Information Inequalities for Joint Distributions, with interpretations and applications

Mokshay Madiman

Yale University

Joint work with Prasad Tetali, Georgia Tech

Heriot–Watt University
20 June 2008

Outline

- The Main Inequality for Joint Entropies
- A simple proof
- Duality, and connections to game theory
- Remarks on entropy inequalities for sums
- New determinantal inequalities
- Counting independent sets
- Concluding remarks

Entropy

For random element X, entropy $H(X) = H(p) = E[-\log p(X)]$

where p is the p.m.f of X (if X is discrete),
or the p.d.f of X (if X is continuous)

Conditional entropy of X given Y is

$$H(X|Y) = \sum_y p_Y(y) H(X|Y = y)$$

where $H(X|Y = y)$ is the entropy of $p(x|Y = y)$.

Two useful facts

- Shannon's Chain Rule:

$$H(X, Y) = H(Y) + H(X|Y)$$

- Conditioning reduces entropy:

$$H(X) - H(X|Y) = D(p_{X,Y} \| p_X \times p_Y) \geq 0$$

Hypergraphs

We wish to consider various subsets of the random variables $X_1, \ldots, X_n$

Conventions

- $[n]$ is the index set $\{1, 2, \ldots, n\}$
- A collection $\mathcal{C}$ of subsets of $[n]$ is a *hypergraph*, and the sets in $\mathcal{C}$ are hyperedges

  E.g.: $\big\{\{1,2\}, \{1,3\}, \{2,3,4\}\big\}$
- For any $\mathbf{s} \in \mathcal{C}$, $X_{\mathbf{s}}$ stands for the collection of random variables $(X_i : i \in \mathbf{s})$, with the indices taken in their increasing order
- For any index i in $[n]$, define the *degree* of i in $\mathcal{C}$ as

$$r(i) = \left|\{\mathbf{t} \in \mathcal{C} : i \in \mathbf{t}\}\right|$$

- We will always implicitly assume that all hypergraphs we consider have *minimum degree* $r_- := \min_{i \in [n]} r(i) > 0$, i.e., every index i appears in at least one hyperedge

Weak form of main inequality

The Inequality

For any hypergraph $\mathcal{C}$ on $[n]$,

$$\sum_{\mathbf{s}\in\mathcal{C}} \frac{H(X_{\mathbf{s}}|X_{\mathbf{s}^c})}{r_+(\mathbf{s})} \leq H(X_{[n]}) \leq \sum_{\mathbf{s}\in\mathcal{C}} \frac{H(X_{\mathbf{s}})}{r_-(\mathbf{s})}$$

where $r_+(\mathbf{s}) = \max_{i\in\mathbf{s}} r(i)$ (max degree within $\mathbf{s}$)
and $r_-(\mathbf{s}) = \min_{i\in\mathbf{s}} r(i)$ (min degree within $\mathbf{s}$)

History

- Shannon '48, Han '78, Chung–Frankl–Graham–Shearer '86
- Recent Applns: Friedgut–Kahn '98, Radhakrishnan '01, Kahn '01, Brightwell–Tetali '03, Friedgut '04, Galvin–Tetali '04

Now unified. . .

- Both upper and lower bounds for general hypergraphs $\mathcal{C}$
- Coefficients improved

Special Cases of the Weak Form

- Applying to the class $\mathcal{C}_1$ of singletons,

$$\sum_{i=1}^{n} H(X_i|X_{[n]\setminus i}) \leq H(X_{[n]}) \leq \sum_{i=1}^{n} H(X_i)$$

Upper bound is Shannon's, and lower bound is related to erasure entropy

- Applying to the class $\mathcal{C}_{n-1}$ of all sets of $n-1$ elements,

$$\frac{1}{n-1}\sum_{i=1}^{n} H(X_{[n]\setminus i}|X_i) \leq H(X_{[n]}) \leq \frac{1}{n-1}\sum_{i=1}^{n} H(X_{[n]\setminus i})$$

This is Han's inequality

- Using $r_- \leq r_-(\mathbf{s})$ and $r_+ \leq r_+(\mathbf{s})$, we have

$$\frac{1}{r_+}\sum_{\mathbf{s}\in\mathcal{C}} H(X_{\mathbf{s}}|X_{\mathbf{s}^c}) \leq H(X_{[n]}) \leq \frac{1}{r_-}\sum_{\mathbf{s}\in\mathcal{C}} H(X_{\mathbf{s}})$$

Upper bound is Shearer's lemma and lower bound is new

Strong (degree) form of main inequality

Theorem 1: [Strong degree form]

For any hypergraph $\mathcal{C}$ on $[n]$,

$$\sum_{\mathbf{s}\in\mathcal{C}} \frac{H(X_{\mathbf{s}}|X_{\mathbf{s}^c\setminus>\mathbf{s}})}{r_+(\mathbf{s})} \leq H(X_{[n]}) \leq \sum_{\mathbf{s}\in\mathcal{C}} \frac{H(X_{\mathbf{s}}|X_{<\mathbf{s}})}{r_-(\mathbf{s})}$$

where $<\mathbf{s}$ is the set of indices less than every index in $\mathbf{s}$
and $>\mathbf{s}$ is the set of indices greater than every index in $\mathbf{s}$

Strong vs. weak form for $\mathcal{C}_1$

- The weak form says

$$\sum_{i=1}^{n} H(X_i|X_{[n]\setminus i}) \leq H(X_{[n]}) \leq \sum_{i=1}^{n} H(X_i)$$

- The strong form is Shannon's chain rule!

$$\sum_{i=1}^{n} H(X_i|X_{[n]\setminus\geq i}) \leq H(X_{[n]}) \leq \sum_{i=1}^{n} H(X_i|X_{<i})$$

Strong (fractional) form of main inequality

Fractional coverings and packings

$\alpha : \mathcal{C} \to \mathbb{R}_+$ is a fractional covering of $[n]$ using $\mathcal{C}$ if

$$\sum_{\mathbf{s} \in \mathcal{C}, \mathbf{s} \ni j} \alpha(\mathbf{s}) \geq 1 \quad \text{for every } j \in [n]$$

Idea: If $\mathbf{s} \ni i$, think of $\mathbf{s}$ as containing a fraction $\alpha(\mathbf{s})$ of i

Similarly, β is a fractional packing if

$$\sum_{\mathbf{s} \in \mathcal{C}, \mathbf{s} \ni j} \beta(\mathbf{s}) \leq 1 \quad \text{for every } j \in [n]$$

Theorem 1': [Strong fractional form]

For any fractional covering α and fractional packing β,

$$\sum_{\mathbf{s} \in \mathcal{C}} \beta(\mathbf{s}) H(X_{\mathbf{s}} | X_{\mathbf{s}^c \setminus > \mathbf{s}}) \leq H(X_{[n]}) \leq \sum_{\mathbf{s} \in \mathcal{C}} \alpha(\mathbf{s}) H(X_{\mathbf{s}} | X_{< \mathbf{s}})$$

Proof of the main inequality

$$\sum_{\mathbf{s}\in\mathcal{C}} \alpha(\mathbf{s}) H(X_{\mathbf{s}}|X_{<\mathbf{s}}) = \sum_{\mathbf{s}\in\mathcal{C}} \alpha(\mathbf{s}) \sum_{j\in\mathbf{s}} H(X_j|X_{<j\cap\mathbf{s}}, X_{<\mathbf{s}}) \,.$$

Proof of the main inequality

$$\sum_{\mathbf{s} \in \mathcal{C}} \alpha(\mathbf{s}) H(X_{\mathbf{s}} | X_{<\mathbf{s}}) = \sum_{\mathbf{s} \in \mathcal{C}} \alpha(\mathbf{s}) \sum_{j \in \mathbf{s}} H(X_j | X_{<j \cap \mathbf{s}}, X_{<\mathbf{s}})$$

$$\geq \sum_{\mathbf{s} \in \mathcal{C}} \alpha(\mathbf{s}) \sum_{j \in \mathbf{s}} H(X_j | X_{<j})$$

Proof of the main inequality

$$\begin{aligned}
\sum_{\mathbf{s}\in\mathcal{C}} \alpha(\mathbf{s}) H(X_{\mathbf{s}}|X_{<\mathbf{s}}) &= \sum_{\mathbf{s}\in\mathcal{C}} \alpha(\mathbf{s}) \sum_{j\in\mathbf{s}} H(X_j|X_{<j\cap\mathbf{s}}, X_{<\mathbf{s}}) \\
&\geq \sum_{\mathbf{s}\in\mathcal{C}} \alpha(\mathbf{s}) \sum_{j\in\mathbf{s}} H(X_j|X_{<j}) \\
&= \sum_{j\in[n]} H(X_j|X_{<j}) \sum_{\mathbf{s}\in\mathcal{C}} \alpha(\mathbf{s}) \mathbf{1}_{\{j\in\mathbf{s}\}}
\end{aligned}$$

Proof of the main inequality

$$\begin{aligned}
\sum_{\mathbf{s}\in\mathcal{C}} \alpha(\mathbf{s}) H(X_{\mathbf{s}}|X_{<\mathbf{s}}) &= \sum_{\mathbf{s}\in\mathcal{C}} \alpha(\mathbf{s}) \sum_{j\in\mathbf{s}} H(X_j|X_{<j\cap\mathbf{s}}, X_{<\mathbf{s}}) \\
&\geq \sum_{\mathbf{s}\in\mathcal{C}} \alpha(\mathbf{s}) \sum_{j\in\mathbf{s}} H(X_j|X_{<j}) \\
&= \sum_{j\in[n]} H(X_j|X_{<j}) \sum_{\mathbf{s}\in\mathcal{C}} \alpha(\mathbf{s}) \mathbf{1}_{\{j\in\mathbf{s}\}} \\
&\geq \sum_{j\in[n]} H(X_j|X_{<j}) \\
&= H(X_{[n]})\,.
\end{aligned}$$

Proof of lower bound is similar.

The degree packing and degree covering

Fractional form $\Rightarrow$ degree form

For any hypergraph $\mathcal{C}$, the numbers $\alpha(\mathbf{s}) = \frac{1}{r_-(\mathbf{s})}$ provide a fractional covering:

$$\sum_{\mathbf{s}\in\mathcal{C},\mathbf{s}\ni i} \frac{1}{r_-(\mathbf{s})} = \sum_{\mathbf{s}\in\mathcal{C}} \frac{\mathbf{1}_{\{i\in\mathbf{s}\}}}{r_-(\mathbf{s})} \geq \sum_{\mathbf{s}\in\mathcal{C}} \frac{\mathbf{1}_{\{i\in\mathbf{s}\}}}{r(i)} = 1$$

Similarly the numbers $\beta(\mathbf{s}) = \frac{1}{r_+(\mathbf{s})}$ provide a fractional packing

Degree form $\Rightarrow$ fractional form

Consider repetitions of sets

Outline

- The Main Inequality for Joint Entropies
- A simple proof
- Duality, and connections to game theory
- Remarks on entropy inequalities for sums
- New determinantal inequalities
- Counting independent sets
- Concluding remarks

Other approaches to the main inequality

- Friedgut '04 proved the upper bound of the weak form
  - Expressed in terms of hypergraph projections rather than entropies
  - Proof more complicated, and cannot be used to get lower bounds

- Can also prove the weak form using folklore results from game theory. Elements (ignoring subtleties):
  - A set function $f : 2^{[n]} \to \mathbb{R}_+$ is increasing if $f(\mathbf{s}) \leq f(\mathbf{t})$ when $\mathbf{s} \subset \mathbf{t}$, and submodular if $f(\mathbf{s} \cup \mathbf{t}) + f(\mathbf{s} \cap \mathbf{t}) \leq f(\mathbf{s}) + f(\mathbf{t})$ for any $\mathbf{s}, \mathbf{t}$
  - Fujishige '78: Entropy as a set function [i.e., $f(\mathbf{s}) = H(X_{\mathbf{s}})$] is increasing and submodular
  - Shapley '71: Games defined by submodular, increasing set functions have a non-empty core
  - Bondareva-Shapley theorem (1960's): A game defined by a set function f has a non-empty core iff f is fractional subadditive

Additional remarks

Duality

The collection of lower bounds for $H(X_{[n]})$ using all $\mathcal{C}$ and β is equivalent to the collection of upper bounds using all $\mathcal{C}$ and α. In fact,

$$\mathsf{Gap}_U(\mathcal{C},\alpha) = \Big(\sum_{\mathbf{s}\in\mathcal{C}} \alpha(\mathbf{s}) - 1\Big)\mathsf{Gap}_L(\bar{\mathcal{C}},\bar{\beta})$$

where $\bar{\beta}$ is a fractional packing of the (complementary) hypergraph $\bar{\mathcal{C}} = \{\mathbf{s}^c : \mathbf{s}\in\mathcal{C}\}$ defined as $\bar{\beta}(\mathbf{s}^c) = \frac{\alpha(\mathbf{s})}{\sum_{\mathbf{s}\in\mathcal{C}}\alpha(\mathbf{s})-1}$

Fractional Form of Entropy Power Inequality

Theorem 2: For independent random variables with densities and finite variances,

$$\mathcal{N}(\text{sum}_{[n]}) \geq \sum_{\mathbf{s} \in \mathcal{C}} \beta(\mathbf{s}) \mathcal{N}(\text{sum}_{\mathbf{s}}) \qquad \text{[M.M. '08]}$$

where $\mathcal{N}(X) = e^{2h(X)}$ and $\text{sum}_{\mathbf{s}} = \sum_{i \in \mathbf{s}} X_i$

Connection to Entropic CLT

- History: Shannon '48, Stam '59, Artstein-Ball-Barthe-Naor '04, M.M.–Barron '07
- Gaussian is MaxEnt: $N(0, \sigma^2)$ has maximum entropy among all densities on $\mathbb{R}$ with variance $\leq \sigma^2$
- Entropic Convergence: Let X_i be drawn i.i.d. from a density. If $EX_1 = 0$ and $EX_1^2 = \sigma^2$, and $S_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$, then as $n \to \infty$,

$$D(S_n \| N(0, \sigma^2)) \downarrow 0 \qquad \text{or equivalently,} \qquad h(S_n) \uparrow h(N(0, \sigma^2))$$

[Barron '86, Artstein et al. '04]

Upper bounds for entropy of sums

Theorem 4: [Upper Bound for Discrete Entropy]

If X_i are independent discrete random variables, then

$$H(\mathsf{sum}_{[n]}) \leq \sum_{\mathbf{s} \in \mathcal{C}} \alpha(\mathbf{s}) H(\mathsf{sum}_{\mathbf{s}})$$

for any fractional covering α using any collection $\mathcal{C}$ of subsets of $[n]$

Theorem 4': [Upper Bound for Differential Entropy]

Let X and $X_1, \ldots, X_n$ be independent $\mathbb{R}^d$-valued random vectors with densities. Suppose α is a fractional covering for the collection $\mathcal{C}$ of subsets of $[n]$. Then

$$h(X + \mathsf{sum}_{[n]}) \leq \sum_{\mathbf{s} \in \mathcal{C}} \alpha(\mathbf{s}) h(X + \mathsf{sum}_{\mathbf{s}}) - \left(\sum_{\mathbf{s} \in \mathcal{C}} \alpha(\mathbf{s}) - 1 \right) h(X)$$

If $h(X) \geq 0$, then Theorem 4' is identical to Theorem 4 except that the random variable X is an additional summand in every sum

Outline

- The Main Inequality for Joint Entropies
- A simple proof
- Duality, and connections to game theory
- Remarks on entropy inequalities for sums
- New determinantal inequalities
- Counting independent sets
- Concluding remarks

Determinantal Inequalities

Let K be a positive definite $n \times n$ matrix and let $\mathcal{C}$ be a hypergraph on $[n]$. Let $K(\mathbf{s})$ denote the submatrix corresponding to the rows and columns indexed by elements of $\mathbf{s}$. Then for any fractional partition α^*,

$$\prod_{\mathbf{s} \in \mathcal{S}} \left(\frac{|K|}{|K(\mathbf{s}^c)|} \right)^{\alpha^*(\mathbf{s})} \leq |K| \leq \prod_{\mathbf{s} \in \mathcal{S}} |K(\mathbf{s})|^{\alpha^*(\mathbf{s})}$$

Special cases

- Hadamard's inequality ($\mathcal{C}_1$ with degree partition):

$$|K| \leq \prod_{i \in [n]} K_{ii}$$

- Fischer inequality:

$$|K| \leq |K(\mathbf{s})| \cdot |K(\mathbf{s}^c)|$$

- Szasz inequality, basic form ($\mathcal{C}_{n-1}$ with degree partition):

$$|K| \leq \prod_{\mathbf{s} \in \mathcal{C}_{n-1}} |K(\mathbf{s})|^{\frac{1}{n-1}}$$

Independent sets in graphs

Goal

Let $G = (V, E)$ be an arbitrary graph on n vertices. An independent set is a subset of vertices such that no two of them are connected by an edge of G. How many independent sets does G have?

E.g.: Complete bipartite graph $K_{m,n}$ has $2^m + 2^n - 1$ independent sets

A Probabilistic Method

- Consider a randomly drawn independent set, represented by the 0-1 random variables $X_{[n]}$. Then the entropy $H(X_{[n]}) = \log |\mathcal{I}(G)|$.
- Apply an entropy inequality to bound $H(X_{[n]})$ using certain subset entropies for a clever choice of hypergraph (new element: ordering is key)
- Perform an estimation of the resulting bound

Counting independent sets

Main Result

Let $G = (V, E)$: an arbitrary graph on n vertices
$\prec$: the ordering on V according to decreasing order of degrees
$p(v)$: the number of neighbors of v which precede v

Then

$$|\mathcal{I}(G)| \leq \prod_{v \in V} 2^{(p(v)+1)\frac{1}{d(v)}}$$

History

- Kahn (2001) proved the result for d-regular, bipartite graphs, and subsequently extended it to all d-regular graphs
- Kahn's conjecture (open): For any graph $G = (V, E)$,

$$|\mathcal{I}(G)| \leq \prod_{uv \in E} \left|\mathcal{I}(K_{d(u),d(v)})\right|^{\frac{1}{d(u)d(v)}} = \prod_{uv \in E} \left(2^{d(u)} + 2^{d(v)} - 1\right)^{\frac{1}{d(u)d(v)}}$$

Construction of hypergraph

Let $\prec$ denote the ordering on vertices according to the decreasing order of their degrees. For each $i \in V$, let

$$P(i) = \{j \in V : \{i, j\} \in E \text{ and } j \prec i\}$$

and define $p(i) = |P(i)|$

Consider the hypergraph $\mathcal{C}$ to be

- the collection of sets $P(i)$ and
- $p(i)$ copies of singleton sets $\{i\}$, for each i

Each i appears in

- $d(i) - p(i)$ sets of the form $P(j)$, corresponding to each j such that $i \prec j$ and $\{i, j\} \in E$, and
- once in each of the $p(i)$ singleton sets $\{i\}$

Thus $r(i) = d(i)$ (degree of i in $\mathcal{C}$=degree of i in G)

Proof of bound on number of independent sets

$$H(X) \leq \sum_{i \in V} \frac{1}{\min_{j \in P(i)} d(j)} H\Big(X_{P(i)} | X_{\prec P(i)}\Big) + \sum_{i \in V} \frac{p(i)}{d(i)} H(X_i | X_{\prec i})$$

Proof of bound on number of independent sets

$$H(X) \leq \sum_{i \in V} \frac{1}{\min_{j \in P(i)} d(j)} H\Big(X_{P(i)} | X_{\prec P(i)}\Big) + \sum_{i \in V} \frac{p(i)}{d(i)} H(X_i | X_{\prec i})$$

$$\leq \sum_{i \in V} \Big(\frac{1}{d(i)} H(X_{P(i)}) + \frac{p(i)}{d(i)} H(X_i | X_{P(i)})\Big)$$

using property of ordering, and since conditioning reduces entropy

Proof of bound on number of independent sets

$$H(X) \leq \sum_{i \in V} \frac{1}{\min_{j \in P(i)} d(j)} H\left(X_{P(i)} | X_{\prec P(i)}\right) + \sum_{i \in V} \frac{p(i)}{d(i)} H(X_i | X_{\prec i})$$

$$\leq \sum_{i \in V} \left(\frac{1}{d(i)} H(X_{P(i)}) + \frac{p(i)}{d(i)} H(X_i | X_{P(i)}) \right)$$

$$= \sum_{i \in V} \frac{1}{d(i)} \sum_{x_{P(i)} \in R_i} \left(q(x_{P(i)}) \log \frac{1}{q(x_{P(i)})} + p(i) q(x_{P(i)}) H(X_i | X_{P(i)} = x_{P(i)}) \right)$$

where $R_i = \{x_{P(i)} : x_{[n]} \text{ represents an independent set}\}$
and $q(x_{P(i)}) = \Pr\{X_{P(i)} = x_{P(i)}\}$

Proof of bound on number of independent sets

$$
\begin{aligned}
H(X) &\leq \sum_{i \in V} \frac{1}{\min_{j \in P(i)} d(j)} H\Big(X_{P(i)} | X_{\prec P(i)}\Big) + \sum_{i \in V} \frac{p(i)}{d(i)} H(X_i | X_{\prec i}) \\
&\leq \sum_{i \in V} \Big(\frac{1}{d(i)} H(X_{P(i)}) + \frac{p(i)}{d(i)} H(X_i | X_{P(i)}) \Big) \\
&= \sum_{i \in V} \frac{1}{d(i)} \sum_{x_{P(i)} \in R_i} \Big(q(x_{P(i)}) \log \frac{1}{q(x_{P(i)})} + p(i) q(x_{P(i)}) H(X_i | X_{P(i)} = x_{P(i)}) \Big) \\
&= \sum_{i \in V} \frac{1}{d(i)} \sum_{x_{P(i)} \in R_i} q(x_{P(i)}) \log \frac{|\mathrm{Range}(x_i | x_{P(i)})|^{p(i)}}{q(x_{P(i)})}
\end{aligned}
$$

Proof of bound on number of independent sets

$$
\begin{aligned}
H(X) &\le \sum_{i\in V} \frac{1}{\min_{j\in P(i)} d(j)} H\left(X_{P(i)} | X_{\prec P(i)}\right) + \sum_{i\in V} \frac{p(i)}{d(i)} H(X_i | X_{\prec i}) \\
&\le \sum_{i\in V} \left(\frac{1}{d(i)} H(X_{P(i)}) + \frac{p(i)}{d(i)} H(X_i | X_{P(i)})\right) \\
&= \sum_{i\in V} \frac{1}{d(i)} \sum_{x_{P(i)}\in R_i} \left(q(x_{P(i)}) \log \frac{1}{q(x_{P(i)})} + p(i) q(x_{P(i)}) H(X_i | X_{P(i)} = x_{P(i)})\right) \\
&\le \sum_{i\in V} \frac{1}{d(i)} \sum_{x_{P(i)}\in R_i} q(x_{P(i)}) \log \frac{|\mathrm{Range}(x_i | x_{P(i)})|^{p(i)}}{q(x_{P(i)})} \\
&\le \sum_{i\in V} \frac{1}{d(i)} \left[q_0 \log \frac{2^{p(i)}}{q_0} + \sum_{x_{P(i)}\in \mathcal{X}_i \setminus \{0_{P(i)}\}} q(x_{P(i)}) \log \frac{1}{q(x_{P(i)})} \right]
\end{aligned}
$$

where $q_0 = q(0_{P(i)})$

Proof of bound on number of independent sets

$$
\begin{aligned}
H(X) &\le \sum_{i \in V} \frac{1}{\min_{j \in P(i)} d(j)} H\left(X_{P(i)} | X_{\prec P(i)}\right) + \sum_{i \in V} \frac{p(i)}{d(i)} H(X_i | X_{\prec i}) \\
&\le \sum_{i \in V} \left(\frac{1}{d(i)} H(X_{P(i)}) + \frac{p(i)}{d(i)} H(X_i | X_{P(i)}) \right) \\
&= \sum_{i \in V} \frac{1}{d(i)} \sum_{x_{P(i)} \in R_i} \left(q(x_{P(i)}) \log \frac{1}{q(x_{P(i)})} + p(i) q(x_{P(i)}) H(X_i | X_{P(i)} = x_{P(i)}) \right) \\
&\le \sum_{i \in V} \frac{1}{d(i)} \sum_{x_{P(i)} \in R_i} q(x_{P(i)}) \log \frac{|\mathrm{Range}(x_i | x_{P(i)})|^{p(i)}}{q(x_{P(i)})} \\
&\le \sum_{i \in V} \frac{1}{d(i)} \left[q_0 \log \frac{2^{p(i)}}{q_0} + \sum_{x_{P(i)} \in \mathcal{X}_i \setminus \{0_{P(i)}\}} q(x_{P(i)}) \log \frac{1}{q(x_{P(i)})} \right] \\
&\le \sum_{i \in V} \frac{1}{d(i)} \log(2^{p(i)} + 2^{p(i)})
\end{aligned}
$$

Relative entropy version

If Q is a product measure,

$$D(P_{X_{[n]}} \| Q_{X_{[n]}}) \leq \sum_{\mathbf{s} \in \mathcal{C}} \frac{D(P_{X_{\mathbf{s}}} \| Q_{X_{\mathbf{s}}})}{r_{-}(\mathbf{s})}$$

Remarks

- Equivalent to the main entropy inequality (if we put in conditioning)
- Hypothesis testing interpretation
- Gives a generalized tensorization property of the entropy functional

Summary

- A new entropy inequality for joint distributions, unifying and extending much previous work
- Related: entropy inequalities for sums
- Applications: counting, determinantal inequalities, hypothesis testing etc.
- Open problems abound (e.g.: can we deploy the lower bound for counting? Kahn's conjecture?)

Thank you for your attention!

○ — ○ — ○

Extras

Extensions: Zero-error source-channel codes

For any N-vertex (source confusability) graph G and
any (channel characteristic) graph F,

one may estimate the number of homomorphisms from G to H:

$$|Hom(G,F)| \leq \prod_{v \in V} |Hom(K_{p(v),p(v)}, F)|^{\frac{1}{d(v)}},$$

where $p(v)$ denotes the number of vertices preceding v in an ordering induced by decreasing degrees.

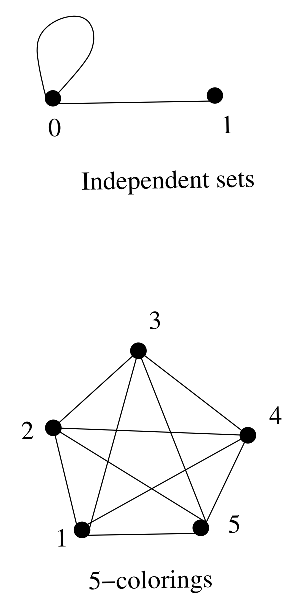

Figure 1: The graphs F relevant for counting independent sets and number of 5-colorings.