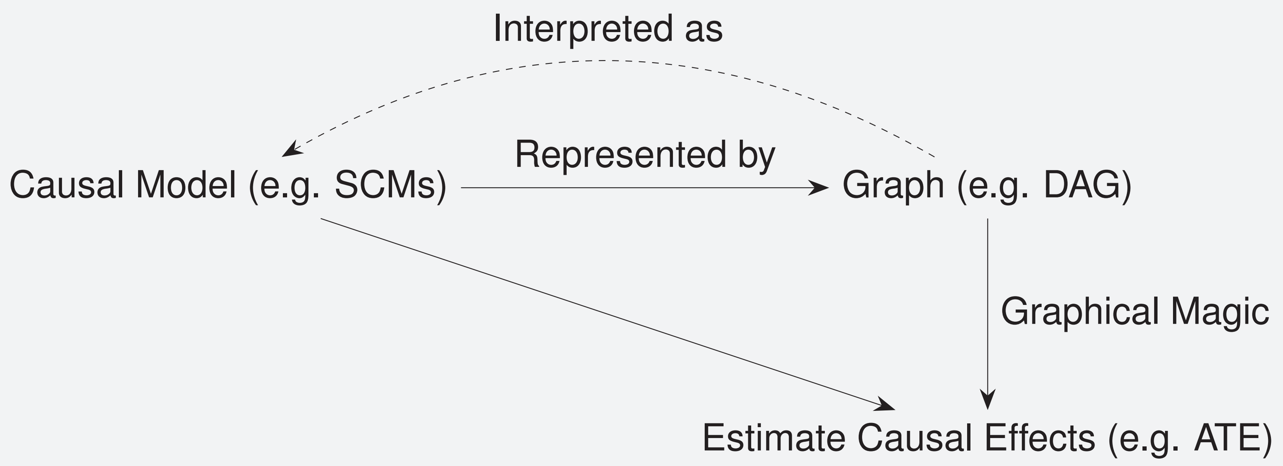
Causal Interpretations of Anterial Graphs

Kai Teh, Kayvan Sadeghi, Terry Soo
Department of Statistical Science, University College London

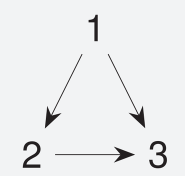
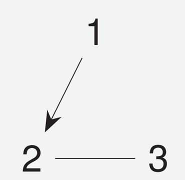
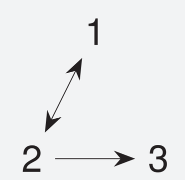
Keywords Causal Inference, Graphical Models

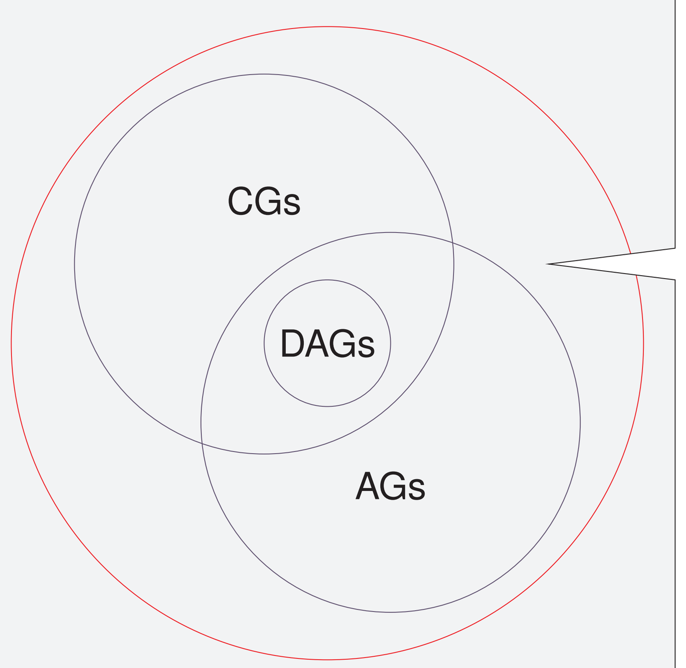
Motivation

In graphical causal inference,



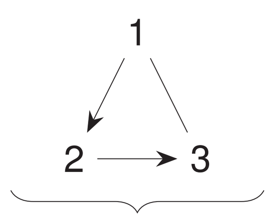
Specifically:

Graphical Model	Causal Interpretation
Directed Acyclic Graphs (DAGs)	Structural Causal Models (SCMs) with independent noise
	$X_1 = f_1(\epsilon_1)$ $X_2 = f_2(X_1, \epsilon_2)$ $X_3 = f_3(X_1, X_2, \epsilon_3)$ ϵ_i i.i.d.
Chain Graphs (CGs)	Nested Gibbs Dynamics (NGs) with independent noise (Richardson and Lauritzen 2002)
	$X_1 = f_1(\epsilon_1)$ $X_{2,3} = f_{2,3}(X_1, \epsilon_{2,3})$ where $\epsilon_1 \perp\!\!\!\perp \epsilon_{2,3}$ $f_1(\epsilon_1) \sim J(X_1)$ $f_{2,3}(X_1, \epsilon_{2,3}) \sim J(X_{2,3} \mid X_1)$ $J(X_1), J(X_{2,3} \mid X_1)$ equilibrium of Gibbs dynamics
Ancestral Graphs (AGs)	Structural Causal Models with potentially dependent noise (Richardson and Spirtes 2002, Sadeghi and Soo 2022)
	$X_1 = f_1(\epsilon_1)$ $X_2 = f_2(\epsilon_2)$ $X_3 = f_3(X_2, \epsilon_3)$ where $\epsilon_i \perp\!\!\!\perp \epsilon_j$ if $i \not\leftrightarrow j$

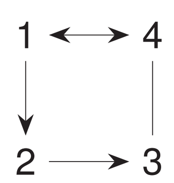


Anterial Graphs (Lauritzen and Sadeghi 2018)

Not allowed ×



Semi-directed cycle

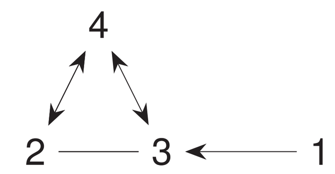


Semi-directed path connecting bidirected edges

We have graphical magic for this!

But what is the causal interpretation??

Interpretations



Chain-connected anterial graph

SCM interpretation

$$X_1 = f_1(\epsilon_1) \qquad X_{2,3} = f_{2,3}(X_1, \epsilon_{2,3}) \qquad X_4 = f_4(\epsilon_4)$$

$$\epsilon_1 \perp\!\!\!\perp \epsilon_{2,3}, \quad \epsilon_1 \perp\!\!\!\perp \epsilon_4$$

$$f_1(\epsilon_1) \sim J(X_1) \qquad f_{2,3}(X_1, \epsilon_{2,3}) \sim J(X_{2,3} \mid X_1) \qquad f_4(\epsilon_4) \sim J(X_4)$$

$J(X_1), J(X_{2,3} \mid X_1), J(X_4)$ equilibrium of Gibbs dynamics

$\epsilon_{2,3}$ and ϵ_4 are allowed to be dependent!

Theorem

Let distribution P be induced from the SCM-interpretation of chain-connected anterial graph $\mathcal{G}$.
If P is a *compositional graphoid***, then P is Markovian to $\mathcal{G}$.

Joint Dynamical Model (JDM) interpretation

JDM interpretation

$$m_t = m_{t-1} + 1 \bmod |V|$$
$$X^t_{V\setminus m_t} = X^{t-1}_{V\setminus m_t}$$
$$X^t_{m_t} = h_{m_t}(X^{t-1}_{\mathrm{ne}(\{m_t\})}, X^t_{\mathrm{pa}(\{m_t\})}, \epsilon^t_{\tau(m_t)})$$

NG (Richardson and Lauritzen 2002)

$$m_t = m_{t-1} + 1 \bmod |\tau|$$
$$X^t_{V\setminus m_t} = X^{t-1}_{V\setminus m_t}$$
$$X^t_{m_t} = h_{m_t}(X^{t-1}_{\mathrm{ne}(\{m_t\})}, X_{\mathrm{pa}(\{m_t\})}, \epsilon^t_{\tau})$$

$\epsilon_{2,3}$ and ϵ_4 are allowed to be dependent!

Updated jointly using current values of parents vs NG which updates chain component in order

PROBLEM: equilibrium may not exist
SOLUTION: consider the equilibrium of convergent subsequences in time

Theorem

Let distribution P be induced from the equilibrium of some convergent subsequence from the JDM-interpretation of chain-connected anterial graph $\mathcal{G}$.
Under some *conditions**, if P is a *compositional graphoid***, then P is Markovian to $\mathcal{G}$.

Conditions* are satisfied when the JDM interpretation
1. takes discrete values, or
2. satisfies time-equicontinuity

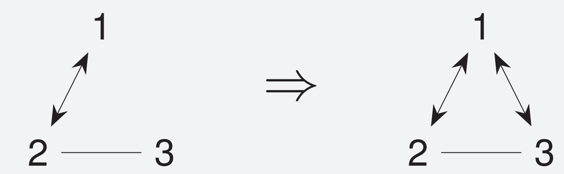
Compositional graphoids** are satisfied by distributions such that
1. full support, and
2. pairwise independence implies joint independence

e.g. Gaussians.

Both are equivalent graphically, (manipulated in the same way).

Chain-connected anterial graphs

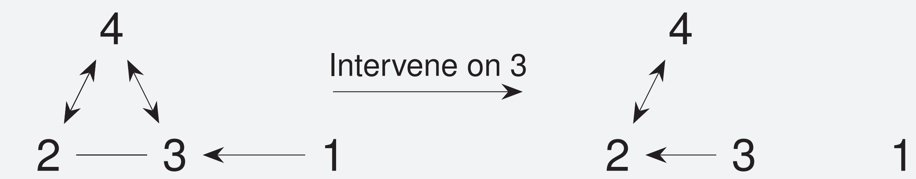
Furthermore, restrict anterial graphs via the **chain-connectedness** condition.



Still contains DAGs, CGs, and AGs!

Interventions

Graphically:



From the interpretations:

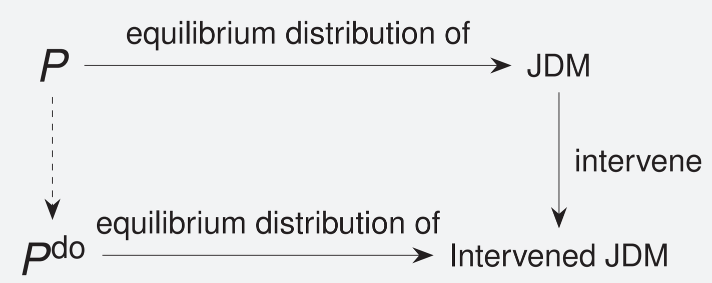
SCM interpretation

$$X_1 = f_1(\epsilon_1) \qquad X_2 = f'_2(\epsilon_{2,3}) \qquad X_3 = x_3 \qquad X_4 = f_4(\epsilon_4)$$

$$f_1(\epsilon_1) \sim J(X_1) \qquad f'_2(\epsilon_{2,3}) \sim J^{do}(X_{2,3} \mid X_1) \qquad f_4(\epsilon_4) \sim J(X_4)$$

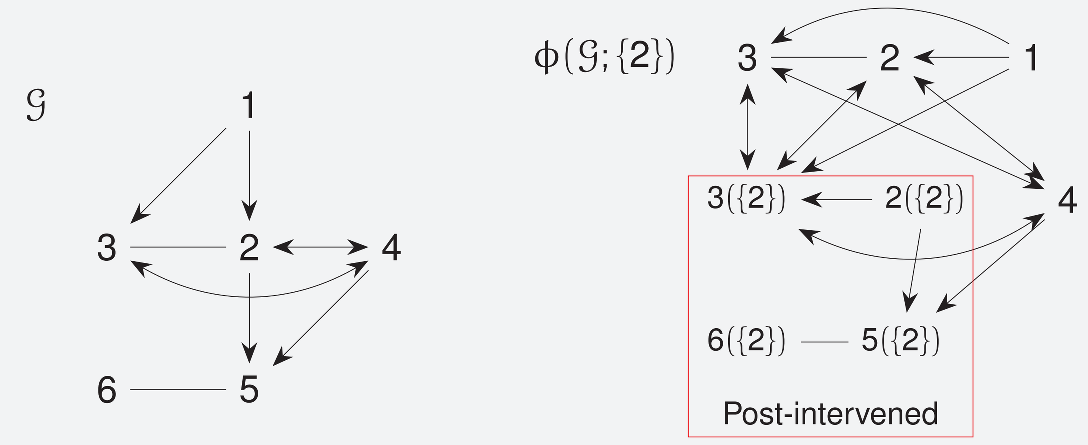
$J^{do}(X_{2,3} \mid X_1)$ equilibrium of intervened Gibbs dynamics

JDM interpretation



Single-World Interpretation

Given a DAG $\mathcal{G}$ and intervention targets C, conditional independence between pre and post-intervened variables can be represented via a single-world intervention graph (SWIG) $\phi(\mathcal{G}; C)$ obtained after transforming said DAG (Richardson and Robins 2013).



Pre-intervened graph   Single World Anterial Intervention Graph (SWAIG)

SWIG {

What ϕ does	
If in $\mathcal{G}$	Then in $\phi(\mathcal{G}; C)$
$i \to j$ and $i \in \tau(C)$	delete $i \to j$, add $i(C) \to j$
$i - j$ and $i \in C, j \notin C$	add $i(C) \to j(C)$
$j \to i$ and $i \in \tau(C) \setminus C$	add $j \to i(C)$
$i - j$ and $i, j \in \tau(C) \setminus C$	add $i(C) - j(C)$
$i \leftrightarrow j$ and $i \in \tau(C) \setminus C$	add $i(C) \leftrightarrow j$
$i \leftrightarrow j$ and $i, j \in \tau(C) \setminus C$	add $i(C) \leftrightarrow j(C)$
$i \in \tau(C)$ and $j \in \tau(i) \setminus C$	add $j(C) \leftrightarrow j$ and $j(C) \leftrightarrow i$