Measures of maximal entropy of bounded density shifts

Felipe García-Ramos * † Ronnie Pavlov ‡ Carlos Reyes *

Abstract

We find sufficient conditions for bounded density shifts to have a unique measure of maximal entropy. We also prove that every measure of maximal entropy of a bounded density shift is fully supported. As a consequence of this, we obtain that bounded density shifts are surjunctive.

1 Introduction

The concept of entropy is of particular interest when trying to define formally how a system behaves at equilibrium. Given a dynamical system, we say that an invariant measure is a uniform equilibrium state if it achieves the maximal possible entropy. It has been of interest to physicists and mathematicians to determine whether a system has a unique equilibrium state or not. When this happens, mathematicians say the system is intrinsically ergodic and physicists sometimes say that the system does not have a phase transition.

In this paper we are interested in trying to determine if bounded density shifts are intrinsically ergodic. Bounded density shifts were introduced by Stanley in [16]. These subshifts are defined somewhat similarly to the classical β-shifts in that they both are hereditary ([9]), meaning that membership in the shift is preserved under coordinatewise reduction of letters. Whereas β-shifts are 'bounded from above' by a specific sequence coming from a β-expansion, bounded density shifts are restricted by length-dependent bounds on the sums of letters in subwords.

*Institute of Physics, Universidad Autónoma de San Luis Potosí
†Institute of Mathematics, Uniwersytet Jagielloński
‡Department of Mathematics, University of Denver

Stanley proved characterizations of when bounded density shifts are shifts of finite type, sofic, or specified which are remarkably similar to those proved in [15] for β-shifts.

A very effective way of proving that a transitive β-shift is intrinsically ergodic is using the Climenhaga-Thompson decomposition [5] (see Section 2.3), which uses specification of a sub-language. Using this powerful result one can prove that β-shifts (and their factors) are intrinsically ergodic in a few lines (see [5, Section 3.1]).

Proving that bounded density shifts are intrinsically ergodic seems much more mysterious. In this paper we also use Climenhaga-Thompson's theorem to prove a fairly general sufficient condition (Theorem 3.5), though checking the conditions is more complicated than for β-shifts.

It is not difficult to find examples satisfying our sufficient condition (Corollary 3.7), and in fact we do not know if any bounded density subshift fails to satisfy it (Question 3.6). We conjecture that the answer of this question is positive at least for binary subshifts and that every bounded density shift is intrinsically ergodic.

This is not the first paper to study intrinsic ergodicity of bounded density shifts. This has been done in [4, 14]. We are able to prove intrinsic ergodicity under different assumptions than those in those papers. Our hypotheses are also much simpler, and provide proofs of intrinsic ergodicity for new classes of bounded density shifts.

Furthermore, we prove that every measure of maximal entropy of a bounded density shift (with positive entropy) is fully supported. This property is sometimes known as entropy minimality because it is equivalent to having lower topological entropy on every proper subshift. As a consequence of this we prove that synchronized bounded density shifts are always intrinsically ergodic, and we also obtain surjunctivity of bounded density shifts.

Acknowledgments: The authors would like to thank Dominik Kwietniak for rewarding conversations and insights. The first author was supported by the Excellence Initiative Strategic Program of the Jagiellonian University with grant number U1U/W16/NO/01.03. The second author gratefully acknowledges the support of a Simons Collaboration Grant.

2 Definitions and preliminary results

2.1 Subshifts

We devote this section to collect some basic definitions in symbolic dynamics. For a broader introduction to subshifts, languages and their properties, see [10].

Let $\mathcal{A}$ a finite set of symbols. We say that w is a *word* if there exists $n \in \mathbb{N}$ such that $w \in \mathcal{A}^n$ and we denote the *length* of w by $|w|$. Let ϵ denote the empty word, i.e. the word with no symbols. A word u is a *subword* of w if $u = w_k w_{k+1} \dots w_l$ for some $1 \leq k \leq l \leq |w|$. For words $w^{(1)}, \dots, w^{(n)}$, we use $w^{(1)} \dots w^{(n)}$ to represent their concatenation. We say that a word u is a *prefix* of w if $u = w_1 \dots w_k$ for some $1 \leq k \leq |w|$ and a suffix if $u = w_k \dots w_{|w|}$ for some $1 \leq k \leq |w|$, denote by $\text{Suf}(w)$ and $\text{Pre}(w)$ the sets of nonempty suffixes and prefixes respectively for w.

We endow $\mathcal{A}^{\mathbb{Z}}$ with the product topology. When describing a point $x \in \mathcal{A}^{\mathbb{Z}}$ as a sequence, we use a dot to indicate the central position as follows, $x = \dots x_{-1}.x_0 x_1 \dots$, where x_i to represent the ith coordinate of x. We represent intervals of integers with $[i, j]$, and $x_{[i,j]} = x_i x_{i+1} ... x_j$.

The shift map $\sigma : \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ is defined by $\sigma(x) = \dots x_{-1} x_0 . x_1 x_2 \dots$. We say that a set $X \subseteq \mathcal{A}^{\mathbb{Z}}$ is a *subshift* if it is closed and invariant under σ.

For any subshift X,let

$$\mathcal{L}_n(X) = \{w \in \mathcal{A}^n : \exists x \in X \text{ and } i, j \in \mathbb{Z} \text{ s.t } x_{[i,j]} = w\}.$$

We define $\mathcal{L}(X) = \bigcup_{i=0}^{\infty} \mathcal{L}_n(X)$ as the *language* of the subshift X. Given a word w, we define its *cylinder set* as $[w] = \{x \in X : x_{[0,|w|-1]} = w\}$. The cylinder sets form a basis of the topology of $\mathcal{A}^{\mathbb{Z}}$.

2.2 Specification properties

A subshift X is *specified* if there exists $M \in \mathbb{N}$ such that for all $u, w \in \mathcal{L}(X)$, there is a $v \in \mathcal{L}_M(X)$ such that $uvw \in \mathcal{L}(X)$. Following [5], we also define specification for subsets of the language.

Definition 2.1. *Let X be a subshift, $\mathcal{G} \subset \mathcal{L}(X)$ and $n, t \in \mathbb{N}_0$. We say that $\mathcal{G}$ has* ***specification (with gap size*** t***)*** *if for all $m \in \mathbb{N}$ and $w^{(1)}, \dots, w^{(m)} \in \mathcal{G}$, there exist $v^{(1)}, \dots, v^{(m-1)} \in \mathcal{L}_t(X)$ such that*

$$w = w^{(1)} v^{(1)} w^{(2)} v^{(2)} \cdots v^{(m-1)} w^{(m)} \in \mathcal{L}(X).$$

2.3 Measures of maximal entropy

For any subshift X, we denote by $M(X)$ the set of Borel probability measures on X. Equipped with the weak* topology $M(X)$ is a compact topological space.

For any $\mu \in M(X)$ and any finite measurable partition ξ of X, the *entropy of* ξ (with respect to μ), denoted by $H_\mu(\xi)$, is defined by

$$H_\mu(\xi) = -\sum_{A\in\xi} \mu(A) \log \mu(A),$$

where terms with $\mu(A) = 0$ are omitted.

Given a subshift X we denote the σ-invariant Borel probability measures with $M(X, \sigma)$. For $\mu \in M(X, \sigma)$, the *entropy of* μ (for the shift map σ) is defined by

$$h_\mu(X) = \lim_{n\to\infty} \frac{-1}{n} \sum_{w\in\mathcal{L}_n(X)} \mu([w]) \log \mu([w]) = \lim_{n\to\infty} \frac{-1}{n} H_\mu(\xi^{(n)}), \tag{1}$$

where $\xi^{(n)}$ represents the partition of X into cylinder sets from the first n letters, i.e. $\xi^{(n)} = \{[w] \ : \ w \in \mathcal{A}^n\}$.

We note for future reference that $\xi^{(n)} = \bigvee_{i=0}^{n-1} \sigma^{-i}\xi^{(1)}$, where $\xi^{(1)}$ is the partition based on x_0 and $\vee$ is the join of partitions. We will later need to make use of the following basic facts about entropy; for proofs and general introduction to entropy theory, see [18].

Theorem 2.2 (Theorem 4.3 [18]). *For any subshift X, $\mu \in M(X)$, and ξ, η finite partitions of X, $H_\mu(\xi \vee \eta) \le H_\mu(\xi) + H_\mu(\eta)$.*

Theorem 2.3 (Corollary 4.2.1 [18]). *For any subshift X and $\mu \in M(X)$, if ξ is a finite measurable partition of X with k sets, then $H_\mu(\xi) \le \log(k)$, with equality only when $\mu(A) = k^{-1}$ for all $A \in \xi$.*

Theorem 2.4 ([18], p. 184). *For any subshift X, finite measurable partition ξ of X, measures $\mu_i \in M(X)$, and $p_i \ge 0$ $(1 \le i \le n)$ with $\sum_{i=1}^n p_i = 1$, $H_{\sum_{i=1}^n p_i\mu_i}(\xi) \ge \sum_{i=1}^n p_i H_{\mu_i}(\xi)$.*

By the well-known Variational Principle, the supremum of $h_\mu(X)$ over all $\mu \in M(X, \sigma)$ is the *topological entropy* $h_{top}(X)$ of X. For any subshift X, we have that

$$h_{top}(X) = \lim_{n\to\infty} \frac{1}{n} \log |\mathcal{L}_n(X)|. \tag{2}$$

For general topological dynamical systems, the supremum above may not be achieved. However, every subshift has at least one measure of maximal entropy, that is $\nu \in M(X,\sigma)$ achieving the supremum above, meaning that $h_\nu(X) = h_{top}(X)$.

We say a subshift is *intrinsically ergodic* if there is only one (probability) measure of maximal entropy.

Every specified subshift is intrinsically ergodic ([1]). This result has been generalized in several works, including [5] and [13]. Before stating the result we need some extra definitions.

Given a collection of words $\mathcal{D} \subseteq \mathcal{L}(X)$ and $n \geq 1$, we define $\mathcal{D}_n = \mathcal{D} \cap \mathcal{L}_n(X)$. We denote the *growth rate* of $\mathcal{D}$ by

$$h(\mathcal{D}) = \limsup_{n\to\infty} \frac{1}{n} \log |\mathcal{D}_n|. \tag{3}$$

Note that $h(\mathcal{L}(X)) = h_{top}(X)$.

Following [5], we say that $\mathcal{L}(X)$ *admits a decomposition* $\mathcal{C}^p\mathcal{G}\mathcal{C}^s$ for $\mathcal{C}^p, \mathcal{G}, \mathcal{C}^s \subset \mathcal{L}(X)$ if every $w \in \mathcal{L}(X)$ can be written as uvw for some $u \in \mathcal{C}^p$, $v \in \mathcal{G}$, $w \in \mathcal{C}^s$. For such a decomposition, we define the collection of words $\mathcal{G}(M)$ for each $M \in \mathbb{N}$ by

$$\mathcal{G}(M) = \{uvw \,:\, u \in \mathcal{C}^p, v \in \mathcal{G}, w \in \mathcal{C}^s, |u| \leq M, |w| \leq M\}. \tag{4}$$

Theorem 2.5 (Climenhaga and Thompson [5])**.** *Let X be a subshift whose language $\mathcal{L}(X)$ admits a decomposition $\mathcal{C}^p\mathcal{G}\mathcal{C}^s$, and suppose that the following conditions are satisfied:*

1. *$\mathcal{G}$ has specification.*
2. *$h(\mathcal{C}^p \cup \mathcal{C}^s) < h_{top}(X)$.*
3. *For every $M \in \mathbb{N}$, there exists τ such that given $v \in \mathcal{G}(M)$, there exist words u, w with $|u| \leq \tau$, $|w| \leq \tau$ for which $uvw \in \mathcal{G}$.*

Then X is intrinsically ergodic.

Remark. *Using results from [12], Climenhaga explained in a blog post [3] that condition 3 is actually not required to prove uniqueness of the measure of maximal entropy. However, this condition is not difficult to check for bounded density shifts with positive entropy (Lemma 3.4) and so we verify it regardless.*

2.4 Bounded density shifts

Bounded density shifts were introduced in [16].

Let $f : \mathbb{N}_0 \to [0, \infty)$ be a function. We say f is **canonical** if

- $f(0) = 0$,
- $f(m+1) \geq f(m)$ for all $m \geq 0$, and
- $f(m+n) \leq f(m) + f(n)$ for all $n, m \in \mathbb{N}$.

The **bounded density shift** associated to a canonical function, f, is defined as follows:

$$X_f = \left\{ x \in (\mathbb{N}_0)^{\mathbb{Z}} : \forall p \in \mathbb{N} \text{ and } \forall i \in \mathbb{Z} \sum_{r=i}^{i+p-1} x_r \leq f(p) \right\}. \tag{5}$$

Note that X_f is a subshift on the alphabet $\mathcal{A} = \{0, 1, ..., \lfloor f(1) \rfloor\}$.

Actually, bounded density shifts can be defined for any function $f : \mathbb{N}_0 \to [0, \infty)$, but it was shown in [16] that every bounded density shift can be defined by some canonical f.

Definition 2.6. *Let X_f be a bounded density shift, the limit*

$$\lim_{n \to \infty} \frac{f(n)}{n} \tag{6}$$

*is called the **limiting gradient** and is denoted by α.*

The existence of the limit is given by Fekete's lemma and the definition of canonical function; furthermore, the limit is an infimum, and so $f(n) \geq \alpha n$ for all n.

There exist bounded density shifts with $\alpha = 0$ but they are fairly trivial systems where the upper density of non-zero coordinates is always 0. A bounded density shift has positive topological entropy if and only if $\alpha > 0$ (see [9, Theorem 12]) if and only if it is coded (determined by a labeled irreducible graph with possibly countably many vertices) ([16, Theorem 3.1]).

As we mentioned in the previous section, the specification property guarantees intrinsic ergodicity. For bounded density shifts, X_f is specified with specification constant M if and only if 0^M is intrinsically synchronizing (see Definition 3.12 in [16, Theorem 5.1]). Bounded density shifts with positive topological entropy without specification can easily be constructed ([16]).

A subshift X with alphabet $\{0, 1, ..., n\}$ is *hereditary* if every time there is $x \in X$ and $y \in A^{\mathbb{Z}}$ with $y_i \leq x_i$ $\forall i \in \mathbb{Z}$, then $y \in X$. It is not difficult to check that bounded density shifts are hereditary.

3 Intrinsic ergodicity

In this section we fix a bounded density shift X_f with $\alpha > 0$. We define

$$\mathcal{G} = \left\{ w \in \mathcal{L}(X_f) \ : \ \text{if } u \in \text{Pre}(w) \cup \text{Suf}(w), \text{ then } \ \frac{1}{|u|}\sum_{i=1}^{|u|} u_i < \alpha \right\}, \text{ and}$$

$$\mathcal{B} = \mathcal{C}^p = \mathcal{C}^s = \left\{ v \in \mathcal{L}(X_f) \ : \ \frac{1}{|v|}\sum_{i=1}^{|v|} v_i \geq \alpha \right\} \cup \{\epsilon\},$$

where ϵ denotes the empty word.

Lemma 3.1. *The language $\mathcal{L}(X_f)$ admits a decomposition $\mathcal{BGB}$.*

Proof. Let $z \in \mathcal{L}(X_f)$. Define u to be the prefix of z in $\mathcal{B}$ of maximal length (which may be the empty word ϵ), and denote its length by $M \geq 0$. Similarly, define w to be the suffix of z in $\mathcal{B}$ of maximal length (which may be the empty word ϵ), and denote its length by $N \geq 0$.

Suppose that $z_{[M+1,N-1]} \notin \mathcal{G}$. Then by definition, there exists a word $v \in \text{Pre}\left(z_{[M+1,N-1]}\right) \cup \text{Suf}\left(z_{[M+1,N-1]}\right)$ with

$$\frac{1}{|v|}\sum_{i=1}^{|v|} v_i \geq \alpha.$$

Without loss of generality we can assume that $v \in \text{Pre}\left(z_{[M+1,N-1]}\right)$, it means that $v \in \mathcal{B}$, but it is not possible because then uv would be a prefix of z in $\mathcal{B}$ longer than u. Therefore $v \in \mathcal{G}$. $\square$

Lemma 3.2. *The set $\mathcal{G}$ has specification.*

Proof. We will show that $\mathcal{G}$ has specification with gap size $t = 0$. Let $m \in \mathbb{N}$, $w^{(1)}, \ldots, w^{(m)} \in \mathcal{G}$, and $z = w^{(1)} \cdots w^{(m)}$. We compute

$$\begin{aligned}
\sum_{i=1}^{n} z_i &= \sum_{i=1}^{n_1} w_i^{(1)} + \sum_{i=1}^{n_2} w_i^{(2)} + \ldots + \sum_{i=1}^{n_m} w_i^{(m)} \\
&< n_1\alpha + n_2\alpha + \ldots + n_m\alpha \\
&= \alpha\left(\sum_{i=1}^{m} n_i\right) \\
&= \alpha n \\
&\leq f(n).
\end{aligned}$$

This implies that $z \in \mathcal{L}(X_f)$. $\square$

Proposition 3.3. *There exists $\mu \in M(X_f, \sigma)$ with $\sum_{i=0}^{\lfloor f(1) \rfloor} i\mu([i]) \geq \alpha$ and $h(\mathcal{B}) \leq h_\mu(X_f)$.*

Proof. For each $n \in \mathbb{N}$ and $w \in \mathcal{L}_n(X_f) \cap \mathcal{B}$, consider the set:

$$K_n = \{{}^\infty 0.w0^\infty : w \in \mathcal{L}_n(X_f) \cap \mathcal{B}\}.$$

By construction $|K_n| = |\mathcal{L}_n(X_f) \cap \mathcal{B}|$. Let $\nu_n \in M(X_f)$ be the atomic measure concentrated uniformly on the points of K_n, i.e.

$$\nu_n = \frac{1}{|K_n|} \sum_{x \in K_n} \delta_x.$$

Let $\mu_n \in M(X_f)$ be defined by

$$\mu_n = \frac{1}{n} \sum_{j=0}^{n-1} \nu_n \circ \sigma^{-j}.$$

Note that

$$\begin{aligned}
\sum_{i=0}^{\lfloor f(1) \rfloor} i\mu_n([i]) &= \sum_{i=0}^{\lfloor f(1) \rfloor} \frac{i}{n} \sum_{j=0}^{n-1} \nu_n \circ \sigma^{-j}([i]) \\
&= \sum_{i=0}^{\lfloor f(1) \rfloor} \frac{i}{n} \sum_{j=1}^{n} \frac{|\{w \in \mathcal{L}_n(X_f) \cap \mathcal{B} : w_i = i\}|}{|K_n|} \\
&= \frac{1}{|K_n|} \sum_{w \in \mathcal{L}_n(X_f) \cap \mathcal{B}} \left(\frac{1}{n} \sum_{j=1}^{n} w_j \right) \\
&\geq \alpha.
\end{aligned}$$

Since $M(X_f)$ is compact (in the weak* topology), we can choose a subsequence such that

$$\lim_{j \to \infty} \frac{1}{n_j} \log |\mathcal{L}_{n_j}(X_f) \cap \mathcal{B}| = \limsup_{n \to \infty} \frac{1}{n} |\mathcal{L}_n(X_f) \cap \mathcal{B}| = h(\mathcal{B}), \tag{7}$$

and $\mu_{n_j} \to \mu \in M(X_f)$. By the definition of μ_n, it is routine to check that $\mu \in M(X_f, \sigma)$, i.e. μ is σ-invariant.

We will use techniques from the proof of the variational principle in [11] to prove that

$$h_\mu(X_f) \geq \limsup_{n\to\infty} \frac{1}{n} \log |\mathcal{L}_n(X_f) \cap \mathcal{B}| = h(\mathcal{B}). \tag{8}$$

Firstly, since $\sum_{i=0}^{\lfloor f(1) \rfloor} i\mu_{n_j}([i]) \geq \alpha$ and $\mu_{n_j} \to \mu$, we also have that $\sum_{i=0}^{\lfloor f(1) \rfloor} i\mu([i]) \geq \alpha$. Consider the partition given by the alphabet $\xi = \{[0], \dots, [\lfloor f(1) \rfloor]\}$. Since all $w \in \mathcal{L}_{n_j}(X_f) \cap \mathcal{B}$ have equal measure $\nu_{n_j}([w]) = |K_{n_j}|^{-1}$ and all other $w \in \mathcal{A}_j^n$ have $\nu_{n_j}([w]) = 0$, by Theorem 2.3,

$$H_{\nu_{n_j}}\left(\bigvee_{i=0}^{n_j-1} \sigma^{-i}\xi\right) = -\sum_{w\in\mathcal{L}_{n_j}(X_f)\cap\mathcal{B}} \nu_{n_j}([w]) \log \nu_{n_j}([w]) = \log|\mathcal{L}_{n_j}(X_f) \cap \mathcal{B}|. \tag{9}$$

Let $q, n \in \mathbb{N}$ with $1 < q < n$ and define $a(t) = \lfloor \frac{n-t}{q} \rfloor$ for $0 \leq t < q$. Note that $a(0) \geq a(1) \geq \cdots \geq a(q-1)$. For every $0 \leq t \leq q-1$, we define

$$S_t = \{0, 1, \dots, t-1, t + a(t)q, t + a(t)q + 1, \dots, n-1\}.$$

So, for any such t, we can rewrite $\{0, 1, \dots, n-1\}$ as follows

$$\{0, 1, \dots, n-1\} = \{t + rq + i | 0 \leq r < a(t), 0 \leq i < q\} \cup S_t. \tag{10}$$

Observe that

$$t + a(t)q = t + \left\lfloor \frac{n-t}{q} \right\rfloor q \geq t + \left(\frac{n-t}{q} - 1\right) q = t + n - t - q = n - q.$$

Thus, the cardinality of S_t is at most $2q$.

Using (10) we get

$$\bigvee_{i=0}^{n_j-1} \sigma^{-i}\xi = \left(\bigvee_{r=0}^{a(t)-1} \sigma^{-(rq+t)} \bigvee_{i=0}^{q-1} \sigma^{-i}\xi\right) \vee \bigvee_{l\in S_t} \sigma^{-l}\xi. \tag{11}$$

Combining (9), (11) and Theorem 2.2 we obtain

$$\begin{aligned}
\log|\mathcal{L}_{n_j}(X_f) \cap \mathcal{B}| &= H_{\nu_{n_j}}\left(\bigvee_{i=0}^{n_j-1} \sigma^{-i}\xi\right) \\
&\leq \sum_{r=0}^{a(t)-1} H_{\nu_{n_j}}\left(\sigma^{-(rq+t)} \bigvee_{i=0}^{q-1} \sigma^{-i}\xi\right) + \sum_{l\in S_t} H_{\nu_{n_j}}\left(\sigma^{-l}\xi\right) \\
&\leq \sum_{r=0}^{a(t)-1} H_{\nu_{n_j}\circ\sigma^{-(rq+t)}}\left(\bigvee_{i=0}^{q-1} \sigma^{-i}\xi\right) + 2q\log(l).
\end{aligned} \tag{12}$$

For the inequality $\sum_{l\in S_t} H_{\nu_{n_j}}(\sigma^{-l}\xi) \leq 2q\log(l)$ we apply Theorem 2.3. We note that for each $0 \leq t \leq q-1$, we have

$$(a(t)-1)\,q+t \leq \left\lfloor \frac{n-t}{q} - 1 \right\rfloor q+t = n-q. \tag{13}$$

Summing the first term in the last line of (12) over t from 0 to $q-1$, and using that the numbers $\{t+rq : 0 \leq t \leq q-1, 0 \leq r \leq a(t)-1\}$ are all distinct and are all no greater than $n-q$, yields

$$\begin{aligned}
&\sum_{t=0}^{q-1}\left(\sum_{r=0}^{a(t)-1} H_{\nu_{n_j}\circ\sigma^{-(rq+t)}}\left(\bigvee_{i=0}^{q-1}\sigma^{-i}\xi\right)\right) = \sum_{r=0}^{a(0)-1} H_{\nu_{n_j}\circ\sigma^{-(rq)}}\left(\bigvee_{i=0}^{q-1}\sigma^{-i}\xi\right) + \cdots \\
&\cdots + \sum_{r=0}^{a(q-1)-1} H_{\nu_{n_j}\circ\sigma^{-(rq+q-1)}}\left(\bigvee_{i=0}^{q-1}\sigma^{-i}\xi\right) \\
&= \sum_{p=0}^{n_j-1} H_{\nu_{n_j}\circ\sigma^{-p}}\left(\bigvee_{i=0}^{q-1}\sigma^{-i}\xi\right).
\end{aligned} \tag{14}$$

Using (12) and (14) we get

$$q\log|\mathcal{L}_{n_j}(X_f)\cap\mathcal{B}| \leq \sum_{p=0}^{n_j-1} H_{\nu_{n_j}\circ\sigma^{-p}}\left(\bigvee_{i=0}^{q-1}\sigma^{-i}\xi\right) + \frac{2q^2}{n_j}\log(l).$$

Now, we divide by n_j and apply Theorem 2.4 (with $p_i = \frac{1}{n_j}$), to obtain

$$\frac{q}{n_j}\log|\mathcal{L}_{n_j}(X_f)\cap\mathcal{B}| \leq H_{\mu_{n_j}}\left(\bigvee_{i=0}^{q-1}\sigma^{-i}\xi\right) + \frac{2q^2}{n_j^2}\log(l). \tag{15}$$

We will also use that

$$\lim_{k\to\infty} H_{\mu_{n_{j_k}}}\left(\bigvee_{i=0}^{q-1}\sigma^{-i}\xi\right) = H_{\mu}\left(\bigvee_{i=0}^{q-1}\sigma^{-i}\xi\right), \tag{16}$$

which is obtained using the definition of weak* convergence. Then, combining (15) and (16) yields

$$
\begin{aligned}
qh(\mathcal{B}) &= \lim_{k\to\infty} \frac{q}{n_{j_k}} \log |\mathcal{L}_{n_{j_k}}(X_f) \cap \mathcal{B}| \\
&\leq \lim_{k\to\infty} H_{\mu_{n_{j_k}}} \left(\bigvee_{i=0}^{q-1} \sigma^{-i}\xi \right) + \lim_{k\to\infty} \frac{2q^2}{n_{j_k}} \log(l) \\
&= H_\mu \left(\bigvee_{i=0}^{q-1} \sigma^{-i}\xi \right).
\end{aligned}
$$

Now, by definition of $h_\mu(X_f)$,

$$
h(B) \leq \lim_{q\to\infty} \frac{1}{q} H_\mu \left(\bigvee_{i=0}^{q-1} \sigma^{-i}\xi \right) = h_\mu(X_f).
$$

$\square$

Lemma 3.4. *For every $M \in \mathbb{N}$, there exists τ such that given $v \in \mathcal{G}(M)$, there exist words u, w with $|u| \leq \tau$, $|w| \leq \tau$ for which $uvw \in \mathcal{G}$.*

Proof. Let $M \in \mathbb{N}$ and $v \in \mathcal{G}(M)$. This implies that there exist $u', w' \in \mathcal{B}, v' \in \mathcal{G}$ such that $v = u'v'w'$ and $|u'| \leq M, |w'| \leq M$. Choose $u = w = 0^\tau$, with $\tau = \left\lceil \frac{2M\lfloor f(1) \rfloor}{\alpha} \right\rceil$.

Let $z \in \mathrm{Pre}(0^\tau u'v'w'0^\tau)$. Consider the following sets, $N_1 = [1, \tau]$, $N_2 = [\tau + 1, \tau + |u'|] \cup [\tau + |u'v'| + 1, \tau + |u'v'w'|]$ and $N_3 = [\tau + |u'| + 1, \tau + |u'v'|]$. Note that N_2 corresponds to the section where u' and w' appear and N_3 where v' appears. Also, we can assume that $|z| \geq \tau$ (otherwise we are considering that $z \in \mathrm{Pre}(0^\tau)$), then

$$\begin{aligned}
\frac{1}{|z|}\sum_{i=1}^{|z|} z_i = \frac{1}{|z|}&\left(\sum_{i\in N_1\cap[1,|z|]} z_i + \sum_{i\in N_2\cap[1,|z|]} z_i + \sum_{i\in N_3\cap[1,|z|]} z_i\right)\\
= \quad & \frac{1}{|z|}\left(\frac{|N_1\cap[1,|z|]|}{|N_1\cap[1,|z|]|}\sum_{i\in N_2\cap[1,|z|]} z_i \; + \; \frac{|N_3\cap[1,|z|]|}{|N_3\cap[1,|z|]|}\sum_{i\in N_3\cap[1,|z|]} z_i\right)\\
\leq \quad & \frac{1}{|z|}\left(\frac{|N_1\cap[1,|z|]|}{|N_1\cap[1,|z|]|}2M\lfloor f(1)\rfloor + \alpha|N_3\cap[1,|z|]|\right)\\
= \quad & \frac{1}{|z|}\left(|N_1\cap[1,|z|]|\frac{2M\lfloor f(1)\rfloor}{\tau} + \alpha|N_3\cap[1,|z|]|\right)\\
\leq \quad & \frac{1}{|z|}\left(\alpha|N_1\cap[1,|z|]| + \alpha|N_3\cap[1,|z|]|\right)\\
= \quad & \alpha\left(\frac{|N_1\cap[1,|z|]| + |N_3\cap[1,|z|]|}{|z|}\right)\\
\leq \quad & \alpha
\end{aligned}$$

Here, the first inequality holds since $v' \in \mathcal{G}$, the second equality holds because $|N_1\cap[1,|z|]| = \tau$ (using $|z| \geq \tau$), and the second inequality holds since $\tau \geq \frac{2M\lfloor f(1)\rfloor}{\alpha}$.

The proof for $z \in \mathrm{Suf}(0^\tau u'v'w'0^\tau)$ is similar. □

Our main result is the following.

Theorem 3.5. *Let X_f be a bounded density shift. If for every measure of maximal entropy μ, we have that $\sum_{i=0}^{\lfloor f(1)\rfloor} i\mu([i]) \geq \alpha$, then X_f is intrinsically ergodic.*

Proof. If $\alpha = 0$, then since all sequences have frequency 0 of non-0 symbols, the unique invariant measure is the delta measure of ${}^\infty 0^\infty$.

If $\alpha > 0$ we will obtain the result using Theorem 2.5. First note that $\mathcal{B} = \mathcal{C}^p = \mathcal{C}^s$. Using Lemma 3.1 we obtain $\mathcal{L}(X) = \mathcal{C}^p\mathcal{G}\mathcal{C}^s$. Now we will check the numbered hypotheses of Theorem 2.5.

1. Lemma 3.2 gives us that $\mathcal{G}$ has specification.

2. Let μ' be the measure constructed in Proposition 3.3. By hypothesis it cannot be a measure of maximal entropy. Thus, $h(\mathcal{C}^p \cup \mathcal{C}^s) = h(\mathcal{B}) \leq h_{\mu'}(X_f) < h_{top}(X_f)$.

3. We obtain this property using Lemma 3.4.

□

Question 3.6. *Is it true that for every binary bounded density shift we have that $\mu([1]) \geq \alpha$ for every measure of maximal entropy?*

Is it true that for every bounded density shift we have that

$$\sum_{i=0}^{\lfloor f(1) \rfloor} i\mu([i]) \geq \alpha$$

for every measure of maximal entropy?

A simple condition that easily provides several examples is the following.

Corollary 3.7. *Let X_f be a bounded density shift. If $\alpha \geq \sum_{i=1}^{\lfloor f(1) \rfloor} \frac{i}{i+1}$ then X_f is intrinsically ergodic.*

Proof. Using [8, Corollary 4.6] and the fact that bounded density shifts are hereditary we have that for any measure of maximal entropy $\mu([i]) \leq \mu([i-1])$. Since μ is a probability measure this implies that $\mu([i]) \leq 1/(i+1)$. Thus,

$$\sum_{i=1}^{\lfloor f(1) \rfloor} i \cdot \mu([i]) \leq \sum_{i=1}^{\lfloor f(1) \rfloor} \frac{i}{i+1}.$$

We obtain the result using Theorem 3.5. □

Remark. *In particular, every binary bounded density shift with $\alpha \geq 1/2$ is intrinsically ergodic.*

We will now prove a property called entropy minimality for all bounded density shifts for $\alpha > 0$ using results from [8]. We first need some definitions.

Definition 3.8. *A subshift X is **entropy minimal** if every subshift strictly contained in X has lower topological entropy.*

Equivalently, X is entropy minimal if every measure of maximal entropy on X is fully supported.

Definition 3.9. *Let X be a subshift and $v \in \mathcal{L}(X)$. The **extender set** of v in X is defined by*

$$E_X(v) = \{y \in \{0, 1, \ldots, \lfloor f(1) \rfloor\}^{\mathbb{Z}} : y_{(-\infty,0]} v y_{[1,\infty)} \in X\}.$$

Theorem 3.10 (García-Ramos and Pavlov [8]). *Let X be a subshift with $h_{top}(X) > 0$, μ a measure of maximal entropy and $v, w \in \mathcal{L}(X)$. If $E_X(v) \subseteq E_X(w)$ then*

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

Theorem 3.11. *Every bounded density shift (with $\alpha > 0$) is entropy minimal.*

Proof. Let X_f be a bounded density shift, $\mu \in M(X_f, \sigma)$ a measure of maximal entropy and $w \in \mathcal{L}(X_f)$. Since the topological entropy of X_f is positive then $1 \in \mathcal{L}(X_f)$, and $\mu([1]) > 0$ (otherwise $\mu([0]) = 1$ and the entropy cannot be positive). By Poincaré's recurrence theorem, there exists $v' \in \mathcal{L}(X_f)$ for which $\mu([v']) > 0$ and

$$\sum_{i=1}^{|v'|} v'_i > \sum_{i=1}^{|w|} w_i.$$

We can then define v which is coordinatewise less than or equal to w with

$$\sum_{i=1}^{|v|} v_i = \sum_{i=1}^{|w|} w_i.$$

By the fact that X_f is hereditary, $E_{X_f}(v') \subset E_{X_f}(v)$, and so by Theorem 3.10, $\mu([v]) \geq \mu([v']) > 0$.

We want to prove that $E_{X_f}(v) \subseteq E_{X_f}(0^{|v|} w 0^{|v|})$. Let $y \in E_{X_f}(v)$, with $x = y_{(-\infty,0]}.v y_{[1,\infty)} \in X_f$, and $x' = y_{(-\infty,0]}.0^{|v|} w 0^{|v|} y_{[1,\infty)}$. Let $n < m \in \mathbb{Z}$. We consider two cases, when $x'_{[n,m]}$ is a subword of $0^{|v|} w 0^{|v|}$ and when it is not. If $x'_{[n,m]}$ is subword of $0^{|v|} w 0^{|v|}$, then $x'_{[n,m]} \in \mathcal{L}(X_f)$ since $w \in \mathcal{L}(X_f)$ ([16, Lemma 2.3]). Otherwise, there exists $p \in \mathbb{Z}$ such that

$$\sum_{i=n}^{m} x'_i \leq \sum_{i=n+p}^{m+p} x_i \leq f(m-n).$$

This implies that $x'_{[n,m]} \in \mathcal{L}(X_f)$. Thus, $x' \in X_f$, and so $y \in E_{X_f}(0^{|v|} w 0^{|v|})$. Since y was arbitrary, $E_{X_f}(v) \subseteq E_{X_f}(0^{|v|} w 0^{|v|})$. Using Theorem 3.10 we conclude that

$$\mu([w]) \geq \mu([0^{|v|} w 0^{|v|}]) \geq \mu([v]) e^{-h_{top}(X)(|w|-|v|)} > 0.$$

Therefore, μ is fully supported.

$\square$

Definition 3.12. *Let X be a subshift. A word $v \in \mathcal{L}(X)$ is* ***intrinsically synchronizing*** *if $uv, vw \in \mathcal{L}(X)$ then $uvw \in \mathcal{L}(X)$.*

A subshift is ***synchronized*** *if there exists $v \in \mathcal{L}(X)$ such that v is an intrinsically synchronizing word.*

Every entropy minimal synchronized subshift is intrinsically ergodic ([17, 8]) and every synchronized subshift is coded ([7]). Hence, we obtain the following corollary.

Corollary 3.13. *Every synchronized bounded density shift is intrinsically ergodic.*

3.1 Surjunctivity

Another application of entropy minimality is surjunctivity. Given a subshift X, we say $\phi : X \to X$ is a *shift-endomorphism* if it's continuous and it commutes with the shift. If a shift-endomorphism is bijective we say it is a *shift-automorphism*.

A subshift X is said to be *surjunctive* if every injective shift-endomorphism of X is a shift-automorphism. Every full shift is surjunctive ([6, Chapter 3]. The following result is known (e.g. see [2]) but it is not explicitly stated. We write the proof since the argument is simple.

Lemma 3.14. *Every entropy minimal subshift is surjunctive.*

Proof. Let X be a subshift and $\phi : X \to X$ an injective shift-endomorphism. This implies that $\phi(X)$ is a subshift which is topologically conjugate to X. Since topological entropy is conjugacy-invariant, $\phi(X)$ has the same topological entropy as X. If X is entropy minimal then $\phi(X) = X$. $\square$

Using this and Theorem 3.11 we obtain the following.

Corollary 3.15. *Every bounded density shift with positive topological entropy is surjunctive.*

References

[1] Rufus Bowen. Some systems with unique equilibrium states. Mathematical systems theory, 8(3):193–202, 1974.

[2] Tullio Ceccherini-Silberstein, Michel Coornaert, and Hanfeng Li. Expansive actions with specification of sofic groups, strong topological markov property, and surjunctivity. arXiv preprint arXiv:2107.12047, 2021.

[3] Vaughn Climenhaga. An improved non-uniform specification result. https://vaughnclimenhaga.wordpress.com/2022/06/06/an-improved-non-uniform-specification-result/. Accessed: 2023-01-11.

[4] Vaughn Climenhaga and Ronnie Pavlov. One-sided almost specification and intrinsic ergodicity. Ergodic Theory and Dynamical Systems, 39(9):2456–2480, 2019.

[5] Vaughn Climenhaga and Daniel J. Thompson. Intrinsic ergodicity beyond specification: β-shifts, S-gap shifts, and their factors. Israel Journal of Mathematics, 192(2):785–817, 2012.

[6] Michel Coornaert. Cellular Automata and Groups. Springer-Verlag Berlin Heidelberg, 2010.

[7] Doris Fiebig and Ulf-Rainer Fiebig. Covers for coded systems. Contemporary Mathematics, 135:139–179, 1992.

[8] Felipe García-Ramos and Ronnie Pavlov. Extender sets and measures of maximal entropy for subshifts. Journal of the London Mathematical Society, 100(3):1013–1033, 2019.

[9] Dominik Kwietniak. Topological entropy and distributional chaos in hereditary shifts with applications to spacing shifts and beta shifts. Discrete and Continuous Dynamical Systems, 33(6):2451–2467, 2013.

[10] Doug Lind and Brian Marcus. An introduction to Symbolic Dynamics and Coding. Cambridge University Press, 1995.

[11] Michal Misiurewicz. A short proof of the variational principle for a $\mathbb{Z}_+^n$ action on a compact space. Asterisque, 40:147–157, 1976.

[12] Maria Jose Pacifico, Fan Yang, and Jiagang Yang. Existence and uniqueness of equilibrium states for systems with specification at a fixed scale: an improved climenhaga–thompson criterion. Nonlinearity, 35(12):5963, 2022.

[13] Ronnie Pavlov. On controlled specification and uniqueness of the equilibrium state in expansive systems. Nonlinearity, 32(7):2441–2466, 2019.

[14] Ronnie Pavlov. On subshifts with slow forbidden word growth. Ergodic Theory and Dynamical Systems, 42(4):1487–1516, 2022.

[15] Jörg Schmeling. Symbolic dynamics for β-shifts and self-normal numbers. Ergodic Theory Dynam. Systems, 17(3):675–694, 1997.

[16] Brett Stanley. Bounded density shifts. Ergodic Theory and Dynamical Systems, 33(6):1891–1928, 2013.

[17] Klaus Thomsen. On the ergodic theory of synchronized systems. Ergodic Theory and Dynamical Systems, 26(4):1235–1256, 2006.

[18] Peter Walters. An Introduction to Ergodic Theory. Springer, 1975.