

**Title:** Lecture - Causal Inference, PHYS 777

**Speakers:** Robert Spekkens

**Collection/Series:** Causal Inference (Elective), PHYS 777, March 31 - May 2, 2025

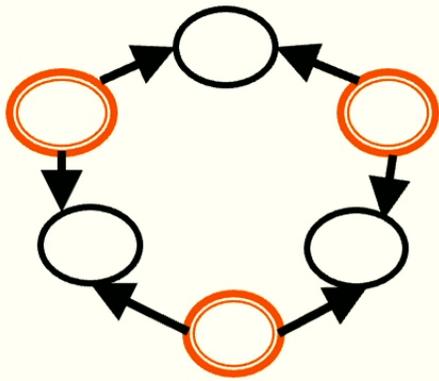
**Subject:** Quantum Foundations

**Date:** April 15, 2025 - 11:30 AM

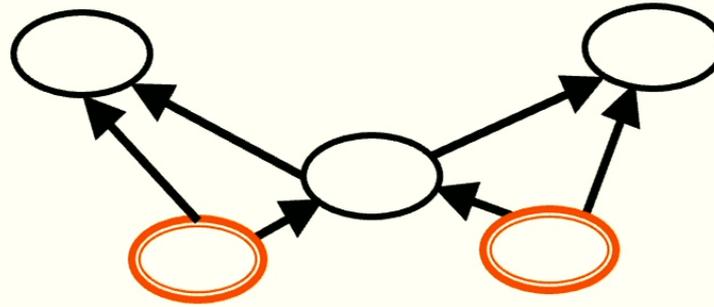
**URL:** <https://pirsa.org/25040042>

# Inequality constraints for causal models

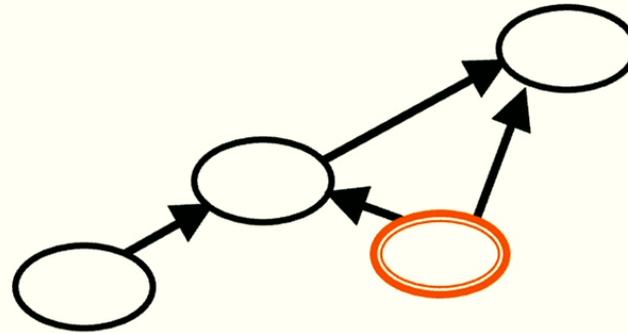
Triangle



Evans

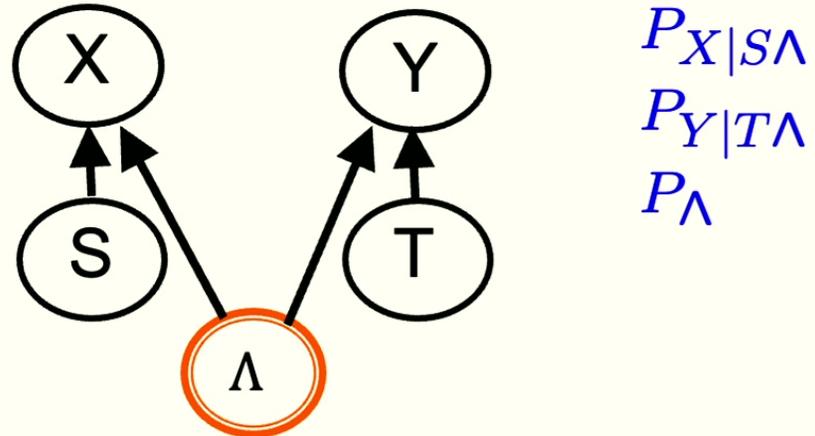


Instrumental



# Inequality constraints for causal models by brute-force quantifier elimination

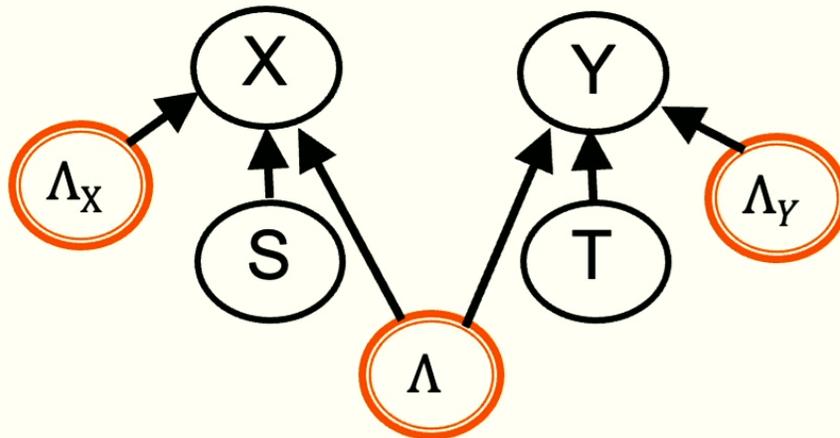
## Bell scenario



$$P_{X|S\Lambda}$$
$$P_{Y|T\Lambda}$$
$$P_{\Lambda}$$

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

## Bell scenario



$$X = f_X(S, \Lambda, \Lambda_X)$$

$$Y = f_Y(T, \Lambda, \Lambda_Y)$$

$$P_\Lambda$$

$$P_{\Lambda_X}$$

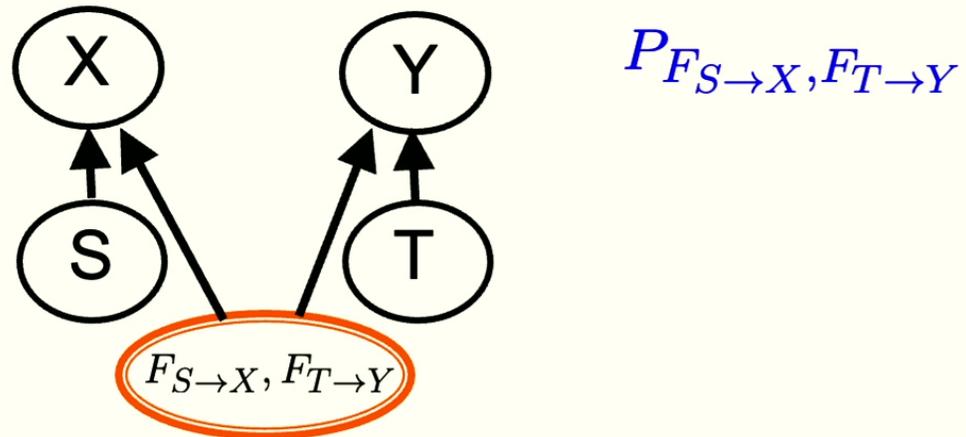
$$P_{\Lambda_Y}$$

$$P_{XY|ST} = \sum_{\Lambda, \Lambda_X, \Lambda_Y} \delta_{X, f_X(S, \Lambda, \Lambda_X)} \delta_{Y, f_Y(T, \Lambda, \Lambda_Y)} P_\Lambda P_{\Lambda_X} P_{\Lambda_Y}$$

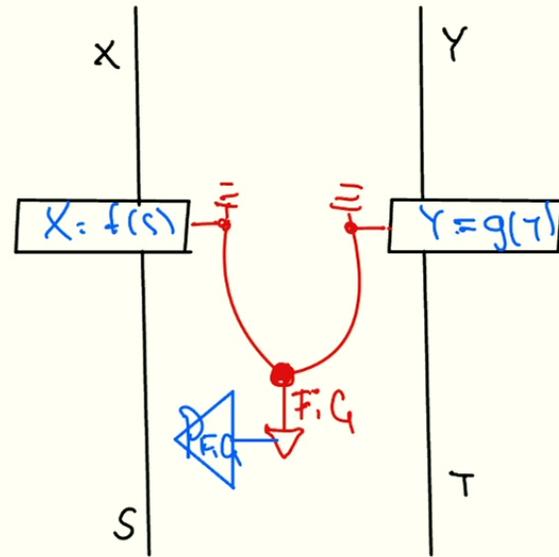
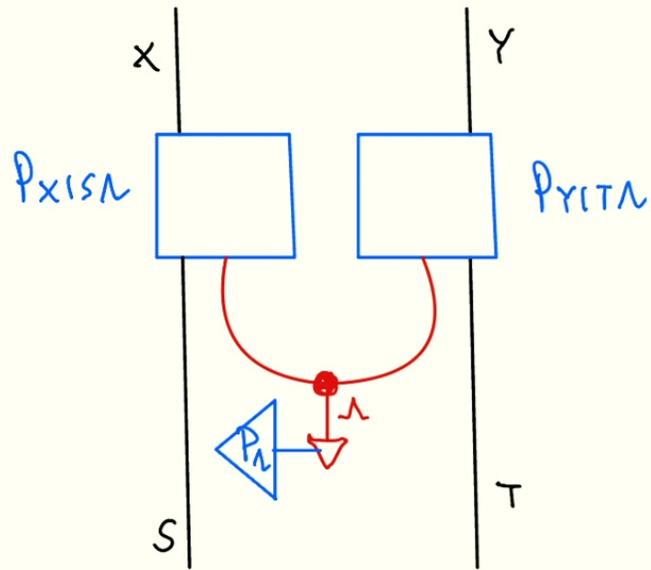
$$P_{X|S\Lambda} = \sum_{\Lambda_X} \delta_{X, f_X(S, \Lambda, \Lambda_X)} P_{\Lambda_X}$$

$$P_{Y|T\Lambda} = \sum_{\Lambda_Y} \delta_{Y, f_Y(T, \Lambda, \Lambda_Y)} P_{\Lambda_Y}$$

## Bell scenario



$$P_{XY|ST} = \sum_{f, f'} \delta_{X, f(S)} \delta_{Y, f'(T)} P_{F_{S \rightarrow X, T \rightarrow Y}}(f, f')$$



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

If X,Y,S,T are binary,  $\Lambda$  can have cardinality 16

$$p_{xy|st} := P_{XY|ST}(xy|st)$$

$$x, y, s, t \in \{0, 1\}$$

$$q_{fg} := P_{F,G}(f, g)$$

$$f, g \in \{I, F, R_0, R_1\}$$

$$p_{00|00} = q_{R_0, R_0} + q_{R_0, I} + q_{I, R_0} + q_{I, I}$$

$$p_{00|01} = q_{R_0, R_0} + q_{R_0, F} + q_{I, R_0} + q_{I, F}$$

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

•  
•  
•

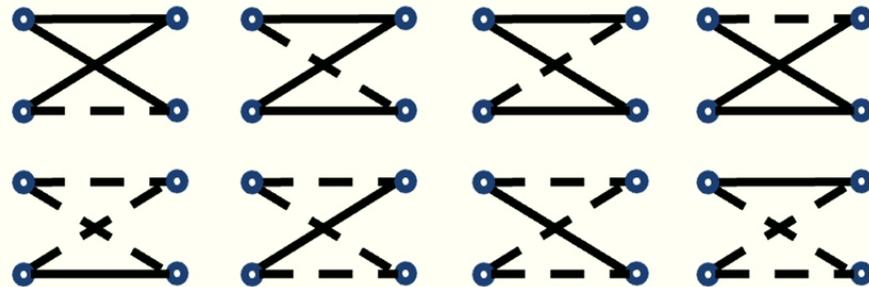
16 linear equalities + inequalities

Do linear quantifier elimination on the 16 q's.

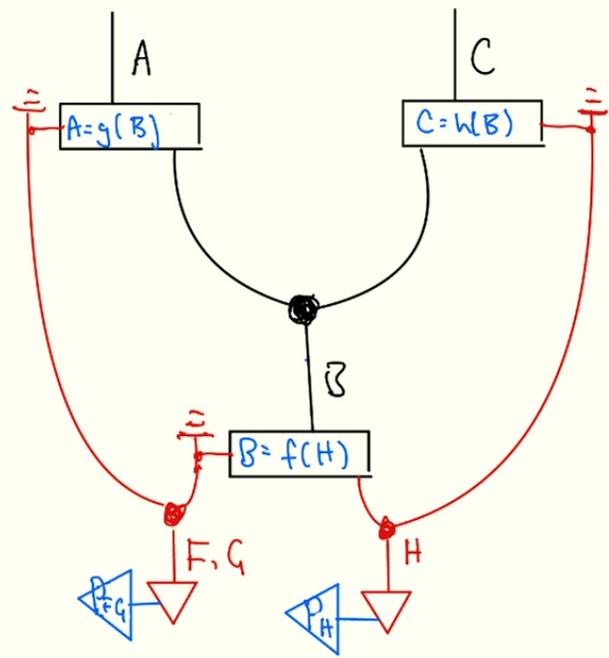
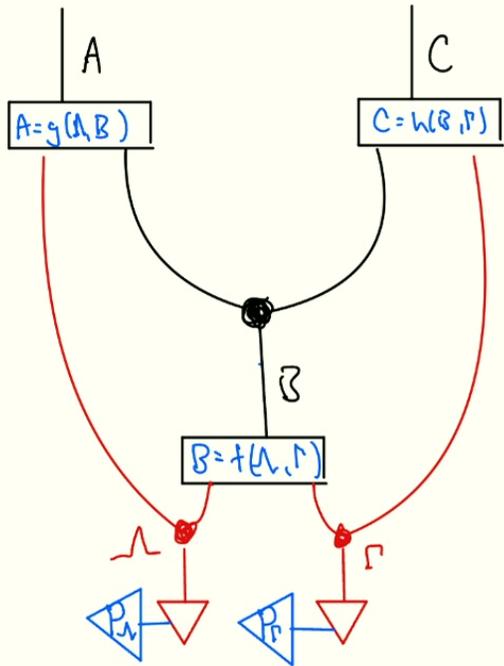
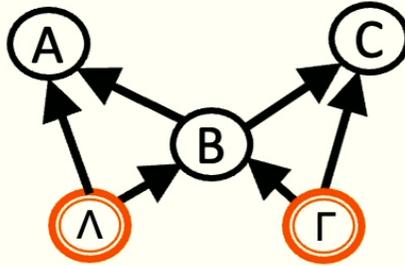
This yields the conditional independence relations and  
the 8 CHSH inequalities

$$\begin{aligned}
 P_{X|ST} &= P_{X|S} \\
 P_{Y|ST} &= P_{Y|T}
 \end{aligned}
 \quad
 \begin{aligned}
 &\frac{1}{4} \sum_{x=y} P_{XY|ST}(xy|00) + \frac{1}{4} \sum_{x=y} P_{XY|ST}(xy|01) \\
 &\frac{1}{4} \sum_{x=y} P_{XY|ST}(xy|10) + \frac{1}{4} \sum_{x \neq y} P_{XY|ST}(xy|11) \leq \frac{3}{4} \\
 &\quad + 7 \text{ others}
 \end{aligned}$$

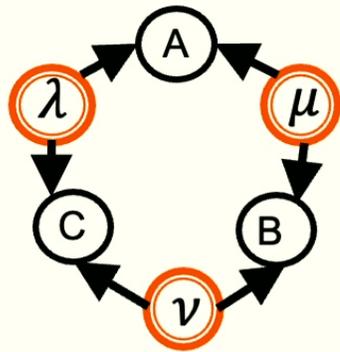
Corresponding to the 8 frustrated four-node networks



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



DAGs for which one can deduce the cardinalities of latent variables in this manner are termed **gearable** (R. Evans, Annals of Statistics, **46**, 2623 (2018))



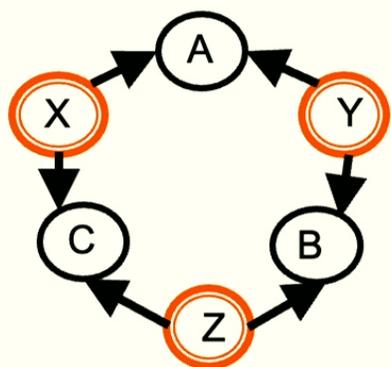
Triangle

The triangle scenario is an example of a DAG that is not gearable

## Techniques for determining upper bounds on cardinalities of the latent variables in more general causal structures

R. Evans, *Annals of Statistics*, 46, 2623 (2018)

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



A, B, C binary →

Sufficient for  
X, Y, Z to be 6-valued

With more than one latent variable,  
we require **nonlinear** quantifier elimination  
which scales badly

# Entropic inequalities by quantifier elimination

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

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

**Def'n: A and B are conditionally independent given C**

$$P_{AB|C} = P_{A|C}P_{B|C}$$

$$P_{B|AC} = P_{B|C}$$

$$P_{A|BC} = P_{A|C}$$

$$P_{ABC} = P_{A|C}P_{B|C}P_C$$

$$I(A : B|C) = 0$$

**Denote this**  
 **$(A \perp B|C)$**

# Entropy vector

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

$$(H(X_1), H(X_2), \dots, H(X_n), H(X_1 X_2), H(X_1 X_3), \dots, H(X_1, X_2, \dots, X_n))$$

# Entropy cone

The closure of this set of vectors is called the  
**entropy cone**

It is a convex cone  
therefore characterized by **linear** inequalities

# 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

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

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

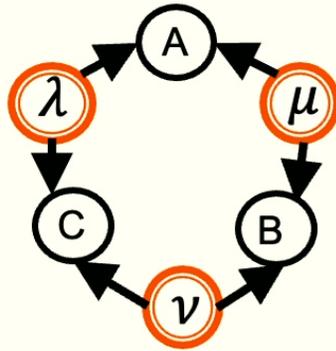
Add the linear inequalities of Shannon-type

Implicitize all entropic quantities **that refer to latent variables**

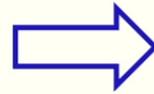
The result is the **marginal Shannon cone**, described by linear inequalities on entropies over observed variables only

## Entropic constraint for the triangle scenario

T. Fritz, 2012



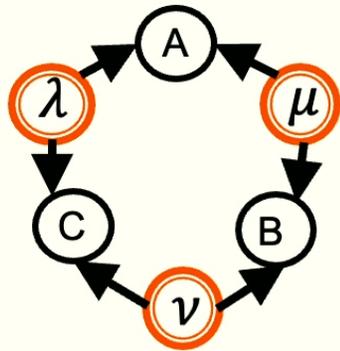
Triangle



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

## Entropic constraint for the triangle scenario

T. Fritz, 2012



Triangle

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

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

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

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

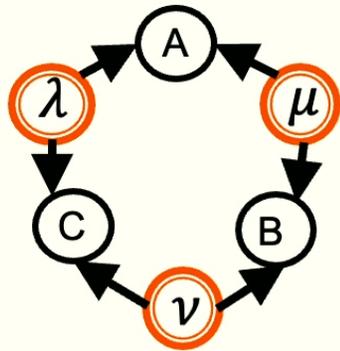
$$A \perp C | \lambda \implies I(A : C | \lambda) = 0 \implies I(A : C) \leq I(A : \lambda)$$

By Shannon-type  
inequalities

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

## Entropic constraint for the triangle scenario

T. Fritz, 2012



Triangle

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

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

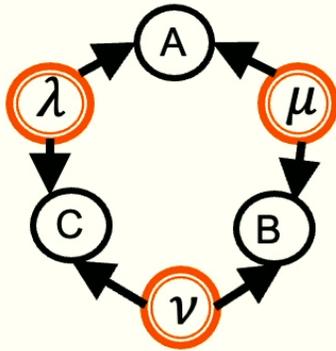
$$B \perp C | \nu \quad \nu \perp \lambda \mu$$

$$\begin{aligned} A \perp B | \mu &\implies I(A : B | \mu) = 0 \implies I(A : B) \leq I(A : \mu) \\ A \perp C | \lambda &\implies I(A : C | \lambda) = 0 \implies I(A : C) \leq I(A : \lambda) \end{aligned} \quad \begin{array}{l} \text{By Shannon-type} \\ \text{inequalities} \end{array}$$

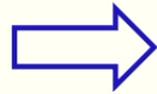
$$\begin{aligned} I(A : B) + I(A : C) &\leq I(A : \mu) + I(A : \lambda) \\ &\leq H(A) + I(\mu : \lambda) \end{aligned} \quad \begin{array}{l} \text{By Shannon-type} \\ \text{inequalities} \end{array}$$

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

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



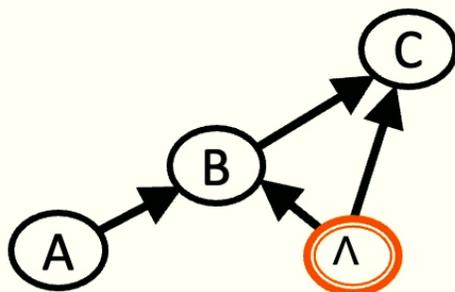
Triangle



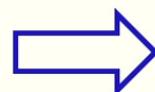
$$I(A : B) + I(A : C) \leq H(A)$$

Note that this inequality detects  
the incompatibility of

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



Instrumental



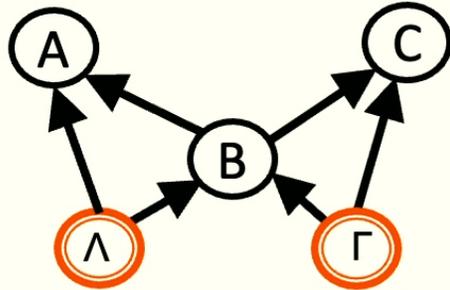
$$I(A : BC) \leq H(B)$$

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

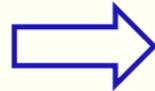
$$P_{ABC}^{\text{pinch}} = \left( \frac{1}{2}[00]_{AC} + \frac{1}{2}[11]_{AC} \right) [0]_B$$

But not our example separating Evans and instrumental

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



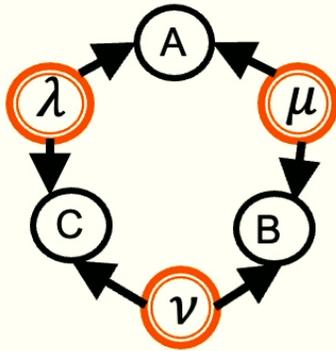
Evans



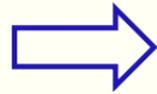
$$I(A : C|B) \leq H(B)$$

Note that this inequality detects the incompatibility of

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



Triangle



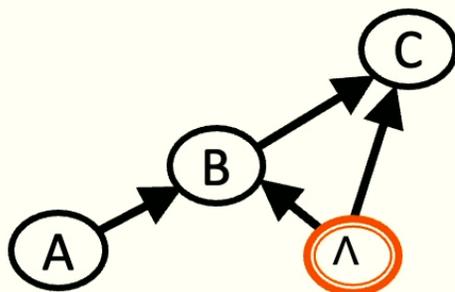
$$I(A : B) + I(A : C) \leq H(A)$$

Note that this inequality detects the incompatibility of

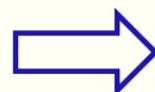
$$P_{ABC} = \frac{1}{2}[000] + \frac{1}{2}[111]$$

But it fails to detect the incompatibility of

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



Instrumental



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

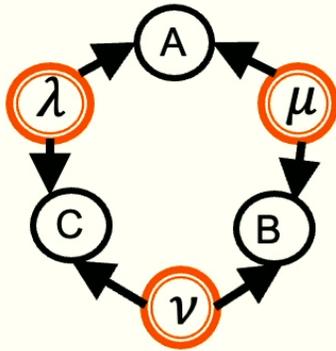
Note that this inequality **also** detects the incompatibility of

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

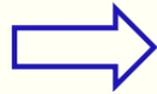
But not our example separating Evans and instrumental

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

(Other examples of the separation **are** detected by the inequality)



Triangle



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

Note that this inequality detects the incompatibility of

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

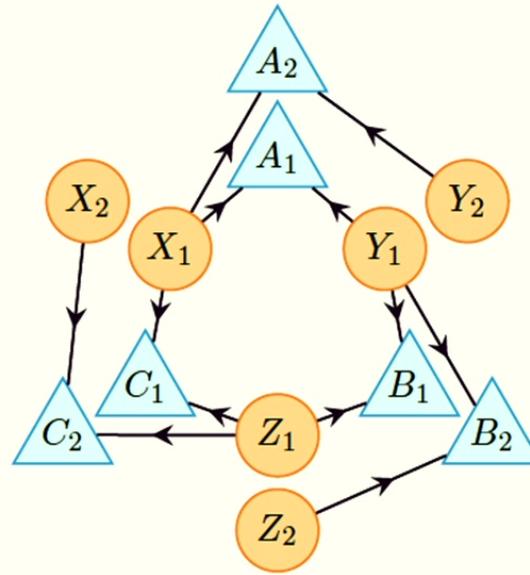
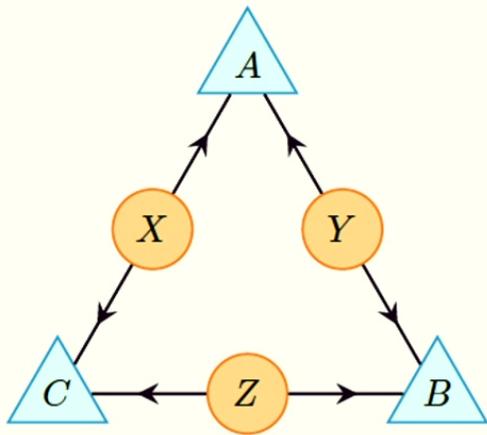
But it fails to detect the incompatibility of

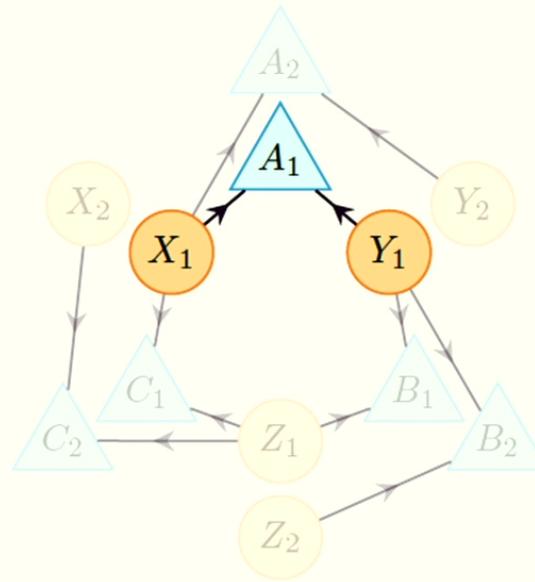
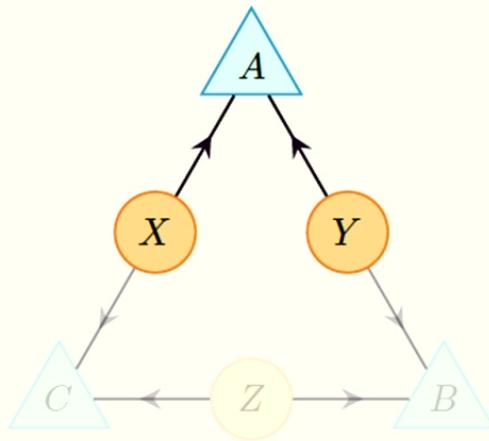
$$P_{ABC} = \frac{1}{3}[001] + \frac{1}{3}[010] + \frac{1}{3}[100]$$

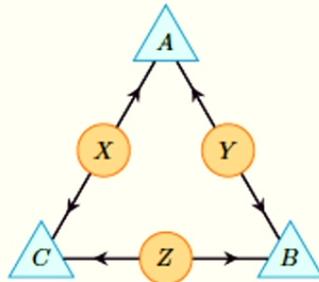
The move to entropies has  
thrown away too much  
information to witness  
certain incompatibilities

# The inflation technique

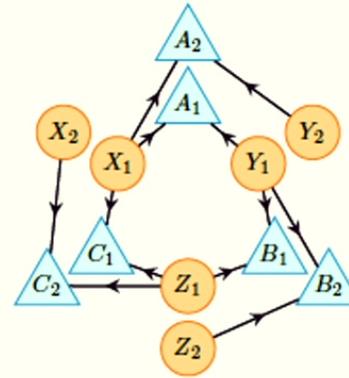




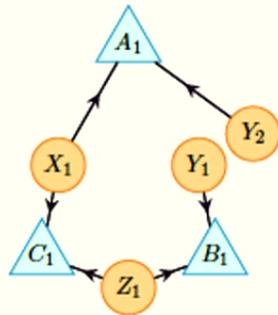




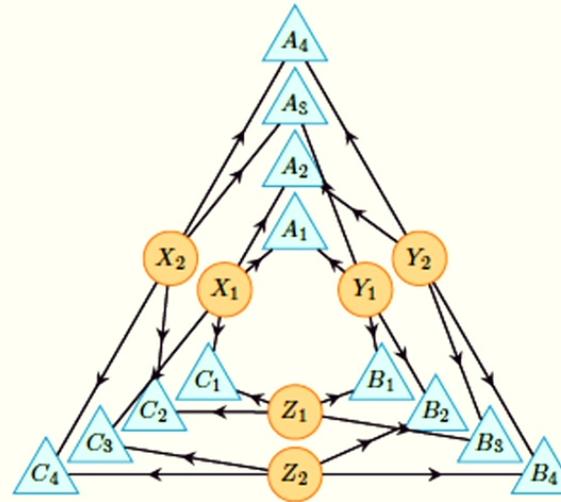
The Triangle Scenario



Spiral Inflation

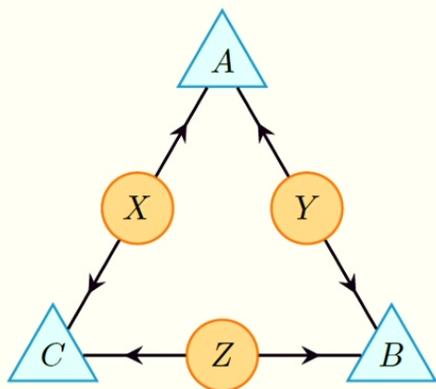


Cut Inflation



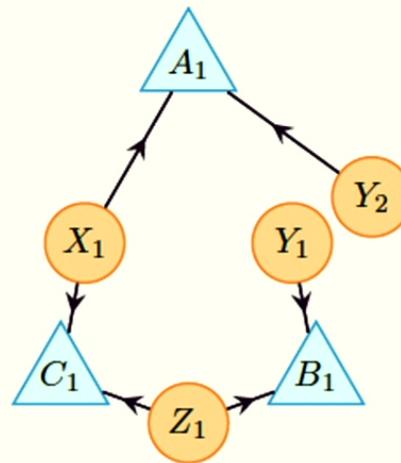
Large Inflation

model M on DAG G



$P_{A|XY}$   
 $P_{B|YZ}$   
 $P_{C|XZ}$   
 $P_X$   
 $P_Y$   
 $P_Z$

$M' = G \rightarrow G'$  Inflation of M

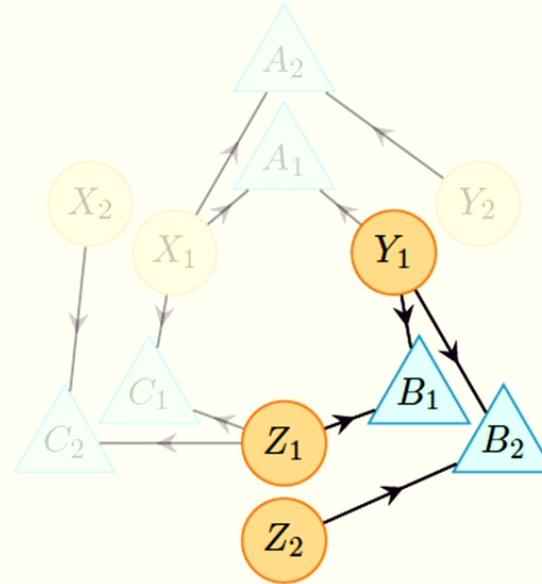
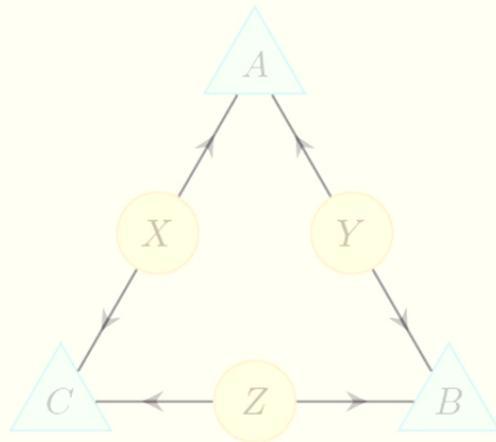


$P_{A_1|X_1Y_2}$   
 $P_{B_1|Y_1Z_1}$   
 $P_{C_1|X_1Z_1}$   
 $P_{X_1}$   
 $P_{Y_1}$   
 $P_{Y_2}$   
 $P_{Z_1}$

with symmetry constraint:

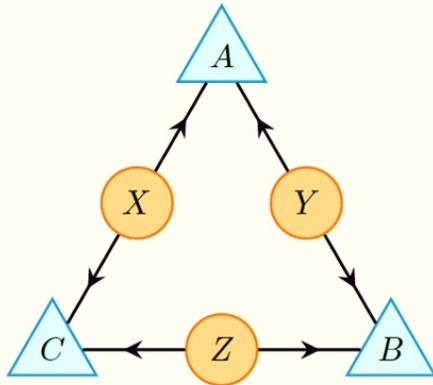
$$P_{Y_1} = P_{Y_2}$$

## **Injectable sets of observed variables in the inflation DAG**



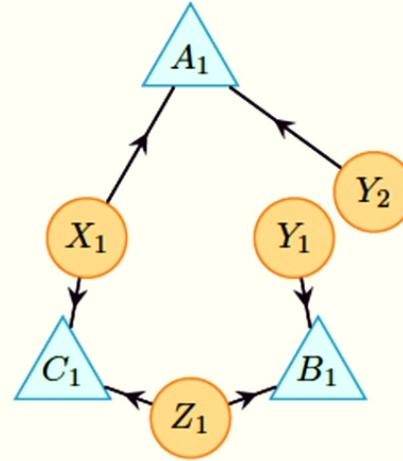
$\{B_1 B_2\}$  is *not* an injectable set

model M on DAG G



- $P_{A|XY}$
- $P_{B|YZ}$
- $P_{C|XZ}$
- $P_X$
- $P_Y$
- $P_Z$

$M' = G \rightarrow G'$  Inflation of M



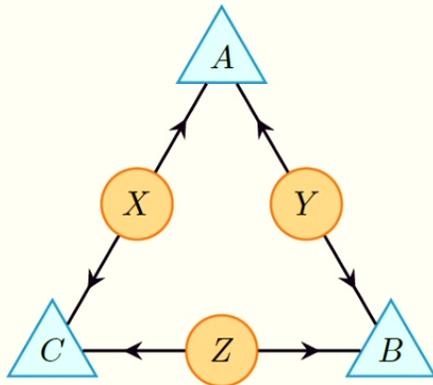
- $P_{A_1|X_1Y_2}$
- $P_{B_1|Y_1Z_1}$
- $P_{C_1|X_1Z_1}$
- $P_{X_1}$
- $P_{Y_1}$
- $P_{Y_2}$
- $P_{Z_1}$

$\{A_1C_1\}$  is an injectable set

$$P_{A_1C_1} = \sum_{X_1Y_2Z_1} P_{A_1|X_1Y_2} P_{C_1|X_1Z_1} P_{X_1} P_{Y_2} P_{Z_1}$$

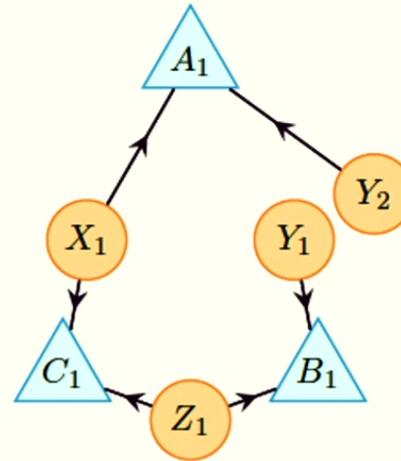
$$P_{AC} = \sum_{XYZ} P_{A|XY} P_{C|XZ} P_X P_Y P_Z$$

model M on DAG G



$P_{A|XY}$   
 $P_{B|YZ}$   
 $P_{C|XZ}$   
 $P_X$   
 $P_Y$   
 $P_Z$

$M' = G \rightarrow G'$  Inflation of M

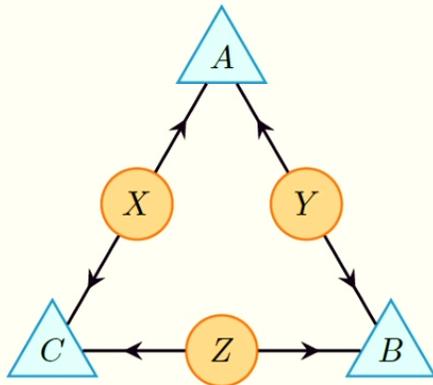


$P_{A_1|X_1Y_2}$   
 $P_{B_1|Y_1Z_1}$   
 $P_{C_1|X_1Z_1}$   
 $P_{X_1}$   
 $P_{Y_1}$   
 $P_{Y_2}$   
 $P_{Z_1}$

$\{A_1B_1\}$  is *not* an injectable set

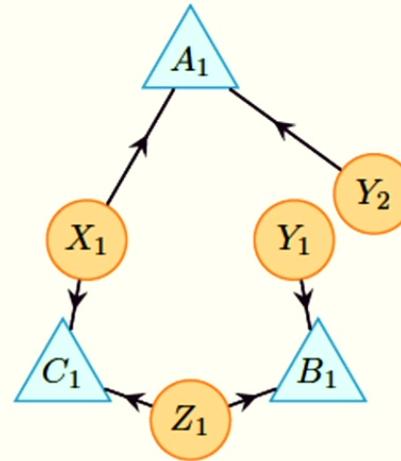
$$P_{A_1B_1} = \left( \sum_{X_1Y_2} P_{A_1|X_1Y_2} P_{Y_2} P_{X_1} \right) \left( \sum_{Z_1Y_1} P_{B_1|Y_1Z_1} P_{Y_1} P_{Z_1} \right)$$

model  $M$  on DAG  $G$



$P_{A|XY}$   
 $P_{B|YZ}$   
 $P_{C|XZ}$   
 $P_X$   
 $P_Y$   
 $P_Z$

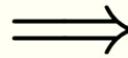
$M' = G \rightarrow G'$  Inflation of  $M$



$P_{A_1|X_1Y_2}$   
 $P_{B_1|Y_1Z_1}$   
 $P_{C_1|X_1Z_1}$   
 $P_{X_1}$   
 $P_{Y_1}$   
 $P_{Y_2}$   
 $P_{Z_1}$

Injectable sets:  $\{A_1\}, \{B_1\}, \{C_1\}, \{A_1C_1\}, \{B_1C_1\}$

$(P_A, P_B, P_C, P_{AC}, P_{BC})$   
 compatible with  $M$



$(P_{A_1}, P_{B_1}, P_{C_1}, P_{A_1C_1}, P_{B_1C_1})$   
 compatible with  $M'$

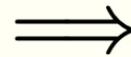
where  $P_{A_1} = P_A$      $P_{A_1C_1} = P_{AC}$   
 $P_{B_1} = P_B$      $P_{B_1C_1} = P_{BC}$   
 $P_{C_1} = P_C$

$M' = G \rightarrow G'$  Inflation of  $M$

$\mathcal{S} \subseteq \text{ImagesInjectableSets}(G)$

$\mathcal{S}' \subseteq \text{InjectableSets}(G')$

$\{P_{\mathbf{V}} : \mathbf{V} \in \mathcal{S}\}$   
is compatible with  $M$



$\{P_{\mathbf{V}'} : \mathbf{V}' \in \mathcal{S}'\}$   
where  $P_{\mathbf{V}'} = P_{\mathbf{V}}$  for  $\mathbf{V}' \sim \mathbf{V}$   
is compatible with  $M'$

# Deriving causal compatibility inequalities by the inflation technique

$M' = G \rightarrow G'$  Inflation of  $M$

$\mathcal{S} \subseteq \text{ImagesInjectableSets}(G)$

$\mathcal{S}' \subseteq \text{InjectableSets}(G')$

$I_{\mathcal{S}'}$  is a **causal compatibility inequality** for model  $M'$

$\{P_{\mathbf{V}} : \mathbf{V} \in \mathcal{S}\}$   
is compatible with  $M$

$\implies$

$\{P_{\mathbf{V}'} : \mathbf{V}' \in \mathcal{S}'\}$   
where  $P_{\mathbf{V}'} = P_{\mathbf{V}}$  for  $\mathbf{V}' \sim \mathbf{V}$   
is compatible with  $M'$

$\Downarrow$

$I_{\mathcal{S}'}$  is **satisfied** for  
 $\{P_{\mathbf{V}'} : \mathbf{V}' \in \mathcal{S}'\}$   
where  $P_{\mathbf{V}'} = P_{\mathbf{V}}$  for  $\mathbf{V}' \sim \mathbf{V}$

$M' = G \rightarrow G'$  Inflation of  $M$

$\mathcal{S} \subseteq \text{ImagesInjectableSets}(G)$        $\mathcal{S}' \subseteq \text{InjectableSets}(G')$

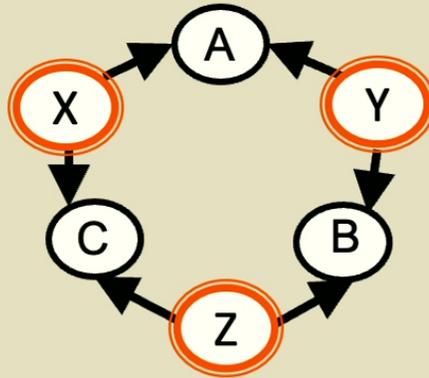
$I_{\mathcal{S}'}$  is a **causal compatibility inequality** for model  $M'$

$\{P_{\mathbf{V}} : \mathbf{V} \in \mathcal{S}\}$

is compatible with  $M$

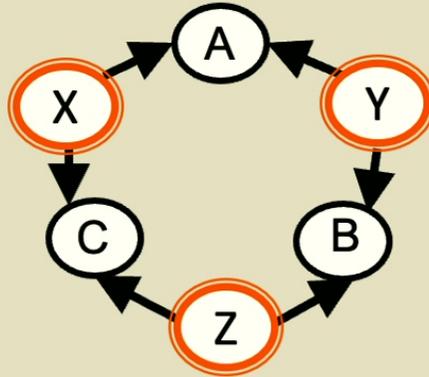


$I_{\mathcal{S}}$  is **satisfied** for  
 $\{P_{\mathbf{V}} : \mathbf{V} \in \mathcal{S}\}$



$$P_A + P_B + P_C - P_A P_B - P_{BC} - P_{AC} \leq 1$$

causal compatibility inequality

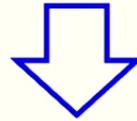


$$P_A + P_B + P_C - P_A P_B - P_{BC} - P_{AC} \leq 1$$

causal compatibility inequality

Consider binary X, Y and Z

$\exists P_{XYZ}$  with  $P_X, P_Y, P_Z, P_{XY}, P_{YZ}, P_{XZ}$  as marginals



Marginal compatibility constraints

$$P_X = \sum_Y P_{XY} = \sum_Z P_{XZ}$$

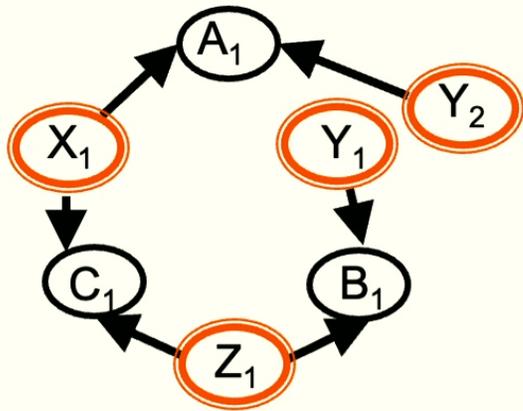
$$P_Y = \sum_X P_{XY} = \sum_Z P_{YZ}$$

$$P_Z = \sum_X P_{XZ} = \sum_Y P_{YZ}$$

$$P_X + P_Y + P_Z - P_{XY} - P_{YZ} - P_{XZ} \leq 1$$

$(P_{A_1 C_1}, P_{B_1 C_1}, P_{A_1 B_1})$  is a valid set of marginals  $\implies (P_{A_1 C_1}, P_{B_1 C_1}, P_{A_1 B_1})$  satisfy  $P_{A_1} + P_{B_1} + P_{C_1} - P_{A_1 B_1} - P_{B_1 C_1} - P_{A_1 C_1} \leq 1$

**Linear quantifier elimination**

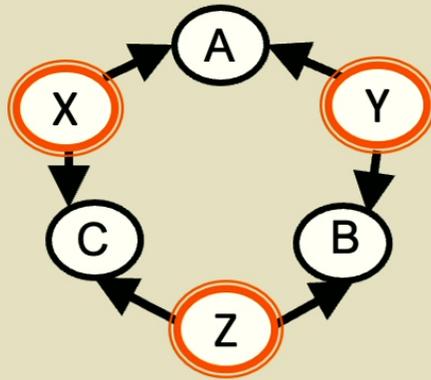


$(P_{A_1 C_1}, P_{B_1 C_1}, P_{A_1 B_1})$  is compatible with  $M'$

$$A_1 \perp_d B_1 \implies P_{A_1 B_1} = P_{A_1} P_{B_1}$$

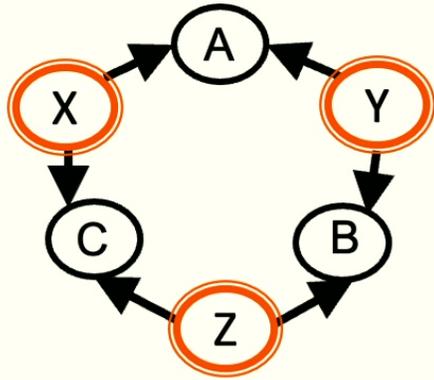
$(P_{A_1 C_1}, P_{B_1 C_1}, P_{A_1 B_1})$  is compatible with  $M'$   $\implies P_{A_1} + P_{B_1} + P_{C_1} - P_{A_1} P_{B_1} - P_{B_1 C_1} - P_{A_1 C_1} \leq 1$

This is a causal compatibility inequality for  $M'$

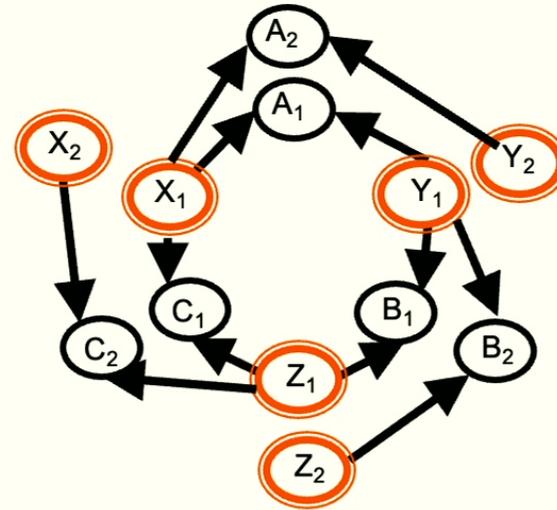


$$\begin{aligned} &P_A(1)P_B(1)P_C(1) \\ &\leq P_{AB}(11)P_C(1) + P_{BC}(11)P_A(1) \\ &+ P_{AC}(11)P_B(1) + P_{ABC}(000) \end{aligned}$$

**causal compatibility inequality**



$$\begin{aligned}
 &P_A(1)P_B(1)P_C(1) \\
 &\leq P_{AB}(11)P_C(1) + P_{BC}(11)P_A(1) \\
 &+ P_{AC}(11)P_B(1) + P_{ABC}(000)
 \end{aligned}$$



$$\begin{aligned}
 &P_{A_2}(1)P_{B_2}(1)P_{C_2}(1) \\
 &\leq P_{A_1B_2}(11)P_{C_2}(1) + P_{B_1C_2}(11)P_{A_2}(1) \\
 &+ P_{A_2C_1}(11)P_{B_2}(1) + P_{A_1B_1C_1}(000)
 \end{aligned}$$

**Polynomial inequality constraints** for causal compatibility with the original DAG



**Linear inequality constraints** from marginal compatibility  
(from linear quantifier elimination)

+

**Polynomial equality constraints** from causal compatibility with the inflated DAG  
(e.g., from d-separation relations)

The technique defines an algorithm for deriving causal compatibility inequalities and for testing compatibility

Proof that this provides a convergent hierarchy of tests:  
Navascués & Wolfe, J. Causal Inf. 8(1) 70 (2020)

Approaches to Bell arguments that follow essentially  
the logic of the inflation technique:

Fine's proof of CHSH inequalities

Hardy's proof of Bell's theorem

The Greenberger-Horne-Zeilinger proof of Bell's theorem

Bell's theorem by way of Kochen-Specker noncontextuality

Braunstein-Caves entropic inequalities

Symmetric extensions / shareability of local correlations

...

**Polynomial inequality constraints** for causal compatibility with the original DAG



**Linear inequality constraints** from marginal compatibility  
(from linear quantifier elimination)

+

**Polynomial equality constraints** from causal compatibility with the inflated DAG  
(e.g., from d-separation relations)

The technique defines an algorithm for deriving causal compatibility inequalities and for testing compatibility

Proof that this provides a convergent hierarchy of tests:  
Navascués & Wolfe, J. Causal Inf. 8(1) 70 (2020)

Approaches to Bell arguments that follow essentially  
the logic of the inflation technique:

Fine's proof of CHSH inequalities

Hardy's proof of Bell's theorem

The Greenberger-Horne-Zeilinger proof of Bell's theorem

Bell's theorem by way of Kochen-Specker noncontextuality

Braunstein-Caves entropic inequalities

Symmetric extensions / shareability of local correlations

...

Approaches to Bell arguments that follow essentially  
the logic of the inflation technique:

Fine's proof of CHSH inequalities

Hardy's proof of Bell's theorem

The Greenberger-Horne-Zeilinger proof of Bell's theorem

Bell's theorem by way of Kochen-Specker noncontextuality

Braunstein-Caves entropic inequalities

Symmetric extensions / shareability of local correlations

...