Polyhedra in information theory

Tobias Boege

Department of Mathematics and Statistics
UiT The Arctic University of Norway

Mørketidens Mattemøte
Tromsø, 24 January 2025

Entropy

Let $X \in \Delta(d)$ be a random variable taking finitely many values $\{1, \ldots, d\}$ with non-negative probabilities $p_1, \ldots, p_d$. Its *Shannon entropy* is

$$H(X) := \sum_{i=1}^{d} p_i \log(1/p_i) \quad \left[\text{with } 0 \cdot \log(1/0) := 0\right]$$

Entropy

Let $X \in \Delta(d)$ be a random variable taking finitely many values $\{1, \ldots, d\}$ with non-negative probabilities $p_1, \ldots, p_d$. Its *Shannon entropy* is

$$H(X) := \sum_{i=1}^{d} p_i \log(1/p_i) \quad [\text{with } 0 \cdot \log(1/0) := 0]$$

- H is continuous on $\Delta(d)$ and analytic on the interior.

Entropy

Let $X \in \Delta(d)$ be a random variable taking finitely many values $\{1, \ldots, d\}$ with non-negative probabilities $p_1, \ldots, p_d$. Its *Shannon entropy* is

$$H(X) := \sum_{i=1}^{d} p_i \log(1/p_i) \quad \left[\text{with } 0 \cdot \log(1/0) := 0\right]$$

- H is continuous on $\Delta(d)$ and analytic on the interior.
- A random vector $X \in \Delta(d_i : i \in N)$ is a random variable in $\Delta(\prod_{i \in N} d_i)$, so the definition of H extends to vectors.

Entropy

Let $X \in \Delta(d)$ be a random variable taking finitely many values $\{1, \ldots, d\}$ with non-negative probabilities $p_1, \ldots, p_d$. Its *Shannon entropy* is

$$H(X) := \sum_{i=1}^{d} p_i \log(1/p_i) \quad [\text{with } 0 \cdot \log(1/0) := 0]$$

- H is continuous on $\Delta(d)$ and analytic on the interior.
- A random vector $X \in \Delta(d_i : i \in N)$ is a random variable in $\Delta(\prod_{i \in N} d_i)$, so the definition of H extends to vectors.
- For a random vector $X = (X_i : i \in N)$ we have 2^N marginals and we collect their entropies in an entropy profile $h_X : 2^N \to \mathbb{R}$.
  - For example (X, Y) has entropy profile $(0, H(X), H(Y), H(X, Y)) \in \mathbb{R}^4$.

Entropy as information

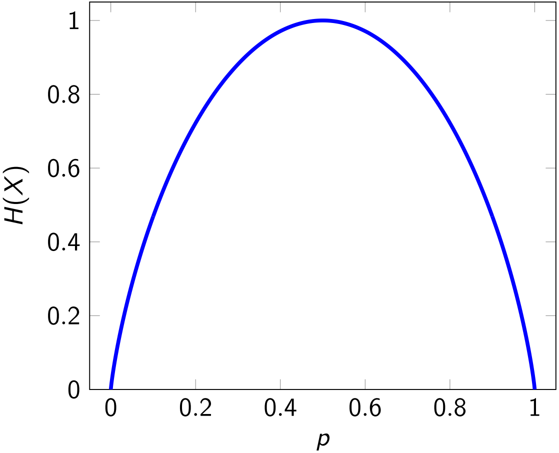


Figure: Entropy of a binary random variable X as a function of $p = \Pr[X = \text{heads}]$.

Synthetic geometry for random variables

Entropy profile encodes qualitative information about the system of random variables:

Synthetic geometry for random variables

Entropy profile encodes qualitative information about the system of random variables:

- Subvector X_I is functionally dependent on X_K if and only if $\boldsymbol{h_X(I \cup K) = h_X(K)}$.

Synthetic geometry for random variables

Entropy profile encodes qualitative information about the system of random variables:

- Subvector X_I is functionally dependent on X_K if and only if $\boldsymbol{h_X(I \cup K) = h_X(K)}$.
- Subvectors X_I and X_J are conditionally independent given X_K if and only if $\boldsymbol{h_X(I \cup K) + h_X(J \cup K) = h_X(I \cup J \cup K) + h_X(K)}$.

Synthetic geometry for random variables

Entropy profile encodes qualitative information about the system of random variables:

- Subvector X_I is functionally dependent on X_K if and only if $\boldsymbol{h_X(I \cup K) = h_X(K)}$.
- Subvectors X_I and X_J are conditionally independent given X_K if and only if $\boldsymbol{h_X(I \cup K) + h_X(J \cup K) = h_X(I \cup J \cup K) + h_X(K)}$.

Many applications deal with random vectors only through their entropy profiles:

Synthetic geometry for random variables

Entropy profile encodes qualitative information about the system of random variables:

- Subvector X_I is functionally dependent on X_K if and only if $\boldsymbol{h_X(I \cup K) = h_X(K)}$.
- Subvectors X_I and X_J are conditionally independent given X_K if and only if $\boldsymbol{h_X(I \cup K) + h_X(J \cup K) = h_X(I \cup J \cup K) + h_X(K)}$.

Many applications deal with random vectors only through their entropy profiles:

- Graphical models in statistics and causality are defined by CI assumptions (e.g., Bayesian networks and d-separation in graphs).

Synthetic geometry for random variables

Entropy profile encodes qualitative information about the system of random variables:

- Subvector X_I is functionally dependent on X_K if and only if $\boldsymbol{h_X(I \cup K) = h_X(K)}$.
- Subvectors X_I and X_J are conditionally independent given X_K if and only if $\boldsymbol{h_X(I \cup K) + h_X(J \cup K) = h_X(I \cup J \cup K) + h_X(K)}$.

Many applications deal with random vectors only through their entropy profiles:

- Graphical models in statistics and causality are defined by CI assumptions (e.g., Bayesian networks and d-separation in graphs).
- Cryptographic protocols use FD and CI constraints to specify operation and information-theoretic security (e.g., secret sharing).

Synthetic geometry for random variables

Entropy profile encodes qualitative information about the system of random variables:

- Subvector X_I is functionally dependent on X_K if and only if $\boldsymbol{h_X(I \cup K) = h_X(K)}$.
- Subvectors X_I and X_J are conditionally independent given X_K if and only if $\boldsymbol{h_X(I \cup K) + h_X(J \cup K) = h_X(I \cup J \cup K) + h_X(K)}$.

Many applications deal with random vectors only through their entropy profiles:

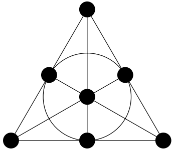
- Graphical models in statistics and causality are defined by CI assumptions (e.g., Bayesian networks and d-separation in graphs).
- Cryptographic protocols use FD and CI constraints to specify operation and information-theoretic security (e.g., secret sharing).
- Quantities in information theory are defined by linear optimization over entropy profiles with FD and CI constraints (e.g., common information).

Example: Perfect secret sharing

- Given: participants $N = \{1, \ldots, n\}$ and a set of qualified subsets $\mathscr{Q} \subseteq 2^N$.

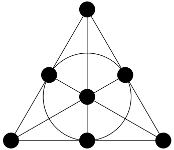
Example: Perfect secret sharing

- Given: participants $N = \{1, \ldots, n\}$ and a set of qualified subsets $\mathscr{Q} \subseteq 2^N$.
- Devise a scheme (a system of random variables) to distribute shares s_p of a randomly generated secret s to the participants such that



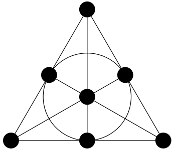
Example: Perfect secret sharing

- Given: participants $N = \{1, \ldots, n\}$ and a set of qualified subsets $\mathscr{Q} \subseteq 2^N$.
- Devise a scheme (a system of random variables) to distribute shares s_p of a randomly generated secret s to the participants such that
  - s_p is a function of s,



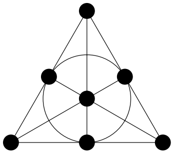
Example: Perfect secret sharing

- Given: participants $N = \{1, \ldots, n\}$ and a set of qualified subsets $\mathscr{Q} \subseteq 2^N$.
- Devise a scheme (a system of random variables) to distribute shares s_p of a randomly generated secret s to the participants such that
  - s_p is a function of s,
  - s is a function of $s_A = (s_p : p \in A)$ whenever $A \in \mathscr{Q}$,



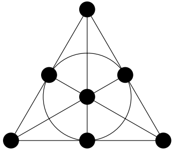
Example: Perfect secret sharing

- Given: participants $N = \{1, \ldots, n\}$ and a set of qualified subsets $\mathscr{Q} \subseteq 2^N$.
- Devise a scheme (a system of random variables) to distribute shares s_p of a randomly generated secret s to the participants such that
  - s_p is a function of s,
  - s is a function of $s_A = (s_p : p \in A)$ whenever $A \in \mathscr{Q}$,
  - s is independent of s_B whenever $B \notin \mathscr{Q}$.



Example: Perfect secret sharing

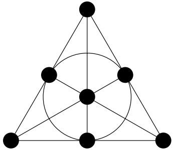
- Given: participants $N = \{1, \ldots, n\}$ and a set of qualified subsets $\mathscr{Q} \subseteq 2^N$.
- Devise a scheme (a system of random variables) to distribute shares s_p of a randomly generated secret s to the participants such that
  - s_p is a function of s,
  - s is a function of $s_A = (s_p : p \in A)$ whenever $A \in \mathscr{Q}$,
  - s is independent of s_B whenever $B \notin \mathscr{Q}$.



- The information ratio is $\sigma(h) = {}^1/_{h(s)} \max \{h(p) : p \in N\}$.

Example: Perfect secret sharing

- Given: participants $N = \{1, \ldots, n\}$ and a set of qualified subsets $\mathscr{Q} \subseteq 2^N$.
- Devise a scheme (a system of random variables) to distribute shares s_p of a randomly generated secret s to the participants such that
  - s_p is a function of s,
  - s is a function of $s_A = (s_p : p \in A)$ whenever $A \in \mathscr{Q}$,
  - s is independent of s_B whenever $B \notin \mathscr{Q}$.



- The information ratio is $\sigma(h) = {}^{1}\!/_{h(s)} \max \{ h(p) : p \in N \}$.
- The optimal information ratio $\sigma(\mathscr{Q}) = \inf \{\sigma(h) : h \models \mathscr{Q}\}$ can be determined by linear optimization over the set of entropy profiles.

The entropy region and information inequalities

Let $\mathbf{H}_N^* \subseteq \mathbb{R}^{2^N}$ consist of all h_X where X is an N-variate discrete random vector. $\mathbf{H}_N^*$ is the image of $\bigcup_{d_1=1}^{\infty} \cdots \bigcup_{d_n=1}^{\infty} \Delta(d_1, \ldots, d_n)$ under the transcendental map $X \mapsto h_X$.

The entropy region and information inequalities

Let $\mathbf{H}_N^* \subseteq \mathbb{R}^{2^N}$ consist of all h_X where X is an N-variate discrete random vector. $\mathbf{H}_N^*$ is the image of $\bigcup_{d_1=1}^{\infty} \cdots \bigcup_{d_n=1}^{\infty} \Delta(d_1, \ldots, d_n)$ under the transcendental map $X \mapsto h_X$.

Theorem

$\overline{\mathbf{H}_N^*}$ *is a convex cone of dimension* $2^N - 1$. *Furthermore* $\operatorname{relint}(\overline{\mathbf{H}_N^*}) \subseteq \mathbf{H}_N^*$.

The entropy region and information inequalities

Let $\mathbf{H}_N^* \subseteq \mathbb{R}^{2^N}$ consist of all h_X where X is an N-variate discrete random vector. $\mathbf{H}_N^*$ is the image of $\bigcup_{d_1=1}^{\infty} \cdots \bigcup_{d_n=1}^{\infty} \Delta(d_1, \ldots, d_n)$ under the transcendental map $X \mapsto h_X$.

Theorem

$\overline{\mathbf{H}_N^*}$ is a convex cone of dimension $2^N - 1$. Furthermore $\operatorname{relint}(\overline{\mathbf{H}_N^*}) \subseteq \mathbf{H}_N^*$.

- Linear optimization is well-behaved! Elements of the dual cone (linear information inequalities) can give bounds for optimization problems.

The entropy region and information inequalities

Let $\mathbf{H}_N^* \subseteq \mathbb{R}^{2^N}$ consist of all h_X where X is an N-variate discrete random vector. $\mathbf{H}_N^*$ is the image of $\bigcup_{d_1=1}^{\infty} \cdots \bigcup_{d_n=1}^{\infty} \Delta(d_1, \ldots, d_n)$ under the transcendental map $X \mapsto h_X$.

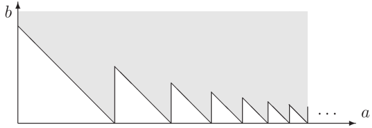
Theorem

$\overline{\mathbf{H}_N^*}$ *is a convex cone of dimension* $2^N - 1$*. Furthermore* $\operatorname{relint}(\overline{\mathbf{H}_N^*}) \subseteq \mathbf{H}_N^*$.

- Linear optimization is well-behaved! Elements of the dual cone (linear information inequalities) can give bounds for optimization problems.

Problem

Find a description of the boundary of $\mathbf{H}_3^*$.



Shannon inequalities

- A function $h : 2^N \to \mathbb{R}$ is a polymatroid if
  - $h(\emptyset) = 0$,
  - $h(I \mid K) := h(I \cup K) - h(K) \geq 0$ ("$=$" is FD).
  - $h(I : J \mid K) := h(I \cup K) + h(J \cup K) - h(I \cup J \cup K) - h(K) \geq 0$ ("$=$" is CI).

Shannon inequalities

- A function $h : 2^N \to \mathbb{R}$ is a polymatroid if
  - $h(\emptyset) = 0$,
  - $h(I \mid K) := h(I \cup K) - h(K) \geq 0$ ("=" is FD).
  - $h(I : J \mid K) := h(I \cup K) + h(J \cup K) - h(I \cup J \cup K) - h(K) \geq 0$ ("=" is CI).

- The set $\mathbf{P}_N$ of polymatroids is a polyhedral cone in $\mathbb{R}^{2^N}$ and $\mathbf{P}_N \supseteq \overline{\mathbf{H}_N^*} \to$ ITIP.

Shannon inequalities

- A function $h : 2^N \to \mathbb{R}$ is a polymatroid if
  - $h(\emptyset) = 0$,
  - $h(I \mid K) := h(I \cup K) - h(K) \geq 0$ ("$=$" is FD).
  - $h(I : J \mid K) := h(I \cup K) + h(J \cup K) - h(I \cup J \cup K) - h(K) \geq 0$ ("$=$" is CI).
- The set $\mathbf{P}_N$ of polymatroids is a polyhedral cone in $\mathbb{R}^{2^N}$ and $\mathbf{P}_N \supseteq \overline{\mathbf{H}_N^*} \to$ ITIP.
- Elements of the dual cone of $\mathbf{P}_N$ are the Shannon inequalities.

Shannon inequalities

- A function $h : 2^N \to \mathbb{R}$ is a polymatroid if
  - $h(\emptyset) = 0$,
  - $h(I \mid K) := h(I \cup K) - h(K) \geq 0$ ("$=$" is FD).
  - $h(I : J \mid K) := h(I \cup K) + h(J \cup K) - h(I \cup J \cup K) - h(K) \geq 0$ ("$=$" is CI).
- The set $\mathbf{P}_N$ of polymatroids is a polyhedral cone in $\mathbb{R}^{2^N}$ and $\mathbf{P}_N \supseteq \overline{\mathbf{H}_N^*} \to$ ITIP.
- Elements of the dual cone of $\mathbf{P}_N$ are the Shannon inequalities.
- FD and CI constraints correspond to faces of $\mathbf{P}_N \leftarrow$ LP over $\overline{\mathbf{H}_N^*}$.

Shannon inequalities

- A function $h : 2^N \to \mathbb{R}$ is a polymatroid if
  - $h(\emptyset) = 0$,
  - $h(I \mid K) := h(I \cup K) - h(K) \geq 0$ ("$=$" is FD).
  - $h(I : J \mid K) := h(I \cup K) + h(J \cup K) - h(I \cup J \cup K) - h(K) \geq 0$ ("$=$" is CI).

- The set $\mathbf{P}_N$ of polymatroids is a polyhedral cone in $\mathbb{R}^{2^N}$ and $\mathbf{P}_N \supseteq \overline{\mathbf{H}_N^*} \to$ ITIP.
- Elements of the dual cone of $\mathbf{P}_N$ are the Shannon inequalities.
- FD and CI constraints correspond to faces of $\mathbf{P}_N \leftarrow$ LP over $\overline{\mathbf{H}_N^*}$.

Theorem ([Mat07])

$\overline{\mathbf{H}_N^*}$ *is not polyhedral for* $|N| \geq 4$.

Conditional Ingleton inequalities

A conditional information inequality is an inequality valid only on a linear slice of $\mathbf{H}_N^*$.

Conditional Ingleton inequalities

A conditional information inequality is an inequality valid only on a linear slice of $\mathbf{H}_N^*$.

Theorem ([KR13] & [Stu21] & [Boe23])

Up to symmetry there are precisely ten minimal sets of conditional independence assumptions on four random variables which ensure $\texttt{Ingleton} \geq 0$.

Conditional Ingleton inequalities

A conditional information inequality is an inequality valid only on a linear slice of $\mathbf{H}_N^*$.

Theorem ([KR13] & [Stu21] & [Boe23])

Up to symmetry there are precisely ten minimal sets of conditional independence assumptions on four random variables which ensure $\texttt{Ingleton} \geq 0$.

Corollary (Which faces of $\mathbf{P}_N$ have entropic points on them?)

On four discrete random variables there are precisely 18 478 *realizable conditional independence structures. (For general N this problem is undecidable!)*

Conditional Ingleton inequalities

A conditional information inequality is an inequality valid only on a linear slice of $\mathbf{H}_N^*$.

Theorem ([KR13] & [Stu21] & [Boe23])

Up to symmetry there are precisely ten minimal sets of conditional independence assumptions on four random variables which ensure $\texttt{Ingleton} \geq 0$.

Corollary (Which faces of $\mathbf{P}_N$ have entropic points on them?)

On four discrete random variables there are precisely 18 478 *realizable conditional independence structures. (For general N this problem is undecidable!)*

Problem

Which of these inequalities hold on $\overline{\mathbf{H}_4^*}$*? (Some do, some don't . . .)*

Generating new inequalities: Extension properties

All widely used polyhedral outer approximations to $\overline{\mathbf{H}_N^*}$ which improve upon $\mathbf{P}_N$ are derived from an extension property

Generating new inequalities: Extension properties

All widely used polyhedral outer approximations to $\overline{\mathbf{H}_N^*}$ which improve upon $\mathbf{P}_N$ are derived from an extension property which is a theorem of the form:

- If $h \in \overline{\mathbf{H}_N^*}$, then there exists $\overline{h} \in \overline{\mathbf{H}_M^*}$ for some $M \supseteq N$ such that $\overline{h}|_N = h$ and some other linear conditions $\varphi(\overline{h}) \geq 0$ hold.

Generating new inequalities: Extension properties

All widely used polyhedral outer approximations to $\overline{\mathbf{H}_N^*}$ which improve upon $\mathbf{P}_N$ are derived from an extension property which is a theorem of the form:

- If $h \in \overline{\mathbf{H}_N^*}$, then there exists $\overline{h} \in \overline{\mathbf{H}_M^*}$ for some $M \supseteq N$ such that $\overline{h}|_N = h$ and some other linear conditions $\varphi(\overline{h}) \geq 0$ hold.
- The extension property is encapsulated in its cone $E_N^M = \{ \varphi(\overline{h}) \geq 0 \}$.

Generating new inequalities: Extension properties

All widely used polyhedral outer approximations to $\overline{\mathbf{H}^*_N}$ which improve upon $\mathbf{P}_N$ are derived from an extension property which is a theorem of the form:

- If $h \in \overline{\mathbf{H}^*_N}$, then there exists $\overline{h} \in \overline{\mathbf{H}^*_M}$ for some $M \supseteq N$ such that $\overline{h}|_N = h$ and some other linear conditions $\varphi(\overline{h}) \geq 0$ hold.
- The extension property is encapsulated in its cone $E^M_N = \{\, \varphi(\overline{h}) \geq 0 \,\}$.

Extension principle: Let E^M_N be the cone of an extension property and $\pi^M_N : \mathbb{R}^{2^M} \to \mathbb{R}^{2^N}$ the canonical projection. Then $\overline{\mathbf{H}^*_N} = \pi^M_N(E^M_N \cap \boxed{\overline{\mathbf{H}^*_M}})$.

Generating new inequalities: Extension properties

All widely used polyhedral outer approximations to $\overline{\mathbf{H}^*_N}$ which improve upon $\mathbf{P}_N$ are derived from an extension property which is a theorem of the form:

- If $h \in \overline{\mathbf{H}^*_N}$, then there exists $\overline{h} \in \overline{\mathbf{H}^*_M}$ for some $M \supseteq N$ such that $\overline{h}|_N = h$ and some other linear conditions $\varphi(\overline{h}) \geq 0$ hold.
- The extension property is encapsulated in its cone $E^M_N = \{\, \varphi(\overline{h}) \geq 0 \,\}$.

Extension principle: Let E^M_N be the cone of an extension property and $\pi^M_N : \mathbb{R}^{2^M} \to \mathbb{R}^{2^N}$ the canonical projection. Then $\overline{\mathbf{H}^*_N} = \pi^M_N(E^M_N \cap \boxed{\overline{\mathbf{H}^*_M}})$.

Relax: $\overline{\mathbf{H}^*_N} \subseteq \pi^M_N(E^M_N \cap \boxed{\mathbf{P}_M})$.

Extension properties: Copy lemma

- Consider $h \in \mathbf{P}_N$ and pick any $L \subseteq N$.
- An *L-copy* of N is a set M with $|N| = |M|$ and $N \cap M = L$ with a bijection $\sigma : N \to M$ fixing L pointwise.

Extension properties: Copy lemma

- Consider $h \in \mathbf{P}_N$ and pick any $L \subseteq N$.
- An L-copy of N is a set M with $|N| = |M|$ and $N \cap M = L$ with a bijection $\sigma : N \to M$ fixing L pointwise. This induces an L-copy of h: $\sigma(h) \in \mathbf{P}_M$.

Extension properties: Copy lemma

- Consider $h \in \mathbf{P}_N$ and pick any $L \subseteq N$.
- An L-copy of N is a set M with $|N| = |M|$ and $N \cap M = L$ with a bijection $\sigma : N \to M$ fixing L pointwise. This induces an L-copy of h: $\sigma(h) \in \mathbf{P}_M$.

The Copy lemma states:

Extension properties: Copy lemma

- Consider $h \in \mathbf{P}_N$ and pick any $L \subseteq N$.
- An L-copy of N is a set M with $|N| = |M|$ and $N \cap M = L$ with a bijection $\sigma : N \to M$ fixing L pointwise. This induces an L-copy of h: $\sigma(h) \in \mathbf{P}_M$.

The Copy lemma states:

- Let $h \in \overline{\mathbf{H}^*_N}$ and $L \subseteq N$, fix an L-copy $\sigma : N \to M$ of N.

Extension properties: Copy lemma

- Consider $h \in \mathbf{P}_N$ and pick any $L \subseteq N$.
- An L-copy of N is a set M with $|N| = |M|$ and $N \cap M = L$ with a bijection $\sigma : N \to M$ fixing L pointwise. This induces an L-copy of h: $\sigma(h) \in \mathbf{P}_M$.

The Copy lemma states:

- Let $h \in \overline{\mathbf{H}_N^*}$ and $L \subseteq N$, fix an L-copy $\sigma : N \to M$ of N.
- There exists $\overline{h} \in \boxed{\overline{\mathbf{H}_{NM}^*}}$ such that

$$\overline{h}|_N = h, \quad \overline{h}|_M = \sigma(h), \quad \overline{h}(N : M \mid L) = 0.$$

Extension properties: Copy lemma

- Consider $h \in \mathbf{P}_N$ and pick any $L \subseteq N$.
- An L-copy of N is a set M with $|N| = |M|$ and $N \cap M = L$ with a bijection $\sigma : N \to M$ fixing L pointwise. This induces an L-copy of h: $\sigma(h) \in \mathbf{P}_M$.

The Copy lemma states:

- Let $h \in \overline{\mathbf{H}^*_N}$ and $L \subseteq N$, fix an L-copy $\sigma : N \to M$ of N.
- There exists $\overline{h} \in \boxed{\overline{\mathbf{H}^*_{NM}}}$ such that

$$\overline{h}|_N = h, \quad \overline{h}|_M = \sigma(h), \quad \overline{h}(N : M \mid L) = 0.$$

- Relaxation: only require $\overline{h} \in \boxed{\mathbf{P}_{NM}}$! This gives a tighter outer bound than $\mathbf{P}_N$:

$$\mathbf{P}_N \supseteq \bigcap_{L \subseteq N} \mathbf{Copy}^L_N \supseteq \overline{\mathbf{H}^*_N}.$$

Using extension properties

To derive new information inequalities [DFZ11] and many more:

Using extension properties

To derive new information inequalities [DFZ11] and many more:

- Take any polyhedral cone $\mathbf{Q}_N \supseteq \overline{\mathbf{H}_N^*}$.

Using extension properties

To derive new information inequalities [DFZ11] and many more:

- Take any polyhedral cone $\mathbf{Q}_N \supseteq \overline{\mathbf{H}_N^*}$.
- Replace $\overline{\mathbf{H}_\bullet^*}$ in an extension property with $\mathbf{Q}_\bullet$.

Using extension properties

To derive new information inequalities [DFZ11] and many more:

- Take any polyhedral cone $\mathbf{Q}_N \supseteq \overline{\mathbf{H}_N^*}$.
- Replace $\overline{\mathbf{H}_\bullet^*}$ in an extension property with $\mathbf{Q}_\bullet$.
- Project to obtain tighter polyhedral cone $\mathbf{Q}'_N \supseteq \overline{\mathbf{H}_N^*}$.

Using extension properties

To derive new information inequalities [DFZ11] and many more:

- Take any polyhedral cone $\mathbf{Q}_N \supseteq \overline{\mathbf{H}_N^*}$.
- Replace $\overline{\mathbf{H}_\bullet^*}$ in an extension property with $\mathbf{Q}_\bullet$.
- Project to obtain tighter polyhedral cone $\mathbf{Q}'_N \supseteq \overline{\mathbf{H}_N^*}$.
- Mix and iterate different extension properties.

Using extension properties

To derive new information inequalities [DFZ11] and many more:

- Take any polyhedral cone $\mathbf{Q}_N \supseteq \overline{\mathbf{H}_N^*}$.
- Replace $\overline{\mathbf{H}_\bullet^*}$ in an extension property with $\mathbf{Q}_\bullet$.
- Project to obtain tighter polyhedral cone $\mathbf{Q}'_N \supseteq \overline{\mathbf{H}_N^*}$.
- Mix and iterate different extension properties.
- Exact polyhedral computations certify validity of new inequalities.

Using extension properties

To derive new information inequalities [DFZ11] and many more:

- Take any polyhedral cone $\mathbf{Q}_N \supseteq \overline{\mathbf{H}^*_N}$.
- Replace $\overline{\mathbf{H}^*_\bullet}$ in an extension property with $\mathbf{Q}_\bullet$.
- Project to obtain tighter polyhedral cone $\mathbf{Q}'_N \supseteq \overline{\mathbf{H}^*_N}$.
- Mix and iterate different extension properties.
- Exact polyhedral computations certify validity of new inequalities.

To disprove information inequalities [KR13]:

Using extension properties

To derive new information inequalities [DFZ11] and many more:

- Take any polyhedral cone $\mathbf{Q}_N \supseteq \overline{\mathbf{H}_N^*}$.
- Replace $\overline{\mathbf{H}_\bullet^*}$ in an extension property with $\mathbf{Q}_\bullet$.
- Project to obtain tighter polyhedral cone $\mathbf{Q}'_N \supseteq \overline{\mathbf{H}_N^*}$.
- Mix and iterate different extension properties.
- Exact polyhedral computations certify validity of new inequalities.

To disprove information inequalities [KR13]:

- Take an entropy profile h.

Using extension properties

To derive new information inequalities [DFZ11] and many more:

- Take any polyhedral cone $\mathbf{Q}_N \supseteq \overline{\mathbf{H}_N^*}$.
- Replace $\overline{\mathbf{H}_\bullet^*}$ in an extension property with $\mathbf{Q}_\bullet$.
- Project to obtain tighter polyhedral cone $\mathbf{Q}'_N \supseteq \overline{\mathbf{H}_N^*}$.
- Mix and iterate different extension properties.
- Exact polyhedral computations certify validity of new inequalities.

To disprove information inequalities [KR13]:

- Take an entropy profile h.
- Apply a sequence of extension properties to $h \to$ polyhedron Q.

Using extension properties

To derive new information inequalities [DFZ11] and many more:

- Take any polyhedral cone $\mathbf{Q}_N \supseteq \overline{\mathbf{H}_N^*}$.
- Replace $\overline{\mathbf{H}_\bullet^*}$ in an extension property with $\mathbf{Q}_\bullet$.
- Project to obtain tighter polyhedral cone $\mathbf{Q}'_N \supseteq \overline{\mathbf{H}_N^*}$.
- Mix and iterate different extension properties.
- Exact polyhedral computations certify validity of new inequalities.

To disprove information inequalities [KR13]:

- Take an entropy profile h.
- Apply a sequence of extension properties to $h \to$ polyhedron Q.
- If every point in Q violates an inequality, it cannot be valid.

Using extension properties

To derive new information inequalities [DFZ11] and many more:

- Take any polyhedral cone $\mathbf{Q}_N \supseteq \overline{\mathbf{H}^*_N}$.
- Replace $\overline{\mathbf{H}^*_\bullet}$ in an extension property with $\mathbf{Q}_\bullet$.
- Project to obtain tighter polyhedral cone $\mathbf{Q}'_N \supseteq \overline{\mathbf{H}^*_N}$.
- Mix and iterate different extension properties.
- Exact polyhedral computations certify validity of new inequalities.

To disprove information inequalities [KR13]:

- Take an entropy profile h.
- Apply a sequence of extension properties to $h \to$ polyhedron Q.
- If every point in Q violates an inequality, it cannot be valid.
- Exact Farkas certificate for invalidity.

Outlook

- There exist more extension properties: Ahlswede–Körner, Slepian–Wolf, ...

Outlook

- There exist more extension properties: Ahlswede–Körner, Slepian–Wolf, ...
- The same concept applies to algebraic matroids (subset of $\overline{\mathbf{H}^*}$): Dress–Lovász.

Outlook

- There exist more extension properties: Ahlswede–Körner, Slepian–Wolf, ...
- The same concept applies to algebraic matroids (subset of $\overline{\mathbf{H}^*}$): Dress–Lovász.
- Over 200 information inequalities and several infinite families are derived from the Copy lemma alone. They have been tabulated but are not reusable data.

Rule [43] Given:

$$
\begin{aligned}
& aI(A;B) \\
\leq\ & bI(A;B|C) + cI(A;C|B) + zI(B;C|A) \\
+\ & eI(A;B|D) + fI(A;D|B) \\
+\ & (b' + d' + z)I(B;D|A) + hI(C;D) \\
+\ & iI(C;D|A) + zI(C;D|B)
\end{aligned}
$$

and

$$
\begin{aligned}
& a'I(A;B) \\
\leq\ & b'I(A;B|C) + c'I(A;C|B) + d'I(B;C|A) \\
+\ & e'I(A;B|D) + f'I(A;D|B) + g'I(B;D|A) \\
+\ & h'I(C;D) + i'I(C;D|A) + j'I(C;D|B)
\end{aligned}
$$

Get:

$$
\begin{aligned}
& (a + a' + z)I(A;B) \\
\leq\ & (a + b + c + f + b' + 2z)I(A;B|C) \\
+\ & (-a + b + c + e + c' + z)I(A;C|B) \\
+\ & (d' + z)I(B;C|A) + (e + e' + z)I(A;B|D) \\
+\ & (f + f')I(A;D|B) \\
+\ & (-a' + b' + e' + g' + i')I(B;D|A) \\
+\ & (h + h' + z)I(C;D) + (i + i')I(C;D|A) \\
+\ & (j')I(C;D|B)
\end{aligned}
$$

Using: RS is copy of CD over AB
Substitutions: A C R S; AD B R S

Abbreviated Proof of (75): T: D-copy of A over BCRS.
L1: -a.c. +c.d. +r.cd.a +c.s.a +b.d.s +a.bs.d +2a.cr.bs +a.bs.cr +d.r.abcs +d.s.abcr
SL1: d.t.a +c.d.t +a.t.cd +c.r.t +a.t.cr +d.r.act +b.t.acdr +a.t.bs +c.s.at +b.t.acs +d.t.s +a.s.dt +b.d.ast +c.t.abds +a.r.bcst +r.ad.bcst +s.ad.bcrt +d.t.abcrs C2L1: 3t.ad.bcrs
S: C-copy of A over BDR.
L2: -2a.c. +2c.d. +a.b.cr +2a.c.br +c.ar.b +a.b.dr +4a.d.br +2a.br.d +2d.br.a +2r.cd.a +d.r.abc
SL2: c.s.b +a.b.cs +c.d.s +a.s.cd +d.s.abc +3a.s.br +3c.s.br +c.r.abs +d.r.s +a.s.dr +d.r.abs +d.br.as +c.r.ads +b.s.acdr +2c.s.abdr +2d.s.abcr
C2L2: 7s.ac.bdr
R: D-copy of C over AB.
S: c.r.a +3c.r.b +d.r.a +7d.r.b +c.d.r +2b.r.acd +r.ab.cd +9c.r.abd +3d.r.abc
C2: 16r.cd.ab

Outlook

- There exist more extension properties: Ahlswede–Körner, Slepian–Wolf, ...
- The same concept applies to algebraic matroids (subset of $\overline{\mathbf{H}^*}$): Dress–Lovász.
- Over 200 information inequalities and several infinite families are derived from the Copy lemma alone. They have been tabulated but are not reusable data.

Outlook

- There exist more extension properties: Ahlswede–Körner, Slepian–Wolf, ...
- The same concept applies to algebraic matroids (subset of $\overline{\mathbf{H}^*}$): Dress–Lovász.
- Over 200 information inequalities and several infinite families are derived from the Copy lemma alone. They have been tabulated but are not reusable data.
- Want a framework to combine and iterate extension properties based on polyhedra and linear programming and certificates for the validity of information inequalities.

Outlook

- There exist more extension properties: Ahlswede–Körner, Slepian–Wolf, ...
- The same concept applies to algebraic matroids (subset of $\overline{\mathbf{H}^*}$): Dress–Lovász.
- Over 200 information inequalities and several infinite families are derived from the Copy lemma alone. They have been tabulated but are not reusable data.
- Want a framework to combine and iterate extension properties based on polyhedra and linear programming and certificates for the validity of information inequalities.

Thank you!

Supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101110545.

Funded by the European Union

References

[BFP24] Michael Bamiloshin, Oriol Farràs, and Carles Padró. *A Note on Extension Properties and Representations of Matroids*. 2024. arXiv: 2306.15085 [math.CO].

[Boe23] Tobias Boege. "No Eleventh Conditional Ingleton Inequality". In: *Experimental Mathematics* (2023). DOI: 10.1080/10586458.2023.2294827.

[DFZ11] Randall Dougherty, Chris Freiling, and Kenneth Zeger. *Non-Shannon Information Inequalities in Four Random Variables*. 2011. arXiv: 1104.3602v1 [cs.IT].

[KR13] Tarik Kaced and Andrei Romashchenko. "Conditional information inequalities for entropic and almost entropic points". In: *IEEE Trans. Inf. Theory* 59.11 (2013), pp. 7149–7167. DOI: 10.1109/TIT.2013.2274614.

[Mat06] František Matúš. "Piecewise linear conditional information inequality". In: *IEEE Trans. Inf. Theory* 52.1 (2006), pp. 236–238. DOI: 10.1109/TIT.2005.860438.

[Mat07] František Matúš. "Infinitely many information inequalities". In: *Proc. IEEE ISIT 2007.* 2007, pp. 41–44.

[Stu21] Milan Studený. "Conditional independence structures over four discrete random variables revisited: conditional ingleton inequalities". In: *IEEE Trans. Inf. Theory* 67.11 (2021), pp. 7030–7049. DOI: 10.1109/TIT.2021.3104250.