The Ingleton inequality for random variables

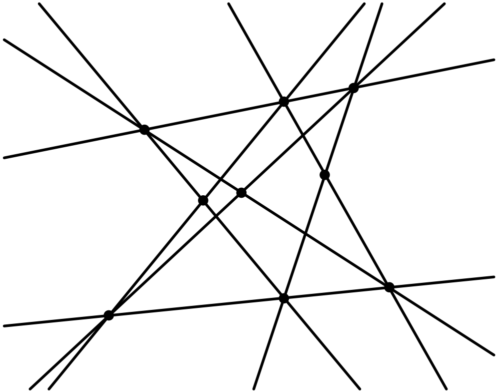
Tobias Boege

Department of Mathematics
KTH Royal Institute of Technology, Sweden

Combinatorial Coworkspace
Kleinwalsertal, 22 March 2024

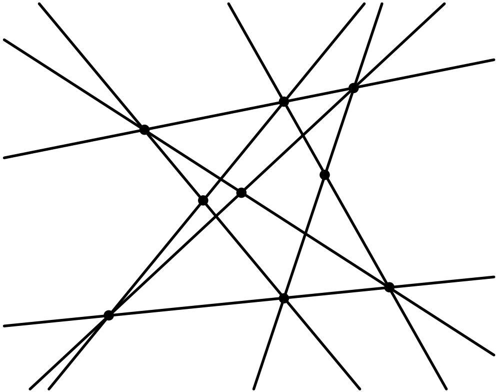
Matroids

- Matroids are combinatorial structures which model “special position” relations in geometry.



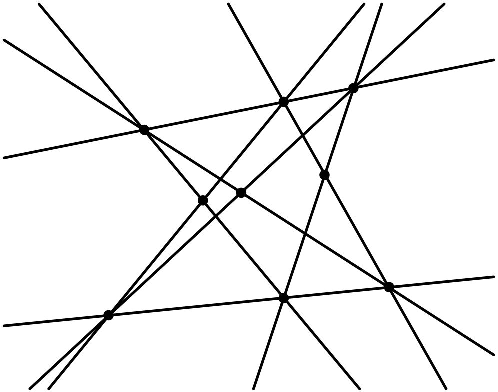
Matroids

- Matroids are combinatorial structures which model "special position" relations in geometry.
  - For example the matroid of a set of points in the projective plane records which triples of points lie on a line.

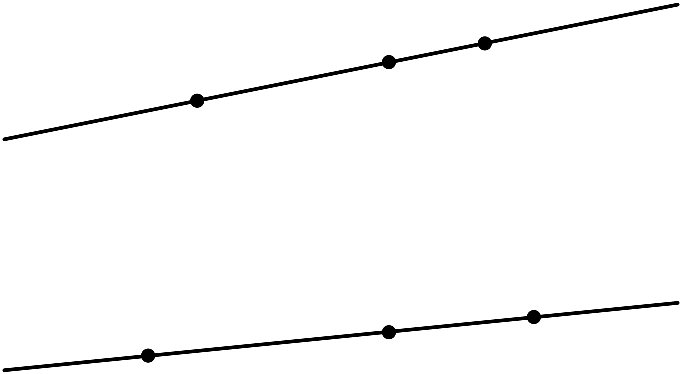


Matroids

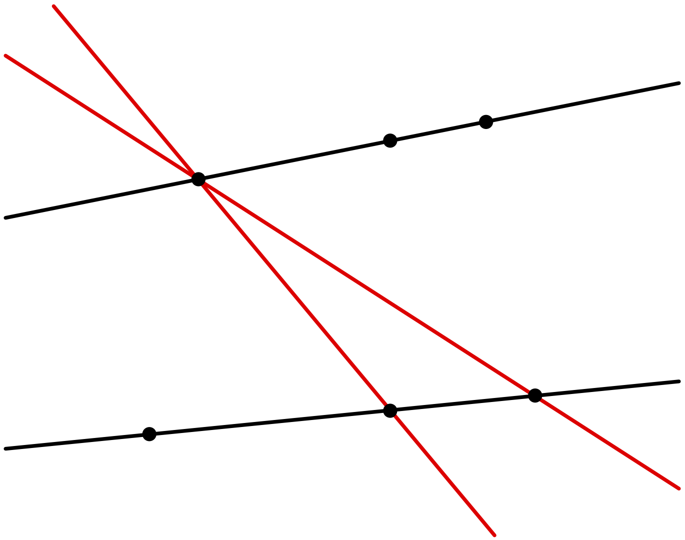
- Matroids are combinatorial structures which model "special position" relations in geometry.
  - For example the matroid of a set of points in the projective plane records which triples of points lie on a line.
- Non-realizability of matroids captures the (non-obvious) laws of projective geometry.



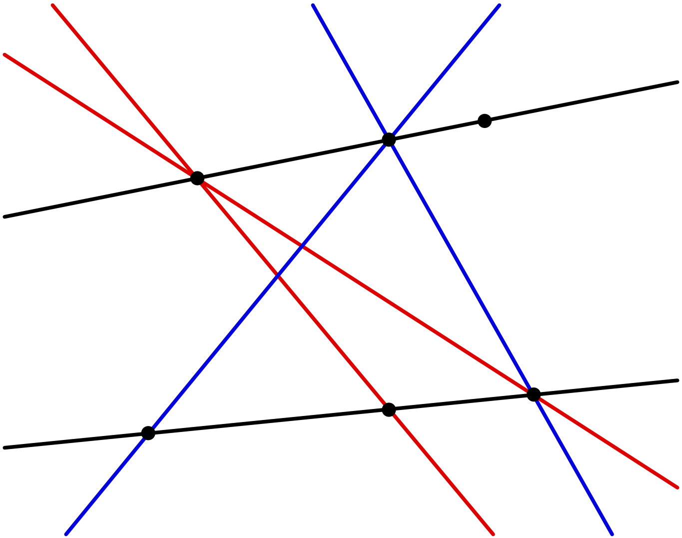
Laws of geometry



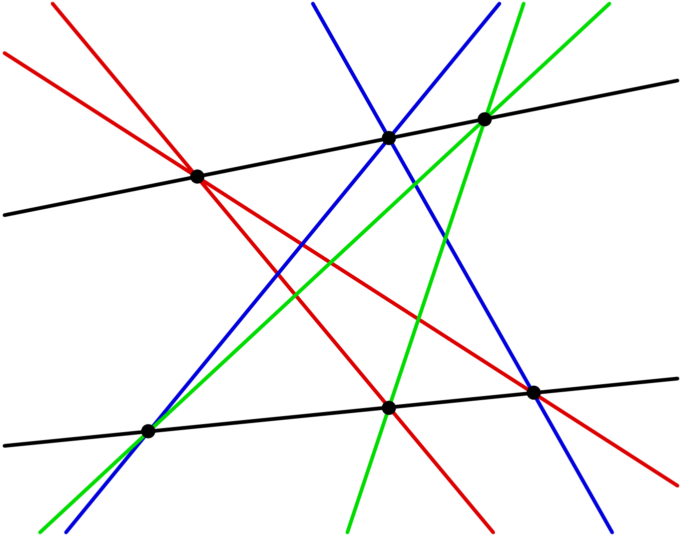
Laws of geometry



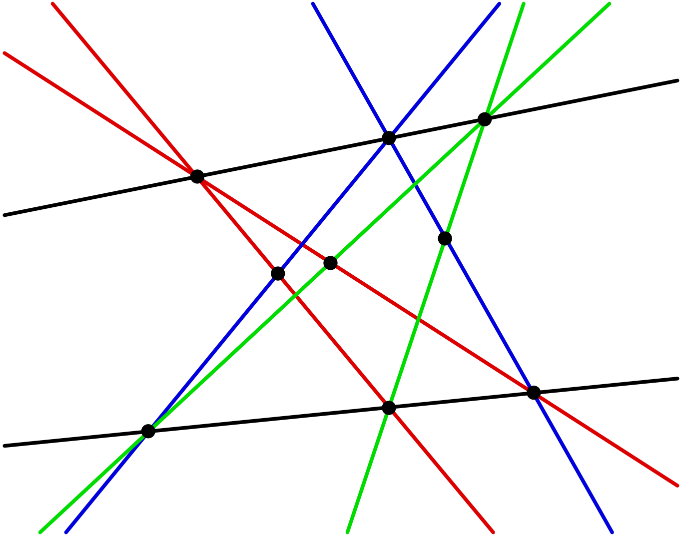
Laws of geometry



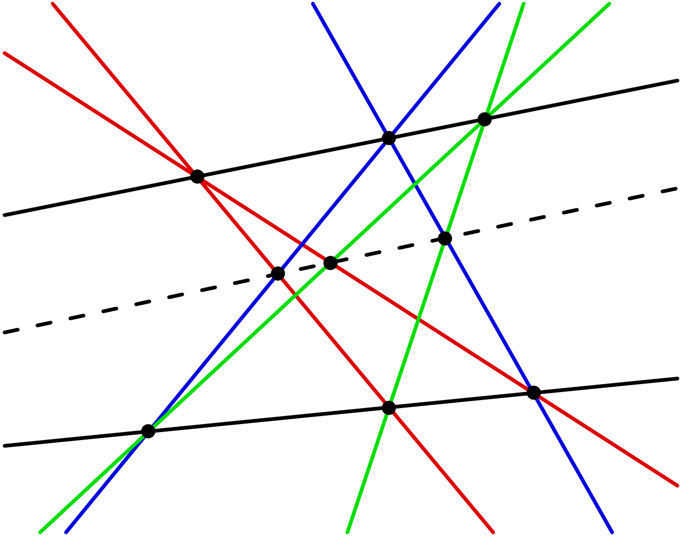
Laws of geometry



Laws of geometry



Laws of geometry



Entropy

Let X be a random variable taking finitely many values $\{1, \ldots, d\}$ with positive probabilities. Its *Shannon entropy* is

$$H(X) := \sum_{i=1}^{d} p(X = i) \log {}^{1}\!/_{p(X = i)}.$$

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

Entropy

Let X be a random variable taking finitely many values $\{1, \ldots, d\}$ with positive probabilities. Its *Shannon entropy* is

$$H(X) := \sum_{i=1}^{d} p(X = i) \log {}^{1}/_{p(X=i)}.$$

- H is continuous on $\Delta(d)$ and analytic on the interior.
- A random vector $X \in \Delta(d_1, \ldots, d_n)$ is a random variable in $\Delta(\prod_{i=1}^{n} d_i)$, so the definition of H extends to vectors.
- For a random vector $X = (X_1, \ldots, X_n)$ 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

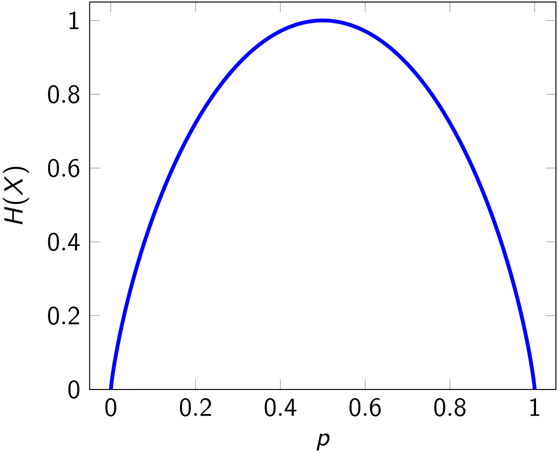


Figure: Entropy of a binary random variable X as a function of $p = p(X = \text{heads})$.

Independence: geometry $\leftrightarrow$ information theory

Information-theoretical "special position" properties of discrete random variables can be formulated in terms of linear functionals on the entropy profile h_X:

Independence: geometry $\leftrightarrow$ information theory

Information-theoretical "special position" properties of discrete random variables can be formulated in terms of linear functionals on the entropy profile h_X:

Rank condition	Matroid concept	Information theory concept
$h(x_i) = 0$	loop	constant random variable
$h(x_i, X_{[n]\setminus i}) = h(x_i) + h(X_{[n]\setminus i})$	coloop	max. private information

Independence: geometry $\leftrightarrow$ information theory

Information-theoretical "special position" properties of discrete random variables can be formulated in terms of linear functionals on the entropy profile h_X:

Rank condition	Matroid concept	Information theory concept
$h(x_i) = 0$	loop	constant random variable
$h(x_i, X_{[n]\setminus i}) = h(x_i) + h(X_{[n]\setminus i})$	coloop	max. private information
$h(x_i, X_K) = h(X_K)$	closure operator	functional dependence
$h(X_{K\cup L}) = h(X_K) + h(X_L)$	independent set	total independence
$h(x_i, x_j \mid X_K) = h(x_i \mid X_K) + h(x_j \mid X_K)$	modular pair	conditional independence

Independence: geometry $\leftrightarrow$ information theory

Information-theoretical “special position” properties of discrete random variables can be formulated in terms of linear functionals on the entropy profile h_X:

Rank condition	Matroid concept	Information theory concept
$h(x_i) = 0$	loop	constant random variable
$h(x_i, X_{[n]\setminus i}) = h(x_i) + h(X_{[n]\setminus i})$	coloop	max. private information
$h(x_i, X_K) = h(X_K)$	closure operator	functional dependence
$h(X_{K\cup L}) = h(X_K) + h(X_L)$	independent set	total independence
$h(x_i, x_j \mid X_K) = h(x_i \mid X_K) + h(x_j \mid X_K)$	modular pair	conditional independence

Even though entropy is a transcendental function, many of these conditions are **polynomial** in the probabilities $\rightarrow$ algebraic statistics.

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 $X \mapsto h_X$.

- $\mathbf{H}_n^*$ is a finite-dimensional space which captures special position information for all discrete random vectors of a fixed length (but unbounded state spaces).

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 $X \mapsto h_X$.

- $\mathbf{H}_n^*$ is a finite-dimensional space which captures special position information for all discrete random vectors of a fixed length (but unbounded state spaces).
- Applications in cryptography, coding theory, engineering want to optimize linear functions over $\mathbf{H}_n^*$.

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 $\bigcup_{d_1=1}^{\infty} \cdots \bigcup_{d_n=1}^{\infty} \Delta(d_1, \dots, d_n)$ under the transcendental map $X \mapsto h_X$.

- $\mathbf{H}_n^*$ is a finite-dimensional space which captures special position information for all discrete random vectors of a fixed length (but unbounded state spaces).
- Applications in cryptography, coding theory, engineering want to optimize linear functions over $\mathbf{H}_n^*$.

Theorem ([Mat07])

$\overline{\mathbf{H}_n^*}$ *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 $X \mapsto h_X$.

- $\mathbf{H}_n^*$ is a finite-dimensional space which captures special position information for all discrete random vectors of a fixed length (but unbounded state spaces).
- Applications in cryptography, coding theory, engineering want to optimize linear functions over $\mathbf{H}_n^*$.

Theorem ([Mat07])

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

- Elements of the dual cone (linear information inequalities) can give bounds for optimization problems.

Ingleton inequality

Let A, B, C, D be subspaces in a finite-dimensional vector space.
Then the Ingleton inequality holds for $h = \dim$:

$$\begin{aligned}\mathtt{I}(A, B \mid C, D) :=\, & h(A, C) + h(B, C) + h(A, D) + h(B, D) + h(C, D) - \\ & h(A, B) - h(C) - h(D) - h(A, C, D) - h(B, C, D) \quad \geq \quad 0.\end{aligned}$$

Ingleton inequality

Let A, B, C, D be subspaces in a finite-dimensional vector space.
Then the Ingleton inequality holds for $h = \dim$:

$$\begin{aligned}\mathtt{I}(A, B \mid C, D) :=& h(A, C) + h(B, C) + h(A, D) + h(B, D) + h(C, D) - \\ & h(A, B) - h(C) - h(D) - h(A, C, D) - h(B, C, D) \;\; \geq \;\; 0.\end{aligned}$$

The Ingleton inequality fails in general for $h = h_X$ but it has been discovered that certain special position assumptions make it true even in the entropic setting, e.g.,

Ingleton inequality

Let A, B, C, D be subspaces in a finite-dimensional vector space.
Then the Ingleton inequality holds for $h = \dim$:

$$\begin{aligned}\mathrm{I}(A, B \mid C, D) := & h(A, C) + h(B, C) + h(A, D) + h(B, D) + h(C, D) - \\ & h(A, B) - h(C) - h(D) - h(A, C, D) - h(B, C, D) \geq 0.\end{aligned}$$

The Ingleton inequality fails in general for $h = h_X$ but it has been discovered that certain special position assumptions make it true even in the entropic setting, e.g.,

- If $C \perp\!\!\!\perp D$ then $\mathrm{I}(A, B \mid C, D) \geq 0$.

Ingleton inequality

Let A, B, C, D be subspaces in a finite-dimensional vector space.
Then the Ingleton inequality holds for $h = \dim$:

$$\mathtt{I}(A, B \mid C, D) := h(A, C) + h(B, C) + h(A, D) + h(B, D) + h(C, D) - h(A, B) - h(C) - h(D) - h(A, C, D) - h(B, C, D) \geq 0.$$

The Ingleton inequality fails in general for $h = h_X$ but it has been discovered that certain special position assumptions make it true even in the entropic setting, e.g.,

- If $C \perp\!\!\!\perp D$ then $\mathtt{I}(A, B \mid C, D) \geq 0$.
- If $A \perp\!\!\!\perp C \mid D$ and $A \perp\!\!\!\perp D \mid C$ then $\mathtt{I}(A, B \mid C, D) \geq 0$.

Ingleton inequality

Let A, B, C, D be subspaces in a finite-dimensional vector space.
Then the Ingleton inequality holds for $h = \dim$:

$$\mathtt{I}(A, B \mid C, D) := h(A, C) + h(B, C) + h(A, D) + h(B, D) + h(C, D) - h(A, B) - h(C) - h(D) - h(A, C, D) - h(B, C, D) \geq 0.$$

The Ingleton inequality fails in general for $h = h_X$ but it has been discovered that certain special position assumptions make it true even in the entropic setting, e.g.,

- If $C \perp\!\!\!\perp D$ then $\mathtt{I}(A, B \mid C, D) \geq 0$.
- If $A \perp\!\!\!\perp C \mid D$ and $A \perp\!\!\!\perp D \mid C$ then $\mathtt{I}(A, B \mid C, D) \geq 0$.
- ...

Ingleton inequality

Let A, B, C, D be subspaces in a finite-dimensional vector space.
Then the Ingleton inequality holds for $h = \dim$:

$$\mathtt{I}(A, B \mid C, D) := h(A, C) + h(B, C) + h(A, D) + h(B, D) + h(C, D) - h(A, B) - h(C) - h(D) - h(A, C, D) - h(B, C, D) \geq 0.$$

The Ingleton inequality fails in general for $h = h_X$ but it has been discovered that certain special position assumptions make it true even in the entropic setting, e.g.,

- If $C \perp\!\!\!\perp D$ then $\mathtt{I}(A, B \mid C, D) \geq 0$.
- If $A \perp\!\!\!\perp C \mid D$ and $A \perp\!\!\!\perp D \mid C$ then $\mathtt{I}(A, B \mid C, D) \geq 0$.
- ...

These are conditional linear information inequalities and they can tell apart honest boundary parts of $\mathbf{H}_n^*$ from fake boundary parts on $\overline{\mathbf{H}_n^*}$.

Conditional Ingleton inequalities

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

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

Check out ⋗https://mathrepo.mis.mpg.de/ConditionalIngleton/⋖ for non-linear algebra and numerical optimization techniques used in part of the proof.

Conditional Ingleton inequalities

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

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

Check out https://mathrepo.mis.mpg.de/ConditionalIngleton/ for non-linear algebra and numerical optimization techniques used in part of the proof.

Corollary

On four discrete random variables there are precisely 18 478 realizable conditional independence structures. (Combinatorial laws of information theory)

Conditional Ingleton inequalities

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

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

Check out ⩾`https://mathrepo.mis.mpg.de/ConditionalIngleton/`⩽ for non-linear algebra and numerical optimization techniques used in part of the proof.

Corollary

On four discrete random variables there are precisely 18 478 realizable conditional independence structures. (Combinatorial laws of information theory)

Problem

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

Challenges

Problem

Find/sample positive points from conditional independence varieties.

Challenges

Problem

Find/sample positive points from conditional independence varieties.

Problem

Let V be a $\mathbb{Z}$-defined variety. The distribution on $\mathbb{F}_q^n$ which is supported and uniform on $V(\mathbb{F}_q)$ has an entropy profile. How to compute it as $q = p^e$ with $p, e \to \infty$?

Challenges

Problem

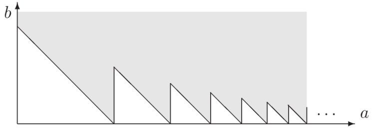
Find/sample positive points from conditional independence varieties.

Problem

Let V be a $\mathbb{Z}$-defined variety. The distribution on $\mathbb{F}_q^n$ which is supported and uniform on $V(\mathbb{F}_q)$ has an entropy profile. How to compute it as $q = p^e$ with $p, e \to \infty$?

Problem

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



Challenges

Problem

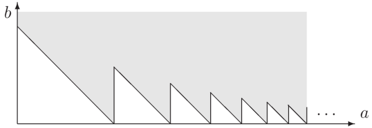
Find/sample positive points from conditional independence varieties.

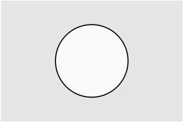
Problem

Let V be a $\mathbb{Z}$-defined variety. The distribution on $\mathbb{F}_q^n$ which is supported and uniform on $V(\mathbb{F}_q)$ has an entropy profile. How to compute it as $q = p^e$ with $p, e \to \infty$?

Problem

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





Challenges

Problem

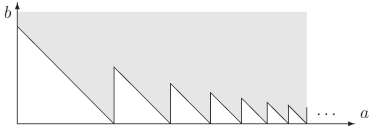
Find/sample positive points from conditional independence varieties.

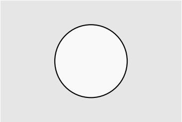
Problem

Let V be a $\mathbb{Z}$-defined variety. The distribution on $\mathbb{F}_q^n$ which is supported and uniform on $V(\mathbb{F}_q)$ has an entropy profile. How to compute it as $q = p^e$ with $p, e \to \infty$?

Problem

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





References

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

[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. ISSN: 0018-9448. DOI: 10.1109/TIT.2013.2274614.

[Mat06] Frantisek Matús. “Piecewise linear conditional information inequality”. In: *IEEE Trans. Inf. Theory* 52.1 (2006), pp. 236–238. ISSN: 0018-9448. DOI: 10.1109/TIT.2005.860438.

[Mat07] František Matúš. “Two constructions on limits of entropy functions.”. In: *IEEE Trans. Inf. Theory* 53.1 (2007), pp. 320–330. ISSN: 0018-9448. DOI: 10.1109/TIT.2006.887090.

[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. ISSN: 0018-9448. DOI: 10.1109/TIT.2021.3104250.