $H \subset G$ be a finite index normal subgroup. Suppose there exists $x \in X$ whose H -orbit is dense in X . Then the action of H on X is CAM.*

Proof. Let Z denote an H -invariant set in X on which the action of H is CAM, which exists by Theorem 3.3. By Lemma 3.2 we have $\bigcup_{i=1}^m g_i(Z) = X$. Moreover, it is straightforward to check that each $g_i(Z)$ is H -invariant and the action of H on any $g_i(Z)$ is also CAM. Now if $x \in X$ has dense H -orbit we may choose g_i such that $x \in g_i(Z)$, and then $X = g_i(Z)$. \square

It follows immediately from Theorem 3.3 and Example 2.5 that any finite index subgroup of the automorphism group of a mixing shift of finite type is CAM. Combining the same example with Corollary 3.6 we obtain the stronger statement for shifts of finite type.

Corollary 3.7. *If H is a finite index subgroup of the automorphism group of a mixing shift of finite type (X, σ) , then the H -subaction on X is a CAM system.*

4. EXISTENCE OF A \mathbb{Z} -CAM SYSTEM

4.1. Construction of a CAM \mathbb{Z} -system with multiple nonatomic, ergodic invariant measures. In this section we show that a \mathbb{Z} -CAM system exists. Moreover, we construct a \mathbb{Z} -CAM system that has two different, non-atomic ergodic measures.

4.1.1. Setting notation. We build the system as a subshift on the alphabet $\mathcal{A} = \{0, 1\}$. The system is defined recursively, defining the language of the system in levels. The first three levels play a special role in starting the inductive process of defining the words on higher levels. Beginning with level four, we use a recursive formula to define the words, and the words constructed are of two types: those used to form the periodic points, denoted by the letter w with appropriate subscripts, and those designed to produce dense orbits, denoted by the letters a and b with appropriate subscripts. Throughout the construction, the periodic words have two subscripts, with the second coordinate denoting the level at which the word is introduced, whereas the density words come in two types (a and b) but only have a single subscript, also indicating the level at which such a word is introduced. We remark that, in our construction, we build a single infinite word that has the desired property, and its orbit closure is the CAM system. We do not explicitly build the language in our construction.

We fix a sequence $\{\varepsilon_k\}_{k=1}^{\infty}$ for use throughout our construction, which we refer to as a *frequency sequence*: assume that $\{\varepsilon_k\}_{k=1}^{\infty}$ is a sequence of non-negative real

numbers that satisfy $\sum_{k=1}^{\infty} \varepsilon_k < 1/3$, and further assume that for any $N \geq 1$, we have

$$(3) \quad \sum_{n=N}^{\infty} \varepsilon_n < \frac{1}{3^N}.$$

To keep track of frequency counts, for words $u, v \in \{0, 1\}^*$ of finite length, define $\mathcal{N}(u, v)$ to be the number of times u occurs as a subword of v .

4.1.2. *Level 1 words (introduction of the words w).* There are 2 words on this level,

$$w_{(1,1)} = 0 \quad \text{and} \quad w_{(2,1)} = 1,$$

and we call these the *level-1 words*. Set $\mathcal{W}_1 = \{w_{(1,1)}, w_{(2,1)}\}$.

4.1.3. *Level 2 words (introduction of the words a and b).* There are 4 words on this level, playing two distinct roles as distinguished by two types: we have periodic words w and density words a, b . This level also introduces a parameter $n_2 > 1$ that controls the densities (note that this is the first level with such a parameter and so there is no n_1 ; the meaning of this is clarified momentarily). We define:

$$w_{(1,2)} = w_{(1,1)}^{n_2+1}$$

$$w_{(2,2)} = w_{(2,1)}^{n_2+1}$$

$$a_2 = w_{(1,1)} w_{(2,1)}^{n_2}$$

$$b_2 = w_{(1,1)}^{n_2} w_{(2,1)}$$

The parameter n_2 is chosen sufficiently large such that $\mathcal{N}(0, a_2)/|a_2| < \varepsilon_1$ and $\mathcal{N}(1, b_2)/|b_2| < \varepsilon_1$. In other words, a_2 consists almost entirely of 1's, b_2 almost entirely of 0's, and all of the words constructed on this level have the same length. We call the four words introduced at this point the *level-2 words* and we write $\mathcal{W}_2 = \{w_{(1,2)}, w_{(2,2)}, a_2, b_2\}$. For use in verifying properties of the words constructed at higher levels, we describe the basic properties of the words constructed at level 2.

Proposition 4.1. *For all distinct words $u, v \in \mathcal{W}_2$, the word u does not occur as a subword of vv .*

Proof. First we establish the statement when $v \in \{w_{(1,2)}, w_{(2,2)}\}$ and $u \in \mathcal{W}_2$ is distinct from v . Neither a_2 nor b_2 occurs as a subword of $w_{(i,2)}w_{(i,2)}$ for $i = 1, 2$, because those words are 1-periodic whereas a_2 and b_2 are not. The word $w_{(1,2)}$ does not occur as a subword of $w_{(2,2)}w_{(2,2)}$, since the symbol 0 does not occur in $w_{(2,2)}$ but does occur in $w_{(1,2)}$ and similarly $w_{(2,2)}$ does not occur as a subword of $w_{(1,2)}w_{(1,2)}$. This establishes the statement when $v \in \{w_{(1,2)}, w_{(2,2)}\}$.

Next we establish the statement when $v \in \{a_2, b_2\}$ and $u \in \mathcal{W}_2$ is distinct from v . Neither $w_{(1,2)}$ nor $w_{(2,2)}$ occurs as a subword of a_2a_2 , as the longest 1-periodic subword of a_2a_2 has length $n_2 \cdot |w_{(2,1)}| < |w_{(1,2)}|$ for $i = 1, 2$. Similarly, neither can occur as a subword of b_2b_2 . We claim that the word a_2 does not occur as a subword of b_2b_2 (the argument that b_2 does not occur as a subword of a_2a_2 is analogous). By construction, $\mathcal{N}(0, a_2)/|a_2| < \varepsilon_1$ and so $\mathcal{N}(0, a_2) < \varepsilon_1|a_2|$.

Therefore $\mathcal{N}(1, a_2) \geq (1 - \varepsilon_1)|a_2|$. Similarly, $\mathcal{N}(0, b_2b_2) \geq (1 - \varepsilon_1)|b_2b_2|$. If a_2 occurs as a subword of b_2b_2 then, using the fact that $|a_2| = \frac{1}{2}|b_2b_2|$, we obtain

$$\mathcal{N}(1, b_2b_2) \geq \mathcal{N}(1, a_2) \geq (1 - \varepsilon_1)|a_2| = \frac{1 - \varepsilon_1}{2}|b_2b_2|.$$

This implies that $\mathcal{N}(0, b_2b_2) \leq (1 - \frac{1 - \varepsilon_1}{2})|b_2b_2| = \frac{1 + \varepsilon_1}{2}|b_2b_2|$. Combined with our earlier observation that $\mathcal{N}(0, b_2b_2) \geq (1 - \varepsilon_1)|b_2b_2|$, we get $1 - \varepsilon_1 \leq \frac{1 + \varepsilon_1}{2}$. However, this can not occur because we assumed that $\varepsilon_1 < \sum_{k=1}^{\infty} \varepsilon_k < 1/3$. Therefore the statement holds when $v \in \{a_2, b_2\}$. \square

4.1.4. *Level 3 words (recursive definitions of periodic and density words)*. There are 6 words on this level, again playing distinct roles distinguished by two types: periodic words w and density words a, b . Again we introduce a parameter $n_3 > 1$ to control the densities. Define the words by setting

$$w_{(1,3)} = w_{(1,2)}^{5n_3+4}$$

$$w_{(2,3)} = w_{(2,2)}^{5n_3+4}$$

$$w_{(3,3)} = a_2^{5n_3+4}$$

$$w_{(4,3)} = b_2^{5n_3+4}$$

$$a_3 = a_2^{n_3} w_{(1,2)} a_2^{n_3} w_{(2,2)} a_2^{n_3} a_2 a_2^{n_3} b_2 a_2^{n_3}$$

$$b_3 = b_2^{n_3} w_{(1,2)} b_2^{n_3} w_{(2,2)} b_2^{n_3} a_2 b_2^{n_3} b_2 b_2^{n_3}$$

and choose n_3 such that $\mathcal{N}(0, a_3)/|a_3| < \varepsilon_1 + \varepsilon_2$ and $\mathcal{N}(1, b_3)/|b_3| < \varepsilon_1 + \varepsilon_2$ (we note that this is possible because $\mathcal{N}(0, a_2)/|a_2| < \varepsilon_1$ and a_3 is mainly made of concatenated copies of a_2 , similarly $\mathcal{N}(1, b_2)/|b_2| < \varepsilon_1$). We further assume that n_3 has been chosen sufficiently large such that for $u \in \{w_{(1,2)}, w_{(2,2)}, b_2\}$, we have

$$(4) \quad \frac{\mathcal{N}(u, a_3a_3)}{|a_3a_3|} < \frac{\varepsilon_2}{|u|(2|u| - 1)},$$

and note that this can be done since a_3a_3 is mainly made of concatenated copies of a_2a_2 and u does not arise as a subword of a_2a_2 . Note that if $|u| = 1$ then the denominator on the right-hand side of this inequality is 1, so this formula generalizes the analogous inequality (and assumptions about the size of n_2) in the previous level. Using the same reasoning, this choice of n_3 ensures that if $u \in \{w_{(1,2)}, w_{(2,2)}, a_2\}$, then we have

$$(5) \quad \frac{\mathcal{N}(u, b_3b_3)}{|b_3b_3|} < \frac{\varepsilon_2}{|u|(2|u| - 1)}.$$

Finally note that by choosing n_3 sufficiently large, we can guarantee that $|a_2|/|a_3|$ is as small as desired. For use in the proof of the next proposition describing occurrences of words and subwords, we have a specific constant we would like it to be smaller than. Let p be the minimal period of the word a_2a_2 which, by symmetry

of the construction, is also the minimal period of b_2b_2 . We assume that n_3 is chosen to be sufficiently large such that

$$(6) \quad \frac{|a_2|}{|a_3|} < \frac{1}{6p} - \frac{1}{9|a_2|}.$$

We call the words constructed on this level the *level-3 words* and set \mathcal{W}_3 to be the set of all level-3 words.

Proposition 4.2. *For any choice of distinct words $u, v \in \mathcal{W}_3$, the word u does not occur as a subword of vv .*

Proof. We check that for each $v \in \mathcal{W}_3$ and $u \in \mathcal{W}_3 \setminus \{v\}$, the word u does not occur as a subword of vv , checking cases depending on the type of word v .

First consider when $v = w_{(i,3)}$ for some $i \in \{1, 2\}$. We consider two cases, depending on the choice of u . First suppose that $u \in \{w_{(3,3)}, w_{(4,3)}, a_3, b_3\}$. By Proposition 4.1, neither a_2 nor b_2 occurs as a subword of $w_{(i,2)}w_{(i,2)}$. The word $vv = w_{(i,3)}w_{(i,3)}$ is the self-concatenation of a large number of copies of $w_{(i,2)}$ and so neither a_2 nor b_2 occurs as a subword of vv . But at least one of a_2 and b_2 occurs as a subword of u , and so u cannot be a subword of vv . Next suppose $u = w_{(j,3)}$ where $j \in \{1, 2\}$ and $j \neq i$. The word $u = w_{(j,3)}$ is not a subword of $vv = w_{(i,3)}w_{(i,3)}$ because $w_{(j,2)}$ is not a subword of $w_{(i,2)}w_{(i,2)}$ and, as noted, vv is the self-concatenation of copies of $w_{(i,2)}$ while $w_{(j,3)}$ is the self-concatenation of copies of $w_{(j,2)}$. Thus u does not occur as a subword of vv for any $u \in \mathcal{W}_3 \setminus \{v\}$. Thus the statement holds when $v \in \{w_{(1,3)}, w_{(2,3)}\}$.

Next consider when $v = w_{(i,3)}$ for some $i \in \{3, 4\}$. Again we have two cases, depending on the choice of u . First suppose $u \in \{w_{(1,3)}, w_{(2,3)}, a_3, b_3\}$. By Proposition 4.1, neither $w_{(1,2)}$ nor $w_{(2,2)}$ occurs as a subword of a_2a_2 or of b_2b_2 . The word $vv = w_{(i,3)}w_{(i,3)}$ is the self-concatenation of a large number of copies of a_2 or of b_2 , and so neither $w_{(1,2)}$ nor $w_{(2,2)}$ occurs as a subword of vv . However at least one of $w_{(1,2)}$ and $w_{(2,2)}$ occurs in u , and so u is not a subword of vv . We next consider when $j \in \{3, 4\} \setminus \{i\}$ and let $u = w_{(j,3)}$. Then one of the words u and v is the self-concatenation of many copies of a_2 and the other is the self-concatenation of many copies of b_2 . By Proposition 4.1, a_2 is not a subword of b_2b_2 and b_2 is not a subword of a_2a_2 , and so u is not a subword of vv . This establishes the statement when $v \in \{w_{(3,3)}, w_{(4,3)}\}$.

Finally consider when $v \in \{a_3, b_3\}$. We give the argument when $v = a_3$, and the argument for b_3 is similar. Again, we have two cases, depending on the choice of u . First suppose $u \in \{w_{(1,3)}, w_{(2,3)}, w_{(4,3)}, b_3\}$. By (4), we know that $\mathcal{N}(x, a_3a_3)/|a_3a_3| < \varepsilon_2/|x|(2|x| - 1)$ for $x \in \mathcal{W}_2 \setminus \{a_2\}$. For any such x and any particular occurrence of x in a_3a_3 , there are $2|x| - 1$ subwords of a_3a_3 that have length $|x|$ and partially (or completely) overlap this occurrence of x . This means that, for any $x \in \mathcal{W}_2 \setminus \{a_2\}$, if we look at the collection of all locations within a_3a_3 where x occurs, we have

$$\frac{\text{subwords of length } |x| \text{ in } a_3a_3 \text{ that partially overlap an occurrence of } x}{|a_3a_3|} < \frac{\varepsilon_2}{|x|}.$$

Let p be the minimal period of the word a_2a_2 . Note that $p \leq |a_2|$. Since all four words in \mathcal{W}_2 have the same length, we deduce that

$$(7) \quad \frac{\mathcal{N}(a_2, a_3a_3)}{|a_3a_3|} \geq \frac{1}{p} - 3 \cdot \frac{\varepsilon_2}{|a_2|}$$

because a_3a_3 is made by concatenating words in \mathcal{W}_2 . On the other hand, we claim that $\mathcal{N}(a_2, u) \leq 6|a_2|$. To check this, note that if $u = b_3$, any occurrence of a_2 in u must partially overlap $w_{(1,2)}$, $w_{(2,2)}$ or a_2 in the definition of b_3 , because it does not occur as a subword of b_2b_2 , and there are only $6|a_2|$ locations that have such overlaps. If $u \neq b_3$ then a_2 does not occur in u , by Proposition 4.1 and the claim follows. If u occurred as a subword of a_3a_3 , such an occurrence would account for half of the letters in a_3a_3 and we could write

$$a_3a_3 = xuy$$

where $|x| + |y| = \frac{|a_3a_3|}{2}$. We refer to this decomposition of a_3a_3 in what follows. Any occurrence of a_2 in a_3a_3 is either entirely within u , entirely within x , entirely within y , or partially overlaps x (or y) and partially overlaps u . There are at most $|x| + |y| = |a_3a_3|/2$ many possible locations for a_2 to occur in a_3a_3 that are not entirely within u , and these locations consist of two disjoint intervals of starting points. Because the minimal period of a_2a_2 is p , within either of these intervals, the density with which a_2 occurs is at most $1/p$ as any higher density would force two occurrences of a_2 whose starting points differ by less than p and therefore contradicting minimality of p . This implies that

$$\begin{aligned} \frac{\mathcal{N}(a_2, a_3a_3)}{|a_3a_3|} &\leq \frac{\# \text{ times } a_2 \text{ occurs not entirely in } u}{|a_3a_3|} + \frac{\# \text{ times } a_2 \text{ occurs in } u}{|a_3a_3|} \\ &\leq \frac{1}{2p} + \frac{6|a_2|}{|a_3a_3|} = \frac{1}{2p} + 3 \cdot \frac{|a_2|}{|a_3|}. \end{aligned}$$

But, by construction, by Equation (6) we have that $|a_2|/|a_3| < \frac{1}{6p} - \frac{1}{9|a_2|}$, and so $\mathcal{N}(a_2, a_3a_3)/|a_3a_3| < \frac{1}{p} - \frac{1}{3|a_2|}$. By our choice of frequency sequence in (3), it follows that $\varepsilon_2 < \sum_{k=2}^{\infty} \varepsilon_k < 1/9$ and so equation (7) implies that $\mathcal{N}(a_2, a_3a_3)/|a_3a_3| \geq \frac{1}{p} - \frac{1}{3|a_2|}$. This is a contradiction, and so u cannot occur as a subword of $vv = a_3a_3$. Lastly, suppose that $u = w_{(3,3)}$. By construction, u is the self-concatenation of $5n_3 + 4$ copies of a_2 . Then u cannot occur as a subword of $vv = a_3a_3$ because

$$a_3a_3 = a_2^{n_3}w_{(1,2)}a_2^{n_3}w_{(2,2)}a_2^{n_3}a_2a_2^{n_3}b_2a_2^{n_3}a_2^{n_3}w_{(1,2)}a_2^{n_3}w_{(2,2)}a_2^{n_3}a_2a_2^{n_3}b_2a_2^{n_3}$$

and so the largest power m such that a_2^m occurs as a subword of a_3a_3 satisfies $m \leq 2n_3 + 2$, by Proposition (4.1) (any larger power would force an occurrence of an element of $\mathcal{W}_2 \setminus \{a_2\}$ as a subword of a_2^m). This establishes the statement when $v = a_3$. The argument for $v = b_3$ is similar, with the roles played by a_2 and b_2 switched. \square

4.1.5. *Preparing to build the level $k + 1$ words (inductive assumptions).* For $k \geq 3$, we construct the words on level $k + 1$ using the words on level k . Again they come in two varieties (periodic words and density words), and again we have a parameter $n_{k+1} > 1$ chosen to guarantee that the words have the desired properties. Inductively, we assume that we have constructed the level- k words: these are words $w_{(i,k)}$ for all $1 \leq i \leq 2k - 2$, words a_k and b_k , with all of these words having equal length, and we denote the collection of all level- k words by \mathcal{W}_k . We further assume that we have defined the parameter $n_k \in \mathbb{N}$, and these constructions satisfy the following properties:

- (i) For $1 \leq i \leq 2k - 4$, we have

$$w_{(i,k)} = w_{(i,k-1)}^{(2k-1)n_k+2(k-1)}.$$

(ii) We have

$$w_{(2k-3,k)} = a_{k-1}^{(2k-1)n_k+2(k-1)}.$$

(iii) We have

$$w_{(2k-2,k)} = b_{k-1}^{(2k-1)n_k+2(k-1)}.$$

(iv) We have

$$a_k = \left(\prod_{i=1}^{2k-4} a_{k-1}^{n_k} w_{(i,k-1)} \right) \cdot a_{k-1}^{n_k} \cdot (a_{k-1}) \cdot a_{k-1}^{n_k} \cdot (b_{k-1}) \cdot a_{k-1}^{n_k}.$$

(v) We have

$$b_k = \left(\prod_{i=1}^{2k-4} b_{k-1}^{n_k} w_{(i,k-1)} \right) \cdot b_{k-1}^{n_k} \cdot (a_{k-1}) \cdot b_{k-1}^{n_k} \cdot (b_{k-1}) \cdot b_{k-1}^{n_k}.$$

We further assume that for any word u constructed on some level m with $m < k$, other than when $u = a_m$, we have

$$(8) \quad \frac{\mathcal{N}(u, a_k a_k)}{|a_k a_k|} < \frac{1}{|u|(2|u|-1)} \sum_{j=m}^{k-1} \varepsilon_j$$

and that for any $m < k$ and any level- m word u other than $u = b_m$, we have

$$\frac{\mathcal{N}(u, b_k b_k)}{|b_k b_k|} < \frac{1}{|u|(2|u|-1)} \sum_{j=m}^{k-1} \varepsilon_j.$$

Finally we assume that

(9) for any $u, v \in \mathcal{W}_k$ with $u \neq v$, the word u does not occur as a subword of vv .

4.1.6. *Level- $(k+1)$ words (recursive definition).* We apply the analogous procedure used for level 3 and construct the periodic and density words, this time assuming the properties given in ((i)) through ((v)).

(vi) For $1 \leq i \leq 2k-2$, define the first group of periodic words by setting

$$w_{(i,k+1)} = w_{(i,k)}^{(2k+1)n_{k+1}+2k}.$$

(vii) Define the next periodic word by setting

$$w_{(2k-1,k+1)} = a_k^{(2k+1)n_{k+1}+2k}.$$

(viii) Define the last periodic word by setting

$$w_{(2k,k+1)} = b_k^{(2k+1)n_{k+1}+2k}.$$

(ix) Define the first type of density word by setting

$$a_{k+1} = \left(\prod_{i=1}^{2(k+1)-4} a_k^{n_{k+1}} w_{(i,k)} \right) \cdot a_k^{n_{k+1}} \cdot (a_k) \cdot a_k^{n_{k+1}} \cdot (b_k) \cdot a_k^{n_{k+1}}.$$

(x) Define the second type of density word by setting

$$b_{k+1} = \left(\prod_{i=1}^{2(k+1)-4} b_k^{n_{k+1}} w_{(i,k)} \right) \cdot b_k^{n_{k+1}} \cdot (a_k) \cdot b_k^{n_{k+1}} \cdot (b_k) \cdot b_k^{n_{k+1}}.$$

We choose the parameter n_{k+1} sufficiently large such that for every $m < k + 1$ and every word $u \in \mathcal{W}_m \setminus \{a_m\}$, we have

$$(10) \quad \frac{\mathcal{N}(u, a_{k+1}a_{k+1})}{|a_{k+1}a_{k+1}|} < \frac{1}{|u|(2|u| - 1)} \sum_{j=m}^k \varepsilon_j.$$

When $m < k$, recall that $\mathcal{N}(u, a_k a_k)/|a_k a_k| < (\sum_{j=m}^{k-1} \varepsilon_j)/|u|(2|u| - 1)$ by (8). Therefore it is possible to satisfy (10) because, for n_{k+1} large, the word a_{k+1} primarily consists of concatenated copies of a_k and so we can make the left hand side of (10) as close as we want to $(\sum_{j=m}^{k-1} \varepsilon_j)/|u|(2|u| - 1)$; in particular we can make it less than the right hand side of (10). When $m = k$ it is possible to satisfy (10) because u does not occur as a subword of $a_k a_k$ and, again, a_{k+1} is primarily made of concatenated copies of a_k . Similarly we choose n_{k+1} sufficiently large such that if $u \in \mathcal{W}_m \setminus \{b_m\}$, then we have

$$(11) \quad \frac{\mathcal{N}(u, b_{k+1}b_{k+1})}{|b_{k+1}b_{k+1}|} < \frac{1}{|u|(2|u| - 1)} \sum_{j=m}^k \varepsilon_j.$$

For our next estimate, let p_k be the minimal period of the word $a_k a_k$. For use in a future argument, we finally require that n_{k+1} is sufficiently large such that

$$(12) \quad \frac{|a_k|}{|a_{k+1}|} < \frac{1}{(4k - 2)p_k} - \frac{1}{3^k \cdot |a_k|}.$$

This is possible as long as the right-hand side of the inequality is positive. Note that $p_k \leq |a_k|$ and that $4k - 2 < 3^k$ because $k \geq 3$, and that all of the words constructed on level $k + 1$ have equal length. We call the words constructed in ((vi)) through ((x)) the *level-(k + 1) words* and let \mathcal{W}_{k+1} denote the set of all level-(k + 1) words. The only thing that is required now to complete our inductive construction is to prove the following proposition.

Proposition 4.3. *For any choice of distinct words $u, v \in \mathcal{W}_{k+1}$, the word u does not occur as a subword of vv .*

Proof. As on level three, fixing some $v \in \mathcal{W}_{k+1}$, we argue that for all $u \in \mathcal{W}_{k+1} \setminus \{v\}$, the statement holds. Again, we split the argument into cases depending on the type of the word v .

First suppose that $v \in \{w_{(i,k+1)} : 1 \leq i \leq 2k - 2\}$ for some $1 \leq i \leq 2k - 2$. We start by considering the case that $u \in \{w_{(2k-1,k+1)}, w_{(2k,k+1)}, a_{k+1}, b_{k+1}\}$. By (9), neither a_k nor b_k occurs as a subword of $w_{(i,k)}w_{(i,k)}$, and by construction, at least one of a_k or b_k occurs as a subword of u . Since $vv = w_{(i,k+1)}w_{(i,k+1)}$ is the self-concatenation of a large number of copies of $w_{(i,k)}$, u cannot be a subword of vv . Next consider when $u = w_{(j,k+1)}$ for some $j \in \{1, 2, \dots, 2k - 2\}$ with $j \neq i$. Again by (9), $w_{(i,k)}$ is not a subword of $w_{(j,k)}w_{(j,k)}$. But $w_{(i,k+1)}$ is the concatenation of many copies of $w_{(i,k)}$ and $w_{(j,k+1)}$ is the concatenation of many copies of $w_{(j,k)}$, and so an occurrence of u in vv forces an occurrence of $w_{(i,k)}$ in the word $w_{(j,k)}w_{(j,k)}$, a contradiction. Thus u cannot occur as a subword of vv . This establishes the statement for $v \in \{w_{(i,k+1)} : 1 \leq i \leq 2k - 2\}$.

Next assume $v = w_{(i,k+1)}$ for some $i \in \{2k - 1, 2k\}$. We first consider when $u \in \{w_{(j,k+1)} : 1 \leq j \leq 2k - 2\} \cup \{a_{k+1}, b_{k+1}\}$. By (9), none of the words in the set $\{w_{(j,k)} : 1 \leq j \leq 2k - 2\}$ occurs as a subword of $a_k a_k$ or of $b_k b_k$. But by construction, the word $w_{(i,k+1)}w_{(i,k+1)}$ is the self-concatenation of a large number

of copies of a_k or b_k , and we know that at least one of the words in $\{w_{(j,k)} : 1 \leq j \leq 2k - 2\}$ occurs as a subword of u . Therefore u cannot occur as a subword of vv . Now instead assume $j \in \{2k - 1, 2k\}$ with $j \neq i$. This time, one of the words $w_{(i,k+1)}$ and $w_{(j,k+1)}$ is the concatenation of many copies of a_k and the other is the concatenation of many copies of b_k . By (9), a_k is not a subword of $b_k b_k$ and b_k is not a subword of $a_k a_k$. Thus u cannot be a subword of vv , proving the statement for $v \in \{w_{(2k-1,k+1)}, w_{(2k-2,k+1)}\}$.

We are left with checking the case that $v \in \{a_{k+1}, b_{k+1}\}$. As on level three, we only include the argument when $v = a_{k+1}$, as the other case is similar. As before, there are two possibilities and we treat them separately. First suppose $u \in \{w_{(j,k+1)} : 1 \leq j \leq 2k - 2\} \cup \{w_{(2k,k+1)}, b_{k+1}\}$. By (10), for any $w \in \mathcal{W}_k \setminus \{a_k\}$ we have that

$$\frac{\mathcal{N}(w, a_{k+1}a_{k+1})}{|a_{k+1}a_{k+1}|} < \frac{\varepsilon_k}{|w|(2|w| - 1)}.$$

For each particular occurrence of w in $a_{k+1}a_{k+1}$, there are $2|w| - 1$ subwords of length $|w|$ that partially overlap this occurrence. This means that

$$\frac{\# \text{ of length } |w| \text{ subwords of } a_{k+1}a_{k+1} \text{ that partially overlap an occurrence of } w}{|a_{k+1}a_{k+1}|} < \frac{\varepsilon_k}{|w|}.$$

Recall that p_k is the minimal period of the word $a_k a_k$. Note that $p_k \leq |a_k|$. Since $|\mathcal{W}_k \setminus \{a_k\}| = 2k - 1$ and all words in \mathcal{W}_k have the same length, we deduce from (3) that

$$(13) \quad \frac{\mathcal{N}(a_k, a_{k+1}a_{k+1})}{|a_{k+1}a_{k+1}|} \geq \frac{1}{p_k} - \frac{(2k - 1)\varepsilon_k}{|a_k|}.$$

On the other hand, by (9), a_k does not occur as a subword of $w_{(i,k+1)}$ for any $k \leq 2k - 2$ (because it is not a subword of $w_{(i,k)}w_{(i,k)}$), a_k does not occur as a subword of $w_{(2k,k+1)}$ (because it is not a subword of $b_k b_k$), and a_k occurs as a subword of b_{k+1} at most $(4k - 2) \cdot |a_k|$ times (because it does not occur as a subword of $b_k b_k$ and so all occurrences have to partially overlap b_k and some $w_{(i,k)}$ or partially overlap the copy of a_k that appears in the definition of b_{k+1}). Therefore,

$$\mathcal{N}(a_k, u) \leq (4k - 2) \cdot |a_k|.$$

If u occurred as a subword of $a_{k+1}a_{k+1}$, such an occurrence would account for half of the letters in $a_{k+1}a_{k+1}$ and we could write

$$a_{k+1}a_{k+1} = xuy$$

where $|x| + |y| = |a_{k+1}a_{k+1}|/2$. Any occurrence of a_k in $a_{k+1}a_{k+1}$ would either be entirely inside of u , entirely inside of x , entirely inside of y , or partially overlap x (or y) and u . Any occurrence of a_k that is not entirely inside u must either start in x or end in y (but never both because $|u| > |a_k|$), so there are at most $|a_{k+1}a_{k+1}|/2$ locations where a_k could occur that are not entirely inside u and the density of occurrences of a_k in concatenated copies of $a_k a_k$ is $1/p_k$. Thus we can estimate the

relation between the number of occurrences and the length:

$$\begin{aligned} \frac{\mathcal{N}(a_k, a_{k+1}a_{k+1})}{|a_{k+1}a_{k+1}|} &\leq \frac{\# \text{ times } a_k \text{ occurs not entirely in } u}{|a_{k+1}a_{k+1}|} + \frac{\# \text{ times } a_k \text{ occurs in } u}{|a_{k+1}a_{k+1}|} \\ &\leq \frac{1}{2p_k} + \frac{(4k-2) \cdot |a_k|}{|a_{k+1}a_{k+1}|} \\ &< \frac{1}{2p_k} + (2k-1) \cdot \frac{|a_k|}{|a_{k+1}|}. \end{aligned}$$

By equation (12), we have that $\frac{|a_k|}{|a_{k+1}|} < \frac{1}{(4k-2)p_k} - \frac{1}{3^k \cdot |a_k|}$ and so $\frac{\mathcal{N}(a_k, a_{k+1}a_{k+1})}{|a_{k+1}a_{k+1}|} < \frac{1}{p_k} - \frac{2k-1}{3^k |a_k|}$. On the other hand, by equation (3), we have $\varepsilon_k < \sum_{m=k}^{\infty} \varepsilon_m < \frac{1}{3^k}$ and so, by equation (13), we also have $\frac{\mathcal{N}(a_k, a_{k+1}a_{k+1})}{|a_{k+1}a_{k+1}|} \geq \frac{1}{p_k} - \frac{2k-1}{3^k |a_k|}$, a contradiction. Therefore u cannot occur as a subword of $vv = a_{k+1}a_{k+1}$. The last possibility we need to consider is when $u = w_{(2k-2, k+1)}$. Then u is the concatenation of $(2k+1)n_{k+1} + 2(k+1)$ copies of a_k . But the largest power m such that a_k^m occurs as a subword of $a_{k+1}a_{k+1}$ satisfies $m \leq 2n_{k+1} + 2$, because any larger power would force an occurrence of an element of $\mathcal{W}_k \setminus \{a_k\}$ (that occurs in $a_{k+1}a_{k+1}$) as a subword of a_k^m , contradicting (9). Thus we have the statement for $v = a_{k+1}$, and the argument for $v = b_{k+1}$ is similar, reversing the roles played by a_k and b_k . \square

4.1.7. *Construction of the shift (X, σ) .* For all $k > 2$, note that a_{k+1} begins and ends with the word a_k . Therefore there is a unique $x \in \{0, 1\}^{\mathbb{Z}}$ such that for all $k > 2$ we have

$$x_1 x_2 \dots x_{|a_k|} = x_{-|a_k|+1} x_{-|a_k|+2} \dots x_0 = a_k.$$

Set X to be the orbit closure of x in $\{0, 1\}^{\mathbb{Z}}$ and for the remainder of this section, we study the properties of the shift (X, σ) .

Proposition 4.4. *If $k > 2$, any $y \in X$ can be written as a bi-infinite concatenation of words constructed on level k . In the special case that $y = x$, more holds: whenever $n \equiv 1 \pmod{|a_k|}$ the subword $x_n x_{n+1} \dots x_{n+|a_k|-1}$ is a level- k word.*

Furthermore, if $n \equiv 1 \pmod{|a_k|}$ and $x_n x_{n+1} \dots x_{n+2|a_k|-1} = uv$ is the concatenation of two level- k words, then either $u = v$ or uv is one of the following:

- (1) $a_k w_{(i, k)}$ or $w_{(i, k)} a_k$ for some $1 \leq i \leq 2k-2$;
- (2) $b_k w_{(i, k)}$ or $w_{(i, k)} b_k$ for some $1 \leq i \leq 2k-2$;
- (3) $a_k b_k$;
- (4) $b_k a_k$.

Moreover, for any $y \in X$ there exists $0 \leq r < |a_k|$ and a sequence $\{n_m\}_{m=1}^{\infty}$ such that $n_m \equiv r \pmod{|a_k|}$ for all $m \in \mathbb{N}$ and $\lim_{m \rightarrow \infty} \sigma^{n_m} x = y$. With this value of r fixed, whenever $n \equiv r+1 \pmod{|a_k|}$ then $y_n y_{n+1} \dots y_{n+2|a_k|-1} = uv$ is a concatenation of two level- k words and either $u = v$ or uv is one of the forms listed in (1), (2), (3), (4).

Proof. It suffices to prove this for the transitive point x used to define (X, σ) , as then the result for general $y \in X$ follows by noting that every level- k word has the same length and y is in the orbit closure of x .

By construction, all level k words have the same length and for any $s > 0$, a_{k+s} is a concatenation of level- $(k+s-1)$ words. Therefore, a_{k+s} is a concatenation of level- k words. For each s , the word $x_{-|a_{k+s}|+1} \dots x_{|a_{k+s}|} = a_{k+s} a_{k+s}$ and so $x_{-|a_{k+s}|+1} \dots x_{|a_{k+s}|}$ can be written as a concatenation of level- k words and

$x_n x_{n+1} \dots x_{n+|a_k|-1}$ is a level- k word for all $n \equiv 1 \pmod{|a_k|}$ with $-|a_{k+s}| < n \leq |a_{k+s}|$. This holds for all s , and so x can be written as a concatenation of level- k words and $x_n x_{n+1} \dots x_{n+|a_k|-1}$ is such a word whenever $n \equiv 1 \pmod{|a_k|}$.

We next turn to the second statement on the form of concatenated level- k words. We consider a level- $(k+s)$ word w and proceed by induction on $s > 0$. It suffices to show that when w is written as a concatenation of level- k words and uv is the concatenation of two adjacent level- k words that occur when w is parsed in this way, then either $u = v$ or uv one of forms given in (1)–(4). By the definition of the level- $(k+1)$ words, the result clearly holds for $s = 1$. For $s = 2$, this follows immediately by writing the level-2 words as concatenations of level-1 words. Inductively, assume this holds for all level- $(k+s)$ words for some $s > 1$, and we consider the level- $(k+s+1)$ words. When $1 \leq i \leq 2(k+s+1) - 4$, the word $w_{(i,k+s+1)}$ is the self-concatenation of $w_{(i,k+s)}$ a large number of times. But, since $s > 1$, $w_{(i,k+s)}$ is itself a self-concatenation of level- $(k+s-1)$ words, this means that any two adjacent level- k words that occur in $w_{(i,k+s)} w_{(i,k+s)}$ already occurred within $w_{(i,k+s-1)}$ (if $i \leq 2(k+s) - 3$) or within a_{k+s-1} or b_{k+s-1} . By induction, the result holds for all $s \geq 1$.

For the final statement of this proposition, note that if $y \in X$, then $y \in \overline{\mathcal{O}(x)}$ and so exists a sequence $\{n_m\}_{m=1}^\infty$ such that $\lim_{m \rightarrow \infty} \sigma^{n_m} x = y$. Without loss of generality, passing to a subsequence if necessary, we can assume that there exists $0 \leq r < |a_k|$ such that $n_m \equiv r \pmod{|a_k|}$ for all m . So whenever $n \equiv r+1 \pmod{|a_k|}$ then $y_n y_{n+1} \dots y_{n+2|a_k|-1}$ is the same as a subword of x that begins at some location congruent to $1 \pmod{|a_k|}$. \square

Proposition 4.5. *The system (X, σ) is a CAM subshift.*

Proof. Let $x \in X$ denote the transitive point used to define the system. We start by showing that the periodic points are dense in X . For any $k > 2$, the word $a_k a_k = x_{-|a_k|+1} \dots x_{|a_k|}$ and so the orbit of any periodic point that contains $a_k a_k$ as a subword is within distance $1/2^{|a_k|+1} < 1/2^{2^{k-1}+1}$ of the point x . For each fixed k and any $s > 0$, the word $w_{(2k-1,k+s)}$ is the self-concatenation of many copies of a_k . Therefore there is a periodic point $p_k \in X$ that is the bi-infinite self-concatenation of copies of a_k . It follows that $x = \lim_{k \rightarrow \infty} p_k$. Since x is a transitive point and is approximated arbitrarily well by periodic points, it then follows that the periodic points are dense in X .

We next check that if $y \in X$ is not periodic, then it has a dense orbit. If, for all $k > 2$, the word $a_k a_k$ occurs as a subword of y , then the orbit of y is dense since points in its orbit would approximate x arbitrarily well. We proceed by contradiction, assuming the existence of some $y \in X$ that is not periodic and does not have a dense orbit. Thus we can assume that there is some $k > 2$ such that $a_k a_k$ does not occur as a subword of y .

The word a_{k+1} does not occur as a subword of y , because this would force an occurrence of $a_k a_k$ in y . By Proposition 4.4, y can be written as a bi-infinite concatenation of words constructed on level $k+2$ and, therefore, can be written as a bi-infinite concatenation of $w_{(1,k+2)}, \dots, w_{(2k-2,k+2)}$ as well as $w_{(2k,k+2)}$ and $w_{(2k+2,k+2)}$. We omit the words $w_{(2k-1,k+2)}, w_{(2k+1,k+2)}, a_{k+2}$, and b_{k+2} from the list of possible level $k+2$ words because:

- $w_{(2k-1,k+2)}$ is the self-concatenation of many copies of a_k ;

- $w_{(2k+1,k+2)}$ is the self-concatenation of many copies of a_{k+1} , which contains $a_k a_k$ as a subword;
- a_{k+2} has a_{k+1} as a subword and a_{k+1} has $a_k a_k$ as a subword;
- b_{k+2} has a_{k+1} as a subword and a_{k+1} has $a_k a_k$ as a subword.

Again applying to Proposition 4.4 to the words constructed on level $k+2$, whenever uv are two adjacent level $k+2$ words that occur in x , we must have $u = v$ since none of the other forms listed in Proposition 4.4 can be written only using the words $w_{(1,k+2)}, \dots, w_{(2k-2,k+2)}$ and $w_{(2k+2,k+2)}$. But then it follows that y is periodic, a contradiction. Thus we have shown that (X, σ) is weakly CAM and is expansive, and so by Proposition 2.17, the system (X, σ) is CAM. \square

4.2. Invariant measures on the system. We continue letting (X, σ) denote the system defined in Section 4.1.7.

Proposition 4.6. *The system (X, σ) supports two nonatomic ergodic measures.*

Proof. Our goal is to construct two (not necessarily ergodic) measures, μ_a and μ_b , supported on X , and show the following:

- (1) $\mu_a \neq \mu_b$ (which we do by showing $\mu_a([0]) \neq \mu_b([0])$);
- (2) almost every measure in the ergodic decomposition of μ_a is nonatomic;
- (3) almost every measure in the ergodic decomposition of μ_b is nonatomic.

It follows from these facts that there exist two distinct, ergodic, nonatomic measures supported on X and, therefore, that CAM systems are capable of having this property.

Let $x \in X$ be the infinite word defined by the condition that for all $k \geq 1$, we have

$$a_k = x_1 \dots x_{|a_k|} = x_{-|a_k|+1} \dots x_0$$

and let $x' \in X$ be the infinite word defined by the condition that for all $k \geq 1$, we have

$$b_k = x'_1 \dots x'_{|a_k|} = x'_{-|a_k|+1} \dots x'_0.$$

For each $k \geq 3$, set

$$\nu_{k,a} := \frac{1}{2^{|a_k|}} \sum_{m=-|a_k|+1}^{|a_k|} \delta_{\sigma^m(x)}$$

and set

$$\nu_{k,b} := \frac{1}{2^{|b_k|}} \sum_{m=-|b_k|+1}^{|b_k|} \delta_{\sigma^m(x')}.$$

Let μ_a be a weak*-limit point of $\{\nu_{k,a}\}_{k=1}^\infty$ and let μ_b be a weak*-limit point of $\{\nu_{k,b}\}_{k=1}^\infty$. Then for all $k \geq 1$, we have

$$\nu_{k,a}([0]) = \frac{\mathcal{N}(0, a_k)}{|a_k|} \leq \sum_{j=1}^k \varepsilon_j < \frac{1}{3} \quad \text{and} \quad \nu_{k,b}([1]) = \frac{\mathcal{N}(1, b_k)}{|b_k|} \leq \sum_{j=1}^k \varepsilon_j < \frac{1}{3}$$

and so in particular it follows that $\mu_a([0]) \leq \frac{1}{3}$ and $\mu_b([1]) \leq \frac{1}{3}$. Let

$$\mu_a = \int_X \mu_a^x d\mu_a(x) \quad \text{and} \quad \mu_b = \int_X \mu_b^x d\mu_b(x)$$

be the ergodic decompositions of μ_a and μ_b , respectively. Then we have

$$(14) \quad \mu_a \left(\left\{ x : \mu_a^x([1]) \geq \frac{2}{3} \right\} \right) > 0;$$

$$(15) \quad \mu_b \left(\left\{ x : \mu_b^x([0]) \geq \frac{2}{3} \right\} \right) > 0.$$

Let $p \in X$ be a periodic point. By construction of X , there is some $k \in \mathbb{N}$ and a level- k word, u , such that p is a shift of the word $\cdots uuuu \cdots$ (because arbitrarily long segments of p have to occur in a_{k+t} as $t \rightarrow \infty$). Fix $\varepsilon > 0$ and find $m > k$ such that $\sum_{j=m}^{\infty} \varepsilon_j < \varepsilon$. Let \tilde{u} be the level- m word that is $|u|$ -periodic and is the self-concatenation of a large number of copies of u . Since $m > k$, we have that $\tilde{u} \neq a_m$ and so for all $t \geq 0$, it follows from Equation (8) that

$$\frac{\mathcal{N}(\tilde{u}, a_{m+t}a_{m+t})}{|a_{m+t}a_{m+t}|} < \frac{1}{|\tilde{u}|(2|u| - 1)} \sum_{j=m}^{\infty} \varepsilon_j < \frac{\varepsilon}{|\tilde{u}|(2|u| - 1)}.$$

Therefore

$$\nu_{a,m+t}([\tilde{u}]) < \frac{\varepsilon}{|\tilde{u}|(2|u| - 1)} + \frac{|\tilde{u}| - 1}{|a_{m+t}a_{m+t}|}$$

for all $t \geq 0$, as the second term on the right hand side is from possible occurrences of \tilde{u} in x that are within distance $|\tilde{u}|$ of the right edge of the rightmost copy of a_{m+t} . Thus it follows that $\mu_a([\tilde{u}]) \leq \varepsilon/|\tilde{u}|(2|u| - 1)$. Similarly, using the same argument but replacing a_m by b_m , it follows that $\mu_b([\tilde{u}]) \leq \varepsilon/|\tilde{u}|(2|u| - 1)$. Repeating this argument for any $\varepsilon > 0$, we conclude that

$$\lim_{c \rightarrow \infty} \mu_a([u^c]) = 0 \quad \text{and} \quad \lim_{c \rightarrow \infty} \mu_b([u^c]) = 0.$$

This implies that $\mu_a(\{x : \mu_a^x \text{ is atomic and concentrated on } p\}) = 0$ and, similarly, that $\mu_b(\{x : \mu_b^x \text{ is atomic and concentrated on } p\}) = 0$. Since the choice of p was arbitrary, it follows that μ_a -almost every ergodic component of μ_a is nonatomic and μ_b -almost every ergodic component of μ_b is nonatomic. However by (14), there is a set of ergodic components of μ_a of positive μ_a -measure that give measure at least $2/3$ to the cylinder set $[1]$, and hence there is at least one nonatomic ergodic measure that does. Similarly, by (15) there is a nonatomic ergodic component of μ_b that gives measure at most $1/3$ to the cylinder set $[1]$, and so there is at least one nonatomic ergodic measure that does. It follows that these two nonatomic, ergodic measures are distinct. \square

Combining Propositions 4.5 and 4.6, we have thus completed the proof of Theorem 1.1 Some open questions and directions for exploration about the system defined in 4.1.7 are contained in Section 6.

5. \mathbb{Z}^d -CAM SYSTEM

5.1. The setup. In Section 4, we showed the existence of a symbolic \mathbb{Z} -CAM system and further showed that such a system can support two distinct nonatomic, ergodic probability measures. In this section we generalize this result to \mathbb{Z}^d . The inductive construction is similar to that in one dimension, but the added dimensions complicate the counting of ways in which configurations can overlap, and so in the initial levels of the construction we include all details. When the proofs become close enough that changes are mainly changing notation, we omit the details.

w	w	u_1	w	u_2	w	w	w	w	w	w	w	w
-----	-----	-------	-----	-------	-----	-----	-----	-----	-----	-----	-----	-----

FIGURE 1. The postcard $\mathcal{P}[u_1, u_2 \mid w, 6]: \{1, \dots, 13|w\} \rightarrow \{0, 1\}$

Similar to the one dimensional case, we construct a \mathbb{Z}^d -subshift on the alphabet $\{0, 1\}$, and the higher dimensional setting necessitates changes in some of the definitions. If $\mathcal{S} \subseteq \mathbb{Z}^d$, we refer to \mathcal{S} as a *shape* and define an \mathcal{S} -*word* to be a function $w: \mathcal{S} \rightarrow \{0, 1\}$. When w is an \mathcal{S} -word, we let $[w]$ denote any function whose domain is a translate of \mathcal{S} and we say $[w]$ *has the same shape* as w ; we let $|w|$ denote the number of integer points in \mathcal{S} .

For the special case that \mathcal{S} is rectangular, we make use of a notion of subwords. Assume that $\mathcal{S} = \prod_{i=1}^d \{1, 2, \dots, m_i\}$ for some $m_1, \dots, m_d \in \mathbb{N}$, $\mathcal{T} = \prod_{i=1}^d \{1, 2, \dots, m'_i\}$ with $m'_i \leq m_i$ for $i = 1, \dots, d$, let $w: \mathcal{S} \rightarrow \mathcal{A}$ be an \mathcal{S} -word and let $v: \mathcal{T} \rightarrow \mathcal{A}$ be a \mathcal{T} -word. We say that v occurs as a *subword* of w if there exist $k_1 \leq m_1 - m'_1, \dots, k_d \leq m_d - m'_d$ such that $v(x_1, \dots, x_d) = w(x_1 + k_1, \dots, x_d + k_d)$ for all $(x_1, \dots, x_d) \in \mathcal{T}$.

For fixed $d, e_1, \dots, e_d, n \geq 1$ and function $w: \{1, 2, \dots, n\}^d \rightarrow \{0, 1\}$, define the $(e_1 \times e_2 \times \dots \times e_d)$ -*fold self-concatenation* of w to be the function

$$w^{(e_1, \dots, e_d)}: \prod_{i=1}^d \{1, 2, \dots, e_i n\} \rightarrow \{0, 1\}$$

defined by

$$w^{(e_1, \dots, e_d)}(x_1, \dots, x_d) := w(x_1 \pmod{n}, x_2 \pmod{n}, \dots, x_d \pmod{n}).$$

For convenience, when d is clear from the context, we write $w^{(e)}$ as shorthand for $w^{(e, e, \dots, e)}$.

If $u_1, \dots, u_k: \{1, 2, \dots, n\}^d \rightarrow \{0, 1\}$ and $e \geq 2k + 4$, define

$$\mathcal{P}[u_1, \dots, u_k \mid w, e]: \{1, 2, \dots, (2e + 1)n\}^d \rightarrow \{0, 1\}$$

to be the *postcard function*, where we consider u_1, \dots, u_k to be the *stamp* (thought of as in the lower left corner) defined by:

Case 1 (the stamp): if $(x_1, \dots, x_d) \in \{2mn + 1, \dots, (2m + 1)n\} \times \{2n + 1, \dots, 3n\}^{d-1}$ for some $1 \leq m \leq k$, define

$$\mathcal{P}[u_1, \dots, u_k \mid w, e](x_1, \dots, x_d) := u_m(x_1 - 2mn, x_2 - 2n, \dots, x_d - 2n).$$

Case 2 (the rest of the postcard): if $(x_1, \dots, x_d) \notin \{2mn + 1, \dots, (2m + 1)n\} \times \{2n + 1, \dots, 3n\}^{d-1}$ for any $1 \leq m \leq k$, we define

$$\mathcal{P}[u_1, \dots, u_k \mid w, e](x_1, \dots, x_d) := w^{(e)}(x_1, \dots, x_d)$$

Examples of postcards in dimensions 1 and 2 are given in Figures 1 and 2.

We define $\mathcal{N}(u, v)$ to be the number of times u occurs as a subword of v . A useful feature of the postcard function is that if u never occurs as a subword of $w^{(2)}$, then for all $1 \leq i \leq k$ we have

$$\lim_{e \rightarrow \infty} \frac{\mathcal{N}(u, \mathcal{P}[u_1, \dots, u_k \mid w, e])}{|\mathcal{P}[u_1, \dots, u_k \mid w, e]|} = 0.$$

w	w	w	w	w	w	w	w	w
w	w	w	w	w	w	w	w	w
w	w	w	w	w	w	w	w	w
w	w	w	w	w	w	w	w	w
w	w	w	w	w	w	w	w	w
w	w	w	w	w	w	w	w	w
w	w	w	w	w	w	w	w	w
w	w	u_1	w	u_2	w	w	w	w
w	w	w	w	w	w	w	w	w
w	w	w	w	w	w	w	w	w

FIGURE 2. The postcard $\mathcal{P}[u_1, u_2 | w, 4]: \{1, \dots, 9|w|\}^2 \rightarrow \{0, 1\}$

5.2. The construction. This construction proceeds recursively, in levels, and is highly reminiscent of the construction of a \mathbb{Z} -CAM subshift. The alphabet is $\mathcal{A} = \{0, 1\}$, and again, all parameters are tuned after the fact to give the resulting system its properties. The first three levels are special and so we construct them explicitly, and the recursive formula starts with level four. As previously, the second coordinate of a subscript always denotes the level.

5.2.1. *The frequency sequence.* Pick a sequence $\{\varepsilon_k\}_{k=1}^{\infty}$ of non-negative real numbers that satisfy

$$\sum_{k=N}^{\infty} \varepsilon_k < \frac{1}{3^{d+N-1}}.$$

for all $N \in \mathbb{N}$. We use this sequence throughout our construction.

5.2.2. *Level 1 (introduction of the words w).* There are 2 words on this level which are both functions from $\{1\}^d$ to \mathcal{A} , and we set

$$w_{(1,1)}(1, \dots, 1) = 0 \quad \text{and} \quad w_{(2,1)}(1, \dots, 1) = 1$$

5.2.3. *Level 2 words (introduction of the words a and b).* There are 4 words on this level, playing two distinct roles as distinguished by two types: we have periodic words w and density words a, b . This level also introduces a parameter $n_2 > 1$ that controls the densities (we explain the meaning of this momentarily). All words at this level are functions $\{1, 2, \dots, n_2\}^d \rightarrow \mathcal{A}$. We define:

$$w_{(1,2)} = w_{(1,1)}^{(n_2)} \quad \text{self-concatenation of copies of } w_{(1,1)}$$

$$w_{(2,2)} = w_{(2,1)}^{(n_2)} \quad \text{self-concatenation of copies of } w_{(2,1)}$$

$$a_2 = \mathcal{P}[w_{(2,1)} | w_{(1,1)}, n_2] \quad \text{all } w_{(1,1)} \text{ except a single copy of } w_{(2,1)}$$

$$b_2 = \mathcal{P}[w_{(1,1)} | w_{(2,1)}, n_2] \quad \text{all } w_{(2,1)} \text{ except a single copy of } w_{(1,1)}$$

The parameter n_2 is chosen sufficiently large such that $\mathcal{N}(1, a_2)/|a_2| < \varepsilon_1$ and $\mathcal{N}(0, b_2)/|b_2| < \varepsilon_1$. In other words, a_2 consists almost entirely of 0's and b_2 almost entirely of 1's. More explicitly, $a_2, b_2: \{1, 2, \dots, n_2\}^d \rightarrow \mathcal{A}$ and the only

point (x_1, \dots, x_d) in the domain of a_2 where $a_2(x_1, \dots, x_d) = 1$ is $(x_1, \dots, x_d) = (3, 3, \dots, 3)$. Similarly for $b_2(x_1, \dots, x_d) = 0$.

We call the four words introduced at this point the *level-2 words* and we write $\mathcal{W}_2 = \{w_{(1,2)}, w_{(2,2)}, a_2, b_2\}$. For use in verifying properties of the words constructed at higher levels, we describe the basic properties of the words constructed at level 2.

Proposition 5.1. *For all distinct words $u, v \in \mathcal{W}_2$, the word u does not occur as a subword of $v^{(2)}$.*

Proof. First we establish the statement when $v \in \{w_{(1,2)}, w_{(2,2)}\}$ and $u \in \mathcal{W}_2$ is distinct from v . Neither a_2 nor b_2 occurs as a subword of $w_{(i,2)}^{(2)}$ for $i = 1, 2$, because $w_{(i,2)}^{(2)}$ is a constant function, whereas a_2 and b_2 are not. The word $w_{(1,2)}$ does not occur as a subword of $w_{(2,2)}^{(2)}$, since the symbol 0 does not occur in $w_{(2,2)}$ but does occur in $w_{(1,2)}$ and similarly $w_{(2,2)}$ does not occur as a subword of $w_{(1,2)}^{(2)}$. This establishes the statement when $v \in \{w_{(1,2)}, w_{(2,2)}\}$.

Next we establish the statement when $v \in \{a_2, b_2\}$ and $u \in \mathcal{W}_2$ is distinct from v . Neither $w_{(1,2)}$ nor $w_{(2,2)}$ occurs as a subword of $a_2^{(2)}$, as the largest m for which there is a constant subword of shape $\{1, 2, \dots, m\}^d$ in $a_2^{(2)}$ satisfies $m < n_2$, whereas any occurrence of $w_{(i,2)}$ forces a constant subword of shape $\{1, 2, \dots, n_2\}^d$. Similarly, neither can occur as a subword of $b_2^{(2)}$. We claim that the word a_2 does not occur as a subword of $b_2^{(2)}$ (the argument that b_2 does not occur as a subword of $a_2^{(2)}$ is analogous). By construction, $\mathcal{N}(1, a_2)/|a_2| < \varepsilon_1$ and so $\mathcal{N}(1, a_2) < \varepsilon_1 |a_2|$. Therefore $\mathcal{N}(0, a_2) \geq (1 - \varepsilon_1)|a_2|$. Similarly, using the fact that $b_2^{(2)}$ consists of concatenated copies of b_2 , we get $\mathcal{N}(1, b_2^{(2)}) \geq (1 - \varepsilon_1)|b_2^{(2)}|$. If a_2 occurs as a subword of $b_2^{(2)}$ then, using the fact that $|a_2| = \frac{1}{2^d}|b_2^{(2)}|$, we obtain

$$\mathcal{N}(0, b_2^{(2)}) \geq \mathcal{N}(0, a_2) \geq (1 - \varepsilon_1)|a_2| = \frac{1 - \varepsilon_1}{2^d}|b_2^{(2)}|.$$

This implies that $\mathcal{N}(1, b_2^{(2)}) \leq (1 - \frac{1 - \varepsilon_1}{2^d})|b_2^{(2)}| = \frac{2^d - 1 + \varepsilon_1}{2^d}|b_2^{(2)}|$. Combined with our earlier observation that $\mathcal{N}(1, b_2^{(2)}) \geq (1 - \varepsilon_1)|b_2^{(2)}|$, we have that $1 - \varepsilon_1 \leq \frac{2^d - 1 + \varepsilon_1}{2^d}$. However, this contradicts our assumption that $\varepsilon_1 < \sum_{k=1}^{\infty} \varepsilon_k < \frac{1}{3^d} \leq \frac{1}{2^d + 1}$. Therefore the statement holds when $v \in \{a_2, b_2\}$. \square

5.2.4. *Level 3 words (recursive definitions of periodic and density words).* There are 6 words on this level, again playing distinct roles distinguished by two types: periodic words w and density words a, b . Again we introduce a parameter $n_3 > 1$

to control the densities. Define the words by setting

$$w_{(1,3)} = w_{(1,2)}^{(5n_3+4)}$$

$$w_{(2,3)} = w_{(2,2)}^{(5n_3+4)}$$

$$w_{(3,3)} = a_2^{(5n_3+4)}$$

$$w_{(4,3)} = b_2^{(5n_3+4)}$$

$$a_3 = \mathcal{P}[w_{(1,2)}, w_{(2,2)}, a_2, b_2 | a_2, n_3]$$

$$b_3 = \mathcal{P}[w_{(1,2)}, w_{(2,2)}, a_2, b_2 | b_2, n_3],$$

meaning that the first four words periodize the words $w_{(1,2)}$, $w_{(2,2)}$, a_2 , and b_2 (in order), the second to last word is mostly concatenated a_2 's with the words \mathcal{W}_2 as a stamp, and the last word is mostly concatenated b_2 's with the words \mathcal{W}_2 as a stamp. We choose n_3 such that $\mathcal{N}(1, a_3)/|a_3| < \varepsilon_1 + \varepsilon_2$ and $\mathcal{N}(0, b_3)/|b_3| < \varepsilon_1 + \varepsilon_2$ (we note that this is possible because $\mathcal{N}(1, a_2)/|a_2| < \varepsilon_1$ and a_3 is mainly made of concatenated copies of a_2 , similarly $\mathcal{N}(0, b_2)/|b_2| < \varepsilon_1$). We further assume that n_3 has been chosen sufficiently large such that for $u \in \{w_{(1,2)}, w_{(2,2)}, b_2\}$, we have

$$(16) \quad \frac{\mathcal{N}(u, a_3^{(2)})}{|a_3^{(2)}|} < \frac{\varepsilon_2}{|u|(2|u| - 1)^d},$$

and note that this can be done since $a_3^{(2)}$ is mainly made of concatenated copies of $a_2^{(2)}$ and u does not arise as a subword of $a_2^{(2)}$. Using the same reasoning, this choice of n_3 ensures that if $u \in \{w_{(1,2)}, w_{(2,2)}, a_2\}$, then we have

$$(17) \quad \frac{\mathcal{N}(u, b_3^{(2)})}{|b_3^{(2)}|} < \frac{\varepsilon_2}{|u|(2|u| - 1)^d}.$$

Finally note that by choosing n_3 sufficiently large, we can guarantee that $|a_2|/|a_3|$ is as small as desired, and for use in the proof of the next proposition describing occurrences of words and subwords, we need a specific bound. Let $a_2^{(\infty)}: \mathbb{Z}^d \rightarrow \mathcal{A}$ be the infinite self-concatenation of a_2 given by:

$$a_2^{(\infty)}(x_1, \dots, x_d) := a_2(x_1 \pmod{n_2}, \dots, x_d \pmod{n_2}).$$

Let $\mathcal{V} \subseteq \mathbb{Z}^d$ be the set of all period vectors of $a_2^{(\infty)}$: $(v_1, \dots, v_d) \in \mathcal{V}$ if and only if $a_2^{(\infty)}(x_1, \dots, x_d) = a_2^{(\infty)}(x_1 + v_1, \dots, x_d + v_d)$ for all $(x_1, \dots, x_d) \in \mathbb{Z}^d$. Then \mathcal{V} is a finite-index subgroup of \mathbb{Z}^d , let p be the index of \mathcal{V} in \mathbb{Z}^d . Note that at any location where a_2 occurs as a subword of $a_2^{(2)}$, say on the set $\prod_{i=1}^d \{y_i + 1, y_i + 2, \dots, y_i + n_2\}$, we get $(y_1, \dots, y_d) \in \mathcal{V}$. Therefore, the number of times a_2 occurs as a subword of

$a_2^{(2)}$ satisfies:

$$(18) \quad \frac{n_2^d}{p} \leq |\{1, \dots, 2n_2\}^d \cap \mathcal{V}| \leq |\{1, \dots, 2n_2 - 1\}^d \cap \mathcal{V}| + dn_2^{d-1} \\ = \frac{n_2^d}{p} + dn_2^{d-1} = n_2^d \cdot \left(\frac{1}{p} + \frac{d}{n_2} \right).$$

We call the words constructed on this level the *level-3 words* and set \mathcal{W}_3 to be the set of all level-3 words.

Proposition 5.2. *For any choice of distinct words $u, v \in \mathcal{W}_3$, the word u does not occur as a subword of $v^{(2)}$.*

Proof. We check that for each $v \in \mathcal{W}_3$ and $u \in \mathcal{W}_3 \setminus \{v\}$, the word u does not occur as a subword of $v^{(2)}$, checking cases depending on the type of word v .

First consider when $v = w_{(i,3)}$ for some $i \in \{1, 2\}$. We consider two cases, depending on the choice of u . First suppose that $u \in \{w_{(3,3)}, w_{(4,3)}, a_3, b_3\}$. By Proposition 5.1, neither a_2 nor b_2 occurs as a subword of $w_{(i,2)}^{(2)}$. The word $v^{(2)} = w_{(i,3)}^{(2)}$ is the self-concatenation of a large number of copies of $w_{(i,2)}$ and so neither a_2 nor b_2 occurs as a subword of $v^{(2)}$. But at least one of a_2 and b_2 occurs as a subword of u , and so u cannot be a subword of $v^{(2)}$. Next suppose $u = w_{(j,3)}$ where $j \in \{1, 2\}$ and $j \neq i$. The word $u = w_{(j,3)}$ is not a subword of $v^{(2)} = w_{(i,3)}^{(2)}$ because $w_{(j,2)}$ is not a subword of $w_{(i,2)}^{(2)}$ and, as noted, $v^{(2)}$ is the self-concatenation of copies of $w_{(i,2)}$ while $w_{(j,3)}$ is the self-concatenation of copies of $w_{(j,2)}$. We deduce that u does not occur as a subword of $v^{(2)}$ for any $u \in \mathcal{W}_3 \setminus \{v\}$. Thus the statement holds when $v \in \{w_{(1,3)}, w_{(2,3)}\}$.

Next consider when $v = w_{(i,3)}$ for some $i \in \{3, 4\}$. Again we have two cases, depending on the choice of u . First suppose $u \in \{w_{(1,3)}, w_{(2,3)}, a_3, b_3\}$. By Proposition 5.1, neither $w_{(1,2)}$ nor $w_{(2,2)}$ occurs as a subword of $a_2^{(2)}$ or of $b_2^{(2)}$. The word $v^{(2)} = w_{(i,3)}^{(2)}$ is the self-concatenation of a large number of copies of a_2 or of b_2 , and so neither $w_{(1,2)}$ nor $w_{(2,2)}$ occurs as a subword of $v^{(2)}$. However at least one of $w_{(1,2)}$ and $w_{(2,2)}$ occurs in u , and so u is not a subword of $v^{(2)}$. We next consider when $j \in \{3, 4\} \setminus \{i\}$ and let $u = w_{(j,3)}$. Then one of the words u and v is the self-concatenation of many copies of a_2 and the other is the self-concatenation of many copies of b_2 . By Proposition 5.1, a_2 is not a subword of $b_2^{(2)}$ and b_2 is not a subword of $a_2^{(2)}$, and so u is not a subword of $v^{(2)}$. This establishes the statement when $v \in \{w_{(3,3)}, w_{(4,3)}\}$.

Finally consider when $v \in \{a_3, b_3\}$. We give the argument when $v = a_3$, and the argument for b_3 is similar. Again, we have two cases, depending on the choice of u . First suppose $u \in \{w_{(1,3)}, w_{(2,3)}, w_{(4,3)}, b_3\}$. By (16), we know that $\mathcal{N}(x, a_3^{(2)})/|a_3^{(2)}| < \varepsilon_2/|x|(2|x| - 1)^d$ for $x \in \mathcal{W}_2 \setminus \{a_2\}$. For any such x and any particular occurrence of x in $a_3^{(2)}$, there are $(2|x| - 1)^d$ subwords of $a_3^{(2)}$ that have the shape $[x]$ and partially (or completely) overlap this occurrence of x . This means that, for any $x \in \mathcal{W}_2 \setminus \{a_2\}$, if we look at the collection of all locations within $a_3^{(2)}$

where x occurs, we have

$$\frac{\# \text{ of subwords with shape } [x] \text{ in } a_3^{(2)} \text{ that overlap an occurrence of } x}{|a_3^{(2)}|} < \frac{\varepsilon_2}{|x|}.$$

Recall that p is the index, in \mathbb{Z}^d , of the stabilizer subgroup of $a_2^{(\infty)}: \mathbb{Z}^d \rightarrow \{0, 1\}$ (when acted on by \mathbb{Z}^d translations) and the number of occurrences of a_2 as a subword of $a_2^{(2)}$ satisfies (18). Since all four words in \mathcal{W}_2 have the same size, we deduce from (16) that

$$(19) \quad \frac{\mathcal{N}(a_2, a_3^{(2)})}{|a_3^{(2)}|} \geq \frac{1}{p} - 3 \cdot \frac{\varepsilon_2}{|a_2|}$$

because $a_3^{(2)}$ is made by concatenating words in \mathcal{W}_2 . On the other hand, we claim that $\mathcal{N}(a_2, u) \leq 3 \cdot (2|a_2| - 1)^d$. To check this, note that if $u = b_3$, any occurrence of a_2 in u must partially overlap $w_{(1,2)}$, $w_{(2,2)}$ or a_2 in the definition of b_3 , because it does not occur as a subword of $b_2^{(2)}$, and there are only $3 \cdot (2|a_2| - 1)^d$ locations that have such overlaps. If $u \neq b_3$ then a_2 does not occur in u , by Proposition 5.1 and the claim follows.

Now we return to showing that u does not occur as a subword of $a_3^{(2)}$. For contradiction, suppose u does occur as a subword of $a_3^{(2)}$. Recall, from the definition, that $a_3^{(2)}$ is primarily made of concatenated copies of a_2 and there are three specific locations where the words $w_{(1,2)}$, $w_{(2,2)}$, b_2 occur instead. Since u occurs as a subword of $a_3^{(2)}$ this means u is primarily made of concatenated copies of a_2 . In fact, allowing for the possibility that u occurs within $a_3^{(2)}$ in a location that overlaps the copies of $w_{(1,2)}$, $w_{(2,2)}$, b_2 and that there may be no copies of a_2 that partially overlap these words, we can still use (18) to estimate that

$$\mathcal{N}(a_2, u) \geq ((n_2 n_3 - 2n_2)^d \cdot \frac{1}{p} - 3 \cdot (2n_2 - 1)^d).$$

In more detail, recall that $u: \{1, 2, \dots, n_2 n_3\}^d \rightarrow \{0, 1\}$ has a domain which is a cube in dimension d . The number $(n_2 n_3 - 2n_2)^d$ is the volume of the subcube obtained by removing the border of length n_2 from each side. Then, using (18), $((n_2 n_3 - 2n_2)^d \cdot \frac{1}{p})$ gives the minimum number of occurrences of a_2 within this subcube if we assume that the subcube is entirely populated by occurrences of $a_2^{(2)}$. Finally we subtract off $3 \cdot (2n_2 - 1)^d$ to account for the fact that this occurrence of u might overlaps the words $w_{(1,2)}$, $w_{(2,2)}$, b_2 in the definition of a_3 and that no word of size $|a_2|$ that partially overlaps these occurrences could be a_2 . Nonetheless, note that this lower bound on $\mathcal{N}(a_2, u)$ is larger than our previous upper bound of $3 \cdot (2|a_2| - 1)^d$ provided $|a_2|/|a_3|$ is sufficiently small (and we make this assumption).

Lastly, suppose that $u = w_{(3,3)}$. By construction, u is the self-concatenation of many copies of a_2 . If u occurred within $v^{(2)} = a_3^{(2)}$ then it would either contain or partially overlap an occurrence of $w_{(1,2)}$ which is used in the definition of a_3 . Since $w_{(1,2)}$ does not occur as a subword of $a_2^{(2)}$, this occurrence of u cannot contain $w_{(1,2)}$ as a subword so it can only occur in a location that partially overlaps $w_{(1,2)}$. But $|u| = |v|$ and u is the periodic concatenation of copies of a_2 , so this occurrence of u overlaps the copy of $w_{(1,2)}$ in all 2^d copies of a_3 that appear in the definition of $a_3^{(2)}$. By periodicity of u , $w_{(1,2)}$ occurs as a subword of $a_2^{(2)}$, which is impossible.

This establishes the statement when $v = a_3$. The argument for $v = b_3$ is similar, with the roles played by a_2 and b_2 switched. \square

5.2.5. *Preparing to build the level $k + 1$ words (inductive assumptions).* For $k \geq 3$, we construct the words on level $k + 1$ using the words on level k . Again they come in two varieties (periodic words and density words), and again we have a parameter $n_{k+1} > 1$ chosen to guarantee that the words have the desired properties. Inductively, we assume that we have constructed the level- k words: these are words $w_{(i,k)}$ for all $1 \leq i \leq 2k - 2$, words a_k and b_k , with all of these words having the same domain: $\{1, 2, \dots, \prod_{i=1}^k n_i\}^d$, where $n_1, \dots, n_k \in \mathbb{N}$ are parameters we assume have been defined already, and we denote the collection of all level- k words by \mathcal{W}_k . We further assume that these constructions satisfy the following properties:

- (1) For $1 \leq i \leq 2k - 4$, we have

$$w_{(i,k)} = w_{(i,k-1)}^{(n_k)}.$$

- (2) We have

$$w_{(2k-3,k)} = a_{k-1}^{(n_k)}.$$

- (3) We have

$$w_{(2k-2,k)} = b_{k-1}^{(n_k)}.$$

- (4) We have

$$a_k = \mathcal{P}[w_{(1,k-1)}, w_{(2,k-1)}, \dots, w_{(2k-4,k-1)}, a_{k-1}, b_{k-1} | a_{k-1}, n_k].$$

- (5) We have

$$b_k = \mathcal{P}[w_{(1,k-1)}, w_{(2,k-1)}, \dots, w_{(2k-4,k-1)}, a_{k-1}, b_{k-1} | b_{k-1}, n_k].$$

We further assume that for any word u constructed on some level m with $m < k$, other than when $u = a_m$, we have

$$(20) \quad \frac{\mathcal{N}(u, a_k^{(2)})}{|a_k^{(2)}|} < \frac{1}{|u|(2|u| - 1)^d} \sum_{j=m}^{k-1} \varepsilon_j$$

and that for any $m < k$ and any level- m word u other than $u = b_m$, we have

$$\frac{\mathcal{N}(u, b_k^{(2)})}{|b_k^{(2)}|} < \frac{1}{|u|(2|u| - 1)^d} \sum_{j=m}^{k-1} \varepsilon_j.$$

Finally we assume that

(21)

for any $u, v \in \mathcal{W}_k$ with $u \neq v$, the word u does not occur as a subword of $v^{(2)}$.

5.2.6. *Level- $(k + 1)$ words (recursive definition).* We apply the analogous procedure used for level 3 and construct the periodic and density words, this time assuming the properties given in (1) through (5). We define a parameter, n_{k+1} , whose properties are discussed below and we set the following notation.

- (1) For $1 \leq i \leq 2k - 2$, define the first group of periodic words by setting

$$w_{(i,k+1)} = w_{(i,k)}^{(n_{k+1})}.$$

- (2) Define the next periodic word by setting

$$w_{(2k-1,k+1)} = a_k^{(n_{k+1})}.$$

(3) Define the last periodic word by setting

$$w_{(2k,k+1)} = b_k^{(n_{k+1})}.$$

(4) Define the first type of density word by setting

$$a_{k+1} = \mathcal{P}[w_{(1,k-1)}, w_{(2,k-1)}, \dots, w_{(2k-2,k-1)}, a_k, b_k | a_k, n_{k+1}].$$

(5) Define the second type of density word by setting

$$b_{k+1} = \mathcal{P}[w_{(1,k-1)}, w_{(2,k-1)}, \dots, w_{(2k-2,k-1)}, a_k, b_k | b_k, n_{k+1}].$$

We choose the parameter n_{k+1} sufficiently large such that for every $m < k + 1$ and every word $u \in \mathcal{W}_m \setminus \{a_m\}$, we have

$$(22) \quad \frac{\mathcal{N}(u, a_{k+1}^{(2)})}{|a_{k+1}^{(2)}|} < \frac{1}{|u|(2|u| - 1)^d} \sum_{j=m}^k \varepsilon_j.$$

When $m < k$, recall that $\mathcal{N}(u, a_k^{(2)})/|a_k^{(2)}| < (\sum_{j=m}^{k-1} \varepsilon_j)/|u|(2|u| - 1)^d$ by (20). Therefore it is possible to satisfy (22) because, for n_{k+1} large, the word a_{k+1} primarily consists of concatenated copies of a_k and so we can make the left hand side of (22) as close as we want to $(\sum_{j=m}^{k-1} \varepsilon_j)/|u|(2|u| - 1)^d$; in particular we can make it less than the right hand side of (22). When $m = k$ it is possible to satisfy (22) because u does not occur as a subword of $a_k^{(2)}$ and, again, a_{k+1} is primarily made of concatenated copies of a_k . Similarly we choose n_{k+1} sufficiently large such that if $u \in \mathcal{W}_m \setminus \{b_m\}$, then we have

$$(23) \quad \frac{\mathcal{N}(u, b_{k+1}^{(2)})}{|b_{k+1}^{(2)}|} < \frac{1}{|u|(2|u| - 1)^d} \sum_{j=m}^k \varepsilon_j.$$

For our next estimate, let p_k be the index, in \mathbb{Z}^d , of the group of periods of the word $a_k^{(\infty)}$ defined by

$$a_k^{(\infty)}(x_1, \dots, x_d) := a_k(x_1 \pmod{n_k}, \dots, x_d \pmod{n_k}).$$

We finally require that n_{k+1} is sufficiently large such that

$$(24) \quad \frac{|a_k|}{|a_{k+1}|} < \frac{1}{(4k-2)p_k} - \frac{1}{3^k \cdot |a_k|}$$

This is possible as long as the right-hand side of the inequality is positive. Note that $p_k \leq |a_k|$ and that $4k - 2 < 3^k$ because $k \geq 3$. Note that all of the words constructed on level $k + 1$ have the same domain: $\{1, 2, \dots, \prod_{i=1}^{k+1} n_i\}^d$. We call the words constructed in (1) through (5) the *level- $(k + 1)$ words* and let \mathcal{W}_{k+1} denote the set of all level- $(k + 1)$ words. Therefore, the only thing that is required to complete our inductive construction is to prove the following proposition.

Proposition 5.3. *For any choice of distinct words $u, v \in \mathcal{W}_{k+1}$, the word u does not occur as a subword of $v^{(2)}$.*

Proof. The proof of this proposition is analogous to that of Proposition 5.2. Fixing some $v \in \mathcal{W}_{k+1}$, we argue that for all $u \in \mathcal{W}_{k+1} \setminus \{v\}$, the statement holds. Again, we split the argument into cases depending on the type of the word v .

First consider when $v \in \{w_{(i,k+1)} : 1 \leq i \leq 2k - 2\}$. We consider two cases, depending on the choice of u . First suppose that $u \in \{w_{(2k-1,k+1)}, w_{(2k,k+1)}, a_{k+1}, b_{k+1}\}$. By Proposition 5.1, neither a_k nor b_k occurs as a subword of $w_{(i,k)}^{(2)}$. The word

$v^{(2)} = w_{(i,k+1)}^{(2)}$ is the self-concatenation of a large number of copies of $w_{(i,k)}$ and so neither a_k nor b_k occurs as a subword of $v^{(2)}$. But at least one of a_k and b_k occurs as a subword of u , and so u cannot be a subword of $v^{(2)}$. Next suppose $u = w_{(j,k+1)}$ where $j \in \{1, 2, \dots, 2k-2\}$ and $j \neq i$. The word $u = w_{(j,k+1)}$ is not a subword of $v^{(2)} = w_{(i,k+1)}^{(2)}$ because $w_{(j,k)}$ is not a subword of $w_{(i,k)}^{(2)}$ and, as noted, $v^{(2)}$ is the self-concatenation of copies of $w_{(i,k)}$ while $w_{(j,k+1)}$ is the self-concatenation of copies of $w_{(j,k)}$. Thus u does not occur as a subword of $v^{(2)}$ for any $u \in \mathcal{W}_{k+1} \setminus \{v\}$. Thus the statement holds when $v \in \{w_{(i,k+1)} : 1 \leq i \leq 2k-2\}$.

Next consider when $v = w_{(i,k+1)}$ for some $i \in \{2k-1, 2k\}$. Again we have two cases, depending on the choice of u . First suppose $u \in \{w_{(j,k+1)} : 1 \leq j \leq 2k-2\} \cup \{a_{k+1}, b_{k+1}\}$. By Proposition 5.1, none of the words in the set $\{w_{(j,k)} : 1 \leq j \leq 2k-2\}$ occurs as a subword of $a_k^{(2)}$ or of $b_k^{(2)}$. But by construction, the word $w_{(i,k+1)}^{(2)}$ is the self-concatenation of a large number of copies of a_k or of b_k , whereas at least one of the words in $\{w_{(j,k)} : 1 \leq j \leq 2k-2\}$ occurs as a subword of u . Therefore u occurs as a subword of $v^{(2)}$. We next consider when $j \in \{2k-1, 2k\} \setminus \{i\}$ and let $u = w_{(j,k+1)}$. Then one of the words u and v is the self-concatenation of many copies of a_k and the other is the self-concatenation of many copies of b_k . By Proposition 5.1, a_k is not a subword of $b_k^{(2)}$ and b_k is not a subword of $a_k^{(2)}$, and so u is not a subword of $v^{(2)}$. This establishes the statement when $v \in \{w_{(2k-1,k+1)}, w_{(2k,k+1)}\}$.

Finally consider when $v \in \{a_{k+1}, b_{k+1}\}$. We give the argument when $v = a_{k+1}$, and the argument for b_{k+1} is similar. Again, we have two cases, depending on the choice of u . First suppose $u \in \mathcal{W}_{k+1} \setminus \{w_{(2k-1,k+1)}, a_{k+1}\}$. By (22), we know that $\mathcal{N}(x, a_{k+1}^{(2)})/|a_{k+1}^{(2)}| < \frac{1}{|x|(2|x|-1)^d} \sum_{j=m}^k$ for $x \in \mathcal{W}_k \setminus \{a_k\}$. For any such x and any particular occurrence of x in $a_{k+1}^{(2)}$, there are $(2|x|-1)^d$ subwords of $a_{k+1}^{(2)}$ that have shape $|x|$ and partially (or completely) overlap this occurrence of x . This means that, for any $x \in \mathcal{W}_k \setminus \{a_k\}$, if we look at the collection of all locations within $a_{k+1}^{(2)}$ where x occurs, we have

$$\frac{\# \text{ of subwords of shape } |x| \text{ in } a_{k+1}^{(2)} \text{ that overlap an occurrence of } x}{|a_{k+1}^{(2)}|} < \frac{1}{|x|} \sum_{j=m}^k.$$

Recall that p_k is the index, in \mathbb{Z}^d , of the stabilizer subgroup of $a_k^{(\infty)} : \mathbb{Z}^d \rightarrow \{0, 1\}$ (when acted on by \mathbb{Z}^d translations) and the number of occurrences of a_k as a subword of $a_k^{(2)}$ satisfies (18). Since all $2k-2$ words in \mathcal{W}_k have the same size, we deduce from (16) that

$$(25) \quad \frac{\mathcal{N}(a_k, a_{k+1}^{(2)})}{|a_{k+1}^{(2)}|} \geq \frac{1}{p_k} - (2k-2) \cdot \frac{1}{|x|} \sum_{j=m}^k$$

because $a_{k+1}^{(2)}$ is made by concatenating words in \mathcal{W}_k . On the other hand, we claim that $\mathcal{N}(a_k, u) \leq (2k-1) \cdot (2|a_k|-1)^d$. To check this, note that if $u = b_{k+1}$, any occurrence of a_k in u must partially overlap elements of $\{w_{(i,k)} : 1 \leq i < 2k-2\} \cup \{a_k\}$, in the postcard function definition of b_{k+1} , because it does not occur as a subword of $b_k^{(2)}$, and there are only $(2k-1) \cdot (2|a_k|-1)^d$ locations that

have such overlaps. If $u \neq b_{k+1}$ then a_k does not occur in u , by Proposition 5.1 and the claim follows.

Now we return to showing that u does not occur as a subword of $a_{k+1}^{(2)}$. For contradiction, suppose u does occur as a subword of $a_{k+1}^{(2)}$. Recall, from the definition, that $a_{k+1}^{(2)}$ is primarily made of concatenated copies of a_k and there are $2k - 1$ specific locations where the words in the set $\{w_{(i,k)} : 1 \leq i < 2k - 2\} \cup \{b_k\}$ occur instead. Since u occurs as a subword of $a_{k+1}^{(2)}$ this means u is primarily made of concatenated copies of a_k . In fact, allowing for the possibility that u occurs within $a_{k+1}^{(2)}$ in a location that overlaps the copies of $\{w_{(i,k)} : 1 \leq i < 2k - 2\} \cup \{b_k\}$ and that there may be no copies of a_k that partially overlap these words, we can still use (18) to estimate that

$$\mathcal{N}(a_k, u) \geq ((n_{k+1} - 2) \prod_{i=2}^k n_i)^d \cdot \frac{1}{p_k} - (2k - 1) \cdot (2n_k - 1)^d.$$

The reasoning for this is analogous to that of Level 3.

Lastly, suppose that $u = w_{(2k-3, k+1)}$. By construction, u is the self-concatenation of many copies of a_k . If u occurred within $v^{(2)} = a_{k+1}^{(2)}$ then it would either contain or partially overlap an occurrence of $w_{(1,k)}$ which is used in the postcard function definition of a_{k+1} . Since $w_{(1,k)}$ does not occur as a subword of $a_k^{(2)}$, this occurrence of u cannot contain $w_{(1,k)}$ as a subword so it can only occur in a location that partially overlaps $w_{(1,k)}$. But $|u| = |v|$ and u is the periodic concatenation of copies of a_k , so this occurrence of u overlaps the copy of $w_{(1,k)}$ in all 2^d copies of a_{k+1} that appear in the definition of $a_{k+1}^{(2)}$. By periodicity of u , $w_{(1,k)}$ occurs as a subword of $a_k^{(2)}$, which is impossible. This establishes the statement when $v = a_{k+1}$. The argument for $v = b_{k+1}$ is similar, with the roles played by a_k and b_k switched. \square

5.2.7. Summarizing the construction and its properties. The remainder of the construction is parallel to that given in the case when $d = 1$. We have built a unique function $\mathbf{x}: \mathbb{Z}^d \rightarrow \{0, 1\}$ whose restriction to the cube $\{1 - \prod_{i=1}^k n_i, \prod_{i=1}^k n_i\}^d$ is $a_k^{(2)}$, and the \mathbb{Z}^d -subshift we define is the orbit closure of this function in $\{0, 1\}^{\mathbb{Z}^d}$. The proof that this system is a weakly CAM system is parallel to that given when $d = 1$: any element of this system is either periodic (meaning has finite orbit under the action of \mathbb{Z}^d) or contains a_k as a subword for all k , and so has dense orbit. The fact that it is CAM again follows from Proposition 2.17 because the system is expansive and weakly CAM. The argument that this system has two non-atomic, ergodic measures is again parallel to that for $d = 1$: passing to a weak*-convergent subsequence of the empirical measures associated to the function \mathbf{x} , and doing the same with the analogous function $\{1 - \prod_{i=1}^k n_i, \prod_{i=1}^k n_i\}^d$ is $b_k^{(2)}$. As in the case when $d = 1$, almost every ergodic component of each of these measures are nonatomic and these measures are distinct because they give different measure to the cylinder set $\{x \in X : x(0, 0, \dots, 0) = 0\}$.

6. FURTHER DIRECTIONS

We conclude with several questions about properties of CAM systems.

Question 6.1. What are the automorphisms of a CAM G -system (X, T) , meaning what are all homeomorphisms $S: X \rightarrow X$ such that $S \circ T_g = T_g \circ S$ for all $g \in G$?

As an example, for the \mathbb{Z} -system defined in 4.1.7, we can easily check that the automorphism group of this particular system contains the bit flip map (interchanging the letters 0 and 1). However, we have little further insight into the automorphism group and its properties for this system or, more generally, whether there exist nontrivial constraints on the automorphism group of a CAM system.

Question 6.2. What are the possible entropies for a CAM system? What are the possible growth rates of complexity for CAM \mathbb{Z}^k -subshifts?

In particular, this question is of interest for \mathbb{Z} -CAM systems and we do not know the entropy of the system defined in 4.1.7. This leads to a related question about chaotic systems:

Question 6.3. Does every chaotic system with positive entropy have a proper infinite subsystem?

If there exists a positive entropy CAM system, then the answer is no, in contrast to the behavior along subsequences as shown in [13]. However, we do not know if any of our examples have positive entropy.

If a group G admits a faithful action with dense periodic points, then we show in Proposition 3.1 that the group is necessarily residually finite. It seems plausible that this condition is in fact sufficient, and so we propose the following:

Conjecture 6.4. Every residually finite group is CAM.

In Proposition 2.17, we show that an expansive and weakly CAM system is a CAM system. However the converse we show is only that a CAM system is weakly CAM, and we ask if this can be improved:

Question 6.5. Is every CAM system expansive?

Note that it is shown in [1, 10] that any \mathbb{Z} -CAM system satisfies sensitivity to initial conditions. Still, we suspect that the answer to the above question is no, but do not have a counterexample.

REFERENCES

- [1] J. BANKS, J. BROOKS, G. CAIRNS, G. DAVIS, AND P. STACEY. On Devaney’s definition of chaos. *Amer. Math. Monthly* **99** (1992), no. 4, 332–334.
- [2] Y. BENOIST AND J. -F. QUINT. Mesures stationnaires et fermés invariants des espaces homogènes. *Ann. of Math. (2)* **174** (2011), no. 2, 1111–1162.
- [3] D. BEREND. Multi-invariant sets on tori. *Trans. Amer. Math. Soc.* **280** (1983), no. 2, 509–532.
- [4] J. BOURGAIN, A. FURMAN, E. LINDENSTRAUSS, AND S. MOZES. Stationary measures and equidistribution for orbits of nonabelian semigroups on the torus. *J. Amer. Math. Soc.* **24** (2011), no. 1, 231–280.
- [5] M. BOYLE, D. LIND, AND D. RUDOLPH. The automorphism group of a shift of finite type. *Trans. Amer. Math. Soc.* **306** (1988), no. 1, 71–114.
- [6] R. L. DEVANEY. An introduction to chaotic dynamical systems. Second edition. Addison-Wesley Studies in Nonlinearity. Addison-Wesley Publishing Company, Advanced Book Program, Redwood City, CA, 1989.
- [7] C. DONG. On density of infinite subsets I. *Discrete Contin. Dyn. Syst.* **39** (2019), no. 5, 2343–2359.
- [8] T. DOWNAROWICZ AND X. YE. When every point is either transitive or periodic. *Colloq. Math.* **93** (2002), no. 1, 137–150.
- [9] H. FURSTENBERG. Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Systems Theory* **1** (1967), 1–49.

- [10] E. GLASNER AND B. WEISS. Sensitive dependence on initial conditions. *Nonlinearity* **6** (1993), no. 6, 1067–1075.
- [11] Y. GUIVARC'H AND A. N. STARKOV. Orbits of linear group actions, random walks on homogeneous spaces and toral automorphisms. *Ergodic Theory Dynam. Systems* **24** (2004), no. 3, 767–802.
- [12] K. HIRAIDE. Expansive homeomorphisms of compact surfaces are pseudo Anosov *Osaka J. Math.* **27** (1990), no. 1, 117–162.
- [13] W. HUANG, J. LI, AND X. YE. Positive entropy implies chaos along any infinite sequence *Trans. Moscow Math. Soc.* **82** (2021), 1–14.
- [14] A. JOHNSON AND D. RUDOLPH. Commuting endomorphisms of the circle. *Ergodic Theory Dynam. Systems* **12** (1992), no. 4, 743–748.
- [15] H. KATO. Continuum-wise expansive homeomorphisms. *Canad. J. Math.* **45** (1993), no. 3, 576–598.
- [16] B. KITCHENS. Expansive dynamics on zero-dimensional groups. *Ergodic Theory Dynam. Systems* **7** (1987), no. 2, 249–261.
- [17] W. KRIEGER. On the subsystems of topological Markov chains. *Ergodic Theory Dynam. Systems* **2** (1982), no. 2, 195–202.
- [18] D. LIND AND B. MARCUS. An introduction to symbolic dynamics and coding. Second edition. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2021
- [19] R. MAÑÉ. Expansive homeomorphisms and topological dimension. *Trans. Amer. Math. Soc.* **252** (1979), 313–319.
- [20] R. MUCHNIK. Semigroup actions on \mathbb{T}^n . *Geom. Dedicata* 110 (2005), 1–47.
- [21] R. PAVLOV AND S. SCHMIEDING. On the structure of generic subshifts. To appear, *Nonlinearity*.
- [22] I. PUTNAM. Cantor minimal systems. University Lecture Series, **70**. American Mathematical Society, Providence, RI, 2018.
- [23] M. RATNER. Raghunathan's topological conjecture and distributions of unipotent flows. *Duke Math. J.* **63** (1991), no. 1, 235–280.
- [24] D. RUDOLPH. $\times 2$ and $\times 3$ invariant measures and entropy. *Ergodic Theory Dynam. Systems* **10** (1990), no. 2, 395–406.
- [25] K. SCHMIDT. Dynamical systems of algebraic origin. Modern Birkhäuser Classics. Birkhäuser/Springer Basel AG, Basel, 1995.