EXTENDER SETS AND MEASURES OF MAXIMAL ENTROPY FOR SUBSHIFTS

FELIPE GARCÍA-RAMOS AND RONNIE PAVLOV

Abstract. For countable amenable G, we prove inequalities relating $\mu(v)$ and $\mu(w)$ for any measure of maximal entropy μ on a G-subshift and any pair of words v, w where the extender set of v is contained in the extender set of w. These results generalize the main result of [16]. When $G = \mathbb{Z}$, we prove a stronger result, and present several applications to the classes of synchronizing and hereditary subshifts.

1. Introduction

In this work, we study what the relationship between the so-called extender sets ([7], [18]) of two words on a subshift can tell us about the relationship of the measures of maximal entropy of the words.

Subshifts are symbolically defined dynamical systems in which points are given by elements of $\mathcal{A}^{\mathbb{G}}$ for some finite alphabet $\mathcal{A}$ and a countable amenable group $\mathbb{G}$, and the dynamics are given by the $\mathbb{G}$-action of translation/shift maps $\{\sigma_g\}_{g\in\mathbb{G}}$.

Though we postpone a formal definition of extender sets to Section 2, extender sets are defined for words (which are just finite configurations of letters from $\mathcal{A}$), and the extender set $E_X(v)$ of a word with shape F $v \in \mathcal{A}^F$ in a subshift X is just the set of configurations in $\mathcal{A}^{F^c}$ which can be combined with v to create a legal point of X. Our results treat pairs of words v, w where $E_X(v) \subseteq E_X(w)$. Informally, this means that in any point of X, if one replaces an occurrence of v by w, then the resulting point is still in X. The starting point of our work is the following result of Meyerovitch.

Theorem 1.1 (Theorem 3.1, [16])**.** *If X is a $\mathbb{Z}^d$-subshift and $v, w \in \mathcal{A}^F$ satisfy $E_X(v) = E_X(w)$, then for every measure of maximal entropy μ on X, $\mu(v) = \mu(w)$.*

Remark 1.2. In fact the theorem from [16] is more general; it treats equilibrium states for a class of potentials ϕ with a property called d-summable variation, and the statement here for measures of maximal entropy corresponds to the $\phi = 0$ case only. Also, in [16], this result is presented as a result about conditional measures; however, it is not hard to show that result is equivalent to the version presented here.

One of our two main results strengthens Theorem 1.1 by requiring the weaker hypothesis of $E_X(v) \subseteq E_X(w)$. In the class of so-called hereditary subshifts, this property holds for many pairs of words even though $E_X(v) = E_X(w)$ may rarely

2010 *Mathematics Subject Classification.* 37B10, 37B40, 37D35.

Key words and phrases. Symbolic dynamics, measure of maximal entropy, extender set, synchronized subshift.

The second author gratefully acknowledges the support of NSF grant DMS-1500685.

or never hold; see Section 3.1.2 for details. We have the following result for finitely generated $\mathbb{G}$.

Theorem 4.4. *Let X be a $\mathbb{G}-$subshift, μ a measure of maximal entropy of X, $F \Subset \mathbb{G}$, and $w, v \in \mathcal{A}^F$. If $E(v) \subseteq E(w)$ then*

$$\mu(v) \leq \mu(w).$$

In general, it is impossible to compare extender sets for words v, w in the language $L(X)$ with different shapes $F \neq F'$, since $E_X(v)$ consists of configurations on F^c and $E_X(w)$ consists of configurations on F'^c. However, in the specific case where $\mathbb{G} = \mathbb{Z}$ and F and F' are contiguous intervals of integers, there is a natural bijection between F^c and F'^c (see Section 2), which allows one to treat containments. Our second main result treats this more general setting (here, $|u|$ represents the length of a word u.)

Theorem 3.8. *Let X be a $\mathbb{Z}$-subshift, μ a measure of maximal entropy of X, and $w, v \in L(X)$. If $E_X(v) \subseteq E_X(w)$, then*

$$\mu(v) \leq \mu(w) e^{h_{top}(X)(|w|-|v|)}.$$

This result is less intuitive than the previous ones, but in some sense can be thought of as saying that the amount of information required to "make up" for the difference in $|v|$ and $|w|$ is $e^{h_{top}(X)}$ per letter. The following corollary, which can be thought of as a different generalization of Theorem 1.1, is immediate.

Corollary 3.9. *Let X be a $\mathbb{Z}$-subshift, μ a measure of maximal entropy of X, and $w, v \in L(X)$. If $E_X(v) = E_X(w)$, then for every measure of maximal entropy of X,*

$$\mu(v) = \mu(w) e^{h_{top}(X)(|w|-|v|)}.$$

In the class of synchronized subshifts (see Section 3.1 for the definition), $E_X(v) = E_X(w)$ holds for many pairs of words of different lengths, in which case Corollary 3.9 gives significant restrictions on μ, including a new proof of uniqueness under the hypothesis of entropy minimality (see Theorem 3.13). We give several applications of Corollary 3.9 to such subshifts in Section 3.1.1. For the subclass of S-gap subshifts, our results yield quite a lot of information; for instance we prove the following, which verifies a conjecture of Climenhaga ([2]).

Corollary 3.17. *Let X_S be an $S-$gap subshift with $\gcd(S+1) = 1$. Then for μ, the unique MME on X_S,*

$$\lim_{n\to\infty} \frac{|L_n(X_S)|}{e^{n h_{top}(X_S)}} = \frac{\mu(1) e^{h_{top}(X_S)}}{(e^{h_{top}(X_S)} - 1)^2}.$$

In fact we show that the existence of the limit holds for a much more general class of subshifts (synchronized subshifts where the unique measure of maximal entropy is mixing).

Section 2 contains definitions and results needed throughout our proofs, Section 3 contains our results for $\mathbb{Z}$-subshifts (including various applications in Section 3.1), and Section 4 contains our results for $\mathbb{G}$-subshifts.

2. General definitions and preliminaries

We will use $\mathbb{G}$ to refer to a countable discrete group. We write $F \Subset \mathbb{G}$ to mean that F is a finite subset of $\mathbb{G}$, and unless otherwise stated, F always refers to such an object.

A sequence $\{F_n\}_{n\in\mathbb{N}}$ with $F_n \Subset \mathbb{G}$ is said to be **Følner** if for every $K \Subset \mathbb{G}$, we have that $|(K \cdot F_n)\Delta F_n|/|F_n| \to 0$. We say that $\mathbb{G}$ is **amenable** if it admits a Følner sequence. In particular, $\mathbb{Z}$ is an amenable group, since any sequence $\{F_n\} = [a_n, b_n] \cap \mathbb{Z}$ with $b_n - a_n \to \infty$ is Følner.

Let $\mathcal{A}$ be any finite set (usually known as the alphabet). We call $\mathcal{A}^{\mathbb{G}}$ the **full $\mathcal{A}$-shift on** $\mathbb{G}$, and endow it with the product topology (using the discrete topology on $\mathcal{A}$). For $x \in \mathcal{A}^{\mathbb{G}}$, we use x_i to represent the ith coordinate of x, and x_F to represent the restriction of x to any $F \Subset \mathbb{G}$.

For any $g \in \mathbb{G}$, we use σ_g to denote the left translation by g on $\mathcal{A}^{\mathbb{G}}$, also called the **shift by** g; note that each σ_g is an automorphism. We say $X \subseteq \mathcal{A}^{\mathbb{G}}$ is a $\mathbb{G}$-**subshift** if it is closed and $\sigma_g(X) = X$ for all $g \in \mathbb{G}$; when $\mathbb{G} = \mathbb{Z}$ we simply call it a subshift.

For $F \Subset \mathbb{G}$, we call an element of $\mathcal{A}^F$ a **word with shape** F. For w a word with shape F and x either a point of $\mathcal{A}^{\mathbb{G}}$ or a word with shape $F' \supset F$, we say that w is a **subword of** x if $x_{g+F} = w$ for some $g \in \mathbb{G}$.

For any F, the F-**language of** X is the set $L_F(X) \subseteq \mathcal{A}^F = \{x_F \ : \ x \in X\}$ of words with shape F that appear as subwords of points of X. When $\mathbb{G} = \mathbb{Z}$, we use $L_n(X)$ to refer to $L_{\{0,\ldots,n-1\}}(X)$ for $n \in \mathbb{N}$. We define

$$L(X) := \bigcup_{F \Subset \mathbb{G}} L_F(X) \text{ if } \mathbb{G} \neq \mathbb{Z} \text{ and}$$
$$L(X) := \bigcup_{n\in\mathbb{N}} L_n(X) \text{ if } \mathbb{G} = \mathbb{Z}.$$

For any $\mathbb{G}$-subshift X and $w \in L_F(X)$, we define the **cylinder set of** w as

$$[w] := \{x \in X : x_F = w\}.$$

Whenever we refer to an interval in $\mathbb{Z}$, it means the intersection of that interval with $\mathbb{Z}$. So, for instance, if $x \in \mathcal{A}^{\mathbb{Z}}$ and $i < j$, $x_{[i,j]}$ represents the subword of x that starts in position i and ends in position j. Unless otherwise stated, a word $w \in \mathcal{A}^n$ is taken to have shape $[0, n)$. Every word $w \in L(\mathcal{A}^{\mathbb{Z}})$ is in some $\mathcal{A}^n$ by definition; we refer to this n as the **length** of w and denote it by $|w|$.

For any amenable $\mathbb{G}$ with Følner sequence $\{F_n\}_{n\in\mathbb{N}}$ and any $\mathbb{G}$-subshift X, we define the **topological entropy of** X as

$$h_{top}(X) = \lim_{n\to\infty} \frac{1}{|F_n|} \log |L_{F_n}(X)|$$

(this definition is in fact independent of the Følner sequence used.)

For any $w \in L(X)$, we define the **extender set of** w as

$$E_X(w) := \{x|_{F^c} \ : \ x \in [w]\}.$$

Example 2.1. For any $G \neq \mathbb{Z}$, if X is the full shift on two symbols, $\{0,1\}^G$, then for any F, all words in $\{0,1\}^F$ have the same extender set, namely $\{0,1\}^{F^c}$.

Example 2.2. Take $G = \mathbb{Z}^2$ and X the hard-square shift on $\{0,1\}$ in which adjacent 1s are forbidden horizontally and vertically. Then if we take $F = \{(0,0)\}$, we see that $E(0)$ is the set of all configurations on $\mathbb{Z}^2 \setminus F$ which are legal, i.e. which contain no adjacent 1s. Similarly, $E(1)$ is the set of all legal configurations on $\mathbb{Z}^2 \setminus F$ which also contain 0s at $(0, \pm 1)$ and $(\pm 1, 0)$. In particular, we note that here $E(1) \subsetneq E(0)$.

In the specific case $\mathbb{G} = \mathbb{Z}$ and $w \in L_n(X)$, we may identify $E_X(w)$ with the set of sequences which are concatenations of the left side and the right side, i.e. $\{(x_{(-\infty,0)} x_{[n,\infty)}) \ : \ x \in [w]\}$, and in this way can relate extender sets even for v, w with different lengths. All extender sets in $\mathbb{Z}$ will be interpreted in this way.

Example 2.3. If X is the full shift on two symbols, $\{0,1\}^{\mathbb{Z}}$, then all words in $L(X)$ have the same extender set, since they are all identified with $\{0,1\}^{\mathbb{Z}}$.

Example 2.4. If X is the golden mean $\mathbb{Z}-$subshift on $\{0,1\}$ where adjacent 1s are prohibited, then $E(000)$ is the set of all legal configurations on $\mathbb{Z} \setminus \{0,1,2\}$, which is identified with the set of all $\{0,1\}$ sequences x which have no adjacent 1s, with the exception that $x_0 = x_1 = 1$ is allowed. This is because 000 may be preceded by a one-sided sequence ending with 1 and followed by a one-sided sequence beginning with 1, and after the identification with $\{0,1\}^{\mathbb{Z}}$, those 1s could become adjacent.

Similarly, $E(01)$ is identified with the set of all x on $\mathbb{Z}$ which have no adjacent 1s and satisfy $x_0 = 0$, and $E(1)$ is identified with the set of all x on $\mathbb{Z}$ which have no adjacent 1s and satisfy $x_0 = x_1 = 0$.

Therefore, even though they have different lengths, we can say here that $E(1) \subsetneq E(01) \subsetneq E(000) = E(0)$.

The next few definitions concern measures. Every measure in this work is assumed to be a Borel probability measure μ on a $\mathbb{G}-$subshift X which is invariant under all shifts σ_g. By a generalization of the Bogolyubov-Krylov theorem, every $\mathbb{G}-$subshift X has at least one such measure. For any such μ and any $w \in L(X)$, we will use $\mu(w)$ to denote $\mu([w])$.

For any Følner sequence $\{F_n\}$, we define the **entropy** of any such μ as

$$h_\mu(X) := \lim_{n\to\infty} \frac{1}{|F_n|} \sum\nolimits_{w \in \mathcal{A}^{F_n}} -\mu(w) \log \mu(w).$$

Again, this limit does not depend on the choice of Følner sequence (see [9] for proofs of this property and of other basic properties of entropy of amenable group actions).

It is always the case that $h_\mu(X) \leq h(X)$, and so a measure μ is called a **measure of maximal entropy** (or **MME**) if $h_\mu(X) = h_{top}(X)$. For amenable $\mathbb{G}$, every $\mathbb{G}-$subshift has at least one measure of maximal entropy [17].

We briefly summarize some classical results from ergodic theory. A measure μ is **ergodic** if every set which is invariant under all σ_g has measure 0 or 1. In fact, every measure μ can be written as a generalized convex combination (really an integral) of ergodic measures; this is known as the **ergodic decomposition** (e.g. see Section 8.7 of [6]). The entropy map $\mu \mapsto h_\mu$ is linear and so the ergodic decomposition extends to measures of maximal entropy as well; every MME can be written as a generalized convex combination of ergodic MMEs.

Theorem 2.5 (Pointwise ergodic theorem [12])**.** *For any ergodic measure μ on a $\mathbb{G}-$subshift X, there exists a Følner sequence $\{F_n\}$ such that for every $f \in L^1(\mu)$,*

$$\mu\left(\left\{x : \lim_{n\to\infty} \frac{1}{|F_n|} \sum_{g\in F_n} f(\sigma_g x) = \int f \, d\mu\right\}\right) = 1.$$

Theorem 2.6 (Shannon-Macmillan-Breiman theorem for amenable groups [23])**.** *For any ergodic measure μ on a $\mathbb{G}-$subshift X, there exists a Følner sequence $\{F_n\}$ such that*

$$\mu\left(\left\{x : \lim_{n\to\infty} -\frac{1}{|F_n|}\mu(x_{F_n}) = h_\mu(X)\right\}\right) = 1.$$

The classical pointwise ergodic and Shannon-Macmillan-Breiman theorems were originally stated for $\mathbb{G} = \mathbb{Z}$ and the Følner sequence $[0, n]$. We only need Theorem 2.6 for the following corollary.

Corollary 2.7. *Let μ be an ergodic measure of maximal entropy on a $\mathbb{G}$-subshift X. There exists a Følner sequence $\{F_n\}$ such that for every $S_n \subseteq L_{F_n}(X)$ such that $\mu(S_n) \to 1$, then*

$$\lim_{n\to\infty} \frac{1}{|F_n|} \log |S_n| = h_{top}(X).$$

Proof. Take X, μ as in the theorem, $\{F_n\}$ a Følner sequence that satisfies the Shannon-Macmillan-Breiman theorem, and S_n as in the theorem. Fix any $\epsilon > 0$. By the definition of topological entropy,

$$\limsup_{n\to\infty} \frac{1}{|F_n|} \log |S_n| \leq \lim_{n\to\infty} \frac{1}{|F_n|} \log |L_{F_n}(X)| = h_{top}(X).$$

For every n, define

$$T_n = \{w \in \mathcal{A}^{F_n} \ : \ \mu(w) < e^{-|F_n|(h_{top}(X)-\epsilon)}.$$

By the Shannon-Macmillan-Breiman theorem, $\mu\left(\bigcup_N \bigcap_{n=N}^{\infty} T_n\right) = 1$, and so $\mu(T_n) \to 1$. Therefore, $\mu(S_n \cap T_n) \to 1$, and by definition of T_n,

$$|S_n \cap T_n| \geq \mu(S_n \cap T_n) e^{|F_n|(h_{top}(X)-\epsilon)}.$$

Therefore, for sufficiently large n, $|S_n| \geq |S_n \cap T_n| \geq 0.5e^{|F_n|(h_{top}(X)-\epsilon)}$. Since $\epsilon > 0$ was arbitrary, the proof is complete. □

Finally, several of our main arguments rely on the following elementary combinatorial lemma, whose proof we leave to the reader.

Lemma 2.8. *If S is a finite set, $\{A_s\}$ is a collection of finite sets, $m = \min\{|A_s|\}$, and $M = \max_{a\in\bigcup A_s} |\{s \mid a \in A_s\}|$, then*

$$\left|\bigcup_{s\in S} A_s\right| \geq |S|\frac{m}{M}.$$

3. Results on $\mathbb{Z}-$Subshifts

In this section we present the results for $\mathbb{G} = \mathbb{Z}$, and must begin with some standard definitions about $\mathbb{Z}-$subshifts.

For words $v \in \mathcal{A}^m$ and $w \in \mathcal{A}^n$ with $m \leq n$, we say that v is a **prefix** of w if $w_{[0,m)} = v$, and v is a **suffix** of w if $w_{[n-m,n)} = v$.

We now need some technical definitions about replacing one or more occurrences of a word v by a word w inside a larger word u, which are key to most of our arguments in this section. First, for any $v \in L(\mathcal{A}^{\mathbb{Z}})$, we define the function $O_v : L(\mathcal{A}^{\mathbb{Z}}) \to \mathcal{P}(\mathbb{N})$ which sends any word u to the set of locations where v occurs as a subword in u, i.e.

$$O_v(u) := \{i \in \mathbb{N} : \sigma_i(u) \in [v]\}.$$

For any $w \in L(\mathcal{A}^{\mathbb{Z}})$, we may then define the function $R_u^{v\to w} : O_v(u) \to L(\mathcal{A}^{\mathbb{Z}})$ which replaces the occurrence of v within u at some position in $O_v(u)$ by the word w. Formally, $R_u^{v\to w}(i)$ is the word u' of length $|u|-|v|+|w|$ defined by $u'_{[0,i)} = u_{[0,i)}$, $u'_{[i,i+|w|)} = w$, and $u'_{[i+|w|,|u|-|v|+|w|)} = u_{[i+|v|,|u|)}$.

Our arguments in fact require replacing many occurrences of v by w within a word u, at which point some technical obstructions appear. For instance, if several occurrences of v overlap in u, then replacing one by w may destroy the other. The following defines conditions on v and w which obviate these and other problems which would otherwise appear in our counting arguments.

Definition 3.1. For $v, w \in L(\mathcal{A}^{\mathbb{Z}})$, we say that v **respects the transition to** w if, for any $u \in L(\mathcal{A}^{\mathbb{Z}})$ and any $i \in O_v(u)$,

(i) $j + |w| - |v| \in O_v(R_u^{v\to w}(i))$ for any $j \in O_v(u)$ with $i < j$,

(ii) $j \in O_v(R_u^{v\to w}(i))$ for any $j \in O_v(u)$ with $i > j$,

(iii) $j \in O_w(R_u^{v\to w}(i))$ for any $j \in O_w(u)$ with $i > j$,

(iv) $j + |w| - |v| > i$ for any $j \in O_v(u)$ with $i < j$.

Informally, v respects the transition to w if, whenever a single occurrence of v is replaced by w in a word u, all other occurrences of v in u are unchanged, all occurrences of w in u to the left of the replacement are unchanged, and all occurrences of v in u which were to the right of the replaced occurrence remain on that side of the replaced occurrence.

When v respects the transition to w, we are able to meaningfully define replacement of a set of occurrences of v by w, even when those occurrences of v overlap, as long as we move from left to right. For any $u, v, w \in L(X)$, we define a function $R_u^{v\to w} : \mathcal{P}(O_v(u)) \to L(X)$ as follows. For any $S := \{s_1, ..., s_n\} \subseteq O_v(u)$ (where we always assume $s_1 < s_2 < \ldots < s_n$), we define sequential replacements $\{u^m\}_{m=1}^{n+1}$ by

1) $u = u^1$.

2) $u^{m+1} = R_{u^m}^{v\to w}(s_m + (m-1)(|w| - |v|))$.

Finally, we define $R_u^{v\to w}(S)$ to be u^{n+1}.

We first need some simple facts about $R_u^{v\to w}$ which are consequences of Definition 3.1.

Lemma 3.2. *For any $u, v, w \in L(X)$ where v respects the transition to w and any $S = \{s_1, ..., s_n\} \subseteq O_v(u)$, all replacements of v by w persist throughout, i.e.* $\{s_1, s_2 + (|w| - |v|), s_3 + 2(|w| - |v|), \ldots, s_n + (n-1)(|w| - |v|)\} \subseteq O_w(R_u^{v \to w}(S))$.

Proof. Choose any v, w, u, S as in the lemma, and any $s_i \in S$. Using the terminology above, clearly $s_i + (i-1)(|w| - |v|) \in O_w(u^{(i+1)})$. By property (iv) of a respected transition, $s_1 < s_2 + |w| - |v| < \ldots < s_n + (n-1)(|w| - |v|)$. Then, since $s_i + (i-1)(|w| - |v|) < s_j + (j-1)(|w| - |v|)$ for $j > i$, by property (iii) of respected transition, $s_i + (i-1)(|w| - |v|) \in O_w(u^{(j+1)})$ for all $j > i$, and so $s_i + (i-1)(|w| - |v|) \in O_w(R_u^{v \to w}(S))$. Since i was arbitrary, this completes the proof. □

Lemma 3.3. *For any $u, v, w \in L(X)$ where v respects the transition to w and any $S = \{s_1, ..., s_n\} \subseteq O_v(u)$, any occurrence of v not explicitly replaced in the construction of $R_u^{v \to w}$ also persists, i.e. if $m \in O_v(u) \setminus S$ and $s_i < m < s_{i+1}$, then $m + i(|w| - |v|) \in O_v(R_u^{v \to w}(S))$.*

Proof. Choose any v, w, u, S as in the lemma, and any $m \in O_v(u) \cap (s_i, s_{i+1})$ for some i. Using property (i) of a respected transition, a simple induction implies that $m + j(|w| - |v|) \in O_v(u^{(j+1)})$ for all $j \leq i$. By property (iv) of a respected transition, $m + i(|w| - |v|) < s_{i+1} + i(|w| - |v|) < \ldots < s_n + (n-1)(|w| - |v|)$. Therefore, using property (ii) of a respected transition allows a simple induction which implies that $m + i(|w| - |v|) \in O_v(u^{(j+1)})$ for all $j > i$, and so $m + i(|w| - |v|) \in O_v(R_u^{v \to w}(S))$. □

We may now prove injectivity of $R_u^{v \to w}$ under some additional hypotheses, which is key for our main proofs.

Lemma 3.4. *Let $v, w \in L(X)$ such that v respects the transition to w, v is not a suffix of w, and w is not a prefix of v. For any $u \in L(X)$ and m, $R_u^{v \to w}$ is injective on the set of m-element subsets of $O_v(u)$.*

Proof. Assume that v, w, u are as in the lemma, and choose $S = \{s_1, ..., s_n\} \neq S' = \{s'_1, ..., s'_n\} \subseteq O_v(u)$ with $|S| = |S'| = m$.

We first treat the case where $|v| \geq |w|$, and recall that w is not a prefix of v. Since $S \neq S'$, we can choose i maximal so that $s_j = s'_j$ for $j < i$. Then $s_i \neq s'_i$; we assume without loss of generality that $s_i < s'_i$. Since $s_i \in S$, we know that $s_i \in O_v(u)$. Since $s'_{i-1} = s_{i-1} < s_i < s'_i$, by Lemma 3.3 $s_i + (i-1)(|w| - |v|) \in O_v(R_u^{v \to w}(S'))$. Also, by Lemma 3.2, $s_i + (i-1)(|w| - |v|) \in O_w(R_u^{v \to w}(S))$. Since w is not a prefix of v, this means that $R_u^{v \to w}(S) \neq R_u^{v \to w}(S')$, completing the proof of injectivity in this case.

We now treat the case where $|v| \leq |w|$, and recall that v is not a suffix of w. Since $S \neq S'$, we can choose i maximal so that $s_{m-j} = s'_{m-j}$ for $j < i$. Then $s_{m-i} \neq s'_{m-i}$; we assume without loss of generality that $s_{m-i} < s'_{m-i}$. Since $s'_{m-i} \in S'$, we know that $s'_{m-i} \in O_v(u)$. Since $s_{m-i} < s'_{m-i} < s'_{m-i+1} = s_{m-i+1}$, by Lemma 3.3 $s'_{m-i} + (m-i)(|w| - |v|) \in O_v(R_u^{v \to w}(S))$. Also, by Lemma 3.2, $s'_{m-i} + (m-i-1)(|w| - |v|) \in O_w(R_u^{v \to w}(S'))$. Since v is not a suffix of w, this means that $R_u^{v \to w}(S) \neq R_u^{v \to w}(S')$, completing the proof of injectivity in this case and in general. □

Lemma 3.5. *Let $v, w \in L(X)$ such that v respects the transition to w, v is not a suffix of w, and w is not a prefix of v. Then for any u' and any $m \leq |O_w(u')|$,*

$$|\{(u,S) \ : \ |S| = m, S \subseteq O_v(u), u' = R_u^{v\to w}(S)\}| \leq \binom{|O_w(u')|}{m}.$$

Proof. Assume that v, w, u' are as in the lemma, and denote the set above by $f(u')$. For any $(u,S) \in f(u')$ we define $g(S) = \{s_1, s_2 + |w| - |v|, \ldots, s_m + (m-1)(|w|-|v|)\}$; note that by Lemma 3.2, $g(S) \subseteq O_w(u')$.

We claim that for any S, there is at most one u for which $(u,S) \in f(u')$. One can find this u by simply reversing each of the replacements in the definition of $R_u^{v\to w}(S)$. Informally, the only such u is $u = R_{u'}^{v\leftarrow w}(g(S))$, where $R_{u'}^{v\leftarrow w}$ is defined analogously to $R_u^{v\to w}$ with replacements of w by v made from right to left instead of v by w made from left to right.

Finally, since $g(S) \subseteq O_w(u')$, and since g is clearly injective, there are less than or equal to $\binom{|O_w(u')|}{m}$ choices for S with $(u,S) \in f(u')$ for some u, completing the proof. □

We may now prove the desired relation for v, w with $E(v) \subseteq E(w)$ under additional assumptions on v and w.

Proposition 3.6. *Let X be a subshift, μ a measure of maximal entropy of X, and $v, w \in L(X)$. If v respects the transition to w, v is not a suffix of w, w is not a prefix of v, and $E_X(v) \subseteq E_X(w)$, then*

$$\mu(v) \leq \mu(w) e^{h_{top}(X)(|w|-|v|)}.$$

Proof. Let $\delta, \varepsilon \in \mathbb{Q}_+$. We may assume without loss of generality that μ is an ergodic MME, since proving the desired inequality for ergodic MMEs implies it for all MMEs by ergodic decomposition.

For every $n \in \mathbb{Z}_+$, we define

$$S_n := \{u \in L_n(X) : |O_v(u)| \geq n(\mu(v) - \delta) \text{ and } |O_w(u)| \leq n(\mu(w) + \delta)\}.$$

By the pointwise ergodic theorem (applied to $\chi_{[v]}$ and $\chi_{[w]}$), $\mu(S_n) \to 1$. Then, by Corollary 2.7, there exists N so that for $n > N$,

$$|S_n| > e^{n(h_{top}(X) - \delta)}. \tag{1}$$

For each $u \in S_n$, we define

$$A_u := \{R_u^{v\to w}(S) : S \subseteq O_v(u) \text{ and } |S| = \varepsilon n\}$$

(without loss of generality we may assume εn is an integer by taking a sufficiently large n.)

Since $E_X(v) \subseteq E_X(w)$ we have that $A_u \subset L(X)$. Also, by Lemma 3.4,

$$|A_u| = \binom{n(\mu(v) - \delta)}{\varepsilon n}$$

for every u.

On the other hand, for every $u' \in \bigcup_{u \in S_n} A_u$ we have that

$$|O_w(u')| \leq n(\mu(w) + \delta) + n\varepsilon(2|w| + 1)$$

(here, we use the fact that any replacement of v by w can create no more than $2|w|$ new occurrences of w.) Therefore, by Lemma 3.5,

$$|\{u \in S_n : u' \in A_u\}| \leq \binom{n\left(\mu(w) + \delta + (2|w|+1)\varepsilon\right)}{\varepsilon n}.$$

Then, by Lemma 2.8, we see that for $n > N$,

$$(2) \quad |L_m(X)| \geq \left| \bigcup_{u \in S_n} A_u \right| \geq |S_n| \binom{n(\mu(v) - \delta)}{\varepsilon n} \binom{n(mu(w) + \delta + (2|w|+1)\varepsilon)}{\varepsilon n}^{-1}$$
$$\geq e^{n(h_{top}(X) - \delta)} \binom{n(\mu(v) - \delta)}{\varepsilon n} \binom{n(mu(w) + \delta + (2|w|+1)\varepsilon)}{\varepsilon n}^{-1}.$$

For readability, we define $x = \mu(v) - \delta$ and $y = \mu(w) + \delta$. Now, we take logarithms of both sides, divide by n, and let n approach infinity. Stirling's approximation and the definition of entropy yield

$$h_{top}(X)(1 + \varepsilon(|w| - |v|)) \geq h_{top}(X) - \delta + x \log x - (x - \varepsilon)\log(x - \varepsilon)$$
$$- (y + (2|w|+1)\varepsilon)\log(y + (2|w|+1)\varepsilon) + (y + 2|w|\varepsilon)\log(y + 2|w|\varepsilon).$$

We subtract $h_{top}(X)$ from both sides, let $\delta \to 0$, and simplify to obtain

$$h_{top}(X)\varepsilon(|w| - |v|) \geq \varepsilon \log \mu(v) + (\mu(v) - \varepsilon)\left(\log \frac{\mu(v)}{\mu(v) - \varepsilon}\right)$$
$$- \varepsilon \log(\mu(w) + (2|w|+1)\varepsilon) - (\mu(w) + 2|w|\varepsilon)\log\left(\frac{\mu(w) + (2|w|+1)\varepsilon}{\mu(w) + 2|w|\varepsilon}\right).$$

We have that

$$\lim_{\varepsilon \to 0} \frac{\mu(v) - \varepsilon}{\varepsilon} \log \frac{\mu(v)}{\mu(v) - \varepsilon}$$
$$= \lim_{\varepsilon \to 0} \frac{\mu(v)}{\varepsilon} \log \frac{\mu(v)}{\mu(v) - \varepsilon}$$
$$= 1,$$

and

$$\lim_{\varepsilon \to 0} -\frac{\mu(w) + 2|w|\varepsilon}{\varepsilon} \log\left(\frac{\mu(w) + (2|w|+1)\varepsilon}{\mu(w) + 2|w|\varepsilon}\right)$$
$$= \lim_{\varepsilon \to 0} -\frac{\mu(w)}{\varepsilon} \log\left(\frac{\mu(w) + (2|w|+1)\varepsilon}{\mu(w) + 2|w|\varepsilon}\right)$$
$$= \lim_{t \to 0} -\frac{1}{t} \log\left(\frac{1 + (2|w|+1)t}{1 + 2|w|t}\right)$$
$$= -1.$$

This implies (by dividing by ε and taking limit on the previous estimate) that

$$h_{top}(X)(|w| - |v|) \geq \log \mu(v) - \log \mu(w).$$

Exponentiating both sides and solving for $\mu(v)$ completes the proof. □

Our strategy is now to show that any pair v, w, the cylinder sets $[v]$ and $[w]$ may each be partitioned into cylinder sets of the form $[\alpha v \beta]$ and $[\alpha w \beta]$ where the additional hypotheses of Theorem 3.6 hold on corresponding pairs. For this, we make the additional assumption that X has positive entropy to avoid some pathological examples (for instance, note that if $X = \{0^\infty\}$, then it's not even possible to satisfy the hypotheses of Theorem 3.6!)

Proposition 3.7. *Consider any subshift X with positive entropy, μ an ergodic measure of maximal entropy of X, and $v, w \in L(X)$. For almost every $x \in [v]$ there exists $N, M \in \mathbb{Z}_+$ such that $\alpha v \beta = x_{[-N,M]}$ respects the transition to $\alpha w \beta$, $\alpha v \beta$ is not a suffix of $\alpha w \beta$, and $\alpha w \beta$ is not a prefix of $\alpha v \beta$.*

Proof. Define

$$Q := \{\gamma \in L(X) : \mu(\gamma) > 0\}$$

and, for all $n \in \mathbb{N}$, define $Q_n := Q \cap A^n$.

Recall that

$$h_\mu(X) = \lim_{n\to\infty} \frac{1}{n} \sum\nolimits_{w \in A^n} -\mu(w) \log \mu(w).$$

The only positive terms of this sum are those corresponding to $w \in Q_n$, and it's a simple exercise to show that when $\sum_{i=1}^t \alpha_i = 1$, $\sum_{i=1}^t -\alpha_i \log \alpha_i$ has a maximum value of $\log t$. Therefore,

$$h_\mu(X) \leq \liminf_{n\to\infty} \frac{1}{n} \log |Q_n|.$$

Since $h_\mu(X) > 0$, $|Q_n|$ grows exponentially. Therefore, there exists $n_2 \in \mathbb{Z}_+$ such that for every $n \geq n_2$ we have that $|Q_n| \geq 2n$.

Let

$$\begin{aligned}
&N := \max\{n_2, |v|\} + 1,\\
&P := \left\{x \in X \ : \ x_{(-\infty,0)} \text{ periodic with period less than } |w|\right\},\\
&S := \left\{x \in X \ : \ \forall \gamma \in Q, \gamma \text{ is a subword of } x_{[0,\infty)}\right\}, \text{ and}\\
&R := [v] \cap S \diagdown P.
\end{aligned}$$

Assume that $v \neq w$ otherwise the result is trivial. Since μ has positive entropy, it is not supported on points with period less than $|w|$, and so for each $i \leq |w|$, there exists a word $u_i \in L_{i+1}(X)$ with different first and last letters. Then the pointwise ergodic theorem (applied to $\chi_{[u_1]}, \ldots, \chi_{[u_{|w|-1}]}$ with $F_n = [-n, 0)$) implies that $\mu(P) = 0$. The pointwise ergodic theorem (applied to $\chi_{[\gamma]}$ for $\gamma \in Q$ with $F_n = [0, n]$) shows that $\mu(S) = 1$, and so $\mu(R) = \mu(v)$. Choose arbitrary $x \in R$. If $x_{(-\infty,0)}v$ is a suffix of $x_{(-\infty,0)}w$, then clearly $|w| > |v|$, and for any $i > 0$, the $(i + |w|)$th letters from the end of $x_{(-\infty,0)}v$ and $x_{(-\infty,0)}w$ must be the same, i.e. $x(-i) = x(-i - |w| + |v|)$. This would imply $x \in P$, which is not the case, and so $x_{(-\infty,0)}v$ is not a suffix of $x_{(-\infty,0)}w$.

We can therefore define $N' \geq N$ to be minimal so that for $\alpha = x_{[-N',0]}$, αv is not a suffix of αw. (Obviously if $|v| \geq |w|$, then $N' = N$.)

Since $x \in S$, we can define the minimal M so that all N'-letter words of positive μ-measure are subwords of $x_{[-N',M)}$; for brevity we write this as $Q_{N'} \sqsubset x_{[-N',M)}$.

Since $N' \geq N \geq n_2$, $|Q_{N'}| \geq 2N'$, and so $M > 2N'$. Also, since M is the first natural with $Q_{N'} \sqsubset x_{[-N',M)}$, then

$$\left| O_{x_{[M-N',M)}}(x_{[-N',M)}) \right| = 1,$$

i.e. the N'-letter suffix of $x_{[-N',M)} = \alpha v \beta$ appears only at the end of $\alpha v \beta$.

First, it is clear that $\alpha v\beta$ is not a suffix of $\alpha w\beta$, since αv was not a suffix of αw by definition of α. Since the N'-letter suffix of $\alpha w\beta$ appears only once within $\alpha v\beta$, we see that $\alpha w\beta$ cannot be a prefix of $\alpha v\beta$ either.

It remains to show that $\alpha v\beta = x_{[-N',M)}$ respects the transition to $\alpha w\beta$. Suppose that a word $u \in L(X)$ contains overlapping copies of $\alpha v\beta$, i.e. we have $i, j \in O_{\alpha v\beta}(u)$ with $j > i$. Since $\left|O_{x_{[M-N',M)}}(x_{[-N',M)})\right| = 1$ we have that $j > i+|v|+M$; otherwise the N'-letter suffix of $\alpha v\beta = x_{[i,i+N'+|v|+M)}$ would be a non-terminal subword of $\alpha v\beta = x_{[j,j+N'+|v|+M)}$. Then $j + |w| - |v| > i + |w| + M > i$, and so property (iv) is verified. Since $j > i+|v|+M$, the central v within $x_{[i,i+N'+|v|+M)}$ is disjoint from $x_{[j,j+N'+|v|+M)}$, and so $j+|w|-|v| \in O_v(R_u^{v\to w}(i))$, verifying property (i).

For property (ii), the same argument as above shows that when $i, j \in O_{\alpha v\beta}(u)$ with $i > j$, $i > j+|v|+M$. Again this means that the central v within $x_{[i,i+N'+|v|+M)}$ is disjoint from $x_{[j,j+N'+|v|+M)}$, and so $j \in O_v(R_u^{v\to w}(i))$, verifying property (ii) and completing the proof.

For property (iii), we simply note that the proof of (ii) is completely unchanged if we instead assumed $j \in O_{\alpha w\beta}(u)$, since the N'-letter suffixes of $\alpha w\beta$ and $\alpha v\beta$ are the same.

□

Finally, we can use Proposition 3.7 to prove the main result of this section.

Theorem 3.8. *Consider any X a subshift with positive entropy, μ a measure of maximal entropy of X, and $v, w \in L(X)$. If $E_X(v) \subseteq E_X(w)$ then*

$$\mu(v) \leq \mu(w) e^{h_{top}(X)(|w|-|v|)}.$$

Proof. Consider X, μ, v, w as in the theorem. We may prove the result for only ergodic μ, since it then follows for all μ by ergodic decomposition.

By Proposition 3.7, there exists $R \subseteq [v]$ with $\mu(R) = \mu(v)$ so that for every $x \in R$, we can define $g(x) := \alpha v\beta$ and $g'(x) := \alpha w\beta$ satisfying the conclusion of Proposition 3.7. Since $E_X(v) \subseteq E_X(w)$, $g'(x) \in L(X)$ and $E_X(g(x)) \subseteq E_X(g'(x))$ for every $x \in R$.

Now, using Proposition 3.6 and Proposition 3.7 we have that

$$\mu(g(x)) \leq \mu(g'(x)) e^{h_{top}(X)(|g'(x)|-|g(x)|)} = \mu(g'(x)) e^{h_{top}(X)(|w|-|v|)}. \tag{3}$$

We claim that for $g(x) \neq g(x') \in g(R)$, $[g(x)]$ and $[g(x')]$ are disjoint. To see this, assume for a contradiction that $g(x) = \alpha v\beta$, $g(x') = \alpha' v\beta'$, and $[\alpha v\beta] \cap [\alpha' v\beta'] \neq \varnothing$. Then either $\alpha \neq \alpha'$ or $\beta \neq \beta'$, we assume the former for now. Either α is a suffix of α' or vice versa, again we assume the former without loss of generality. But then αv satisfies the definition of α in Proposition 3.7, and so we have a contradiction; α' was not the shortest possible choice in the definition of α' and β' for $g(x')$. The argument when $\beta \neq \beta'$ is similar; again the contradiction comes because one of β or β' was not in fact the minimal possible choice when it was defined.

Therefore, $\{[g(x)]\}$ forms a partition of R. Since all $[g(x)]$ are disjoint, all $\{[g'(x)]\}$ are also disjoint, and so

$$\sum_{g(x)\in g(R)} \mu(g(x)) = \mu(R) = \mu(v) \text{ and}$$

$$\sum_{g'(x)\in g'(R)} \mu(g'(x)) \le \mu(w).$$

(In fact the final inequality is an equality since the collection $\{[g'(x)]\}$ actually partitions $[w] \cap S \setminus P$, using the language of Proposition 3.7; we will not need this fact though.) We may then sum (3) over $x \in R$ yielding

$$\begin{aligned}\mu(v) &= \sum_{g(x)\in g(R)} \mu(g(x)) \\ &\le e^{h_{top}(X)(|w|-|v|)} \sum_{g'(x)\in g'(R)} \mu(g'(x)) \\ &\le \mu(w)e^{h_{top}(X)(|w|-|v|)},\end{aligned}$$

as desired.

□

The following corollary is immediate.

Corollary 3.9. *Let X be a $\mathbb{Z}$-subshift, μ a measure of maximal entropy of X, and $w, v \in L(X)$. If $E_X(v) = E_X(w)$, then for every measure of maximal entropy of X,*

$$\mu(v) = \mu(w)e^{h_{top}(X)(|w|-|v|)}.$$

3.1. **Applications.** We will now present some corollaries and applications of Theorem 3.8 and Corollary 3.9 to various classes of subshifts.

3.1.1. *Synchronized subshifts.*

Definition 3.10. For a subshift X, we say that $v \in L(X)$ is **synchronizing** if for every $uv, vw \in L(X)$, it is true that $uvw \in L(X)$. A subshift X is **synchronized** if $L(X)$ contains a synchronizing word.

The following fact is immediate from the definition of synchronizing word.

Lemma 3.11. *If w is a synchronizing word for a subshift X, then for any $v \in L(X)$ which contains w as both a prefix and suffix, $E_X(v) = E_X(w)$.*

Definition 3.12. A subshift X is **entropy minimal** if every subshift strictly contained in X has lower topological entropy. Equivalently, X is entropy minimal if every MME on X is fully supported.

The following result was first proved in [21], but we may also derive it as a consequence of Corollary 3.9 with a completely different proof.

Theorem 3.13. *Let X be a synchronized subshift. If X is entropy minimal then X has a unique measure of maximal entropy.*

Proof. Let μ be an ergodic measure of maximal entropy of such an X. Let w be a synchronizing word, $u \in L(X)$ and

$$R_u := \left\{x \in [u] : \left|O_w(x_{(-\infty,0]})\right| \ge 1 \text{ and } \left|O_w(x_{[(|u|,\infty])})\right| \ge 1\right\}.$$

Since X is entropy minimal, $\mu(w) > 0$, and so by the pointwise ergodic theorem (applied to $\chi_{[w]}$ with $F_n = [-n, 0]$ or $(|u|, n])$, $\mu(R_u) = \mu(u)$.

For every $x \in R_u$ we define minimal $n \geq |w|$ and $m \geq |w| + |u|$ so that $g_u(x) := x_{[-n,m]}$ contains w as both a prefix and a suffix. Then $\{[g_u(x)]\}$ forms a partition of R_u.

By Lemma 3.11, $E(w) = E(wvw)$ for all v s.t. $wvw \in L(X)$. Then by Corollary 3.9 we have that

$$\mu(g_u(x)) = \mu(w) e^{h_{top}(X)(|w| - |g_u(x)|)}.$$

Since $g_u(R_u)$ is countable we can write

$$\mu(u) = \mu(R_u) = \mu(w) \sum_{g_u(x) \in g_u(R_u)} e^{h_{top}(X)(|w| - |g_u(x)|)}.$$

This implies that

$$1 = \sum_{a \in \mathcal{A}} \mu(a) = \mu(w) \sum_{a \in \mathcal{A}} \sum_{g_a(x) \in g_a(R_a)} e^{h_{top}(X)(|w| - |g_a(x)|)}.$$

We combine the two equations to yield

$$\begin{aligned} \mu(u) &= \frac{\sum_{g_u(x) \in g_u(R_u)} e^{h_{top}(X)(|w| - |g_u(x)|)}}{\sum_{a \in \mathcal{A}} \sum_{g_a(x) \in g_a(R_a)} e^{h_{top}(X)(|w| - |g_a(x)|)}} \\ &= \frac{\sum_{g_u(x) \in g_u(R_u)} e^{-h_{top}(X)|g_u(x)|}}{\sum_{a \in \mathcal{A}} \sum_{g_a(x) \in g_a(R_a)} e^{-h_{top}(X)|g_a(x)|}}. \end{aligned}$$

Since the right-hand side is independent of the choice of the measure we conclude there can only be one ergodic measure of maximal entropy, which implies by ergodic decomposition that there is only one measure of maximal entropy. $\square$

In [4], one of the main tools used in proving uniqueness of the measure of maximal entropy for various subshifts was boundedness of the quantity $\frac{|L_n(X)|}{e^{nh_{top}(X)}}$. One application of our results is to show that this quantity in fact converges to a limit for a large class of synchronized shifts.

Definition 3.14. A measure μ on a subshift X is **mixing** if, for all measurable A, B,

$$\lim_{n \to \infty} \mu(A \cap \sigma_{-n} B) = \mu(A)\mu(B).$$

Theorem 3.15. *Let X be a synchronized entropy minimal subshift such that the measure of maximal entropy is mixing. We have that*

$$\lim_{n \to \infty} \frac{|L_n(X)|}{e^{nh_{top}(X)}} \text{ exists.}$$

Proof. We denote $\lambda := e^{h_{top}(X)}$ and define μ to be the unique measure of maximal entropy for X. Let $w \in L(X)$ be a synchronizing word and

$$R_n := \{u \in L_n(X) : w \text{ is a prefix and a suffix of } u\}.$$

Lemma 3.11 and Corollary 3.9 imply that for every $u \in R_n$,

$$\mu(u) = \mu(w) \lambda^{|w| - n}.$$

This implies that

$$\sum_{u \in R_n} \mu(u) = |R_n| \, \mu(w) \lambda^{|w| - n}$$

On the other hand

$$\sum_{u \in R_n} \mu(u) = \mu([w] \cap \sigma_{|w|-n}[w]).$$

Since the measure is mixing we obtain that

$$\lim_{n\to\infty} \mu([w] \cap \sigma_{|w|-n}[w]) = \mu([w])^2.$$

Combining the three equalities above yields

$$\lim_{n\to\infty} \frac{|R_n|}{\lambda^n} = \frac{\mu(w)}{\lambda^{|w|}}.$$

For all $n \in \mathbb{N}$, we define

$$P_n := \left\{u \in L_{n+|w|}(x) : w \text{ is a prefix of } u, |O_w|(u) = 1\right\} \text{ and}$$
$$S_n := \left\{u \in L_{n+|w|}(x) : w \text{ is a suffix of } u, |O_w|(u) = 1\right\}$$

to be the sets of $(n + |w|)$-letter words in $L(X)$ containing w exactly once as a prefix/suffix respectively. We also define

$$K_n := \{u \in L_n(x) : |O_w(u)| = 0\}$$

to be the set of n-letter words in $L(X)$ not containing w. Then partitioning words in $L_n(X) \setminus K_n$ by the first and last appearance of w, recalling that w is synchronized, gives the formula

$$|L_n(X)| = |K_n| + \sum_{0\le i<j\le n} |S_i||R_{j-i}||P_{n-j}|,$$

thus

$$\frac{|L_n(X)|}{\lambda^n} = \frac{|K_n|}{\lambda^n} + \sum_{0\le i<j\le n} \frac{|S_i|}{\lambda^i}\frac{|R_{j-i}|}{\lambda^{j-i}}\frac{|P_{n-j}|}{\lambda^{n-j}}. \tag{4}$$

We now wish to take the limit as $n \to \infty$ of both sides of (4). First, we note that since X is entropy minimal, $h_{top}(X_w) < h_{top}(X)$, where X_w is the subshift of points of X not containing w. Therefore,

$$\limsup_{n\to\infty} \frac{1}{n} \log|K_n| < h_{top}(X).$$

Since all words in P_n and S_n are the concatenation of w with a word in K_n, $|P_n|, |S_n| \le |K_n|$, and so

$$\limsup_{n\to\infty} \frac{1}{n} \log|P_n|, \limsup_{n\to\infty} \frac{1}{n} \log|S_n| < h_{top}(X),$$

implying that the infinite series

$$\sum_{n=0}^{\infty} \frac{|P_n|}{\lambda^n} \text{ and } \sum_{n=0}^{\infty} \frac{|S_n|}{\lambda^n} \text{ converge.}$$

We now take the limit of the right-hand side of (4).

$$\lim_{n\to\infty} \frac{|K_n|}{\lambda^n} + \sum_{0\le i<j\le n} \frac{|S_i|}{\lambda^i}\frac{|R_{j-i}|}{\lambda^{j-i}}\frac{|P_{n-j}|}{\lambda^{n-j}} = \lim_{n\to\infty} \sum_{0\le k\le n} \left(\frac{|R_k|}{\lambda^k}\left(\sum_{i=0}^{n-k} \frac{|S_i|}{\lambda^i}\frac{|P_{n-k-i}|}{\lambda^{n-k-i}}\right)\right).$$

Since $\frac{|R_k|}{\lambda^k}$ converges to the limit $\frac{\mu(w)}{\lambda^{|w|}}$ and the series $\sum_{m=0}^{\infty}\sum_{i=0}^{m}\frac{|S_i|}{\lambda^i}\frac{|P_{m-i}|}{\lambda^{n-k-i}}$ converges, the above can be rewritten as

$$\lim_{n\to\infty}\sum_{0\le k\le n}\left(\frac{|R_k|}{\lambda^k}\left(\sum_{i=0}^{n-k}\frac{|S_i|}{\lambda^i}\frac{|P_{n-k-i}|}{\lambda^{n-k-i}}\right)\right)=\frac{\mu(w)}{\lambda^{|w|}}\lim_{m\to\infty}\sum_{m=0}^{\infty}\sum_{i=0}^{m}\frac{|S_i|}{\lambda^i}\frac{|P_{m-i}|}{\lambda^{n-k-i}}$$
$$=\frac{\mu(w)}{\lambda^{|w|}}\sum_{n=0}^{\infty}\frac{|P_n|}{\lambda^n}\sum_{n=0}^{\infty}\frac{|S_n|}{\lambda^n}.$$

Recalling (4), we see that $\lim_{n\to\infty}\frac{|L_n(X)|}{\lambda^n}$ converges to this limit as well, completing the proof. □

We will be able to say even more about a class of synchronized subshifts called the S-gap subshifts.

Definition 3.16. Let $S\subseteq\mathbb{N}\cup\{0\}$. We define the $S-$gap subshift X_S by the set of forbidden words $\{10^n1 \ : \ n\notin S\}$. Alternately, X_S is the set of bi-infinite $\{0,1\}$ sequences where the gap between any two nearest 1s has length in S.

It is immediate from the definition that 1 is a synchronizing word for every $S-$gap subshift. Also, all S-gap subshifts are entropy minimal (see Theorem C, Remark 2.4 of [5]), and as long as $\gcd(S+1)=1$, their unique measure of maximal entropy is mixing (in fact Bernoulli) by Theorem 1.6 of [3]. (This theorem guarantees that the unique MME is Bernoulli up to period d given by the gcd of periodic orbit lengths, and it's clear that $S+1$ is contained in the set of periodic orbit lengths.)

In this case Climenhaga [2] conjectured that the limit $\lim_{n\to\infty}\frac{|L_n(X_S)|}{e^{nh_{top}(X_S)}}$ existed; we prove this and we give an explicit formula for the limit.

Corollary 3.17. *Let X_S be an $S-$gap subshift with* $\gcd(S+1)=1$. *Then for μ the unique MME on X_S,*

$$\lim_{n\to\infty}\frac{|L_n(X_S)|}{e^{nh_{top}(X_S)}}=\frac{\mu(1)e^{h_{top}(X_S)}}{(e^{h_{top}(X_S)}-1)^2}.$$

Proof. Using the notation of Theorem 3.15, we define $w=1$ and write $\lambda=e^{h_{top}(X_S)}$. It is easy to see that $|P_i|=|S_i|=1$ for all i. As noted above, X_S is entropy minimal and its unique measure of maximal entropy is mixing, and so the proof of Theorem 3.15 implies that

$$\lim_{n\to\infty}\frac{|L_n(X_S)|}{e^{nh_{top}(X_S)}}=\frac{\mu(1)}{\lambda}\left(\sum_{i=0}^{\infty}\frac{1}{\lambda^i}\right)^2=\frac{\mu(1)}{\lambda}\left(\frac{1}{1-\lambda^{-1}}\right)^2=\frac{\mu(1)\lambda}{(\lambda-1)^2},$$

completing the proof.

□

As noted in [2], a motivation for proving the existence of this limit is to fill a gap from [19] for a folklore formula for the topological entropy of X_S. Two proofs of this formula are presented in [2], and Corollary 3.9 yields yet another proof.

Corollary 3.18. *Let X_S be an S-gap subshift. Then $h_{top}(X_S)=\log\lambda$, where λ is the unique solution of*

$$1=\sum_{n\in S}\lambda^{-n-1}.$$

Proof. For any S-gap shift X_S, we can write

$$[1] = \left(\bigsqcup_{n=0}^{\infty} [10^n 1] \right) \cup \{x \in X_S \ : \ x_0 = 1 \text{ and } \forall n > 0, x_n = 0\}.$$

By shift-invariance, $\mu(10^\infty) = 0$, and so by Lemma 3.11 and Corollary 3.9,

$$\mu(1) = \sum_{n \in S} \mu(10^n 1) = \sum_{n \in S} \mu(1) e^{h_{top}(X_S)(-n-1)}.$$

Dividing both sides by $\mu(1)$ completes the proof.

□

We also prove that for every $S-$gap subshift, the unique measure of maximal entropy has highly constrained values, which are very similar to those of the Parry measure for shifts of finite type.

Theorem 3.19. *Let X_S be an $S-$gap subshift and μ the measure of maximal entropy. Then $\mu(1) = \frac{1}{\sum_{n \in S}(n+1)e^{-h_{top}(X_S)(n+1)}}$, and for every $w \in L(X_S)$, there exists a polynomial f_w with integer coefficients so that $\mu(w) = k_w + \mu(1) f_w(e^{-h_{top}(X_S)})$ for some integer k_w.*

Proof. As noted above, S-gap shifts are synchronized and entropy minimal, and so have unique measures of maximal entropy.

Denote by μ the unique measure of maximal entropy for some $S-$gap subshift X_S, and for readability we define

$$t = e^{-h_{top}(X)}.$$

Since X_S is entropy minimal, $\mu(1) > 0$, and so by the pointwise ergodic theorem (applied to $\chi_{[1]}$), μ-a.e. point of X_S contains infinitely many 1s. Therefore, we can partition points of X_S according to the closest 1 symbols to the left and right of the origin, and represent X_S (up to a null set) as the disjoint union $\bigcup_{n \in S} \bigcup_{i=0}^{n} \sigma_i [10^n 1]$. Then by Lemma 3.11 and Corollary 3.9,

$$\begin{aligned} 1 &= \sum_{n \in S} (n+1) \mu(10^n 1) \\ &= \sum_{n \in S} (n+1) \mu(1) t^{n+1}, \end{aligned}$$

yielding the claimed formula for $\mu(1)$.

Now we prove the general formula for $\mu(w)$, and will proceed by induction on the length n of w. For the base case $n = 1$, $\mu(0) = 1 - \mu(1)$, verifying the theorem.

Now, assume that the theorem holds for every $n \leq N$ for some $N \geq 1$. Let $w \in L_{N-1}(X_S)$, and we will verify the theorem for $1w1$, $1w0$, $0w1$, and $0w0$. If $1w1 \notin L(X_S)$, then

$$\begin{aligned} &\mu(1w1) = 0, \\ &\mu(1w0) = \mu(1w) - \mu(1w1) = \mu(1w), \\ &\mu(0w1) = \mu(w1) - \mu(1w1) = \mu(w1), \text{ and} \\ &\mu(0w0) = 1 - \mu(1w1) - \mu(1w0) - \mu(0w1) = 1 - \mu(1w) - \mu(w1). \end{aligned}$$

The theorem now holds by the inductive hypothesis.

If $1w1 \in L(X_S)$, then as before $E_{X_S}(1w1) = E_{X_S}(1)$, implying

$\mu(1w1) = \mu(1)t^{1+|w|}$,

$\mu(1w0) = \mu(1w) - \mu(1w1) = \mu(1w) - \mu(1)t^{1+|w|}$,

$\mu(0w1) = \mu(w1) - \mu(1w1) = \mu(w1) - \mu(1)t^{1+|w|}$, and

$\mu(0w0) = 1 - \mu(1w1) - \mu(1w0) - \mu(0w1) = 1 - \mu(1w) - \mu(w1) + \mu(1)t^{1+|w|}$,

again implying the theorem by the inductive hypothesis and completing the proof. □

3.1.2. *Hereditary subshifts.* One particular class of subshifts with many pairs of words such that $E_X(v) \subsetneq E_X(w)$ are the hereditary subshifts (definition introduced in [8]). Some examples are $\beta-$shifts [11], $\mathscr{B}-$free shifts ([10]), spacing shifts ([14]), multi-choice shifts ([13]) and bounded density shifts ([20]). Many of these examples have a unique measure of maximal entropy, but not every hereditary subshift has this property (see [10]) .

A partial order $\leq$ on a finite set $\mathcal{A}$ induces a partial order on $\mathcal{A}^n$ and $\mathcal{A}^{\mathbb{Z}}$ (coordinatewise) which will also be denoted by $\leq$. When $\mathcal{A} = \{0, 1..., m\}$ we will always use the linear order $0 \leq 1 \leq \ldots \leq m$.

Definition 3.20. Let $X \subseteq \mathcal{A}^{\mathbb{Z}}$ be a subshift and $\leq$ a partial order on $\mathcal{A}$. We say X is $\leq -$**hereditary (or simply hereditary)** if for every $x \in \mathcal{A}^{\mathbb{Z}}$ such that there exists $y \in X$ such that $x \leq y$ then $x \in X$.

This definition immediately implies that whenever $x \leq y$ for $x, y \in L(X)$, $E_X(y) \subseteq E_X(x)$, yielding the following corollary of Theorem 3.8.

Corollary 3.21. *Let X be a $\leq -$hereditary subshift, μ a measure of maximal entropy, and $v, w \in L_n(X)$ for some $n \in \mathbb{N}$. If $u \leq v$ then $\mu(v) \leq \mu(u)$.*

Having $u \leq v$ is sufficient but not necessary for $E(v) \subseteq E(w)$. In particular, for $\beta-$shifts and bounded density shifts, there are many other pairs (with different lengths) where this happens. This is due to an additional property satisfied by these hereditary shifts.

Definition 3.22. Let $X \subseteq \{0, 1, ..., m\}^{\mathbb{Z}}$ be a hereditary subshift. We say X is i**-hereditary** if for every $u \in L_n(X)$ and u' obtained by inserting a 0 somewhere in u, it is the case that $u' \in L_{n+1}(X)$.

In particular, $\beta-$shifts and bounded density shifts are i-hereditary, but not every spacing shift is i-hereditary. It's immediate that any i-hereditary shift satisfies $E_X(0^j) \subseteq E_X(0^k)$ whenever $j \geq k$. We can get equality if we assume the additional property of specification.

Definition 3.23. A subshift X has the **specification property (at distance** N**)** if for every $u, w \in L(X)$ there exists $v \in L_N(X)$ such that $uvw \in L(X)$.

Clearly, if X is hereditary and has specification property at distance N, then $u0^N w$ and $u0^{N+1}w \in L(X)$ for all $u, w \in L(X)$, and so in this case $E_X(0^N) = E_X(0^{N+1})$. We then have the following corollary of Theorem 3.8.

Corollary 3.24. *Let $X \subseteq \{0, 1, ..., m\}^{\mathbb{Z}}$ be a i-hereditary subshift. Then for every $n \in \mathbb{Z}_+$*

$$h_{top}(X) \geq \log \frac{\mu(0^n)}{\mu(0^{n+1})}.$$

Furthermore, if X has the specification property at distance N, then
$$h_{top}(X) = \log \frac{\mu(0^N)}{\mu(0^{N+1})}.$$

We note that if X has the specification property at distance N, then it also has it at any larger distance. Therefore, the final formula can be rewritten as

$$\begin{aligned} h_{top}(X) = \lim_{N\to\infty} \log \frac{\mu(0^N)}{\mu(0^{N+1})} &= \lim_{N\to\infty} -\log \mu(x(0) = 0 \mid x_{[-N,-1]} = 0^N) \\ &= -\log \mu(x(0) = 0 \mid x_{(-\infty,-1]} = 0^\infty), \end{aligned}$$

recovering a formula (in fact a more general one for topological pressure of $\mathbb{Z}^d$ SFTs) proved under different hypotheses in [15].

4. $\mathbb{G}-$SUBSHIFTS

Throughout this section, $\mathbb{G}$ will denote a countable amenable group generated by a finite set $G = \{g_1, ..., g_d\}$ which is torsion-free, i.e. $g^n = e$ if and only if $n = 0$. For any $N = (N_1, ..., N_d) \in \mathbb{Z}_+^d$, we define $\mathbb{G}_N$ to be the subgroup generated by $\left\{g_1^{N_1}, ..., g_d^{N_d}\right\}$, and use ${}^{\mathbb{G}}\!/\!_{\mathbb{G}_N}$ to represent the collection $\{g \cdot \mathbb{G}_N : g \in \mathbb{G}\}$ of left cosets of $\mathbb{G}_N$. Clearly, $\left|{}^{\mathbb{G}}\!/\!_{\mathbb{G}_N}\right| = N_1 N_2 \cdots N_d$.

We again must begin with some relevant facts and definitions. The following structural lemma is elementary, and we leave the proof to the reader.

Lemma 4.1. *For any amenable $\mathbb{G}$ and $F \Subset \mathbb{G}$, there exists $N = (N_1, ..., N_d) \in \mathbb{Z}_+^d$ such that for every nonidentity $g \in \mathbb{G}_N$, $g \cdot F \cap F = \varnothing$.*

As in the $\mathbb{Z}$ case, if $v, w \in L_F(\mathcal{A}^{\mathbb{G}})$ for some $F \Subset \mathbb{G}$, we define the function $O_v : L(\mathcal{A}^{\mathbb{G}}) \to \mathcal{P}(\mathbb{N})$ which sends a word to the set of locations where v appears as a subword, i.e.
$$O_v(u) := \{g \in \mathbb{G} : \sigma_g(u) \in [v]\}.$$
We also define the function $R_u^{v\to w} : O_v(u) \to L(\mathcal{A}^{\mathbb{G}})$, where $R_u^{v\to w}(g)$ is the word you obtain by replacing the occurrence of v at $g \cdot F$ within u by w.

We now again must define a way to replace many occurrences of v by w within a word u, but will do this via restricting the sets of locations where the replacements occur rather than the pairs (v, w). We say $S \subset \mathbb{G}$ is $F-$**sparse** if $g \cdot F \cap g' \cdot F = \varnothing$ for every unequal pair $g, g' \in S$. When $v, w \in L_F(X)$ and S is $F-$sparse, we may simultaneously replace occurrences of v by w at locations $g \cdot F$, $g \in S$ by w without any of the complications dealt with in the one-dimensional case, and we denote the resulting word by $R_u^{v\to w}(S)$. Formally, $R_u^{v\to w}(S)$ is just the image of u under the composition of $R_u^{v\to w}(s)$ over all $s \in S$.

The following lemmas are much simpler versions of Lemmas 3.4 and 3.5 for F-sparse sets.

Lemma 4.2. *For any F, $v, w \in L_F(X)$, and F-sparse set $T \subseteq O_v(u)$, $R_u^{v\to w}$ is injective on subsets of T.*

Proof. Fix F, u, v, w, T as in the lemma. If $S \neq S' \subseteq T$, then either $S \setminus S'$ or $S' \setminus S$ is nonempty; assume without loss of generality that it is the former. Then, if $s \in S \setminus S'$, by definition $(R_u^{v\to w}(S))_{s+F} = w$ and $(R_u^{v\to w}(S'))_{s+F} = v$, and so $R_u^{v\to w}(S) \neq R_u^{v\to w}(S')$. $\square$

Lemma 4.3. *For any F and $v, w \in L_F(X)$, any F-sparse set $T \subseteq O_v(u)$, any u', and any $m \leq |T \cap O_w(u')|$,*

$$|\{(u, S) \ : \ S \text{ is } F\text{-sparse}, |S| = m, S \subseteq T, u' = R_u^{v \to w}(S)\}| \leq \binom{|T \cap O_w(u')|}{m}.$$

Proof. Fix any such F, u', v, w, T, m as in the lemma. Clearly, for any S, $S \subseteq O_w(R_u^{v \to w}(S))$, and so if $R_u^{v \to w}(S) = u'$, then $S \subseteq O_w(u')$. There are only $\binom{|T \cap O_w(u')|}{m}$ choices for $S \subseteq T \cap O_w(u')$ with $|S| = m$, and an identical argument to that of Lemma 3.5 shows that for each such S, there is only one u for which $R_u^{v \to w}(S) = u'$. □

Whenever $v, w \in L_F(X)$ and $E_X(v) \subseteq E_X(w)$, clearly $R_u^{v \to w}(S) \in L(X)$ for any F-sparse set $S \subseteq O_v(u)$; this, along with the use of Lemma 4.1, will be the keys to the counting arguments used to prove our main result for $\mathbb{G}$-subshifts.

Theorem 4.4. *Let X be a $\mathbb{G}-$subshift, μ a measure of maximal entropy of X, $F \Subset \mathbb{G}$, and $v, w \in L_F(X)$. If $E(v) \subseteq E(w)$ then*

$$\mu(v) \leq \mu(w).$$

Proof. Take $\mathbb{G}$, X, μ, F, v, and w as in the theorem, and suppose for a contradiction that $\mu(v) > \mu(w)$. Choose any $\delta \in \mathbb{Q}_+$ with $\delta < \frac{\mu(v) - \mu(w)}{5}$. Let F_n be a Følner sequence satisfying Theorem 2.6. For every $n \in \mathbb{Z}_+$, we define

$$S_n := \{u \in L_{F_n}(X) : |O_v(u)| \geq |F_n|\left(\mu(v) - \delta\right) \text{ and } |O_w(u)| \leq |F_n|\left(\mu(w) + \delta\right)\}.$$

By the pointwise ergodic theorem (applied to $\chi_{[v]}$ and $\chi_{[w]}$), $\mu(S_n) \to 1$, and then by Corollary 2.7,

$$\lim_{n \to \infty} \frac{\log |S_n|}{n} = h_{top}(X). \tag{5}$$

Let $N \in \mathbb{Z}_+^d$ be a number obtained by Lemma 4.1 that is minimal in the sense that if any of the coordinates is decreased then it will not satisfy the property of the lemma.

We note that for every $u \in S_n$, $|O_v(u)| - |O_w(u)| > 3\delta|F_n|$. Therefore, for every $u \in S_n$, there exists $h(u) \in {}^{\mathbb{G}}\!/_{\mathbb{G}_N}$ such that

$$|O_v(u) \cap h(u)| - |O_w(u) \cap h(u)| > \frac{3\delta}{M}|F_n|, \tag{6}$$

where $M = \left|{}^{\mathbb{G}}\!/_{\mathbb{G}_N}\right|$.

For every $u \in S_n$, define $k_n(u) \in \mathbb{N}$ satisfying $|O_v(u) \cap h(u)| \in [k_n(u)|F_n|\frac{\delta}{M}, (k_n(u)+1)|F_n|\frac{\delta}{M}]$.

Using $M = \left|{}^{\mathbb{G}}\!/_{\mathbb{G}_N}\right|$ and the fact that $3 \leq k_n(u) \leq \frac{M}{\delta}$, we may choose $S'_n \subseteq S_n$ with $|S'_n| \geq \frac{|S_n|}{M^2/\delta}$, $h_n \in {}^{\mathbb{G}}\!/_{\mathbb{G}_N}$ and $k_n \in \mathbb{N}$ such that for every $u \in S'_n$ we have $h(u) = h_n$ and $k_n(u) = k_n$. This implies that for every $u \in S'_n$

$$|O_v(u) \cap h_n(u)| \geq (k_n + 1)|F_n|\frac{\delta}{M}, \text{ and hence}$$

$$|O_w(u) \cap h_n(u)| \leq (k_n - 2)|F_n|\frac{\delta}{M} \text{ (using (6))}.$$

By the pigeonhole principle, we may pass to a sequence on which $h_n = h$ and $k_n = k$ are constant, and for the rest of the proof consider only n in this sequence. Let $\varepsilon \in \mathbb{Q}_+$ with $\varepsilon < \frac{\delta}{|F \cdot F^{-1}|}$. For each $u \in S_n'$, we define

$$A_u := \{R_u^{v \to w}(S) : S \subseteq O_v(u) \cap h \text{ and } |S| = \varepsilon\, |F_n|\, /M\}$$

(without loss of generality we may assume $\varepsilon\, |F_n|\, /M$ is an integer by taking a sufficiently large n) .

Since $E_X(v) \subseteq E_X(w)$, we have that $A_u \subset L(X)$. By Lemma 4.2,

$$|A_u| \geq \binom{|O_v(u) \cap h|}{\varepsilon\, |F_n|\, /M} \geq \binom{\delta k |F_n|/M}{\varepsilon\, |F_n|\, /M}.$$

On the other hand, for every $u' \in \bigcup_{u \in S_n} A_u$, we have that

$$|O_w(u') \cap h| \leq \frac{|F_n|}{M} \left((k_n - 2)\delta + \varepsilon |F \cdot F^{-1}| \right) \leq \frac{\delta |F_n|}{M} (k_n - 1).$$

(here, we use $|O_w(u) \cap h(u)| \leq (k_n - 2)|F_n| \frac{\delta}{M}$ plus $|S| = \varepsilon\, |F_n|\, /M$ and the simple fact that a replacement of v by w in u can create at most $|F \cdot F^{-1}|$ new occurrences of w.) Therefore, by Lemma 4.3,

$$|\{u \in S_n' : u' \in A_u\}| \leq \binom{\delta(k_n - 1)|F_n|/M}{\varepsilon\, |F_n|\, /M}.$$

By combining the two inequalities, we see that

$$|L_n(X)| \geq \left| \bigcup_{u \in S_n'} A_u \right| \geq |S_n'| \binom{\delta k_n |F_n|/M}{\varepsilon\, |F_n|\, /M} \binom{\delta(k_n - 1)|F_n|/M}{\varepsilon\, |F_n|\, /M}^{-1}. \tag{7}$$

Now, we take logarithms of both sides, divide by $|F_n|$, and let n approach infinity (along the earlier defined sequence). Then we use the definition of entropy, the inequality $|S_n'| \geq \frac{|S_n|}{M^2/\delta}$, (5), and Stirling's approximation to yield

$$\begin{aligned} h_{top}(X) \geq h_{top}(X) + \frac{\varepsilon}{M} \Bigg[& \left(\frac{\delta k}{\varepsilon} \log \frac{\delta k}{\varepsilon} - \left(\frac{\delta k}{\varepsilon} - 1 \right) \log \left(\frac{\delta k}{\varepsilon} - 1 \right) \right) \\ & - \left(\frac{\delta(k-1)}{\varepsilon} \log \frac{\delta(k-1)}{\varepsilon} - \left(\frac{\delta(k-1)}{\varepsilon} - 1 \right) \log \left(\frac{\delta(k-1)}{\varepsilon} - 1 \right) \right) \Bigg]. \end{aligned}$$

Since the function $x \log x - (x-1)\log(x-1)$ is strictly increasing for $x > 1$, the right-hand side of the above is strictly greater than $h_{top}(X)$, a contradiction. Therefore, our original assumption does not hold and hence $\mu(v) \leq \mu(w)$. □

References

[1] A. Bertrand-Mathis. *Specification, synchronisation, average length.* In International Colloquium on Coding Theory and Applications, Springer, Berlin, Heidelberg, (1986), 86-95.

[2] V. Climenhaga. *Entropy of S-gap shifts.* https://vaughnclimenhaga.wordpress.com/2014/09/08/entropy-of-s-gap-shifts/, Vaughn Climenhaga's Math Blog, posted on September 8, 2014.

[3] V. Climenhaga. *Specification and towers in shift spaces.* arXiv:1502.00931.

[4] V. Climenhaga and D. J. Thompson. *Intrinsic ergodicity beyond specification: β-shifts, S-gap shifts, and their factors.* Israel Journal of Mathematics, 192.2 (2012), 785-817.

[5] V. Climenhaga and D. J. Thompson. *Equilibrium states beyond specification and the Bowen property.* Journal of the London Mathematical Society, 87.2 (2013), 401-427.

[6] E. Glasner, *Ergodic theory via joinings.* American Mathematical Society, No. 101 (2003).

[7] S. Kass and K. Madden. *A sufficient condition for non-soficness of higher-dimensional subshifts.* Proceedings of the American Mathematical Society, 141.11 (2013), 3803-3816.

[8] D. Kerr and H. Li. *Independence in topological and C*-dynamics*. Mathematische Annalen, 338.4 (2007), 869-926.
[9] D. Kerr and H. Li. *Ergodic theory*. Springer Monographs in Mathematics, Springer, (2016).
[10] J. Kułaga-Przymus, M. Lemańczyk, and B. Weiss. *On invariant measures for B-free systems*. Proceedings of the London Mathematical Society, 110.6 (2015), 1435-1474.
[11] D. Kwietniak. *Topological entropy and distributional chaos in hereditary shifts with applications to spacing shifts and beta shifts*. Discrete and Continuous Dynamical Systems, 33 (2013), 2451-2467.
[12] E. Lindenstrauss. *Pointwise theorems for amenable groups*. Inventiones mathematicae, 146.2 (2001), 259-295.
[13] E. Louidor, B. Marcus, and R. Pavlov. *Independence entropy of $\mathbb{Z}^d$-shift spaces*. Acta Applicandae Mathematicae, 126.1 (2013), 297-317.
[14] K. Lau and A. Zame. *On weak mixing of cascades*. Math. Systems Theory, 6 (1972), 307-311.
[15] B. Marcus and R. Pavlov. *An integral representation for topological pressure in terms of conditional probabilities*. Israel Journal of Mathematics, 207.1 (2015), 395-433.
[16] T. Meyerovitch. *Gibbs and equilibrium measures for some families of subshifts*. Ergodic Theory and Dynamical Systems, 33.3 (2013), 934-953.
[17] M. Misiurewicz. *Topological conditional entropy*. Studia Math., 55 (1976), 175-200.
[18] N. Ormes and R. Pavlov. *Extender sets and multidimensional subshifts*. Ergodic Theory and Dynamical Systems, 36.3 (2016), 908-923.
[19] C. Spandl. *Computing the topological entropy of shifts*. MLQ Math. Log. Q., 53.4 (2007), 493-510.
[20] B. Stanley. *Bounded density shifts*. Ergodic Theory and Dynamical Systems, 33.6 (2013), 1891-1928.
[21] K. Thomsen. *On the ergodic theory of synchronized systems*. Ergodic Theory and Dynamical Systems, 26.4 (2006), 1235-1256.
[22] P. Walters. *An introduction to ergodic theory*. Vol. 79. Springer Science & Business Media, (2000).
[23] B. Weiss. *Actions of amenable groups*. Topics in dynamics and ergodic theory., Vol. 310, Cambridge Univ. Press Cambridge, (2003), 226-262.

Felipe García-Ramos, CONACyT & Physics Institute of the Universidad Autónoma de San Luis Potosí, Av. Manuel Nava #6, Zona Universitaria, C.P. 78290, San Luis Potosí, S.L.P., Mexico
Email address: fgramos@conacyt.mx

Ronnie Pavlov, Department of Mathematics, University of Denver, 2390 S. York St., Denver, CO 80208, USA
Email address: rpavlov@du.edu