Title: Causal Inference Lecture - 230320

Speakers: Robert Spekkens

Collection: Causal Inference: Classical and Quantum

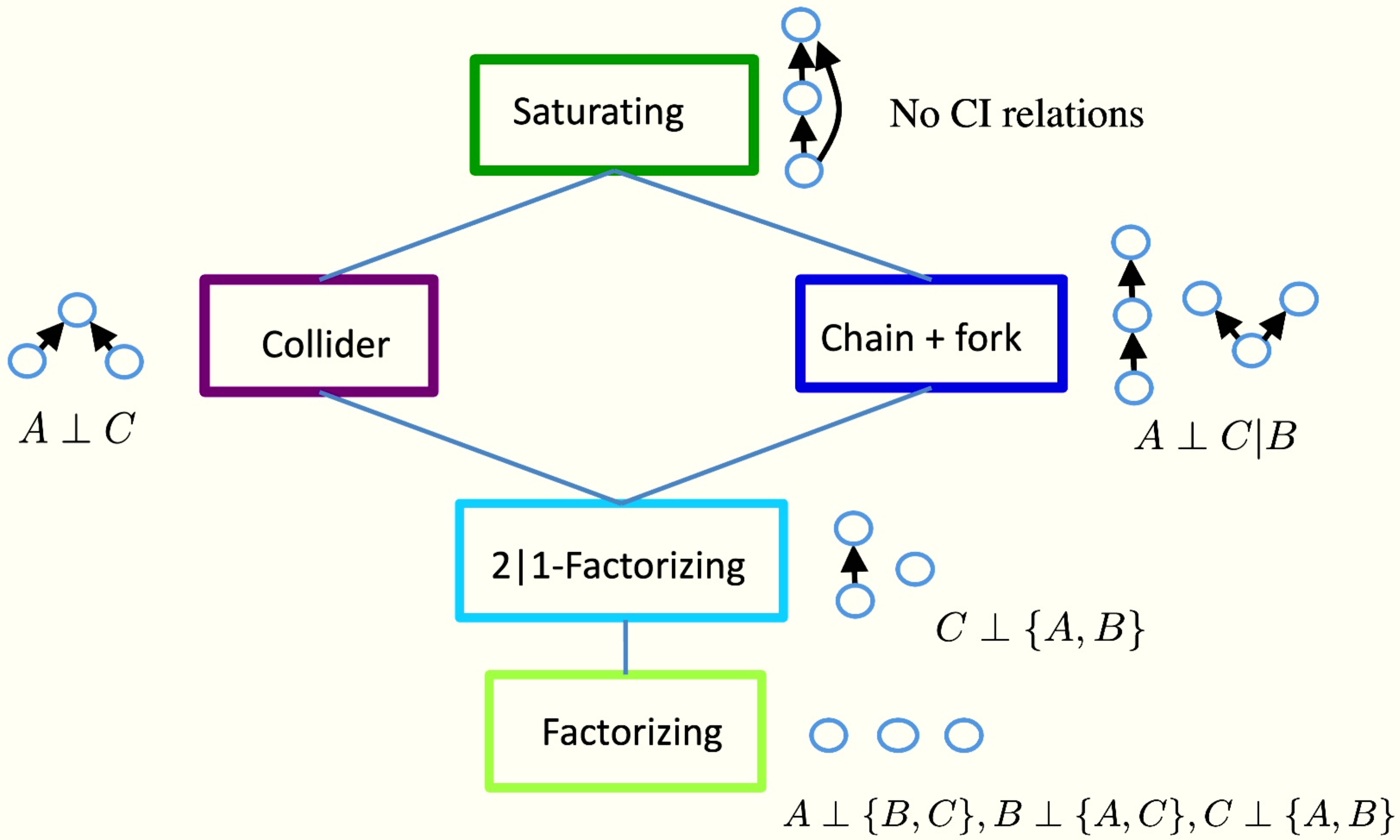
Date: March 20, 2023 - 10:00 AM

URL: https://pirsa.org/23030073

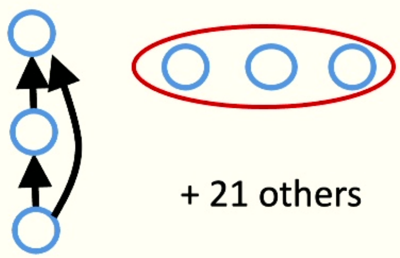
Abstract: zoom link: https://pitp.zoom.us/j/94143784665?pwd=VFJpajVIMEtvYmRabFYzYnNRSVAvZz09

Inequality constraints for latent-permitting causal models

Seeing that there must be inequality constraints

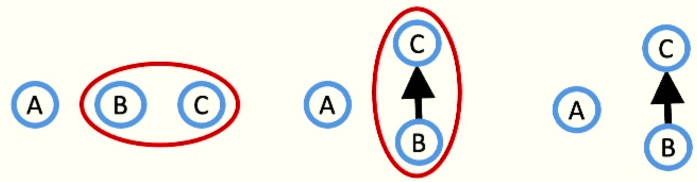

Saturating
No CI relations
Collider
$A \perp C$
Chain + fork
$A \perp C|B$
2|1-Factorizing
$C \perp \{A, B\}$
Factorizing
$A \perp \{B, C\}, B \perp \{A, C\}, C \perp \{A, B\}$

Saturating



+ 21 others

1|2-Factorizing

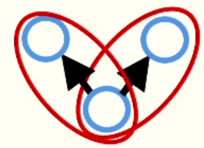


$$A \perp CB$$

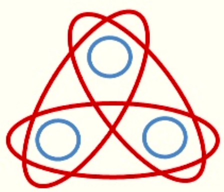
Factorizing

$$A \perp CB, B \perp AC, C \perp AB$$

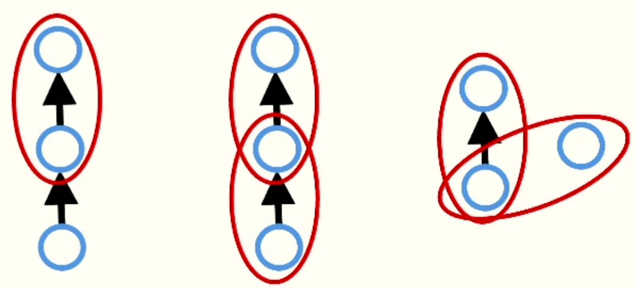
Evans

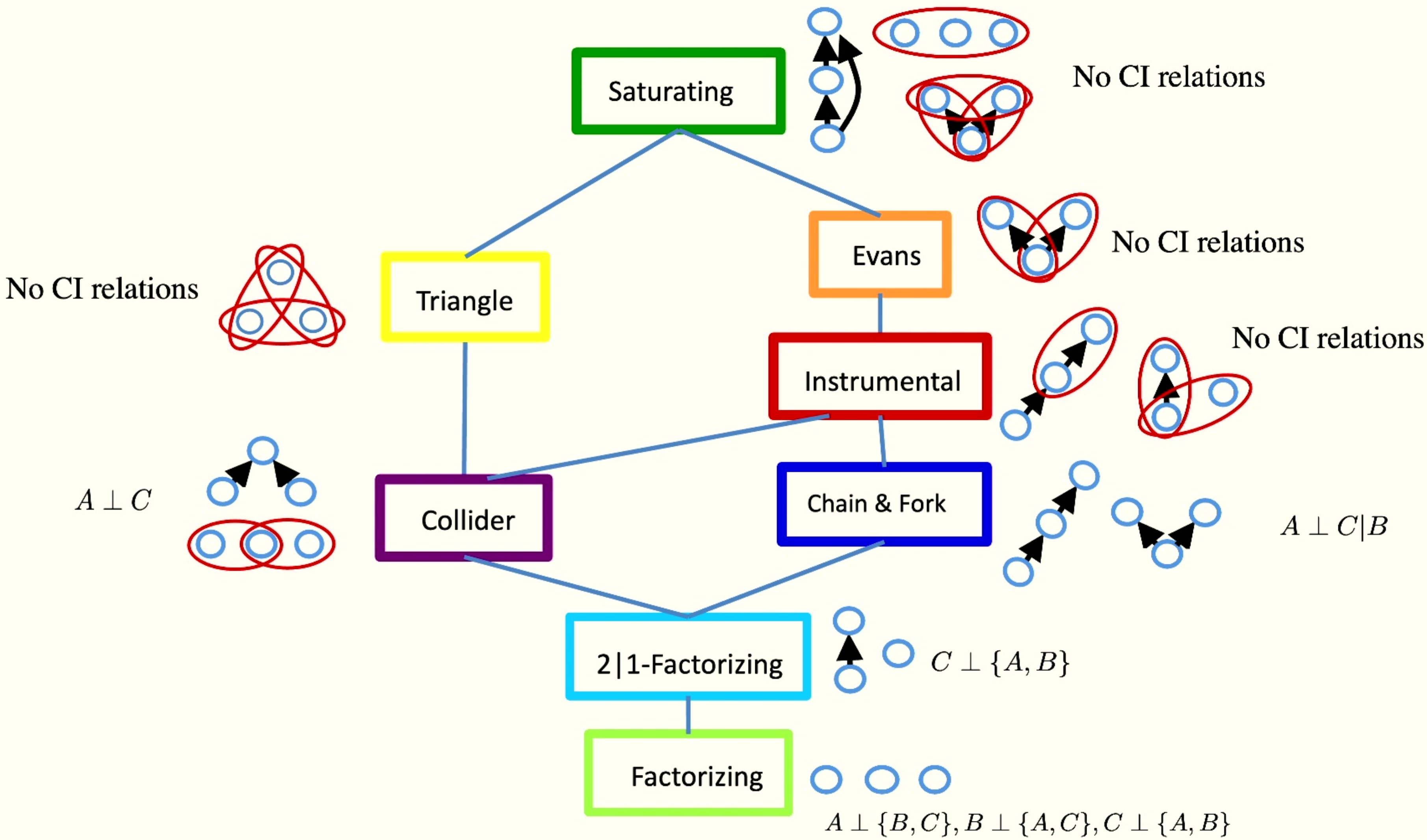


Triangle



Instrumental

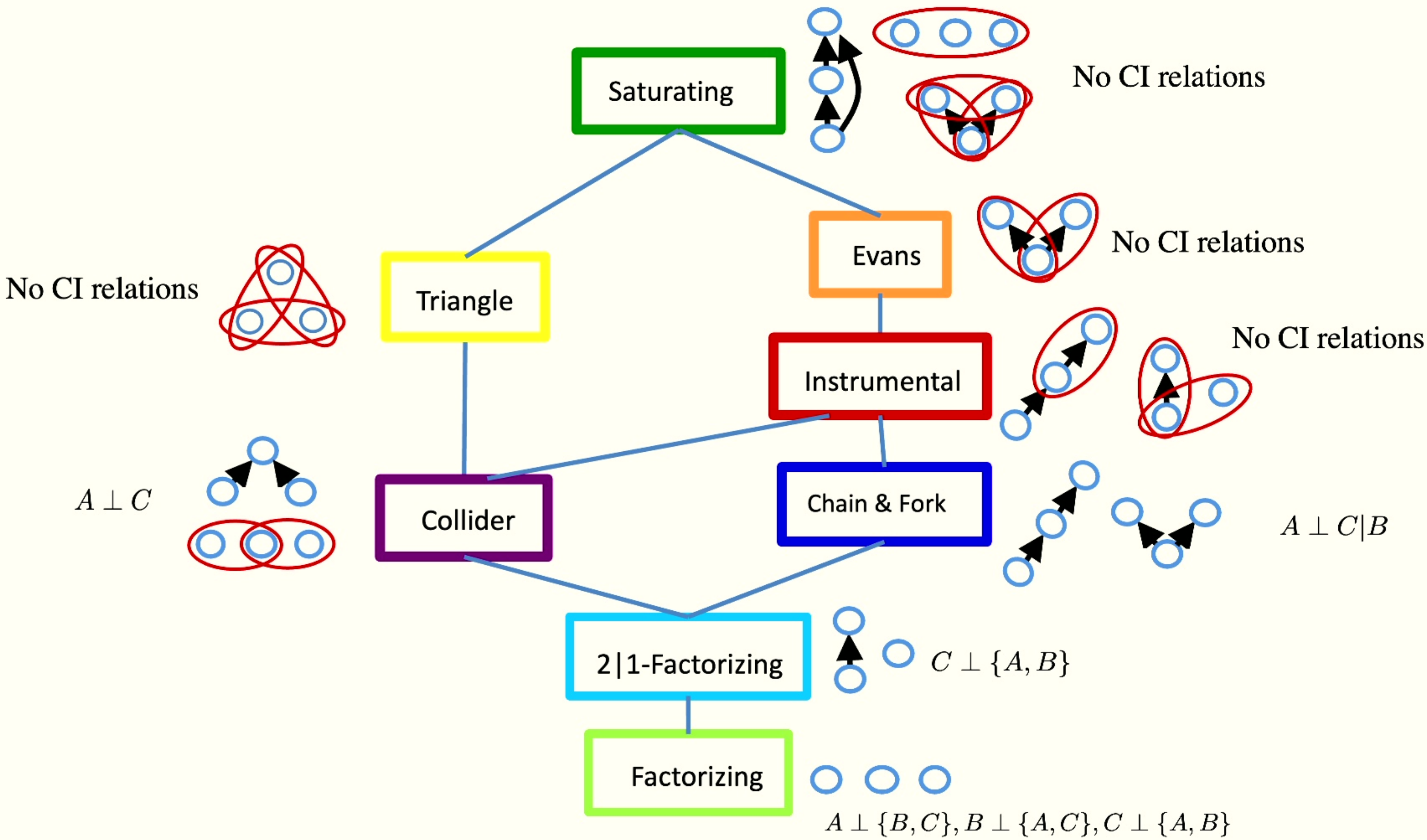




Saturating
No CI relations
Evans
No CI relations
No CI relations
Triangle
Instrumental
No CI relations
$A \perp C$
Collider
Chain & Fork
$A \perp C|B$
2|1-Factorizing
$C \perp \{A, B\}$
Factorizing
$A \perp \{B, C\}, B \perp \{A, C\}, C \perp \{A, B\}$

Depicted dominances are consistent with edge and hyperedge dropping, but one needs to show that they are not strict

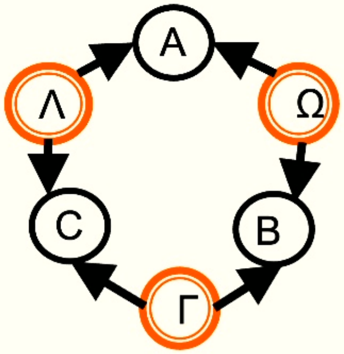
That is, one needs to show observational *inequivalences*


Saturating
No CI relations
Evans
No CI relations
No CI relations
Triangle
Instrumental
No CI relations
$A \perp C$
Collider
Chain & Fork
$A \perp C | B$
2|1-Factorizing
$C \perp \{A, B\}$
Factorizing
$A \perp \{B, C\}, B \perp \{A, C\}, C \perp \{A, B\}$

Note:
There are no nested Markov constraints beyond conditional independence relations at the level of 3 nodes

Therefore, the constraints that separate these observational equivalence classes must be inequalities

Some simple examples establishing the separations

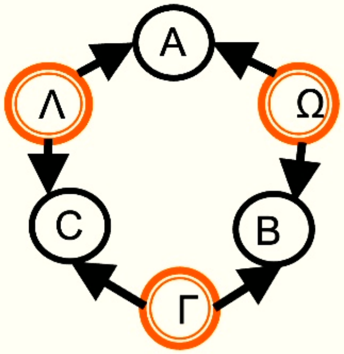


Triangle

Compatible set does not include:

$$P_{ABC}^{\text{corr}} = \frac{1}{2}[000] + \frac{1}{2}[111]$$

Steudel and Ay, 2012



Compatible set does not include:

$$P^{\text{corr}}_{ABC} = \tfrac{1}{2}[000] + \tfrac{1}{2}[111]$$

Proof: The marginals of the target state are:

$$P^{\text{corr}}_{AB} = P^{\text{corr}}_{AC} = P^{\text{corr}}_{BC} = \tfrac{1}{2}[00] + \tfrac{1}{2}[11]$$

$$P_{ABC} = \sum_{\Lambda\Omega\Gamma} \mathrm{P}_{A|\Lambda\Omega}\mathrm{P}_{B|\Omega\Gamma}\mathrm{P}_{C|\Gamma\Lambda}\mathrm{P}_{\Lambda}\mathrm{P}_{\Omega}\mathrm{P}_{\Gamma}$$

$$P_{AB} = \sum_{\Lambda\Omega\Gamma} \mathrm{P}_{A|\Lambda\Omega}\mathrm{P}_{B|\Omega\Gamma}\mathrm{P}_{\Lambda}\mathrm{P}_{\Omega}\mathrm{P}_{\Gamma}$$

Require:

$$\mathrm{P}_{A|\Lambda\Omega} = \delta_{A,\Omega}$$

$$\mathrm{P}_{B|\Gamma\Omega} = \delta_{B,\Omega}$$

$$P_{\Omega} = \tfrac{1}{2}[0] + \tfrac{1}{2}[1]$$

So that

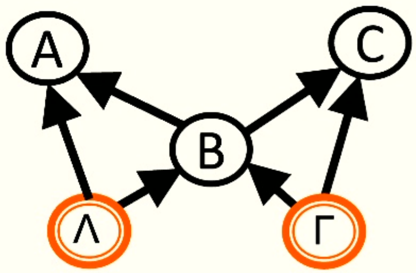
$$P_{AB} = \left(\sum_{\Omega} \delta_{A,\Omega}\delta_{B,\Omega}\mathrm{P}_{\Omega}\right)\sum_{\Lambda}\mathrm{P}_{\Lambda}\sum_{\Gamma}\mathrm{P}_{\Gamma}$$

$$= \tfrac{1}{2}[00] + \tfrac{1}{2}[11]$$

But then:

$$P_{AC} = \sum_{\Lambda\Omega\Gamma} \mathrm{P}_{A|\Lambda\Omega}\mathrm{P}_{C|\Lambda\Gamma}\mathrm{P}_{\Lambda}\mathrm{P}_{\Omega}\mathrm{P}_{\Gamma}$$

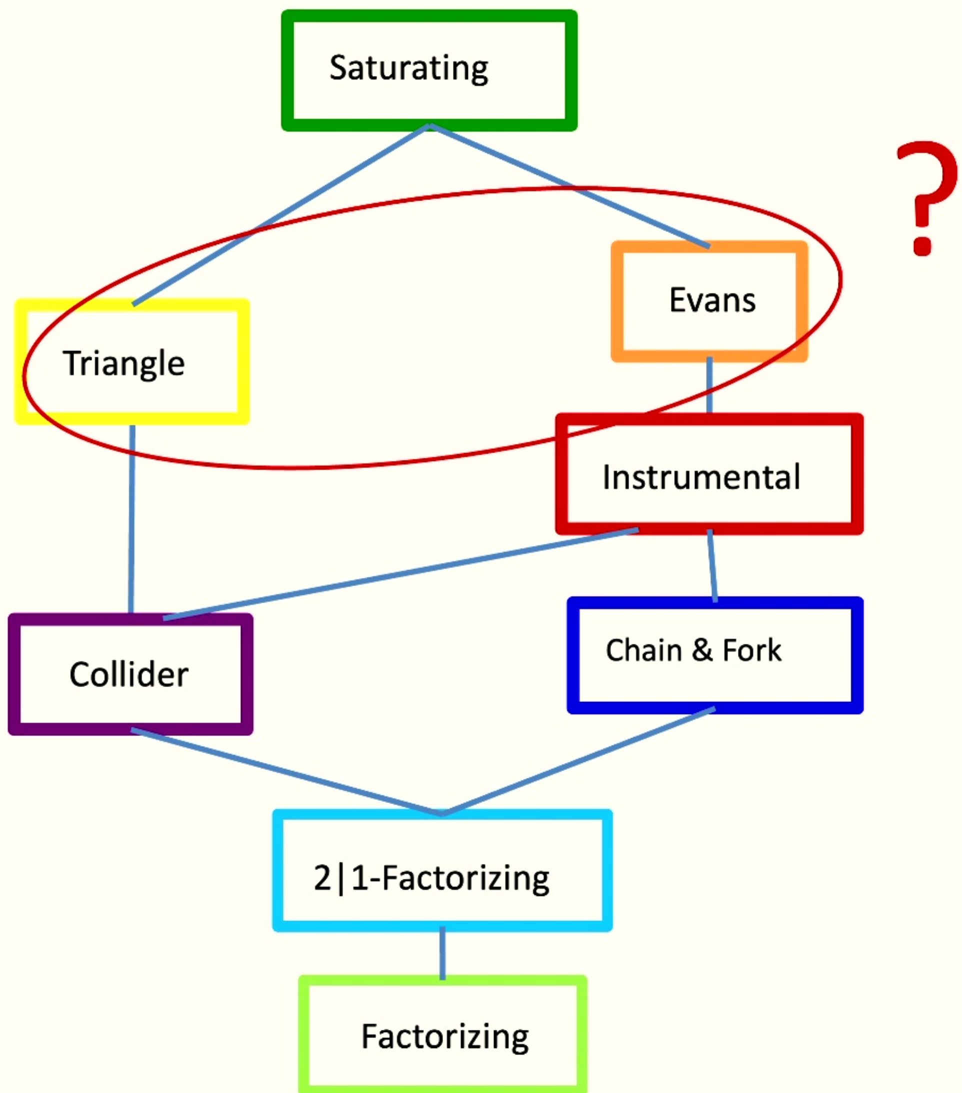
$$= \left(\sum_{\Omega} \delta_{A,\Omega}\mathrm{P}_{\Omega}\right)\left(\sum_{\Lambda\Gamma}\mathrm{P}_{C|\Lambda\Gamma}\mathrm{P}_{\Lambda}\mathrm{P}_{\Gamma}\right)$$

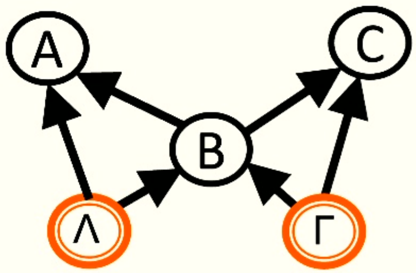


Evans

Compatible set **does not** include:

$$P_{ABC}^{\text{pinch}} = \left(\tfrac{1}{2}[00]_{AC} + \tfrac{1}{2}[11]_{AC}\right)[0]_B$$


Saturating
?
Evans
Triangle
Instrumental
Collider
Chain & Fork
2|1-Factorizing
Factorizing



Evans

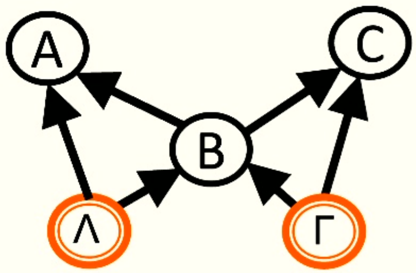
Compatible set **does not** include:

$$P_{ABC}^{\text{pinch}} = \left(\tfrac{1}{2}[00]_{AC} + \tfrac{1}{2}[11]_{AC}\right)[0]_B$$

Proof:

Require: $P_{B|\Lambda\Gamma} = [0]_B$

Which implies
$$P_{ABC} = \sum_{\Lambda\Gamma} \mathrm{P}_{A|\Lambda B}\mathrm{P}_{B|\Lambda\Gamma}\mathrm{P}_{C|\Gamma B}\mathrm{P}_{\Lambda}\mathrm{P}_{\Gamma}$$
$$= \left(\sum_{\Lambda} \mathrm{P}_{A|\Lambda B}\mathrm{P}_{\Lambda}\right)\left(\sum_{\Gamma} \mathrm{P}_{C|\Gamma B}\mathrm{P}_{\Gamma}\right)[0]_B$$



Evans

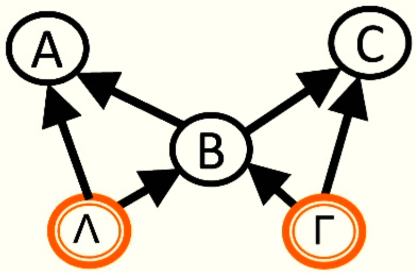
Compatible set **does not** include:

$$P^{\text{pinch}}_{ABC} = \left(\tfrac{1}{2}[00]_{AC} + \tfrac{1}{2}[11]_{AC}\right)[0]_B$$

Proof:

Require: $P_{B|\Lambda\Gamma} = [0]_B$

Which implies
$$P_{ABC} = \sum_{\Lambda\Gamma} \mathrm{P}_{A|\Lambda B}\mathrm{P}_{B|\Lambda\Gamma}\mathrm{P}_{C|\Gamma B}\mathrm{P}_{\Lambda}\mathrm{P}_{\Gamma}$$
$$= \left(\sum_{\Lambda} \mathrm{P}_{A|\Lambda B}\mathrm{P}_{\Lambda}\right)\left(\sum_{\Gamma} \mathrm{P}_{C|\Gamma B}\mathrm{P}_{\Gamma}\right)[0]_B$$
$$= P_A P_C\ [0]_B$$



Evans

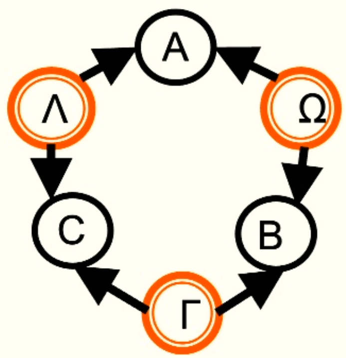
Compatible set **does** include:

$$P^{\mathrm{corr}}_{ABC} = \frac{1}{2}[000] + \frac{1}{2}[111]$$

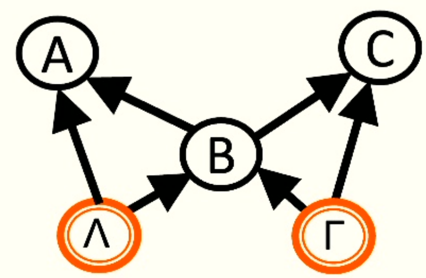
Λ is random
A and B copy Λ
C copies B

$$P_{ABC}^{\text{pinch}} = \left(\tfrac{1}{2}[00]_{AC} + \tfrac{1}{2}[11]_{AC}\right)[0]_B$$

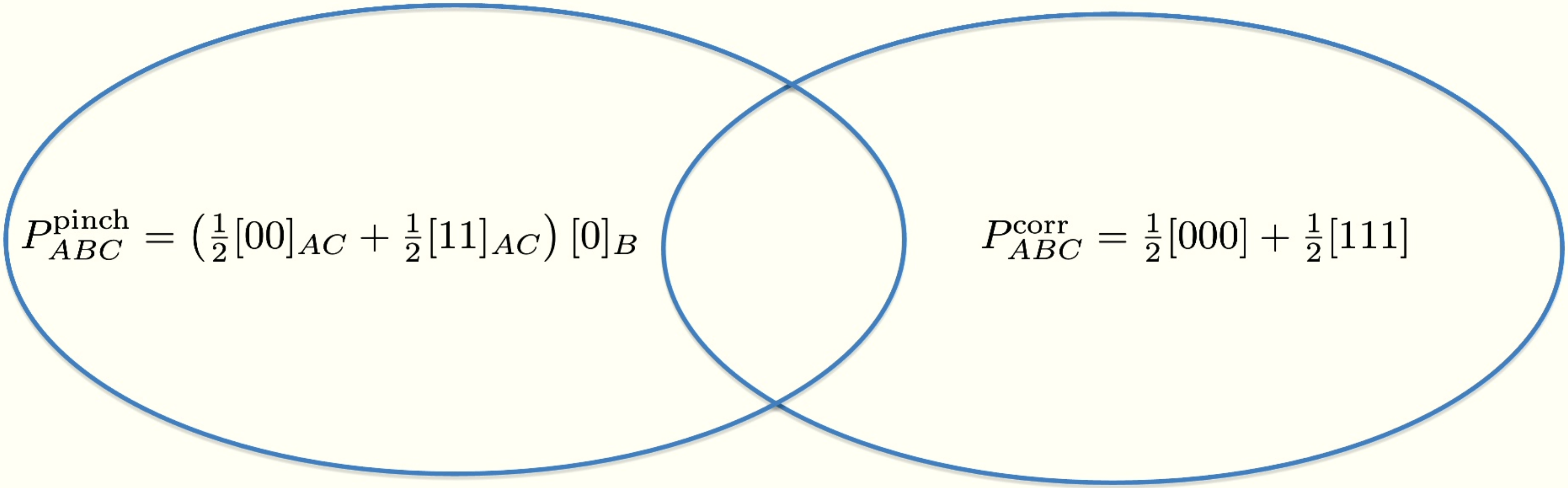
$$P_{ABC}^{\text{corr}} = \tfrac{1}{2}[000] + \tfrac{1}{2}[111]$$



Triangle

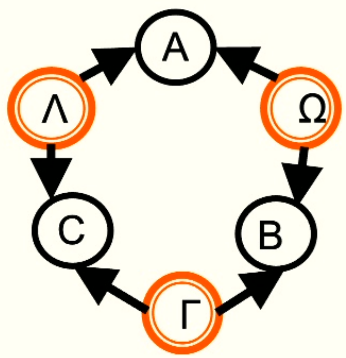


Evans

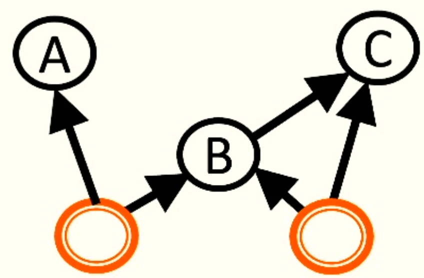


$$P^{\text{pinch}}_{ABC} = \left(\tfrac{1}{2}[00]_{AC} + \tfrac{1}{2}[11]_{AC}\right)[0]_B$$

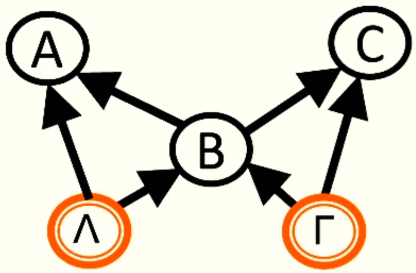
$$P^{\text{corr}}_{ABC} = \tfrac{1}{2}[000] + \tfrac{1}{2}[111]$$



Triangle



Instrumental



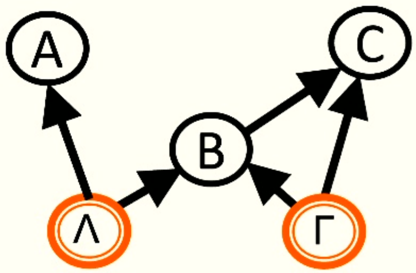
Evans

Compatible set **does** include:

$$P^{\text{pinch2}}_{ABC} = \frac{1}{2}\left(\frac{1}{2}[00]_{AC} + \frac{1}{2}[11]_{AC}\right)[0]_B + \frac{1}{2}[00]_{AC}[1]_B$$

Proof:

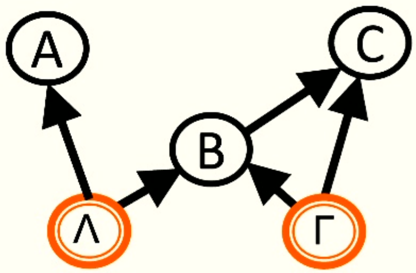
$$P_\Lambda = \frac{1}{2}[0]_\Lambda + \frac{1}{2}[1]_\Lambda$$
$$P_\Gamma = \frac{1}{2}[0]_\Gamma + \frac{1}{2}[1]_\Gamma$$
$$P_{B|\Lambda\Gamma} = \delta_{B,\Lambda\oplus\Gamma}$$
$$P_{A|B\Lambda} = \delta_{A,\Lambda}\delta_{B,0} + \delta_{A,0}\delta_{B,1}$$
$$P_{C|B\Gamma} = \delta_{C,\Gamma}\delta_{B,0} + \delta_{C,0}\delta_{B,1}$$



Instrumental

Compatible set **does not** include:

$$P_{ABC}^{\text{pinch2}} = \frac{1}{2}\left(\frac{1}{2}[00]_{AC} + \frac{1}{2}[11]_{AC}\right)[0]_B + \frac{1}{2}[00]_{AC}[1]_B$$



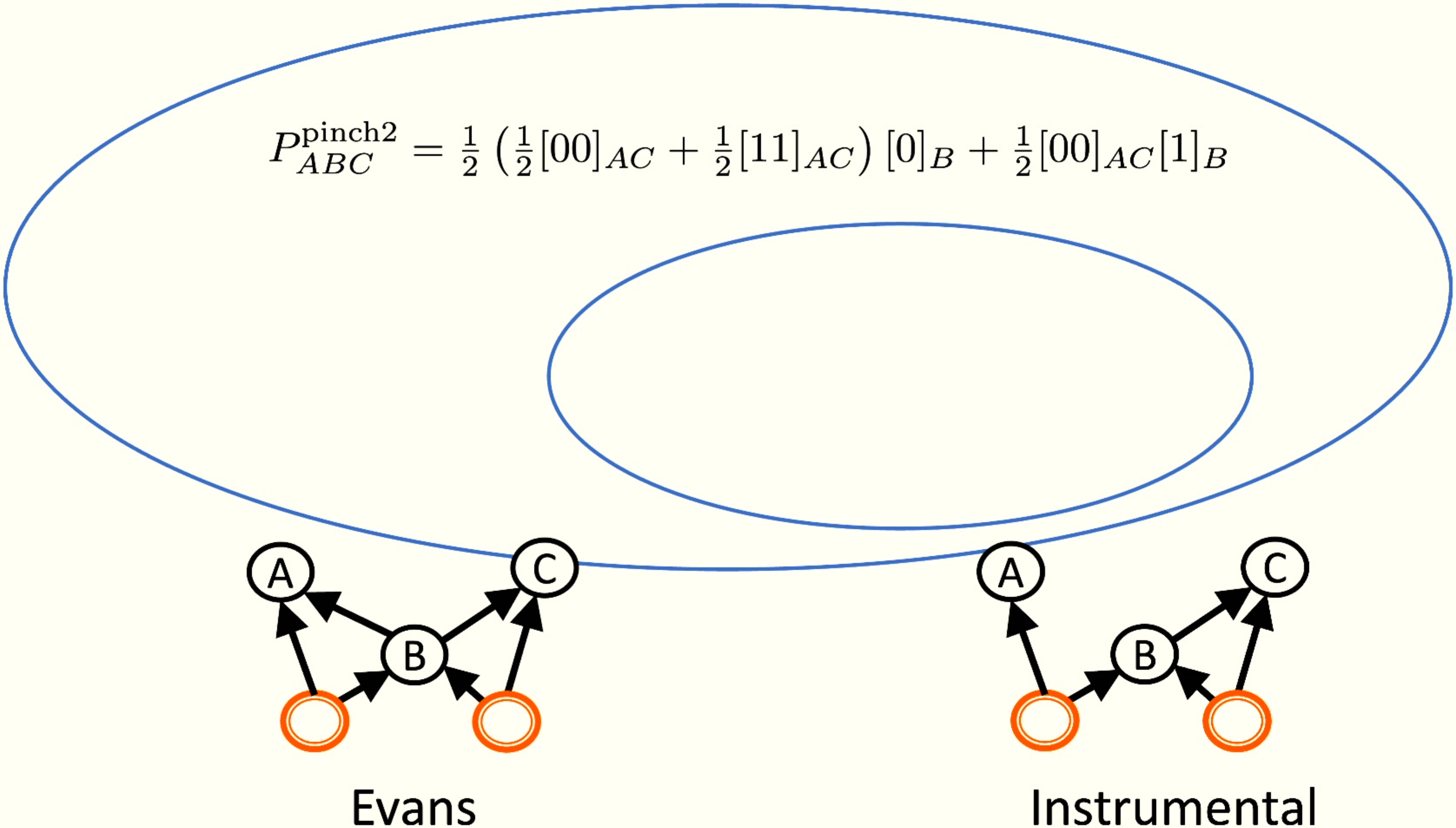
Instrumental

Compatible set **does not** include:

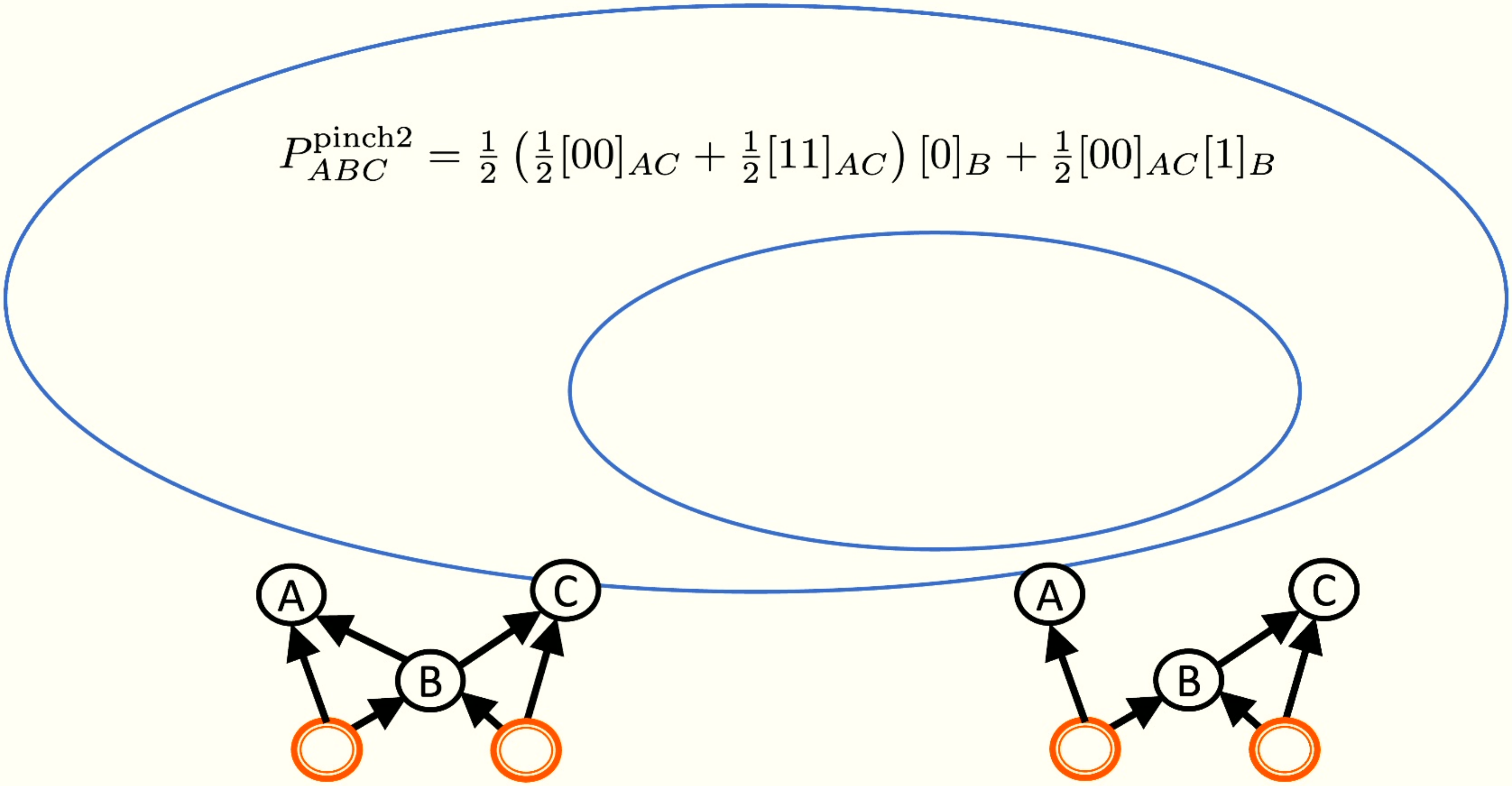
$$P_{ABC}^{\text{pinch2}} = \frac{1}{2}\left(\frac{1}{2}[00]_{AC} + \frac{1}{2}[11]_{AC}\right)[0]_B + \frac{1}{2}[00]_{AC}[1]_B$$

I don't know of a simple way to prove this.

$$P_{ABC}^{\text{pinch2}} = \frac{1}{2}\left(\frac{1}{2}[00]_{AC} + \frac{1}{2}[11]_{AC}\right)[0]_B + \frac{1}{2}[00]_{AC}[1]_B$$

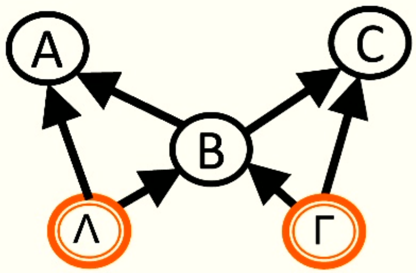


$$P_{ABC}^{\text{pinch2}} = \frac{1}{2}\left(\frac{1}{2}[00]_{AC} + \frac{1}{2}[11]_{AC}\right)[0]_B + \frac{1}{2}[00]_{AC}[1]_B$$



Evans

Instrumental



Evans

Compatible set **does** include:

$$P_{ABC}^{\text{pinch2}} = \frac{1}{2}\left(\frac{1}{2}[00]_{AC} + \frac{1}{2}[11]_{AC}\right)[0]_B + \frac{1}{2}[00]_{AC}[1]_B$$

Proof:

$$P_\Lambda = \frac{1}{2}[0]_\Lambda + \frac{1}{2}[1]_\Lambda$$

$$P_\Gamma = \frac{1}{2}[0]_\Gamma + \frac{1}{2}[1]_\Gamma$$

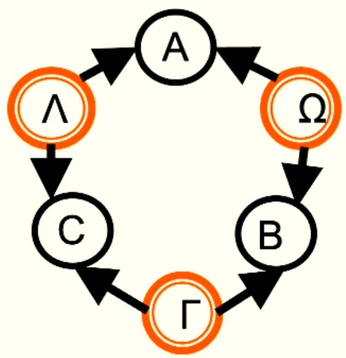
$$P_{B|\Lambda\Gamma} = \delta_{B,\Lambda\oplus\Gamma}$$

$$P_{A|B\Lambda} = \delta_{A,\Lambda}\delta_{B,0} + \delta_{A,0}\delta_{B,1}$$

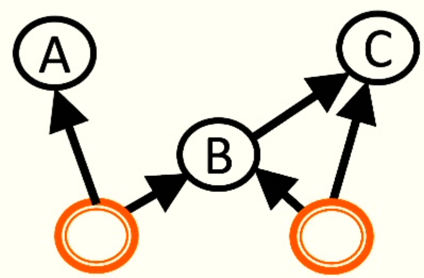
$$P_{C|B\Gamma} = \delta_{C,\Gamma}\delta_{B,0} + \delta_{C,0}\delta_{B,1}$$

$$P_{ABC}^{\text{pinch}} = \left(\tfrac{1}{2}[00]_{AC} + \tfrac{1}{2}[11]_{AC}\right)[0]_B$$

$$P_{ABC}^{\text{corr}} = \tfrac{1}{2}[000] + \tfrac{1}{2}[111]$$



Triangle

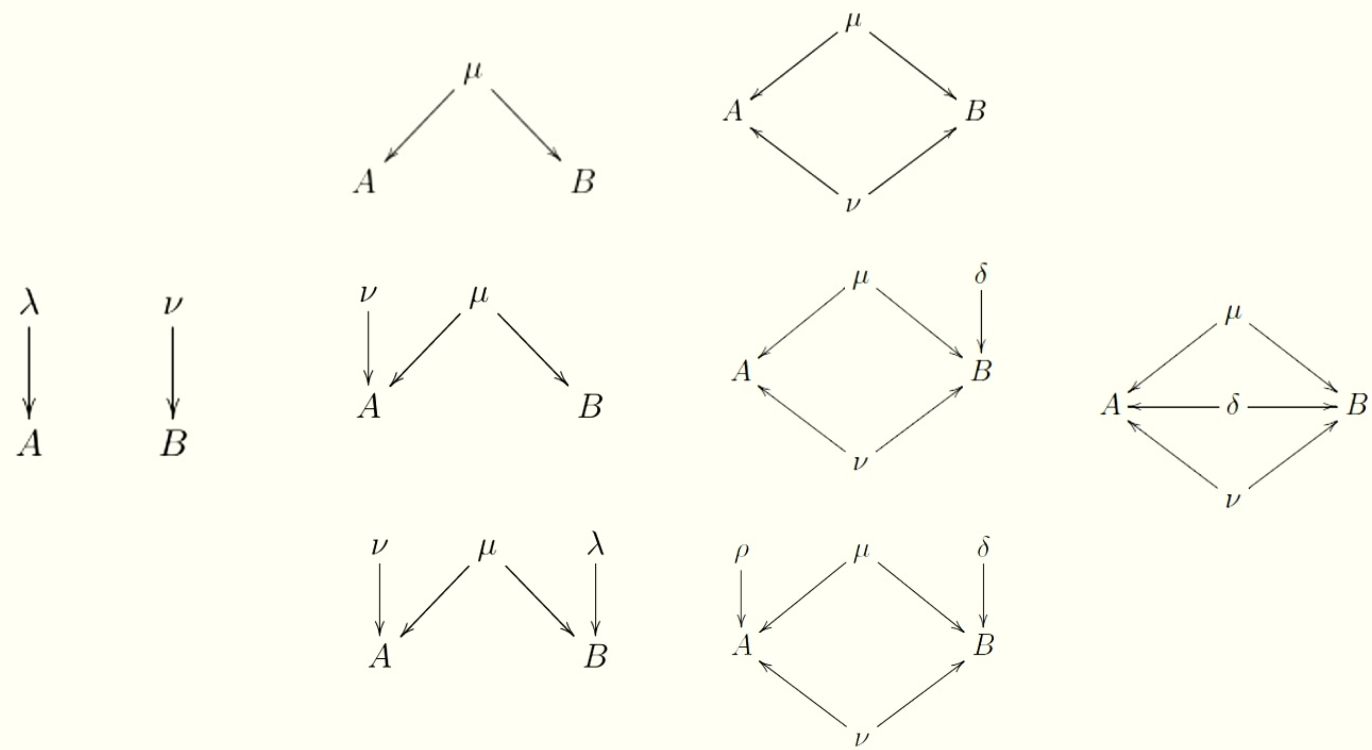


Instrumental

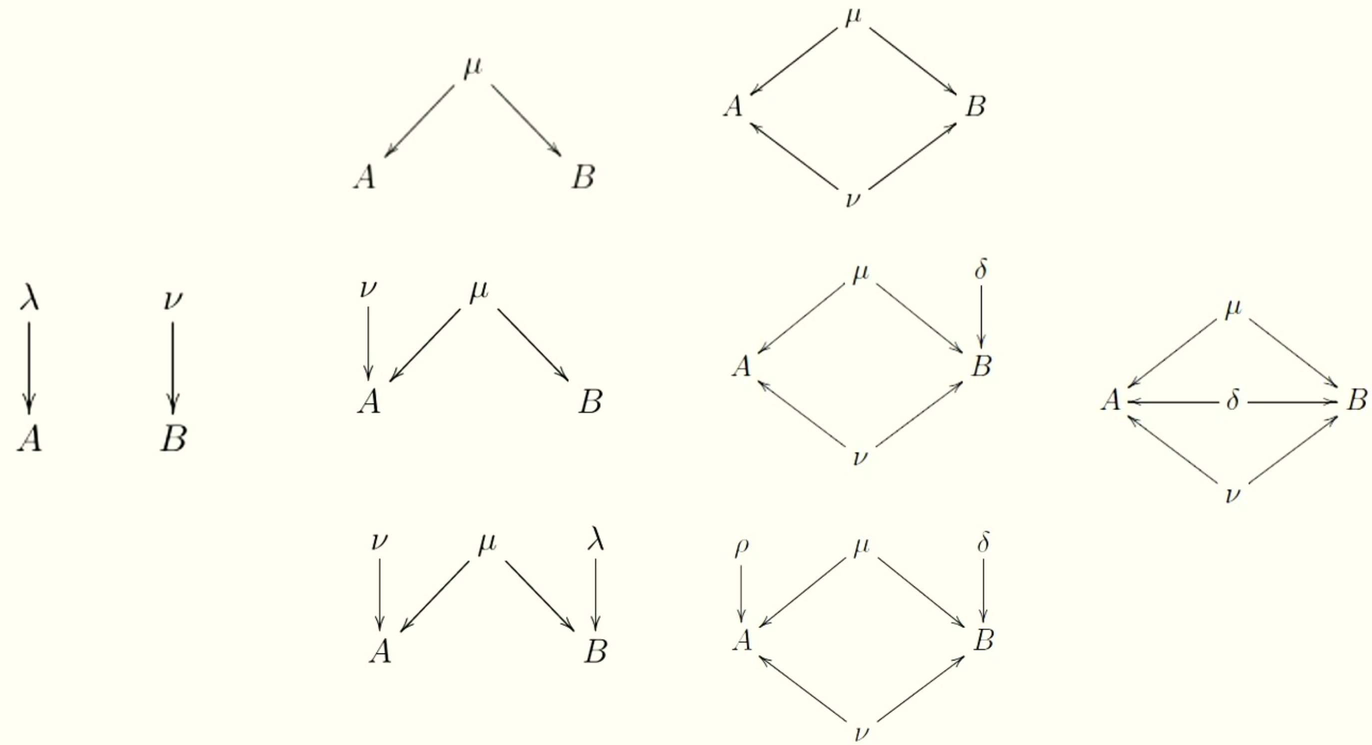
Inequality constraints for causal models

Inequality constraints for structural causal models on two binary observed variables with binary latent variables by quantifier elimination

Observed variables are binary
All latent variables are binary

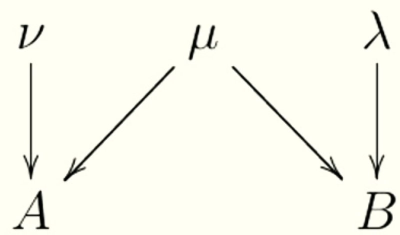


Observed variables are binary
All latent variables are binary



Note: conditional independence techniques can only determine whether A and B are causally connected or not

Different choices of the functional dependences of A and B on the latent variables



$$A = \nu \oplus \mu \qquad B = \lambda \oplus \mu$$

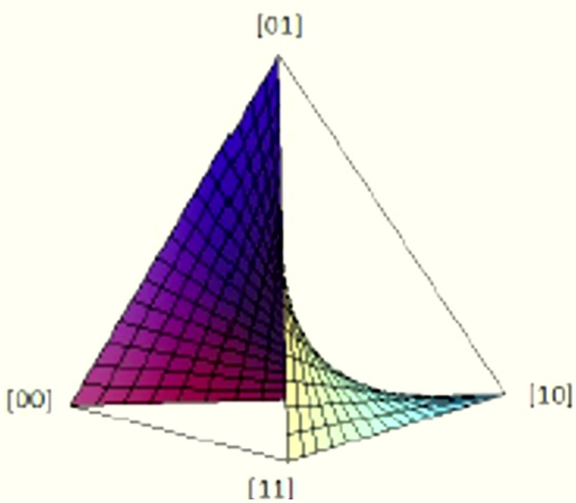
$$A = \nu \oplus \mu \oplus 1 \qquad B = \lambda \oplus \mu \oplus 1$$

$$A = \nu\mu \qquad B = \lambda\mu$$

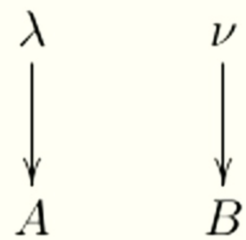
$$A = \nu\mu \oplus 1 \qquad B = \lambda\mu \oplus 1$$

$$\begin{array}{ccc} \lambda & & \nu \\ \downarrow & & \downarrow \\ A & & B \end{array}$$

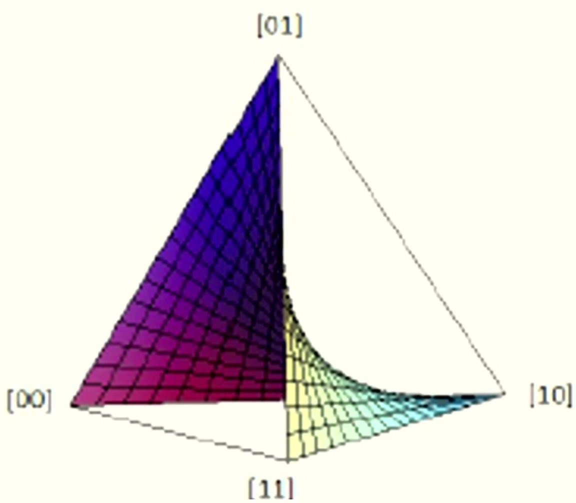
$$A = \lambda \qquad\qquad B = \nu$$



$$p_{00}p_{11} = p_{01}p_{10}$$



$$A = \lambda \qquad\qquad B = \nu$$



$$p_{00}p_{11} = p_{01}p_{10}$$

$$\begin{array}{ccc} \lambda & & \nu \\ \downarrow & & \downarrow \\ A & & B \end{array}$$

$$A = \lambda \qquad\qquad B = \nu$$

$$q_1 \equiv p(\lambda = 0)$$
$$q_2 \equiv p(\nu = 0)$$

$$\bar{q}_1 \equiv 1 - q_1$$
$$\bar{q}_2 \equiv 1 - q_2$$

$$p_{00} = q_1 q_2$$
$$p_{01} = q_1 \bar{q}_2$$
$$p_{10} = \bar{q}_1 q_2$$
$$p_{10} = \bar{q}_1 \bar{q}_2$$

$$p_{00} = q_1 q_2$$
$$p_{01} = q_1 \bar{q}_2$$
$$p_{10} = \bar{q}_1 q_2$$
$$p_{10} = \bar{q}_1 \bar{q}_2$$

This system of polynomial equations defines an algebraic variety in 6d

$$p_{00} = q_1 q_2$$
$$p_{01} = q_1 \bar{q}_2 \qquad 0 \leq q_1 \leq 1$$
$$p_{10} = \bar{q}_1 q_2 \qquad 0 \leq q_2 \leq 1$$
$$p_{10} = \bar{q}_1 \bar{q}_2$$

This system of polynomial equalities and inequalities defines a semi-algebraic set in 6d

$$p_{00} = q_1 q_2$$
$$p_{01} = q_1 \bar{q}_2$$
$$p_{10} = \bar{q}_1 q_2$$
$$p_{10} = \bar{q}_1 \bar{q}_2$$

This system of polynomial equations defines an algebraic variety in 6d

$$p_{00} = q_1 q_2$$
$$p_{01} = q_1 \bar{q}_2 \qquad 0 \leq q_1 \leq 1$$
$$p_{10} = \bar{q}_1 q_2 \qquad 0 \leq q_2 \leq 1$$
$$p_{10} = \bar{q}_1 \bar{q}_2$$

This system of polynomial equalities and inequalities defines a semi-algebraic set in 6d

Project onto 4d subspace of $p_{00}, p_{01}, p_{10}, p_{11}$

By Tarski-Seidenberg theorem, this is also a semi-algebraic set

$$p_{00} = q_1 q_2$$
$$p_{01} = q_1 \bar{q}_2$$
$$p_{10} = \bar{q}_1 q_2$$
$$p_{10} = \bar{q}_1 \bar{q}_2$$

This system of polynomial equations defines an algebraic variety in 6d

$$p_{00} = q_1 q_2$$
$$p_{01} = q_1 \bar{q}_2$$
$$p_{10} = \bar{q}_1 q_2$$
$$p_{10} = \bar{q}_1 \bar{q}_2$$

$$0 \leq q_1 \leq 1$$
$$0 \leq q_2 \leq 1$$

This system of polynomial equalities and inequalities defines a semi-algebraic set in 6d

Project onto 4d subspace of $p_{00}, p_{01}, p_{10}, p_{11}$

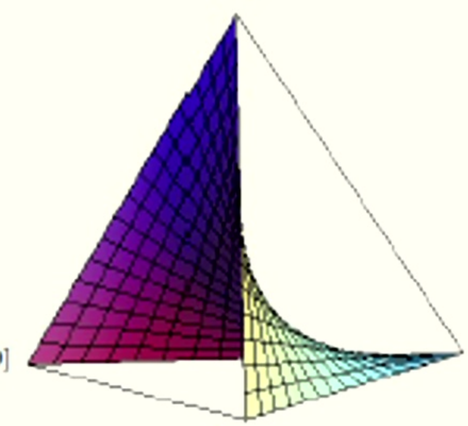
By Tarski-Seidenberg theorem, this is also a semi-algebraic set

Implicitization

a.k.a

Quantifier Elimination

$$p_{00} p_{11} = p_{01} p_{10}$$
$$0 \leq p_{ab} \leq 1 \; \forall a, b$$
$$\sum_{ab} p_{ab} = 1$$



$$p_{00} = q_1 q_2$$
$$p_{01} = q_1 \bar{q}_2$$
$$p_{10} = \bar{q}_1 q_2$$
$$p_{10} = \bar{q}_1 \bar{q}_2$$

imply

$$p_{00} p_{11} = p_{01} p_{10}$$

because

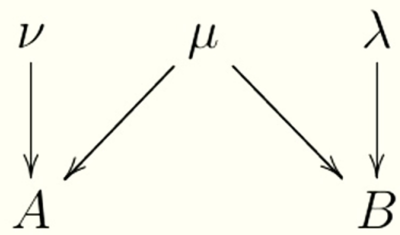
$$(q_1 q_2)(\bar{q}_1 \bar{q}_2) = (q_1 \bar{q}_2)(\bar{q}_1 q_2)$$

and $\sum_{ab} p_{ab} = 1$

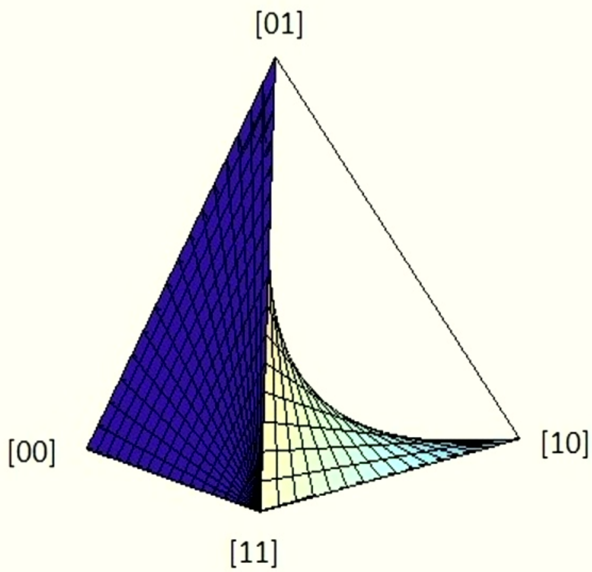
$$0 \leq q_1 \leq 1$$
$$0 \leq q_2 \leq 1$$

imply only

$$0 \leq p_{ab} \leq 1 \; \forall a, b$$

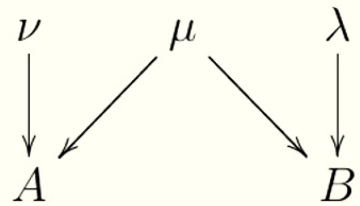


$$A = \mu\nu \qquad\qquad B = \mu\lambda$$



$$p_{00}p_{11} > p_{01}p_{10}$$

Inequality constraint



$$A = \mu\nu \qquad B = \mu\lambda$$

μ	ν	λ	$A = \mu\nu$	$B = \mu\lambda$
0	0	0	0	0
0	0	1	0	0
0	1	0	0	0
0	1	1	0	0
1	0	0	0	0
1	0	1	0	1
1	1	0	1	0
1	1	1	1	1

$$q_1 \equiv p(\mu = 0)$$
$$q_2 \equiv p(\nu = 0)$$
$$q_3 \equiv p(\lambda = 0)$$

$$p_{00} = q_1 + \bar{q}_1 q_2 q_3$$
$$p_{01} = \bar{q}_1 q_2 \bar{q}_3$$
$$p_{10} = \bar{q}_1 \bar{q}_2 q_3$$
$$p_{10} = \bar{q}_1 \bar{q}_2 \bar{q}_3$$

$$p_{00} = q_1 + \bar{q}_1 q_2 q_3$$
$$p_{01} = \bar{q}_1 q_2 \bar{q}_3$$
$$p_{10} = \bar{q}_1 \bar{q}_2 q_3$$
$$p_{11} = \bar{q}_1 \bar{q}_2 \bar{q}_3$$

This system of polynomial equations defines an algebraic variety in 7d

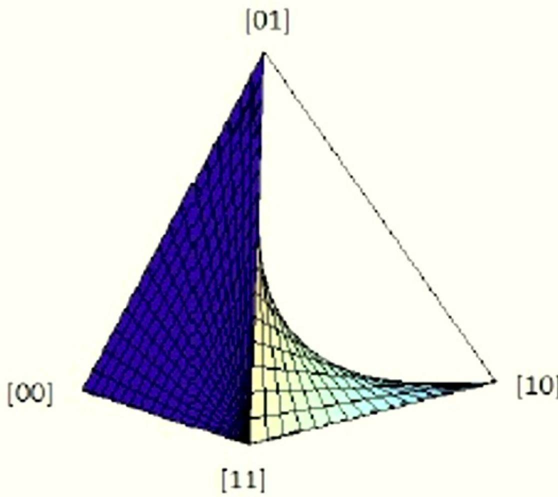
$$p_{00} = q_1 + \bar{q}_1 q_2 q_3$$
$$p_{01} = \bar{q}_1 q_2 \bar{q}_3$$
$$p_{10} = \bar{q}_1 \bar{q}_2 q_3$$
$$p_{11} = \bar{q}_1 \bar{q}_2 \bar{q}_3$$

$$0 \leq q_1 \leq 1$$
$$0 \leq q_2 \leq 1$$
$$0 \leq q_3 \leq 1$$

This system of polynomial equalities and inequalities defines a semi-algebraic set in 7d

Project onto 4d subspace of $p_{00}, p_{01}, p_{10}, p_{11}$
By Tarski-Seidenberg theorem, this is also a semi-algebraic set

$$p_{00} p_{11} \geq p_{01} p_{10}$$
$$0 \leq p_{ab} \leq 1 \ \forall a, b$$
$$\sum_{ab} p_{ab} = 1$$



Let $k[x_1, \ldots, x_n]$ denote the polynomials on $x_1, \ldots, x_n$ over the field k

$$f = \sum_\alpha c_\alpha x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}, \quad \text{where } \alpha \equiv (\alpha_1, \alpha_2, \ldots, \alpha_n)$$

An algebraic variety is the solution set of a system of polynomial equations

$$\mathbf{V}(f_1, \ldots, f_s) = \{(a_1, \ldots, a_n) \in k^n : \\ f_i(a_1, \ldots, a_n) = 0, \forall, 1 \leq i \leq s\}$$

A subset $I \subset k[x_1, \ldots, x_n]$ is an ideal *if it satisfies:*

1. $0 \in I$,

2. If $f, g \in I$, then $f + g \in I$,

3. If $f \in I$ and $h \in k[x_1, \ldots, x_n]$, then $hf \in I$.

Let $k[x_1, \dots, x_n]$ denote the polynomials on $x_1, \dots, x_n$ over the field k

$$f = \sum_\alpha c_\alpha x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}, \quad \text{where} \;\; \alpha \equiv (\alpha_1, \alpha_2, \dots, \alpha_n)$$

An algebraic variety is the solution set of a system of polynomial equations

$$\mathbf{V}(f_1, \dots, f_s) = \{(a_1, \dots, a_n) \in k^n : \\ f_i(a_1, \dots, a_n) = 0, \forall, 1 \leq i \leq s\}$$

A subset $I \subset k[x_1, \dots, x_n]$ is an ideal *if it satisfies:*

1. $0 \in I$,

2. If $f, g \in I$, then $f + g \in I$,

3. If $f \in I$ and $h \in k[x_1, \dots, x_n]$, then $hf \in I$.

An ideal defines an algebraic variety

$$\mathbf{V}(I) = \{(a_1, \dots, a_n) \in k^n : f(a_1, \dots, a_n) = 0, \quad \forall f \in I\}$$

ideal generated by $f_1, \dots, f_s$

$$\langle f_1, \dots, f_s \rangle = \Big\{ \sum_{i=1}^{s} h_i f_i : h_1, \dots, h_s \in k[x_1, \dots, x_n] \Big\}.$$

$$\mathbf{V}(I) = \mathbf{V}(f_1, \dots, f_s)$$

$$p_{00} = q_1 + \bar{q}_1 q_2 q_3$$
$$p_{01} = \bar{q}_1 q_2 \bar{q}_3$$
$$p_{10} = \bar{q}_1 \bar{q}_2 q_3$$
$$p_{11} = \bar{q}_1 \bar{q}_2 \bar{q}_3$$

The ideal is $\langle p_{00} - q_1 - \overline{q}_1 q_2 q_3, p_{01} - \overline{q}_1 q_2 \overline{q}_3, p_{10} - \overline{q}_1 \overline{q}_2 q_3, p_{11} - \overline{q}_1 \overline{q}_2 \overline{q}_3 \rangle.$

The Groebner basis for the ideal [6] is given by

$$g_1 = q_2 q_1 - q_1 - q_2 - p_{10} - p_{11} + 1$$
$$g_2 = q_3 q_1 - q_1 - q_3 - p_{01} - p_{11} + 1$$
$$g_3 = q_3 p_{10} + q_3 p_{11} - p_{10}$$
$$g_4 = q_2 p_{01} + q_3 p_{11} - p_{01}$$
$$g_5 = p_{00} + p_{01} + p_{10} + p_{11} - 1$$
$$g_6 = p_{11}^2 + p_{01} p_{10} + p_{11} p_{10} - p_{11} + p_{01} p_{11} + p_{11} q_1.$$

The semi-algebraic set we seek are the sol'ns to:

$$g_1 = g_2 = g_3 = g_4 = g_5 = g_6 = 0$$

Under the constraint that $0 < q_i < 1$

$$g_1 = q_2 q_1 - q_1 - q_2 - p_{10} - p_{11} + 1$$
$$g_2 = q_3 q_1 - q_1 - q_3 - p_{01} - p_{11} + 1$$
$$g_3 = q_3 p_{10} + q_3 p_{11} - p_{10}$$
$$g_4 = q_2 p_{01} + q_3 p_{11} - p_{01}$$
$$g_5 = p_{00} + p_{01} + p_{10} + p_{11} - 1$$
$$g_6 = p_{11}^2 + p_{01} p_{10} + p_{11} p_{10} - p_{11} + p_{01} p_{11} + p_{11} q_1.$$

$$g_5 = 0 \implies p_{00} + p_{01} + p_{10} + p_{11} = 1$$

$$g_6 = 0 \implies p_{11}\left(p_{10} + p_{01} + p_{11} - 1\right) + p_{01} p_{10} + p_{11} q_1 = 0$$

$$\implies q_1 = \frac{p_{11} p_{00} - p_{10} p_{01}}{p_{11}}.$$

$$q_1 \geq 0 \implies p_{00} p_{11} \geq p_{01} p_{10}$$

$$p_{00} = q_1 + \bar{q}_1 q_2 q_3$$
$$p_{01} = \bar{q}_1 q_2 \bar{q}_3$$
$$p_{10} = \bar{q}_1 \bar{q}_2 q_3$$
$$p_{11} = \bar{q}_1 \bar{q}_2 \bar{q}_3$$

This system of polynomial equations defines an algebraic variety in 7d

$$p_{00} = q_1 + \bar{q}_1 q_2 q_3$$
$$p_{01} = \bar{q}_1 q_2 \bar{q}_3$$
$$p_{10} = \bar{q}_1 \bar{q}_2 q_3$$
$$p_{11} = \bar{q}_1 \bar{q}_2 \bar{q}_3$$

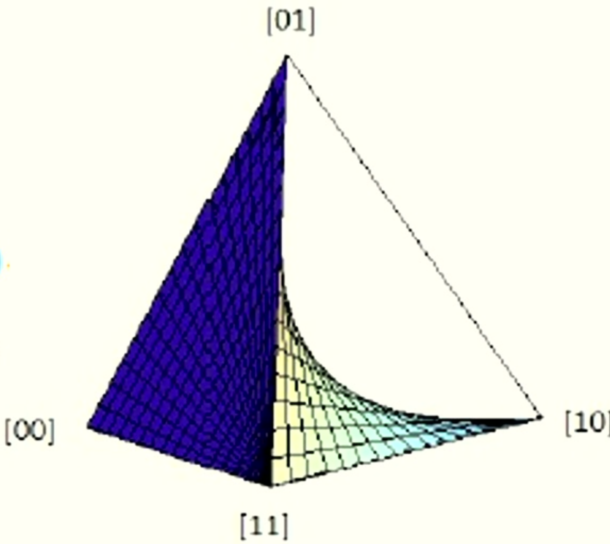
$$0 \leq q_1 \leq 1$$
$$0 \leq q_2 \leq 1$$
$$0 \leq q_3 \leq 1$$

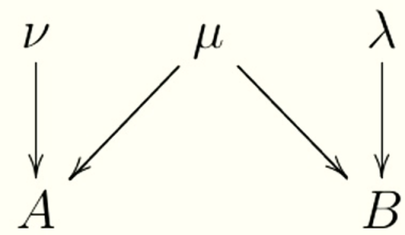
This system of polynomial equalities and inequalities defines a semi-algebraic set in 7d

Project onto 4d subspace of $p_{00}, p_{01}, p_{10}, p_{11}$
By Tarski-Seidenberg theorem, this is also a semi-algebraic set

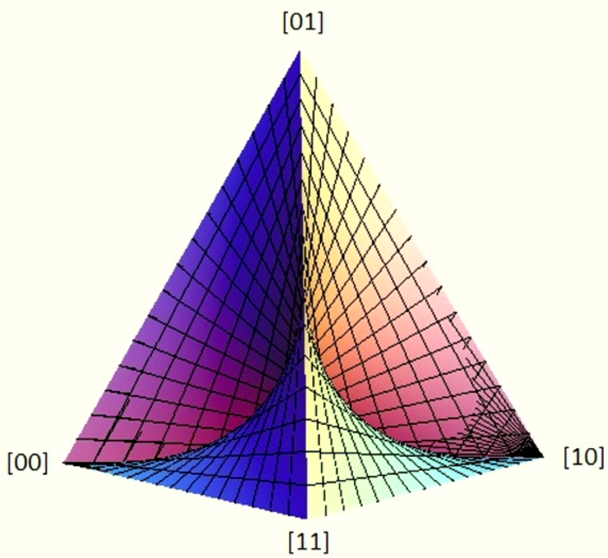
$$p_{00}p_{11} \geq p_{01}p_{10}$$
$$0 \leq p_{ab} \leq 1 \ \forall a, b$$
$$\sum_{ab} p_{ab} = 1$$

The nontrivial constraint is a polynomial **inequality**
We must intersect the tetrahedron with a **semi-algebraic set**

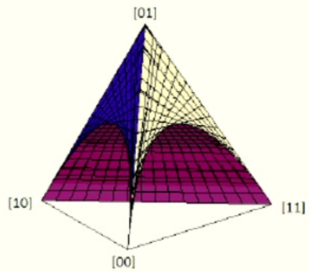
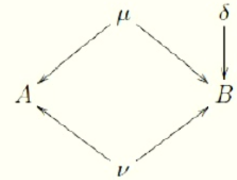
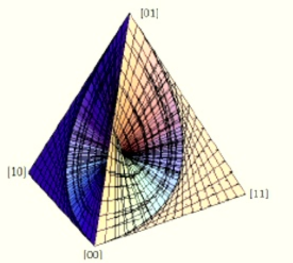
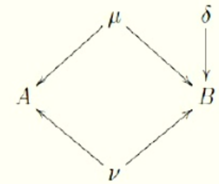
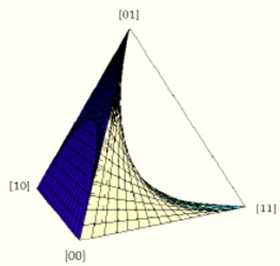
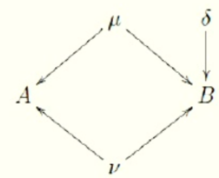




$$A = \nu \oplus \mu \qquad B = \mu\lambda$$

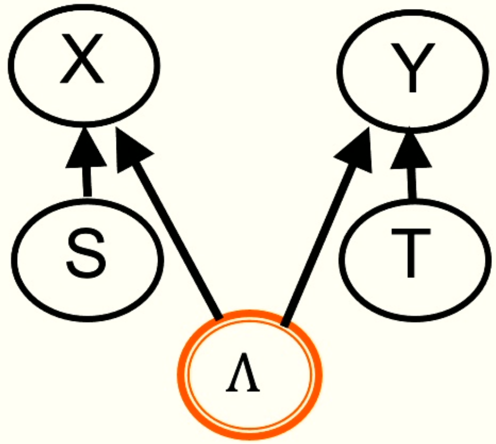


$$(p_{01} - p_{11})(p_{00}p_{11} - p_{01}p_{10}) < 0$$
$$(p_{01} - p_{11})(p_{10}p_{11} - p_{01}p_{00}) < 0$$

#	Algebraic variety generated by the class	Test for feasibility of the class	Minimal causal model of the class
#14		$\left(p_{01} + 2p_{10} + 2p_{11} - 2\right)^2 \geq 4p_{00}$	 $A = \mu \oplus \nu, \quad B = \mu\nu\delta \oplus \mu \oplus \nu$
Id, f_B, f_A, f_{AB}, $f_{AB}SX$, f_BXS, XS, f_AXS, SX, f_BSX, f_ASX, $f_{AB}SX$			
#15		$4(p_{10} - p_{11})(p_{00}p_{10} - p_{01}p_{11}) \leq \left(p_{11}(2p_{01} + p_{11}) - p_{10}(2p_{00} + p_{10})\right)^2$, $4(p_{10} - p_{11})(p_{01}p_{10} - p_{00}p_{11}) \leq \left(p_{11}(2p_{01} + p_{11}) - p_{10}(2p_{00} + p_{10})\right)^2$	 $A = \mu \oplus \nu, \quad B = \mu\nu \oplus \delta$
Id, $f_{AB}S$, $f_{AB}X$, f_BX, f_AS			
#16		$\lvert 4(p_{10} - p_{11})(p_{00}p_{10} - p_{01}p_{11})\rvert \leq \left(p_{11}(2p_{01} + 2p_{10} + p_{00}) - p_{10}(2p_{00} + 2p_{11} + p_{01})\right)^2$ $p_{00}p_{11} > p_{01}p_{10}$	 $A = \nu\mu, \quad B = \mu \oplus \nu\delta$
Same as #13			

BUT: Nonlinear quantifier elimination scales very badly with the size of the problem

Bell scenario



$P_{X|S\Lambda}$

$P_{Y|T\Lambda}$

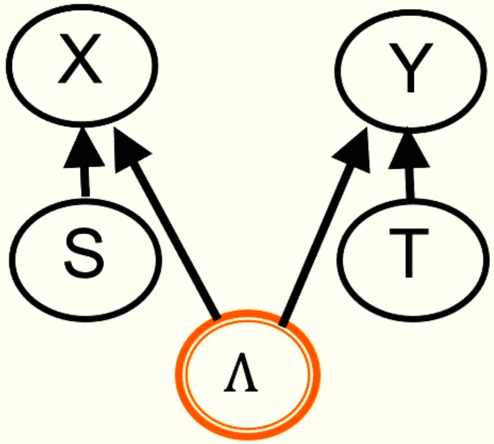
P_{Λ}

P_S

P_T

$$P_{XY|ST} = \sum_{\Lambda} P_{Y|T\Lambda} P_{X|S\Lambda} P_{\Lambda}$$

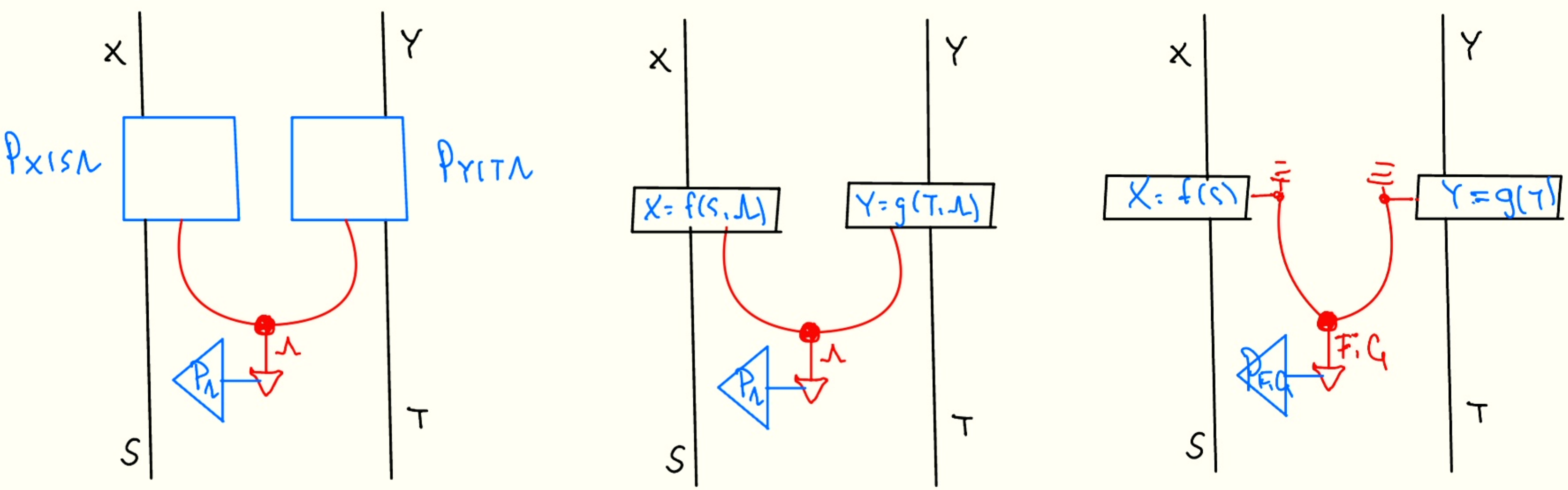
Bell scenario



$P_{X|S\Lambda}$
$P_{Y|T\Lambda}$
P_{Λ}
P_S
P_T

$$P_{XY|ST} = \sum_{\Lambda} P_{Y|T\Lambda} P_{X|S\Lambda} P_{\Lambda}$$

But the cardinality of Λ is unrestricted!


X
Y
S
T
$P_{X|S\Lambda}$
$P_{Y|T\Lambda}$
Λ
P_Λ
X = f(S, Λ)
Y = g(T, Λ)
X = f(S)
Y = g(T)
F, G
$P_{F,G}$

$$P_{XY|ST} = \sum_{f,g} \delta_{X,f(S)} \delta_{Y,g(T)} P_{F,G}(f,g)$$

If X,Y,S,T are binary, Λ can have cardinality 16

$$p_{xy|st} := P_{XY|ST}(xy|st) \qquad q_{fg} := P_{F,G}(f,g)$$

$$x,y,s,t \in \{0,1\} \qquad f,g \in \{\mathrm{id},\mathrm{fp},\mathrm{r}_0,\mathrm{r}_1\}$$

$$p_{00|00} = q_{\mathrm{r}_0,\mathrm{r}_0} + q_{\mathrm{r}_0,\mathrm{id}} + q_{\mathrm{id},\mathrm{r}_0} + q_{\mathrm{id},\mathrm{id}}$$

$$p_{00|01} = q_{\mathrm{r}_0,\mathrm{r}_1} + q_{\mathrm{r}_0,\mathrm{fp}} + q_{\mathrm{id},\mathrm{r}_1} + q_{\mathrm{id},\mathrm{fp}}$$

$$0 \leq q_{fg} \leq 1 \;\forall f,g$$

$$\vdots$$

16 linear equalities + inequalities

Do linear quantifier elimination on the 16 q's.

This yields the conditional independence relations and the 8 CHSH inequalities

$$P_{X|ST} = P_{X|S}$$
$$P_{Y|ST} = P_{Y|T}$$

$$\frac{1}{4}\sum_{x=y} P_{XY|ST}(xy|00) + \frac{1}{4}\sum_{x=y} P_{XY|ST}(xy|01)$$
$$\frac{1}{4}\sum_{x=y} P_{XY|ST}(xy|10) + \frac{1}{4}\sum_{x\neq y} P_{XY|ST}(xy|11) \leq \frac{3}{4}$$

+ 7 others

Clauser, Horne, Shimony and Holte, Phys. Rev. Lett.23, 880 (1967)

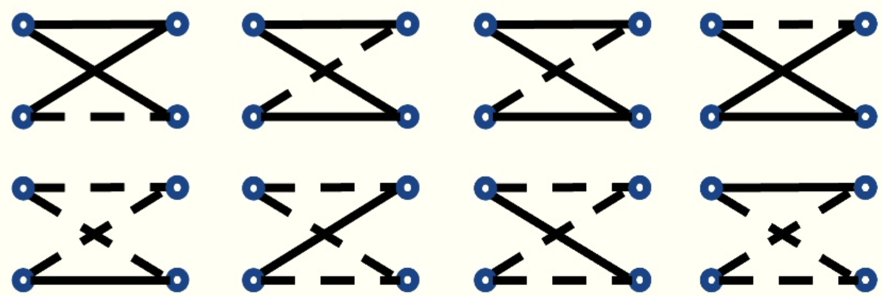
This yields the conditional independence relations and the 8 CHSH inequalities

$$P_{X|ST} = P_{X|S}$$
$$P_{Y|ST} = P_{Y|T}$$

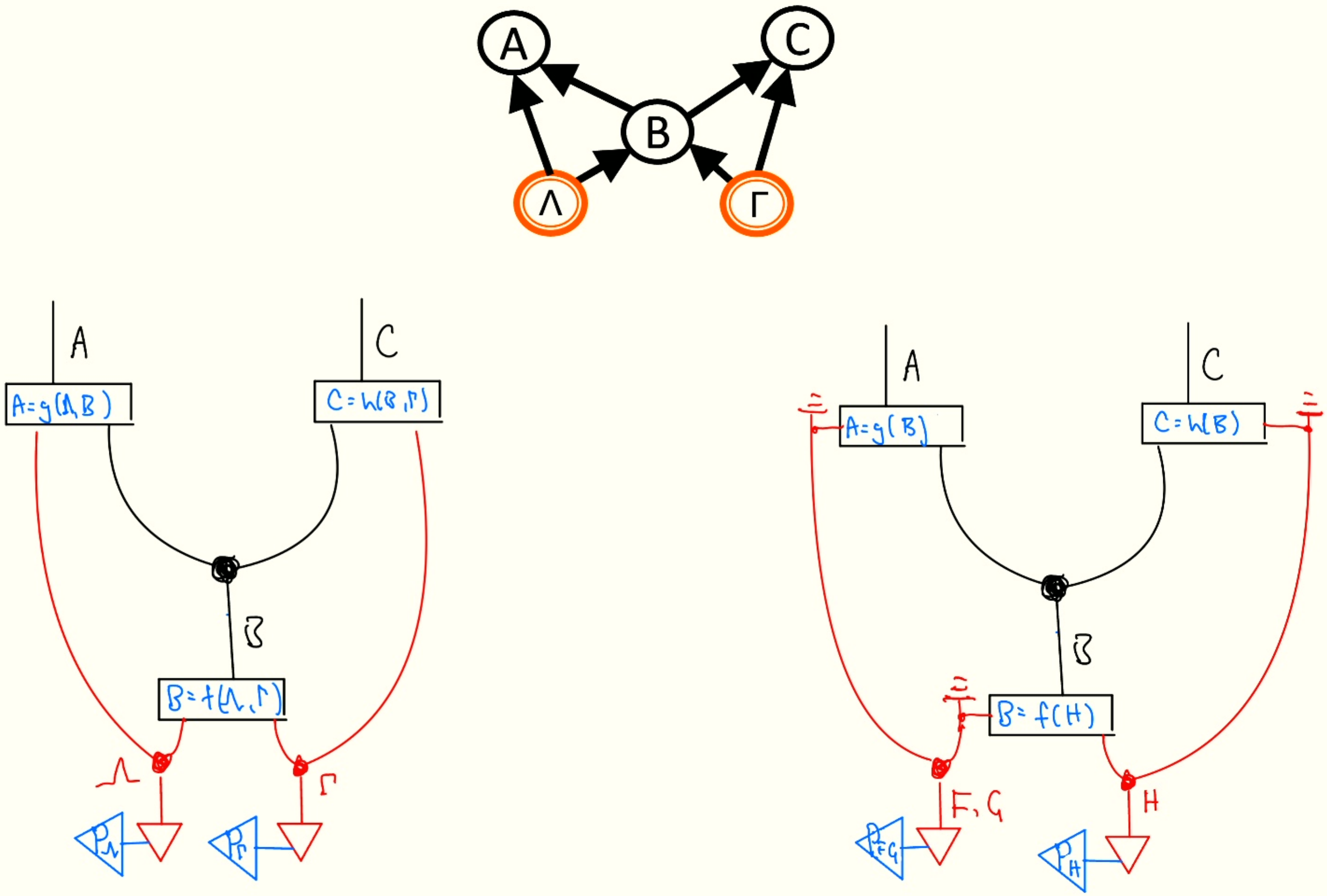
$$\frac{1}{4}\sum_{x=y} P_{XY|ST}(xy|00) + \frac{1}{4}\sum_{x=y} P_{XY|ST}(xy|01)$$
$$\frac{1}{4}\sum_{x=y} P_{XY|ST}(xy|10) + \frac{1}{4}\sum_{x\neq y} P_{XY|ST}(xy|11) \leq \frac{3}{4}$$

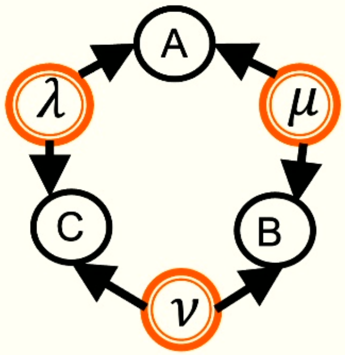
+ 7 others

Corresponding to the 8 frustrated four-node networks



Clauser, Horne, Shimony and Holte, Phys. Rev. Lett.23, 880 (1967)


A
B
C
Λ
Γ
A=g(Λ,B)
C=h(B,Γ)
B=f(Λ,Γ)
A=g(B)
C=h(B)
B=f(H)
F, G
H



Triangle

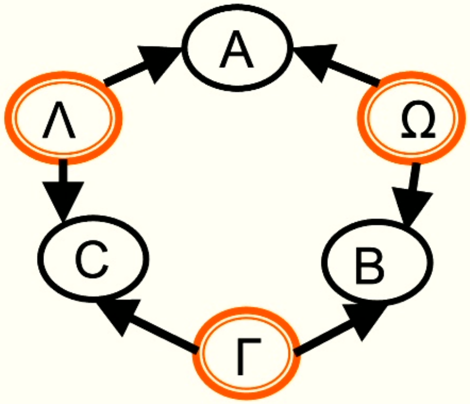
The triangle scenario is not gearable

Nonetheless, there are other techniques for determining sufficient cardinalities of the latent variables

See: D. Rosset, N. Gisin, and E. Wolfe. Quantum Inf. & Comp. **18,** 0910 (2018)

E.g. If A, B, C are binary, then it is sufficient if the the latent variables are 6-valued

Causal structure



Parameters

$\mathrm{P}_{A|\Lambda\Omega}$

$\mathrm{P}_{B|\Omega\Gamma}$

$\mathrm{P}_{C|\Gamma\Lambda}$

P_{Λ}

P_{Ω}

P_{Γ}

$$P_{ABC} = \sum_{\Lambda\Omega\Gamma} \mathrm{P}_{A|\Lambda\Omega}\mathrm{P}_{B|\Omega\Gamma}\mathrm{P}_{C|\Gamma\Lambda}\mathrm{P}_{\Lambda}\mathrm{P}_{\Omega}\mathrm{P}_{\Gamma}$$

Shannon entropy

$$H(X) := -\sum_x P_X(x) \log P_X(x)$$

Conditional entropy

$$H(X|Y) := H(XY) - H(Y)$$

Mutual information

$$I(X:Y) = H(X) + H(Y) - H(XY)$$

Conditional mutual information

$$I(X:Y|Z) = H(XZ) + H(YZ) - H(XYZ) - H(Z)$$

Entropy vector

For the joint distribution of the random variables X_1, …X_n , the components of the entropy vector are the entropies of the marginals for all possible subsets of variables:

$$(H(X_1), H(X_2), \ldots, H(X_n), H(X_1X_2), H(X_1X_3), \ldots, H(X_1, X_2, \cdots, X_n))$$

Entropy vector

For the joint distribution of the random variables X_1, ...X_n , the components of the entropy vector are the entropies of the marginals for all possible subsets of variables:

$$(H(X_1), H(X_2), \ldots, H(X_n), H(X_1X_2), H(X_1X_3), \ldots, H(X_1, X_2, \cdots, X_n))$$

The entropy vector is a coarse-grained description of the joint probability distribution

An outer approximation to the entropy cone: the Shannon cone

Monotonicity

$H(XA) \geq H(X)$

for every variable A and sets of variables X

Submodularity

$H(X) + H(XAB) \leq H(XA) + H(XB)$

where A and B are variables not in the set X

Inequalities describing the Shannon cone are termed **Shannon-type**

Valid inequalities for the entropy cone that are not Shannon-type are termed **non-Shannon-type** (R. Yeung, IEEE Trans. Inf. Th., 43, 1997)

Example: for distributions on X, Y, Z, the linear equalities defining the Shannon cone are

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 \\ 1 & 1 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 & -1 \\ 0 & -1 & 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & -1 & 0 & 1 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} H(X) \\ H(Y) \\ H(Z) \\ H(XY) \\ H(XZ) \\ H(YZ) \\ H(XYZ) \end{pmatrix} \geq 0$$

So far, these are statements about a joint distribution, with no causal content

What are the constraints on the entropies of observed variables for a given causal structure?

The entropic technique for deriving inequality constraints:

The set of all constraints on observed and latent variables are the conditional independence relations among these

These imply linear equalities on the components of the entropy vector for observed and latent variables

Add the linear inequalities of Shannon-type

Implicitize all entropic quantities that refer to latent variables

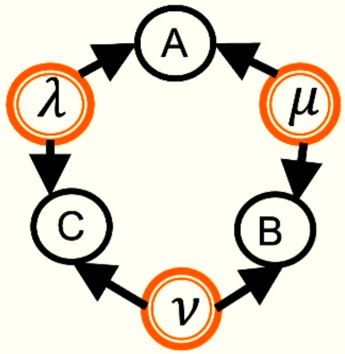
The entropic technique for deriving inequality constraints:

The set of all constraints on observed and latent variables are the conditional independence relations among these
These imply linear equalities on the components of the entropy vector for observed and latent variables

Add the linear inequalities of Shannon-type

Implicitize all entropic quantities that refer to latent variables

The result is the **marginal Shannon cone,** described by linear inequalities on entropies over observed variables only



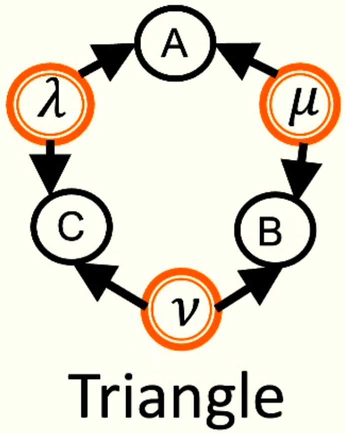
Triangle

Entropic constraint for the triangle scenario

T. Fritz, 2012

$$A \perp B|\mu \quad \lambda \perp \mu\nu$$
$$A \perp C|\lambda \quad \mu \perp \lambda\nu$$
$$B \perp C|\nu \quad \nu \perp \lambda\mu$$

Entropic constraint for the triangle scenario

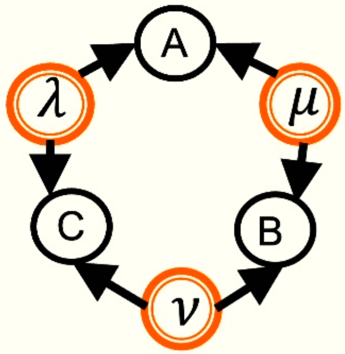


T. Fritz, 2012

$$A \perp B|\mu \quad \lambda \perp \mu\nu$$
$$A \perp C|\lambda \quad \mu \perp \lambda\nu$$
$$B \perp C|\nu \quad \nu \perp \lambda\mu$$

$$A \perp B|\mu \implies I(A:B|\mu) = 0 \implies I(A:B) \leq I(A:\mu)$$

By Shannon-type inequalities



Triangle

Entropic constraint for the triangle scenario

T. Fritz, 2012

$$A \perp B|\mu \quad \lambda \perp \mu\nu$$
$$A \perp C|\lambda \quad \mu \perp \lambda\nu$$
$$B \perp C|\nu \quad \nu \perp \lambda\mu$$

$$A \perp B|\mu \implies I(A:B|\mu) = 0 \implies I(A:B) \leq I(A:\mu)$$
$$A \perp C|\lambda \implies I(A:C|\lambda) = 0 \implies I(A:C) \leq I(A:\lambda)$$

By Shannon-type inequalities

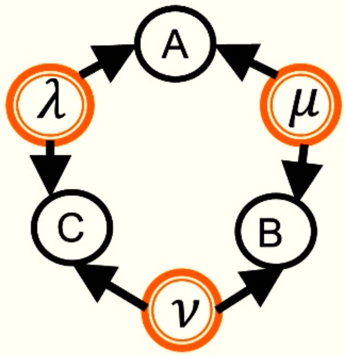
$$I(A:B) + I(A:C) \leq I(A:\mu) + I(A:\lambda)$$
$$\leq H(A) + I(\mu:\lambda)$$

By Shannon-type inequalities

$$\mu \perp \lambda \implies I(\mu:\lambda) = 0$$

This is **linear** quantifier elimination
(Fourier-Motzkin elimination)

The key: *products* of probabilities
become *sums* of entropies

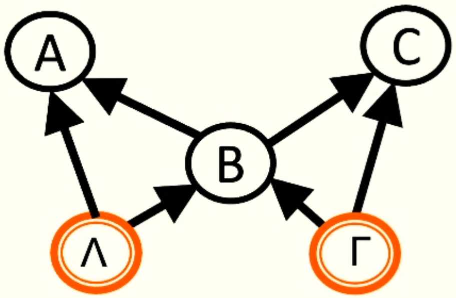


Triangle

$$I(A:B) + I(A:C) \leq H(A)$$

Note that this inequality detects the incompatibility of

$$P_{ABC} = \frac{1}{2}[000] + \frac{1}{2}[111]$$

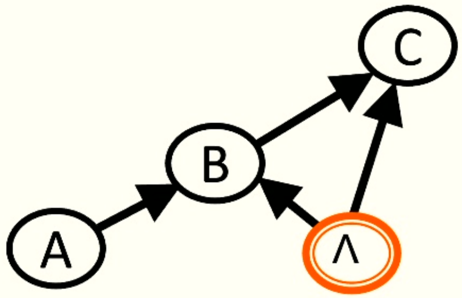


Evans

$$I(A:C) \leq H(B)$$

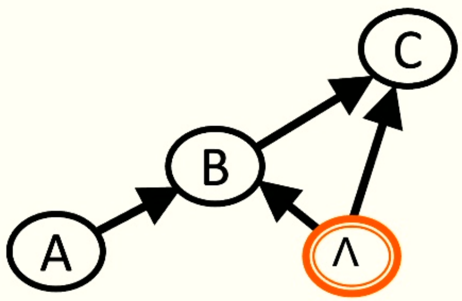
Note that this inequality detects the incompatibility of

$$P^{\text{pinch}}_{ABC} = \left(\tfrac{1}{2}[00]_{AC} + \tfrac{1}{2}[11]_{AC}\right)[0]_B$$



Instrumental

$$\Longrightarrow \quad I(A:BC) \leq H(B)$$



Instrumental

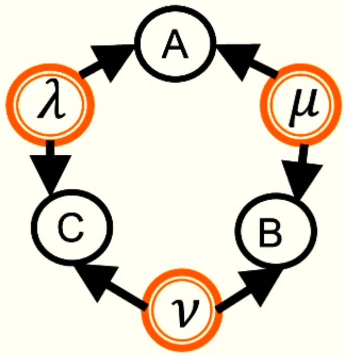
$$\Longrightarrow \quad I(A:BC) \leq H(B)$$

Note that this inequality **also** detects the incompatibility of

$$P_{ABC}^{\text{pinch}} = \left(\tfrac{1}{2}[00]_{AC} + \tfrac{1}{2}[11]_{AC}\right)[0]_B$$

But not our example separating Evans and instrumental

$$P_{ABC}^{\text{pinch2}} = \tfrac{1}{2}\left(\tfrac{1}{2}[00]_{AC} + \tfrac{1}{2}[11]_{AC}\right)[0]_B + \tfrac{1}{2}[00]_{AC}[1]_B$$



Triangle

$$\Rightarrow \quad I(A:B) + I(A:C) \leq H(A)$$

Note that this inequality detects the incompatibility of

$$P_{ABC} = \tfrac{1}{2}[000] + \tfrac{1}{2}[111]$$

But it fails to detect the incompatibility of

$$P_{ABC} = \tfrac{1}{3}[001] + \tfrac{1}{3}[010] + \tfrac{1}{3}[100]$$

The move to entropies has thrown away too much information to witness certain incompatibilities

Next Lecture:
The inflation technique