SUBSYSTEMS OF SHIFTS OF FINITE TYPE OVER COUNTABLE AMENABLE GROUPS

by

Robert Bland

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Applied Mathematics

Charlotte

2024

Approved by:

Dr. Kevin McGoff

Dr. William Brian

Dr. Oleg Safronov

Dr. Erik Saule

ABSTRACT

ROBERT BLAND. Subsystems of shifts of finite type over countable amenable groups. (Under the direction of DR. KEVIN MCGOFF)

This dissertation is broadly concerned with the subsystem problem for subshifts on countable amenable groups: given a subshift X, what are its subsystems and can they be classified or characterized by a simple criterion? In particular, we focus on the restricted case where X is a subshift of finite type (SFT). We pursue a variety of approaches to the problem. Firstly, we utilize the theory of topological entropy to demonstrate that an SFT with positive entropy exhibits a ubiquity of subsystems. Specifically, we prove that for any countable amenable group G, if X is a G-SFT with positive topological entropy $h(X) > 0$, then the entropies of the SFT subsystems of X are dense in the interval $[0, h(X)]$. In fact, we prove a "relative" version of the same result: if X is a G-SFT and $Y \subset X$ is a subshift such that $h(Y) < h(X)$, then the entropies of the SFTs Z for which $Y \subset Z \subset X$ are dense in $[h(Y), h(X)]$. We also establish analogous results for sofic subshifts. These results generalize the results of Desai for $G = \mathbb{Z}^d$. Secondly, we present an embedding theorem which provides conditions under which a given subshift may be realized as a subsystem of a given SFT. Namely, the result we obtain is as follows. Let G be a countable amenable group with the comparison property. Let X be a strongly aperiodic subshift over G. Let Y be a strongly irreducible shift of finite type over G which has no global period, meaning that the shift action is faithful on Y. If $h(X) < h(Y)$ and Y contains at least one factor of X, then X embeds into Y. This result partially extends the classical result of Krieger for $G = \mathbb{Z}$ and the results of Lightwood for $G = \mathbb{Z}^d$ for $d \geq 2$. Our proofs rely on recent developments in the theory of tilings and quasi-tilings of amenable groups due to Downarowicz, Huczek, and Zhang.

ACKNOWLEDGEMENTS

This dissertation represents many hundreds of hours of work across several years, and it would not have been possible were it not for the support and guidance of many. First, I would like to express my gratitude to my advisor Dr. Kevin McGoff for his invaluable instruction, advice, and mentorship throughout my graduate experience. I would also like to thank the members of my dissertation committee, Dr. William Brian, Dr. Oleg Safronov, and Dr. Erik Saule for their consideration and guidance. I also extend my thanks to my research colleagues and mentors, especially Dr. Ronnie Pavlov for his collaboration and support, as well as Dr. Mike Boyle for his many helpful readings and comments on this work. I also appreciate the many wonderful faculty of the Department of Mathematics and Statistics at UNC Charlotte for their instruction and mentorship throughout my student experience, particularly Ms. Debra Shafer, Dr. Kim Harris, Dr. Charles Burnap, Dr. Evan Houston, Dr. Shaozhong Deng, and Ms. Jeanne-Marie Linker. I acknowledge the financial support of UNC Charlotte via the Graduate Assistant Support Plan (GASP), as well as the National Science Foundation grant DMS-1847144. Finally, I would like to express my deepest appreciation to my friends and family for their support throughout my academic experience. This work only exists today thanks to all of those mentioned here and more. Thank you.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

In dynamical systems, one studies the long term behavior and properties of systems which evolve over time. In the most general framework, one has a pair (X, T) where X is the state space of the system of interest and T is "the dynamics", which is generally an action of some group G on X by structure-preserving maps (continuous in the case that X is a topological space, and measure-preserving in the case that X is a measure space). A particular state may be mapped by T to any number of other states, indexed by the group G; this collection is termed the *orbit* of the given state. A common strategy for analyzing dynamical systems is to pass to a symbolic framework in which a state is replaced by a symbolic encoding of its orbit. In the case where $G = \mathbb{Z}$, this means encoding the evolution of a given state with a sequence of symbols, in which case the dynamics corresponds to the action of the *shift map* which shifts the sequence of symbols one position forward.

Let us establish this framework precisely. Let G be a countable group and let $\mathcal{A}$ be a finite alphabet of symbols, endowed with the discrete topology. A function $x : G \to \mathcal{A}$ is a labeling of the group G by symbols from $\mathcal{A}$. The set of all such labelings is the product space $\mathcal{A}^G$, which is compact and metrizable in the product topology. In this topology, for each $g \in G$ the map $\sigma^g : \mathcal{A}^G \to \mathcal{A}^G$ given by

$$\sigma^g(x)(h) = x(hg) \quad \forall x \in \mathcal{A}^G \ \forall h \in G$$

is a homeomorphism of $\mathcal{A}^G$. The effect of σ^g is to translate or "shift" a given labeling x by a given group element g. The collection $\sigma = (\sigma^g)_{g \in G}$ is an action of G on $\mathcal{A}^G$ which is termed the *shift action*, and the pair $(\mathcal{A}^G, \sigma)$ is a dynamical system termed the *full shift (on $\mathcal{A}$)*. A subsystem of $\mathcal{A}^G$ is given by a subset $X \subset \mathcal{A}^G$ which is closed and *shift-invariant*, meaning $\sigma^g(X) = X$ for all $g \in G$, together with the restricted shift action $\sigma|_X = (\sigma^g|_X)_{g \in G}$. The pair $(X, \sigma|_X)$ (often referring solely to X, since the dynamics are always given by the shift map) is a dynamical system called a *subshift*.

Symbolic dynamics is the subfield of dynamical systems theory which focuses on subshifts as the central objects of interest. The symbolic framework is clean, nearly universal, and allows many definitions as well as the statements of many theorems to be simplified (compare e.g. the definition of topological entropy for general topological dynamical systems to Definition 2.2.14, the entropy of a subshift when G is an amenable group). The symbolic framework also allows one to leverage intuitive, combinatorial arguments for the analysis of dynamical systems. Symbolic dynamics is deeply connected with the broader field of dynamical systems, and it has also found many important independent connections and applications to various fields, from formal logic and theoretical computer science to statistical mechanics and stochastic processes. The field of symbolic dynamics is old and mature, with many of its most classical and fundamental results now over fifty years old, but at the same time many of the most natural and fundamental questions about subshifts remain open.

In this work, we are centrally interested in the *subsystem problem* for subshifts.

Problem. Let G be a countable group and let X be a subshift over G. What are the subsystems of X? Can they be classified or characterized by a simple criterion?

This is a natural, fundamental question which has analogues in many other areas of math. In algebra, one is interested to know what are the subgroups of a given group; etc. The general problem remains open, as it is too broad to be tractable, but in this dissertation we pursue various weakenings of the problem and establish various results in certain restricted cases.

One common restriction is to assume that X is a *subshift of finite type (SFT)*, which is our primary setting here. An SFT is a subshift which is specified by finitely many conditions about which configurations of symbols are "forbidden" from appearing in points of the subshift (Definition 2.2.12). SFTs are among the most well studied and well understood subshifts in dynamics, as their finitary nature makes them particularly amenable to combinatorial arguments. In the case where $G = \mathbb{Z}$, one has the interpretation of an SFT as the set of all bi-infinite walks in a connected directed graph (equivalently: an SFT is the set of all sequences of symbols accepted by a given finite state machine). In the case where $G = \mathbb{Z}^d$, one has the interpretation of covering the group in all possible ways with square "tiles", subject to local constraints governing which tiles are allowed to lay adjacent to which other tiles. All subshifts are specified by local contraints on symbols in this way, but an SFT is specified by only finitely many such constraints.

It is also common to take some restrictions on the group G. The classical theory of symbolic dynamics corresponds to the case where $G = \mathbb{Z}$, which has been studied in some form for the past seventy years. Interest for the case where $G = \mathbb{Z}^d$ has increased in the past thirty years, as connections to computability and recursion theory have been explored. Even more recently, the study of subshifts has been taken beyond $\mathbb{Z}^d$ to more general classes of groups, in particular to *amenable* groups (Definition 2.2.2). An amenable group is one which may in some sense be "well approximated" by its large finite subsets, in the same sense that $[-n, n]$ is a good finitary approximation of $\mathbb{Z}$. These large finite subsets provide "approximately shift-invariant" domains upon which the dynamics of a given subshift may be analyzed and related to the dynamics on the whole group. In this work, we assume throughout that the group G is amenable.

One of the main technical tools we employ in our analysis of subshifts is the theory of *entropy* (Definition 2.2.14). The entropy of a subshift X is a nonnegative real number $h(X)$ defined by the condition that, if one takes larger and larger finite subsets $F \subset G$, then the number of ways to label F with symbols while satisfying the constraints of X is asymptotically $\exp(h(X)|F|)$. In this way, $h(X)$ is a measure of the complexity of X; if X and Y are subshifts with $h(X) < h(Y)$, then at a sufficiently large scale Y exhibits more "dynamical variety" than X does. This definition coincides with the usual definition of entropy for general topological dynamical systems. The number $h(X)$ may also be interpreted as the average amount of information per symbol needed to distinguish a typical point of the subshift; this points to the connection between topological entropy and measure-theoretic entropy.

The other main technology exploited in our proofs is that of *tilings and quasi-tilings* of the group G (Definition 3.2.14). A quasi-tiling of G is a collection of translates ("tiles") of one of finitely many finite subsets of G ("shapes"). In an ideal case, the tiles are disjoint and cover G (i.e., the tiles form a partition of G). Such tilings are easy to obtain for $G = \mathbb{Z}^d$, where one may select the shapes to be large hypercubes. The existence of quasi-tilings exhibiting some form of "near-covering" and "near-disjointness" has been known for over thirty years, due first to Ornstein and Weiss [1]. Recently, there have been many developments in the theory of quasi-tilings of amenable groups due to Downarowicz, Huczek, and Zhang [2, 3, 4]. These have allowed for arguments and approaches typical in the analysis of $\mathbb{Z}^d$-subshifts to be extended to the analysis of G-subshifts for arbitrary amenable G. Tilings are useful for analyzing the dynamics of subshifts because they split the group G into large pieces of a uniformly bounded size; knowledge of the dynamics on the finite subdomains can be stitched together to make conclusions about the dynamics on the whole group. In this work, we typically use tilings to construct maps from one subshift to another; the map "reads" the labeling of the input point in each tile, then "writes" a labeling in the corresponding tile in the output point according to some finite rule. Because the map is "locally determined" in this way, the map is continuous and shift-commuting.

We now state our main results. In Chapter 2, we investigate the question of *how many* subsystems a given SFT has. We demonstrate that an SFT with positive entropy exhibits a ubiquity of subsystems. Specifically, the result we obtain is as follows.

Theorem 2.4.2. *Let G be a countable amenable group, let X be a G-SFT, and let $Y \subset X$ be any subsystem such that $h(Y) < h(X)$. Then*

$$\{h(Z) : Y \subset Z \subset X \text{ and } Z \text{ is an SFT}\}$$

is dense in $[h(Y), h(X)]$.

This has the consequence that an SFT X over a countable amenable group with positive entropy exhibits a subsystem Y (not necessarily an SFT) with $h(Y) = r$ for any choice of $r \in [0, h(X)]$ (Corollary 2.5.3). This provides a partial answer to Problem 1, demonstrating that the collection of subsystems of a positive entropy SFT is very rich and well populated. We also obtain analogous results for *sofic* subshifts (Definition 2.2.13), which are those subshifts obtained by taking a *factor* (an image under a continuous and shift-commuting map) of an SFT. We obtain the same result as above, with the keyword "SFT" replaced everywhere by "sofic" (Theorem 2.5.2). These results generalize those of Desai [5] which address the case where $G = \mathbb{Z}^d$.

In Chapter 3, we consider the finer problem of characterizing the subsystems of a given SFT. We present an embedding theorem which provides sufficient conditions under which in a restricted case a given subshift may be realized as a subsystem of a given SFT. Specifically, the result we obtain is as follows.

Theorem 3.3.5. *Let G be a countable amenable group with the comparison property. Let X be a nonempty strongly aperiodic subshift over G. Let Y be a strongly irreducible SFT over G with no global period. If $h(X) < h(Y)$ and Y contains at least one factor of X, then X embeds into Y.*

Notice we make a further restriction on the group G here, namely that it exhibits the *comparison property.* We do not attempt to define the comparison property here, but we note that the class of amenable groups satisfying the comparison property includes all elementary amenable groups (those obtained by taking subgroups, quotients, extensions, and directed unions of finite and abelian groups) [6, Theorem A] and all subexponential groups (those for which every finitely generated subgroup is of subexponential growth) [4, Theorem 6.33]. It is conjectured that all amenable groups satisfy the comparison property. This result partially extends the classical embedding theorem of Krieger [7] in the case where $G = \mathbb{Z}$; it also generalizes the results and approach of Lightwood [8] in the case where $G = \mathbb{Z}^d$.

This dissertation is adapted from the articles [9] (Chapter 2) and [10] (Chapter 3) which were published open-access by Cambridge University Press in the Journal of Ergodic Theory and Dynamical Systems. For this reason, notational convention may differ slightly between the two chapters. We refer the reader to sections §2.2 and §3.2 which establish notation and other preliminary concepts for Chapters 2 and 3 respectively.

CHAPTER 2: SUBSYSTEM ENTROPIES OF SHIFTS OF FINITE TYPE AND SOFIC SHIFTS ON COUNTABLE AMENABLE GROUPS

2.1 Introduction

Let G be a countable group and let $\mathcal{A}$ be a finite alphabet of symbols. In symbolic dynamics, the central objects of study are the subsystems of the so-called *full shift*, the dynamical system $(\mathcal{A}^G, \sigma)$, where σ denotes the action of G on $\mathcal{A}^G$ by translations (Definition 2.2.5). *Shifts of finite type* (Definition 2.2.12) and *sofic shifts* (Definition 2.2.13) are the most widely studied and well understood examples of symbolic dynamical systems. In each of these cases, the system of interest is completely specified by a finite amount of information. This allows for combinatorial, finitary arguments to be applied to the analysis of the dynamics of such systems.

Entropy is one of the most fundamental invariants of a topological dynamical system. Many fundamental results from classical entropy theory (i.e., in the case where $G = \mathbb{Z}$) only generalize if G is an *amenable* group (Definition 2.2.2). Amenability allows one to "approximate" the group by a sequence of finite subsets in a way that is useful for studying dynamics. See Definition 2.2.14 for the definition of the entropy of a symbolic dynamical system on an amenable group.

In general, one would like to understand the structure of the collection of subsystems of a given subshift. In this chapter we study the entropies of the SFT subsystems of a given SFT, as well as the entropies of the sofic subsystems of sofic shifts. There are many existing results in the literature in the case where $G = \mathbb{Z}$. For example, the Krieger Embedding Theorem [7] characterizes the irreducible SFT subsystems of a given irreducible $\mathbb{Z}$-SFT. Additionally, Lind [11] has provided an algebraic characterization of the real numbers that are realized as the entropy of a $\mathbb{Z}$-SFT.

However, the situation is very different in cases where $G \neq \mathbb{Z}$. Even in the case where $G = \mathbb{Z}^d$ for $d > 1$, the classes of SFTs and sofic shifts behave quite differently. For example, Boyle, Pavlov, and Schraudner [12] have shown by example that the subsystems of $\mathbb{Z}^d$ sofic shifts can be badly behaved for $d > 1$ (in contrast with the case where $d = 1$). Moreover, Hochman and Meyerovitch [13] have characterized the real numbers that are realized as entropy of a $\mathbb{Z}^d$-SFT (with $d > 1$), but in contrast to the result of Lind mentioned above, the characterization is in algorithmic terms and unavoidably involves concepts from computability and recursion theory. Nonetheless, Desai [5] has shown that a $\mathbb{Z}^d$-SFT with positive entropy has a wealth of SFT subsystems (sharpening an earlier result of Quas and Trow [14]).

Theorem 2.1.1 ([5])**.** *Let $G = \mathbb{Z}^d$ for some $d \in \mathbb{N}$ and let X be a G-SFT such that $h(X) > 0$. Then*

$$\{h(Y) : Y \subset X \text{ and } Y \text{ is an SFT}\}$$

is dense in $[0, h(X)]$.

In recent years, several results of the $G = \mathbb{Z}$ and $G = \mathbb{Z}^d$ cases have seen extensions to larger classes of groups, especially amenable groups. To name a few: Barbieri [15] has classified the real numbers that are realized as the entropy of a G-SFT for many types of amenable G (extending the result of Hochman and Meyerovitch mentioned above); Frisch and Tamuz [16] have investigated the (topologically) generic properties of G-subshifts for arbitrary amenable G; Barbieri and Sablik [17] have shown how an arbitrary effective G-subshift, where G is finitely generated, may be simulated by a G'-SFT, where G' is the semidirect product $G' = \mathbb{Z}^2 \rtimes G$; and Huczek and Kopacz [18] have recently obtained a partial generalization of Boyle's lower entropy factor theorem [19] to countable amenable groups with the comparison property. In this vein, we prove the following generalization of Theorem 2.1.1 to arbitrary countable amenable groups.

Theorem 2.4.2. *Let G be a countable amenable group, let X be a G-SFT, and let $Y \subset X$ be any subsystem such that $h(Y) < h(X)$. Then*

$$\{h(Z) : Y \subset Z \subset X \text{ and } Z \text{ is an SFT}\}$$

is dense in $[h(Y), h(X)]$.

Choosing $G = \mathbb{Z}^d$ and $Y = \varnothing$ in the above theorem recovers the result of Desai (Theorem 2.1.1 above). Note that a shift space $X \subset \mathcal{A}^G$ has at most countably many SFT subsystems, and therefore the set of entropies of SFT subsystems is at most countable. In this sense, Theorem 2.4.2 is “the most one could hope for" in terms of the ubiquity of SFT subsystems that a given SFT may exhibit.

Remark 2.1.2. After a preprint of this work was made public, the authors of [16] made us aware that a short alternate proof of Theorem 2.4.2 can be derived from their main results. Specifically, they prove there that for any countable amenable group G and any real $c \geq 0$, the set of G-subshifts with entropy c is dense (in fact residual) within the space of G-subshifts with entropy at least c with respect to the Hausdorff topology. This result immediately implies that for any G-SFT X, there exist G-subshifts contained in X that achieve all possible entropies in $[0, h(X)]$; then, some simple approximations with G-SFTs (in the sense of our Theorem 2.2.12) can be used to obtain a proof of Theorem 2.4.2.

For sofic shifts, we obtain the following result.

Theorem 2.5.2. *Let G be a countable amenable group, let W be a sofic G-shift, and let $V \subset W$ be any subsystem such that $h(V) < h(W)$. Then*

$$\{h(U) : V \subset U \subset W \text{ and } U \text{ is sofic}\}$$

is dense in $[h(V), h(W)]$.

From this result, we can quickly derive the fact (Corollary 2.5.3) that if X is a sofic G-shift, then each real number in $[0, h(X)]$ can be realized as the entropy of some (not necessarily sofic) subsystem of X. (Recall that the alternate proof of Theorem 2.4.2 described in Remark 2.1.2 above relies on a version of this result requiring X to be an SFT.) The tool for proving Theorem 2.5.2 (from Theorem 2.4.2) is provided by the following theorem, which may be of independent interest. We note that this result generalizes another theorem of Desai [5, Proposition 4.3], which addressed the case $G = \mathbb{Z}^d$.

Theorem 2.5.1. *Let G be a countable amenable group and let W be a sofic G-shift. For every $\varepsilon > 0$, there exists an SFT $\tilde{X}$ and a one-block code $\tilde{\phi} : \tilde{X} \to W$ such that the maximal entropy gap of $\tilde{\phi}$ satisfies $\mathcal{H}(\tilde{\phi}) < \varepsilon$.*

The maximal entropy gap $\mathcal{H}(\tilde{\phi})$ is defined in §2.2 (Definition 2.2.16). In particular, this result implies that if Y is sofic and $\varepsilon > 0$, then there is an SFT X that factors onto Y and satisfies $h(X) < h(Y) + \varepsilon$.

Our proofs of Theorems 2.4.2, 2.5.1, and 2.5.2 take the same general approach as the arguments given by Desai for the $G = \mathbb{Z}^d$ case. However, the extension to the general amenable setting requires substantial new techniques. Indeed, our proofs are made possible by the existence of *exact tilings* (Definition 2.3.1) of the group G that possess nice dynamical properties. Such exact tilings are trivial to find for $\mathbb{Z}^d$ (by tiling the group using large hypercubes), but for arbitrary amenable groups were only recently constructed by Downarowicz, Huczek, and Zhang [3]; their construction is the main technical tool employed in this chapter.

As mentioned in Remark 2.1.2 above, Theorem 2.4.2 can be alternately derived from results in [16]. We present a self-contained proof here for two reasons. Firstly, we would like to present a direct adaptation of the techniques from [5], since it demonstrates the power of the improved tiling results of [3]. Secondly, this presentation provides a unified approach to all of our proofs, since our proofs in the sofic setting (where we are not aware of alternative proofs) also rely on tiling-based constructions that are similar to those in our proof of Theorem 2.4.2.

This chapter is organized as follows. In §2.2 we discuss basic notions and elementary theorems of symbolic dynamics, set in terms appropriate for countable amenable groups. In §2.3 we define and explore the concept of tilings and exact tilings of amenable groups, appealing to Downarowicz, Huczek, and Zhang for the existence of certain desirable tilings. In §2.4 we prove our main results for G-SFTs, and in §2.5 we prove our main results for sofic G-shifts. Finally, in §2.6 we provide a example of a $\mathbb{Z}^2$ sofic shift whose only SFT subsystem is a fixed point.

2.2 Basics of symbolic dynamics

2.2.1 Amenable groups

We begin with a brief overview of amenable groups.

Definition 2.2.1 (Group theory notations)**.** Let G be a group and let K, $F \subset G$ be subsets. We employ the following notations.

i. The *group identity* is denoted by the symbol $e \in G$,

ii. $KF = \{kf : k \in K \text{ and } f \in F\}$,

iii. $K^{-1} = \{k^{-1} : k \in K\}$,

iv. $Kg = \{kg : k \in K\}$ for each $g \in G$,

v. $K \sqcup F$ expresses that K and F are disjoint, and is their *(disjoint) union*,

vi. $K \triangle F = (K \setminus F) \sqcup (F \setminus K)$ is the *symmetric difference* of K and F, and

vii. $|K|$ is the *cardinality* of the (finite) set K.

Definition 2.2.2 (Følner condition for amenability)**.** Let G be a countable group. A *Følner sequence* is a sequence $(F_n)_n$ of finite subsets $F_n \subset G$ which *exhausts* G (in the sense that for each $g \in G$, we have $g \in F_n$ for all sufficiently large n) and for which it holds that

$$\lim_{n\to\infty} \frac{|KF_n \triangle F_n|}{|F_n|} = 0$$

for every finite subset $K \subset G$. If such a sequence exists, then G is said to be an *amenable* group.

Throughout this chapter, G denotes a fixed countably infinite amenable group and $(F_n)_n$ is a fixed Følner sequence for G.

Definition 2.2.3 (Invariance)**.** Let K, $F \subset G$ be finite subsets, and let $\varepsilon > 0$. We say F is (K, ε)-*invariant* if

$$\frac{|KF \triangle F|}{|F|} < \varepsilon.$$

If $e \in K$ and F is (K, ε)-invariant, then F is also (K', ε')-invariant for any $\varepsilon' > \varepsilon$ and any $K' \subset K$ such that $e \in K'$. If F is (K, ε)-invariant, then so is the translate Fg for each fixed $g \in G$. Invariance is the primary way by which we say a large finite subset $F \subset G$ is a "good finite approximation" of G, according to the finitary quantifiers K and ε. The amenability of G provides a wealth of nearly invariant sets, which enables such approximation for the purpose of studying the dynamics of G-actions.

Next we develop concepts related to the geometry of finite subsets of G.

Definition 2.2.4 (Boundary and interior)**.** Let K, $F \subset G$ be finite subsets. The K-*boundary* of F is the set

$$\partial_K F = \{f \in F : Kf \not\subset F\},$$

and the K-*interior* of F is the set

$$\operatorname{int}_K F = \{f \in F : Kf \subset F\}.$$

Observe that $F = (\partial_K F) \sqcup (\operatorname{int}_K F)$.

If F is sufficiently invariant with respect to K, then the K-boundary of F is a small subset of F (proportionally), by the following lemma.

Lemma 2.2.1. *Suppose K, $F \subset G$ are nonempty finite subsets and $e \in K$. Then*

$$\frac{1}{|K|}|KF \triangle F| \leq |\partial_K F| \leq |K||KF \triangle F|.$$

In particular, if F is (K, ε)-invariant then $|\partial_K F| < \varepsilon |K||F|$.

Proof. If $e \in K$, then $KF \triangle F = KF \setminus F$. If $g \in KF \setminus F$, then $g = kf$ for some $k \in K$ and $f \in \partial_K F$, by Definition 2.2.4. Therefore $KF \setminus F \subset K\partial_K F$, in which case $|KF \setminus F| \leq |K||\partial_K F|$.

For the second inequality, note that $f \in \partial_K F$ implies $\exists k \in K$ such that $kf \notin F$, therefore $g = kf \in KF \setminus F$ is a point such that $f \in K^{-1}g \subset K^{-1}(KF \setminus F)$. Consequently $\partial_K F \subset K^{-1}(KF \setminus F)$, in which case $|\partial_K F| \leq |K||KF \setminus F|$.

Finally if F is (K, ε)-invariant, then $|\partial_K F| \leq |K||KF \setminus F| < \varepsilon |K||F|$. □

Given finite subsets K, $F \subset G$, in this chapter we focus on the KK^{-1}-boundary and KK^{-1}-interior of F (rather than the K-boundary and K-interior), and we make use of the following lemma.

Lemma 2.2.2. *Let K, $F \subset G$. For any translate Kg of K (for any $g \in G$), either $Kg \subset F$ or $Kg \subset \left(\operatorname{int}_{KK^{-1}} F\right)^c$ (or both are true).*

Proof. Suppose $Kg \not\subset \left(\operatorname{int}_{KK^{-1}} F\right)^c$. Then $\exists f \in \operatorname{int}_{KK^{-1}} F$ such that $f \in Kg$, which implies $g \in K^{-1}f$ and hence $Kg \subset KK^{-1}f \subset F$. □

2.2.2 Shift spaces

Here we present necessary definitions from symbolic dynamics. See Lind and Marcus [20] for an introductory treatment of these concepts.

Definition 2.2.5 (Shifts and subshifts)**.** Let $\mathcal{A}$ be a finite set of symbols equipped with the discrete topology. A function $x : G \to \mathcal{A}$ is called an *$\mathcal{A}$-labelling* of G. By convention, we write x_g for the symbol $x(g) \in \mathcal{A}$ which is placed by x at $g \in G$. The set of all $\mathcal{A}$-labellings of G is denoted $\mathcal{A}^G$, which we equip with the product topology. For each $g \in G$, let $\sigma^g : \mathcal{A}^G \to \mathcal{A}^G$ denote the map given by

$$(\sigma^g x)_h = x_{hg} \quad \forall h \in G$$

for each $x \in \mathcal{A}^G$. The collection $\sigma = (\sigma^g)_{g \in G}$ is an action of G on $\mathcal{A}^G$ by homeomorphisms. The pair $(\mathcal{A}^G, \sigma)$ is a dynamical system called the *full shift* over the alphabet $\mathcal{A}$. A subset $X \subset \mathcal{A}^G$ is called *shift-invariant* if $\sigma^g x \in X$ for each $x \in X$ and $g \in G$. A closed, shift-invariant subset $X \subset \mathcal{A}^G$ is called a *subshift* or a *shift space*. For a given $x \in \mathcal{A}^G$, the *orbit* of x is the subset $\mathcal{O}(x) = \{\sigma^g x : g \in G\} \subset \mathcal{A}^G$. The subshift *generated by* x is the topological closure of $\mathcal{O}(x)$ as a subset of $\mathcal{A}^G$, and is denoted $\overline{\mathcal{O}}(x) \subset \mathcal{A}^G$.

Definition 2.2.6 (Codes and factors)**.** Let $\mathcal{A}_X$, $\mathcal{A}_W$ be finite alphabets and let $X \subset \mathcal{A}_X^G$ and $W \subset \mathcal{A}_W^G$ be subshifts. A map $\phi : X \to W$ is *shift-commuting* if $\phi \circ \sigma^g = \sigma^g \circ \phi$ for each $g \in G$; the map ϕ is said to be a *sliding block code* if it is continuous and shift-commuting; and ϕ is said to be a *factor map* if it is a surjective sliding block code. If a factor map exists from X to W, then W is said to be a *factor* of X and X is said to *factor onto* W. If a sliding block code $\phi : X \to W$ is invertible and bi-continuous, then ϕ is said to be a *topological conjugacy*, in which case X and W are said to be *topologically conjugate*.

Definition 2.2.7 (Products of shifts)**.** If $\mathcal{A}$ and Σ are finite alphabets, then $\mathcal{A} \times \Sigma$ is also a finite alphabet (of ordered pairs). If $X \subset \mathcal{A}^G$ and $T \subset \Sigma^G$ are subshifts, then we view the *dynamical direct product* $X \times T$ as a subshift of $(\mathcal{A} \times \Sigma)^G$, defined by $(x, t) \in X \times T$ if and only if $x \in X$ and $t \in T$. The shift space $X \times T$ factors onto both X and T via the projection maps π_X and π_T, given by $\pi_X(x,t) = x$ and $\pi_T(x,t) = t$ for each $(x,t) \in X \times T$.

Remark 2.2.3. Definition 2.2.7 above introduces an abuse of notation, as technically we have $(x,t) \in \mathcal{A}^G \times \Sigma^G \neq (\mathcal{A} \times \Sigma)^G$. However, if equipped with the G-action ς given by $\varsigma^g(x,t) = (\sigma^g x, \sigma^g t)$, then $\mathcal{A}^G \times \Sigma^G$ becomes a dynamical system that is topologically conjugate to $(\mathcal{A} \times \Sigma)^G$.

2.2.3 Patterns

In this section we describe *patterns* and their related combinatorics.

Definition 2.2.8 (Patterns)**.** Let $\mathcal{A}$ be a finite alphabet and let $F \subset G$ be a finite set. A function $p : F \to \mathcal{A}$ is called a *pattern*, said to be of *shape* F. The set of all patterns of shape F is denoted $\mathcal{A}^F$. The set of all patterns of any finite shape is denoted $\mathcal{A}^* = \bigcup_F \mathcal{A}^F$, where the union is taken over all finite subsets $F \subset G$.

Remark 2.2.4. Given a point $x \in \mathcal{A}^G$ and a finite subset $F \subset G$, we take $x(F)$ to mean the *restriction* of x to F, which is itself a pattern of shape F. Usually this is denoted $x|_F \in \mathcal{A}^F$, but we raise F from the subscript for readability.

Definition 2.2.9 (One-block code)**.** Let $\mathcal{A}_X$ and $\mathcal{A}_W$ be finite alphabets and let $X \subset \mathcal{A}_X^G$ and $W \subset \mathcal{A}_W^G$ be subshifts. A factor map $\phi : X \to W$ is said to be a *one-block code* if there exists a function $\Phi : \mathcal{A}_X \to \mathcal{A}_W$ with the property that

$$\phi(x)_g = \Phi(x_g), \quad \forall g \in G$$

for each $x \in X$.

Definition 2.2.10 (Occurrence)**.** Let $\mathcal{A}$ be a finite alphabet and let $F \subset G$ be a finite set. A pattern $p \in \mathcal{A}^F$ is said to *occur* in a point $x \in \mathcal{A}^G$ if there exists an element $g \in G$ such that $(\sigma^g x)(F) = p$. If $X \subset \mathcal{A}^G$ is a subshift, then the collection of all patterns of shape F occurring in any point of X is denoted by

$$\mathcal{P}(F, X) = \{(\sigma^g x)(F) \in \mathcal{A}^F : x \in X \text{ and } g \in G\}.$$

If $X \subset \mathcal{A}^G$ is a subshift and $F \subset G$ is a finite subset, then $|\mathcal{P}(F, X)| \leq |\mathcal{A}|^{|F|}$. If $F' \subset G$ is another finite subset, then $|\mathcal{P}(F \cup F', X)| \leq |\mathcal{P}(F, X)| \cdot |\mathcal{P}(F', X)|$. If $F' \subset F$ and $X' \subset X$, then $|\mathcal{P}(F', X')| \leq |\mathcal{P}(F, X)|$.

Definition 2.2.11 (Forbidden patterns)**.** Let $\mathcal{A}$ be a finite alphabet, let $F \subset G$ be a finite set and let $X \subset \mathcal{A}^G$ be a subshift. A pattern $p \in \mathcal{A}^F$ is said to be *allowed* in X if $p \in \mathcal{P}(F, X)$ (if p occurs in at least one point of X).

Given a (finite or infinite) collection of patterns $\mathcal{F} \subset \mathcal{A}^*$, a new subshift $X' \subset X$ may be constructed by expressly *forbidding* the patterns in $\mathcal{F}$ from occurring in points of X. We denote this by

$$X' = \mathcal{R}(X, \mathcal{F}) = \{x \in X : \forall p \in \mathcal{F},\ p \text{ does not occur in } x\}.$$

For a single pattern p, we abbreviate $\mathcal{R}(X, \{p\})$ as $X \setminus p$. The shift X is said to be *specified* by the collection $\mathcal{F}$ if $X = \mathcal{R}(\mathcal{A}^G, \mathcal{F})$.

2.2.4 Shifts of finite type

In this section, we define shifts of finite type and sofic shifts over G. We also discuss many related elementary facts.

Definition 2.2.12 (SFTs)**.** A subshift $X \subset \mathcal{A}^G$ is a *shift of finite type (SFT)* if there is a finite collection $\mathcal{F} \subset \mathcal{A}^*$ such that $X = \mathcal{R}(\mathcal{A}^G, \mathcal{F})$. For an SFT, it is always possible to take $\mathcal{F}$ in the form $\mathcal{F} = \mathcal{A}^K \setminus \mathcal{P}(K, X)$ for some large finite subset $K \subset G$. In this case, we say X is *specified* by (patterns of shape) K.

If $X \subset \mathcal{A}^G$ is an SFT specified by a finite subset $K \subset G$, then it holds that

$$x \in X \iff \forall g \in G \left((\sigma^g x)(K) \in \mathcal{P}(K, X)\right)$$

for each $x \in \mathcal{A}^G$. If K specifies X, then so does K' for any (finite) subset $K' \supset K$. If X and T are SFTs, then so is the dynamical direct product $X \times T$.

Definition 2.2.13 (Sofic shifts)**.** A subshift W is *sofic* if there exists an SFT X which factors onto W.

The following elementary facts are needed; we abbreviate the proofs as they are similar to the well-known the proofs in the case where $G = \mathbb{Z}$ (see [20]).

Proposition 2.2.5. *Let X be an SFT, let W be a sofic shift, and let $\phi : X \to W$ be a factor map. Then there exists an SFT $\tilde{X}$ and a topological conjugacy $\tilde{\phi} : \tilde{X} \to X$ such that the composition $\phi \circ \tilde{\phi} : \tilde{X} \to W$ is a one-block code.*

Proof. Because ϕ is continuous and shift-commuting, there exists a large finite subset $K \subset G$ such that for each x, $x' \in X$ and each $g \in G$, it holds that

$$(\sigma^g x)(K) = (\sigma^g x')(K) \implies \phi(x)_g = \phi(x')_g.$$

Suppose that $e \in K$ and that $\mathcal{P}(K, X)$ specifies X as an SFT. Let $\tilde{\mathcal{A}} = \mathcal{P}(K, X)$ be a new finite alphabet, and let $\tilde{X} \subset \tilde{\mathcal{A}}^G$ be the set of all points $\tilde{x} \in \tilde{\mathcal{A}}^G$ such that

$$\exists x \in X, \forall g \in G, \tilde{x}_g = (\sigma^g x)(K).$$

Then $\tilde{X}$ is an SFT specified by patterns of shape $K^{-1}K$. The map $\tilde{\phi} : \tilde{X} \to X$ desired for the theorem is given by

$$\tilde{\phi}(\tilde{x})_g = (\tilde{x}_g)_e \in \mathcal{A}, \quad \forall g \in G, \forall \tilde{x} \in \tilde{X}.$$

□

Proposition 2.2.6. *For any subshift $X \subset \mathcal{A}^G$, there is a descending family of SFTs $(X_n)_n$ such that $X = \bigcap_n X_n$.*

Proof. Let $(p_n)_n$ enumerate $\{p \in \mathcal{A}^* : p \text{ does not occur in } X\}$, and for each n let

$$X_n = \mathcal{R}\big(\mathcal{A}^G, \{p_1, p_2, \ldots, p_n\}\big).$$

Then $(X_n)_n$ witnesses the result. □

Proposition 2.2.7. *Let $X \subset \mathcal{A}^G$ be a subshift and let $X_0 \subset \mathcal{A}^G$ be an SFT such that $X \subset X_0$. If $(X_n)_n$ is any descending family of subshifts such that $X = \bigcap_n X_n$, then $X_n \subset X_0$ for all sufficiently large n.*

Proof. Take $K \subset G$ to specify X_0 as an SFT. Note $\big(\mathcal{P}(K, X_n)\big)_n$ is a descending family of finite sets, and it is therefore eventually constant. In particular, we have

$$\mathcal{P}(K, X_n) = \mathcal{P}(K, X) \subset \mathcal{P}(K, X_0)$$

for all sufficiently large n. □

When $G = \mathbb{Z}^d$, SFTs are often reduced via conjugacy to so-called *1-step* SFTs, in which the allowed patterns are specified by allowed adjacent pairs of symbols. Such SFTs are often desired because they allow for a kind of “surgery" of patterns. If two patterns occur in two different labellings from a 1-step SFT, and yet they agree on their 1-boundaries, then the first may be *excised* and *replaced* by the second. This yields a new labelling which also belongs to the 1-step SFT. Although there is no obvious notion of 1-step SFTs when $G \neq \mathbb{Z}^d$, we do have the following result which allows for this sort of excision and replacement of patterns.

Lemma 2.2.8. *Let $X \subset \mathcal{A}^G$ be an SFT specified by $K \subset G$, let $F \subset G$ be a finite subset, and let x, $y \in X$ be two points such that x and y agree on $\partial_{KK^{-1}}F$. Then the point z, defined by $z_g = y_g$ if $g \in F$ and $z_g = x_g$ if $g \notin F$, also belongs to X.*

Proof. Let $g \in G$. By Lemma 2.2.2, either $Kg \subset F$ or $Kg \subset \big(\mathrm{int}_{KK^{-1}} F\big)^c$. In the first case, we have $(\sigma^g z)(K) = (\sigma^g y)(K)$ which is an allowed pattern in X. In the second case, we have $Kg \subset (F^c) \sqcup (\partial_{KK^{-1}}F)$. Since x and y agree on $\partial_{KK^{-1}}F$, we have $(\sigma^g z)(K) = (\sigma^g x)(K)$ which is again an allowed pattern in X. In either case, $(\sigma^g z)(K)$ is allowed in X for every g, hence $z \in X$. □

2.2.5 Entropy

Let $X \subset \mathcal{A}^G$ be a nonempty subshift. Recall that for a given large finite set $F \subset G$, the number of patterns of shape F that occur in any point of X is $|\mathcal{P}(F, X)|$, which is at most $|\mathcal{A}|^{|F|}$. As this grows exponentially (with respect to $|F|$), we are interested in the *exponential growth rate* of $|\mathcal{P}(F, X)|$ as F becomes very large and approaches the whole group G. For nonempty finite sets $F \subset G$, we let

$$h(F, X) = \frac{1}{|F|} \log |\mathcal{P}(F, X)|.$$

If F, $F' \subset G$ are disjoint finite subsets, then $h(F \sqcup F', X) \leq h(F, X) + h(F', X)$. This is because $|\mathcal{P}(F \sqcup F', X)| \leq |\mathcal{P}(F, X)| \cdot |\mathcal{P}(F', X)|$ and

$$\frac{1}{|F \sqcup F'|} = \frac{1}{|F| + |F'|} \leq \min\Big(\frac{1}{|F|}, \frac{1}{|F'|}\Big).$$

Definition 2.2.14 (Entropy)**.** Let X be a nonempty subshift. The *(topological) entropy* of X is the nonnegative real number $h(X)$ given by the limit

$$h(X) = \lim_{n\to\infty} h(F_n, X),$$

where $(F_n)_n$ is again the Følner sequence of G. For the empty subshift, we adopt the convention that $h(\varnothing) = 0$.

It is well-known that the limit above exists, does not depend on the choice of Følner sequence for G, and is an invariant of topological conjugacy (see [21]).

For any subshift $X \subset \mathcal{A}^G$ and any finite subset $F \subset G$ it holds that $h(F, X) \leq \log |\mathcal{A}|$, and consequently $h(X) \leq \log |\mathcal{A}|$. More generally, if X and X' are subshifts such that $X \subset X'$, then $h(F, X) \leq h(F, X')$ for every finite subset $F \subset G$ and consequently $h(X) \leq h(X')$. If X and X' are subshifts over $\mathcal{A}$, then so is $X \cup X'$ and $h(X \cup X') = \max\big(h(X), h(X')\big)$.

The following proposition is a classical fact; a proof is given in [21].

Proposition 2.2.9. *Let G be a countable amenable group. If a G-shift W is a factor of a G-shift X, then $h(W) \leq h(X)$.*

Frequently in this chapter we refer to "measuring" or *approximating* the entropy of a subshift via a large set F. We give a precise definition as follows.

Definition 2.2.15 (Entropy approximation)**.** Let $X \subset \mathcal{A}^G$ be a subshift, and let $\delta > 0$. A finite subset $F \subset G$ is said to *δ-approximate* the entropy of X if

$$h(X) - \delta < h(F, X) < h(X) + \delta.$$

We shall more commonly write $h(X) < h(F, X) + \delta < h(X) + 2\delta$.

Infinitely many such sets exist for any δ, as provided by the Følner sequence and the definition of $h(X)$. We introduce this notion so that we may layer invariance conditions and entropy-approximating conditions as needed.

Proposition 2.2.10. *For finitely many choices of i, let $K_i \subset G$ be any finite subsets, and let $\varepsilon_i > 0$ be any positive constants. For finitely many choices of j, let $X_j \subset \mathcal{A}_j^G$*

be any subshifts over any finite alphabets, and let $\delta_j > 0$ be any positive constants. There exists a finite subset $F \subset G$ which is (K_i, ε_i)-invariant for every i, and which δ_j-approximates the entropy of X_j for every j.

Proof. Choose $F = F_n$ for sufficiently large n. □

The following theorem is an elementary generalization of a classical statement (see [20] for a proof in the case where $G = \mathbb{Z}$). We omit the proof here for brevity.

Proposition 2.2.11. *Let $(X_n)_n$ be a descending family of subshifts, and let $X = \bigcap_n X_n$. Then*

$$h(X) = \lim_{n\to\infty} h(X_n).$$

It is desirable to work with SFTs as much as possible while preserving (or, in our case, approximating) relevant dynamical quantities. We shall make frequent use of the next theorem, which we justify with several of the above results.

Theorem 2.2.12. *Let $X \subset \mathcal{A}^G$ be a subshift and suppose that $X_0 \subset \mathcal{A}^G$ is an SFT such that $X \subset X_0$. For any $\varepsilon > 0$, there exists an SFT $Z \subset \mathcal{A}^G$ such that $X \subset Z \subset X_0$ and $h(X) \leq h(Z) < h(X) + \varepsilon$.*

Proof. By Proposition 2.2.6, there is a descending family of SFTs $(X_n)_n$ such that $X = \bigcap_n X_n$. By Proposition 2.2.7, we have $X_n \subset X_0$ for all sufficiently large n. By Proposition 2.2.11, we have $h(X) \leq h(X_n) < h(X) + \varepsilon$ for all sufficiently large n. Choose $Z = X_n$ for n large enough to meet both conditions. □

If $\phi : X \to W$ is a factor map of subshifts, then we have already seen that $h(W) \leq h(X)$. The “entropy drop" or *entropy gap* between X and W is the quantity $h(X) - h(W)$. A subsystem $X' \subset X$ induces a corresponding subsystem $\phi(X') = W' \subset W$, and later in this chapter we will want a uniform bound for the entropy gap between every X' and W' pair. We make this idea precise in the following definition.

Definition 2.2.16 (Maximal entropy gap)**.** Suppose $\phi : X \to W$ is a factor map. The *maximal entropy gap* of ϕ is the quantity

$$\mathcal{H}(\phi) = \sup_{X'} \big(h(X') - h(\phi(X'))\big),$$

where the supremum is taken over all subshifts $X' \subset X$. In particular, it holds that

$$h(W) \leq h(X) \leq h(W) + \mathcal{H}(\phi).$$

Recall that if X and T are subshifts, then the dynamical direct product $X \times T$ factors onto both X and T via the projection map(s) $\pi_X(x,t) = x$ and $\pi_T(x,t) = t$.

Proposition 2.2.13. *Let X and T be shift spaces. The maximal entropy gap of the projection map $\pi_X : X \times T \to X$ is*

$$\mathcal{H}(\pi_X) = h(T).$$

Proof. It is classically known that $h(X \times T) = h(X) + h(T)$, in which case $h(T) = h(X \times T) - h(X) \leq \mathcal{H}(\pi_X)$. For the converse inequality, suppose $Z \subset X \times T$ is any subshift. Note by Definition 2.2.7 that $z \in Z$ implies $z = (z^X, z^T)$, where $z^X = \pi_X(z) \in \pi_X(Z) \subset X$ and $z^T \in T$. Therefore $Z \subset \pi_X(Z) \times T$, in which case it follows that $h(Z) \leq h(\pi_X(Z)) + h(T)$. Since Z was arbitrary, we have

$$h(T) \leq \mathcal{H}(\pi_X) = \sup_{Z} \big(h(Z) - h(\pi_X(Z))\big) \leq h(T),$$

where the supremum is taken over all subshifts $Z \subset X \times T$. □

A quick corollary is that when $h(T) = 0$, we have $h(Z) = h(\pi_X(Z))$ for any subsystem $Z \subset X \times T$.

2.3 Tilings of amenable groups

2.3.1 Definition and encoding

In this section we consider the notion of *tilings* of G. The existence of tilings of G with certain properties is essential in our constructions in subsequent sections.

Definition 2.3.1 (Quasi-tilings and exact tilings)**.** A *quasi-tiling* of G is a pair $(\mathcal{S}, C)$, where $\mathcal{S}$ is a finite collection of finite subsets of G (called the *shapes* of the tiling) and C is a function that assigns each shape $S \in \mathcal{S}$ to a subset $C(S) \subset G$, called the set of *centers* or *center-set* attributed to S. We require that e is in S for each $S \in \mathcal{S}$. The following properties are also required.

i. For distinct shapes S, $S' \in \mathcal{S}$, the subsets $C(S)$ and $C(S')$ are disjoint.

ii. The shapes in $\mathcal{S}$ are "translate-unique", in the sense that

$$S \neq S' \implies Sg \neq S', \quad \forall g \in G,$$

for each S, $S' \in \mathcal{S}$.

iii. The map $(S, c) \mapsto Sc \subset G$ defined on the domain $\{(S, c) : S \in \mathcal{S} \text{ and } c \in C(S)\}$ is injective.

We may refer to both the pair $(\mathcal{S}, C)$ and the collection

$$\mathcal{T} = \mathcal{T}(\mathcal{S}, C) = \{Sc \subset G : S \in \mathcal{S} \text{ and } c \in C(S)\}$$

as "*the* quasi-tiling." Each subset $\tau = Sc \in \mathcal{T}$ is called a *tile*. For a quasi-tiling $\mathcal{T}$, we denote the union of all the tiles by $\bigcup \mathcal{T}$. A quasi-tiling $\mathcal{T}$ may not necessarily cover G in the sense that $\bigcup \mathcal{T} = G$; nor is it necessary for any two distinct tiles τ, $\tau' \in \mathcal{T}$ to be disjoint. However, if both of these conditions are met (that is, if $\mathcal{T}$ is a *partition* of G), then $\mathcal{T}$ is called an *exact* tiling of G.

Ornstein and Weiss [1] previously constructed quasi-tilings of G with good dynamical properties, and this construction has become a fundamental tool for analyzing the dynamics of G-actions. Downarowicz, Huczek, and Zhang [3] sharpened this construction, showing that a countable amenable group exhibits many *exact* tilings with good dynamical properties, as we describe below (see Theorem 2.3.3).

A quasi-tiling $\mathcal{T}$ of G may be encoded in symbolic form, allowing for dynamical properties to be attributed to and studied for quasi-tilings. The encoding method presented here differs from the one presented in [3], as we will only require exact tilings in this chapter. See Remark 2.3.2 below for further discussion of the relation between our encoding and the encoding given in [3].

Definition 2.3.2 (Encoding)**.** Let $\mathcal{S}$ be a finite collection of finite shapes, and let

$$\Sigma(\mathcal{S}) = \{(S, s) : s \in S \in \mathcal{S}\},$$

which we view as a finite alphabet. If $\mathcal{T}$ is an exact tiling of G over $\mathcal{S}$, then it corresponds to a unique point $t \in \Sigma(\mathcal{S})^G$ as follows. For each $g \in G$, there is a unique tile $Sc \in \mathcal{T}$ containing g; let $s = gc^{-1} \in S$ and set $t_g = (S, s)$.

In the above definition, note that s is the “relative position" of g in the translate Sc of S. In other words, t labels each element g of G with both the type of shape of the tile containing g and the relative position of g within that tile. In particular, $g \in C(S) \iff t_g = (S, e)$.

Note that the correspondence $\mathcal{T} \mapsto t \in \Sigma(\mathcal{S})^G$, when regarded as a map on the set of all exact tilings of G over $\mathcal{S}$, is injective. However, the correspondence is not surjective in general. Let $\Sigma_E(\mathcal{S}) \subset \Sigma(\mathcal{S})^G$ be the set of all encodings of exact tilings of G over $\mathcal{S}$. It may be the case that *no* exact tiling of G over $\mathcal{S}$ exists, in which case $\Sigma_E(\mathcal{S}) = \varnothing$. In general, we have the following useful theorem.

Proposition 2.3.1. *Let $\mathcal{S}$ be a finite collection of finite shapes drawn from G. Then $\Sigma_E(\mathcal{S}) \subset \Sigma(\mathcal{S})^G$ is an SFT.*

Proof. Let $\Sigma_1(\mathcal{S})$ be the set of all points $t \in \Sigma(\mathcal{S})^G$ that satisfy the following local rule: for each $g \in G$, if $t_g = (S_0, s_0) \in \Sigma(\mathcal{S})$ then

$$t_{sc} = (S_0, s), \quad \forall s \in S_0, \tag{R1}$$

where $c = s_0^{-1} g$. It is easy to see that $\Sigma_1(\mathcal{S})$ is an SFT, and from Definition 2.3.2 it is immediate that $\Sigma_E(\mathcal{S}) \subset \Sigma_1(\mathcal{S})$.

For the reverse inclusion, let $t \in \Sigma_1(\mathcal{S})$ be an arbitrary point satisfying the local rule (R1) everywhere. For each $S \in \mathcal{S}$, let $C(S) = \{g \in G : t_g = (S, e)\}$. Then $\mathcal{T} = \mathcal{T}(\mathcal{S}, C)$ is a quasi-tiling. To complete the proof, it suffices to show that $\mathcal{T}$ is exact and encoded by t, since that would give $t \in \Sigma_E(\mathcal{S})$ and then $\Sigma_E(\mathcal{S}) = \Sigma_1(\mathcal{S})$.

Let $g \in G$, suppose $t_g = (S, s)$, and let $c = s^{-1} g$. By rule (R1) and the fact that $e \in S$, we have $t_c = t_{ec} = (S, e)$ and therefore $c \in C(S)$. Hence, $g = sc \in Sc \in \mathcal{T}$. This demonstrates that $\bigcup \mathcal{T} = G$. Next, suppose Sc, $S'c' \in \mathcal{T}$ are not disjoint and let $g \in Sc \cap S'c'$. Then $g = sc = s'c'$ for some $s \in S$ and $s' \in S'$. From $c \in C(S)$ we have $t_c = (S, e)$, and by the rule (R1) we have

$$t_g = t_{sc} = (S, s).$$

By identical proof we have $t_g = (S', s')$, from which it follows that $S = S'$ and $s = s'$. The latter implies that

$$c = s^{-1} g = s'^{-1} g = c',$$

and hence Sc and $S'c'$ are the same tile. This demonstrates that $\mathcal{T}$ is a partition of G, and therefore $\mathcal{T}$ is an exact tiling of G over S. Finally, we note that it is straightforward to check that $\mathcal{T}$ is encoded by t, which completes the proof. $\square$

Remark 2.3.2. Before we move on, we note here that the encoding method presented above (Definition 2.3.2) differs from the one presented in [3]. The encoding method in that work gives symbolic encodings for all quasi-tilings, which is not necessary for our present purposes. Indeed, the encoding in [3] uses the alphabet $\Lambda(\mathcal{S}) = \mathcal{S} \cup \{0\}$, and a point $\lambda \in \Lambda(\mathcal{S})^G$ encodes a quasi-tiling $(\mathcal{S}, C)$ when $\lambda_g = S \iff g \in C(S)$ and $\lambda_g = 0$ otherwise. This is a prudent encoding method for the study of general quasi-tilings, as any quasi-tiling may be encoded in this manner. Our encoding method works only for exact tilings, but is well-suited to our purposes. In fact, if one is only interested in exact tilings, then the two encodings are equivalent. Indeed, if $\Lambda_E(\mathcal{S}) \subset \Lambda(\mathcal{S})^G$ is the collection of all encodings of exact tilings of G over $\mathcal{S}$, then there is a topological conjugacy $\phi : \Sigma_E(\mathcal{S}) \to \Lambda_E(\mathcal{S})$ given by $\phi(t)_g = S \iff t_g = (S, e)$ and $\phi(t)_g = 0$ otherwise.

Next we turn our attention to the dynamical properties of tilings, as derived from their encodings.

Definition 2.3.3 (Dynamical tiling system)**.** Let $\mathcal{S}$ be a finite collection of finite shapes, let $\mathcal{T}$ be an exact tiling of G over $\mathcal{S}$, and let $\mathcal{T}$ be encoded by the point $t \in \Sigma_E(\mathcal{S})$. The *dynamical tiling system* generated by $\mathcal{T}$ is the subshift generated by t in $\Sigma(\mathcal{S})^G$, denoted $\Sigma_{\mathcal{T}} = \overline{\mathcal{O}}(t) \subset \Sigma_E(\mathcal{S})$.

This allows for the dynamical properties (e.g., entropy) of $\Sigma_{\mathcal{T}}$ as a subshift of $\Sigma(\mathcal{S})^G$ to be ascribed to $\mathcal{T}$. The *tiling entropy* of $\mathcal{T}$ is $h(\mathcal{T}) = h(\Sigma_{\mathcal{T}})$, the entropy of $\Sigma_{\mathcal{T}}$ as a subshift of $\Sigma(\mathcal{S})^G$.

The tiling entropy of $\mathcal{T}$ is a measure of the “complexity" of tile patterns that occur in large regions of G. In particular, when $\mathcal{T}$ has entropy zero, the number of ways to cover a large region $F \subset G$ by tiles in $\mathcal{T}$ grows subexponentially (with respect to $|F|$).

The following theorem is quickly deduced from the main result of Downarowicz, Huczek, and Zhang [3], which we state in this form for convenience. It is this result that allows us to utilize exact tilings of G in this chapter.

Theorem 2.3.3 ([3])**.** *Let $K \subset G$ be a finite subset, and let $\varepsilon > 0$. Then there exists a finite collection of finite shapes $\mathcal{S}$ with the following properties.*

i. Each shape $S \in \mathcal{S}$ is (K, ε)-invariant.

ii. $K \subset S$ and $|S| > \varepsilon^{-1}$ for each shape $S \in \mathcal{S}$.

iii. There exists a point $t_0 \in \Sigma_E(\mathcal{S})$ such that $h(\overline{\mathcal{O}}(t_0)) = 0$.

The point t_0 encodes an exact tiling $\mathcal{T}_0$ of G over $\mathcal{S}$ with tiling entropy $h(\mathcal{T}_0) = 0$.

2.3.2 Approximating sets with tiles

Entropy and other dynamical properties of G-shifts are well measured by sets with strong invariance properties (the Følner sequence F_n provides a wealth of such sets). However, we would instead like to utilize an (appropriately selected) exact tiling $\mathcal{T}$ for this purpose. In this section, we build good *tile approximations* of sets: finite collections of tiles $\mathcal{T}^* \subset \mathcal{T}$ attributed to large, suitably invariant subsets $F \subset G$ that are *good* in the sense that the symmetric difference $F \triangle \bigcup \mathcal{T}^*$ is small (as a proportion of $|F|$).

Definition 2.3.4 (Tile approximation)**.** Let $F \subset G$ be a finite subset. An exact tiling $\mathcal{T}$ of G induces two finite collections of tiles: the *outer approximation* of F by $\mathcal{T}$, denoted

$$\mathcal{T}^\times(F) = \{\tau \in \mathcal{T} : \tau \cap F \neq \varnothing\},$$

and the *inner approximation* of F by $\mathcal{T}$, denoted

$$\mathcal{T}^\circ(F) = \{\tau \in \mathcal{T} : \tau \subset F\}.$$

Denote $F^\times(\mathcal{T}) = \bigcup \mathcal{T}^\times(F)$ and $F^\circ(\mathcal{T}) = \bigcup \mathcal{T}^\circ(F)$. Observe that $F^\circ(\mathcal{T}) \subset F \subset F^\times(\mathcal{T})$.

Lemma 2.3.4. *Let $\mathcal{S}$ be a finite collection of shapes from G, and let $U = \bigcup \mathcal{S}$. Let $\varepsilon > 0$, and choose $\delta > 0$ such that $\delta |U||UU^{-1}| < \varepsilon$. Let $F \subset G$ be a finite subset that is (UU^{-1}, δ)-invariant. For any exact tiling $\mathcal{T}$ of G over $\mathcal{S}$, the following statements hold:*

i. $|F^\times(\mathcal{T}) \setminus F^\circ(\mathcal{T})| < \varepsilon |F|$,

ii. $(1-\varepsilon)|F| < |F^\circ(\mathcal{T})| \leq |F|$, *and*

iii. $|F| \leq |F^\times(\mathcal{T})| < (1+\varepsilon)|F|$.

Proof. First, we observe that each tile $\tau \in \mathcal{T}$ is contained in a translate Ug for some $g \in G$; indeed, we have $\tau = Sc$ for some $S \in \mathcal{S}$ and $c \in C(S) \subset G$, then $S \subset U$ implies $\tau \subset Uc$. This fact also gives that $|\tau| \leq |U|$ for every tile $\tau \in \mathcal{T}$.

We claim that every tile $\tau \in \mathcal{T}^\times(F) \setminus \mathcal{T}^\circ(F)$ intersects $\partial_{UU^{-1}} F$. To establish the claim, we first note that for each such tile τ it holds that $\tau \cap F \neq \varnothing$ and $\tau \not\subset F$. So let $f \in \tau \cap F$, and note that $f \in \tau \subset Ug$ for some $g \in G$. From $\tau \not\subset F$ we also have $Ug \not\subset F$. By Lemma 2.2.2 we have $f \in Ug \subset \left(\operatorname{int}_{UU^{-1}} F\right)^c$, and hence $f \in F \setminus \left(\operatorname{int}_{UU^{-1}} F\right) = \partial_{UU^{-1}} F$, which establishes our claim.

By the claim in the previous paragraph, there is a map $\gamma : \mathcal{T}^\times(F) \setminus \mathcal{T}^\circ(F) \to \partial_{UU^{-1}} F$ with the property that $\gamma(\tau) \in \tau$ for each τ. Observe that γ is injective, as distinct tiles are disjoint, and therefore $|\mathcal{T}^\times(F) \setminus \mathcal{T}^\circ(F)| \leq |\partial_{UU^{-1}} F|$. We also have that

$$|\partial_{UU^{-1}} F| < \delta |UU^{-1}||F|,$$

by the invariance hypothesis on F and Lemma 2.2.1. Then

$$\begin{aligned}|F^{\times}(\mathcal{T}) \setminus F^{\circ}(\mathcal{T})| &= \sum_{\tau \in \mathcal{T}^{\times}(F) \setminus \mathcal{T}^{\circ}(F)} |\tau| \\ &\leq |\mathcal{T}^{\times}(F) \setminus \mathcal{T}^{\circ}(F)||U| \\ &\leq |\partial_{UU^{-1}} F||U| \\ &< \delta |U||UU^{-1}||F| \\ &< \varepsilon |F|.\end{aligned}$$

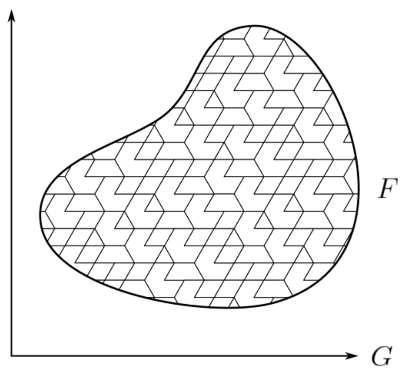
This establishes statement $(i.)$. The remaining two statements are easy to check using the fact that $F^{\circ}(\mathcal{T}) \subset F \subset F^{\times}(\mathcal{T})$ and statement $(i.)$. □

One more notion is necessary to develop before moving on from tilings: the *frame* of a given subset with respect to a given tiling.

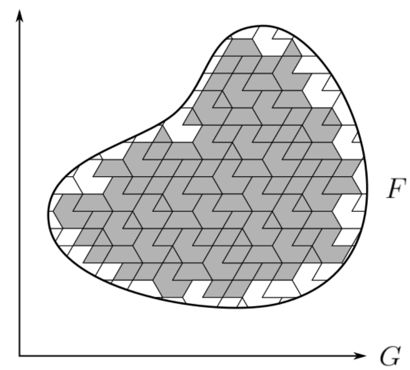
Definition 2.3.5 (Frame of a tiling)**.** Let F, $K \subset G$, and let $\mathcal{T}$ be an exact tiling of G. The inner $(\mathcal{T}, K)$-*frame* of F is the subset

$$\mathrm{fr}_{\mathcal{T},K}(F) = \bigcup_{\tau} \partial_K(\tau),$$

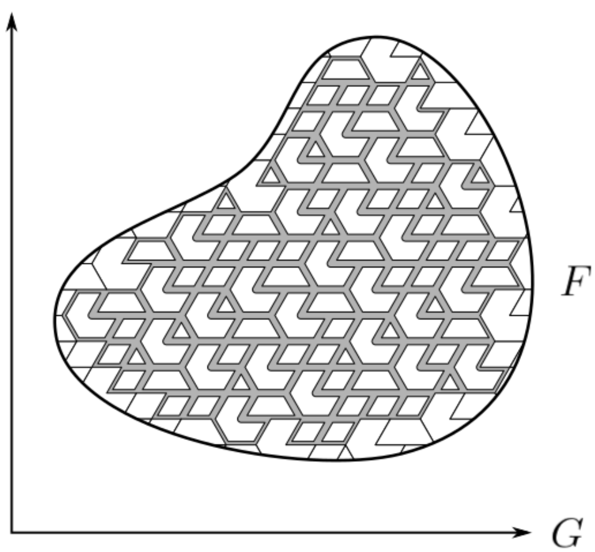
where the union ranges over all $\tau \in \mathcal{T}^{\circ}(F)$. See Figure 2.1 for an illustration.



(a) A hypothetical region $F \subset G$ with illustrated tiling $\mathcal{T}$.



(b) The $\mathcal{T}$-interior of F is shaded.



(c) The inner frame of F (with respect to $\mathcal{T}$, K) is shaded. The K-boundary of each tile inside F is taken.

Figure 2.1: A sketch of the construction of $\mathrm{fr}_{\mathcal{T},K}(F)$.

2.4 Results for SFTs

Having discussed everything about tilings relevant for our purposes, we are now ready to begin discussing our main results. In this section we present our results for SFTs, and in the following section we turn our attention to sofic shifts.

Theorem 2.4.1. *Let G be a countable amenable group, and let X be a G-SFT such that $h(X) > 0$. Then*

$$\{h(Y) : Y \subset X \text{ and } Y \text{ is an SFT}\}$$

is dense in $[0, h(X)]$.

Before we begin the proof, let us give a short outline of the main ideas. The broad strokes of this proof come from Desai [5], whose argument in the case where $G = \mathbb{Z}^d$ we are able to extend to the case where G is an arbitrary countable amenable group. This is possible by utilizing the exact tilings of G constructed by Downarowicz, Huczek and Zhang [3].

Given an arbitrary $\varepsilon > 0$, we produce a family of SFT subshifts of X whose entropies are 2ε-dense in $[0, h(X)]$. We accomplish this by first selecting an exact, zero entropy tiling $\mathcal{T}_0$ of G with suitably large, invariant tiles. Then we build subshifts with strongly controlled entropies inside the product system $Z_0 = X \times \Sigma_0$, where Σ_0 is the dynamical tiling system generated by $\mathcal{T}_0$.

To construct these subshifts from Z_0, we control which patterns in the X layer can appear in the "interior" of the tiles in the Σ_0 layer. We are able to finely comb away entropy from Z_0 by forbidding these patterns one at a time. This process generates a descending family of subsystems for which the entropy drop between consecutive subshifts is less than ε. After enough such patterns have been forbidden, the overall entropy is less than ε. This collection of subshifts therefore has entropies that are ε-dense in $[0, h(Z_0)]$. Then we project the subshifts into X and utilize Theorem 2.2.12

to produce SFTs subsystems of X with entropies that are 2ε-dense in $[0, h(X)]$.

Proof. Let $X \subset \mathcal{A}^G$ be an SFT such that $h(X) > 0$, let $K \subset G$ be a large finite subset such that $\mathcal{P}(K, X)$ specifies X as an SFT, and let ε be any constant such that $0 < \varepsilon < h(X)$. Choose $\delta > 0$ such that

$$2\delta + \delta \log 2 + 2\delta \log |\mathcal{A}| < \varepsilon.$$

By Theorem 2.3.3, there exists a finite collection $\mathcal{S}$ of finite subsets of G with the following properties.

i. Each shape $S \in \mathcal{S}$ is (KK^{-1}, η)-invariant, where $\eta > 0$ is a constant such that $\eta|KK^{-1}| < \delta$. By Lemma 2.2.1, this implies that $|\partial_{KK^{-1}} S| < \delta|S|$ for each shape $S \in \mathcal{S}$.

ii. $KK^{-1} \subset S$ and $|S| > \delta^{-1}$ for each $S \in \mathcal{S}$.

iii. There is a point $t_0 \in \Sigma_E(\mathcal{S})$ such that $h(\overline{\mathcal{O}}(t_0)) = 0$. Consequently, t_0 encodes an exact tiling $\mathcal{T}_0$ of G over $\mathcal{S}$ with tiling entropy zero.

For the remainder of this proof, these are all fixed. We shall abbreviate $\partial F = \partial_{KK^{-1}} F$ for any finite subset $F \subset G$. For a pattern p on F, we take ∂p to mean $p(\partial F)$ and call this the *border* of p (with respect to KK^{-1}).

Let $\Sigma_0 = \overline{\mathcal{O}}(t_0) \subset \Sigma_E(\mathcal{S})$ be the dynamical tiling system generated by the tiling $\mathcal{T}_0$, which has entropy zero. Of central importance to this proof is the product system $X \times \Sigma_0$, which factors onto X via the projection map $\pi : X \times \Sigma_0 \to X$ given by $\pi(x, t) = x$ for each $(x, t) \in X \times \Sigma_0$. Let us establish some terminology for certain patterns of interest which occur in this system.

Given a shape $S \in \mathcal{S}$, we shall refer to a pattern $b = (b^X, b^{\mathcal{T}}) \in \mathcal{P}(S, X \times \Sigma_0)$ as a *block* (to distinguish from patterns of any general shape). If a block $b \in (\mathcal{A} \times \Sigma(\mathcal{S}))^S$

satisfies $b_s^{\mathcal{T}} = (S, s)$ for every $s \in S$, then we shall say b is *aligned*. See Figure 2.2 for an illustration of the aligned property.

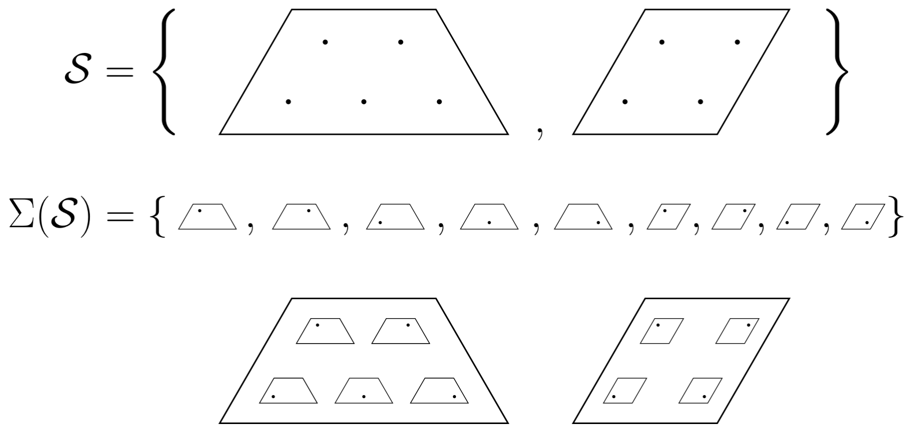


Figure 2.2: A hypothetical collection of shapes $\mathcal{S}$, the appropriate alphabet $\Sigma(\mathcal{S})$, and (the $\mathcal{T}$-layer of) two aligned blocks are pictured. Each point of each block is labelled with the correct shape type and relative displacement within that shape.

For a subshift $Z \subset X \times \Sigma_0$, we denote the subcollection of aligned blocks of shape S that occur in Z by

$$\mathcal{P}^a(S, Z) \subset \mathcal{P}(S, Z) \subset (\mathcal{A} \times \Sigma(\mathcal{S}))^S,$$

where the superscript a identifies the subcollection. Given a shape $S \in \mathcal{S}$ and an aligned block b of shape S, consider the border $\partial b \in (\mathcal{A} \times \Sigma(\mathcal{S}))^{\partial S}$. We are interested in the number of ways that the border ∂b may be extended to all of S - that is, the number of allowed (and in particular, aligned) *interiors* for S which agree with ∂b on the boundary ∂S. For a subshift $Z \subset X \times \Sigma_0$, we denote this collection by

$$\text{ints}^a(\partial b, Z) = \{b' \in \mathcal{P}^a(S, Z) : \partial b' = \partial b\}.$$

We shall extend all the same terminology described above (blocks, aligned blocks, borders, interiors) to tiles $\tau = Sc \in \mathcal{T}_0$, as there is a bijection between $\mathcal{P}(S, Z)$ and $\mathcal{P}(\tau, Z) = \mathcal{P}(Sc, Z)$. For a given tile $\tau \in \mathcal{T}_0$, a block $b \in (\mathcal{A} \times \Sigma(\mathcal{S}))^\tau$ is aligned if $b^{\mathcal{T}}_{sc} = (S, s)$ for each $sc \in Sc = \tau$. The subcollection of aligned blocks of shape τ occurring in a shift $Z \subset X \times \Sigma_0$ is denoted $\mathcal{P}^a(\tau, Z)$. Given a border $\partial b \in (\mathcal{A} \times \Sigma(\mathcal{S}))^{\partial\tau}$, the collection of aligned blocks of shape τ occurring in Z agreeing with ∂b on $\partial\tau$ is also denoted $\operatorname{ints}^a(\partial b, Z) \subset \mathcal{P}^a(\tau, Z)$.

For the theorem, we shall inductively construct a descending family of subshifts $(Z_n)_n$ of $X \times \Sigma_0$ as follows. Begin with $Z_0 = X \times \Sigma_0$, then assume Z_n has been constructed for $n \geq 0$. If there exists a shape $S_n \in \mathcal{S}$ and an aligned block $\beta_n \in \mathcal{P}^a(S_n, Z_n)$ such that

$$|\operatorname{ints}^a(\partial\beta_n, Z_n)| > 1,$$

then let $Z_{n+1} = Z_n \setminus \beta_n$. If no such block exists on any shape $S \in \mathcal{S}$, then Z_n is the final subshift in the chain and the chain is finite in length.

Let us first argue that in fact, the chain *must* be finite in length. For each $n \geq 0$ we have $Z_{n+1} \subset Z_n$, in which case $\mathcal{P}^a(S, Z_{n+1}) \subset \mathcal{P}^a(S, Z_n)$ for every shape $S \in \mathcal{S}$. Moreover, for the distinguished shape S_n (the shape of the forbidden block β_n), it holds that $\mathcal{P}^a(S_n, Z_{n+1}) \sqcup \{\beta_n\} \subset \mathcal{P}^a(S_n, Z_n)$. This implies that

$$\sum_{S \in \mathcal{S}} |\mathcal{P}^a(S, Z_n)|$$

strictly decreases with n. There is no infinite strictly decreasing sequence of positive integers, hence the descending chain must be finite in length. Let $N \geq 0$ be the index of the terminal subshift, and note by construction that the shift Z_N satisfies

$$|\operatorname{ints}^a(\partial b, Z_N)| = 1$$

for every aligned block $b \in \mathcal{P}^a(S, Z_N)$ on any shape $S \in \mathcal{S}$.

Most of the rest of the proof aims to establish the following two statements:

$$h(Z_{n+1}) \leq h(Z_n) < h(Z_{n+1}) + \varepsilon \quad \text{for each } n < N, \text{ and} \tag{U1}$$

$$h(Z_N) < \varepsilon. \tag{U2}$$

To begin, let $F \subset G$ be a finite subset satisfying the following two conditions:

(F1) F is (UU^{-1}, ϑ)-invariant, where $U = \bigcup \mathcal{S}$ and ϑ is a positive constant such that $\vartheta |U||UU^{-1}| < \delta$. Note this implies that F may be well approximated by tiles from any exact tiling of G over $\mathcal{S}$, in the sense of Lemma 2.3.4.

(F2) F is large enough to δ-approximate (Definition 2.2.15) the entropy of Σ_0 and Z_n for every $n \leq N$. This implies in particular that $h(F, \Sigma_0) < \delta$.

Such a set exists by Proposition 2.2.10. We fix F for the remainder of this proof.

Now for each $n \leq N$, we claim that

$$|\mathcal{P}(F, Z_n)| \geq \sum_t \sum_f \prod_\tau |\operatorname{ints}^a(f(\partial\tau), Z_n)|, \quad \text{and} \tag{E1}$$

$$|\mathcal{P}(F, Z_n)| \leq |\mathcal{A}|^{\delta|F|} \cdot \sum_t \sum_f \prod_\tau |\operatorname{ints}^a(f(\partial\tau), Z_n)|, \tag{E2}$$

where the indices t, f, and τ are as follows. The variable t ranges over $\mathcal{P}(F, \Sigma_0)$, and therefore t is the restriction to F of an encoding of an exact, zero entropy tiling $\mathcal{T}_t$ of G over $\mathcal{S}$. The variable f ranges over all $(\mathcal{A} \times \Sigma(\mathcal{S}))$-labellings of the $(\mathcal{T}_t,\ KK^{-1})$-frame of F (Definition 2.3.5) that are allowed in Z_0 and for which $f^{\mathcal{T}}$ agrees with t. Lastly, the variable τ ranges over the tiles in $\mathcal{T}_t^\circ(F)$.

To begin the argument towards the claims (E1) and (E2), let $n \leq N$ be arbitrary. To count patterns $p \in \mathcal{P}(F, Z_n)$, write $p = (p^X, p^{\mathcal{T}})$ and sum over all possible labellings in the tiling component. We have

$$|\mathcal{P}(F, Z_n)| = \sum_t |\{p \in \mathcal{P}(F, Z_n) : p^{\mathcal{T}} = t\}|, \tag{2.1}$$

where the sum ranges over all $t \in \mathcal{P}(F, \Sigma_0)$. This is valid because $Z_n \subset Z_0 = X \times \Sigma_0$, hence any $z = (z^X, z^{\mathcal{T}}) \in Z_n$ must have $z^{\mathcal{T}} \in \Sigma_0$.

Next, let $t \in \mathcal{P}(F, \Sigma_0)$ be fixed. The pattern t extends to/encodes an exact, zero entropy tiling $\mathcal{T}_t$ of G over $\mathcal{S}$ (possibly distinct from the original selected tiling $\mathcal{T}_0$, but as Σ_0 is generated by $\mathcal{T}_0$, one may take $\mathcal{T}_t$ to be a translation of $\mathcal{T}_0$ that agrees with t on F).

Recall that $F^\circ(\mathcal{T}_t) = \bigcup \mathcal{T}_t^\circ(F) \subset F$ is the inner tile approximation of F by the tiling $\mathcal{T}_t$ (Definition 2.3.4), which we shall abbreviate here as F_t°. Recall also that the $(\mathcal{T}_t,\, KK^{-1})$-frame of F is the subset $\bigcup_\tau \partial\tau$ where the union is taken over all $\tau \in \mathcal{T}_t^\circ(F)$ (Definition 2.3.5). Since K is fixed for this proof, we shall abbreviate the frame as $\mathrm{fr}_t(F)$. From Equation (2.1), we now split over all allowed labellings of $\mathrm{fr}_t(F)$. We have

$$|\mathcal{P}(F, Z_n)| = \sum_t \sum_f |\{p \in \mathcal{P}(F, Z_n) : p^{\mathcal{T}} = t \text{ and } p(\mathrm{fr}_t(F)) = f\}|, \tag{2.2}$$

where the first sum is taken over all $t \in \mathcal{P}(F, \Sigma_0)$, and the second sum is taken over all $f \in \mathcal{P}(\mathrm{fr}_t(F), Z_0)$ for which $f^{\mathcal{T}}$ agrees with t.

We have the pattern $t \in \mathcal{P}(F, \Sigma_0)$ fixed from before; next we fix a frame pattern $f \in \mathcal{P}(\mathrm{fr}_t(F), Z_0)$ such that $f^{\mathcal{T}}$ agrees with t. We wish to count the number of patterns $p \in \mathcal{P}(F, Z_n)$ such that $p^{\mathcal{T}} = t$ and $p(\mathrm{fr}_t(F)) = f$. Let this collection be denoted by $D_n = D(t, f;\, Z_n) \subset \mathcal{P}(F, Z_n)$. Observe that each D_n is finite and $D_{n+1} \subset D_n$ for each $n < N$. Consider the map $\gamma : D_0 \to \prod_\tau \mathcal{P}(\tau, Z_0)$ given by $\gamma(p) = (p(\tau))_\tau$, which sends a pattern $p \in D_0$ to a vector of blocks indexed by $\mathcal{T}_t^\circ(F)$. We claim the map γ

is at most $|\mathcal{A}|^{\delta|F|}$-to-1, and for each $n \leq N$ we have

$$\gamma(D_n) = \prod_{\tau} \operatorname{ints}^a(f(\partial\tau), Z_n). \tag{2.3}$$

Together, these claims will provide a bound for $|D_n|$ from above and below, which combine with Equation (2.2) to yield the claims (E1) and (E2).

First we argue that γ is at most $|\mathcal{A}|^{\delta|F|}$-to-1. This is where we first invoke the invariance of F. Suppose $(b_\tau)_\tau \in \prod_\tau \mathcal{P}(\tau, Z_0)$ is a fixed vector of blocks. If $p \in D_0$ is a pattern such that $\gamma(p) = (b_\tau)_\tau$, then $p^{\mathcal{T}}$ is determined by t and $p(\tau) = b_\tau$ for each tile $\tau \in \mathcal{T}_t^\circ(F)$. Therefore, p is uniquely determined by $p^X(F \setminus F_t^\circ)$, hence $|\gamma^{-1}(b_\tau)_\tau| \leq |\mathcal{A}|^{|F \setminus F_t^\circ|}$. By property (F1) of the set F and by Lemma 2.3.4 we have $|F \setminus F_t^\circ| < \delta|F|$, and thus the map γ is at most $|\mathcal{A}|^{\delta|F|}$-to-1.

Next, we shall prove the set equality (2.3). Let $n \leq N$, and let $p \in D_n$. For each tile $\tau \in \mathcal{T}_t^\circ(F)$, the block $p(\tau) \in \mathcal{P}(\tau, Z_n) \subset (\mathcal{A} \times \Sigma(\mathcal{S}))^\tau$ is aligned; this is because $p^{\mathcal{T}} = t$ and t encodes the tiling $\mathcal{T}_t$ itself. Moreover, p agrees with f on $\operatorname{fr}_t(F)$ by assumption that $p \in D_n = D(t, f; Z_n)$, in which case $p(\partial\tau) = f(\partial\tau)$ for each tile $\tau \in \mathcal{T}_t^\circ(F)$. This demonstrates that

$$\gamma(D_n) \subset \prod_{\tau} \operatorname{ints}^a(f(\partial\tau), Z_n).$$

We shall prove the reverse inclusion by induction on n. For the $n = 0$ case, let $(b_\tau)_\tau$ be a vector of blocks such that $b_\tau \in \mathcal{P}^a(\tau, Z_0)$ and $\partial b_\tau = f(\partial\tau)$ for each $\tau \in \mathcal{T}_t^\circ(F)$. To construct a γ-preimage of $(b_\tau)_\tau$ in D_0, begin with a point $x \in X$ such that $x(\operatorname{fr}_t(F)) = f^X$. Such a point exists because f occurs in some point of $Z_0 = X \times \Sigma_0$. Note that

$$x(\partial\tau) = f^X(\partial\tau) = \partial b_\tau^X$$

for each $\tau \in \mathcal{T}_t^\circ(F)$, because $\partial\tau \subset \operatorname{fr}_t(F)$ for each τ. Moreover, for each τ it holds that the block b_τ^X occurs in a point of X, as each block $b_\tau = (b_\tau^X, b_\tau^{\mathcal{T}})$ occurs in a point of $Z_0 = X \times \Sigma_0$. Because X is an SFT specified by patterns of shape K, we may repeatedly apply Lemma 2.2.8 to excise the block $x(\tau)$ and replace it with b_τ^X for every $\tau \in \mathcal{T}_t^\circ(F)$. Every tile is disjoint, so the order in which the blocks are replaced does not matter. After at most finitely many steps, we obtain a new point $x' \in X$ such that $x'(\mathrm{fr}_t(F)) = f^X$ and $x'(\tau) = b_\tau^X$ for each $\tau \in \mathcal{T}_t^\circ(F)$.

Recall that the point $t \in \Sigma_0$ is fixed from before. The point $(x', t) \in X \times \Sigma_0$ is therefore allowed in $Z_0 = X \times \Sigma_0$. Let $p = (x', t)(F) \in \mathcal{P}(F, Z_0)$. We have that $p^{\mathcal{T}} = t$ and $p(\mathrm{fr}_t(F)) = f$ by the selection of x'. This implies that $p \in D_0$. It also holds that $p(\tau) = b_\tau$ for each $\tau \in \mathcal{T}_t^\circ(F)$, as t itself encodes the tiling $\mathcal{T}_t$ from which the tiles $\tau \in \mathcal{T}_t^\circ(F)$ are drawn (and each block b_τ is aligned, by assumption). We then finally have $\gamma(p) = (b_\tau)_\tau$, which settles the case $n = 0$.

Now suppose the set equality (2.3) holds for some fixed $n < N$, and let $(b_\tau)_\tau \in \prod_\tau \mathrm{ints}^a(f(\partial\tau), Z_{n+1})$. From the inclusion $Z_{n+1} \subset Z_n$ and the inductive hypothesis, it follows there is a pattern $p \in D_n$ such that $\gamma(p) = (b_\tau)_\tau$. Suppose $p = (x, t)(F)$ for some $(x, t) \in Z_n$ (by induction, t is the point fixed from before). We need to modify p only slightly to find a γ-preimage of $(b_\tau)_\tau$ which occurs in Z_{n+1} (and hence belongs to D_{n+1}).

Consider the block β_n determined at the beginning of this proof, which is forbidden in the subshift Z_{n+1}. If β_n occurs anywhere in the point (x, t), then (by the assumption that β_n is aligned) it must occur on a tile $\tau \in \mathcal{T}_t$. It does *not* occur on any of the tiles from $\mathcal{T}_t^\circ(F)$, because for each tile $\tau \in \mathcal{T}_t^\circ(F)$ we have $(x, t)(\tau) = b_\tau$ which is allowed in Z_{n+1} by assumption.

Yet, β_n may occur in (x, t) outside of F_t°. By the construction of Z_{n+1}, we have

$$|\,\mathrm{ints}^a(\partial\beta_n, Z_n)| > 1,$$

and therefore there is an aligned block $\tilde{b}$ which occurs in Z_n such that $\tilde{b} \neq \beta_n$ and $\partial\tilde{b} = \partial\beta_n$. Apply Lemma 2.2.8 at most countably many times to excise β_n^X wherever it may occur in x, replacing it with $\tilde{b}^X$. This yields a new point $x' \in X$.

Then $(x', t) \in Z_0$ also belongs to Z_{n+1}. It was already the case that none of the blocks $\beta_0, \ldots, \beta_{n-1}$ could occur anywhere in (x, t) by the assumption $(x, t) \in Z_n$, and now neither does β_n occur anywhere in (x', t). The pattern $p' = (x', t)(F)$ may be distinct from $p = (x, t)(F)$ (the labelling may change on $F \setminus F_t^\circ$), but we did not replace any of the blocks within F_t°. We still have $p'(\tau) = b_\tau$ for each tile $\tau \in \mathcal{T}_t^\circ(F)$, and hence $p' \in D_{n+1}$ and $\gamma(p') = (b_\tau)_\tau$.

This completes the induction, and we conclude that the set equality (2.3) holds for each $n \leq N$. From this equality and the fact that γ is at most $|\mathcal{A}|^{\delta|F}$-to-1, we obtain

$$\prod_\tau |\operatorname{ints}^a(f(\partial\tau), Z_n)| \leq |D(t, f\,; Z_n)| \leq |\mathcal{A}|^{\delta|F|} \cdot \prod_\tau |\operatorname{ints}^a(f(\partial\tau), Z_n)|.$$

Notice that the above inequalites hold for each fixed $t \in \mathcal{P}(F, \Sigma_0)$, each fixed $f \in \mathcal{P}(\operatorname{fr}_t(F), X \times \Sigma_0)$ such that $f^{\mathcal{T}} = t(\operatorname{fr}_t(F))$, and each $n \leq N$. From these inequalities and Equation (2.2), we conclude that (E1) and (E2) hold, i.e.,

$$|\mathcal{P}(F, Z_n)| \geq \sum_t \sum_f \prod_\tau |\operatorname{ints}^a(f(\partial\tau), Z_n)|, \quad \text{and} \tag{E1}$$

$$|\mathcal{P}(F, Z_n)| \leq |\mathcal{A}|^{\delta|F|} \cdot \sum_t \sum_f \prod_\tau |\operatorname{ints}^a(f(\partial\tau), Z_n)|, \tag{E2}$$

where the first sum is taken over all $t \in \mathcal{P}(F, \Sigma_0)$, the second sum over all $f \in \mathcal{P}(\operatorname{fr}_t(F), Z_0)$ for which $f^{\mathcal{T}}$ agrees with t, and the product over all $\tau \in \mathcal{T}_t^\circ(F)$.

Property (F2) of the set F implies that $h(F, \Sigma_0) < \delta$, in which case $|\mathcal{P}(F, Z_0)| < e^{\delta|F|}$. Consequently, the variable t in (E1) and (E2) ranges over fewer than $e^{\delta|F|}$ terms. Moreover, by the selection of $\mathcal{S}$, we have $|\partial\tau| < \delta|\tau|$ for each $\tau \in \mathcal{T}_t^\circ(F)$, in which case it follows that

$$|\operatorname{fr}_t(F)| = \Big|\bigcup_\tau \partial\tau\Big| = \sum_\tau |\partial\tau| < \sum_\tau \delta|\tau| = \delta\Big|\bigcup_\tau \tau\Big| = \delta|F_t^\circ| \le \delta|F|.$$

Here we have used that distinct tiles from $\mathcal{T}_t$ are disjoint. From this estimate, we deduce that there are fewer than $|\mathcal{A}|^{\delta|F|}$ labellings of the frame of F that agree with a fixed t on the $\mathcal{T}$-layer. Consequently, the variable f in (E1) and (E2) ranges over fewer than $|\mathcal{A}|^{\delta|F|}$ terms. Observe also that the size of $\mathcal{T}_t^\circ(F)$ as a collection is small compared to F. Indeed, $|S| > \delta^{-1}$ for each shape $S \in \mathcal{S}$, in which case

$$|F| \ge |F_t^\circ| = \sum_\tau |\tau| \ge |\mathcal{T}_t^\circ(F)| \cdot \big(\min_{S\in\mathcal{S}} |S|\big) > |\mathcal{T}_t^\circ(F)|\delta^{-1},$$

and therefore $|\mathcal{T}_t^\circ(F)| < \delta|F|$. Consequently, the variable τ in (E1) and (E2) ranges over fewer than $\delta|F|$ terms.

Before returning to (U1) and (U2), one more estimate is necessary. For each $n < N$, each shape $S \in \mathcal{S}$, and each aligned block $b \in \mathcal{P}^a(S, Z_n)$, we claim that

$$|\operatorname{ints}^a(\partial b, Z_n)| \le 2\,|\operatorname{ints}^a(\partial b, Z_{n+1})|. \tag{2.4}$$

Let $S \in \mathcal{S}$, and let $b \in \mathcal{P}^a(S, Z_n)$ be an aligned block occurring in Z_n, distinct from the forbidden block β_n. Say $b = (x,t)(S)$ for some $(x,t) \in Z_n$. We claim that b occurs in a point of Z_{n+1}. Note that t extends to/encodes an exact, zero entropy tiling $\mathcal{T}_t$ of G over $\mathcal{S}$, and note that S is a tile of $\mathcal{T}_t$. This is because b is aligned by assumption, in which case $t_e = b_e^{\mathcal{T}} = (S, e)$, and therefore $e \in C_t(S)$.

Suppose the forbidden block β_n occurs anywhere in (x,t). Because β_n is aligned, it must occur on a tile $\tau \in \mathcal{T}_t$. It does *not* occur on S, because $b \ne \beta_n$. By the assumption that $|\operatorname{ints}^a(\partial\beta_n, Z_n)| > 1$, we know there is an aligned block $\tilde{b}_n \ne \beta_n$ that occurs in Z_n such that $\partial\tilde{b}_n = \partial\beta_n$.

Recall X is an SFT specified by patterns of shape K, and $\tilde{b}_n^X$ is allowed in X. Again we may apply Lemma 2.2.8 at most countably many times, excising β_n^X wherever it may occur in x and replacing it with $\tilde{b}_n^X$. At the end we receive a new point $x' \in X$, within which β_n^X does not occur. Then (x', t) is allowed in Z_{n+1} and $(x', t)(S) = b$, hence $b \in \mathcal{P}^a(S, Z_{n+1})$.

The conclusion is that β_n is the only aligned block lost from Z_n to Z_{n+1}. For each $b \in \mathcal{P}^a(S, Z_n)$, we have either $\mathrm{ints}^a(\partial b, Z_n) = \mathrm{ints}^a(\partial b, Z_{n+1})$ or $\mathrm{ints}^a(\partial b, Z_n) = \mathrm{ints}^a(\partial b, Z_{n+1}) \sqcup \{\beta_n\}$. If two positive integers differ by at most 1 then their ratio is at most 2, hence the inequality (2.4) follows.

Finally, we shall use the estimates (E1) and (E2) to argue for the ultimate claims (U1) and (U2) made before. For the first, consider a fixed $n < N$. It is clear that $h(Z_{n+1}) \leq h(Z_n)$ by inclusion. For the second inequality in (U1), we have

$$\begin{aligned}
|\mathcal{P}(F, Z_n)| &\leq |\mathcal{A}|^{\delta|F|} \cdot \sum_t \sum_f \prod_\tau |\,\mathrm{ints}^a(f(\partial\tau), Z_n)| \\
&\leq |\mathcal{A}|^{\delta|F|} \cdot \sum_t \sum_f \prod_\tau 2\,|\,\mathrm{ints}^a(f(\partial\tau), Z_{n+1})| \\
&< |\mathcal{A}|^{\delta|F|} \cdot 2^{\delta|F|} \cdot \sum_t \sum_f \prod_\tau |\,\mathrm{ints}^a(f(\partial\tau), Z_{n+1})| \\
&\leq |\mathcal{A}|^{\delta|F|} \cdot 2^{\delta|F|} \cdot |\mathcal{P}(F, Z_{n+1})|,
\end{aligned}$$

where the inequalities are justified by (E2), (2.4), the fact that $|\mathcal{T}_t^\circ(F)| < \delta|F|$, and

(E1), respectively. Taking logs and dividing through by $|F|$, we obtain

$$\begin{aligned} h(Z_n) &< h(F, Z_n) + \delta \\ &< \big(\delta \log|\mathcal{A}| + \delta \log 2 + h(F, Z_{n+1})\big) + \delta \\ &< \delta \log|\mathcal{A}| + \delta \log 2 + \big(h(Z_{n+1}) + \delta\big) + \delta \\ &= h(Z_{n+1}) + 2\delta + \delta \log 2 + \delta \log|\mathcal{A}| \\ &< h(Z_{n+1}) + \varepsilon, \end{aligned}$$

where we have used the property (F2) of F, the previous display, the property (F2) again, and our choice of δ. This inequality establishes (U1). For (U2), recall that the terminal shift Z_N has the property that any aligned border $\partial b \in \mathcal{P}^a(\partial S, Z_N)$ on any shape $S \in \mathcal{S}$ has exactly 1 allowed aligned interior. Hence, we see that

$$\begin{aligned} |\mathcal{P}(F, Z_N)| &\leq |\mathcal{A}|^{\delta|F|} \cdot \sum_t \sum_f \prod_\tau |\operatorname{ints}^a(f(\partial\tau), Z_N)| \\ &= |\mathcal{A}|^{\delta|F|} \cdot \sum_t \sum_f \prod_\tau 1 \\ &< |\mathcal{A}|^{\delta|F|} \cdot e^{\delta|F|} \cdot |\mathcal{A}|^{\delta|F|} \cdot 1, \end{aligned}$$

where the first inequality is justified by (E2) and the last inequality is justified by our bounds on the number of terms in the sums (established previously). Taking logs and dividing through by $|F|$, we finally have

$$\begin{aligned} h(Z_N) &< h(F, Z_N) + \delta \\ &< \big(\delta + 2\delta \log|\mathcal{A}|\big) + \delta \\ &= 2\delta + 2\delta \log|\mathcal{A}| \\ &< \varepsilon, \end{aligned}$$

where we have used the property (F2) of the set F, the previous display, and our choice of δ. We have now established (U2).

With (U1) and (U2) in hand, the rest of the proof is easy. By (U1) and (U2), we have that $(Z_n)_{n\leq N}$ is a family of subshifts of $X \times \Sigma_0$ such that $(h(Z_n))_{n\leq N}$ is ε-dense in $[0, h(X \times \Sigma_0)]$.

For each $n \leq N$, let $X_n = \pi(Z_n) \subset X$, where π is the projection map $\pi(x,t) = x$. From Proposition 2.2.13, $\mathcal{H}(\pi) = h(\Sigma_0) = 0$, hence $h(X_n) = h(Z_n)$ for every $n \leq N$.

Then $(X_n)_{n\leq N}$ is a descending family of subshifts of X such that $(h(X_n))_{n\leq N}$ is ε-dense in $[0, h(X)]$. Though each X_n may not be an SFT, we do know that X is an SFT. One may therefore apply Theorem 2.2.12 to construct a family of SFTs $(Y_n)_{n\leq N}$ such that for each $n \leq N$, we have $X_n \subset Y_n \subset X$ and $h(X_n) \leq h(Y_n) < h(X_n) + \varepsilon$. Hence $(h(Y_n))_{n\leq N}$ is 2ε-dense in $[0, h(X)]$. As ε was arbitrary, we conclude that the entropies of the SFT subsystems of X are dense in $[0, h(X)]$. □

The following "relative" version of Theorem 2.4.1 is stronger and easily obtained as a consequence of Theorem 2.4.1.

Theorem 2.4.2. *Let G be a countable amenable group, let X be a G-SFT, and let $Y \subset X$ be any subsystem such that $h(Y) < h(X)$. Then*

$$\{h(Z) : Y \subset Z \subset X \text{ and } Z \text{ is an SFT}\}$$

is dense in $[h(Y), h(X)]$.

Proof. We prove the density directly. Suppose $(a, b) \subset [h(Y), h(X)]$ for positive reals $a < b$, and let $\varepsilon < (b - a)/2$. By Theorem 2.4.1, there exists an SFT $Z_0 \subset X$ such that $a < h(Z_0) < a + \varepsilon$. Note that these inequalities give $h(Y) < h(Z_0)$. Consider the subshift $Y \cup Z_0 \subset X$, which has entropy

$$h(Y \cup Z_0) = \max\big(h(Y), h(Z_0)\big) = h(Z_0) \in (a, a + \varepsilon).$$

Because X is an SFT and $Y \cup Z_0 \subset X$, by Theorem 2.2.12 there is an SFT Z such that $Y \cup Z_0 \subset Z \subset X$ and $h(Y \cup Z_0) \leq h(Z) < h(Y \cup Z_0) + \varepsilon$. Thus we have

$$a < h(Y \cup Z_0) \leq h(Z) < h(Y \cup Z_0) + \varepsilon < a + 2\varepsilon < b.$$

Since (a, b) was arbitrary, the proof is complete. □

2.5 Results for sofic shifts

2.5.1 An extension result for sofic shifts

In order to address the case of sofic shifts, we seek to leverage our results on SFTs. In particular, given a sofic shift W, we would like an SFT X such that W is a factor of X and such that the maximal entropy drop across the factor map is very small. The following theorem guarantees the existence of such SFTs.

Theorem 2.5.1. *Let $W \subset \mathcal{A}_W^G$ be a sofic shift. For every $\varepsilon > 0$, there exists an SFT $\tilde{X}$ and a one-block code $\tilde{\phi} : \tilde{X} \to W$ such that the maximal entropy gap of $\tilde{\phi}$ satisfies $\mathcal{H}(\tilde{\phi}) < \varepsilon$.*

Proof. Since W is sofic, there exists an SFT $X \subset \mathcal{A}_X^G$ and a factor map $\phi : X \to W$. Without loss of generality, we assume that

i. ϕ is a one-block code, witnessed by the function $\Phi : \mathcal{A}_X \to \mathcal{A}_W$, and

ii. $\mathcal{A}_X$ and $\mathcal{A}_W$ are disjoint.

We abbreviate $\mathcal{A}_{XW} = \mathcal{A}_X \sqcup \mathcal{A}_W$. Let $\varepsilon > 0$, and select $\delta > 0$ such that

$$4\delta + \delta(1 + \delta) \log |\mathcal{A}_X| < \varepsilon/2.$$

Let $K \subset G$ be a large finite subset that specifies X as an SFT. The set K is fixed for the remainder of this proof, and thus we shall denote $\partial_{KK^{-1}} F$ by ∂F and $\text{int}_{KK^{-1}} F$

by $\operatorname{int} F$ for any finite set $F \subset G$. By Theorem 2.3.3, there exists a finite set of finite shapes $\mathcal{S}$ such that the following conditions are met.

i. Each shape $S \in \mathcal{S}$ is (KK^{-1}, η)-invariant, where $\eta > 0$ is a constant such that $\eta|KK^{-1}| < \delta$. By Lemma 2.2.1, this implies $|\partial S| < \delta|S|$ for each $S \in \mathcal{S}$.

ii. $KK^{-1} \subset S$ and $|S| > \delta^{-1}$ for each $S \in \mathcal{S}$.

iii. There is a point $t_0 \in \Sigma_E(\mathcal{S})$ such that $h(\overline{\mathcal{O}}(t_0)) = 0$.

Recall by Proposition 2.3.1 that $\Sigma_E(\mathcal{S})$ is an SFT. By Theorem 2.2.12, there is an SFT T such that $\overline{\mathcal{O}}(t_0) \subset T \subset \Sigma_E(\mathcal{S})$ and $h(T) < h(\overline{\mathcal{O}}(t_0)) + \delta$. Consequently, each point $t \in T$ is an encoding of an exact tiling $\mathcal{T}_t$ of G over $\mathcal{S}$ (possibly distinct from the original tiling $\mathcal{T}_0$), with tiling system entropy

$$h(\mathcal{T}_t) = h(\overline{\mathcal{O}}(t)) \leq h(T) < h(\overline{\mathcal{O}}(t_0)) + \delta = 0 + \delta.$$

Because X and T are SFTs, we have that $X \times T \subset (\mathcal{A}_X \times \Sigma(\mathcal{S}))^G$ is also an SFT.

Let $t \in T$ be arbitrary, and recall that $\mathcal{T}_t$ is a partition of G. Thus, for each $g \in G$, there is a unique tile $\tau \in \mathcal{T}_t$ such that $g \in \tau$. We define the notation $\mathcal{T}_t(g)$ by setting $\mathcal{T}_t(g) = \tau$. Next we define a map $\phi_t : X \to \mathcal{A}_{XW}^G$ by the following rule: for each $g \in G$ and $x \in X$,

$$\phi_t(x)_g = \begin{cases} x_g & \text{if } g \in \partial\mathcal{T}_t(g), \\ \Phi(x_g) & \text{if } g \in \operatorname{int}\mathcal{T}_t(g). \end{cases}$$

This map is well-defined, as $\tau = \partial\tau \sqcup \operatorname{int}\tau$. The map ϕ_t applies the one-block code ϕ to “most" of a point x, by relabelling the *interiors* of each tile $\tau \in \mathcal{T}_t$.

We now define a sliding block code $\varphi : X \times T \to (\mathcal{A}_{XW} \times \Sigma(\mathcal{S}))^G$ by applying the map(s) ϕ_t fiber-wise: for each point $(x, t) \in X \times T$, let

$$\varphi(x, t) = (\phi_t(x), t).$$

It is straightforward to check that φ is indeed a sliding block code (Definition 2.2.6). For the theorem, the desired shift $\tilde{X}$ is identified with the range of this map. Let

$$\tilde{X} = \varphi(X \times T) \subset (\mathcal{A}_{XW} \times \Sigma(\mathcal{S}))^G.$$

See Figure 2.3 for an illustration of the construction. It remains to show that there is a one-block code $\tilde{\phi} : \tilde{X} \to W$, that the shift $\tilde{X}$ is an SFT, and that $\mathcal{H}(\tilde{\phi}) < \varepsilon$.

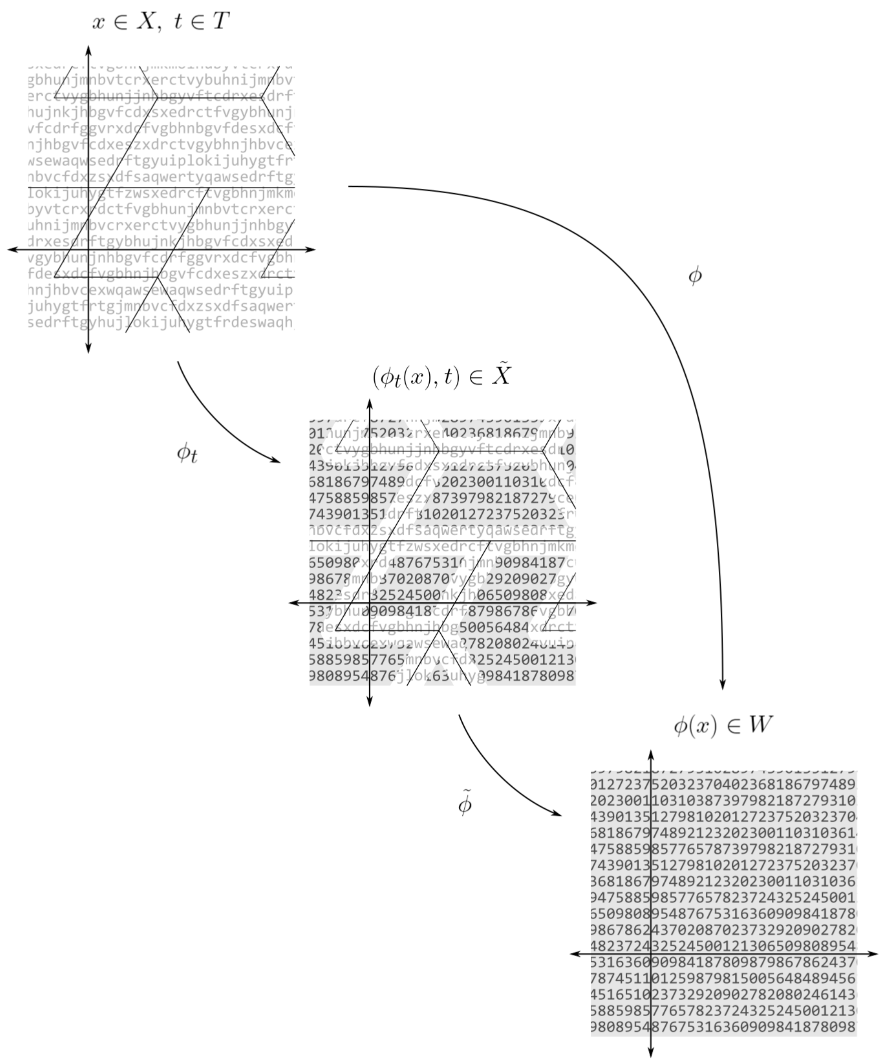


Figure 2.3: A hypothetical point $x \in X$ with a tiling $t \in T$ overlayed; the partially-transformed point $\phi_t(x)$ is pictured, which is labelled with symbols from both X and W; finally, the wholly-transformed image point $\phi(x) \in W$ is reached.

First, let us show that $\tilde{X}$ factors onto W. The factor map is induced by the function $\tilde{\Phi} : \mathcal{A}_{XW} \to \mathcal{A}_W$, which is an extension of Φ, defined by the following rule: $\tilde{\Phi}(\alpha) = \alpha$ if $\alpha \in \mathcal{A}_W$, and $\tilde{\Phi}(\alpha) = \Phi(\alpha)$ if $\alpha \in \mathcal{A}_X$. Let $\tilde{\phi} : \tilde{X} \to \mathcal{A}_W^G$ be given by

$$\tilde{\phi}(\tilde{x}, t)_g = \tilde{\Phi}(\tilde{x}_g), \quad \forall g \in G \text{ and } \forall (\tilde{x}, t) \in \tilde{X}.$$

It is clear that $\tilde{\phi}$ is a one-block code. Let us now show that $\tilde{\phi}(\tilde{X}) = W$. Let $x \in X$ and $t \in T$, in which case $(\phi_t(x), t) \in \tilde{X}$ is an arbitrary point. The effect of applying the map ϕ_t to x is to apply the one-block code ϕ to "most" of x. The map $\tilde{\phi}$ then "completes" the relabelling, via the extended function $\tilde{\Phi}$. In fact, we have that $\tilde{\phi}(\phi_t(x), t) = \phi(x) \in W$, hence $\tilde{\phi}(\tilde{X}) \subset W$. For the reverse inclusion, let $w \in W$. Since $\phi : X \to W$ is onto, there exists a point $x \in X$ such that $\phi(x) = w$. Choose $t \in T$ arbitrarily; then $(\phi_t(x), t) \in \tilde{X}$ and $\tilde{\phi}(\phi_t(x), t) = \phi(x) = w$. We conclude that $\tilde{\phi} : \tilde{X} \to W$ is a genuine factor map (and a one-block code).

Let us now show that $\tilde{X}$ is an SFT. We repeat that the shift $\tilde{X}$ can be written in the following instructive form:

$$\tilde{X} = \big\{(\phi_t(x), t) : x \in X \text{ and } t \in T\big\} \subset (\mathcal{A}_{XW} \times \Sigma(\mathcal{S}))^G.$$

In order to show that $\tilde{X}$ is an SFT, we will construct an SFT $\tilde{X}_1 \subset (\mathcal{A}_{XW} \times \Sigma(\mathcal{S}))^G$ and then prove that $\tilde{X} = \tilde{X}_1$. Recall that $K \subset G$ specifies X as an SFT. Let $K_T \subset G$ be a finite subset such that $\mathcal{P}(K_T, T)$ specifies T. We define $\tilde{X}_1$ to be the set of points $(\tilde{x}, t) \in (\mathcal{A}_{XW} \times \Sigma(\mathcal{S}))^G$ that satisfy the following local rules.

(R1) Any pattern of shape K_T that occurs in t must belong to $\mathcal{P}(K_T, T)$, and any pattern of shape K that occurs in $\tilde{x}$ and belongs to $\mathcal{A}_X^K$ must also belong to $\mathcal{P}(K, X)$ (recall $\mathcal{P}(K, \tilde{X}) \subset \mathcal{A}_{XW}^K = (\mathcal{A}_X \sqcup \mathcal{A}_W)^K$ in general). Note by Definition 2.2.10 that this condition is shift-invariant.

(R2) For any shape $S \in \mathcal{S}$ and any $c \in G$, if t satisfies $(\sigma^c t)_s = (S, s)$ for each $s \in S$, then $\exists b \in \mathcal{P}(S, X)$ such that $(\sigma^c \tilde{x})_s = b_s \in \mathcal{A}_X$ for all $s \in \partial S$ and $(\sigma^c \tilde{x})_s = \Phi(b_s) \in \mathcal{A}_W$ for all $s \in \operatorname{int} S$.

As these are local rules, they define an SFT; call it $\tilde{X}_1 \subset (\mathcal{A}_{XW} \times \Sigma(\mathcal{S}))^G$. Moreover, it is easily checked that any point $(\phi_t(x), t) \in \tilde{X}$ satisfies these rules everywhere (by construction of $\tilde{X}$), and so we have $\tilde{X} \subset \tilde{X}_1$.

For the reverse inclusion, consider a point $(\tilde{x}, t) \in \tilde{X}_1$. From (R1) it follows that $t \in T$, as T is an SFT specified by K_T. Therefore, t encodes an exact tiling $\mathcal{T}_t$ of G over $\mathcal{S}$ with $h(\mathcal{T}_t) < \delta$. Let $(\tau_n)_n$ enumerate the tiles of $\mathcal{T}_t$, and for each n let $\tau_n = S_n c_n$ for some $S_n \in \mathcal{S}$ and $c_n \in G$. Recall $\{\tau_n : n \in \mathbb{N}\}$ is a partition of G.

Let $n \in \mathbb{N}$, and consider $c = c_n$ and $S = S_n$. Observe that, because t encodes the tiling $\mathcal{T}_t$, we have $(\sigma^c t)_s = (S, s)$ for each $s \in S$. Then by (R2), there exists a block $b = b_n \in \mathcal{P}(S, X)$ such that $(\sigma^c \tilde{x})_s = b_s$ for all $s \in \partial S$ and $(\sigma^c \tilde{x})_s = \Phi(b_s)$ for all $s \in \operatorname{int} S$.

Define a point $x \in \mathcal{A}_X^G$ by setting $x(\tau_n) = b_n$ for each $n \in \mathbb{N}$. We claim that x is an allowed point of X and that $\phi_t(x) = \tilde{x}$. Toward this, let $g \in G$ be arbitrary, and consider the translate Kg (recall that K specifies X as an SFT).

If Kg intersects the *interior* of any tile $\tau_n = S_n c_n$, then $Kg \subset \tau_n$ by Lemma 2.2.2. In this case, the pattern $(\sigma^g x)(K)$ is a subpattern of b_n, and must therefore be allowed in X as $b_n \in \mathcal{P}(S_n, X)$. The alternative is that Kg is disjoint from the interior of every tile, in which case $Kg \subset \bigcup_n \partial \tau_n$. By (R2), we also have $\tilde{x}_g \in \mathcal{A}_X$ for every $g \in \bigcup_n \operatorname{int} \tau_n$. In this case we have $(\sigma^g x)(K) = (\sigma^g \tilde{x})(K)$, which is again allowed in X by (R1).

In either case we have that $(\sigma^g x)(K)$ is allowed in X for any $g \in G$, and hence $x \in X$. Then by the definition of ϕ_t, we see that $\phi_t(x) = \tilde{x}$. Thus, we have found a point $(x, t) \in X \times T$ such that $\varphi(x, t) = (\phi_t(x), t) = (\tilde{x}, t)$, and hence $(\tilde{x}, t) \in \tilde{X}$. We conclude that $\tilde{X} = \tilde{X}_1$, and therefore $\tilde{X}$ is an SFT.

Finally, let us show that $\mathcal{H}(\tilde{\phi}) < \varepsilon$. Towards this end, let $\tilde{X}' \subset \tilde{X}$ be any subsystem of $\tilde{X}$, and let $W' = \tilde{\phi}(\tilde{X}') \subset W$. We will show that $h(\tilde{X}') - h(W') < \varepsilon/2$.

Let $F \subset G$ be a finite subset such that the following conditions are met.

(F1) F is (UU^{-1}, ϑ)-invariant, where $U = \bigcup \mathcal{S}$ and $\vartheta > 0$ is a constant such that $\vartheta |U||UU^{-1}| < \delta$ (recall δ was selected at the beginning of this proof). Note this implies that F may be well approximated by tiles from any exact tiling of G over $\mathcal{S}$, in the sense of Lemma 2.3.4.

(F2) F is large enough to δ-approximate (Definition 2.2.15) the entropy of the shifts X', W' and T (recall that $h(T) < \delta$, in which case $h(F, T) < 2\delta$).

Such a set exists by Proposition 2.2.10. This set is fixed for the remainder of this proof. Recall that $\tilde{\phi}$ is a one-block code, and therefore there is a well defined map $\tilde{\Phi}_F : \mathcal{P}(F, \tilde{X}') \to \mathcal{P}(F, W')$ which takes a pattern $p \in \mathcal{P}(F, \tilde{X}')$ and applies the one-block code to p (at each element of F).

Recall also that a pattern $p \in \mathcal{P}(F, \tilde{X}')$ is of the form $p = (\phi_t(x), t)(F)$ for some points $x \in X$ and $t \in T$. The point t encodes an exact tiling $\mathcal{T}_t$ of G over $\mathcal{S}$. For each tile $\tau \in \mathcal{T}_t$, the definition of ϕ_t implies that

$$\phi_t(x)(\operatorname{int} \tau) \in \mathcal{A}_W^*, \quad \text{and} \quad \phi_t(x)(\partial \tau) \in \mathcal{A}_X^*. \tag{2.5}$$

Let $q = \tilde{\Phi}_F(p) \in \mathcal{P}(F, W')$. Recall that every element $g \in F$ belongs to a unique tile $\tau = \mathcal{T}_t(g) \in \mathcal{T}_t^{\times}(F)$, where $\mathcal{T}_t^{\times}(F) \subset \mathcal{T}_t$ is the outer approximation of F by the tiling $\mathcal{T}_t$ (Definition 2.3.4). By (2.5), we have that

$$q_g = \tilde{\Phi}_F(p)_g = \tilde{\Phi}(p_g^{\tilde{X}}) = \begin{cases} \Phi(p_g^{\tilde{X}}) & \text{if } g \in \partial \mathcal{T}_t(g) \\ p_g^{\tilde{X}} & \text{if } g \in \operatorname{int} \mathcal{T}_t(g) \end{cases}$$

In particular, we have $q_g = p_g^{\tilde{X}}$ whenever g belongs to the set

$$F \cap \Big(\bigcup_{\tau} \operatorname{int} \tau \Big),$$

where the union is taken over all $\tau \in \mathcal{T}_t^{\times}(F)$.

In light of these observations, we are ready to estimate $|\mathcal{P}(F, \tilde{X}')|$ in terms of $|\mathcal{P}(F, W')|$. We first use $\tilde{\Phi}_F$ to split over $\mathcal{P}(F, W')$, and then we split again over all possible T-layers. Indeed, we have

$$|\mathcal{P}(F, \tilde{X}')| = \sum_{q} \sum_{t} |\{p \in \tilde{\Phi}_F^{-1}(q) : p^T = t\}| \tag{2.6}$$

where the sums are taken over all patterns $q \in \mathcal{P}(F, W')$ and $t \in \mathcal{P}(F, T)$. Choose and fix patterns q and t. If $p \in \mathcal{P}(F, \tilde{X}')$ is a pattern such that $\tilde{\Phi}_F(p) = q$ and $p^T = t$, then the observations above imply that p is uniquely determined by

$$p^{\tilde{X}} \Big(F \cap \Big(\bigcup_{\tau} \partial \tau \Big) \Big) \in \mathcal{A}_X^*$$

where the union is taken over all tiles $\tau \in \mathcal{T}_t^{\times}(F)$. Moreover, our choice of $\mathcal{S}$ and the property (F1) of F together yield that

$$\Big| \bigcup_{\tau} \partial \tau \Big| < \delta \Big| \bigcup_{\tau} \tau \Big| = \delta |F_t^{\times}| < \delta(1+\delta)|F|.$$

Therefore, there are at most $|\mathcal{A}_X|^{\delta(1+\delta)|F|}$ patterns p such that $\tilde{\Phi}_F(p) = q$ and $p^T = t$. From this and Equation (2.6), we have

$$|\mathcal{P}(F, \tilde{X}')| \leq |\mathcal{P}(F, W')| \cdot |\mathcal{P}(F, T)| \cdot |\mathcal{A}_X|^{\delta(1+\delta)|F|}.$$

By taking logs and dividing through by $|F|$, we obtain the following:

$$\begin{aligned} h(\tilde{X}') &< h(F, \tilde{X}') + \delta \\ &\leq \big(h(F, W') + h(F, T) + \delta(1+\delta)\log|\mathcal{A}_X|\big) + \delta \\ &< \big(h(W') + \delta\big) + \big(2\delta\big) + \delta(1+\delta)\log|\mathcal{A}_X| + \delta \\ &= h(W') + 4\delta + \delta(1+\delta)\log|\mathcal{A}_X| \\ &< h(W') + \varepsilon/2, \end{aligned}$$

where we have used property (F2) of the set F, the above inequality, property (F2) again, and our choice of δ respectively. Since $\tilde{X}' \subset \tilde{X}$ was arbitrary, we have that

$$\mathcal{H}(\tilde{\phi}) = \sup_{\tilde{X}' \subset \tilde{X}} \big(h(\tilde{X}') - h(\tilde{\phi}(\tilde{X}'))\big) \leq \varepsilon/2 < \varepsilon,$$

which completes the proof. □

2.5.2 Subsystem entropies for sofic shifts

Here we present our main result concerning subsystem entropies for sofic shifts. The proof follows easily by combining our extension result (Theorem 2.5.1) with our result for SFTs (Theorem 2.4.2).

Theorem 2.5.2. *Let G be a countable amenable group, let W be a sofic G-shift and let $V \subset W$ be any subsystem such that $h(V) < h(W)$. Then*

$$\{h(U) : V \subset U \subset W \textit{ and } U \textit{ is sofic}\}$$

is dense in $[h(V), h(W)]$.

Proof. We prove the density directly. Let $(a, b) \subset [h(V), h(W)]$ for some real numbers $a < b$. Let $\varepsilon < (b-a)/2 < h(W) - h(V)$. By Theorem 2.5.1, there exists an SFT X and a factor map $\phi : X \to W$ such that $\mathcal{H}(\phi) < \varepsilon$.

Consider the preimage $Y = \phi^{-1}(V) \subset X$, which is a subshift. Note that $\phi(Y) = V$ because ϕ is surjective, hence $\phi|_Y : Y \to V$ is a factor map. We then have that

$$h(Y) \leq h(V) + \mathcal{H}(\phi) < h(V) + \varepsilon < h(W) \leq h(X).$$

Note also that $b \leq h(W) \leq h(X)$ and that $a \geq h(V) > h(Y) - \varepsilon$, which together yield that $(a + \varepsilon, b) \subset [h(Y), h(X)]$. By Theorem 2.4.2, there exists an SFT Z such that $Y \subset Z \subset X$ and $h(Z) \in (a + \varepsilon, a + 2\varepsilon) \subset (a, b)$. It follows that $U = \phi(Z)$ is a sofic shift for which $V \subset U \subset W$ and $h(U) \leq h(Z) < h(U) + \varepsilon$. Then we have

$$a < h(Z) - \varepsilon < h(U) \leq h(Z) < a + 2\varepsilon < b.$$

Thus $h(U) \in (a, b)$, which completes the proof. □

If one selects $V = \varnothing$ for the above theorem, then one recovers the statement that the entropies of the sofic subsystems of W are dense in $[0, h(W)]$. Next, we present our result concerning the entropies of arbitrary subsystems of sofic shifts.

Corollary 2.5.3. *Let W be a sofic shift. For every nonnegative real $r \leq h(W)$, there exists a subsystem $R \subset W$ for which $h(R) = r$.*

Proof. If $h(W) = 0$, then $r = 0$, in which case one may simply select $R = W$. If $h(W) > 0$, then let $W_0 = W$ and let $(\varepsilon_n)_n$ be a sequence of positive real numbers converging to zero. We have that W_0 is sofic and $r \leq h(W_0)$, and without loss of generality we assume that $h(W_0) < r + \varepsilon_0$.

Inductively construct a descending sequence of sofic shifts as follows. If $W_n \subset W$ is a sofic shift such that $r \leq h(W_n) < r + \varepsilon_n$, then by Theorem 2.5.2 there exists a sofic shift $W_{n+1} \subset W_n$ for which $r \leq h(W_{n+1}) < r + \varepsilon_{n+1}$.

Then $R = \bigcap_n W_n \subset W$ is a subshift such that $h(R) = \lim_n h(W_n) = r$ by Proposition 2.2.11. □

2.6 A counter-example

Theorem 2.5.2 implies that the entropies of the *sofic* subsystems of a sofic shift space W are dense in $[0, h(W)]$. One may wonder if this can be somehow “sharpened"; that is, one may wonder whether

$$\{h(W') : W' \subset W \text{ and } W' \text{ is an SFT}\}$$

is dense in $[0, h(W)]$. However, this statement is nowhere close to true in general, as we illustrate in this section by counterexample. This example is an adaptation of a construction of Boyle, Pavlov, and Schraudner [12].

Proposition 2.6.1. *There exists a sofic $\mathbb{Z}^2$-shift with positive entropy whose only SFT subsystem is a singleton.*

Proof. We first construct a certain point in $\{0,1\}^{\mathbb{Z}}$ as the limit of a sequence of finite words, then consider the subshift it generates. Let $\delta = 0.1$ and let $(T_n)_n$ be the sequence of natural numbers given by

$$T_n = 2n \cdot 2^n \cdot \delta^{-1} + 1$$

for each n. Let $w^1 = 010 \in \{0,1\}^3$, and for each n define the word

$$w^{n+1} = w^n w^n \cdots w^n w^n \, 0^n 1 0^n, \tag{2.7}$$

where the w^n term is repeated exactly T_n times. The limit word $w^\infty \in \{0,1\}^{\mathbb{N}_0}$ is an infinite one-sided sequence. Define a two-sided sequence $x^* \in \{0,1\}^{\mathbb{Z}}$ by $x^*_i = w^\infty_{|i|}$ for each $i \in \mathbb{Z}$. Let $X = \overline{\mathcal{O}}(x^*) \subset \{0,1\}^{\mathbb{Z}}$ be the subshift generated by x^*. We claim that X exhibits the following three properties.

(P1) X is *effective*, meaning that there exists a finite algorithm which enumerates a set of words $\mathcal{F} \subset \{0,1\}^*$ such that $X = \mathcal{R}\big(\{0,1\}^{\mathbb{Z}}, \mathcal{F}\big)$.

(P2) There exists a point $x \in X$ such that

$$\limsup_{n\to\infty} \frac{|\{k \in [-n,n] : x_k = 1\}|}{2n+1} > 0.1.$$

(P3) For each $x \in X$, either $x = 0^{\mathbb{Z}}$ or x contains the word $0^n 1 0^n$ for every n.

For (P1), let N be arbitrary. Note that because $X = \overline{\mathcal{O}}(x^*)$, any word of length N occurring in any point $x \in X$ is also a word occurring in x^*. By the recursive definition (2.7) and the fact that the sequence $\{T_n\}_{n=1}^{\infty}$ is recursive, there is an algorithm which, upon input N, prints all the words of length N that do *not* appear as subwords of x^*. The shift X is therefore effective.

For (P2), we argue that x^* satisfies the condition. For each n, let L_n be the length of the word w^n. Note that by the recurrence (2.7), we have

$$L_{n+1} = T_n L_n + 2n + 1 \quad \forall n.$$

For each n, let f_n be the *frequency* of 1s in w^n, given by

$$f_n = \frac{|\{i : w_i^n = 1\}|}{L_n}.$$

Observe that $f_n \leq 1$ for each n and $f_1 = \frac{1}{3}$. It follows from the recurrence (2.7) that

$$f_{n+1} = \frac{f_n T_n L_n + 1}{T_n L_n + 2n + 1}$$

for each n. This implies that

$$\begin{aligned} f_n - f_{n+1} &= \frac{f_n T_n L_n + f_n(2n+1) - f_n T_n L_n - 1}{T_n L_n + 2n + 1} \\ &\leq \frac{1(2_n+1) - 1}{T_n} \end{aligned}$$

in which case $f_n - f_{n+1} \leq \frac{2n}{T_n} < \frac{\delta}{2^n}$ for each n. Hence, we have that

$$\begin{aligned} f_1 - f_n &= (f_1 - f_2) + (f_2 - f_3) + \cdots + (f_{n-1} - f_n) \\ &< \frac{\delta}{2} + \frac{\delta}{4} + \cdots + \frac{\delta}{2^{n-1}} \\ &< \delta \end{aligned}$$

in which case $\frac{1}{3} - \delta < f_n$ for each n. By the recurrence (2.7), we therefore have that

$$\limsup_{n\to\infty} \frac{|\{k \in [-n,n] : x_k^* = 1\}|}{2n+1} \geq \frac{1}{3} - \delta > 0.1.$$

and the subsequence along $(L_n)_n$ is a witness.

For (P3), let n be fixed. First, observe that the infinite sequence w^∞ may be decomposed as a concatenation of blocks, where each block is either the word w^n or 0^m10^m for some $m \geq n$. Moreover, each w^n begins with 0 and ends with 0^n. This implies that 1s belonging to distinct blocks (of the form w^n or 0^m10^m for $m \geq n$) in the decomposition of w^∞ mentioned above are separated by at least $n+1$ appearances of the symbol 0. Therefore, if for any $k \leq n$ we have that 10^k1 appears anywhere in w^∞, then it must appear as a subword of a single block (rather than overlapping two distinct blocks), and that block must be w^n.

Next, let $x \in X$ be arbitrary. If the symbol 1 appears in x at most one time, then (P3) trivially holds. Otherwise, assume that 10^k1 appears somewhere in x for some $k \geq 1$. Without loss of generality, suppose $x_0 = x_{k+1} = 1$ and $x_i = 0$ for $i \in [1,k]$.

Now consider the subword $\omega = x([-L_n, L_n])$ for any n such that $k < L_n$. Because $X = \overline{\mathcal{O}}(x^*)$, the word ω must be a subword of x^*. Then, either ω is a subword of $x^*([-2L_n, 2L_n])$, or ω is a subword of w^∞ or a mirror reflection of one. In the first case, the definitions of x^* and w^∞ imply that ω contains the word w^n or its mirror. In the latter two cases, the observation of the previous paragraph implies that ω must contain w^n or its mirror. In any case, 0^n10^n is a subword of x. As n can be made arbitrarily large, this proves (P3).

We now use the shift X to construct the $\mathbb{Z}^2$-shift which is desired for the theorem. For each point $x \in X$, let $x^{\mathbb{Z}} \in \{0,1\}^{\mathbb{Z}^2}$ denote the $\mathbb{Z}^2$-labelling given by

$$\left(x^{\mathbb{Z}}\right)_{(i,j)} = x_i$$

for each $(i,j) \in \mathbb{Z}^2$. That is, $x^{\mathbb{Z}}$ is a $\mathbb{Z}^2$-labelling such that the symbols along each column are constant, and each row is equal to x itself. We shall also denote

$$X^{\mathbb{Z}} = \{x^{\mathbb{Z}} : x \in X\} \subset \{0,1\}^{\mathbb{Z}^2}.$$

It is a theorem of Aubrun and Sablik [22] that if X is effective, then $X^{\mathbb{Z}}$ is sofic.

Next, consider the alphabet $\{0, 1, 1'\}$, where we have artificially created two independent 1 symbols. Let $\pi : \{0,1,1'\}^{\mathbb{Z}^2} \to \{0,1\}^{\mathbb{Z}^2}$ be the one-block code which collapses 1 and $1'$. Let $Y = \pi^{-1}(X^{\mathbb{Z}}) \subset \{0,1,1'\}^{\mathbb{Z}^2}$. The shift Y is a copy of the shift $X^{\mathbb{Z}}$, in which the 1 symbols of every point have been replaced either by 1 or $1'$ in every possible combination.

We claim that the shift Y is the desired subshift for the theorem. Specifically, we claim that Y is sofic, that Y has positive entropy, and that the only nonempty SFT subsystem of Y is the singleton $\{0^{\mathbb{Z}^2}\}$.

To prove that Y is sofic, we construct an SFT S' and a factor map $\phi' : S' \to Y$ to witness the soficity of Y. Since $X^{\mathbb{Z}}$ is sofic, there is an SFT $S \subset \mathcal{A}^{\mathbb{Z}^2}$ and a factor map

$\phi : S \to X^{\mathbb{Z}}$. Without loss of generality, assume that ϕ is a one-block code induced by the function $\Phi : \mathcal{A} \to \{0, 1\}$.

Define a new finite alphabet $\mathcal{A} \times \{1, 1'\}$ and a one-block code ϕ' induced by the function $\Phi' : \mathcal{A} \times \{1, 1'\} \to \{0, 1, 1'\}$ which is given by

$$\Phi'(a, b) = \begin{cases} 0 & \text{if } \Phi(a) = 0 \\ b & \text{if } \Phi(a) = 1 \end{cases}$$

for each $(a, b) \in \mathcal{A} \times \{1, 1'\}$.

Let $S' = S \times \{1, 1'\}^{\mathbb{Z}^2}$, which we regard as a subshift of $(\mathcal{A} \times \{1, 1'\})^{\mathbb{Z}^2}$. Note that S' is an SFT, because both S and $\{1, 1'\}^{\mathbb{Z}^2}$ (the full $\mathbb{Z}^2$-shift on two symbols) are SFTs. A point $s' \in S'$ is of the form $s' = (s, \iota)$, where s is a point of S and $\iota \in \{1, 1'\}^{\mathbb{Z}^2}$ is an arbitrary 2-coloring of $\mathbb{Z}^2$. The reader may easily check that $(\pi \circ \phi')(s, \iota) = \phi(s) \in X^{\mathbb{Z}}$, from which it follows that $\phi'(S') = \pi^{-1}(X^{\mathbb{Z}}) = Y$. Then $\phi' : S' \to Y$ is a factor map. Since S' is an SFT, we conclude that Y is sofic.

Next, we will show that $h(Y) > 0$. From property (P2), the point $x^* \in X$ exhibits 1s in more than 10% of the positions in each of infinitely many symmetric intervals, say of the form $[-\ell_n, \ell_n]$ for an increasing sequence of natural numbers $(\ell_n)_n$. Therefore, the point $(x^*)^{\mathbb{Z}}$ exhibits 1s in more than 10% of the positions in each square $F_n = [-\ell_n, \ell_n]^2$. Each 1 in the pattern $(x^*)^{\mathbb{Z}}(F_n)$ may be replaced by 1 or $1'$ independently to yield an allowed pattern of Y, which implies that

$$|\mathcal{P}(F_n, Y)| \geq 2^{0.1|F_n|} \quad \forall n.$$

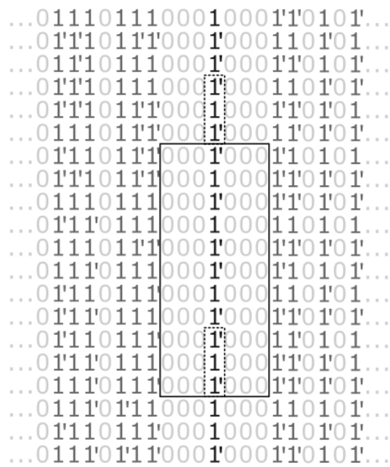
As $(F_n)_n$ is a Følner sequence for $\mathbb{Z}^2$, we then have $h(Y) \geq 0.1 \log 2 > 0$.

It remains to show that the only nonempty SFT subsystem of Y is the singleton $\{0^{\mathbb{Z}^2}\}$. Suppose to the contrary that $Z \subset Y$ is an SFT subsystem of Y which contains a nonzero point. Since Z is an SFT, we may find a constant $k \in \mathbb{N}$ such that the allowed patterns of Z are specified by the shape $K = [0,k)^2 \subset \mathbb{Z}^2$.

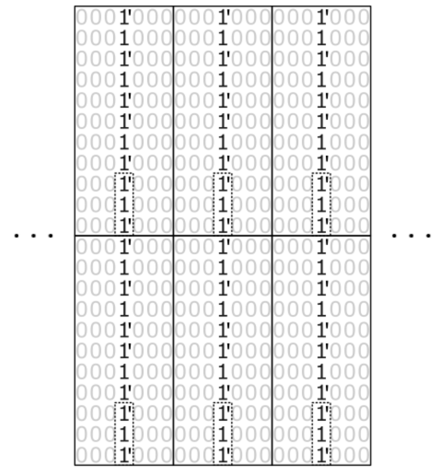
Let $z \in Z$ be a point different from $0^{\mathbb{Z}^2}$ and note $\pi(z) = x^{\mathbb{Z}} \in X^{\mathbb{Z}}$ for some $x \in X$ with $x \neq 0^{\mathbb{Z}}$. By property (P1), the string 0^n10^n appears in x for every n. Let $n > k$ be fixed. Suppose without loss of generality that 0^n10^n appears centered at the origin of x (with $x_0 = 1$ and $x_i = 0$ for $0 < |i| \leq n$). Thus we have $z_{(0,0)} = 1$ or $1'$. In fact, by the definition of Y, we have $z_{(0,j)} \in \{1, 1'\}$ for every $j \in \mathbb{Z}$.

Consider the $i = 0$ column of the point z. Starting with each index $\ell \in \mathbb{Z}$ and looking up, there is a corresponding vertically oriented word $\omega^\ell \in \{1, 1'\}^n$ given by $\omega^\ell_j = z_{(0,\ell+j)}$ for each $j \in [0,n)$. By the pigeonhole principle, there must exist a word $\varsigma \in \{1,1'\}^n$ such that $\varsigma = \omega^\ell$ for infinitely many choices of ℓ. That is, for infinitely many choices of ℓ, we have $z_{(0,\ell+j)} = \varsigma_j$ for each $j \in [0,n)$.

Let $\ell_1 < \ell_2$ be two such indices where a repetition occurs, with $\ell_2 - \ell_1 > n$. That is, we have $z_{(0,\ell_1+j)} = z_{(0,\ell_2+j)} = \varsigma_j$ for every $j \in [0,n)$. Now consider the rectangle $r = z\big([-n,n] \times [\ell_1, \ell_2)\big)$. Tile $\mathbb{Z}^2$ with infinitely many translated copies of r to obtain a new point $z' \in \{0,1,1'\}^{\mathbb{Z}^2}$. Figure 2.4 illustrates the construction.



(a) A hypothetical point z is illustrated around $[-n,n]\times[\ell_1,\ell_2)$. The repeated vertical word ς is indicated by the dotted box, and the rectangle r by the solid box.



(b) The rectangle r is used to tile $\mathbb{Z}^2$ and thereby construct z'. Every $k\times k$ block which occurs in z' also occurs in z.

Figure 2.4: An illustration of the construction of the contradictory point z', in a hypothetical case where $n=3$ and $k=2$.

Every pattern of shape $K=[0,k)^2$ which occurs in z' is a pattern which occurs in z (including the pattern of all zeroes), hence they are all allowed in Z. Because Z is an SFT specified by K, it then follows that $z'\in Z$. Because $Z\subset Y=\pi^{-1}(X^{\mathbb{Z}})$, there must exist a point $x'\in X$ such that $\pi(z')=(x')^{\mathbb{Z}}$. We obtain a contradiction, as the point x' cannot satisfy the property (P3) of X. For instance, the word $0^{3n}10^{3n}$ cannot appear in x' (as each row of z' is periodic in the horizontal direction with period $2n+1<3n$). This demonstrates that if Z is an SFT, then it contains no nonzero point. Therefore, the only nonempty SFT subsystem of Y is $\{0^{\mathbb{Z}^2}\}$. $\square$

CHAPTER 3: AN EMBEDDING THEOREM FOR SUBSHIFTS OVER AMENABLE GROUPS WITH THE COMPARISON PROPERTY

3.1 Introduction

In this chapter, our central question is as follows. Given subshifts X and Y over a countable amenable group G, under what conditions does X *embed* into Y? That is, under what conditions is X isomorphic to a subsystem of Y? One necessary condition is that $h(X) \leq h(Y)$, where $h(X)$ is the *topological entropy* of the subshift X (Definition 3.2.11), because topological entropy is preserved under isomorphism and non-increasing under taking subsystems. When $G = \mathbb{Z}$, the classical embedding theorem of Krieger [7, Theorem 3] provides a complete answer in the case that Y is a mixing shift of finite type (SFT) and $h(X) < h(Y)$. Namely, a certain necessary condition about periodic points turns out to be sufficient for an embedding $\psi : X \to Y$ to exist. In particular, the condition is automatically satisfied if X is *strongly aperiodic* (Definition 3.2.8), meaning that no point of X exhibits a non-identity element of G as a period (in other words, the shift action is *free* on X). Krieger's embedding theorem has become a cornerstone of the structure theory of SFTs over $\mathbb{Z}$.

Much less is known about the embedding problem for groups other than $G = \mathbb{Z}$. In the case where $G = \mathbb{Z}^d$ for $d \geq 2$, one result is given by Lightwood: suppose X is a strongly aperiodic subshift and suppose Y is an SFT which satisfies a mixing condition (“square mixing", known elsewhere as the “uniform filling" property) and contains a point with a finite orbit. If $h(X) < h(Y)$ and Y contains at least one *factor* of X (an image of X under a continuous and shift-commuting map), then X embeds into Y [8, Theorem 2.5]. Lightwood also proved that in the case where $G = \mathbb{Z}^2$, a square mixing SFT automatically contains a point with a finite orbit [8, Lemma 9.2]. Later, Lightwood proved that if Y is an SFT over $\mathbb{Z}^2$ satisfying a slightly stronger mixing condition ("square-filling mixing", which implies square mixing), then Y automatically contains at least one factor of X [23, Theorem 2.8]. These results together provide a partial extension of Krieger's embedding theorem to $G = \mathbb{Z}^2$.

In this chapter, we obtain an embedding theorem for subshifts over countable amenable groups with the *comparison property.* We do not define the comparison property here, but we appeal to it in the form of Theorem 3.2.10, a consequence of the comparison property due to Downarowicz and Zhang [24, 4]. We also note that the class of amenable groups with the comparison property includes every countable group containing no finitely generated subgroup of exponential growth (proof given in both [24, Theorem 5.11] and [4, Theorem 6.33]). In particular, this includes every countable abelian group. It is unknown whether there exists a countable amenable group without the comparison property.

Our main result is as follows.

Theorem 3.3.5. *Let G be a countable amenable group with the comparison property. Let X be a nonempty strongly aperiodic subshift over G. Let Y be a strongly irreducible SFT over G with no global period. If $h(X) < h(Y)$ and Y contains at least one factor of X, then X embeds into Y.*

If one selects $G = \mathbb{Z}^d$ for $d \geq 2$ in the above statement, then one does not immediately recover the theorem of Lightwood. We assume here a slightly stronger mixing condition on Y (strong irreducibility in place of square mixing). But, we do *not* assume that Y contains a point with a finite orbit; instead, we assume that Y has no global period, a condition which is automatically true for strongly irreducible subshifts over $\mathbb{Z}^d$.

In the following paragraphs, we briefly discuss the hypotheses of the above theorem and how they are invoked in the proof.

The condition that Y has no global period means that the shift action is faithful on Y. This condition is examined in detail in §3.3.1. This condition is necessary for the theorem; in particular, if a nonempty strongly aperiodic subshift X embeds into Y, then Y exhibits an aperiodic point and therefore has no global period.

The condition that X is strongly aperiodic allows us to derive systems of useful *quasi-tilings* (Definition 3.2.14) of the group G as factors of X, by appealing to a theorem of Downarowicz and Huczek [2, Lemma 3.4]. The strong aperiodicity is necessary for this; indeed, if X factors onto systems of quasi-tilings with arbitrarily large, disjoint tiles, then no point of X can exhibit a non-trivial period. The comparison property allows us to go one step further and derive from X systems of useful *exact tilings* of the group G. We accomplish this by adapting a construction of Downarowicz and Zhang (presented in both [24, Theorem 6.3] and [4, Theorem 7.5]).

The condition that $h(X) < h(Y)$ allows us to deduce that if a finite subset (namely, the shape of a tile in a given quasi-tiling) is large enough, then there are more patterns of that shape appearing in points of Y than in X. This implies that there is an injective map from tile patterns in X to tile patterns in Y. The condition that Y is a strongly irreducible SFT allows us to mix those tile patterns together into a single point of Y.

The condition that Y contains a factor of X means that there exists a continuous and shift-commuting map (a *homomorphism*) $\phi : X \to Y$, though not necessarily one that is injective. This is of course necessary for the theorem, since an embedding is a homomorphism, but the existence of a possibly non-injective homomorphism is a useful condition and one that is much easier to attain. It is automatically satisfied if, for instance, Y contains a fixed point. We utilize the homomorphism ϕ in our construction to code the boundary regions of the tiles in a given tiling, as well as the region of G not covered by any tile.

The condition that Y has no global period allows *marker* patterns to be constructed in Y (Theorem 3.3.4), which are used to encode the locations of the centers of the tiles of a given quasi-tiling within a controllably-sparse subset of the symbols of a point of Y. This extends the approach of Lightwood [8] who constructed marker patterns for $G = \mathbb{Z}^d$; the concept originates in the work of Krieger and Boyle for $G = \mathbb{Z}$ [19, 7]. The marker patterns allow one to uniquely decode, from a given image point $y = \psi(x)$, which quasi-tiling $t = \mathcal{T}(x)$ was used to construct y to begin with. Then, the tile pattern injections from earlier allow one to uniquely reconstruct the preimage x.

We note that the class of strongly aperiodic subshifts is well populated. It is known, for instance, that every countable group exhibits a strongly aperiodic subshift on the alphabet $\{0, 1\}$ [25, Theorem 2.4]. Moreover, if X_1 is a strongly aperiodic subshift and X_2 is an arbitrary subshift, then the direct product (Definition 3.2.13) $X_1 \times X_2$ is also strongly aperiodic; hence, there are at least as many strongly aperiodic subshifts as there are subshifts, in the sense of cardinality.

We also mention here the similarity of hypotheses between the present work and contemporary work by Huczek and Kopacz [18] which considers the "factor problem" (under what hypotheses on X and Y does there exist a surjective homomorphism from X to Y?) for subshifts over discrete amenable groups, which is complementary to the embedding problem. In particular, [18, Theorem 2.12] utilizes the comparison property of G in a form similar to Theorem 3.2.11. Additionally, [18, Theorem 5.1] assumes that the domain X is strongly irreducible and that X exhibits, for each finite subset $F \subset G$, a pattern which exhibits no element of F as a period. This latter condition is (under assumption of strong irreducibility) equivalent to our "separating elements" condition (Definition 3.3.1), which is the form in which we appeal to the faithfulness of the shift action of G on Y.

This chapter is organized as follows. In §3.2, we review some preliminary material about countable amenable groups, symbolic dynamics, and quasi-tilings over countable amenable groups with and without the comparison property. In §3.3 we present our main results. We first construct a subsystem $Y_0 \subset Y$ which we pass to in the construction of our embedding (Theorem 3.3.2). We then construct marker patterns for Y (Theorem 3.3.4). Finally, we present the construction of our embedding of X into Y (Theorem 3.3.5). In §3.4, we discuss various ways in which Theorem 3.3.5 could potentially be strengthened and associated obstacles.

3.2 Prelimaries

3.2.1 Amenable groups

In this section, we briefly review the theory of countable amenable groups and state a few lemmas which shall be needed later. For a more thorough introduction to the theory of dynamics on amenable groups, see [26].

Definition 3.2.1 (Invariance, amenability, and Følner sequences)**.** Let G be a countable group, let $K \subset G$ be a finite subset and let $\varepsilon > 0$. A finite subset $F \subset G$ is said to be (K, ε)-*invariant* if

$$|KF \triangle F| < \varepsilon |F|.$$

The group G is *amenable* if, for any finite subset $K \subset G$ and $\varepsilon > 0$, there exists a finite subset $F \subset G$ which is (K, ε)-invariant. Equivalently, G is amenable if there exists a sequence $(F_n)_n$ of finite subsets of G such that for each fixed finite subset $K \subset G$, it holds that

$$\lim_{n \to \infty} \frac{|KF_n \triangle F_n|}{|F_n|} = 0.$$

In this case, we say that $(F_n)_n$ is a *(right) Følner sequence.*

Throughout this chapter, G denotes a fixed countable amenable group with identity element e. It is classically known [27, Theorem 5.2] that G exhibits a Følner sequence $(F_n)_n$ that is ascending ($F_n \subset F_{n+1}$ for each n), that satisfies $\bigcup_n F_n = G$, and such that each F_n is *symmetric* ($F_n^{-1} = F_n$). Throughout this chapter, $(F_n)_n$ denotes a fixed Følner sequence with all of the above properties. The symmetry property implies that $(F_n)_n$ is both left Følner and right Følner, because for each finite subset $K \subset G$ we have that

$$|KF_n \triangle F_n| = |(KF_n \triangle F_n)^{-1}| = |F_n^{-1}K^{-1} \triangle F_n^{-1}| = |F_n K^{-1} \triangle F_n|.$$

In this setting, when a finite subset $F \subset G$ is described as “large", it is implied that F is (K, ε)-invariant for some finite subset $K \subset G$ and $\varepsilon > 0$, which may be clear from context or chosen arbitrarily beforehand. This sort of terminology is common but vague; a precise formulation is as follows. Let $\phi(F)$ be a property of finite subsets $F \subset G$. We shall say $\phi(F)$ holds for all *sufficiently large* F if there is a (K, ε) such that if F is (K, ε)-invariant then $\phi(F)$ is true. We shall say $\phi(F)$ holds for *arbitrarily large* subsets F if for every (K, ε) there is a set F for which F is (K, ε)-invariant and $\phi(F)$ is true.

We shall need the following elementary lemma; we omit the proof for brevity.

Lemma 3.2.1. *Let $K \subset G$ be a finite subset with $e \in K$. For any two finite subsets F_0, $F_1 \subset G$, it holds that*

$$|KF_1 \setminus F_1| \leq |KF_0 \setminus F_0| + |K||F_0 \triangle F_1|.$$

Next, we review concepts relating to the geometry of finite subsets of G.

Definition 3.2.2 (Boundary and interior)**.** Let $F \subset G$ and $K \subset G$ be finite subsets. The K-*boundary* of F is the subset

$$\partial_K F = \{f \in F : Kf \not\subset F\}$$

and the K-*interior* of F is the subset

$$\operatorname{int}_K F = \{f \in F : Kf \subset F\}.$$

Observe $F = \partial_K F \sqcup \operatorname{int}_K F$.

We shall need the following elementary lemmas; see [9, Lemma 2.1, Lemma 2.2].

Lemma 3.2.2. *Let F, $K \subset G$ be finite subsets and let $g \in G$ be arbitrary. If Kg intersects $\operatorname{int}_{KK^{-1}} F$, then $Kg \subset F$.*

Lemma 3.2.3. *Let F, $K \subset G$ be finite subsets. It holds that*

$$|\partial_K F| \leq |K||KF \triangle F|.$$

Next we review the notion of *density* for infinite subsets of G. Roughly speaking, an infinite subset $C \subset G$ has density $\rho \in [0, 1]$ if, whenever a finite subset F is sufficiently large, it holds for all $g \in G$ that

$$|F \cap Cg| \sim \rho|F|.$$

We use the Følner sequence to make this notion precise.

Definition 3.2.3 (Banach density)**.** Given a subset $C \subset G$, the *upper Banach density* of C is

$$\overline{D}(C) = \liminf_{n\to\infty} \sup_{g\in G} \frac{|F_n \cap Cg|}{|F_n|}$$

and the *lower Banach density* of C is

$$\underline{D}(C) = \limsup_{n\to\infty} \inf_{g\in G} \frac{|F_n \cap Cg|}{|F_n|}.$$

These definitions have also appeared in the recent work on quasi-tilings of amenable groups due to Downarowicz, Huczek, and Zhang [2, 3, 24, 4]. The value of the upper (resp. lower) density does not depend on the choice of Følner sequence [3, Lemma 2.9]. Note that $\overline{D}(C) = 1 - \underline{D}(G \setminus C)$ holds for any subset $C \subset G$.

Next we review a notion regarding how an infinite subset $C \subset G$ may be distributed throughout the group G.

Definition 3.2.4 (Separation)**.** Given a finite subset $L \subset G$, we say an infinite subset $C \subset G$ is *L-separated* if $Lc_1 \cap Lc_2 = \varnothing$ for every distinct pair $c_1 \neq c_2 \in C$.

Note that if C is L-separated, then so is Cg for each fixed $g \in G$. Using this fact, one may easily check that if C is L-separated then $\overline{D}(C) \leq 1/|L|$. The following lemma states something slightly stronger.

Lemma 3.2.4. *Let M, $L \subset G$ be finite subsets with $e \in M \subset L$. For any nonempty finite subset $F \subset G$ and any L-separated subset $C \subset G$, we have that*

$$\frac{|F \cap MC|}{|F|} \leq \frac{|M|}{|L|} + |M|\frac{|\partial_L F|}{|F|} + |M|\frac{|M^{-1}F \setminus F|}{|F|}.$$

Proof. Note that C is also M-separated by inclusion. Let $C^\times = \{c \in C : F \cap Mc \neq \varnothing\}$. Observe that $C^\times$ is finite, as $C^\times \subset M^{-1}F$. By the fact that C is M-separated, we have

$$|F \cap MC| \leq \sum_{c \in C^\times} |M| = |M||C^\times|.$$

Now let $C^\circ = \{c \in C : Lc \subset F\}$. Observe that $C^\circ \subset C^\times$. We emphasize that $C^\times$ is given in terms of M, while C° is given in terms of L. By definition, we have $LC^\circ \subset F$, in which case

$$|F| \geq |LC^\circ| = |L||C^\circ|$$

where the equality is a consequence of the fact that C is L-separated. Hence $|C^\circ| \leq |F|/|L|$.

Let $c \in C^\times \setminus C^\circ$ be fixed, in which case $F \cap Mc \neq \varnothing$ and $Lc \not\subset F$. Therefore, if $c \in F$ then $c \in \partial_L F$, while if $c \notin F$ then $c \in M^{-1}F \setminus F$. This demonstrates that

$$C^\times \setminus C^\circ \subset (\partial_L F) \cup (M^{-1}F \setminus F).$$

From this, we see that

$$
\begin{aligned}
|F \cap MC| &\leq |M||C^{\times}| \\
&= |M||C^{\circ}| + |M||C^{\times} \setminus C^{\circ}| \\
&\leq |M|\frac{|F|}{|L|} + |M||\partial_L F| + |M||M^{-1}F \setminus F|.
\end{aligned}
$$

After dividing by $|F|$, we obtain the conclusion. □

With M and L fixed as in the above lemma, if one chooses $F = F_n$ and lets n approach infinity, then one may easily see that $\overline{D}(MC) \leq |M|/|L|$ whenever $C \subset G$ is L-separated. However, what is especially significant for our purposes here is that the density of MC can be estimated by sets F which are sufficiently large with respect to M and L alone, and there is no dependence on C other than the fact that C is L-separated. If C were an arbitrary subset satisfying $\overline{D}(C) < \varepsilon$, then that density is not in general approximated by finite subsets except for very large ones, depending on C.

3.2.2 Shifts and subshifts

In this section, we briefly review symbolic dynamics for amenable groups and state a few useful lemmas. For a more thorough introduction to symbolic dynamics and dynamics on amenable groups, see [26, 28].

Definition 3.2.5 (Labelings and patterns)**.** Let $\mathcal{A}$ be a finite alphabet of symbols, endowed with the discrete topology. A function $x : G \to \mathcal{A}$ is an $\mathcal{A}$-*labeling* of G. The set of all such labelings is denoted $\mathcal{A}^G$, which is endowed with the product topology. Given a finite subset $F \subset G$, a function $p : F \to \mathcal{A}$ is called a *pattern*, said to be of *shape* F. The set of all patterns of shape F is denoted $\mathcal{A}^F$, and the set of all patterns of any shape is denoted $\mathcal{A}^*$.

Given a point $x \in \mathcal{A}^G$ and a finite subset $F \subset G$, in this chapter we shall take

$x(F)$ to mean the restriction of x to F, which is itself a pattern of shape F. This is normally denoted $x|_F \in \mathcal{A}^F$, but we raise the F from the subscript for readability. We shall also extend this notation to patterns where suitable; given a subset $F' \subset F$ and a pattern $p \in \mathcal{A}^F$, we shall take $p(F')$ to mean the restriction of p to F'.

Definition 3.2.6 (Shifts and subshifts)**.** The group G acts on $\mathcal{A}^G$ by way of (right) translations; for each fixed $g \in G$ we have a homeomorphism $\sigma^g : \mathcal{A}^G \to \mathcal{A}^G$ given by $\sigma^g(x)(g_1) = x(g_1 g)$ for every $g_1 \in G$ and $x \in X$. The action $\sigma = (\sigma^g)_g$ is called the *shift action* of G on $\mathcal{A}^G$, and together $(\mathcal{A}^G, \sigma)$ is a dynamical system called the *full shift* over $\mathcal{A}$. A *subshift* is a subset $X \subset \mathcal{A}^G$ which is σ-invariant and closed in the topology of $\mathcal{A}^G$.

A *fixed point* is a point $x \in \mathcal{A}^G$ such that $x(g_1) = x(g_2)$ for every g_1, $g_2 \in G$, in which case $X = \{x\}$ is a trivial subshift. We shall say a subshift $X \subset \mathcal{A}^G$ is *nontrivial* if it contains at least two points. We shall also assume for each subshift $X \subset \mathcal{A}^G$ that the alphabet $\mathcal{A}$ is taken to be *minimal*, in the sense that every symbol $a \in \mathcal{A}$ appears in at least one point $x \in X$. Note that for a nontrivial subshift X, it necessarily holds that $|\mathcal{A}| \geq 2$.

Definition 3.2.7 (Patterns in subshifts)**.** A pattern $p \in \mathcal{A}^F$ is said to *appear* in a point $x \in \mathcal{A}^G$ at an element $g \in G$ if $\sigma^g(x)(F) = p$. The set of all patterns of shape F appearing in any point of X is denoted $\mathcal{P}(F, X) \subset \mathcal{A}^F$. The set of all patterns of any shape appearing in any point of X is denoted $\mathcal{P}(X) \subset \mathcal{A}^*$.

Given a subshift $X \subset \mathcal{A}^G$ and a (finite or infinite) collection of patterns $\mathcal{F} \subset \mathcal{A}^*$, one may construct a subshift $X_0 \subset X$ by expressly *forbidding* the patterns in $\mathcal{F}$ from appearing in the points of X. That is,

$$X_0 = \{x \in X : \text{no pattern from } \mathcal{F} \text{ appears in } x\}.$$

We denote the subshift X_0 by $\langle X \mid \mathcal{F}\rangle$. Every subshift may be realized in this form; indeed, $X = \langle \mathcal{A}^G \mid \mathcal{A}^* \setminus \mathcal{P}(X)\rangle$ holds for every subshift $X \subset \mathcal{A}^G$.

Next we review some special classes of subshifts.

Definition 3.2.8 (Strongly aperiodic subshifts)**.** A point x is *aperiodic* if $\sigma^g(x) = x$ only when $g = e$. A subshift X is *strongly aperiodic* if every $x \in X$ is aperiodic. In other words, the action σ is *free* on X.

Definition 3.2.9 (Shifts of finite type)**.** A subshift $Y \subset \mathcal{A}^G$ is a *shift of finite type (SFT)* if there exists a finite collection of patterns $\mathcal{F} \subset \mathcal{A}^*$ such that $Y = \langle \mathcal{A}^G \mid \mathcal{F}\rangle$. For such a subshift, it is always possible to take $\mathcal{F}$ in the form $\mathcal{A}^K \setminus \mathcal{P}(K, Y)$ for some finite subset $K \subset G$. In this case, we say that K *witnesses* Y as an SFT.

We will need the following elementary lemma; see [9, Lemma 2.8].

Lemma 3.2.5. *Let Y be an SFT witnessed by $K \subset G$, let y_1, $y_2 \in Y$ be arbitrary points and let $F \subset G$ be a finite subset. If $y_1(\partial_{KK^{-1}}F) = y_2(\partial_{KK^{-1}}F)$, then the point y given by $y(g) = y_1(g)$ if $g \in F$ and $y(g) = y_2(g)$ if $g \in G \setminus F$ is a point belonging to Y.*

Definition 3.2.10 (Strong irreducibility)**.** A subshift $Y \subset \mathcal{A}^G$ is *strongly irreducible* if there exists a finite subset $K \subset G$ with $e \in K$ such that for any finite subsets F_1, $F_2 \subset G$ and allowed patterns $p_1 \in \mathcal{P}(F_1, Y)$ and $p_2 \in \mathcal{P}(F_2, Y)$, if KF_1 is disjoint from F_2 then there is a point $y \in Y$ such that $y(F_1) = p_1$ and $y(F_2) = p_2$. In this case, we say that K *witnesses* Y as strongly irreducible or that Y is a strongly irreducible subshift of *parameter* K.

Note that strong irreducibility is preserved under taking factors (Definition 3.2.12) and products (Definition 3.2.13) of subshifts. Strong irreducibility is equivalent to a seemingly stronger mixing condition, as the following lemma demonstrates. The proof relies on a standard compactness argument; we omit it for brevity.

Lemma 3.2.6. *Suppose Y is a strongly irreducible subshift of parameter $K \subset G$. For any finite subset $F \subset G$ and points y_1, $y_2 \in Y$, there is a point $y \in Y$ such that*

$$y(F) = y_1(F) \text{ and } y(G \setminus KF) = y_2(G \setminus KF).$$

Next we review topological entropy of subshifts.

Definition 3.2.11 (Entropy)**.** The *(topological) entropy* of a subshift X is given by

$$h(X) = \lim_{n\to\infty} h(F_n, X),$$

where $h(F, X) = \frac{1}{|F|} \log |\mathcal{P}(F, X)|$ for each nonempty finite subset $F \subset G$ and $(F_n)_n$ is a Følner sequence for G.

It is classically known [26, Theorem 4.38] that the limit above exists and does not depend on the choice of Følner sequence. Indeed, more recently it has been shown that

$$h(X) = \inf_F h(F, X)$$

where the infimum is taken over all finite subsets $F \subset G$ [29, Corollary 6.3].

Next we review homomorphisms between subshifts.

Definition 3.2.12 (Homomorphisms)**.** Let $\mathcal{A}_X$ and $\mathcal{A}_Y$ be finite alphabets and let $X \subset \mathcal{A}_X^G$ and $Y \subset \mathcal{A}_Y^G$ be subshifts. A *homomorphism* between X and Y is a map $\phi : X \to Y$ that is both continuous and shift-commuting. By the Curtis-Lyndon-Hedlund theorem [30, Theorem 1.8.1], a map $\phi : X \to Y$ is a homomorphism if and only if there exists a finite subset $F \subset G$ and a function $\Phi : \mathcal{P}(F, X) \to \mathcal{A}_Y$ such that for every $g \in G$ and $x \in X$ it holds that

$$\phi(x)(g) = \Phi(\sigma^g(x)(F)).$$

If ϕ is surjective, then ϕ is said to be a *factor map*, X is said to *factor onto* Y, and Y is said to be a *factor* of X. If ϕ is injective, then ϕ is said to be an *embedding* and X is said to *embed into* Y.

If X embeds into Y, then $h(X) \leq h(Y)$. If X factors onto Y, then $h(X) \geq h(Y)$.

Next we review the primary way by which we join two subshifts together.

Definition 3.2.13 (Product systems)**.** Let $\mathcal{A}$ and Λ be finite alphabets and let $X \subset \mathcal{A}^G$ and $T \subset \Lambda^G$ be subshifts. The product system $X \times T$ equipped with the action ς given by $\varsigma^g(x,t) = (\sigma^g(x), \sigma^g(t))$ is isomorphic to a subshift over the alphabet $\mathcal{A} \times \Lambda$, with (x,t) corresponding to the point $\tilde{x}$ such that $\tilde{x}(g) = (x(g), t(g))$ for each $g \in G$. Abusing notation, we regard $X \times T$ as a subshift of $(\mathcal{A} \times \Lambda)^G$.

The following lemma is a consequence of the fact that for a finite subset $F \subset G$ and subshifts X and T, $|\mathcal{P}(F, X \times T)| = |\mathcal{P}(F, X)| \cdot |\mathcal{P}(F, T)|$.

Lemma 3.2.7. *Let X and T be subshifts. Then $h(X \times T) = h(X) + h(T)$.*

3.2.3 Quasi-tilings

In this section we review quasi-tilings of amenable groups, which were originally introduced and studied by Ornstein and Weiss [31]. We also state a theorem of Downarowicz and Huczek which is essential for our main result.

Definition 3.2.14 (Quasi-tilings)**.** Let $\mathcal{S} = \{S_1, \ldots, S_r\}$ be a collection of finite, nonempty subsets of G which is "shift-irreducible" in the sense that there is no pair of distinct subsets S_1, $S_2 \in \mathcal{S}$ and element $g \in G$ for which $S_1 g = S_2$. We refer to these subsets as *shapes*. A *quasi-tiling* of G over $\mathcal{S}$ is an assignment of each shape $S \in \mathcal{S}$ to a (generally infinite) subset $C_S \subset G$ (called the set of *centers* for S) such that the sets $\{C_S : S \in \mathcal{S}\}$ are pairwise disjoint and the map $(S, c) \mapsto Sc$ is injective over $\{(S, c) : S \in \mathcal{S} \text{ and } c \in C_S\}$.

A quasi-tiling of G over $\mathcal{S}$ may be encoded as a point of the symbolic space $\Lambda(\mathcal{S})^G$, where $\Lambda(\mathcal{S}) = \mathcal{S} \cup \{\varnothing\}$ is thought of as an alphabet of $r+1$ symbols. A point $t \in \Lambda(\mathcal{S})^G$ encodes the quasi-tiling when $t(c) = S$ if and only if $c \in C_S$ for each $S \in \mathcal{S}$ and $c \in G$, and $t(g) = \varnothing$ otherwise. Here we shall identify t with the quasi-tiling it formally encodes. We shall write $C(t) = \{g \in G : t(g) \neq \varnothing\}$, which is precisely the set $\bigcup_{S \in \mathcal{S}} C_S$. A *tile* of a quasi-tiling t is a subset of G of the form $t(c)c$ where $c \in C(t)$. A *system of quasi-tilings* is a subshift $T \subset \Lambda(\mathcal{S})^G$ such that every $t \in T$ encodes a quasi-tiling of G over $\mathcal{S}$.

A quasi-tiling t is *disjoint* if $t(g_1)g_1 \cap t(g_2)g_2 = \varnothing$ for every $g_1 \neq g_2 \in G$. A quasi-tiling t is said to *cover* the group G if $\bigcup_g t(g)g = G$. An *exact* tiling is one which is both disjoint and covers G (in other words, the tiles of t form a partition of G). For most applications, it is not necessary that quasi-tilings be exactly disjoint or exactly covering; it is often sufficient to have a quasi-tiling whose tiles are "nearly disjoint" and which "nearly covers" G. The following definition formalizes the "nearly disjoint" condition.

Definition 3.2.15 (Retractions and ε-disjointness)**.** Given a quasi-tiling t, a *retraction* of t is any quasi-tiling $\mathrm{ret}(t)$ (which is in general given over a different collection of shapes than t) such that $\mathrm{ret}(t)(g) \subset t(g)$ for every $g \in G$. Given $\varepsilon > 0$, a quasi-tiling t is said to be *ε-disjoint* if it has a disjoint retraction $\mathrm{ret}(t)$ such that, for every $c \in C(t)$, it holds that

$$|t(c) \setminus \mathrm{ret}(t)(c)| < \varepsilon |t(c)|.$$

In words, every tile of t may be "retracted" to a subset of proportion at least $1 - \varepsilon$ such that the retracted subsets are all pairwise disjoint.

The following definition formalizes the "nearly covering" condition.

Definition 3.2.16 (ρ-covering)**.** Given $\rho \in (0,1)$, a quasi-tiling t is *ρ-covering* if

$$\underline{D}\Big(\bigcup_{g \in G} t(g)g\Big) \geq \rho$$

where $\underline{D}$ is the lower Banach density (Definition 3.2.3).

We shall need the following elementary lemma; see [3, Lemma 3.4].

Lemma 3.2.8. *Let ρ_0, $\rho_1 \in (0,1)$ be fixed. Suppose t_0 is a ρ_0-covering quasi-tiling and suppose t_1 is a disjoint retraction of t_0 such that $|t_1(c)| \geq \rho_1 |t_0(c)|$ for each $c \in C(t_0)$. Then t_1 is $\rho_0\rho_1$-covering.*

The existence of useful quasi-tilings over countable amenable groups (i.e., with arbitrarily large shapes and arbitrarily good near-disjointness and near-covering properties) has been known in some form since 1987, due first to Ornstein and Weiss [31, I.§2 Theorem 6]. This construction was sharpened in 2015 by Downarowicz, Huczek and Zhang who demonstrated that a countable amenable group exhibits an *exact* tiling with arbitrarily large shapes; moreover, one can find a system of such tilings which has topological entropy zero [3, Theorem 5.2].

For our purposes, we require not just that a system of nice quasi-tilings exists, but also that one may be obtained as a topological factor of a given subshift X. We have the following theorem of Downarowicz and Huczek [2, Lemma 3.4]. Not every property claimed here was stated in their theorem (here we state property (5) and the fact that the map $t \mapsto \mathrm{ret}(t)$ in (4) is a homomorphism), but a close reading of their proof reveals that it may be minorly modified to conclude this slightly stronger result. Here we provide a short argument which fills in the gaps, appealing to the construction in [2] as required.

Theorem 3.2.9. *Let X be a strongly aperiodic subshift, let $\varepsilon \in (0, 1/3)$ be arbitrary, and suppose that $r \in \mathbb{N}$ satisfies $(1-\varepsilon/2)^r < \varepsilon$. For any $n_0 \in \mathbb{N}$ and finite subset*

$L \subset G$, there is a collection of shapes $\mathcal{S} = \{F_{n_1}, \dots, F_{n_r}\}$ and a system of quasi-tilings $T \subset \Lambda(\mathcal{S})^G$ such that

1. *$n_0 < n_1 < \cdots < n_r$,*
2. *there is a factor map $\mathcal{T} : X \to T$,*
3. *every $t \in T$ is $(1-\varepsilon)$-covering,*
4. *every $t \in T$ is ε-disjoint as witnessed by a continuous and shift-commuting retraction map $t \mapsto \mathrm{ret}(t)$, and*
5. *for every $t \in T$, the set $C(t)$ is L-separated.*

Proof. In [2], n_1 is chosen as n_0+1 and n_i is inductively chosen so that F_{n_i} is (F_{n_j}, δ_j)-invariant for every $j < i$, where $\delta_j > 0$ is specified in the construction. From this, we infer that $n_i - n_{i-1}$ may be arbitrarily large for each $i = 1, \dots, r$. We therefore additionally assume that for each $i = 1, \dots, r$, we have the following properties.

(P1) $|F_{n_i} \cap F_{n_i}\ell| \geq (1-\varepsilon)|F_{n_i}|$ for every $\ell \in L^{-1}L$.

(P2) $F_{n_{i-1}}L^{-1}L \subset F_{n_i}$.

This is possible because the sequence $(F_n)_n$ was chosen to be both left and right Følner, to be ascending in n, and to satisfy $\bigcup_n F_n = G$ (Definition 3.2.1).

Let $S_i = F_{n_i}$ for each $i = 1, \dots, r$ and let $\mathcal{S} = \{S_1, \dots, S_r\}$. Our modifications to the choice of $\mathcal{S}$ preserve the invariance conditions assumed by [2]. We now define and construct everything else as in [2], summarized below.

Let $x \in X$ be fixed. In [2], the quasi-tiling $t = \mathcal{T}(x) \in \Lambda(\mathcal{S})^G$ is constructed inductively, with the tiles of shape S_r chosen first, then S_{r-1}, and so on, down to S_1. Consequently, for each $c \in C(t)$ there is a well-defined subset $C(t)_{<c} \subset C(t)$ which denotes the centers of all the tiles laid *before* the tile at c in the inductive process. Some tiles of the same shape are laid simultaneously, so the implied ordering given by

$c_1 < c_2$ if $c_1 \in C(t)_{<c_2}$ is not total. But, tiles laid simultaneously in the construction of [2] are necessarily disjoint.

This in hand, the retraction map $t \mapsto \mathrm{ret}(t)$ is given by

$$\mathrm{ret}(t)(c) = t(c) \setminus \Big(\bigcup_{c_0 \in C(t)_{<c}} t(c_0)c_0 \Big) c^{-1}$$

for each $c \in C(t)$, and $\mathrm{ret}(t)(g) = \varnothing$ otherwise. In [2] it is shown by induction that this is a disjoint retraction satisfying $|t(c) \setminus \mathrm{ret}(t)(c)| < \varepsilon |t(c)|$ for each $c \in C(t)$. Moreover, it is quick to check that the map $t \mapsto \mathrm{ret}(t)$ is continuous and shift-commuting, as the elements $c_0 \in C(t)_{<c}$ with $t(c)c \cap t(c_0)c_0 \neq \varnothing$ are determined by $\sigma^c(t)(F)$ for some (possibly very large) finite subset $F \subset G$. The construction in [2] also gives that t is $(1-\varepsilon)$-covering.

Let $T = \mathcal{T}(X) \subset \Lambda(\mathcal{S})^G$. From the observations of the previous paragraph, we see that properties (1) through (4) hold. For the theorem, it remains to check property (5): that $C(t)$ is L-separated for each $t \in T$. Let $t \in T$ be fixed and suppose to the contrary that $Lc_1 \cap Lc_2 \neq \varnothing$ for some distinct $c_1 \neq c_2 \in C(t)$, in which case $c_1 c_2^{-1} \in L^{-1}L$. Let $t(c_1) = S_i$ and $t(c_2) = S_j$ and suppose without loss of generality that $i \leq j$.

If $i = j$ then let $S = S_i = S_j$, in which case $|Sc_1 \cap Sc_2| = |S \cap Sc_1 c_2^{-1}| \geq (1-\varepsilon)|S|$ by the assumed property (P1). Note that $R_1 = \mathrm{ret}(t)(c_1) \subset S$ and $R_2 = \mathrm{ret}(t)(c_2) \subset S$ are subsets such that $|S \setminus R_1| < \varepsilon |S|$ and $|S \setminus R_2| < \varepsilon |S|$ by construction. Moreover, by construction $R_1 c_1 \cap R_2 c_2 = \varnothing$, in which case $Sc_1 \cap Sc_2 \subset (Sc_1 \setminus R_1 c_1) \cup (Sc_2 \setminus R_2 c_2)$. This implies that

$$|Sc_1 \cap Sc_2| \leq |Sc_1 \setminus R_1 c_1| + |Sc_2 \setminus R_2 c_2| < 2\varepsilon |S|.$$

We thus obtain $(1-\varepsilon)|S| \leq |Sc_1 \cap Sc_2| < 2\varepsilon |S|$, which contradicts the assumption that $\varepsilon < 1/3$.

In the second case, suppose $i < j$. Then, $c_1 c_2^{-1} \in L^{-1}L$ implies that $S_i c_1 c_2^{-1} \subset S_i L^{-1} L \subset S_j$ by the assumed property (P2). Hence, $t(c_1)c_1 \subset t(c_2)c_2$. This doesn't immediately contradict the ε-disjointness of t, as it may be the case that $t(c_1)c_1 \subset t(c_2)c_2 \setminus \mathrm{ret}(t)(c_2)c_2$. However, here we appeal to a property of the retraction map which is easily checked by induction. For each $c \in C(t)$, we have that

$$\bigcup_{c_0 \in C(t)_{<c}} t(c_0)c_0 = \bigcup_{c_0 \in C(t)_{<c}} \mathrm{ret}(t)(c_0)c_0. \tag{R1}$$

Moreover, the assumption that $i < j$ implies that $c_2 \in C(t)_{<c_1}$, as the tiles are placed in order of largest to smallest. In that case, we have that

$$\mathrm{ret}(t)(c_1)c_1 \subset t(c_1)c_1 \subset t(c_2)c_2 \subset \bigcup_{c_0 \in C(t)_{<c_1}} t(c_0)c_0.$$

This, together with property (R1), implies that $\mathrm{ret}(t)(c_1)c_1$ intersects $\mathrm{ret}(t)(c_0)c_0$ for some $c_0 \in C(t)_{<c_1}$, which contradicts the disjointness of $\mathrm{ret}(t)$.

This covers all cases, so we conclude that $C(t)$ must be L-separated for every $t \in T$. With this, we have verified all properties for the theorem not already proved in [2]. □

One sees as a consequence of the above theorem that X factors directly onto $\mathrm{ret}(T)$, a system of disjoint quasi-tilings with arbitrarily large shapes and near-covering of G. This is stated by Downarowicz and Huczek [2, Corollary 3.5]. However, one has to give up control of the number of tile shapes in exchange for perfect disjointness of the tiles.

For our purposes, we shall make use of the intermediate factor T. Each $t \in T$ carries all of the information needed to construct a disjoint quasi-tiling (by way of taking the retraction $\mathrm{ret}(t)$); we retain control of the “density" of that information by controlling the number of tile shapes and distributing the centers of the tiles arbitrarily sparsely throughout the group.

3.2.4 Comparison property

In this section, we turn our attention to the case where G has the comparison property. In short, with the comparison property one may demonstrate that, if ε is sufficiently small and if the shapes in $\mathcal{S}$ are sufficiently large, then the subshift T in Theorem 3.2.9 factors onto a system T_1 of exact tilings.

For a thorough discussion of the comparison property for countable amenable groups and its consequences, see [4]. Here, we only repeat that the class of groups with the comparison property includes all countable groups with no finitely generated subgroup of exponential growth, and it is still unknown whether there exists a countable amenable group without the comparison property.

The following theorem is a consequence of the main results of [24, Proposition 4.3, Theorem 4.7], also appearing in [4, Proposition 6.10, Theorem 6.12]. We state the result in this form for convenience; this is the form in which we shall appeal to the comparison property later in our construction.

Given a subshift T, suppose we assign to each $t \in T$ a subset $G_t \subset G$. We can encode each subset G_t by its indicator function $\chi_{G_t} \in \{0,1\}^G$. We say that the assignment $t \mapsto G_t$ is continuous and shift-commuting if the map $t \mapsto \chi_{G_t}$ is continuous and shift-commuting in the usual sense (Definition 3.2.12). Equivalently, there is a finite subset $F \subset G$ and a collection of patterns $\mathcal{G} \subset \mathcal{P}(F,T)$ such that for every $g \in G$ and $t \in T$, we have $g \in G_t$ if and only if $\sigma^g(t)(F) \in \mathcal{G}$.

Theorem 3.2.10. *Suppose G has the comparison property. Let T be a subshift over G and suppose for every $t \in T$ we have corresponding disjoint subsets A_t, $B_t \subset G$ such that*

1. *the assignments $t \mapsto A_t$, B_t are continuous and shift commuting, and*
2. *there exists an $\varepsilon > 0$ such that $\underline{D}(B_t) - \overline{D}(A_t) > \varepsilon$ for every $t \in T$.*

Then there is a family of injections $\phi_t : A_t \to B_t$ induced by a block code, in the sense that there is a finite subset $F \subset G$ and a function $\Phi : \mathcal{P}(F, T) \to F$ such that for every $t \in T$ and every $g \in A_t$ it holds that

$$\phi_t(g) = \Phi(\sigma^g(t)(F))g.$$

The following theorem is implicitly proved in both [24, Theorem 6.3] and [4, Theorem 7.5]. The proof, a construction which we adapt in part of our Theorem 3.3.5, relies on the characterization of the comparison property given in the previous theorem. We appeal to each of these theorems later, utilizing Theorem 3.2.10 in Theorem 3.3.5 and utilizing Theorem 3.2.11 in Theorem 3.3.2.

Theorem 3.2.11. *Let G be a countable amenable group with the comparison property. For every finite subset $K \subset G$ and $\varepsilon > 0$, there is a $\delta > 0$ such that, if T is a system of disjoint, $(1-\delta)$-covering quasi-tilings with (K, δ)-invariant shapes, then T factors onto a system T_1 of exact tilings with (K, ε)-invariant shapes, by way of a factor map $\mathrm{ex} : T \to T_1$ such that $C(t) = C(\mathrm{ex}(t))$ for every $t \in T$ and $t(c) \subset \mathrm{ex}(t)(c)$ for every $c \in C(t)$.*

3.3 Theorems

3.3.1 Target system

In order to construct an embedding from a given subshift X into a given subshift Y, it will first be necessary for our construction to pass to a subsystem $Y_0 \subset Y$ in a way that preserves most of the conditions on Y. In this section, we construct that subshift Y_0.

In the following definition, we borrow a phrase from functional analysis.

Definition 3.3.1 (Separating elements)**.** Let G be a discrete group. We say that a subshift Y over G *separates* elements of G if for every pair of distinct elements g_1,

$g_2 \in G$, there is a point $y \in Y$ such that $y(g_1) \neq y(g_2)$. Because Y is shift-invariant, this is true if and only if it holds that for each $g \neq e$, there is a point $y \in Y$ such that $y(e) \neq y(g)$.

This condition is not invariant under topological conjugacy. However, a subshift Y_1 is conjugate to a subshift Y which separates elements of G if and only if there is a finite subset $F \subset G$ such that, for every $g \neq e$, there is a point $y_1 \in Y_1$ with $\sigma^g(y_1)(F) \neq y_1(F)$. The forward implication follows from the Curtis-Lyndon-Hedlund theorem, and the converse implication follows from passing to a higher block presentation of Y.

This condition is similar to the condition that Y has no *global period*, meaning that there is no element $g \neq e$ such that $\sigma^g(y) = y$ for every $y \in Y$ (equivalently, the shift action σ is *faithful* on Y). If a subshift Y separates elements of G, then Y necessarily has no global period. The converse holds when G is abelian, but not in general. When Y is strongly irreducible, we have the following partial converse.

Lemma 3.3.1. *Let G be a discrete group and let Y be a nontrivial strongly irreducible subshift over G. If Y has no global period, then Y is conjugate to a subshift which separates elements of G.*

Proof. Let $K \subset G$ be a finite subset containing e which witnesses the strong irreducibility of Y and write $K = \{e, k_1, k_2, \ldots, k_n\}$. Because Y has no global period, for each $i \leq n$ there exists a point $y_i \in Y$ and an element $g_i \in G$ such that $\sigma^{k_i}(y_i)(g_i) \neq y_i(g_i)$.

Let $F = \{e, g_1, g_2, \ldots, g_n\}$ and let $g \in G \setminus \{e\}$ be chosen arbitrarily. Because Y is nontrivial and strongly irreducible, if $g \in G \setminus K$ then we may construct a point $y \in Y$ such that $y(e) \neq y(g)$, in which case $y(F) \neq \sigma^g(y)(F)$. If instead $g = k_i$ for some $i \leq n$, then the fact that $y_i(g_i) \neq \sigma^{k_i}(y_i)(g_i)$ implies that $y_i(F) \neq \sigma^g(y_i)(F)$.

We have shown that for every $g \neq e$ there is a point $y \in Y$ for which $y(F) \neq \sigma^g(y)(F)$, in which case Y is conjugate to a subshift which separates elements of

G by the observation in the paragraph following Definition 3.3.1, and the lemma is completed. □

The previous lemma demonstrates that when Y is strongly irreducible, we may use conjugacy to pass back and forth between the condition that Y has no global period and the condition that Y separates elements of G. Indeed, in the proof of Theorem 3.3.5 we shall appeal to the fact that Y has no global period in order to replace Y with a conjugate subshift which separates elements of G.

It is not in general true that a strongly irreducible subshift automatically has no global period. However, as noted in the proof of the above lemma, if a subshift Y is nontrivial and strongly irreducible as witnessed by $K \subset G$, then for each $g \notin K$ there is a point $y \in Y$ such that $y \neq \sigma^g(y)$. This implies that any element which is a global period of Y must belong to K (moreover, the subgroup of G consisting of all global periods of Y must be contained in K). In this sense, to assume that a strongly irreducible subshift Y also has no global period only imposes finitely many additional conditions on Y.

From this, we also see that when G is a torsion-free group (such as $\mathbb{Z}^d$), then a nontrivial strongly irreducible subshift over G automatically has no global period. In fact, it is not difficult to show that a strongly irreducible subshift over a group G with no element of finite order must necessarily separate elements of G.

We now proceed with the main construction for this section. Given a strongly irreducible SFT Y which separates elements of G and has positive entropy, the following theorem produces a subsystem $Y_0 \subset Y$ which is also strongly irreducible, separates elements of G, and has entropy in any arbitrary subinterval of $[0, h(Y)]$. The construction presented below is a modification of the construction presented in [9, Theorem 4.1]. Here we invoke the comparison property in order to construct a strongly irreducible system of exact tilings of G.

Theorem 3.3.2. *Let G be a countable amenable group with the comparison property, let Y be a strongly irreducible SFT over G which separates elements of G, and let $\tilde{Y} \subset Y$ be a subshift satisfying $h(\tilde{Y}) < h(Y)$. For every subinterval $(a,b) \subset [h(\tilde{Y}), h(Y)]$, there is a strongly irreducible subshift Y_0 which separates elements of G and satisfies $\tilde{Y} \subset Y_0 \subset Y$ and $a < h(Y_0) < b$.*

Proof. Let $\mathcal{A}$ be a finite alphabet such that $Y \subset \mathcal{A}^G$. Let $K \subset G$ be a finite subset with $e \in K$ chosen to witness Y as a strongly irreducible SFT. We shall abbreviate $\operatorname{int}^n F = \operatorname{int}_{K^n} F$ and $\partial^n F = \partial_{K^n} F$ for each natural $n \in \mathbb{N}$ and finite subset $F \subset G$ for the remainder of this proof. We shall also abbreviate $\partial^n p = p(\partial^n F)$ for each pattern p of shape F.

Choose $\varepsilon > 0$ such that

$$\varepsilon < \min\left(\frac{b-a}{3+\log 2 + 2\log|\mathcal{A}|}, \frac{b-a}{5+4\log|\mathcal{A}|}\right).$$

It is a theorem of Frisch and Tamuz [32, Theorem 2.1] that for any (K,δ) there exists a strongly irreducible system of disjoint, $(1-\delta)$-covering quasi-tilings of G whose every shape is (K,δ)-invariant and whose entropy is less than δ. This, in combination with Theorem 3.2.11 and the fact that strong irreducibility is preserved under factor maps, implies the following. There exists a finite collection of shapes $\mathcal{S}$ and a strongly irreducible system $T \subset \Lambda(\mathcal{S})^G$ of exact tilings of G such that $h(T) < \varepsilon$ and every shape $S \in \mathcal{S}$ satisfies the following.

(S1) $K \subset \operatorname{int}^2 S$,

(S2) $|S| > \varepsilon^{-1}$ and $2|S| < e^{\varepsilon|S|}$,

(S3) $|\partial^2 S| < \varepsilon|S|$, and

(S4) $h(S, \tilde{Y}) < h(\tilde{Y}) + \varepsilon$.

For the majority of this proof, we operate primarily in the product system $Z_0 = Y \times T$. For each finite subset $F \subset G$ and pattern $p \in \mathcal{P}(F, Z_0)$, we shall write $p = (p^Y, p^T)$ where $p^Y \in \mathcal{P}(F, Y)$ and $p^T \in \mathcal{P}(F, T)$. If $F = S$ for some $S \in \mathcal{S}$, then we shall describe p as a “block".

A block $b \in \mathcal{P}(S, Z_0)$ is called *aligned* if $b^T(e) = S$ (note $e \in K$ and $K \subset S$ by (S1)). Note that if b is aligned then $b^T(s) = \varnothing$ for each $s \in S \setminus \{e\}$, by the fact that every tiling $t \in T$ is disjoint. For a given subshift $Z \subset Z_0$, we denote the subcollection of all aligned blocks of shape S allowed in Z by

$$\mathcal{P}^a(S, Z) \subset \mathcal{P}(S, Z) \subset (\mathcal{A} \times \Lambda(\mathcal{S}))^S$$

where the superscript a identifies the subcollection.

Let $\pi : Z_0 \to Y$ be the projection map defined by $\pi(y, t) = y$ for each $(y, t) \in Z_0$, which is a homomorphism. For each subshift $Z \subset Z_0$ and each fixed $z = (y, t) \in Z$, we have $y = \pi(z) \in \pi(Z)$ and $t \in T$, thus $Z \subset \pi(Z) \times T$. Consequently, for each subshift $Z \subset Z_0$ it holds that

$$h(Z) \leq h(\pi(Z) \times T) = h(\pi(Z)) + h(T) < h(\pi(Z)) + \varepsilon$$

where above we have used Lemma 3.2.7 and the fact that $h(T) < \varepsilon$.

Let $S \in \mathcal{S}$ be fixed. Here we choose and fix a collection of aligned blocks $\mathcal{W}(S) \subset \mathcal{P}^a(S, Z_0)$. We shall refer to these as “witness patterns", because they will later allow us to demonstrate (“witness") for each $g \in G$ a point $z \in Z_0$ such that $z^Y(e) \neq z^Y(g)$. These witness patterns shall be of three types.

For the first type, let $s \in S \setminus \{e\}$ be arbitrary. Because Y separates elements of G, there is a pattern $b^Y \in \mathcal{P}(S, Y)$ such that $b^Y(e) \neq b^Y(s)$. If $b^T \in \mathcal{P}(S, T)$ is given by $b^T(e) = S$ and $b^T(s) = \varnothing$ for each $s \in S \setminus \{e\}$, then $b = (b^Y, b^T)$ is an allowed block in Z_0 which satisfies $b^Y(e) \neq b^Y(s)$. For each $s \in S \setminus \{e\}$, pick and save one such block

b to the collection $\mathcal{W}(S)$.

For the second type, let $0 \in \mathcal{A}$ be a distinguished symbol of the alphabet. For every $y \in Y$, we can find a block $b \in \mathcal{P}(S, Z_0)$ such that $b(e) = (0, S)$ and $b^Y(\partial^2 S) = y(\partial^2 S)$. This is because Y is strongly irreducible of parameter K and property (S1). For each pattern $y(\partial^2 S) \in \mathcal{P}(\partial^2 S, Y)$, pick and save one such block b to the collection $\mathcal{W}(S)$.

For the third type, let $1 \in \mathcal{A}$ be a distinguished symbol different from 0. For each $s \in S$, we can find a block $b \in \mathcal{P}(S, Y)$ such that $b^T(e) = S$ and $b^Y(s) = 1$. For each $s \in S$, pick and save one such block b to the collection $\mathcal{W}(S)$. This completes the description of the collection $\mathcal{W}(S)$. Note that $|\mathcal{W}(S)| \leq 2|S| + \mathcal{A}^{|\partial^2 S|} < e^{\varepsilon|S|} \cdot |\mathcal{A}|^{\varepsilon|S|}$ by properties (S2) and (S3).

We now construct a descending chain of subshifts $(Z_n)_n$ of Z_0 and claim that the subshift Y_0 desired for the theorem shall be given by $Y_0 = \pi(Z_n)$ for some $n \leq N$. Suppose for induction that $Z_n \subset Z_0$ has been constructed for $n \geq 0$. If there exists a shape $S_n \in \mathcal{S}$ and an aligned block $\beta_n \in \mathcal{P}^a(S_n, Z_n)$ such that

(B1) β_n^Y does not appear in $\tilde{Y}$ ($\beta_n^Y \notin \mathcal{P}(S_n, \tilde{Y})$),

(B2) β_n is not one of the reserved witness patterns ($\beta_n \notin \mathcal{W}(S_n)$), and

(B3) there exists an aligned block $b \in \mathcal{P}^a(S_n, Z_n)$ with $b \neq \beta_n$ and $\partial^2 b = \partial^2 \beta_n$,

then let $Z_{n+1} = \langle Z_n \mid \beta_n \rangle$. If no such block exists for any shape, then the descending chain is finite in length and Z_n is the terminal subshift.

In fact, the chain *must* be finite in length. This is because $Z_{n+1} \subset Z_n$ implies $\mathcal{P}(S, Z_{n+1}) \subset \mathcal{P}(S, Z_n)$ for each $S \in \mathcal{S}$, and $\mathcal{P}(S_n, Z_{n+1}) \sqcup \{\beta_n\} \subset \mathcal{P}(S_n, Z_n)$ for each $n \geq 0$. Hence we have that

$$\sum_{S \in \mathcal{S}} |\mathcal{P}(S, Z_n)|$$

is a nonnegative integer sequence which strictly decreases with n, and therefore must terminate. Let $N \geq 0$ be the index of the terminal subshift, and note by construction that for each $S \in \mathcal{S}$ and each aligned block $b \in \mathcal{P}^a(S, Z_N)$, either b is uniquely determined by $\partial^2 b$, or $b \in \mathcal{W}(S)$, or $b^Y \in \mathcal{P}(S, \tilde{Y})$.

We note the following intermediate lemma which shall be referenced multiple times in the remainder of this proof.

Lemma 3.3.3. *For each $n < N$ and $(y,t) \in Z_n$, there is a point $(y^*, t) \in Z_{n+1}$ satisfying $(y,t)(g) = (y^*,t)(g)$ for every $g \notin \bigcup_c \operatorname{int} t(c)c$, where the union is taken over all $c \in G$ with $\sigma^c(y,t) = \beta_n$.*

Proof. By property (B3), there is a block $b \in \mathcal{P}^a(S_n, Z_n)$ such that $b(\partial^2 S_n) = \beta_n(\partial^2 S_n)$. In words, we simply replace every appearance of β_n in (y,t) with b to construct the point (y^*, t).

Precisely, let $(c_k)_k$ enumerate the group elements c for which $\sigma^c(y,t) = \beta_n$, which is necessarily a subset of the elements c for which $t(c) = S_n$ because β_n is aligned. Let $y^* \in \mathcal{A}^G$ be given by $y^*(g) = b^Y(gc_k^{-1})$ whenever $g \in S_n c_k$, and $y^*(g) = y(g)$ for every $g \notin \bigcup_k S_n c_k$. This point is well defined because t is a disjoint tiling.

Because Y is an SFT of parameter K and by Lemma 3.2.5, we see that y^* is an allowed point of Y. By construction, for every $S \in \mathcal{S}$ and $c \in G$ with $t(c) = S$, either $\sigma^c(y,t)(S) = \sigma^c(y^*,t)(S)$ or $\sigma^c(y,t) = \beta_n$ and $\sigma^c(y^*,t) = b$. Consequently, no forbidden block β_i for any $i < n$ can appear in (y^*,t), else that would force an appearance of β_i in (y,t), where β_i is already forbidden. Moreover, the block β_n cannot appear in (y^*,t) by construction, thus $(y^*,t) \in Z_{n+1}$. Finally, we see by construction that $(y,t)(g) = (y^*,t)(g)$ for every $g \notin \bigcup_c \operatorname{int} t(c)c$ where the union is taken over all $c \in G$ with $\sigma^c(y,t) = \beta_n$ as desired, hence we are done. □

Now continuing the proof of Theorem 3.3.2, we claim that for each n, the subshift $\pi(Z_n)$ is strongly irreducible, separates elements of G, and satisfies $\tilde{Y} \subset \pi(Z_n) \subset Y$. Moreover, we claim that $h(\pi(Z_n)) - h(\pi(Z_{n+1})) < b - a$ for each $n < N$, and $h(\pi(Z_N)) - h(\tilde{Y}) < b - a$.

For $\tilde{Y} \subset \pi(Z_n) \subset Y$, let $\tilde{y} \in \tilde{Y}$ and $t \in T$ be arbitrary, in which case $(\tilde{y}, t) \in Z_0$. Note that, for each $n < N$, the block β_n cannot appear in $(\tilde{y}, t)$, else that would imply that β_n^Y appears in $\tilde{y}$, contradicting property (B1). Thus, $(\tilde{y}, t) \in Z_n$ for each $n \leq N$. This demonstrates that $\tilde{Y} \times T \subset Z_n$, and in particular $\tilde{Y} \subset \pi(Z_n) \subset Y$, for each $n \leq N$.

For the strong irreducibility, note that $Z_0 = Y \times T$ is strongly irreducible because both Y and T are strongly irreducible. For induction, suppose Z_n is strongly irreducible of parameter $K_n \subset G$ (with $e \in K_n$) for a fixed $n < N$. Let $U = \bigcup_{S \in \mathcal{S}} S$. We claim that Z_{n+1} is strongly irreducible of parameter $UU^{-1}K_nUU^{-1}$.

Indeed, let (y_1, t_1), $(y_2, t_2) \in Z_{n+1}$ be arbitrary points, and let F_1, $F_2 \subset G$ be arbitrary finite subsets with $K_nUU^{-1}F_1 \cap UU^{-1}F_2 = \varnothing$. As $Z_{n+1} \subset Z_n$ and Z_n is strongly irreducible of parameter K_n, there is a point $(y, t) \in Z_n$ with $(y, t)(UU^{-1}F_1) = (y_1, t_1)(UU^{-1}F_1)$ and $(y, t)(UU^{-1}F_2) = (y_2, t_2)(UU^{-1}F_2)$.

Now consider the forbidden block β_n of shape $S_n \in \mathcal{S}$. Suppose $\sigma^g(y, t)(S_n) = \beta_n$ for some $g \in G$ with $S_ng \cap F_1 \neq \varnothing$. It follows that $g \in S_n^{-1}F_1$, hence $S_ng \subset S_nS_n^{-1} \subset UU^{-1}F_1$, hence $\sigma^g(y_1, t_1)(S_n) = \sigma^g(y, t)(S_n) = \beta_n$, contradicting the fact that $(y_1, t_1) \in Z_{n+1}$ and β_n is forbidden in Z_{n+1}. From this observation (and by an identical argument for F_2), we see that if β_n appears anywhere in (y_0, t) then it does *not* appear on any tile of t which intersects either F_1 or F_2.

Let $(y^*, t) \in Z_{n+1}$ be the point delivered by Lemma 3.3.3 as applied to (y, t). From the previous paragraph, we have $(y^*, t)(F_i) = (y, t)(F_i) = (y_i, t_i)(F_i)$ for each $i = 1, 2$. We conclude that Z_{n+1} is strongly irreducible of parameter $UU^{-1}K_nUU^{-1}$, therefore completing the induction. As strong irreducibility is preserved under taking factors, it follows that $\pi(Z_n)$ is strongly irreducible for each $n \leq N$.

To see that each $\pi(Z_n)$ separates elements of G, let $n \leq N$ and $g \neq e$ be fixed. We proceed by cases on g.

If $g \in S$ for some $S \in \mathcal{S}$, then there is a witness block $b \in \mathcal{W}(S)$ such that $b^Y(e) \neq$

$b^Y(g)$. Choose and fix $(y_0, t) \in Z_0$ such that $(y_0, t)(S) = b$ and apply Lemma 3.3.3 at most N times to produce the point $(y_n, t) \in Z_n$. Note $(y_n, t)(S) = (y_0, t)(S) = b$, as $b \neq \beta_n$ for every $n < N$ by property (B2). Consequently, $y_n \in \pi(Z_n)$ satisfies $y_n(e) = b^Y(e) \neq b^Y(g) = y_n(g)$.

If $g \notin S$ for any $S \in \mathcal{S}$, then pick $S_0 \in \mathcal{S}$ arbitrarily and fix $t \in T$ with $t(e) = S_0$. Because t is an exact tiling, there is a unique $c \neq e$ such that $g \in t(c)c$. Let $S_1 = t(c)$ and write $g = s_1 c$ for some $s_1 \in S_1 \in \mathcal{S}$. Recall $0, 1 \in \mathcal{A}$ are two distinguished symbols determined in the construction of $\mathcal{W}(S)$. By construction, there is a witness block $b_1 \in \mathcal{W}(S_1)$ with $b_1^Y(s_1) = 1$. Pick any $y \in Y$ with $\sigma^c(y)(S_1) = b_1$. By construction, there is a witness block $b_0 \in \mathcal{W}(S_0)$ with $b_0^Y(e) = 0$ and $\partial^2 b_0^Y = y(\partial^2 S_0)$. Because Y is an SFT of parameter K and by property (S1) and Lemma 3.2.5, there is a point $y^* \in Y$ given by $y^*(S_0) = b_0^Y$ and $y^*(g) = y(g)$ otherwise. In particular, $\sigma^c(y^*)(S_1) = b_1^Y$. Then, apply Lemma 3.3.3 at most N times to the initial point $(y^*, t) \in Z_0$, thus yielding $(y_n, t) \in Z_n$. Observe that $(y_n, t)(S_0) = (y^*, t)(S_0) = b_0$ and $\sigma^c(y_n, t)(S_1) = \sigma^c(y^*, t)(S_1) = b_1$, because $b_0 \neq \beta_n \neq b_1$ for every $n < N$ by property (B2). Consequently, $y_n \in \pi(Z_n)$ satisfies $y_n(e) = 0 \neq 1 = y_n(g)$. This finishes all cases and demonstrates that $\pi(Z_n)$ separates elements of G for each $n \leq N$.

The proofs that $h(\pi(Z_n)) - h(\pi(Z_{n+1})) < b - a$ for every $n < N$ and $h(\pi(Z_N)) - h(\tilde{Y}) < b - a$ are nearly identical to arguments appearing in [9, Theorem 4.1]. We proceed quickly through the argument here. Choose a finite subset $F \subset G$ such that $|h(Z_n) - h(F, Z_n)| < \varepsilon$ for every $n \leq N$, $|h(F, T) - h(T)| < \varepsilon$ (this implies in particular that $h(F, T) < 2\varepsilon$ by choice of T), and $|F \setminus U^{-1}F| < \varepsilon |F|$, where $U = \bigcup_{S \in \mathcal{S}} S$. For each $n \leq N$, let

$$P(n) = \sum_{t(F)} \prod_{c} |\mathcal{P}^a(\operatorname{int} t(c), Z_n)|$$

where the sum is taken over all patterns $t(F) \in \mathcal{P}(F, T)$ and the product is taken over all $c \in C(t) \cap U^{-1}F \cap F$. For each $n \leq N$, it holds that

$$P(n) \leq |\mathcal{P}(F, Z_n)| \leq |\mathcal{A}|^{2\varepsilon|F|} \cdot P(n).$$

The first inequality follows from the strong irreducibility of Y in combination with Lemma 3.3.3. The latter inequality follows from projecting a pattern $p \in \mathcal{P}(F, Z_n)$ to the interiors of the tiles described by $p^T = t(F) \in \mathcal{P}(F, T)$ which are contained in F. This projection determines p up to the portion of F not covered by those tile interiors, a subset of F of size at most $2\varepsilon|F|$ by the combination of the fact that the tiling is exact, the assumed invariance condition on F, and property (S3).

Moreover, for each $n < N$ it holds that $P(n) \leq 2^{\varepsilon|F|} \cdot P(n+1)$. This follows from the fact that $|\mathcal{P}^a(\operatorname{int} S, Z_n)| - |\mathcal{P}^a(\operatorname{int} S, Z_{n+1})| \leq 1$ (at most one aligned block is removed as one passes from Z_n to Z_{n+1}), in combination with the assumed invariance condition on F.

The above, in combination with the assumed entropy estimating properties of F, implies that

$$h(Z_n) < h(Z_{n+1}) + 2\varepsilon + \varepsilon \log 2 + 2\varepsilon \log |\mathcal{A}|.$$

This, together with the fact that $h(Z_n) < h(\pi(Z_n)) + \varepsilon$ for each $n \leq N$ and the choice of ε, finally gives that $h(\pi(Z_n)) - h(\pi(Z_{n+1}) < b - a$ holds for each $n < N$.

Next, consider the terminal subshift Z_N. Recall that for every shape $S \in \mathcal{S}$, every aligned block $b \in \mathcal{P}^a(S, Z_N)$ either belongs to $\mathcal{W}(S)$, is uniquely determined by $\partial^2 b$, or satisfies $b^Y \in \mathcal{P}(S, \tilde{Y})$ by construction. Recall also that $|\mathcal{W}(S)| \leq e^{\varepsilon|S|} \cdot |\mathcal{A}|^{\varepsilon|S|}$. This, together with property (S4), imply that for every shape $S \in \mathcal{S}$ it holds that

$$|\mathcal{P}^a(\operatorname{int} S, Z_N)| \leq e^{h(\tilde{Y})|S|} \cdot e^{2\varepsilon|S|} \cdot |\mathcal{A}|^{2\varepsilon|S|}.$$

This, together with the facts that $h(F, T) < 2\varepsilon$ and $|\mathcal{P}(F, Z_N)| \leq |\mathcal{A}|^{2\varepsilon|F|} \cdot P(N)$

argued earlier, gives us that

$$|\mathcal{P}(F, Z_N)| \leq e^{h(\tilde{Y})|F|} \cdot e^{4\varepsilon|F|} \cdot |\mathcal{A}|^{4\varepsilon|F|}$$

in which case it follows that $h(Z_N) < h(\tilde{Y}) + 5\varepsilon + 4\varepsilon \log |\mathcal{A}|$. This and our choice of ε finally give that $h(\pi(Z_N)) - h(\tilde{Y}) < b - a$.

Recall that $h(\tilde{Y}) < a < b < h(Y)$, and recall also that $\pi(Z_0) = Y$ and thus $h(\pi(Z_0)) > b$. We have demonstrated that $h(\pi(Z_n)) - h(\pi(Z_{n+1})) < b - a$ for every $n < N$ and $h(\pi(Z_N)) - h(\tilde{Y}) < b - a$. There must therefore exist at least one $n \leq N$ for which $\pi(Z_n)$ satisfies $a < h(\pi(Z_n)) < b$, thus completing the proof. $\square$ $\square$

3.3.2 Marker patterns

Let Y be a subshift over G. In this section, we construct *marker patterns* for Y. Marker patterns are patterns $m \in \mathcal{P}(Y)$ for which appearances of m in any $y \in Y$ are separated by arbitrarily large displacements. Ideally, distinct appearances of a marker pattern appear on disjoint regions of G, but in practice, there is potentially some small overlap. The concept of marker patterns originates in the seminal work of Krieger and Boyle which considers the $G = \mathbb{Z}$ case [19, 7] and the concept has been utilized by many authors since. For instance, marker patterns were constructed for $G = \mathbb{Z}^d$ by Lightwood [8, Lemma 6.3] in the case that Y is an SFT with the uniform filling property which contains a point with a finite orbit. Here we generalize this construction to the case that G is an arbitrary countable amenable group and Y is a strongly irreducible SFT which separates elements of G and has positive entropy. In particular, we do *not* invoke the comparison property in this construction. This construction, in combination with Theorem 3.3.2, provides the marker patterns for Y which we need in the proof of our main result.

We construct a marker pattern m here by first constructing an aperiodic point $y \in Y$, then taking m to be the pattern of a large-enough shape appearing in y at e.

We construct aperiodic points as follows: begin by passing to a subsystem $Y_1 \subset Y$ such that $h(Y_1) < h(Y)$, in which case one may find arbitrarily many patterns which do not appear in Y_1 but which do appear in Y. By beginning with a point in Y_1 and mixing in one of these "forbidden" patterns from Y (via the strong irreducibility of Y), we obtain a point which lacks all but possibly finitely many periods. Then, we pass to a subsystem $Y_0 \subset Y_1$ (which is also assumed to be strongly irreducible and to separate elements of G) from which an auxiliary pattern can be found which lacks precisely that finite set of periods. Mixing this via Y_0 into our point from before yields the desired aperiodic point.

Theorem 3.3.4. *Let $Y_0 \subset Y_1 \subset Y$ be subshifts and suppose that Y_0 and Y are strongly irreducible, Y_0 separates elements of G, and $h(Y_1) < h(Y)$. For any $r \in \mathbb{N}$, there exists a finite subset $M \subset G$ such that for any fixed $y_0 \in Y_0$, there are patterns $m_1, \ldots, m_r \in \mathcal{P}(M, Y) \setminus \mathcal{P}(M, Y_1)$ and points $y_1, \ldots, y_r \in Y$ satisfying*

1. *m_i appears in y_i only at e,*

2. *m_i does not appear in y_j for any $j \neq i$, and*

3. *$y_i(g) = y_0(g)$ for every $g \notin M$*

for each $i = 1, \ldots, r$.

Proof. Because $h(Y_1) < h(Y)$, there exists a finite subset $F \subset G$ with $e \in F$ such that $|\mathcal{P}(F, Y) \setminus \mathcal{P}(F, Y_1)| > r$. Choose and fix r distinct patterns $a_1, \ldots, a_r \in \mathcal{P}(F, Y) \setminus \mathcal{P}(F, Y_1)$.

Choose a finite subset $K \subset G$ with $e \in K$ to witness the strong irreducibility of Y and Y_0. Let $\{e, g_1, \ldots, g_N\}$ be an enumeration of $F^{-1}KF$. Let $M_0 = KF$ and suppose for induction that $M_{n-1} \subset G$ has been constructed for a fixed $n \in [1, N]$. Because G is infinite, by the pigeonhole principle there exists an element $h_n \in G$ such that $Kh_n \cup Kh_n g_n$ is disjoint from M_{n-1}. Let $M_n = M_{n-1} \sqcup (Kh_n \cup Kh_n g_n)$. We claim that $M = M_N$ is the subset desired for the theorem. Note that

$$M = KF \sqcup \Big(\bigsqcup_{n=1}^{N} Kh_n \cup Kh_n g_n \Big)$$

by induction.

Let $y_0 \in Y_0$ be arbitrary. Let $y_0^{(0)} = y_0$ and suppose for induction that $y_0^{(n-1)} \in Y_0$ has been constructed for a fixed $n \in [1, N]$. Because Y_0 separates elements of G, there exists a point $y^* \in Y_0$ with $y^*(h_n) \neq y^*(h_n g_n)$. Because Y_0 is strongly irreducible of parameter K, we may find a point $y_0^{(n)} \in Y_0$ such that $y_0^{(n)}(\{h_n, h_n g_n\}) = y^*(\{h_n, h_n g_n\})$ and $y_0^{(n)}(g) = y_0^{(n-1)}(g)$ for every $g \notin Kh_n \cup Kh_n g_n$.

By induction, the point $y_0^{(N)} \in Y_0$ satisfies $y_0^{(N)}(h_n) \neq y_0^{(N)}(h_n g_n)$ for each $n = 1, \ldots, N$ and $y_0^{(N)}(g) = y_0(g)$ for every $g \notin \bigcup_{n=1}^{N} Kh_n \cup Kh_n g_n$.

For each $i = 1, \ldots, r$, because Y is strongly irreducible of parameter K, we may find a point $y_i \in Y$ such that $y_i(F) = a_i$ and $y_i(g) = y^{(N)}(g)$ for each $g \notin KF$. Hence by construction, $y_i(h_n) \neq y_i(h_n g_n)$ for each $n = 1, \ldots, N$, and $y_i(g) = y_0(g)$ for each $g \notin M$. Finally, let $m_i = y_i(M)$. We claim that these are the patterns and points desired for the theorem. See Figure 3.1 for an illustration of the construction.

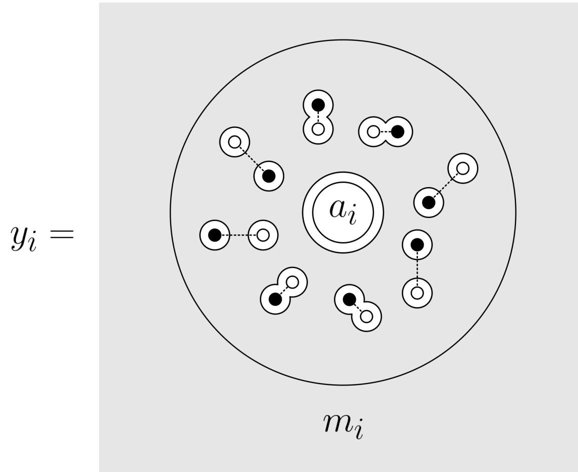


Figure 3.1: An illustration of the construction of the marker pattern m_i. The pattern a_i is selected from Y to be forbidden in Y_1. Thus, y_i has no element of $G \setminus F^{-1}KF$ as a period. Then, small pairs of differing symbols are inductively mixed in (via Y_0) to prevent y_i from having any period $g \in F^{-1}KF \setminus \{e\}$. The shaded exterior is the base point $y_0 \in Y_0$.

For property (1), let $i \leq r$ be fixed. Observe that there is no $g \in F^{-1}KF \setminus \{e\}$ such that $\sigma^g(y_i)(M) = y_i(M)$, as otherwise would contradict the fact that every $g \in F^{-1}KF$ has a corresponding $h \in M \cap Mg^{-1}$ such that $y_i(h) \neq y_i(hg)$. Moreover, if $g \notin F^{-1}KF$ then Fg is disjoint from KF, in which case $\sigma^g(y_i)(F) = \sigma^g(y_0^{(N)})(F)$. We must therefore have $\sigma^g(y_i)(F) \neq a_i$, as otherwise would contradict the fact that a_i is forbidden in Y_1 and hence also forbidden in Y_0. Consequently, $\sigma^g(y_i)(M) = y_i(M)$ only when $g = e$.

For property (2), let $i \neq j$ be fixed. Observe that $a_j \neq a_i$ implies that $y_j(M) \neq y_i(M)$. Moreover, $\sigma^g(y_j)(M) \neq y_i(M)$ for any $g \neq e$ by an identical argument as in the previous paragraph. This is because $y_j(g) = y_0^{(N)}(g) = y_i(g)$ for every $g \notin KF$, and both a_i and a_j are forbidden in Y_1, and hence also forbidden in Y_0.

Property (3) is true by construction. $\square$

One may wonder why we mention the subshift Y_1 at all, instead of merely assuming $h(Y_0) < h(Y)$ directly and stating that the marker patterns are forbidden in Y_0. In fact, later on we shall have some additional structure on Y_1 which shall prove useful for our construction. Namely, in the chain of inclusions $Y_0 \subset Y_1 \subset Y$, we shall have that Y_0 is strongly irreducible, Y_1 is an SFT, and Y is a strongly irreducible SFT. Then, having the marker patterns forbidden in not only Y_0 but also in Y_1 shall be significant.

3.3.3 Main result

We are now ready to present our main result. First we briefly outline the proof to come.

Beginning with a strongly aperiodic subshift X, we derive a chain of factors of X in the form

$$X \xrightarrow{\mathcal{T}} T_0 \xrightarrow{\text{ret}} T_1 \xrightarrow{\text{ex}} T_2$$

where each T_i for $i = 0, 1, 2$ is a system of quasi-tilings of G. The foremost system T_0 is one in which the number of shapes is controlled, delivered by Theorem 3.2.9. The system T_1 (consisting of retractions of points of T_0) is one in which each quasi-tiling exhibits large, disjoint tiles which nearly cover G. The lattermost system T_2 is a system of exact tilings, and we invoke the comparison property to construct it. Our construction of T_2 adapts part of the proof of [24, Theorem 6.3] (proof also appears in [4, Theorem 7.5]). We include the full construction here for the sake of completeness, and because we leverage our control over the foremost system T_0, which is not explicitly referenced in the constructions found in [24, 4].

In principle, we wish to embed all the information of an arbitrary point $x \in X$ into a point $y \in Y$ injectively and continuously (i.e., in such a way that it can be uniquely and "locally" decoded). From x, we derive quasi-tilings t_0, t_1, and t_2 as above. On the side of X, the fact that t_2 is exact (and therefore *covers* G) allows us to partition all of the information in x into local "blocks" of a bounded size. The fact that t_2 covers G is especially significant for our construction; it would be fine if t_2 was not perfectly disjoint, so long as we had full covering and only a small (controllable) amount of "redundancy". Then, by choosing our quasi-tilings with shapes sufficiently large, we may exhibit an injective map from patterns of shapes from T_2 in X to patterns of shapes from T_1 in Y (we describe this as a "block injection"). Here we invoke the hypothesis $h(X) < h(Y)$, as well as the fact that each shape from T_2 is only slightly (controllably) larger in cardinality than a shape from T_1, as careful estimation shows. We thereby construct a point of Y by taking blocks from x, passing them through block injections to obtain blocks of Y, and gluing them together within Y via a mixing condition. On this side, we see that the fact that t_1 is *disjoint* is essential, so that there are no conflicts in laying these blocks together within Y. It is permissible for t_1 to not completely cover G, which is indeed the situation we grapple with in the coming proof.

That leaves open the issue of how to decode the point x if only given y. If one knew which tilings t_1 and t_2 were used to construct y from x, then it would be easy: simply look at the patterns appearing in y on each of the tiles of t_1, pass these backwards through the block injections described above, then lay these new blocks upon the tiles of t_2 and thereby reconstruct x. Here we utilize the quasi-tiling t_0, which has a controlled number of shapes and tile centers spread arbitrarily sparsely throughout the group. This forces the "information density" of t_0 to be arbitrarily low, hence we are able to encode t_0 within y (with the use of marker patterns) by giving up only a subset of symbols of controllably small density.

Given y, one is therefore able to decode the point t_0 by looking at the marker patterns, thereby deriving both t_1 and t_2, thereby decoding x by the algorithm outlined above.

Theorem 3.3.5. *Let G be a countable amenable group with the comparison property. Let X be a nonempty strongly aperiodic subshift over G. Let Y be a strongly irreducible SFT over G with no global period. If $h(X) < h(Y)$ and Y contains at least one factor of X, then X embeds into Y.*

Proof. Let $\mathcal{A}_X$ and $\mathcal{A}_Y$ be finite alphabets such that $X \subset \mathcal{A}_X^G$ and $Y \subset \mathcal{A}_Y^G$. Suppose $\phi : X \to Y$ is a homomorphism, not necessarily injective, in which case $\tilde{Y} = \phi(X) \subset Y$ is a factor of X. Note that $h(\tilde{Y}) \le h(X) < h(Y)$. Without loss of generality (by Lemma 3.3.1), we may assume that Y separates elements of G (Definition 3.3.1). In that case, by Theorem 3.3.2 there exists a strongly irreducible subshift Y_0 which separates elements of G and satisfies $\tilde{Y} \subset Y_0 \subset Y$ and $h(X) < h(Y_0) < h(Y)$. By [9, Theorem 4.2], there exists an SFT Y_1 such that $Y_0 \subset Y_1 \subset Y$ and $h(Y_1) < h(Y)$. It is significant for our proof that these inequalities are strict.

Choose a finite subset $K \subset G$ with $e \in K$ such that $K^{-1} = K$, K witnesses Y and Y_1 as SFTs, and K witnesses Y and Y_0 as strongly irreducible. As in the proof of Theorem 3.3.2, we shall abbreviate $\operatorname{int}^n F = \operatorname{int}_{K^n} F$ and $\partial^n F = \partial_{K^n} F = F \setminus \operatorname{int}^n F$ for each natural $n \in \mathbb{N}$ and finite subset $F \subset G$ for the remainder of this proof. We shall also abbreviate $\partial^n p = p(\partial^n F)$ for each pattern p of shape F.

Choose $\varepsilon > 0$ such that $\varepsilon < 1/3$ and

$$\varepsilon < \frac{h(Y_0) - h(X)}{1 + 5\log|\mathcal{A}_X| + (5 + 4|K|^6)\log|\mathcal{A}_Y|}.$$

Choose $r = 1 + \lceil (2/\varepsilon)\log(1/\varepsilon) \rceil$, in which case $(1 - \varepsilon/2)^r < \varepsilon$. Let $M \subset G$ be the subset delivered by Theorem 3.3.4 for Y_0, Y_1, Y, and r as chosen here. By passing to a superset of M if necessary, we may assume that $K \subset M$ and that $M^{-1} = M$. Choose a finite subset $L \subset G$ such that $M^6 \subset L$ and $|M^6|/|L| < \varepsilon$.

Let $T_0 \subset \Lambda(\mathcal{S}_0)^G$ be the quasi-tiling system delivered by Theorem 3.2.9 for X, ε, r, and L as chosen here, with n_0 chosen so that every $S_0 \in \mathcal{S}_0$ satisfies the following hypotheses.

(H1) $1/|S_0| < \varepsilon(1 - \varepsilon)$.

(H2) $\big(|K^3||M^6||L|\big)|LS_0 \setminus S_0| < \varepsilon|S_0|$.

(H3) $h(S_0, X) < h(X) + \varepsilon$.

Note that Theorem 3.2.9 gives that $|\mathcal{S}_0| \leq r$, so choose and fix an enumeration $(S_0^{(i)})_i$ of $\mathcal{S}_0$ where $i = 1, \ldots, r$. Theorem 3.2.9 also gives that every quasi-tiling $t_0 \in T_0$ is ε-disjoint as witnessed by a continuous and shift-commuting retraction map $t_0 \mapsto \mathrm{ret}(t_0)$. Let $\mathcal{S}_1$ be the set of all shapes realized as tiles of $\mathrm{ret}(t_0)$ for each $t_0 \in T_0$. Each shape $S_0 \in \mathcal{S}_0$ gives rise to a subcollection of shapes $S_1 \in \mathcal{S}_1$, all of which satisfy $S_1 \subset S_0$ and $|S_0 \setminus S_1| < \varepsilon |S_0|$. Let $T_1 = \mathrm{ret}(T_0) \subset \Lambda(\mathcal{S}_1)^G$ be the system of all quasi-tilings obtained by taking retractions of the quasi-tilings in T_0, in which case the map ret is a factor map from T_0 to T_1. Note that every $t_1 \in T_1$ is disjoint by Theorem 3.2.9. Moverover, every $t_1 \in T_1$ is $(1-\varepsilon)$-covering, since the retraction map ret has the property that $\bigcup_g t_0(g)g = \bigcup_g \mathrm{ret}(t_0)(g)g$ for each $t_0 \in T_0$, and every $t_0 \in T_0$ is $(1-\varepsilon)$-covering by Theorem 3.2.9.

We aim to go one step further and construct a factor T_2 of T_1, which shall be a system of exact tilings. It is in this step that we appeal to the comparison property of G. To begin, note that for each pair of shapes $S_0 \in \mathcal{S}_0$ and $S_1 \in \mathcal{S}_1$ with $S_1 \subset S_0$ and $|S_0 \setminus S_1| < \varepsilon |S_0|$, it holds that

$$|S_1| = |S_0| - |S_0 \setminus S_1| > (1-\varepsilon)|S_0| > 1/\varepsilon$$

where above we invoke the hypothesis (H1). This implies that there exists an integer in the interval $[2\varepsilon|S_1|, 3\varepsilon|S_1|)$. Therefore, for each shape $S_1 \in \mathcal{S}_1$, we may find and fix an arbitrary subset $B(S_1) \subset S_1$ such that $2\varepsilon|S_1| \leq |B(S_1)| < 3\varepsilon|S_1|$.

To each $t_1 \in T_1$, we assign two disjoint subsets A_{t_1}, B_{t_1} of G in the following way. Let $A_{t_1} = G \setminus \bigcup_g t_1(g)g$ and let $B_{t_1} = \bigcup_g B(t_1(g))g$. Observe that the assignments $t_1 \mapsto A_{t_1}$, B_{t_1} are continuous and shift-commuting. Observe also that $\overline{D}(A_{t_1}) \leq \varepsilon$ because t_1 is $(1-\varepsilon)$-covering, and that $\underline{D}(B_{t_1}) \geq 2\varepsilon(1-\varepsilon)$ by Lemma 3.2.8 (with $\rho_0 = 1-\varepsilon$ and $\rho_1 = 2\varepsilon$). Consequently,

$$\underline{D}(B_{t_1}) - \overline{D}(A_{t_1}) \geq 2\varepsilon(1-\varepsilon) - \varepsilon > 0$$

where above we have used the fact that $\varepsilon \in (0, 1/2)$.

Therefore, by Theorem 3.2.10, there exists a family of injections $\phi_{t_1} : A_{t_1} \to B_{t_1}$ which is induced by a block code, in the sense that there is a finite subset $F \subset G$ and a function $\Phi : \mathcal{P}(F, T_1) \to F$ such that for every $t_1 \in T_1$ and $g \in A_{t_1}$, it holds that

$$\phi_{t_1}(g) = \Phi(\sigma^g(t_1)(F))g.$$

With this, we are ready to construct T_2. For each $t_1 \in T_1$, let $t_2 = \mathrm{ex}(t_1)$ be the quasi-tiling obtained according to the rule

$$t_2(c) = t_1(c) \sqcup \phi_{t_1}^{-1}(B(t_1(c))c)c^{-1}$$

for every $c \in C(t_1)$, and $t_2(g) = \varnothing$ otherwise. In words, we expand each tile $t_1(c)c$ by including all group elements which map into $B(t_1(c))c \subset B_{t_1}$ under the injective map ϕ_{t_1}. Consequently, $|t_2(c) \setminus t_1(c)| \leq |B(t_1(c))| < 3\varepsilon|t_1(c)|$ for every $c \in C(t_2) = C(t_1)$.

Let $\mathcal{S}_2$ be the set of all shapes realized as tiles of $\mathrm{ex}(t_1)$ for each $t_1 \in T_1$. There are at most finitely many because $\mathrm{ex}(t_1)(g) \subset F^{-1}t_1(g)$ for every $g \in G$. Each shape $S_1 \in \mathcal{S}_1$ gives rise to a subcollection of shapes $S_2 \in \mathcal{S}_2$, all of which satisfy $S_1 \subset S_2$ and $|S_2 \setminus S_1| < 3\varepsilon|S_1|$.

Let $T_2 = \mathrm{ex}(T_1) \subset \Lambda(\mathcal{S}_2)^G$. We claim that the map ex is a homomorphism and that every $t_2 \in T_2$ is an exact tiling of G. To see that the map ex is continuous and shift-commuting, let $t_1 \in T_1$ be arbitrary and let $t_2 = \mathrm{ex}(t_1)$. We note that for each $g \in G$ and $c \in C(t_2) = C(t_1)$, the definition given above implies that $g \in t_2(c)$ if and only if

$$g \in t_1(c) \text{ or } \Phi(\sigma^{gc}(t_1)(F))g \in B(t_1(c)).$$

Moreover, for each $f \in F$ we have $\sigma^{gc}(t_1)(f) = t_1(fgc) = \sigma^c(t_1)(fg)$. Because $t_2(c) \subset F^{-1}t_1(c)$, we see that $t_2(c)$ depends only on $\sigma^c(t_1)(FF^{-1}U_1)$, where $U_1 = \bigcup_{S_1 \in \mathcal{S}_1} S_1$.

This demonstrates that the map ex is induced by a block code, which is sufficient to demonstrate that ex is continuous and shift-commuting.

For the claimed exactness, let $t_2 \in T_2$ and $g \in G$ be arbitrary. Choose $t_1 \in T_1$ such that $t_2 = \mathrm{ex}(t_1)$. If there exists a $c \in C(t_1)$ such that $g \in t_1(c)c$, then the c is necessarily unique by the disjointness of t_1 and also it follows that $g \in t_2(c)c$. Otherwise, $g \in A_{t_1}$, in which case $\phi_{t_1}(g) \in B_{t_1} = \bigcup_c B(t_1(c))c$. Therefore there exists a $c \in C(t_1)$ such that $\phi_{t_1}(g) \in B(t_1(c))c \subset t_1(c)c$. The c must again be unique by the disjointness of t_1. Then $g \in \phi_{t_1}^{-1}(B(t_1(c)c)) \subset t_2(c)c$. We see that for each $g \in G$ there exists a unique $c \in C(t_2)$ such that $g \in t_2(c)c$, hence t_2 is an exact tiling of G.

There is one last quasi-tiling system we shall need. Let $\mathcal{S}_1^*$ denote the collection of all shapes obtained in the form $\mathrm{int}^3(S_1 \setminus M^6 C)$ for any $S_1 \in \mathcal{S}_1$ and any L-separated subset $C \subset G$. Each shape $S_1 \in \mathcal{S}_1$ gives rise to a subcollection of shapes $S_1^* \in \mathcal{S}_1^*$, all of which satisfy $S_1^* \subset S_1$. For each $t_1 \in T_1$, let t_1^* be the retraction of t_1 such that, for each $c \in C(t_1)$, it holds that

$$t_1^*(c) = \mathrm{int}^3(t_1(c) \setminus M^6 C(t_1)c^{-1})$$

and $t_1^*(g) = \varnothing$ otherwise. Observe that each such quasi-tiling belongs to $\Lambda(S_1^*)^G$. Let $T_1^* = \{t_1^* \in \Lambda(\mathcal{S}_1^*)^G : t_1 \in T_1\}$. It is quick to see that the map $t_1 \mapsto t_1^*$ is a factor map from T_1 to T_1^*, because the assignment $t_1 \mapsto C(t_1)$ is continuous and shift-commuting. Moreover, for any choice of $t_0 \in T_0$ with corresponding t_1, t_2, and t_1^*, it holds that $C(t_0) = C(t_1) = C(t_2) = C(t_1^*)$, thus all of these subsets are L-separated.

With our quasi-tiling systems constructed, we now aim to construct an injection which will carry patterns on tiles in X to patterns on tiles in Y_0. On the side of X, we use patterns of those shapes belonging to $\mathcal{S}_2$. On the side of Y_0, we shall use patterns of those shapes belonging to $\mathcal{S}_1^*$. We separate out the construction (and its rather involved estimations) into the following lemma.

Lemma 3.3.6. *Let (S_0, S_1, S_2, S_1^*) be any tuple of shapes from $\mathcal{S}_0$, $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{S}_1^*$ respectively such that $S_1^* \subset S_1 \subset S_0 \cap S_2$, $|S_0 \setminus S_1| < \varepsilon|S_0|$, $|S_2 \setminus S_1| < 3\varepsilon|S_1|$, and $S_1^* = \mathrm{int}^3(S_1 \setminus M^6 C)$ for some L-separated subset $C \subset G$. Then*

$$|\mathcal{P}(S_2, X)| < |\mathcal{P}(S_1^*, Y_0)|.$$

Proof. We begin on the side of X. By the hypotheses, a quick calculation shows that $|S_0 \triangle S_2| < 5\varepsilon|S_0|$. It therefore holds that

$$|\mathcal{P}(S_2, X)| \leq |\mathcal{P}(S_0, X)| \cdot |\mathcal{A}_X|^{|S_0 \triangle S_2|} < e^{(h(X)+\varepsilon)|S_0|} \cdot |\mathcal{A}_X|^{5\varepsilon|S_0|} \tag{X1}$$

where above we invoke the hypothesis (H3) in the latter inequality.

Now we shall estimate the number of patterns in Y_0 of shape S_1^*. We begin by estimating the size of S_1^* in terms of the size of S_0. We proceed in stages, beginning with S_1. We first note that

$$\begin{aligned}
|S_1 \cap M^6 C| &\leq |S_0 \cap M^6 C| \\
&\leq \frac{|M^6|}{|L|}|S_0| + |M^6||\partial_L S_0| + |M^6||(M^6)^{-1} S_0 \setminus S_0| \\
&< \varepsilon|S_0| + \varepsilon|S_0| + \varepsilon|S_0| \\
&= 3\varepsilon|S_0|
\end{aligned}$$

where above we have used the fact that $S_1 \subset S_0$, Lemma 3.2.4, the choice of L, Lemma 3.2.3, and the hypothesis (H2) in conjunction with the fact that $(M^6)^{-1} = M^6 \subset L$. It follows that

$$|S_1 \setminus M^6 C| = |S_1| - |S_1 \cap M^6 C| \geq (1-\varepsilon)|S_0| - 3\varepsilon|S_0| = (1-4\varepsilon)|S_0|.$$

This, together with the fact that $S_1 \setminus M^6C \subset S_1 \subset S_0$, gives us

$$\begin{aligned} |\partial^3(S_1 \setminus M^6C)| &\leq |K^3||K^3(S_1 \setminus M^6C) \setminus (S_1 \setminus M^6C)| \\ &\leq |K^3||K^3S_0 \setminus S_0| + |K^3|^2|S_0 \setminus (S_1 \setminus M^6C)| \\ &< \varepsilon|S_0| + |K|^6 \cdot 4\varepsilon|S_0| \\ &= \varepsilon(1 + 4|K|^6)|S_0| \end{aligned}$$

where above we have used Lemma 3.2.3, Lemma 3.2.1, the hypothesis (H2) in conjuction with the fact that $K^3 \subset L$, and the calculation of the previous display.

The combination of the previous two displays gives us

$$\begin{aligned} |\operatorname{int}^3(S_1 \setminus M^6C)| &= |S_1 \setminus M^6C| - |\partial^3(S_1 \setminus M^6C)| \\ &\geq (1 - 4\varepsilon)|S_0| - \varepsilon(1 + 4|K|^6)|S_0| \\ &= |S_0| - \varepsilon(5 + 4|K|^6)|S_0|. \end{aligned}$$

Recall that $S_1^* = \operatorname{int}^3(S_1 \setminus M^6C)$ and note that $|\mathcal{P}(S_0, Y_0)| \leq |\mathcal{P}(S_1^*, Y_0)| \cdot |\mathcal{A}_Y|^{|S_0 \setminus S_1^*|}$. This, in combination with the previous display, gives that

$$\begin{aligned} |\mathcal{P}(S_1^*, Y_0)| &\geq |\mathcal{P}(S_0, Y_0)| \cdot |\mathcal{A}_Y|^{-|S_0 \setminus S_1^*|} \\ &> e^{h(Y_0)|S_0|} \cdot |\mathcal{A}_Y|^{-\varepsilon(5+4|K|^6)|S_0|} \end{aligned}$$

where above we have also used the fact that $h(S_0, Y_0) \geq h(Y_0)$ (see Definition 3.2.11). Our choice of ε gives us

$$h(X) + \varepsilon + 5\varepsilon \log |\mathcal{A}_X| < h(Y_0) - \varepsilon(5 + 4|K|^6) \log |\mathcal{A}_Y|$$

in which case the previous calculation, in combination with (X1), finally gives that

$$|\mathcal{P}(S_2, X)| < |\mathcal{P}(S_1^*, Y_0)|.$$

$\square$

The previous lemma implies that, for each tuple (S_0, S_1, S_2, S_1^*) satisfying the hypotheses of the previous lemma, there exists an injective map

$$\Psi(S_2, S_1^*) : \mathcal{P}(S_2, X) \to \mathcal{P}(S_1^*, Y_0).$$

Let these injections be chosen and fixed. We now continue the proof of Theorem 3.3.5.

Here we choose our marker patterns. Pick a point $y^{(m)} \in Y_0$ arbitrarily to serve as the "substrate" of the marker patterns. Let $y_1, \ldots, y_r \in Y$ and $m_i = y_i(M) \in \mathcal{P}(M, Y) \setminus \mathcal{P}(M, Y_1)$ be the points and patterns delivered by Theorem 3.3.4 for the choice of substrate $y^{(m)}$.

Now we appeal to the strong irreducibility of Y_0 in order to construct certain "mixing" patterns. We do this twice: once for patterns of shape M^6, and once for patterns of each shape $S_1^* \in \mathcal{S}_1^*$.

Note that $K^2 \cdot KM^3 \subset M^6$ because $K \subset M$, thus $KM^3 \subset \mathrm{int}^2(M^6)$ and thus KM^3 is disjoint from $\partial^2 M^6$. The subshift Y_0 is strongly irreducible as witnessed by K, therefore for any pair of patterns $u \in \mathcal{P}(M^3, Y_0)$ and $\partial^2 v \in \mathcal{P}(\partial^2 M^6, Y_0)$, there is a pattern $w \in \mathcal{P}(M^6, Y_0)$ such that $w(g) = u(g)$ for each $g \in M^3$ and $w(g) = \partial^2 v(g)$ for each $g \in \partial^2 M^6$. Choose and fix one such pattern w for each choice of u and $\partial^2 v$; we shall denote it by $w = u \cup \partial^2 v$.

Let $S_1^* \in \mathcal{S}_1^*$ be arbitrary and suppose $S_1^* = \mathrm{int}^3(S_1 \setminus M^6 C)$ for some shape $S_1 \in \mathcal{S}_1$ and some L-separated subset $C \subset G$. Observe that $K^2 \cdot KS_1^* \subset S_1 \setminus M^6 C$, thus KS_1^* is disjoint from $\partial^2(S_1 \setminus M^6 C)$. Therefore, for each pair of patterns $u \in \mathcal{P}(S_1^*, Y_0)$ and $\partial^2 v \in \mathcal{P}(\partial^2(S_1 \setminus M^6 C), Y_0)$, there is a pattern $w \in \mathcal{P}(S_1 \setminus M^6 C, Y_0)$ such that

$w(g) = u(g)$ for each $g \in S_1^*$ and $w(g) = \partial^2 v(g)$ for each $g \in \partial^2(S_1 \setminus M^6 C)$. Choose and fix one such pattern w for each choice of u and $\partial^2 v$; we shall again denote it by $w = u \cup \partial^2 v$.

We are now ready to construct the map $\psi : X \to Y$ desired for the theorem. To begin, let $x \in X$ be fixed, let $t_0 = \mathcal{T}(x) \in T_0$, let $t_1 = \mathrm{ret}(t_0) \in T_1$, let $t_2 = \mathrm{ex}(t_1) \in T_2$, let $t_1^* \in T_1^*$ be delivered by t_1, and let $y_0 = \phi(x) \in Y_0$ be the point delivered by the homomorphism $\phi : X \to Y$ assumed to exist at the beginning of the proof. Let $C = C(t_0) = C(t_1) = C(t_2) = C(t_1^*) \subset G$.

We construct the point $y = \psi(x) \in Y$ in two stages, first by constructing a point $y_1 \in Y_1$ which locally looks like a point of Y_0, and then modifying y_1 into a point of Y by placing down the marker patterns.

Note that for each $c \in C$, the tuple of shapes $(t_0(c), t_1(c), t_2(c), t_1^*(c))$ satisfies the hypotheses of Lemma 3.3.6 by construction. Let y_1 be the point satisfying the following two conditions for every $c \in C$ when $(S_1, S_2, S_1^*) = (t_1(c), t_2(c), t_1^*(c))$.

$$\sigma^c(y_1)(M^6) = y^{(m)}(M^3) \cup \sigma^c(y_0)(\partial^2 M^6) \tag{C1}$$

$$\sigma^c(y_1)(S_1 \setminus M^6 C c^{-1}) = \Psi(S_2, S_1^*)(\sigma^c(x)(S_2)) \cup \sigma^c(y_0)(\partial^2(S_1 \setminus M^6 C)) \tag{C2}$$

Everywhere else, let $y_1(g) = y_0(g)$. Here we utilize the block injection(s) $\Psi(S_2, S_1^*)$ delivered by Lemma 3.3.6, as well as the mixing boundary patterns $w = u \cup \partial^2 v$ constructed earlier. The construction of y_1 is illustrated in Figure 3.2.

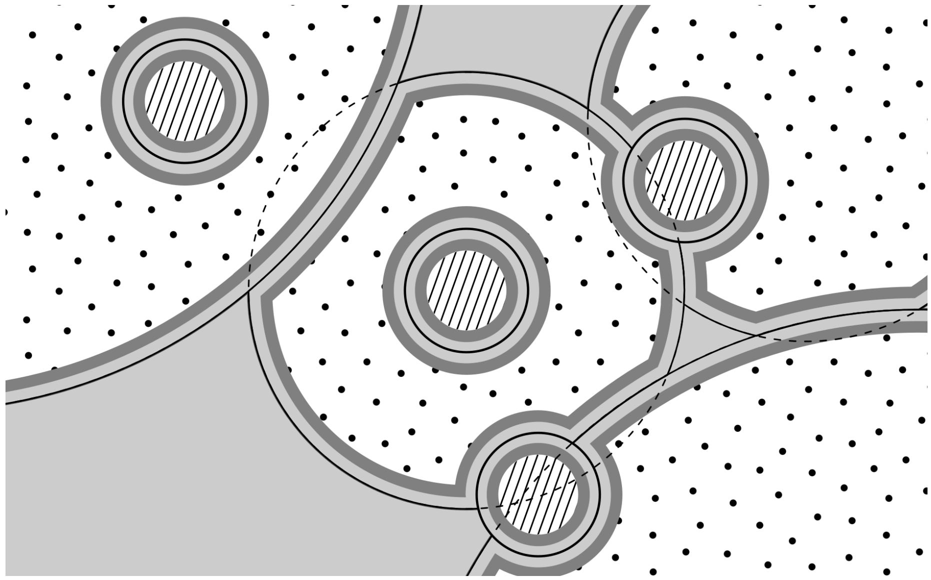


Figure 3.2: An illustration of the construction of y_1 for a hypothetical tiling of $\mathbb{Z}^2$ using different sized circles. The largest solid circles indicate the tiles of t_0 (the overlap, indicated by dashed lines, is removed in t_1). The smaller solid circles indicate the translates of M^6, wherein the marker patterns will later be placed. The stippled tile interiors are the patterns given by the block injections from Lemma 3.3.6. The darkened boundaries are delivered by the strong irreducibility of Y_0, to mix each tile interior pattern with its respective boundary pattern from y_0. The small hatched circles are each labeled with the pattern drawn from the marker substrate, $y^{(m)}(M^3)$. The shaded exterior is the base point y_0.

First we argue that y_1 is well-defined. Let $g \in G$ be fixed; we split over three cases. If $g \in M^6C$, then there is a unique $c \in C$ such that $g \in M^6c$ because C is L-separated and $M^6 \subset L$. In this case, $y_1(g)$ is determined by condition (C1). If $g \in \left(\bigcup_c t_1(c)c\right) \setminus M^6C$, then again there is a unique $c \in C$ such that $g \in t_1(c)c \setminus M^6C = (t_1(c) \setminus M^6Cc^{-1})c$ by the disjointness of t_1. In this case, $y_1(g)$ is determined by condition (C2). Otherwise, $y_1(g) = y_0(g)$ is again uniquely determined.

Next we argue that y_1 belongs to Y_1. Recall that Y_1 is an SFT as witnessed by K. Let $g \in G$ be arbitrary and consider the translate Kg. If Kg intersects $\operatorname{int}^2(M^6)c$ for some $c \in C$, then $Kg \subset M^6c$ by Lemma 3.2.2 and the fact that $K = K^{-1}$. Moreover, the c must be unique. In this case, $\sigma^g(y_1)(K)$ is given by condition (C1) and therefore $\sigma^g(y_1)(K) \in \mathcal{P}(K, Y_0)$. If Kg intersects $\operatorname{int}^2(t_1(c)c \setminus M^6C)$ for some

$c \in C$, then $Kg \subset t_1(c)c \setminus M^6C$ also by Lemma 3.2.2, in which case the c must again be unique. In this case, $\sigma^g(y_1)(K)$ is given by condition (C2) and therefore $\sigma^g(y_1)(K) \in \mathcal{P}(K, Y_0)$ again. If neither of these cases hold, then

$$Kg \subset (G \setminus M^6C) \cup \partial^2(M^6)C \cup \Big(G \setminus \bigcup_c (t_1(c)c \setminus M^6C)\Big) \cup \bigcup_c \partial^2(t_1(c)c \setminus M^6C).$$

In this case, $\sigma^g(y_1)(K) = \sigma^g(y_0)(K)$ and therefore $\sigma^g(y_1)(K) \in \mathcal{P}(K, Y_0)$ again. We see that every pattern of shape K appearing in y_1 is allowed in Y_0. Since $Y_0 \subset Y_1$ and Y_1 is an SFT witnessed by K, we conclude that $y_1 \in Y_1$.

Next we argue that the map $x \mapsto y_1$ is continuous and shift-commuting. For a fixed $g \in G$, in order to determine the symbol $y_1(g)$, one must know first the tiles from t_0, t_1, t_2, and t_1^* to which g belongs. This requires looking only at the symbols of the involved quasi-tilings within a finite neighborhood of g. Then, one either applies condition (C1) or condition (C2) or returns $y_0(g)$. As every involved quasi-tiling and y_0 are derived from x in a continuous and shift-commuting manner, it is evident that $y_1(g)$ depends only on $\sigma^g(x)(F_1)$, where F_1 is some (possibly very large but) finite subset of G. We conclude that the map $x \mapsto y_1$ is continuous and shift-commuting (but possibly non-injective).

Finally, we construct $y = \psi(x) \in Y$ from y_1 as follows. For every $c \in C$, let y satisfy $\sigma^c(y)(M) = m_i$, where $i \in [1, r]$ is the index of the shape $t_0(c) = S_0^{(i)} \in \mathcal{S}_0$ which was fixed at the beginning of the proof, and m_i is the corresponding marker pattern. Everywhere else, let $y(g) = y_1(g)$.

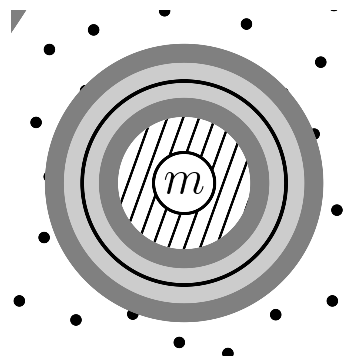


Figure 3.3: An illustration of the construction of y near a tile center. The marker pattern m_i corresponding to the shape $S_0^{(i)}$ is placed directly over the marker substrate, the pattern $y^{(m)}(M^3)$.

The point y is well-defined by identical argument as (and as a consequence of) the fact that y_1 is well-defined. Before proceeding, we note a property of y which is critical to later arguments. Let $c \in C$ be arbitrary and suppose that $t_0(c) = S_0^{(i)}$ for some unique index $i \in [1, r]$. Recall $y_i \in Y$ is the point from which the marker pattern m_i is drawn. Then

$$\sigma^c(y)(M^3) = y_i(M^3). \tag{Y1}$$

This is true because $\sigma^c(y)(M) = m_i = y_i(M)$ by construction, together with the fact that

$$\sigma^c(y)(M^3 \setminus M) = \sigma^c(y_1)(M^3 \setminus M) = y^{(m)}(M^3 \setminus M) = y_i(M^3 \setminus M)$$

where above we have used the construction of y, condition (C2), and Theorem 3.3.4.

Next we argue that y belongs to Y. Recall that Y is an SFT witnessed by K. Let $g \in G$ be arbitrary and consider the translate Kg. If Kg intersects Mc for some $c \in C$, then $Kg \subset M^3c$ by the fact that $K = K^{-1}$ and $K \subset M$. Moreover, the c must be unique. Then, if $t_0(c) = S_0^{(i)}$ for some unique index $i \in [1, r]$, property (Y1) implies that $\sigma^g(y)(K) = \sigma^{gc^{-1}}(y_i)(K)$. As $y_i \in Y$, we therefore see that $\sigma^g(y)(K) \in \mathcal{P}(K, Y)$. In the opposite case, $Kg \subset G \setminus MC$, in which case $\sigma^g(y)(K) = \sigma^g(y_1)(K)$ by construction. Because $y_1 \in Y_1 \subset Y$, we again see that $\sigma^g(y)(K) \in \mathcal{P}(K, Y)$. Therefore, every pattern of shape K appearing in y is allowed in Y, thus $y \in Y$.

Next we argue that the map ψ is continuous and shift-commuting. This follows by identical argument as (and as a consequence of) the fact that the map $x \mapsto y_1$ is continuous and shift-commuting.

Before proving that the map ψ is injective, we note one more property in advance. Let $x \in X$ be fixed, let $t_0 = \mathcal{T}(x) \in T_0$, let $C = C(t_0) \subset G$ and let $y = \psi(x) \in Y$. We claim that for each $g \in G$ and each index $i \in [1, r]$, we have

$$\sigma^g(y)(M) = m_i \text{ if and only if } t_0(g) = S_0^{(i)} \in \mathcal{S}_0. \tag{Y2}$$

The reverse implication is obvious by construction. For the forward implication, let $g \in G$ be arbitrary and suppose $\sigma^g(y)(M) = m_i$ for some index $i \in [1, r]$. If Mg is disjoint from MC, then $\sigma^g(y)(M) = \sigma^g(y_1)(M)$, in which case $\sigma^g(y)(M) = m_i$ contradicts the fact that m_i is forbidden in Y_1. Therefore Mg must intersect MC, in which case $Mg \subset M^3c$ for some $c \in C$ by the fact that $M = M^{-1}$. Moreover, the c must be unique; indeed, if $Mg \subset M^3c_1$ and $Mg \subset M^3c_2$ for distinct $c_1 \neq c_2$, then $g \in M^4c_1 \cap M^4c_2 \neq \varnothing$, contradicting the fact that C is L-separated and $M^4 \subset L$. Then suppose $t_0(c) = S_0^{(j)}$ for some index $j \in [1, r]$. By (Y1) we have that $\sigma^c(y)(M^3) = y_j(M^3)$. Moreover, the fact that $\sigma^g(y)(M) = m_i$ implies that $m_i = \sigma^{gc^{-1}}(y_j)(M)$. By Theorem 3.3.4, this only happens in the case where $j = i$ and $g = c$. The claim follows.

Finally, we argue that ψ is injective. Let x_1, $x_2 \in X$ be arbitrary and suppose that $\psi(x_1) = \psi(x_2)$. Property (Y2) implies that $\mathcal{T}(x_1) = \mathcal{T}(x_2) = t_0 \in T_0$. Then, let t_1, t_1^*, and t_2 be derived from t_0 as before and let $C = C(t_0) \subset G$. The condition (C2) and the fact that Ψ is an injective map implies that $\sigma^c(x_1)(t_2(c)) = \sigma^c(x_2)(t_2(c))$ for every $c \in C$. Because t_2 is an exact tiling of G, it follows that $x_1 = x_2$. □ □

3.4 Discussion

In Theorem 3.3.5, can the assumption that Y contains a factor of X be dropped? That is, under what conditions on X and Y does there necessarily exist a homomorphism $\phi : X \to Y$? This is true if for example Y contains a fixed point, because every subshift factors onto a fixed point. A homomorphism was constructed by Lightwood [23, Theorem 2.8] for $G = \mathbb{Z}^2$ in the case that X is strongly aperiodic and Y is an SFT which satisfies the “square-filling mixing" condition, therefore providing an extension of Krieger’s embedding theorem to $\mathbb{Z}^2$. But, the existence of such homomorphisms in general remains open, even for $G = \mathbb{Z}^d$ where $d \geq 3$.

Can the mixing condition on Y be weakened? One might wish to replace the “uniform" mixing condition of strong irreducibility with a non-uniform mixing condition, such as topological mixing. However, such conditions are too weak for the construction given here, as we rely on the fact that we can cut and paste patterns on tiles in Y that are “packed together" relatively tightly, with the tiling known to cover a fraction of the group which can be chosen arbitrarily close to 1. Furthermore, it is known that topological mixing is too weak for a general embedding theorem (i.e., for a general X) even for $G = \mathbb{Z}^2$; indeed, Quas and Şahin [33, Theorem 1.1] have constructed an example of a topologically mixing SFT Y and a nonnegative constant $h_0 < h(Y)$ such that if X is any subshift with the uniform filling property (a mixing condition stronger than topological mixing) which satisfies $h_0 \leq h(X) \leq h(Y)$, then X canot be embedded in Y.

Can the assumption that G has the comparison property be dropped? If there are no amenable groups without the comparison property, then this point is moot. We invoke the comparison property in two places in this proof, in each case to construct a desirable system of quasi-tilings. In the first case, as part of the construction of the subshift Y_0, we construct a strongly irreducible system of exact tilings of G by way of Theorem 3.2.11 (due to Downarowicz and Zhang [4, 24]) in combination with a construction of Frisch and Tamuz [32]. In the second case, we adapt the proof of Theorem 3.2.11 to construct a system T_2 of exact tilings as a factor of a suitable system of disjoint quasi-tilings T_1. As mentioned before, the most significant aspect of T_2 is that its tilings completely cover G, and disjointness here could in principle be traded for near-disjointness.

If one refuses the comparison property and utilizes instead either T_0 or T_1 on the side of X to construct the map $\psi : X \to Y$, then ψ has the property that for every x_1, $x_2 \in X$, if $\psi(x_1) = \psi(x_2)$ then $\mathcal{T}(x_1) = \mathcal{T}(x_2) = t_0 \in T_0$ and $x_1(g) = x_2(g)$ for every $g \in \bigcup_c t_0(c)c$. This map is therefore not necessarily injective (unless t_0 covers G), but one does have that the difference in entropy between X and $\psi(X) \subset Y$ can be made arbitrarily small (based on the covering density of t_0). Therefore in this case, if Y contains at least one factor of X then necessarily Y contains many "nontrivial" factors of X, in particular with entropy arbitrarily close to X.

One aspect of the quasi-tilings given by Theorem 3.2.9 that we have not exploited is that they are actually *maximal*, in the sense that no one additional tile of any shape could be inserted anywhere without breaking the ε-disjointness property (this is seen if one closely inspects the proof of [2, Lemma 3.4]). This condition is similar to that of the ρ-covering condition (Definition 3.2.16); while the ρ-covering is only in general witnessed at some "scale" F (depending on the quasi-tiling, and possibly much larger than the tiles themselves), the maximality implies that the covering is somehow witnessed on the scale of the tiles. This leverage could be useful for some applications.

REFERENCES

[1] D. Ornstein and B. Weiss, “Entropy and isomorphism theorems for actions of amenable groups,” *J. Analyse Math.*, vol. 48, pp. 1–141, 1987.

[2] T. Downarowicz and D. Huczek, “Dynamical quasitilings of amenable groups,” *Bull. Pol. Acad. Sci. Math.*, vol. 66, no. 1, pp. 45–55, 2018.

[3] T. Downarowicz, D. Huczek, and G. Zhang, “Tilings of amenable groups,” *J. Reine Angew. Math.*, vol. 747, pp. 277–298, 2019.

[4] T. Downarowicz and G. Zhang, “Symbolic extensions of amenable group actions and the comparison property,” *Preprint*, 2019. https://arxiv.org/abs/2106.10687.

[5] A. Desai, “Subsystem entropy for $\mathbb{Z}^d$ sofic shifts,” *Indag. Math. (N.S.)*, vol. 17, no. 3, pp. 353–359, 2006.

[6] D. Kerr and P. Naryshkin, “Elementary amenability and almost finiteness,” 2022.

[7] W. Krieger, “On the subsystems of topological Markov chains,” *Ergodic Theory Dynam. Systems*, vol. 2, no. 2, pp. 195–202 (1983), 1982.

[8] S. Lightwood, “Morphisms from non-periodic $\mathbb{Z}^2$-subshifts. I. Constructing embeddings from homomorphisms,” *Ergodic Theory Dynam. Systems*, vol. 23, no. 2, pp. 587–609, 2003.

[9] R. Bland, K. McGoff, and R. Pavlov, “Subsystem entropies of shifts of finite type and sofic shifts on countable amenable groups,” *Ergodic Theory and Dynamical Systems*, vol. 43, no. 9, pp. 2881–2914, 2023.

[10] R. Bland, “An embedding theorem for subshifts over amenable groups with the comparison property,” *Ergodic Theory and Dynamical Systems*, pp. 1–31, 2024.

[11] D. Lind, “The entropies of topological Markov shifts and a related class of algebraic integers,” *Ergodic Theory Dynam. Systems*, vol. 4, no. 2, pp. 283–300, 1984.

[12] M. Boyle, R. Pavlov, and M. Schraudner, “Multidimensional sofic shifts without separation and their factors,” *Trans. Amer. Math. Soc.*, vol. 362, no. 9, pp. 4617–4653, 2010.

[13] M. Hochman and T. Meyerovitch, “A characterization of the entropies of multidimensional shifts of finite type,” *Ann. of Math. (2)*, vol. 171, no. 3, pp. 2011–2038, 2010.

[14] A. Quas and P. Trow, “Subshifts of multi-dimensional shifts of finite type,” *Ergodic Theory Dynam. Systems*, vol. 20, no. 3, pp. 859–874, 2000.

[15] S. Barbieri, "On the entropies of subshifts of finite type on countable amenable groups," *Groups Geom. Dyn.*, vol. 15, no. 2, pp. 607–638, 2021.

[16] J. Frisch and O. Tamuz, "Symbolic dynamics on amenable groups: the entropy of generic shifts," *Ergodic Theory Dynam. Systems*, vol. 37, no. 4, pp. 1187–1210, 2017.

[17] S. Barbieri and M. Sablik, "A generalization of the simulation theorem for semidirect products," *Ergodic Theory Dynam. Systems*, vol. 39, no. 12, pp. 3185–3206, 2019.

[18] D. Huczek and S. Kopacz, "Factoring strongly irreducible group shift actions onto full shifts of lower entropy," *Preprint*, 2021. https://arxiv.org/abs/2106.10687.

[19] M. Boyle, "Lower entropy factors of sofic systems," *Ergodic theory and dynamical systems*, vol. 3, no. 4, pp. 541–557, 1983.

[20] D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding.* Cambridge University Press, Cambridge, 1995.

[21] D. Kerr and H. Li, *Ergodic theory.* Springer Monographs in Mathematics, Springer, Cham, 2016. Independence and dichotomies.

[22] N. Aubrun and M. Sablik, "Simulation of effective subshifts by two-dimensional subshifts of finite type," *Acta Appl. Math.*, vol. 126, pp. 35–63, 2013.

[23] S. Lightwood, "Morphisms from non-periodic $\mathbb{Z}^2$ subshifts. II. Constructing homomorphisms to square-filling mixing shifts of finite type," *Ergodic Theory Dynam. Systems*, vol. 24, no. 4, pp. 1227–1260, 2004.

[24] T. Downarowicz and G. Zhang, "The comparison property of amenable groups," *Preprint*, 2017. https://arxiv.org/abs/2106.10687.

[25] N. Aubrun, S. Barbieri, and S. Thomassé, "Realization of aperiodic subshifts and uniform densities in groups," *Groups, geometry and dynamics*, vol. 13, no. 1, pp. 107–129, 2019.

[26] D. Kerr and H. Li, *Ergodic theory.* Springer Monographs in Mathematics, Springer, Cham, 2016. Independence and dichotomies.

[27] I. Namioka, "Følner's conditions for amenable semi-groups," *Math. Scand.*, vol. 15, pp. 18–28, 1964.

[28] D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding.* Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2021. Second edition [of 1369092].

[29] T. Downarowicz, B. Frej, and P. P. Romagnoli, "Shearer's inequality and infimum rule for shannon entropy and topological entropy," *Dynamics and numbers*, pp. 63–75, 2016.

[30] T. Ceccherini-Silberstein and M. Coornaert, *Cellular automata and groups.* Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2010.

[31] D. Ornstein and B. Weiss, “Entropy and isomorphism theorems for actions of amenable groups,” *J. Analyse Math.*, vol. 48, pp. 1–141, 1987.

[32] J. Frisch and O. Tamuz, “Symbolic dynamics on amenable groups: the entropy of generic shifts,” *Ergodic Theory Dynam. Systems*, vol. 37, no. 4, pp. 1187–1210, 2017.

[33] A. Quas and A. Şahin, “Entropy gaps and locally maximal entropy in $\mathbb{Z}^d$ subshifts,” *Ergodic Theory Dynam. Systems*, vol. 23, no. 4, pp. 1227–1245, 2003.