Non-Shannon type conditional information inequalities: proofs and application

Andrei Romashchenko (CNRS, LIRMM)

August 19, Prague

Subject of this talk: conditional (constraint) information inequalities

Subject of this talk: conditional (constraint) information inequalities

linear inequalities for the Shannon entropy

that are valid under **linear** constraints

Subject of this talk: conditional (constraint) information inequalities

linear inequalities for the Shannon entropy
that are valid under **linear** constraints

i.e.,
[some **linear** constraints] $\Longrightarrow$ [some **linear** inequality]

Subject of this talk: conditional (constraint) information inequalities

linear inequalities for the Shannon entropy

that are valid under **linear** constraints

i.e.,
[some **linear** constraints] $\Longrightarrow$ [some **linear** inequality]

e.g., $\mathrm{I}(x:y) = \mathrm{I}(x:y|a) = 0 \Longrightarrow \mathrm{I}(a:b) \le \mathrm{I}(a:b\,|\,x) + \mathrm{I}(a:b\,|\,y)$
[Zhang–Yeung'97]

Not in this talk

Not in this talk

- **piecewise-linear** conditional inequalities
  [Matúš 2006]

Not in this talk

- **piecewise-linear** conditional inequalities
  [Matúš 2006]
- **non-linear** information inequalities
  [Chan–Grant 2008, based on Matúš 2007]

Not in this talk

- **piecewise-linear** conditional inequalities
  [Matúš 2006]
- **non-linear** information inequalities
  [Chan–Grant 2008, based on Matúš 2007]
- conditional information **equalities**
  [*conditional independence* properties, Studený, Matúš]

Outline

Conditional information inequalities

(a) **Trivial, Shannon-type**:

if $\mathrm{I}(x:y)=0$ then $\mathrm{H}(a) \leq \mathrm{H}(a\,|\,x)+\mathrm{H}(a\,|\,y)$

Conditional information inequalities

(a) **Trivial, Shannon-type**:

if $\mathrm{I}(x:y)=0$ then $\mathrm{H}(a) \le \mathrm{H}(a\,|\,x) + \mathrm{H}(a\,|\,y)$

this is true since

$\mathrm{H}(a) \le \mathrm{H}(a\,|\,x) + \mathrm{H}(a\,|\,y) + \mathrm{I}(x:y)$ [Shannon-type unconditional inequalitiy]

Conditional information inequalities

(b) **Trivial, non Shannon-type**:

if $\mathrm{I}(a : b \,|\, z) = \mathrm{I}(a : z \,|\, b) = \mathrm{I}(b : z \,|\, a) = 0$ then

$$\mathrm{I}(a : b) \leq \mathrm{I}(a : b \,|\, x) + \mathrm{I}(a : b \,|\, y) + \mathrm{I}(x : y)$$

Conditional information inequalities

(b) **Trivial, non Shannon-type**:

if $\mathrm{I}(a:b\,|\,z) = \mathrm{I}(a:z\,|\,b) = \mathrm{I}(b:z\,|\,a) = 0$ then

$$\mathrm{I}(a:b) \le \mathrm{I}(a:b\,|\,x) + \mathrm{I}(a:b\,|\,y) + \mathrm{I}(x:y)$$

this is true since

$$\begin{aligned}\mathrm{I}(a:b) \quad \le \quad & \mathrm{I}(a:b\,|\,x) + \mathrm{I}(a:b\,|\,y) + \mathrm{I}(x:y) \\ & + \mathrm{I}(a:b\,|\,z) + \mathrm{I}(a:z\,|\,b) + \mathrm{I}(b:z\,|\,a)\end{aligned}$$

[non Shannon-type unconditional inequalitiy]

Conditional information inequalities

(c) **Non-trivial**, e.g.:

$$\underbrace{\mathrm{I}(x:y)=\mathrm{I}(x:y|a)=0}_{\text{[Z.Zhang–R.W.Yeung'97]}} \quad \underbrace{\mathrm{I}(x:a|b)=\mathrm{I}(x:b|a)=0}_{\text{[F.Matúš'99]}} \quad \underbrace{\mathrm{H}(a|x,y)=\mathrm{I}(x:y|a)=0}_{\text{[T.Kaced and A.R.'11]}}$$

$$\searrow \qquad \downarrow \qquad \swarrow$$

Ingleton's inequality

$$\mathrm{I}(a:b) \le \mathrm{I}(a:b\,|\,x) + \mathrm{I}(a:b\,|\,y) + \mathrm{I}(x:y)$$

Conditional information inequalities

(c) **Non-trivial**, e.g.:

$$\underbrace{\mathrm{I}(x:y)=\mathrm{I}(x:y|a)=0}_{\text{[Z.Zhang–R.W.Yeung'97]}} \quad \underbrace{\mathrm{I}(x:a|b)=\mathrm{I}(x:b|a)=0}_{\text{[F.Matúš'99]}} \quad \underbrace{\mathrm{H}(a|x,y)=\mathrm{I}(x:y|a)=0}_{\text{[T.Kaced and A.R.'11]}}$$

$$\searrow \qquad \downarrow \qquad \swarrow$$

Ingleton's inequality

$$\mathrm{I}(a:b) \le \mathrm{I}(a:b\,|\,x) + \mathrm{I}(a:b\,|\,y) + \mathrm{I}(x:y)$$

Claim: These three implications are *essentially* conditional inequalities.

Theorem

The inequality

$$\mathrm{H}(a|x,y) = \mathrm{I}(x:y|a) = 0 \Rightarrow \mathrm{I}(a:b) \leq \mathrm{I}(a:b|x) + \mathrm{I}(a:b|y) + \mathrm{I}(x:y)$$

is essentially conditional.

Theorem

The inequality

$$\mathrm{H}(a|x,y) = \mathrm{I}(x:y|a) = 0 \Rightarrow \mathrm{I}(a:b) \le \mathrm{I}(a:b|x) + \mathrm{I}(a:b|y) + \mathrm{I}(x:y)$$

is essentially conditional.

We cannot reduce it to an unconditional inequality!

Theorem

The inequality

$$\mathrm{H}(a|x,y) = \mathrm{I}(x:y|a) = 0 \Rightarrow \mathrm{I}(a:b) \le \mathrm{I}(a:b|x) + \mathrm{I}(a:b|y) + \mathrm{I}(x:y)$$

is essentially conditional.

We cannot reduce it to an unconditional inequality!

That is, for all λ_1, λ_2 the inequality

$$\mathrm{I}(a:b) \le \mathrm{I}(a:b|x) + \mathrm{I}(a:b|y) + \mathrm{I}(x:y) + \lambda_1 \mathrm{H}(a|x,y) + \lambda_2 \mathrm{I}(x:y \mid a)$$

does not hold.

Theorem

The inequality

$$\mathrm{H}(a|x,y) = \mathrm{I}(x:y|a) = 0 \Rightarrow \mathrm{I}(a:b) \leq \mathrm{I}(a:b|x) + \mathrm{I}(a:b|y) + \mathrm{I}(x:y)$$

is essentially conditional.

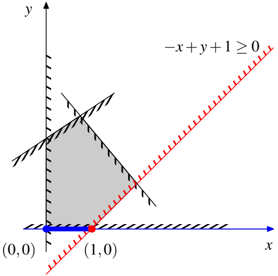
We cannot reduce it to an unconditional inequality!

More precisely, for all λ_1, λ_2 there exist (a, b, x, y) such that

$$\mathrm{I}(a:b) \not\leq \mathrm{I}(a:b|x) + \mathrm{I}(a:b|y) + \mathrm{I}(x:y) + \lambda_1 \mathrm{H}(a|x,y) + \lambda_2 \mathrm{I}(x:y \mid a)$$

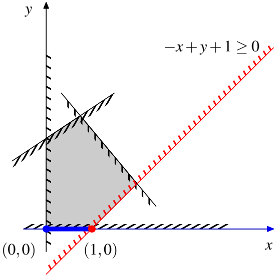
Outline

A geometric view on conditional inequalities:



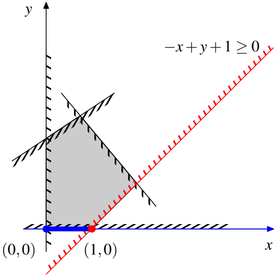
if $y = 0$ then $x \leq 1$

A geometric view on conditional inequalities:



if $y = 0$ then $x \leq 1$ $\Longleftarrow$ $x \leq 1 + y$

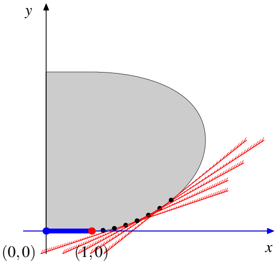
A geometric view on conditional inequalities:



if $y = 0$ then $x \leq 1 \;\Longleftarrow\; x \leq 1 + y$

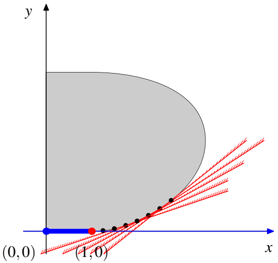
NOT essentially conditional

A geometric view on conditional inequalities:



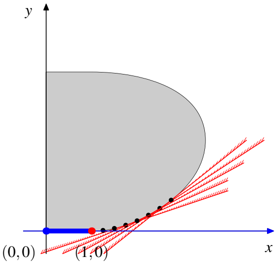
if $y = 0$ then $x \leq 1$

A geometric view on conditional inequalities:



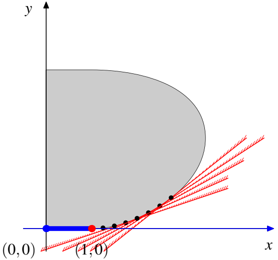
if $y = 0$ then $x \leq 1$ $\Longleftarrow$ follows from an *infinite* family of linear inequalities

A geometric view on conditional inequalities:



if $y = 0$ then $x \leq 1 \longleftarrow$ this inequality is essentially conditional

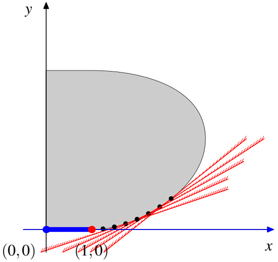
A geometric view on conditional inequalities:



if $y = 0$ then $x \leq 1 \longleftarrow$ this inequality is essentially conditional

NO unconditional inequality $x \leq 1 + \lambda y$

A geometric view on conditional inequalities:

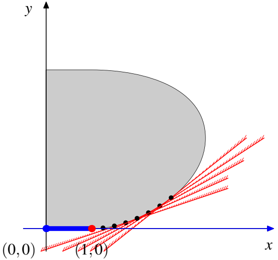


$\color{blue}{\text{if } y = 0 \text{ then } x \leq 1}$ $\longleftarrow$ this inequality is essentially conditional

NO unconditional inequality $x \leq 1 + \lambda y$

$\exists$ one essentially conditional inequality $\Longrightarrow$ the grey area is not polyhedral

A geometric view on conditional inequalities:



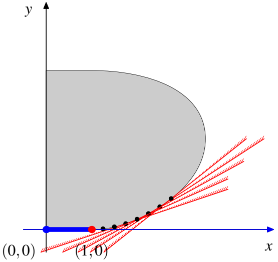
if $y = 0$ then $x \leq 1$ ⟵ this inequality is essentially conditional

NO unconditional inequality $x \leq 1 + \lambda y$

$\exists$ one essentially conditional inequality $\Longrightarrow$ the grey area is not polyhedral

$\Longrightarrow$ $\exists$ infinitely many independent tangent lines

A geometric view on conditional inequalities:



if $y = 0$ then $x \leq 1 \longleftarrow$ this inequality is essentially conditional

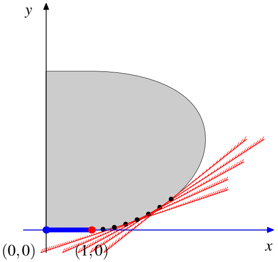
NO unconditional inequality $x \leq 1 + \lambda y$

$\exists$ one essentially conditional inequality $\Longrightarrow$ the grey area is not polyhedral

$\Longrightarrow \exists$ infinitely many independent tangent lines

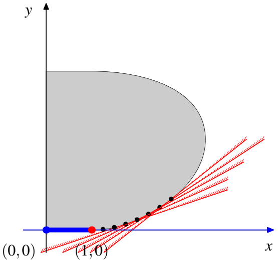
formal proof: Farkas lemma

A geometric view on conditional inequalities:



if $y = 0$ then $x \leq 1$ $\longleftarrow$ this inequality is essentially conditional

A geometric view on conditional inequalities:



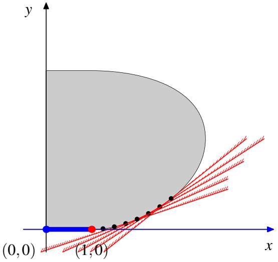
if $y = 0$ then $x \leq 1$ $\longleftarrow$ this inequality is essentially conditional

one essentially conditional inequality for *the grey area*

$\Downarrow$

infinitely many independent unconditional inequality for *the grey area*

A geometric view on conditional inequalities:

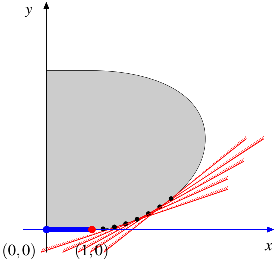


one essentially conditional inequality for *(almost) entropic points* ($\geq$ 4 r.v.)

$\Downarrow$

infinitely many unconditional information inequalities

A geometric view on conditional inequalities:

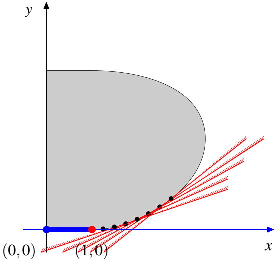


one essentially conditional inequality for *(almost) entropic points* ($\geq$ 4 r.v.)

$\Downarrow$

infinitely many unconditional information inequalities (in $\mathbb{R}^{15}$)

A geometric view on conditional inequalities:



one essentially conditional inequality for *(almost) entropic points* ($\geq$ 4 r.v.)

$\Downarrow$

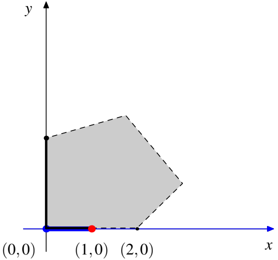
infinitely many unconditional information inequalities (in $\mathbb{R}^{15}$)

$\Downarrow$

Theorem (Matúš)

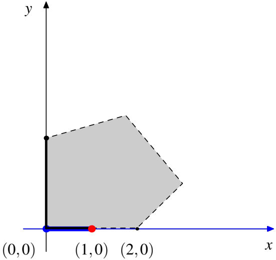
There exist infinitely many independent linear information inequalities.

A geometric view on conditional inequalities:



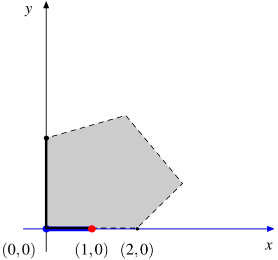
if $y = 0$ then $x \leq 1$

A geometric view on conditional inequalities:



if $y = 0$ then $x \leq 1$ $\Longleftarrow$ from a complex structure of the borderline

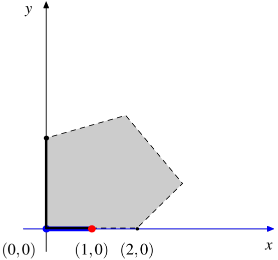
A geometric view on conditional inequalities:



if $y = 0$ then $x \leq 1$ $\Longleftarrow$ from a complex structure of the borderline

NO unconditional inequality $x \leq 1 + \lambda y$

A geometric view on conditional inequalities:



if $y = 0$ then $x \leq 1$ $\Longleftarrow$ from a complex structure of the borderline

NO unconditional inequality $x \leq 1 + \lambda y$

this inequality is also essentially conditional

Outline

Digression:
How people prove unconditional information inequalities

Digression:

How people prove unconditional information inequalities

People use **conditional** inequalities with **delusive** constraints.

Digression:

How people prove unconditional information inequalities

People use **conditional** inequalities with **delusive** constraints.

Simplified Example:

If $I(a, b : z|xy) = 0$ then

$$\mathrm{I}(x : y) \leq \mathrm{I}(x : y|a) + \mathrm{I}(x : y|b) + \mathrm{I}(a : b) + \mathrm{I}(x : y|z) + \mathrm{I}(x : z|y) + \mathrm{I}(y : z|x)$$

[Shannon-type conditional inequality]

Digression:
How people prove unconditional information inequalities

People use **conditional** inequalities with **delusive** constraints.

Simplified Example:

If $I(a, b : z|xy) = 0$ then

$$\mathrm{I}(x : y) \leq \mathrm{I}(x : y|a) + \mathrm{I}(x : y|b) + \mathrm{I}(a : b) + \mathrm{I}(x : y|z) + \mathrm{I}(x : z|y) + \mathrm{I}(y : z|x)$$

[Shannon-type conditional inequality]

We forget the **constraint** and obtain a non-Shannon type unconditional inequality.

Digression:
How people prove unconditional information inequalities

People use **conditional** inequalities with **delusive** constraints.

More "physical" example: Ahlswede–Körner Lemma

in more detail: the talk of **Carles Padró**

Digression:
How people prove unconditional information inequalities

People use **conditional** inequalities with **delusive** constraints.

Classical argument [Zhang–Yeung]:

Copy Lemma

For all (a, b, x, y) *there is* a' *(clone of* a *conditional on* (x, y)*) such that*

- $\mathrm{H}(a') = \mathrm{H}(a)$,
  $\mathrm{H}(a', x) = \mathrm{H}(a, x)$, $\mathrm{H}(a', y) = \mathrm{H}(a, y)$,
  $\mathrm{H}(a', x, y) = \mathrm{H}(a, x, y)$
- a' *and* (a, b) *are independent conditional on* (x, y)

Digression:

How people prove unconditional information inequalities

People use **conditional** inequalities with **delusive** constraints.

Classical argument [Zhang–Yeung]:

Copy Lemma

For all (a, b, x, y) there is a' (clone of a conditional on (x, y)) such that

- $\mathrm{H}(a') = \mathrm{H}(a)$,
  $\mathrm{H}(a', x) = \mathrm{H}(a, x)$, $\mathrm{H}(a', y) = \mathrm{H}(a, y)$,
  $\mathrm{H}(a', x, y) = \mathrm{H}(a, x, y)$
- *a' and (a, b) are independent conditional on (x, y)*

If a' satisfies these constraints then

$\mathrm{I}(x : y) \leq \mathrm{I}(x : y|a) + \mathrm{I}(x : y|b) + \mathrm{I}(a : b) + \mathrm{I}(x : y|a) + \mathrm{I}(x : a|y) + \mathrm{I}(y : a|x)$

[Shannon-type conditional inequality]

Digression:

How people prove unconditional information inequalities

People use **conditional** inequalities with **delusive** constraints.

Classical argument [Zhang–Yeung]:

Copy Lemma

For all (a, b, x, y) *there is* a' *(clone of* a *conditional on* (x, y)*) such that*

- $\mathrm{H}(a') = \mathrm{H}(a)$,
  $\mathrm{H}(a', x) = \mathrm{H}(a, x)$, $\mathrm{H}(a', y) = \mathrm{H}(a, y)$,
  $\mathrm{H}(a', x, y) = \mathrm{H}(a, x, y)$
- a' *and* (a, b) *are independent conditional on* (x, y)

If a' satisfies these constraints then

$\mathrm{I}(x : y) \leq \mathrm{I}(x : y|a) + \mathrm{I}(x : y|b) + \mathrm{I}(a : b) + \mathrm{I}(x : y|a) + \mathrm{I}(x : a|y) + \mathrm{I}(y : a|x)$

[Shannon-type conditional inequality]

We forget the **constraint** and obtain a non-Shannon type unconditional inequality.

Digression:

How people prove unconditional information inequalities

People use **conditional** inequalities with **delusive** constraints.

Classical example [Zhang–Yeung]:

Copy Lemma

For all (a, b, x, y) *there is* a' *(clone of a conditional on* (x, y)*) such that*

- $\mathrm{H}(a') = \mathrm{H}(a)$,
  $\mathrm{H}(a', x) = \mathrm{H}(a, x)$, $\mathrm{H}(a', y) = \mathrm{H}(a, y)$,
  $\mathrm{H}(a', x, y) = \mathrm{H}(a, x, y)$
- a' *and* (a, b) *are independent conditional on* (x, y)

All known proofs of non-Shannon type unconditional inequalities can be translated in the language of the Copy Lemma [observed by T. Kaced].

Outline

How to prove a conditional inequality: a toy example

Proposition

If $\mathrm{I}(x:a|y) = \mathrm{I}(y:a|x) = 0$ *then* $\mathrm{I}(x:y) \leq 2\mathrm{I}(x:y|a) + I(x:y|b) + \mathrm{I}(a:b)$

How to prove a conditional inequality: a toy example

Proposition

If $\mathrm{I}(x:a|y) = \mathrm{I}(y:a|x) = 0$ *then* $\mathrm{I}(x:y) \leq 2\mathrm{I}(x:y|a) + I(x:y|b) + \mathrm{I}(a:b)$

Lazy proof: We know from [Zhang-Yeung 98] that for **all** (a, b, x, y)

$$\mathrm{I}(x:y) \leq 2\mathrm{I}(x:y|a) + I(x:y|b) + \mathrm{I}(a:b) + \mathrm{I}(x:a|y) + \mathrm{I}(y:a|x)$$

This *universal* inequality implies our *conditional* inequality.

How to prove a conditional inequality: a toy example

Proposition

If $\mathrm{I}(x:a|y) = \mathrm{I}(y:a|x) = 0$ *then* $\mathrm{I}(x:y) \leq 2\mathrm{I}(x:y|a) + I(x:y|b) + \mathrm{I}(a:b)$

Direct application of the Copy Lemma (from the proof of [Zhang-Yeung 98]):
Every tuple **all** (a, b, x, y) can be extended to (a, b, x, y, a') such that

- (a', x, y) has the same distribution as (a, x, y)
- a' and (a, b) are independent conditional on (x, y)

[we have made a **clone** of a conditional on (x, y)]

How to prove a conditional inequality: a toy example

Proposition

If $\mathrm{I}(x:a|y) = \mathrm{I}(y:a|x) = 0$ *then* $\mathrm{I}(x:y) \leq 2\mathrm{I}(x:y|a) + I(x:y|b) + \mathrm{I}(a:b)$

Direct application of the Copy Lemma (from the proof of [Zhang-Yeung 98]):
Every tuple **all** (a, b, x, y) can be extended to (a, b, x, y, a') such that

- (a', x, y) has the same distribution as (a, x, y)
- a' and (a, b) are independent conditional on (x, y)

[we have made a **clone** of a conditional on (x, y)]

[Shannon-type inequalities + our constraints + definition of a'] $\Longrightarrow$ our inequality.

How to prove a conditional inequality: a toy example

Proposition

If $\mathrm{I}(x : a|y) = \mathrm{I}(y : a|x) = 0$ *then* $\mathrm{I}(x : y) \leq 2\mathrm{I}(x : y|a) + I(x : y|b) + \mathrm{I}(a : b)$

Direct application of the Copy Lemma (from the proof of [Zhang-Yeung 98]):
Every tuple **all** (a, b, x, y) can be extended to (a, b, x, y, a') such that

- (a', x, y) has the same distribution as (a, x, y)
- a' and (a, b) are independent conditional on (x, y)

[we have made a **clone** of a conditional on (x, y)]

There is a Shannon type inequality

$$\mathrm{I}(x : y) \leq \mathrm{I}(x : y|a) + I(x : y|b) + \mathrm{I}(a : b) + \mathrm{I}(x : a'|y) + \mathrm{I}(y : a'|x) + \mathrm{I}(x : y|a') + 3\mathrm{I}(a' : a, b|x, y)$$

[this inequality + our constraints + definition of a'] $\Longrightarrow$ our inequality.

How to prove a conditional inequality: a toy example

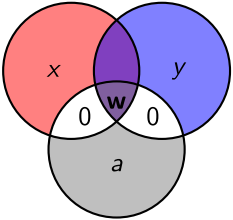
Proposition

If $\mathrm{I}(x:a|y) = \mathrm{I}(y:a|x) = 0$ *then* $\mathrm{I}(x:y) \leq 2\mathrm{I}(x:y|a) + I(x:y|b) + \mathrm{I}(a:b)$

Materialization of the mutual information:

Lemma on Double Markov Property.
For all (a, x, y), if $\mathrm{I}(x:a|y) = \mathrm{I}(y:a|x) = 0$ then there exists a w such that

- $\mathrm{H}(w) = \mathrm{I}(x, y : a)$,
- $\mathrm{H}(w|x) = \mathrm{H}(w|y) = 0$.



How to prove a conditional inequality: a toy example

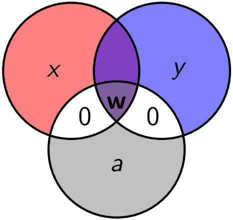
Proposition

If $\mathrm{I}(x : a|y) = \mathrm{I}(y : a|x) = 0$ *then* $\mathrm{I}(x : y) \leq 2\mathrm{I}(x : y|a) + I(x : y|b) + \mathrm{I}(a : b)$

Materialization of the mutual information:

Lemma on Double Markov Property.
For all (a, x, y), if $\mathrm{I}(x : a|y) = \mathrm{I}(y : a|x) = 0$ then there exists a w such that

- $\mathrm{H}(w) = \mathrm{I}(x, y : a)$,
- $\mathrm{H}(w|x) = \mathrm{H}(w|y) = 0$.



[Shannon-type inequalities + our constraints + definition of w] $\Longrightarrow$ our inequality.

How to prove a conditional inequality: a toy example

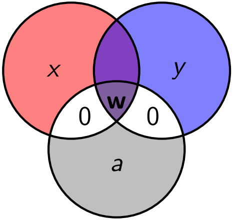
Proposition

If $\mathrm{I}(x:a|y) = \mathrm{I}(y:a|x) = 0$ *then* $\mathrm{I}(x:y) \leq 2\mathrm{I}(x:y|a) + I(x:y|b) + \mathrm{I}(a:b)$

Materialization of the mutual information:

Lemma on Double Markov Property.
For all (a, x, y), if
$\mathrm{I}(x:a|y) = \mathrm{I}(y:a|x) = 0$ then there exists a w such that

- $\mathrm{H}(w) = \mathrm{I}(x, y : a)$,
- $\mathrm{H}(w|x) = \mathrm{H}(w|y) = 0$.



For all a, b, x, y, w we have the following Shannon type inequality

$$\mathrm{H}(w) \leq 2\mathrm{H}(w|x) + 2\mathrm{H}(w|y) + \mathrm{I}(x:y|a) + \mathrm{I}(x:y|b) + \mathrm{I}(a:b)$$

How to prove a conditional inequality: a toy example

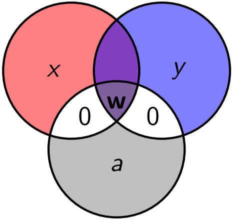
Proposition

If $\mathrm{I}(x:a|y) = \mathrm{I}(y:a|x) = 0$ *then* $\mathrm{I}(x:y) \leq 2\mathrm{I}(x:y|a) + I(x:y|b) + \mathrm{I}(a:b)$

Materialization of the mutual information:

Lemma on Double Markov Property.
For all (a, x, y), if $\mathrm{I}(x:a|y) = \mathrm{I}(y:a|x) = 0$ then there exists a w such that

- $\mathrm{H}(w) = \mathrm{I}(x, y : a)$,
- $\mathrm{H}(w|x) = \mathrm{H}(w|y) = 0$.



For all a, b, x, y, w we have the following Shannon type inequality

$$\underbrace{\mathrm{H}(w)}_{=\,\mathrm{I}(x,y:a)} \leq \underbrace{2\mathrm{H}(w|x)}_{=\,0} + \underbrace{2\mathrm{H}(w|y)}_{=\,0} + \mathrm{I}(x:y|a) + \mathrm{I}(x:y|b) + \mathrm{I}(a:b)$$

How to prove a conditional inequality: a toy example

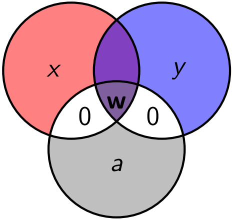
Proposition

If $\mathrm{I}(x:a|y) = \mathrm{I}(y:a|x) = 0$ *then* $\mathrm{I}(x:y) \leq 2\mathrm{I}(x:y|a) + I(x:y|b) + \mathrm{I}(a:b)$

Materialization of the mutual information:

Lemma on Double Markov Property.
For all (a, x, y), if
$\mathrm{I}(x:a|y) = \mathrm{I}(y:a|x) = 0$ then there exists a w such that

- $\mathrm{H}(w) = \mathrm{I}(x, y : a)$,
- $\mathrm{H}(w|x) = \mathrm{H}(w|y) = 0$.



For all a, b, x, y, w we have the following Shannon type inequality

$$\begin{array}{ccccccc} \mathrm{H}(w) & \leq & 2\mathrm{H}(w|x) & + & 2\mathrm{H}(w|y) & + & \mathrm{I}(x:y|a) + \mathrm{I}(x:y|b) + \mathrm{I}(a:b) \\ \| & & \| & & \| & & \\ \mathrm{I}(x,y:a) & & 0 & & 0 & & \\ \| & & & & & & \\ \mathrm{I}(x:y) & - & \mathrm{I}(x:y|a) & & & & \end{array}$$

How to prove a conditional inequality: a toy example

Proposition

If $\mathrm{I}(x : a|y) = \mathrm{I}(y : a|x) = 0$ *then* $\mathrm{I}(x : y) \leq 2\mathrm{I}(x : y|a) + I(x : y|b) + \mathrm{I}(a : b)$

How to prove a conditional inequality: a toy example

Proposition

If $\mathrm{I}(x : a|y) = \mathrm{I}(y : a|x) = 0$ *then* $\mathrm{I}(x : y) \leq 2\mathrm{I}(x : y|a) + I(x : y|b) + \mathrm{I}(a : b)$

How to prove a conditional inequality: a more serious approach

Theorem (Matúš)

If $\mathrm{I}(x : a|y) = \mathrm{I}(y : a|x) = 0$ *then* $\mathrm{I}(x : y) \leq \mathrm{I}(x : y|a) + \mathrm{I}(x : y|b) + \mathrm{I}(a : b)$

How to prove a conditional inequality: a more serious approach

Theorem (Matúš)

If $\mathrm{I}(x : a|y) = \mathrm{I}(y : a|x) = 0$ *then* $\mathrm{I}(x : y) \leq \mathrm{I}(x : y|a) + \mathrm{I}(x : y|b) + \mathrm{I}(a : b)$

Idea of the proof:

Approximate **this** inequality by infinitely many non-Shannon type inequalities.

How to prove a conditional inequality: a more serious approach

Theorem (Matúš)

If $\mathrm{I}(x:a|y) = \mathrm{I}(y:a|x) = 0$ *then* $\mathrm{I}(x:y) \leq \mathrm{I}(x:y|a) + \mathrm{I}(x:y|b) + \mathrm{I}(a:b)$

Sketch of the proof: For each integer $k > 0$ we can prove the following *non-Shannon type* inequality

$$\begin{aligned}\mathrm{I}(x:y) \quad \leq \quad & \mathrm{I}(x:y|a) + I(x:y|b) + \mathrm{I}(a:b) \\ & \tfrac{1}{k}\mathrm{I}(x:y|a) + \tfrac{k+1}{2}\big(\mathrm{I}(x:a|y) + \mathrm{I}(y:a|x)\big)\end{aligned}$$

How to prove a conditional inequality: a more serious approach

Theorem (Matúš)

If $\mathrm{I}(x:a|y) = \mathrm{I}(y:a|x) = 0$ *then* $\mathrm{I}(x:y) \leq \mathrm{I}(x:y|a) + \mathrm{I}(x:y|b) + \mathrm{I}(a:b)$

Sketch of the proof: For each integer $k > 0$ we can prove the following *non-Shannon type* inequality

$$\begin{aligned} \mathrm{I}(x:y) \leq\ & \mathrm{I}(x:y|a) + I(x:y|b) + \mathrm{I}(a:b) \\ & \tfrac{1}{k}\mathrm{I}(x:y|a) + \tfrac{k+1}{2}\big(\mathrm{I}(x:a|y) + \mathrm{I}(y:a|x)\big) \end{aligned}$$

It remains to let $k \to \infty$.

Ad hoc proof of an essentially conditional inequality

Theorem

If $\mathrm{I}(x:y|a) = \mathrm{H}(a|x,y) = 0$ *then* $\mathrm{I}(x:y) \leq \mathrm{I}(x:y|a) + \mathrm{I}(x:y|b) + \mathrm{I}(a:b)$

Ad hoc proof of an essentially conditional inequality

Theorem

If $\mathrm{I}(x:y|a) = \mathrm{H}(a|x,y) = 0$ *then* $\mathrm{I}(x:y) \leq \mathrm{I}(x:y|a) + \mathrm{I}(x:y|b) + \mathrm{I}(a:b)$

Idea of the proof: augmented Copy Lemma

Ad hoc proof of an essentially conditional inequality

Theorem

If $\mathrm{I}(x:y|a) = \mathrm{H}(a|x,y) = 0$ *then* $\mathrm{I}(x:y) \leq \mathrm{I}(x:y|a) + \mathrm{I}(x:y|b) + \mathrm{I}(a:b)$

Sketch of the proof (augmented Copy Lemma):

- make independent **clones** x' and y' for x and y respectively conditional on (a,b)

Ad hoc proof of an essentially conditional inequality

Theorem

If $\mathrm{I}(x:y|a) = \mathrm{H}(a|x,y) = 0$ *then* $\mathrm{I}(x:y) \leq \mathrm{I}(x:y|a) + \mathrm{I}(x:y|b) + \mathrm{I}(a:b)$

Sketch of the proof (augmented Copy Lemma):

- make independent **clones** x' and y' for x and y respectively conditional on (a,b)
- observation 1:

$$\begin{aligned}\mathrm{H}(x',y',a,b) &= \mathrm{H}(a,b) + \mathrm{H}(x'|a,b) + \mathrm{H}(y'|a,b) \\ &= \mathrm{H}(a,b) + \mathrm{H}(x|a,b) + \mathrm{H}(y|a,b)\end{aligned}$$

Ad hoc proof of an essentially conditional inequality

Theorem

If $\mathrm{I}(x:y|a) = \mathrm{H}(a|x,y) = 0$ *then* $\mathrm{I}(x:y) \leq \mathrm{I}(x:y|a) + \mathrm{I}(x:y|b) + \mathrm{I}(a:b)$

Sketch of the proof (augmented Copy Lemma):

- make independent **clones** x' and y' for x and y respectively conditional on (a,b)
- observation 1:

$$\begin{aligned} \mathrm{H}(x',y',a,b) &= \mathrm{H}(a,b) + \mathrm{H}(x'|a,b) + \mathrm{H}(y'|a,b) \\ &= \mathrm{H}(a,b) + \mathrm{H}(x|a,b) + \mathrm{H}(y|a,b) \end{aligned}$$

- observation 2:

$$\begin{aligned} \mathrm{H}(x',y',a,b) &\leq \mathrm{H}(b) + \mathrm{H}(x'|b) + \mathrm{H}(y'|b) + \mathrm{H}(a|x',y') \\ &= \mathrm{H}(b) + \mathrm{H}(x|b) + \mathrm{H}(y|b) + 0 \end{aligned}$$

Ad hoc proof of an essentially conditional inequality

Theorem

If $\mathrm{I}(x:y|a) = \mathrm{H}(a|x,y) = 0$ *then* $\mathrm{I}(x:y) \leq \mathrm{I}(x:y|a) + \mathrm{I}(x:y|b) + \mathrm{I}(a:b)$

Sketch of the proof (augmented Copy Lemma):

- make independent **clones** x' and y' for x and y respectively conditional on (a,b)
- observation 1:

$$\begin{aligned}\mathrm{H}(x',y',a,b) &= \mathrm{H}(a,b) + \mathrm{H}(x'|a,b) + \mathrm{H}(y'|a,b) \\ &= \mathrm{H}(a,b) + \mathrm{H}(x|a,b) + \mathrm{H}(y|a,b)\end{aligned}$$

- observation 2:

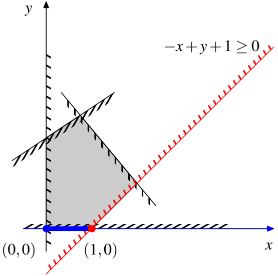
$$\begin{aligned}\mathrm{H}(x',y',a,b) &\leq \mathrm{H}(b) + \mathrm{H}(x'|b) + \mathrm{H}(y'|b) + \mathrm{H}(a|x',y') \\ &= \mathrm{H}(b) + \mathrm{H}(x|b) + \mathrm{H}(y|b) + 0\end{aligned}$$

- observation 3: $\mathrm{H}(a,b) + \mathrm{H}(x|a,b) + \mathrm{H}(y|a,b) \leq \mathrm{H}(b) + \mathrm{H}(x|b) + \mathrm{H}(y|b)$ is equivalent to Ingleton's inequality

Outline

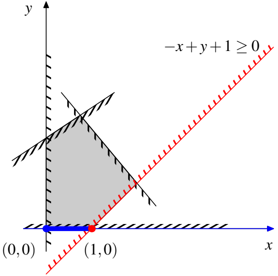
Outline

Applications (1):
Non-essentially conditional inequalities



if $y = 0$ then $x \leq 1 \;\; \Longleftarrow x \leq 1 + y$

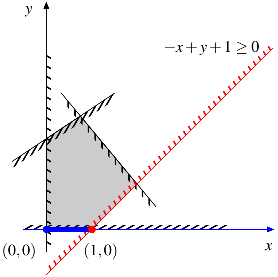
Applications (1):
Non-essentially conditional inequalities



if $y = 0$ then $x \leq 1 \Longleftarrow x \leq 1 + y$

This case looks simple and boring.

Applications (1):
Non-essentially conditional inequalities



if $y = 0$ then $x \leq 1 \;\; \Longleftarrow x \leq 1 + y$

This case looks simple and boring. **But it is not!**

Applications (1):
Non-essentially conditional inequalities

Archetypical example: lower bounds in **secret sharing**.

Applications (1):
Non-essentially conditional inequalities

Archetypical example: lower bounds in **secret sharing**.

[constraints of a secret sharing scheme] $\Longrightarrow$ [some bounds for the size of shares]

Secret sharing, reminder (1)

- secret S_0 (e.g., uniformly distributed on $\{0, 1\}^k$)
- n participants
- **access structure**: a family of authorized groups $C_1, \ldots, C_m$

Secret sharing, reminder (1)

- secret S_0 (e.g., uniformly distributed on $\{0,1\}^k$)
- n participants
- **access structure**: a family of authorized groups $C_1, \ldots, C_m$

perfect secret sharing scheme: a distribution $(S_0, S_1, \ldots, S_n)$ such that

- a collection of shares S_i from each authorized group gives **all** information on S_0
- a collection of shares S_i from any *non-authorized* group gives **no** information on S_0

Secret sharing, reminder (2)

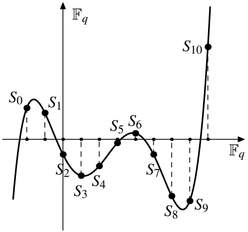
secret key: S_0 uniformly distributed on $\{0,1\}^k$

Standard example:

- any group of $\geq t$ participants knows the secret
- any group of $< t$ participants know nothing about the secret

Classical solution (Shamir scheme):

- fix points $x_0, x_1, \ldots, x_n$ in $\mathbb{F}_{2^k}$ (public information)
- choose a secret random polynomial $Q(x)$ of degree $\leq t-1$
- the i-th participant obtains $S_i = Q(x_i)$, $i = 1, \ldots, n$
- let the **secret** $S_0 = Q(x_0)$



Secret sharing, reminder (2)

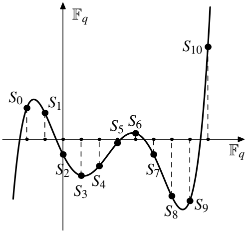
secret key: S_0 uniformly distributed on $\{0,1\}^k$

Standard example:

- any group of $\geq t$ participants knows the secret
- any group of $< t$ participants know nothing about the secret

Classical solution (Shamir scheme):

- fix points $x_0, x_1, \ldots, x_n$ in $\mathbb{F}_{2^k}$ (public information)
- choose a secret random polynomial $Q(x)$ of degree $\leq t-1$
- the i-th participant obtains $S_i = Q(x_i)$, $i = 1, \ldots, n$
- let the **secret** $S_0 = Q(x_0)$



Given $\geq t$ pairs $(x_i, Q(x_i))$ we reconstruct $Q(x)$ and S_0.

Secret sharing, reminder (2)

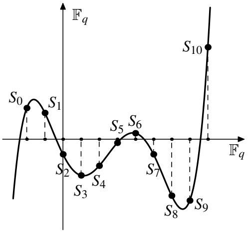
secret key: S_0 uniformly distributed on $\{0,1\}^k$

Standard example:

- any group of $\geq t$ participants knows the secret
- any group of $< t$ participants know nothing about the secret

Classical solution (Shamir scheme):

- fix points $x_0, x_1, \ldots, x_n$ in $\mathbb{F}_{2^k}$ (public information)
- choose a secret random polynomial $Q(x)$ of degree $\leq t-1$
- the i-th participant obtains $S_i = Q(x_i)$, $i = 1, \ldots, n$
- let the **secret** $S_0 = Q(x_0)$



Given $< t$ pairs $(x_i, Q(x_i))$ we know nothing about S_0:
all values of S_0 remain **possible** and even **equiprobable**.

Secret sharing, reminder (3)

Information ratio of a secret sharing scheme: $\frac{\max \mathrm{H}(S_i)}{\mathrm{H}(S_0)}$.

Fundamental problem: minimize information ratio for a given access structure.

Secret sharing, reminder (3)

Information ratio of a secret sharing scheme: $\frac{\max \mathrm{H}(S_i)}{\mathrm{H}(S_0)}$.

Fundamental problem: minimize information ratio for a given access structure.

Very simple example:

- 4 participants
- **minimal** authorized groups:
  $\{1,2\}$, $\{2,3\}$, $\{3,4\}$

Question: What is the optimal information ratio for this access structure?

There is a simple construction with information ratio $= 3/2$.

Shannon's inequalities $\Longrightarrow$ we cannot do better.

Secret sharing, reminder (3)

Information ratio of a secret sharing scheme: $\frac{\max \mathrm{H}(S_i)}{\mathrm{H}(S_0)}$.

Fundamental problem: minimize information ratio for a given access structure.

Very simple example:

- 4 participants
- **minimal** authorized groups:
  $\{1, 2\}$, $\{2, 3\}$, $\{3, 4\}$

Question: What is the optimal information ratio for this access structure?

There is a simple construction with information ratio $= 3/2$.

Shannon's inequalities $\Longrightarrow$ we cannot do better.
[This is a conditional information inequality!]

Secret sharing: computing the information ratio

Very simple example:

- 4 participants
- **minimal** authorized groups:
  $\{1,2\}$, $\{2,3\}$, $\{3,4\}$

Question: What is the optimal information ratio for this access structure?

Shannon's inequalities: information ratio $\geq 3/2$.

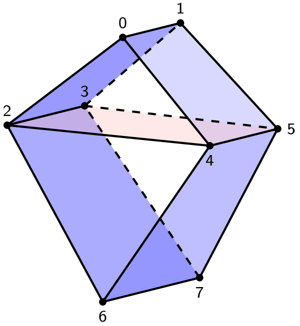
Computer-assisted proof:

- write down all equations that define the access structure
- write down all *basic inequalities* for Shannon's entropy of $(S_0, S_1, S_2, S_3, S_4)$
- write that $\mathrm{H}(S_i) \leq T$ for $i = 1, 2, 3, 4$
- ask your favorite **linear programming solver** to find $\min(T)$

The answer: minimal $T = (3/2)\mathrm{H}(S_0)$.

Vámos matroid

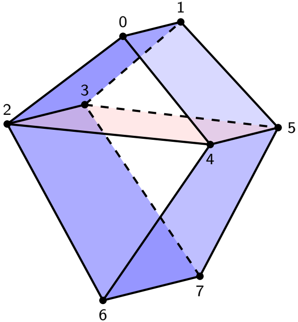
ground set $= \{0, 1, 2, 3, 4, 5, 6, 7\}$



$\mathrm{rk}(\text{one point}) = 1$
$\mathrm{rk}(\text{two points}) = 2$
$\mathrm{rk}(\text{three points}) = 3$
$\mathrm{rk}(\{0, 1, 2, 3\}) = \mathrm{rk}(\{0, 1, 4, 5\}) = \mathrm{rk}(\{2, 3, 6, 7\}) = \mathrm{rk}(\{4, 5, 6, 7\}) = \mathrm{rk}(\{2, 3, 4, 5\}) = 3$
$\mathrm{rk}(\text{other sets}) = 4$

Vámos matroid

ground set $= \{0, 1, 2, 3, 4, 5, 6, 7\}$



$\mathrm{rk}(\text{one point}) = 1$
$\mathrm{rk}(\text{two points}) = 2$
$\mathrm{rk}(\text{three points}) = 3$
$\mathrm{rk}(\{0, 1, 2, 3\}) = \mathrm{rk}(\{0, 1, 4, 5\}) = \mathrm{rk}(\{2, 3, 6, 7\}) = \mathrm{rk}(\{4, 5, 6, 7\}) = \mathrm{rk}(\{2, 3, 4, 5\}) = 3$
$\mathrm{rk}(\text{other sets}) = 4$

An access structure on this matroid: participants $\{1, \ldots, 7\}$, and

$i_1, \ldots, i_s$ **know the secret** if and only if $\mathrm{rk}(i_1, \ldots, i_s) = \mathrm{rk}(0, i_1, \ldots, i_s)$

secret sharing on matroids: why do we care ?

Matroids:
a structure with a **rank** function generalizing ranks of linear (sub)spaces

secret sharing on matroids: why do we care ?

Matroids:
a structure with a **rank** function generalizing ranks of linear (sub)spaces

[Brickell–Davenport]: The access structure of every ideal secret sharing scheme can be defined on a matroid.

Natural conjecture: For every access structure on a matroid there is an ideal secret sharing scheme

secret sharing on matroids: why do we care ?

Matroids:
a structure with a **rank** function generalizing ranks of linear (sub)spaces

[Brickell–Davenport]: The access structure of every ideal secret sharing scheme can be defined on a matroid.

Natural conjecture: For every access structure on a matroid there is an ideal secret sharing scheme

The conjecture looks plausible: This is true for **linear** access structures.

very plausible: Shannon's inequalities cannot disprove it.

secret sharing on matroids: why do we care ?

Matroids:
a structure with a **rank** function generalizing ranks of linear (sub)spaces

[Brickell–Davenport]: The access structure of every ideal secret sharing scheme can be defined on a matroid.

Natural conjecture: For every access structure on a matroid there is an ideal secret sharing scheme

The conjecture looks plausible: This is true for **linear** access structures.

very plausible: Shannon's inequalities cannot disprove it.

But there is a counter-example [Seymour]: Vámos matroid

our toy problem: secret sharing on Vámos matroid

Problem:
Find the optimal information ratio for a secret sharing on this access structure.

our toy problem: secret sharing on Vámos matroid

Problem:
Find the optimal information ratio for a secret sharing on this access structure.

upper bound: information ratio $\leq 4/3$

our toy problem: secret sharing on Vámos matroid

Problem:
Find the optimal information ratio for a secret sharing on this access structure.

upper bound: information ratio $\leq 4/3$

lower bound:

Seymour 1992	> 1

our toy problem: secret sharing on Vámos matroid

Problem:
Find the optimal information ratio for a secret sharing on this access structure.

upper bound: information ratio $\leq 4/3$

lower bound:

Seymour 1992	> 1
Beimel–Livne 2006	$\geq 1 + \Omega(1/\sqrt{k})$ for a secret of size k

our toy problem: secret sharing on Vámos matroid

Problem:
Find the optimal information ratio for a secret sharing on this access structure.

upper bound: information ratio $\leq 4/3$

lower bound:

Seymour 1992	> 1
Beimel–Livne 2006	$\geq 1 + \Omega(1/\sqrt{k})$ for a secret of size k
Beimel–Livne–Padró 2008	$\geq 11/10$

our toy problem: secret sharing on Vámos matroid

Problem:
Find the optimal information ratio for a secret sharing on this access structure.

upper bound: information ratio $\leq 4/3$

lower bound:

Seymour 1992	> 1
Beimel–Livne 2006	$\geq 1 + \Omega(1/\sqrt{k})$ for a secret of size k
Beimel–Livne–Padró 2008	$\geq 11/10$
Metcalf-Burton 2011	$\geq 9/8 = 1.125$

our toy problem: secret sharing on Vámos matroid

Problem:
Find the optimal information ratio for a secret sharing on this access structure.

upper bound: information ratio $\leq 4/3$

lower bound:

Seymour 1992	> 1
Beimel–Livne 2006	$\geq 1 + \Omega(1/\sqrt{k})$ for a secret of size k
Beimel–Livne–Padró 2008	$\geq 11/10$
Metcalf-Burton 2011	$\geq 9/8 = 1.125$
Hadian 2013	$\geq 67/59 \approx 1.135593$

our toy problem: secret sharing on Vámos matroid

Problem:
Find the optimal information ratio for a secret sharing on this access structure.

upper bound: information ratio $\leq 4/3$

lower bound:

Seymour 1992	> 1
Beimel–Livne 2006	$\geq 1 + \Omega(1/\sqrt{k})$ for a secret of size k
Beimel–Livne–Padró 2008	$\geq 11/10$
Metcalf-Burton 2011	$\geq 9/8 = 1.125$
Hadian 2013	$\geq 67/59 \approx 1.135593$
Farràs–Kaced–Martín–Padró 2018	$\geq 33/29 \approx 1.137931$

our toy problem: secret sharing on Vámos matroid

Problem:
Find the optimal information ratio for a secret sharing on this access structure.

upper bound: information ratio $\leq 4/3$

lower bound:

Seymour 1992	> 1
Beimel–Livne 2006	$\geq 1 + \Omega(1/\sqrt{k})$ for a secret of size k
Beimel–Livne–Padró 2008	$\geq 11/10$
Metcalf-Burton 2011	$\geq 9/8 = 1.125$
Hadian 2013	$\geq 67/59 \approx 1.135593$
Farràs–Kaced–Martín–Padró 2018	$\geq 33/29 \approx 1.137931$
Gürpınar-R. 2019	$\geq 561/491 \approx 1.142566$

our toy problem: secret sharing on Vámos matroid

Problem:
Find the optimal information ratio for a secret sharing on this access structure.

upper bound: information ratio $\leq 4/3$

lower bound:

Seymour 1992	> 1
Beimel–Livne 2006	$\geq 1 + \Omega(1/\sqrt{k})$ for a secret of size k
Beimel–Livne–Padró 2008	$\geq 11/10$
Metcalf-Burton 2011	$\geq 9/8 = 1.125$
Hadian 2013	$\geq 67/59 \approx 1.135593$
Farràs–Kaced–Martín–Padró 2018	$\geq 33/29 \approx 1.137931$
Gürpınar-R. 2019	$\geq 561/491 \approx 1.142566$

The last two bounds follow from new (unknown!) inequalities for Shannon's entropy. They remain undiscovered, but we have already applied them.

Classical approach

Write a **linear program** as follows.

Constraints:

- equations from the definition of a **perfect secret sharing**
- all **Shannon-type** inequalities for entropy, $I(* : * \mid *) \geq 0$
- (optional) symmetry conditions

Objective function:

minimize $\left[\max_i \frac{\mathrm{H}(\text{secret share}_i)}{\mathrm{H}(\text{secret})} \right]$

Classical approach

Write a **linear program** as follows.

Constraints:

- equations from the definition of a **perfect secret sharing**
- all **Shannon-type** inequalities for entropy, $I(* : * \mid *) \geq 0$
- (optional) symmetry conditions

Objective function:

minimize $\left[\max_i \frac{\mathrm{H}(\text{secret share}_i)}{\mathrm{H}(\text{secret})} \right]$

Answer: trivial, information ratio ≥ 1 [for secret sharing on matroids]

Modern approach

Write a **linear program** as follows

Constraints:

- equations from the definition of a **perfect secret sharing**
- all **Shannon-type** inequalities $I(* : * \mid *) \geq 0$
- some **known non-Shannon-type** inequalities
- (optional) symmetry conditions

Objective function:

minimize $\left[\max_i \frac{\mathrm{H}(\text{secret share}_i)}{\mathrm{H}(\text{secret})} \right]$

Modern approach

Write a **linear program** as follows

Constraints:

- equations from the definition of a **perfect secret sharing**
- all **Shannon-type** inequalities $I(* : * \mid *) \geq 0$
- some **known non-Shannon-type** inequalities
- (optional) symmetry conditions

Objective function:

minimize $\left[\max_i \frac{\mathrm{H}(\text{secret share}_i)}{\mathrm{H}(\text{secret})} \right]$

Answer: some non-trivial bounds!
[Beimel-Livne-Padró 2008], [Metcalf-Burton 2011], [Hadian 2013]

PostModern approach

Write a **linear program** as follows

Constraints:

- equations from the definition of a **perfect secret sharing**
- all **Shannon-type** inequalities $I(* : * \mid *) \geq 0$
- ~~some **known non-Shannon-type** inequalities~~
- new variables and constraints borrowed from proofs of non-Shannon-type inequalities [Ahlswede-Körner or Copy lemma]
- (optional) symmetry conditions

Objective function:

minimize $\left[\max_i \frac{\mathrm{H}(\text{secret share}_i)}{\mathrm{H}(\text{secret})}\right]$

PostModern approach

Write a **linear program** as follows

Constraints:

- equations from the definition of a **perfect secret sharing**
- all **Shannon-type** inequalities $I(* : * \,|\, *) \geq 0$
- ~~some **known non-Shannon-type** inequalities~~
- new variables and constraints borrowed from proofs of non-Shannon-type inequalities [Ahlswede-Körner or Copy lemma]
- (optional) symmetry conditions

Objective function:

minimize $\left[\max_i \frac{\mathrm{H}(\text{secret share}_i)}{\mathrm{H}(\text{secret})} \right]$

[Farràs-Kaced-Martín-Padró 2018] and [Gürpınar-R.]

PostModern approach

Write a **linear program** as follows

Constraints:

- equations from the definition of a **perfect secret sharing**
- all **Shannon-type** inequalities $I(* : * \,|\, *) \geq 0$
- ~~some **known non-Shannon-type** inequalities~~
- oversimplified technical explanation:
  make **clones** of (S_0, S_1, S_6, S_7) conditional on (S_2, S_3, S_4, S_5) (twice!)
- (optional) symmetry conditions

Objective function:

minimize $\left[\max_i \mathrm{H}(\text{secret share}_i)\right]$

Answer: **information ratio** $\geq 561/491 \approx 1.142566$

Modern approach vs. PostModern approach

Modern approach:

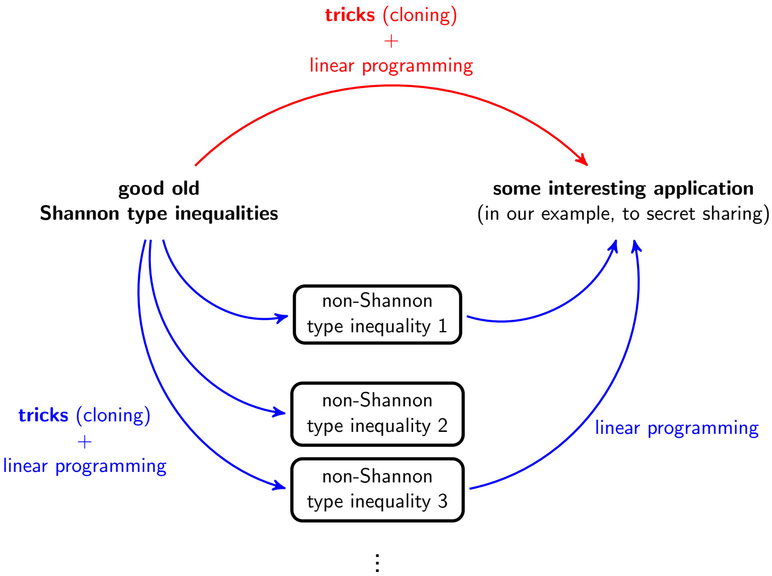
Stage 1: computer-aided search of non-Shannon type inequalities
[**materializing info (Ahlswede-Körner) or cloning (Copy Lemma)**
+ linear programming]

Stage 2: computer-aided linear programming for secret sharing involving inequalities found on **Stage 1**

PostModern approach:

One Shot: computer-aided linear programming for a secret sharing problem involving **cloning**

In one picture: postmodern vs. modern approaches



Once again: not in this talk

Once again: not in this talk

- sharp lower bounds in secret sharing

Once again: not in this talk

- sharp lower bounds in secret sharing: **Carles Padró**

Once again: not in this talk

- sharp lower bounds in secret sharing: **Carles Padró**
- use of symmetries in the entropy space

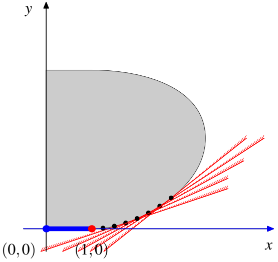
Once again: not in this talk

- sharp lower bounds in secret sharing: **Carles Padró**
- use of symmetries in the entropy space: **Qi Chen**

Outline

Applications (2)

the cone of almost entropic points is not polyhedral



one essentially conditional inequality for *(almost) entropic points* (≥ 4 r.v.)

$\Downarrow$

infinitely many unconditional information inequalities (in $\mathbb{R}^{15}$)

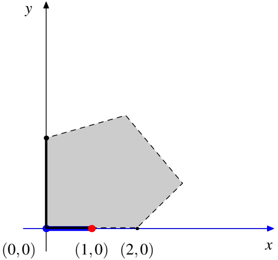
$\Downarrow$

Theorem (Matúš)

There exist infinitely many independent linear information inequalities.

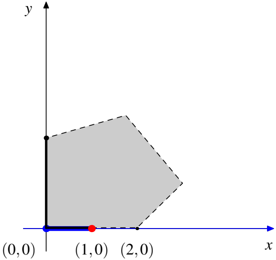
Outline

Applications of essentially conditional inequalities for strictly entropic points: combinatorics (work in progress)



if $y = 0$ then $x \leq 1$ $\Longleftarrow$ from a complex structure of the borderline

Applications of essentially conditional inequalities for strictly entropic points: combinatorics (work in progress)



if $y = 0$ then $x \leq 1 \Longleftarrow$ from a complex structure of the borderline

What is it all about?

What is it all about? What is it for?

What is it all about? What is it for?

- not in this talk: conditional independence properties

What is it all about? What is it for?

- not in this talk: conditional independence properties,
  talk by **Milan Studený**

What is it all about? What is it for?

- not in this talk: conditional independence properties,
  talk by **Milan Studený**

- this talk: combinatorial applications

Theorem

$$\mathbf{H}(\mathsf{a}|\mathsf{x},\mathsf{y}) = \mathbf{I}(\mathsf{x}:\mathsf{y}|\mathsf{a}) = \mathbf{0} \Rightarrow \mathbf{I}(\mathsf{a}:\mathsf{b}) \leq \mathbf{I}(\mathsf{a}:\mathsf{b}\,|\,\mathsf{x}) + \mathbf{I}(\mathsf{a}:\mathsf{b}\,|\,\mathsf{y}) + \mathbf{I}(\mathsf{x}:\mathsf{y})$$

Theorem

$$\mathbf{H}(\mathsf{a}|\mathsf{x},\mathsf{y}) = \mathbf{I}(\mathsf{x}:\mathsf{y}|\mathsf{a}) = \mathbf{0} \Rightarrow \mathbf{I}(\mathsf{a}:\mathsf{b}) \le \mathbf{I}(\mathsf{a}:\mathsf{b}\,|\,\mathsf{x}) + \mathbf{I}(\mathsf{a}:\mathsf{b}\,|\,\mathsf{y}) + \mathbf{I}(\mathsf{x}:\mathsf{y})$$

What is the intuition behind it?

Theorem

$$\mathbf{H}(\mathsf{a}|\mathsf{x},\mathsf{y}) = \mathbf{I}(\mathsf{x}:\mathsf{y}|\mathsf{a}) = \mathbf{0} \Rightarrow \mathbf{I}(\mathsf{a}:\mathsf{b}) \le \mathbf{I}(\mathsf{a}:\mathsf{b}\,|\,\mathsf{x}) + \mathbf{I}(\mathsf{a}:\mathsf{b}\,|\,\mathsf{y}) + \mathbf{I}(\mathsf{x}:\mathsf{y})$$

We relax the constraint and make the statement stronger:

Theorem

$$\mathbf{H}(\mathsf{a}|\mathsf{x},\mathsf{y}) = \mathbf{I}(\mathsf{x}:\mathsf{y}|\mathsf{a}) = \mathbf{0} \Rightarrow \mathbf{I}(\mathsf{a}:\mathsf{b}) \leq \mathbf{I}(\mathsf{a}:\mathsf{b}\,|\,\mathsf{x}) + \mathbf{I}(\mathsf{a}:\mathsf{b}\,|\,\mathsf{y}) + \mathbf{I}(\mathsf{x}:\mathsf{y})$$

We relax the constraint and make the statement stronger:

(*) $\forall i, j$ there is at most one k s.t. $(\Pr[X_i \;\&\; A_k] > 0$ and $\Pr[Y_j \;\&\; A_k] > 0)$

Theorem

$$\mathbf{H(a|x,y) = I(x:y|a) = 0 \Rightarrow I(a:b) \leq I(a:b\,|\,x) + I(a:b\,|\,y) + I(x:y)}$$

We relax the constraint and make the statement stronger:

(*) $\forall i, j$ there is at most one k s.t. $(\Pr[X_i \ \&\ A_k] > 0$ and $\Pr[Y_j \ \&\ A_k] > 0)$

Observation [Kaced, R., Vereshchagin]:

$$\mathbf{H(a|x,y) = I(x:y|a) = 0} \implies (*)$$

Theorem

$$\mathbf{H}(\mathsf{a}|\mathsf{x},\mathsf{y}) = \mathbf{I}(\mathsf{x}:\mathsf{y}|\mathsf{a}) = \mathbf{0} \Rightarrow \mathbf{I}(\mathsf{a}:\mathsf{b}) \leq \mathbf{I}(\mathsf{a}:\mathsf{b}\,|\,\mathsf{x}) + \mathbf{I}(\mathsf{a}:\mathsf{b}\,|\,\mathsf{y}) + \mathbf{I}(\mathsf{x}:\mathsf{y})$$

We relax the constraint and make the statement stronger:

(*) $\forall i, j$ there is at most one k s.t. $(\Pr[X_i \ \& \ A_k] > 0$ and $\Pr[Y_j \ \& \ A_k] > 0)$

Observation [Kaced, R., Vereshchagin]:

$$\mathbf{H}(\mathsf{a}|\mathsf{x},\mathsf{y}) = \mathbf{I}(\mathsf{x}:\mathsf{y}|\mathsf{a}) = \mathbf{0} \implies (*) \implies \mathrm{H}(a\,|\,x,b) + \mathrm{H}(a\,|\,y,b) \leq \mathrm{H}(a\,|\,b)$$

Theorem

$\mathbf{H}(\mathbf{a}|\mathbf{x},\mathbf{y}) = \mathbf{I}(\mathbf{x}:\mathbf{y}|\mathbf{a}) = \mathbf{0} \Rightarrow \mathbf{I}(\mathbf{a}:\mathbf{b}) \le \mathbf{I}(\mathbf{a}:\mathbf{b}\,|\,\mathbf{x}) + \mathbf{I}(\mathbf{a}:\mathbf{b}\,|\,\mathbf{y}) + \mathbf{I}(\mathbf{x}:\mathbf{y})$

We relax the constraint and make the statement stronger:

(*) $\forall i, j$ there is at most one k s.t. $(\Pr[X_i \,\&\, A_k] > 0$ and $\Pr[Y_j \,\&\, A_k] > 0)$

Observation [Kaced, R., Vereshchagin]:

$$\mathbf{H}(\mathbf{a}|\mathbf{x},\mathbf{y}) = \mathbf{I}(\mathbf{x}:\mathbf{y}|\mathbf{a}) = \mathbf{0} \implies (*) \implies \mathrm{H}(a\,|\,x,b) + \mathrm{H}(a\,|\,y,b) \le \mathrm{H}(a\,|\,b)$$

$$\Downarrow$$

$$\mathbf{I}(\mathbf{a}:\mathbf{b}) \le \mathbf{I}(\mathbf{a}:\mathbf{b}\,|\,\mathbf{x}) + \mathbf{I}(\mathbf{a}:\mathbf{b}\,|\,\mathbf{y}) + \mathbf{I}(\mathbf{x}:\mathbf{y})$$

(*) $\forall X_i, Y_j$ there is at most one A_k s.t. $(\Pr[X_i \mathbin{\&} A_k] > 0$ and $\Pr[Y_j \mathbin{\&} A_k] > 0)$

Theorem

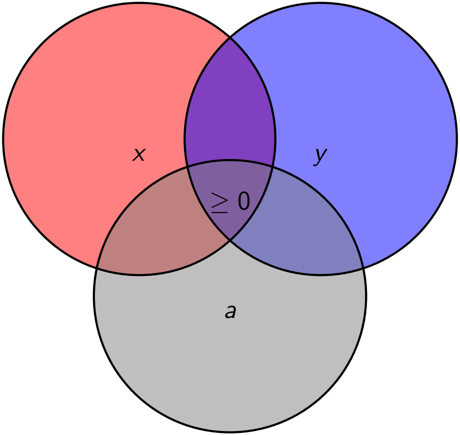
$(*) \implies \mathrm{H}(a \mid x) + \mathrm{H}(a \mid y) \leq \mathrm{H}(a)$

(*) $\forall X_i, Y_j$ there is at most one A_k s.t. $(\Pr[X_i \,\&\, A_k] > 0$ and $\Pr[Y_j \,\&\, A_k] > 0)$

Theorem

$(*) \implies \mathrm{H}(a \,|\, x) + \mathrm{H}(a \,|\, y) \leq \mathrm{H}(a)$

Equivalent form: $(*) \implies \mathrm{I}(a : x : y) \geq 0$



It is about graph coloring!

G: a bi-partite graph with colored edges

It is about graph coloring!

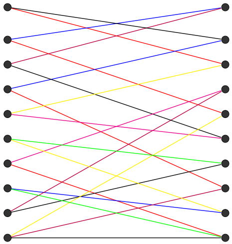
G: a bi-partite graph with colored edges

(*) for any vertices $v \in \mathrm{Left}$ and $w \in \mathrm{Right}$ there exist ≤ 1 **common color**

It is about graph coloring!

G: a bi-partite graph with colored edges

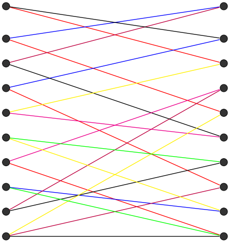
(*) for any vertices $v \in \mathrm{Left}$ and $w \in \mathrm{Right}$ there exist ≤ 1 **common color**



It is about graph coloring!

G: a bi-partite graph with colored edges

(*) for any vertices $v \in \mathrm{Left}$ and $w \in \mathrm{Right}$ there exist ≤ 1 **common color**

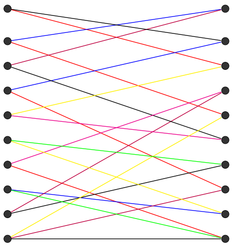


Take a random edge

It is about graph coloring!

G: a bi-partite graph with colored edges

(*) for any vertices $v \in \mathrm{Left}$ and $w \in \mathrm{Right}$ there exist ≤ 1 **common color**



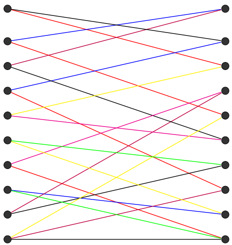
Take a random edge

- $x :=$ the left end of the edge
- $y :=$ the right end of the edge
- $a :=$ the color of the edge

It is about graph coloring!

G: a bi-partite graph with colored edges

(*) for any vertices $v \in \mathrm{Left}$ and $w \in \mathrm{Right}$ there exist ≤ 1 **common color**



Take a random edge

- x := the left end of the edge
- y := the right end of the edge
- a := the color of the edge

Theorem $(*) \implies \mathrm{H}(a\,|\,x) + \mathrm{H}(a\,|\,y) \leq \mathrm{H}(a)$

A toy application: a bound for an edge coloring

G: a bi-partite graph

A toy application: a bound for an edge coloring

G: a bi-partite graph = a union of M *matchings*

A toy application: a bound for an edge coloring

G: a bi-partite graph = a union of M *matchings*

- every vertex $v \in \mathrm{Left}$ is involved in $\geq A$ matchings

A toy application: a bound for an edge coloring

G: a bi-partite graph = a union of M *matchings*

- every vertex $v \in \mathrm{Left}$ is involved in $\geq A$ matchings
- every vertex $w \in \mathrm{Right}$ is involved in $\geq B$ matchings

A toy application: a bound for an edge coloring

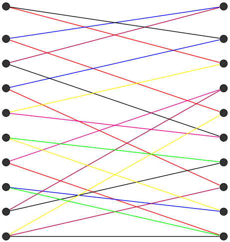
G: a bi-partite graph = a union of M *matchings*

- every vertex $v \in \mathrm{Left}$ is involved in $\geq A$ matchings
- every vertex $w \in \mathrm{Right}$ is involved in $\geq B$ matchings
- any $v \in \mathrm{Left}$ and $w \in \mathrm{Right}$ are involved in ≤ 1 **common** matching

A toy application: a bound for an edge coloring

G: a bi-partite graph = a union of M *matchings*

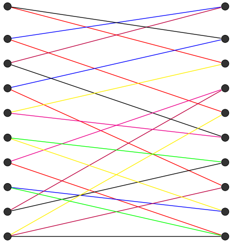
- every vertex $v \in \mathrm{Left}$ is involved in $\geq A$ matchings
- every vertex $w \in \mathrm{Right}$ is involved in $\geq B$ matchings
- any $v \in \mathrm{Left}$ and $w \in \mathrm{Right}$ are involved in ≤ 1 **common** matching



A toy application: a bound for an edge coloring

G: a bi-partite graph = a union of M *matchings*

- every vertex $v \in \mathrm{Left}$ is involved in $\geq A$ matchings
- every vertex $w \in \mathrm{Right}$ is involved in $\geq B$ matchings
- any $v \in \mathrm{Left}$ and $w \in \mathrm{Right}$ are involved in ≤ 1 **common** matching

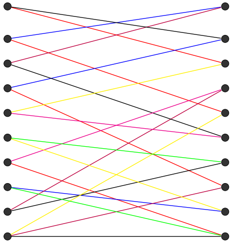


Prove that $M \geq A \cdot B$

A toy application: a bound for an edge coloring

G: a bi-partite graph = a union of M *matchings*

- every vertex $v \in \mathrm{Left}$ is involved in $\geq A$ matchings
- every vertex $w \in \mathrm{Right}$ is involved in $\geq B$ matchings
- any $v \in \mathrm{Left}$ and $w \in \mathrm{Right}$ are involved in ≤ 1 **common** matching



Prove that $M \geq A \cdot B$

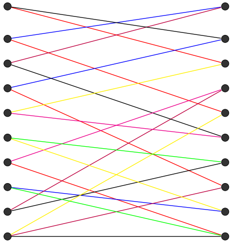
Proof: take a random **edge**,

- $x :=$ the left end of this **edge**
- $y :=$ the right end of this **edge**
- $a :=$ the color of this **edge**

A toy application: a bound for an edge coloring

G: a bi-partite graph = a union of M *matchings*

- every vertex $v \in \mathrm{Left}$ is involved in $\geq A$ matchings
- every vertex $w \in \mathrm{Right}$ is involved in $\geq B$ matchings
- any $v \in \mathrm{Left}$ and $w \in \mathrm{Right}$ are involved in ≤ 1 **common** matching



Prove that $M \geq A \cdot B$

Proof: take a random **edge**,

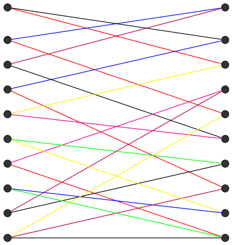
- $x :=$ the left end of this **edge**
- $y :=$ the right end of this **edge**
- $a :=$ the color of this **edge**

Then $\mathrm{H}(a\,|\,x) \;+\; \mathrm{H}(a\,|\,y) \;\leq\; \mathrm{H}(a)$

A toy application: a bound for an edge coloring

G: a bi-partite graph = a union of M *matchings*

- every vertex $v \in \mathrm{Left}$ is involved in $\geq A$ matchings
- every vertex $w \in \mathrm{Right}$ is involved in $\geq B$ matchings
- any $v \in \mathrm{Left}$ and $w \in \mathrm{Right}$ are involved in ≤ 1 **common** matching



Prove that $M \geq A \cdot B$

Proof: take a random **edge**,

- $x :=$ the left end of this **edge**
- $y :=$ the right end of this **edge**
- $a :=$ the color of this **edge**

Then $\underbrace{\mathrm{H}(a\,|\,x)}_{\geq \log A} + \underbrace{\mathrm{H}(a\,|\,y)}_{\geq \log B} \leq \underbrace{\mathrm{H}(a)}_{\leq \log M}$

Slightly different view: Secret key agreement

Slightly different view: Secret key agreement

Alice: knows a random **x**

Slightly different view: Secret key agreement

Alice: knows a random **x**
Bob: knows a random **y**

Slightly different view: Secret key agreement

Alice: knows a random **x**
Bob: knows a random **y**

x and y are correlated

Slightly different view: Secret key agreement

Alice: knows a random $\mathbf{x}$
Bob: knows a random $\mathbf{y}$
$\mathbf{x}$ and $\mathbf{y}$ are correlated

Alice and **Bob**:

- communicate via a public channel
- may use public and private randomness

Slightly different view: Secret key agreement

Alice: knows a random **x**
Bob: knows a random **y**

x and y are correlated

Alice and **Bob**:

- communicate via a public channel
- may use public and private randomness

Aim: construct a common **z** such that
$\mathrm{H}(\mathbf{z} \mid \text{communication transcript}) \approx |\mathbf{z}|$

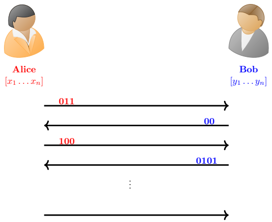
Slightly different view: Secret key agreement

Alice: knows a random **x**
Bob: knows a random **y**
x and y are correlated

Alice and **Bob**:

- communicate via a public channel
- may use public and private randomness

Aim: construct a common **z** such that
$\mathrm{H}(\mathbf{z} \mid \text{communication transcript}) \approx |\mathbf{z}|$



Slightly different view: Secret key agreement

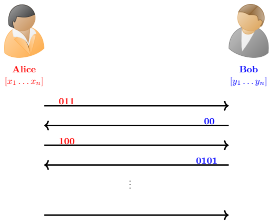
Alice: knows a random **x**
Bob: knows a random **y**
x and y are correlated

Alice and **Bob**:

- communicate via a public channel
- may use public and private randomness

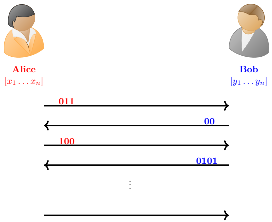
Aim: construct a common **z** such that
$\mathrm{H}(\mathbf{z} \mid \text{communication transcript}) \approx |\mathbf{z}|$

Question: How large can be entropy of **z**?



Slightly different view: Secret key agreement

Alice: knows a random **x**
Bob: knows a random **y**
x and y are correlated



Alice and **Bob**:

- communicate via a public channel
- may use public and private randomness

Aim: construct a common **z** such that
$\mathrm{H}(\mathbf{z} \mid \text{communication transcript}) \approx |\mathbf{z}|$

Question: How large can be entropy of **z**?

Theorem (see Ahlswede–Csiszár, Maurer 93)

1. *There is a protocol that produces a secret key z of size* $\approx I(x : y)$ *w.h.p.*
2. *No protocol can do better.*

Conditional inequality: it is about communication protocols!

Theorem

For all communication protocols $\mathrm{H}(\textbf{key} \mid \textbf{transcript}) \leq \mathbf{I}(\mathbf{x} : \mathbf{y})$

Conditional inequality: it is about communication protocols!

Theorem

For all communication protocols $\mathrm{H}(\textbf{key} \mid \textbf{transcript}) \le \mathbf{I}(\mathbf{x} : \mathbf{y})$

Simple observation: if no communication, then $\mathrm{H}(\textbf{key}) \le \mathrm{I}(\mathbf{x} : \mathbf{y})$

Indeed:

$$\begin{array}{rcccccc} \mathrm{H}(\textbf{key}) & \le & \mathrm{H}(\textbf{key} \mid x) & + & \mathrm{H}(\textbf{key} \mid y) & + & \mathrm{I}(x : y) \\ & = & 0 & + & 0 & + & \mathrm{I}(x : y) \end{array}$$

Conditional inequality: it is about communication protocols!

Theorem

For all communication protocols $\mathrm{H}(\textbf{key} \mid \textbf{transcript}) \leq \mathbf{I}(\mathbf{x} : \mathbf{y})$

Simple observation: if no communication, then $\mathrm{H}(\textbf{key}) \leq \mathrm{I}(\mathbf{x} : \mathbf{y})$

Indeed:

$$\begin{array}{rcccccc} \mathrm{H}(\textbf{key}) & \leq & \mathrm{H}(\textbf{key} \mid x) & + & \mathrm{H}(\textbf{key} \mid y) & + & \mathrm{I}(x : y) \\ & = & 0 & + & 0 & + & \mathrm{I}(x : y) \end{array}$$

Still simple: with a communication,

$$\mathrm{H}(\textbf{key}) \leq \mathbf{I}(\mathbf{x} : \mathbf{y} \mid \textbf{transcript})$$

Conditional inequality: it is about communication protocols!

Theorem

For all communication protocols $\mathrm{H}(\mathbf{key} \mid \mathbf{transcript}) \leq \mathbf{I}(\mathbf{x} : \mathbf{y})$

Simple observation: if no communication, then $\mathrm{H}(\mathbf{key}) \leq \mathrm{I}(\mathbf{x} : \mathbf{y})$

Indeed:

$$\begin{array}{rcccccc} \mathrm{H}(\mathbf{key}) & \leq & \mathrm{H}(\mathbf{key} \mid x) & + & \mathrm{H}(\mathbf{key} \mid y) & + & \mathrm{I}(x : y) \\ & = & 0 & + & 0 & + & \mathrm{I}(x : y) \end{array}$$

Still simple: with a communication,

$$\mathrm{H}(\mathbf{key}) \leq \mathbf{I}(\mathbf{x} : \mathbf{y} \mid \mathbf{transcript})$$

Hard part: $\mathbf{I}(\mathbf{x} : \mathbf{y} \mid \mathbf{transcript}) \leq \mathbf{I}(\mathbf{x} : \mathbf{y})$

Conditional inequality: it is about communication protocols!

Theorem

For all communication protocols $\mathrm{H}(\mathbf{key} \mid \mathbf{transcript}) \leq \mathbf{I}(\mathbf{x} : \mathbf{y})$

Simple observation: if no communication, then $\mathrm{H}(\mathbf{key}) \leq \mathrm{I}(\mathbf{x} : \mathbf{y})$

Indeed:

$$\begin{array}{rcccccc} \mathrm{H}(\mathbf{key}) & \leq & \mathrm{H}(\mathbf{key} \mid x) & + & \mathrm{H}(\mathbf{key} \mid y) & + & \mathrm{I}(x : y) \\ & = & 0 & + & 0 & + & \mathrm{I}(x : y) \end{array}$$

Still simple: with a communication,

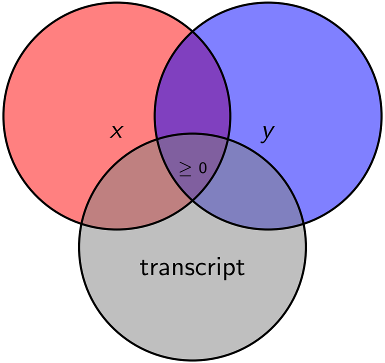
$$\mathrm{H}(\mathbf{key}) \leq \mathbf{I}(\mathbf{x} : \mathbf{y} \mid \mathbf{transcript})$$

Hard part: $\mathbf{I}(\mathbf{x} : \mathbf{y} \mid \mathbf{transcript}) \leq \mathbf{I}(\mathbf{x} : \mathbf{y})$

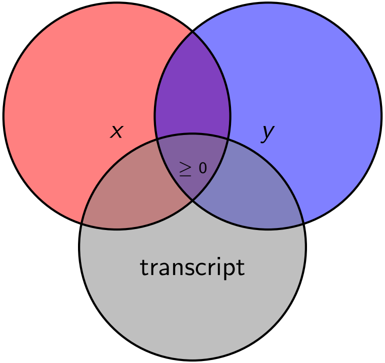
Equivalent form: $\mathbf{I}(\mathbf{x} : \mathbf{y} : \mathbf{transcript}) \geq \mathbf{0}$,

which is true for all **communication transcripts**

the core of the proof:

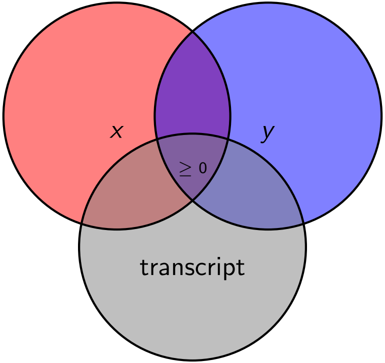


the core of the proof:



external information complexity $\geq$ internal information complexity

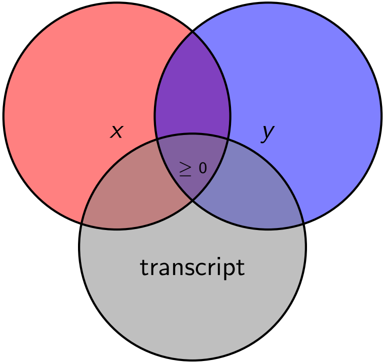
the core of the proof:



external information complexity $\geq$ internal information complexity

advantages of this approach (conditional information inequality) :

the core of the proof:

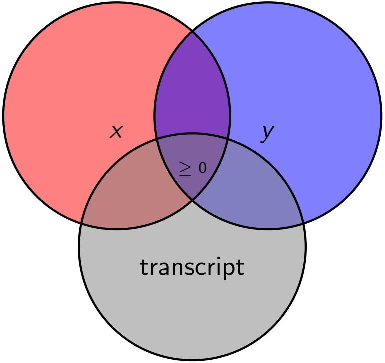


external information complexity $\geq$ internal information complexity

advantages of this approach (conditional information inequality) :

- applies to a light version of non-determinism (*bi-clique cover*)

the core of the proof:



external information complexity $\geq$ internal information complexity

advantages of this approach (conditional information inequality) :

- applies to a light version of non-determinism (*bi-clique cover*)
- translation to the setting of Kolmogorov complexity [R.-Zimand]