ENTROPY CONJUGACY FOR MARKOV MULTI-MAPS OF THE INTERVAL

JAMES P. KELLY AND KEVIN MCGOFF

Abstract. We consider a class $\mathcal{F}$ of Markov multi-maps on the unit interval. Any multi-map gives rise to a space of trajectories, which is a closed, shift-invariant subset of $[0,1]^{\mathbb{Z}_+}$. For a multi-map in $\mathcal{F}$, we show that the space of trajectories is (Borel) entropy conjugate to an associated shift of finite type. Additionally, we characterize the set of numbers that can be obtained as the topological entropy of a multi-map in $\mathcal{F}$.

1. Introduction

Multi-maps, also called set-valued maps, have been studied in the topological dynamics literature for some time, with such notable examples as [1, 21, 22]. In the past decade multi-maps have been studied extensively, with a particular focus on the topological structure of the associated space of trajectories or a related inverse limit space; see [15]. This development has also led to a renewed interest in the dynamics of multi-maps [11, 13, 17, 18]. Additionally, multi-maps are the topological analogues of random maps of the interval, which have received substantial attention, e.g., [4, 10, 14, 23].

In the study of single-valued maps of the interval, Markov maps [7] are particularly well-understood. These maps have a finite invariant set such that the map is strictly monotone on the intervals between elements of that set. This structure allows one to associate to each Markov interval map a corresponding shift of finite type that preserves many aspects of the dynamics.

Recent work [3, 5, 6, 12] has generalized the notion of Markov interval maps to the setting of multi-maps and established some of their basic properties. In particular, [2] proves that under some conditions on the Markov multi-map, one may find upper and lower bounds for its entropy using associated shifts of finite type.

Our main results substantially sharpen this previous work. Under mild conditions on the Markov multi-map, we associate to it a single shift of finite type, and then we establish a close connection (in the form of a Borel entropy conjugacy) between the dynamics of the multi-map and its associated shift of finite type. In particular, for a Markov multi-map in the class $\mathcal{F}$ considered here, the topological entropies of the Markov multi-map and of the associated shift of finite type must be equal. Furthermore, we demonstrate the richness of the class $\mathcal{F}$ by showing that any number that appears as the entropy of a shift of finite type also appears as the entropy of a Markov multi-map in $\mathcal{F}$.

1.1. **Statement of main results.** A multi-map of the unit interval is a function $F : [0,1] \to 2^{[0,1]}$, where $2^{[0,1]}$ is taken to be the set of closed subsets of $[0,1]$. Given such a multi-map, its trajectory space $X = X(F)$ is defined by

$$X(F) = \Big\{ x = (x_n)_{n=0}^{\infty} \in [0,1]^{\mathbb{Z}_+} : \forall n \geq 0, x_{n+1} \in F(x_n) \Big\}.$$

Further, let $\sigma_X : X \to X$ denote the left-shift map $(x_n)_{n=0}^{\infty} \mapsto (x_{n+1})_{n=0}^{\infty}$ on X. We seek to understand the multi-map F by studying the dynamics of the system (X, σ_X).

In Section 3 we introduce a class of multi-maps that we call Markov multi-maps, and in Definition 3.4 we state what it means for a Markov multi-map to be properly parametrized. In this work we focus on a specific class $\mathcal{F}$ of Markov multi-maps (see Definition 3.9): properly parametrized Markov multi-maps with complete sets of coding and avoiding words and positive entropy. For any multi-map F in this class, one may associate to F a square matrix $M = M(F)$ with entries in $\{0,1\}$ (see Section 3.2). The matrix M encodes the combinatorial structure of F. Let Σ_M be the shift of finite type defined by M, with left-shift map σ_M. The following theorem provides a precise correspondence between a "large" subset of the trajectory space X and a "large" subset of the SFT Σ_M, where "large" here refers to a notion of entropy. A precise definition of Borel entropy conjugacy, originally defined by Buzzi [9] under the term "entropy conjugacy," appears in Definition 2.2.

Theorem 1.1. *Let $\mathcal{F}$ be the class of Markov multi-maps specified in Definition 3.9. Let F be in $\mathcal{F}$ with trajectory space X and associated SFT Σ_M. Then (X, σ_X) is Borel entropy conjugate to (Σ_M, σ_M).*

Since Borel entropy conjugacy is known to preserve topological entropy, we immediately obtain the following corollary.

Corollary 1.2. *Let F be in $\mathcal{F}$ with trajectory space X and associated SFT Σ_M. Then $\mathrm{h}_{\mathrm{top}}(X, \sigma_X) = \mathrm{h}_{\mathrm{top}}(\Sigma_M, \sigma_M)$.*

In fact, since Borel entropy conjugacy provides a correspondence between all ergodic measures with large enough entropy, the following corollary is also immediate.

Corollary 1.3. *Let F be in $\mathcal{F}$ with trajectory space X and associated SFT Σ_M. Then (X, σ_X) has the same number of measures of maximal entropy as (Σ_M, σ_M). In particular, if (Σ_M, σ_M) is irreducible, then (X, σ_X) is intrinsically ergodic (i.e., has a unique measure of maximal entropy).*

Remark 1.4. Random maps of the interval have been studied primarily with an eye towards the existence and properties of absolutely continuous invariant measures, e.g., see [4, 7]. While the entropy conjugacy guaranteed by Theorem 1.1 provides a correspondence between ergodic measures of large entropy on X and on Σ_M, it does not address questions about whether any of these measures is absolutely continuous on X.

Let us now answer a question of Karl Petersen (personal communication). Let $\mathcal{H}(\mathcal{F})$ denote the set of real numbers $r > 0$ such that there exists a multi-map $F \in \mathcal{F}$ having trajectory space X with $\mathrm{h}_{\mathrm{top}}(X, \sigma_X) = r$. Recall that Lind has characterized the set of positive real numbers that arise as the entropy of a SFT as the set of all positive rational multiples of logarithms of Perron numbers [19].

Theorem 1.5. *The set $\mathcal{H}(\mathcal{F})$ is equal to the set of all positive rational multiples of logarithms of Perron numbers.*

1.2. **Organization of the paper.** In Section 2, we provide background information and notation concerning shifts of finite type, ergodic theory, and Borel entropy conjugacy. Section 3 introduces Markov multi-maps and the class $\mathcal{F}$ of interest. Taken together, Sections 4 – 7 contain the proof of our main result, Theorem 1.1. In Section 8 we establish some sufficient conditions for a Markov multi-map to be in $\mathcal{F}$, and then in Section 9 we prove the realization result, Theorem 1.5. Finally, Section 10 contains some examples of Markov multi-maps.

2. Background and notation

We denote by $2^{[0,1]}$ the set of all non-empty, closed subsets of $[0, 1]$. A *multi-map* on $[0, 1]$ is a function $F\colon [0, 1] \to 2^{[0,1]}$. The *graph* of a multi-map F is the set $G(F) = \{(x, y) \in [0, 1]^2 \colon y \in F(x)\}$. A *trajectory* for F is a sequence $(x_0, x_1, \ldots) \in [0, 1]^{\mathbb{Z}_+}$ such that for all $n \geq 1$, we have $x_n \in F(x_{n-1})$, or equivalently $(x_{n-1}, x_n) \in G(F)$. We denote by $X = X(F)$ the set of trajectories for F, and we give

X the topology it inherits as a subspace of $[0,1]^{\mathbb{Z}_+}$ with the product topology. We also define the *left-shift* on X, denoted σ_X, by setting $\sigma_X(x_0, x_1, \ldots) = (x_1, x_2, \ldots)$. Observe that σ_X is a continuous mapping on X, and if $G(F)$ is closed in $[0,1]^2$, then X is closed in $[0,1]^{\mathbb{Z}_+}$.

2.1. **Shifts of finite type.** Let $\mathcal{A}$ be a finite set, which we call the *alphabet*. An element $b \in \mathcal{A}^n$ is called a *word* of length n. The full shift on $\mathcal{A}$ is $\Sigma = \mathcal{A}^{\mathbb{Z}_+}$, endowed with the product topology induced by the discrete topology on $\mathcal{A}$. Given a set of words $\mathcal{F}$, we may define $\Sigma_{\mathcal{F}} \subseteq \Sigma$ to be the set of points that do not contain any word in $\mathcal{F}$. We refer to words in $\mathcal{F}$ as *forbidden words*. Then $\Sigma_{\mathcal{F}}$ is closed and invariant under the left-shift on Σ. If $\mathcal{F}$ is finite, then we refer to $\Sigma_{\mathcal{F}}$ as a *shift of finite type (SFT)*. In this work we restrict attention to SFTs for which all the forbidden words have length two, called nearest neighbor SFTs. For more on SFTs, we refer the reader to the book [20].

Any nearest neighbor SFT may be expressed in terms of a directed graph, (V, E), where the set of vertices V is equal to $\mathcal{A}$, and given $a, b \in \mathcal{A}$, there is an edge from a to b in the edge set E if and only if $ab \notin \mathcal{F}$. Furthermore, we associate to any such graph its *adjacency matrix* M, defined as the square matrix indexed by $\mathcal{A}$ such that for $a, b \in \mathcal{A}$, if $ab \notin \mathcal{F}$ then $M(a,b) = 1$, and otherwise $M(a,b) = 0$. Note that any zero-one matrix indexed by $\mathcal{A}$ also defines an associated nearest neighbor SFT (by letting ab be a forbidden word whenever $M(a,b) = 0$). The nearest neighbor SFT defined by a zero-one matrix M is denoted by Σ_M, and the left-shift restricted to Σ_M is denoted by σ_M.

In what follows it is convenient to have some notation for words of arbitrary length that do not contain any forbidden word. For $n \geq 2$, we let $\mathcal{L}_n$ denote the set of words $a_0 \ldots a_{n-1} \in \mathcal{A}^n$ such that $M(a_i, a_{i+1}) = 1$ for each $i = 0, \ldots, n-2$. Then let

$$\mathcal{L} = \bigcup_{n \geq 1} \mathcal{L}_n,$$

where $\mathcal{L}_1 = \mathcal{A}$.

A nearest neighbor SFT defined by the matrix M is *irreducible* if for every pair of non-empty, open sets $U, V \subseteq \Sigma_M$, there exists $n \geq 1$ such that $\sigma_M^n(U) \cap V \neq \emptyset$. Equivalently, Σ_M is irreducible if for each $a, b \in \mathcal{A}$, there exists $n \geq 1$ such that $M^n(a,b) > 0$.

Consider an arbitrary nearest neighbor SFT Σ_M on alphabet $\mathcal{A}$. It has an associated finite directed graph Γ, with vertex set $\mathcal{A}$ and an edge from a to b whenever $M(a,b) = 1$. Let $\mathcal{C}_1, \ldots, \mathcal{C}_K \subset \mathcal{A}$ be the vertex sets of the maximal strongly connected components of Γ, which we call the irreducible components of Γ. For each $\mathcal{C}_k$, the set of points in Σ_M containing only symbols from $\mathcal{C}_k$ forms an irreducible SFT, which we denote by $\Sigma_M(\mathcal{C}_k)$. We refer to $\Sigma_M(\mathcal{C}_k)$ as an irreducible component of Σ_M. Note that the irreducible components $\Sigma_M(\mathcal{C}_1), \dots, \Sigma_M(\mathcal{C}_k)$ are pairwise disjoint, and the set $\Sigma_M \setminus \bigcup_k \Sigma_M(\mathcal{C}_k)$ contains only wandering points. See [20, Chapter 4] for more details on this decomposition. We also denote by $\mathcal{L}(\mathcal{C}_k)$ the set of words of arbitrary length on $\mathcal{C}_k$ that do not contain a forbidden word.

2.2. **Invariant measures and entropy.** In this work a topological dynamical system consists of a pair $(\mathcal{X}, T)$, where $T : \mathcal{X} \to \mathcal{X}$ is a continuous self-map of a compact metrizable space. For any such system, we let $\mathcal{M}(\mathcal{X}, T)$ denote the set of Borel probability measures μ on $\mathcal{X}$ such that $\mu(E) = \mu(T^{-1}E)$ for all Borel sets $E \subset \mathcal{X}$. Note that $\mathcal{M}(\mathcal{X}, T)$ is a nonempty, convex set that is compact in the weak* topology. A measure $\mu \in \mathcal{M}(\mathcal{X}, T)$ is called ergodic if $\mu(E) \in \{0, 1\}$ for all Borel sets E such that $T^{-1}(E) \subset E$. The set of ergodic measures is denoted by $\mathcal{M}_e(\mathcal{X}, T)$. Note that a measure $\mu \in \mathcal{M}(\mathcal{X}, T)$ is an extreme point in $\mathcal{M}(\mathcal{X}, T)$ if and only if μ is ergodic.

The following notation is used in subsequent sections. For any Borel set $E \subset \mathcal{X}$, the union of all of its pre-images is denoted

$$\mathrm{Pre}(E) = \bigcup_{n \geq 0} T^{-n}(E).$$

Note that $T^{-1}(\mathrm{Pre}(E)) \subset \mathrm{Pre}(E)$, and therefore if $\mu \in \mathcal{M}_e(\mathcal{X}, T)$ then $\mu(\mathrm{Pre}(E)) \in \{0, 1\}$.

We also require some elementary facts regarding the entropy theory of dynamical systems. Complete definitions and proofs can be found in [24]. Let $\mathrm{h_{top}}(\mathcal{X}, T)$ denote the topological entropy of the topological system $(\mathcal{X}, T)$. Furthermore, when the system $(\mathcal{X}, T)$ is understood and $\mu \in \mathcal{M}(\mathcal{X}, T)$, we denote the measure-theoretic entropy of μ by $h(\mu)$. The standard variational principle for entropy states that

$$\mathrm{h_{top}}(\mathcal{X}, T) = \sup_{\mu \in \mathcal{M}(\mathcal{X}, T)} h(\mu),$$

and the supremum may be taken over only the ergodic measures. Furthermore, for SFTs it is known that the supremum is achieved, and if the SFT is irreducible, then it has a unique measure of maximal entropy. Furthermore, we note for future use that an irreducible SFT is entropy minimal, i.e., if X is an irreducible SFT of positive entropy and Y is a strict subset of X, then $\mathrm{h_{top}}(Y, \sigma|_Y) < \mathrm{h_{top}}(X, \sigma|_X)$ (see [20] for a proof).

2.3. **Entropy conjugacy.** We adopt the following definition of entropy for Borel sets (following Buzzi [9]).

Definition 2.1. Let $T : \mathcal{X} \to \mathcal{X}$ be a topological dynamical system. For a Borel set $E \subset \mathcal{X}$, let

$$\mathrm{h}_{\mathrm{prob}}(E) = \sup\bigl\{h(\mu) : \mu \in \mathcal{M}_e(\mathcal{X}, T),\ \mu(E) > 0\bigr\}.$$

Now we define a notion of entropy conjugacy, which was previously introduced by Buzzi [9].

Definition 2.2. Suppose that $T_0 : \mathcal{X}_0 \to \mathcal{X}_0$ and $T_1 : \mathcal{X}_1 \to \mathcal{X}_1$ are topological dynamical systems. We say that they are *Borel entropy conjugate* if there exist Borel sets $E_0 \subset \mathcal{X}_0$ and $E_1 \subset \mathcal{X}_1$ and an invertible Borel bi-measurable map $\psi : \mathcal{X}_0 \setminus E_0 \to \mathcal{X}_1 \setminus E_1$ such that

- $\mathrm{h}_{\mathrm{prob}}(E_0) < \mathrm{h}_{\mathrm{top}}(\mathcal{X}_0, T_0)$;
- $\mathrm{h}_{\mathrm{prob}}(E_1) < \mathrm{h}_{\mathrm{top}}(\mathcal{X}_1, T_1)$; and
- $\psi \circ T_0 = T_1 \circ \psi$ on $\mathcal{X}_0 \setminus E_0$.

It is an easy corollary of the variational principle for topological dynamical systems that if $(\mathcal{X}_0, T_0)$ and $(\mathcal{X}_1, T_1)$ are Borel entropy conjugate, then $\mathrm{h}_{\mathrm{top}}(\mathcal{X}_0, T_0) = \mathrm{h}_{\mathrm{top}}(\mathcal{X}_1, T_1)$.

Remark 2.3. In his work on topological entropy for non-compact sets, Bowen introduced a notion that he called entropy conjugacy [8]. Bowen's definition of entropy conjugacy requires that the sets E_0 and E_1 have smaller topological entropy (in the dimension-theoretic sense defined in his paper) than the full system and that the conjugating map ψ is continuous. As such, Bowen's notion of entropy conjugacy is stronger than the notion of Borel entropy conjugacy defined above.

3. Markov multi-maps

We now give a precise definition of Markov multi-maps on the interval $[0, 1]$. This definition is based on the one given in [2], though our definition is slightly less general.

Definition 3.1. A Markov multi-map F of the interval $[0, 1]$ is defined by a tuple $(P, \mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, D, R, \{f_a\}_{a \in \mathcal{A}_0})$ satisfying the following conditions:

(1) $P = \{p_0, \dots, p_r\}$ is a partition of the interval $[0, 1]$ with $0 = p_0 < \cdots < p_r = 1$;
(2) $\mathcal{A} = \mathcal{A}_0 \sqcup \mathcal{A}_1 \sqcup \mathcal{A}_2$ is a finite set;
(3) $D : \mathcal{A} \to 2^{[0,1]}$, and for each $a \in \mathcal{A}$, there exists $p_i \in P$ such that

$$D(a) = \begin{cases} [p_i, p_{i+1}], & \text{if } a \in \mathcal{A}_0 \\ \{p_i\}, & \text{if } a \in \mathcal{A}_1 \cup \mathcal{A}_2; \end{cases}$$

(4) $R : \mathcal{A} \to 2^{[0,1]}$, and for each $a \in \mathcal{A}$, there exists $u \leq v$ in P such that $R(a) = [u, v]$ and

$$\begin{cases} u < v, & \text{if } a \in \mathcal{A}_0 \\ u < v \text{ and } R(a) \cap P = \{u, v\}, & \text{if } a \in \mathcal{A}_1 \\ u = v, & \text{if } a \in \mathcal{A}_2; \end{cases}$$

(5) for each $a \in \mathcal{A}_0$, the map $f_a : D(a) \to R(a)$ is a homeomorphism;

(6) $[0, 1] \subset \bigcup_{a \in \mathcal{A}} D(a)$.

3.1. **The graph of a Markov multi-map.** Let F be a Markov multi-map. For $a \in \mathcal{A}_0$, let $G(a)$ denote the graph of f_a. For $a \in \mathcal{A}_1 \cup \mathcal{A}_2$, let $G(a) = D(a) \times R(a)$. Then the graph of F is

$$G(F) = \bigcup_{a \in \mathcal{A}} G(a).$$

Note that each $G(a)$ is closed in $[0, 1] \times [0, 1]$, and so is $G(F)$. An example of a Markov multi-map and its graph is given in Example 3.6, and more examples are shown in Section 10.

Now we make some additional graph-related definitions that are used repeatedly throughout this work.

Definition 3.2. Let $a \in \mathcal{A}$.

- Suppose $a \in \mathcal{A}_0$ with $D(a) = [p_i, p_{i+1}]$ and $R(a) = [u, v]$. Define $D_0(a) = (p_i, p_{i+1})$ and $R_0(a) = (u, v)$, and let $G_0(a)$ be the graph of $f_a|_{D_0(a)}$.
- Suppose $a \in \mathcal{A}_1$ with $D(a) = \{p\}$ and $R(a) = [u, v]$. Define $D_0(a) = \{p\}$ and $R_0(a) = (u, v)$, and let $G_0(a) = \{p\} \times R_0(a)$.
- Suppose $a \in \mathcal{A}_2$ with $D(a) = \{p\}$ and $R(a) = \{q\}$. Define $D_0(a) = \{p\}$ and $R_0(a) = \{q\}$, and let $G_0(a) = \{(p, q)\}$.

Our results require that F has some additional structure, which we now begin to define.

Definition 3.3. We say that F satisfies the *no crossing property* if the following holds: for all $a, b \in \mathcal{A}_0$, if $G_0(a) \cap G_0(b) \neq \varnothing$ then $a = b$.

The following property strictly implies the no crossing property.

Definition 3.4. We say that F is *properly parametrized* if the collection $\{G_0(a) : a \in \mathcal{A}\}$ forms a partition of $G(F)$.

We think of the no crossing property as a property of the graph $G(F)$ (and the partition P), whereas being properly parametrized depends on the particular parametrization of the Markov multi-map F. However, these properties are related by Lemma 4.1: if F_0 is a Markov multi-map with the no crossing property, then there exists a properly parametrized Markov multi-map F_1 such that $G(F_0) = G(F_1)$.

Remark 3.5. If F is a Markov multi-map, then it possesses the following graph *Markov property*: for all $a, b \in \mathcal{A}$, if $D_0(b) \cap R_0(a) \neq \varnothing$, then $D_0(b) \subset R_0(a)$. This property is used to define the SFT associated with F, which appears in the next section.

3.2. **The SFT associated to a Markov multi-map.** We associate to any Markov multi-map F an SFT as follows. Let M be the square matrix indexed by $\mathcal{A}$ such that for $a, b \in \mathcal{A}$,

$$M(a,b) = \begin{cases} 1, & \text{if } D_0(b) \subset R_0(a) \\ 0, & \text{otherwise.} \end{cases}$$

Let $\Sigma_M \subset \mathcal{A}^{\mathbb{Z}_+}$ be the nearest neighbor SFT with alphabet $\mathcal{A}$ and adjacency matrix M.

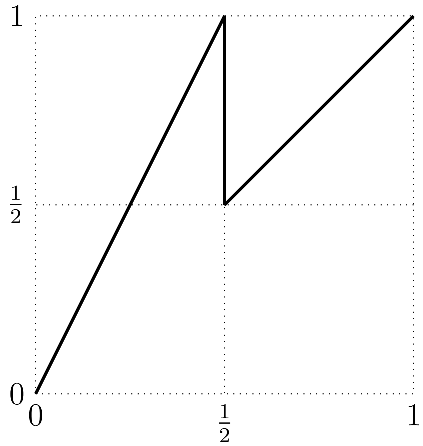
To show how a Markov multi-map may be properly parametrized and how the adjacency matrix is constructed, we present the following example of a fairly simple Markov multi-map with the no-crossing property.

Example 3.6. The graph in Figure 1 shows the graph of a Markov multi-map on $[0,1]$ with the partition $P = \{0, 1/2, 1\}$. To properly parametrize it, we first consider the two diagonal lines, and we index them with the set $\mathcal{A}_0 = \{1, 2\}$. Then the vertical line segment is indexed by $\mathcal{A}_1 = \{3\}$. Finally this leaves the four endpoints of those segments which we index with $\mathcal{A}_2 = \{4, 5, 6, 7\}$. Then we define the following:

$$\begin{array}{llll} D(1) = [0, 1/2] & R(1) = [0, 1] & D(2) = [1/2, 1] & R(2) = [1/2, 1] \\ D(3) = \{1/2\} & R(3) = [1/2, 1] & D(4) = \{0\} & R(4) = \{0\} \\ D(5) = \{1/2\} & R(5) = \{1\} & D(6) = \{1/2\} & R(6) = \{1/2\} \\ D(7) = \{1\} & R(7) = \{1\}. & & \end{array}$$

Then for all $a \in \mathcal{A} = \mathcal{A}_0 \cup \mathcal{A}_1 \cup \mathcal{A}_2$, if $D(a)$ is a singleton, we let $D_0(a) = D(a)$, and if $D(a)$ is an interval of the form $D(a) = [p, q]$, then $D_0(a) = (p, q)$. We define each $R_0(a)$ likewise.

Next we construct the adjacency matrix corresponding to this graph. Note that $R_0(1) = (0, 1)$ which contains as subsets all of the following: $D_0(1)$, $D_0(2)$, $D_0(3)$, $D_0(5)$, $D_0(6)$. However $D_0(4)$ and $D_0(7)$ are not subsets of $R_0(1)$. This produces the first row of the adjacency matrix which has a 1 in every column except for the 4th and 7th. We may



$$M = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

FIGURE 1. The graph of a Markov multi-map and its corresponding adjacency matrix

do the same for each element of $\mathcal{A}$ to produce the adjacency matrix pictured in Figure 1.

In our main results, we relate the SFT Σ_M to the trajectory space X. In particular, Theorem 1.1 establishes sufficient conditions for these systems to be Borel entropy conjugate.

3.3. **Nested intervals.** Let F be a properly parametrized Markov multi-map with associated matrix M. Here we associate to each sequence $\mathbf{a} \in \Sigma_M$ a nonempty closed (possibly degenerate) interval in $[0, 1]$. To begin, for each $a \in \mathcal{A}_0$, we let f_a^{-1} be the standard inverse function (which exists since f_a is assumed to be a homeomorphism). For $a \in \mathcal{A}_1 \cup \mathcal{A}_2$, we let f_a^{-1} be the unique map such that $f_a^{-1} : R(a) \to D(a)$ (which exists since $R(a)$ is non-empty and $D(a)$ is a singleton in this case).

Let $u = a_0 \ldots a_n \in \mathcal{L}_{n+1}$. Define the set

$$I_u = f_{a_0}^{-1} \circ \cdots \circ f_{a_{n-1}}^{-1}(D(a_n))$$

We make the following elementary observations.

- I_u is non-empty. (Since $M(a_i, a_{i+1}) = 1$, we have that $D(a_{i+1}) \subset R(a_i)$, so $f_{a_i}^{-1}$ maps $D(a_{i+1})$ into $D(a_i)$.)
- I_u is a closed (possibly degenerate) interval, since I_u is the image of the closed interval $D(a_n)$ under the continuous, monotone map $f_{a_0}^{-1} \circ \cdots \circ f_{a_{n-1}}^{-1}$.
- $I_{a_0 \ldots a_{n+1}} = f_{a_0}^{-1}(I_{a_1 \ldots a_{n+1}})$.

- $I_{a_0\dots a_{n+1}} \subset I_{a_0\dots a_n}$ (since $f_{a_n}^{-1}(D(a_{n+1})) \subset D(a_n)$).

Now consider $\mathbf{a} = (a_n)_{n=0}^\infty \in \Sigma_M$. Then $\{I_{a_0\dots a_n}\}_{n=1}^\infty$ is a nested sequence of non-empty, closed intervals in $[0,1]$. Let

$$I_{\mathbf{a}} = \bigcap_{n=1}^\infty I_{a_0\dots a_n}.$$

Then $I_{\mathbf{a}}$ is a non-empty, closed interval. Additionally, we note that

$$I_{\mathbf{a}} = f_{a_0}^{-1}(I_{\sigma(\mathbf{a})}).$$

These intervals appear in the next section in the definitions that characterize the class of Markov multi-maps in our main results.

3.4. **Definition of the class $\mathcal{F}$.** In this section we define the class $\mathcal{F}$ of Markov multi-maps that appears in our main results. Let F be a properly parametrized Markov multi-map with associated matrix M.

Definition 3.7. Suppose $\mathcal{C} \subset \mathcal{A}$ is an irreducible component of the graph with adjacency matrix M. We say that $\mathcal{C}$ has a *coding word* if there exists $u \in \mathcal{L}(\mathcal{C})$ such that if $\mathbf{a} \in \Sigma_M(\mathcal{C})$ and $\{n \geq 0 : \sigma^n(\mathbf{a}) \in [u]\}$ is infinite, then $I_{\mathbf{a}}$ is a singleton. Furthermore, we say that F has a *complete set of coding words* if each irreducible component with positive entropy has a coding word.

Definition 3.8. Suppose $\mathcal{C} \subset \mathcal{A}$ is an irreducible component of the graph with adjacency matrix M. We say that $\mathcal{C}$ has an *avoiding word* if there exists $u \in \mathcal{L}(\mathcal{C})$ such that if $\mathbf{a} \in [u]$, then $I_{\mathbf{a}} \cap P = \varnothing$. Furthermore, we say that F has a *complete set of avoiding words* if the following condition holds: if $\mathcal{C}$ is an irreducible component with positive entropy that is entirely contained in $\mathcal{A}_0$, then $\mathcal{C}$ has an avoiding word.

Now we are prepared to give a precise definition of the class of Markov multi-maps that appears in our main results.

Definition 3.9. The class $\mathcal{F}$ consists of all properly parametrized Markov multi-maps F such that F has a complete set of coding words, F has a complete set of avoiding words, and the associated SFT Σ_M has positive entropy.

3.5. **Finite labeled trajectories.** Here we define some additional terminology that is useful in the following sections.

Definition 3.10. Let $F \in \mathcal{F}$. Let $m \geq 1$. We say that $x = x_0, \dots, x_m \in [0,1]^{m+1}$ is a *finite trajectory* of F if

$$(x_n, x_{n+1}) \in G(F), \quad \forall n = 0, \dots, m-1.$$

Next, we say that $(x, b) \in [0,1]^{m+1} \times \mathcal{L}_m$ is a *finite labeled trajectory* (of length $m+1$) if

$$(x_n, x_{n+1}) \in G(b_n), \quad \forall n = 0, \ldots, m-1.$$

Let $\mathcal{T}_m$ be the set of finite labeled trajectories of length $m+1$. We endow $\mathcal{T}_m$ with the subspace topology inherited from $[0,1]^{m+1} \times \mathcal{L}_m$ (which has the product of the usual topology on $[0,1]^{m+1}$ and the discrete topology on $\mathcal{L}_m$).

Finally, we say that $(x, b) \in \mathcal{T}_m$ is a *special finite labeled trajectory* of length $m+1$ if

$$(x_n, x_{n+1}) \in G_0(b_n), \quad \forall n = 0, \ldots, m-1.$$

Let $\mathcal{S}_m$ denote the set of special finite labeled trajectories of length $m+1$, and we let $\mathcal{S}_m$ inherit the subspace topology inherited from $\mathcal{T}_m$.

Remark 3.11. Let F be in $\mathcal{F}$, and let x be a finite trajectory of F of length $m+1$. Since $G(F)$ is the union of the sets $\{G(a)\}_{a\in\mathcal{A}}$, there exists $b \in \mathcal{L}_m$ such that (x, b) is in $\mathcal{T}_m$. In fact, since F is properly parametrized, the sets $\{G_0(a)\}_{a\in\mathcal{A}}$ form a partition of $G(F)$, and therefore there exists a *unique* element $b \in \mathcal{L}_m$ such that (x, b) is in $\mathcal{S}_m$.

4. Preliminary results

4.1. **Parametrization lemma.** The following simple result states that any Markov multi-map with the no-crossing property can be properly parametrized without changing its graph. Since the space of trajectories of a Markov multi-map depends only on its graph, this reparametrization also preserves the space of trajectories.

Lemma 4.1. *Suppose that F_0 is a Markov multi-map with the no-crossing property. Then there exists a properly parametrized Markov multi-map F_1 with $G(F_0) = G(F_1)$.*

Proof. Let F be a Markov multi-map with the no-crossing property. Let $\mathcal{B}_0 = \mathcal{A}_0$. Let

$$\mathcal{B}_1 = \left\{ \{p_i\} \times [p_j, p_{j+1}] : p_i, p_j \in P \text{ and } \exists a \in \mathcal{A}_1, \{p_i\} \times [p_j, p_{j+1}] \subset G(a) \right\}.$$

Let

$$\mathcal{B}_2 = \left\{ (p, q) : p, q \in P, (p, q) \in G(F) \right\}.$$

Then let F_1 be the Markov multi-map defined by $\mathcal{B}_0$, $\mathcal{B}_1$, and $\mathcal{B}_2$. $\square$

4.2. **Graph lemmas.** In this section we prove a few facts about graphs of Markov multi-maps. Throughout the remainder of this section, we consider $F \in \mathcal{F}$. Since F is properly parametrized, we know that if $a, b \in \mathcal{A}$ are distinct elements, then $G_0(a) \cap G_0(b) = \varnothing$. However, it is possible that $a \neq b$ and yet $G(a)$ has nontrivial intersection with $G(b)$. The following lemma shows that any such intersection must be contained in $P \times P$.

Lemma 4.2. *Let F be in $\mathcal{F}$. Suppose $(x, y) \in G(F) \setminus (P \times P)$. Then there is a unique $a \in \mathcal{A}$ such that $(x, y) \in G(a)$, and furthermore $(x, y) \in G_0(a)$.*

Proof. Since $(x, y) \in G(F) = \cup_a G(a)$, there must exist some $a \in \mathcal{A}$ such that $(x, y) \in G(a)$. For uniqueness, suppose that $(x, y) \in G(a) \cap G(b)$. By the no-crossing property, for any $a \neq b$, we have $G(a) \cap G(b) \subset P \times P$. Since $(x, y) \notin P \times P$, we conclude that $a = b$.

Since $(x, y) \notin P \times P$ and $(x, y) \in G(a)$, we see that $a \in \mathcal{A}_0 \cup \mathcal{A}_1$, and we must have $(x, y) \in G_0(a)$. □

The next lemma asserts that $G(F)$ cannot accumulate along a horizontal line to any point of $G(F) \cap (P \times P)$.

Lemma 4.3. *Let $(p, q) \in G(F) \cap (P \times P)$. Then there exists an open set $U \subset [0, 1] \times [0, 1]$ such that $(p, q) \in U$ and if $(y, q) \in G_0(a) \cap U$, then $y = p$ and a is the unique element of $\mathcal{A}_2$ such that $G(a) = \{(p, q)\}$.*

Proof. Let $L_q = [0, 1] \times \{q\}$. By our definition of Markov multi-map, $L_q \cap G(F)$ is a finite set containing (p, q). Then there exists a relatively open interval I in $[0, 1]$ such that $I \times \{q\} \cap G(F) = \{(p, q)\}$. Let $U = I \times [0, 1]$. Then U is open in $[0, 1] \times [0, 1]$ and if $(y, q) \in G_0(a)$, then $y = p$ and a must be the unique element of $\mathcal{A}_2$ such that $G(a) = \{(p, q)\}$. □

The next two lemmas address the convergence of sequences in the space of finite labeled trajectories.

Lemma 4.4. *Let $m \geq 1$. Suppose that $x = x_0, \dots, x_m \in [0, 1]^{m+1}$ is a finite trajectory of F such that*

- *$x_0, \dots, x_{m-1} \in P$, and*
- *$x_m \in [0, 1] \setminus P$.*

Let w be the unique element of $\mathcal{L}_m$ such that $(x, w) \in \mathcal{S}_m$. If the sequence $\{(y^k, b^k)\}_{k=1}^{\infty}$ is in $\mathcal{S}_m$ and converges to (x, b) in $[0, 1]^{m+1} \times \mathcal{A}^m$, then $b = w$.

Proof. By our hypotheses on $x_0, \ldots, x_m$, we have that $w_{m-1} \in \mathcal{A}_1$ and if $m \geq 2$, then $w_n \in \mathcal{A}_2$ for $n = 0, \ldots, m-2$. Let us now show that $b_0 \ldots b_{m-1} = w$.

First, note that since $\mathcal{A}^m$ has the discrete topology, for all large enough k, we have $b^k = b$. Then for all large enough k, we have $(y_n^k, y_{n+1}^k) \in G(b_n)$ for all $n = 0, \ldots, m-1$. Since $G(b_n)$ is closed and $\{(y_n^k, y_{n+1}^k)\}_{k=1}^{\infty}$ converges to (x_n, x_{n+1}), we see that $(x_n, x_{n+1}) \in G(b_n)$ for each $n = 0, \ldots, m-1$.

Since $x_m \notin P$, Lemma 4.2 gives that there is a unique $a \in \mathcal{A}$ such that $(x_{m-1}, x_m) \in G(a)$, and therefore we must have $b_{m-1} = a = w_{m-1}$. Furthermore, since $x_{m-1} \in P$ and $x_m \notin P$, we see that $w_{m-1} \in \mathcal{A}_1$, which implies that $D_0(w_{m-1}) = \{x_{m-1}\}$. Then since $(y_{m-1}^k, y_m^k) \in G_0(b_{m-1}^k) = G_0(b_{m-1}) = G_0(w_{m-1})$ for all large enough k and $D_0(w_{m-1}) = \{x_{m-1}\}$, we see that $y_{m-1}^k = x_{m-1}$ for all large enough k.

We claim by backwards induction that for each $j = 0, \ldots, m-1$, we have $b_j = w_j$ and $y_j^k = x_j$ for all large enough k. We have established the base case ($j = m-1$) in the preceding paragraph. Now suppose it holds for some $j+1$. Let U be given by Lemma 4.3 for the point (x_j, x_{j+1}). By the inductive hypothesis, for all large enough k, we have $y_{j+1}^k = x_{j+1} \in P$. Also, for all large enough n, we must have $(y_j^k, y_{j+1}^k) \in U$ (since $\{(y_j^k, y_{j+1}^k)\}_{k=1}^{\infty}$ converges to (x_j, x_{j+1})). Then for all large enough k, we have $(y_j^k, x_{j+1}) \in U \cap G_0(b_j)$. By our choice of U, we must have that $b_j = w_j$ and $y_j^k = x_j$ for all large enough k, which completes the induction. □

Lemma 4.5. *Let $m \geq 1$. Suppose that $x = x_0, \ldots, x_m \in [0,1]^{m+1}$ is a finite trajectory of F such that $x_m \in [0,1] \setminus P$. Let w be the unique element of $\mathcal{L}_m$ such that $(x, w) \in \mathcal{S}_m$. If the sequence $\{(y^k, b^k)\}_{k=1}^{\infty}$ is in $\mathcal{S}_m$ and converges to (x, b) in $[0,1]^{m+1} \times \mathcal{A}^m$, then $b = w$.*

Proof. As $\mathcal{A}^m$ has the discrete topology, we must have that $b^k = b$ for all large enough k. Then for all large enough k, we have $(y_n^k, y_{n+1}^k) \in G(b_n)$, which is closed, and therefore $(x_n, x_{n+1}) \in G(b_n)$.

Observe that if $(x_n, x_{n+1}) \in G(F) \setminus (P \times P)$, then $b_n = w_n$ by Lemma 4.2. Now suppose that we have some n such that $(x_n, x_{n+1}) \in P \times P$. Then there exists $N \in [n+2, m]$ such that $x_j \in P$ for all $j = n, \ldots, N-1$ and $x_N \notin P$. Thus $x_n, \ldots, x_N$ satisfies the conditions of Lemma 4.4, and we conclude that $b_n = w_n$. □

5. Construction of the joint system and factor maps

Let F be a Markov multi-map in $\mathcal{F}$ with associated trajectory space X and SFT Σ_M. In the following section, we introduce a topological dynamical system by taking limits of special finite labeled trajectories. We call this system the *joint system*. Then in Sections 5.2 and 5.3, we show that the joint system is in fact a common extension of X and Σ_M. The joint system and its factor maps onto X and Σ_M are central to the construction of the Borel entropy conjugacy in our proof of Theorem 1.1. We establish their key properties in this section.

5.1. The joint system.

Let F be a Markov multi-map in $\mathcal{F}$ with associated SFT Σ_M. Here we define a subset V of the product space $[0,1]^{\mathbb{Z}_+} \times \Sigma_M$, which will serve as a common extension of the trajectory space X and the SFT Σ_M.

Definition 5.1. Let F be in $\mathcal{F}$ with associated trajectory space X and SFT Σ_M. Define a set $V = V(F) \subset [0,1]^{\mathbb{Z}_+} \times \Sigma_M$ as follows. A pair $(x, \mathbf{a}) \in [0,1]^{\mathbb{Z}_+} \times \Sigma_M$ is in V if there exists a sequence $\{\ell_k\}_{k=1}^{\infty}$ of natural numbers tending to infinity and a sequence $\{(y^k, a^k)\}_{k=1}^{\infty}$ of special finite labeled trajectories, with $(y^k, a^k) \in \mathcal{S}_{\ell_k}$, such that for each $n \geq 0$, the sequence $\{(y_n^k, a_n^k)\}_{k=1}^{\infty}$ converges to (x_n, a_n) in $[0,1] \times \mathcal{A}$.

Proposition 5.2. *V is closed and invariant under the left shift.*

Proof. Suppose that $(x^m, \mathbf{a}^m)$ is a sequence in V that converges in $[0,1]^{\mathbb{Z}_+} \times \Sigma_M$ to $(x, \mathbf{a})$. For each m, we have that $(x^m, \mathbf{a}^m) \in V$, and therefore there exists a sequence of natural numbers $\{\ell(m,k)\}_{k=1}^{\infty}$ tending to infinity and a sequence of special finite labeled trajectories $(y^{m,k}, b^{m,k}) \in \mathcal{S}_{\ell(m,k)}$ such that for each m and n,

$$\lim_k b_n^{m,k} = a_n^m, \quad \text{and} \quad \lim_k y_n^{m,k} = x_n^m.$$

To complete the proof, we exhibit a sequence $\{\ell_j\}_{j=1}^{\infty}$ of natural numbers and a sequence $\{(z^j, c^j)\}_{j=1}^{\infty}$ of special finite labeled trajectories to demonstrate that $(x, \mathbf{a}) \in V$.

Let $j \geq 1$. First choose m_j such that for all $n = 0, \ldots, j$, we have $a_n^{m_j} = a_n$ and

$$\left| x_n^{m_j} - x_n \right| < \frac{1}{2j}.$$

Next choose k_j (depending on m_j) such that $\ell(m_j, k_j) \geq j$ and for all $n = 0, \ldots, j$, we have $b_n^{m_j, k_j} = a_n^{m_j}$ and

$$\left| y_n^{m_j, k_j} - x_n^{m_j} \right| < \frac{1}{2j}.$$

Finally, let $\ell_j = j$, and define

$$z^j = y_0^{m_j,k_j}, \ldots, y_j^{m_j,k_j}$$

and

$$c^j = b_0^{m_j,k_j} \ldots b_j^{m_j,k_j}.$$

Then $\{\ell_j\}_{j=1}^\infty$ tends to infinity and $\{(z^j, c^j)\}_{j=1}^\infty$ is a sequence of special finite labeled trajectories. Furthermore, for each n, we have that $\{(z_n^j, c_n^j)\}_{j=1}^\infty$ converges to (x_n, a_n) in $[0,1] \times \mathcal{A}$. We have thus exhibited the necessary sequences to establish that $(x, \mathbf{a}) \in V$.

To establish shift invariance, suppose that $(x, \mathbf{a}) \in V$. Let us show that $(\sigma(x), \sigma(\mathbf{a})) \in V$. There exists natural numbers $\{\ell_k\}_{k=1}^\infty$ and special finite trajectories $\{(y^k, a^k)\}_{k=1}^\infty$ that witness the fact that $(x, \mathbf{a}) \in V$. Define $z_n^k = y_{n+1}^k$ and $b_n^k = a_{n+1}^k$. Then the sequences $\{\ell_k - 1\}_{k=1}^\infty$ and $\{(z^k, b^k)\}_{k=1}^\infty$ establish that $(\sigma(x), \sigma(\mathbf{a})) \in V$. ☐

As V is invariant under the left shift, we define $\sigma_V : V \to V$ by letting $\sigma_V(x, \mathbf{a}) = (\sigma(x), \sigma(\mathbf{a}))$.

In the proof of our main results, we use the joint space V as an intermediary between the spaces X and Σ_M. To make this connection precise, we define factor maps from V onto each of X and Σ_M.

5.2. **Factoring onto Σ_M.** Here we show that the joint space V from Definition 5.1 factors onto Σ_M.

Definition 5.3. Let F be in $\mathcal{F}$ with associated trajectory space X, SFT Σ_M, and joint space V. Define the map $\phi : V \to \Sigma_M$ by the rule $\phi(x, \mathbf{a}) = \mathbf{a}$.

It is clear that ϕ is continuous and commutes with the left shift.

Remark 5.4. Note that if $M(a,b) = 1$, then $f_a^{-1}(D_0(b)) \subset D_0(a)$. Furthermore, if $y \in D_0(b)$ and $x = f_a^{-1}(y)$, then $(x, y) \in G_0(a)$. Thus, if $a_0 \ldots a_\ell \in \mathcal{L}_{\ell+1}$ and $y_\ell \in D_0(a_\ell)$, then for each $n = 0, \ldots, \ell-1$, we have

$$y_n = f_{a_n}^{-1} \circ \cdots \circ f_{a_{\ell-1}}^{-1}(y_\ell) \in D_0(a_n),$$

and $(y_n, y_{n+1}) \in G_0(a_n)$.

Proposition 5.5. *ϕ is surjective.*

Proof. Let $\mathbf{a} \in \Sigma_m$. For each $\ell \geq 1$, let $y_\ell^\ell \in D_0(a_\ell)$ be arbitrary. For $n = 0, \ldots, \ell-1$, let $y_n^\ell = f_{a_n}^{-1} \circ \cdots \circ f_{a_{\ell-1}}^{-1}(y_\ell^\ell)$. By Remark 5.4, for each $n = 0, \ldots, \ell-1$, we have $(y_n^\ell, y_{n+1}^\ell) \in G_0(a_n)$. Also, for each n, we have that $y_n^\ell \in [0,1]$, which is sequentially compact. Thus, by a diagonal argument, there exists a subsequence $\{\ell_k\}_{k=1}^{\infty}$ tending to infinity such that for each n, there exists $x_n \in [0,1]$ such that

$$\lim_{k\to\infty} y_n^{\ell_k} = x_n.$$

Setting $x = (x_n)_{n=0}^{\infty}$, we see that $(x, \mathbf{a}) \in V$ and $\phi(x, \mathbf{a}) = \mathbf{a}$. □

The following proposition asserts that ϕ preserves the entropy of ergodic measures. Its proof is an adaptation of the proof of [2, Theorem 4.1], and we provide it in Appendix A for completeness.

Proposition 5.6. *Let $\phi : V \to \Sigma_M$ be as in Definition 5.3. Furthermore, let $\mu \in \mathcal{M}_e(V, \sigma_V)$ and $\nu \in \mathcal{M}_e(\Sigma_M, \sigma_M)$ be such that $\nu = \mu \circ \phi^{-1}$. Then $h(\nu) = h(\mu)$.*

5.3. **Factoring onto X.** Here we show that the joint space V from Definition 5.1 factors onto X.

Definition 5.7. Let F be in $\mathcal{F}$ with associated trajectory space X, SFT Σ_M, and joint space V. Define the map $\pi : V \to [0,1]^{\mathbb{Z}_+}$ by the rule $\pi(x, \mathbf{a}) = x$.

It is clear that π is continuous and commutes with the left shift. The following result shows that the image of π is contained in X.

Proposition 5.8. *Suppose that $(x, \mathbf{a}) \in V$. Then $x \in X$.*

Proof. Let $n \geq 0$. Since $(x, \mathbf{a}) \in V$, there exists y_n^k and y_{n+1}^k such that $\lim_k y_n^k = x_n$, $\lim_k y_{n+1}^k = x_{n+1}$, and $(y_n^k, y_{n+1}^k) \in G(a_n)$. Since $G(a_n)$ is closed, we see that $(x_n, x_{n+1}) \in G(a_n)$ for each $n \geq 0$. Then by the definition of X, we have $x \in X$. □

By Proposition 5.8, we have $\pi : V \to X$. Next we establish that π in fact maps *onto* X. First, let $V_0 \subset V$ be the set of points $(x, \mathbf{a}) \in V$ such that for each $n \geq 0$, we have $(x_n, x_{n+1}) \in G_0(a_n)$.

Proposition 5.9. *For each $x \in X$, there exists a unique $\mathbf{a} \in \Sigma_M$ such that $(x, \mathbf{a}) \in V_0$. In particular, $\pi : V \to X$ is surjective.*

Proof. Let $x = (x_n)_{n=0}^{\infty} \in X$. Let $n \geq 0$. Since $(x_n, x_{n+1}) \in G(F)$ and $\{G_0(a) : a \in \mathcal{A}\}$ is a partition of $G(F)$, there is a unique element $a_n \in \mathcal{A}$ such that $(x_n, x_{n+1}) \in G_0(a_n)$. This uniquely defines a sequence $\mathbf{a} = (a_n)_{n=0}^{\infty}$.

Let us show that $\mathbf{a} \in \Sigma_M$. Since Σ_M is a SFT defined by the matrix M, it suffices to show that for each $n \geq 1$, we have $M(a_{n-1}, a_n) = 1$. Let $n \geq 1$. By construction, we have that $(x_{n-1}, x_n) \in G_0(a_{n-1})$ and $(x_n, x_{n+1}) \in G_0(a_n)$. Then $x_n \in R_0(a_{n-1})$ and $x_n \in D_0(a_n)$, and therefore $D_0(a_n) \cap R_0(a_{n-1}) \neq \varnothing$. By the Markov property, we conclude that $D_0(a_n) \subset R_0(a_{n-1})$, and therefore $M(a_{n-1}, a_n) = 1$, as desired.

Finally, note that $(x, \mathbf{a}) \in V$. Indeed, for each $k \geq 1$, let $\ell_k = k$, $y_n^k = x_n$ and $b_n^k = a_n$. Then we have exhibited the necessary sequences to establish that $(x, \mathbf{a}) \in V$. □

6. Constructing the bad sets

Our aim is to show that under certain conditions, we can construct a Borel entropy conjugacy between X and Σ_M. In this construction, we identify "bad sets", on which the Borel entropy conjugacy map will not be defined. The main source of difficulty in constructing our Borel entropy conjugacy arises from the fact that points in the trajectory space that stay in the critical set P can have multiple symbolic codings. In order to deal with this difficulty, we group such symbolic codings into the "bad sets" and show that we have only removed sets of strictly smaller entropy that the full system. In fact, we carry out this process for each irreducible component of Σ_M separately. In the following section, we define the critical set of points in X that cause us difficulty. Then in the following sections we analyze the irreducible components in detail and construct their bad sets.

6.1. **The critical system.** Let F be in $\mathcal{F}$ with trajectory space X, SFT σ_M, and joint system V. Consider the set of trajectories contained in the critical set P:

$$X_P = \big\{x \in X : \forall n \geq 0, x_n \in P\big\}.$$

Note that X_P is closed and invariant under σ_X. We refer to X_P as the critical system. Now let $Z = \pi^{-1}(X_P) \subset V$, and note that Z is closed and invariant under σ_V. As we mentioned above, one of the main difficulties in relating X and Σ_M lies in the fact that π may not be injective on Z (or its pre-images under the shift).

6.2. **Irreducible components.** We find it useful to distinguish between the following types of irreducible components for Markov multi-maps.

Definition 6.1. Let F be in $\mathcal{F}$ with associated SFT Σ_M. Let $\mathcal{C} \subset \mathcal{A}$ be an irreducible component of the M-graph. We say that

- $\mathcal{C}$ is of Type I if $\mathcal{C} \subset \mathcal{A}_0$;
- $\mathcal{C}$ is of Type II if $\mathcal{C} \subset \mathcal{A}_2$;
- $\mathcal{C}$ is of Type III if it is not Type I or Type II.

Remark 6.2. Suppose $\mathcal{C}$ is of Type III. Then for each $i \in \{0,1,2\}$, we must have $\mathcal{C} \cap \mathcal{A}_i \neq 0$. In fact, there must exist allowable transitions in $\mathcal{C}$ from $\mathcal{A}_0$ to $\mathcal{A}_2$ (cross-over), from $\mathcal{A}_2$ to $\mathcal{A}_1$ (into a vertical line), and from $\mathcal{A}_1$ to $\mathcal{A}_0$ (out of vertical line).

Let $\Sigma_M(\mathcal{C})$ denote the irreducible component of Σ_M corresponding to $\mathcal{C}$. Note that $\Sigma_M(\mathcal{C})$ is an SFT contained in Σ_M. Also, for distinct irreducible components $\mathcal{C}_1$ and $\mathcal{C}_2$, we have that $\Sigma_M(\mathcal{C}_1)$ and $\Sigma_M(\mathcal{C}_2)$ are disjoint. Let $V(\mathcal{C})$ denote $\phi^{-1}(\Sigma_M(\mathcal{C})) \subset V$, where V is the joint system. Also, let $Z(\mathcal{C}) = Z \cap V(\mathcal{C})$, where $Z = \pi^{-1}(X_P) \subset V$.

6.3. **Constructing the bad sets: Types I and III.** We now show the existence of our "bad sets" off of which ϕ and π are injective. In the proof of the following proposition, we use the following immediate consequence of Lemma 4.5: if $(x, \mathbf{a}), (x, \mathbf{b}) \in V$ and $x_m \notin P$, then for each $n < m$, we have $a_n = b_n$. Also, for notation, for any word $w \in \mathcal{L}$ and any $\mathbf{a} \in \Sigma_M$, let

$$N_w(\mathbf{a}) = \big|\big\{n \geq 0 : \sigma^n(\mathbf{a}) \in [w]\big\}\big|.$$

Proposition 6.3. *Let F be in $\mathcal{F}$ with SFT Σ_M and joint system V. Suppose that $\mathcal{C}$ is a Type I or Type III irreducible component of the M-graph such that* $\mathrm{h}_{\mathrm{top}}(\Sigma_M(\mathcal{C}), \sigma_M|_{\Sigma_M(\mathcal{C})}) > 0$. *Then there exists words u^c and u^a in $\mathcal{L}(\mathcal{C})$ such that if*

$$B_0 = \{\mathbf{a} \in \Sigma_M(\mathcal{C}) : N_{u^c}(\mathbf{a}) < \infty \text{ or } N_{u^a}(\mathbf{a}) < \infty\},$$

and $B = \phi^{-1}(B_0)$, then

(1) $\mathrm{Pre}(Z(\mathcal{C})) \subset B$,
(2) $\mathrm{h}_{\mathrm{prob}}(B) < \mathrm{h}_{\mathrm{top}}(V(\mathcal{C}), \sigma|_{V(\mathcal{C})})$, *and*
(3) *both π and ϕ are injective on $V(\mathcal{C}) \setminus B$.*

Proof. First, suppose that $\mathcal{C}$ is Type I. Since F is in $\mathcal{F}$, it has a complete set of coding words, and we may select a coding word u^c for $\mathcal{C}$. Similarly, since F is in $\mathcal{F}$, it has a complete set of avoiding words, and then since $\mathcal{C}$ is of Type I, we may select an avoiding word u^a for $\mathcal{C}$.

Now suppose that $\mathcal{C}$ is Type III. Since $\mathcal{C}$ is of Type III, it contains a word $u^a = w_0 w_1$ such that $w_0 \in \mathcal{A}_0$ and $w_1 \in \mathcal{A}_1 \cup \mathcal{A}_2$. Note that u^a is an avoiding word. Furthermore, since $\mathcal{C}$ is of Type III, it contains a symbol $u^c \in \mathcal{A}_1$. Note that u^c is a coding word for $\mathcal{C}$.

For the remainder of the proof, we do not distinguish between whether $\mathcal{C}$ is Type I or Type III.

Then let

$$B_0 = \{\mathbf{a} \in \Sigma_M(\mathcal{C}) : N_{u^c}(\mathbf{a}) < \infty \text{ or } N_{u^a}(\mathbf{a}) < \infty\},$$

and let $B = \phi^{-1}(B_0)$. (Note that B_0 and B are invariant.)

To establish (1), let $(x, \mathbf{a}) \in \mathrm{Pre}(Z(\mathcal{C}))$. Then there exists N such that for each $n \geq N$, we have $x \in I_{\sigma^n(\mathbf{a})} \cap P$. Since u^a is an avoiding word, we see that $N_{u^a}(\mathbf{a}) < \infty$. It follows that $\mathbf{a} \in B_0$, and therefore $(x, \mathbf{a}) \in B$.

Now we establish (2). Let $Y \subset \Sigma_M$ be the SFT obtained by forbidding u^c and u^a. Since $\Sigma_M(\mathcal{C})$ is irreducible, it is entropy minimal. Then $\mathrm{h_{top}}(Y, \sigma|_Y) < \mathrm{h_{top}}(\Sigma_M(\mathcal{C}), \sigma_M|_{\Sigma_M(\mathcal{C})})$, as Y is a strict subsystem of $\Sigma_M(\mathcal{C})$. Suppose μ is an ergodic measure on $\Sigma_M(\mathcal{C})$ such that $\mu(B_0) = 1$. As μ is ergodic and the words u^c and u^a appear only finitely for points in B_0, we must have $\mu([u^c]) = \mu([u^a]) = 0$, and therefore $\mu(Y) = 1$. Then by the variational principle, we see that $h(\mu) \leq \mathrm{h_{top}}(Y, \sigma|_Y)$. Taking the supremum over all such μ, we obtain that $\mathrm{h_{prob}}(B_0) \leq \mathrm{h_{top}}(Y, \sigma|_Y) < \mathrm{h_{top}}(\Sigma_M(\mathcal{C}), \sigma_M|_{\Sigma_M(\mathcal{C})})$. Furthermore, since ϕ preserves the entropy of ergodic measures (by Proposition 5.6), we obtain that $\mathrm{h_{prob}}(B) = \mathrm{h_{prob}}(B_0) < \mathrm{h_{top}}(\Sigma_M(\mathcal{C}), \sigma_M|_{\Sigma_M(\mathcal{C})}) = \mathrm{h_{top}}(V(\mathcal{C}), \sigma_V|_{V(\mathcal{C})})$.

To show that ϕ is injective on $V(\mathcal{C}) \setminus B$, let $(x, \mathbf{a}) \in V(\mathcal{C}) \setminus B$, and suppose $(y, \mathbf{a}) \in V(\mathcal{C}) \setminus B$. Then $\mathbf{a} \notin B_0$, and in fact $\sigma^n(\mathbf{a}) \notin B_0$ for all $n \geq 0$. Let $n \geq 0$. Then $\sigma^n(\mathbf{a})$ contains the word u^c infinitely many times, and therefore $I_{\sigma^n(\mathbf{a})}$ is a singleton (since u^c is a coding word). Since we must have both $x_n \in I_{\sigma^n(\mathbf{a})}$ and $y_n \in I_{\sigma^n(\mathbf{a})}$, we conclude that $x_n = y_n$. As $n \geq 0$ was arbitrary, we have shown that ϕ is injective on $V(\mathcal{C}) \setminus B$.

To show that π is injective on $V(\mathcal{C}) \setminus B$, let $(x, \mathbf{a}), (x, \mathbf{b}) \in V(\mathcal{C}) \setminus B$. Let $T = \{m \geq 0 : \sigma^m(\mathbf{a}) \in [u^a]\}$. For each $m \in T$, we have that $x_m \in I_{\sigma^m(\mathbf{a})} \subset [0,1] \setminus P$ (since u^a is an avoiding word). Since $(x, \mathbf{a}) \in V(\mathcal{C}) \setminus B$, the set T must be infinite. Let $n \geq 0$. Since T is infinite, there exists $m > n$ such that $m \in T$. Then $x_m \notin P$. By Lemma 4.5, we see that $a_n = b_n$. As $n \geq 0$ was arbitrary, we conclude that π is injective on $V(\mathcal{C}) \setminus B$. □

6.4. **Constructing the bad sets: Type II.** We don't have to remove any bad sets from Type II components. Indeed, the following proposition establishes that π and ϕ are injective on the union of all Type II components.

Proposition 6.4. *Let F be in $\mathcal{F}$, and let*

$$V_P = \bigcup_{\mathcal{C} \text{ of Type II}} V(\mathcal{C}).$$

Then π and ϕ are injective on V_P.

Proof. Suppose $(x, \mathbf{a}), (x, \mathbf{b}) \in V_P$. Then $a_n, b_n \in \mathcal{A}_2$ for all $n \geq 0$, and we must have $G(a_n) = \{(x_n, x_{n+1})\} = G(b_n)$ for all n. Therefore $a_n = b_n$ for all $n \geq 0$, and π is injective on V_P.

Suppose that $(x, \mathbf{a}), (y, \mathbf{a}) \in V_P$. Then $D(a_n)$ is a singleton for each n, and we must have $\{x_n\} = D(a_n) = \{y_n\}$ for all n. Therefore $x_n = y_n$ for all $n \geq 0$, and ϕ is injective on V_P. □

7. Proof of the main result

Now that we have constructed the bad sets for each type of irreducible component, we are ready to prove our main result on Borel entropy conjugacy.

Proof of Theorem 1.1. Let F be in $\mathcal{F}$ with associated trajectory space X and SFT Σ_M. Furthermore, let V be the associated joint space, as in Definition 5.1, and let $\phi : V \to \Sigma_M$ and $\pi : V \to X$ be the maps defined in Definitions 5.3 and 5.7, respectively.

Since $F \in \mathcal{F}$, we have that Σ_M has positive entropy. Enumerate the irreducible components with positive entropy: $\mathcal{C}_1, \dots, \mathcal{C}_J$. For each $\mathcal{C}_j$ of Type I or Type III, let $B_j \subset V(\mathcal{C}_j)$ be the bad set given by Proposition 6.3. For each $\mathcal{C}_j$ of Type II, let $B_j = \varnothing$. Furthermore, let

$$B_0 = \phi^{-1}\left(\Sigma_M \setminus \left(\bigcup_{j=1}^{J} \Sigma_M(\mathcal{C}_j)\right)\right).$$

Then let

$$B = \bigcup_{j=0}^{J} B_j.$$

For each $j = 1, \dots, J$, let $A_j = V(\mathcal{C}_j) \setminus B_j$. Note that

$$V \setminus B = \bigcup_{j=1}^{J} A_j. \tag{7.1}$$

Proposition 7.1. *π is injective on $V \setminus B$.*

Proof. Suppose that $(x, \mathbf{a}), (x, \mathbf{b}) \in V \setminus B$. By (7.1), there exists i, j such that $(x, \mathbf{a}) \in A_i$ and $(x, \mathbf{b}) \in A_j$. If $x_n \notin P$ for infinitely many n, then $\mathbf{a} = \mathbf{b}$ by Lemma 4.5.

Now suppose that $x_n \in P$ for all by finitely many n. Then there exists N such that $\sigma^N(x) \in X_P$. Hence $(x, \mathbf{a}), (x, \mathbf{b}) \in \mathrm{Pre}(Z)$, and therefore $\mathcal{C}_i$ and $\mathcal{C}_j$ must be of Type II (since $A_k \cap Z = \varnothing$ whenever $\mathcal{C}_k$ is of Type I or Type III). Since π is injective on V_P, we conclude that $\mathbf{a} = \mathbf{b}$. □

Proposition 7.2. *ϕ is injective on $V \setminus B$.*

Proof. Suppose that $(x, \mathbf{a}), (y, \mathbf{a}) \in V \setminus B$. By (7.1), there exists i, j such that $(x, \mathbf{a}) \in A_i$ and $(y, \mathbf{a}) \in A_j$. Then $\mathbf{a} \in \Sigma_M(\mathcal{C}_i) \cap \Sigma_M(\mathcal{C}_j)$. Since distinct irreducible components are disjoint, we see that $i = j$. Since ϕ is injective on A_i, we conclude that $x = y$. □

Lemma 7.3. $\mathrm{h_{top}}(X, \sigma_X) = \mathrm{h_{top}}(V, \sigma_V)$.

Proof. First, by Proposition 5.8, we have that π is a factor map. Since entropy cannot increase under a factor map, $\mathrm{h_{top}}(X, \sigma_X) \leq \mathrm{h_{top}}(V, \sigma_V)$.

Now fix $\mathcal{C}_i$ such that $\mathrm{h_{top}}(V, \sigma_V) = \mathrm{h_{top}}(V(\mathcal{C}_i), \sigma|_{V(\mathcal{C}_i)})$, and let μ be an ergodic measure on $V(\mathcal{C}_i)$ such that $h(\mu) = \mathrm{h_{top}}(V(\mathcal{C}_i), \sigma|_{V(\mathcal{C}_i)})$ (which exists $\Sigma_M(\mathcal{C}_i)$ has a measure of maximal entropy and ϕ preserves entropy). Since $\mathrm{h_{prob}}(B_i) < \mathrm{h_{top}}(V(\mathcal{C}_i), \sigma|_{V(\mathcal{C}_i)}) = h(\mu)$, we must have that $\mu(B_i) = 0$. Then π is injective a set of full μ-measure, and therefore π is an isomorphism from μ to $\pi\mu = \mu \circ \pi^{-1}$. In particular, $\mathrm{h_{top}}(V, \sigma|_V) = \mathrm{h_{top}}(V(\mathcal{C}_i), \sigma|_{V(\mathcal{C}_i)}) = h(\mu) = h(\pi\mu) \leq \mathrm{h_{top}}(X, \sigma|_X)$, where the last inequality follows from the Variational Principle. We have now shown that $\mathrm{h_{top}}(X, \sigma_X) = \mathrm{h_{top}}(V, \sigma_V)$. □

Let $A_X = \pi(V \setminus B)$ and $A_\Sigma = \phi(V \setminus B)$.

Proposition 7.4. *A_X is Borel, $\sigma(A_X) = A_X$, and $\mathrm{h_{prob}}(X \setminus A_X) < \mathrm{h_{top}}(X, \sigma|_X)$.*

Proof. Consider A_i. First suppose that $\mathcal{C}_i$ is of Type II. Then $A_i = V(\mathcal{C}_i)$, which is compact. Thus $\pi(A_i)$ is also compact. In particular, $\pi(A_i)$ is closed and hence Borel.

Now suppose that $\mathcal{C}_i$ is of Type I or Type III. Then there exist words u and v in $\mathcal{L}(\mathcal{C}_i)$ (in particular, a coding word and any avoiding word) such that

$$A_i = \bigcap_N \left[\left(\bigcup_{n \geq N} \sigma^{-n}[u] \right) \cap \left(\bigcup_{n \geq N} \sigma^{-n}[v] \right) \right].$$

Note that for each n, the sets $[u]$ and $[v]$ are compact. Hence $\pi[u]$ and $\pi[v]$ are compact and in particular closed. Then

$$\pi(A_i) = \bigcap_N \left[\left(\bigcup_{n \geq N} \sigma^{-n}\pi[u] \right) \cap \left(\bigcup_{n \geq N} \sigma^{-n}\pi[v] \right) \right],$$

which shows that $\pi(A_i)$ is Borel. Finally, since $A_X = \cup_i \pi(A_i)$, we conclude that A_X is Borel.

For each A_i, we have $\sigma(A_i) = A_i$, and therefore $\sigma(\pi(A_i)) = \pi(\sigma(A_i)) = \pi(A_i)$. As $A_X = \sqcup_i \pi(A_i)$, we see that $\sigma(A_X) = A_X$.

Let ν be an ergodic invariant measure on X such that $\nu(X \setminus A_X) > 0$. Since A_X is invariant and ν is ergodic, we have that $\nu(X \setminus A_X) = 1$, and therefore $\nu(A_X) = 0$. Also, ν is supported on some set of the form $\pi(B_i)$, with $1 \leq i \leq J$. Let μ be an ergodic measure on V such that $\pi\mu = \nu$. Then $\mu(A) \leq \mu(\pi^{-1}\pi(A)) = \nu(A_X) = 0$, and $\mu(V(\mathcal{C}_i)) = 1$. Therefore $\mu(B_i) = 1$. Finally, we observe that $h(\nu) \leq h(\mu) \leq \mathrm{h}_{\mathrm{prob}}(B_i) \leq \max_i \mathrm{h}_{\mathrm{prob}}(B_i)$. Since the right hand side is strictly less than $\mathrm{h}_{\mathrm{top}}(V, \sigma_V)$, which equals $\mathrm{h}_{\mathrm{top}}(X, \sigma_X)$ by Lemma 7.3, we conclude that $\mathrm{h}_{\mathrm{prob}}(X \setminus A_X) < \mathrm{h}_{\mathrm{top}}(X, \sigma|_X)$. □

The following proposition may be quite easily deduced from the definitions, and we omit its proof.

Proposition 7.5. *A_Σ is Borel, $\sigma(A_\Sigma) = A_\Sigma$, and* $\mathrm{h}_{\mathrm{prob}}(\Sigma_M \setminus A_\Sigma) < \mathrm{h}_{\mathrm{prob}}(\Sigma_M)$.

Now define $\psi = \pi|_{V\setminus B} \circ \phi|_{V\setminus B}^{-1} : A_\Sigma \to A_X$, which will serve as our Borel entropy conjugacy map.

Proposition 7.6. *ψ is bijective, bi-measurable, and commutes with the left shift.*

Proof. Taken together, Propositions 7.1 and 7.2 yield that ψ is bijective. Let $E \subset A_X$ be Borel. Then $\pi^{-1}(E) \cap (V \setminus B)$ is Borel. Also, since $\phi|_{V\setminus B}$ is an injective continuous map on the Borel set $V \setminus B$, it maps Borel sets to Borel sets. Therefore $\psi^{-1}(E) = \phi|_{V\setminus B}(\pi^{-1}(E) \cap (V \setminus B))$ is Borel measurable. Therefore ψ is Borel measurable. An analogous argument shows that ϕ^{-1} is also Borel measurable. Finally, since π and ϕ commute with the left shift, ψ also commutes with the left shift. □

By the previous propositions, we conclude that ψ is the desired Borel entropy conjugacy between X and Σ_M. □

8. Sufficient conditions for F to be in $\mathcal{F}$

Now that we have proved Theorem 1.1, we wish to highlight its utility by establishing some straightforward conditions that are sufficient for a Markov multi-map F to be in the family $\mathcal{F}$. We focus on the case where there is an irreducible component $\mathcal{C}$ that contains all of $\mathcal{A}_0$. For single-valued functions, this condition amounts to the topological transitivity of the system.

Definition 8.1. Suppose $\mathcal{C}$ is an irreducible component. We say that *F codes for points on $\mathcal{C}$* if

$$\limsup_n \max \{\ell\left(I_{a_0\cdots a_n}\right) : \mathbf{a} \in \Sigma_M(\mathcal{C})\} = 0.$$

Lemma 8.2. *Suppose $\mathcal{C}$ is an irreducible component with $\mathcal{A}_0 \subset \mathcal{C}$. If F codes for points on $\mathcal{C}$, then $\mathcal{C}$ has a coding word and an avoiding word.*

Proof. Since F codes for points on $\mathcal{C}$, by definition, we must have that $I_{\mathbf{a}}$ is a singleton for all $\mathbf{a} \in \Sigma_M(\mathcal{C})$. Therefore every word in $\mathcal{L}(\mathcal{C})$ is a coding word. To see that $\mathcal{C}$ also has an avoiding word, we consider two cases.

Case 1: Suppose $\mathcal{A}_0$ is a strict subset of $\mathcal{C}$. Then $\mathcal{C}$ must be a Type III component, and we showed in the proof of Proposition 6.3 that every Type III component has an avoiding word.

Case 2: Suppose $\mathcal{A}_0 = \mathcal{C}$. We have shown that $\mathcal{C}$ has a coding word, so there must exist $a, b \in \mathcal{A}_0$ such that $ab \in \mathcal{L}(\mathcal{A}_0)$ and I_{ab} is a strict subset of $I_a = D(a)$. This would imply that $D(b)$ is a strict subset of $R(a)$.

By the definition of a Markov multi-map, $D(b)$ is an interval between adjacent elements of the partition P. It follows that P partitions $R(a)$ into at least two intervals, so there exist distinct elements $p_i, p_j \in P$ such that $[p_i, p_{i+1}] \cup [p_j, p_{j+1}] \subseteq R(a)$. Then there must be $b_1, b_2 \in \mathcal{A}_0$ such that $D(b_1) = [p_i, p_{i+1}]$ and $D(b_2) = [p_j, p_{j+1}]$.

The interval I_{ab_1} is a strict subset of I_a, so it contains at most one endpoint of $D(a)$. Since $\mathcal{A}_0$ is irreducible, there exists $u \in \mathcal{L}(\mathcal{A}_0)$ such that $ab_1ua \in \mathcal{L}(\mathcal{A}_0)$. The interval I_{ab_1ua} is contained in I_{ab_1}, so it contains at most one endpoint of I_a. Then $I_{ab_1uab_1}$ and $I_{ab_1uab_2}$ are non-overlapping, so at least one of them is disjoint from P. Therefore $\mathcal{A}_0$ has an avoiding word. $\square$

Next we define what it means for F to be uniformly expanding on $\mathcal{C}$, and we show that if that is the case, then F codes for points on $\mathcal{C}$. Recall that for each $a \in \mathcal{A}$, we have a well-defined function $f_a^{-1} \colon R(a) \to D(a)$, and if $u = u_0 \cdots u_n \in \mathcal{L}$, then we define $f_u^{-1} = f_{u_0}^{-1} \circ \cdots \circ f_{u_n}^{-1}$.

Definition 8.3. Suppose $\mathcal{C}$ is an irreducible component and F is a Markov multi-map such that f_a is a diffeomorphism for each $a \in \mathcal{A}_0$. We say that F is uniformly expanding on $\mathcal{C}$ if there exists $N \in \mathbb{N}$ such that

$$\sup\left\{\left|(f_u^{-1})'(x)\right| : u \in \mathcal{L}_N(\mathcal{C}) \cap (\mathcal{A}_0)^N, x \in D(u_N)\right\} < 1.$$

Lemma 8.4. *Let $\mathcal{C}$ be an irreducible component. If F is uniformly expanding on $\mathcal{C}$, then F codes for points on $\mathcal{C}$.*

Proof. Let $0 < \lambda < 1$ such that $|(f_u^{-1})'(x)| < \lambda$ for all $u = a_0 \cdots a_{N-1} \in \mathcal{L}_N(\mathcal{C})$ and $x \in R(a_{N-1})$. Then $\ell(I_u) \leq \lambda \ell(R(a_{N-1})) \leq \lambda$. It follows that for all $k \geq 1$ and $u \in \mathcal{L}_{kN}(\mathcal{C})$, we have $\ell(I_u) < \lambda^k$. Therefore F codes for points on $\mathcal{C}$. $\square$

By combining these results, we arrive at the following sufficient condition for F to be in $\mathcal{F}$.

Corollary 8.5. *Let F be a properly parametrized Markov multi-map with associated SFT Σ_M. Suppose that $\mathrm{h}_{\mathrm{top}}(\Sigma_M, \sigma_M) > 0$, and furthermore there is an irreducible component $\mathcal{C}$ with $\mathcal{A}_0 \subset \mathcal{C}$. If F is uniformly expanding on $\mathcal{C}$, then $F \in \mathcal{F}$, and hence (X, σ_X) is entropy conjugate to (Σ_M, σ_M).*

Proof. By Lemma 8.4 and Lemma 8.2, the component $\mathcal{C}$ has a coding word and an avoiding word. Since $\mathcal{A}_0 \subset \mathcal{C}$, no irreducible component (except possibly $\mathcal{C}$) could be contained in $\mathcal{A}_0$, so F has a complete set of coding words and a complete set of avoiding words. Thus $F \in \mathcal{F}$, so by Theorem 1.1, (X, σ_X) is entropy conjugate to (Σ_M, σ_M). □

9. Realization of entropies

We now prove Theorem 1.5, which we restate here.

Theorem 1.5. *The set $\mathcal{H}(\mathcal{F})$ is equal to the set of all positive rational multiples of logarithms of Perron numbers.*

Proof. It suffices to show that for any irreducible SFT with positive entropy, there is a Markov multi-map in $\mathcal{F}$ with the same entropy.

Let Σ_M be an irreducible SFT with positive entropy associated with the $n \times n$ matrix M. Since Σ_M has positive entropy, there is one row of M with (at least) two ones. After possibly permuting the alphabet, suppose the first row has a one in columns k and $k+1$.

Now we define a Markov multi-map F on the interval $[1, n+2]$ in terms of its graph. (To illustrate our construction, we give a specific matrix M in Example 9.1, and we show that graph of the corresponding multi-map in Figure 2.)

Let
$$P = \left\{1, 1 + \frac{1}{2}, 2, 2 + \frac{1}{2}, \ldots, n, n + \frac{1}{2}, n+1, n + \frac{3}{2}, n+2\right\}.$$
For each $i, j \in \{1, \ldots, n\}$, we associate the rectangle $[i, i+1/2] \times [j, j+1/2]$ with the matrix entry $M(i,j)$. If $M(i,j) = 1$, we include (in the graph of F) a straight line connecting the bottom left corner (i,j) to the top right corner $(i+1/2, j+1/2)$, and if $M(i,j) = 0$, we leave that rectangle empty. There is one exception to this rule however. We have assumed that $M(1,k) = M(1,k+1) = 1$, and instead of including two separate graphs in those two rectangles, we include one line connecting $(1,k)$ to $(1+1/2, k+3/2)$.

In this way we would have the graph of a multi-map with domain $\bigcup_{i=1}^{n}[i, i+1/2]$. We need F to be defined on all of $[1, n+2]$, so we next define the graph in each rectangle of the form $[i+1/2, i+1] \times [n+3/2, n+2]$, where $i \in \{1, \ldots, n\}$. In these rectangles, we include a straight line connecting the points $(i+1/2, n+3/2)$ and $(i+1, n+2)$. Then finally, in each of the rectangles $[n+1, n+3/2] \times [n+3/2, n+2]$ and $[n+3/2, n+2] \times [n+3/2, n+2]$, we include a straight line connecting the bottom left corner to the top right corner.

We have described the graph of F, but in order to show it is a Markov multi-map in $\mathcal{F}$ we should specify the indexing set $\mathcal{A}$ and identify a coding word and an avoiding word. Let $\mathcal{C}_0$ be a labeling of all of the straight lines that correspond to ones in the matrix M. (Recall that the cardinality of $\mathcal{C}_0$ will be one less than the number of ones in M, because the ones in the $(1, k)$ and $(1, k+1)$ entries correspond to just one straight line in the graph.) Then let $\mathcal{B}_0$ be the additional straight lines whose ranges were all $[n+3/2, n+2]$, and define $\mathcal{A}_0 = \mathcal{C}_0 \cup \mathcal{B}_0$.

Each of the straight lines we considered have a bottom left endpoint and a top right endpoint. Let $\mathcal{C}_2$ and $\mathcal{B}_2$ be the collections of these left and right endpoints, respectively, and let $\mathcal{A}_2 = \mathcal{C}_2 \cup \mathcal{B}_2$. Finally let $\mathcal{A}_1 = \varnothing$.

Then $\mathcal{C}_0$ is a Type I irreducible component whose corresponding SFT has the same entropy as Σ_M. To complete the proof, we show that if $\Sigma(\mathcal{A})$ and $\Sigma(\mathcal{C}_0)$ are the SFTs associated with $\mathcal{A}$ and $\mathcal{C}_0$ respectively, then $\mathrm{h}_{\mathrm{top}}(\Sigma(\mathcal{A})) = \mathrm{h}_{\mathrm{top}}(\Sigma(\mathcal{C}_0))$. Towards this end, we show that the symbols in $\mathcal{B}_0, \mathcal{C}_2$, and $\mathcal{B}_2$ do not increase the entropy.

Let $b_0 \in \mathcal{B}_0$ represent the straight line in $[n+3/2, n+2] \times [n+3/2, n+2]$, then any $b \in \mathcal{B}_0$ can only be followed by b_0. This means $\mathrm{h}_{\mathrm{prob}}(\Sigma(\mathcal{A}_0)) = \mathrm{h}_{\mathrm{prob}}(\Sigma(\mathcal{C}_0))$. Now we consider $\mathcal{C}_2$ and $\mathcal{B}_2$. Each of these individually follows nearly the same pattern as $\mathcal{A}_0$ with only one difference. Let $a^* \in \mathcal{A}_0$ correspond to the straight line in $[1, 1/2] \times [k, k+3/2]$, and let $c^* \in \mathcal{C}_2$ and $b^* \in \mathcal{B}_2$ correspond to the respective endpoints of this line. The symbol a^* can be followed by any symbol whose domain is $[k, k+1/2]$ or $[k+1, k+3/2]$. On the other hand c^* can only be followed by points whose first coordinate is k, and b^* can only be followed by points whose first coordinate is $k+3/2$. The SFTs corresponding to $\mathcal{C}_2$ and $\mathcal{B}_2$ are disjoint from one another and are invariant. It follows that the entropy contributed by these sets is less than or equal to the entropy from $\mathcal{C}_0$.

All of this shows that F is a Markov multi-map whose associated SFT has the same entropy as (Σ_M, σ_M). It only remains to show that $F \in \mathcal{F}$. The only irreducible component in $\mathcal{A}_0$ with positive entropy is $\mathcal{C}_0$. We must show it has a coding word and an avoiding word. Once again let

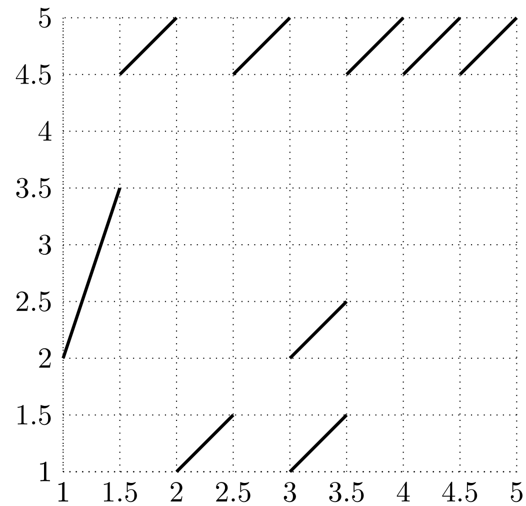


FIGURE 2. Markov multi-map from Example 9.1

$a^* \in \mathcal{A}_0$ correspond to the straight line in $[1, 1/2] \times [k, k + 3/2]$. The range $R(a^*)$ is partitioned by P into three non-overlapping intervals, so every occurrence of a^* in a word u decreases the length of I_u by a factor of 3. It follows that a^* is a coding word.

We can also use a^* to construct an avoiding word. Let $u = u_1 \cdots u_m \in \mathcal{L}(\mathcal{C}_0)$ be any word such that $a^*ua^* \in \mathcal{L}(\mathcal{C}_0)$. The interval I_{a^*u} is a strict subset of $[1, 1 + 1/2]$, so it contains at most one point of P. There is then an element $b \in \mathcal{C}_0$ such that $I_{a^*ua^*b}$ is disjoint from P and hence an avoiding word. Therefore $F \in \mathcal{F}$. $\square$

Example 9.1. Consider the 3×3 matrix

$$M = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

Using the method outlined in the proof of Theorem 1.5, this matrix would yield the graph pictured in Figure 2. Recall we stipulated that there must be at least two adjacent 1s in the first row of the matrix. These appear in the second and third columns of M. This gives us a line in the graph connecting the points $(1, 2)$ and $(1.5, 3.5)$.

For the rest of the graph, note that if we rotate the matrix counter-clockwise ninety degrees, then the pattern of 1s in the matrix matches the pattern of lines in the lower portion of the graph.

10. Examples

We show various examples demonstrating the utility of our results. We begin by showing that Theorem 1.1 generalizes the well-known result for the case that F is single-valued.

Example 10.1. Suppose F is any uniformly expanding (single-valued) Markov map. Then Corollary 8.5 recovers the well-known fact that F is entropy conjugate to its combinatorial SFT.

Next we give an example of a Markov multi-map that is not uniformly expanding but still satisfies the hypotheses of Theorem 1.1.

Example 10.2. Let $P = \{0, 1/3, 2/3, 1\}$, $\mathcal{A}_0 = \{1, 2, 3, 4\}$, $\mathcal{A}_1 = \emptyset$, and $\mathcal{A}_2 = \{5, \ldots, 10\}$. Let

$$\begin{aligned} D(1) &= [0, 1/3] & R(1) &= [1/3, 2/3] \\ D(2) &= [1/3, 2/3] & R(2) &= [0, 2/3] \\ D(3) &= [1/3, 2/3] & R(3) &= [2/3, 1] \\ D(4) &= [2/3, 1] & R(4) &= [1/3, 2/3]. \end{aligned}$$

For each $a \in \{1, 2, 3\}$, let $G(a)$ be a straight line from the bottom left corner to the top right corner of $D(a) \times R(a)$, and let $G(4)$ be a straight line from the top left to the bottom right of $D(4) \times R(4)$. Then we define $G(5) = \{(0, 1/3)\}$, $G(6) = \{(1/3, 2/3)\}$, $G(7) = \{(2/3, 1)\}$, $G(8) = \{(1/3, 0)\}$, $G(9) = \{(2/3, 1/3)\}$, $G(10) = \{(1, 1/3)\}$ (all of the endpoints of $G(1), \ldots, G(4)$). This defines a Markov multi-map whose graph is pictured in Figure 3.

Then $\mathcal{A}_0$ and $\mathcal{A}_2$ are both irreducible components with $\mathcal{A}_0$ (a Type I component) having greater entropy. The graph $G(3)$ has slope 2, but the rest of the graphs $G(1), G(2)$, and $G(4)$ have slope 1, so F is not uniformly expanding on $\mathcal{A}_0$. However, $3 \in \mathcal{L}_1$ is a coding word, because each time the symbol 3 appears in a word $u \in \mathcal{L}$, the length of the interval I_u is divided in half. Also $331 \in \mathcal{L}_3$ is an avoiding word, because $I_{331} = [3/6, 7/12]$. Thus $F \in \mathcal{F}$, so by Theorem 1.1 (X, σ_X) is entropy conjugate to (Σ_M, σ_M).

Next we show an example with a Type III irreducible component.

Example 10.3. Let $P = \{0, 1/2, 1\}$, $\mathcal{A}_0 = \{1, 2\}$, $\mathcal{A}_1 = \{3, 4\}$, and $\mathcal{A}_2 = \{5, 6, 7, 8, 9\}$. Let

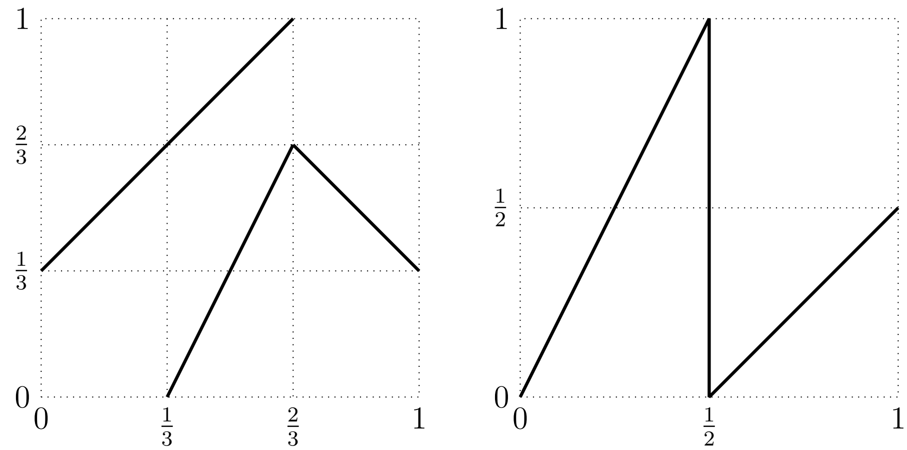


FIGURE 3. Markov multi-maps from Example 10.2 (left) and Example 10.3 (right)

$$\begin{aligned} D(1) &= [0, 1/2] & R(1) &= [0, 1] \\ D(2) &= [1/2, 1] & R(2) &= [0, 1/2] \end{aligned}$$

Let $G(1)$ be the straight line connecting $(0, 0)$ and $(1/2, 1)$, and let $G(2)$ be the straight line connecting $(1/2, 0)$ and $(1, 1/2)$. Then let $G(3)$ and $G(4)$ be the vertical lines $\{1/2\} \times [0, 1/2]$ and $\{1/2\} \times [1/2, 1]$ respectively. Finally, define $G(5), \dots, G(9)$ so that they are the endpoints of the graphs of $G(1), \dots, G(4)$. The graph of this Markov multi-map is pictured in Figure 3.

In this case, two symbols from $\mathcal{A}_2$ represent the points $\{(0, 0)\}$ and $\{(1/2, 0)\}$. For simplicity, say these are $G(8)$ and $G(9)$. Then $\mathcal{C} = \{1, \dots, 7\}$ is a Type III irreducible component which means it must have a coding and an avoiding word. In this case, we can use $u = 13 \in \mathcal{L}_2$ as both a coding and an avoiding word, because $I_{13} = \{1/4\}$. Since there is no Type I component, we automatically have $F \in \mathcal{F}$.

Finally we give an example that does not satisfy our hypotheses, and for which $\mathrm{h}_{\mathrm{top}}(\Sigma_M, \sigma_M)$ is strictly greater than $\mathrm{h}_{\mathrm{top}}(X, \sigma_X)$.

Example 10.4. Define a Markov multi-map as follows. Let $P = \{0, 1\}$, $\mathcal{A}_0 = \{1, 2\}$, $\mathcal{A}_1 = \varnothing$, and $\mathcal{A}_2 = \{3, 4\}$. Let $D(1) = D(2) = R(1) = R(2) = [0, 1]$. Let $f_1, f_2 \colon [0, 1] \to [0, 1]$ be defined by $f_1(x) = x^2$ and $f_2(x) = x^3$. Let $G(3) = \{(0, 0)\}$, and $G(4) = \{(1, 1)\}$.

Then the only non-trivial irreducible component is $\mathcal{A}_0 = \{1,2\}$, and $\Sigma_M(\mathcal{A}_0)$ is the full shift on two symbols, which has entropy $\log 2$. However, the only non-wandering points of (X, σ_X) are the fixed points $(0,0,\dots)$ and $(1,1,\dots)$, so $\mathrm{h}_{\mathrm{top}}(X, \sigma_X) = 0$.

Note that $I_u = [0,1]$ for all $u \in \mathcal{A}_0$, so this multi-map has neither coding words nor avoiding words. Thus $F \notin \mathcal{F}$, and Theorem 1.1 does not apply.

Acknowledgments

The authors wish to thank the anonymous referees for helpful comments regarding the presentation of the paper. KM gratefully acknowledges support from NSF grants DMS-1613261 and DMS-1847144.

References

[1] Ethan Akin. *The general topology of dynamical systems*, volume 1 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1993.

[2] Lori Alvin and James P. Kelly. Topological entropy of Markov set-valued functions. *to appear in Ergodic Theory and Dynamical Systems.*

[3] Lori Alvin and James P. Kelly. Markov set-valued functions and their inverse limits. *Topology Appl.*, 241:102–114, 2018.

[4] Wael Bahsoun, Christopher Bose, and Anthony Quas. Deterministic representation for position-dependent random maps. *Discrete & Continuous Dynamical Systems - A*, 22, 2008.

[5] Iztok Banič and Tjaša Lunder. Inverse limits with generalized Markov interval functions. *Bull. Malays. Math. Sci. Soc.*, 39(2):839–848, 2016.

[6] Iztok Banič and Matevž Črepnjak. Markov pairs, quasi Markov functions and inverse limits. *Houston J. Math.*, 44(2):695–707, 2018.

[7] Rufus Bowen. Invariant measures for Markov maps of the interval. *Comm. Math. Phys.*

[8] Rufus Bowen. Topological entropy for noncompact sets. *Transactions of the American Mathematical Society*, 184:125–136, 1973.

[9] Jérôme Buzzi. Intrinsic ergodicity of smooth interval maps. *Israel J. Math.*, 100(1):125–161, 1997.

[10] Jérôme Buzzi. Exponential decay of correlations for random lasota–yorke maps. *Communications in mathematical physics*, 208(1):25–54, 1999.

[11] Welington Cordeiro and Maria José Pacífico. Continuum-wise expansiveness and specification for set-valued functions and topological entropy. *Proc. Amer. Math. Soc.*, 144(10):4261–4271, 2016.

[12] Matevž Črepnjak and Tjaša Lunder. Inverse limits with countably Markov interval functions. *Glas. Mat. Ser. III*, 51(71)(2):491–501, 2016.

[13] Goran Erceg and Judy Kennedy. Topological entropy on closed sets in $[0,1]^2$. *Topology Appl.*, 246:106–136, 2018.

[14] Gary Froyland. Ulam's method for random interval maps. *Nonlinearity*, 12(4):1029, 1999.

[15] W. T. Ingram. *An introduction to inverse limits with set-valued functions.* SpringerBriefs in Mathematics. Springer, New York, 2012.
[16] Anatole Katok. Lyapunov exponents, entropy and periodic orbits for diffeomorphisms. *Publications Mathématiques de l'IHÉS*, 51:137–173, 1980.
[17] James P. Kelly and Tim Tennant. Topological entropy of set-valued functions. *Houston J. Math.*, 43(1):263–282, 2017.
[18] Judy Kennedy and Van Nall. Dynamical properties of shift maps on inverse limits with a set valued function. *Ergodic Theory Dynam. Systems*, 38(4):1499–1524, 2018.
[19] D. A. Lind. The entropies of topological markov shifts and a related class of algebraic integers. *Ergodic Theory and Dynamical Systems*, 4(2):283?300, 1984.
[20] Douglas Lind and Brian Marcus. *An introduction to symbolic dynamics and coding*. Cambridge university press, 1995.
[21] Richard McGehee. Attractors for closed relations on compact Hausdorff spaces. *Indiana Univ. Math. J.*, 41(4):1165–1209, 1992.
[22] Walter Miller and Ethan Akin. Invariant measures for set-valued dynamical systems. *Trans. Amer. Math. Soc.*, 351(3):1203–1225, 1999.
[23] Stephan Pelikan. Invariant densities for random maps of the interval. *Transactions of the American Mathematical Society*, 281(2):813–825, 1984.
[24] Peter Walters. *An introduction to ergodic theory*, volume 79. Springer Science & Business Media, 2000.

Appendix A. Proof of Proposition 5.6

Here we aim to prove Proposition 5.6. First, we recall the result of Katok [16] relating the measure-theoretic entropy of an ergodic measure to Bowen balls. Consider a compact metric space $(\mathcal{X}, d)$ and a continuous transformation $T : \mathcal{X} \to \mathcal{X}$. For $n \geq 1$, define the metric d_n on $\mathcal{X}$ by setting

$$d_n(x, y) = \max\bigl\{d\bigl(T^k(x), T^k(y)\bigr) : k = 0, \dots, n-1\bigr\}.$$

For $\epsilon > 0$, an (n, ϵ)-ball is a ball of radius ϵ with respect to the metric d_n. Now let μ be in $\mathcal{M}_e(\mathcal{X}, T)$. For $\alpha \in (0, 1)$, $\epsilon > 0$, and $n \geq 1$, let $s(T, n, \epsilon, \alpha)$ denote the minimal cardinality of a collection of (n, ϵ)-balls whose union has μ-measure at least α. Katok showed that

$$h(\mu) = \lim_{\epsilon \to 0^+} \limsup_n \frac{1}{n} \log s(T, n, \epsilon, \alpha).$$

Let us now prove that the factor map $\phi : V \to \Sigma_M$ preserves the entropy of all ergodic measures.

Proof of Proposition 5.6. As entropy cannot increase under factor maps, we have $h(\nu) \leq h(\mu)$. To complete the proof, we establish the reverse inequality. Fix $\alpha \in (0, 1)$. For $n \geq 1$, let $r(n, \alpha)$ denote the minimal cardinality of a set of words $W \subset \mathcal{L}_n$ such that

$$\nu\left(\bigcup_{w\in W}[w]\right) \geq \alpha. \tag{A.1}$$

Let $\epsilon > 0$. For $n \geq 1$, let $s(n,\epsilon,\alpha)$ denote the minimal cardinality of a collection $\mathcal{U}$ of (n,ϵ) balls in V such that

$$\mu\left(\bigcup_{U\in\mathcal{U}} U\right) \geq \alpha.$$

Now let $n \geq 1$. Select a set $W = \{w_1,\dots,w_K\} \subset \mathcal{L}_n$ with cardinality $K = r(n,\alpha)$ and satisfying (A.1). By the construction given in the proof of [2, Theorem 4.1], for each k, there exists a collection $\mathcal{U}_k$ of (n,ϵ) balls in V such that

$$\phi^{-1}([w_k]) \subset \bigcup_{U\in\mathcal{U}_k} U,$$

and

$$|\mathcal{U}_k| \leq (n+1)\left(\left\lceil\frac{1}{\epsilon}\right\rceil + 1\right).$$

Let $\mathcal{U} = \cup_k \mathcal{U}_k$. Note that

$$\mu\left(\bigcup_{U\in\mathcal{U}}\right) \geq \mu\left(\bigcup_{k=1}^{K}\phi^{-1}([w_k])\right) = \nu\left(\bigcup_{k=1}^{K}[w_k]\right) \geq \alpha,$$

and furthermore

$$|\mathcal{U}| \leq \sum_{k=1}^{K}|\mathcal{U}_k| \leq (n+1)\left(\left\lceil\frac{1}{\epsilon}\right\rceil + 1\right)r(n,\alpha).$$

Hence

$$s(n,\epsilon,\alpha) \leq |\mathcal{U}| \leq (n+1)\left(\left\lceil\frac{1}{\epsilon}\right\rceil + 1\right)r(n,\alpha).$$

Taking the limit supremum of this inequality as n tends to infinity yields

$$\limsup_n \frac{1}{n}\log s(n,\epsilon,\alpha) \leq \limsup_n \frac{1}{n}\log r(n,\alpha).$$

By Katok [16], as we let ϵ tend to 0, we obtain

$$h(\mu) \leq h(\nu).$$

$\square$

(J. P. Kelly) Department of Mathematics, Christopher Newport University, Newport News, VA 23606, USA
Email address: james.kelly@cnu.edu

(K. McGoff) Department of Mathematics and Statistics, University of North Carolina at Charlotte, Charlotte, NC 28223, USA
Email address: kmcgoff1@uncc.edu