

ON INTRINSIC ERGODICITY AND WEAKENINGS OF THE SPECIFICATION PROPERTY

RONNIE PAVLOV

ABSTRACT. Since seminal work of Bowen ([2]), it has been known that the specification property implies various useful properties about an expansive topological dynamical system, among them uniqueness of the measure of maximal entropy (often referred to as intrinsic ergodicity). Weakenings of the specification property have been defined and profitably applied in various works such as [6], [9], [11], [16], and [17].

It has been an open question (see p. 798 of [4]) whether two of these properties, which we here call almost specification and non-uniform specification, imply intrinsic ergodicity for expansive topological systems. We answer this question negatively by exhibiting examples of subshifts with multiple measures of maximal entropy with disjoint support which have non-uniform specification with any gap function $f(n) = O(\ln n)$ or almost specification with any mistake function $g(n) \geq 4$. We also show some results in the opposite direction, showing that subshifts with non-uniform specification with gap function $f(n) = o(\ln n)$ or almost specification with mistake function $g(n) = 1$ cannot have multiple measures of maximal entropy with disjoint support.

1. INTRODUCTION

Entropy is one of the most well-studied invariants in the field of dynamical systems. Entropy can be defined both for measure-theoretic dynamical systems (given by a probability space (X, μ) and T a μ -preserving self-map of X) and for topological dynamical systems (given by a compact topological space X and T a continuous self-map of X). A relation between the two notions of entropy is given by the celebrated Variational Principle, which states that for any topological dynamical system (X, T) , the topological entropy is the supremum over all measure-theoretic entropies for Borel measures μ on X which are preserved by T . For this reason, such a measure on X whose entropy achieves this supremum is called a **measure of maximal entropy**.

It is well-known that expansive topological dynamical systems always have at least one measure of maximal entropy, and symbolic systems/subshifts are some of the best-known examples of expansive topological dynamical systems. Subshifts may have multiple measures of maximal entropy even under assumptions such as minimality or almost soficity; see [7] and [13] for some classical examples. A system is called **intrinsically ergodic** when there is a unique measure of maximal entropy, and establishing intrinsic ergodicity is a central problem in both ergodic theory and topological dynamics. A common way of proving intrinsic ergodicity is via the specification property and various weakenings of it. All such properties involve

2010 *Mathematics Subject Classification*. Primary: 37B10; Secondary: 37B40, 37D35.

Key words and phrases. Measure of maximal entropy, specification, subshift.

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

combining segments of orbits into a new orbit in various ways; in subshifts, these orbit segments can be represented by words occurring in points of the subshift. For instance, for subshifts, the traditional **specification** property can be stated as follows: there exists a constant R so that any words in the language can be combined into a new word in the language when separated by a gap of size R . Bowen proved in [2] that for expansive systems (such as subshifts), the specification property implies intrinsic ergodicity. We consider two weakenings of the classical specification property in the symbolic setting; the first (called **non-uniform specification**) allows one to combine arbitrary words in the language into a new word in the language if gaps which are “small” in comparison to the combined words are placed in between (controlled by a function $f(n) = o(n)$), and the second (called **almost specification**) allows one to concatenate arbitrary words in the language into a new word in the language if a “small” number of letters are allowed to change in each word (controlled by a function $g(n) = o(n)$). See Section 2 for formal definitions.

In [4], Climenhaga and Thompson proved, among other results, that all factors of β -shifts are intrinsically ergodic, answering an open question of Klaus Thomsen. They did this by introducing a more nuanced property about decompositions of words in the language, and showing that this property implies intrinsic ergodicity; as it is a little complex, we do not state this property here. However, they did not determine whether the simpler properties of non-uniform specification and/or almost specification (the latter is satisfied by β -shifts and their factors) imply intrinsic ergodicity, and left this as an open question. We answer this question negatively, by exhibiting two different examples of subshifts which have non-uniform specification and almost specification, respectively, and yet have multiple measures of maximal entropy.

Theorem 1.1. *For any positive nondecreasing function $f(n)$ with $\liminf_{n \rightarrow \infty} \frac{f(n)}{\ln n} > 0$, there exists a subshift with non-uniform specification with gap function $f(n)$ with exactly two ergodic measures of maximal entropy, whose supports are disjoint.*

Theorem 1.2. *There exists a subshift with almost specification with mistake function $g(n) = 4$ with exactly two ergodic measures of maximal entropy, whose supports are disjoint.*

It is particularly surprising that even boundedness of the mistake function for systems with almost specification does not imply uniqueness of the measure of maximal entropy, given that previously it was unknown whether $g(n) = o(n)$ might even be sufficient for uniqueness.

The example proving Theorem 1.1 is quite similar to a subshift example of Haydn ([8]), for which he also proved existence of two ergodic measures of maximal entropy with disjoint supports. The alphabet for both our example and Haydn’s consists of three types of symbols: 0s, some number N of positive integers, and N negative integers. Points in Haydn’s subshift consist of (arbitrary) runs of positive symbols and runs of negative symbols, which are separated by runs of 0s, and for which nonzero symbols flanking a run of 0s must have opposite sign. He used a counting argument to prove that this subshift has topological entropy $\ln N$, and since it contains the full shifts on N positive symbols and N negative symbols, this yields multiple measures of maximal entropy with disjoint support, namely the uniform Bernoulli measures on these two full shifts.

Haydn required runs of 0s to have length at least linear in the lengths of the runs of positive and negative symbols flanking it, and so his example did not have the non-uniform specification property. (See Definition 2.14.) However, with some alterations and a slightly more complicated counting argument, we can achieve the same conclusion for a subshift with non-uniform specification. A more direct motivation for our example is the iceberg shift of Burton and Steif ([3]), which is a two-dimensional shift of finite type with specification, but with two ergodic measures of maximal entropy. That example also consists of “islands” of positive and negative symbols with restricted adjacency, and our example was designed to exhibit the same phenomenon in one dimension.

We also prove some results in the opposite direction, proving that if $f(n)$ grows extremely slowly and/or $g(n)$ has an even smaller upper bound, then the subshift cannot have two measures of maximal entropy with disjoint supports. This in some sense precludes a “strong nonuniqueness” of the measures of maximal entropy.

Theorem 1.3. *If a subshift has non-uniform specification with gap function $f(n)$ where $\liminf_{n \rightarrow \infty} \frac{f(n)}{\ln n} = 0$, then it cannot have two measures of maximal entropy with disjoint support.*

Theorem 1.4. *If a subshift has almost specification with mistake function $g(n) = 1$, then it cannot have two measures of maximal entropy with disjoint support.*

We have then completely answered the question of whether non-uniform specification for a particular gap function $f(n)$ can coexist with multiple measures of maximal entropy with disjoint supports, and leave open only the case $1 \leq g(n) < 4$ for the corresponding question for almost specification. This suggests that for both non-uniform specification and almost specification, there is a “phase transition” significantly below n for the relevant gap or mistake function for how the property influences the measures of maximal entropy.

It is still plausible that an extremely slow growth rate for $f(n)$ and/or upper bound less than 4 on $g(n)$ may imply uniqueness of the measure of maximal entropy, but we do not know whether this is true or not.

Question 1.5. *Do there exist positive functions $F(n)$ and/or $G(n)$ so that non-uniform specification with a gap function $f(n) \leq F(n)$ and/or almost specification with a mistake function $g(n) \leq G(n)$ forces intrinsic ergodicity?*

We’ll now give a very informal overview of the techniques used to prove these results; most are based on some elementary counting arguments. Specifically, given any collection \mathcal{W} of finite words, one can define a subshift X as (limits of) biinfinite concatenations of the words in \mathcal{W} . (Such systems are called **coded** subshifts and are well-studied, though we will not need any advanced structural facts about such systems.) We mainly need relationships between the sizes of the various sets $\mathcal{W}_n := A^n \cap \mathcal{W}$ and the topological entropy of X . For instance, both Theorems 1.1 and 1.2 use the fact that

$$\sum_{n=1}^{\infty} |\mathcal{W}_n| e^{-n \ln N} < 1$$

to prove that $h(X) \leq \ln N$. Similarly, Theorem 1.4 involves constructing \mathcal{W} from the supports of two ergodic MMEs supposed (for contradiction) to have disjoint

supports, and using the fact that

$$\sum_{n=1}^{\infty} |\mathcal{W}_n| e^{-nh(X)} > 1$$

to create a greater topological entropy than $h(X)$, a contradiction. (The proof of Theorem 1.3 is a simpler version, where only two words from \mathcal{W} need to be concatenated, rather than arbitrarily many.)

In the example proving Theorem 1.1, the collections \mathcal{W}_n consist of long words of constant (negative or positive) sign, followed by a run of 0s of length controlled by the desired gap function $f(n)$. In this way, $f(n)$ determines the relation of $|\mathcal{W}_n|$ to $e^{n \ln N}$. In the example proving Theorem 1.2, we instead use Hamming codes to create collections \mathcal{W}_n for which any word of constant sign can be changed to a word in some \mathcal{W}_n with no more than 2 changes (this yields the desired almost specification with $g(n) = 4$), and $|\mathcal{W}_n| e^{-n \ln N}$ behaves like a constant times n^{-2} (so that the sum $\sum_{n=1}^{\infty} |\mathcal{W}_n| e^{-n \ln N}$ can be made smaller than 1.)

Finally, we quickly summarize the structure of the paper. Section 2 gives formal definitions of all relevant concepts, Section 3 gives proofs of our results concerning non-uniform specification (namely Theorems 1.1 and 1.3), and Section 4 gives proofs of our results concerning almost specification (namely Theorems 1.2 and 1.4).

Remark 1.6. We would like to point out that the question of whether non-uniform/almost specification implies intrinsic ergodicity has been independently answered negatively by Kwietniak, Oprocha, and Rams ([10]). They also prove other interesting results about implications of and relations between non-uniform specification (there called weak specification), almost specification, and the so-called Climenhaga-Thompson decomposition from [4].

Remark 1.7. We note that the property that we call non-uniform specification has gone by different names in different works: in [6] and [16] it is called almost weak specification, in [10] it is called weak specification, and in [20] it is called almost specification. We also note that the term ‘non-uniform specification’ is used for a measure-theoretic notion of similar flavor in [18]; we hope that the quite different systems studied there (diffeomorphisms on manifolds) mean that we are unlikely to have caused much additional confusion.

2. DEFINITIONS AND PRELIMINARIES

Definition 2.1. For any finite alphabet A , the **full shift** over A is the set $A^{\mathbb{Z}} = \{\dots x_{-1}x_0x_1\dots : x_i \in A\}$, which is viewed as a compact topological space with the (discrete) product topology.

Definition 2.2. A **word** over A is a member of $A^{\{i, i+1, \dots, j\}}$ for some $i < j$, whose **length** $j - i + 1$ is denoted by $|w|$. The set $\bigcup_{i, j \in \mathbb{Z}, i < j} A^{\{i, i+1, \dots, j\}}$ of all words over A is denoted by A^* . For any n , we use A^n to denote the set $A^{\{1, \dots, n\}}$.

Definition 2.3. The **shift action**, denoted by $\{\sigma^n\}_{t \in \mathbb{Z}}$, is the \mathbb{Z} -action on a full shift $A^{\mathbb{Z}}$ defined by $(\sigma^n x)_m = x_{m+n}$ for $m, n \in \mathbb{Z}$.

Definition 2.4. A **subshift** is a closed subset of a full shift $A^{\mathbb{Z}}$ which is invariant under the shift action, which is a compact space with the induced topology from $A^{\mathbb{Z}}$.

The single shift $\sigma := \sigma^1$ is an automorphism on any subshift, and so for any subshift X , (X, σ) is a topological dynamical system. An alternate definition for a subshift is in terms of a list of forbidden words; for any set $\mathcal{F} \subset A^*$, one can define the set $X(\mathcal{F}) := \{x \in A^{\mathbb{Z}} : x_i x_{i+1} \dots x_j \notin \mathcal{F} \forall i, j \in \mathbb{Z}, i < j\}$. It is well known that any $X(\mathcal{F})$ is a subshift, and all subshifts are representable in this way.

Definition 2.5. The **language** of a subshift X , denoted by $\mathcal{L}(X)$, is the set of all words which appear in points of X . For any $n \in \mathbb{Z}$, $\mathcal{L}_n(X) := \mathcal{L}(X) \cap A^n$, the set of words in the language of X with length n .

In the previous definition, we dealt only with words from A^n rather than $A^{\{i, \dots, j\}}$ for arbitrary $i < j$; this is because any word in $A^{\{i, \dots, j\}}$ can clearly be thought of as a word in A^{j-i+1} by simply shifting it. We will generally consider two words to be the same if they are shifts of each other.

Definition 2.6. For any subshift and word $w \in \mathcal{L}_n(X)$, the **cylinder set** $[w]$ is the set of all $x \in X$ with $x_1 x_2 \dots x_n = w$.

Definition 2.7. For any subshift $X \subset A^{\mathbb{Z}}$ and any $k \in \mathbb{N}$, the **k th higher-power shift** associated to X , denoted X^k , is a subshift with alphabet $\mathcal{L}_k(X)$ defined by the following rule: $y \in (\mathcal{L}_k(X))^{\mathbb{Z}}$ is an element of X^k if and only if the point x defined by concatenating the “letters” of y is in X . (Formally, $\forall n \in \mathbb{Z}$, the n th letter of x is defined to be the $(n \bmod k)$ th letter of $y_{\lfloor n/k \rfloor}$.)

It is well-known that the dynamical systems (X^k, σ) and (X, σ^k) are topologically isomorphic.

Definition 2.8. The **topological entropy** of a subshift X is

$$h(X) := \lim_{n \rightarrow \infty} \frac{1}{n} \ln |\mathcal{L}_n(X)|.$$

We also need some definitions from measure-theoretic dynamics; all measures considered in this paper will be Borel probability measures on a full shift $A^{\mathbb{Z}}$.

Definition 2.9. A measure μ on $A^{\mathbb{Z}}$ is **ergodic** if any measurable set C which is shift-invariant, meaning $\mu(C \Delta \sigma C) = 0$, has measure 0 or 1.

Not all σ -invariant measures are ergodic, but a well-known result called the ergodic decomposition shows that any non-ergodic measure can be written as a “weighted average” (formally, an integral) of ergodic measures; see Chapter 6 of [19] for more information.

Definition 2.10. For any σ -invariant measure μ on a full shift $A^{\mathbb{Z}}$, the **measure-theoretic entropy** of μ is

$$h(\mu) := \lim_{n \rightarrow \infty} \frac{-1}{n} \sum_{w \in A^n} \mu([w]) \ln \mu([w]),$$

where terms with $\mu([w]) = 0$ are omitted from the sum.

In Definitions 2.8 and 2.10, a standard subadditivity argument shows that the limits can be replaced by infimums; i.e. for any n , $h(X) \leq \frac{1}{n} \ln |\mathcal{L}_n(X)|$ and $h(\mu) \leq \frac{-1}{n} \sum_{w \in A^n} \mu([w]) \ln \mu([w])$. This implies the following fact.

Lemma 2.11. For any σ -invariant measure μ on a subshift X ,

$$|\{w \in \mathcal{L}_n(X) : \mu([w]) > 0\}| \geq e^{nh(\mu)}.$$

Proof. Choose any such X , μ , and n , and denote the set in the lemma by S . It is easily checked that for any probability vector (x_1, \dots, x_n) , $-\sum_{i=1}^n x_i \ln x_i \leq \ln n$, and equality is achieved if and only if all x_i are equal to $\frac{1}{n}$ (see [19], Corollary 4.2.1 for a proof). Therefore,

$$h(\mu) \leq \frac{-1}{n} \sum_{w \in A^n} \mu([w]) \ln \mu([w]) \leq \frac{\ln |S|}{n}.$$

This implies that $|S| \geq e^{nh(\mu)}$. □

We also need the following fact, which is essentially just an application of Birkhoff's ergodic theorem; a proof (of a more general version) can be found as Lemma 4.8 in [12].

Lemma 2.12. *For any subshift X , any ergodic measure μ on X , any word $w \in \mathcal{L}_n(X)$, any $n \in \mathbb{N}$, and any $\epsilon > 0$, define the set $C_{n,\epsilon,w}(X)$ to be the set of all $w \in \mathcal{L}_n(X)$ which have between $n(\mu([w]) - \epsilon)$ and $n(\mu([w]) + \epsilon)$ occurrences of w . Then,*

$$\liminf_{n \rightarrow \infty} \frac{\ln |C_{n,\epsilon,w}(X)|}{n} \geq h(\mu).$$

Definition 2.13. For any subshift X , a **measure of maximal entropy** on X is a measure μ with support contained in X for which $h(\mu) = h(X)$.

It is well-known that every subshift has at least one measure of maximal entropy, and the ergodic decomposition and affineness of the entropy map (see Theorem 8.7(ii) in [19]) then imply the existence of an ergodic measure of maximal entropy as well. For full shifts in particular, it is well-known that there is only one measure of maximal entropy, namely the uniform Bernoulli measure μ defined by $\mu([w]) = |A|^{-n}$ for all n and $w \in A^n$.

As noted in the introduction, the specification property of Bowen ([2]) implies uniqueness of the measure of maximal entropy for subshifts. We consider two weakenings of Bowen's property, which we call **non-uniform specification** and **almost specification**. Though these properties can be defined for arbitrary topological dynamical systems, we here restrict our attention to subshifts, giving definitions specific to that case which are slightly simpler.

The following property was originally defined in [11], Lemma 2.1, but was not there given a name.

Definition 2.14. A subshift X has **non-uniform specification with gap function** $f(n)$ if

- $f(n)$ is positive and nondecreasing
- $\frac{f(n)}{n} \rightarrow 0$
- For any words $w^{(1)}, w^{(2)}, \dots, w^{(k)} \in \mathcal{L}(X)$, and for any integers n_1, \dots, n_{k-1} where $n_i \geq f(|w^{(i)}|)$ for all i , there exist words $v^{(1)} \in \mathcal{L}_{n_1}(X), \dots, v^{(k-1)} \in \mathcal{L}_{n_{k-1}}(X)$ so that the word $w^{(1)}v^{(1)}w^{(2)}v^{(2)} \dots w^{(k-1)}v^{(k-1)}w^{(k)} \in \mathcal{L}(X)$.

The assumption that $f(n)$ is nondecreasing is not explicitly required in the literature, but prevents some pathological cases which would make our proofs more complicated. We require $f(n)$ to be positive since, for any n , $f(n) = 0$ would imply that the higher power shift X^n is just a full shift.

Definition 2.15. A subshift X has **specification with gap g** if it has non-uniform specification with the constant gap function $f(n) = g$.

We note that Definition 2.15 is slightly different from some definitions of specification in the literature, which often assume not only that the created word $w^{(1)}v^{(1)}w^{(2)}v^{(2)} \dots w^{(k-1)}v^{(k-1)}w^{(k)}$ is in $\mathcal{L}(X)$, but that it is also a subword of a periodic point in X . However, our version of specification implies that all words in $\mathcal{L}(X)$ are subwords of periodic points (see [1]), and so the definitions are equivalent for subshifts.

A second weakening of specification was defined by Pfister and Sullivan in [15], and was there called the g -almost product property. We follow the convention of [17] and call this property **almost specification**.

Definition 2.16. A subshift X has **almost specification with mistake function $g(n)$** if

- $g(n)$ is positive and nondecreasing
- $\frac{g(n)}{n} \rightarrow 0$
- For any words $w^{(1)}, w^{(2)}, \dots, w^{(k)} \in \mathcal{L}(X)$, there exist words $v^{(1)}, v^{(2)}, \dots, v^{(k)} \in \mathcal{L}(X)$ so that $|w^{(i)}| = |v^{(i)}|$ for every i , $w^{(i)}$ and $v^{(i)}$ differ on at most $g(|w^{(i)}|)$ letters for every i , and the concatenation $v^{(1)}v^{(2)} \dots v^{(k)}$ is in $\mathcal{L}(X)$.

There are many subshifts known to satisfy almost specification; for instance, any β -shift has almost specification with mistake function $g(n) = 1$ (see [14]), and many of the so-called S -gap shifts also satisfy almost specification (see Remark 5.1 from [5]). For subshifts, non-uniform specification implies almost specification, but it's shown in [10] that this implication does not hold for (properly defined versions of) these properties for general topological dynamical systems.

Remark 2.17. The fact that non-uniform specification implies almost specification for subshifts was the motivation for our use of the term “non-uniform specification” rather than “almost weak specification” in this paper; otherwise, almost weak specification would imply almost specification, which would be quite confusing given the names.

Remark 2.18. In [20], Yamamoto also studies various weakenings of specification and their implications. The property that he calls almost specification is our non-uniform specification, and the property that he calls the almost product property is essentially our almost specification.

3. NON-UNIFORM SPECIFICATION

Proof of Theorem 1.1. We define X to have alphabet $A = \{-N, \dots, -1, 0, 1, \dots, N\}$, with $N > e^6 \approx 403.4$, and list of forbidden words \mathcal{F} consisting of:

- All adjacent pairs ij whose product is negative, i.e. consisting of one negative and one positive letter
- All words $v_1v_2 \dots v_j0^mv_{j+1}$, where all $v_i \neq 0$ and $m \leq \ln j$.

X then contains all points which look like $\dots w^{(-1)}0^{m_{-1}}w^{(0)}0^{m_0}w^{(1)}0^{m_1} \dots$, where each $w^{(i)}$ consists of a “run” of nonzero letters of the same sign, and for every i , $m_i > \ln|w^{(i)}|$. All other points of X are “degenerate” cases which have

either an infinite or biinfinite string of 0s or positives or negatives. We also note for future reference that, given any word $w \in A^*$ which does not contain a forbidden word as described in the above list, the point $\dots 000w000\dots$ is clearly in X , and so $w \in \mathcal{L}(X)$.

First, we will show that X has non-uniform specification with gap function $1 + \lfloor \ln n \rfloor$. Consider any words $w^{(1)}, \dots, w^{(k)}$ in $\mathcal{L}(X)$ and any positive m_1, \dots, m_{k-1} with $m_i \geq 1 + \lfloor \ln |w^{(i)}| \rfloor$ for each i . Define $w := w^{(1)}0^{m_1}w^{(2)}0^{m_2}\dots 0^{m_{k-1}}w^{(k)}$; we claim that $w \in \mathcal{L}(X)$, which will demonstrate the desired non-uniform specification. Firstly, since all m_i are positive, introducing the runs of 0 letters between the words $w^{(i)}$ could not have possibly introduced an adjacent pair of nonzero letters of opposite sign. All that's left is to show that w does not contain any word of the form $v_1v_2\dots v_j0^mv_{j+1}$ with all v_i nonzero and $m \leq \ln j$. Suppose for a contradiction that w does contain such a word, call it u . Clearly u cannot be contained in any of the $w^{(i)}$, since they were assumed to be in $\mathcal{L}(X)$. Just as clearly, the central 0^m in u must contain an entire 0^{m_i} from w , and the letters $v_1\dots v_j$ must all be from the suffix of some $w^{(i)}$. However, this means that $m \geq 1 + \lfloor \ln |w^{(i)}| \rfloor > \ln |w^{(i)}|$, and since we assumed $m \leq \ln j$, it must be the case that $j > |w^{(i)}|$. But this is impossible; $w^{(i)}$ is preceded in w by a 0 for $i > 1$, and by nothing for $i = 1$. Therefore, w contains no forbidden words, and so is in $\mathcal{L}(X)$, proving non-uniform specification of X with gap function $1 + \lfloor \ln n \rfloor$.

We will now show that for any $C > 0$, there exists a higher-power shift X^k with non-uniform specification with gap function $\max(1, \lceil C \ln n \rceil)$. Since any positive f with $\liminf_{n \rightarrow \infty} \frac{f(n)}{\ln n} > 0$ is bounded from below by $\max(1, \lceil C \ln n \rceil)$ for some $C > 0$, this will imply that for any such f , some X^k has non-uniform specification with gap function f .

Choose $C > 0$ and define any $k > 4C^{-1}$. Consider X^k , the k th higher-power shift of X . The reader may check that the non-uniform specification of X with gap function $1 + \lfloor \ln n \rfloor$ implies non-uniform specification of X^k with gap function $\lceil (1 + \lfloor \ln(kn) \rfloor)/k \rceil$. We claim that

$$(1) \quad \lceil (1 + \lfloor \ln(kn) \rfloor)/k \rceil \leq \max(1, \lceil C \ln n \rceil)$$

for all n , which would imply that X^k has non-uniform specification with gap function $\max(1, \lceil C \ln n \rceil)$ as desired.

To verify (1), we first note that both sides are always positive, and so we do not need to prove anything in the case where the left-hand side is 1. Suppose that the left-hand side is at least 2. Then, $(1 + \lfloor \ln(kn) \rfloor)/k > 1$, which implies that $\ln(kn) \geq k$, and so $n \geq e^k/k > k$, since k is a positive integer. Finally, we note that $\lfloor \ln(kn) \rfloor \geq 1$, and so

$$\frac{1 + \lfloor \ln(kn) \rfloor}{k} \leq \frac{2}{k} \ln(kn) < \frac{2}{k} \ln(n^2) \leq \frac{4}{k} \ln n < C \ln n.$$

This implies (1), and therefore that for any positive $f(n)$ with $\liminf_{n \rightarrow \infty} \frac{f(n)}{\ln n} > 0$, there exists k for which X^k has non-uniform specification with gap function $f(n)$.

Now we will show that X has exactly two ergodic measures of maximal entropy, whose supports are disjoint, and that this property holds for all higher-power shifts X^k as well, which will complete the proof of Theorem 1.1. We begin by using a counting argument to show that $h(X) = \ln N$. Since X contains the full shifts on

N positive symbols and N negative symbols respectively, clearly $h(X) \geq \ln N$, and so it suffices to show that $h(X) \leq \ln N$.

For this, we will just bound $\mathcal{L}_n(X)$ from above for all n . Every $w \in \mathcal{L}_n(X)$ can be decomposed as $w^{(1)}0^{m_1}w^{(2)}0^{m_2} \dots w^{(k)}0^{m_k}$, where each $w^{(i)}$ consists of a ‘‘run’’ of nonzero letters of the same sign. In addition, for every $i > 1$, $|w^{(i)}| > 0$, and for every $i < k$, $m_i > \ln |w^{(i)}|$ (note that we have no control over the last run of 0s). This means that $n_i \geq 2$ for $1 < i < k$. We define notation $\ell_i = |w^{(i)}|$ and auxiliary parameters $n_i = \ell_i + m_i$ for $1 \leq i \leq k$. Then clearly $\sum n_i = n$. For fixed choice of k and n_i , w is completely determined by the choice of $|w^{(i)}|$, the choice of whether all symbols of $w^{(i)}$ are positive or negative, and the letters of $w^{(i)}$, for $1 \leq i \leq k$.

For each $1 < i < k$, we know that $m_i > \ln \ell_i$, so $\ell_i + \ln \ell_i < n_i$. We wish to use this to bound ℓ_i from above. Since $x + \ln x$ is an increasing function with range \mathbb{R} , there exists $x \in \mathbb{R}$ for which $x + \ln x = n_i$. Clearly such x is less than n_i , and so $x = n_i - \ln x > n_i - \ln n_i$. But then $x + \ln(n_i - \ln n_i) < n_i$, so $x < n_i - \ln(n_i - \ln n_i)$. Since $\ell_i + \ln \ell_i < n_i$, $\ell_i < x < n_i - \ln(n_i - \ln n_i)$.

This means that for each n_i , $1 < i < k$, the number of choices for $w^{(i)}0^{m_i}$ is bounded from above by

$$2 \left(\sum_{\ell_i=1}^{\lfloor n_i - \ln(n_i - \ln n_i) \rfloor} N^{\ell_i} \right) < 4N^{n_i - \ln(n_i - \ln n_i)}.$$

For simplicity, we define $M_j = 4N^{j - \ln(j - \ln j)}$ for $j \geq 2$. For $i = 1$ and $i = k$, we more simply bound the number of choices for $w^{(i)}0^{m_i}$ by

$$2 \left(\sum_{\ell_i=0}^{n_i} N^{\ell_i} \right) < 4N^{n_i}.$$

We then see that

$$\begin{aligned} (2) \quad |\mathcal{L}_n(X)| &< 4N^n + 16(n-1)N^n + \sum_{k=3}^n \sum_{\substack{n_1, \dots, n_k \\ \sum n_i = n}} 4N^{n_1} \left(\prod_{i=2}^{k-1} M_{n_i} \right) 4N^{n_k} \\ &= 4N^n + 16(n-1)N^n + 16N^n \sum_{k=3}^n \sum_{\substack{n_1, \dots, n_k \\ \sum n_i = n}} \left(\prod_{i=2}^{k-1} M_{n_i} N^{-n_i} \right), \end{aligned}$$

where the first two terms correspond to the cases $k = 1$ and $k = 2$ respectively. We can bound the third term from above as follows:

$$\begin{aligned}
(3) \quad & 16N^n \sum_{k=3}^n \sum_{\substack{n_1, \dots, n_k \\ \sum n_i = n}} \left(\prod_{i=2}^{k-1} M_{n_i} N^{-n_i} \right) \\
& = 16N^n \sum_{k=3}^n \sum_{\substack{n_2, \dots, n_{k-1} \\ \sum n_i < n}} \left(n - \sum_{j=2}^{k-1} n_j - 1 \right) \left(\prod_{i=2}^{k-1} M_{n_i} N^{-n_i} \right) \\
& < 16nN^n \sum_{k=3}^n \sum_{\substack{n_1, \dots, n_{k-1} \\ \sum n_i < n}} \left(\prod_{i=2}^{k-1} \frac{4}{N^{\ln(n_i - \ln n_i)}} \right) \leq 16nN^n \sum_{k=3}^n \left(\sum_{t=2}^{\infty} \frac{4}{N^{\ln(t - \ln t)}} \right)^{k-2}.
\end{aligned}$$

We move to bounding $\sum_{t=2}^{\infty} \frac{4}{N^{\ln(t - \ln t)}} = \sum_{t=2}^{\infty} \frac{4}{(t - \ln t)^{\ln N}}$ from above:

$$\sum_{t=2}^{\infty} \frac{4}{(t - \ln t)^{\ln N}} < \frac{4}{(2 - \ln 2)^{\ln N}} + \frac{4}{(3 - \ln 3)^{\ln N}} + \int_3^{\infty} \frac{1}{(x - \ln x)^{\ln N}} dx.$$

Since $2 - \ln 2 > 1.3$, $3 - \ln 3 > 1.8$, and $x - \ln x > 0.5x$ for $x \geq 3$, we can bound this from above by

$$\frac{4}{1.3^{\ln N}} + \frac{4}{1.8^{\ln N}} + \int_3^{\infty} \frac{1}{(0.5x)^{\ln N}} = \frac{4}{1.3^{\ln N}} + \frac{4}{1.8^{\ln N}} + \frac{3}{((\ln N) - 1) \cdot 1.5^{\ln N}}.$$

Since $N > e^6$, the reader can check that this quantity is less than 1, and we denote it by α . Then, by (2) and (3),

$$(4) \quad |\mathcal{L}_n(X)| < 4N^n + 16(n-1)N^n + 16nN^n \sum_{k=3}^n \alpha^{k-2} < \frac{16n}{1-\alpha} N^n.$$

Now taking logs, dividing by n , and letting $n \rightarrow \infty$ shows that $h(X) \leq \ln N$, and since X contains a full shift on N symbols, that $h(X) = \ln N$ as well.

Now, consider any ergodic measure of maximal entropy μ of X . Our goal is to show that $\mu([a0]) = 0$. We begin by showing that $\mu([a0]) = 0$ for every $a \in \{1, \dots, N\}$. For a contradiction, assume that $\mu([a0]) = \beta > 0$ for some such a . Define, for every n , the set $C_{n,0.5\beta,a0}(X)$ of n -letter words with between $0.5n\beta$ and $1.5n\beta$ occurrences of $a0$. By Lemma 2.12, $\liminf_{n \rightarrow \infty} \frac{\ln |C_{n,0.5\beta,a0}(X)|}{n} \geq h(\mu) = h(X) = \ln N$. However, in the decomposition of n -words in X given above, k is at least equal to the number of occurrences of $a0$, and so the same argument used above to show (4) implies that

$$|C_{n,0.5\beta,a0}(X)| < 16nN^n \sum_{k=\lceil 0.5n\beta \rceil}^n \alpha^{k-2} \leq \frac{16n}{1-\alpha} N^n \alpha^{0.5n\beta}.$$

Then $\limsup_{n \rightarrow \infty} \frac{\ln |C_{n,0.5\beta,a0}|}{n} \leq \ln N + 0.5\beta \ln \alpha < \ln N$, a contradiction. Therefore, $\mu([a0]) = 0$, and a similar proof shows that $\mu([b0]) = 0$ as well for all $b \in \{-N, \dots, -1\}$. Therefore, μ is entirely supported either on $0^{\mathbb{Z}}$, $\{1, \dots, N\}^{\mathbb{Z}}$, or $\{-N, \dots, -1\}^{\mathbb{Z}}$. Clearly the first case is impossible, as it would imply $h(X) = h(\mu) = 0$, and so $\mu([0]) = 0$. The only such measures with entropy $\ln N$ are the uniform Bernoulli measures over $\{1, \dots, N\}$ and $\{-N, \dots, -1\}$, and so X has exactly two ergodic measures of maximal entropy, with disjoint supports. By the ergodic

decomposition, any measure of maximal entropy can be written as a “weighted average” of ergodic measures of maximal entropy, and so the property that $\mu([0]) = 0$ is shared by non-ergodic measures of maximal entropy on X as well.

We now prove the same about any higher power shift X^k . Consider any ergodic measure of maximal entropy ν on X^k . Then ν clearly induces a measure ν^* on X by defining $\nu^*([w^{(1)} \dots w^{(n)}]) = \nu([w^{(1)} \dots w^{(n)}])$ for all choices of $w^{(1)}, \dots, w^{(n)} \in \mathcal{L}_k(X)$; the $w^{(i)}$ are interpreted as concatenated k -letter words on the left-hand side and as letters in the alphabet of X^k on the right-hand side. The measure ν^* may not be σ -invariant, but it is invariant under σ^k since ν was σ -invariant on X^k . Therefore, $\mu := \frac{1}{k} \sum_{i=0}^{k-1} \sigma^i \nu^*$ is a σ -invariant measure on X , and it is a measure of maximal entropy on X since ν was a measure of maximal entropy on X^k and the entropy map $\mu \mapsto h(\mu)$ is affine. Therefore, $\mu([0]) = 0$, which clearly implies that $\nu([u]) = 0$ for every $u \in \mathcal{L}_k(X)$ containing a 0 letter. Now, since ν is ergodic as a measure on X^k , this implies that ν is supported either entirely on the full shift on $\{1, \dots, N\}^k$ or the full shift on $\{-1, \dots, -N\}^k$, and as before this means that there are exactly two choices for ν , whose supports are disjoint. \square

Remark 3.1. We note that the alphabet size for our examples were quite large, and point out that this is to some extent unavoidable for our construction. Specifically, suppose that we want to engineer a subshift with non-uniform specification for gap function g with $g(n) \leq m$ for $n \leq k$. Then, our construction would require allowing m 0s to follow any run of nonzero letters of length less than or equal to k . In particular, this would mean that any $k+m$ -letter words of the form $w0v0^m$, where w and v are any (possibly empty) words with all letters positive, would be freely concatenated to make new words in X . But there are at least kN^{k-1} such words, implying that $h(X) \geq \frac{\ln(kN^{k-1})}{k+m}$. Since we wish for $h(X)$ to be equal to $\ln N$, this means that $kN^{k-1} \leq N^{k+m}$, so $k \leq N^{m+1}$ and $N \geq k^{1/(m+1)}$. In other words, if we wish to realize a gap function which will stay small for quite some time, it enforces a lower bound on the alphabet size in our construction. In particular, this shows that we could not use some constant N for all of the gap functions satisfying the hypotheses of Theorem 1.1.

Proof of Theorem 1.3. Suppose for a contradiction that X is a subshift with non-uniform specification with a gap function $f(n)$ satisfying $\liminf_{n \rightarrow \infty} \frac{f(n)}{\ln n} = 0$ and which possesses two measures of maximal entropy μ, ν with disjoint supports.

We first use μ and ν to give a lower bound on $|\mathcal{L}_n(X)|$. For every n , define $\mathcal{M}_n(X) = \{w \in \mathcal{L}_n(X) : \mu([w]) > 0\}$ and $\mathcal{N}_n(X) = \{w \in \mathcal{L}_n(X) : \nu([w]) > 0\}$. (For $n = 0$, we define both $\mathcal{M}_n(X)$ and $\mathcal{N}_n(X)$ to be singletons consisting of the empty word \emptyset .) For any $k < n$, any k -letter subword of a word in $\mathcal{M}_n(X)$ must be in $\mathcal{M}_k(X)$, and a similar statement holds for $\mathcal{N}_n(X)$ and $\mathcal{N}_k(X)$. Lemma 2.11 implies that $|\mathcal{M}_n(X)| \geq e^{nh(\mu)} = e^{nh(X)}$ and $|\mathcal{N}_n(X)| \geq e^{nh(\nu)} = e^{nh(X)}$ for all n . Also, since μ and ν have disjoint supports, there exists N so that $\mathcal{M}_n(X) \cap \mathcal{N}_n(X) = \emptyset$ for all $n \geq N$.

We choose any n for which $n > f(n)$ (possible since $\frac{f(n)}{n} \rightarrow 0$) and will use non-uniform specification to bound $|\mathcal{L}_n(X)|$ from below as follows: for any $i \in [0, n - f(n)]$ which is a multiple of $f(n) + N$, choose words $w \in \mathcal{M}_i(X)$ and $v \in \mathcal{N}_{n-f(n)-i}(X)$. Then, since $i, n - i - f(n) \leq n$ and since $f(n)$ is nondecreasing, non-uniform specification of X implies that there exists u with length $f(n)$ so

that $wuv \in \mathcal{L}_n(X)$. We claim that the map from (i, w, v) to wuv (choose u to be the lexicographically minimal option to make this map a function) is one-to-one. To see this, suppose for a contradiction that for choices $(i, w, v) \neq (i', w', v')$, $wuv = w'u'v'$. If $i = i'$, then either $w \neq w'$ or $v \neq v'$, and we have an obvious contradiction. But, if $i \neq i'$, then $|i - i'| \geq f(n) + N$, implying that either w and v' share an N -letter subword or w' and v share an N -letter subword, both contradictions since that word would be in $\mathcal{M}_N(X) \cap \mathcal{N}_N(X)$. Therefore, the map is one-to-one, and so generates at least

$$\sum_{i \in [0, n-f(n)], (f(n)+N)|i} |\mathcal{M}_i(X)| |\mathcal{N}_{n-f(n)-i}(X)| \geq \sum_{i \in [0, n-f(n)], (f(n)+N)|i} e^{ih(X)} e^{(n-f(n)-i)h(X)} \geq \frac{n-f(n)}{f(n)+N} e^{(n-f(n))h(X)}$$

words in $\mathcal{L}_n(X)$. Therefore,

$$(5) \quad |\mathcal{L}_n(X)| \geq \frac{n-f(n)}{f(n)+N} e^{(n-f(n))h(X)}.$$

We also note that for any t and any $w^{(1)}, \dots, w^{(t)} \in \mathcal{L}_n(X)$, we can use non-uniform specification to create a word $w^{(1)}v^{(1)}w^{(2)}v^{(2)} \dots w^{(t)}v^{(t)}$ in $\mathcal{L}_{t(n+f(n))}(X)$, where all $v^{(t)}$ are of length $f(n)$. This clearly implies that

$$|\mathcal{L}_{t(n+f(n))}(X)| \geq |\mathcal{L}_n(X)|^t,$$

and we can take logarithms, divide by t , and let t approach infinity to see that

$$(6) \quad (n+f(n))h(X) \geq \ln |\mathcal{L}_n(X)|.$$

Combining (5) and (6) implies that for large enough n ,

$$(n+f(n))h(X) \geq \ln |\mathcal{L}_n(X)| \geq \ln(n-f(n)) - \ln(f(n)+N) + (n-f(n))h(X).$$

We rewrite as

$$2f(n)h(X) + \ln(f(n)+N) \geq \ln(n-f(n)).$$

However, if we choose a sequence n_k along which $\frac{f(n_k)}{\ln n_k} \rightarrow 0$ and let $k \rightarrow \infty$, then all terms on the left-hand side are $o(\ln(n_k))$, and the right side gets arbitrarily close to $\ln(n_k)$. Therefore, our original assumption was false, completing the proof. \square

4. ALMOST SPECIFICATION

We must begin with some lemmas related to coding theory, namely constructions of small sets which are n -spanning with respect to the Hamming distance.

Definition 4.1. For any alphabet A and $n \in \mathbb{N}$, the **Hamming distance** d on A^n is given by $d(v, w) := |\{i : v_i \neq w_i\}|$, the number of locations at which v and w differ.

Lemma 4.2. For every alphabet A and positive integer n , there exists a set $T_{A,n} \subset A^n$ such that $|T_{A,n}| \leq \frac{1}{2^{\lfloor \log_2 n \rfloor}} |A|^n$ and $T_{A,n}$ is 1-spanning with respect to the Hamming distance d , i.e. for any $w \in A^n$, there exists $t \in T_{A,n}$ s.t. $d(t, w) \leq 1$.

Proof. Choose any A and n , and assume without loss of generality that $A = \{0, \dots, |A| - 1\}$. Define $m = \lfloor \log_2 n \rfloor$, so that $2^m \leq n < 2^{m+1}$. Then, for any $v = v_0 \dots v_{m-1} \in \{0, 1\}^m$, define $T_{A,n,v}$ to be the set of all $w = w_0 \dots w_{n-1} \in A^n$ such that for every $j \in [0, m)$, the sum of w_i over all $i \in [0, 2^m - 1]$ whose binary expansion has a 0 in the 2^j place is equal to $v_j \pmod{2}$. For example, take $A = \{0, 1, 2\}$ and $n = 10$ (so $m = 3$), and $v = 010$. Then, $T_{A,n,v}$ is the set of all $w \in A^n$ for which $w_0 + w_2 + w_4 + w_6 = 0 \pmod{2}$, $w_0 + w_1 + w_4 + w_5 = 1 \pmod{2}$, and $w_0 + w_1 + w_2 + w_3 = 0 \pmod{2}$, and so $0121200111 \in T_{A,n,v}$ and $0211221100 \notin T_{A,n,v}$.

We claim that any set $T_{A,n,v}$ is 1-spanning. To see this, consider any $w \in A^n$. Then, for some values of $j \in [0, m)$, the sum of w_i over all i whose binary expansion has a 0 in the 2^j place is already equal to $v_j \pmod{2}$, and for some it is not. Define $J \subseteq [0, m)$ to be the set of j for which the aforementioned sum is equal to $v_j \pmod{2}$. Then, choose $i \in [0, 2^m)$ so that the binary expansion of i has 0s precisely in 2^j -indexed places for $j \notin J$, i.e. $i = \sum_{j \in J} 2^j$. Note that $0 \leq i \leq 2^m - 1$. We can then define w' to be any word obtained by changing w_i to any letter of A with the opposite parity; then we claim that $w' \in T_{A,n,v}$. This is because the sum of w'_i over all i whose binary expansion has a 0 in the 2^j place is equal to the corresponding sum of w_i if and only if $j \in J$, and so this sum will now always equal $v_j \pmod{2}$.

For instance, continuing the example above: $w = 0211221100 \notin T_{A,n,v}$. In this case, $w_0 + w_2 + w_4 + w_6 = 0 \pmod{2}$, $w_0 + w_1 + w_4 + w_5 \neq 1 \pmod{2}$, and $w_0 + w_1 + w_2 + w_3 = 0 \pmod{2}$, so $J = \{0, 2\}$. Then, we would define $i = 2^0 + 2^2 = 5$, and define w' by changing $w_5 = 2$ to a letter of A with opposite parity, so $w' = 0211211100$. Then, $w' \in T_{A,n,v}$.

Since $T_{A,n,v}$ is a partition of A^n with 2^m choices for v , there exists $T_{A,n,v}$ with cardinality less than or equal to $\frac{1}{2^m}|A|^n$; define $T_{A,n}$ to be that set. \square

Remark 4.3. The sets $T_{A,n}$ are essentially truncated Hamming codes on general alphabets. To say a bit more, the case where $A = \{0, 1\}$ and n is a power of 2 (say $n = 2^m$) is special; it is one of the few cases where a “perfect” code is known to exist, i.e. a set C which is 1-spanning and for which every $w \in A^n$ has a unique $t \in C$ for which $d(w, t) = 1$. This is the Hamming code, and it coincides with our construction exactly for such n and A with $v = 0 \dots 0$.

Since we need such sets for all lengths and alphabets, we simply chose, for any n , the largest power of 2 less than or equal to n (i.e. 2^m), and used a Hamming code on the first 2^m digits. We also used the same parity check idea even for larger alphabets where it is not nearly as efficient, since it still suffices for our purposes.

Lemma 4.4. *For every alphabet A and positive integer n , there exists a set $U_{A,n} \subseteq A^n$ such that $|U_{A,n}| \leq \frac{16}{n^2}|A|^n$ and $U_{A,n}$ is 2-spanning with respect to the Hamming distance d , i.e. for all $w \in A^n$, there exists $u \in U_{A,n}$ s.t. $d(u, w) \leq 2$.*

Proof. For any n , we simply define $T_{A, \lfloor 0.5n \rfloor}$ and $T_{A, \lceil 0.5n \rceil}$ as above, and define $U_{A,n} = \{uv : u \in T_{A, \lfloor 0.5n \rfloor}, v \in T_{A, \lceil 0.5n \rceil}\}$. Then $U_{A,n}$ is 2-spanning; for any $w \in A^n$, at most one change is required to change its prefix of length $\lfloor 0.5n \rfloor$ to a word in $T_{A, \lfloor 0.5n \rfloor}$, and at most one change is required to change its suffix of length $\lceil 0.5n \rceil$ to a word in $T_{A, \lceil 0.5n \rceil}$. It's not hard to check that $2^{\lfloor \log_2(\lfloor 0.5n \rfloor) \rfloor} \geq 0.25n$.

Then, by Lemma 4.2,

$$|U_{A,n}| = |T_{A,\lfloor 0.5n \rfloor}| |T_{A,\lceil 0.5n \rceil}| \leq \frac{1}{0.25n} |A|^{\lfloor 0.5n \rfloor} \cdot \frac{1}{0.25n} |A|^{\lceil 0.5n \rceil} = \frac{16}{n^2} |A|^n.$$

□

Lemma 4.5. *For every alphabet A and positive integer $n > 1$, there exists a set $V_{A,n} \subset A^n$ such that $|V_{A,n}| = \frac{1}{|A|^2} |A|^n$ and $V_{A,n}$ is 2-spanning with respect to the Hamming distance d .*

Proof. Simply define $V_{A,n} = \{w = w_1 w_2 \dots w_n \in A^n : w_1 = w_2 = 1\}$. The reader may check that $V_{A,n}$ has the desired properties. □

Proof of Theorem 1.2. We define X to have alphabet $A = \{-N, \dots, -1, 1, \dots, N\}$, where N will be defined later. For a parameter ℓ , also to be determined later, we define

$$P_n = \begin{cases} \{1\} & n = 1 \\ V_{\{1, \dots, N\}, n} & 1 < n \leq \ell \\ U_{\{1, \dots, N\}, n} & n > \ell \end{cases}$$

and

$$N_n = \begin{cases} \{-1\} & n = 1 \\ V_{\{-N, \dots, -1\}, n} & 1 < n \leq \ell \\ U_{\{-N, \dots, -1\}, n} & n > \ell \end{cases}$$

where $U_{A,n}$ and $V_{A,n}$ are defined as in Lemmas 4.4 and 4.5. Clearly then $|N_n| = |P_n|$ for all n ; we denote their common value by M_n . We note that by construction, all N_n and P_n are 2-spanning in $\{-N, \dots, -1\}^n$ and $\{1, \dots, N\}^n$ respectively.

Then, we define X via the list of forbidden words \mathcal{F} consisting of

- All words nPn' where $n, n' < 0$, P consists of positive letters, and $P \notin \bigcup P_n$ and
- All words pNp' where $p, p' < 0$, N consists of negative letters, and $N \notin \bigcup N_n$.

Points of X then consist of biinfinite concatenations of words of constant sign, each of which must be in either some P_n or N_n , depending on its sign and length. There are also transient points of X which have one or more infinite words of constant sign, on which there are no restrictions. We note for future reference that, given any word $w \in A^*$ which does not contain a forbidden word as described in the above list, $w \in \mathcal{L}(X)$, since w can clearly be extended to a point of X by appending arbitrary letters with the same sign as the first letter of w to the left and arbitrary letters with the same sign as the last letter of w to the right.

We first show that X has almost specification with gap function $g(n) = 4$. For this purpose, consider any words $w^{(1)}, \dots, w^{(k)} \in \mathcal{L}(X)$. Then, the concatenation $w = w^{(1)}w^{(2)} \dots w^{(k)}$ might not be in $\mathcal{L}(X)$. However, we can turn this into a word in $\mathcal{L}(X)$ by making changes to each maximal word of constant sign within w which place them in either some P_n or N_n . By the 2-spanning property of all P_n and N_n , we can change at most 2 letters in each maximal word of constant sign within w and create a new word $w' \in \mathcal{L}(X)$. No maximal word of constant sign which is not a prefix or suffix of $w^{(i)}$ would have required a change, since $w^{(i)} \in \mathcal{L}(X)$, implying that any such word would have been in some P_n or N_n anyway.

Therefore, when w was changed to w' , no more than 4 changes would have been made in any $w^{(i)}$, those being only in the words of constant sign at the beginning and end at $w^{(i)}$. This completes the proof of almost specification with $g(n) = 4$.

We wish to show that X has exactly two ergodic measures of maximal entropy (with disjoint supports), and again we proceed by showing that $h(X) = \ln N$. Since X contains the full shifts on N positive symbols and N negative symbols respectively, $h(X) \geq \ln N$, and so it suffices to show that $h(X) \leq \ln N$. For this, we will just bound $\mathcal{L}_n(X)$ from above for all n . Every $w \in \mathcal{L}_n(X)$ consists of a concatenation of words of constant sign. Therefore, we can parametrize elements of $\mathcal{L}_n(X)$ by the number $k \geq 1$ of such concatenated words and their lengths n_1, n_2, \dots, n_k , which clearly must sum to n . We then see that

$$(7) \quad |\mathcal{L}_n(X)| = 2N^n + 2(n-1)N^n + \sum_{k=3}^n \sum_{\substack{n_1, \dots, n_k \\ \sum n_i = n}} 2N^{n_1} \left(\prod_{j=2}^{k-1} M_{n_j} \right) N^{n_k}.$$

Here, the first two terms correspond to the cases $k = 1, 2$. In each term, the factor of 2 comes from choosing the sign of the first word of constant sign, after which all signs are forced. The N^{n_1} and N^{n_k} in the third term are from the prefix and suffix of w of constant sign, on which there are no restrictions, and the M_j represent the number of choices for the other subwords of constant sign. We now bound the third term of (7) from above.

$$(8) \quad \sum_{k=3}^n \sum_{\substack{n_1, \dots, n_k \\ \sum n_i = n}} 2N^{n_1} \left(\prod_{j=2}^{k-1} M_{n_j} \right) N^{n_k} = 2N^n \sum_{k=3}^n \sum_{\substack{n_1, \dots, n_k \\ \sum n_i = n}} \prod_{j=2}^{k-1} (M_{n_j} N^{-n_j}) = \\ 2N^n \sum_{k=3}^n \sum_{\substack{n_2, \dots, n_{k-1} \\ \sum n_i < n}} \left(n - \sum_{j=2}^{k-1} n_j - 1 \right) \prod_{j=2}^{k-1} (M_{n_j} N^{-n_j}) \leq \\ 2nN^n \sum_{k=3}^n \sum_{\substack{n_2, \dots, n_{k-1} \\ \sum n_i < n}} \prod_{j=2}^{k-1} (M_{n_j} N^{-n_j}) \leq 2nN^n \sum_{k=3}^n \left(\sum_{t=1}^{\infty} M_t N^{-t} \right)^{k-2}.$$

We now bound $\sum_{t=1}^{\infty} M_t N^{-t}$ by using the bounds of Lemmas 4.4 and 4.5:

$$\sum_{t=1}^{\infty} M_t N^{-t} = \frac{1}{N} + \sum_{t=2}^{\ell} |V_{\{1, \dots, N\}, t}| N^{-t} + \sum_{t=\ell+1}^{\infty} |U_{\{1, \dots, N\}, t}| N^{-t} \leq \\ N^{-1} + \sum_{t=2}^{\ell} N^{-2} + \sum_{t=\ell+1}^{\infty} 16t^{-2} \leq N^{-1} + (\ell-1)N^{-2} + 16\ell^{-1}.$$

Choose N and ℓ to be any pair for which this expression is less than 1, for instance $N = 10$ and $\ell = 32$ (then $N^{-1} + (\ell-1)N^{-2} + 16\ell^{-1} = 0.1 + 0.31 + 0.5 = 0.91 < 1$), and denote $\sum_{t=1}^{\infty} M_t N^{-t}$ by $\alpha < 1$. Then, (7) and (8) imply that

$$(9) \quad |\mathcal{L}_n(X)| \leq 2N^n + 2(n-1)N^n + 2nN^n \sum_{k=3}^n \alpha^{k-2} = \frac{2n}{1-\alpha} N^n.$$

This clearly implies that $h(X) \leq \ln N$, and since again X contains a full shift on N symbols, that $h(X) = \ln N$ as well.

Now, consider any ergodic measure of maximal entropy μ of X . Our goal is to show that $\mu([ij]) = 0$ for every i, j of opposite sign. For a contradiction, assume that $\mu([ij]) = \beta > 0$ for some such i, j . Define, for every n , the set $C_{n,0.5\beta,ij}(X)$ of n -letter words with between $0.5n\beta$ and $1.5n\beta$ occurrences of ij . Again we apply Lemma 2.12 to see that $\liminf_{n \rightarrow \infty} \frac{\ln |C_{n,0.5\beta,ij}|}{n} \geq h(\mu) = h(X) = \ln N$. However, in the decomposition above, k must be at least the number of such occurrences, and so the same argument used above to show (9) implies that

$$|C_{n,0.5\beta,ij}(X)| < 2nN^n \sum_{k=[0.5n\beta]}^{\infty} \alpha^{k-2} \leq \frac{2n}{1-\alpha} N^n \alpha^{0.5n\beta}.$$

Then $\liminf_{n \rightarrow \infty} \frac{\ln |C_{n,0.5\beta,ij}(X)|}{n} \leq \ln N + 0.5\beta \ln \alpha < \ln N$, a contradiction. Therefore, $\mu([ij]) = 0$ for all i, j of opposite sign, which clearly implies that μ is supported entirely on $\{1, \dots, N\}^{\mathbb{Z}}$ or $\{-N, \dots, -1\}^{\mathbb{Z}}$. Again, the only such measures with entropy $\ln N$ are the uniformly distributed Bernoulli measures over $\{1, \dots, N\}$ and $\{-N, \dots, -1\}$, and so X has exactly two ergodic measures of maximal entropy, with disjoint supports. \square

For the proof of Theorem 1.4, we require one more lemma related to coding theory.

Lemma 4.6. *For every alphabet A , positive integer n , and set $W \subset A^n$, there exists a set $S \subset W$ such that $|S| \geq \frac{|W|}{4n|A|^2}$ and W is 3-separated with respect to the Hamming distance d , i.e. for all $w, w' \in W$, $d(w, w') \geq 3$.*

Proof. Choose any A and n , and again assume without loss of generality that $A = \{0, \dots, |A| - 1\}$. Then, for any $i \in [0, 2|A|)$ and $j \in [0, 2|A|n)$, define $S_{n,i,j} = \{w = w_1 \dots w_n \in W : \sum_{k=1}^n w_k = i \pmod{2|A|}, \sum_{k=1}^n kw_k = j \pmod{2|A|n}\}$. We claim that each $S_{n,i,j}$ is 3-separated with respect to the Hamming distance. It is obvious that changing a single letter of a word in $S_{n,i,j}$ cannot yield another word in $S_{n,i,j}$ since changing a single letter must change the sum $\sum_{k=1}^n w_k \pmod{2|A|}$. Suppose for a contradiction that there exist $w \neq w' \in S_{n,i,j}$ differing on exactly two letters. Then $\sum_{k=1}^n w_k$ and $\sum_{k=1}^n w'_k$ are both equal to $i \pmod{2|A|}$, and differ by at most $2(|A| - 1)$, and so must be equal. Similarly, $\sum_{k=1}^n kw_k$ and $\sum_{k=1}^n kw'_k$ are both equal to $j \pmod{2|A|n}$, and differ by at most $n(|A| - 1) + (n - 1)(|A| - 1)$, and are therefore also equal. But recall that w and w' differ on exactly two letters, say those indexed by s and t . Then, $w_s + w_t = w'_s + w'_t$ and $sw_s + tw_t = sw'_s + tw'_t$, which implies that $w_s = w'_s$ and $w_t = w'_t$, a contradiction. We have then shown that each $S_{n,i,j}$ is 3-separated.

Since the sets $S_{n,i,j}$ clearly partition W , and there are $4n|A|^2$ choices for the pair i, j , one of the $S_{n,i,j}$ has cardinality at least $\frac{|W|}{4n|A|^2}$; define S to be that set. \square

Proof of Theorem 1.4. Suppose for a contradiction that X is a subshift with almost specification with gap function $g(n) = 1$ and two measures of maximal entropy μ, ν with disjoint supports. For every n , as we did in Theorem 1.3, again define $\mathcal{M}_n(X) = \{w \in \mathcal{L}_n(X) : \mu([w]) > 0\}$ and $\mathcal{N}_n(X) = \{w \in \mathcal{L}_n(X) : \nu([w]) > 0\}$.

We again note that by Lemma 2.11, $|\mathcal{M}_n(X)| \geq e^{nh(\mu)} = e^{nh(X)}$ and $|\mathcal{N}_n(X)| \geq e^{nh(\mu)} = e^{nh(X)}$ for all n , and that there exists N so that $\mathcal{M}_n(X) \cap \mathcal{N}_n(X) = \emptyset$ for all $n \geq N$.

For every n , we use Lemma 4.6 to define sets $\mathcal{M}'_n \subseteq \mathcal{M}_n(X)$ and $\mathcal{N}'_n \subseteq \mathcal{N}_n(X)$ which are 3-separated in the Hamming distance and for which $|\mathcal{M}'_n|, |\mathcal{N}'_n| \geq \frac{e^{nh(X)}}{4n|A|^2}$. We make the notation $S_n := \min(|\mathcal{M}'_n|, |\mathcal{N}'_n|)$. We now proceed somewhat as in the proof of Theorem 1.3, in that we will make many words in $\mathcal{L}_n(X)$ by using almost specification to nearly concatenate words in \mathcal{M}'_j and \mathcal{N}'_j for various $j < n$. However, instead of concatenating only two words, we now will need arbitrarily many. First, we choose t such that

$$(10) \quad \sum_{j=1}^t \frac{1}{12iN|A|^2} > 1,$$

and we denote this sum by α .

Now, we choose any $n > 3tN$ and create words in $\mathcal{L}_n(X)$ in the following way: define $k = \lfloor n/3tN \rfloor$, and define any $n_i \in [1, t]$ for $1 \leq i \leq k$. We note that $k > n/6tN$ since $n > 3tN$. Then, choose any words $w_1 \in \mathcal{M}'_{3Nn_1}$, $w_2 \in \mathcal{N}'_{3Nn_2}$, and so on, alternating between the sets, until finishing with w_k in either \mathcal{M}'_{3Nn_k} or \mathcal{N}'_{3Nn_k} , depending on whether k is odd or even, respectively. Finally, choose w_{k+1} in either $\mathcal{N}'_{n-\sum_{i=1}^k 3Nn_i}$ or $\mathcal{M}'_{n-\sum_{i=1}^k 3Nn_i}$, whichever is the opposite of what was chosen for w_k . For whichever words were chosen, use the assumed almost specification of X with $g(n) = 1$ to make a word $f(w_1, \dots, w_{k+1}) = v_1 v_2 \dots v_{k+1} \in \mathcal{L}_n(X)$, where each v_i differs from w_i on at most one letter.

We claim that this operation is injective, i.e. $f(w_1, \dots, w_{k+1}) \neq f(w'_1, \dots, w'_{k+1})$ unless $w_i = w'_i$ for $1 \leq i \leq k$. Assume for a contradiction that k -tuples (w_i) and (w'_i) exist for which $f(w_1, \dots, w_{k+1}) = f(w'_1, \dots, w'_{k+1})$. There are two cases.

The first case is where $n_i = n'_i$ for all i . Then, there exists j so that $w_j \neq w'_j$. Also, since $f(w_1, \dots, w_{k+1}) = f(w'_1, \dots, w'_{k+1})$, both w_j and w'_j become the same word v with at most one changed letter in each. Since v and w_j differ on at most one letter and v and w'_j differ on at most one letter, w_j and w'_j differ on at most two letters. This contradicts the 3-separated property of \mathcal{M}'_{n_j} and \mathcal{N}'_{n_j} .

The second case is where $n_j \neq n'_j$ for some j . Choose j minimal so that $n_j \neq n'_j$, and assume without loss of generality that $n_j < n'_j$. Then since all w_i and w'_i have lengths which are multiples of $3N$, the subwords u and u' of $w_1 w_2 \dots w_{k+1}$ and $w'_1 w'_2 \dots w'_{k+1}$ respectively of length $3N$ beginning at index $\sum_{i=1}^j 3Nn_i + 1$ are subwords of w_{j+1} and w'_{j+1} respectively. Then either $u \in \mathcal{M}'_{n_{j+1}}$ and $u' \in \mathcal{N}'_{n'_j}$ or $u \in \mathcal{N}'_{n_{j+1}}$ and $u' \in \mathcal{M}'_{n'_j}$. Also, since $f(w_1, \dots, w_{k+1}) = f(w'_1, \dots, w'_{k+1})$, u and u' become the same word u'' after making at most one change to each. Since $|u''| = 3N$, u'' must contain a subword of length at least N which was unchanged in both u and u' , and therefore is a subword of each, a contradiction since no word in $\bigcup_i \mathcal{M}_i(X)$ can share an N -letter subword with a word in $\bigcup_i \mathcal{N}_i(X)$.

We have shown that f is one-to-one, and so generates at least

$$\sum_{\substack{n_1, \dots, n_k \\ 1 \leq n_i \leq t}} \left(\prod_{i=1}^k S_{3Nn_i} \right) S_{n-\sum_{i=1}^k 3Nn_i}$$

words in $\mathcal{L}_n(X)$. Then, by Lemma 4.6,

$$\begin{aligned} & \sum_{\substack{n_1, \dots, n_k \\ 1 \leq n_i \leq t}} \left(\prod_{i=1}^k S_{3Nn_i} \right) S_{n - \sum_{i=1}^k 3Nn_i} \geq \\ & \sum_{\substack{n_1, \dots, n_k \\ 1 \leq n_i \leq t}} \left(\prod_{i=1}^k \frac{e^{3Nn_i h(X)}}{12N|A|^2 n_i} \right) \frac{e^{(n - \sum_{i=1}^k 3Nn_i)h(X)}}{4|A|^2 (n - \sum_{i=1}^k 3Nn_i)} \geq \frac{e^{nh(X)}}{4n|A|^2} \sum_{\substack{n_1, \dots, n_k \\ 1 \leq n_i \leq t}} \prod_{i=1}^k \frac{1}{12N|A|^2 n_i} \\ & = \frac{e^{nh(X)}}{4n|A|^2} \left(\sum_{j=1}^t \frac{1}{12N|A|^2 j} \right)^k \geq \frac{e^{nh(X)}}{4n|A|^2} \alpha^{n/6tN}. \end{aligned}$$

Therefore, $\mathcal{L}_n(X) \geq (e^{nh(X)}/(4n|A|^2))\alpha^{n/6tN}$ for all $n > 3tN$. However, taking logarithms, dividing by n , and letting $n \rightarrow \infty$ would imply that $h(X) \geq h(X) + \frac{1}{6tN} \ln \alpha$, a contradiction since $\alpha > 1$ by (10). Therefore, our original assumption was false, and measures of maximal entropy μ and ν on X with disjoint supports cannot exist, completing the proof. \square

ACKNOWLEDGMENTS

The author would like to thank the anonymous referee for some very useful comments and suggestions, one of which greatly simplified the proof of Theorem 1.1. The author would also like to thank Dan Thompson for several profitable discussions about almost specification, and Anthony Quas for pointing out the connection between our examples and those from [8].

REFERENCES

- [1] Anne Bertrand. Specification, synchronisation, average length. In *Coding theory and applications (Cachan, 1986)*, volume 311 of *Lecture Notes in Comput. Sci.*, pages 86–95. Springer, Berlin, 1988.
- [2] Rufus Bowen. Periodic points and measures for Axiom A diffeomorphisms. *Trans. Amer. Math. Soc.*, 154:377–397, 1971.
- [3] Robert Burton and Jeffrey E. Steif. New results on measures of maximal entropy. *Israel J. Math.*, 89(1-3):275–300, 1995.
- [4] Vaughn Climenhaga and Daniel J. Thompson. Intrinsic ergodicity beyond specification: β -shifts, S -gap shifts, and their factors. *Israel J. Math.*, 192(2):785–817, 2012.
- [5] Vaughn Climenhaga, Daniel J. Thompson, and Kenichiro Yamamoto. Large deviations for systems with non-uniform structure. *Trans. Amer. Math. Soc.*, to appear.
- [6] Masahito Dateyama. The almost weak specification property for ergodic group automorphisms of abelian groups. *J. Math. Soc. Japan*, 42(2):341–351, 1990.
- [7] Manfred Denker, Christian Grillenberger, and Karl Sigmund. *Ergodic theory on compact spaces*. Lecture Notes in Mathematics, Vol. 527. Springer-Verlag, Berlin-New York, 1976.
- [8] Nicolai T. A. Haydn. Phase transitions in one-dimensional subshifts. *Discrete Contin. Dyn. Syst.*, 33(5):1965–1973, 2013.
- [9] Marcin Kulczycki, Dominik Kwietniak, and Piotr Oprocha. On almost specification and average shadowing properties. *Fund. Math.*, 224(3):241–278, 2014.
- [10] Dominik Kwietniak, Piotr Oprocha, and Michal Rams. On entropy of dynamical systems with almost specification. Submitted, 2014.
- [11] Brian Marcus. A note on periodic points for ergodic toral automorphisms. *Monatsh. Math.*, 89(2):121–129, 1980.
- [12] Ronnie Pavlov. Perturbations of multidimensional shifts of finite type. *Ergodic Theory Dynam. Systems*, 31(2):483–526, 2011.

- [13] Karl Petersen. Chains, entropy, coding. *Ergodic Theory Dynam. Systems*, 6(3):415–448, 1986.
- [14] Charles-Edouard Pfister and Wayne G. Sullivan. Large deviations estimates for dynamical systems without the specification property. Applications to the β -shifts. *Nonlinearity*, 18(1):237–261, 2005.
- [15] Charles-Edouard Pfister and Wayne G. Sullivan. On the topological entropy of saturated sets. *Ergodic Theory Dynam. Systems*, 27(3):929–956, 2007.
- [16] Anthony Quas and Terry Soo. Ergodic universality of some topological dynamical systems. *Trans. Amer. Math. Soc.*, 368(6):4137–4170, 2016.
- [17] Daniel J. Thompson. Irregular sets, the β -transformation and the almost specification property. *Trans. Amer. Math. Soc.*, 364(10):5395–5414, 2012.
- [18] Paulo Varandas. Non-uniform specification and large deviations for weak Gibbs measures. *J. Stat. Phys.*, 146(2):330–358, 2012.
- [19] Peter Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.
- [20] Kenichiro Yamamoto. On the weaker forms of the specification property and their applications. *Proc. Amer. Math. Soc.*, 137(11):3807–3814, 2009.

RONNIE PAVLOV, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF DENVER, 2280 S. VINE ST., DENVER, CO 80208

E-mail address: rpavlov@du.edu

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