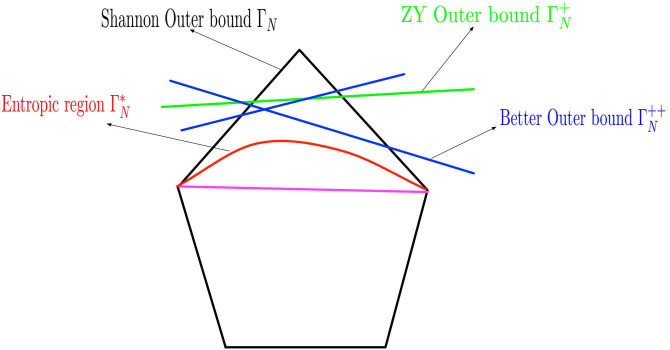
Non-Shannon Information Inequalities in Four Random Variables

Yunshu Liu

ASPITRG Research Group

2013-03-13

Questions



Shannon outer Bound:
Shannon Type Inequalities

Better outer Bound:
Shannon Type + ZY-Non Shannon
Shannon Type + other Non Shannon?

References:

[1]. R. Dougherty, C. Freiling, K. Zeger, Non-Shannon Information Inequalities in Four Random Variables, http://arxiv.org/abs/1104.3602v1

[2]. Weidong Xu, Jia Wang, Jun Sun, A projection method for derivation of non-Shannon-type information inequalities, IEEE Int. Symp. Information Theory (ISIT), 2008, page 2116-2120

[3]. F. Matúš and M. Studený, Conditional Independences among Four Random Variables I, Combinatorics, Probability and Computing, 1995, page 269-278.

Outline

Preliminaries on Information Inequalities

D-copy and Projection Methods

The Region of Entropic Vectors Γ_4^* and 4-atom conjecture

Outline

- Preliminaries on Information Inequalities
- D-copy and Projection Methods
- The Region of Entropic Vectors Γ_4^* and 4-atom conjecture

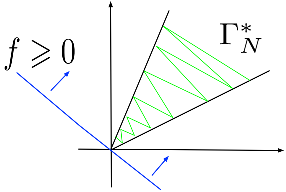
Information Inequalities

Information Inequalities: Definition

If we consider the function of all $(2^n - 1)$ joint entropies associated with $\mathbf{X}$:

$$f = \sum_i \lambda_{\mathcal{A}_i} h_{\mathcal{A}_i} \geqslant 0$$

where $\lambda_{\mathcal{A}_i}$ are real coefficients. If a inequality of this form is true for any random variables $\{X_1, \cdots X_n\}$, we call it an Information inequality.



Information Inequalities

Shannon-Type Information inequality

A Shannon-Type Information inequality is any information inequality of the form

$$\sum_i \alpha_i I(A_i; B_i | C_i) \geq 0$$

where each $\alpha_i \geq 0$.
Shannon-Type Information inequalities form a region Γ_N in $\mathbb{R}^{\mathcal{P}(N)}$

$$\Gamma_N = \left\{ \boldsymbol{h} \in \mathbb{R}^{\mathcal{P}(N)} \;\middle|\; \begin{array}{c} h_{\mathcal{A}} + h_{\mathcal{B}} \geq h_{\mathcal{A}\cap\mathcal{B}} + h_{\mathcal{A}\cup\mathcal{B}} \; \forall \mathcal{A}, \mathcal{B} \subseteq \mathcal{N} \\ h_{\mathcal{P}} \geq h_{\mathcal{Q}} \geq 0 \quad \forall \mathcal{Q} \subseteq \mathcal{P} \subseteq \mathcal{N} \end{array} \right\}$$

Information Inequalities

Phrase $(i, j|K)$

Let $N = \{1, 2, \cdots n\}$ and $\mathcal{S}$ be the family of all couples $(i, j|K)$, where $K \subset N$ and ij is the union of two singletons i and j of $N - K$.

Conditional Independence Relations

$$I(X_i; X_j|\mathbf{X_K}) = 0$$

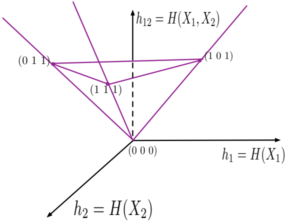
$$I(X_i; X_j|\mathbf{X_K}) = I(X_i; X_j) \qquad \text{If K = } \emptyset$$
$$I(X_i; X_j|\mathbf{X_K}) = H(X_i|\mathbf{X_K}) \qquad \text{If i = j}$$

Information Inequalities

Relationship between Γ_n^* and Γ_n

$\Gamma_2^* = \Gamma_2$ and $\bar{\Gamma}_3^* = \Gamma_3$:
All the Information inequalities on $N \leqslant 3$ variables are Shannon-Type.

For N = 2, Use $h_1 + h_2 \geqslant h_{12}$, $h_{12} \geqslant h_1$ and $h_{12} \geqslant h_2$, we get



Non-Shannon-Type Information inequalities for $N \geqslant 4$

Zhang-Yeung Inequality: The first Non-Shannon-Type Information inequality

$$2I(X_1;X_2) \leq I(X_3;X_4)+I(X_3;X_1,X_2)+3I(X_1;X_2|X_3)+I(X_1;X_2|X_4)$$

Ingleton Inequality

$$\begin{aligned} & Ingleton_{12} \\ &= I(X_1;X_2|X_3) + I(X_1;X_2|X_4) + I(X_3;X_4|\emptyset) - I(X_1;X_2|\emptyset) \end{aligned}$$

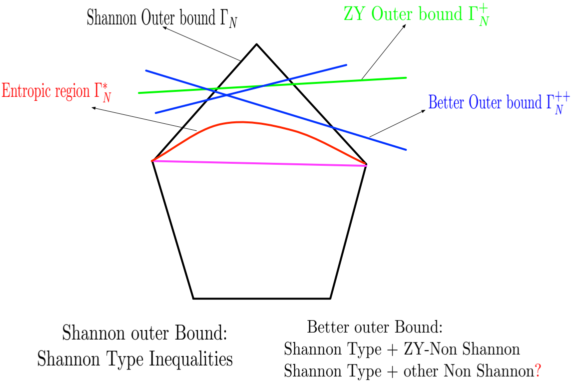
Rewrite Zhang-Yeung Inequality

$$Ingleton_{12} + I(X_1;X_2|X_3) + I(X_1;X_3|X_2) + I(X_2;X_3|X_1) \geqslant 0$$

Information Inequalities

For $N \geq 4$ we have $\Gamma_N \neq \bar{\Gamma}^*_N$

Non-Shannon-Type Information inequalities exist for $N \geqslant 4$



Outline

- Preliminaries on Information Inequalities
- D-copy and Projection Methods
- The Region of Entropic Vectors Γ_4^* and 4-atom conjecture

D-copy and Projection Methods

D-copy

Let A, B, C, D be jointly distributed random variables. Then there is another random variable E, jointly distributed with A, B, C, D with the following properties:

- The marginal distributions of (A, B, C) and (A, B, E) are the same with E replacing C.
- $I(CD; E|AB) = 0$

In this case we say that E is a D-copy of C over (A, B)

D-copy and Projection Methods

Generalized Copy Lemma

Given an m-dimension random vector $x = \{x_1, \cdots, x_m\}$, there exists a random vector $\mathbf{x}_{\hat{\mathbf{K}}}$ such that

- The marginal distributions of $(\mathbf{x}_{\hat{\mathbf{K}}}, \mathbf{x}_{\mathbf{I}})$ and $(\mathbf{x}_{\mathbf{K}}, \mathbf{x}_{\mathbf{I}})$ are the same with $\mathbf{x}_{\hat{\mathbf{K}}}$ replacing $\mathbf{x}_{\mathbf{K}}$.
- Markov chain $\mathbf{x}_{\{\mathbf{K}\}\cup\mathbf{J}} \leftrightarrow \mathbf{x}_{\mathbf{I}} \leftrightarrow \mathbf{x}_{\hat{\mathbf{K}}}$

where K, I, J are disjoint subsets of M. In this case we say that $\mathbf{x}_{\hat{\mathbf{K}}}$ is a $\mathbf{x}_{\mathbf{J}}$-copy of $\mathbf{x}_{\mathbf{K}}$ over $\mathbf{x}_{\mathbf{I}}$.

D-copy and Projection Methods

Example of different ways to generate copy variables

For $x = \{x_1, x_2, x_3, x_4\}$, we want to generate $y = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$.

- One step: generate x_5, x_6 and x_7 together
- Two steps: generate x_5, x_6 together then generate x_7(include permutations)
- Three steps: generate x_5, x_6 and x_7 separately

D-copy and Projection Methods

Constructing an n-dimension random vector $y = \{x_1, \cdots, x_m, x_{m+1}, \cdots, x_n\}$ from m-dimension random vector $x = \{x_1, \cdots, x_m\}$

Start with $x = \{x_1, \cdots, x_m\}$, choose a partition of the new variables $x_{m+1}, \cdots, x_n$, add auxiliary random variables for each partition according to copy lemma for some parameters K, I and J, then we will get a random vector $y = \{x_1, \cdots, x_m, x_{m+1}, \cdots, x_n\}$.

D-copy and Projection Methods

Projection down to initial m variables from higher dimension N

Let Γ_N be the shannon outer bound for N random variables, denote C_N the set of all additional linear inequalities required by the copy lemma, denote $\pi_{X^{\mathcal{P}(M)}}(H_N)$ be the projection of a polyhedral cone H_N onto H_M which is related to the initial m variables, then we have:

$$\Gamma_M^* \subseteq \pi_{X^{\mathcal{P}(M)}}(\Gamma_N \cap C_N) \subseteq \Gamma_M$$

If $\pi_{X^{\mathcal{P}(M)}}(\Gamma_N \cap C_N) \subsetneq \Gamma_M$, we get a better outer bound than Γ_M.

Iterative way to generate better outer bound

If $\pi_{X^{\mathcal{P}(M)}}(\Gamma_N \cap C_N) \subsetneq \Gamma_M$, let $\Gamma_M^+ = \pi_{X^{\mathcal{P}(M)}}(\Gamma_N \cap C_N)$, then we get:

$$\Gamma_M^* \subseteq \pi_{X^{\mathcal{P}(M)}}(\Gamma_N^+ \cap C_N) \subseteq \Gamma_M^+$$

D-copy and Projection Methods

Projection Methods for 4-variables Non-Shannon Information Inequalities

Zhang-Yeung Inequality(One copy variable with one copy step):

$$Ingleton_{12} + I(X_1; X_2|X_3) + I(X_1; X_3|X_2) + I(X_2; X_3|X_1) \geqslant 0$$

DFZ Inequality(Two copy variables with at most two copy steps):

$$2Ingleton_{12} + 3I(X_1; X_2|X_3) + 3I(X_1; X_3|X_2) + I(X_2; X_3|X_1) \geqslant 0$$

D-copy and Projection Methods

Non-Shannon Information Inequalities use at most 4 copy variables with at most 3 copy steps

$$\begin{aligned}
&Ingleton_{12} + a * I(X_1; X_2|X_3) + b * I(X_1; X_3|X_2) \\
&+c * I(X_2; X_3|X_1) + d * I(X_1; X_2|X_4) + e * I(X_1; X_4|X_2) \\
&+f * I(X_2; X_4|X_1) + g * I(X_3; X_4|X_1) \geqslant 0
\end{aligned}$$

Computational wall

Calculation for 4 copy variables with at most 3 copy steps took an estimated 50-100 CPU-years to complete.

Outline

- Preliminaries on Information Inequalities
- D-copy and Projection Methods
- The Region of Entropic Vectors Γ_4^* and 4-atom conjecture

The Region of Entropic Vectors Γ_4^*

Shannon outer bound Γ_4: 28 faces(H-representation) and 41 vertices(V-representation)

$$\Gamma_N = \left\{ \boldsymbol{h} \in \mathbb{R}^{\mathcal{P}(N)} \;\middle|\; \begin{array}{c} h_{\mathcal{A}} + h_{\mathcal{B}} \geq h_{\mathcal{A}\cap\mathcal{B}} + h_{\mathcal{A}\cup\mathcal{B}} \; \forall \mathcal{A}, \mathcal{B} \subseteq \mathcal{N} \\ h_{\mathcal{P}} \geq h_{\mathcal{Q}} \geq 0 \quad \forall \mathcal{Q} \subseteq \mathcal{P} \subseteq \mathcal{N} \end{array} \right\}$$

H-representation of Γ_4

$I(A;B) = 0$ include 6 permutations

$I(A;B|C) = 0$ include 12 permutations

$I(A;B|CD) = 0$ include 6 permutations

$H(A|BCD) = 0$ include 4 permutations

The Region of Entropic Vectors Γ_4^*

$\Gamma_4 \backslash R_4$:the gap between Γ_4 and R_4

Six permutations of Ingleton Inequalities corresponding to six pyramid-like gap between Γ_4 and R_4

H-representation of $\Gamma_4 \backslash R_4$

$$I(A;B|C) \geqslant 0 \quad I(A;C|B) \geqslant 0 \quad I(B;C|A) \geqslant 0 \quad I(C;D|A) \geqslant 0$$
$$I(A;B|D) \geqslant 0 \quad I(A;D|B) \geqslant 0 \quad I(B;D|A) \geqslant 0 \quad I(C;D|B) \geqslant 0$$
$$H(A|BCD) \geqslant 0 \quad H(B|ACD) \geqslant 0 \quad H(C|ABD) \geqslant 0$$
$$H(D|ABC) \geqslant 0 \quad I(C;D) \geqslant 0 \quad I(A;B|CD) \geqslant 0 \quad Ingleton_{12} \leqslant 0$$

14 = 1+8+1+4

The Region of Entropic Vectors Γ_4^*

V-representation of $\Gamma_4 \backslash R_4$

For $N = \{1, 2, 3, 4\}$, with $I \subseteq N$ and $0 \leq t \leq |N \backslash I|$, define

$$r_t^I(J) = \min\{t, |J \backslash I|\} \textit{ with } J \subseteq N$$

$$g_i^{(2)}(J) = \begin{cases} 2 & \text{if } J = i \\ min\{2, |J|\} & \text{if } J \neq i \end{cases}$$

$$g_i^{(3)}(J) = \begin{cases} |J| & \text{if } i \notin J \\ min\{3, |J| + 1\} & \text{if } i \in J \end{cases}$$

$$f_{ij}(K) = \begin{cases} 3 & \text{if } K \in \{ik, jk, il, jl, kl\} \\ min\{4, 2|K|\} & \text{otherwise} \end{cases}$$

The Region of Entropic Vectors Γ_4^*

V-representation of $\Gamma_4 \backslash R_4$

The V-representation of $\Gamma_4 \backslash R_4$ are generated by the 15 linearly independent functions f_{ij}, r_1^{ijk}, r_1^{ijl}, r_1^{ikl}, r_1^{jkl}, $r_1^{\emptyset}$, $r_3^{\emptyset}$, r_1^{i}, r_1^{j}, r_1^{ik}, r_1^{jk}, r_1^{il}, r_1^{jl}, r_2^{k}, r_2^{l}.

Notes:

No point in Γ_4 can simultaneously fail two Ingleton inequalities, thus the six ”pyramid” have disjoint interiors.

The Region of Entropic Vectors Γ_4^*

Projection results using D-copy

- No copy: 28 faces and 41 vertices
- One copy variable: 40 faces and 89 vertices
- Two copy variables: 160 faces and 299 vertices
- Three copy variables: 796 faces and 10361 vertices
- Four copy variables with at most three copy steps: 4924 faces and 224801 vertices

The Region of Entropic Vectors Γ_4^*

The percent of pyramid left(in volume)

- No copy: 100%
- One copy variable: 98.4568%
- Two copy variables: 97.7040%
- Three copy variables: 96.7214%
- Four copy variables with at most three copy steps: 96.4682%
- $\Gamma_4^* \backslash R_4$ is at least 53.4815%

Four-atom conjecture

Ingleton score

Given a probability distribution, we define the Ingleton score of the distribution to be

$$\begin{aligned} Ingleton \quad score &= \frac{Ingleton_{12}}{H(ABCD)} \\ &= \frac{I(A;B|C) + I(A;B|D) + I(C;D) - I(A;B)}{H(ABCD)} \end{aligned}$$

Ingleton score determine how much a distribution violate one of the six Ingleton Inequality.

Four-atom conjecture

Lowest possible Ingleton score

The lowest possible Ingleton score is approximately -0.0894. It is attained by a four-variable binary distribution, given by

$$P(0,0,0,0) = P(1,1,1,1) \approx 0.35$$
$$P(0,1,0,1) = P(0,1,1,0) \approx 0.15$$

Four-atom conjecture

Ingleton score of other 4-atom distributions

C and D independent, A = CD and B =(1-C)(1-D)

$P(0,0,0,1) = P(0,0,1,0) = P(0,1,0,0) = P(1,0,1,1) = 1/4$

Ingleton score $= -0.0613$

C and D independent, A = C(1-D) and B =(1-C)D

$P(0,0,0,0) = P(0,0,1,1) = P(0,1,0,1) = P(1,0,1,0) = 1/4$

Ingleton score $= -0.0613$

Four-atom conjecture

Minimum Ingleton score of Outer bound

- No copy: -0.2500
- One copy variable: -0.1667
- Two copy variables: -0.1667
- Three copy variables: -0.1590
- Four copy variables with at most three copy steps: -0.1579
- compare to -0.0894

Thanks!

Questions!