Entropy inequalities and linear rank inequalities

Randall Dougherty
(Center for Communications Research, San Diego)
Chris Freiling
(California State University, San Bernardino)
Ken Zeger
(University of California, San Diego)

Entropy inequalities: inequalities such as

$$H(A,C) + H(B,C) \geq H(A,B,C) + H(C)$$

which are true for joint entropies of any random variables in a finite probability distribution.

Linear rank inequalities: inequalities such as

$$\dim(A+C) + \dim(B+C) \geq \dim(A+B+C) + \dim(C)$$

which are true for joint dimensions of any subspaces of a finite vector space.

Fact [CORRECTED]: Any entropy inequality is a linear rank inequality.

To show this, it suffices to turn any configuration of subspaces of a finite vector space into a configuration of random variables whose joint entropies match the joint dimensions of the subspaces (up to a scalar factor).

Given subspaces $A, B, \ldots$ of vector space V over the field F_q with q elements, let f be chosen uniformly at random from the set of linear functions from V to F_q, and define the random variables $x_A = f|A$, $x_B = f|B$, etc.; then

$$H(x_A) = (\log q) \dim(A)$$
$$H(x_A, x_B) = (\log q) \dim(A + B)$$

and so on.

Shannon inequalities

$$H(\varnothing) = 0$$

$$H(A) \geq 0$$

$$H(A|B) = H(A,B) - H(B) \geq 0$$

$$I(A;B) = H(A) + H(B) - H(A,B) \geq 0$$

$$I(A;B|C) = H(A,C) + H(B,C) - H(C) - H(A,B,C) \geq 0$$

Or any nonnegative linear combination of these.

Shannon inequalities

$$H(\varnothing) = 0$$

$$H(A) \geq 0$$

$$H(A) = \dim A$$

$$H(A|B) = H(A,B) - H(B) \geq 0$$

$$H(A|B) = \dim\left((A+B)/B\right)$$

$$I(A;B) = H(A) + H(B) - H(A,B) \geq 0$$

$$I(A;B) = \dim\left(A \cap B\right)$$

$$I(A;B|C) = H(A,C) + H(B,C) - H(C) - H(A,B,C) \geq 0$$

$$I(A;B|C) = \dim\left(((A+C) \cap (B+C))/C\right)$$

Or any nonnegative linear combination of these.

A vector (assigning a real number to each subset of $\{A, B, \ldots\}$) is *entropic* if it is the list of joint entropies of some random variables over a finite probability distribution.

A vector is *(linearly) representable* if it is the list of joint subspace dimensions (ranks) of some subspaces of a finite vector space.

A vector is a *polymatroid* if it satisfies the Shannon inequalities.

The closure of the set of entropic vectors is a convex cone.

The closure of the set of entropic vectors is a convex cone.

The closure of the set of nonnegative scalar multiples of vectors representable over a fixed finite field is a convex cone.

For up to three variables, all entropy inequalities (and hence all linear rank inequalities) are Shannon.

The Zhang-Yeung entropy inequality (1998)

$$2I(A;B) \leq I(C;D) + I(C;A,B) + 3I(A;B|C) + I(A;B|D)$$

The Zhang-Yeung entropy inequality (1998)

$$2I(A;B) \leq I(C;D) + I(C;A,B) + 3I(A;B|C) + I(A;B|D)$$

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

The Ingleton linear rank inequality (1969)

$$H(A)+H(B)+H(C,D)+H(A,B,C)+H(A,B,D)$$
$$\leq H(A,B)+H(A,C)+H(A,D)+H(B,C)+H(B,D)$$

The Ingleton linear rank inequality (1969)

$$H(A) + H(B) + H(C,D) + H(A,B,C) + H(A,B,D)$$
$$\leq H(A,B) + H(A,C) + H(A,D) + H(B,C) + H(B,D)$$

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

Copying random variables

If A, B, and C are sets of random variables, then one can construct a new set of random variables R such that:

- (A, B) and (R, B) are identically distributed;
- $I(R; AC|B) = 0$.

We say that "R is a C-copy of A over B."

$$p(A = a, B = b, C = c, R = r) = \frac{p(A = a, B = b, C = c)p(A = r, B = b)}{p(B = b)}$$

The Zhang-Yeung inequality

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

is a consequence of “R is a D-copy of C over AB”.

The Zhang-Yeung inequality

$$
\begin{aligned}
I(A;B) \le\ & 2I(A;B|C) + I(A;C|B) + I(B;C|A) + I(A;B|D) + I(C;D) \\
& + 3I(R;CD|AB) - (H(R) - H(C)) + 2(H(RA) - H(CA)) \\
& + 2(H(RB) - H(CB)) - 3(H(RAB) - H(CAB))
\end{aligned}
$$

is a Shannon inequality.

The Zhang-Yeung inequality

$$
\begin{aligned}
&I(C;D|R) + I(C;R|A) + I(C;R|B) + I(C;R|ABD) + I(D;R|A) + \\
&I(D;R|B) + I(D;R|ABC) + I(A;B|CR) + I(A;B|DR) + I(AB;R|CD) + \\
&\quad I(A;B) = 2I(A;B|C) + I(A;C|B) + I(B;C|A) + I(A;B|D) + I(C;D) \\
&\qquad + 3I(R;CD|AB) - (H(R) - H(C)) + 2(H(RA) - H(CA)) \\
&\qquad + 2(H(RB) - H(CB)) - 3(H(RAB) - H(CAB))
\end{aligned}
$$

is an entropy identity.

Inequalities obtained from two copy variables

$$2I(A;B) \le 5I(A;B|C) + 3I(A;C|B) + I(B;C|A) + 2I(A;B|D) + 2I(C;D)$$

$$2I(A;B) \le 4I(A;B|C) + 2I(A;C|B) + I(B;C|A) + 3I(A;B|D) + I(A;D|B) + 2I(C;D)$$

$$2I(A;B) \le 4I(A;B|C) + 4I(A;C|B) + I(B;C|A) + 2I(A;B|D) + I(A;D|B) + I(B;D|A) + 2I(C;D)$$

$$2I(A;B) \le 3I(A;B|C) + 3I(A;C|B) + 3I(B;C|A) + 2I(A;B|D) + 2I(C;D)$$

$$2I(A;B) \le 3I(A;B|C) + 4I(A;C|B) + 2I(B;C|A) + 3I(A;B|D) + I(A;D|B) + 2I(C;D)$$

$$2I(A;B) \le 3I(A;B|C) + 2I(A;C|B) + 2I(B;C|A) + 2I(A;B|D) + I(A;D|B) + I(B;D|A) + 2I(C;D)$$

Common form of inequalities

$$
\begin{aligned}
aI(A;B) \le\ & bI(A;B|C) + cI(A;C|B) + dI(B;C|A) \\
& + eI(A;B|D) + fI(A;D|B) + gI(B;D|A) \\
& + hI(C;D) + iI(C;D|A)
\end{aligned}
$$

Inequalities obtained from three copy variables

$$
\begin{aligned}
2I(A;B) \le\ & 3I(A;B|C) + 3I(A;C|B) + 2I(B;C|A) \\
& + 2I(A;B|D) + 2I(C;D) \\
2I(A;B) \le\ & 5I(A;B|C) + 2I(A;C|B) + I(B;C|A) \\
& + 2I(A;B|D) + 2I(C;D) \\
2I(A;B) \le\ & 4I(A;B|C) + I(A;C|B) + 3I(B;C|A) \\
& + 2I(A;B|D) + 2I(C;D)
\end{aligned}
$$

etc.

Inequalities obtained from three copy variables

34 new inequalities (744 counting permutations)

plus 1 previous inequality (24 counting permutations)

Inequalities obtained from four copy variables

(using at most three copy steps)

203 new inequalities (4632 counting permutations)

plus 11 previous inequalities (264 counting permutations)

Generating new inequalities from known ones

$$2I(A;B) \leq 5I(A;B|C) + 3I(A;C|B) + I(B;C|A) + 2I(A;B|D) + 2I(C;D)$$

follows from "R is a D-copy of A over BC" and instances of the Zhang-Yeung inequality.

$$2I(A;B) \leq 5I(A;B|C) + 3I(A;C|B) + I(B;C|A) + 2I(A;B|D) + 2I(C;D)$$

follows from "R is a D-copy of A over BC" and the following instance of the Zhang-Yeung inequality:

$$\begin{aligned} I(AR;BR) \leq\ & 2I(AR;BR|CR) + I(AR;CR|BR) + I(BR;CR|AR) \\ & + I(AR;BR|D) + I(CR;D) \end{aligned}$$

$$
\begin{aligned}
2I(A;B) \leq\ & 5I(A;B|C) + 3I(A;C|B) + I(B;C|A) + 2I(A;B|D) + 2I(C;D) \\
& + I(AR;BR) - 2I(AR;BR|CR) - I(AR;CR|BR) \\
& - I(BR;CR|AR) - I(AR;BR|D) - I(CR;D)
\end{aligned}
$$

follows from “R is a D-copy of A over BC”.

$$
\begin{aligned}
I(A;B) + I(A;B) \leq\ & 2I(A;B|C) + 2I(A;B|C) + I(A;B|C) \\
& + I(A;C|B) + I(A;C|B) + I(A;C|B) + I(B;C|A) \\
& + I(A;B|D) + I(A;B|D) + I(C;D) + I(C;D) \\
& + I(AR;BR) - 2I(AR;BR|CR) - I(AR;CR|BR) \\
& - I(BR;CR|AR) - I(AR;BR|D) - I(CR;D)
\end{aligned}
$$

follows from “R is a D-copy of A over BC”.

$$
\begin{aligned}
I(A;B) + I(A;B) \leq\ & 2I(A;B|C) + 2I(A;B|C) + I(A;B|C) \\
& + I(A;C|B) + I(A;C|B) + I(A;C|B) + I(B;C|A) \\
& + I(A;B|D) + I(A;B|D) + I(C;D) + I(C;D) \\
& + I(AR;BR) - 2I(AR;BR|CR) - I(AR;CR|BR) \\
& - I(BR;CR|AR) - I(AR;BR|D) - I(CR;D)
\end{aligned}
$$

follows from “R is a D-copy of A over BC”.

$$
\begin{aligned}
aI(A;B) + dI(A;B) \leq\ & bI(A;B|C) + 2dI(A;B|C) + hI(A;B|C) \\
& + aI(A;C|B) + cI(A;C|B) + dI(A;C|B) + dI(B;C|A) \\
& + aI(A;B|D) + dI(A;B|D) + dI(C;D) + hI(C;D) \\
& + aI(AR;BR) - bI(AR;BR|CR) - cI(AR;CR|BR) \\
& - dI(BR;CR|AR) - aI(AR;BR|D) - hI(CR;D)
\end{aligned}
$$

follows from “R is a D-copy of A over BC”.

$$
\begin{aligned}
(a+d)I(A;B) \le\ & (b+2d+h)I(A;B|C) + (a+c+d)I(A;C|B) + dI(B;C|A) \\
& + (a+d)I(A;B|D) + (d+h)I(C;D) \\
& + aI(AR;BR) - bI(AR;BR|CR) - cI(AR;CR|BR) \\
& - dI(BR;CR|AR) - aI(AR;BR|D) - hI(CR;D)
\end{aligned}
$$

follows from “R is a D-copy of A over BC”.

$$(a+d)I(A;B) \leq (b+2d+h)I(A;B|C) + (a+c+d)I(A;C|B) + dI(B;C|A) \\ + (a+d)I(A;B|D) + (d+h)I(C;D)$$

follows from "R is a D-copy of A over BC" and:

$$aI(AR;BR) \leq bI(AR;BR|CR) + cI(AR;CR|BR) + dI(BR;CR|AR) \\ + aI(AR;BR|D) + hI(CR;D)$$

$$(a+d)I(A;B) \leq (b+2d+h)I(A;B|C) + (a+c+d)I(A;C|B) + dI(B;C|A)$$
$$+ (a+d)I(A;B|D) + (d+h)I(C;D)$$

follows from "R is a D-copy of A over BC" and an instance of:

$$aI(A;B) \leq bI(A;B|C) + cI(A;C|B) + dI(B;C|A)$$
$$+ aI(A;B|D) + hI(C;D)$$

If

$$aI(A;B) \leq bI(A;B|C) + cI(A;C|B) + dI(B;C|A) \\ + aI(A;B|D) + hI(C;D)$$

is an entropy inequality, then

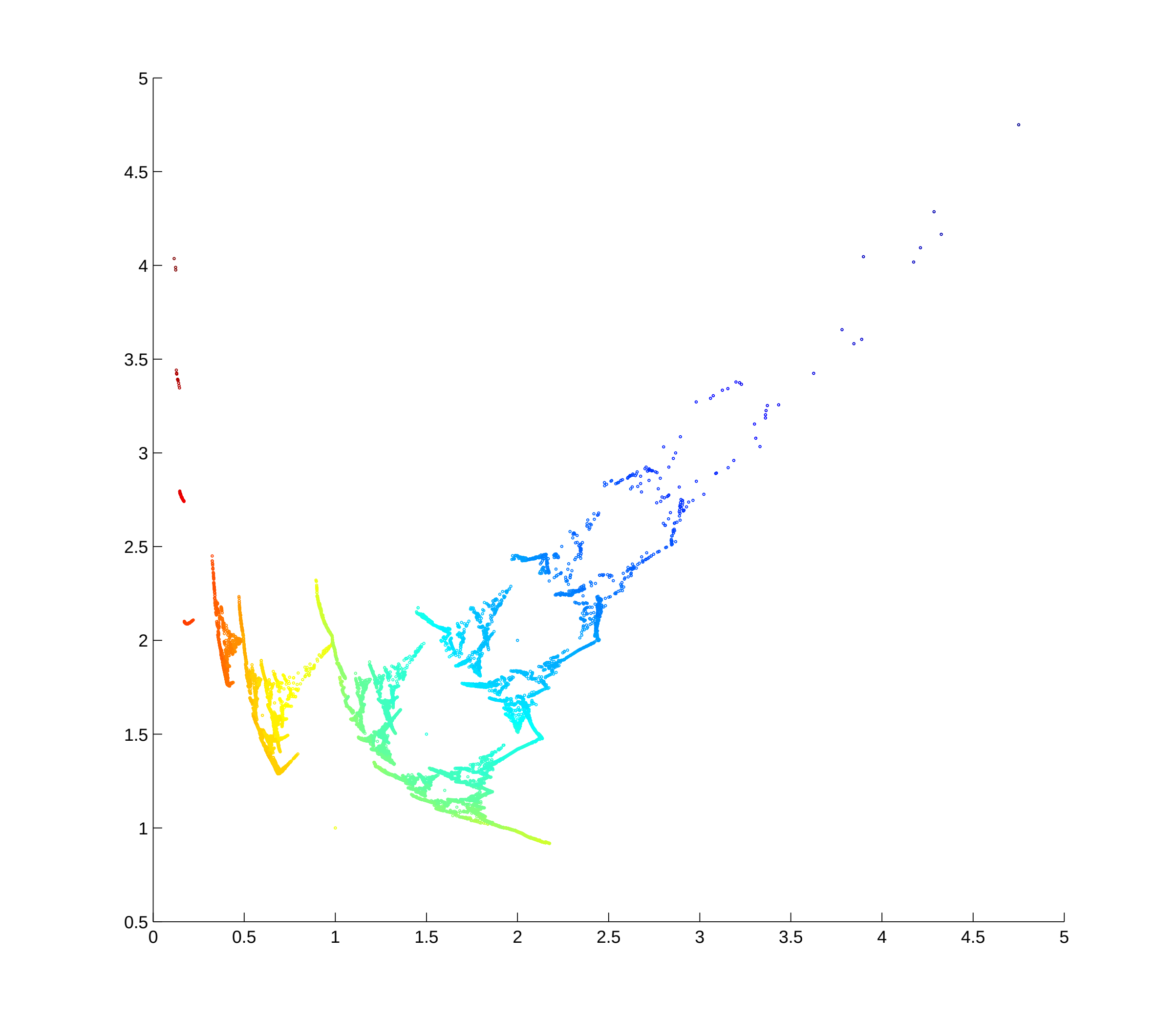
$$(a+d)I(A;B) \leq (b+2d+h)I(A;B|C) + (a+c+d)I(A;C|B) + dI(B;C|A) \\ + (a+d)I(A;B|D) + (d+h)I(C;D)$$

is an entropy inequality.

Iterating this rule gives:

$$
\begin{aligned}
I(A;B) &\leq 2I(A;B|C) + I(A;C|B) + I(B;C|A) + I(A;B|D) + I(C;D) \\
2I(A;B) &\leq 5I(A;B|C) + 3I(A;C|B) + I(B;C|A) + 2I(A;B|D) + 2I(C;D) \\
3I(A;B) &\leq 9I(A;B|C) + 6I(A;C|B) + I(B;C|A) + 3I(A;B|D) + 3I(C;D) \\
4I(A;B) &\leq 14I(A;B|C) + 10I(A;C|B) + I(B;C|A) + 4I(A;B|D) + 4I(C;D) \\
5I(A;B) &\leq 20I(A;B|C) + 15I(A;C|B) + I(B;C|A) + 5I(A;B|D) + 5I(C;D) \\
6I(A;B) &\leq 27I(A;B|C) + 21I(A;C|B) + I(B;C|A) + 6I(A;B|D) + 6I(C;D) \\
7I(A;B) &\leq 35I(A;B|C) + 28I(A;C|B) + I(B;C|A) + 7I(A;B|D) + 7I(C;D) \\
&\ldots
\end{aligned}
$$

This is another way of generating the list of inequalities that Matúš proved and showed could not follow from finitely many entropy inequalities.



$$I(A;B) \leq (x+1)I(A;B|C) + yI(A;C|B) + zI(B;C|A) + I(A;B|D) + I(C;D)$$

The Ingleton inequality

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

This inequality and the Shannon inequalities generate all linear rank inequalities on four variables. (Hammer-Romashchenko-Shen-Vereshchagin)

For four variables, the Shannon region consists of the representable regions together with six (conified) simplices, one for each of the permuted-variable forms of the Ingleton inequality. The entropy region consists of the representable region together with a proper subset of each of the six simplices.

Common informations

Random variable Z is a *common information* of random variables A and B if:

$$H(Z|A) = 0$$
$$H(Z|B) = 0$$
$$H(Z) = I(A; B)$$

Common informations do not always exist, but they do in the case that A and B come from vector subspaces (Z will correspond to the intersection of these subspaces).

24 new five-variable linear rank inequalities

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

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

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

$$\ldots$$

$$\begin{aligned} I(A;C,D) + I(B;C,D) \leq{} & I(B;D) + I(B;C|E) + I(C;E|D) + I(A;E) \\ & + I(A;C|B,D) + I(A,B;D|C) + I(A;D|B,E) + I(A;B|D,E) \end{aligned}$$

Five-variable inequalities

These 24 (1700 counting permutations) inequalities, together with 4 (120) instances of the Ingleton inequality and 5 (85) elemental Shannon inequalities, give the complete nonredundant list of five-variable linear rank inequalities.

Extreme rays for five-variable linear ranks

162 extreme rays (7943 counting permutations)

All have been verified to be representable over any sufficiently large field.

The following are Shannon inequalities:

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

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

Actually, one can improve the latter inequality to:

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

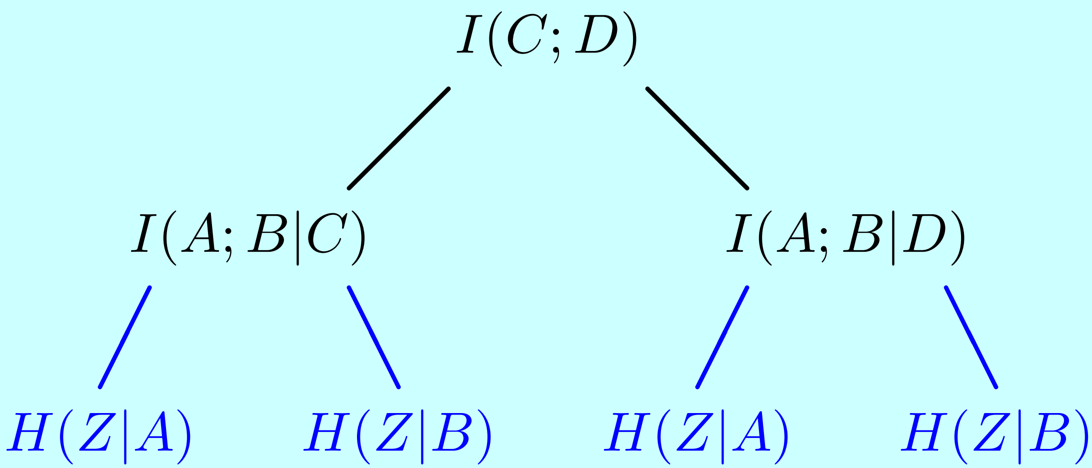
Proof of the Ingleton inequality

$$I(C;D)$$

$I(A;B|C)$ $I(A;B|D)$

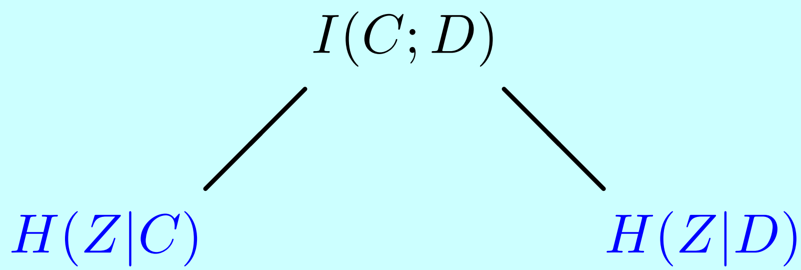
Z is a common information of A and B

Proof of the Ingleton inequality



Z is a common information of A and B

Proof of the Ingleton inequality



Z is a common information of A and B

Proof of the Ingleton inequality

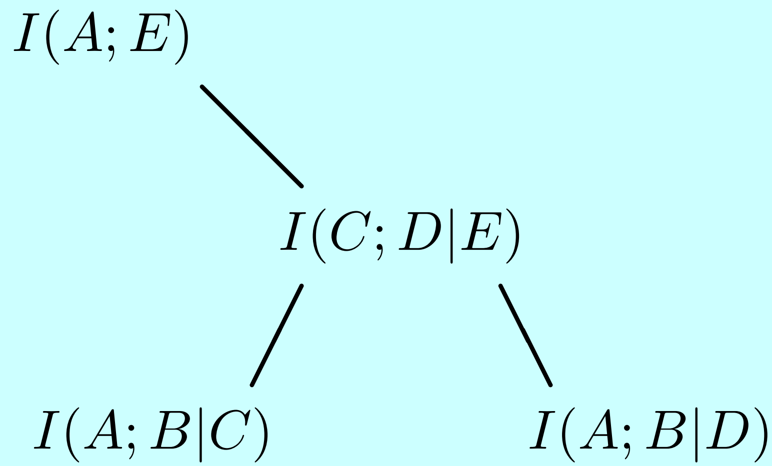
$$H(Z)$$

Z is a common information of A and B

Proof of the Ingleton inequality

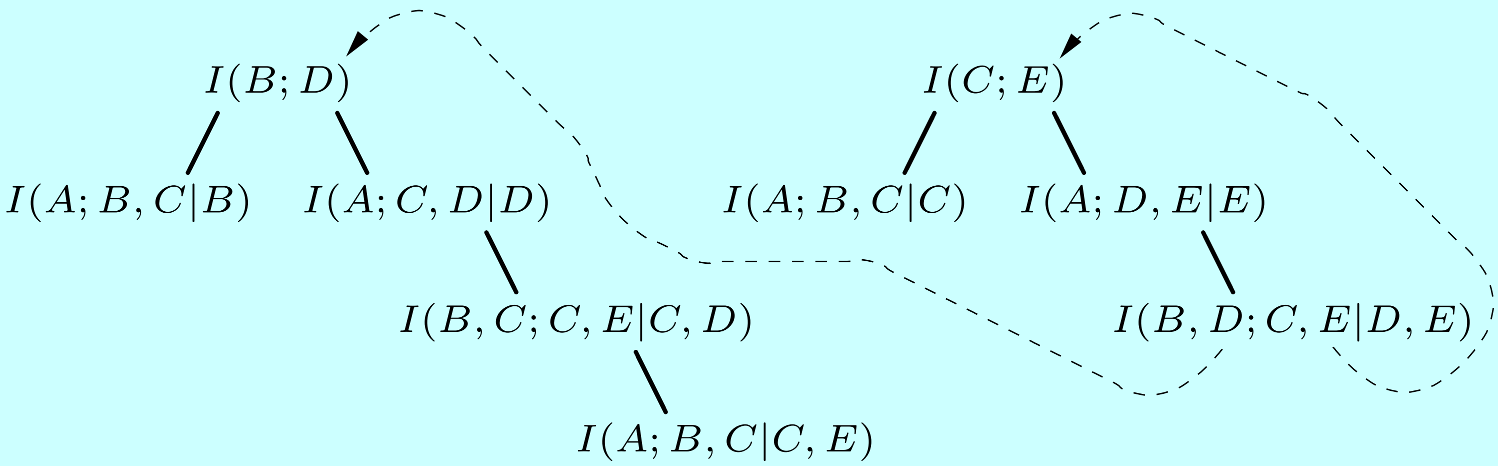
$$I(A;B)$$

Z is a common information of A and B



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

Another linear rank inequality



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

Six-variable inequalities — partial results

The sharp six-variable inequalities include 6 (246) elemental Shannon inequalities, 12 (1470) instances of the Ingleton inequality, and 167 (61740) instances of the new five-variable inequalities. In addition to these, we currently have 746458 (531344600) true six-variable inequalities, proved using one, two, or three common informations.

We also have a stockpile of 27907 (17251597) six-variable polymatroids which have been verified to be representable and give extreme rays of the linear rank region.

Differences in computational methods:

Higher dimension

Sharp inequalities

Actual representable polymatroids

(generic vector selection)

Yet to come:

Dependence on characteristic

Can methods that have so far been used only on one side (entropy inequalities or linear rank inequalities) be used on the other side as well?

The End.