On the Intersection and Composition properties for discrete random variables

Tobias Boege

Department of Mathematics and Statistics
UiT The Arctic University of Norway

Seminar on Algorithmic Aspects of Information Theory
5 March 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

Conditional independence

- N fixed finite ground set indexing jointly distributed random variables.
- Identify $i \in N$ with $\{i\} \subseteq N$ and for $I, J \subseteq N$ abbreviate $IJ = I \cup J$.

Conditional independence

- N fixed finite ground set indexing jointly distributed random variables.
- Identify $i \in N$ with $\{i\} \subseteq N$ and for $I, J \subseteq N$ abbreviate $IJ = I \cup J$.
- Conditional independence for $I, J, K \subseteq N$ disjoint:

$$[I \perp\!\!\!\perp J \mid K] \iff \text{vanishing of conditional mutual information.}$$

- The CI symbols are symmetric $[I \perp\!\!\!\perp J \mid K] \iff [J \perp\!\!\!\perp I \mid K]$.

Conditional independence

- N fixed finite ground set indexing jointly distributed random variables.
- Identify $i \in N$ with $\{i\} \subseteq N$ and for $I, J \subseteq N$ abbreviate $IJ = I \cup J$.
- Conditional independence for $I, J, K \subseteq N$ disjoint:

$$[I \perp\!\!\!\perp J \mid K] \iff \text{vanishing of conditional mutual information.}$$

- The CI symbols are symmetric $[I \perp\!\!\!\perp J \mid K] \iff [J \perp\!\!\!\perp I \mid K]$.
- A set S of CI symbols is a semigraphoid if it satisfies

$$\begin{aligned}[I \perp\!\!\!\perp JK \mid L] &\iff [I \perp\!\!\!\perp J \mid L] \wedge [I \perp\!\!\!\perp K \mid JL] \\ &\iff [I \perp\!\!\!\perp K \mid L] \wedge [I \perp\!\!\!\perp J \mid KL]\end{aligned}$$

- E.g., conditional independence relation of every system of random variables.

Partial converses of the semigraphoid property

$$[I \perp\!\!\!\perp JK \mid L] \implies \begin{cases} ①[I \perp\!\!\!\perp J \mid L] \wedge ②[I \perp\!\!\!\perp J \mid KL] \wedge \\ ③[I \perp\!\!\!\perp K \mid L] \wedge ④[I \perp\!\!\!\perp K \mid JL] \end{cases}$$

Partial converses of the semigraphoid property

$$[I \perp\!\!\!\perp JK \mid L] \implies \begin{cases} ①[I \perp\!\!\!\perp J \mid L] \wedge ②[I \perp\!\!\!\perp J \mid KL] \wedge \\ ③[I \perp\!\!\!\perp K \mid L] \wedge ④[I \perp\!\!\!\perp K \mid JL] \end{cases}$$

- ① $\wedge$ ④ and ② $\wedge$ ③ are sufficient for $[I \perp\!\!\!\perp JK \mid L]$ by semigraphoid axioms.

Partial converses of the semigraphoid property

$$[I \perp\!\!\!\perp JK \mid L] \implies \begin{cases} ①[I \perp\!\!\!\perp J \mid L] \wedge ②[I \perp\!\!\!\perp J \mid KL] \wedge \\ ③[I \perp\!\!\!\perp K \mid L] \wedge ④[I \perp\!\!\!\perp K \mid JL] \end{cases}$$

- ① $\wedge$ ④ and ② $\wedge$ ③ are sufficient for $[I \perp\!\!\!\perp JK \mid L]$ by semigraphoid axioms.
- Intersection property: ② $\wedge$ ④ $\implies [I \perp\!\!\!\perp JK \mid L]$.

Partial converses of the semigraphoid property

$$[I \perp\!\!\!\perp JK \mid L] \implies \begin{cases} ①[I \perp\!\!\!\perp J \mid L] \;\wedge\; ②[I \perp\!\!\!\perp J \mid KL] \;\wedge \\ ③[I \perp\!\!\!\perp K \mid L] \;\wedge\; ④[I \perp\!\!\!\perp K \mid JL] \end{cases}$$

- ① ∧ ④ and ② ∧ ③ are sufficient for $[I \perp\!\!\!\perp JK \mid L]$ by semigraphoid axioms.
- Intersection property: ② ∧ ④ $\implies [I \perp\!\!\!\perp JK \mid L]$.
- Composition property: ① ∧ ③ $\implies [I \perp\!\!\!\perp JK \mid L]$.

Partial converses of the semigraphoid property

$$[I \perp\!\!\!\perp JK \mid L] \implies \begin{cases} ①[I \perp\!\!\!\perp J \mid L] \wedge ②[I \perp\!\!\!\perp J \mid KL] \wedge \\ ③[I \perp\!\!\!\perp K \mid L] \wedge ④[I \perp\!\!\!\perp K \mid JL] \end{cases}$$

- ① ∧ ④ and ② ∧ ③ are sufficient for $[I \perp\!\!\!\perp JK \mid L]$ by semigraphoid axioms.
- Intersection property: ② ∧ ④ $\implies [I \perp\!\!\!\perp JK \mid L]$.
- Composition property: ① ∧ ③ $\implies [I \perp\!\!\!\perp JK \mid L]$.
- Let's ignore ① ∧ ② and ③ ∧ ④ today . . .

Partial converses of the semigraphoid property

$$[I \perp\!\!\!\perp JK \mid L] \implies \begin{cases} ①[I \perp\!\!\!\perp J \mid L] \wedge ②[I \perp\!\!\!\perp J \mid KL] \wedge \\ ③[I \perp\!\!\!\perp K \mid L] \wedge ④[I \perp\!\!\!\perp K \mid JL] \end{cases}$$

- ① ∧ ④ and ② ∧ ③ are sufficient for $[I \perp\!\!\!\perp JK \mid L]$ by semigraphoid axioms.
- Intersection property: ② ∧ ④ $\implies [I \perp\!\!\!\perp JK \mid L]$.
- Composition property: ① ∧ ③ $\implies [I \perp\!\!\!\perp JK \mid L]$.
- Let's ignore ① ∧ ② and ③ ∧ ④ today . . .

Modulo the semigraphoid axioms Intersection and Composition are logical converses:

Intersection	② ∧ ④	$\implies$	① ∧ ③
Composition	② ∧ ④	$\impliedby$	① ∧ ③

Intersection, Composition and duality

Goal: find sufficient conditions on the distribution which ensure Intersection or Composition.

Intersection, Composition and duality

Goal: find sufficient conditions on the distribution which ensure Intersection or Composition.

- Intersection has received lots of attention. Composition not so much.

Intersection, Composition and duality

Goal: find sufficient conditions on the distribution which ensure Intersection or Composition.

- Intersection has received lots of attention. Composition not so much.
- Recent interest in Composition comes from machine learning. [AAZ22]

Intersection, Composition and duality

Goal: find sufficient conditions on the distribution which ensure Intersection or Composition.

- Intersection has received lots of attention. Composition not so much.
- Recent interest in Composition comes from machine learning. [AAZ22]

- Curiously they are dual to each other via $[I \perp\!\!\!\perp J \mid K]^* := [I \perp\!\!\!\perp J \mid N \setminus IJK]$:

Intersection $\quad [I \perp\!\!\!\perp J \mid KL] \wedge [I \perp\!\!\!\perp K \mid JL] \implies [I \perp\!\!\!\perp J \mid L] \wedge [I \perp\!\!\!\perp K \mid L]$

Intersection, Composition and duality

Goal: find sufficient conditions on the distribution which ensure Intersection or Composition.

- Intersection has received lots of attention. Composition not so much.
- Recent interest in Composition comes from machine learning. [AAZ22]
- Curiously they are dual to each other via $[I \perp\!\!\!\perp J \mid K]^* := [I \perp\!\!\!\perp J \mid N \setminus IJK]$:

$$\text{Intersection} \qquad [I \perp\!\!\!\perp J \mid KL] \wedge [I \perp\!\!\!\perp K \mid JL] \implies [I \perp\!\!\!\perp J \mid L] \wedge [I \perp\!\!\!\perp K \mid L]$$

$$\downarrow \qquad \text{use } \bar{L} = N \setminus IJKL$$

Intersection, Composition and duality

Goal: find sufficient conditions on the distribution which ensure Intersection or Composition.

- Intersection has received lots of attention. Composition not so much.
- Recent interest in Composition comes from machine learning. [AAZ22]
- Curiously they are dual to each other via $[I \perp\!\!\!\perp J \mid K]^* := [I \perp\!\!\!\perp J \mid N \setminus IJK]$:

$$\begin{array}{lc} \text{Intersection} & [I \perp\!\!\!\perp J \mid KL] \wedge [I \perp\!\!\!\perp K \mid JL] \implies [I \perp\!\!\!\perp J \mid L] \wedge [I \perp\!\!\!\perp K \mid L] \\ \downarrow & \text{use } \bar{L} = N \setminus IJKL \\ \text{Intersection}^* & [I \perp\!\!\!\perp J \mid \bar{L}] \wedge [I \perp\!\!\!\perp K \mid \bar{L}] \implies [I \perp\!\!\!\perp J \mid K\bar{L}] \wedge [I \perp\!\!\!\perp K \mid J\bar{L}] \end{array}$$

but this is Composition with L replaced by $\bar{L}$.

Examples

- The conditional independence structures of jointly regular Gaussian random variables satisfy Intersection and Composition.

Studený's question [Stu05, p. 191]

Is every conditional independence relation on regular Gaussians also realizable by discrete (or even binary) random variables?

Examples

- The conditional independence structures of jointly regular Gaussian random variables satisfy Intersection and Composition.

Studený's question [Stu05, p. 191]

Is every conditional independence relation on regular Gaussians also realizable by discrete (or even binary) random variables?

- Various types of graphical models satisfy Intersection and Composition. Proofs are combinatorial or reduce to properties of Gaussians.
  - d-separation, u-separation, m-separation, $*$-separation, ...

Examples

- The conditional independence structures of jointly regular Gaussian random variables satisfy Intersection and Composition.

Studený's question [Stu05, p. 191]

Is every conditional independence relation on regular Gaussians also realizable by discrete (or even binary) random variables?

- Various types of graphical models satisfy Intersection and Composition. Proofs are combinatorial or reduce to properties of Gaussians.
  - d-separation, u-separation, m-separation, $*$-separation, ...
- Positive distributions satisfy Intersection.

Examples

- The conditional independence structures of jointly regular Gaussian random variables satisfy Intersection and Composition.

Studený's question [Stu05, p. 191]

Is every conditional independence relation on regular Gaussians also realizable by discrete (or even binary) random variables?

- Various types of graphical models satisfy Intersection and Composition. Proofs are combinatorial or reduce to properties of Gaussians.
  - d-separation, u-separation, m-separation, $*$-separation, ...
- Positive distributions satisfy Intersection.
- MTP_2 distributions satisfy Composition.

Non-example: matroids

- Let $r\colon 2^N \to \mathbb{Z}$ be a matroid. The set of modular pairs of r is a semigraphoid:

$$\mathscr{S}(r) := \{\, [I \perp\!\!\!\perp J \mid K] : r(IK) + r(JK) = r(IJK) + r(K) \,\}.$$

Non-example: matroids

- Let $r\colon 2^N \to \mathbb{Z}$ be a matroid. The set of modular pairs of r is a semigraphoid:

$$\mathscr{S}(r) := \{\, [I \perp\!\!\!\perp J \mid K] : r(IK) + r(JK) = r(IJK) + r(K) \,\}.$$

Lemma

If S satisfies Composition and $[i \perp\!\!\!\perp j]$ for all $i \neq j$ then S is totally independent.

- *If r is a simple matroid then $\mathscr{S}(r)$ satisfies Composition if and only if r is uniform.*

Non-example: matroids

- Let $r\colon 2^N \to \mathbb{Z}$ be a matroid. The set of modular pairs of r is a semigraphoid:

$$\mathscr{S}(r) := \{\, [I \perp\!\!\!\perp J \mid K] : r(IK) + r(JK) = r(IJK) + r(K) \,\}.$$

Lemma

If S satisfies Composition and $[i \perp\!\!\!\perp j]$ for all $i \neq j$ then S is totally independent.

- *If r is a simple matroid then $\mathscr{S}(r)$ satisfies Composition if and only if r is uniform.*

Lemma*

If S satisfies Intersection and $[i \perp\!\!\!\perp j \mid N \setminus ij]$ for all $i \neq j$ then S is totally independent.

- *If r is a co-simple matroid then $\mathscr{S}(r)$ satisfies Intersection if and only if r is zero.*

Intersection for three binary random variables

$$[I \perp\!\!\!\perp J \mid KL] \wedge [I \perp\!\!\!\perp K \mid JL] \implies [I \perp\!\!\!\perp J \mid L] \wedge [I \perp\!\!\!\perp K \mid L]$$

- By marginalizing to $IJKL$, conditioning on L and viewing I, J, K as single random variables, we can reduce one instance of Intersection to the trivariate case.

Intersection for three binary random variables

$$[I \perp\!\!\!\perp J \mid KL] \wedge [I \perp\!\!\!\perp K \mid JL] \implies [I \perp\!\!\!\perp J \mid L] \wedge [I \perp\!\!\!\perp K \mid L]$$

- By marginalizing to $IJKL$, conditioning on L and viewing I, J, K as single random variables, we can reduce one instance of Intersection to the trivariate case.

```
needsPackage "GraphicalModels";
R = markovRing(3:2);
I = conditionalIndependenceIdeal(R, {{{1},{2},{3}}, {{1},{3},{2}}});
J = conditionalIndependenceIdeal(R, {{{1},{2,3},{}}});
decompose(I:J)
```

Intersection for three binary random variables

$$[I \perp\!\!\!\perp J \mid KL] \wedge [I \perp\!\!\!\perp K \mid JL] \implies [I \perp\!\!\!\perp J \mid L] \wedge [I \perp\!\!\!\perp K \mid L]$$

- By marginalizing to $IJKL$, conditioning on L and viewing I, J, K as single random variables, we can reduce one instance of Intersection to the trivariate case.

```
needsPackage "GraphicalModels";
R = markovRing(3:2);
I = conditionalIndependenceIdeal(R, {{{1},{2},{3}}, {{1},{3},{2}}});
J = conditionalIndependenceIdeal(R, {{{1},{2,3},{}}});
decompose(I:J)
```

$$\langle p_{110}, p_{101}, p_{010}, p_{001} \rangle \cap \langle p_{111}, p_{100}, p_{011}, p_{000} \rangle$$

- Failure of Intersection only on the boundary. Full support implies Intersection.

The characteristic bipartite graph

- Let i, j, k be discrete random variables taking r_i, r_j, r_k states, respectively.
- The characteristic bipartite graph $G(j,k)$ is the bipartite graph on $[r_j] \sqcup [r_k]$ with an edge $y - z$ whenever $\Pr[j = y, k = z] > 0$.

The characteristic bipartite graph

- Let i, j, k be discrete random variables taking r_i, r_j, r_k states, respectively.
- The characteristic bipartite graph $G(j,k)$ is the bipartite graph on $[r_j] \sqcup [r_k]$ with an edge $y - z$ whenever $\Pr[j = y, k = z] > 0$.
- If $[i \perp\!\!\!\perp j \mid k]$ and $[i \perp\!\!\!\perp k \mid j]$ hold, then $G(j,k)$ is admissible: all its connected components are $K_{p,q}$ for $p, q \geq 1$.

The characteristic bipartite graph

- Let i, j, k be discrete random variables taking r_i, r_j, r_k states, respectively.
- The characteristic bipartite graph $G(j, k)$ is the bipartite graph on $[r_j] \sqcup [r_k]$ with an edge $y - z$ whenever $\Pr[j = y, k = z] > 0$.
- If $[i \perp\!\!\!\perp j \mid k]$ and $[i \perp\!\!\!\perp k \mid j]$ hold, then $G(j, k)$ is admissible: all its connected components are $K_{p,q}$ for $p, q \geq 1$.

Theorem (Cartwright–Engström conjecture & Fink's theorem [Fin11])

The conditional independence model $\mathscr{M}([i \perp\!\!\!\perp j \mid k] \wedge [i \perp\!\!\!\perp k \mid j])$ decomposes into irreducible components, one for each admissible bipartite graph on $[r_j] \sqcup [r_k]$. Only the component corresponding to K_{r_j, r_k} is fully contained in $\mathscr{M}([i \perp\!\!\!\perp jk])$.

- Hence, $G(j, k)$ being connected is sufficient for (one instance of) Intersection.

The common information criterion

- $G(j,k)$ was already used by Gács and Körner [GK73]:

The common information criterion

- $G(j,k)$ was already used by Gács and Körner [GK73]:
  - Let $\mathscr{C}$ be the set of connected components of $G(j,k)$ and let g be a function of jk with values in $\mathscr{C}$ such that $g(y,z) = C$ whenever $y, z \in C$.

The common information criterion

- $G(j, k)$ was already used by Gács and Körner [GK73]:
  - Let $\mathscr{C}$ be the set of connected components of $G(j, k)$ and let g be a function of jk with values in $\mathscr{C}$ such that $g(y, z) = C$ whenever $y, z \in C$.
  - g realizes the Gács–Körner common information of j and k.

The common information criterion

- $G(j,k)$ was already used by Gács and Körner [GK73]:
  - Let $\mathscr{C}$ be the set of connected components of $G(j,k)$ and let g be a function of jk with values in $\mathscr{C}$ such that $g(y,z) = C$ whenever $y, z \in C$.
  - g realizes the Gács–Körner common information of j and k.

Theorem

If $[i \perp\!\!\!\perp j \mid k]$ and $[i \perp\!\!\!\perp k \mid j]$, then $[i \perp\!\!\!\perp jk \mid g]$. Hence, if $G(j,k)$ is connected, g is constant and therefore $[i \perp\!\!\!\perp jk]$ is implied.

The common information criterion

- $G(j,k)$ was already used by Gács and Körner [GK73]:
  - Let $\mathscr{C}$ be the set of connected components of $G(j,k)$ and let g be a function of jk with values in $\mathscr{C}$ such that $g(y,z) = C$ whenever $y, z \in C$.
  - g realizes the Gács–Körner common information of j and k.

Theorem

If $[i \perp\!\!\!\perp j \mid k]$ and $[i \perp\!\!\!\perp k \mid j]$, then $[i \perp\!\!\!\perp jk \mid g]$. Hence, if $G(j,k)$ is connected, g is constant and therefore $[i \perp\!\!\!\perp jk]$ is implied.

- The same criterion has been reached independently by Florens, Mouchart, Rolin and San Martín as well as Dawid in the language of σ-algebras and measurable separability. See the historical account in [MMR05].

The common information criterion

- $G(j,k)$ was already used by Gács and Körner [GK73]:
  - Let $\mathscr{C}$ be the set of connected components of $G(j,k)$ and let g be a function of jk with values in $\mathscr{C}$ such that $g(y,z) = C$ whenever $y, z \in C$.
  - g realizes the Gács–Körner common information of j and k.

Theorem

If $[i \perp\!\!\!\perp j \mid k]$ and $[i \perp\!\!\!\perp k \mid j]$, then $[i \perp\!\!\!\perp jk \mid g]$. Hence, if $G(j,k)$ is connected, g is constant and therefore $[i \perp\!\!\!\perp jk]$ is implied.

- The same criterion has been reached independently by Florens, Mouchart, Rolin and San Martín as well as Dawid in the language of σ-algebras and measurable separability. See the historical account in [MMR05].
- Also found by followers of Cartwright and Engström [KRS19].

The common information criterion

- $G(j, k)$ was already used by Gács and Körner [GK73]:
  - Let $\mathscr{C}$ be the set of connected components of $G(j, k)$ and let g be a function of jk with values in $\mathscr{C}$ such that $g(y, z) = C$ whenever $y, z \in C$.
  - g realizes the Gács–Körner common information of j and k.

Theorem

If $[i \perp\!\!\!\perp j \mid k]$ and $[i \perp\!\!\!\perp k \mid j]$, then $[i \perp\!\!\!\perp jk \mid g]$. Hence, if $G(j, k)$ is connected, g is constant and therefore $[i \perp\!\!\!\perp jk]$ is implied.

- The same criterion has been reached independently by Florens, Mouchart, Rolin and San Martín as well as Dawid in the language of σ-algebras and measurable separability. See the historical account in [MMR05].
- Also found by followers of Cartwright and Engström [KRS19].
- Also known as the Double Markov property [CK11, Exercise 16.25].

The conditional Ingleton criterion

- The common information criterion involves an auxiliary variable g.

The conditional Ingleton criterion

- The common information criterion involves an auxiliary variable g.
- All relations among four discrete random variables in terms of conditional independence follow from conditional Ingleton inequalities [Stu21].

The conditional Ingleton criterion

- The common information criterion involves an auxiliary variable g.
- All relations among four discrete random variables in terms of conditional independence follow from conditional Ingleton inequalities [Stu21].

Theorem

The following is an essentially conditional information inequality:

$$[i \perp\!\!\!\perp j \mid k] \wedge [i \perp\!\!\!\perp k \mid j] \implies \mathtt{Ingl}(i : g \mid j : k) \geq 0.$$

The conditional Ingleton criterion

- The common information criterion involves an auxiliary variable g.
- All relations among four discrete random variables in terms of conditional independence follow from conditional Ingleton inequalities [Stu21].

Theorem

The following is an essentially conditional information inequality:

$$[i \perp\!\!\!\perp j \mid k] \wedge [i \perp\!\!\!\perp k \mid j] \implies \mathtt{Ingl}(i : g \mid j : k) \geq 0.$$

It implies the conditional independence rule

$$[i \perp\!\!\!\perp j \mid k] \wedge [i \perp\!\!\!\perp k \mid j] \wedge [j \perp\!\!\!\perp k \mid g] \wedge [i \perp\!\!\!\perp g] \implies [i \perp\!\!\!\perp jk].$$

The conditional Ingleton criterion

- The common information criterion involves an auxiliary variable g.
- All relations among four discrete random variables in terms of conditional independence follow from conditional Ingleton inequalities [Stu21].

Theorem

The following is an essentially conditional information inequality:

$$[i \perp\!\!\!\perp j \mid k] \wedge [i \perp\!\!\!\perp k \mid j] \implies \mathtt{Ingl}(i : g \mid j : k) \geq 0.$$

It implies the conditional independence rule

$$[i \perp\!\!\!\perp j \mid k] \wedge [i \perp\!\!\!\perp k \mid j] \wedge [j \perp\!\!\!\perp k \mid g] \wedge [i \perp\!\!\!\perp g] \implies [i \perp\!\!\!\perp jk].$$

- What is necessary to construct such a g à la Gács–Körner?

Conditional Ingleton vs. Gács–Körner

It is not difficult to parametrize binary distributions which satisfy the conditional Ingleton criterion but fail the common information criterion using Cylindrical Algebraic Decomposition in `Mathematica`, e.g., having $G(j,k) = \{0{-}1, 1{-}0\}$.

i	j	k	g	Pr
0	0	1	1	$1/4$
0	1	0	0	$1/4$
1	0	1	1	$1/4$
1	1	0	0	$1/4$

Conditional Ingleton vs. Gács–Körner

It is not difficult to parametrize binary distributions which satisfy the conditional Ingleton criterion but fail the common information criterion using Cylindrical Algebraic Decomposition in `Mathematica`, e.g., having $G(j,k) = \{0{-}1, 1{-}0\}$.

i	j	k	g	Pr
0	0	1	1	$1/4$
0	1	0	0	$1/4$
1	0	1	1	$1/4$
1	1	0	0	$1/4$

- Note the functional dependencies $g = k = 1 - j$.

Conditional Ingleton vs. Gács–Körner

It is not difficult to parametrize binary distributions which satisfy the conditional Ingleton criterion but fail the common information criterion using Cylindrical Algebraic Decomposition in `Mathematica`, e.g., having $G(j,k) = \{0{-}1, 1{-}0\}$.

i	j	k	g	Pr
0	0	1	1	$1/4$
0	1	0	0	$1/4$
1	0	1	1	$1/4$
1	1	0	0	$1/4$

- Note the functional dependencies $g = k = 1 - j$.
- Gács–Körner common information is maximal with $H(G(j,k)) = \log 2$.

Conditional Ingleton vs. Gács–Körner

It is not difficult to parametrize binary distributions which satisfy the conditional Ingleton criterion but fail the common information criterion using Cylindrical Algebraic Decomposition in `Mathematica`, e.g., having $G(j,k) = \{0{-}1, 1{-}0\}$.

i	j	k	g	Pr
0	0	1	1	$1/4$
0	1	0	0	$1/4$
1	0	1	1	$1/4$
1	1	0	0	$1/4$

- Note the functional dependencies $g = k = 1 - j$.
- Gács–Körner common information is maximal with $H(G(j,k)) = \log 2$.
- Distribution on ijk is quasi-uniform and $[i \perp\!\!\!\perp jk]$ holds.

Composition for three binary random variables

$$[I \perp\!\!\!\perp J \mid L] \wedge [I \perp\!\!\!\perp K \mid L] \implies [I \perp\!\!\!\perp J \mid KL] \wedge [I \perp\!\!\!\perp K \mid JL]$$

Composition for three binary random variables

$$[I \perp\!\!\!\perp J \mid L] \wedge [I \perp\!\!\!\perp K \mid L] \implies [I \perp\!\!\!\perp J \mid KL] \wedge [I \perp\!\!\!\perp K \mid JL]$$

```
needsPackage "GraphicalModels";
R = markovRing(3:2);
I = conditionalIndependenceIdeal(R, {{{1},{2},{}}, {{1},{3},{}}});
J = ideal(sum gens R);
decompose(I:J)
```

Composition for three binary random variables

$$[I \perp\!\!\!\perp J \mid L] \wedge [I \perp\!\!\!\perp K \mid L] \implies [I \perp\!\!\!\perp J \mid KL] \wedge [I \perp\!\!\!\perp K \mid JL]$$

```
needsPackage "GraphicalModels";
R = markovRing(3:2);
I = conditionalIndependenceIdeal(R, {{{1},{2},{}}, {{1},{3},{}}});
J = ideal(sum gens R);
decompose(I:J)
```

Theorem (Kirkup's theorem [Kir07])

There is only one irreducible component of $\mathscr{M}([i \perp\!\!\!\perp j] \wedge [i \perp\!\!\!\perp k])$ on which the sum of all probabilities does not vanish.

Composition for three binary random variables

$$[I \perp\!\!\!\perp J \mid L] \wedge [I \perp\!\!\!\perp K \mid L] \implies [I \perp\!\!\!\perp J \mid KL] \wedge [I \perp\!\!\!\perp K \mid JL]$$

```
needsPackage "GraphicalModels";
R = markovRing(3:2);
I = conditionalIndependenceIdeal(R, {{{1},{2},{}}, {{1},{3},{}}});
J = ideal(sum gens R);
decompose(I:J)
```

Theorem (Kirkup's theorem [Kir07])

There is only one irreducible component of $\mathscr{M}([i \perp\!\!\!\perp j] \wedge [i \perp\!\!\!\perp k])$ on which the sum of all probabilities does not vanish.

- No graphs, no interesting boundary structure.

Composition for three binary random variables

$$[I \perp\!\!\!\perp J \mid L] \wedge [I \perp\!\!\!\perp K \mid L] \implies [I \perp\!\!\!\perp J \mid KL] \wedge [I \perp\!\!\!\perp K \mid JL]$$

```
needsPackage "GraphicalModels";
R = markovRing(3:2);
I = conditionalIndependenceIdeal(R, {{{1},{2},{}}, {{1},{3},{}}});
J = ideal(sum gens R);
decompose(I:J)
```

Theorem (Kirkup's theorem [Kir07])

There is only one irreducible component of $\mathscr{M}([i \perp\!\!\!\perp j] \wedge [i \perp\!\!\!\perp k])$ on which the sum of all probabilities does not vanish.

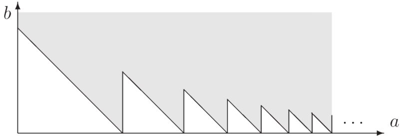
- No graphs, no interesting boundary structure.
- There exist positive distributions violating Composition.

Matúš's criterion

- A distribution on N is tight if each $i \in N$ functionally depends on $N \setminus i$.

Theorem ([Mat06])

The tight entropy profiles with $[i \perp\!\!\!\perp j]$ and $[i \perp\!\!\!\perp k]$ are described by a piecewise linear information inequality $\rightarrow$

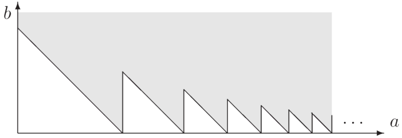


Matúš's criterion

- A distribution on N is tight if each $i \in N$ functionally depends on $N \setminus i$.

Theorem ([Mat06])

The tight entropy profiles with $[i \perp\!\!\!\perp j]$ and $[i \perp\!\!\!\perp k]$ are described by a piecewise linear information inequality $\to$



- If h is the entropy profile of ijk, tight and satisfies $[i \perp\!\!\!\perp j]$ and $[i \perp\!\!\!\perp k]$, then

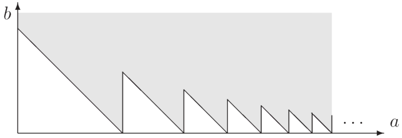
$$h(ijk) = h(ij) = h(ik) = h(jk) = h(i) + h(j) = h(i) + h(k).$$

Matúš's criterion

- A distribution on N is tight if each $i \in N$ functionally depends on $N \setminus i$.

Theorem ([Mat06])

The tight entropy profiles with $[i \perp\!\!\!\perp j]$ and $[i \perp\!\!\!\perp k]$ are described by a piecewise linear information inequality →



- If h is the entropy profile of ijk, tight and satisfies $[i \perp\!\!\!\perp j]$ and $[i \perp\!\!\!\perp k]$, then

$$h(ijk) = h(ij) = h(ik) = h(jk) = h(i) + h(j) = h(i) + h(k).$$

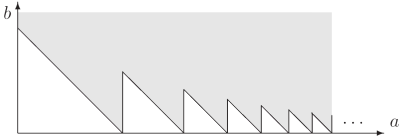
If also $[i \perp\!\!\!\perp jk]$ holds, then $h(ijk) = h(i) + h(jk)$, hence $h(i) = 0$ and $[j \perp\!\!\!\perp k]$.

Matúš's criterion

- A distribution on N is tight if each $i \in N$ functionally depends on $N \setminus i$.

Theorem ([Mat06])

The tight entropy profiles with $[i \perp\!\!\!\perp j]$ and $[i \perp\!\!\!\perp k]$ are described by a piecewise linear information inequality →



- If h is the entropy profile of ijk, tight and satisfies $[i \perp\!\!\!\perp j]$ and $[i \perp\!\!\!\perp k]$, then

$$h(ijk) = h(ij) = h(ik) = h(jk) = h(i) + h(j) = h(i) + h(k).$$

If also $[i \perp\!\!\!\perp jk]$ holds, then $h(ijk) = h(i) + h(jk)$, hence $h(i) = 0$ and $[j \perp\!\!\!\perp k]$.

- How to generalize away from the tightness constraints?

Dual conditional Ingleton criterion

Theorem

The following is an essentially conditional information inequality:

$$[i \perp\!\!\!\perp j \mid g] \wedge [i \perp\!\!\!\perp k \mid g] \implies \mathtt{Ingl}(j : k \mid i : g) \geq 0.$$

Dual conditional Ingleton criterion

Theorem

The following is an essentially conditional information inequality:

$$[i \perp\!\!\!\perp j \mid g] \wedge [i \perp\!\!\!\perp k \mid g] \implies \mathtt{Ingl}(j : k \mid i : g) \geq 0.$$

It implies the conditional independence rule

$$[i \perp\!\!\!\perp j \mid g] \wedge [i \perp\!\!\!\perp k \mid g] \wedge [j \perp\!\!\!\perp k \mid i] \wedge [i \perp\!\!\!\perp g \mid jk] \implies [i \perp\!\!\!\perp jk \mid g].$$

Dual conditional Ingleton criterion

Theorem

The following is an essentially conditional information inequality:

$$[i \perp\!\!\!\perp j \mid g] \wedge [i \perp\!\!\!\perp k \mid g] \implies \mathtt{Ingl}(j : k \mid i : g) \geq 0.$$

It implies the conditional independence rule

$$[i \perp\!\!\!\perp j \mid g] \wedge [i \perp\!\!\!\perp k \mid g] \wedge [j \perp\!\!\!\perp k \mid i] \wedge [i \perp\!\!\!\perp g \mid jk] \implies [i \perp\!\!\!\perp jk \mid g].$$

- This is formally dual to the conditional Ingleton criterion for Intersection.

Dual conditional Ingleton criterion

Theorem

The following is an essentially conditional information inequality:

$$[i \perp\!\!\!\perp j \mid g] \wedge [i \perp\!\!\!\perp k \mid g] \implies \mathtt{Ingl}(j : k \mid i : g) \geq 0.$$

It implies the conditional independence rule

$$[i \perp\!\!\!\perp j \mid g] \wedge [i \perp\!\!\!\perp k \mid g] \wedge [j \perp\!\!\!\perp k \mid i] \wedge [i \perp\!\!\!\perp g \mid jk] \implies [i \perp\!\!\!\perp jk \mid g].$$

- This is formally dual to the conditional Ingleton criterion for Intersection.
- The Composition property is obtained conditionally on g.

Dual conditional Ingleton criterion

Theorem

The following is an essentially conditional information inequality:

$$[i \perp\!\!\!\perp j \mid g] \wedge [i \perp\!\!\!\perp k \mid g] \implies \mathit{Ingl}(j : k \mid i : g) \geq 0.$$

It implies the conditional independence rule

$$[i \perp\!\!\!\perp j \mid g] \wedge [i \perp\!\!\!\perp k \mid g] \wedge [j \perp\!\!\!\perp k \mid i] \wedge [i \perp\!\!\!\perp g \mid jk] \implies [i \perp\!\!\!\perp jk \mid g].$$

- This is formally dual to the conditional Ingleton criterion for Intersection.
- The Composition property is obtained conditionally on g.
- How to use this? Any constructions of suitable g?

The duality paradox

According to Šimeček's database for four random variables [Šim06]:

- There are 1 098 probabilistically representable semigraphoids (up to permutation).

The duality paradox

According to Šimeček's database for four random variables [Šim06]:

- There are 1 098 probabilistically representable semigraphoids (up to permutation).
- 369 satisfy Intersection and 369 satisfy Composition (and 133 satisfy both).

The duality paradox

According to Šimeček's database for four random variables [Šim06]:

- There are 1 098 probabilistically representable semigraphoids (up to permutation).
- 369 satisfy Intersection and 369 satisfy Composition (and 133 satisfy both).
- The definitions of Intersection and Composition properties are dual.

The duality paradox

According to Šimeček's database for four random variables [Šim06]:

- There are 1 098 probabilistically representable semigraphoids (up to permutation).
- 369 satisfy Intersection and 369 satisfy Composition (and 133 satisfy both).
- The definitions of Intersection and Composition properties are dual.
- Conditional Ingleton inequalities provide dual sufficient conditions.

The duality paradox

According to Šimeček's database for four random variables [Šim06]:

- There are 1 098 probabilistically representable semigraphoids (up to permutation).
- 369 satisfy Intersection and 369 satisfy Composition (and 133 satisfy both).
- The definitions of Intersection and Composition properties are dual.
- Conditional Ingleton inequalities provide dual sufficient conditions.
- But the bijection between the two sets of size 369 is **not** duality:

$$\{\, [i \perp\!\!\!\perp g \mid j], [i \perp\!\!\!\perp g \mid k], [i \perp\!\!\!\perp g \mid jk], [j \perp\!\!\!\perp k \mid i], [j \perp\!\!\!\perp k \mid g] \,\} \checkmark \quad \text{[Stu21, Ex. 4]}$$

$$\{\, [i \perp\!\!\!\perp g \mid j], [i \perp\!\!\!\perp g \mid k], [i \perp\!\!\!\perp g \quad\;\;], [j \perp\!\!\!\perp k \mid i], [j \perp\!\!\!\perp k \mid g] \,\} ✗ \quad \text{[Stu21, (I:1)]}$$

The duality paradox

According to Šimeček's database for four random variables [Šim06]:

- There are 1 098 probabilistically representable semigraphoids (up to permutation).
- 369 satisfy Intersection and 369 satisfy Composition (and 133 satisfy both).
- The definitions of Intersection and Composition properties are dual.
- Conditional Ingleton inequalities provide dual sufficient conditions.
- But the bijection between the two sets of size 369 is **not** duality:

$$\{\, [i \perp\!\!\!\perp g \mid j], [i \perp\!\!\!\perp g \mid k], [i \perp\!\!\!\perp g \mid jk], [j \perp\!\!\!\perp k \mid i], [j \perp\!\!\!\perp k \mid g] \,\} \ \checkmark \quad \text{[Stu21, Ex. 4]}$$

$$\{\, [i \perp\!\!\!\perp g \mid j], [i \perp\!\!\!\perp g \mid k], [i \perp\!\!\!\perp g], [j \perp\!\!\!\perp k \mid i], [j \perp\!\!\!\perp k \mid g] \,\} \ \text{✗} \quad \text{[Stu21, (I:1)]}$$

- Is there a (geometric / algebraic) relation between Intersection and Composition on the level of distributions?

References I

[AAZ22] Arash A. Amini, Bryon Aragam, and Qing Zhou. *A non-graphical representation of conditional independence via the neighbourhood lattice*. 2022. arXiv: 2206.05829 [math.ST].

[CK11] Imre Csiszár and János Körner. *Information theory. Coding theorems for discrete memoryless systems*. 2nd ed. Cambridge University Press, 2011. DOI: 10.1017/CBO9780511921889.

[Fin11] Alex Fink. "The binomial ideal of the intersection axiom for conditional probabilities". In: *J. Algebr. Comb.* 33.3 (2011), pp. 455–463. DOI: 10.1007/s10801-010-0253-5.

[GK73] Peter Gács and János Körner. "Common information is far less than mutual information". In: *Probl. Control Inf. Theory* 2 (1973), pp. 149–162.

[KRS19] Thomas Kahle, Johannes Rauh, and Seth Sullivant. "Algebraic aspects of conditional independence and graphical models". In: *Handbook of graphical models*. Ed. by Marloes Maathuis, Mathias Drton, Steffen Lauritzen, and Martin Wainwright. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. CRC Press, 2019, pp. 61–80. ISBN: 978-1-4987-8862-5; 978-0-4298-7424-6.

References II

[Kir07] George A. Kirkup. "Random variables with completely independent subcollections". In: *J. Algebra* 309.2 (2007), pp. 427–454. DOI: 10.1016/j.jalgebra.2006.06.023.

[MMR05] Ernesto San Martín, Michel Mouchart, and Jean-Marie Rolin. "Ignorable common information, null sets and Basu's first theorem". In: *Sankhyā* 67.4 (2005), pp. 674–698.

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

[Šim06] Petr Šimeček. "A short note on discrete representability of independence models". In: *Proccedings of the European Workshop on Probabilistic Graphical Models*. 2006, pp. 287–292.

[Stu05] Milan Studený. *Probabilistic Conditional Independence Structures*. Information Science and Statistics. Springer, 2005.

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