Griffith UNIVERSITY
Centre for Quantum Dynamics

Classical Causal Models cannot faithfully explain Bell nonlocality or Kochen-Specker contextuality in arbitrary scenarios

Jason C. Pearl (joint work with Eric G. Cavalcanti)

Griffith University

Gold Coast, Queensland, Australia

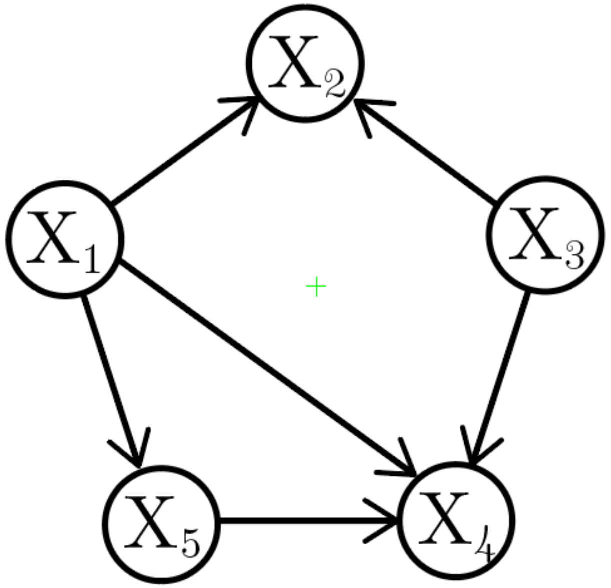
QCQMB’21 workshop

18 May 2021

J. Pearl, *Causality: Models, Reasoning and Inference* (Cambridge University Press, 2000).

Directed Acyclic Graphs (DAG)

Causal structure → Directed Acyclic graph (DAG)



Nodes: Random variables

Arrows: Causal links

Parents of X_4: Set of direct causes of X_4

$$Pa(X_4) = \{X_1, X_3, X_5\}$$

Descendents of X_3: Set of effects of X_3

$$De(X_3) = \{X_2, X_4\}$$

Non-Descendents of X_3: Set of non-effects of X_3

$$Nd(X_3) = \{X_1, X_5\}$$

Causal Markov Condition

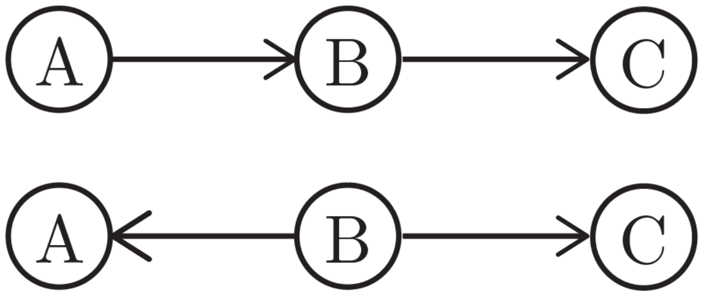
$$(X \perp\!\!\!\perp \mathrm{Nd}(X) \mid \mathrm{Pa}(X)) \longrightarrow P(X_1, \ldots, X_n) = \prod_j P(X_j | \mathrm{Pa}(X_j))$$

The *d*-separation condition for a DAG

Sets X and Y are *d-separated* given a set Z iff Z blocks all paths between X and Y

$$(X \perp Y | Z)_d$$

Chain or Fork



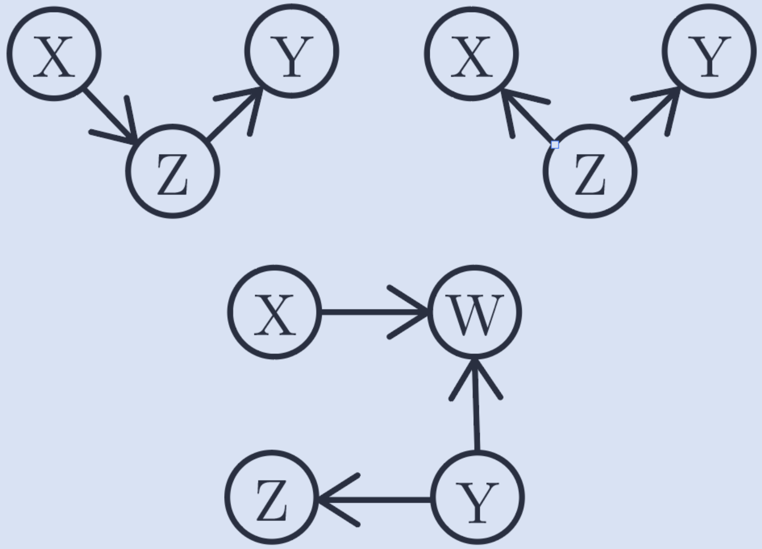
Path is blocked if $\mathrm{B} \in \mathrm{Z}$

Collider (inverted fork)

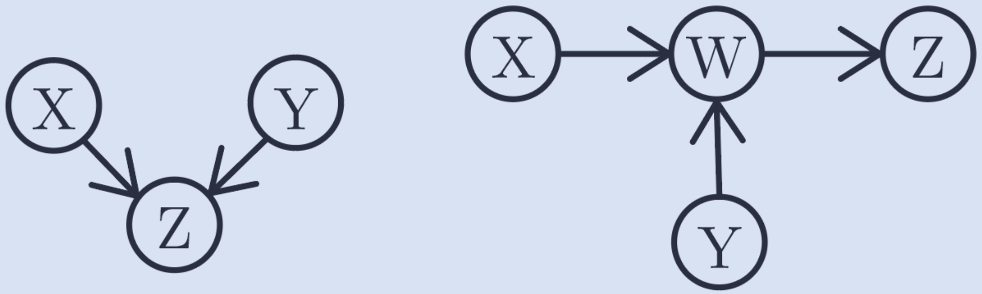
Path is blocked if $\mathrm{B} \notin \mathrm{Z}$ and $\mathrm{De}(\mathrm{B}) \cap Z = \{\}$

d-separation examples

$(X \perp Y | Z)_d$



$\neg(X \perp Y | Z)_d$



d-separation implies Conditional Independence (CI)

If a DAG *G* satisfies a particular *d*-separation condition, any probability distribution *P* compatible with *G* satisfies the associated CI.

Sound: For all *P* compatible with DAG *G*

$$(X \perp Y|Z)_d \Rightarrow (X \perp Y|Z)$$

Complete: If all *P* compatible with *G* satisfy $(X \perp Y|Z)$, then

$$(X \perp Y|Z)_d$$

Semi-Graphoid Axioms

Symmetry $(\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y} \mid \boldsymbol{Z}) \Leftrightarrow (\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{X} \mid \boldsymbol{Z})$

Decomposition $(\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y}\boldsymbol{W} \mid \boldsymbol{Z}) \Rightarrow (\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y} \mid \boldsymbol{Z})$

Weak union $(\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y}\boldsymbol{W} \mid \boldsymbol{Z}) \Rightarrow (\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y} \mid \boldsymbol{Z}\boldsymbol{W})$

Contraction $(\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y} \mid \boldsymbol{Z}) \;\&\; (\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{W} \mid \boldsymbol{Z}\boldsymbol{Y}) \Rightarrow (\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y}\boldsymbol{W} \mid \boldsymbol{Z})$

Causal framework for Bell & KS scenarios

Set of measurements: $\mathcal{M} = \{m_1, \ldots, m_k\}$

Set of measurement outcomes: $\mathcal{O}_m = \mathcal{O} \;\; \forall m$

Measurement contexts: $c \subseteq \mathcal{M}$ IFF $c \in \mathcal{C}$

i.e. m_1, m_2 compatible $\leftrightarrow$ $\{m_1, m_2\} \in \mathcal{C}$

Bell scenario: $\mathcal{M} = \mathcal{M}_1 \cup \mathcal{M}_2 \cup \ldots \cup \mathcal{M}_n$

$\mathcal{M}_i \cap \mathcal{M}_j = \{\,\} \;\; \forall i \neq j$

$x_1 \in \mathcal{M}_1, \; x_2 \in \mathcal{M}_2, \ldots, x_n \in \mathcal{M}_n$

Measurement notation

Measurement settings: $\boldsymbol{X} = \{X_1, X_2, ..., X_n\}$

e.g. $\{X_1 = m_1, ..., X_n = m_n\} \in \mathcal{C}$

Measurement outcomes: $\boldsymbol{A} = \{A_1, A_2, ..., A_n\}$

Measurement-outcome pair: (X_i, A_i) for all $i \in \mathcal{I} = \{1, 2, ..., n\}$

Index subset $\gamma \subseteq \mathcal{I}$ $\quad \boldsymbol{A}_\gamma \subseteq \boldsymbol{A}$ and $\boldsymbol{X}_\gamma \subseteq \boldsymbol{X}$

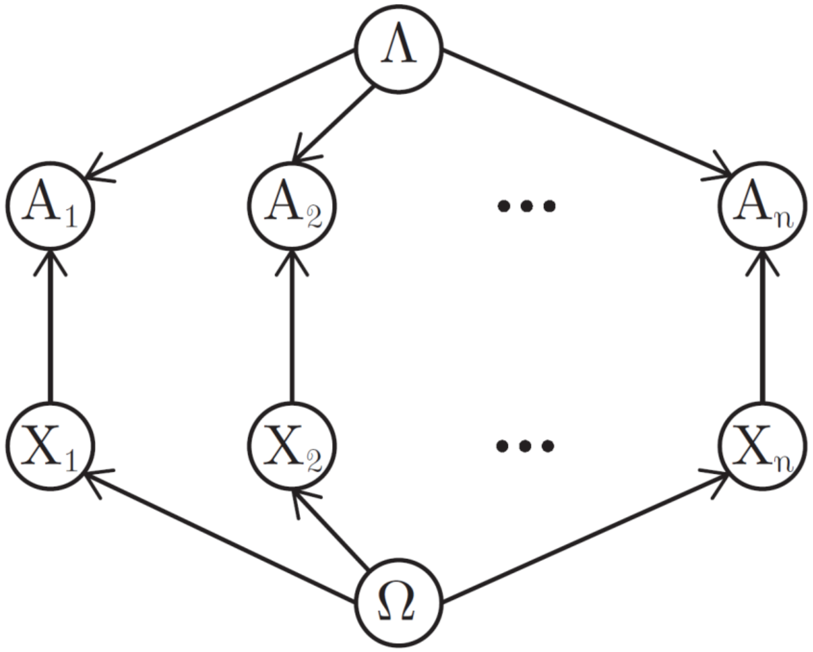
$\boldsymbol{A}_{\setminus\gamma} = \boldsymbol{A} \setminus \boldsymbol{A}_\gamma \qquad \boldsymbol{X}_{\setminus\gamma} = \boldsymbol{X} \setminus \boldsymbol{X}_\gamma$

Classical Causal Model

A (classical) causal model Γ for a phenomenon $\mathcal{P}$ consists of,

$\exists\ \boldsymbol{\Xi}$, DAG G on $\{\mathbf{A}, \mathbf{X}, \boldsymbol{\Xi}\}$ and P compatible with G s.t.

$$\mathcal{P}(\boldsymbol{A}\boldsymbol{X}) = \sum_{\boldsymbol{\Xi}} P(\boldsymbol{A}\boldsymbol{X}\boldsymbol{\Xi})$$



No Disturbance

A phenomenon satisfies no-disturbance iff

(i) $\mathcal{P}(\boldsymbol{A}_\gamma|\boldsymbol{X}) = \mathcal{P}(\boldsymbol{A}_\gamma|\boldsymbol{X}_\gamma)$ $\forall$ $\{A_i, X_i\}$, $\gamma \subseteq \mathcal{I}$ & $i \in \gamma$

A subset of outcomes depends only on the associated subset of settings.

(ii) $\mathcal{P}(A_i|X_i = m) = \mathcal{P}(A_j|X_j = m) \forall i, j$

Marginals for the same measurement are independent of the index.

No-disturbance for a scenario with 3 measurements

Three constraints of the form: $\mathcal{P}(A_1|X_1X_2X_3) \quad = \quad \mathcal{P}(A_1|X_1)$

Three constraints of the form: $\mathcal{P}(A_1A_2|X_1X_2X_3) = \mathcal{P}(A_1A_2|X_1X_2)$

Causal model notation: $(\boldsymbol{A}_\gamma \perp\!\!\!\perp \boldsymbol{X}_{\backslash\gamma} \mid \boldsymbol{X}_\gamma)$

Factorisability

A causal model for a phenomenon is factorisable IFF

$$\mathcal{P}(\boldsymbol{A}|\boldsymbol{X}) \;\; = \textstyle\sum_{\boldsymbol{\Lambda}} P(\boldsymbol{\Lambda}) \prod_i P(A_i|\boldsymbol{\Lambda} X_i)$$

For Kochen-Specker scenarios

$$P(A_i|\boldsymbol{\Lambda} X_i = m) = P(A_j|\boldsymbol{\Lambda} X_j = m)$$

A causal model for a Bell scenario is ***Bell-local*** IFF it is factorisable.

A causal model for a contextuality scenario satisfies ***KS-noncontextuality*** IFF it satisfies measurement noncontextuality, outcome determinism and freedom of choice.

Fine-Abramsky-Brandenburger Theorem:

A phenomenon satisfies KS-noncontextuality IFF it has a factorisable model.

Faithfulness (no fine-tuning)

A causal model Γ is faithful relative to a phenomenon $\mathcal{P}$ IFF

1. All CI's $(C \perp D|E)$ in $\mathcal{P}$ correspond to $(C \perp D|E)_d$ in *G* of Γ.

   i.e. if $\mathcal{P}$ satisfies $(C \perp D|E)$, then any faithful DAG satisfies $(C \perp D|E)_d$

2. Operational symmetries of $\mathcal{P}$ are reflected by the model, rather than holding only for special values of the model parameters.

$$\mathcal{P}(A_i|X_i = m) = \mathcal{P}(A_j|X_j = m) \longrightarrow P(A_i|\mathbf{\Lambda}X_i = m) = P(A_j|\mathbf{\Lambda}X_j = m)$$

Results

Theorem 1: *Every phenomenon satisfying no-disturbance in an arbitrary contextuality scenario that has a faithful causal model is factorisable.*

Corollary 1: *No fine-tuning and no-disturbance (no-signalling) imply KS noncontextuality (Bell locality) in arbitrary scenarios.*

Corollary 2: *Every classical causal model that reproduces the violation of a Bell-KS inequality for a no-disturbance phenomenon in an arbitrary Bell-KS scenario requires fine-tuning.*

Outline of the Proof

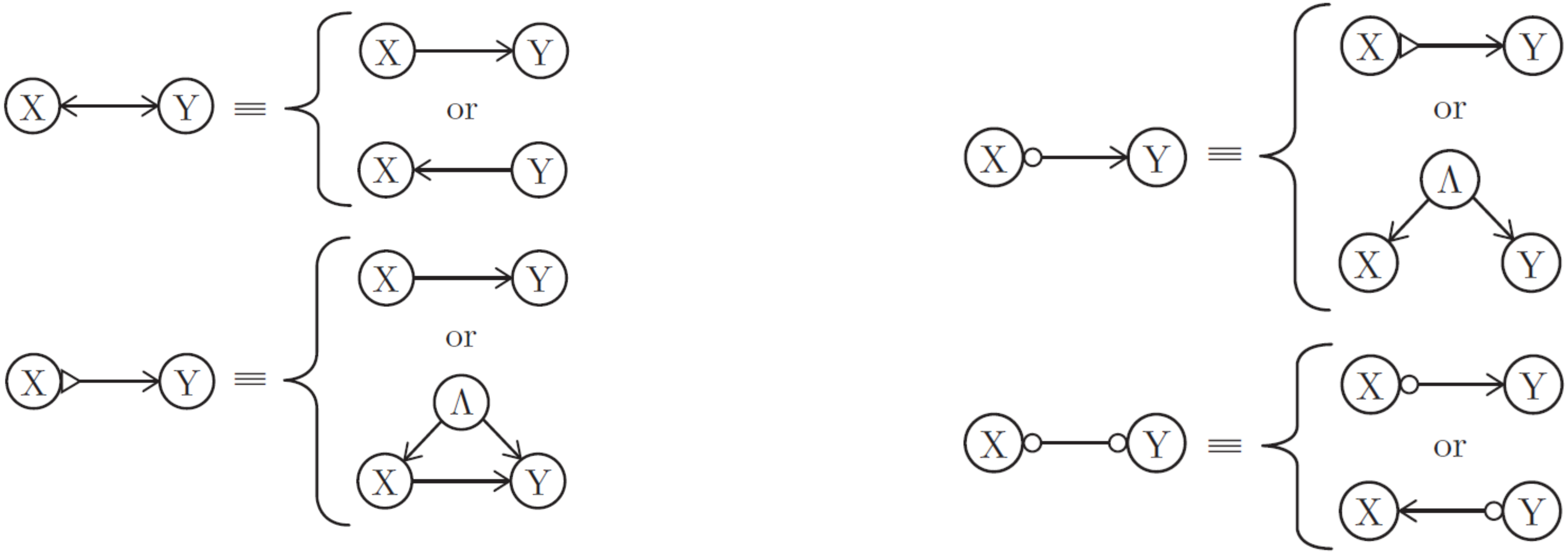
From conditional probability, any phenomenon can be written as

$$\mathcal{P}(\boldsymbol{A} \mid \boldsymbol{X}) = \sum_{\boldsymbol{\Xi}} P(\boldsymbol{A} \mid \boldsymbol{X}\boldsymbol{\Xi}) P(\boldsymbol{\Xi} \mid \boldsymbol{X})$$

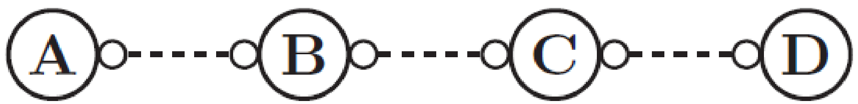
No-disturbance + NFT $\longrightarrow$ additional constraints on the model

These constraints lead to factorisability of the model.

Graphical shortcut notation



Lemma 1



A chained graph $\mathcal{V}_c$.

- **A, B, C, D** represent *sets of vertices.*
- Connections indicated represent possible connections between elements in **A, B, C, D**.
- Dashed connections represent the possibility of no causal connection.

$$(A \perp\!\!\!\perp C \mid B)_d \longrightarrow (A \perp\!\!\!\perp CD \mid B)_d, \; (A \perp\!\!\!\perp D \mid BC)_d$$

Proof

$(A \perp\!\!\!\perp C \mid B)_d$ implies that **B** blocks all paths between **A** and **C.** So **B** blocks all paths from **A** to **D.** Thus **B** blocks all paths between **A** and **CD**. From the weak union axiom,

$$(A \perp\!\!\!\perp CD \mid B)_d \longrightarrow (A \perp\!\!\!\perp D \mid BC)_d$$

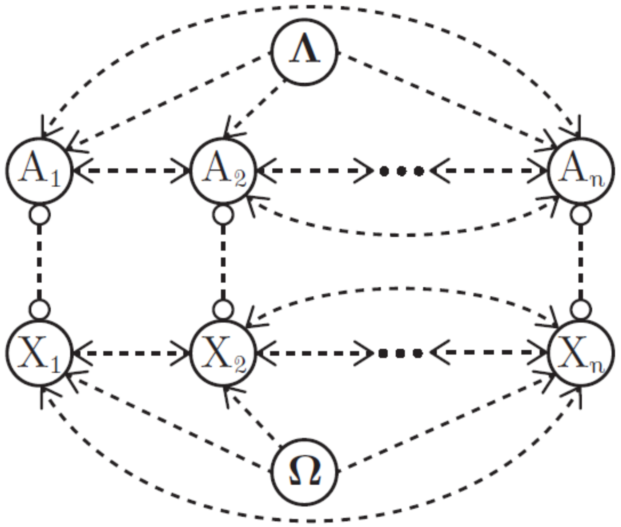
Proof

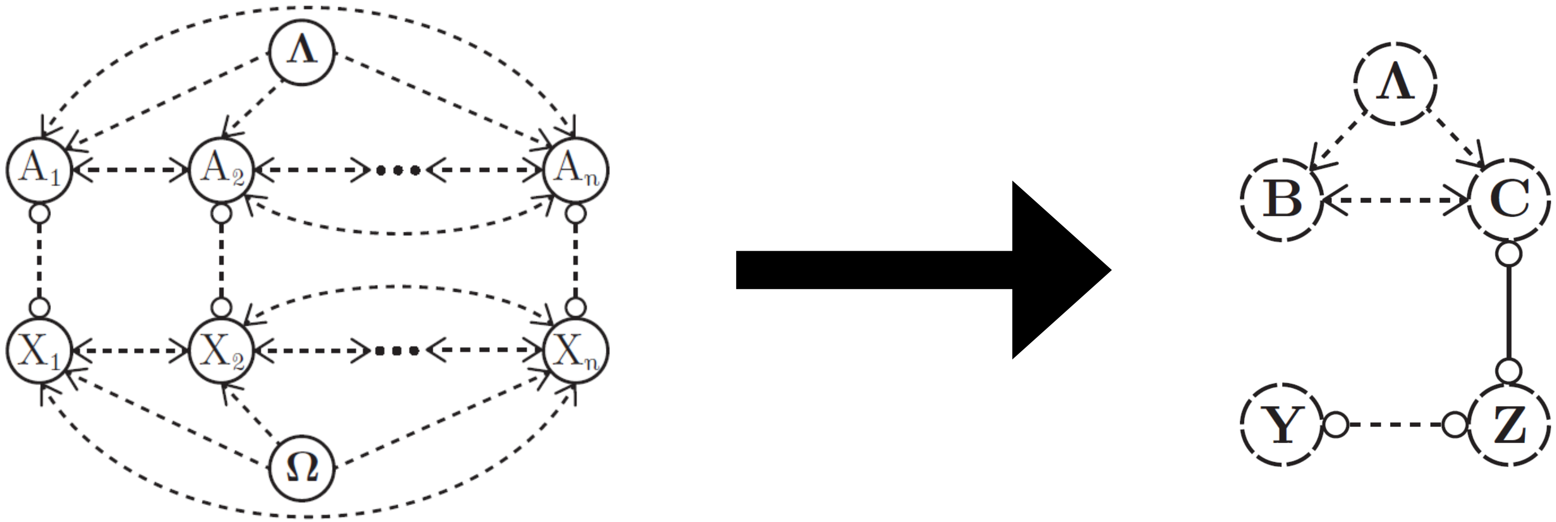
- DAG structure for a no-disturbance phenomenon.
- Arbitrary number of parties or measurements per context.
- Latent variables as common causes between observables.

$(\boldsymbol{A}_\gamma \perp\!\!\!\perp \boldsymbol{X}_{\setminus\gamma} \mid \boldsymbol{X}_\gamma)_d$ $\longrightarrow$ **No direct or common cause between** $\{A_i, \boldsymbol{X}_{\setminus i}\}$

Any causal connection remains between

$\{A_i, X_i\}$, $\{A_i, \boldsymbol{A}_{\setminus i}\}$, $\{X_i, \boldsymbol{X}_{\setminus i}\}$





All members of $\boldsymbol{A} = \{A_1, \ldots, A_n\}$ and $\boldsymbol{X} = \{X_1, \ldots, X_n\}$ are grouped into subsets **B, C, Y, Z**

$\boldsymbol{B} \subseteq \boldsymbol{A}$ have no causal connection to $\boldsymbol{X}$.

$\boldsymbol{C} \subseteq \boldsymbol{A}$ have some causal connection to $\boldsymbol{X}$.

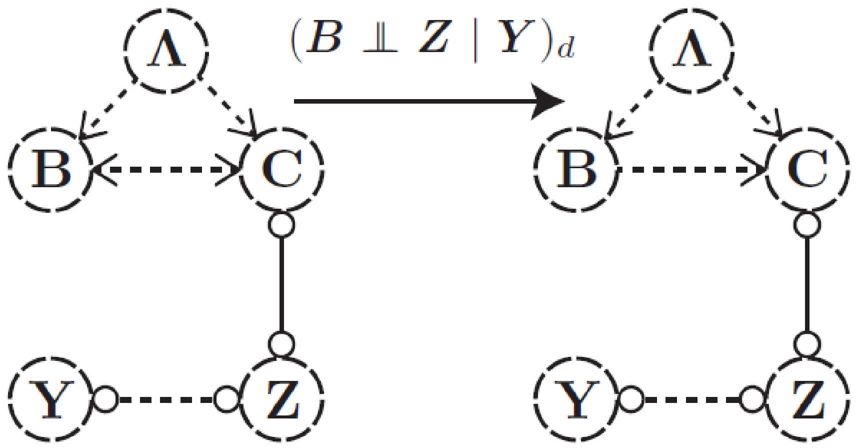
Dashed nodes represent the possibility of an empty set.

i.e. $\boldsymbol{B} = \{\} \implies \boldsymbol{C} = \boldsymbol{A},\ \boldsymbol{Y} = \{\}$ and $\boldsymbol{Z} = \boldsymbol{X}$. All pairs $\{A_i, X_i\}$ have some causal link.

From no-disturbance + NFT, $(B \perp\!\!\!\perp Z \mid Y)_d$

Any path between ***B*** and ***Z*** must pass through one element of ***C.***

Since ***C*** is not in ***Y,*** it must act as a collider. Direct links from ***C*** to ***B*** would therefore violate $(B \perp\!\!\!\perp Z \mid Y)_d$.



Elimination of direct links from C to B.

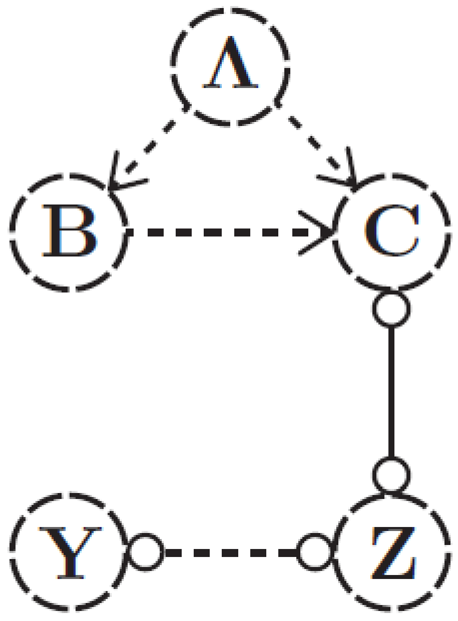
Y cannot act as a middle node between ***B*** and ***Z***.

$(\boldsymbol{B} \perp\!\!\!\perp \boldsymbol{Z} \mid \boldsymbol{Y})_d$ implies $(\boldsymbol{B} \perp\!\!\!\perp \boldsymbol{Z})_d$

B and ***Z*** are *d*-separated given any non-collider $\Longrightarrow$ $\boxed{(\boldsymbol{B} \perp\!\!\!\perp \boldsymbol{Z} \mid \boldsymbol{\Lambda})_d}$

No-disturbance + NFT $\Longrightarrow$ $(\boldsymbol{C} \perp\!\!\!\perp \boldsymbol{Y} \mid \boldsymbol{Z})_d \Longrightarrow (\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{C} \mid \boldsymbol{Z})_d$

Lemma 1: $(\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{C} \mid \boldsymbol{Z})_d \Longrightarrow (\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{C}\boldsymbol{B}\boldsymbol{\Lambda} \mid \boldsymbol{Z})_d$



Weak union: $(\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{C}\boldsymbol{B} \mid \boldsymbol{Z}\boldsymbol{\Lambda})_d$

Symmetry and $\boldsymbol{BC} = \boldsymbol{A}$: $\boxed{(\boldsymbol{A} \perp\!\!\!\perp \boldsymbol{Y} \mid \boldsymbol{Z}\boldsymbol{\Lambda})_d}$

Consider the pair $\{C_i, C_j\} \in \mathbf{C}$

$(\boldsymbol{A}_\gamma \perp\!\!\!\perp \boldsymbol{X}_{\setminus\gamma} \mid \boldsymbol{X}_\gamma)_d$ + decomposition $\Longrightarrow$ $(C_j \perp\!\!\!\perp Z_i \mid Z_j)_d$

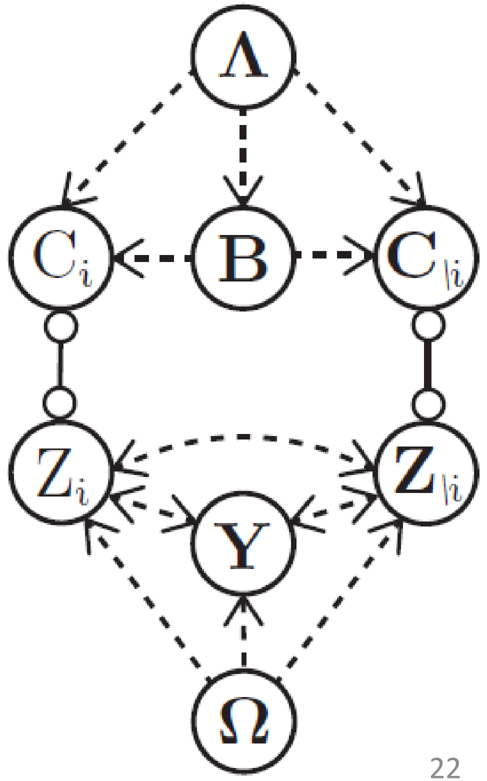
For a path $(Z_i - C_i - C_j)$ to be blocked by Z_i, C_i must be a collider.

This eliminates a direct link from C_i to C_j

Similarly, for $(Z_j - C_j - C_i)$ and $(C_i \perp\!\!\!\perp Z_j \mid Z_i)_d$

Direct links from C_j to C_i are eliminated.

No pair $\{C_i, C_j\} \in \mathbf{C}$ can have a direct causal link.



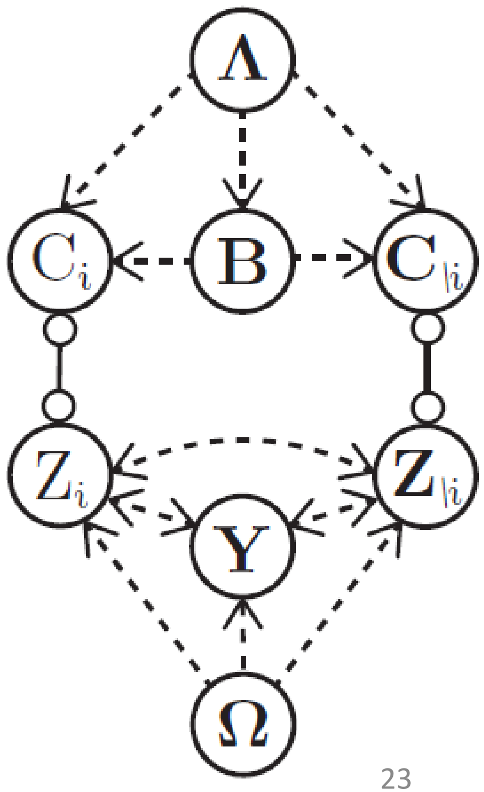
Consider paths between C_i and $\boldsymbol{C}_{\setminus i}$ that go through Z_i

$(\boldsymbol{A}_\gamma \perp\!\!\!\perp \boldsymbol{X}_{\setminus\gamma} \mid \boldsymbol{X}_\gamma)_d$ + decomposition $\Longrightarrow$ $(C_i \perp\!\!\!\perp \boldsymbol{Z}_{\setminus i}\boldsymbol{Y} \mid Z_i)_d$

Z_i cannot be a collider between C_i and $\boldsymbol{Z}_{\setminus i}\boldsymbol{Y}$

Z_i must be the middle node of a chain or a fork

Z_i blocks all paths between C_i and $\boldsymbol{C}_{\setminus i}$ through Z_i



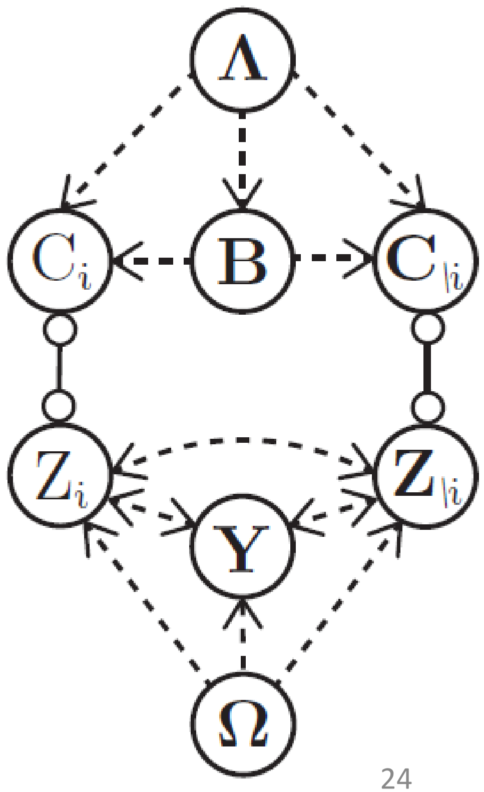
Consider paths between C_i and $\boldsymbol{C}_{\setminus i}$ that go through $\boldsymbol{B\Lambda}$

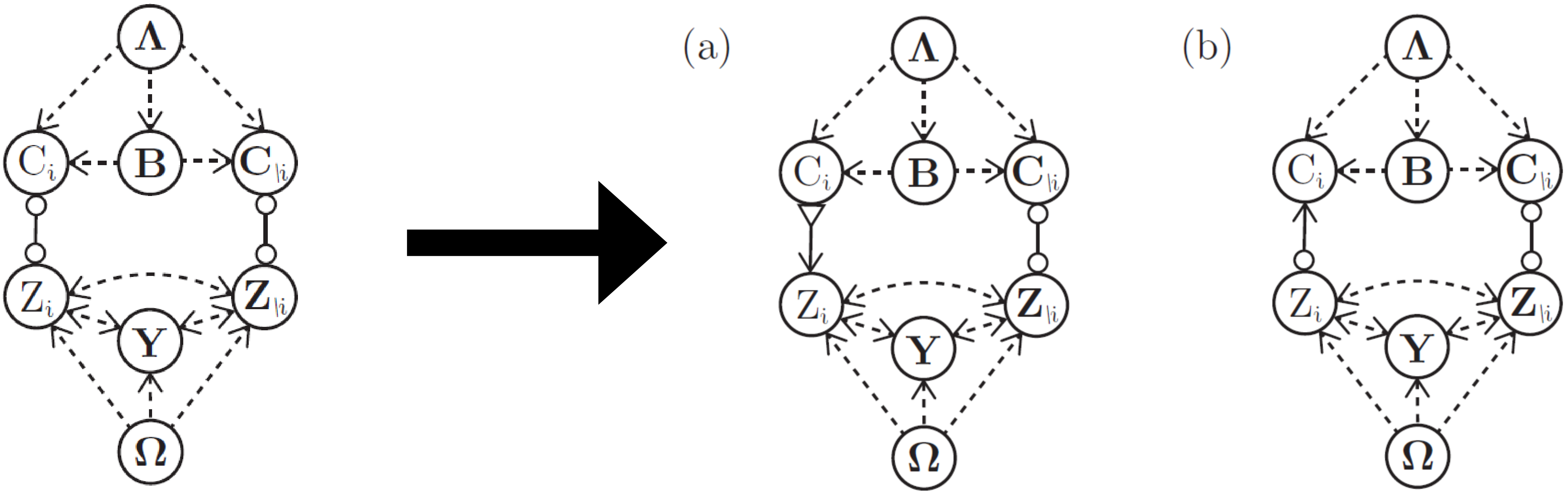
$\boldsymbol{B\Lambda}$ acts as the middle node of a chain or fork

Thus $\boldsymbol{B\Lambda}$ blocks all paths through $\boldsymbol{B\Lambda}$

No-disturbance + NFT $\Longrightarrow$ $(C_i \perp\!\!\!\perp \boldsymbol{C}_{\setminus i} \mid \boldsymbol{ZB\Lambda})_d$

$(C_i \perp\!\!\!\perp \boldsymbol{Z}_{\setminus i} \mid Z_i)_d \quad \Longrightarrow \quad (C_i \perp\!\!\!\perp \boldsymbol{Z}_{\setminus i} \mid Z_i\boldsymbol{\Lambda})_d$





(a) Direct link from C_i to Z_i with or without a common cause
(b) Excludes a direct link from C_i to Z_i

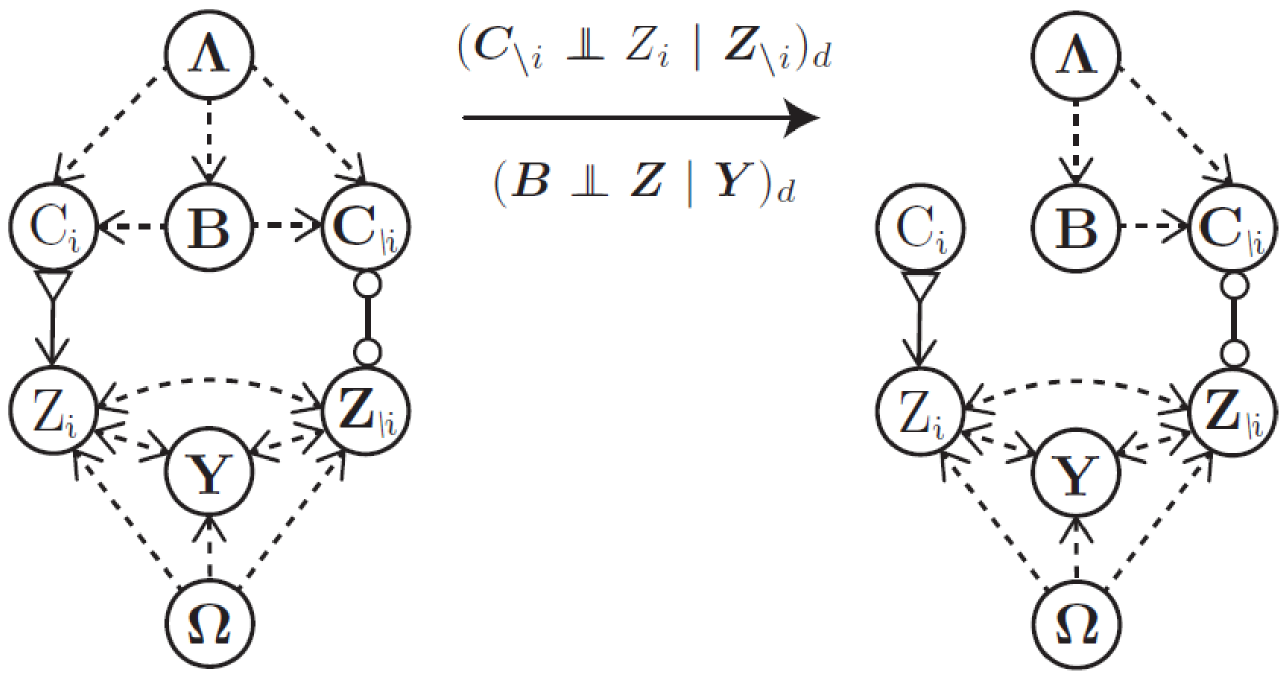
No-disturbance + NFT $\implies (\boldsymbol{B} \perp\!\!\!\perp \boldsymbol{Z} \mid \boldsymbol{Y})_d$ and $(\boldsymbol{C}_{\setminus i} \perp\!\!\!\perp Z_i \mid \boldsymbol{Z}_{\setminus i})_d$

C_i and any member of $\boldsymbol{B}$ or $\boldsymbol{C}_{\setminus i}$ cannot share a common cause

For graphs of this type,

there are no paths of type

$\boldsymbol{\Lambda} - C_i - Z_i$



Consider paths of type $\boldsymbol{\Lambda} - C_i - Z_i$

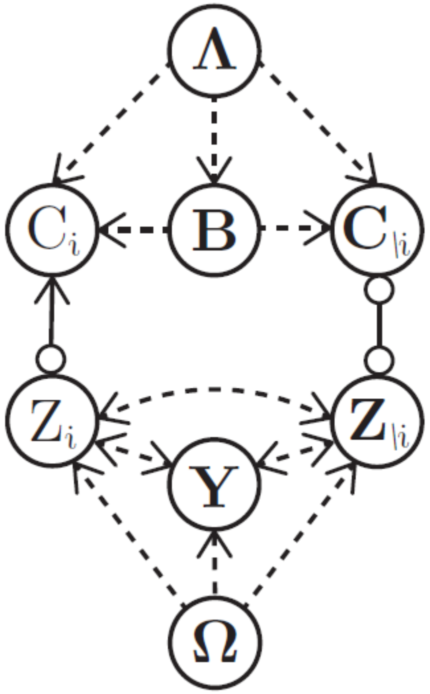
C_i must always act as a collider, where Z_i is not a descendent.

Paths of this type are blocked by the empty set.

Every path between $\boldsymbol{\Lambda}$ and $\boldsymbol{Z}$ includes a subpath of this form in (a) or (b).

$$\boxed{(\boldsymbol{\Lambda} \perp\!\!\!\perp \boldsymbol{X})_d}$$

(b)



The derived *d*-separation conditions imply the corresponding conditional independence (CI) relations

The joint distribution can be written as

$$\mathcal{P}(\boldsymbol{A} \mid \boldsymbol{X}) = \sum_{\boldsymbol{\Lambda}\,\boldsymbol{\Omega}} P(\boldsymbol{A}\boldsymbol{\Omega} \mid \boldsymbol{X}\boldsymbol{\Lambda}) P(\boldsymbol{\Lambda} \mid \boldsymbol{X})$$

Summing over $\boldsymbol{\Omega}$**,** using (1) and writing $\boldsymbol{X} = \boldsymbol{YZ}$

$$\mathcal{P}(\boldsymbol{A} \mid \boldsymbol{X}) = \sum_{\boldsymbol{\Lambda}} P(\boldsymbol{A} \mid \boldsymbol{Y}\boldsymbol{Z}\boldsymbol{\Lambda}) P(\boldsymbol{\Lambda})$$

From $\boldsymbol{A} = \boldsymbol{BC}$ and using (2) and (3)

$$\mathcal{P}(\boldsymbol{A} \mid \boldsymbol{X}) = \sum_{\boldsymbol{\Lambda}} P(\boldsymbol{C} \mid \boldsymbol{Z}\boldsymbol{B}\boldsymbol{\Lambda}) P(\boldsymbol{B} \mid \boldsymbol{\Lambda}) P(\boldsymbol{\Lambda})$$

Derived CI relations

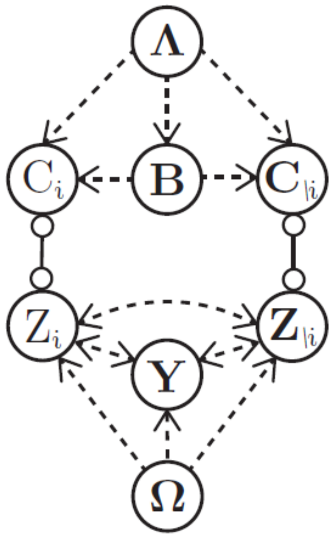
$(\boldsymbol{\Lambda} \perp\!\!\!\perp \boldsymbol{X})$ (1)

$(\boldsymbol{B} \perp\!\!\!\perp \boldsymbol{Z} \mid \boldsymbol{\Lambda})$ (2)

$(\boldsymbol{A} \perp\!\!\!\perp \boldsymbol{Y} \mid \boldsymbol{Z}\boldsymbol{\Lambda})$ (3)

$(C_i \perp\!\!\!\perp \boldsymbol{C}_{\setminus i} \mid \boldsymbol{Z}\boldsymbol{B}\boldsymbol{\Lambda})$ (4)

$(C_i \perp\!\!\!\perp \boldsymbol{Z}_{\setminus i} \mid Z_i\boldsymbol{\Lambda})$ (5)



No variables outside $\boldsymbol{B}$ can have a direct causal link to $\boldsymbol{B}$

Let $\boldsymbol{\Lambda}$ determine $\boldsymbol{B}$

$$P(\boldsymbol{C} \mid \boldsymbol{Z}\boldsymbol{B}\boldsymbol{\Lambda}) = P(\boldsymbol{C} \mid \boldsymbol{Z}\boldsymbol{\Lambda})$$

$$P(\boldsymbol{C} \mid \boldsymbol{Z}\boldsymbol{\Lambda}) = \prod_j P(C_j \mid \boldsymbol{C}\backslash\{C_1, C_2, \ldots, C_j\}\boldsymbol{Z}\boldsymbol{\Lambda})$$

From (4) and (5) then, and similarly for $P(\boldsymbol{B}|\boldsymbol{\Lambda})$

$$P(\boldsymbol{C} \mid \boldsymbol{Z}\boldsymbol{\Lambda}) = \prod_j P(C_j \mid Z_j\boldsymbol{\Lambda})$$

$$P(\boldsymbol{B} \mid \boldsymbol{\Lambda}) = \prod_k P(B_k \mid \boldsymbol{\Lambda})$$

$$\mathcal{P}(\boldsymbol{A} \mid \boldsymbol{X}) = \sum_{\boldsymbol{\Lambda}} P(\boldsymbol{\Lambda}) \prod_j P(C_j \mid Z_j\boldsymbol{\Lambda}) \prod_k P(B_k \mid \boldsymbol{\Lambda})$$

Derived CI relations

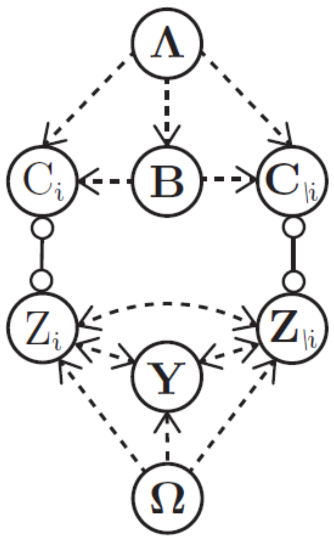
$$(\boldsymbol{\Lambda} \perp\!\!\!\perp \boldsymbol{X}) \quad (1)$$

$$(\boldsymbol{B} \perp\!\!\!\perp \boldsymbol{Z} \mid \boldsymbol{\Lambda}) \quad (2)$$

$$(\boldsymbol{A} \perp\!\!\!\perp \boldsymbol{Y} \mid \boldsymbol{Z}\boldsymbol{\Lambda}) \quad (3)$$

$$(C_i \perp\!\!\!\perp \boldsymbol{C}_{\backslash i} \mid \boldsymbol{Z}\boldsymbol{B}\boldsymbol{\Lambda}) \quad (4)$$

$$(C_i \perp\!\!\!\perp \boldsymbol{Z}_{\backslash i} \mid Z_i\boldsymbol{\Lambda}) \quad (5)$$



This completes the proof for Bell scenarios.

Since a no-disturbance phenomenon satisfies the operational symmetry,

$$\mathcal{P}(A_i|X_i = m) = \mathcal{P}(A_j|X_j = m)$$

No fine-tuning requires that

$$P(A_i|\mathbf{\Lambda}X_i = m) = P(A_j|\mathbf{\Lambda}X_j = m)$$

Which completes the proof for KS scenarios.

Wrapping up

Bell-nonlocality and Kochen-Specker contextuality as violations of the classical framework of causality.

Generalises previous results

C. J. Wood and R. W. Spekkens, New Journal of Physics **17**, 033002 (2015).

E. G. Cavalcanti, Physical Review X **8**, 021018 (2018).