Title: Which causal scenarios might support non-classical correlations?

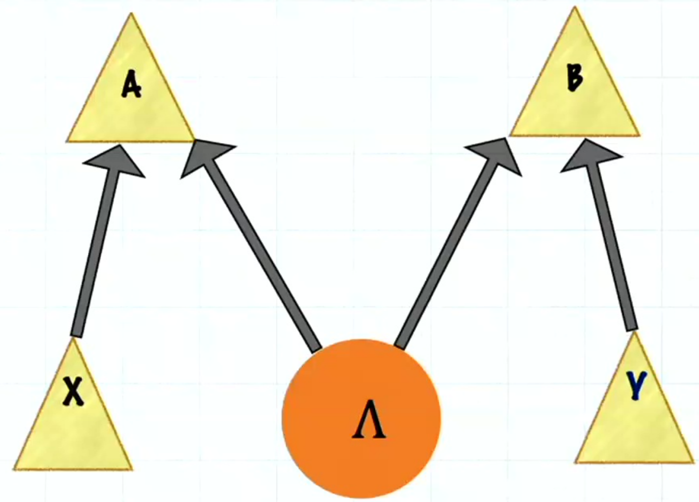
Speakers: Shashaank Khanna

Series: Quantum Foundations, Quantum Information

Date: September 16, 2024 - 2:35 PM

URL: https://pirsa.org/24090109

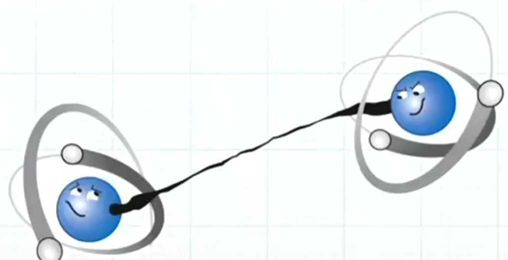
Which causal scenarios might support Non-Classical correlations?



- Shashaank Khanna, Marina Maceil Ansanelli, Matthew F. Pusey, Elie Wolfe

PERIMETER INSTITUTE FOR THEORETICAL PHYSICS

UNIVERSITY of York



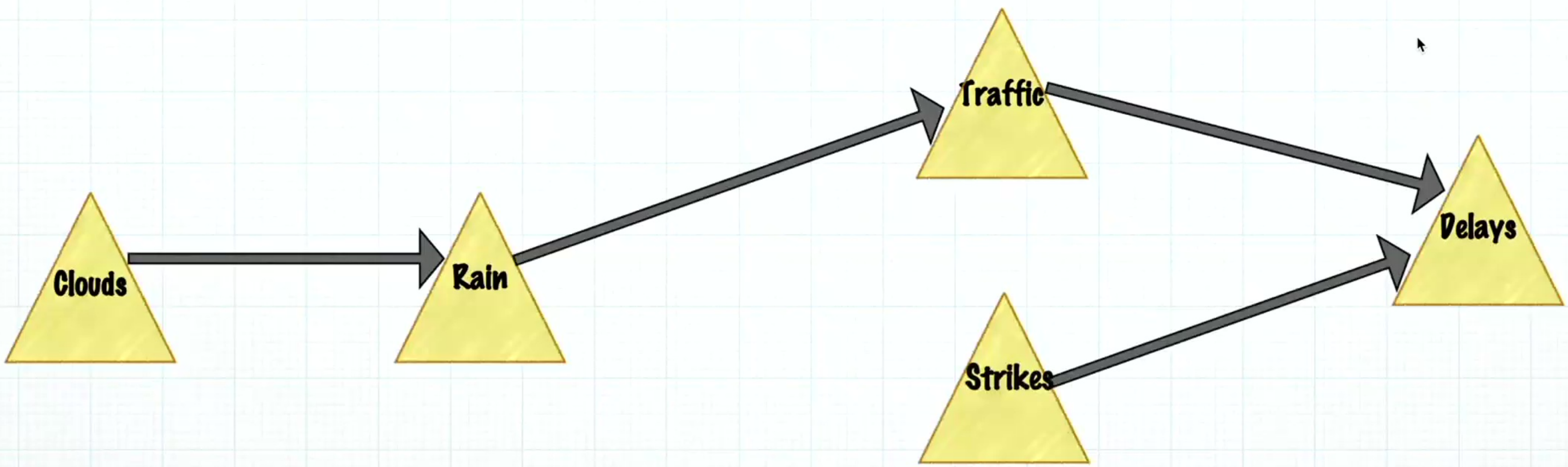
What are causal scenarios (DAGs) ?

Generalized way to represent cause and effect relations among observed events .

Events modelled as random variables.

No directed cycles -> Directed Acyclic Graphs (DAGs)

Edges indicate direct causation.



Causal Markov condition for DAGs

If a probability distribution P over the variables in a DAG G can be factorised as:

$$P(x_1, \ldots x_n) = \prod_i P(x_i | PA_G(x_i))$$

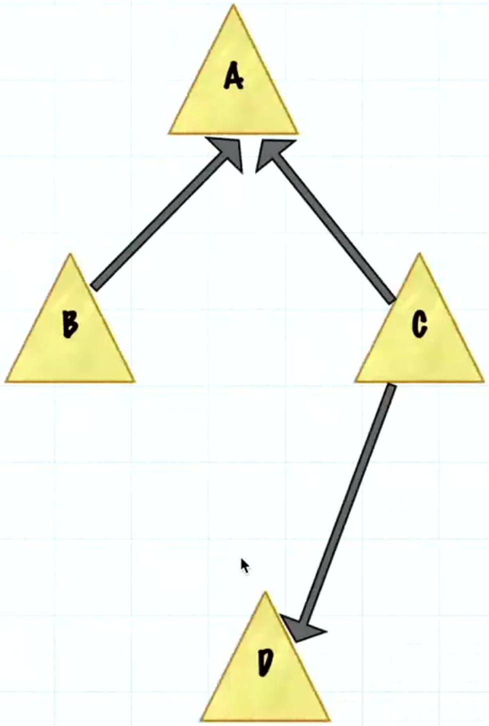
$PA_G(x_i)$ -> parents of x_i in G,

then P is Markov with respect to G

and G is a classically causal explanation of P.

Notion of d-separation in DAGs

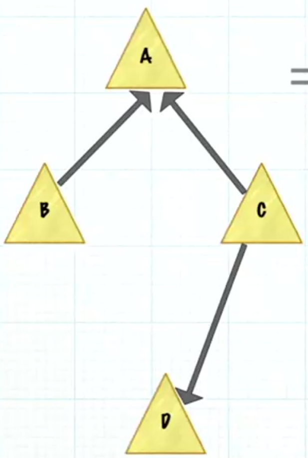
d-separation -> a graphical condition to read off conditional independences.



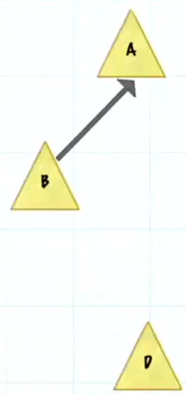
Towards e-separation ?

If two sets A, B are d-separated by Z after deletion of a set of nodes W in the graph then A and B are e-separated by Z.

For eg:

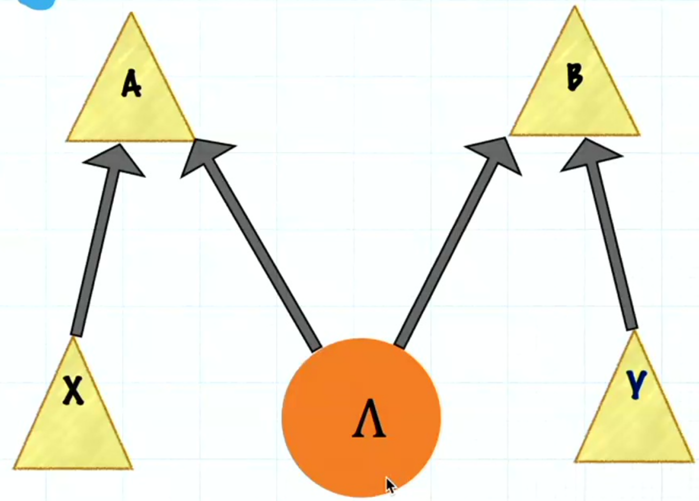


$\implies A \perp_d D$ is false but after deletion of C we have



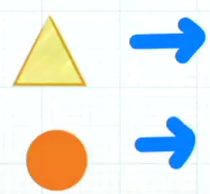
$\implies A \perp_d D$ is true

Bell's Theorem recast using DAGs



The causal Markov condition for the Bell DAG encodes the notion of Local Causality.

$$P(A,B,X,Y) = \sum_{\Lambda} P(A|X,\Lambda)P(B|Y,\Lambda)P(X)P(Y)P(\Lambda)$$



Observed Variables

Latent Variables

Different theories allow different types of distributions !

$C = \{P(x_1 \ldots x_n): P\ follows\ Causal\ Markov\ condition\}$

$Q = \{P(x_1 \ldots x_n): P\ can\ be\ obtained\ from\ Quantum\ theory\ by\ Born\ rule\}$

$G = \{P(x_1 \ldots x_n): P\ can\ be\ obtained\ from\ Generalized\ Probabilistic\ Theories\}$

Different theories allow different types of distributions !

$C = \{P(x_1 \ldots x_n): P\ follows\ Causal\ Markov\ condition\}$

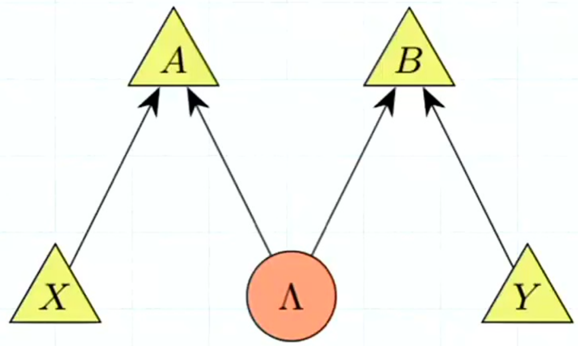
$Q = \{P(x_1 \ldots x_n): P\ can\ be\ obtained\ from\ Quantum\ theory\ by\ Born\ rule\}$

$G = \{P(x_1 \ldots x_n): P\ can\ be\ obtained\ from\ Generalized\ Probabilistic\ Theories\}$

$I = \{P(x_1, \ldots. x_n): P\ respects\ all\ observed\ conditional\ independences\}$

Quantum vs Classical: Allowed Probabilities

For Bell DAG:



$$C = \{P(A,B,X,Y): P(A,B,X,Y) = \sum_{\Lambda} P(A|X,\Lambda)P(B|Y,\Lambda)P(X)P(Y)P(\Lambda)\}$$

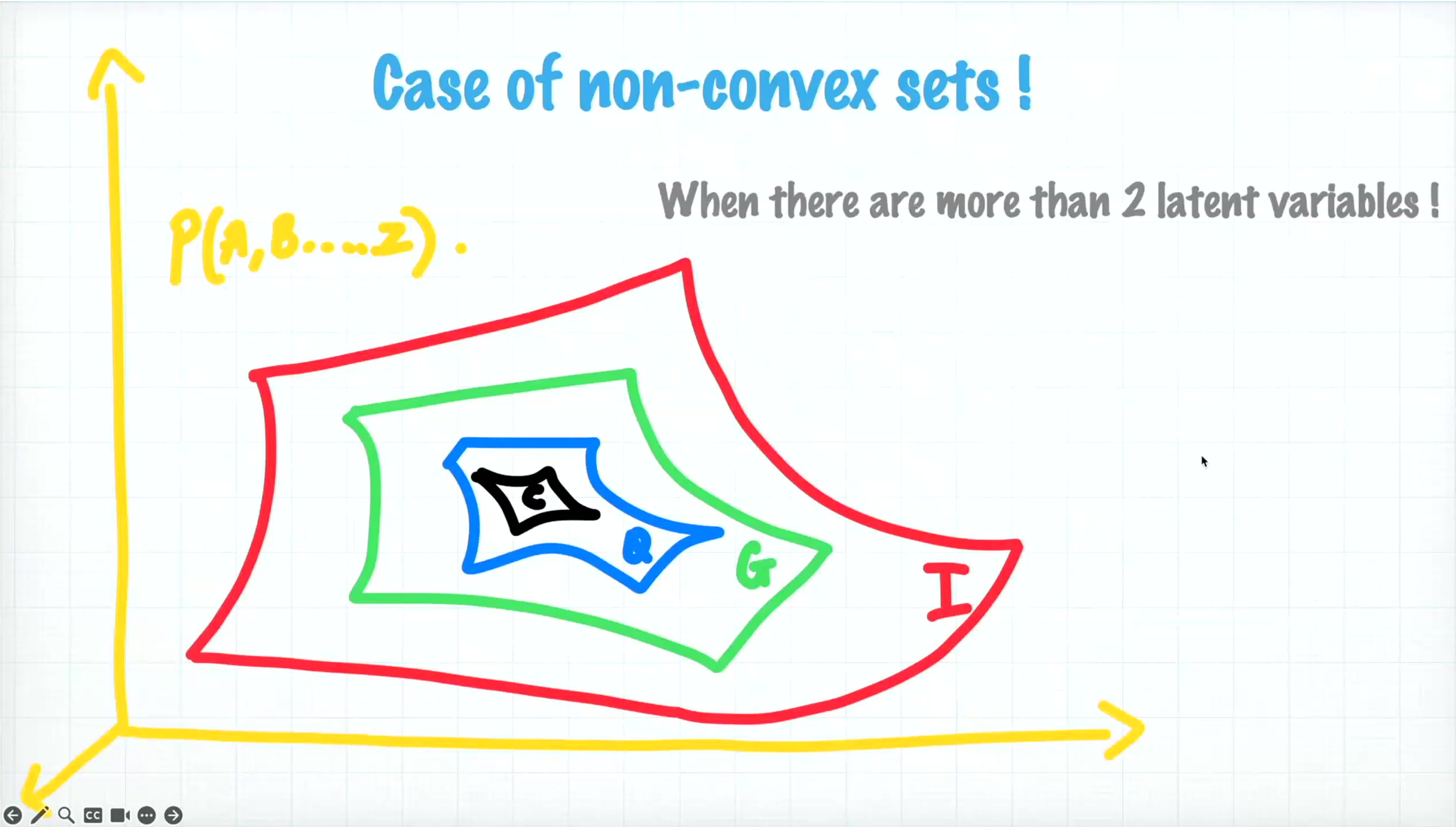
$$Q = \{P(A,B,X,Y): P(A,B,X,Y) = tr[(E_X^A \otimes E_Y^B)\rho_{\Lambda_{AB}}]P(X)P(Y)\}$$

What happens when there are no latent variables in the DAG ?

For a DAG, G, without latent variables, a probability distribution P is Markov with respect to G **if and only if** P satisfies all the observed d-separation relations.

Hence for a latent free DAG,

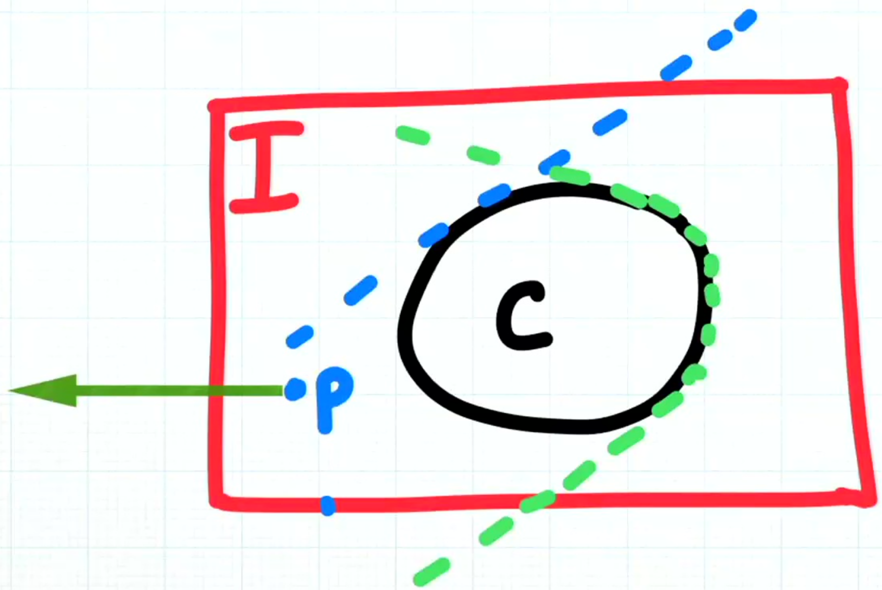
$$C_{LF} = Q_{LF} = G_{LF} = I_{LF}$$


Case of non-convex sets !
When there are more than 2 latent variables !
P(A,B.....Z) .
C
Q
G
I

"Interesting (Non-Algebraic) DAGs"

Only those DAGs which have $C \subset I$ can possibly support "Non-Classical" correlations and are termed "Interesting" or "Non-Algebraic" Otherwise they are "Non-interesting" or "Algebraic".

Non-Classical Probability Distribution



Henson, Lal and Pusey (HLP): Sufficient condition for "non-interestingness"

* Provided a series of graphical transformations which when met were proof of "non-interestingness".
* When not met the DAG could be "interesting" or not.
* Characterized all DAGs up to 6 nodes as "interesting" or not.
* Couldn't characterise DAGs of 7, 8.. nodes

HLP Conjecture !

That these transformations introduced by HLP are both sufficient and necessary to certify "non-interestingness".

That is,

If using these transformations and nothing more one can get an mDAG that is "non-interesting", then the original mDAG is "non-interesting" as well, otherwise it is "interesting".

Introduction to mDAGs

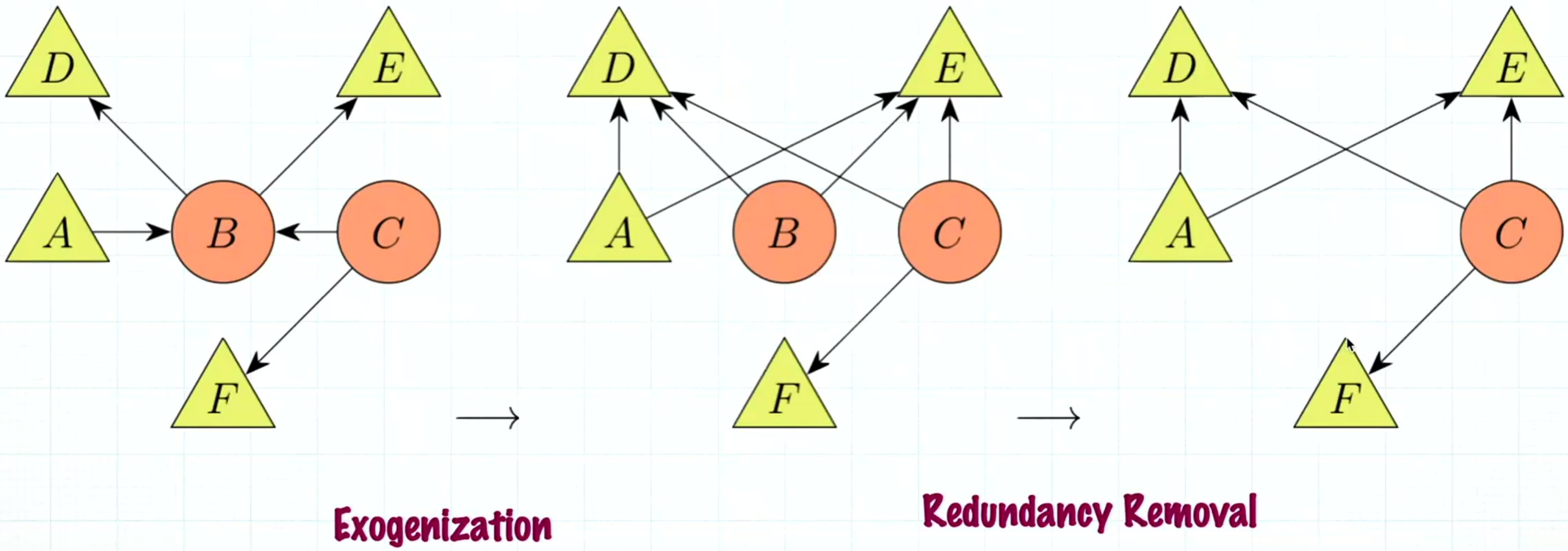
1. **Exogenization:** In a DAG G, with set of latent nodes $\{\lambda_i\}$, $\forall\lambda_i$ add edges $m \rightarrow n \forall m \in PA_G(\lambda_i)$ to every $n \in CH_G(\lambda_i)$ and delete the edges $m \rightarrow \lambda_i \forall m \in PA_G(\lambda_i)$

2. **Redundancy Removal:** Delete all latent variables λ_i for which $CH_G(\lambda_i) \subseteq CH_G(\lambda_j)$ where λ_j is another latent variable s.t $\lambda_i \neq \lambda_j$ and $PA_G(\lambda_i) = PA_G(\lambda_j) = \phi$

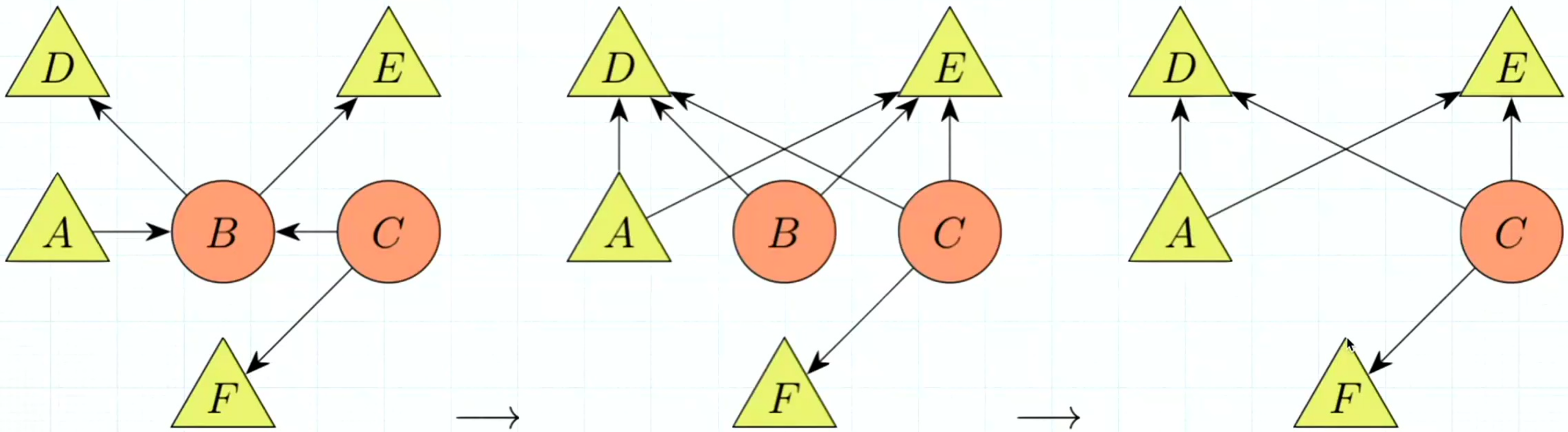
These lead to another DAG G' s.t $C_G = C_{G'}$

G' will be called an mDAG. : Graphs for Margins of Bayesian Networks (Evans 2016)

Example of an mDAG transformation



Example of an mDAG transformation



Exogenization

Redundancy Removal

C_G = $C_{G'}$

Evan's result on mDAGs

Any mDAG, G is "non-interesting" if and only if $\exists$ another mDAG H that does not have any latent variables and for which $C_G = C_H$.

Because for the if part we have,

$C_G \subseteq I_G$ and $C_G = C_H = I_H$ where $C_G = C_H \Longrightarrow I_G = I_H$

and thus, $C_G = I_G$

For the only if part refer: Latent Free Equivalent mDAGS, Evans(2023)

Can we find other graphical conditions ?

Yes, we can !

Maximality,

d-separation,

e-separation,

Infeasible supports of probability distributions

Can we find other graphical conditions ?

Yes, we can !

Maximality,

d-separation,

e-separation,

Infeasible supports of probability distributions

$\Longrightarrow$ They show "Interestingness"

Using d-separation to certify "interestingness"

If an mDAG G has a set of observed d-separation relations that cannot be produced by ANY latent free DAG, then G is "interesting".

Proof: $C_G = C_H \Longrightarrow I_G = I_H$, the contrapositive leading to

$$I_G \neq I_H \Longrightarrow C_G \neq C_H \quad \forall \text{ possible latent free } H$$

Hence by Evan's result G is "interesting".

Using e-separation to certify "interestingness"

Firstly, if for any 2 mDAGs, G and H, $C_G = C_H$ then their sets of observed e-separation relations must be identically the same (just like for d-separation).

If the observed e-separation relations in a mDAG, G cannot be reproduced by ANY latent free mDAG H, then G is "interesting" (again by Evan's result).

Supports of a probability distribution

Given a probability distribution $P(X_1, \ldots. X_n)$ its support is defined as:

$$S(P(X_1, \ldots X_n)) = \{\{x_1, \ldots x_n\} | P(X_1 = x_1, \ldots. X_n = x_n) \geq 0\}$$

If there exists a $P \in C_G$ s.t $S(P) = S$, where S is a set of events, then we say that S is **classical** w.r.t G.

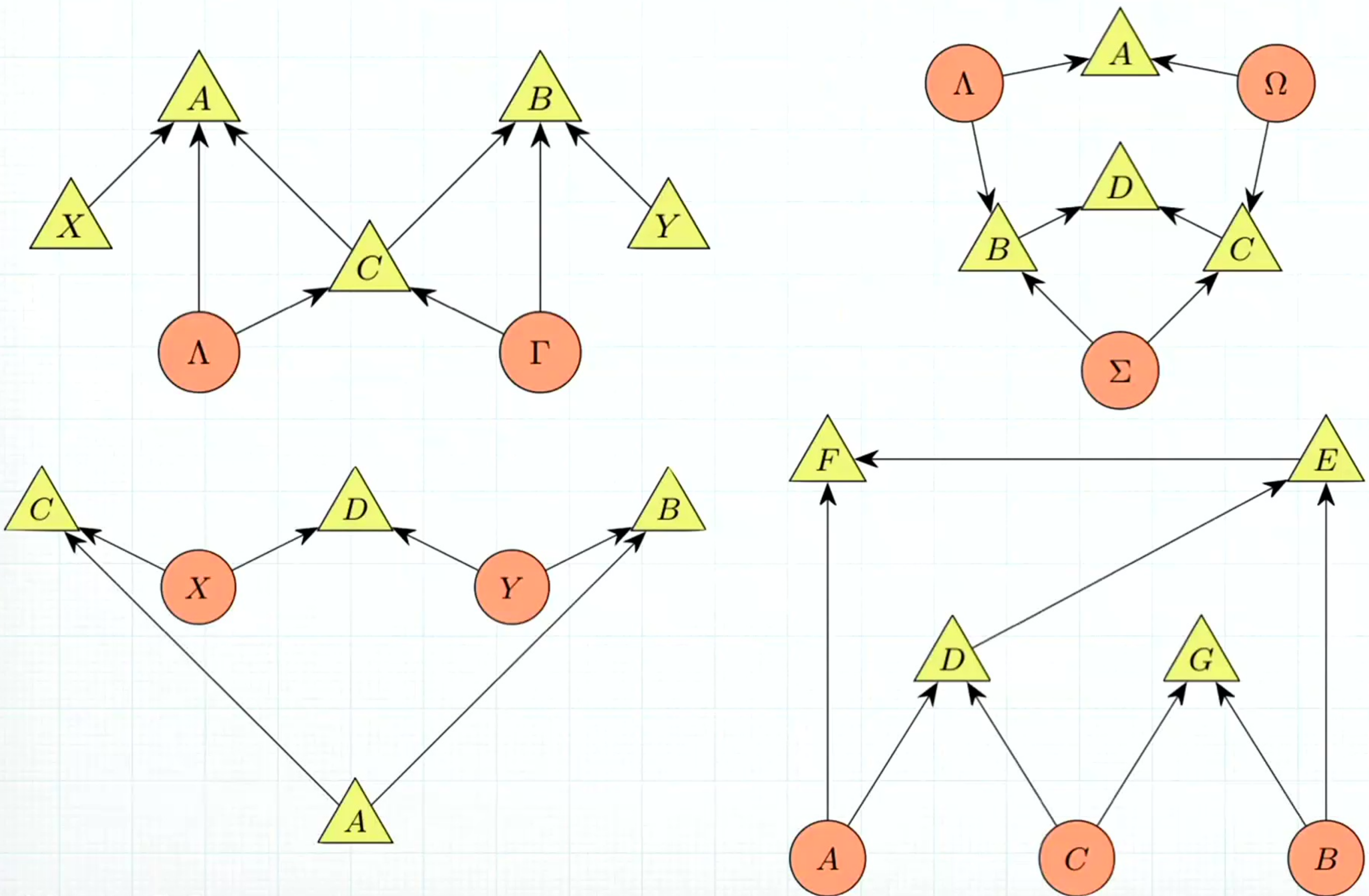
If there exists a $P \in I_G$ s.t $S(P) = S$, where S is a set of events, then we say that S is **classical-up-to-observed conditional independences** w.r.t G.

Classically infeasible supports for "interestingness"

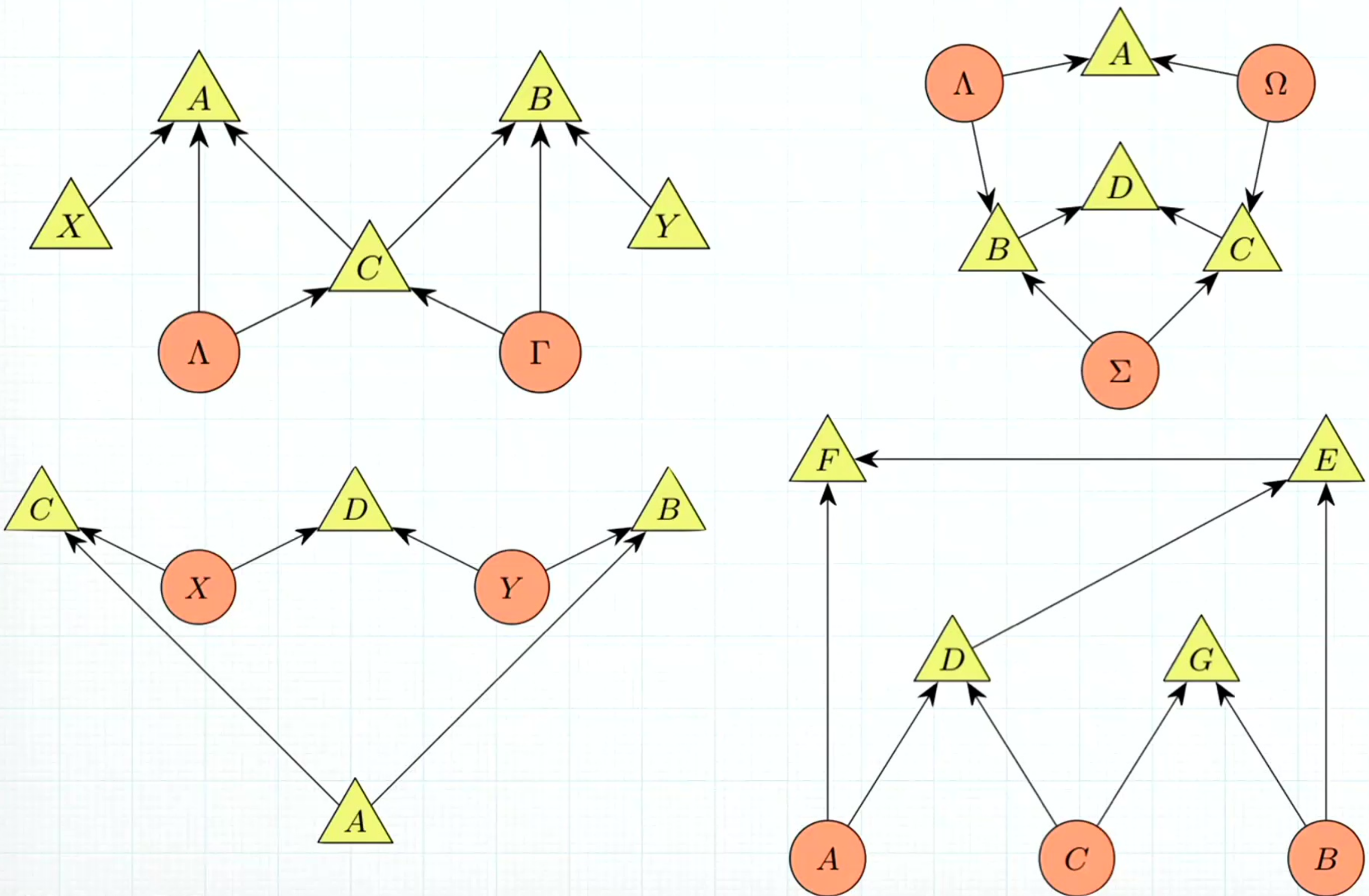
If two mDAGs G and H s.t $C_G = C_H$ then their sets of classical supports must be identical (unknown if this could be only-if as well).

If an mDAG, G has a set of classical supports that cannot be reproduced in ANY latent free mDAG, then G is "interesting" (by invoking Evan's result).

Some "interesting" DAGs we found



Some "interesting" DAGs we found

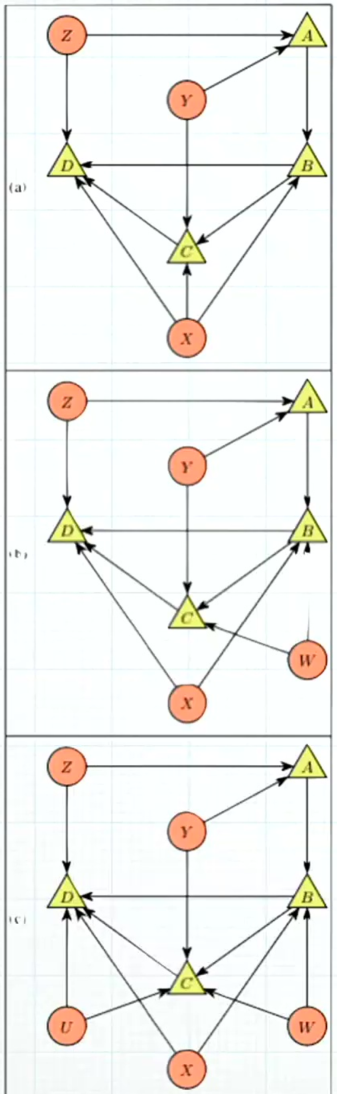


We not only find that they are "interesting" but find the exact probability distributions that are non-classical !

Computational Results

Category	DAGs with 3 observed nodes	DAGs with 4 observed nodes	DAGs with 5 observed nodes
Total Count of DAGs	46	2809	1,718,596
DAGs remaining after HLP condition (since it is only a sufficient condition)	5	996	1,009,961
DAGs remaining after various graphical criteria, like Maximality, d-separation, e-separation, Infeasible supports of Probability distributions…..	0	3	< 12,834

$\approx$ 99% reduction of uncharacterised DAGs $\Longrightarrow$ HLP condition looks to be necessary as well !



3 unclassified mDAGs of 4 observed nodes

Shannon cones corresponding to sets C and I are the same for these 3 mDAGs, so no difference can be found at the level of Shannon entropic inequalities.

What to do-: Explore Non Shannon type inequalities or accelerate supports algorithm to solve these 3.

Summary and Future work

- Evidence towards HLP condition being necessary as well.
- Several graphical criteria to check "interestingness".
- Explicit construction of "Non-Classical" distributions.
- These scenarios can exhibit classical-quantum or post quantum gap.
- Potential candidates for exhibiting quantum or post quantum advantage.
- Importance for classical causal inference (in ML, AI)
- Attacking specific scenarios to confirm classical-quantum advantage.

Thank You !