Partially Observable Systems and Quotient Entropy via Graphs

by

Leonhard Horstmeyer and Sharwin Rezagholi

Preprint no.: 46 2019
Partially Observable Systems and Quotient Entropy via Graphs

Leonhard Horstmeyer

Basic Research Community for Physics
Complexity Science Hub Vienna

Sharwin Rezagholi

Max Planck Institute for Mathematics in the Sciences, Leipzig

July 11, 2019

Abstract

We consider the category of partially observable dynamical systems, to which the entropy theory of dynamical systems extends functorially. This leads us to introduce quotient-topological entropy. We discuss the structure that emerges. We show how quotient entropy can be explicitly computed by symbolic coding. To do so, we make use of the relationship between the category of dynamical systems and the category of graphs, a connection mediated by Markov partitions and topological Markov chains.

(37B10 Symbolic dynamics, 37B40 Topological entropy, 37C15 Topological equivalence, conjugacy, invariants, 54H20 Topological dynamics)

Contents

1 Introduction 2
2 Dynamical systems and partial observability 2
3 Entropy and quotient entropy 6
4 Symbolic dynamics and the category of graphs 10
5 Quotient entropy and Markov partitions 15

References 20

horstmeyer [at] csh [dot] ac [dot] at
1 Introduction

Science may be conceived as the attempt to understand dynamical systems that we may only partially observe: Scientists must understand dynamical systems via their actions on quotients of their domain. (See [ABB+16] for a survey of problems of this kind.) In this context, the most basic complexity indicator, topological entropy, can be extended to partially observable dynamical systems as quotient-topological entropy, which can be used to quantify the complexity reduction due to partial observability. Just as topological entropy, quotient-topological entropy is hard to compute explicitly. For partially observable dynamical systems which admit Markov partitions, we provide a scheme to calculate their quotient-topological entropy. This scheme is the computation of the entropy of a topological Markov chain whose entropy is equal to the respective quotient-entropy. These chains are generated by images of the graphs that generate the symbolic representations of the base systems.

Acknowledgements. SR benefited immensely from discussions with Renan Assimos, Florio Ciaglia, Jürgen Jost, Slava Matveev, and Paolo Perrone.

2 Dynamical systems and partial observability

The basic mathematical entities of this work are from the category $\textbf{Comp}$ of compact Hausdorff topological spaces and continuous surjective maps. We are interested in dynamical systems obtained by the iteration of endomorphisms. A dynamical system is a tuple $(X, f)$ where $X \in \text{ob}(\textbf{Comp})$ and $f \in \text{hom}_{\text{Comp}}(X, X)$. It generates the continuous monoid action $X \times \mathbb{N}_0 \to X$ given by $(x, t) \mapsto f^t(x)$. These are the objects of the category $\textbf{Sys}$. A morphism $m : (X, f) \to (Y, g)$ is an $m \in \text{hom}_{\text{Comp}}(X, Y)$ such that the following diagram commutes.

$$
\begin{array}{ccc}
X & \xrightarrow{f} & X \\
m & & m \\
\downarrow & & \downarrow \\
Y & \xrightarrow{g} & Y
\end{array}
$$

We now define the category $\textbf{PSys}$ of partially observable systems. The objects of $\textbf{PSys}$ are triples $(X, f, q_X)$ where $(X, f) \in \text{ob}(\textbf{Sys})$ and $q_X \in \text{hom}_{\text{Comp}}(X, -)$. A morphism in $\textbf{PSys}$ is a tuple $(m, n)$, where $m \in \text{hom}_{\text{Sys}}((X, f), (Y, g))$ and $n \in \text{hom}_{\text{Comp}}(\tilde{X}, \tilde{Y})$, such that the following diagram commutes.

$$
\begin{array}{ccc}
X & \xrightarrow{f} & X & \xrightarrow{q_X} & \tilde{X} \\
m & & m & & n \\
\downarrow & & \downarrow & & \downarrow \\
Y & \xrightarrow{g} & Y & \xrightarrow{q_Y} & \tilde{Y}
\end{array}
$$
Remark 1 (\(\text{PSys}\) is a comma-category). There is a forgetful functor \(F : \text{Sys} \to \text{Comp}\) given by the assignment \((X, f) \mapsto X\) on objects and the assignment \(m \mapsto m\) on morphisms. The functor \(F\) forgets the dynamics.

\[
\begin{array}{ccc}
  X & \xrightarrow{f} & X \\
  \downarrow m & & \downarrow m \\
  Y & \xrightarrow{g} & Y
\end{array}
\]

The pair of functors \((F, \text{id}_{\text{Comp}})\) generates the comma-category \(\frac{F}{\text{id}}\). Its objects are triples of the form \(((X, f), \tilde{X}, q_X)\) where \((X, f)\) is a dynamical system, \(\tilde{X}\) is a compact Hausdorff space, and \(q_X : X \to \tilde{X}\) is a continuous surjection. A morphism between \(((X, f), \tilde{X}, q_X)\) and \(((Y, g), \tilde{Y}, q_Y)\) is a tuple \((m, n)\), where \(m : (X, f) \to (Y, g)\) is a morphism in \(\text{Sys}\) and \(n : \tilde{X} \to \tilde{Y}\) is a morphism in \(\text{Comp}\). We obtain the category \(\text{PSys}\) by switching to compact notation.

There is a forgetful functor \(U : \text{PSys} \to \text{Sys}\) which forgets the partial observability. On objects we have \(((X, f), q) \mapsto (X, f)\) and on morphisms we have \(((m, n) \mapsto m)\). There is a functorial embedding \(E : \text{Sys} \to \text{PSys}\) where on objects we have \((X, f) \mapsto (X, f, \text{id})\), and on morphisms we have \(m \mapsto (m, m)\). The partially observable systems \((X, f, q)\) where \(q : X \to \tilde{X}\) is a morphism of \(\text{Sys}\) admit the projection \((X, f, q) \mapsto (\tilde{X}, \tilde{f})\). The archetype of systems which admit these projections are the product systems.

Remark 2. We have an adjunction \(U \vdash E\). The monad generated by this adjunction is trivial since \(U \circ E = \text{id}_{\text{Sys}}\). The comonad generated by the endofunctor \(E \circ U : \text{PSys} \to \text{PSys}\) is more interesting. The coalgebras of this comonad consist of objects in \(\text{PSys}\) of the form \((X, f, q)\) where \(q : X \to \tilde{X}\) is a morphism of \(\text{Sys}\) admit the projection \((X, f, q) \mapsto (\tilde{X}, \tilde{f})\). The archetype of systems which admit these projections are the product systems.

We close this section by characterizing quotient maps which are morphisms of dynamical systems.

Definition 1. Let \(X\) be a set. The canonical order induced onto \(2^{2^X}\) by set inclusion in \(2^X\) is such that for \(\mathcal{X}, \mathcal{Y} \in 2^{2^X}\) we have \(\mathcal{X} \sqsubseteq \mathcal{Y}\) if and only if for every \(A \in \mathcal{X}\) there exists \(B \in \mathcal{Y}\) fulfilling \(A \subseteq B\).

Let \(\mathcal{A}, \mathcal{B} \in 2^{2^X}\). Their infimum is \(\mathcal{A} \sqcap \mathcal{B} = \{(A \cap B)\}_{A \in \mathcal{A}, B \in \mathcal{B}}\). Recall that the kernel of a map \(f : X \to Y\) is the partition of \(X\) given as

\[
\ker(f) = \{(f^{-1}(y))\}_{y \in Y}.
\]

The finest partition of a set \(X\) is the partition into points \(\{\{x\}\}_{x \in X}\), the kernel of a homeomorphism. We will abuse notation by simply writing \(\ker(f)\) for \(\frac{X}{\ker(f)}\). The following obvious lemma is the topological analogon to the isomorphism theorem for vector spaces. The latter is crucial in the characterization of morphisms of evolution equations in Banach spaces [AR17] and in turn for flows on finite-dimensional manifolds [HA16].

3
Lemma 3. Let $f : X \to Y$ be a continuous map between compact Hausdorff spaces. Then $\ker(f) \cong \text{im}(f)$, where we equipped $\ker(f)$ with the quotient topology.

Proof. Preimages of points under $f$ are closed and therefore the induced quotient space is homeomorphic to $\ker(f)$. Consider the map $h : \ker(f) \to \text{im}(f)$ where $h([x]) = f(x)$. The map $h$ is a continuous bijection by construction. A continuous bijection between compact Hausdorff spaces is a homeomorphism.

Proposition 1 (Characterization of morphisms). Let $(X, f)$ be a dynamical system and consider a continuous surjection $q_X : X \to \tilde{X}$. Then $q_X$ is a morphism of dynamical systems if and only if

$$\ker(q_X) \subseteq \ker(q_X \circ f).$$

Proof. Suppose that $q_X$ is a morphism. Let $A \in \ker(q_X \circ f)$. Hence there exists a $y_0 \in Y$ such that $A = (q_X \circ f)^{-1}(y_0)$. Since $q_X$ is a morphism there exists $g : \tilde{X} \to \tilde{X}$ such that

$$A = (g \circ q_X)^{-1}(y_0) = q_X^{-1}(g^{-1}(y_0)).$$

Recall that $g^{-1}(y_0) = \{y_1 \in \tilde{X} : g(y_1) = y_0\}$. Hence $A = \bigcup_{y_1 \in g^{-1}(y_0)} q_X^{-1}(y_1)$, which is a union of cells in $\ker(q_X)$. Hence every $A \in \ker(q_X \circ f)$ contains at least one $B \in \ker(q_X)$ and therefore $\ker(q_X) \subseteq \ker(q_X \circ f)$.

Now, suppose that eq. (1) holds. By lemma 3 we have $\ker(q_X) \cong \tilde{X}$. Let $h$ be the respective homeomorphism. Consider the commutative diagram

$$\begin{array}{ccc}
X & \xrightarrow{f} & X \\
\downarrow{q_X} & & \downarrow{q_X} \\
\tilde{X} & \xrightarrow{g} & \tilde{X} \\
\downarrow{h} & & \downarrow{h} \\
\ker(q_X) & \xrightarrow{f'} & \ker(q_X)
\end{array}$$

where $f' : \ker(q_X) \to \ker(q_X)$ is the assignment $f'([x]) = [f(y)]$ for some $y \in [x]$. This construction is possible by the axiom of choice, and is well-defined since it is independent of the choice of $y \in [x]$. Define $g : \tilde{X} \to \tilde{X}$ to be $g = h^{-1} \circ f' \circ h$.

Remark 4. Consider the partially observable system $(X, f, q)$. The following two statements are equivalent.

(i) $\ker(q) \subseteq \ker(q \circ f)$

(ii) $\prod_{i=0}^n f^{-1}(\ker(q)) = \ker(q)$ for all $n \in \mathbb{N}$.

Proof. Note that $f^{-1}(\ker(q)) = f^{-1}\left((q^{-1}(y))_{y \in Y}\right) = ((q \circ f)^{-1}(y))_{y \in Y} = \ker(q \circ f)$ and therefore $\ker(q) \subseteq \ker(q \circ f)$ is equivalent to $\ker(q) \subseteq f^{-1}(\ker(q))$ which is equivalent to
\( \ker(q) \cap f^{-1} \ker(q) = \ker(q) \). Hence the claim holds for \( n = 1 \). We proceed by induction. Suppose the claim holds for some \( n \in \mathbb{N} \). Then \( \ker(q) \subseteq \bigcap_{i=1}^{n} f^{-i} \ker(q) \). We conclude
\[
\ker(q) \cap f^{- (n+1)} \ker(q) \subseteq \bigcap_{i=1}^{n+1} f^{-i} \ker(q)
\]
which proves the claim.

A morphism of systems requires that the target system is equivalent to the action of the source system on the fibers of the morphism.

**Remark 5.** Let \((X, f, q_X)\) and \((Y, g, q_Y)\) be isomorphic partially observable systems and suppose that \( (X, f) \xrightarrow{q_X} (\tilde{X}, \tilde{f}) \) is a morphism of dynamical systems. Then there exists a morphism \( (Y, g) \xrightarrow{q_Y} (\tilde{Y}, \tilde{g}) \) of dynamical systems and an isomorphism of dynamical systems \( (\tilde{X}, \tilde{f}) \simeq (\tilde{Y}, \tilde{g}) \).

**Proof.** Suppose \( (X, f, q_X) \xrightarrow{(m,n)} (Y, g, q_Y) \) is an isomorphism and consider the following diagram.

Consider the map \( \tilde{g} : \tilde{Y} \to \tilde{Y} \) given as \( \tilde{g} = n \circ \tilde{f} \circ n^{-1} \). By inspection of the diagram one verifies that \( q_Y \circ g = \tilde{g} \circ q_Y \) and \( n \circ \tilde{f} = \tilde{g} \circ n \).

**Example 1 (Quaternion rotations and the Hopf map).** We write \( S_{n-1} \subset \mathbb{R}^n \) for the sphere \( S_{n-1} = \{ x \in \mathbb{R}^n : \| x \|_2 = 1 \} \) and \( C_{n-1} \subset \mathbb{R}^n \) for the unit hypercube. We exhibit an isomorphism of partially observable systems, according to the following diagram.

We now define the maps in the above diagram. \( h_m : (S_m \subset \mathbb{R}^{m+1}) \to (C_m \subset \mathbb{R}^{m+1}) \) is given by \( h_m(x) = \| x \|_{m+1}^{-1} x \). The map \( h_m^{-1} : (C_m \subset \mathbb{R}^{m+1}) \to (S_m \subset \mathbb{R}^{m+1}) \) is given by \( h_m^{-1}(y) = \| y \|_{m+1}^{-1} y \). The quotient map \( q_S : (S_3 \subset \mathbb{R}^4) \to (S_2 \subset \mathbb{R}^3) \) is the Hopf map
\[
q_S(x) = (x_1^2 + x_2^2 - x_3^2 - x_4^2, 2(x_1x_4 + x_2x_3), 2(x_2x_4 - x_1x_3)) .
\]
We set \( q_C : (C_3 \subset \mathbb{R}^4) \rightarrow (C_2 \subset \mathbb{R}^2) \) to
\[
q_C(x) = h_2 \circ q_S \circ h_3^{-1}(x) = \frac{q_S \left( \|x\|_{2}^{-1}x \right)}{\|q_S \left( \|x\|_{2}^{-1}x \right)\|}. 
\]

It remains to specify the dynamics. There are several rotations that can play this role. A possible choice are the rotations given by the following linear maps.
\[
R_3 = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & -1 \\
0 & 0 & 1 & 0
\end{pmatrix} \quad R_2 = \begin{pmatrix}
1 & 0 & 0 \\
0 & 0 & -1 \\
0 & 1 & 0
\end{pmatrix}
\]

Note that \( R_i : S_i \rightarrow S_i \) and \( R_i : C_i \rightarrow C_i \) formally coincide as linear operators in the Euclidean ambient space. We have \((S_3, R_3, q_S) \simeq (C_3, R_3, q_C)\).

## 3 Entropy and quotient entropy

The complexity of a topological dynamical system is quantified by its entropy. Loosely speaking, the entropy of a system is the asymptotic rate of divergence of infinitesimally close orbits. A zero-entropy system exhibits such divergence at subexponential rate. A positive-entropy system does so at exponential rates, and a system of infinite entropy at superexponential rates. Sometimes systems of positive entropy are called chaotic systems. The notion of topological entropy was developed in the 60s, starting with Adler et al. [AKM65, Din69, Bow71]. The topological entropy of the system \((X, f)\) is
\[
h(X, f) = \sup \left\{ \lim_{n \uparrow \infty} \frac{1}{n} \ln \left( \square^n_f(\mathcal{U}) \right) : \mathcal{U} \text{ is an open cover of } X \right\},
\]
where \( \square^n_f(\mathcal{U}) \) denotes the minimal cardinality of a finite subcover of \( X \) extracted from the open cover \( \bigcap_{i=1}^n f^{-i} \mathcal{U} \). This defines a map \( h : \text{ob}(\text{Sys}) \rightarrow [0, \infty] \).

A well-known property of entropy is that it is functorial: If there is a diagram \((X, f) \rightarrow (Y, g)\), then \( h(X, f) \geq h(Y, g) \). The assignment \( \text{Ent} : \text{Sys} \rightarrow ([0, \infty], \geq) \) that assigns \( (X, f) \mapsto h(X, f) \) on objects is a functor. (We consider the ordered set \(([0, \infty], \geq)\) as a thin category in the usual way.)

Although the entropy is a number assigned to a system, we may use it to assign a number to morphisms: the difference in entropy between source and target, its complexity-reduction. (This approach, in the context of information-theoretic Shannon entropy, originates with [BFL11].) Unfortunately, the assignment \( \text{Ent} : \text{Arr}(\text{Sys}) \rightarrow ([0, \infty], \geq) \) given by
\[
\text{Ent} \left( (X, f) \quad \overset{m}{\rightarrow} \quad (Y, g) \right) = h(X, f) - h(Y, g),
\]
where \( \text{Arr}(\text{Sys}) \) denotes the Arrow category, is not functorial, as the following example shows.
Example 2. Pick a system \((X, f)\) with \(\text{Ent}(X, f) = r > 0\) and consider the following morphism in \(\text{Arr}(\text{Sys})\).

\[
\begin{array}{ccc}
(X, f) & \xrightarrow{id} & (X, f) \\
\downarrow{\text{id}} & & \downarrow{!} \\
(X, f) & \xrightarrow{!} & ([1], \text{id})
\end{array}
\]

The string of inequalities \(\text{Ent}(\text{id}) = 0 \geq \text{Ent}(!) = \text{Ent}(X, f) = r > 0\) cannot be fulfilled.

Yet \(\text{Ent}\) is functorial on a subcategory of \(\text{Arr}(\text{Sys})\), the category \(\text{Fac}\) of factorizations of dynamical systems. The objects are morphisms of \(\text{Sys}\). The morphisms are diagrams of the following form.

\[
\begin{array}{ccc}
(X, f) & \xrightarrow{m} & (Y, g) \\
\downarrow{n} & & \downarrow{r} \\
(Z, h) & & \end{array} = \begin{array}{ccc}
X & \xrightarrow{f} & X \\
\downarrow{m} & & \downarrow{m} \\
Y & \xrightarrow{g} & Y \\
\downarrow{r} & & \downarrow{r} \\
Z & \xrightarrow{h} & Z
\end{array}
\]

In the above case, we say that \(r\) is a morphism from \(m\) to \(n\). Composition of morphisms is concatenation of diagrams.

\[
\begin{array}{ccc}
(X, f) & \xrightarrow{m} & (Y, g) \\
\downarrow{n} & & \downarrow{r} \\
(Z, h) & \xrightarrow{o} & (A, i)
\end{array} \sim \begin{array}{ccc}
(X, f) & \xrightarrow{m} & (Y, g) \\
\downarrow{o} & & \downarrow{sor} \\
(A, i) & \xrightarrow{o} & \end{array}
\]

Note that objects of \(\text{Fac}\) (which are morphisms of \(\text{Sys}\)) do not admit a morphism between them whenever they start from different objects (as morphisms in \(\text{Sys}\)).

There is the functorial embedding \(\text{Fac} \hookrightarrow \text{PSys}\) where \(\left( (X, f) \xrightarrow{m} (Y, g) \right) \mapsto (X, f, m)\) on objects and

\[
\begin{array}{ccc}
(X, f) & \xrightarrow{m} & (Y, g) \\
\downarrow{n} & & \downarrow{r} \\
(Z, h) & & \end{array} \mapsto (\text{id, r}) \in \text{hom}_{\text{PSys}} ((X, f, m), (X, f, n))
\]

on morphisms. The functoriality of entropy on \(\text{Sys}\) directly implies the functoriality of \(\text{Ent}\) on \(\text{Fac}\).

We want to quantify the loss of complexity due to partial observability. The quantity that we use for this purpose is quotient-topological entropy. We define the quotient-topological entropy of the partially observable system \((X, f, q)\) as

\[
\tilde{h}(X, f, q) = \sup \left\{ \lim_{n \to \infty} \frac{1}{n} \ln \left( \square^q_n(U) \right) : U \text{ is open cover of } X \text{ in the } q\text{-induced topology} \right\},
\]
where the topology induced by $q$ is the topology that makes $X$ homeomorphic to $\ker(q)$. It may not be immediately clear that the above definition is sensical. To see that it is, it suffices to use well-known arguments [AKM65]. For any fixed open cover $\mathcal{U}$ of $X$ the quantity

$$
\lim_{n \to \infty} \frac{1}{n} \ln \left( \sqcap_f^n(\mathcal{U}) \right)
$$

exists and is finite [AKM65, Property 8]. Since the open covers of $X$ with the relation $\sqsubseteq$ are a downward-directed set the (quotient-)entropy exists as a nonnegative extended real number.

**Remark 6.** If $X$ is a metric space, the quotient-entropy may be computed by taking a limit along covers by quotient-metric balls of vanishing diameter. In fact, Bowen’s construction [Bow71] may be applied to the quotient metric.

We have the commutative diagram

$$
\begin{array}{ccc}
\text{Fac} & \xrightarrow{\text{Ent}} & \text{PSys} \\
\downarrow & & \downarrow \text{QEnt} \\
([0, \infty], \geq) & & \end{array}
$$

where $\text{QEnt} : \text{PSys} \to [0, \infty]$ assigns $\text{QEnt}(X, f, q) = h(X, f) - \tilde{h}(X, f, q)$. Note that $\text{QEnt}$ is *not* functorial on all of $\text{PSys}$, but it is functorial on the image of the above embedding.

**Example 3.** Let $(X, f)$ be a dynamical system such that $\text{Ent}(X, f) = r > 0$ and consider the following morphism in $\text{PSys}$.

$$
\begin{array}{ccc}
X & \xrightarrow{f} & X \\
\downarrow \text{id} & & \downarrow \text{id} \\
X & \xrightarrow{f} & X
\end{array}
$$

The string of inequalities $\text{QEnt}(X, f, \text{id}) = 0 \geq \text{QEnt}(X, f, !) = r > 0$ cannot be fulfilled.

We interpret $\text{Ent}$ as measuring the complexity-reduction of a morphism. In this light, $\text{QEnt}$ extends $\text{Ent}$ to partially observable systems: it quantifies the complexity reduction due to partial observability. Quotient-topological entropy shares the functoriality of entropy.

**Proposition 2.** The assignment $\text{QEnt} : \text{PSys} \to ([0, \infty], \geq)$ where $\text{QEnt}(X, f, q) = \tilde{h}(X, f, q)$ on objects is a functor.

**Proof.** Let $(X, f, q_X) \to (Y, g, q_Y)$. We have the following commutative diagram.

$$
\begin{array}{ccc}
X & \xrightarrow{f} & X \\
\downarrow m & & \downarrow m \\
Y & \xrightarrow{g} & Y
\end{array}
$$
Open covers of $X$ in the topology induced by $q_X$ are equivalently open covers of $\tilde{X}$. Open covers of $Y$ in the topology induced by $q_Y$ are equivalently open covers of $\tilde{Y}$. Let $\mathcal{U}$ be such an open cover of $\tilde{X}$. The assignment $\mathcal{U} \mapsto \{n(U)\}_{U \in \mathcal{U}}$ yields an open cover of $\tilde{Y}$. Furthermore every open cover of $\tilde{Y}$ arises in this way. This assignment respects the fibers of $m$ and does not increase the number of elements in the cover. By Proposition 1, $g$ respects the fibers of $m$, therefore

$$f^{-i}(\mathcal{U}) \subseteq g^{-i}(n(\mathcal{U}))$$

$$\bigcap_{i=1}^{n} f^{-i}(\mathcal{U}) \subseteq \bigcap_{i=1}^{n} g^{-i}(n(\mathcal{U}))$$

$$\square^p_f(\mathcal{U}) \geq \square^p_g(n(\mathcal{U}))$$

$$\lim_{n \to \infty} \frac{1}{n} \ln \left( \square^p_f(\mathcal{U}) \right) \geq \lim_{n \to \infty} \frac{1}{n} \ln \left( \square^p_g(n(\mathcal{U})) \right)$$

$$\sup_{\mathcal{U}} \lim_{n \to \infty} \frac{1}{n} \ln \left( \square^p_f(\mathcal{U}) \right) \geq \sup_{\mathcal{U}} \lim_{n \to \infty} \frac{1}{n} \ln \left( \square^p_g(n(\mathcal{U})) \right)$$

where $\mathcal{U}$ is an open cover of $X$ in the topology induced by $q_X$. □

The following statements show that the observable complexity is at most equal to the complexity of the base system (Corollary 8), and that, whenever two quotient maps are ordered with respect to their resolution, the respective quotient entropies behave monotonically (Corollary 9).

**Lemma 7** (Property 10 in [AKM65]). Let $(X, f)$ be a dynamical system and let $\mathcal{A}, \mathcal{B}$ be open covers of $X$. Then

$$\lim_{n \to \infty} \frac{1}{n} \ln \left( \square^p_f(\mathcal{A}) \right) \leq \lim_{n \to \infty} \frac{1}{n} \ln \left( \square^p_g(\mathcal{B}) \right).$$

**Corollary 8.** Let $(X, f, q)$ be a partially observable system. Then $\hat{h}(X, f, q_X) \leq h(X, f)$.

**Corollary 9** (Quotient entropy is antitone). Consider the system $(X, f)$. Let $q_0 : X \to Y_0$ and $q_1 : X \to Y_1$ be quotient maps such that $\ker(q_0) \subseteq \ker(q_1)$. Then $\hat{h}(X, f, q_0) \geq \hat{h}(X, f, q_1)$.

If the quotient map is a morphism, the quotient-entropy of the partially observable system equals the entropy of the target-system.

**Proposition 3.** Consider $(X, f, q_X)$ where $(X, f) \xrightarrow{q_X} (\tilde{X}, g)$. Then

$$h(\tilde{X}, g) = \hat{h}(X, f, q_X).$$

**Proof.** This follows from the observation that the $\mathbb{N}$-action by $g$ on $\tilde{X}$ is isomorphic to the $\mathbb{N}$-action by $f$ on fibers of $q_X$. □
Remark 10. In the light of Proposition 3 and Corollary 9, one may interpret \( \overline{QEnt} \) as a closure indicator. If the quotient-map is a morphism, we have \( \overline{QEnt} = 0 \) since the kernel of the quotient map is preserved (Remark 4). It is not a precise closure indicator: If the kernel is scrambled with subexponential speed, this is not witnessed by entropy.

The basic properties of entropy, additivity on products (the Bernoulli property) and multiplicativity under iteration, are shared by quotient-entropy.

Proposition 4. The following two statements hold.

(i) Let \( m \in \mathbb{N} \). The \( m \)-step partially observable system \((X, f^m, q_X)\) fulfills
\[
\tilde{h}(X, f^m, q_X) = m \cdot \tilde{h}(X, f, q_X).
\]

(ii) Let \((X, f, q_X)\) and \((Y, g, q_Y)\) be partially observable systems.
Then \( \tilde{h}(X \times Y, f \times g, q_X \times q_Y) = \tilde{h}(X, f, q_X) + \tilde{h}(Y, g, q_Y) \).

Proof. (i): See the proof of Theorem 2 in [AKM65]. (ii): See the proof of Theorem 3 in [AKM65]. \( \square \)

4 Symbolic dynamics and the category of graphs

In the theory of dynamical systems, topological Markov chains provide a link between continuous-topological approaches and discrete-algebraic ones. We recall their construction. We denote by \([m] = \{1, \ldots, m\}\) with the discrete topology. We equip the sequence space \([m]^{\mathbb{N}}\) with the product topology, which is metrized by the family of product metrics
\[
\text{dist}_\lambda(s, s') = \sum_{i \in \mathbb{N}} \frac{1}{\lambda^i} 1_{\{s_i \neq s'_i\}}
\]
for some \( \lambda \in (1, \infty) \). A matrix \( A \in \text{Mat}_{n \times n}(\{0, 1\}) \), which may be interpreted as a graph with vertex set \([n]\), generates the word space
\[
W_A = \{ w \in [n]^{\mathbb{N}} : A_{w_t, w_{t+1}} = 1 \text{ for all } t \in \mathbb{N} \},
\]
which we equip with the subspace topology of the sequence space \([n]^{\mathbb{N}}\). We assume, without loss of generality, that the generating matrix \( A \) is nontrivial: for all \( i \in [n] \) there exists \( j \in [n] \) such that \( A_{ij} = 1 \). The shift on the sequence space \([n]^{\mathbb{N}}\) is the map \( \sigma : [n]^{\mathbb{N}} \to [n]^{\mathbb{N}} \) given as \( (\sigma(w))_t = w_{t+1} \). A topological Markov chain is the system obtained by iterating the shift map on an invariant word space, a system \((W_A, \sigma_A)\) where \( A \in \text{Mat}_{n \times n}(\{0, 1\}) \) and \( \sigma(W_A) \subseteq W_A \). Although the shift is defined on the entire sequence space, we index it on the generating matrix to keep the invariant word space in mind.

We recall the spectral formula for topological Markov chains, probably first explicitly stated by Parry [Par64]: \((W_A, \sigma_A)\) has topological entropy \( \ln(\lambda_A^\ast) \) where \( \lambda_A^\ast \) denotes the largest eigenvalue of \( A \).
We are interested in quotient maps of state spaces and morphisms of systems. The related isomorphism problem for topological Markov chains has been solved by Williams, who gave an algebraic condition on matrices encoding edge-shifts [Wil73]. We treat topological Markov chains as vertex-shifts.

**Definition 2** (Letter-by-letter morphism). Let \((W_A, \sigma_A)\) and \((W_B, \sigma_B)\) be topological Markov chains and assume \(W_A \subseteq [n]^N\), \(W_B \subseteq [m]^N\), where \(n \geq m\). We call a morphism \(\pi : (W_A, \sigma_A) \to (W_B, \sigma_B)\) letter-by-letter if there exists \(c : [n] \to [m]\) such that \((\pi(w))_t = c(w_t)\) for all \(t \in \mathbb{N}\).

**Remark 11.** One may wonder whether defining letter-by-letter morphisms using a family \(c_t : [n] \to [m]\) for \(t \in \mathbb{N}\) leads to a more general theory. This is not so. The theorem of Curtis-Lyndon-Hedlund [Hed69], states that any equivariant continuous map between topological Markov chains is a cellular automaton, hence locally generated by a finite-to-one map between the respective alphabets. This is a consequence of choosing the product topology on sequence space. (Hence any morphism between topological Markov chains can be considered to be letter-by-letter in an appropriate block representation.) For us the following statement will suffice:

Let \(\pi : [n]^N \to [m]^N\) be a letter-by-letter map that is generated by a family \(c_t : [n] \to [m]\) for \(t \in \mathbb{N}\). Suppose that \(\pi\) is equivariant with respect to the shift. Then there exists \(c : [n] \to [m]\) such that \(c_t = c\) for all \(t \in \mathbb{N}\).

**Proof.** By hypothesis \(\pi\) is such that \(\pi \circ \sigma = \sigma \circ \pi\). We have

\[
(\pi \circ \sigma(w))_t = \pi((\sigma(w))_t) = c_t((\sigma(w))_t) = c_t(w_{t+1})
\]

and

\[
(\sigma \circ \pi(w))_t = \pi((\pi(w))_t) = (\pi(w))_{t+1} = c_{t+1}(w_{t+1}),
\]

whereby we conclude that \(c_t(w_{t+1}) = c_{t+1}(w_{t+1})\) for all \(t \in \mathbb{N}\) and any \(w \in [n]^N\). For any \(x \in [n]\) we can pick \(w \in [n]^N\) such that \(w_2 = x\). Hence \(c_1(x) = c_2(x)\) for all \(x \in [n]\). Suppose that \(c_t = c_1\). For all \(v \in [n]^N\) such that \(v_{i+1} = x\) the same argument implies \(c_{i+1}(x) = c_1(x)\). Whenever a test-sequence for \(y \in [n]\) does not exist, setting \(c_{i+1}(y) = c_1(y)\) does no harm. The claim follows by induction. \(\Box\)

**Lemma 12.** Let \(c : [n] \to [m]\), where \(n \geq m\), be surjective. Then the induced map \(C : [n]^N \to [m]^N\) where \((C(w))_t = c(w_t)\) is continuous and surjective.

**Proof.** Fix \(\lambda \in (1, \infty)\). The metrics of the form \(\text{dist}_\lambda\) metrize the topologies of \([n]^N\) and \([m]^N\). Note that \(1_{\{v \neq w\}} \geq 1_{\{c(v) \neq c(w)\}}\) for any \(v, w \in [n]^N\) by the surjectivity of \(c\), and hence

\[
\text{dist}_\lambda(v, w) = \sum_{i \in \mathbb{N}} \frac{1}{\lambda^i} 1_{\{v_i \neq w_i\}} \geq \sum_{i \in \mathbb{N}} \frac{1}{\lambda^i} 1_{\{c(v_i) \neq c(w_i)\}} = \text{dist}_\lambda(C(v), C(w)).
\]

Under this choice of metric \(C\) is 1-Lipschitz, hence continuous. The surjectivity of \(C\) follows from the surjectivity of \(c\). \(\Box\)
The sequence space is a Cantor space: it is compact, metrizable, totally disconnected, and has no isolated point. Hence the following lemma.

**Lemma 13** (See Paragraph 3d in [KH97]). The sequence space \([m]^\mathbb{N}\) with the metric

\[
\text{dist}_{10m}(w, w') = \sum_{i=1}^{\infty} \frac{1}{(10m)^i} I\{w_i \neq w'_i\}
\]

is such that for any \(x \in [m]^\mathbb{N}\) the set \(O^l_x := \{w \in [m]^\mathbb{N} : w_i = x_i, i \leq l\}\) is the open ball of radius \((10m)^{-l}\) around each of its points. There are exactly \(m^{l+1}\) different balls of this form. A cover of \([m]^\mathbb{N}\) by such balls is minimal. The minimal cardinality of a cover of \(S \subseteq [m]^\mathbb{N}\) by such balls equals the number of sets of the form \(C^l_x\) having nonempty intersection with \(S\).

The above lemma allows us to estimate quotient-entropy by a simple procedure.

**Proposition 5.** Let \((W_A, \sigma_A, C)\) be a partially observable topological Markov chain where \(C : W_A \to [m]^\mathbb{N}\). Suppose that \(B_0, B_1 \in \text{Mat}_{m \times m}(\{0, 1\})\) and \(W_{B_0} \subseteq C(W_A) \subseteq W_{B_1}\). Then

\[
\ln(\lambda_{B_0}^\text{max}) \leq \hat{h}(W_A, \sigma_A, C) \leq \ln(\lambda_{B_1}^\text{max})
\]

**Proof.** A minimal \(\epsilon\)-quotient-cover of \(W_A \subseteq [n]^\mathbb{N}\) is the pullback of a minimal \(\epsilon\)-cover of \(C(W_A) \subseteq [m]^\mathbb{N}\). We may assume, without loss of generality, that minimal covers of \(W_{B_0}, C(W_A), W_{B_1} \subseteq [m]^\mathbb{N}\) are inherited from a minimal cover of \([m]^\mathbb{N}\) as in Lemma 13. Since \(W_{B_0} \subseteq C(W_A) \subseteq W_{B_1}\), the cardinalities of these minimal \(\epsilon\)-covers fulfill the same inequality. We have shown that \(h(B_0) \leq \hat{h}(A) \leq h(B_1)\), it remains to apply the spectral formula. \(\square\)

Let \((W_A \subseteq [n]^\mathbb{N}, \sigma_A)\) be a topological Markov shift. Consider \(c : [n] \to [m], n \geq m\), and the induced map \(C : [n]^\mathbb{N} \to [m]^\mathbb{N}\). The following diagram always commutes.

\[
\begin{array}{ccc}
W_A & \xrightarrow{\sigma_A} & W_A \\
\downarrow{C} & & \downarrow{C} \\
C(W_A) & \xrightarrow{\sigma} & C(W_A)
\end{array}
\]

This can be seen from

\[
\sigma \circ C(w) = C \circ \sigma_A(w)
\]

\[
\sigma(\{c(w_i)\}) = C(\{w_{i+1}\})
\]

\[
\{c(w_{i+1})\} = \{c(w_{i+1})\}.
\]

The interesting question is whether the subspace \(C(W_A) \subseteq [m]^\mathbb{N}\) is the word space of some matrix \(B \in \text{Mat}_{m \times m}(\{0, 1\})\).

The following proposition characterizes shift-equivariance of letter-by-letter maps between sequence spaces. Note that this is not a characterization of morphisms but of embeddings of the respective \(\mathbb{N}\)-actions.
Proposition 6. Let \( A \in \text{Mat}_{n \times n}(\{0,1\}) \) and \( B \in \text{Mat}_{m \times m}(\{0,1\}) \) where \( n \geq m \). Let \( e : [n] \to [m] \) be a surjection inducing \( E : [n]^N \to [m]^N \). Then the following two statements are equivalent.

(i) \( \sigma_B \circ E = E \circ \sigma_A \)

(ii) \( A_{ij} = 1 \) implies \( B_{e(i)e(j)} = 1 \).

Proof. (ii) \( \implies \) (i): We have \( A_{w_tw_{t+1}} = 1 \implies B_{e(w_t)e(w_{t+1})} = 1 \) for all \( w \in W_A \) and \( t \in \mathbb{N} \). Hence

\[(\sigma_B \circ E(w))_t = \sigma_B( e(w_t) ) = e(w_{t+1}) = (E \circ \sigma_A(w))_t.\]

(i) \( \implies \) (ii): Suppose that (ii) does not hold. Hence there exists a word \( v \in W_A \) such that \( A_{v_iv_{i+1}} = 1 \) and \( B_{e(v_i)e(v_{i+1})} = 0 \) for some \( i \in \mathbb{N} \). This implies

\[(\sigma_B \circ E(v))_i = \sigma_B( e(v_i) ) \neq e(v_{i+1}) = (E \circ \sigma_A(v))_i.\]

A topological Markov chain is fully specified by its underlying graph, which is fully specified by the respective adjacency matrix. (This as an isomorphism of categories.)

Definition 3 (The category \( \text{Graph} \)). The objects are matrices \( A \in \text{Mat}_{n \times n}(\{0,1\}) \) for \( n \in \mathbb{N} \). A morphism \( A \xrightarrow{f} B \) where \( A = \{A_{ij}\}_{i,j \in [n]} \) and \( B = \{B_{ij}\}_{i,j \in [m]} \) is a surjection \( f : [n] \to [m] \) such that \( A_{ij} = 1 \) implies \( B_{f(i)f(j)} = 1 \).

The composition of morphisms is the composition of the underlying maps. The associativity- and identity-properties are inherited from maps between finite sets. Via topological Markov chains, the category \( \text{Graph} \) is functorially embedded into the category \( \text{Sys} \).

Proposition 7. We have a functorial embedding \( S : \text{Graph} \hookrightarrow \text{Sys} \) given by \( A \mapsto (W_A, \sigma_A) \) on objects and \( (A \xrightarrow{f} B) \mapsto \left( (W_A, \sigma_A) \xrightarrow{F} (W_B, \sigma_B) \right) \) on morphisms, where \( F : W_A \to W_B \) is given as \( (F(w))_i = f(w_i) \).

Sketch of proof. Clearly \( S \) is faithful. Suppose that \( A \xrightarrow{f} B \) in \( \text{Graph} \). We want to show that \( SA \xrightarrow{Sf} SB \), hence \( (W_A, \sigma_A) \xrightarrow{F} (W_B, \sigma_B) \) in \( \text{Sys} \), which is equivalent to the commutativity of the following diagram.

\[
\begin{array}{ccc}
W_A & \xrightarrow{\sigma_A} & W_A \\
\downarrow F & & \downarrow F \\
W_B & \xrightarrow{\sigma_B} & W_B
\end{array}
\]

Suppose that \( A \) is \( n \times n \) and that \( B \) is \( m \times m \). Let \( w \in W_A \). Hence \( A_{w_tw_{t+1}} = 1 \) for all \( t \in \mathbb{N} \). Since \( f : [n] \to [m] \) is a morphism, \( A_{w_tw_{t+1}} = 1 \) implies \( B_{f(w_t)f(w_{t+1})} = 1 \) and therefore \( Sf(w) \in W_B \) by Proposition 6. \( \Box \)
Remark 14. The categorical products in \textbf{Graph} correspond to the Kronecker-products of adjacency matrices. The embedding of graphs into dynamical systems respects finite products. We have \((W_{A \otimes B}, \sigma_{A \otimes B}) \simeq (W_A, \sigma_A) \otimes (W_B, \sigma_B)\).

Whenever a graph may be obtained from another by adding edges there is a morphism. In particular, for any \(A \in \text{Mat}_{n \times n}(\{0, 1\})\) we have \(A \to (1)_{n \times n}\). The existence of a graph-morphism implies embeddability of the actions (Proposition 6), not the existence of a morphism. A necessary condition for the existence of a morphism can be obtained as follows. Let \(A \xrightarrow{f} B\) in \textbf{Graph}. Suppose that \(g : B \to A\) is right-inverse to \(f\), hence \(f \circ g = \text{id}\). Then \(Sg : SB \to SA\) is right-inverse to \(Sf : SA \to SB\). In particular, \(Sf\) is surjective. The existence of a section from the generating graph is a sufficient condition for topological Markov chains to be related by a morphism. The following proposition is immediate.

**Proposition 8.** Let \(A \in \text{Mat}_{n \times n}(\{0, 1\})\) and \(B \in \text{Mat}_{m \times m}(\{0, 1\})\). Consider a surjection \(c : [n] \to [m]\). The induced map \(C : W_A \to W_B\) is a morphism of dynamical systems if \(c\) is a graph-morphism and admits a right-inverse graph-morphism.

In general, it is nontrivial to verify whether a graph is the quotient of another. In the following example the quotient-property is obvious since any inverse of the set-map which induces the graph morphism does induce an inverse graph-morphism.

**Example 4.**

\[
\begin{pmatrix}
0 & 1 & 1 \\
1 & 1 & 1 \\
1 & 1 & 1
\end{pmatrix}
\sim
\begin{pmatrix}
0 & 1 \\
1 & 1
\end{pmatrix}
\]

Note that the condition of Proposition 8 is not necessary. In fact, we can give an independent second sufficient condition, which is itself not necessary (as can be seen from Example 8).

**Proposition 9.** Suppose that \(c : [n] \to [m]\) induces a graph-morphism between \(A \in \text{Mat}_{n \times n}(\{0, 1\})\) and \(B \in \text{Mat}_{m \times m}(\{0, 1\})\). If all subgraphs of \(A\) induced by \(c^{-1}(i)\) for \(i \in [m]\) are either totally disconnected or have no sinks, then \(Sc : W_A \to W_B\) is a letter-by-letter morphism of topological Markov chains.

**Proof.** It suffices to show that \(Sc(W_A) = W_B\), hence that \(Sc\) maps words in \(W_A\) to words in \(W_B\) and that every word in \(W_B\) is the image of a word in \(W_A\) under \(Sc\).

Fix \(w \in W_A\). Define \(U_k := \{x \in [m]^N : x_i = c(w_i), i \leq k\}\). Note that \(U_1 \cap W_B \neq \emptyset\).

Suppose that \(U_k \cap W_B \neq \emptyset\). We claim that \(U_{k+1} \cap W_B \neq \emptyset\). Consider the pair \((w_k, w_{k+1})\).

We have \(A_{w_k, w_{k+1}} = 1\). Since \(c\) is a graph-morphism this implies \(B_{c(w_k), c(w_{k+1})} = 1\).

Therefore \(U_{k+1} \cap W_B \neq \emptyset\) and, by induction, \(U_n \cap W_B \neq \emptyset\) for all \(n \in \mathbb{N}\). We conclude \(Sc(w) \in W_B\).
Fix \( v \in W_B \). Define \( V_k := \{ x \in [n]^N : x_i \in c^{-1}(v_i), i \leq k \} \). Note that \( V_1 \cap W_A \neq \emptyset \). Suppose that \( V_k \cap W_A \neq \emptyset \) for some \( k \in \mathbb{N} \). Consider the pair \((v_k, v_{k+1})\), hence \( B_{v_k,v_{k+1}} = 1 \). By construction there exists \( A_{ij} = 1 \) where \( i \in c^{-1}(v_k) \) and \( j \in c^{-1}(v_{k+1}) \) and \( \{ x \in V_k : x_k = i \} \neq \emptyset \) by hypothesis. Therefore \( V_{k+1} \cap W_A \neq \emptyset \) and, by induction, \( V_n \cap W_A \neq \emptyset \) for all \( n \in \mathbb{N} \). We conclude \( Sc^{-1}(v) \cap W_A \neq \emptyset \). \( \square \)

## 5 Quotient entropy and Markov partitions

There are several equivalent ways to define Markov partitions. We follow Adler [Adl98] and Gromov [Gro16].

**Definition 4** (Markov partition). A Markov partition for the system \( (X,f) \) is a finite collection \( \{U_i\}_{i=1}^n, n \geq 2 \), of open sets \( U_i \subset X \) such that the following three conditions hold.

1. \( U_i \cap U_j = \emptyset \) whenever \( i \neq j \).
2. \( \bigcup_{i=1}^n U_i \) is dense in \( X \).
3. For all \( n, m \in \mathbb{N} \), if simultaneously \( f^n(U_i) \cap U_j \neq \emptyset \) and \( f^m(U_j) \cap U_k \neq \emptyset \) hold, then \( f^{n+m}(U_i) \cap U_k \neq \emptyset \).

Markov partitions are not quite partitions, Adler [Adl98] refers to them as topological partitions. The existence of a Markov partition for a dynamical system is a demanding property. Bowen [Bow70] showed that a Markov partition exists for any system satisfying Axiom A. Explicit constructions go back to Sinai [Sin68]. The usefulness of Markov partitions originates in the following construction.

**Definition 5** (Hadamard construction [Had98]). The Hadamard construction assigns to a dynamical system \( (X,f) \) with Markov partition \( \{U_i\}_{i=1}^n \) the graph \( A \in \text{Mat}_{n \times n}(\{0, 1\}) \) where

\[
A_{ij} = \begin{cases} 1 & \text{if } f(U_i) \cap U_j \neq \emptyset \\ 0 & \text{otherwise.} \end{cases}
\]

The Hadamard construction and the embedding \( S : \text{Graph} \hookrightarrow \text{Sys} \) have some properties of a retraction. Let \( (X,f) \) be a dynamical system with Markov partition \( \{U_i\}_{i=1}^n \). Denote by \( \phi_i(x) \) the index of the partition element that contains the image of \( x \in X \) under \( f^i \) up to closure, hence \( f^i(x) \in \overline{U}_{\phi_i(x)} \). The assignment \( \Phi : X \to [n]^N \) is defined as \( \Phi(x) = \{ \phi_i(x) \}_{i \in \mathbb{N}} \). This assignment may be one-to-many on some points: We suppose that some choice is made. The pseudo-inverse assignment \( \Phi^{-1} : [m]^N \to 2^X \) is obtained in the following manner. For any \( w \in W_A \), define \( R_0(w) = \{ x \in X_{w_0} \} \), and, iteratively for \( t \in \mathbb{N} \),

\[
R_t(w) = \{ x \in R_{t-1}(w) : f(x) \in X_{w_t} \}.
\]
Then set \( \Phi^{-1}(w) \coloneqq \bigcap_{t=0}^{\infty} R_t(w) \subset X \). Often, for example if \( f \) is a minimal homeomorphism, we have \( |\Phi^{-1}(w)| = 1 \).

We are interested in systems that are \textit{finitely presented}, in the sense that they admit a Markov partition such that the Hadamard construction followed by the embedding \( S \) yields a homeomorphism. In that case the Hadamard construction yields an isomorphism between the original system and a topological Markov chain.

These structural relationships, as well as others appearing in this work, are illustrated in the following diagram.

\[
\begin{array}{c}
\text{PSys} \\
\text{Sys} \\
\text{Graph}
\end{array} \xrightarrow{E} \begin{array}{c}
\text{QEnt} \\
\text{Ent} \\
\ln \lambda_{\text{max}}
\end{array} \xrightarrow{[0, \infty], \geq} \begin{array}{c}
E_n : [0,1]_{[0,1]} \rightarrow [0,1]_{[0,1]} \\
\text{where } E_n(x) = nx \mod 1
\end{array}
\]

\textbf{Example 5 (Circle rotations).} Consider \([0,1]_{[0,1]}\), the unit interval with identified endpoints. The family of maps \( E_n : [0,1]_{[0,1]} \rightarrow [0,1]_{[0,1]} \) where \( E_n(x) = nx \mod 1 \) is finitely presented via the Markov partitions

\[
\left\{ \left( \frac{i-1}{n}, \frac{i}{n} \right) \right\}_{i=1}^{n}.
\]

These partitions correspond to \((1)_{n \times n}\), the complete graph with \( n \) vertices. (See, for example, [BS02, Paragraph 1.3].) Consider the map \( E_4 \) with the Markov partition into quadrants. We consider the quotient by vertical projection \( x \mapsto \cos(2\pi x) \) and the map \( \{1, 4\} \mapsto 1, \{2, 3\} \mapsto 2 \).

The Markov graph is the complete graph on four vertices and hence \( \text{Ent}(E_4) = \ln(4) \). Clearly no projection of the circle onto the interval is a morphism of a rotation. We are tempted to say that the quotient map reduces the symbolic dynamics to the full shift on two symbols and that this should imply \( \hat{h}(E_4) = \ln(2) \) and \( \text{QEnt}(E_4) = \ln(4) - \ln(2) = \ln(2) \).

We proceed by formalizing the above example and by proving its correctness.
Definition 6 (Compatible topological partition). Let \((X, f, q)\) be a partially observable system with Markov partition \(\mathcal{X} = \{U_i\}_{i=1}^n\) for \((X, f)\). Consider a surjection \(s : [n] \to [m]\) such that \(Y := \{q(U_{s(i)})\}_{s(i) \in [m]} = \{V_j\}_{j \in [m]}\) fulfills \(\mathcal{X} \subseteq q^{-1}Y\). Then \(Y\) is a compatible topological partition.

Corollary 15 (Corollary of the definition). We have a graph-morphism \(s : A \to B\) where \(A\) is the Markov graph of \((X, f)\) given by

\[
A_{ij} = \begin{cases} 
1 & \text{if } f(U_i) \cap U_j \neq \emptyset \\
0 & \text{otherwise}, 
\end{cases}
\]

and \(B\) is defined by

\[
B_{ij} = \begin{cases} 
1 & \text{if } f(q^{-1}V_i) \cap q^{-1}V_j \neq \emptyset \\
0 & \text{otherwise}.
\end{cases}
\]

Proof. Suppose \(A_{ij} = 1\), which is equivalent to \(f(U_i) \cap U_j \neq \emptyset\). We have \(U_i \subseteq q^{-1}V_{s(i)}\), which is equivalent to \(f(U_i) \subseteq f(q^{-1}V_{s(i)})\), and \(U_j \subseteq q^{-1}V_{s(j)}\). We conclude \(f(q^{-1}V_{s(i)}) \cap q^{-1}V_{s(j)} \neq \emptyset\), which is equivalent to \(B_{s(i)s(j)} = 1\). Hence \(s\) is a graph-morphism.

Proposition 10. Let \((X, f)\) be a finitely presented dynamical system with Markov partition \(\mathcal{X} = \{U_i\}_{i=1}^n\). Consider the partially observable system \((X, f, q)\) with quotient map \(q : X \to Y\). Suppose that the compatible topological partition \(Y = \{V_i\}_{i=1}^m\) of \(Y\) induces the graph-morphism \(s : A \to B\). Suppose that \(s\) admits a right-inverse graph-morphism. Then \((W_B, \sigma_B)\) is a dynamical system such that the diagram

\[
\begin{array}{ccc}
X & \xrightarrow{f} & X \\
\downarrow{\Phi} & & \downarrow{\Phi} \\
W_A & \xrightarrow{\sigma_A} & W_A \\
\downarrow{S} & & \downarrow{S} \\
W_B & \xrightarrow{\sigma_B} & W_B \\
& & \downarrow{h}
\end{array}
\]

commutes where \(h : Y \to W_B\) is a homeomorphism and \(S : [n]^N \to [m]^N\) is pointwise application of \(s\). In particular, this implies \(\hat{h}(X, f, q) = \ln (\lambda_B)^{\max}\).

Proof. By Proposition 8, the map \(S\) is a letter-by-letter morphism of topological Markov chains since \(s\) admits a right-inverse graph-morphism. Since \(S : W_A \to W_B\) is a morphism, the respective square in the above diagram commutes.

We now construct the map \(h\). By hypothesis, a trajectory through the cells of \(\mathcal{X}\) uniquely identifies a point in \(X\). A trajectory through \(\mathcal{Y}\) is equivalent to a trajectory through its pullback onto \(X\). Such a trajectory only identifies subsets of \(X\) and not points. We extend \(\Phi^{-1} : 2^{W_A} \to 2^X\) by direct images. The restriction \(\Phi^{-1}|_{\ker(S)} : \ker(S) \to \ker(q)\) is a continuous bijection of compact Hausdorff spaces and therefore a homeomorphism.
By Proposition 2 we know that \( \tilde{h}(X, f, q) = \tilde{h}(W_A, \sigma_A, S) \). But we have also shown that \((W_A, \sigma_A) \rightarrow (W_B, \sigma_B)\). By Proposition 3 we have \( \tilde{h}(W_A, \sigma_A, S) = h(W_B, \sigma_B) \). It remains to apply the spectral formula.

Note that in Example 5 the hypotheses of Proposition 10 are fulfilled: The existence of a right-inverse graph morphism is obvious since the Markov graph of the circle rotation is complete.

**Example 6.** Consider the circle-rotation \( E_3 \) and the corresponding Markov partition.

![Circle Rotation Diagram]

The dynamic of the compatible Markov partition of the interval is the full shift on two symbols. Hence \( \tilde{h}(E_3) = \ln(2) \) and \( Q\text{Ent}(E_3) = \ln(3) - \ln(2) = \ln(3/2) < Q\text{Ent}(E_4) \). (Combining with Example 5, we have obtained a case where \( \tilde{h}(X, f, q) = \tilde{h}(X, g, q) \) for \( f \neq g \).)

**Example 7.** Consider the circle rotation \( E_3 \) and the horizontal projection \( x \mapsto \sin(2\pi x) \). The images of the cells of the Markov partition overlap in the unit interval in a way that makes a compatible selection impossible.
We close by discussing a class of examples of our scheme for calculating quotient-entropy. Every standard map of the interval is finitely presented [KH97, Paragraph 15.1]. Under an appropriate restriction of admissible quotient maps, there always is a quotient chain.

**Definition 7** (Standard map). A continuous map \( f : [0, 1] \to [0, 1] \) is a standard map if the following three conditions hold.

(i) There exists \( n \in \mathbb{N} \), \( n \geq 2 \), such that the set \( \{0, 1/n, 2/n, \ldots, (n-1)/n, 1\} \) is \( f \)-invariant.

(ii) The restriction \( f|_{((i-1)/n, i/n]} \) is nonconstant and affine for every \( i \in [n] \).

(iii) There are no sinks in the Markov graph \( F \in \text{Mat}_{n \times n} (\{0,1\}) \) given by

\[
F_{ij} = \begin{cases} 
1 & \text{if } f \left( \left( \frac{i-1}{n}, \frac{i}{n} \right) \right) \cap \left( \frac{i-1}{n}, \frac{i}{n} \right) \neq \emptyset \\
0 & \text{otherwise}
\end{cases}
\]

A Markov partition for an iterated standard-map is given by \( \left\{ \left( \frac{i-1}{n}, \frac{i}{n} \right) \right\}_{i=1}^{n} \). If we only consider certain quotient maps of the interval, we obtain a class of partially observable systems whose properties are straightforward.

**Definition 8** (Standard quotient). Let \( ([0, 1], f) \) be a dynamical system generated by iterations of a map with standard structure \( \{0, 1/n, 2/n, \ldots, (n-1)/n, 1\} \). The quotient map \( q : [0, 1] \to [0, 1] \) is standard if the following three conditions hold.

(i) \( q \left( \{0, 1/n, 2/n, \ldots, (n-1)/n, 1\} \right) \subseteq \{0, 1/n, 2/n, \ldots, (n-1)/n, 1\} \)

(ii) The restriction \( q|_{((i-1)/n, i/n]} \) is affine for every \( i \in [n] \).

(iii) \( q \) is monotone.

A nontrivial standard quotient-map must be constant on at least one interval. Monotonicity and surjectivity imply that \( q(0) = 0 \) and \( q(1) = 1 \).

**Proposition 11.** Consider \( ([0, 1], f, q) \) where \( f : [0, 1] \to [0, 1] \) is a standard map and \( q : [0, 1] \to [0, 1] \) is a standard quotient. Then there is a compatible selection corresponding to a right-invertible graph-morphism.

**Proof.** Denote by \( \mathcal{X} = \{U_i\}_{i=1}^{n} \) the Markov partition of \([0, 1]\) corresponding to \( f \). Consider the image of the partition under the quotient map \( q(\mathcal{X}) := \{q(U_i)\}_{i=1}^{n} \). Extract the subset \( \mathcal{Y} := \{V \in q(\mathcal{X}) : V \neq \emptyset\} \). Clearly \( |\mathcal{Y}| \leq |\mathcal{X}| \). Note that \( \mathcal{Y} \) consists of mutually disjoint open sets whose union is dense in \([0, 1]\). We have \( \mathcal{X} \subseteq q^{-1}\mathcal{Y} \). Since \( q \) is monotonous, we may enumerate the subsets \( C_j = q^{-1}(V_j) \) increasingly on \( j \in [m] \) where \( m \leq n \). Define \( s : [n] \to [m] \) by setting \( s(i) = \min_{j \in [m]} \{ j : V_i \subseteq C_j \} \). For all \( C_j \) there exists a unique \( V_i \) such that \( V_i \subseteq (C_j \setminus \cup_{k \neq j} C_k) \). Let \( g : [m] \to [n] \) be this assignment. We have obtained \( \mathcal{Y} = \{V_j\}_{j=1}^{m} \) and \( s : [n] \to [m] \) such that Proposition 10 applies. \( \square \)
Example 8. Consider the following standard map $f$ and standard quotient $q$.

\[
\begin{pmatrix}
1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 0 & 0
\end{pmatrix}
\]

We have $h([0, 1], f) = \ln(\lambda_{\text{max}}^3) = \ln 3$. The map given by $\{1, 2\} \mapsto 1$, $\{3, 4\} \mapsto 2$ is a graph-morphism with right-inverse $1 \mapsto 2$, $2 \mapsto 3$. We have $\mathbb{Q}\text{Ent}([0, 1], f, q) = \ln(3/2)$.

References


