Non-Shannon Information Inequalities in Four Random Variables

Yunshu Liu

ASPIRG 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:
Outline

Preliminaries on Information Inequalities

D-copy and Projection Methods

The Region of Entropic Vectors $\Gamma_4^*$ and 4-atom conjecture
Outline

- Preliminaries on Information Inequalities
- D-copy and Projection Methods
- The Region of Entropic Vectors $\Gamma_4^*$ and 4-atom conjecture
Information Inequalities

Information Inequalities: Definition

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

\[
f = \sum_i \lambda_{A_i} h_{A_i} \geq 0
\]

where \(\lambda_{A_i}\) 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 \( \Gamma_N \) in \( \mathbb{R}^\mathcal{P}(N) \)

$$\Gamma_N = \left\{ h \in \mathbb{R}^\mathcal{P}(N) \mid h_A + h_B \geq h_{A \cap B} + h_{A \cup B} \quad \forall A, B \subseteq N \\
\quad h_P \geq h_Q \geq 0 \quad \forall Q \subseteq P \subseteq N \right\}$$
Information Inequalities

Phrase \((i, j|K)\)
Let \(N = \{1, 2, \cdots, n\}\) and \(S\) 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

\[
I(X_i; X_j|X_K) = 0
\]

\[
I(X_i; X_j|X_K) = I(X_i; X_j) \quad \text{if } K = \emptyset
\]

\[
I(X_i; X_j|X_K) = H(X_i|X_K) \quad \text{if } i = j
\]
Information Inequalities

Relationship between $\Gamma_n^*$ and $\Gamma_n$

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

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

\[
\begin{align*}
    h_1 &= H(X_1) \\
    h_2 &= H(X_2) \\
    h_{12} &= H(X_1, X_2)
\end{align*}
\]
Non-Shannon-Type Information inequalities for $N \geq 4$

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

$$Ingleton_{12} = I(X_1; X_2|X_3) + I(X_1; X_2|X_4) + I(X_3; X_4|∅) - I(X_1; X_2|∅)$$

Rewrite Zhang-Yeung Inequality

$$Ingleton_{12} + I(X_1; X_2|X_3) + I(X_1; X_3|X_2) + I(X_2; X_3|X_1) \geq 0$$
For $N \geq 4$ we have $\Gamma_N \neq \Gamma_N^*$

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

- Preliminaries on Information Inequalities
- D-copy and Projection Methods
- The Region of Entropic Vectors $\Gamma_4^*$ and 4-atom conjecture
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)
Generalized Copy Lemma
Given an m-dimension random vector $x = \{x_1, \cdots, x_m\}$, there exists a random vector $x^\hat{K}$ such that

- The marginal distributions of $(x^\hat{K}, x_I)$ and $(x_K, x_I)$ are the same with $x^\hat{K}$ replacing $x_K$.
- Markov chain $x_{\{K\} \cup J} \leftrightarrow x_I \leftrightarrow x^\hat{K}$

where $K$, $I$, $J$ are disjoint subsets of $M$. In this case we say that $x^\hat{K}$ is a $x_J$-copy of $x_K$ over $x_I$. 
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
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 $\Gamma_N$ 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 $\pi_{xP(M)}(H_N)$ be the projection of a polyhedral cone $H_N$ onto $H_M$ which is related to the initial m variables, then we have:

$$\Gamma^*_M \subseteq \pi_{xP(M)}(\Gamma_N \cap C_N) \subseteq \Gamma_M$$

If $\pi_{xP(M)}(\Gamma_N \cap C_N) \subsetneq \Gamma_M$, we get a better outer bound than $\Gamma_M$.

Iterative way to generate better outer bound

If $\pi_{xP(M)}(\Gamma_N \cap C_N) \subsetneq \Gamma_M$, let $\Gamma^+_M = \pi_{xP(M)}(\Gamma_N \cap C_N)$, then we get:

$$\Gamma^*_M \subseteq \pi_{xP(M)}(\Gamma^+_N \cap C_N) \subseteq \Gamma^+_M$$
D-copy and Projection Methods

Projection Methods for 4-variables Non-Shannon Information Inequalities

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

\[ \text{Ingleton}_{12} + I(X_1; X_2|X_3) + I(X_1; X_3|X_2) + I(X_2; X_3|X_1) \geq 0 \]

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

\[ 2\text{Ingleton}_{12} + 3I(X_1; X_2|X_3) + 3I(X_1; X_3|X_2) + I(X_2; X_3|X_1) \geq 0 \]
D-copy and Projection Methods

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

\[
\text{Ingleton}_{12} + a \times I(X_1; X_2|X_3) + b \times I(X_1; X_3|X_2) \\
+ c \times I(X_2; X_3|X_1) + d \times I(X_1; X_2|X_4) + e \times I(X_1; X_4|X_2) \\
+ f \times I(X_2; X_4|X_1) + g \times I(X_3; X_4|X_1) \geq 0
\]

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 $\Gamma_4^*$ and 4-atom conjecture
The Region of Entropic Vectors $\Gamma_4^*$

Shannon outer bound $\Gamma_4$: 28 faces($H$-representation) and 41 vertices($V$-representation)

$$\Gamma_N = \left\{ h \in \mathbb{R}^{\mathcal{P}(N)} \mid \begin{array}{l} h_A + h_B \geq h_{A \cap B} + h_{A \cup B} \quad \forall A, B \subseteq N \\ h_P \geq h_Q \geq 0 \quad \forall Q \subseteq P \subseteq N \end{array} \right\}$$

$H$-representation of $\Gamma_4$

$I(A; B) = 0$ 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 $\Gamma^*_4$

$\Gamma_4 \setminus R_4$: the gap between $\Gamma_4$ and $R_4$

Six permutations of Ingleton Inequalities corresponding to six pyramid-like gap between $\Gamma_4$ and $R_4$

H-representation of $\Gamma_4 \setminus R_4$

\[
\begin{align*}
I(A; B|C) &\geq 0 & I(A; C|B) &\geq 0 & I(B; C|A) &\geq 0 & I(C; D|A) &\geq 0 \\
I(A; B|D) &\geq 0 & I(A; D|B) &\geq 0 & I(B; D|A) &\geq 0 & I(C; D|B) &\geq 0 \\
H(A|BCD) &\geq 0 & H(B|ACD) &\geq 0 & H(C|ABD) &\geq 0 \\
H(D|ABC) &\geq 0 & I(C; D) &\geq 0 & I(A; B|CD) &\geq 0 & \text{Ingleton}_{12} &\leq 0
\end{align*}
\]

$14 = 1+8+1+4$
The Region of Entropic Vectors $\Gamma_4^*$

V-representation of $\Gamma_4 \setminus R_4$

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

$$r^I_t(J) = \min\{t, |J \setminus I|\} \quad \text{with } J \subseteq N$$

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

$$g_i^{(3)}(J) = \begin{cases} |J| & \text{if } i \not\in J \\ \min\{3, |J| + 1\} & \text{if } i \in J \end{cases}$$

$$f_{ij}(K) = \begin{cases} 3 & \text{if } K \in \{ik, jk, il, jl, kl\} \\ \min\{4, 2|K|\} & \text{otherwise} \end{cases}$$
The Region of Entropic Vectors $\Gamma_4^*$

**V-representation of $\Gamma_4 \setminus R_4$**

The V-representation of $\Gamma_4 \setminus R_4$ are generated by the 15 linearly independent functions $f_{ij}, r_{ijk}^1, r_{ijl}^1, r_{ikl}^1, r_{jkl}^1, r_{\emptyset}^1, r_{ij}^3, r_{ij}^1, r_{ik}^1, r_{jk}^1, r_{il}^1, r_{jl}^1, r_{kl}^2, r_{kl}^1$.  

**Notes:**

No point in $\Gamma_4$ can simultaneously fail two Ingleton inequalities, thus the six ”pyramid” have disjoint interiors.
The Region of Entropic Vectors $\Gamma^*_4$

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 $\Gamma_4^*$

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%
- $\Gamma_4^* \setminus R_4$ is at least 53.4815%
Ingleton score

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

\[
\text{Ingleton score} = \frac{\text{Ingleton}_{12}}{H(ABCD)} = \frac{I(AB|C) + I(AB|D) + I(CD) - I(AB)}{H(ABCD)}
\]

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
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!