

SUBSYSTEMS OF TRANSITIVE SUBSHIFTS WITH LINEAR COMPLEXITY

ANDREW DYKSTRA, NICHOLAS ORMES, AND RONNIE PAVLOV

ABSTRACT. We bound the number of distinct minimal subsystems of a given transitive subshift of linear complexity, continuing work of Ormes and Pavlov [7]. We also bound the number of generic measures such a subshift can support based on its complexity function. Our measure-theoretic bounds generalize those of Boshernitzan [1] and are closely related to those of Cyr and Kra [2].

1. INTRODUCTION

In this work, we study symbolically defined dynamical systems called subshifts. A subshift is defined by a finite set \mathcal{A} (called an alphabet), the (left) shift action σ on $\mathcal{A}^{\mathbb{Z}}$, and a set $X \subset \mathcal{A}^{\mathbb{Z}}$ of sequences that is closed in the product topology and σ -invariant. For convenience, we will refer to a subshift (X, σ) only as X since the dynamics are always understood to come from σ . (See Section 2.1 for more details.)

Given a subshift X , let $c_X(n)$ denote the number of words of length n that appear in X , i.e., the complexity function of X . Assuming that X is transitive and $c_X(n)$ grows linearly, we ask: what is the interplay between $c_X(n)$ and the structure of the sub-dynamical systems of X ? We study this question in both the topological and measure-theoretic categories.

In the topological category, we provide bounds on how large $c_X(n)$ must be in order to accommodate a given number of minimal subsystems in X . If X contains only one minimal subsystem, then, by the Morse-Hedlund Theorem (Theorem 3.2), either X is periodic or $c_X(n) \geq n + 1$ for every n . In [7], Ormes and Pavlov show that, if X is

2010 *Mathematics Subject Classification*. Primary: 37B10; Secondary: 37A25, 68R15.

Key words and phrases. Symbolic dynamics, linear complexity, generic measures.

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

transitive and not minimal, then

$$(1.1) \quad \limsup_{n \rightarrow \infty} (c_X(n) - 1.5n) = \infty.$$

Moreover they show that the bound (1.1) is *sharp* in the sense that the threshold $1.5n$ cannot be increased by any nondecreasing, unbounded function $g : \mathbb{N} \rightarrow \mathbb{N}$; for any such g , there is an example of a transitive non-minimal subshift X where

$$\limsup_{n \rightarrow \infty} (c_X(n) - (1.5n + g(n))) < \infty.$$

For X containing two or more minimal subsystems, in Section 3 we establish the following.

Theorem 1.2. *Let X be a transitive subshift which is the orbit closure of a recurrent point x , where X has $j \geq 2$ proper minimal subsystems, exactly i of which are infinite ($0 \leq i \leq j$). Then the following bounds hold and are sharp:*

- (1) $\limsup_{n \rightarrow \infty} (c_X(n) - (j + i + 1)n) = \infty$, and
- (2) $\liminf_{n \rightarrow \infty} (c_X(n) - (j + i)n) = \infty$.

The notion that the bounds in Theorem 1.2 are *sharp* is the same as for (1.1), namely, that for any nondecreasing unbounded $g : \mathbb{N} \rightarrow \mathbb{N}$, there exists X satisfying the hypotheses of the theorem for which the bounds do not hold when g is added to $(j+i+1)n$ or $(j+i)n$ respectively.

In Section 4 we consider the case of a general (not necessarily recurrent) transitive subshift. We establish bounds on the growth rate of $c_X(n)$ in the special cases where X contains one or two minimal subsystems and then prove the following.

Theorem 1.3. *Let X be a transitive subshift where X has $j \geq 3$ minimal subsystems, exactly i of which are infinite ($0 \leq i \leq j$). Then*

$$\liminf_{n \rightarrow \infty} (c_X(n) - (j + i)n) = \infty.$$

Of course, Theorem 1.3 implies that

$$\limsup_{n \rightarrow \infty} (c_X(n) - (j + i)n) = \infty$$

for transitive X as well, and we show that this bound is sharp in the same sense as above.

Turning our attention to the measure-theoretic category, we consider the well-studied problem of bounding the number of ergodic measures that a given subshift can support. For example, in [1], Boshernitzan shows that if X is minimal and

$$\liminf_{n \rightarrow \infty} (c_X(n) - Kn) = -\infty,$$

then X can support at most $K - 1$ ergodic measures. He also shows, again assuming minimality, that X is uniquely ergodic provided that

$$\limsup_{n \rightarrow \infty} \frac{c_X(n)}{n} < 3.$$

Cyr and Kra, motivated by work of Katok [4] and Veech [9] on interval exchange transformations, extended Boshernitzan's work by considering arbitrary (not necessarily minimal) subshifts and nonatomic generic (not necessarily ergodic) measures [2]. Note that because they are working in such a general setting, their results cannot establish bounds on the number of atomic measures. Indeed, given any subshift X , one can always union X with a fixed point to obtain a new subshift Y with an additional (atomic) ergodic measure, where $c_Y(n) = c_X(n) + 1$.

In this paper we consider transitive systems and obtain bounds on the number of generic measures. In particular, we show the following in Section 5.

Theorem 1.4. *Let $X = \overline{\mathcal{O}(x)}$ be a transitive subshift where x is recurrent and aperiodic. If*

$$\limsup_{n \rightarrow \infty} \frac{c_X(n)}{n} < 3,$$

then X is uniquely ergodic.

The following result is similar to a result obtained in [2]; see Section 5 for more details.

Theorem 1.5. *Let $X = \overline{\mathcal{O}(x)}$ be a transitive subshift where x is not eventually periodic in both directions. If*

$$\liminf_{n \rightarrow \infty} (c_X(n) - gn) = -\infty$$

for $g \in \mathbb{N}$, then X has at most $g - 1$ generic measures.

Note that Theorem 1.5 does not imply Theorem 1.3. Indeed, if X contains g minimal subsystems, then there are least g generic measures on X . Theorem 1.5 would then imply that

$$\liminf_{n \rightarrow \infty} (c_X(n) - gn) > -\infty,$$

whereas Theorem 1.3 gives the stronger conclusion

$$\liminf_{n \rightarrow \infty} (c_X(n) - gn) = \infty.$$

2. PRELIMINARIES

2.1. Subshifts. We recall some basic definitions; for more information, see [5].

A *full shift* is a pair $(\mathcal{A}^{\mathbb{Z}}, \sigma)$ where \mathcal{A} is a finite alphabet, $\mathcal{A}^{\mathbb{Z}}$ has the product of the discrete topology on \mathcal{A} , and $\sigma : \mathcal{A}^{\mathbb{Z}} \rightarrow \mathcal{A}^{\mathbb{Z}}$ is the left-shift defined by $\sigma(x)_i = x_{i+1}$ for each $x = (x_i)_{i \in \mathbb{Z}} \in \mathcal{A}^{\mathbb{Z}}$. A *subshift* is a pair (X, σ) where X is a closed and σ -invariant subset of some $\mathcal{A}^{\mathbb{Z}}$. To conserve notation, we will often refer to the subshift (X, σ) as simply X .

A subshift X is *transitive* if there exists $x \in X$ such that $X = \overline{\mathcal{O}(x)}$, the closure of the orbit $\mathcal{O}(x) = \{\sigma^n(x) : n \in \mathbb{Z}\}$. We call such a point $x \in X$ a *transitive point*. If $X = \overline{\mathcal{O}(x)}$ for every $x \in X$, then X is *minimal*. A transitive subshift X is *periodic* if it has a transitive point x which is periodic, meaning that there exists $p \in \mathbb{Z}$ such that $\sigma^p(x) = x$. Note that a transitive subshift X is periodic if and only if X has finite cardinality.

Given a subshift X , a *word of length n in X* is a block of symbols $w = w_1 w_2 \cdots w_n$ that occurs in some point $x \in X$, i.e., $w = x_i x_{i+1} \cdots x_{i+n-1}$ for some $i \in \mathbb{Z}$. Let $\mathcal{L}_n(X)$ denote the set of all words of length n occurring in some point in X , and $\mathcal{L}(X) = \bigcup_{n=1}^{\infty} \mathcal{L}_n(X)$. The *complexity function* of X is the function $c_X(n) : \mathbb{N} \rightarrow \mathbb{N}$ that gives the cardinality of $\mathcal{L}_n(X)$. If X is transitive, then $X = \overline{\mathcal{O}(x)}$ for some $x \in X$, and in this case $c_X(n)$ is equal to the number of words of length n in x .

For a symbol $a \in \mathcal{A}$ and $n \geq 1$, the expression a^n denotes the word of length n formed by concatenating a with itself n times. Correspondingly, a^∞ denotes the infinite concatenation of a with itself. Depending on the situation, a^∞ may denote a bi-infinite sequence, a left-infinite sequence, or a right-infinite sequence. The choice of meaning should be clear from context.

A point $x \in X$ is *recurrent* if every word in x occurs at least twice (equivalently, infinitely often). If every word in x occurs infinitely often with uniformly bounded gaps between occurrences, then x is *uniformly recurrent*. Note that x is uniformly recurrent if and only if $X = \overline{\mathcal{O}(x)}$ is minimal.

If X and Y are subshifts, then for any continuous $f : X \rightarrow Y$ such that $f \circ \sigma = \sigma \circ f$, there exist $m, a \in \mathbb{N}$ such that for every $x \in X$, $f(x)_0$ is determined by the word $x_{-m} \cdots x_0 \cdots x_a$. In this case f is called an $(m + a + 1)$ -*block map* with *memory m* and *anticipation a* . If such a map f is surjective then it is a *factor map*. In this paper, when we define a factor map f on a subshift X , it is understood to have codomain $f(X)$. In addition, any $(m + a + 1)$ -block map has an

obvious associated action on finite words as well (for any n -letter word w , $f(n)$ has length $n - m - a$); we use f to refer to this function also since usage should always be clear from context.

A word w in $\mathcal{L}(X)$ is *right-special* if there exist $a, b \in \mathcal{A}$ such that $wa, wb \in \mathcal{L}(X)$ with $a \neq b$. Similarly, w is *left-special* if there exist $a, b \in \mathcal{A}$ such that $aw, bw \in \mathcal{L}(X)$ with $a \neq b$. Let $RS_X(n)$ (or just $RS(n)$ if X is understood) denote the set of right-special words in X .

Let x be an element of a subshift X . By the *omega-limit set* of x , we mean the set

$$\omega(x) = \bigcap_{N \geq 1} \overline{\{\sigma^n(x) : n \geq N\}}$$

For any $x \in X$, the set $\omega(x)$ is a closed and shift-invariant subset of X , so is itself a subshift.

We say that a point x in a subshift X is *eventually periodic to the right* if there exist integers $p > 0$ and $N > 0$ such that for all $i > N$, $x_i = x_{i+p}$. Similarly, we say that x is *eventually periodic to the left* if there exist integers $p > 0$ and $N > 0$ such that for all $i < -N$, $x_i = x_{i-p}$.

2.2. Sturmian subshifts. There are several different approaches to defining Sturmian subshifts (see [3] for an introduction). We outline one such approach here.

For any irrational β , define the map $R_\beta : [0, 1) \rightarrow [0, 1)$ by $R_\beta(x) = x + \beta \pmod{1}$. For any $x \in (0, 1)$, define the sequence $s(x) \in \{0, 1\}^{\mathbb{Z}}$ by

$$s_n(x) = \begin{cases} 1 & \text{if } R_\beta^n(x) \in [0, \beta) \\ 0 & \text{if } R_\beta^n(x) \in [\beta, 1). \end{cases}$$

The bi-infinite sequence $s(x)$ is called a *Sturmian sequence* for β . For any two-element set $\{a, b\}$, we call a subshift $X \subset \{a, b\}^{\mathbb{Z}}$ a *Sturmian subshift* if X can be obtained as the orbit closure of a Sturmian sequence with 0 replaced by a and 1 by b .

2.3. Bounds on the Complexity Function. Here we introduce shorthand notation for bounds on the complexity function.

Definition 2.1. Given a function $f : \mathbb{N} \rightarrow \mathbb{R}$, write $c_X(n) \succsim f(n)$ if

$$\limsup_{n \rightarrow \infty} (c_X(n) - f(n)) = \infty.$$

Definition 2.2. Given a function $f : \mathbb{N} \rightarrow \mathbb{R}$, write $c_X(n) \succeq f(n)$ if

$$\liminf_{n \rightarrow \infty} (c_X(n) - f(n)) = \infty.$$

Note that the bounds in Theorem 1.2 can be rewritten using this notation as

- (1) $c_X(n) \asymp (j+i+1)n$ and
- (2) $c_X(n) \asymp (j+i)n$,

and the bound in Theorem 1.3 can be rewritten as $c_X(n) \asymp (j+i)n$.

As mentioned in the introduction, throughout this work we will say that a bound of the form $c_X(n) \asymp f(n)$ (or $c_X(n) \asymp f(n)$) which holds under some hypotheses on X is **sharp** if it fails when any nondecreasing unbounded $g : \mathbb{N} \rightarrow \mathbb{N}$ is added to $f(n)$, i.e., if for any such g there exists a subshift X satisfying the relevant hypotheses for which $c_X(n) \asymp f(n) + g(n)$ (or $c_X(n) \asymp f(n) + g(n)$) is false.

3. TRANSITIVE SYSTEMS WITH A RECURRENT TRANSITIVE POINT

The main goal of this section is to prove Theorem 1.2, which assumes that X contains two or more minimal subsystems. But first we recall some existing results that can be used to treat X containing a single minimal subsystem.

3.1. Single minimal subsystem. If X is itself minimal and not a periodic orbit, then the following two results establish that

$$c_X(n) \geq n + 1$$

for all $n \geq 1$ and that this bound cannot be improved.

Theorem 3.1 ([6]). *If X is a Sturmian subshift, then*

$$c_X(n) = n + 1 \text{ for all } n \geq 1.$$

Theorem 3.2 ([6]). *Let X be any subshift. If there exists $n \geq 1$ such that $c_X(n) \leq n$, then X is a finite set of periodic points.*

Thus the Sturmian subshifts are examples of the lowest complexity subshifts that are not periodic (see [8] for more). If X is transitive and not minimal, then the following (sharp) bound was proved in [7].

Theorem 3.3 ([7]). *Suppose X is a transitive subshift with a recurrent transitive point x . If X is not minimal, then*

$$c_X(n) \asymp 1.5n,$$

and this bound is sharp.

Combining the results of [7] and techniques of the proof of Theorem 1.2, we will be able to establish the following.

Theorem 3.4. *Suppose X is a transitive subshift with a recurrent transitive point x . If X properly contains an infinite minimal subsystem, then*

$$c_X(n) \asymp 2.5n,$$

and this bound is sharp.

We postpone the proof of Theorem 3.4 until after the proof of Theorem 1.2.

3.2. Multiple minimal subsystems. For $j \geq 2$ and $0 \leq i \leq j$, consider the set of subshifts $S = S(j, i)$, where $X \in S(j, i)$ if and only if X is transitive with a recurrent transitive point, and X has j distinct minimal subsystems, exactly i of which are infinite. To prove Theorem 1.2 for $X \in S(j, i)$, we first reduce to the $i = 0$ case via factor maps.

3.2.1. Reduction to the $i = 0$ case. We define a factor map π that maps the j distinct minimal subsystems of X to j distinct points, which are fixed by σ . Lemma 3.5 will provide an inequality between the complexity sequences for X and $\pi(X)$, which will allow us to simply work with $\pi(X)$ moving forward.

Let M_1, \dots, M_j denote the minimal subsystems of X , where, without loss of generality, M_1, \dots, M_i are infinite and M_{i+1}, \dots, M_j are finite. Since the sets M_1, \dots, M_j are pairwise disjoint closed subsets of X , there is an $r \geq 1$ such that the sets $\mathcal{L}_r(M_1), \dots, \mathcal{L}_r(M_j)$ are pairwise disjoint. Fix such a value of r ; pick j distinct symbols a_1, \dots, a_j that do not occur in x ; and define an r -block map ϕ with domain X , memory zero, and anticipation $(r - 1)$ as follows:

$$\phi(y)_q = \begin{cases} a_p & \text{if } y_q \cdots y_{q+r-1} \in \mathcal{L}_r(M_p); \\ y_q & \text{otherwise.} \end{cases}$$

Now pick a symbol b that does not occur in $\phi(x)$, and define a 2-block map ψ with domain $\phi(X)$, memory zero, and anticipation 1 as follows:

$$\psi(z)_q = \begin{cases} b & \text{if } z_q z_{q+1} \in \{a_p s, s a_p\} \text{ for some } p \text{ and } s \neq a_p; \\ y_q & \text{otherwise.} \end{cases}$$

Note that post-composing ϕ with ψ has the effect of ensuring that words of the form a_p^n are always preceded and followed by the ‘‘marker’’ symbol b . Define $\pi = \psi \circ \phi$. Then the minimal subsystems of $\pi(X)$ are simply the one-point sets $\pi(M_p) = \{a_p^\infty\}$. If $X \in S(j, i)$, it follows that $\pi(X) \in S(j, 0)$.

Lemma 3.5. *If $X \in S(j, i)$ with $j \geq 2$ and π is the factor map defined above, then, for every $n > r$, $c_X(n) \geq c_{\pi(X)}(n - r) + in$.*

Proof. Note that for any q -block factor map from a subshift X onto a subshift Y , a word of length m in X determines a word of length $m - q + 1$ in Y . It follows that $c_X(m) \geq c_Y(m - q + 1)$.

Applying this to our situation, since ψ is a 2-block map, $c_{\phi(X)}(n) \geq c_{\pi(X)}(n - 1)$. To complete the proof, it is enough to show that $c_X(n) \geq$

$c_{\phi(X)}(n - r + 1) + in$. For any n , let W_0 denote the set of words in $\phi(X)$ of the form a_p^{n-r+1} where $n \geq r$ and $1 \leq p \leq i$, and let $W_1 = \mathcal{L}_{n-r+1}(\phi(X)) \setminus W_0$. Each word in W_1 has at least one ϕ -preimage in $\mathcal{L}_n(X)$, but each $a_p^{n-r+1} \in W_0$ has at least $n+1$ preimages in $\mathcal{L}_n(M_p)$, since any word in $\mathcal{L}_n(M_p)$ is a preimage of a_p^{n-r+1} , and $c_{M_p}(n) \geq n+1$ for all n , using Theorem 3.2 and the fact that M_p is infinite. \square

Lemma 3.6. *For each $p \in \{1, \dots, j\}$ and $n \in \mathbb{N}$, a_p^n is both right- and left-special in $\pi(X)$.*

Proof. Since $a_p^\infty \in \pi(X)$, a_p^n can be followed by a_p . Assume for a contradiction that a_p^n is not right-special, i.e., it can only be followed by a_p . Since $\pi(X) \neq \{a_p^\infty\}$, there must exist $m \in \mathbb{Z}$ so that

$$\sigma^m(x) = \cdots y_{-3}y_{-2}y_{-1} \cdot a_p^\infty,$$

where $y_{-1} \neq a_p$. Since x is recurrent, $y_{-k} \cdots y_{-k+n} = y_{-1}a_p^n$ for some $k \geq n+2$. Then $-k+n \leq -2$, so there is some smallest $q \geq 1$ such that $y_{-k+n+q} \neq a_p$. Then $y_{-k+q} \cdots y_{-k+n+q-1} = a_p^n$ and $y_{-k+n+q} \neq a_p$, implying that a_p^n is right-special, a contradiction. Our original assumption was then false, i.e., a_p^n is right-special; a symmetric argument shows that a_p^n is left-special. \square

To establish that (1) and (2) from Theorem 1.2 hold, first observe that, by Lemma 3.5, it is enough to show only that (1) and (2) hold for $\pi(X)$. Indeed, if we prove that $c_{\pi(X)}(n) \succ (j+1)n$, then there is a strictly increasing sequence (n_k) such that $c_{\pi(X)}(n_k - r) > (j+1)(n_k - r) + k$ for all k . So by Lemma 3.5,

$$\begin{aligned} c_X(n_k) &\geq c_{\pi(X)}(n_k - r) + in_k \\ &> (j+1)(n_k - r) + k + in_k \\ &= (j+i+1)n_k + k - r(j+1), \end{aligned}$$

which implies $c_X(n) \succ (j+i+1)n$. By a similar argument, $c_{\pi(X)}(n) \succeq jn$ implies $c_X(n) \succeq (j+i)n$.

Because of this, to simplify notation we will replace $\pi(X)$ with X and make the following assumptions:

- **(A1):** $X \in \mathcal{S}(j, 0)$;
- **(A2):** the minimal subsystems of X are the one-point systems $M_1 = \{a_1^\infty\}, \dots, M_j = \{a_j^\infty\}$; and
- **(A3):** for any word in X of the form $a_p s$ or $s a_p$, $s = b$.

3.2.2. *Proof that (1) and (2) from Theorem 1.2 hold when $i = 0$.* Assume X is as above. Then there is a transitive point x for X that is recurrent, and that therefore cannot be both eventually periodic to the

left and eventually periodic to the right. Without loss of generality, assume that x is not eventually periodic to the right. Our proof will involve bounding from below the number of right-special words in X of various lengths (were x eventually periodic to the right, we would instead count left-special words). The following elementary lemma will be used to verify that $c_X(n) \succ (j+1)n$.

Lemma 3.7. *For each $n \geq 2$, $c_X(n) \geq c_X(n-1) + \#RS(n-1)$. Therefore, for $m < n$, $c_X(n) \geq c_X(m) + \sum_{\ell=m}^{n-1} \#RS(\ell)$.*

Proof. For every n , each word of length $n-1$ can be extended to at least one word of length n , while each right-special word of length $n-1$ can be extended to at least two words of length n . This yields $c_X(n) \geq c_X(n-1) + \#RS(n-1)$. Applying this recursively, we obtain the inequality $c_X(n) \geq c_X(m) + \sum_{\ell=m}^{n-1} \#RS(\ell)$ for $m < n$. \square

We begin by considering the set of right-special words provided by Lemma 3.6, which we will call B :

$$B := \{a_p^n : 1 \leq p \leq j \text{ and } n \in \mathbb{N}\}.$$

By just considering the elements of B , we see that $\#RS(n) \geq j$ for all n .

By combining the inequality $\#RS(n) \geq j$ with Lemma 3.7, we can show that the bound $c_X(n) \succ (j+1)n$ would imply $c_X(n) \succ jn$. Indeed, note that $c_X(n) \succ (j+1)n$ implies that there exists an increasing sequence of integers (n_k) such that $c_X(n_k) \geq (j+1)n_k$. Thus for $n > n_k$,

$$\begin{aligned} c_X(n) &\geq c_X(n_k) + \sum_{\ell=n_k}^{n-1} \#RS(\ell) \\ &\geq (j+1)n_k + j(n - n_k) \\ &\geq jn + n_k, \end{aligned}$$

which gives $c_X(n) \succeq jn$.

Thus we will be done if we establish the bound $c_X(n) \succ (j+1)n$. We divide the proof into cases, beginning with the simplest case.

Case (i): For all sufficiently large n , $\#RS(n) \geq j+1$ and there is a strictly increasing sequence (n_k) where $\#RS(n_k) \geq j+2$ for all k .

The assumptions imply that, for $n > n_k$,

$$\sum_{\ell=1}^n \#RS(\ell) \geq (j+1)n + k.$$

This implies that $c_X(n) \succ (j+1)n$, which completes the proof of Theorem 1.2 in Case (i).

We now assume that the hypotheses of Case (i) do not hold. That is, precisely one of the following two conditions holds.

Case (ii): There is a strictly increasing sequence (n_k) such that $\#RS(n_k) = j$, i.e., $RS(n_k) = B$, for all k ;

Case (iii): for all sufficiently large n , $\#RS(n) = j+1$.

Lemma 3.8. *In Case (ii) or (iii), for each k , after re-indexing the minimal subsystems M_1, \dots, M_j if necessary, there exist words $w_{12}^{(k)}$, $w_{23}^{(k)}$, \dots , $w_{j1}^{(k)}$, each of which begins and ends with the symbol b , such that the transitive point $x \in X$ has one of the following forms:*

$$x = \cdots a_1^{\geq n_k} w_{12}^{(k)} a_2^{\geq n_k} w_{23}^{(k)} \cdots a_j^{\geq n_k} w_{j1}^{(k)} a_1^{\geq n_k} w_{12}^{(k)} \cdots$$

$$x = a_1^\infty w_{12}^{(k)} a_2^{\geq n_k} w_{23}^{(k)} \cdots a_j^{\geq n_k} w_{j1}^{(k)} a_1^{\geq n_k} w_{12}^{(k)} \cdots,$$

where each $a_p^{\geq n_k}$ represents a word of the form a_p^n for some $n \geq n_k$.

Proof. Recall that x is a transitive point that is not eventually periodic to the right. For all $p \in \{1, \dots, j\}$ and all $n \geq 1$, a_p^n occurs in x . Since x is not eventually periodic to the right, $a_p^n b$ must occur as well. Since x is recurrent, $a_p^n b$ occurs infinitely many times in x . It follows that ba_p^n also occurs infinitely many times in x .

Assume the hypothesis of Case (ii), and fix $k \geq 1$ and $p \in \{1, \dots, j\}$. Define $w_1 = b$. Since we are in Case (ii), $a_p^{n_k-1} w_1 \notin RS(n_k)$, so there is only one symbol, call it w_2 , that can appear after $a_p^{n_k-1} w_1$. Similarly there is only one symbol that can appear after $a_p^{n_k-2} w_1 w_2$. Continuing in this way, each successive symbol w_i is forced (at least) until some $w_i \cdots w_{i+n_k-1} \in RS(n_k)$, i.e., $w_i \cdots w_{i+n_k-1} = a_r^{n_k}$ for some r . There must exist some smallest i for which this is true, since the omega-limit set $\omega(x)$ must contain one of the minimal subsystems M_1, \dots, M_j . Set $w(p) = w_1 w_2 \cdots w_{i-1}$. Note that $w(p)$ is the only word that begins with b that can follow $a_p^{n_k}$ in x , and that $a_r^{n_k}$ is the only word that can follow $w(p)$. Set $f(p) = r$.

Since $1 \leq p \leq j$ was arbitrary, we obtain a function $f : \{1, \dots, j\} \rightarrow \{1, \dots, j\}$. Since $ba_p^{n_k}$ appears in x for each p , each $p \in \{1, \dots, j\}$ is equal to $f(q)$ for some $q \in \{1, \dots, j\}$, i.e. f is a bijection and thereby a composition of cyclic permutations. If f were the composition of two or more cyclic permutations, then x would not contain all $a_p^{n_k}$ for $1 \leq p \leq j$, a contradiction to transitivity of x . Therefore, f must cyclically permute the elements of $\{1, \dots, j\}$. Re-indexing if necessary, we may assume that $f(p) = p+1$ for $p < j$ and $f(j) = 1$. For

$p \in \{1, \dots, j\}$, set $w_{pf(p)}^{(k)} = w(p)$. After re-indexing again in the case where x is eventually left-asymptotic, it follows that x has one of the two prescribed forms.

Now assume the hypothesis of Case (iii). By the assumptions on x , there must exist a strictly increasing sequence (n_k) such that $ba_1^{n_k}b$ occurs in x for all k . Because $ba_1^{n_k+1}$ also occurs in x , we know that $ba_1^{n_k}$ is right-special for every k . By deleting the first few terms of the sequence (n_k) if necessary, this together with the assumption of Case (3) implies that

$$RS(n_k + 1) = \{ba_1^{n_k}, a_1^{n_k+1}, \dots, a_j^{n_k+1}\}$$

for every k . The word $a_1^{n_k}b \notin RS(n_k + 1)$, so, as in Case (ii), the word $a_1^{n_k}b$ forces a transition word $w(1)$ (whose first symbol is b). But note that the only symbols that can follow $ba_1^{n_k}$ are also a_1 or b , which similarly implies that the word $ba_1^{n_k}b$ forces the same transition word $w(1)$. The rest of the argument from Case (ii) carries through with n_k replaced by $n_k + 1$. We obtain the same possible forms of x as in Case (ii). \square

We proceed assuming the conclusion of Lemma 3.8. For each $p \in \{1, \dots, j\}$ and $k \in \mathbb{N}$, define

$$\ell_{k,p} := \min\{\ell \geq n_k : ba_p^\ell b \in \mathcal{L}(X)\},$$

and set $\ell_{0,p} := 0$. Consider the following set of words:

$$E(k) := \{a_1^{\ell_{k,1}} w_{12}^{(k)} a_2^{\ell_{k,2}}, \dots, a_j^{\ell_{k,j}} w_{j1}^{(k)} a_1^{\ell_{k,1}}\}.$$

It follows from Lemma 3.8 that, if $u \in E(k)$, then any suffix of u is right-special. Some of these suffixes, namely, the *constant* suffixes $a_2^{\ell_{k,2}}, \dots, a_1^{\ell_{k,1}}$, are in B . But many of these suffixes are not constant and therefore not in B .

For example, let $E_1(k)$ denote the set of suffixes of $a_1^{\ell_{k,1}} w_{12}^{(k)} a_2^{\ell_{k,2}}$ that are longer than $\ell_{k,2}$ but no longer than $\ell_{k,2} + \ell_{k,1}$. Then each $u \in E_1(k)$ is right-special and not in B .

In general, for $1 \leq p \leq j$, let $E_p(k)$ denote the set of suffixes of $a_p^{\ell_{k,p}} w_{pq}^{(k)} a_q^{\ell_{k,q}}$ that are longer than $\ell_{k,q}$ but no longer than $\ell_{k,q} + \ell_{k,p}$, where $q = p + 1$ if $2 \leq p < j$ and $q = 1$ if $p = j$. Note that $\#E_p(k) = \ell_{k,p}$, and therefore

$$(3.9) \quad \sum_{p=1}^j \#E_p(k) = \ell_{k,1} + \dots + \ell_{k,j}.$$

Also, each $E_p(k) \cap B = \emptyset$, and, for $r \neq p$, $E_p(k) \cap E_r(k) = \emptyset$.

Proposition 3.10. *For each $k \in \mathbb{N}$,*

$$c_X(\ell_{k,1} + \cdots + \ell_{k,j}) \geq (j+1)(\ell_{k,1} + \cdots + \ell_{k,j}) + (\ell_{k-1,1} + \cdots + \ell_{k-1,j}).$$

Proof. To simplify notation, let $L_k = \ell_{k,1} + \cdots + \ell_{k,j}$. We proceed by induction, where, since $L_0 = 0$, the base case is the assertion that $c_X(L_1) \geq (j+1)L_1$. This assertion follows from Lemma 3.7 together with the following observations:

- for $1 \leq \ell \leq L_1 - 1$, there are j right-special words of length ℓ in B ;
- there are L_1 (distinct) right-special words in $\bigcup_{p=1}^j E_p(1)$, none of which are in B .

Now assume $c_X(L_{k-1}) \geq (j+1)L_{k-1} + L_{k-2}$, and observe that, by Lemma 3.7 together with Equation (3.9),

$$\begin{aligned} c_X(L_k) &\geq c_X(L_{k-1}) + \sum_{\ell=L_{k-1}}^{L_k-1} \#RS(\ell) \\ &\geq [(j+1)L_{k-1} + L_{k-2}] + [j(L_k - L_{k-1}) + L_k] \\ &\geq (j+1)L_k + L_{k-1}. \end{aligned}$$

□

Since $(\ell_{k-1,1} + \cdots + \ell_{k-1,j}) \rightarrow \infty$ as $k \rightarrow \infty$, it follows from Proposition 3.10 that $c_X(n) \succ (j+1)n$, which by our earlier discussion completes the proof of Theorem 1.2 in Cases (ii) and (iii).

3.2.3. Sharpness of Theorem 1.2 when $i = 0$. Here we define a family of examples that demonstrates the sharpness of the bounds in Theorem 1.2. We will consider the $i = 0$ cases first, and then show how to modify the argument for $i > 0$. Let $g : \mathbb{N} \rightarrow \mathbb{N}$ be a given nondecreasing unbounded function.

Define a sequence $\omega = \omega_1\omega_2\omega_3\cdots$ via the rule that $\omega_{2^m k + m} = m$ for $m \geq 1$ and $k \geq 0$, so that $\omega = 1213121412131215\cdots$. Then, define a doubly-infinite sequence

$$x = j^\infty \cdot (1^{n_1^1} 2^{n_1^2} \cdots j^{n_1^j}) (1^{n_2^1} 2^{n_2^2} \cdots j^{n_2^j}) (1^{n_3^1} 2^{n_3^2} \cdots j^{n_3^j}) \cdots,$$

for a doubly-indexed sequence n_k^p satisfying

$$n_1^1 \ll n_1^2 \ll \cdots \ll n_1^j \ll n_2^1 \ll n_2^2 \cdots \ll n_2^j \ll \cdots.$$

(The pattern of n_k^p within x is as follows: the superscript p is always the same as the letter being repeated, and the subscript k comes from ω , in that it is $\omega_1 = 1$ for the first j exponents, then $\omega_2 = 2$ for the next j , then $\omega_3 = 1$ for the next j , and so on.) Note that while ω

is fixed, different sequences (n_k^p) give rise to different sequences x . In other words, x represents a family of examples parametrized by (n_k^p) .

Regardless of choice of (n_k^p) , the transitive subshift $X = \overline{\mathcal{O}(x)}$ has j minimal subsystems: $M_1 = \{1^\infty\}, \dots, M_j = \{j^\infty\}$. Also, we can enumerate the right-special words in X as follows.

All words in the previously-defined set B are again right-special, and any right-special word not in B must end with a word from the set

$$C = \{12^{n_k^2}, 23^{n_k^3}, \dots, j1^{n_k^j} : k \in \mathbb{N}\}.$$

For $1 \leq p \leq j-1$ and for any given k , the word

$$p^{n_k^p}(p+1)^{n_k^{p+1}}$$

is *maximally right-special* in the sense that:

- any suffix of $p^{n_k^p}(p+1)^{n_k^{p+1}}$ is right-special; and
- for any symbol s , the word $sp^{n_k^p}(p+1)^{n_k^{p+1}}$ fails to be right-special.

Let $D = D(k)$ denote the set of words that are not in B and are suffixes of one of these. Each n in an interval of the form $(n_k^{p+1}, n_k^{p+1} + n_k^p]$ corresponds to the length of a word in D . Moreover, we can ensure that these intervals are disjoint by requiring that

$$n_k^3 > n_k^2 + n_k^1, \quad n_k^4 > n_k^3 + n_k^2, \quad \dots, \quad n_k^j > n_k^{j-1} + n_k^{j-2}.$$

Now consider right-special words ending in $j1^{n_k^1}$. In ω , the left-most occurrence of k is ω_{2^k-1} ; any occurrence of a symbol $m \geq k$ in ω is directly preceded by the word $\omega_1 \cdots \omega_{2^k-1-1}$; and any occurrence of the word $\omega_1 \cdots \omega_{2^k-1-1}$ is directly followed by a symbol $m \geq k$. Moreover, any occurrence of the word $k\omega_1 \cdots \omega_{2^k-1-1}$ must be followed by a symbol $m > k$, and any occurrence of the word $m\omega_1 \cdots \omega_{2^k-1-1}$ for $m > k$ must be followed by k .

It follows from these observations about ω that, in x , there is a unique word $u = u(k)$ of length

$$L = L(k) = \sum_{i=1}^j \left(n_{\omega_1}^i + n_{\omega_2}^i + \cdots + n_{\omega_{2^k-1-1}}^i \right),$$

namely,

$$u := x_0 \cdots x_{L-1} = \underbrace{(1^{n_1^1} \cdots j^{n_1^j})(1^{n_2^1} \cdots j^{n_2^j}) \cdots (1^{n_1^1} \cdots j^{n_1^j})}_{2^k-1 \text{ parenthetical blocks}}$$

that directly precedes any occurrence of $1^{n_k^1}$. It follows that $j^{n_k^j}u1^{n_k^1}$ is maximally right-special.

Let $F = F(k)$ denote the set of words that are not in B and are suffixes of $j n_k^j u 1 n_k^1$, and note that $F \cap D = \emptyset$. Then, for each $n \in (n_k^1, n_k^1 + n_k^j + L(k)]$, there is exactly one word of length n in F , and that word is right-special.

We have established the following.

Proposition 3.11. *The set of right-special words in X is*

$$\bigcup_{n=1}^{\infty} RS(n) = B \cup \bigcup_{k=1}^{\infty} (D(k) \cup F(k)).$$

Moreover,

$$\#RS(n) \leq \begin{cases} j & \text{if } n_k^j + n_k^1 + L(k) < n \leq n_{k+1}^1 \text{ for some } k; \\ j + 2 & \text{if } n_k^p < n \leq n_k^p + n_k^{p-1} \text{ for some } k \text{ and } 2 \leq p \leq j; \\ j + 1 & \text{otherwise.} \end{cases}$$

Set $n_0^j = 0$. Since g is nondecreasing and unbounded, we can choose the sequence $n_1^1 < n_1^2 < \dots < n_1^j < n_2^1 < \dots$ to grow fast enough so that for each $k \in \mathbb{N}$, and $p \in \{1, \dots, j\}$, $g(n_k^p)$ is larger than the sum of all n_ℓ^q smaller than n_k^p . More specifically, choose (n_k^p) so that for all $k \geq 1$ and $p \in \{1, \dots, j\}$

$$g(n_k^p) > \sum_{\ell=1}^{k-1} \sum_{q=1}^j n_\ell^q + \sum_{q=1}^{p-1} n_k^q.$$

Note that the right-hand side above provides an upper bound on the number of $n \in [1, n_k^{p+1})$ with $\#RS(n) = j + 2$. Lemma 3.7 then implies that for $n \in [n_k^p, n_k^{p+1})$,

$$c_X(n) \leq (j + 1)n + g(n_k^p) \leq (j + 1)n + g(n).$$

Similarly, if $n \in [n_{k-1}^j, n_k^1)$ for some $k \geq 1$, then

$$c_X(n) \leq (j + 1)n + g(n_{k-1}^j) \leq (j + 1)n + g(n).$$

We've shown that $c_X(n) \leq (j + 1)n + g(n)$ for all n . Since g was arbitrary, this shows that the bound $c_X(n) \asymp (j + 1)n$ is sharp.

To see that the bound $c_X(n) \geq jn$ is sharp, we consider the complexity along the subsequence (n_k^1) . If we choose (n_k^p) to grow fast enough, then for all $k \geq 1$,

$$g(n_{k+1}^1) > (j + 2)(n_k^j + n_k^1 + L(k)).$$

Then Lemma 3.7 and Proposition 3.11 imply that

$$\begin{aligned} c_X(n_{k+1}^1) &\leq c_X(n_k^j + n_k^1 + L(k)) + j(n_{k+1}^1 - (n_k^j + n_k^1 + L(k))) \\ &\leq (j+2)(n_k^j + n_k^1 + L(k)) + jn_{k+1}^1 - j(n_k^j + n_k^1 + L(k)) \\ &\leq g(n_{k+1}^1) + jn_{k+1}^1 - j(n_k^j + n_k^1 + L(k)). \end{aligned}$$

Since $c_X(n) \leq jn + g(n)$ along the sequence n_k^1 and g was arbitrary, the bound $c_X(n) \succeq jn$ is sharp.

3.2.4. Sharpness of Theorem 1.2 when $i > 0$. We now wish to show that the sharp examples constructed in Section 3.2.3 can be extended to the $i > 0$ case. For this, first consider a system $X \in S(j, 0)$ constructed using the form from Section 3.2.3, i.e., X is the orbit closure of a sequence of the form

$$x = j^\infty \cdot (1^{n_1^1} 2^{n_1^2} \dots j^{n_1^j}) (1^{n_2^1} 2^{n_2^2} \dots j^{n_2^j}) (1^{n_3^1} 2^{n_3^2} \dots j^{n_3^j}) \dots$$

Now let $1 \leq i \leq j$, and let S_{j-i+1}, \dots, S_j be arbitrary Sturmian subshifts with alphabets disjoint from each other and from $\{1, \dots, j-i\}$. Our goal is to replace constant strings of symbols from $\{j-i+1, \dots, j\}$ with sequences chosen from S_{j-i+1}, \dots, S_j to create $X' \in S(j, i)$ in such a way that the complexity is increased by exactly in . We begin with an elementary observation about minimal subshifts, which applies in particular to S_{j-i+1}, \dots, S_j .

Lemma 3.12. *Given a minimal subshift S and a word $w \in \mathcal{L}(S)$, there exist arbitrarily long words v with the property that $wvw \in \mathcal{L}(S)$.*

Proof. By minimality, every point in $x \in S$ contains w infinitely many times. Therefore, we can find two instances of w in x occurring at indices that are arbitrarily far apart. \square

To stitch S_{j-i+1}, \dots, S_j into X , we need to impose a further assumption on the sequences $(n_m^i)_{m \in \mathbb{N}}$. By Lemma 3.12, we can recursively define $n_1^1, n_1^2, \dots, n_1^j, n_2^1, n_2^2, \dots, n_2^j, n_3^1, \dots$ in such a way that, associated to each $p \in \{j-i+1, \dots, j\}$ and $k \in \mathbb{N}$, there is a word $w_k^p \in \mathcal{L}_{n_k^p}(S_p)$, and every such w_k^p is both a prefix and suffix of w_{k+1}^p . The proof of sharpness only required rapid growth of the sequence (n_k^p) , and Lemma 3.12 ensures that we may recursively choose n_k^p with arbitrarily rapid growth such that the words w_k^p have the desired conditions. Since each w_k^j is a suffix of w_{k+1}^j , the sequence w_k^j has a left-infinite limit (as $k \rightarrow \infty$), which we denote by w_∞^j .

Now define

$$x' = w_\infty^j \cdot (1^{n_1^1} \dots (j-i)^{n_1^{j-i}} w_1^{j-i+1} \dots w_1^j) (1^{n_2^1} \dots (j-i)^{n_2^{j-i}} w_2^{j-i+1} \dots w_2^j) \dots,$$

the sequence obtained by replacing each p^{n_k} in x by w_k^p , and replacing j^∞ by w_j^∞ . Then each S_p for $p \in \{j-i+1, \dots, j\}$ is an infinite minimal subsystem of $X' := \overline{\mathcal{O}(x')}$, and each $\{p^\infty\}$ for $p \in \{1, \dots, j-i\}$ is a finite minimal subsystem of X' . It is not hard to check that X' contains no other minimal subsystems, and so $X' \in S(j, i)$. It remains to show that $c_{X'}(n) = c_X(n) + in$.

For this, consider the following 1-block factor map ϕ applied to X' . Since the alphabets of the S_p are disjoint, we may map any letter in the alphabet of S_p to p , and leave other letters (for $p \in \{1, \dots, j-1\}$) unchanged. The map ϕ induces a surjection from $\mathcal{L}_n(X')$ to $\mathcal{L}_n(X)$ for all n . We claim that every word in $\mathcal{L}(X)$ which is not constant has only a single ϕ -preimage. To see this, consider $w \in \mathcal{L}(X)$ containing multiple letters. Without loss of generality, we can extend w on the left and right so that w contains some a^{n_a} as a prefix and b^{n_b} as a suffix; if the extension has only one preimage, then of course w did as well. Then, by construction of x' , the only subword of x' mapping to w under ϕ is obtained by replacing every maximal subword of the form p^{n_k} in x by w_k^p for $p \in \{j-i+1, \dots, j\}$. (The only possible ambiguity comes from a^{n_a} and b^{n_b} , but recall that w_k^p is a prefix and suffix of $w_{k'}^p$ for all $k' > k$, and so even if a^{n_a} and/or b^{n_b} were portions of longer runs of a 's or b 's in x , the corresponding word in x' still contains w_m^a and/or w_m^b at those locations if a and/or b are in $\{j-i+1, \dots, j\}$.)

On the other hand, for $k \in \{j-i+1, \dots, j\}$, any constant word of the form k^n has every word in $\mathcal{L}_n(S_k)$ as a preimage, since $S_k \subset X'$, and all words in $\mathcal{L}_n(S_k)$ map to k^n under ϕ . Since $c_{S_k}(n) = n+1$ for all n , this means that all such words have $n+1$ preimages under ϕ . Combining this yields $c_{X'}(n) = 1 \cdot (c_X(n) - i) + (n+1)i = c_X(n) + in$ for all n .

Now, the proof from Section 3.2.3 provides examples of $X \in S(j, 0)$ demonstrating sharpness of the bounds

$$c_X(n) \overline{\succ} (j+1)n \quad \text{and} \quad c_X(n) \succeq jn$$

in the $i = 0$ case. The procedure above yields, for any $0 \leq i \leq j$, $X' \in S(j, i)$ with $c_{X'}(n) = c_X(n) + in$, and so such X' demonstrate the sharpness of the more general bounds (1) and (2) from Theorem 1.2.

3.3. Proof of Theorem 3.4. We sketch the proof of Theorem 3.4 here. Suppose X is a transitive subshift X with a recurrent transitive point x such that X properly contains an infinite minimal subshift M . Because $M \neq X$, there is an $r \geq 1$ such that $\mathcal{L}_r(M) \neq \mathcal{L}_r(X)$. Then

define a factor map π on X such that

$$\pi(z)_k = \begin{cases} 0 & \text{if } z_k \cdots z_{k+r-1} \in \mathcal{L}_r(M) \\ 1 & \text{otherwise.} \end{cases}$$

The image $\pi(x)$ is a recurrent transitive point for $\pi(X)$, and the subshift $\pi(X)$ contains a unique minimal subshift $\pi(M) = \{0^\infty\}$. Since $\pi(X) \neq \{0^\infty\}$, Theorem 3.3 gives $c_{\pi(X)}(n) \succ 1.5n$. Using an estimate as in Lemma 3.5 with $i = 1$ yields $c_X(n) \succ 2.5n$. Finally, Theorem 3.3 also implies that there exist examples of such $\pi(X)$ demonstrating sharpness of $c_{\pi(X)}(n) \succ 1.5n$, and the reader may check that the examples constructed in [7] were similar to those from Section 3.2.3, relying only on rapid growth of an auxiliary sequence. Therefore, the same technique used in Section 3.2.4 yields X demonstrating sharpness of $c_X(n) \succ 2.5n$.

4. GENERAL TRANSITIVE SYSTEMS

Our main theorem in the recurrent case, Theorem 1.2, provides bounds in terms of i and j for subshifts in the sets $S(j, i)$. Our main theorem in the general transitive case, Theorem 1.3, offers similar bounds in terms of sets that we will refer to as $T(j, i)$. Here $T(j, i) \supset S(j, i)$ is the set of all transitive subshifts with $j \geq 1$ minimal subsystems, exactly i of which are infinite where $0 \leq i \leq j$.

4.1. Two or fewer subsystems. We first tackle the cases where $j \leq 2$. When $j = 1$, the results of Theorems 3.1 and 3.2 yield the minimal complexity sequence of $c_X(n) = n + 1$ for X not periodic.

When $j = 2$, the orbit closure of the sequence

$$x = \dots 0000.11111 \dots$$

produces a transitive system in $T(2, 0)$ satisfying $c_X(n) = n + 1$ for all n .

Theorem 4.1. *Let $X \in T(2, i)$ where $i > 0$. Then*

$$\liminf_{n \rightarrow \infty} (c_X(n) - (i + 1)n) > -\infty.$$

Moreover, this bound is optimal in that the $-\infty$ cannot be replaced by any integer.

Proof. Suppose X contains two minimal subsystems M_1 and M_2 . Then there is an $r > 0$ such that $\mathcal{L}_r(M_1) \cap \mathcal{L}_r(M_2) = \emptyset$. Define a factor map π on X such that

$$\pi(z)_k = \begin{cases} i & \text{if } z_k \cdots z_{k+r-1} \in \mathcal{L}_r(M_i), \text{ for } i = 1, 2 \\ 0 & \text{otherwise.} \end{cases}$$

Let $y = \pi(x)$. Then, for all $n \geq 1$, the words 1^n and 2^n occur in y . This means that y is not periodic, so, by the Morse-Hedlund Theorem, $c_Y(n) \geq n + 1$ for all n . Since $i > 0$, we may assume without loss of generality that M_1 is infinite, and so again by Morse-Hedlund, $c_{M_1}(n) \geq n + 1$ for all n . Since π is an r -block map, for $n \geq r$, the word 1^{n-r+1} has at least $c_{M_1}(n) \geq n + 1$ π -preimages, so we obtain

$$c_X(n) \geq c_{\pi(X)}(n - r + 1) + n > 2n - r + 1.$$

If M_2 is also infinite, then both 1^{n-r+1} and 2^{n-r+1} have at least $n + 1$ preimages, and all of those preimages are distinct. Therefore,

$$c_X(n) \geq c_{\pi(X)}(n - r + 1) + 2n > 3n - r + 1.$$

We also claim that the $-\infty$ in Theorem 4.1 cannot be replaced by any integer. We first treat the $i = 1$ case: choose any $N > 3$ and define a Sturmian subshift $Z \subset \{0, 1\}^{\mathbb{Z}}$ created by β where $\frac{1}{N+1} < \beta < \frac{1}{N}$. Then $10^k 1 \in \mathcal{L}(Z)$ if and only if $k = N - 1$ or N . Consider a right-infinite word s in Z , let

$$x = 0^\infty . s,$$

and let $X = \overline{\mathcal{O}(x)}$. Then, for $1 \leq n < N$, there is only one right-special word in $\mathcal{L}_n(X)$, namely, 0^n . Therefore $c_X(N) = N + 1$. Sturmian systems contain exactly one right-special word of length n for every $n \geq 1$. Therefore, when $n \geq N$, there are two right-special words in X : 0^n and a different one that is a subword of s . Lemma 3.7 then implies that $c_X(n) = 2n - N + 1$ for $n \geq N$, so $\liminf c_X(n) - 2n = -N + 1$. Since N could be arbitrarily large, there is no uniform lower bound in the $i = 1$ case of Theorem 4.1.

For the $i = 2$ case, consider two Sturmian subshifts $Z_1, Z_2 \subset \{0, 1\}^{\mathbb{Z}}$ created by distinct $\beta_1, \beta_2 \in (\frac{1}{N+1}, \frac{1}{N})$ for any $N > 3$. Consider a left-infinite word $r \in Z_1$ ending in 1 and a right-infinite word $s \in Z_2$ beginning with 1. Let

$$x = r.s,$$

and let $X = \overline{\mathcal{O}(x)}$.

There are three types of words in x : subwords of r , subwords of s , and words that contain the word 11. For $n \geq 2$ there are exactly $n - 1$ subwords of x that contain 11. For $1 \leq n \leq N$, $\mathcal{L}_n(Z_1) = \mathcal{L}_n(Z_2)$; both of these equal the set of words of length n in $\{0, 1\}^n$ that contain at most one 1. Therefore $c_X(n) = (n - 1) + (n + 1) = 2n$ for $1 \leq n \leq N$.

For $n > N$, there are at most three right-special words: a single word w_1 in $\mathcal{L}_n(Z_1)$ that can be extended in two ways in $\mathcal{L}(Z_1)$, a single word w_2 in $\mathcal{L}_n(Z_2)$ that can be extended in two ways in $\mathcal{L}(Z_2)$, and the n -letter suffix w_3 of r . Therefore, by Lemma 3.7, $c_X(n) = 3n - N$

for $n > N$, and so $\liminf c_X(n) - 3n = -N$, implying that there is no uniform lower bound in the $i = 2$ case of Theorem 4.1. \square

4.2. Three or more subsystems. We now proceed with the proof of Theorem 1.3, which gives the bound $c_X(n) \succeq (j + i)n$ for $X \in T(j, i)$ with $j \geq 3$. For such an X , let $x \in X$ be a transitive point. Consider the same r -block factor map π as constructed in Section 3.2.1. Then $\pi(X) \in T(j, 0)$ for $j \geq 3$. Set $y = \pi(x)$ and note that y is a transitive point for $\pi(X)$. We also note that y cannot be eventually periodic in both directions, or else $\pi(X)$ would have at most two minimal subsystems (the periodic alpha-limit and omega-limit sets of y), a contradiction.

We then assume without loss of generality that y is not eventually periodic to the right. (If y were eventually periodic to the right, then below we would consider left-special, as opposed to right-special words, and arrive at the same conclusion.)

For each $p \in \{1, \dots, j\}$ and $n \in \mathbb{N}$, we claim a_p^n is right-special. Indeed, a_p^{n+1} must occur in y for all p since $\{a_p^\infty\}$ is a minimal subsystem of $\pi(X)$. Since y is not eventually periodic to the right, the word $a_p^n b$ must also occur in y .

Therefore $c_{\pi(X)}(n) \geq jn$ for all $n \geq 1$. By Lemma 3.7, we may establish $c_{\pi(X)}(n) \succeq jn$ by finding an infinite set of n for which $\#RS_{\pi(X)}(n) \geq j+1$. Fix p so that $\{a_p^\infty\}$ is in the omega-limit set $\omega(y)$. Then there exist infinitely many n for which $ba_p^n b$ occurs in y . Recall that ba_p^{n+1} occurs in y for all $n \geq 1$, so $ba_p^n \in RS_{\pi(x)}(n+1)$, i.e. $\#RS_{\pi(x)}(n+1) \geq j+1$, for infinitely many n .

We then have $c_{\pi(X)}(n) \succeq jn$. Finally, as in Lemma 3.5, for every p with $\pi^{-1}\{a_p^\infty\}$ infinite, the number of π -preimages of a_p^{n-r+1} is at least n , establishing

$$c_X(n) \succeq (j + i)n.$$

Remark 4.2. The proof above also works if $X \in S(2, i)$ has a transitive point x such that $\pi(x)$ is not both eventually periodic to the right and eventually periodic to the left.

Theorem 4.3. *Let $X \in T(j, i)$ where $j \geq 3$. Then the bound*

$$c_X(n) \overline{\succ} (j + i)n$$

holds and is sharp.

Proof. The bound itself follows immediately from Theorem 1.3. To see that the bound is sharp, let $g : \mathbb{N} \rightarrow \mathbb{N}$ be any nondecreasing unbounded function. Consider a point of the form

$$(4.4) \quad x = 0^\infty . 1^{n_1} 2^{n_2} \dots (j-1)^{n_{j-1}} 1^{n_j} 2^{n_{j+1}} \dots (j-1)^{n_{2j-2}} \dots ,$$

where $n_1 \ll n_2 \ll n_3 \ll \dots$, and set $X = \overline{\mathcal{O}(x)}$.

For any $n \geq 1$, $0^n, 1^n, \dots, (j-1)^n$ are all right-special in X . No word in X containing 01 is right-special, nor is any word that contains three distinct symbols. The only other right-special words are of the form

$$p^m(p+1)^{n_k} \text{ where } k \equiv (p+1) \pmod{j-1}, 1 \leq m \leq n_{k-1}$$

or

$$(j-1)^m 1^{n_k} \text{ where } k \equiv 1 \pmod{j-1}, 1 \leq m \leq n_{k-1}.$$

In other words, we have $\#RS(n) = j+1$ only for $n \in (n_k, n_k + n_{k-1}]$. Therefore, if $n \in [n_k, n_{k+1})$, the number of right-special words of length

less than n which are not of the form a^n is at most $\sum_{i=1}^{k-1} n_i$. If the sequence (n_k) grows sufficiently fast, then $g(n_k) > \sum_{i=1}^{k-1} n_i$, and then

Lemma 3.7 implies that

$$c_X(n) < jn + g(n)$$

for all n .

For the $T(j, i)$ bound where $j \geq 3$ and $i > 0$, consider the family of examples obtained by replacing the blocks $0^\infty, 1^{n_1}, 2^{n_2}, \dots, (i-1)^{n_{i-1}}, 1^{n_j}, 2^{n_{j+1}}, \dots, (i-1)^{n_{j+i-2}}, \dots$ with blocks from Sturmian sequences as in Section 3.2.4. Then the estimate in Lemma 3.5 gives the sharpness of the bound. \square

5. GENERIC MEASURES

We say that a point x in a subshift X is *generic* for a measure μ if the measures $\nu_n(x) := \frac{1}{n} \sum_{i=0}^{n-1} \delta_{\sigma^i x}$ converge to μ in the weak topology, i.e., if

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} f(\sigma^i(x)) = \int f d\mu$$

for all continuous $f : X \rightarrow \mathbb{R}$. The pointwise ergodic theorem implies that whenever μ is ergodic, μ -almost every $x \in X$ is generic for μ . We say that a measure μ is a *generic measure* on X if there exists $x \in X$ that is generic for μ . All ergodic measures are clearly generic by the pointwise ergodic theorem, but generic measures are not necessarily ergodic; for instance, it is easily checked that

$$x = 0^\infty.011000111100000111111\dots$$

is generic for $\mu = \frac{\delta_{0^\infty} + \delta_{1^\infty}}{2}$.

Our goal is to provide bounds on the number of generic measures a transitive subshift can support. One of the more general results in this vein is the following theorem of Cyr and Kra, which does not assume transitivity of the subshift, but also does not control for the number of atomic measures supported on periodic subshifts.

Theorem 5.1 ([2]). *Suppose X is a subshift, and there exists $k \geq 3$ such that*

$$\limsup_{n \rightarrow \infty} \frac{c_X(n)}{n} < k.$$

If X has a generic measure μ and there is a generic point $z \in X$ for μ such that the subshift $Z = \overline{\mathcal{O}(z)}$ is not uniquely ergodic, then X has at most $k - 2$ distinct, nonatomic, generic measures.

Combining the Cyr-Kra result above with others in this paper we obtain the same conclusion under a different hypothesis.

Theorem 5.2. *Suppose X is a transitive subshift which is the orbit closure of a recurrent point and there exists $k \geq 3$ such that*

$$\limsup_{n \rightarrow \infty} \frac{c_X(n)}{n} < k.$$

Then X has at most $k - 2$ distinct, nonatomic, generic measures.

Proof. Let X be a transitive subshift with a recurrent transitive point. Fix a generic measure μ for X , and let z be a generic point for μ . If $\overline{\mathcal{O}(z)}$ is not uniquely ergodic, then we are done by Theorem 5.1. Thus we may assume that $Z = \overline{\mathcal{O}(z)}$ is uniquely ergodic. Since Z is uniquely ergodic, it follows that every point $x \in Z$ is generic for μ (see, for instance, [10]).

Now assume for a contradiction that X has $(k - 1) \geq 2$ generic measures μ_1, \dots, μ_{k-1} , with respective generic points z_1, \dots, z_{k-1} . Then by the preceding paragraph, we obtain $(k - 1)$ disjoint subsystems Z_1, \dots, Z_{k-1} such that for all $x \in Z_i$, x is generic for μ_i . Each Z_i contains a distinct minimal subsystem. Therefore, by Theorem 1.2,

$$\limsup_{n \rightarrow \infty} (c_X(n) - kn) = \infty,$$

a contradiction. □

In general, there does not appear to be a simple way to combine our results with those of [2] in order to use an upper bound on $\limsup_{n \rightarrow \infty} \frac{c_X(n)}{n}$ to bound the number of generic (not necessarily nonatomic) measures in the transitive case. However, with some effort, we are able to show

that in the case where $k = 3$, i.e., when X is transitive with a recurrent transitive point and

$$\limsup_{n \rightarrow \infty} \frac{c_X(n)}{n} < 3,$$

then X is uniquely ergodic; this is our Theorem 1.4.

5.1. Proof of Theorem 1.4. If X is minimal, then this is Theorem 1.5 from [1]. If X is not minimal, then it cannot be the case that X contains two or more minimal subsystems, since our Theorem 1.2 would imply that $c_X(n) \succ 3n$, which would contradict $\limsup_{n \rightarrow \infty} (\frac{c_X(n)}{n}) < 3$.

So X contains a unique minimal subsystem M_1 . By Theorem 1.5 from [1], there is a unique (ergodic) measure μ supported on M_1 . Consider the factor map π (as defined in sub-section 3.2.1); then μ pushes forward under π to $\delta_{a_1^\infty}$ in $\pi(X)$, and it is the only measure which does so. So, if we are able to prove that $\pi(X)$ is uniquely ergodic, then its unique measure is $\delta_{a_1^\infty}$, which implies that μ is the unique measure on X .

Since the hypotheses of Theorem 1.4 are preserved under application of a factor map, we can assume without loss of generality that X has a unique minimal subsystem $\{a_1^\infty\}$. We can further reduce (by applying a 1-block factor map sending a_1 to 0 and all other letters to 1) to the case where $X \subseteq \{0, 1\}^{\mathbb{Z}}$ and that X has unique minimal subsystem $\{0^\infty\}$. Toward a contradiction, suppose that such an X has an ergodic $\mu \neq \delta_{0^\infty}$.

Lemma 5.3. *Let $(n_k) \subseteq \mathbb{N}$ be a strictly increasing sequence and x a generic point for μ . Then for all sufficiently large k , $x_{[0, \infty)}$ has the form*

$$x_{[0, \infty)} = .w_0^{(k)} 0^{\geq n_k} w_1^{(k)} 0^{\geq n_k} w_2^{(k)} \dots,$$

where every $w_i^{(k)}$ begins and ends with 1 and does not contain 0^{n_k} . Moreover, $\frac{|w_0^{(k)}|}{n_k} \rightarrow \infty$.

Proof. We first note that since x is generic for $\mu \neq \delta_{0^\infty}$, $x_{[0, \infty)}$ contains infinitely many 1s. Moreover, since 0^∞ is the only minimal subsystem of X , the omega-limit set of x must contain 0^∞ , i.e. $x_{[0, \infty)}$ contains 0^n for arbitrarily large n . Therefore, $x_{[0, \infty)}$ has the claimed form, and it remains only to show that $\frac{|w_0^{(k)}|}{n_k} \rightarrow \infty$.

By genericity,

$$\begin{aligned}
 \mu([0]) &= \int \chi_{[0]} d\mu \\
 &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} \chi_{[0]}(\sigma^i(x)) \\
 &= \lim_{N \rightarrow \infty} \frac{\# \text{ zeros in } x_0 \cdots x_{N-1}}{N}.
 \end{aligned}$$

Observe that $\mu([0]) < 1$ since we are assuming $\mu \neq \delta_{0^\infty}$. It follows that $|w_0^{(k)}| \rightarrow \infty$. Therefore, if we let $r_k = |w_0^{(k)}|$ and z_k denote the number of zeros in $w_0^{(k)}$, then

$$\lim_{k \rightarrow \infty} \frac{z_k}{r_k} = \mu([0]) \quad \text{and} \quad \lim_{k \rightarrow \infty} \frac{z_k + n_k}{r_k + n_k} = \mu([0]).$$

Through some algebraic manipulation,

$$0 = \lim_{k \rightarrow \infty} \left(\frac{z_k}{r_k} - \frac{z_k + n_k}{r_k + n_k} \right) = \lim_{k \rightarrow \infty} \left(\left(\frac{n_k}{n_k + r_k} \right) \left(\frac{z_k - r_k}{r_k} \right) \right).$$

Now,

$$\lim_{k \rightarrow \infty} \left(\frac{z_k - r_k}{r_k} \right) = \lim_{k \rightarrow \infty} \left(\frac{z_k}{r_k} - 1 \right) = \mu([0]) - 1 \neq 0,$$

so $\lim_{k \rightarrow \infty} \left(\frac{n_k}{n_k + r_k} \right) = 0$, which implies that $\frac{r_k}{n_k} \rightarrow \infty$. \square

Lemma 5.4. *If x is a sequence which is not eventually periodic to the right, $n \in \mathbb{N}$, and w is a subword of x with length at least $c_X(n) + n$, then w contains an n -letter right-special word.*

Proof. Assume that x and w are as in the lemma. By assumption, w contains more than $c_X(n)$ subwords of length n , and therefore one is repeated, call it u . If no n -letter subword of w is right-special, then for every n -letter subword of w , there is only one choice of a letter which may follow it in x . However, this would mean that the portion of x between the two occurrences of u would have to repeat indefinitely to the right, a contradiction to the assumption that x is not eventually periodic to the right. Therefore, w contains an n -letter right-special word. \square

Lemma 5.5. *There does not exist a strictly increasing sequence (n_k) such that $\#RS(n_k) = 1$ for all k .*

Proof. If such a sequence (n_k) exists, then for every k , the only right-special word of length n_k is 0^{n_k} . Let x be generic for μ . We note that x

cannot be right-asymptotic to a periodic point: if that periodic point were 0^∞ , then by genericity $\mu = \delta_{0^\infty}$; and if it were not 0^∞ , then X would contain a minimal subsystem other than $\{0^\infty\}$.

Consider $w_0^{(k)}$ as defined in Lemma 5.3. We note that $w_0^{(k)}$ cannot contain a right-special word of length n_k , since by definition it would not be 0^{n_k} , a contradiction. Therefore, by Lemma 5.4, $|w_0^{(k)}| < c_X(n_k) + n_k$.

Next we observe that there exists $K \in \mathbb{N}$ such that $c_X(n) < Kn$ for all n . To see this, note that, since $\limsup_{n \rightarrow \infty} (\frac{c_X(n)}{n}) < 3$, there exists $N \in \mathbb{N}$ such that $c_X(n) < 3n$ for all $n \geq N$. And for $1 \leq n \leq N-1$, we can simply let $K(n) \in \mathbb{N}$ be any value such that $K(n) > c_X(n)$. Then just define

$$K := \max\{3, K(1), \dots, K(N-1)\}.$$

Then $|w_0^{(k)}| < c_X(n_k) + n_k < (K+1)n_k$ for all k , which contradicts the conclusion from Lemma 5.3 that $\frac{|w_0^{(k)}|}{n_k} \rightarrow \infty$. □

Following [1], we define a strictly increasing sequence (n_k) to be *logarithmically syndetic* if there exists M such that $\frac{n_{k+1}}{n_k} < M$ for all k .

Lemma 5.6. *There is a logarithmically syndetic sequence (n_k) such that $\#RS(n_k) \leq 2$ for all k . Moreover, if $RS(n_k) = \{0^{n_k}, b_k\}$ for some k and $b_k \neq 0^{n_k}$, then b_k cannot be followed by three or more symbols.*

Proof. Let $\varepsilon = 3 - \limsup_{n \rightarrow \infty} (\frac{c_X(n)}{n})$. Then there exists N such that $\frac{c_X(n)}{n} < 3 - \varepsilon/2$ for all $n \geq N$. We claim that any interval of the form $[i, (\frac{6}{\varepsilon}) \cdot i]$ where $i \geq N$ must contain a value of n such that $c(n+1) - c(n) \leq 2$. The statement of Lemma 5.6 then follows immediately by Lemma 3.7.

To prove the claim, let $i \geq N$ and suppose $c_X(n+1) - c_X(n) \geq 3$ for all $n \in [i, j]$, where $j = (\frac{6}{\varepsilon}) \cdot i$. Then

$$\begin{aligned} c_X(j) &\geq c_X(i) + 3(j-i) \\ &> 3(j-i) \\ &= j(3 - \varepsilon/2). \end{aligned}$$

But $\frac{c_X(j)}{j} < 3 - \varepsilon/2$ since $j \geq N$. □

Let (n_k) be a fixed logarithmically syndetic sequence as given by Lemma 5.6, and let $M > 1$ be a constant such that $\frac{n_{k+1}}{n_k} < M$. By Lemma 5.5, by deleting a finite number of terms in (n_k) if necessary,

we can assume that $\#RS(n_k) = 2$ for all k . Let b_k denote the only right-special word of length n_k other than 0^{n_k} . By passing to a subsequence and increasing M to $10M$ if necessary, we can assume that

$$(5.7) \quad 10 < \frac{n_{k+1}}{n_k} < M \quad \text{for all } k.$$

Since $\frac{|w_0^{(k)}|}{n_k} \rightarrow \infty$ and $\limsup_{k \rightarrow \infty} \frac{c_X(n_k)}{n_k} \leq \limsup_{n \rightarrow \infty} \frac{c_X(n)}{n} < 3$, we see that $|w_0^{(k)}| > c_X(n_k) + n_k$ for sufficiently large k . By Lemma 5.4, $w_0^{(k)}$ contains a right-special word of length n_k , which cannot be 0^{n_k} , so $w_0^{(k)}$ contains at least one occurrence of b_k .

Define a_k to be the shortest (possibly empty) word such that

$$(5.8) \quad w_0^{(k)} = a_k b_k r_k$$

for some word r_k , and define d_k to be the shortest word such that

$$(5.9) \quad w_0^{(k)} = \ell_k b_k d_k$$

for some word ℓ_k . Therefore we can write

$$w_0^{(k)} = a_k b_k u_k d_k$$

for some (possibly empty) word u_k . Observe that $|a_k|, |d_k| < c_X(n_k) + n_k < 4n_k$ for all sufficiently large k , since otherwise, by Lemma 5.4, a_k or d_k would contain b_k , contradicting minimality of length in their definitions. By definition, b_k is a suffix of $a_k b_k u_k$, but since $\frac{|w_0^{(k)}|}{n_k} \rightarrow \infty$, for sufficiently large k we can assume that

$$\begin{aligned} |w_0^{(k)}| &> 13n_k \\ &= 4n_k + n_k + 4n_k + 4n_k \\ &> |a_k| + |b_k| + |d_k| + c_X(n_k) + n_k, \end{aligned}$$

which implies that $|u_k| > c_X(n_k) + n_k$. Therefore, by Lemma 5.4, u_k contains at least one copy of b_k .

Remark 5.10. Let $p > 0$ be an arbitrary positive integer. As above, for sufficiently large k we have

$$\begin{aligned} |w_0^{(k)}| &> (13 + 4p)n_k \\ &> |a_k| + |b_k| + |d_k| + p(c_X(n_k) + n_k). \end{aligned}$$

Applying Lemma 5.4, we see that for sufficiently large k , u_k contains at least p copies of b_k .

Observe that b_k is a suffix of $a_k b_k u_k$, and b_k is the only word in $RS(n_k)$ that appears in $w_0^{(k)}$. Also, a_k and d_k are the *shortest* words that satisfy

(5.8) and (5.9), and (by Lemma 5.6), b_k can only be followed by two distinct symbols. Therefore $w_0^{(k)}$ must have the form

$$(5.11) \quad w_0^{(k)} = a_k b_k (e_k b_k)^{m_k} d_k,$$

where e_k is the shortest (possibly empty) word such that $w_0^{(k)} = a_k b_k e_k b_k q_k$ for some word q_k .

In the form (5.11), observe (by Lemma 5.4) that

$$(5.12) \quad |e_k| < c_X(n_k) + n_k < 4n_k$$

for all sufficiently large k since e_k does not contain any right-special word of length n_k . Also, $m_k \rightarrow \infty$ by Remark 5.10. Therefore, by deleting a finite number of terms from the beginning of (n_k) if necessary, we can assume that

$$(5.13) \quad m_k > M^2 + 1 \quad \text{for all } k \geq 1.$$

Observe that the first symbols of e_k and d_k must be different: if they were the same, then, since b_k is the only word in $RS(n_k)$ that appears in $w_0^{(k)}$, we would have $e_k = d_k$ and $x_{[0,\infty)} = a_k b_k (e_k b_k)^\infty$. As noted, for example, in the proof of Lemma 5.5, this cannot happen because x cannot be eventually periodic to the right.

Now observe that all suffixes of $(e_k b_k)^{m_k-1}$ of length at least n_k are right-special: this is due to the fact that any such suffix can be followed by either e_k or d_k , and, as we just noted, the first symbols of e_k and d_k must be different. By construction, none of these suffixes have the form 0^n . Therefore $\#RS(n) \geq 2$ for each $n \in (n_k, (m_k - 1) \cdot |e_k b_k|]$.

Lemma 5.14. *For all k , both b_{k+1} and b_{k+2} are suffixes of $(e_k b_k)^{m_k-1}$.*

Proof. Using (5.7) and (5.13), observe that

$$\begin{aligned} |b_{k+1}| &= n_{k+1} \\ &< M \cdot n_k \\ &< (m_k - 1) \cdot n_k \\ &\leq (m_k - 1) \cdot |e_k b_k|. \end{aligned}$$

Therefore $|b_{k+1}| \in (n_k, (m_k - 1) \cdot |e_k b_k|)$. And b_{k+1} is the only right-special word of length n_{k+1} other than $0^{n_{k+1}}$, which implies that b_{k+1} is a suffix of $(e_k b_k)^{m_k-1}$.

Next, observe that by (5.7) and (5.13),

$$\begin{aligned}
 |b_{k+2}| &= n_{k+2} \\
 &< M \cdot n_{k+1} \\
 &< M^2 \cdot n_k \\
 &< (m_k - 1) \cdot n_k \\
 &\leq (m_k - 1) \cdot |e_k b_k|.
 \end{aligned}$$

Therefore b_{k+2} is also a suffix of $(e_k b_k)^{m_k-1}$. \square

Lemma 5.15. *For all k , $e_k b_k e_k b_k$ is a suffix of b_{k+1} .*

Proof. By Lemma 5.14, b_{k+1} is a suffix of $(e_k b_k)^{m_k-1}$, so it is enough to show that $|b_{k+1}| \geq |e_k b_k e_k b_k|$. To see this, observe that, using (5.7) and (5.12),

$$\begin{aligned}
 |b_{k+1}| &= n_{k+1} \\
 &> 10n_k \\
 &= 4n_k + n_k + 4n_k + n_k \\
 &> |e_k| + |b_k| + |e_k| + |b_k| \\
 &= |e_k b_k e_k b_k|.
 \end{aligned}$$

\square

Proposition 5.16. *For all k , $e_{k+1} b_{k+1} e_{k+1} b_{k+1}$ is a suffix of $(e_k b_k)^{m_k-1}$.*

Proof. If we apply Lemma 5.15 to $k + 1$ instead of k , we see that $e_{k+1} b_{k+1} e_{k+1} b_{k+1}$ is a suffix of b_{k+2} , which, by Lemma 5.14, is a suffix of $(e_k b_k)^{m_k-1}$. \square

Finally, to arrive at a contradiction, we will use the words $e_k b_k e_k b_k$ to construct a minimal subsystem of X other than $\{0^\infty\}$. To do this, first define L_k to be the length of the longest block of zeros occurring in $e_k b_k e_k b_k$. Then L_k is also the length of the longest block of zeros in $(e_k b_k)^{m_k-1}$ (assuming $m_k \geq 3$, which is trivial since $m_k \rightarrow \infty$).

Proposition 5.17. *(L_k) is a constant sequence.*

Proof. By Proposition 5.16, $e_{k+1} b_{k+1} e_{k+1} b_{k+1}$ is a suffix of $(e_k b_k)^{m_k-1}$, so $L_{k+1} \leq L_k$. On the other hand, by Lemma 5.15, $e_k b_k e_k b_k$ is a suffix of b_{k+1} , so $L_k \leq L_{k+1}$. \square

For each k , define $y^{(k)}$ to be any point in X of the form

$$y^{(k)} = \ell_k e_k b_k \cdot e_k b_k r_k,$$

where ℓ_k and r_k are left- and right-infinite sequences. By compactness, a subsequence of $(y^{(k)})$ converges to a point $y \in X$. But by construction

(and applying Proposition 5.17), $\overline{\mathcal{O}(y)}$ is a subsystem of X that is disjoint from $\{0^\infty\}$. Therefore $\overline{\mathcal{O}(y)}$ contains a minimal subsystem other than $\{0^\infty\}$, a contradiction, completing the proof of Theorem 1.4. \square

Remark 5.18. The assumption of recurrence was only used above to establish that X has a single minimal subsystem. Our results in this paper show that the only transitive X with more than one minimal subsystem and $\limsup_{n \rightarrow \infty} \left(\frac{c_X(n)}{n} \right) < 3$ are of the following types.

- X has two periodic minimal subsystems (e.g. the orbit closure of $x = \dots 000.1111\dots$ or $x = 1^\infty.0^{n_1}10^{n_2}10^{n_3}1\dots$)
- $\pi(X)$ is eventually periodic in both directions and X has one infinite subsystem (e.g., $x = \dots 000.s\dots$ where s is a one-sided sequence from a Sturmian system).

We now conclude with a proof of Theorem 1.5, which establishes an upper bound of $g - 1$ on the number of generic measures when $X = \overline{\mathcal{O}(x)}$ for some x that is not eventually periodic in both directions and

$$\liminf(c_X(n) - gn) = -\infty.$$

5.2. Proof of Theorem 1.5. Assume that X is as above and x is not eventually periodic to the right. Since $\liminf(c_X(n) - gn) = -\infty$, Lemma 3.7 implies that the number of n -letter right-special words is strictly less than g for infinitely many n . Therefore, there exists $C \leq g - 1$ and a strictly increasing sequence (n_k) such that there are exactly C right-special words of length n_k ; call them $b_k^{(i)}$ for $1 \leq i \leq C$. We may further assume that

$$c_X(n_k) < 2gn_k$$

for every k by considering values of n where $c_X(n) - gn$ is smaller than all previous values $c_X(i) - gi$ for $i < n$; see the discussion following Theorem 2.2 in [1] for details. Therefore, by Lemma 5.4, every word of length $(2g + 1)n_k$ contains at least one word $b_k^{(i)}$.

For each k and i , choose a point $x_k^{(i)}$ with the word $b_k^{(i)}$ appearing in coordinates 0 through $n_k - 1$, and define the measure $\nu_k^{(i)} := \nu_{n_k}^{x_k^{(i)}}$ as discussed at the start of Section 5, i.e.,

$$\nu_k^{(i)} = \frac{1}{n_k} \sum_{j=0}^{n_k-1} \delta_{\sigma^j x_k^{(i)}}.$$

By compactness, we can pass to a subsequence and assume without loss of generality that for all i , $(\nu_k^{(i)})_k$ converges to a limit $\nu^{(i)}$. We claim that every generic measure is some $\nu^{(i)}$. For this, assume that μ is an arbitrary generic measure, and let $x \in X$ be generic for μ .

For each k , the word $x_0 \cdots x_{(2g+1)n_k-1}$ contains some $b_k^{(i_k)}$ by Lemma 5.4, and by again passing to a subsequence, we can assume that i_k is always equal to some fixed i . For each k , define $m_k \leq 2gn_k$ so that $x_{m_k} \cdots x_{m_k+n_k-1} = b_k^{(i)}$.

Since x is generic for μ , both $\nu_{m_k}^x$ and $\nu_{m_k+n_k-1}^x$ converge to μ in the weak topology. This topology is induced by the metric

$$d(\mu, \nu) := \sum_{n \in \mathbb{N}} 2^{-n} |\mu([w_n]) - \nu([w_n])|,$$

where $\{w_n\}$ is an arbitrary enumeration of the set of finite words on $\{0, 1\}$. Then for any $\epsilon > 0$, we may choose K so that for all $k > K$, $d(\nu_{m_k}^x, \mu), d(\nu_{m_k+n_k}^x, \mu) < \epsilon$. This implies that

$$d(m_k \nu_{m_k}^x, m_k \mu) < m_k \epsilon \quad \text{and} \quad d((m_k+n_k) \nu_{m_k+n_k}^x, (m_k+n_k) \mu) < (m_k+n_k) \epsilon,$$

which together imply that

$$d((m_k + n_k) \nu_{m_k+n_k}^x - m_k \nu_{m_k}^x, n_k \mu) < (2m_k + n_k) \epsilon.$$

Therefore

$$d(\nu_{n_k}^{\sigma^{m_k} x}, \mu) < \frac{2m_k + n_k}{n_k} \epsilon \leq (4g + 1) \epsilon.$$

Since $\epsilon > 0$ was arbitrary, $\nu_{n_k}^{\sigma^{m_k} x} \rightarrow \mu$. Recall that $\sigma^{m_k} x$ begins with $b_k^{(i)}$, and so $\nu_{n_k}^{\sigma^{m_k} x} \rightarrow \nu^{(i)}$ by definition of $\nu^{(i)}$. Therefore, $\mu = \nu^{(i)}$, and since μ was an arbitrary generic measure, every generic measure is one of the $\nu^{(i)}$. Since there are $C \leq g - 1$ such measures, the proof is complete. □

REFERENCES

- [1] M. Boshernitzan, *A unique ergodicity of minimal symbolic flows with linear block growth*, Journal d'Analyse Math. **44**, no. 1 (1984) 77-96.
- [2] V. Cyr and B. Kra, *Counting generic measures for a subshift of linear growth*, J. Eur. Math. Soc. **21**, no. 2 (2019) 355-380.
- [3] N. Pytheas Fogg, *Substitutions in Dynamics, Arithmetics and Combinatorics*, Lecture Notes in Mathematics **1794**, Springer, Berlin, 2002.
- [4] A. Katok, *Invariant measures of flows on orientable surfaces*, Dokl. Akad. Nauk SSSR **211** (1973) 775-778.
- [5] D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding*, Cambridge University Press, Cambridge, 1995.

- [6] M. Morse and G. A. Hedlund, *Symbolic dynamics II. Sturmian trajectories*, Amer. J. Math. **62** (1940) 1-42.
- [7] N. Ormes and R. Pavlov, *On the complexity function for sequences which are not uniformly recurrent*, Contemp. Math., to appear.
- [8] M. E. Paul, *Minimal symbolic flows having minimal block growth*, Math. Systems Theory **8(4)** (1974/75) 309-315.
- [9] W. Veech, *Interval exchange transformations*, Journal d'Analyse Math. **33** (1978) 222-272.
- [10] P. Walters, *An introduction to ergodic theory*, Springer-Verlag, New York-Berlin, 1982.

ANDREW DYKSTRA, DEPARTMENT OF MATHEMATICS, HAMILTON COLLEGE,
CLINTON, NY 13323, USA

E-mail address: adykstra@hamilton.edu

NIC ORMES, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF DENVER, 2390
S. YORK ST., DENVER, CO 80208

E-mail address: nic.ormes@du.edu

RONNIE PAVLOV, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF DENVER,
2390 S. YORK ST., DENVER, CO 80208

E-mail address: rpavlov@du.edu

URL: www.math.du.edu/~rpavlov/