# Inequalities for Shannon Entropy, Kolmogorov Complexity and Linear Ranks of Subspaces

Congduan Li

ASPITRG
Drexel University

February 1, 2013

# Outline

1. Kolmogorov complexity and Shannon inequalities
2. Inequalities for ranks of subspaces
3. Ingleton inequalities
4. Results for $N = 4$
5. Relationships
6. For $N \geq 5$

# Shannon inequalities

For random variables $A, B, C$, we have

$$\begin{aligned} H(A,B) &= H(B) + H(A|B) \\ I(A;B) &= H(B) - H(B|A) \\ I(A;B) &= H(A) + H(B) - H(A,B) \\ I(A;B|C) &= H(A|C) + H(B|C) - H(A,B|C). \end{aligned}$$

General form:

$$I(A;B|C) \geq 0.$$

# Kolmogorov complexity

- The Kolmogorov complexity $K(a)$ of a binary string $a$ is defined as the minimal length of a program that generates $a$.
- Conditional complexity $K(a|b)$: minimal length of a program that produces $a$ having $b$ as input
- $K(a|b) = K(a, b) - k(a)$
- $I(a; b) = K(b) - k(b|a) = K(a) + K(b) - K(a, b)$
- $I(a; b|c) = K(a|c) + K(b|c) - K(a, b|c)$
- General: $I(a, b|c) \geq 0$.

# Relation between Shannon and Kolmogorov

**Theorem**

*Any linear inequality that is true for Kolmogorov complexity is also true for Shannon entropy, and vice versa.*

# $\bar{\Gamma}_N^*$ and $\Gamma_N$

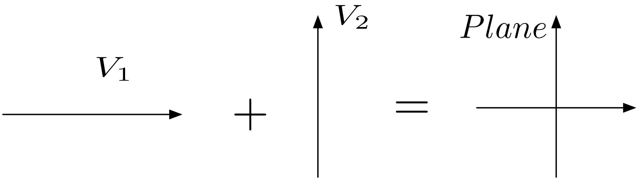
- Let $(X_1, \ldots, X_N)$ be an arbitrary collection of discrete random variables. To each of the $2^N - 1$ non-empty subsets of the collection of random variables, $X_{\mathcal{A}} := (X_i \,|\, i \in \mathcal{A})$ with $\mathcal{A} \subseteq \{1, \ldots, N\}$, there is associated a joint Shannon entropy $H(X_{\mathcal{A}})$. Stacking these subset entropies for different subsets into a $2^N - 1$ dimensional vector we form an entropic vector

$$\mathbf{h} = [H(X_{\mathcal{A}}) \,|\mathcal{A} \subseteq \{1, \ldots, N\}, \mathcal{A} \neq \emptyset] \tag{1}$$

- Region of entropic vectors $\bar{\Gamma}_N^*$ is a convex cone.
- Convex cone $\Gamma_N$ obtained by basic inequalities $I(A; B|C) \geq 0, \forall A, B, C \subseteq [N]$ is an outer bound on $\bar{\Gamma}_N^*$, while $\bar{\Gamma}_N^* = \Gamma_N$ when $N = 1, 2, 3$.

# Ranks of subspaces

- A subspace arrangement is a collection of subspaces $V = V_1, ..., V_n$ of some finite dimensional vector space.
- Define $rk_V(A) = \dim(\sum_{i \in A} V_i)$
- Representable in various field. Unqualified case, representable in some field. This is the case considered here.



# Mapping

- Map to entropic vector region: Suppose the space has rank $k$ with field $\mathbb{F}$. Define auxilary random variables $u_1, \ldots u_k$ uniformly distributed over $\mathbb{F}$. Then define random varibles $X_1, \ldots X_N = [u_1, \ldots u_k] \times [V_1 \ldots V_N]$ corresponding to the vector subspaces, then $X_1, \ldots X_N$ have entropy the same as ranks of the corresponding vector subspaces, if the entropy is taken using $\log_{|\mathbb{F}|}(\cdot)$.
- Conic hull of ranks of subspaces forms an inner bound: $\Gamma_N^{space} \subseteq \bar{\Gamma}_N^*$.

**Theorem**

*Any linear inequlity implied by basic inequality is valid for ranks of linear subspaces.* $\Gamma_N^{space} \subseteq \Gamma_N$

Further more,

**Theorem**

*Any linear inequality valid for Shannon entropy is valid for ranks (dimensions) in any linear space.* $\Gamma_N^{space} \subseteq \bar{\Gamma}_N^*$.

Up to $N = 3$, we have

**Theorem**

*For n=1, 2, 3 any inequality valid for ranks (dimensions) is a consequence (linear combination with nonnegative coefficients) of basic inequalities.*

# Ingleton's inequality

Ingleton established an necessary condition for a matroid with ground set S and rank function r to be representable over a field: for any subsets $A, B, C, D$ of $S$ there must hold:

$$
\begin{aligned}
r(A) + r(B) + r(C \cup D) + r(A \cup B \cup C) + r(A \cup B \cup D) \quad &\leq \\
r(A \cup B) + r(A \cup C) + r(A \cup D) + r(B \cup C) + r(B \cup D)
\end{aligned}
$$

Rewritten as:

$$I(A; B) \leq I(A; B|C) + I(A; B|D) + I(C; D).$$

# Counter example

Ingleton's inequality is not always true for Shannon entropy.

**Theorem**

*There exist four random variables $A, B, C, D$ such that*

$$\begin{aligned} I(A;B) &> 0 \\ I(A;B|C) &= 0 \\ I(A;B|D) &= 0 \\ I(C;D) &= 0. \end{aligned}$$

Let $C, D$ independent uniformly distributed on $\{0, 1\}$, $A = C(1 - D), B = D(1 - C)$. Given $C$ or $D$, $A$ or $B$ is 0, so $A, B$ independent. However, $A, B$ are not unconditionally independent, since they cannot be equal to 1 sumultaneously.

# Case $N = 4$

The results we know now for $N = 4$ are:

**Theorem**

*For $N = 4$, all inequalities that are valid for ranks are implied by basic inequalities and Ingleton inequalities.*

For $N = 4$, there are 41 extrme rays in $\Gamma_4$, 35 are in $\Gamma_4^{space}$ (also Ingleton inner bound) and 6 violate Ingleton.
For those obey Ingleton's inequalities, 27 are ranks of some representable matroids (including $U_{2,4}$ which is an exclusion minor for binary representale matroids, but can be representable on ternay field). The other 8 are projections (pair up 8 variables to get 4 variables, and project on these 4 variables) of representable matroids.

Those Ingleton violators are represented by Vamos matroid if we project Vamos matroid ranks to 4 paired random variables. So, we have

**Theorem**

*For $N = 4$, all inequalities that are valid for ranks in arbitrary matroids (including projections of matroids) are consequences of basic inequalities.*

# One way to prove Ingleton

**Definition**

We call a random variable $C$ common information for $A, B$ if

$$
\begin{aligned}
H(C|A) &= 0 \\
H(C|B) &= 0 \\
H(C) &= I(A;B)
\end{aligned}
$$

Will be also important to prove some other inequalities for $N \geq 5$.

# One Shannon inequality

**Theorem**

*For any random variables* $A, B, C, D, E$

$$H(E) \leq 2H(E|A) + 2H(E|B) + I(A;B|C) + I(A;B|D) + I(C;D).$$

This is implied by basic inequlities, so is Shannon type. However, if we assume $E$ is a common information between $A, B$ and apply to the inequality, we get Ingleton:

$$I(A;B) \leq I(A;B|C) + I(A;B|D) + I(C;D).$$

(For ranks of linear subspaces, common information always exists, so Ingleton always is true for ranks of subspaces.)

# Relationships so far

In general, basic inequalities valid for Shannon, and Shannon valid for ranks of linear spaces.

In general, $\Gamma_N^{space} \subseteq \bar{\Gamma}_N^* \subseteq \Gamma_N$.

For at most 3, they are equal. For $N \geq 4$, they subset relationships are strict, since we have already known Ingleton inequality, Non-Shannon type inequalities.

Also, we know Shannon + Ingleton fully characterize $\Gamma_4^{space}$. Now the question is if we need more extra inequalities to characterize $\Gamma_N^{space}$ for $N \geq 5$. The answer is YES.

On $N = 5$, DFZ gave 24 new inequalities together with Shannon and Ingleton inequalities to fully characterize $\Gamma_5^{space}$.
For $N \geq 6$, DFZ and Kinser showed (independently) a general form to generate new inequalities valid for linear spaces but not for Shannon entropy.

# To prove new inequalities for $N = 5$

Basic idea to replace random variable(s) with common information of some other variables in Shannon inequalities.

**Fact**

*The inequality $H(Z|R) + I(R;S|T) \geq I(Z;S|T)$ is a Shannon inequality.*

**Proof.**

Use Shannon inequalities, we have

$$\begin{aligned} H(Z|R) + H(S|Z,T) &\geq H(Z|R,T) + H(S|Z,T) \\ &\geq I(S;Z|R,T) + H(S|Z,T) \\ &\geq I(S;Z|R,T) + H(S|R,Z,T) \\ &= H(S|R,T) \end{aligned}$$

Add $H(S|T)$ to both sides and reshape we get the desired one. □

# Proof idea cont.

**Corollary**

*If $H(Z|R) = 0$, then $I(R; S|T) \geq I(Z; S|T)$.*

**Proof of Ingleton inequality**

Let $Z$ be a common information of $A$ and $B$, so that $H(Z|A) = H(Z|B) = 0$ and $H(Z) = I(A; B)$. Then

$$\begin{aligned}
& I(A; B|C) + I(A; B|D) + I(C; D) \\
\geq\ & I(Z; B|C) + I(Z; B|D) + I(C; D) \\
\geq\ & I(Z; Z|C) + I(Z; Z|D) + I(C; D) \\
=\ & H(Z|C) + H(Z|D) + I(C; D) \\
\geq\ & H(Z|C) + I(Z; C) \\
\geq\ & I(Z; Z) \\
=\ & H(Z) = I(A; B)
\end{aligned}$$

# Proof examples

Inequality (1):
$I(A;B) \leq I(A;B|C) + I(A;B|D) + I(C;D|E) + I(A;E)$

## Proof of inequality (1)

Let $Z$ be a common information of $A$ and $B$, so that $H(Z|A) = H(Z|B) = 0$ and $H(Z) = I(A;B)$. Then

$$
\begin{aligned}
& I(A;B|C) + I(A;B|D) + I(C;D) \\
\geq\ & I(Z;B|C) + I(Z;B|D) + I(C;D) \\
\geq\ & I(Z;Z|C) + I(Z;Z|D) + I(C;D) \\
=\ & H(Z|C) + H(Z|D) + I(C;D) \\
\geq\ & H(Z|C) + I(Z;C) \\
\geq\ & I(Z;Z) \\
=\ & H(Z) = I(A;B)
\end{aligned}
$$

# Proof examples

Inequality (2):
$I(A;B) \leq I(A;B|C) + I(A;C|D) + I(A;D|E) + I(B;E)$

## Proof of inequality (2)

Let $Z$ be a common information of $A$ and $B$, so that $H(Z|A) = H(Z|B) = 0$ and $H(Z) = I(A;B)$. Then

$$\begin{aligned}
& I(A;B|C) + I(A;C|D) + I(A;D|E) + I(B;E) \\
\geq \; & I(Z;Z|C) + I(Z;C|D) + I(Z;D|E) + I(Z;E) \\
\geq \; & I(Z;Z|D) + I(Z;D|D) + I(Z;E) \\
= \; & H(Z|D) + I(Z;D|D) + I(Z;E) \\
\geq \; & I(Z;Z|E) + I(Z;E) \\
= \; & H(Z) = I(A;B)
\end{aligned}$$

Generalize the inequality (2) and follow the proof pattern, we see that if $A_0$ and $B_0$ have a common information, then

$$
\begin{aligned}
I(A_0; B_0) \leq\ & I(A_0; B_0|B_1) \\
& + I(A_0; B_1|B_2) \\
& \dots \\
& + I(A_0; B_{n-1}|B_n) \\
& + I(B_0; B_n)
\end{aligned}
$$

This is essentially the main results in Kinser's paper. He shows that this is irreducible for general $n$, i.e this inequality cannot be implied by those for up to $n-1$.

# Another general new inequality

**Fact**

*The inequality $H(Z|R) + I(R;S|T) \geq I(Z;S|T) + H(Z|R,S,T)$ is a Shannon inequality.*

**Proof.**

Use Shannon inequalities, we have

$$\begin{aligned} & H(Z|R) + H(S|Z,T) \\ \geq\ & H(Z|R,T) + H(S|Z,T) \\ =\ & H(Z|R,S,T) + I(S;Z|R,T) + H(S|Z,T) \\ \geq\ & H(Z|R,S,T) + I(S;Z|R,T) + H(S|R,Z,T) \\ \geq\ & H(Z|R,S,T) + H(S|R,T) \end{aligned}$$

Add $H(S|T)$ to both sides and reshape we get the desired one. □

# Another general new inequality cont.

**Corollary**

$H(Z|R) + H(Z|S) + I(R;S|T) \geq H(Z|T) + H(Z|R,S,T)$

In the case of $T$ is null variable, it becomes

$$H(Z|R) + H(Z|S) + I(R;S) \geq H(Z) + H(Z|R,S)$$

# Proof examples

Inequality (8):
$2I(A;B) \leq I(A;B|C)+I(A;B|D)+I(A;B|E)+I(C;D)+I(C,D;E)$

## Proof of inequality (2)

Let $Z$ be a common information of $A$ and $B$, so that $H(Z|A) = H(Z|B) = 0$ and $H(Z) = I(A;B)$. Then

$$\begin{aligned}
& I(A;B|C) + I(A;B|D) + I(A;B|E) + I(C;D) + I(C,D;E) \\
\geq\ & I(Z;Z|C) + I(Z;Z|D) + I(Z;Z|E) + I(C;D) + I(C,D;E) \\
\geq\ & H(Z) + H(Z|C,D) + H(Z|E) + I(C;D) + I(C,D;E) \\
\geq\ & 2H(Z) = 2I(A;B)
\end{aligned}$$

# Another general inequality

Generalize the inequality (8) and follow the proof pattern, we see that if $A$ and $B$ have a common information, then

$$\begin{aligned}(n-1)I(A;B) \quad \leq \quad & I(A;B|C_1) + I(A;B|C_2) + \ldots I(A;B|C_n) \\ & + I(C_1;C_2) + I(C_1,C_2;C_3) \\ & + ... + I(C_1,\ldots,C_{n-1};C_n)\end{aligned}$$

This is in DFZ's paper. They also show that this is irreducible for general $n$, i.e this inequality cannot be implied by those for up to $n-1$. The proof of irreducibility is similar in both papers: construct random variables that satisfy inequalities for fewer variables but not for this for $n+2$.

# Completeness for $N = 5$

- DFZ show that the new 24 new inequalities + Shannon + Ingleton fully characterize linear subspace ranks, i.e the linear inner bound for region of entropic vectors
- For $N \geq 6$, not fully characterized yet. But know that there will be new inequalities with increase of $N$.