CLASS DEGREE AND RELATIVE MAXIMAL ENTROPY

MAHSA ALLAHBAKHSHI AND ANTHONY QUAS

Abstract. Given a factor code \( \pi \) from a one-dimensional shift of finite type \( X \) onto an irreducible sofic shift \( Y \), if \( \pi \) is finite-to-one there is an invariant called the degree of \( \pi \) which is defined as the number of preimages of a typical point in \( Y \). We generalize the notion of the degree to the class degree which is defined for any factor code on a one-dimensional shift of finite type. Given an ergodic measure \( \nu \) on \( Y \), we find an invariant upper bound on the number of ergodic measures on \( X \) which project to \( \nu \) and have maximal entropy among all measures in the fibre \( \pi^{-1}\{\nu\} \). We show that this bound and the class degree of the code agree when \( \nu \) is ergodic and fully supported. One of the main ingredients of the proof is a uniform distribution property for ergodic measures of relative maximal entropy.

1. Introduction

It is a well-known result that a one-dimensional irreducible shift of finite type on a finite alphabet has a unique measure of maximal entropy, the measure developed by Shannon and Parry \[15, 20\]. In contrast, we consider the relative case in which one is given a factor code \( \pi : X \to Y \) from a shift of finite type \( X \) to a sofic shift \( Y \), and a measure \( \nu \) on \( Y \). In this case ergodic measures on \( X \) in the fibre \( \pi^{-1}\{\nu\} \) having maximal entropy in the fibre, so-called measures of relative maximal entropy, are not well understood.

Measures of relative maximal entropy appear frequently in different topics in symbolic dynamics, and have been applied to the problem of computing the Hausdorff dimension of certain sets by Gatzouras and Peres \[7\]. There are connections of measures of relative maximal entropy with functions of Markov chains \[1, 3, 4, 14\], measures that maximize a weighted entropy functional \[6, 21\], the theory of pressure and equilibrium states \[8, 9, 19\], relative pressure and relative equilibrium states \[11, 12, 23\], and compensation functions \[3, 23\]. Other uses of such measures arise from their application in the mathematics of information transfer \[16\] and information-compressing channels \[14\].

Unlike in the case considered by Shannon and Parry where the measure of maximal entropy is unique, it is known that measures of relative maximal entropy may not be unique \[16, 17\]. It is natural to ask how many of such measures there may be for a particular factor code \( \pi : X \to Y \) and an ergodic invariant measure on \( Y \). A previous result on this problem is the result of Petersen, Quas, and Shin \[17\] which established an explicit finite upper bound on the number of measures of relative maximal entropy. They considered the case (always achievable by recoding)
that \( \pi \) is a one-block code from a one-step shift of finite type. In this case, the main result of [17] states that the number of measures of relative maximal entropy is bounded above by the number of preimages of any symbol \( w \) of \( Y \) which has positive probability with respect to the measure on \( Y \).

The bound in [17] suffers from not being invariant under conjugacy and becomes arbitrarily large simply by recoding to a higher block presentation. To avoid this issue one possibility is to take the minimum of this bound over all shifts of finite type \( \tilde{X} \) which are conjugate to \( X \). However, no algorithm for computing this minimum has been found. On the other hand, in this paper we find a simpler conjugacy-invariant upper bound which can be strictly better and is never worse than the minimum-over-conjugates bound mentioned above.

Given a finite-to-one factor code \( \pi \) from a shift of finite type \( X \) to a sofic shift \( Y \), the degree is defined to be the minimal cardinality of the set of preimages of a point in \( Y \). This is widely-studied and known to be invariant under recoding [13]. One can show that the number of preimages of every transitive point in \( Y \) is exactly the degree of \( \pi \). Previous efforts to find an analogue of this when \( \pi \) is infinite-to-one have proved unsuccessful [2].

Given \( y \) in \( Y \), we define an equivalence relation on \( \pi^{-1}(y) \) which is respected by recoding. The relation is motivated by communicating classes in Markov chains. Loosely, two points \( x \) and \( x' \) in \( \pi^{-1}(y) \) lie in the same equivalence class, “transition class”, if one can find an element of \( \pi^{-1}(y) \) matching either \( x \) or \( x' \) from \( -\infty \) to an arbitrarily large positive coordinate \( n \) and matching the other at \( +\infty \). The minimal number of transition classes (always finite) in \( \pi^{-1}(y) \) when \( y \) runs over the points in \( Y \) is called the class degree of the code. We show that the number of transition classes over every transitive point of \( Y \) is exactly the class degree. If \( \pi \) is finite-to-one the degree and the class degree of \( \pi \) agree.

We prove an analogue of Theorem 1 in [17]: given an ergodic measure \( \nu \) on \( Y \), for distinct ergodic measures \( \mu_1 \) and \( \mu_2 \) of relative maximal entropy, the relative independent joining of \( \mu_1 \) and \( \mu_2 \) assigns measure zero to the set of points \( (u,v) \) such that \( u,v \) are in a common transition class.

Using a combinatorial characterization of class degree and further arguments, we show that the number of transition classes above a typical point of \( \nu \) is an upper bound on the number of ergodic measures of maximal entropy in the fibre \( \pi^{-1}\{\nu}\). This bound is equal to the class degree of the code when \( \nu \) is ergodic and fully supported.

One main ingredient to this work is a relative version of the uniform conditional distribution property for measures of maximal entropy. We follow techniques developed by Burton and Steif [5] (which were related to an earlier work of Lanford and Ruelle [10]) to show a uniform conditional distribution property for measures of relative maximal entropy.

2. Background

Throughout the paper, \( X \) will denote a one-dimensional shift of finite type on a finite alphabet (SFT) and the shift map will be denoted by \( T_X \), or simply by \( T \) when there is no ambiguity. We will deal with a factor code \( \pi \) from \( X \) to a sofic shift \( Y \). A detailed description of these notions is given in [13]. A triple \((X,Y,\pi)\) is called a factor triple when \( \pi : X \to Y \) is a factor code from an SFT \( X \) to a sofic shift \( Y \). The alphabet of a shift space \( X \) is denoted by \( A(X) \) and the \( \sigma \)-algebra
on \(X\) generated by cylinder sets of \(X\) is denoted by \(\mathcal{B}_{X}\). The set of all \(n\)-blocks that occur in points of \(X\) is denoted by \(\mathcal{B}_{n}(X)\), and the language of \(X\), \(\mathcal{L}(X)\), is the collection \(\bigcup_{n=0}^{\infty} \mathcal{B}_{n}(X)\). Let \(x\) be in \(X\) and \(G \subseteq \mathbb{Z}\). The configuration which occurs in \(x\) on \(G\) is denoted by \(x_{G}\). If \(G = [i, j]\) is a \(\mathbb{Z}\)-interval we sometimes denote \(x_{G}\) by \(x_{[i,j]}\).

We say two factor triples \((X, Y, \pi)\) and \((\tilde{X}, \tilde{Y}, \tilde{\pi})\) are conjugate, and denote it by \((X, Y, \pi) \cong (\tilde{X}, \tilde{Y}, \tilde{\pi})\), if \(X\) is conjugate to \(\tilde{X}\) under a conjugacy \(\phi\), \(Y\) is conjugate to \(\tilde{Y}\) under a conjugacy \(\psi\), and \(\tilde{\pi} \circ \phi = \psi \circ \pi\). Let \((X, Y, \pi)\) be a factor triple, where \(\pi\) is a one-block factor code induced by the map \(\pi_{b}: \mathcal{A}(X) \to \mathcal{A}(Y)\) (b stands for block). The map \(\pi_{b}\) naturally extends to elements of \(\mathcal{L}(X)\). Above every \(Y\)-block \(W\) of length \(n\) there is a set of \(X\)-blocks \(W'\) of length \(n\) which are sent to \(W\) by \(\pi_{b}\); i.e., \(\pi_{b}(W') = W\). Given \(0 \leq i \leq n - 1\), define

\[
d(W, i) = |\{a \in \mathcal{A}(X) : \exists W' \text{ with } \pi_{b}(W') = W, W'_i = a\}|,
\]

and let

\[
d_{\pi} = \min\{d(W, i) : W \in \mathcal{L}(Y), 0 \leq i \leq |W| - 1\}.
\]

A magic block is a block \(W\) such that \(d(W, i) = d_{\pi}\) for some \(0 \leq i \leq |W| - 1\). Such an index \(i\) is called a magic coordinate of \(W\). A factor code \(\pi\) has a magic symbol if there is a magic block of \(\pi\) of length 1.

**Theorem 2.1** (\cite{13} Theorem 9.1.7). Let \((X, Y, \pi)\) be a factor triple. There is a factor triple \((\tilde{X}, \tilde{Y}, \tilde{\pi})\) conjugate to \((X, Y, \pi)\) such that \(\tilde{\pi}\) is a one-block code with a magic symbol.

### 3. Uniform conditional distribution

By a well-known result of Parry \cite{15}, generalizing an earlier result of Shannon \cite{20}, in one dimension every irreducible shift of finite type on a finite alphabet has a unique measure of maximal entropy. Burton and Steif \cite{51} give a counterexample to this statement in higher dimensions. However, they show that such measures all have the uniform conditional distribution property stated in Theorem \ref{3.1} (which is related to an earlier work of Lanford and Ruelle \cite{10}). Given a finite set \(G \subseteq \mathbb{Z}^{d}\), the boundary of the complement of \(G\) is \(\partial G^{c} = \{i \in G^{c} : \exists j \in G \text{ with } ||i - j|| = 1\}\).

**Theorem 3.1** (\cite{5} Proposition 1.19)]. Let \(X\) be a one-step \(d\)-dimensional SFT and \(\mu\) be a measure of maximal entropy on \(X\). Then the conditional distribution of \(\mu\) on any finite set \(G \subseteq \mathbb{Z}^{d}\) given the configuration on \(G^{c}\) is \(\mu\)-a.s. uniform over all configurations on \(G\) which extend the configuration on \(\partial G^{c}\).

Given a factor triple \((X, Y, \pi)\) and an ergodic measure \(\nu\) on \(Y\), there can exist more than one ergodic measure of relative maximal entropy over \(\nu\); i.e., there can be more than one ergodic measure on \(X\) which projects to \(\nu\) and has maximal entropy among all measures in the fibre \(\pi^{-1}\{\nu\}\); see \cite{17} Example 3.3. We use Lemma \ref{3.2} and follow techniques developed by Burton and Steif in the proof of Theorem \ref{3.1} to show the uniform conditional distribution property for measures of relative maximal entropy in Theorem \ref{3.3}

**Lemma 3.2** (\cite{22} Theorem 4.7]). Let \((X, \mathcal{B}, \mu)\) be a probability space. Let \(\mathcal{A}\) be a finite subalgebra of \(\mathcal{B}\) and let \((\mathcal{F}_{n})_{n=1}^{\infty}\) be an increasing sequence of \(\sigma\)-subalgebras of \(\mathcal{B}\) with \(\bigvee_{n=1}^{\infty} \mathcal{F}_{n} = \mathcal{B}\). Then \(H(\mathcal{A}|\mathcal{F}_{n}) \to H(\mathcal{A}|\mathcal{F})\).
Theorem 3.3. Let \( \pi : X \to Y \) be a one-block factor code from a one-step, one-dimensional SFT \( X \) to a sofic shift \( Y \). Let \( \nu \) be an invariant measure on \( Y \), and \( \mu \) be an invariant measure of relative maximal entropy over \( \nu \). Given a finite set \( G \subseteq \mathbb{Z} \) and \( y \) in \( Y \) the conditional distribution of \( \mu \) on \( \pi^{-1}(y) \) restricted to \( G \) given the configuration on \( G^c \) is \( \mu \)-a.s. uniform over all configurations on \( G \) which extend the configuration on \( \partial G^c \) and map to the same configuration in \( Y \) under the factor code \( \pi \).

Proof. Let \( G \) be a finite subset of \( \mathbb{Z} \). Let \( \Delta \) be a configuration of \( Y \) on \( G \). Pick a configuration \( \eta \) of \( X \) on \( \partial G^c \) such that \( \mu(\eta \cap \pi^{-1}(\Delta)) > 0 \). Starting from \( \mu \), we define a measure \( \tilde{\gamma} \) on \( X \) by uniformizing over preimages of \( \Delta \) that have \( \eta \) on the boundary. We then show that if \( \mu \) does not have the required uniform conditional distribution property, then \( \tilde{\gamma} \) has greater entropy than \( \mu \), but still is an element of the fibre \( \pi^{-1}\{\nu\} \). This is a contradiction and will therefore establish the required uniform conditional distribution property of \( \mu \).

Let \( D = \{\alpha_1, \ldots, \alpha_L\} \) be the set of all configurations on \( G \) which extend \( \eta \) and map to \( \Delta \) under the factor code \( \pi \). Let the \( \mathbb{Z} \)-interval \( R = [-m, m - 1] \) be large enough so that \( G \cup \partial G^c \subseteq R \) and let \((R_n)_{n \in \mathbb{Z}}\) be the partition of \( \mathbb{Z} \) by translates of \( R \) (\( R_0 = R \)): i.e. \( R_n = [(2n - 1)m, (2n + 1)m - 1] \). Let \( G_n \) and \( \partial G_n^c \) be the corresponding translates of \( G \) and \( \partial G^c \) in \( R_n \). Given \( S \subseteq \mathbb{Z} \) let \( \mathcal{P}(X_S) \) be the partition of \( X \) generated by the configurations of \( X \) on \( S \), and \( \sigma(X_S) \) be the \( \sigma \)-algebra generated by \( \mathcal{P}(X_S) \). When \( S = [a, b] \) is a \( \mathbb{Z} \)-interval we sometimes denote \( X_S \) by \( X_a^b \). Considering \( R \) above, \( \sigma(X_{-m}^{-1}) \) is the finite \( \sigma \)-algebra generated by the partition

\[
\mathcal{P}(X_{-m}^{-1}) = \{ -m[x_m x_{m-1} \ldots x_{m-1}]_{m-1} : x_m x_{m-1} \ldots x_{m-1} \in \mathcal{L}(X) \},
\]

and \( \sigma(X_{-m}^{-1}) = \sigma(X_{-3m}^{-1}) \lor \sigma(X_{-5m}^{-1}) \lor \ldots \).

We obtain the measure \( \gamma \) from the measures \( \mu \) and \( \eta \) as follows. Define the map \( \Phi : X \times D^\mathbb{Z} \to X \) by

\[
\Phi(x, \zeta)_{G_n} = \begin{cases} 
\zeta_n & \text{if } x_{\partial G_n^c} = \eta \text{ and } x_{G_n} \in D \\
x_{G_n} & \text{otherwise,} 
\end{cases}
\]

and \( \Phi(x, \zeta)_{G_n \cap R_n} = x_{G_n \cap R_n} \) for each integer \( n \). Since \( \zeta_n \) is in \( D \) and each element of \( D \) extends \( \eta \), the assumption that \( X \) is a one-step SFT implies that \( \Phi(x, \zeta) \) lies in \( X \). For each \( \zeta \) in \( D^\mathbb{Z} \) we have \( \pi(x) = \pi(\Phi(x, \zeta)) \) since \( \Phi(x, \zeta) \) and \( x \) are the same except having alternative \( \alpha_i \)'s in the same positions. Let \( C \) be a cylinder set of \( X \). Define \( \gamma(C) = (\mu \times \lambda)\Phi^{-1}(C) \), where \( \lambda \) is the Bernoulli \( (1/L, \ldots, 1/L) \) measure on \( D^\mathbb{Z} \). The measure \( \gamma \) is not necessarily invariant under \( T \); however, for each cylinder set \( C \) we have \( \gamma(C) = \gamma(T^{-2m}C) \). So the new measure \( \tilde{\gamma} \) on \( X \) defined by \( \tilde{\gamma}(C) = \frac{1}{2m}(\gamma(C) + \ldots + \gamma(T^{-2m+1}C)) \) is \( T \)-invariant. Since for each cylinder set \( E \) of \( Y \) we have

\[
\gamma(\pi^{-1}E) = (\mu \times \lambda)(\pi^{-1}E \times D^\mathbb{Z}) = \nu(E),
\]

we deduce that both measures \( \gamma \) and \( \tilde{\gamma} \) are in the fibre \( \pi^{-1}\{\nu\} \).

Define an equivalence relation on \( X \) as follows. Given \( x, x' \in X \), say \( x \sim_0 x' \) if either \( x_R = x'_R \) or else \( x_{R \cap G^c} = x'_{R \cap G^c} \) and \( x_{\partial G^c} = x'_{\partial G^c} = \eta \), and \( x_G, x'_G \) in \( D \). Denote the equivalence class containing \( x \) by \( C_0(x) \). Such equivalence classes form a subpartiton of \( \mathcal{P}(X_{-m}^{-1}) \). Let \( \mathcal{A} \) be the \( \sigma \)-algebra generated by these equivalence
classes. We show $h_\gamma(T) \geq h_\mu(T)$ as follows, using Lemma 3.4 which appears below:

(3.1)

$$h_\gamma(T) = \frac{1}{2m} h_{\gamma}(T^{2m})$$

$$\geq \frac{1}{2m} H_\gamma(A|\sigma(X^{-\infty}_{m-1})) + \frac{1}{2m} H_\gamma(\sigma(X^{-\infty}_{m-1}) \cup A)$$

	\hspace{1cm} \text{Lem 3.4 (a)}

$$\geq \frac{1}{2m} H_\mu(A|\sigma(X^{-\infty}_{m-1})) + \frac{1}{2m} H_\mu(\sigma(X^{-\infty}_{m-1}) \cup A)$$

	\hspace{1cm} \text{Lem 3.4 (b)}

$$\geq \frac{1}{2m} H_\mu(A|\sigma(X^{-\infty}_{m-1})) + \frac{1}{2m} H_\mu(\sigma(X^{-\infty}_{m-1}) \cup A)$$

	\hspace{1cm} \text{Lem 3.4 (c)}

$$= \frac{1}{2m} H_\mu(\sigma(X^{-\infty}_{m-1}) | A)$$

$$= \frac{1}{2m} h_\mu(T^{2m})$$

$$= h_\mu(T).$$

**Lemma 3.4.** Reusing previous notation, we have

(a) $H_\gamma(\sigma(X^{-\infty}_{m-1}) \cup A) = H_\gamma(\sigma(X^{-\infty}_{m-1}) | A)$.

(b) $H_\gamma(\sigma(X^{-\infty}_{m-1}) | A) \geq H_\mu(\sigma(X^{-\infty}_{m-1}) | A)$. Equality occurs if and only if

$$\frac{\mu(A \cap \tilde{A})}{\mu(A)} = 1/L,$$

for each $\alpha$ in $D$ and $\tilde{A}$ in $\mathcal{P}(A)$ such that for each $x$ in $\tilde{A}$ we have $x_{\partial G^c} = \eta$ and $x_G$ in $D$.

(c) $H_\gamma(A|\sigma(X^{-\infty}_{m-1})) \geq H_\mu(A|\sigma(X^{-\infty}_{m-1}))$.

**Proof.** By definition, for a measure $\rho$ in $\{\mu, \gamma\}$ we have

$$H_\rho(\sigma(X^{-\infty}_{m-1}) | A) = -\sum_{i,k} \rho(O_i \cap A_k) \log \frac{\rho(O_i \cap A_k)}{\rho(A_k)},$$

where $O_i$ is in $\mathcal{P}(X^{-\infty}_{m-1})$ and $A_k$ is in $\mathcal{P}(A)$ (if $\rho(O_i \cap A_k)$ is zero we define $\rho(O_i \cap A_k) \log \frac{\rho(O_i \cap A_k)}{\rho(A_k)} = 0$). Let $x$ be in $A_k$. If $x_{\partial G^c} \neq \eta$ or $x_G \notin D$, then for every $O_i$ in $\mathcal{P}(X^{-\infty}_{m-1})$ we have either $O_i \cap A_k = \emptyset$ or $O_i \cap A_k = A_k$, which both imply $\rho(O_i \cap A_k) \log \frac{\rho(O_i \cap A_k)}{\rho(A_k)} = 0$. Let $\{A_1, \ldots, A_M\}$ be the set consisting of elements in the partition $\mathcal{P}(A)$ so that for any point $x$ in $\tilde{A}$ we have $x_{\partial G^c} = \eta$ and $x_G$ in $D$. Then

$$H_\rho(\sigma(X^{-\infty}_{m-1}) | A) = -\sum_{i,k} \rho(O_i \cap \tilde{A}_k) \log \frac{\rho(O_i \cap \tilde{A}_k)}{\rho(A_k)}.$$
There are exactly \( L \) pairwise disjoint sets \( O_i \) in \( \mathcal{P}(X_{m-1}^m) \), defined by blocks which agree everywhere except on \( G_i \) and form a partition of \( A_k \). Let these sets be denoted by \( O_{k,1}, \ldots, O_{k,L} \) where \( O_{k,i} = \alpha_i \cap A_k \) for each \( 1 \leq i \leq L \). It follows that

\[
H_{\mu}(\sigma(X_{m-1}^m)|A) = -\sum_{k,i} \rho(O_{k,i} \cap A_k) \log \frac{\rho(O_{k,i} \cap A_k)}{\rho(A_k)}
\]

(3.2)

\[
= -\sum_{k,i} \rho(\alpha_i \cap A_k) \log \frac{\rho(\alpha_i \cap A_k)}{\rho(A_k)}.
\]

Let \( n \geq 1 \). By definition of \( \gamma \), for each configuration \( \alpha \) in \( D \), each cylinder set \( P \) in \( \mathcal{P}(X_{m-1}^m) \), and each \( 1 \leq k \leq M \) we have

\[
\frac{\gamma(\alpha \cap P \cap A_k)}{\gamma(P \cap A_k)} = \frac{\gamma(\alpha \cap A_k)}{\gamma(A_k)} = \frac{1}{L}.
\]

It follows that

\[
H_{\gamma}\left(\sigma(X_{m-1}^m) \mid \sigma(X_{m-1}^m - (2n+1)m) \lor A\right) = -\sum_{i,j,k} \gamma(\alpha_i \cap P_j \cap A_k) \log \frac{\gamma(\alpha_i \cap P_j \cap A_k)}{\gamma(P_j \cap A_k)}
\]

\[
= \log L \sum_{i,k} \gamma(\alpha_i \cap A_k)
\]

\[
= H_{\gamma}\left(\sigma(X_{m-1}^m) \mid A\right).
\]

Then (a) follows from Lemma 3.2 and the fact that \( \left(\sigma\left(X_{m-1}^m - (2n+1)m\right)\right)_{n=1}^{\infty} \) is an increasing sequence of \( \sigma \)-algebras with \( \bigvee_{n=1}^{\infty} \sigma\left(X_{m-1}^m - (2n+1)m\right) = \sigma(X_{m-1}^m) \).

(b) is a standard property of entropy.

To show (c) let \( A_{m-1}^{-m} \) denote the \( \sigma \)-algebra \( \bigvee_{i=1}^{n} T^{-2mi}A \) (the choice of notation is intended to remind the reader that \( A_{m-1}^{-m} \) is a \( \sigma \)-subalgebra of \( \sigma(X_{m-1}^m) \)).

Any set \( B \) in \( A_{m-1}^{-m} \) splits up into \( L^n \) elements of \( \sigma(X_{m-1}^m - (2n+1)m) \), \( B_1, \ldots, B_{L^k} \) of equal measure (where \( K \) is the number of occurrences in \( B \) of blocks that are randomized under \( \gamma \)). Each of these sets has the property that \( \gamma(A \cap B_i)/\gamma(B_i) = \gamma(A \cap B)/\gamma(B) \). Accordingly we see that \( H_{\gamma}(A|A_{m-1}^{-m}) = H_{\gamma}(A|\sigma(X_{m-1}^m - (2n+1)m)) \).

Taking a limit we obtain

\[
H_{\gamma}(A|\sigma(X_{m-1}^m)) = H_{\gamma}(A|A_{m-1}^{-m}).
\]

We then notice that \( H_{\gamma}(A|A_{m-1}^{-m}) = H_{\mu}(A|A_{m-1}^{-m}) \) since the measures \( \mu \) and \( \gamma \) agree on elements of \( \bigvee_{i=0}^{\infty} T^{-2mi}A \).

Finally it is clear that \( H_{\mu}(A|A_{m-1}^{-m}) \geq H_{\mu}(A|\sigma(X_{m-1}^m)) \). Combining these three equalities and inequalities we obtain \( H_{\gamma}(A|\sigma(X_{m-1}^m)) \geq H_{\mu}(A|\sigma(X_{m-1}^m)) \) as required.

\[\text{Proof of Theorem 3.3 (continued).}\] Since \( \mu \) is a measure of relative maximal entropy it follows that all of the inequalities in equations (3.1) are forced to be equalities. In particular, we have \( H_{\mu}\left(\sigma\left(X_{m-1}^m\right) \mid A\right) = H_{\gamma}\left(\sigma\left(X_{m-1}^m\right) \mid A\right) \). Then Lemma 3.3(b)
implies that
\begin{equation}
\frac{\mu(x \cap \tilde{A})}{\mu(\tilde{A})} = 1/L,
\end{equation}
for each configuration \( x \) in \( D \) and \( \tilde{A} \) in \( \mathcal{P}(A) \) such that for each \( x \) in \( \tilde{A} \) we have \( x_{\partial G^c} = \eta \) and \( x_G \) in \( D \). Note that given a finite set \( G \subseteq \mathbb{Z} \), both configurations \( \Delta \in \mathcal{B}_Y \) on \( G \) and \( \eta \in \mathcal{B}_X \) on \( \partial G^c \) with \( \mu(\eta \cap \pi^{-1}(\Delta)) > 0 \) are chosen arbitrarily. By choosing different configurations and noting that \( \tilde{A} \in \mathcal{P}((\pi^{-1}(\sigma(Y_G)) \cup \sigma(X_{R\cap G^c})) = \mathcal{P}(\pi_{1}(\eta(Y_R)) \cup \sigma(X_{R\cap G^c})) \), equation (3.3) implies that for any configuration \( x \) of \( X \) occurring at \( G \) we have
\[
\mathbb{E}(1_{\alpha} | \pi^{-1}(\sigma(Y_R)) \cup \sigma(X_{R\cap G^c})) = \begin{cases} 
\frac{1}{L(x,G)} & \text{\( \alpha \) extends} x_{\partial G^c}, \ x_G \in \pi_{b}^{-1}(\pi_b(\alpha)) \\
0 & \text{otherwise,}
\end{cases}
\]
where \( L(x,G) \) is the number of configurations of \( X \) occurring at \( G \) which extend \( x_{\partial G^c} \) and project to \( \pi_b(x_G) \).

Now for a positive integer \( t \) let \( R^{(t)} = [-m + t, m + t - 1] \), and let \( (R^{(t)_n})_{n \in \mathbb{Z}} \) be the partition of \( \mathbb{Z} \) by translates of \( R^{(t)} \), where \( R^{(t)_0} = R^{(t)} \); i.e., \( R^{(t)_n} = [(2n - 1)(m + t), (2n + 1)(m + t) - 1] \). Let \( R^{(0)} = R \). Since \( (\sigma(Y_{R^{(t)}}))_{t \to \infty} \) and \( (\sigma(X_{R^{(t)})\cap G^c}))_{t \to \infty} \) are increasing sequences with \( \bigvee_{t=0}^\infty \sigma(Y_{R^{(t)}}) = \mathcal{B}_Y \) and \( \bigvee_{t=0}^\infty \sigma(X_{R^{(t)}\cap G^c}) = \sigma(X_{G^c}) \), it follows from Lemma 3.2 that
\[
\mathbb{E}(1_{\alpha} | \pi^{-1}(\mathcal{B}_Y) \cup \sigma(X_{G^c})) = \begin{cases} 
\frac{1}{L(x,G)} & \text{\( \alpha \) extends} x_{\partial G^c}, \ x_G \in \pi_{b}^{-1}(\pi_b(\alpha)) \\
0 & \text{otherwise.}
\end{cases}
\]

\[
\square
\]

4. Class degree of a factor code

When \( \pi \) is a finite-to-one factor code from an SFT \( X \) to a sofic shift \( Y \), there is a uniform upper bound on the number of preimages of points in \( Y \) [13]. The minimal number of \( \pi \)-preimages of points in \( Y \) is called the degree of the code and denoted by \( d_{\pi} \).

**Theorem 4.1** ([13]). Let \( \pi : X \to Y \) be a finite-to-one factor code from an SFT \( X \) to an irreducible sofic shift \( Y \). Then every doubly transitive point of \( Y \) has exactly \( d_{\pi} \) preimages.

Theorem 9.1.11 in [13] shows this result when \( X \) is irreducible. However, the proof only depends on the irreducibility of \( Y \).

**Corollary 4.2.** Let \( (X,Y,\pi) \) and \( (\tilde{X},\tilde{Y},\tilde{\pi}) \) be conjugate factor triples. Then we have \( d_{\pi} = d_{\tilde{\pi}} \).

**Definition 4.3.** A family \( x^{(1)}, \ldots, x^{(k)} \) of sequences in an SFT \( X \) is mutually separated if, for each integer \( i \), \( x_{i}^{(1)}, \ldots, x_{i}^{(k)} \) are all distinct.

**Proposition 4.4** ([13]). Let \( \pi : X \to Y \) be a finite-to-one one-block factor code from an SFT to an irreducible sofic shift \( Y \). The preimages of a transitive point are mutually separated. If \( \pi \) has a magic symbol \( w \), then \( d_{\pi} = |\pi_{b}^{-1}(w)| \).
For a proof, the reader is referred to Proposition 9.1.9 and Exercise 9.1.3 of [13]. We will find a quantity analogous to the degree when $\pi$ is an infinite-to-one factor code. This will be done by developing the following equivalence relation on $X$. Figure 1 illustrates Definition 4.5.

**Definition 4.5.** Suppose $(X, Y, \pi)$ is a factor triple and $x, x' \in X$. We say there is a transition from $x$ to $x'$ and denote it by $x \rightarrow x'$ if for each integer $n$, there exists a point $v$ in $X$ so that

1. $\pi(v) = \pi(x) = \pi(x')$, and
2. $v_n^n = x^n_n$, $v_i^{\infty} = x_i^{\infty}$ for some $i \geq n$.

Write $x \not\rightarrow x'$ if the above conditions do not hold. We write $x \sim x'$, and say $x$ and $x'$ are in the same transition class if $x \rightarrow x'$ and $x' \rightarrow x$. The relation $\sim$ is an equivalence relation. Denote the set of transition classes in $X$ over $y \in Y$ by $\mathcal{C}_\pi(y)$. Sometimes we denote $\mathcal{C}_\pi(y)$ simply by $\mathcal{C}(y)$ when there is no ambiguity in understanding $\pi$. Say $[x] \rightarrow [x']$ if $x \rightarrow x'$ (well-defined). Use the notation $[x] \not\rightarrow [x']$ otherwise.

The following fact is derived from Definition 4.5 immediately.

**Fact 4.6.** Let $\pi : X \rightarrow Y$ be a one-block factor code from a one-step SFT $X$ to a sofic shift $Y$. Let $x, x'$ project to the same point under $\pi$ and $x_{a_i} = x'_{a_i}$, where $(a_i)_{i \in \mathbb{N}}$ is a strictly increasing sequence in $\mathbb{Z}$. Then we have $x \sim x'$.

Note that Fact 4.6 gives an obvious case when two points lie in the same transition class. Simple examples allow one to find equivalent points without having a common symbol at the same time.

**Definition 4.7.** Let $(X, Y, \pi)$ be a factor triple. The minimal number of transition classes over points of $Y$ is called the class degree of $\pi$ and denoted by $c_\pi$.

The following proposition shows that conjugate factor triples have the same class degree.

**Proposition 4.8.** Let $(X, Y, \pi)$ and $(\tilde{X}, \tilde{Y}, \tilde{\pi})$ be conjugate factor triples under conjugacies $\phi : X \rightarrow \tilde{X}$ and $\psi : Y \rightarrow \tilde{Y}$. Then:

(a) For each $y$ in $Y$, $|\mathcal{C}_\pi(y)| = |\mathcal{C}_{\tilde{\pi}}(\psi(y))|$
(b) $c_\pi = c_{\tilde{\pi}}$

**Proof.** (a) is clear by noting that given $u, v$ in $X$, $\phi(u) \sim \phi(v)$ if and only if $u \sim v$, and (b) is a direct corollary of (a).
The following two theorems give an upper bound on the number of transition classes over points of $Y$. Later in Theorem 4.22 we give a finitary characterization of the number of such classes.

**Theorem 4.9.** Let $\pi : X \rightarrow Y$ be a one-block factor code from a one-step SFT $X$ to a sofic shift $Y$. Let $y$ be in $Y$ and let $w$ be a symbol in $Y$ occurring infinitely often in positive coordinates of $y$. Then $|\mathcal{C}(y)| \leq |\pi^{-1}_b(w)|$.

*Proof.* Let $(a_j)_{j \in \mathbb{N}}$ be a strictly increasing sequence of integers such that $y_{a_j} = w$ for each $j \in \mathbb{N}$. Let $|\pi^{-1}_b(w)| = d$ and suppose $\pi^{-1}(y)$ contains $d + 1$ distinct transition classes $C_1, \ldots, C_{d+1}$. Form a set $A = \{x_1, \ldots, x_{d+1}\}$, where $x_i$ is an arbitrary point of $C_i$. Since $\pi^{-1}_b(w)$ contains exactly $d$ symbols, it follows that for each $j \in \mathbb{N}$ there are at least two points in $A$ with the same $a_j$th coordinate. The Pigeonhole Principle implies that there is a subsequence $(b_k)_{k \in \mathbb{N}}$ of $(a_j)_{j \in \mathbb{N}}$ and at least two points $x$ and $x'$ in $A$ with $x_{b_k} = x'_{b_k}$ for each $k \in \mathbb{N}$. Fact 4.4 implies that $x \sim x'$. This contradicts the assumption that $x$ and $x'$ are in different transition classes. \hfill $\square$

A similar proof to the above implies the following theorem.

**Theorem 4.10.** Let $\pi : X \rightarrow Y$ be a one-block factor code from a one-step SFT $X$ to a sofic shift $Y$. Then $|\mathcal{C}(y)| < \infty$ for each $y \in Y$. Moreover, $\min_{w \in A(Y)} \{|\pi^{-1}_b(w)|\}$ is the uniform upper bound on the number of transition classes over each right transitive point of $Y$.

We show when $\pi : X \rightarrow Y$ is a factor code from an SFT $X$ to an irreducible sofic shift $Y$, there are exactly $c_\pi$ transition classes over a transitive point of $Y$. In order to show this, we introduce another quantity $c^*_\pi$ in Definition 4.11 defined concretely in terms of blocks. Since the class degree is invariant under conjugacy of factor triples, by recoding we may focus only on one-step SFTs and one-block factor codes.

**Definition 4.11.** Let $\pi : X \rightarrow Y$ be a one-block factor code from a one-step SFT $X$ to a sofic shift $Y$. Let $W$ be a $Y$-block of length $p+1$. Let $n$ be an integer in $(0, p)$, and $M$ be a subset of $\pi^{-1}_b(W_n)$. We say $U \in \pi^{-1}_b(W)$ is *routable through* $a \in M$ at time $n$ if there is a block $U' \in \pi^{-1}_b(W)$ with $U'_0 = U_0$, $U'_n = a$, and $U'_p = U_p$. A triple $(W, n, M)$ is called a *transition block* of $\pi$ if every block in $\pi^{-1}_b(W)$ is routable through a symbol of $M$ at time $n$. The cardinality of the set $M$ is called the *depth* of the transition block $(W, n, M)$.

Figure 2 illustrates Definition 4.11.

**Definition 4.12.** Let

$$c^*_\pi = \min\{|M| : (W, n, M) \text{ is a transition block of } \pi\}.$$ 

A *minimal transition block* of $\pi$ is a transition block of depth $c^*_\pi$.

**Example 4.13.** In Figure 3 we display an example of a labeled graph which defines an infinite-to-one one-block factor code $\pi$. We see that $(001, 1, (b))$ is a minimal transition block of $\pi$ of depth 1. For example, observe that block $U = aac$ is routable through $b$ at time 1 by considering $U' = abc$. 

Figure 2. \((W, n, M)\) is a transition block with \(M = \{a_1, a_2\}\). The blocks \(U, V, K \in \pi_b^{-1}(W)\) are routable through members of \(M\) at time \(n\) via blocks \(U', V', K' \in \pi_b^{-1}(W')\).

\[\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure2.png}
\caption{Graph for Example 4.13}
\end{figure}\]

We need to develop some definitions and lemmas before proving Theorem 4.22 below. Given a factor triple \((X, Y, \pi)\), a point \(y\) in \(Y\) and a block \(V\) occurring in \(y\), it may be the case that some preimages of \(V\) cannot be extended to a point in \(X\) which projects to \(y\) under \(\pi\). Here we show that in order to recognize those nonextendable blocks we do not need to deal with a bi-infinite sequence. A finite extension of a block is sufficient to determine if the block can be extended to a point above \(y\).

**Definition 4.14.** Let \((X, Y, \pi)\) be a factor triple and \(y\) be in \(Y\). Let \(V\) be a block occurring in \(y\) at a \(Z\)-interval \([m, n]\). A block \(W\) of \(Y\) is called a \(y\)-synchronizing extension of \(V\) if \(W\) occurs in \(y\) at \([m - l, n + l]\) for some integer \(l \geq 0\) and
\[
\{U \in \pi_b^{-1}(V) : \exists x \in \pi^{-1}(y), x_{[m,n]} = U\}
= \{U \in \pi_b^{-1}(V) : \exists B \in \pi_b^{-1}(W), B_{[l,l+n-m]} = U\}.
\]

**Lemma 4.15.** Let \((X, Y, \pi)\) be a factor triple and \(y\) in \(Y\). Let \(V\) be a block occurring in \(y\) at a \(Z\)-interval \([m, n]\). There is a \(y\)-synchronizing extension of \(V\).

**Proof.** Let
\[
S = \{U \in \pi_b^{-1}(V) : \exists x \in \pi^{-1}(y), x_{[m,n]} = U\}.
\]
For \(l \geq 0\) let
\[
S^l = \{U \in \pi_b^{-1}(V) : \exists B^l \in \pi_b^{-1}(y_{[m-l,n+l]}), B^l_{[l,l+n-m]} = U\}.
\]
We show for some $l$ we have $S = S^l$. Then it will imply that $y_{[m-l,n+l]}$ is a $y$-synchronizing extension of $V$.

Clearly for all $l \geq 0$ we have $S \subseteq S^l$. We show $S^l \subseteq S$ for some $l$ as follows. First note that since $\pi$ is a factor code, $S$ is not empty. It follows that the sequence $(S^l)_{l=0}^{\infty}$ is a nonincreasing sequence of nonempty finite sets. Thus there is an $l$ such that for each $r \geq l$ we have $S^l = S^r$. We show $S^l \subseteq S$.

Let $U$ be in $S^l$. Then $U$ is in $S^r$ for each $r \geq l$. It follows that for each $r \geq l$ there is $B_r$ in $\pi_b^{-1}(y_{[m-r,n+r]})$ with $B_{[r,r+n-m]} = U$. For each $r \geq l$ consider the cylinder set $C_{m-r}(B_r)$. The sequence $(C_{m-r}(B_r))_{r \geq l}$ is a nonincreasing sequence of nonempty compact subsets of $X$. It follows that $\bigcap_{r \geq l} C_{m-r}(B_r)$ is not empty. Let $x \in \bigcap_{r \geq l} C_{m-r}(B_r)$. Then $x_{[m,n]} = U$ and $\pi(x) = y$, which implies that $U$ is in $S$.

Here we introduce a relation between symbols and transition classes.

**Definition 4.16.** Let $\pi : X \to Y$ be a one-block factor code from a one-step SFT $X$ to a sofic shift $Y$. Let $y$ be in $Y$. Let $n$ be in $\mathbb{Z}$, $i$ be a symbol in $\pi_b^{-1}(y_n)$, and $C$ be a transition class in $\mathcal{C}(y)$. Say $i$ belongs to $C$ at time $n$, and denote it by $i \in S_n(C)$, if

$$\{D \in \mathcal{C}(y) : \text{there is } x \text{ in } D \text{ with } x_n = i\} = \{D \in \mathcal{C}(y) : C \to D\}.$$ 

In other words, the set of transition classes in $\mathcal{C}(y)$ containing a point with $n$th coordinate being $i$ is the same as the set of transition classes to which there is a transition from the class $C$ (including $C$ itself).

Say $i$ is transient at time $n$ if there is no transition class $C$ in $\mathcal{C}(y)$ for which $i \in S_n(C)$.

Example 4.17 illustrates the definition above.

**Example 4.17.** Consider the directed labeled graph in Figure 4, which presents a one-block factor code $\pi : X \to Y$, where $X \subseteq \{a, b, c, d, e, f, g\}^\mathbb{Z}$ and $Y \subseteq \{0, 1\}^\mathbb{Z}$. Let $y$ be the point $\cdots 0101010 \cdots$ in $Y$. By Definition 4.16 there are three distinct transition classes in $\mathcal{C}(y)$ as follows:

(1) Transition Class $C_1 = [x_1] = \{x_1\}$, where $x_1 = \cdots bababab \cdots$.
(2) Transition Class $C_2 = [x_2]$, where $x_2$ is a point in $\pi^{-1}(y)$ which does not contain symbols $e$ or $g$ but contains only $d$'s and $f$'s from some time onwards, for example: $x_2 = \cdots bab\ast f d \cdots$.
(3) Transition Class $C_3 = [x_3]$, where $x_3$ is a point in $\pi^{-1}(y)$ which does not contain symbols $d$ or $f$ but contains only $e$'s and $g$'s from some time onwards, for example: $x_3 = \cdots e \ast g e g e \cdots$. 

![Figure 4](image)
Clearly $C_1 \to C_2$ and $C_1 \to C_3$, but not vice versa. $C_2 \to C_3$, $C_3 \to C_2$, and symbol $c$ is transient at any time.

Note that the class degree of the factor code $\pi$ is 2 (consider a point of $Y$ in which the block 11 occurs infinitely often in positive coordinates).

**Lemma 4.18.** Let $\pi : X \to Y$ be a one-block factor code from a one-step SFT $X$ to a sofic shift $Y$, and $x$ be a point in $X$. There is $m < \infty$ such that for each $n \geq m$ the symbol $x_n$ is not transient. In fact, $m$ can be found in such a way that for each $n \geq m$, $x_n$ belongs to the class $[x]$ at time $n$; i.e., $x_n \in S_n[x]$.

**Proof.** Let $\pi(x) = y$. Suppose that the transition class $C$ in $\mathcal{C}(y)$ satisfies $[x] \to C$. Then there exists $i < \infty$ such that for any $z \leq \pi^{-1}(y)$ with $z_i = x_i$ we have $z \notin C$. Denote the smallest such $i$ by $i^C$. Note that $i^C$ may be $-\infty$ (but not $+\infty$). Let $m = \max\{i^C : C \in \mathcal{C}(y), [x] \to C\}$. Let $n \geq m$. We show

\begin{equation}
\{C \in \mathcal{C}(y) : [x] \to C\} = \{C \in \mathcal{C}(y) : \text{there is } z \in C \text{ with } z_n = x_n\},
\end{equation}

which will imply the result $x_n \in S_n[x]$.

Let $C$ be a transition class in $\mathcal{C}(y)$ with $[x] \to C$. It follows that $x \to z$ for some $z$ in $C$. Therefore, there is a point $u$ in $\pi^{-1}(y)$ as follows:

$$u_t = \begin{cases} x_t & \text{if } -\infty < t \leq n \\ z_l & \text{if } l \leq t < \infty, \end{cases}$$

for some $l \geq n$. Since $u_t = z_l$ from time $l$ onwards, it follows that $u$ is in the class $C$. Moreover, we have $u_n = x_n$ by the construction. It follows that $C$ belongs to the set on the right-hand side of equation (4.1).

Now let $C$ be an equivalence class in $\mathcal{C}(y)$ with $[x] \to C$. Suppose, for a contradiction, that $C$ contains a point $z$ with $z_n = x_n$. Form the point $u$ as follows:

$$u_t = \begin{cases} x_t & \text{if } -\infty < t \leq n \\ z_l & \text{if } n < t < \infty. \end{cases}$$

Note that since $X$ is a one-step SFT and $z_n = x_n$, the point $u$ is in $X$. Moreover, $u$ maps to $y$ by the construction. Noticing that $u_t = z_l$ from the time $n$ onwards shows that $u$ lies in the class $C$. Recall that $n \geq i^C$. By the argument in the first paragraph of the proof, any point in $\pi^{-1}(y)$ with $i^C$th coordinate equal to $x_{i^C}$ does not belong to $C$. This contradicts that $u$ is a point in $C$. \qed

**Lemma 4.19.** Let $\pi : X \to Y$ be a one-block factor code from a one-step SFT $X$ to a sofic shift $Y$. Let $y \in Y$. There is $m < \infty$ such that for each $n \geq m$ and $C$ in $\mathcal{C}(y)$ there is a symbol $i$ in $\pi_b^{-1}(y_n)$ which belongs to the class $C$ at time $n$; i.e., $i \in S_n(C)$.

**Proof.** Let $\mathcal{C}(y) = \{C_1, \ldots, C_d\}$ for some finite $d$. Let $A = \{x^{(1)}, \ldots, x^{(d)}\}$ be a set containing an arbitrary point $x^{(i)}$ of $C_i$ for each $C_i$. By Lemma 4.18 there is an $n_{C_i} < \infty$ such that for each $n \geq n_{C_i}$ the symbol $x_b^{(i)}$ belongs to $C_i$ at time $n$. Let $m = \sup\{n_{C_i} : 1 \leq i \leq d\}$. Note that $m < \infty$. Moreover, for each $n \geq m$ and $C_i$ the symbol $x_b^{(i)}$ belongs to the class $C_i$ at time $n$. \qed

**Proposition 4.20.** Let $\pi : X \to Y$ be a one-block factor code from a one-step SFT $X$ to a sofic shift $Y$. Let $\nu$ be an invariant measure on $Y$. Then

$$\nu \left( \{y \in Y : \forall C \in \mathcal{C}(y), \forall n \in \mathbb{Z} \text{ there is } i \in \pi_b^{-1}(y_n) \text{ with } i \in S_n(C) \} \right) = 1.$$
Proof. Let $y$ be in $Y$ and let $m(y)$ be the infimum of the set of $m$'s with the properties given in the statement of Lemma 4.19. We show
\[ \nu \left( \{ y \in Y : m(y) = -\infty \} \right) = 1, \]
which will imply the result immediately. Note that $C \in \mathcal{C}(y)$ if and only if $T(C) \in \mathcal{C}(T(y))$. This shows $m(T(y)) = m(y) - 1$. For $k < \infty$ let $A_k = \{ y \in Y : m(y) = k \}$. We have $T(A_k) = A_{k-1}$. Since $\nu$ is $T$-invariant it follows that $\nu(A_k) = 0$. Therefore $m(y) = -\infty$ for $\nu$-a.e. $y$ in $Y$. \hfill \square

Definition 4.21. Let $X$ be a shift space. A point $x$ in $X$ is recurrent if for each block $B$ which occurs in $x$, $B$ occurs infinitely often in the positive coordinates of $x$.

Note that given any ergodic measure $\mu$ on $X$, $\mu$-almost every point of $X$ is recurrent.

Theorem 4.22. Let $\pi$ be a one-block factor code from a one-step SFT $X$ to a sofic shift $Y$. Let $y$ be a point of $Y$ and $d$ be the minimal depth of a transition block occurring in $y$. Then $d \leq |\mathcal{C}(y)|$. Moreover, if $y$ is recurrent, then $d = |\mathcal{C}(y)|$.

Proof. We show $d \leq |\mathcal{C}(y)|$ by finding a transition block of depth $|\mathcal{C}(y)|$ occurring in $y$. We construct such a transition block in the following four stages. Let $h = |\mathcal{C}(y)|$ and suppose $\mathcal{C}(y) = \{ C_1, \ldots, C_h \}$. Choose an integer $n_1$ satisfying the properties given in the statement of Lemma 4.19.

Stage 1. We claim there is $n_2 \in [n_1, \infty)$ such that for each $x \in \pi^{-1}(y)$, there is $n_1 \leq l \leq n_2$ so that $x_l$ is not transient.

Proof. Suppose there is no such $n_2$. It follows that for each $j \geq n_1$ there is a point $x^{(j)}$ in $\pi^{-1}(y)$ such that $x^{(j)}_l$ is transient for all $n_1 \leq l \leq j$. Consider the sequence $(x^{(j)})_{j \in \mathbb{Z}}$, and let $x$ be the limit of a convergent subsequence of it. Clearly $x \in \pi^{-1}(y)$. However, $x_l$ is transient for each $l \geq n_1$, contradicting Lemma 4.18. \hfill \square

Stage 2. We claim there is $n_3 \in [n_2, \infty)$ and a set of symbols
\[ M' = \{ a_1, \ldots, a_h \} \]
with $a_e \in S_{n_3}(C_e)$ such that for each $i$ in $S_{n_2}(C_e)$ there is a block
\[ U \in \pi_{b_1}^{-1}(y_{n_2}y_{n_2+1} \ldots y_{n_3-1}y_{n_3}) \]
which begins with $i$ and ends with $a_e$. See Figure 5.

Proof. Let $C_e$ be a transition class in $\mathcal{C}(y)$ and $x^e$ be a point in $C_e$. For each symbol $i$ in $S_{n_2}(C_e)$ (nonempty by the choice of $n_1$), there is a bi-infinite sequence $z^i$ in the class $C_e$ with $z_{n_2}^i = i$ such that $z^i$ matches $x^e$ from some time $k^i \in [n_2, \infty)$ onwards. Let
\[ k^e = \max\{ k^i : i \in S_{n_2}(C_e) \} \]
and
\[ n_3 = \max\{ k^e : C_e \in \mathcal{C}(y) \}. \]
Note that $n_3 < \infty$. Rename $x^e_{n_3}$ for each $C_e \in \mathcal{C}(y)$ as $a_e$, and let $M' = \{ a_1, \ldots, a_h \}$. \hfill \square

Stage 3. We claim there is $n_4 \in [n_3, \infty)$ such that for each point $x$ in $\pi^{-1}(y)$ there is a point $x' \in \pi^{-1}(y)$ so that $x'_r = x_r$ for every $r \in (-\infty, n_4] \cup [n_4, \infty)$, and $x'_{n_4} \in M'$. See Figure 6.
Figure 5. An example illustrating stage 2. \( \mathcal{C}(y) = \{ C_1, C_2, C_3 \} \), \( i_1, i_2, i_3 \in S_{n_2}(C_1), i_4 \in S_{n_2}(C_2), i_5, i_6 \in S_{n_2}(C_3) \), and \( M' = \{ a_1, a_2, a_3 \} \).

Figure 6. Graph for stage 3. \( M' = \{ a_1, a_2, a_3 \} \). \( x, x' \in \pi^{-1}(y) \), \( x'_r = x_r \) for each \( r \in (-\infty, n_1] \cup [n_4, \infty) \), and \( x'_{n_3} \in M' \). Same for \( z, z' \) and \( u, u' \).

Proof. Let \( x \) be a point in \( \pi^{-1}(y) \). By stage 1, there is \( n_1 \leq t \leq n_2 \) such that \( x_t \) is not transient; i.e., \( x_t \) belongs to some transition class \( C_e \) at time \( t \). It follows that there is a point \( u \) in the class \( C_e \) with \( u_t = x_t \). The new block

\[
(4.2) \quad \ldots x_{n_1} \ldots x_{t} u_{t+1} \ldots u_{n_2}
\]
is an allowable block of \( X \). Moreover, this block maps to the block
\[
\ldots y_t \ldots y_{n_2}.
\]
Note that the symbol \( u_{n_2} \) belongs to the class \( C_e \) at time \( n_2 \). So by stage 2, there is a block starting at \( u_{n_2} \) ending at \( a_e \) which maps to
\[
y_{n_2} \cdots y_{n_3}.
\]
Denote this block by
\[
(4.3) \quad u_{n_2} v_{n_2+1} \ldots v_{n_3-1} a_e.
\]
Connect blocks in equations (1.2) and (4.3) at \( u_{n_2} \) to get the following allowable block of \( X \):
\[
\ldots x_1 \ldots x_t u_{t+1} \ldots u_{n_2} v_{n_2+1} \ldots v_{n_3-1} a_e.
\]
Observe that having \( x_t \) in \( S_t(C_e) \) implies that \( x \) must belong to a transition class \( C_f \) with \( C_e \rightarrow C_f \). On the other hand having \( a_e \) in \( S_{n_3}(C_e) \) implies that there is a point \( b \) in \( C_e \) with
\[
b_{[-\infty,n_3]} = \ldots x_{n_1} \ldots x_t u_{t+1} \ldots u_{n_2} v_{n_2+1} \ldots v_{n_3-1} a_e,
\]
which matches \( x \) from some time \( j \) onwards for some \( n_3 \leq j < \infty \). Let
\[
n_e = \min_{n_3 \leq j < \infty} \{ j : \exists x' \in \pi^{-1}(y) \text{ with } x'_r = x_r \forall r \in (-\infty, n_1] \cup [j, \infty), \ x'_r = x_e \text{ for each } r \in (\infty, n_1] \cup [n_e, \infty) \},
\]
We claim that there is an \( n_e < \infty \) such that for each \( x \in \pi^{-1}(y) \) with \( x_t \in S_t(C_e) \) for some \( n_1 \leq t \leq n_2 \), there exists a point \( x' \in \pi^{-1}(y) \) with \( x'_r = x_r \) for each \( r \in (\infty, n_1] \cup [n_e, \infty) \), and \( x'_r = a_e \). Then letting
\[
n_4 = \max\{n_e : C_e \in \mathcal{C}(y)\} < \infty
\]
will complete the proof of stage 3.

Suppose, for a contradiction, that there does not exist such an \( n_e \). Then there is a sequence \( x^{(i)} \) with \( x^{(i)}_t \in S_t(C_e) \) for some \( n_1 \leq t_i \leq n_2 \), such that \( \lim_{i \to \infty} x^{(i)}_t = \infty \). Let \( x^* \) be the limit of a convergent subsequence of \( x^{(i)} \). Clearly \( x^* \) is in \( \pi^{-1}(y) \) and \( n_{x^*} = \infty \), which contradicts the fact that for each \( x \) in \( \pi^{-1}(y), n_x < \infty \). \( \square \)

**Stage 4.** Let \( W \) be a \( y \)-synchronizing extension of the block
\[
V = y_{n_1} \ldots y_{n_4}
\]
onoccurring in \([n_0, n_5]\) for some \( n_0 \leq n_1 \) and \( n_5 \geq n_4 \). We claim \((W, n_3 - n_0, M')\) is a transition block of depth \( |\mathcal{C}(y)| \).

**Proof.** Let \( U \in \pi^{-1}_b(W) \). By the definition of a \( y \)-synchronizing extension, there is an \( x \) in \( \pi^{-1}(y) \) with \( x_{[n_1, n_4]} = U \). By stages 1, 2, and 3 there is a point \( x' \) in \( \pi^{-1}(y) \) with \( x'_r = x_r \) for each \( r \in (-\infty, n_1] \cup [n_4, \infty) \), and \( x'_{n_3} \in M' \). Having \( x'_{n_1} = x_{n_1}, x'_{n_3} = x_3 \), \( x'_{n_4} = x_{n_4} \) and \( \pi_b(x'_{[n_1, n_4]}) = V \) simply means that \( U \) is routable through a symbol of \( M' \) at time \( n_3 - n_0 \). Since \( U \) in \( \pi^{-1}_b(W) \) is arbitrary it follows that \((W, n_3 - n_0, M')\) is a transition block. Because \( M' \) has \( |\mathcal{C}(y)| \) elements it follows that \( W \) is a transition block of depth \( |\mathcal{C}(y)| \). \( \square \)

Now suppose \( y \) is recurrent. Take a transition block of depth \( d \) which occurs in \( y \). This block occurs infinitely often in the positive coordinates of \( y \). Apply the same method which was used in the proof of Theorem 4.1 to see \( |\mathcal{C}(y)| \leq d \). Since by the first part of the theorem \( d \leq |\mathcal{C}(y)| \) we obtain \( d = |\mathcal{C}(y)| \) when \( y \) is recurrent. \( \square \)
Corollary 4.23. Let \( \pi : X \to Y \) be a factor code from an SFT \( X \) to an irreducible sofic shift \( Y \). Then:

(a) \( c_\pi = c_\pi^* \).
(b) There are exactly \( c_\pi \) transition classes over every right transitive point of \( Y \).
(c) Given an ergodic measure \( \nu \) on \( Y \), \( \nu \)-almost every point of \( y \) has the same number of transition classes over it. We call this number \( |\mathcal{C}(\nu)| \).

Proof. Let \((W,n,M)\) be a transition block of \( \pi \) of depth \( c_\pi^* \) and let \( y \) be a right transitive point of \( Y \). Notice that \( W \) occurs infinitely often in the positive coordinates of \( y \). Then (a) and (b) are clear by Theorem 4.22.

Let \( \nu \) be an ergodic measure on \( Y \). Then \( \nu \)-almost every \( y \) is recurrent. For such a recurrent point \( y \), let \( d_y \) be the depth of a transition block of minimal depth occurring in \( y \). By Theorem 4.22, \( |\mathcal{C}(y)| = d_y \). Since the (measurable) function \( y \to d_y \) is shift-invariant and \( \nu \) is ergodic, it must be almost everywhere constant. \( \square \)

We now show that for a finite-to-one factor code, the degree and the class degree of the code are the same.

Theorem 4.24. Let \( \pi : X \to Y \) be a finite-to-one factor code from an SFT to an irreducible sofic shift \( Y \). Then \( c_\pi = d_\pi \).

Proof. Corollary 4.2 and Proposition 4.8(b) imply that the degree of a code and the class degree of a code are both invariant under conjugacy of factor triples. So by recoding, without loss of generality, we may assume \( X \) is a one-step SFT and \( \pi \) is a one-block factor code. Let \( y \) be a transitive point of \( Y \). By Proposition 4.4 the preimages of \( y \) are mutually separated. It follows that each transition class above \( y \) has only one point. Therefore, \( c_\pi = |\pi^{-1}(y)| = d_\pi \). \( \square \)

5. Bounding the number of ergodic measures of relative maximal entropy

In Section 3 we mentioned that although every one-dimensional irreducible SFT has a unique ergodic measure (Parry measure) of maximal entropy, there can be more than one ergodic measure of relative maximal entropy over an ergodic measure; i.e., given a factor triple \((X,Y,\pi)\) and an ergodic measure \( \nu \) on \( Y \), there can exist more than one ergodic measure on \( X \) that projects to \( \nu \) under \( \pi \) and has maximal entropy among measures in the fibre \( \pi^{-1}\{\nu\} \); see [17, Example 3.3]. In this section we show that the number of such measures over a fully supported ergodic measure can be no more than the class degree.

Theorem 5.1. Let \((X,Y,\pi)\) be a factor triple and \( \nu \) be a fully supported ergodic measure on \( Y \). The number of ergodic measures of relative maximal entropy over \( \nu \) is at most \( c_\pi \).

Since entropy is a conjugacy invariant the following observation follows immediately.

Observation 5.2. Let \((X,Y,\pi)\) and \((\tilde{X},\tilde{Y},\tilde{\pi})\) be conjugate factor triples. Let \( \nu \) be an ergodic measure on \( Y \) and \( \tilde{\nu} \) be its corresponding ergodic measure on \( \tilde{Y} \). The number of ergodic measures of relative maximal entropy over \( \nu \) is the same as the number of ergodic measures of relative maximal entropy over \( \tilde{\nu} \).
In 2003, Petersen, Quas, and Shin \cite{Petersen2003} found an upper bound on the number of measures of relative maximal entropy. Let \( N_\nu(\pi) \) denote the minimum number of symbols in \( \pi^{-1}_b(w) \) as \( w \) runs over the symbols in \( A(Y) \) for which \( \nu|w| > 0 \).

**Theorem 5.3 \((\cite{Petersen2003} \text{ Corollary 1})\).** Let \( \pi : X \to Y \) be a one-block code from a one-step SFT \( X \) to a sofic shift \( Y \). Let \( \nu \) be an ergodic measure on \( Y \). The number of ergodic measures of maximal entropy over \( \nu \) is at most \( N_\nu(\pi) \).

This bound suffers from not being invariant under conjugacy. For example, the full two-shift and any higher block presentation of it give different bounds on the number of ergodic measures of maximal entropy which map to the trivial measure on the full one-shift.

One possibility to avoid having a noninvariant upper bound is to take the minimum of \( \nu \)-generic points over all conjugate factor triples with one-block factor codes. This even improves the original bound. However, there is no known algorithm for computing this minimum.

Here we show that given an ergodic measure \( \nu \) on \( Y \), the number of transition classes over a \( \nu \)-generic point of \( Y \) is an upper bound on the number of measures of relative maximal entropy over \( \nu \). Proposition 4.8(b) verifies that this bound is invariant under conjugacy of factor triples, and Proposition 5.4 shows that it beats the bound mentioned above obtained by minimizing the bound in \( \cite{Petersen2003} \) over conjugate factor triples.

**Proposition 5.4.** Let \( \pi : X \to Y \) be a one-block factor code from an SFT \( X \) to a sofic shift \( Y \). Let \( \nu \) be an ergodic measure on \( Y \). Then we have \( |\mathcal{E}_\pi^\nu(y)| \leq \min\{N_\nu(\tilde{\pi}): (X,Y,\pi) \cong (\tilde{X},\tilde{Y},\tilde{\pi}), \tilde{\pi} \text{ is a one-block}\} \), where \( y \) is a \( \nu \)-generic point of \( Y \). Equality holds if \( \pi \) is finite-to-one and \( \nu \) is fully supported.

**Proof.** Let the minimum of \( N_\nu(\tilde{\pi}) \) over all factor triples \((\tilde{X},\tilde{Y},\tilde{\pi})\) which are conjugate to \((X,Y,\pi)\) and \( \tilde{\pi} \) is a one-block be attained at \((\tilde{X},\tilde{Y},\tilde{\pi})\), where \( X \) is conjugate to \( \tilde{X} \) under a conjugacy \( \phi : X \to \tilde{X} \), \( \tilde{Y} \) is conjugate to \( Y \) under a conjugacy \( \psi : Y \to \tilde{Y} \), and \( \tilde{\pi} = \psi \circ \phi \circ \pi \). Denote the measure \( \nu \circ \psi^{-1} \) by \( \bar{\nu} \). Suppose that the minimum of \( |\pi^{-1}_b(w)| \) over all \( w \) in \( A(\tilde{Y}) \) with \( \bar{\nu}|w| > 0 \) is attained at a symbol \( \bar{w} \). Let \( \bar{y} \) be a \( \bar{\nu} \)-generic point of \( \tilde{Y} \). There is a strictly increasing sequence of integers \( (a_k)_{k \in \mathbb{N}} \) with \( \bar{y}_{a_k} = \bar{w} \). Theorem 4.12 implies that \( |\mathcal{E}_{\tilde{\pi}}^\bar{\nu}(\bar{y})| \leq |\pi^{-1}_b(\bar{w})| \). Let \( y = \psi^{-1}(\bar{y}) \). Then by Proposition 4.8(a) we have \( |\mathcal{E}_{\pi}^\nu(y)| = |\mathcal{E}_{\tilde{\pi}}^\bar{\nu}(\bar{y})| \leq |\pi^{-1}_b(\bar{w})| \), which completes the proof of the first part of the result.

Now suppose \( \pi \) is finite-to-one and \( \nu \) is fully supported. Since \( \nu \) is fully supported, we have

\[ N_\nu(\tilde{\pi}) = \min\{|\pi^{-1}_b(w)|: w \in A(\tilde{Y})\} \]

Therefore the minimum of \( N_\nu(\tilde{\pi}) \) over all \((\tilde{X},\tilde{Y},\tilde{\pi})\) which are conjugate to \((X,Y,\pi)\) and \( \tilde{\pi} \) is a one-block is attained at a factor triple \((\tilde{X},\tilde{Y},\tilde{\pi})\) with a magic symbol \( \bar{w} \) in \( A(\tilde{Y}) \). By Proposition 4.4 we have \( |\pi^{-1}_b(\bar{w})| = d_{\bar{w}} \), which implies

\[ \min\{N_\nu(\tilde{\pi}): (X,Y,\pi) \cong (\tilde{X},\tilde{Y},\tilde{\pi}), \tilde{\pi} \text{ is a one-block}\} = d_{\bar{w}} \]

Since the degree of a code is invariant under conjugacy of factor triples we deduce that

\[ \min\{N_\nu(\tilde{\pi}): (X,Y,\pi) \cong (\tilde{X},\tilde{Y},\tilde{\pi}), \tilde{\pi} \text{ is a one-block}\} = d_{\bar{w}} \]

Recalling that when \( \pi \) is finite-to-one the degree and the class degree of \( \pi \) are equal, the proof is complete. \( \square \)
Note that equality in Proposition 5.4 does not always hold. For example, consider the trivial factor code \( \pi : \{0,1\}^\mathbb{Z} \to \{0\}^\mathbb{Z} \). Then \( c_\pi = 1 \); however, if \( (\hat{X}, \{0\}^\mathbb{Z}, \hat{\pi}) \) is a factor triple conjugate to \( (\{0,1\}^\mathbb{Z}, \{0\}^\mathbb{Z}, \pi) \), then \( \mathcal{A}(\hat{X}) \) must be strictly greater than 1 and therefore \( N_\nu(\hat{\pi}) > 1 \) for any ergodic measure \( \nu \) on \( Y \).

**Definition 5.5.** Let \( \pi : X \to Y \) be a one-block factor code from an SFT \( X \) to a sofic shift \( Y \) and let \( \nu \) be an ergodic measure on \( Y \). Let \( \mu_1, \ldots, \mu_n \) be invariant measures in the fibre \( \pi^{-1}\{\nu\} \). The relatively independent joining \( \hat{\mu} = \mu_1 \otimes \cdots \otimes \mu_n \) of \( \mu_1, \ldots, \mu_n \) over \( \nu \) is defined as follows: if \( A_1, \ldots, A_n \) are measurable subsets of \( X \), then

\[
\hat{\mu}(A_1 \times \cdots \times A_n) = \int_Y \prod_{i=1}^n E_{\mu_i}(1_{A_i}|\pi^{-1}B_Y) \circ \pi^{-1} d\nu.
\]

Writing \( p_i \) for the projection \( X^n \to X \) onto the \( i \)-th coordinate, it is shown that for \( \hat{\mu} \)-almost every \( \hat{x} \in X^n, \pi(p_i(\hat{x})) \) is the same for each \( i \).

We will use Theorem 5.6 below, which is the main theorem from [17], to prove a stronger theorem (Theorem 5.7).

**Theorem 5.6.** [17, Theorem 1]. Let \( \pi : X \to Y \) be a one-block factor code from a one-step SFT \( X \) to a sofic shift \( Y \). Let \( \nu \) be an ergodic measure on \( Y \), and two distinct ergodic measures \( \mu_1 \) and \( \mu_2 \) be measures of relative maximal entropy over \( \nu \). Then \( (\mu_1 \otimes \mu_2)\{(u, v) \in X \times X : u_0 = v_0\} = 0 \).

**Theorem 5.7.** Let \((X, Y, \pi)\) be a factor triple. Let \( \nu \) be an ergodic measure on \( Y \), and two distinct measures \( \mu_1 \) and \( \mu_2 \) be ergodic measures of relative maximal entropy over \( \nu \). Then \( (\mu_1 \otimes \mu_2)\{(u, v) \in X \times X : u \sim v\} = 0 \).

**Proof.** First we show that, without loss of generality, we may assume \( X \) is a one-step SFT and \( \pi \) is a one-block factor code with a magic symbol. Suppose \( (\hat{X}, \hat{Y}, \hat{\pi}) \) is a factor triple conjugate to \((X, Y, \pi)\) and \( \phi : X \to \hat{X} \) is a conjugacy from \( X \) to \( \hat{X} \). By the proof of Proposition 4.8(a) we have

\[
(\phi \times \phi)\{(u, v) \in X \times X : u \sim v\} = \{(\phi(u), \phi(v)) \in \hat{X} \times \hat{X} : \phi(u) \sim \phi(v)\}.
\]

Moreover, the corresponding measure to \( \mu_1 \otimes \mu_2 \) under the conjugacy

\[
\phi \times \phi : X \times X \to \hat{X} \times \hat{X}
\]

is \( \hat{\mu}_1 \otimes \hat{\mu}_2 \), where \( \hat{\mu}_1 \) and \( \hat{\mu}_2 \) are corresponding measures to \( \mu_1 \) and \( \mu_2 \) under \( \phi \). It follows that

\[
(\mu_1 \otimes \mu_2)\{(u, v) \in X \times X : u \sim v\} = (\hat{\mu}_1 \otimes \hat{\mu}_2)\{(\phi(u), \phi(v)) \in \hat{X} \times \hat{X} : \phi(u) \sim \phi(v)\}.
\]

Therefore by recoding, without loss of generality, we may assume \( X \) is a one-step SFT and \( \pi \) is a one-block factor code.

Let \( u, v \) be in \( X \), \( n \) be in \( \mathbb{Z} \), and \( a \) be a symbol in \( \mathcal{A}(X) \). Write \( u \overset{a}{\approx}_n v \) if \( u \sim v \) and there are integers \( m \leq n \) and \( p \geq n \) such that

\[
u_i = \begin{cases} u_i & -\infty < i < m, \quad p < i < \infty \\ a & i = n. \end{cases}
\]

It is worth mentioning that the relation \( \overset{a}{\approx}_n \) is not in general an equivalence relation on \( X \). We have

\[
\{(u, v) : u \sim v\} = \bigcup_{a \in \mathcal{A}(X)} \bigcup_{n \in \mathbb{Z}} \{(u, v) : u \overset{a}{\approx}_n v\}.
\]
Now suppose, for a contradiction, that \((\mu_1 \otimes \mu_2)\{(u, v) : u \sim v\} > 0\). There must be an integer \(n\) and a symbol \(a\) in \(\mathcal{A}(X)\) such that

\[
(\mu_1 \otimes \mu_2)\{(u, v) : u \equiv^n a, v \sim v\} > 0.
\]

By applying \((T \times T)^{-n}\), without loss of generality, we may assume \(n = 0\). Since we have

\[
\{(u, v) : u \equiv^n a, v \sim v\} = \bigcup_{k \in \mathbb{Z}} \{(u, v) : u \equiv^n a, u_k, v_k\}
\]

it follows that there is an integer \(k\) so that

\[
(\mu_1 \otimes \mu_2)\{(u, v) : u \equiv^n a, v_k, v_k\} > 0.
\]

Considering only blocks of length \(2k + 1\), there must be blocks \(A, B\) in \(X\) such that

\[
(5.1) \quad (\mu_1 \otimes \mu_2)\{(u, v) : u \equiv^n a, v_{[-k, k]} = A, v_{[-k, k]} = B\} > 0.
\]

Both \(A\) and \(B\) are routable through \(a\) at time \(n\); i.e., there are blocks \(A', B'\) in \(\pi^{-1}_b(W)\) with \(A'_{-k} = A_{-k}, B'_{-k} = B_{-k}, A'_0 = B'_0 = a, A'_0 = A_k\), and \(B'_0 = B_k\). Let \(G \subseteq \mathbb{Z}\) be the set \(\{-k, \ldots, k\}\). Basic properties of conditional expectation imply that

\[
E_{\mu_1} \left(1_{[A]}|\pi^{-1}(\mathcal{B}_Y)\right) = \mathbb{E} \left(\mathbb{E}_{\mu_1} \left(1_{[-n][A]}|\pi^{-1}(\mathcal{B}_Y) \vee \sigma(X_{G'})\right)|\pi^{-1}(\mathcal{B}_Y)\right)
\]

\[
= \mathbb{E} \left(\mathbb{E}_{\mu_1} \left(1_{[-n][A']}|\pi^{-1}(\mathcal{B}_Y) \vee \sigma(X_{G'})\right)|\pi^{-1}(\mathcal{B}_Y)\right)
\]

\[
= \mathbb{E}_{\mu_1} \left(1_{[A']}|\pi^{-1}(\mathcal{B}_Y)\right),
\]

where the second equality follows from Theorem 5.3. Similarly we have

\[
(5.3) \quad \mathbb{E}_{\mu_2} \left(1_{[B]}|\pi^{-1}(\mathcal{B}_Y)\right) = \mathbb{E}_{\mu_2} \left(1_{[B']}|\pi^{-1}(\mathcal{B}_Y)\right).
\]

Let \(D = \{(u, v) : u \equiv^n a, v_{[-k, k]} = A, v_{[k, k]} = B\}\). Since \(D \subseteq [-k][A] \times [k][B]\) we have

\[
\mu(D) \leq (\mu_1 \otimes \mu_2)([A] \times [B])
\]

\[
= \int_Y \mathbb{E}_{\mu_1} \left(1_{[A]}|\pi^{-1}(\mathcal{B}_Y)\right)\mathbb{E}_{\mu_2} \left(1_{[B]}|\pi^{-1}(\mathcal{B}_Y)\right) \circ \pi^{-1} d\nu
\]

\[
= \int_Y \mathbb{E}_{\mu_1} \left(1_{[A']}|\pi^{-1}(\mathcal{B}_Y)\right)\mathbb{E}_{\mu_2} \left(1_{[B']}|\pi^{-1}(\mathcal{B}_Y)\right) \circ \pi^{-1} d\nu \quad \text{by } (5.2) \text{ and } (5.3)
\]

\[
= (\mu_1 \otimes \mu_2)([A'] \times [B'])
\]

\[
= 0,
\]

where the last equality follows from Theorem 5.0 since \(A'_0 = B'_0 = a\). This contradicts equation (5.1). \(\square\)

The following theorem is a more general case of Theorem 5.1.

**Theorem 5.8.** Let \((X, Y, \pi)\) be a factor triple and \(\nu\) be an ergodic measure on \(Y\). The number of ergodic measures of relative maximal entropy over \(\nu\) is at most \(|\mathcal{G}(\nu)|\).

**Proof.** Let \((\tilde{X}, \tilde{Y}, \tilde{\pi})\) be a factor triple conjugate to \((X, Y, \pi)\) under conjugacies \(\phi : X \to \tilde{X}\) and \(\psi : Y \to \tilde{Y}\). By Observation 5.2 the number of ergodic measures
of relative maximal entropy over \(\nu\) and \(\nu \circ \psi^{-1}\) is the same. Moreover, by Proposition 4.8(a), for each \(y\) in \(Y\) we have \(|\mathcal{C}_\pi(y)| = |\mathcal{C}_\tilde{\pi}(\psi(y))|\). Therefore by recoding, without loss of generality, we may assume \(X\) is a one-step SFT and \(\pi\) is a one-block factor code.

Suppose, for a contradiction, that there are \(n > |\mathcal{C}(\nu)|\) ergodic measures \(\mu_1, \ldots, \mu_n\) on \(X\) of relative maximal entropy over \(\nu\). Form the relatively independent joining \(\hat{\mu}\) on \(X^n\) of the measures \(\mu_1, \ldots, \mu_n\). Note that for \(\hat{\mu}\)-a.e. \(\hat{x} \in X^n\), \(\pi(p_i(\hat{x}))\) has \(|\mathcal{C}(\nu)|\) transition classes over it. The assumption \(n > |\mathcal{C}(\nu)|\) implies that for \(\hat{\mu}\)-a.e. \(\hat{x} = (x_1, \ldots, x_n) \in X^n\) there are distinct \(i, j\) such that \(p_i(\hat{x}) \sim p_j(\hat{x})\); i.e.,
\[
\hat{\mu}\left( \bigcup_{1 \leq i < j \leq n} \{\hat{x} = (x_1, \ldots, x_n): x_i \sim x_j\} \right) = 1.
\]

At least one of the sets \(S_{i,j} = \{(x_1, \ldots, x_n): x_i \sim x_j\}\) must have positive \(\hat{\mu}\)-measure. It follows that
\[
0 < \hat{\mu}(S_{i,j}) = \hat{\mu}(\{(x_1, \ldots, x_n), x_i \sim x_j\}) = (\mu_i \otimes \mu_j)(\{(u, v): u \sim v\}).
\]

This contradicts Theorem 5.7. \(\square\)

6. Open questions

Let \(X\) be a one-sided topologically mixing shift of finite type. An invariant measure \(\mu\) on \(X\) is a Gibbs measure corresponding to \(f \in C(X)\) if there are constants \(C_1, C_2 > 0\) and \(P > 0\) such that
\[
C_1 \leq \frac{\mu([x_0x_1 \ldots x_{n-1}])}{\exp(-Pn + (S_n f)(x))} \leq C_2
\]
for every \(x \in X\) and \(n \geq 1\), where \((S_n f)(x) = \sum_{k=0}^{n-1} f(T^k(x))\).

Walters [24, 25] introduces a class \(\text{Bow}(X)\) of functions that contains the functions with summable variation, all of which have unique equilibrium states. Let
\[
\text{var}_n(f) = \sup\{ |f(x) - f(y)|: x, y \in X, x_i = y_i \text{ for all } 0 \leq i \leq n - 1\}.
\]

Then \(\text{Bow}(X, T) = \{ f \in C(X): \sup_{n \geq 1} \text{var}_n(S_n f) < \infty \}\).

**Theorem 6.1.** ([24, Theorem 2.16]) Let \(f \in \text{Bow}(X)\). Then \(f\) has a unique equilibrium state \(\mu\) which is a Gibbs measure.

The relative version of this result is an open question. We make the following conjecture.

**Conjecture 6.2.** Let \((X, Y, \pi)\) be a factor triple and \(\nu\) be a fully supported ergodic measure on \(Y\). For any function \(f \in \text{Bow}(X)\) the number of ergodic measures of maximal pressure of \(f\) in the fibre \(\pi^{-1}\{\nu\}\) is at most \(c_\pi\).

**References**


Department of Mathematics and Statistics, University of Victoria, PO Box 3045
STN CSC, Victoria British Columbia, Canada V8W 3P4

Current address: Centro de Modelamiento Matemático, Universidad de Chile, Av. Blanco Encalada 2120 Piso 7, Santiago, Chile

Department of Mathematics and Statistics, University of Victoria, PO Box 3045
STN CSC, Victoria British Columbia, Canada V8W 3P4