III. Measuring Interactions Within a Dynamical System

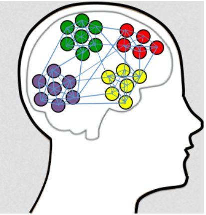
Karl Petersen and Benjamin Wilson

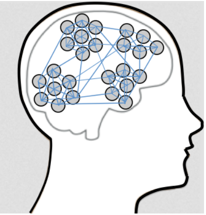
University of North Carolina at Chapel Hill

Cantor Salta
November 4, 2015

Contrasting brain actions in higher vertebrates:

1. Specificity and modularity (functional segregation)
2. Global functions and mass actions (integration in perception and behavior)





Neural Complexity (G. Edelman, O. Sporns, G. Tononi, 1994)

Neural Complexity (G. Edelman, O. Sporns, G. Tononi, 1994)

- Neither specificity nor coordination alone adequately accounts for brain activity.

Neural Complexity (G. Edelman, O. Sporns, G. Tononi, 1994)

- Neither specificity nor coordination alone adequately accounts for brain activity.
- So they propose a general measure—*intricacy*— that encompasses these fundamental aspects of brain organization.

Neural Complexity (G. Edelman, O. Sporns, G. Tononi, 1994)

- Neither specificity nor coordination alone adequately accounts for brain activity.
- So they propose a general measure—*intricacy*— that encompasses these fundamental aspects of brain organization.
- High values are associated with non-trivial organization of the network. This is the case when segregation coexists with integration.

Neural Complexity (G. Edelman, O. Sporns, G. Tononi, 1994)

- Neither specificity nor coordination alone adequately accounts for brain activity.
- So they propose a general measure—*intricacy*— that encompasses these fundamental aspects of brain organization.
- High values are associated with non-trivial organization of the network. This is the case when segregation coexists with integration.
- Low values are associated with systems that are either completely independent (segregated, disordered) or completely dependent (integrated, ordered).

Intricacy is an average of the interaction, measured by mutual information, between sets of sites and their complements.

Mutual Information

Entropy of a random variable X taking values in a discrete set E:

$$H(X) = -\sum_{x \in E} Pr\{X = x\} \log Pr\{X = x\}.$$

Mutual information between random variables X and Y over the same probability space:

$$MI(X, Y) = H(X) + H(Y) - H(X, Y).$$

- $MI(X, Y) \geqslant 0$
- $MI(X, Y) = 0 \Leftrightarrow X$ and Y are independent

- $n^* = \{0, 1 \ldots, n-1\}$
- $X = \{X_i : i \in n^*\}$ a family of random variables representing an isolated neural system with n elementary components (neuronal groups)
- For $S \subset n^*$, $X_S = \{X_i : i \in S\}$
- $S^c = n^* \setminus S$.

Intricacy (or Neural Complexity), C_N

Average of mutual information over subfamilies of a family of random variables

$$C_N(X) = \frac{1}{n+1} \sum_{S \subset n^*} \frac{1}{\binom{n}{|S|}} MI(X_S, X_{S^c}).$$

Intricacies in probability (J. Buzzi, L. Zambotti, 2009)

- Give a general probabilisitic representation of neural complexity.
- Neural complexity belongs to a natural class of functionals: *weighted averages of mutual information* whose weights satisfy certain properties.

Intricacies in probability (J. Buzzi, L. Zambotti, 2009)

- Give a general probabilisitic representation of neural complexity.
- Neural complexity belongs to a natural class of functionals: *weighted averages of mutual information* whose weights satisfy certain properties.

System of coefficients

A *system of coefficients*, c_S^n, is a family of numbers satisfying for all $n \in \mathbb{N}$ and $S \subset n^*$

1. $c_S^n \geqslant 0$;
2. $\sum_{S \subset n^*} c_S^n = 1$;
3. $c_{S^c}^n = c_S^n$.

Mutual information functional

- For a fixed $n \in \mathbb{N}$ let $X = \{X_i : i \in n^*\}$ be a collection of random variables all taking values in the same finite set.
- Given a system of coefficients, c^n_S, the corresponding *mutual information functional*, $\mathcal{J}^c(X)$ is defined by

$$\mathcal{J}^c(X) = \sum_{S \subset n^*} c^n_S MI(X_S, X_{S^c}).$$

Mutual information functional

- For a fixed $n \in \mathbb{N}$ let $X = \{X_i : i \in n^*\}$ be a collection of random variables all taking values in the same finite set.
- Given a system of coefficients, c_S^n, the corresponding *mutual information functional*, $\mathcal{I}^c(X)$ is defined by

$$\mathcal{I}^c(X) = \sum_{S \subset n^*} c_S^n MI(X_S, X_{S^c}).$$

Intricacies: Probabilistic definition

An *intricacy* is a mutual information functional satisfying:

1. Exchangeability: invariance by permutations of n;
2. Weak additivity: $\mathcal{I}^c(X, Y) = \mathcal{I}^c(X) + \mathcal{I}^c(Y)$ for any two independent systems $X = \{X_i : i \in n^*\}$ and $Y = \{Y_j : j \in m^*\}$.

Theorem (Buzzi, Zambotti)

Let c_S^n be a system of coefficients and $\mathcal{I}^c$ the associated mutual information functional. $\mathcal{I}^c$ is an intricacy if and only if there exists a symmetric probability measure λ_c on $[0,1]$ such that

$$c_S^n = \int_{[0,1]} x^{|S|}(1-x)^{n-|S|}\lambda_c(dx)$$

Theorem (Buzzi, Zambotti)

Let c_S^n be a system of coefficients and $\mathcal{I}^c$ the associated mutual information functional. $\mathcal{I}^c$ is an intricacy if and only if there exists a symmetric probability measure λ_c on $[0,1]$ such that

$$c_S^n = \int_{[0,1]} x^{|S|}(1-x)^{n-|S|}\lambda_c(dx)$$

Examples

1. $c_S^n = \dfrac{1}{(n+1)}\dfrac{1}{\binom{n}{|S|}}$ (Edelman-Sporns-Tononi);
2. For $0 < p < 1$,
$$c_S^n = \frac{1}{2}(p^{|S|}(1-p)^{n-|S|} + (1-p)^{|S|}p^{n-|S|}) \text{ (p-symmetric);}$$
3. For $p = 1/2$, $c_S^n = 2^{-n}$ (uniform).

Definitions in Dynamics

Topological dynamical system, (X, T)

- X a compact Hausdorff (often metric) space;
- $T : X \to X$ a homeomorphism.

Definitions in Dynamics

Topological dynamical system, (X, T)

- X a compact Hausdorff (often metric) space;
- $T : X \to X$ a homeomorphism.

For an open cover $\mathcal{U}$ of X, recall that $N(\mathcal{U})$ is the minimum cardinality of any subcover of $\mathcal{U}$.

Definition (Adler, Konheim, McAndrew, 1965)

The *topological entropy* of (X, T) is defined by

$$h_{\text{top}}(X, T) = \sup_{\mathcal{U}} \lim_{n\to\infty} \frac{1}{n} \log N(\mathcal{U} \vee T^{-1}\mathcal{U} \vee \cdots \vee T^{-n+1}\mathcal{U}).$$

Measure of randomness or disorder of a system.

Let (X, T) be a topological dynamical system and $\mathcal{U}$ an open cover of X. Given $n \in \mathbb{N}$ and a subset $S \subset n^*$ define

$$\mathcal{U}_S = \bigvee_{i \in S} T^{-i}\mathcal{U}.$$

Let (X, T) be a topological dynamical system and $\mathcal{U}$ an open cover of X. Given $n \in \mathbb{N}$ and a subset $S \subset n^*$ define

$$\mathcal{U}_S = \bigvee_{i \in S} T^{-i}\mathcal{U}.$$

Definition (P-W)

Let c_S^n be a system of coefficients. Define the *topological intricacy of* (X, T) *with respect to the open cover* $\mathcal{U}$ to be

$$\mathrm{Int}(X, \mathcal{U}, T) := \lim_{n\to\infty} \frac{1}{n} \sum_{S \subset n^*} c_S^n \log\left(\frac{N(\mathcal{U}_S)N(\mathcal{U}_{S^c})}{N(\mathcal{U}_{n^*})}\right).$$

Definition (P-W)

The *topological average sample complexity of T with respect to the open cover $\mathcal{U}$* is defined to be

$$\mathrm{Asc}(X, \mathcal{U}, T) := \lim_{n\to\infty} \frac{1}{n} \sum_{S \subset n^*} c_S^n \log N(\mathcal{U}_S).$$

Definition (P-W)

The *topological average sample complexity of T with respect to the open cover $\mathcal{U}$* is defined to be

$$\mathrm{Asc}(X, \mathcal{U}, T) := \lim_{n\to\infty} \frac{1}{n} \sum_{S \subset n^*} c_S^n \log N(\mathcal{U}_S).$$

So $\quad \mathrm{Int}(X, \mathcal{U}, T) = 2\,\mathrm{Asc}(X, \mathcal{U}, T) - h_{\mathrm{top}}(X, \mathcal{U}, T).$

Theorem

The limits in the definitions of $\mathrm{Int}(X, \mathcal{U}, T)$ *and* $\mathrm{Asc}(X, \mathcal{U}, T)$ *exist.*

Proof based on subadditivity of the sequence

$$b_n := \sum_{S \subset n^*} c_S^n \log N(\mathcal{U}_S)$$

and Fekete's Subadditive Lemma: For every subadditive sequence a_n, the limit $\lim_{n\to\infty} a_n/n$ exists and is equal to $\inf_n a_n/n$.

Theorem

The limits in the definitions of $\mathrm{Int}(X, \mathcal{U}, T)$ *and* $\mathrm{Asc}(X, \mathcal{U}, T)$ *exist.*

Proof based on subadditivity of the sequence

$$b_n := \sum_{S \subset n^*} c_S^n \log N(\mathcal{U}_S)$$

and Fekete's Subadditive Lemma: For every subadditive sequence a_n, the limit $\lim_{n\to\infty} a_n/n$ exists and is equal to $\inf_n a_n/n$.

Proposition

For each open cover $\mathcal{U}$*,*
$\mathrm{Asc}(X, \mathcal{U}, T) \leqslant h_{\mathrm{top}}(X, \mathcal{U}, T) \leqslant h_{\mathrm{top}}(X, T)$*, and hence*
$\mathrm{Int}(X, \mathcal{U}, T) \leqslant h_{\mathrm{top}}(X, \mathcal{U}, T) \leqslant h_{\mathrm{top}}(X, T)$.

In particular, a dynamical system with zero (or relatively low) topological entropy (integrated, ordered) has zero (or relatively low) topological intricacy.

Recall that the set of words of length n in a subshift X is denoted by $\mathcal{L}_n(X)$.

Recall that the set of words of length n in a subshift X is denoted by $\mathcal{L}_n(X)$.

For a subset $S \subset n^*$, $S = \{s_0, s_1, \ldots, s_{|S|-1}\}$, denote the set of words we can see at the places in S for all words in $\mathcal{L}_n(X)$ by $\mathcal{L}_S(X)$:

$$\mathcal{L}_S(X) = \{w_{s_0} w_{s_1} \ldots w_{s_{|S|-1}} : w = w_0 w_1 \ldots w_{n-1} \in \mathcal{L}_n(X)\}.$$

Notice $\mathcal{L}_{n^*}(X) = \mathcal{L}_n(X)$.

Intricacy of a subshift, X

$$\mathrm{Int}(X, \mathcal{U}_0, \sigma) = \lim_{n\to\infty} \frac{1}{n} \sum_{S \subset n^*} c_S^n \log\left(\frac{|\mathcal{L}_S(X)||\mathcal{L}_{S^c}(X)|}{|\mathcal{L}_{n^*}(X)|}\right)$$

Intricacy of a subshift, X

$$\mathrm{Int}(X, \mathcal{U}_0, \sigma) = \lim_{n \to \infty} \frac{1}{n} \sum_{S \subset n^*} c_S^n \log \left(\frac{|\mathcal{L}_S(X)||\mathcal{L}_{S^c}(X)|}{|\mathcal{L}_{n^*}(X)|} \right)$$

Example (Computing $|\mathcal{L}_S(X)|$ for the golden mean sft)

Let $n = 3$, $n^* = \{0, 1, 2\}$.

$S = \{0, 1\}$

0	0	
0	1	
1	0	

$|\mathcal{L}_S(X)| = 3$

$S = \{0, 2\}$

0		0
0		1
1		0
1		1

$|\mathcal{L}_S(X)| = 4$

Example (Computing $|\mathcal{L}_S(X)|$ for the golden mean sft)

S	S^c	$\lvert\mathcal{L}_S(X)\rvert$	$\lvert\mathcal{L}_{S^c}(X)\rvert$
$\emptyset$	$\{0,1,2\}$	1	5
$\{0\}$	$\{1,2\}$	2	3
$\{1\}$	$\{0,2\}$	2	4
$\{2\}$	$\{0,1\}$	2	3
$\{0,1\}$	$\{2\}$	3	2
$\{0,2\}$	$\{1\}$	4	2
$\{1,2\}$	$\{0\}$	3	2
$\{0,1,2\}$	$\emptyset$	5	1

Example (Computing $|\mathcal{L}_S(X)|$ for the golden mean sft)

S	S^c	$\lvert\mathcal{L}_S(X)\rvert$	$\lvert\mathcal{L}_{S^c}(X)\rvert$
$\emptyset$	$\{0,1,2\}$	1	5
$\{0\}$	$\{1,2\}$	2	3
$\{1\}$	$\{0,2\}$	2	4
$\{2\}$	$\{0,1\}$	2	3
$\{0,1\}$	$\{2\}$	3	2
$\{0,2\}$	$\{1\}$	4	2
$\{1,2\}$	$\{0\}$	3	2
$\{0,1,2\}$	$\emptyset$	5	1

$$\frac{1}{3 \cdot 2^3} \sum_{S \subset 3^*} \log\left(\frac{|\mathcal{L}_S(X)||\mathcal{L}_{S^c}(X)|}{|\mathcal{L}_{n^*}(X)|}\right) = \frac{1}{24} \log\left(\frac{6^4 \cdot 8^2}{5^6}\right) \approx 0.070$$

Theorem

Let X be a shift of finite type with adjacency matrix M such that $M^2 > 0$. Let $c_S^n = 2^{-n}$ for all S. Then

$$\mathrm{Asc}(X, \mathcal{U}_0, \sigma) = \frac{1}{4} \sum_{k=1}^{\infty} \frac{\log |\mathcal{L}_{k^*}(X)|}{2^k}.$$

Theorem

Let X be a shift of finite type with adjacency matrix M such that $M^2 > 0$. Let $c_S^n = 2^{-n}$ for all S. Then

$$\mathrm{Asc}(X, \mathcal{U}_0, \sigma) = \frac{1}{4} \sum_{k=1}^{\infty} \frac{\log |\mathcal{L}_{k^*}(X)|}{2^k}.$$

Asc is sensitive to word counts of all lengths, so is a finer measurement than h_{top}, which just gives the asymptotic exponential growth rate.

Theorem

Let X be a shift of finite type with adjacency matrix M such that $M^2 > 0$. Let $c_S^n = 2^{-n}$ for all S. Then

$$\mathrm{Asc}(X, \mathcal{U}_0, \sigma) = \frac{1}{4} \sum_{k=1}^{\infty} \frac{\log |\mathcal{L}_{k^*}(X)|}{2^k}.$$

Asc is sensitive to word counts of all lengths, so is a finer measurement than h_{top}, which just gives the asymptotic exponential growth rate.
Proof idea: Most subsets $S \subset n^*$ are also subsets of $(n-1)^*$.

Theorem

Let X be a shift of finite type with adjacency matrix M such that $M^2 > 0$. Let $c_S^n = 2^{-n}$ for all S. Then

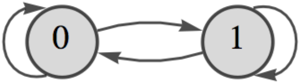
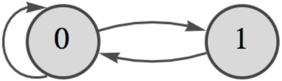
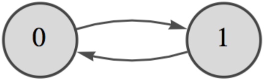
$$\mathrm{Asc}(X, \mathcal{U}_0, \sigma) = \frac{1}{4} \sum_{k=1}^{\infty} \frac{\log |\mathcal{L}_{k^*}(X)|}{2^k}.$$

Asc is sensitive to word counts of all lengths, so is a finer measurement than h_{top}, which just gives the asymptotic exponential growth rate.
Proof idea: Most subsets $S \subset n^*$ are also subsets of $(n-1)^*$.

Corollary

For the full r-shift with $c_S^n = 2^{-n}$ for all S,

$$\mathrm{Asc}(\Sigma_r, \mathcal{U}_0, \sigma) = \frac{\log r}{2} \quad \textit{and} \quad \mathrm{Int}(\Sigma_r, \mathcal{U}_0, \sigma) = 0.$$

	Adjacency Graph	Entropy	Asc	Int
Disordered		0.693	0.347	0
		0.481	0.286	0.090
Ordered		0	0	0

Theorem

Let (X, T) be a topological dynamical system and fix the system of coefficients to be $c_S^n = 2^{-n}$. Then

$$\sup_{\mathcal{U}} \operatorname{Asc}(X, \mathcal{U}, T) = h_{\mathrm{top}}(X, T).$$

Theorem

Let (X, T) *be a topological dynamical system and fix the system of coefficients to be* $c_S^n = 2^{-n}$*. Then*

$$\sup_{\mathcal{U}} \operatorname{Asc}(X, \mathcal{U}, T) = h_{\text{top}}(X, T).$$

- The proof depends on the structure of average subsets of $n^* = \{0, 1, \ldots, n-1\}$.

Theorem

Let (X, T) be a topological dynamical system and fix the system of coefficients to be $c_S^n = 2^{-n}$. Then

$$\sup_{\mathcal{U}} \mathrm{Asc}(X, \mathcal{U}, T) = h_{\mathrm{top}}(X, T).$$

- The proof depends on the structure of average subsets of $n^* = \{0, 1, \ldots, n-1\}$.
- Most $S \subset n^*$ have size about $n/2$, so are not too sparse.

Theorem

Let (X, T) be a topological dynamical system and fix the system of coefficients to be $c_S^n = 2^{-n}$. Then

$$\sup_{\mathcal{U}} \operatorname{Asc}(X, \mathcal{U}, T) = h_{\text{top}}(X, T).$$

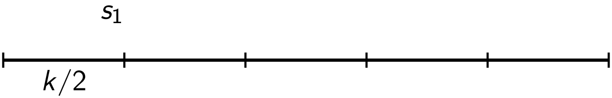
- The proof depends on the structure of average subsets of $n^* = \{0, 1, \ldots, n-1\}$.
- Most $S \subset n^*$ have size about $n/2$, so are not too sparse.
- In ordinary topological entropy of a subshift, using the time-0 partition (or open cover) α, when we replace α by $\alpha_{k^*} = \alpha_0^{k-1}$ in counting the number of cells or calculating the entropy of the refined partition, instead of α_{n^*}, we are looking at $\alpha_{(n+k)^*}$, and when k is fixed, as n grows the result is the same.

Theorem

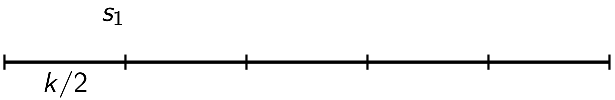
Let (X, T) be a topological dynamical system and fix the system of coefficients to be $c_S^n = 2^{-n}$. Then

$$\sup_{\mathcal{U}} \mathrm{Asc}(X, \mathcal{U}, T) = h_{\mathrm{top}}(X, T).$$

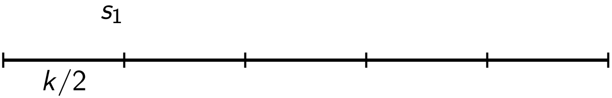
- The proof depends on the structure of average subsets of $n^* = \{0, 1, \dots, n-1\}$.
- Most $S \subset n^*$ have size about $n/2$, so are not too sparse.
- In ordinary topological entropy of a subshift, using the time-0 partition (or open cover) α, when we replace α by $\alpha_{k^*} = \alpha_0^{k-1}$ in counting the number of cells or calculating the entropy of the refined partition, instead of α_{n^*}, we are looking at $\alpha_{(n+k)^*}$, and when k is fixed, as n grows the result is the same.
- When we code by k-blocks, $S \subset n^*$ is replaced by $S + k^*$, and the effect on α_{S+k^*} as compared to α_S is similar, since it acts similarly on each of the long subintervals comprising S.



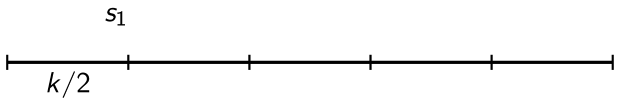
- Fix a k for coding by k-blocks (or looking at $N((\mathcal{U}_k)_S)$ or $H((\alpha_k)_S)$).



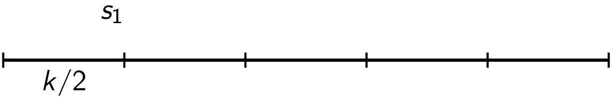
- Fix a k for coding by k-blocks (or looking at $N((\mathcal{U}_k)_S)$ or $H((\alpha_k)_S)$).
- Cut n^* into consecutive blocks of length $k/2$.



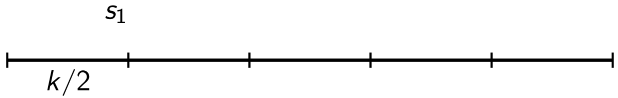
- Fix a k for coding by k-blocks (or looking at $N((\mathcal{U}_k)_S)$ or $H((\alpha_k)_S)$).
- Cut n^* into consecutive blocks of length $k/2$.
- When $s \in S$ is in one of these intervals of length $k/2$, then $s + k^*$ *covers* the next interval of length $k/2$.



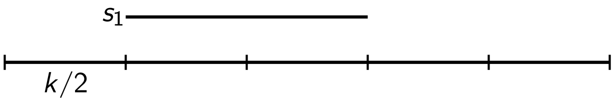
- Fix a k for coding by k-blocks (or looking at $N((\mathcal{U}_k)_S)$ or $H((\alpha_k)_S)$).
- Cut n^* into consecutive blocks of length $k/2$.
- When $s \in S$ is in one of these intervals of length $k/2$, then $s + k^*$ *covers* the next interval of length $k/2$.
- So if S hits many of the intervals of length $k/2$, then $S + k^*$ starts to look like a union of long intervals, say each with $|E_j| > k$.



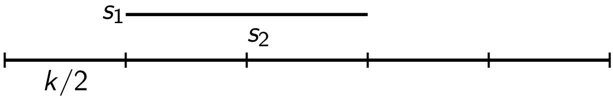
- Fix a k for coding by k-blocks (or looking at $N((\mathcal{U}_k)_S)$ or $H((\alpha_k)_S)$).
- Cut n^* into consecutive blocks of length $k/2$.
- When $s \in S$ is in one of these intervals of length $k/2$, then $s + k^*$ *covers* the next interval of length $k/2$.
- So if S hits many of the intervals of length $k/2$, then $S + k^*$ starts to look like a union of long intervals, say each with $|E_j| > k$.
- By shaving a little off each of these relatively long intervals, we can assume that also the gaps have length at least k.



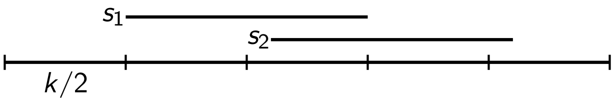
- Fix a k for coding by k-blocks (or looking at $N((\mathcal{U}_k)_S)$ or $H((\alpha_k)_S)$).
- Cut n^* into consecutive blocks of length $k/2$.
- When $s \in S$ is in one of these intervals of length $k/2$, then $s + k^*$ *covers* the next interval of length $k/2$.
- So if S hits many of the intervals of length $k/2$, then $S + k^*$ starts to look like a union of long intervals, say each with $|E_j| > k$.
- By shaving a little off each of these relatively long intervals, we can assume that also the gaps have length at least k.



- Fix a k for coding by k-blocks (or looking at $N((\mathcal{U}_k)_S)$ or $H((\alpha_k)_S)$).
- Cut n^* into consecutive blocks of length $k/2$.
- When $s \in S$ is in one of these intervals of length $k/2$, then $s + k^*$ *covers* the next interval of length $k/2$.
- So if S hits many of the intervals of length $k/2$, then $S + k^*$ starts to look like a union of long intervals, say each with $|E_j| > k$.
- By shaving a little off each of these relatively long intervals, we can assume that also the gaps have length at least k.



- Fix a k for coding by k-blocks (or looking at $N((\mathcal{U}_k)_S)$ or $H((\alpha_k)_S)$).
- Cut n^* into consecutive blocks of length $k/2$.
- When $s \in S$ is in one of these intervals of length $k/2$, then $s + k^*$ *covers* the next interval of length $k/2$.
- So if S hits many of the intervals of length $k/2$, then $S + k^*$ starts to look like a union of long intervals, say each with $|E_j| > k$.
- By shaving a little off each of these relatively long intervals, we can assume that also the gaps have length at least k.



- Fix a k for coding by k-blocks (or looking at $N((\mathcal{U}_k)_S)$ or $H((\alpha_k)_S)$).
- Cut n^* into consecutive blocks of length $k/2$.
- When $s \in S$ is in one of these intervals of length $k/2$, then $s + k^*$ *covers* the next interval of length $k/2$.
- So if S hits many of the intervals of length $k/2$, then $S + k^*$ starts to look like a union of long intervals, say each with $|E_j| > k$.
- By shaving a little off each of these relatively long intervals, we can assume that also the gaps have length at least k.

- Given $\epsilon > 0$, we may assume k is large enough that for every interval $I \subset \mathbb{N}$ with $|I| \geqslant k/2$,

$$0 \leqslant \frac{\log N(I)}{\operatorname{card}(I)} - h_{\text{top}}(X, \sigma) < \epsilon.$$

- Given $\epsilon > 0$, we may assume k is large enough that for every interval $I \subset \mathbb{N}$ with $|I| \geqslant k/2$,

$$0 \leqslant \frac{\log N(I)}{\mathrm{card}(I)} - h_{\mathrm{top}}(X, \sigma) < \epsilon.$$

- We let $\mathfrak{B}$ denote the set of $S \subset n^*$ which miss at least $2n\epsilon/k$ of the intervals of length $k/2$

- Given $\epsilon > 0$, we may assume k is large enough that for every interval $I \subset \mathbb{N}$ with $|I| \geqslant k/2$,

$$0 \leqslant \frac{\log N(I)}{\mathrm{card}(I)} - h_{\mathrm{top}}(X, \sigma) < \epsilon.$$

- We let $\mathfrak{B}$ denote the set of $S \subset n^*$ which miss at least $2n\epsilon/k$ of the intervals of length $k/2$
- and show that $\lim_{n\to\infty} \frac{\mathrm{card}(\mathfrak{B})}{2^n} = 0$.

- Given $\epsilon > 0$, we may assume k is large enough that for every interval $I \subset \mathbb{N}$ with $|I| \geqslant k/2$,

$$0 \leqslant \frac{\log N(I)}{\mathrm{card}(I)} - h_{\mathrm{top}}(X, \sigma) < \epsilon.$$

- We let $\mathfrak{B}$ denote the set of $S \subset n^*$ which miss at least $2n\epsilon/k$ of the intervals of length $k/2$
- and show that $\lim_{n\to\infty} \frac{\mathrm{card}(\mathfrak{B})}{2^n} = 0$.
- If $S \notin \mathfrak{B}$, then S hits many of the intervals of length $k/2$,

- Given $\epsilon > 0$, we may assume k is large enough that for every interval $I \subset \mathbb{N}$ with $|I| \geqslant k/2$,

$$0 \leqslant \frac{\log N(I)}{\operatorname{card}(I)} - h_{\mathrm{top}}(X, \sigma) < \epsilon.$$

- We let $\mathfrak{B}$ denote the set of $S \subset n^*$ which miss at least $2n\epsilon/k$ of the intervals of length $k/2$
- and show that $\lim_{n\to\infty} \frac{\operatorname{card}(\mathfrak{B})}{2^n} = 0$.
- If $S \notin \mathfrak{B}$, then S hits many of the intervals of length $k/2$,
- and hence $S + k^*$ is the union of intervals of length at least k, and we can arrange that the gaps are also long enough to satisfy the above estimate comparing to $h_{\mathrm{top}}(X, \sigma)$.

Measure-theoretic situation

Measure-theoretic dynamical system $(X, \mathcal{B}, \mu, T)$

- X is a measure space
- $\mathcal{B}$ is a σ-algebra of measurable subsets of X
- μ is a probability measure on X, i.e., $\mu(X) = 1$
- $T : X \to X$ is a measure-preserving transformation on X, i.e., T is a one-to-one onto map such that $\mu(T^{-1}E) = \mu(E)$ for all $E \in \mathcal{B}$

Measure-theoretic situation

Measure-theoretic dynamical system $(X, \mathcal{B}, \mu, T)$

- X is a measure space
- $\mathcal{B}$ is a σ-algebra of measurable subsets of X
- μ is a probability measure on X, i.e., $\mu(X) = 1$
- $T : X \to X$ is a measure-preserving transformation on X, i.e., T is a one-to-one onto map such that $\mu(T^{-1}E) = \mu(E)$ for all $E \in \mathcal{B}$

Entropy of a partition

The *entropy of a finite measurable partition* $\alpha = \{A_1, \ldots, A_n\}$ *of* X is defined by

$$H_\mu(\alpha) = -\sum_{i=1}^{n} \mu(A_i) \log \mu(A_i).$$

Definition

The *entropy of X and T with respect to* μ *and a partition* α is

$$h_\mu(X, \alpha, T) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+1}\alpha).$$

The *entropy of the transformation T* is defined to be

$$h_\mu(X, T) = \sup_\alpha h_\mu(X, \alpha, T).$$

For a partition α of X and a subset $S \subset n^*$ define

$$\alpha_S = \bigvee_{i \in S} T^{-i}\alpha.$$

For a partition α of X and a subset $S \subset n^*$ define

$$\alpha_S = \bigvee_{i \in S} T^{-i} \alpha.$$

Definition (P-W)

Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system, $\alpha = \{A_1, \ldots, A_n\}$ a finite measurable partition of X, and c_S^n a system of coefficients. The *measure-theoretic intricacy of* T *with respect to the partition* α is

$$\mathrm{Int}_\mu(X, \alpha, T) = \lim_{n \to \infty} \frac{1}{n} \sum_{S \subset n^*} c_S^n \left[H_\mu(\alpha_S) + H_\mu(\alpha_{S^c}) - H_\mu(\alpha_{n^*}) \right].$$

For a partition α of X and a subset $S \subset n^*$ define

$$\alpha_S = \bigvee_{i \in S} T^{-i}\alpha.$$

Definition (P-W)

Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system, $\alpha = \{A_1, \ldots, A_n\}$ a finite measurable partition of X, and c_S^n a system of coefficients. The *measure-theoretic intricacy of T with respect to the partition α* is

$$\mathrm{Int}_\mu(X, \alpha, T) = \lim_{n\to\infty} \frac{1}{n} \sum_{S \subset n^*} c_S^n \left[H_\mu(\alpha_S) + H_\mu(\alpha_{S^c}) - H_\mu(\alpha_{n^*}) \right].$$

The *measure-theoretic average sample complexity of T with respect to the partition α* is

$$\mathrm{Asc}_\mu(X, \alpha, T) = \lim_{n\to\infty} \frac{1}{n} \sum_{S \subset n^*} c_S^n H_\mu(\alpha_S).$$

Theorem

The limits in the definitions of measure-theoretic intricacy and measure-theoretic average sample complexity exist.

Theorem
The limits in the definitions of measure-theoretic intricacy and measure-theoretic average sample complexity exist.

Theorem
Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system and fix the system of coefficients $c_S^n = 2^{-n}$. Then

$$\sup_{\alpha} \mathrm{Asc}_{\mu}(X, \alpha, T) = h_{\mu}(X, T).$$

Theorem

The limits in the definitions of measure-theoretic intricacy and measure-theoretic average sample complexity exist.

Theorem

Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system and fix the system of coefficients $c_S^n = 2^{-n}$. Then

$$\sup_{\alpha} \mathrm{Asc}_{\mu}(X, \alpha, T) = h_{\mu}(X, T).$$

The proofs are similar to those for the corresponding theorems in topological setting.

Theorem

The limits in the definitions of measure-theoretic intricacy and measure-theoretic average sample complexity exist.

Theorem

Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system and fix the system of coefficients $c_S^n = 2^{-n}$. Then

$$\sup_{\alpha} \mathrm{Asc}_{\mu}(X, \alpha, T) = h_{\mu}(X, T).$$

The proofs are similar to those for the corresponding theorems in topological setting.These observations indicate that there may be a topological analogue of the following result.

Theorem

The limits in the definitions of measure-theoretic intricacy and measure-theoretic average sample complexity exist.

Theorem

Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system and fix the system of coefficients $c_S^n = 2^{-n}$. Then

$$\sup_{\alpha} \mathrm{Asc}_{\mu}(X, \alpha, T) = h_{\mu}(X, T).$$

The proofs are similar to those for the corresponding theorems in topological setting.These observations indicate that there may be a topological analogue of the following result.

Theorem (Ornstein-Weiss, 2007)

If J is a finitely observable functional defined for ergodic finite-valued processes that is an isomorphism invariant, then J is a continuous function of the measure-theoretic entropy.

- The arguments adapt to open covers $(\mathcal{U}_k)$ and partitions α_k.

- The arguments adapt to open covers $(\mathcal{U}_k)$ and partitions α_k.
- So it is better to examine these measures *locally*:

- The arguments adapt to open covers $(\mathcal{U}_k)$ and partitions α_k.
- So it is better to examine these measures *locally*:
- Fix a k and find the topological average sample complexity $Asc(X, \mathcal{U}_k, \sigma) = \lim_{n\to\infty} \frac{1}{n} \sum_{S\subset n^*} c_S^n \log N((\mathcal{U}_k)_S)$,

- The arguments adapt to open covers $(\mathcal{U}_k)$ and partitions α_k.
- So it is better to examine these measures *locally*:
- Fix a k and find the topological average sample complexity $Asc(X, \mathcal{U}_k, \sigma) = \lim_{n\to\infty} \frac{1}{n} \sum_{S\subset n^*} c_S^n \log N((\mathcal{U}_k)_S)$,
- or do not take the limit on n, and study it as a function of n,

- The arguments adapt to open covers $(\mathcal{U}_k)$ and partitions α_k.
- So it is better to examine these measures *locally*:
- Fix a k and find the topological average sample complexity $Asc(X, \mathcal{U}_k, \sigma) = \lim_{n\to\infty} \frac{1}{n} \sum_{S\subset n^*} c^n_S \log N((\mathcal{U}_k)_S)$,
- or do not take the limit on n, and study it as a function of n,
- analogously to the symbolic or topological complexity functions.

- The arguments adapt to open covers $(\mathcal{U}_k)$ and partitions α_k.
- So it is better to examine these measures *locally*:
- Fix a k and find the topological average sample complexity $Asc(X, \mathcal{U}_k, \sigma) = \lim_{n\to\infty} \frac{1}{n} \sum_{S\subset n^*} c_S^n \log N((\mathcal{U}_k)_S)$,
- or do not take the limit on n, and study it as a function of n,
- analogously to the symbolic or topological complexity functions.
- Similarly for the measure-theoretic version: fix a partition α and study the limit, or the function of n.

$$\mathrm{Asc}_\mu(X, T, \alpha) = \lim_{n\to\infty} \frac{1}{n} \sum_{S\subset n^*} c_S^n H_\mu(\alpha_S).$$

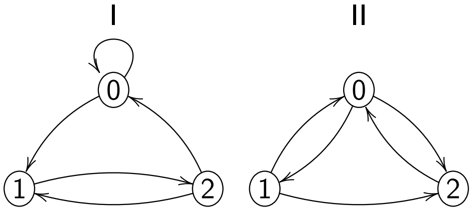
So we begin study of Asc for a fixed open cover as a function of n.

$$\mathrm{Asc}(X, \sigma, \mathcal{U}_k, n) = \frac{1}{n} \sum_{S \subset n^*} c_S^n \log N(S).$$

So we begin study of Asc for a fixed open cover as a function of n.

$$\mathrm{Asc}(X, \sigma, \mathcal{U}_k, n) = \frac{1}{n} \sum_{S \subset n^*} c_S^n \log N(S).$$

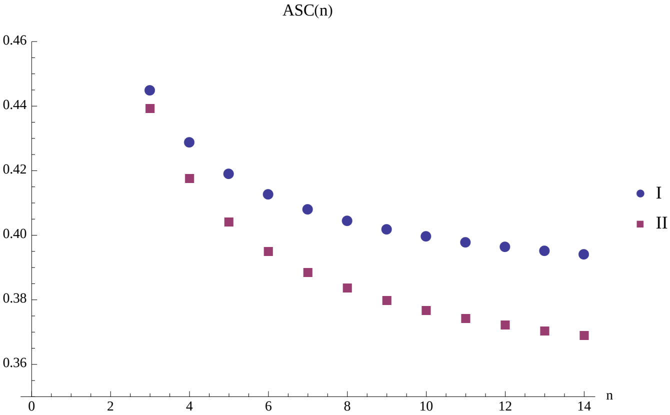
Example



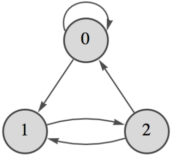
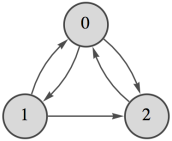
$(\mathrm{M}_I)^3 > 0$ $\quad$ $(\mathrm{M}_{II})^4 > 0$

Figure: Graphs of two subshifts with the same complexity function but different average sample complexity functions.

$$\mathrm{Asc}(n) = \frac{1}{n}\frac{1}{2^n}\sum_{S \subset n^*} \log N(S)$$



Interesting example

Adjacency Graph	h_{top}	Asc(10)	Int(10)
	0.481	0.399	0.254
	0.481	0.377	0.208

These SFTs have the same entropy and complexity functions (words of length n) but different Asc and Int functions.

Results in measure-theoretic setting

For a fixed partition α, we give a relationship between $\mathrm{Asc}_{\mu}(X, \alpha, T)$ and a series summed over i involving the conditional entropies $H_{\mu}(\alpha \mid \alpha_i^{\infty})$.

Results in measure-theoretic setting

For a fixed partition α, we give a relationship between $\mathrm{Asc}_\mu(X, \alpha, T)$ and a series summed over i involving the conditional entropies $H_\mu(\alpha \mid \alpha_i^\infty)$.

Idea

Results in measure-theoretic setting

For a fixed partition α, we give a relationship between $\mathrm{Asc}_\mu(X, \alpha, T)$ and a series summed over i involving the conditional entropies $H_\mu(\alpha \mid \alpha_i^\infty)$.

Idea

- View a subset $S \subset n^*$ as corresponding to a random binary string of length n generated by Bernoulli measure $\mathcal{B}(1/2, 1/2)$ on the full 2-shift.

Results in measure-theoretic setting

For a fixed partition α, we give a relationship between $\mathrm{Asc}_\mu(X, \alpha, T)$ and a series summed over i involving the conditional entropies $H_\mu(\alpha \mid \alpha_i^\infty)$.

Idea

- View a subset $S \subset n^*$ as corresponding to a random binary string of length n generated by Bernoulli measure $\mathcal{B}(1/2, 1/2)$ on the full 2-shift.
- For example $\{0, 2, 3\} \subset 5^* \leftrightarrow 10110$.

Results in measure-theoretic setting

For a fixed partition α, we give a relationship between $\mathrm{Asc}_{\mu}(X, \alpha, T)$ and a series summed over i involving the conditional entropies $H_{\mu}(\alpha \mid \alpha_i^{\infty})$.

Idea

- View a subset $S \subset n^*$ as corresponding to a random binary string of length n generated by Bernoulli measure $\mathcal{B}(1/2, 1/2)$ on the full 2-shift.
- For example $\{0, 2, 3\} \subset 5^* \leftrightarrow 10110$.
- The average entropy, $H_{\mu}(\alpha_S)$, over all $S \subset n^*$, is then an integral and can be interpreted in terms of the entropy of a first-return map to the cylinder $A = [1]$ in a cross product of our system X and the full 2-shift, Σ_2.

Theorem

Let $(X, \mathcal{B}, \mu, T)$ *be a measure-preserving system and* α *a finite measurable partition of* X*. Let* $A = [1] = \{\xi \in \Sigma_2^+ : \xi_0 = 1\}$ *and* $\beta = \alpha \times A$ *the related finite partition of* $X \times A$*. Denote by* $T_{X \times A}$ *the first-return map on* $X \times A$ *and let* $P_A = P/P[1]$ *denote the measure* P *restricted to* A *and normalized. Let* $c_S^n = 2^{-n}$ *for all* $S \subset n^*$*. Then*

$$\mathrm{Asc}_\mu(X, \alpha, T) = \frac{1}{2} h_{\mu \times P_A}(X \times A, \beta, T_{X \times A}).$$

Theorem

Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system and α a finite measurable partition of X. Let $A = [1] = \{\xi \in \Sigma_2^+ : \xi_0 = 1\}$ and $\beta = \alpha \times A$ the related finite partition of $X \times A$. Denote by $T_{X\times A}$ the first-return map on $X \times A$ and let $P_A = P/P[1]$ denote the measure P restricted to A and normalized. Let $c_S^n = 2^{-n}$ for all $S \subset n^*$. Then

$$\mathrm{Asc}_\mu(X, \alpha, T) = \frac{1}{2} h_{\mu\times P_A}(X \times A, \beta, T_{X\times A}).$$

Theorem

Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system and α a finite measurable partition of X. Let $c_S^n = 2^{-n}$ for all $S \subset n^*$. Then

$$\mathrm{Asc}_\mu(X, \alpha, T) \geqslant \frac{1}{2} \sum_{i=1}^{\infty} \frac{1}{2^i} H_\mu\left(\alpha \mid \alpha_i^\infty\right).$$

Equality holds in certain cases (in particular, for Markov shifts)

In the topological case the first-return map $T_{X \times A}$ is not continuous nor expansive nor even defined on all of $X \times A$ in general, so known results about measures of maximal entropy and equilibrium states do not apply.

In the topological case the first-return map $T_{X \times A}$ is not continuous nor expansive nor even defined on all of $X \times A$ in general, so known results about measures of maximal entropy and equilibrium states do not apply. To maximize Int, there is the added problem of the minus sign in

$$\mathrm{Int}(X, \mathcal{U}, T) = 2\,\mathrm{Asc}(X, \mathcal{U}, T) - h_{\mathrm{top}}(X, \mathcal{U}, T).$$

In the topological case the first-return map $T_{X \times A}$ is not continuous nor expansive nor even defined on all of $X \times A$ in general, so known results about measures of maximal entropy and equilibrium states do not apply. To maximize Int, there is the added problem of the minus sign in

$$\mathrm{Int}(X, \mathcal{U}, T) = 2\,\mathrm{Asc}(X, \mathcal{U}, T) - h_{\mathrm{top}}(X, \mathcal{U}, T).$$

Maybe some modern work on local or relative variational principles, almost subadditive potentials, equilibrium states for shifts with infinite alphabets, etc. could apply? (Barreira, Mummert, Yayama, Cao-Feng-Huang, Huang-Ye-Zhang, Huang-Maass-Romagnoli-Ye, Cheng-Zhao-Cao, ...)

In the topological case the first-return map $T_{X \times A}$ is not continuous nor expansive nor even defined on all of $X \times A$ in general, so known results about measures of maximal entropy and equilibrium states do not apply.To maximize Int, there is the added problem of the minus sign in

$$\mathrm{Int}(X, \mathcal{U}, T) = 2\,\mathrm{Asc}(X, \mathcal{U}, T) - h_{\mathrm{top}}(X, \mathcal{U}, T).$$

Maybe some modern work on local or relative variational principles, almost subadditive potentials, equilibrium states for shifts with infinite alphabets, etc. could apply? (Barreira, Mummert, Yayama, Cao-Feng-Huang, Huang-Ye-Zhang, Huang-Maass-Romagnoli-Ye, Cheng-Zhao-Cao, ...)

But the above theorem does give up some information immediately:

In the topological case the first-return map $T_{X \times A}$ is not continuous nor expansive nor even defined on all of $X \times A$ in general, so known results about measures of maximal entropy and equilibrium states do not apply.To maximize Int, there is the added problem of the minus sign in

$$\mathrm{Int}(X, \mathcal{U}, T) = 2\,\mathrm{Asc}(X, \mathcal{U}, T) - h_{\mathrm{top}}(X, \mathcal{U}, T).$$

Maybe some modern work on local or relative variational principles, almost subadditive potentials, equilibrium states for shifts with infinite alphabets, etc. could apply? (Barreira, Mummert, Yayama, Cao-Feng-Huang, Huang-Ye-Zhang, Huang-Maass-Romagnoli-Ye, Cheng-Zhao-Cao, ...)
But the above theorem does give up some information immediately:

Proposition

When $T : X \to X$ is an expansive homeomorphism on a compact metric space (e.g., a subshift), $\mathrm{Asc}_\mu(X, T, \alpha)$ is an affine upper semicontinuous (in the weak topology) function of μ, so the set of maximal measures for $\mathrm{Asc}_\mu(X, T, \alpha)$ is nonempty, compact, and convex and contains ergodic measures (see Walters, p. 198 ff.).*

Markov Shift

Markov Shift

- Consider the measure on the shift space (Σ_n, σ) given by s stochastic matrix $P = (P_{ij})$ and fixed probability vector $p = \begin{pmatrix} p_0 & p_1 & \dots & p_{n-1} \end{pmatrix}$, i.e. $\sum p_i = 1$ and $pP = p$.

Markov Shift

- Consider the measure on the shift space (Σ_n, σ) given by s stochastic matrix $P = (P_{ij})$ and fixed probability vector $p = \begin{pmatrix} p_0 & p_1 & \dots & p_{n-1} \end{pmatrix}$, i.e. $\sum p_i = 1$ and $pP = p$.
- The measure $\mu_{P,p}$ is defined as usual on cylinder sets by $\mu_{p,P}[i_0 i_1 \dots i_k] = p_{i_0} P_{i_0 i_1} \cdots P_{i_{k-1} i_k}$.

Markov Shift

- Consider the measure on the shift space (Σ_n, σ) given by s stochastic matrix $P = (P_{ij})$ and fixed probability vector $p = \begin{pmatrix} p_0 & p_1 & \dots & p_{n-1} \end{pmatrix}$, i.e. $\sum p_i = 1$ and $pP = p$.
- The measure $\mu_{P,p}$ is defined as usual on cylinder sets by $\mu_{p,P}[i_0 i_1 \dots i_k] = p_{i_0} P_{i_0 i_1} \cdots P_{i_{k-1} i_k}$.

Example (1-step Markov measure on the golden mean shift)

Denote by $P_{00} \in [0, 1]$ the probability of going from 0 to 0 in a sequence of $X_{\{11\}} \subset \Sigma_2$. Then

$$P = \begin{pmatrix} P_{00} & 1 - P_{00} \\ 1 & 0 \end{pmatrix}, \quad p = \begin{pmatrix} \frac{1}{2-P_{00}} & \frac{1-P_{00}}{2-P_{00}} \end{pmatrix}$$

Markov Shift

- Consider the measure on the shift space (Σ_n, σ) given by s stochastic matrix $P = (P_{ij})$ and fixed probability vector $p = \begin{pmatrix} p_0 & p_1 & \dots & p_{n-1} \end{pmatrix}$, i.e. $\sum p_i = 1$ and $pP = p$.
- The measure $\mu_{P,p}$ is defined as usual on cylinder sets by $\mu_{p,P}[i_0 i_1 \dots i_k] = p_{i_0} P_{i_0 i_1} \cdots P_{i_{k-1} i_k}$.

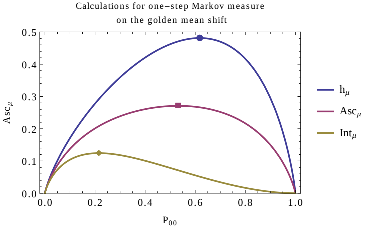
Example (1-step Markov measure on the golden mean shift)

Denote by $P_{00} \in [0,1]$ the probability of going from 0 to 0 in a sequence of $X_{\{11\}} \subset \Sigma_2$. Then

$$P = \begin{pmatrix} P_{00} & 1 - P_{00} \\ 1 & 0 \end{pmatrix}, \quad p = \begin{pmatrix} \frac{1}{2-P_{00}} & \frac{1-P_{00}}{2-P_{00}} \end{pmatrix}$$

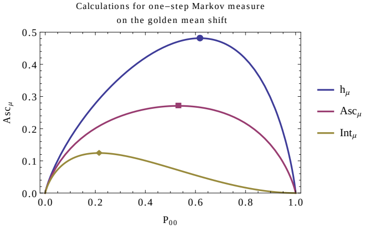
Using the series formula and known equations for conditional entropy, we approximate Asc_μ and Int_μ for Markov measures on SFTs.

1-step Markov measures on the golden mean shift



P_{00}	h_μ	Asc_μ	Int_μ
0.618	**0.481**	0.266	0.051
0.533	0.471	**0.271**	0.071
0.216	0.292	0.208	**0.124**

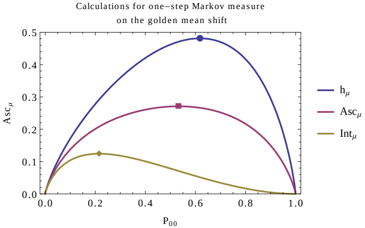
1-step Markov measures on the golden mean shift



P_{00}	h_μ	Asc_μ	Int_μ
0.618	**0.481**	0.266	0.051
0.533	0.471	**0.271**	0.071
0.216	0.292	0.208	**0.124**

- Maximum value of $h_\mu = h_{\text{top}} = \log\phi$ when $P_{00} = 1/\phi$

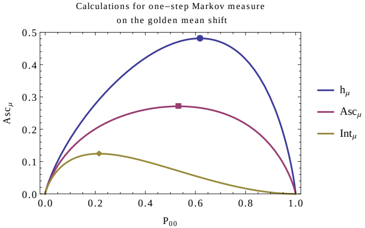
1-step Markov measures on the golden mean shift



P_{00}	h_μ	Asc_μ	Int_μ
0.618	**0.481**	0.266	0.051
0.533	0.471	**0.271**	0.071
0.216	0.292	0.208	**0.124**

- Maximum value of $h_\mu = h_{\text{top}} = \log\phi$ when $P_{00} = 1/\phi$
- Unique maxima among 1-step Markov measures for Asc_μ and Int_μ

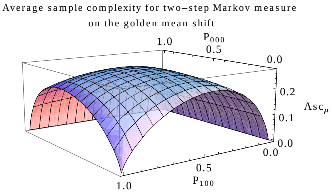
1-step Markov measures on the golden mean shift

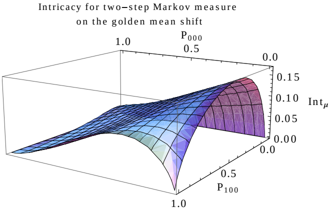


P_{00}	h_μ	Asc_μ	Int_μ
0.618	**0.481**	0.266	0.051
0.533	0.471	**0.271**	0.071
0.216	0.292	0.208	**0.124**

- Maximum value of $h_\mu = h_{\mathrm{top}} = \log \phi$ when $P_{00} = 1/\phi$
- Unique maxima among 1-step Markov measures for Asc_μ and Int_μ
- The maxima for Asc_μ, Int_μ, and h_μ are achieved by different measures

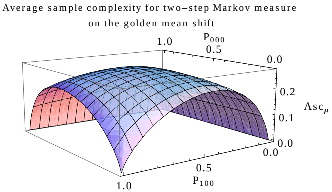
2-step Markov measures on the golden mean shift

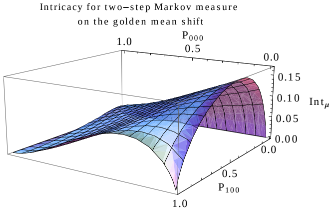




P_{000}	P_{100}	h_{μ}	Asc_{μ}	Int_{μ}
0.618	0.618	**0.481**	0.266	0.051
0.483	0.569	0.466	**0.272**	0.078
0	0.275	0.344	0.221	**0.167**

2-step Markov measures on the golden mean shift

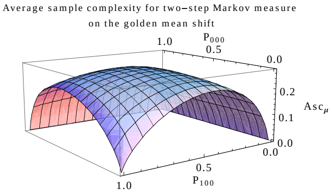


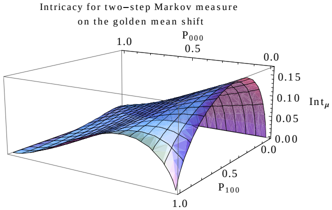


P_{000}	P_{100}	h_μ	Asc_μ	Int_μ
0.618	0.618	**0.481**	0.266	0.051
0.483	0.569	0.466	**0.272**	0.078
0	0.275	0.344	0.221	**0.167**

- Asc_μ appears to be strictly convex, so it would have a unique maximum among 2-step Markov measures

2-step Markov measures on the golden mean shift

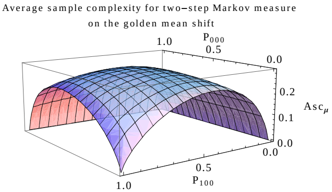


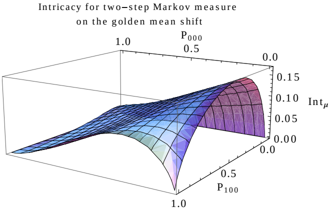


P_{000}	P_{100}	h_μ	Asc_μ	Int_μ
0.618	0.618	**0.481**	0.266	0.051
0.483	0.569	0.466	**0.272**	0.078
0	0.275	0.344	0.221	**0.167**

- Asc_μ appears to be strictly convex, so it would have a unique maximum among 2-step Markov measures
- Int_μ appears to have a unique maximum among 2-step Markov measures on a proper subshift ($P_{000} = 0$)

2-step Markov measures on the golden mean shift

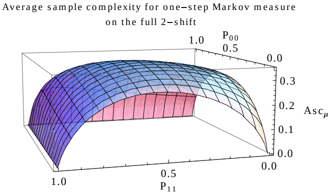


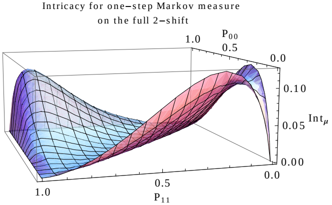


P_{000}	P_{100}	h_μ	Asc_μ	Int_μ
0.618	0.618	**0.481**	0.266	0.051
0.483	0.569	0.466	**0.272**	0.078
0	0.275	0.344	0.221	**0.167**

- Asc_μ appears to be strictly convex, so it would have a unique maximum among 2-step Markov measures
- Int_μ appears to have a unique maximum among 2-step Markov measures on a proper subshift ($P_{000} = 0$)
- The maxima for Asc_μ, Int_μ, and h_μ are achieved by different measures

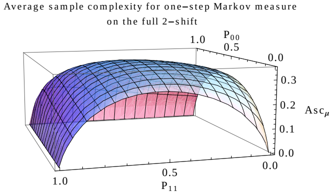
1-step Markov measures on the full 2-shift

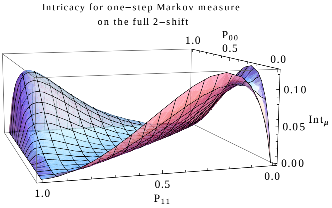




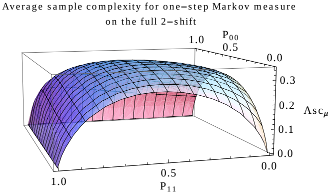
P_{00}	P_{11}	h_μ	Asc_μ	Int_μ
0.5	0.5	**0.693**	**0.347**	0
0.216	0	0.292	0.208	**0.124**
0	0.216	0.292	0.208	**0.124**
0.905	0.905	0.315	0.209	0.104

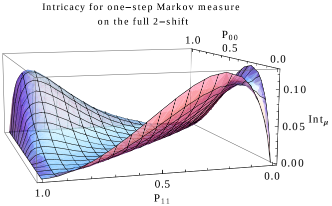
1-step Markov measures on the full 2-shift





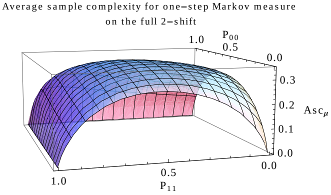
1-step Markov measures on the full 2-shift

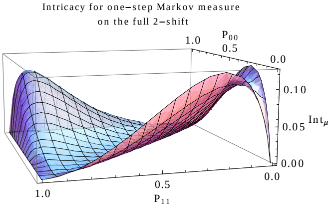




- Asc_{μ} appears to be strictly convex, so it would have a unique maximum among 1-step Markov measures

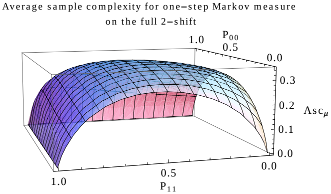
1-step Markov measures on the full 2-shift

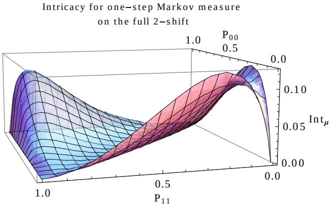




- Asc_μ appears to be strictly convex, so it would have a unique maximum among 1-step Markov measures
- Int_μ appears to have two maxima among 1-step Markov measures on proper subshifts ($P_{00} = 0$ and $P_{11} = 0$).

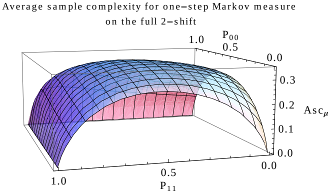
1-step Markov measures on the full 2-shift

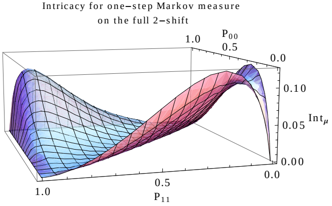




- Asc_μ appears to be strictly convex, so it would have a unique maximum among 1-step Markov measures
- Int_μ appears to have two maxima among 1-step Markov measures on proper subshifts ($P_{00} = 0$ and $P_{11} = 0$).
- There seems to be a 1-step Markov measure that is fully supported and is a local maximum for Int_μ among all 1-step Markov measures.

1-step Markov measures on the full 2-shift



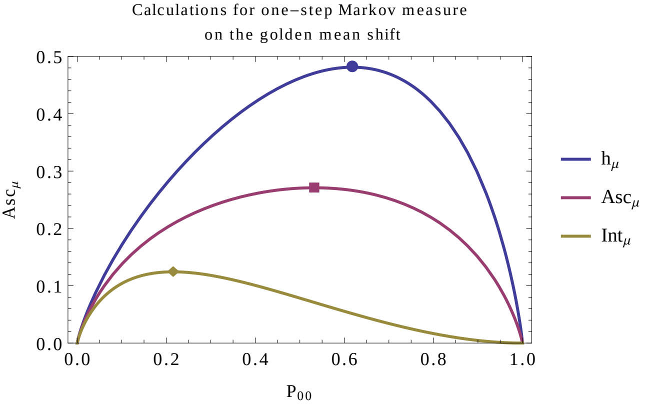


- Asc_μ appears to be strictly convex, so it would have a unique maximum among 1-step Markov measures
- Int_μ appears to have two maxima among 1-step Markov measures on proper subshifts ($P_{00} = 0$ and $P_{11} = 0$).
- There seems to be a 1-step Markov measure that is fully supported and is a local maximum for Int_μ among all 1-step Markov measures.
- The maxima for Asc_μ, Int_μ, and h_μ are achieved by different measures.

We summarize some of the questions generated above.

We summarize some of the questions generated above.
Conj. 1: On the golden mean SFT, for each r there is a unique r-step Markov measure μ_r that maximizes $\mathrm{Asc}_{\mu}(X, \sigma, \alpha)$ among all r-step Markov measures.

We summarize some of the questions generated above.
Conj. 1: On the golden mean SFT, for each r there is a unique r-step Markov measure μ_r that maximizes $\mathrm{Asc}_\mu(X, \sigma, \alpha)$ among all r-step Markov measures.



Conj. 2: $\mu_2 \neq \mu_1$

Conj. 2: $\mu_2 \neq \mu_1$

P_{00}	h_μ	Asc_μ	Int_μ
0.618	**0.481**	0.266	0.051
0.533	0.471	**0.271**	0.071
0.216	0.292	0.208	**0.124**

Table: Calculations for one-step Markov measures on the golden mean shift. Bolded numbers are maxima for given category.

Conj. 2: $\mu_2 \neq \mu_1$

P_{00}	h_μ	Asc_μ	Int_μ
0.618	**0.481**	0.266	0.051
0.533	0.471	**0.271**	0.071
0.216	0.292	0.208	**0.124**

Table: Calculations for one-step Markov measures on the golden mean shift. Bolded numbers are maxima for given category.

P_{000}	P_{100}	h_μ	Asc_μ	Int_μ
0.618	0.618	**0.481**	0.266	0.051
0.483	0.569	0.466	**0.272**	0.078
0	0.275	0.344	0.221	**0.167**

Table: Calculations for two-step Markov measures on the golden mean shift.

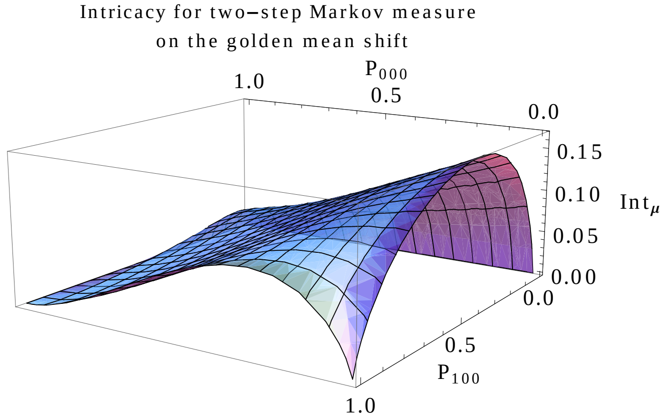
Conj. 3: On the golden mean SFT there is a unique measure that maximizes $\mathrm{Asc}_\mu(X, T, \alpha)$. It is not Markov of any order (and of course is not the same as $\mu_{\max}$).

Conj. 3: On the golden mean SFT there is a unique measure that maximizes $\mathrm{Asc}_\mu(X, T, \alpha)$. It is not Markov of any order (and of course is not the same as $\mu_{\max}$).

Conj. 4: On the golden mean SFT for each r there is a unique r-step Markov measure that maximizes $\mathrm{Int}_\mu(X, T, \alpha)$ among all r-step Markov measures.

Conj. 3: On the golden mean SFT there is a unique measure that maximizes $\mathrm{Asc}_\mu(X, T, \alpha)$. It is not Markov of any order (and of course is not the same as $\mu_{\max}$).

Conj. 4: On the golden mean SFT for each r there is a unique r-step Markov measure that maximizes $\mathrm{Int}_\mu(X, T, \alpha)$ among all r-step Markov measures.



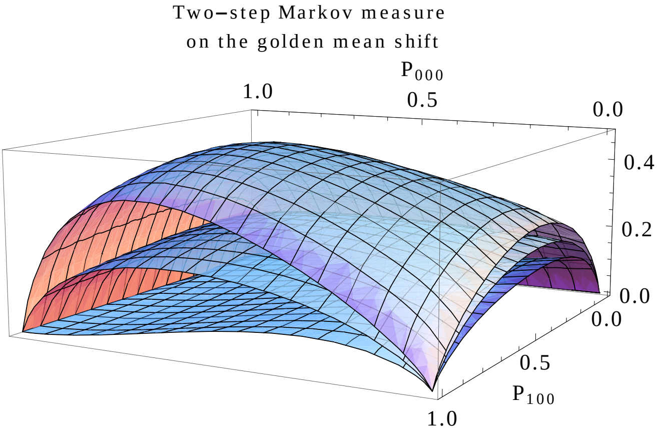
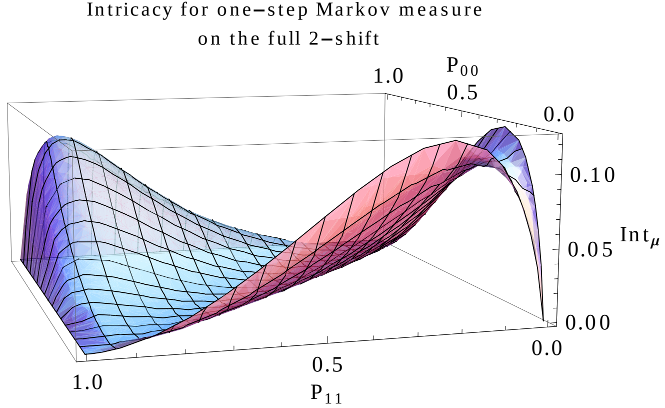


Figure: Combination of the plots of h_μ, Asc_μ, and Int_μ for two-step Markov measures on the golden mean shift.

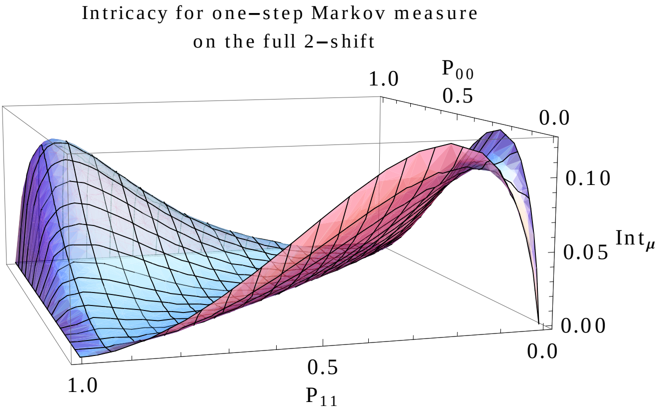
Conj. 5: On the 2-shift there are *two* 1-step Markov measures that maximize $\mathrm{Int}_{\mu}(X, T, \alpha)$ among all 1-step Markov measures. They are supported on the golden mean SFT and its image under the dualizing map $0 \leftrightarrow 1$.

Conj. 5: On the 2-shift there are *two* 1-step Markov measures that maximize $\mathrm{Int}_\mu(X, T, \alpha)$ among all 1-step Markov measures. They are supported on the golden mean SFT and its image under the dualizing map $0 \leftrightarrow 1$.



Conj. 6: On the 2-shift there is a 1-step Markov measure that is *fully supported* and is a local maximum point for $\mathrm{Int}_\mu(X, T, \alpha)$ among all 1-step Markov measures.

Conj. 6: On the 2-shift there is a 1-step Markov measure that is *fully supported* and is a local maximum point for $\mathrm{Int}_{\mu}(X, T, \alpha)$ among all 1-step Markov measures.



- The conjectures extend to arbitrary shifts of finite type and other dynamical systems.

- The conjectures extend to arbitrary shifts of finite type and other dynamical systems.
- We do not know whether there is a variational principle $\sup_{\mu} \mathrm{Asc}_{\mu}(X, T, \alpha) = \mathrm{Asc}_{\mathrm{top}}(X, T)$.

- The conjectures extend to arbitrary shifts of finite type and other dynamical systems.
- We do not know whether there is a variational principle $\sup_{\mu} \mathrm{Asc}_{\mu}(X, T, \alpha) = \mathrm{Asc}_{\mathrm{top}}(X, T)$.
- Analogous definitions, results, and conjectures exist when entropy is generalized to pressure, by including a potential function which measures the energy or cost associated with each configuration.

- The conjectures extend to arbitrary shifts of finite type and other dynamical systems.
- We do not know whether there is a variational principle $\sup_{\mu} \mathrm{Asc}_{\mu}(X, T, \alpha) = \mathrm{Asc}_{\mathrm{top}}(X, T)$.
- Analogous definitions, results, and conjectures exist when entropy is generalized to pressure, by including a potential function which measures the energy or cost associated with each configuration.
- First one can consider a function of just a single coordinate that gives the value of each symbol.

- The conjectures extend to arbitrary shifts of finite type and other dynamical systems.
- We do not know whether there is a variational principle $\sup_\mu \mathrm{Asc}_\mu(X, T, \alpha) = \mathrm{Asc}_{\mathrm{top}}(X, T)$.
- Analogous definitions, results, and conjectures exist when entropy is generalized to pressure, by including a potential function which measures the energy or cost associated with each configuration.
- First one can consider a function of just a single coordinate that gives the value of each symbol.
- Maximum intricacy may be useful for finding areas of high information activity, such as working regions in a brain (Edelman-Sporns-Tononi) or coding regions in genetic material (Koslicki-Thompson).

The end

The end (of this talk)

The end (of this talk)
(and series).