Entropy profiles and algebraic matroids

Tobias Boege

arXiv:2502.20355

Department of Mathematics and Statistics
UiT The Arctic University of Norway

Discrete Mathematics & Geometry seminar,
TU Berlin, 11 June 2025

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

Entropy profiles

Let ξ be a random variable taking finitely many values $\{1, \dots, d\}$ with probabilities p_i.

- Its Shannon entropy is

$$H(\xi) := -\sum_{i=1}^{d} p_i \log p_i, \text{ where } 0 \log 0 := 0.$$

Entropy profiles

Let ξ be a random variable taking finitely many values $\{1, \ldots, d\}$ with probabilities p_i.

- Its Shannon entropy is

$$H(\xi) := -\sum_{i=1}^{d} p_i \log p_i \,, \text{ where } 0 \log 0 := 0.$$

- A random vector $\xi = (\xi_i : i \in N)$ has 2^N marginals.

Entropy profiles

Let ξ be a random variable taking finitely many values $\{1, \ldots, d\}$ with probabilities p_i.

- Its Shannon entropy is

$$H(\xi) := -\sum_{i=1}^{d} p_i \log p_i, \text{ where } 0 \log 0 := 0.$$

- A random vector $\xi = (\xi_i : i \in N)$ has 2^N marginals.
- The collection of all the marginal entropies is the entropy profile $h_\xi : 2^N \to \mathbb{R}$.

Entropy profiles

Let ξ be a random variable taking finitely many values $\{1, \ldots, d\}$ with probabilities p_i.

- Its Shannon entropy is

$$H(\xi) := -\sum_{i=1}^{d} p_i \log p_i, \text{ where } 0 \log 0 := 0.$$

- A random vector $\xi = (\xi_i : i \in N)$ has 2^N marginals.
- The collection of all the marginal entropies is the entropy profile $h_\xi : 2^N \to \mathbb{R}$.
- Entropy profiles are "rank functions": monotone and submodular.

Entropy as information

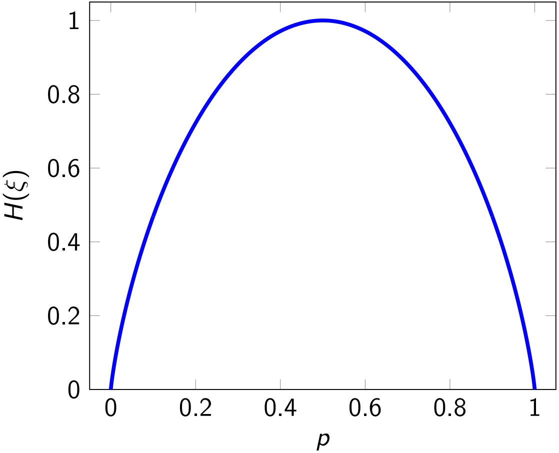


Figure: Entropy of a binary random variable ξ as a function of $p = p(\xi = \text{heads})$.

Special position for random variables

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

Special position for random variables

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

- Subvector ξ_I is functionally dependent on ξ_K if and only if $\boldsymbol{h_\xi(I \cup K) = h_\xi(K)}$.

Special position for random variables

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

- Subvector ξ_I is functionally dependent on ξ_K if and only if $\boldsymbol{h_\xi(I \cup K) = h_\xi(K)}$.
- Subvectors ξ_I and ξ_J are conditionally independent given ξ_K if and only if $\boldsymbol{h_\xi(I \cup K) + h_\xi(J \cup K) = h_\xi(I \cup J \cup K) + h_\xi(K)}$.

Special position for random variables

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

- Subvector ξ_I is functionally dependent on ξ_K if and only if $\boldsymbol{h_\xi(I \cup K) = h_\xi(K)}$.
- Subvectors ξ_I and ξ_J are conditionally independent given ξ_K if and only if $\boldsymbol{h_\xi(I \cup K) + h_\xi(J \cup K) = h_\xi(I \cup J \cup K) + h_\xi(K)}$.

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

Special position for random variables

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

- Subvector ξ_I is functionally dependent on ξ_K if and only if $\boldsymbol{h_\xi(I \cup K) = h_\xi(K)}$.
- Subvectors ξ_I and ξ_J are conditionally independent given ξ_K if and only if $\boldsymbol{h_\xi(I \cup K) + h_\xi(J \cup K) = h_\xi(I \cup J \cup K) + h_\xi(K)}$.

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).

Special position for random variables

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

- Subvector ξ_I is functionally dependent on ξ_K if and only if $\boldsymbol{h_\xi(I \cup K) = h_\xi(K)}$.
- Subvectors ξ_I and ξ_J are conditionally independent given ξ_K if and only if $\boldsymbol{h_\xi(I \cup K) + h_\xi(J \cup K) = h_\xi(I \cup J \cup K) + h_\xi(K)}$.

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).

Special position for random variables

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

- Subvector ξ_I is functionally dependent on ξ_K if and only if $\boldsymbol{h_\xi(I \cup K) = h_\xi(K)}$.
- Subvectors ξ_I and ξ_J are conditionally independent given ξ_K if and only if $\boldsymbol{h_\xi(I \cup K) + h_\xi(J \cup K) = h_\xi(I \cup J \cup K) + h_\xi(K)}$.

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

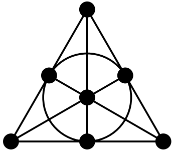
- Graphical models in statistics and causality are defined by CI assumptions (e.g., Bayesian networks and d-separation in graphs).
- Cryptographic protocols use FD and CI constraints to specify operation and information-theoretic security (e.g., secret sharing).
- Quantities in information theory are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

Example: Perfect secret sharing

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

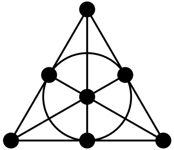
Example: Perfect secret sharing

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



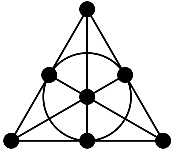
Example: Perfect secret sharing

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



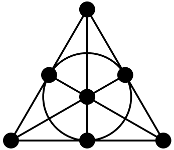
Example: Perfect secret sharing

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



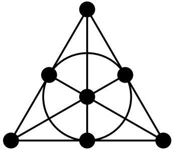
Example: Perfect secret sharing

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

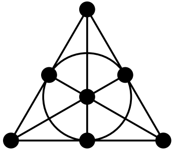
- Given: participants $N = \{1, \dots, n\}$ and a set of qualified subsets $\mathscr{Q} \subseteq 2^N$.
- Devise a scheme to distribute shares s_p of a randomly generated secret s to the participants such that
  - s_p is a function of s,
  - s is a function of $s_A = (s_p : p \in A)$ whenever $A \in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

The entropy region and information inequalities

Let $H_N^* \subseteq \mathbb{R}^{2^N}$ consist of all h_ξ where ξ is an N-variate discrete random vector. 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 $\xi \mapsto h_\xi$.

The entropy region and information inequalities

Let $H_N^* \subseteq \mathbb{R}^{2^N}$ consist of all h_ξ where ξ is an N-variate discrete random vector. 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 $\xi \mapsto h_\xi$.

Theorem ([ZY97], [Mat07b])

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

The entropy region and information inequalities

Let $H_N^* \subseteq \mathbb{R}^{2^N}$ consist of all h_ξ where ξ is an N-variate discrete random vector. 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 $\xi \mapsto h_\xi$.

Theorem ([ZY97], [Mat07b])

$\overline{H_N^*}$ *is a convex cone of dimension* $2^N - 1$. *Furthermore* $\mathrm{relint}(\overline{H_N^*}) \subseteq H_N^*$.

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

The entropy region and information inequalities

Let $H_N^* \subseteq \mathbb{R}^{2^N}$ consist of all h_ξ where ξ is an N-variate discrete random vector. 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 $\xi \mapsto h_\xi$.

Theorem ([ZY97], [Mat07b])

$\overline{H_N^*}$ *is a convex cone of dimension* $2^N - 1$*. Furthermore* $\mathrm{relint}(\overline{H_N^*}) \subseteq H_N^*$.

- Linear optimization works well! Elements of the dual cone (linear information inequalities) can give bounds for optimization problems.
- [DFZ11] contains over 200 inequalities and several parametric families.

The entropy region and information inequalities

Let $H_N^* \subseteq \mathbb{R}^{2^N}$ consist of all h_ξ where ξ is an N-variate discrete random vector. 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 $\xi \mapsto h_\xi$.

Theorem ([ZY97], [Mat07b])

$\overline{H_N^*}$ *is a convex cone of dimension* $2^N - 1$*. Furthermore* $\mathrm{relint}(\overline{H_N^*}) \subseteq H_N^*$.

- Linear optimization works well! Elements of the dual cone (linear information inequalities) can give bounds for optimization problems.
- [DFZ11] contains over 200 inequalities and several parametric families.

Theorem ([Mat07a])

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

A conditional information inequality

A conditional information inequality is an inequality valid for all entropy profiles satisfying some linear equations.

A conditional information inequality

A conditional information inequality is an inequality valid for all entropy profiles satisfying some linear equations.

$$\underline{H(A:B)} = \underline{H(A:B \mid C)} = 0 \implies H(C:D \mid A) + H(C:D \mid B) + H(A:B) \geq H(C:D) \quad (1)$$

A conditional information inequality

A conditional information inequality is an inequality valid for all entropy profiles satisfying some linear equations.

$$\underline{H(A:B)} = \underline{H(A:B \mid C)} = 0 \implies H(C:D \mid A) + H(C:D \mid B) + H(A:B) \geq H(C:D) \quad (1)$$

- This is useful in situations where $A \perp\!\!\!\perp B$ and $A \perp\!\!\!\perp B \mid C$.

A conditional information inequality

A conditional information inequality is an inequality valid for all entropy profiles satisfying some linear equations.

$$\underline{H(A:B)} = \underline{H(A:B \mid C)} = 0 \implies H(C:D \mid A) + H(C:D \mid B) + H(A:B) \geq H(C:D) \quad (1)$$

- This is useful in situations where $A \perp\!\!\!\perp B$ and $A \perp\!\!\!\perp B \mid C$.
- A natural question is whether this inequality can be lifted to an unconditional one by introducing Lagrange multipliers:

$$\lambda \underline{H(A:B)} + \mu \underline{H(A:B \mid C)} + H(C:D \mid A) + H(C:D \mid B) + H(A:B) \geq H(C:D). \quad (2)$$

A conditional information inequality

A conditional information inequality is an inequality valid for all entropy profiles satisfying some linear equations.

$$\underline{H(A:B)} = \underline{H(A:B \mid C)} = 0 \implies H(C:D \mid A) + H(C:D \mid B) + H(A:B) \geq H(C:D) \quad (1)$$

- This is useful in situations where $A \perp\!\!\!\perp B$ and $A \perp\!\!\!\perp B \mid C$.
- A natural question is whether this inequality can be lifted to an unconditional one by introducing Lagrange multipliers:

$$\lambda \underline{H(A:B)} + \mu \underline{H(A:B \mid C)} + H(C:D \mid A) + H(C:D \mid B) + H(A:B) \geq H(C:D). \quad (2)$$

- Kaced and Romashchenko [KR13] proved that (1) is **essentially conditional**, i.e., there are no $\lambda, \mu \in \mathbb{R}$ such that (2) holds.

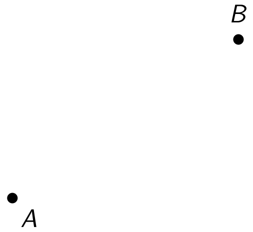
The Kaced–Romashchenko* configuration

Consider the affine plane over a finite field $\mathbb{F}_q$.

The Kaced–Romashchenko* configuration

Consider the affine plane over a finite field $\mathbb{F}_q$.

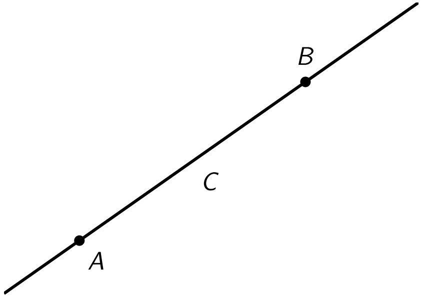
- Choose two points A and B with different x-coordinates uniformly at random.



The Kaced–Romashchenko* configuration

Consider the affine plane over a finite field $\mathbb{F}_q$.

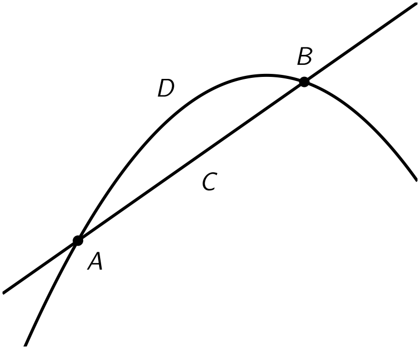
- Choose two points A and B with different x-coordinates uniformly at random.
- Draw the line C through A and B.



The Kaced–Romashchenko* configuration

Consider the affine plane over a finite field $\mathbb{F}_q$.

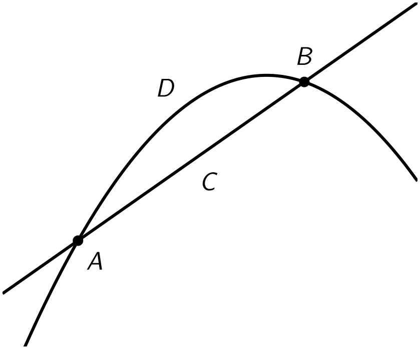
- Choose two points A and B with different x-coordinates uniformly at random.
- Draw the line C through A and B.
- Choose a non-degenerate parabola D through A and B uniformly at random.



The Kaced–Romashchenko* configuration

Consider the affine plane over a finite field $\mathbb{F}_q$.

- Choose two points A and B with different x-coordinates uniformly at random.
- Draw the line C through A and B.
- Choose a non-degenerate parabola D through A and B uniformly at random.



The coordinates of (A, B, C, D) define the support of a distribution on $\mathbb{F}_q^2 \times \mathbb{F}_q^2 \times \mathbb{F}_q^2 \times \mathbb{F}_q^3$ and the distribution is uniform on this set.

The Kaced–Romashchenko* configuration

- Elementary parameter counting yields

$$H(A:B) = H(A:B \mid C) = \log(q) - \log(q-1) \text{ and}$$
$$H(C:D \mid A) = H(C:D \mid B) = \log(q-1) - \log(q-2).$$

The Kaced–Romashchenko* configuration

- Elementary parameter counting yields

$$H(A:B) = H(A:B \mid C) = \log(q) - \log(q-1) \text{ and}$$
$$H(C:D \mid A) = H(C:D \mid B) = \log(q-1) - \log(q-2).$$

- However, $H(C:D) = \log(q) - \log(q-1) + \log(2)$. The $\log(2)$ term reflects that only half of all pairs (C, D) defined over $\mathbb{F}_q$ intersect in two $\mathbb{F}_q$-rational points!

The Kaced–Romashchenko* configuration

- Elementary parameter counting yields

$$H(A:B) = H(A:B \mid C) = \log(q) - \log(q-1) \text{ and}$$
$$H(C:D \mid A) = H(C:D \mid B) = \log(q-1) - \log(q-2).$$

- However, $H(C:D) = \log(q) - \log(q-1) + \log(2)$. The $\log(2)$ term reflects that only half of all pairs (C, D) defined over $\mathbb{F}_q$ intersect in two $\mathbb{F}_q$-rational points!

Hence, for this distribution

$$\lambda H(A:B) + \mu H(A:B \mid C) + H(C:D \mid A) + H(C:D \mid B) + H(A:B) - H(C:D)$$
$$= (\lambda + \mu)\log\left(\frac{q}{q-1}\right) + 2\log\left(\frac{q-1}{q-2}\right) - \log 2$$

The Kaced–Romashchenko* configuration

- Elementary parameter counting yields

$$H(A:B) = H(A:B \mid C) = \log(q) - \log(q-1) \text{ and}$$
$$H(C:D \mid A) = H(C:D \mid B) = \log(q-1) - \log(q-2).$$

- However, $H(C:D) = \log(q) - \log(q-1) + \log(2)$. The $\log(2)$ term reflects that only half of all pairs (C, D) defined over $\mathbb{F}_q$ intersect in two $\mathbb{F}_q$-rational points!

Hence, for this distribution

$$\lambda H(A:B) + \mu H(A:B \mid C) + H(C:D \mid A) + H(C:D \mid B) + H(A:B) - H(C:D)$$
$$= (\lambda + \mu)\log\left(\frac{q}{q-1}\right) + 2\log\left(\frac{q-1}{q-2}\right) - \log 2$$

which becomes negative for any λ, μ as $q \to \infty$.

An algebraic point of view

- Kaced and Romashchenko define an irreducible (quasiaffine) variety V and equip its $\mathbb{F}_q$-rational points with the uniform distribution $\rightarrow \xi(\mathbb{F}_q)$.

An algebraic point of view

- Kaced and Romashchenko define an irreducible (quasiaffine) variety V and equip its $\mathbb{F}_q$-rational points with the uniform distribution $\to \xi(\mathbb{F}_q)$.
- They use dimensions and facts from number theory to compute entropies.

An algebraic point of view

- Kaced and Romashchenko define an irreducible (quasiaffine) variety V and equip its $\mathbb{F}_q$-rational points with the uniform distribution $\to \xi(\mathbb{F}_q)$.
- They use dimensions and facts from number theory to compute entropies.
- Clearly $H(\xi(\mathbb{F}_q)) = \log |V(\mathbb{F}_q)| \to$ point counting!

An algebraic point of view

- Kaced and Romashchenko define an irreducible (quasiaffine) variety V and equip its $\mathbb{F}_q$-rational points with the uniform distribution $\to \xi(\mathbb{F}_q)$.
- They use dimensions and facts from number theory to compute entropies.
- Clearly $H(\xi(\mathbb{F}_q)) = \log|V(\mathbb{F}_q)| \to$ point counting!
- Entropies of marginals are more complicated. We have to deal with a coordinate projection of V in which each point is weighted by the size of its fiber:

$$\Pr[\xi_I(\mathbb{F}_q) = a] = \frac{|V(\mathbb{F}_q) \cap \pi_I^{-1}(a)|}{|V(\mathbb{F}_q)|}.$$

An algebraic point of view

- Kaced and Romashchenko define an irreducible (quasiaffine) variety V and equip its $\mathbb{F}_q$-rational points with the uniform distribution $\to \xi(\mathbb{F}_q)$.
- They use dimensions and facts from number theory to compute entropies.
- Clearly $H(\xi(\mathbb{F}_q)) = \log|V(\mathbb{F}_q)| \to$ point counting!
- Entropies of marginals are more complicated. We have to deal with a coordinate projection of V in which each point is weighted by the size of its fiber:

$$\Pr[\xi_I(\mathbb{F}_q) = a] = \frac{|V(\mathbb{F}_q) \cap \pi_I^{-1}(a)|}{|V(\mathbb{F}_q)|}.$$

Can this be done by computer algebra?

Model theory of finite fields

$\mathbb{F}_q$-definable sets are sets of the form $\varphi(\mathbb{F}_q^n; b) = \{ a \in \mathbb{F}_q^n : \mathbb{F}_q \models \varphi(a, b) \}$ where $\varphi(x_1, \ldots, x_n, y_1, \ldots, y_m)$ is a first-order formula in the language of rings and $b \in \mathbb{F}_q^m$.

Model theory of finite fields

$\mathbb{F}_q$-definable sets are sets of the form $\varphi(\mathbb{F}_q^n; b) = \{ a \in \mathbb{F}_q^n : \mathbb{F}_q \models \varphi(a, b) \}$ where $\varphi(x_1, \ldots, x_n, y_1, \ldots, y_m)$ is a first-order formula in the language of rings and $b \in \mathbb{F}_q^m$.

- More briefly: the smallest set of sets containing all varieties defined over $\mathbb{F}_q$ and closed under complement and projection.

Model theory of finite fields

$\mathbb{F}_q$-definable sets are sets of the form $\varphi(\mathbb{F}_q^n; b) = \{ a \in \mathbb{F}_q^n : \mathbb{F}_q \models \varphi(a, b) \}$ where $\varphi(x_1, \dots, x_n, y_1, \dots, y_m)$ is a first-order formula in the language of rings and $b \in \mathbb{F}_q^m$.

- More briefly: the smallest set of sets containing all varieties defined over $\mathbb{F}_q$ and closed under complement and projection.

Theorem ([CDM92])

Consider a formula $\varphi(x_1, \dots, x_n, y_1, \dots, y_m)$. There exist finitely many formulas $\psi_k(y_1, \dots, y_m)$, indexed by $k \in K$, with accompanying $\mu_k \in \mathbb{Q}$ and $d_k \in \mathbb{N}$ such that

Model theory of finite fields

$\mathbb{F}_q$-definable sets are sets of the form $\varphi(\mathbb{F}_q^n; b) = \{ a \in \mathbb{F}_q^n : \mathbb{F}_q \models \varphi(a, b) \}$ where $\varphi(x_1, \ldots, x_n, y_1, \ldots, y_m)$ is a first-order formula in the language of rings and $b \in \mathbb{F}_q^m$.

- More briefly: the smallest set of sets containing all varieties defined over $\mathbb{F}_q$ and closed under complement and projection.

Theorem ([CDM92])

Consider a formula $\varphi(x_1, \ldots, x_n, y_1, \ldots, y_m)$. There exist finitely many formulas $\psi_k(y_1, \ldots, y_m)$, indexed by $k \in K$, with accompanying $\mu_k \in \mathbb{Q}$ and $d_k \in \mathbb{N}$ such that for every sufficiently large finite field $\mathbb{F}_q$ and every $b \in \mathbb{F}_q^m$:

Model theory of finite fields

$\mathbb{F}_q$-definable sets are sets of the form $\varphi(\mathbb{F}_q^n; b) = \{ a \in \mathbb{F}_q^n : \mathbb{F}_q \models \varphi(a, b) \}$ where $\varphi(x_1, \ldots, x_n, y_1, \ldots, y_m)$ is a first-order formula in the language of rings and $b \in \mathbb{F}_q^m$.

- More briefly: the smallest set of sets containing all varieties defined over $\mathbb{F}_q$ and closed under complement and projection.

Theorem ([CDM92])

Consider a formula $\varphi(x_1, \ldots, x_n, y_1, \ldots, y_m)$. There exist finitely many formulas $\psi_k(y_1, \ldots, y_m)$, indexed by $k \in K$, with accompanying $\mu_k \in \mathbb{Q}$ and $d_k \in \mathbb{N}$ such that for every sufficiently large finite field $\mathbb{F}_q$ and every $b \in \mathbb{F}_q^m$:

- *There exists a unique $k \in K$ such that $\mathbb{F}_q \models \psi_k(b)$.*

Model theory of finite fields

$\mathbb{F}_q$-definable sets are sets of the form $\varphi(\mathbb{F}_q^n; b) = \{ a \in \mathbb{F}_q^n : \mathbb{F}_q \models \varphi(a, b) \}$ where $\varphi(x_1, \ldots, x_n, y_1, \ldots, y_m)$ is a first-order formula in the language of rings and $b \in \mathbb{F}_q^m$.

- More briefly: the smallest set of sets containing all varieties defined over $\mathbb{F}_q$ and closed under complement and projection.

Theorem ([CDM92])

Consider a formula $\varphi(x_1, \ldots, x_n, y_1, \ldots, y_m)$. There exist finitely many formulas $\psi_k(y_1, \ldots, y_m)$, indexed by $k \in K$, with accompanying $\mu_k \in \mathbb{Q}$ and $d_k \in \mathbb{N}$ such that for every sufficiently large finite field $\mathbb{F}_q$ and every $b \in \mathbb{F}_q^m$:

- *There exists a unique $k \in K$ such that $\mathbb{F}_q \models \psi_k(b)$.*
- *$\mathbb{F}_q \models \psi_k(b)$ if and only if $|\varphi(\mathbb{F}_q^n; b)| = \mu_k q^{d_k} + \mathcal{O}(\mu_k q^{d_k - 1/2})$.*

Measure and decomposition

Definition

Let X be an $\mathbb{F}$-definable set. A *fiber decomposition* with respect to $\pi_I(X)$ is a finite family of $\mathbb{F}$-definable sets Y_k, called *cells*,

Measure and decomposition

Definition

Let X be an $\mathbb{F}$-definable set. A *fiber decomposition* with respect to $\pi_I(X)$ is a finite family of $\mathbb{F}$-definable sets Y_k, called *cells*, together with non-negative $\mu_k \in \mathbb{Q}$ and $d_k \in \mathbb{N}$, for $k \in K$, such that for all sufficiently large $\mathbb{G}/\mathbb{F}$:

Measure and decomposition

Definition

Let X be an $\mathbb{F}$-definable set. A *fiber decomposition* with respect to $\pi_I(X)$ is a finite family of $\mathbb{F}$-definable sets Y_k, called *cells*, together with non-negative $\mu_k \in \mathbb{Q}$ and $d_k \in \mathbb{N}$, for $k \in K$, such that for all sufficiently large $\mathbb{G}/\mathbb{F}$:

- $\mathbb{G}^I = \bigsqcup_{k \in K} Y_k(\mathbb{G})$, and

Measure and decomposition

Definition

Let X be an $\mathbb{F}$-definable set. A *fiber decomposition* with respect to $\pi_I(X)$ is a finite family of $\mathbb{F}$-definable sets Y_k, called *cells*, together with non-negative $\mu_k \in \mathbb{Q}$ and $d_k \in \mathbb{N}$, for $k \in K$, such that for all sufficiently large $\mathbb{G}/\mathbb{F}$:

- $\mathbb{G}^I = \bigsqcup_{k \in K} Y_k(\mathbb{G})$, and
- $|X(\mathbb{G}) \cap \pi_I^{-1}(a)| = \mu_k |\mathbb{G}|^{d_k} + \mathcal{O}(\mu_k |\mathbb{G}|^{d_k - 1/2})$ for each $a \in Y_k(\mathbb{G})$.

Measure and decomposition

Definition

Let X be an $\mathbb{F}$-definable set. A *fiber decomposition* with respect to $\pi_I(X)$ is a finite family of $\mathbb{F}$-definable sets Y_k, called *cells*, together with non-negative $\mu_k \in \mathbb{Q}$ and $d_k \in \mathbb{N}$, for $k \in K$, such that for all sufficiently large $\mathbb{G}/\mathbb{F}$:

- $\mathbb{G}^I = \bigsqcup_{k \in K} Y_k(\mathbb{G})$, and
- $|X(\mathbb{G}) \cap \pi_I^{-1}(a)| = \mu_k |\mathbb{G}|^{d_k} + \mathcal{O}(\mu_k |\mathbb{G}|^{d_k - 1/2})$ for each $a \in Y_k(\mathbb{G})$.

Theorem ([FHJ94])

Fiber decompositions are computable. Moreover, one can compute a bound $m \in \mathbb{N}$, numbers $d_k \in \mathbb{N}$ and non-negative $\mu_k \in \mathbb{Q}$ such that for every finite extension $\mathbb{G}/\mathbb{F}$:

$$|X(\mathbb{G})| = \mu_k |\mathbb{G}|^{d_k} + \mathcal{O}(\mu_k |\mathbb{G}|^{d_k - 1/2}), \text{ where } k \equiv [\mathbb{G} : \mathbb{F}] \pmod{m}.$$

Computability of the entropy profiles

Theorem

Let X be an $\mathbb{F}$-definable set in n free variables and $\xi(\mathbb{G})$ the uniform distribution on $X(\mathbb{G})$. For a projection $\pi_I(X)$ let $(Y_k : k \in K)$, be a fiber decomposition and set $X_k = X \cap \pi_I^{-1}(Y_k)$. For large enough $\mathbb{G}/\mathbb{F}$, the entropy profile satisfies

$$h_{\xi(\mathbb{G})}(I) = \sum_{\dim_{\mathbb{G}}(X_k)=\dim_{\mathbb{G}}(X)} \frac{\mu_{\mathbb{G}}(X_k)}{\mu_{\mathbb{G}}(X)} \log\left(\frac{\mu_{\mathbb{G}}(X)\mu_{\mathbb{G}}(Y_k)}{\mu_{\mathbb{G}}(X_k)}|\mathbb{G}|^{\dim_{\mathbb{G}}(Y_k)}\right) + \mathcal{O}\left(\frac{\log|\mathbb{G}|}{\sqrt{|\mathbb{G}|}}\right).$$

Computability of the entropy profiles

Theorem

Let X be an $\mathbb{F}$-definable set in n free variables and $\xi(\mathbb{G})$ the uniform distribution on $X(\mathbb{G})$. For a projection $\pi_I(X)$ let $(Y_k : k \in K)$, be a fiber decomposition and set $X_k = X \cap \pi_I^{-1}(Y_k)$. For large enough $\mathbb{G}/\mathbb{F}$, the entropy profile satisfies

$$h_{\xi(\mathbb{G})}(I) = \sum_{\dim_{\mathbb{G}}(X_k)=\dim_{\mathbb{G}}(X)} \frac{\mu_{\mathbb{G}}(X_k)}{\mu_{\mathbb{G}}(X)} \log\left(\frac{\mu_{\mathbb{G}}(X)\mu_{\mathbb{G}}(Y_k)}{\mu_{\mathbb{G}}(X_k)}|\mathbb{G}|^{\dim_{\mathbb{G}}(Y_k)}\right) + \mathcal{O}\left(\frac{\log|\mathbb{G}|}{\sqrt{|\mathbb{G}|}}\right).$$

The leading term does not vanish, can be effectively computed from a defining formula for X and is periodic in the extension degree $[\mathbb{G} : \mathbb{F}]$.

Computability of the entropy profiles

Theorem

Let X be an $\mathbb{F}$-definable set in n free variables and $\xi(\mathbb{G})$ the uniform distribution on $X(\mathbb{G})$. For a projection $\pi_I(X)$ let $(Y_k : k \in K)$, be a fiber decomposition and set $X_k = X \cap \pi_I^{-1}(Y_k)$. For large enough $\mathbb{G}/\mathbb{F}$, the entropy profile satisfies

$$h_{\xi(\mathbb{G})}(I) = \sum_{\dim_{\mathbb{G}}(X_k)=\dim_{\mathbb{G}}(X)} \frac{\mu_{\mathbb{G}}(X_k)}{\mu_{\mathbb{G}}(X)} \log\left(\frac{\mu_{\mathbb{G}}(X)\mu_{\mathbb{G}}(Y_k)}{\mu_{\mathbb{G}}(X_k)}|\mathbb{G}|^{\dim_{\mathbb{G}}(Y_k)}\right) + \mathcal{O}\left(\frac{\log|\mathbb{G}|}{\sqrt{|\mathbb{G}|}}\right).$$

The leading term does not vanish, can be effectively computed from a defining formula for X and is periodic in the extension degree $[\mathbb{G} : \mathbb{F}]$.

- The sequence $\left(\frac{1}{\log|\mathbb{G}|} h_{\xi(\mathbb{G})} : \mathbb{G} \supseteq \mathbb{F}\right)$ has finitely many convergent subsequences and their (rational!) limits can all be computed.

Algebraic matroids

Theorem

Moreover, if X is an $\mathbb{F}$-irreducible algebraic variety, then there exists a tower of finite fields $\mathbb{F} = \mathbb{G}_0 \subseteq \mathbb{G}_1 \subseteq \dots$ with

$$\lim_{n\to\infty} \frac{1}{\log|\mathbb{G}_n|} h_{\xi(\mathbb{G}_n)}(I) = \dim \pi_I(X(\overline{\mathbb{F}})), \quad \textit{for every } I \subseteq N.$$

Algebraic matroids

Theorem

Moreover, if X is an $\mathbb{F}$-irreducible algebraic variety, then there exists a tower of finite fields $\mathbb{F} = \mathbb{G}_0 \subseteq \mathbb{G}_1 \subseteq \dots$ with

$$\lim_{n\to\infty} \frac{1}{\log|\mathbb{G}_n|} h_{\xi(\mathbb{G}_n)}(I) = \dim \pi_I(X(\overline{\mathbb{F}})), \quad \text{for every } I \subseteq N.$$

Corollary ([Mat24])

Algebraic matroids are almost-entropic.

Algebraic matroids

Theorem

Moreover, if X is an $\mathbb{F}$-irreducible algebraic variety, then there exists a tower of finite fields $\mathbb{F} = \mathbb{G}_0 \subseteq \mathbb{G}_1 \subseteq \dots$ with

$$\lim_{n\to\infty} \frac{1}{\log|\mathbb{G}_n|} h_{\xi(\mathbb{G}_n)}(I) = \dim \pi_I(X(\overline{\mathbb{F}})), \quad \text{for every } I \subseteq N.$$

Corollary ([Mat24])

Algebraic matroids are almost-entropic.

- Algebraic independence in the limit is explained through diminishing stochastic dependence among the coordinate functions.

Algebraic matroids

Theorem

Moreover, if X is an $\mathbb{F}$-irreducible algebraic variety, then there exists a tower of finite fields $\mathbb{F} = \mathbb{G}_0 \subseteq \mathbb{G}_1 \subseteq \dots$ with

$$\lim_{n\to\infty} \frac{1}{\log|\mathbb{G}_n|} h_{\xi(\mathbb{G}_n)}(I) = \dim \pi_I(X(\overline{\mathbb{F}})), \quad \textit{for every } I \subseteq N.$$

Corollary ([Mat24])

Algebraic matroids are almost-entropic.

- Algebraic independence in the limit is explained through diminishing stochastic dependence among the coordinate functions.
- Entropy profile can be seen as a "valuated" refinement of the algebraic matroid.

Example using Galois stratification

- To eliminate x from the variety defined by $x^3 + ax^2 + bx + c = 0$, stratify the triples (a, b, c) according to the number of rational roots of $f(a, b, c) \in \mathbb{F}[x]$.

Example using Galois stratification

- To eliminate x from the variety defined by $x^3 + ax^2 + bx + c = 0$, stratify the triples (a, b, c) according to the number of rational roots of $f(a, b, c) \in \mathbb{F}[x]$.
- Let Ω be the splitting field of f over $\mathbb{F}(a, b, c)$ with Galois group $G = S_3$.

Example using Galois stratification

- To eliminate x from the variety defined by $x^3 + ax^2 + bx + c = 0$, stratify the triples (a, b, c) according to the number of rational roots of $f(a, b, c) \in \mathbb{F}[x]$.
- Let Ω be the splitting field of f over $\mathbb{F}(a, b, c)$ with Galois group $G = S_3$.
- If $f(a, b, c)$ is separable, it defines a Galois extension of $\mathbb{F}$ with cyclic Galois group $G(a, b, c) \subseteq G$.

Example using Galois stratification

- To eliminate x from the variety defined by $x^3 + ax^2 + bx + c = 0$, stratify the triples (a, b, c) according to the number of rational roots of $f(a, b, c) \in \mathbb{F}[x]$.
- Let Ω be the splitting field of f over $\mathbb{F}(a, b, c)$ with Galois group $G = S_3$.
- If $f(a, b, c)$ is separable, it defines a Galois extension of $\mathbb{F}$ with cyclic Galois group $G(a, b, c) \subseteq G$.
- The number of rational roots of $f(a, b, c)$ in $\mathbb{F}$ is determined by the splitting type of $f(a, b, c)$ in $\mathbb{F}[x]$ which corresponds to the conjugacy class of $G(a, b, c)$ in G.

Example using Galois stratification

- To eliminate x from the variety defined by $x^3 + ax^2 + bx + c = 0$, stratify the triples (a, b, c) according to the number of rational roots of $f(a, b, c) \in \mathbb{F}[x]$.
- Let Ω be the splitting field of f over $\mathbb{F}(a, b, c)$ with Galois group $G = S_3$.
- If $f(a, b, c)$ is separable, it defines a Galois extension of $\mathbb{F}$ with cyclic Galois group $G(a, b, c) \subseteq G$.
- The number of rational roots of $f(a, b, c)$ in $\mathbb{F}$ is determined by the splitting type of $f(a, b, c)$ in $\mathbb{F}[x]$ which corresponds to the conjugacy class of $G(a, b, c)$ in G.
- The Chebotarev density theorem computes the density of triples with given conjugacy class $\mathcal{C}$:

$$\frac{|\mathcal{C}|}{[\Omega : \mathbb{F}(a, b, c)]} = \frac{|\mathcal{C}|}{6},$$

Example using Galois stratification

- To eliminate x from the variety defined by $x^3 + ax^2 + bx + c = 0$, stratify the triples (a, b, c) according to the number of rational roots of $f(a, b, c) \in \mathbb{F}[x]$.
- Let Ω be the splitting field of f over $\mathbb{F}(a, b, c)$ with Galois group $G = S_3$.
- If $f(a, b, c)$ is separable, it defines a Galois extension of $\mathbb{F}$ with cyclic Galois group $G(a, b, c) \subseteq G$.
- The number of rational roots of $f(a, b, c)$ in $\mathbb{F}$ is determined by the splitting type of $f(a, b, c)$ in $\mathbb{F}[x]$ which corresponds to the conjugacy class of $G(a, b, c)$ in G.

Splitting type	$[1, 1, 1]$	$[1, 2]$	$[3]$	$[1, 1^2]$	$[1^3]$
Conjugacy class	id	$(1\ 2)$	$(1\ 2\ 3)$	–	–
Density	$1/6$	$3/6$	$2/6$	0	0
Rational roots	3	1	0	2	1

More details on Galois stratification

In greater generality, Galois stratification requires:

- Basic normal (irreducible) decomposition over $\mathbb{F}$.

More details on Galois stratification

In greater generality, Galois stratification requires:

- Basic normal (irreducible) decomposition over $\mathbb{F}$.
- Computing the splitting field Ω and Galois group G of a polynomial over $\mathbb{F}(V)$ where V is an irreducible $\mathbb{F}$-variety.

More details on Galois stratification

In greater generality, Galois stratification requires:

- Basic normal (irreducible) decomposition over $\mathbb{F}$.
- Computing the splitting field Ω and Galois group G of a polynomial over $\mathbb{F}(V)$ where V is an irreducible $\mathbb{F}$-variety.
- Computing the relative algebraic closure of $\mathbb{F}$ in Ω.

More details on Galois stratification

In greater generality, Galois stratification requires:

- Basic normal (irreducible) decomposition over $\mathbb{F}$.
- Computing the splitting field Ω and Galois group G of a polynomial over $\mathbb{F}(V)$ where V is an irreducible $\mathbb{F}$-variety.
- Computing the relative algebraic closure of $\mathbb{F}$ in Ω.
- Perhaps some relative integral closures of coordinate rings.

More details on Galois stratification

In greater generality, Galois stratification requires:

- Basic normal (irreducible) decomposition over $\mathbb{F}$.
- Computing the splitting field Ω and Galois group G of a polynomial over $\mathbb{F}(V)$ where V is an irreducible $\mathbb{F}$-variety.
- Computing the relative algebraic closure of $\mathbb{F}$ in Ω.
- Perhaps some relative integral closures of coordinate rings.
- Working with conjugacy classes of cyclic subgroups in G.

More details on Galois stratification

In greater generality, Galois stratification requires:

- Basic normal (irreducible) decomposition over $\mathbb{F}$.
- Computing the splitting field Ω and Galois group G of a polynomial over $\mathbb{F}(V)$ where V is an irreducible $\mathbb{F}$-variety.
- Computing the relative algebraic closure of $\mathbb{F}$ in Ω.
- Perhaps some relative integral closures of coordinate rings.
- Working with conjugacy classes of cyclic subgroups in G.

Algorithms are given in [FJ23] but with little regard for the state of the art in computer algebra.

More details on Galois stratification

In greater generality, Galois stratification requires:

- Basic normal (irreducible) decomposition over $\mathbb{F}$.
- Computing the splitting field Ω and Galois group G of a polynomial over $\mathbb{F}(V)$ where V is an irreducible $\mathbb{F}$-variety.
- Computing the relative algebraic closure of $\mathbb{F}$ in Ω.
- Perhaps some relative integral closures of coordinate rings.
- Working with conjugacy classes of cyclic subgroups in G.

Algorithms are given in [FJ23] but with little regard for the state of the art in computer algebra. Is it possible to produce an implementation in `Oscar`?

References I

[Boe25] Tobias Boege. *The entropy profiles of a definable set over finite fields*. 2025. arXiv: 2502.20355 [cs.IT].

[CDM92] Zoé Chatzidakis, Lou van den Dries, and Angus Macintyre. "Definable sets over finite fields". In: *J. Reine Angew. Math.* 427 (1992), pp. 107–135.

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

[FHJ94] Michael D. Fried, Dan Haran, and Moshe Jarden. "Effective counting of the points of definable sets over finite fields". In: *Isr. J. Math.* 85.1-3 (1994), pp. 103–133. DOI: 10.1007/BF02758639.

[FJ23] Michael D. Fried and Moshe Jarden. *Field arithmetic*. 4th corrected edition. Vol. 11. Ergeb. Math. Grenzgeb., 3. Folge. Springer, 2023. ISBN: 978-3-031-28019-1; 978-3-031-28022-1; 978-3-031-28020-7. DOI: 10.1007/978-3-031-28020-7.

[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.

References II

[Mat07a] František Matúš. "Infinitely many information inequalities". In: *Proceedings of the 2007 IEEE International Symposium on Information Theory*. Institute of Electrical and Electronics Engineers (IEEE), 2007, pp. 41–44. DOI: 10.1109/ISIT.2007.4557201.

[Mat07b] František Matúš. "Two constructions on limits of entropy functions.". In: *IEEE Trans. Inf. Theory* 53.1 (2007), pp. 320–330. DOI: 10.1109/TIT.2006.887090.

[Mat24] František Matúš. "Algebraic matroids are almost entropic". In: *Proc. Am. Math. Soc.* 152.1 (2024), pp. 1–6. DOI: 10.1090/proc/13846.

[ZY97] Zhen Zhang and Raymond W. Yeung. "A non-Shannon-type conditional inequality of information quantities". In: *IEEE Trans. Inf. Theory* 43.6 (1997), pp. 1982–1986. DOI: 10.1109/18.641561.