Cyclic Causality: D-separation and Nonlinearity

David Reber

3/24 - 3/30 update

Past To-Do's

- 3/24 - 3/25
  - ✓ Relate problem to the literature
  - ✓ Writeup, submit mid-semester report
- 3/27 - 3/30
  - Come up with an example for each of the following:
    - ✓ Cyclic SCMs
    - ✓ Loss of d-separation
    - ✓ Intrinsic SCMs
      - Validity of d-separation via numerics

(It'll be a short week for me as I'll be out of town 3/31 - 4/3)

[Spirtes 2013]

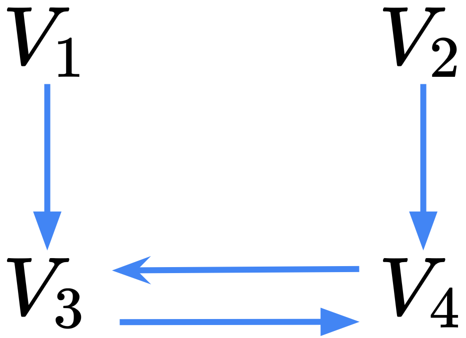
Example: Cyclic SCM

$$M \begin{cases} \mathbf{u}, \mathbf{v} \in \mathbb{R}^4 \\ U_i \sim \mathcal{N}(0,1) \quad iid. \\ f_1 : V_1 \leftarrow U_1 \\ f_2 : V_2 \leftarrow U_2 \\ f_3 : V_3 \leftarrow V_1 V_4 + U_3 \\ f_4 : V_4 \leftarrow V_2 V_3 + U_4 \end{cases}$$

[Spirtes 2013]

Example: Cyclic SCM

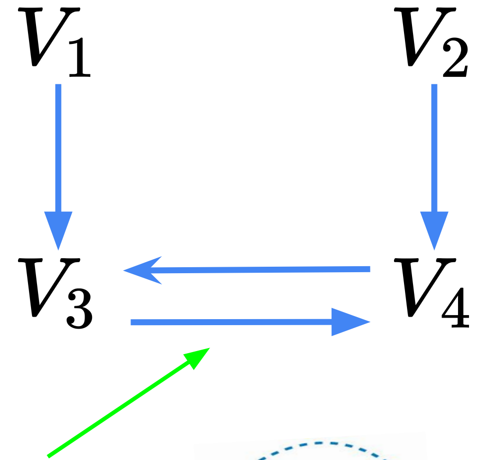
$$
M \left\{
\begin{array}{l}
\mathbf{u}, \mathbf{v} \in \mathbb{R}^4 \\
U_i \sim \mathcal{N}(0,1) \quad iid. \\
f_1 : V_1 \leftarrow U_1 \\
f_2 : V_2 \leftarrow U_2 \\
f_3 : V_3 \leftarrow V_1 V_4 + U_3 \\
f_4 : V_4 \leftarrow V_2 V_3 + U_4
\end{array}
\right.
$$



[Spirtes 2013]

Example: Cyclic SCM

$$
M \begin{cases} \mathbf{u}, \mathbf{v} \in \mathbb{R}^4 \\ U_i \sim \mathcal{N}(0,1) \quad iid. \\ f_1 : V_1 \leftarrow U_1 \\ f_2 : V_2 \leftarrow U_2 \\ f_3 : V_3 \leftarrow V_1 V_4 + U_3 \\ f_4 : V_4 \leftarrow V_2 V_3 + U_4 \end{cases}
$$



Not latent confounder:

[Spirtes 2013]

Example: Observational Distribution

$$
\begin{aligned}
V_1 &\leftarrow U_1 \\
V_2 &\leftarrow U_2 \\
V_3 &\leftarrow V_1 V_4 + U_3 \\
V_4 &\leftarrow V_2 V_3 + U_4
\end{aligned}
$$

[Spirtes 2013]

Example: Observational Distribution

$$\begin{aligned} V_1 &\leftarrow U_1 \\ V_2 &\leftarrow U_2 \\ V_3 &\leftarrow V_1 V_4 + U_3 \\ V_4 &\leftarrow V_2 V_3 + U_4 \end{aligned}$$

For fixed $\boldsymbol{U}=\boldsymbol{u}$, there's a unique equilibrium!

[Spirtes 2013]

Example: Observational Distribution

$$\begin{aligned} V_1 &\leftarrow U_1 \\ V_2 &\leftarrow U_2 \\ V_3 &\leftarrow V_1 V_4 + U_3 \\ V_4 &\leftarrow V_2 V_3 + U_4 \end{aligned}$$

For fixed $\boldsymbol{U}=\boldsymbol{u}$, there's a unique equilibrium!

$$\begin{aligned} V_1(\mathbf{U}) &= U_1 \\ V_2(\mathbf{U}) &= U_2 \\ V_3(\mathbf{U}) &= \tfrac{U_1 U_4 + U_3}{1 - U_1 U_2} \\ V_4(\mathbf{U}) &= \tfrac{U_1 U_4 + U_3}{1 - U_1 U_2} U_2 + U_4 \end{aligned}$$

Potential Responses

[Spirtes 2013]

Example: Observational Distribution

$$\boxed{\begin{aligned} V_1 &\leftarrow U_1 \\ V_2 &\leftarrow U_2 \\ V_3 &\leftarrow V_1 V_4 + U_3 \\ V_4 &\leftarrow V_2 V_3 + U_4 \end{aligned}}$$

For fixed $\boldsymbol{U}$=$\boldsymbol{u}$, there's a unique equilibrium!

$$U_i \sim \mathcal{N}(0,1) \quad iid.$$

$$P(U_i = u_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u_i^2}$$

$$\begin{aligned} V_1(\mathbf{U}) &= U_1 \\ V_2(\mathbf{U}) &= U_2 \\ V_3(\mathbf{U}) &= \frac{U_1 U_4 + U_3}{1 - U_1 U_2} \\ V_4(\mathbf{U}) &= \frac{U_1 U_4 + U_3}{1 - U_1 U_2} U_2 + U_4 \end{aligned}$$

Potential Responses

[Spirtes 2013]

Example: Observational Distribution

$$\begin{aligned} V_1 &\leftarrow U_1 \\ V_2 &\leftarrow U_2 \\ V_3 &\leftarrow V_1 V_4 + U_3 \\ V_4 &\leftarrow V_2 V_3 + U_4 \end{aligned}$$

For fixed $\boldsymbol{U}$=$\boldsymbol{u}$, there's a unique equilibrium!

$$\begin{aligned} V_1(\mathbf{U}) &= U_1 \\ V_2(\mathbf{U}) &= U_2 \\ V_3(\mathbf{U}) &= \frac{U_1 U_4 + U_3}{1 - U_1 U_2} \\ V_4(\mathbf{U}) &= \frac{U_1 U_4 + U_3}{1 - U_1 U_2} U_2 + U_4 \end{aligned}$$

Potential Responses

$$U_i \sim \mathcal{N}(0,1) \quad iid.$$

$$P(U_i = u_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u_i^2}$$

$$P(\mathbf{V}) = (\frac{1}{4\pi^2}) e^{-\frac{v_1^2}{2}} e^{-\frac{v_2^2}{2}} e^{-\frac{(v_3 - v_1 v_4)^2}{2}} e^{-\frac{(v_4 - v_2 v_3)^2}{2}} |(1 - v_1 v_2)^{-1}|$$

[Spirtes 2013]

Example: Observational Distribution

$$V_1 \leftarrow U_1$$
$$V_2 \leftarrow U_2$$
$$V_3 \leftarrow V_1 V_4 + U_3$$
$$V_4 \leftarrow V_2 V_3 + U_4$$

For fixed $\boldsymbol{U}=\boldsymbol{u}$, there's a unique equilibrium!

$$V_1(\mathbf{U}) = U_1$$
$$V_2(\mathbf{U}) = U_2$$
$$V_3(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2}$$
$$V_4(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2} U_2 + U_4$$

Potential Responses

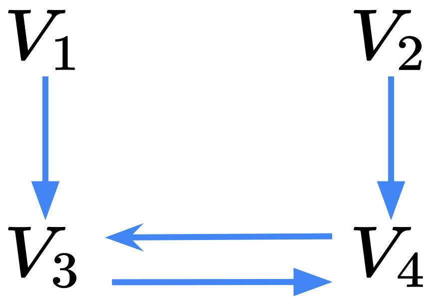
$$U_i \sim \mathcal{N}(0,1) \quad iid.$$
$$P(U_i = u_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u_i^2}$$

$$P(\mathbf{V}) = \left(\frac{1}{4\pi^2}\right) e^{-\frac{v_1^2}{2}} e^{-\frac{v_2^2}{2}} e^{-\frac{(v_3 - v_1 v_4)^2}{2}} e^{-\frac{(v_4 - v_2 v_3)^2}{2}} \left|(1 - v_1 v_2)^{-1}\right|$$

[Spirtes 2013
Bongers et al. 2021
Forr’e et al. 2017]

Example: d-separation

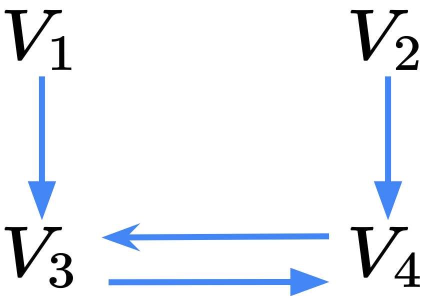
$$
\begin{aligned}
V_1 &\leftarrow U_1 \\
V_2 &\leftarrow U_2 \\
V_3 &\leftarrow V_1 V_4 + U_3 \\
V_4 &\leftarrow V_2 V_3 + U_4
\end{aligned}
$$



[Spirtes 2013
Bongers et al. 2021
Forr’e et al. 2017]

Example: d-separation

$$
\begin{aligned}
V_1 &\leftarrow U_1 \\
V_2 &\leftarrow U_2 \\
V_3 &\leftarrow V_1 V_4 + U_3 \\
V_4 &\leftarrow V_2 V_3 + U_4
\end{aligned}
$$

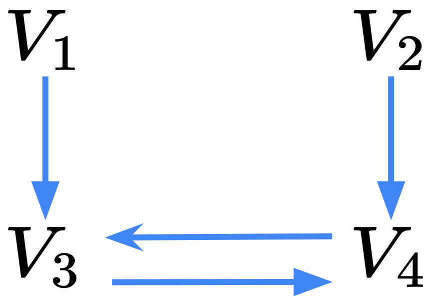


$$(V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_G \text{?}$$

Example: d-separation

[Spirtes 2013
Bongers et al. 2021
Forr’e et al. 2017]

$$\begin{aligned} V_1 &\leftarrow U_1 \\ V_2 &\leftarrow U_2 \\ V_3 &\leftarrow V_1 V_4 + U_3 \\ V_4 &\leftarrow V_2 V_3 + U_4 \end{aligned}$$



2 edges means 2 paths

$$(V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_G \textbf{?}$$

[Spirtes 2013
Bongers et al. 2021
Forr’e et al. 2017]

Example: d-separation

$$
\begin{aligned}
V_1 &\leftarrow U_1 \\
V_2 &\leftarrow U_2 \\
V_3 &\leftarrow V_1 V_4 + U_3 \\
V_4 &\leftarrow V_2 V_3 + U_4
\end{aligned}
$$

$$
\begin{tikzcd}
V_1 \arrow[d] & V_2 \arrow[d] \\
V_3 \arrow[r, shift right] & V_4 \arrow[l, shift right]
\end{tikzcd}
$$

2 edges means 2 paths

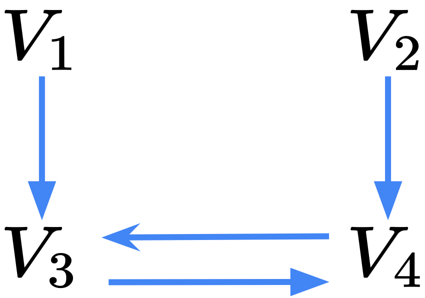
$$(V_1 \perp\!\!\!\perp V_2 \,|\, V_3, V_4)_G\,?$$

Path 1: $V_1 \rightarrow \boxed{V_2} \rightarrow \boxed{V_3} \leftarrow V_4$

Example: d-separation

[Spirtes 2013
Bongers et al. 2021
Forr’e et al. 2017]

$$\boxed{\begin{aligned} V_1 &\leftarrow U_1 \\ V_2 &\leftarrow U_2 \\ V_3 &\leftarrow V_1 V_4 + U_3 \\ V_4 &\leftarrow V_2 V_3 + U_4 \end{aligned}}$$



2 edges means 2 paths

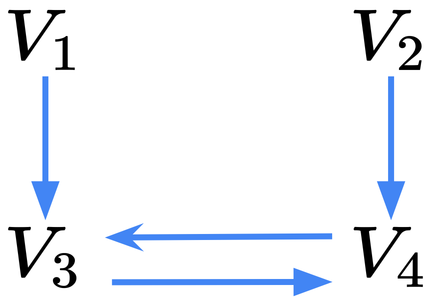
$$(V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_G \textbf{?}$$

Path 1: $V_1 \rightarrow \boxed{V_2} \rightarrow \boxed{V_3} \leftarrow V_4$ *Blocked*

Example: d-separation

[Spirtes 2013
Bongers et al. 2021
Forr’e et al. 2017]

$$V_1 \leftarrow U_1$$
$$V_2 \leftarrow U_2$$
$$V_3 \leftarrow V_1 V_4 + U_3$$
$$V_4 \leftarrow V_2 V_3 + U_4$$



2 edges means 2 paths

$$(V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_G \textbf{?}$$

Path 1: $V_1 \rightarrow \boxed{V_2} \rightarrow \boxed{V_3} \leftarrow V_4$ *Blocked*

Path 2: $V_1 \rightarrow \boxed{V_2} \leftarrow \boxed{V_3} \leftarrow V_4$

[Spirtes 2013
Bongers et al. 2021
Forr'e et al. 2017]

Example: d-separation

$$
\begin{aligned}
V_1 &\leftarrow U_1 \\
V_2 &\leftarrow U_2 \\
V_3 &\leftarrow V_1 V_4 + U_3 \\
V_4 &\leftarrow V_2 V_3 + U_4
\end{aligned}
$$

$$
\begin{tikzcd}
V_1 \arrow[d] & V_2 \arrow[d] \\
V_3 \arrow[r, shift right] & V_4 \arrow[l, shift right]
\end{tikzcd}
$$

2 edges means 2 paths

$$(V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_G \textbf{?}$$

Path 1: $V_1 \rightarrow \boxed{V_2} \rightarrow \boxed{V_3} \leftarrow V_4$ *Blocked*

Path 2: $V_1 \rightarrow \boxed{V_2} \leftarrow \boxed{V_3} \leftarrow V_4$ *Blocked*

[Spirtes 2013
Bongers et al. 2021
Forr’e et al. 2017]

Example: d-separation

$$
\begin{aligned}
V_1 &\leftarrow U_1 \\
V_2 &\leftarrow U_2 \\
V_3 &\leftarrow V_1 V_4 + U_3 \\
V_4 &\leftarrow V_2 V_3 + U_4
\end{aligned}
$$

$$
\begin{tikzcd}
V_1 \arrow[d] & V_2 \arrow[d] \\
V_3 \arrow[r, shift right] & V_4 \arrow[l, shift right]
\end{tikzcd}
$$

2 edges means 2 paths

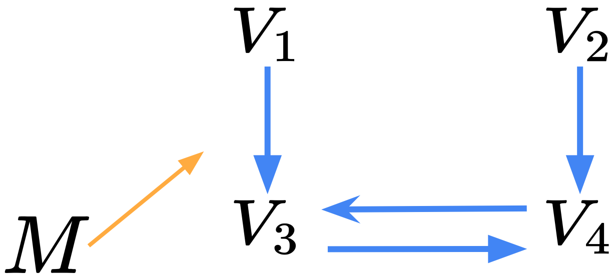
$(V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_G$ ✓

Path 1: $V_1 \rightarrow \boxed{V_2} \rightarrow \boxed{V_3} \leftarrow V_4$ *Blocked*

Path 2: $V_1 \rightarrow \boxed{V_2} \leftarrow \boxed{V_3} \leftarrow V_4$ *Blocked*

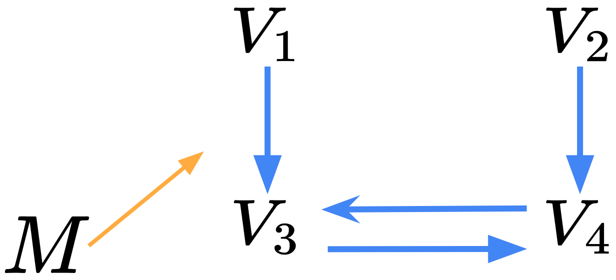
[Spirtes 2013]

Example: Directed Global Markov Property



[Spirtes 2013]

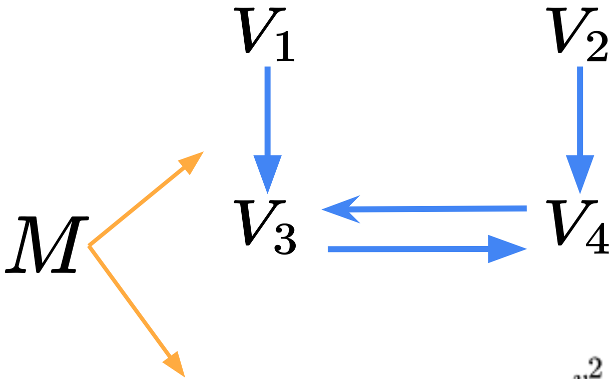
Example: Directed Global Markov Property



$$\xrightarrow{\text{d-separation}} (V_1 \perp\!\!\!\perp V_2 \mid V_3, V_4)_G$$

[Spirtes 2013]

Example: Directed Global Markov Property

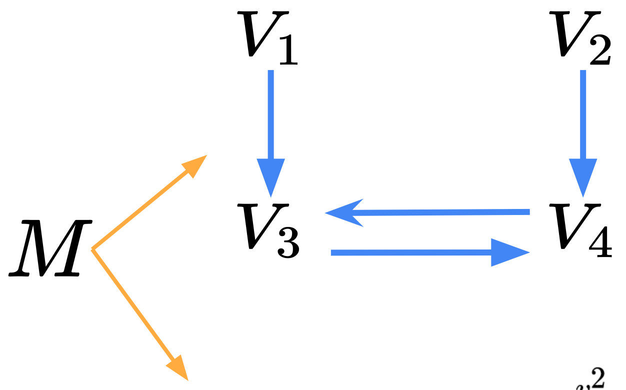


$$\xrightarrow{\text{d-separation}} (V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_G$$

$$P(\mathbf{V}) = (\frac{1}{4\pi^2})e^{-\frac{v_1^2}{2}}e^{-\frac{v_2^2}{2}}e^{-\frac{(v_3-v_1v_4)^2}{2}}e^{-\frac{(v_4-v_2v_3)^2}{2}}|(1-v_1v_2)^{-1}|$$

[Spirtes 2013]

Example: Directed Global Markov Property



$$\xrightarrow{\text{d-separation}} (V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_G$$

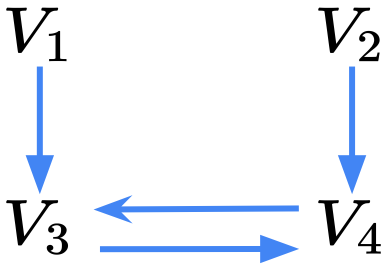
$$P(\mathbf{V}) = \left(\frac{1}{4\pi^2}\right) e^{-\frac{v_1^2}{2}} e^{-\frac{v_2^2}{2}} \underline{e^{-\frac{(v_3 - v_1 v_4)^2}{2}} e^{-\frac{(v_4 - v_2 v_3)^2}{2}} \left|(1 - v_1 v_2)^{-1}\right|}$$

Not Possible to Factor!

[Spirtes 2013]

Example: Directed Global Markov Property

M



d-separation

$$(V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_G$$

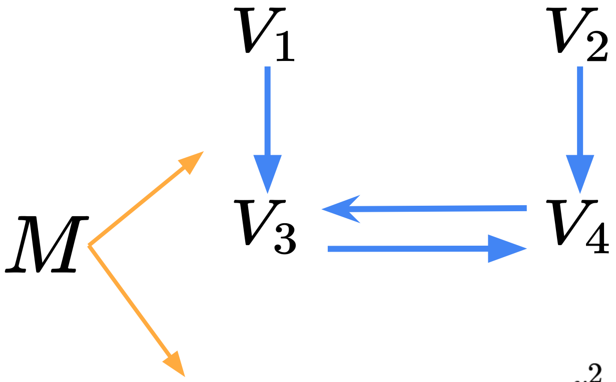
$$P(\mathbf{V}) = (\frac{1}{4\pi^2})e^{-\frac{v_1^2}{2}}\, e^{-\frac{v_2^2}{2}}\, \underline{e^{-\frac{(v_3 - v_1 v_4)^2}{2}}\, e^{-\frac{(v_4 - v_2 v_3)^2}{2}}\, |(1 - v_1 v_2)^{-1}|}$$

Not Possible to Factor!

So $(V_1 \not\perp\!\!\!\perp V_2 | V_3, V_4)_P$

[Spirtes 2013]

Example: Directed Global Markov Property



$$\xrightarrow{\text{d-separation}} (V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_G$$

$$P(\mathbf{V}) = \left(\frac{1}{4\pi^2}\right) e^{-\frac{v_1^2}{2}} e^{-\frac{v_2^2}{2}} \underline{e^{-\frac{(v_3 - v_1 v_4)^2}{2}} e^{-\frac{(v_4 - v_2 v_3)^2}{2}} \left|(1 - v_1 v_2)^{-1}\right|}$$

Not Possible to Factor!

So $(V_1 \not\perp\!\!\!\perp V_2 | V_3, V_4)_P$

Let’s double check with numerics…

Numerics: Directed Global Markov Property

$$V_1(\mathbf{U}) = U_1$$
$$V_2(\mathbf{U}) = U_2$$
$$V_3(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2}$$
$$V_4(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2} U_2 + U_4$$

$$P(\mathbf{U}) \xrightarrow{V_i(\mathbf{U})} P(\mathbf{V})$$

- Monte Carlo: sample from $P(\mathbf{U})$ and apply $V_i(\mathbf{U})$

Numerics: Directed Global Markov Property

$$V_1(\mathbf{U}) = U_1$$
$$V_2(\mathbf{U}) = U_2$$
$$V_3(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2}$$
$$V_4(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2} U_2 + U_4$$

💡 $P(\mathbf{U}) \xrightarrow{V_i(\mathbf{U})} P(\mathbf{V})$

- Monte Carlo: sample from $P(\mathbf{U})$ and apply $V_i(\mathbf{U})$

- Test $(V_1 \overset{?}{\perp\!\!\!\perp} V_2 | V_3, V_4)_P$ numerically…
  - 🤔 Conditional independence + continuous domain? That's hard…

[FCIT: Chalupka et al. 2018]

Numerics: Directed Global Markov Property

$$V_1(\mathbf{U}) = U_1$$
$$V_2(\mathbf{U}) = U_2$$
$$V_3(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2}$$
$$V_4(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2} U_2 + U_4$$

💡 $P(\mathbf{U}) \xrightarrow{V_i(\mathbf{U})} P(\mathbf{V})$

➢ Monte Carlo: sample from $P(\mathbf{U})$ and apply $V_i(\mathbf{U})$

- Test $(V_1 \overset{?}{\perp\!\!\!\perp} V_2 | V_3, V_4)_P$ numerically…

🤔 Conditional independence + continuous domain? That's hard…

- …using FCIT

(a fast independence test which tries to predict V_1 from V_2 after V_3 and V_4 are provided)

[FCIT: Chalupka et al. 2018]

Numerics: Directed Global Markov Property

$$V_1(\mathbf{U}) = U_1$$
$$V_2(\mathbf{U}) = U_2$$
$$V_3(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2}$$
$$V_4(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2} U_2 + U_4$$

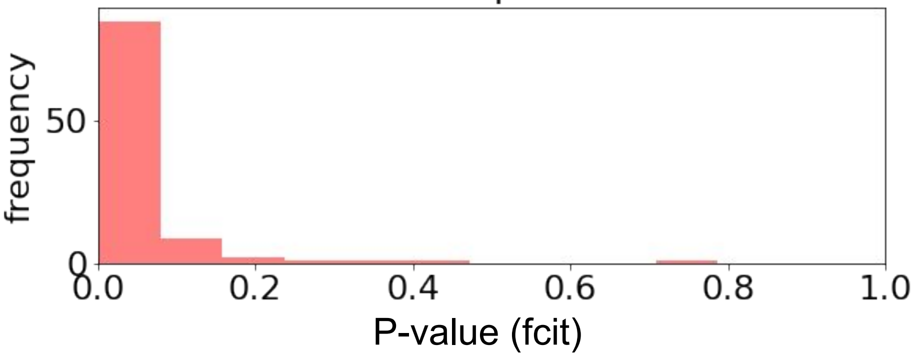
$$P(\mathbf{U}) \xrightarrow{V_i(\mathbf{U})} P(\mathbf{V})$$

➢ Monte Carlo: sample from $P(\mathbf{U})$ and apply $V_i(\mathbf{U})$

- Test $(V_1 \overset{?}{\perp\!\!\!\perp} V_2 | V_3, V_4)_P$ numerically…
  - 🤔 Conditional independence + continuous domain? That's hard…
- …using FCIT

(a fast independence test which tries to predict V_1 from V_2 after V_3 and V_4 are provided)

$(V_1 \not\perp\!\!\!\perp V_2 | V_3, V_4)$



[FCIT: Chalupka et al. 2018]

Numerics: Directed Global Markov Property

$$V_1(\mathbf{U}) = U_1$$
$$V_2(\mathbf{U}) = U_2$$
$$V_3(\mathbf{U}) = \frac{U_1U_4+U_3}{1-U_1U_2}$$
$$V_4(\mathbf{U}) = \frac{U_1U_4+U_3}{1-U_1U_2}U_2 + U_4$$

$$P(\mathbf{U}) \xrightarrow{V_i(\mathbf{U})} P(\mathbf{V})$$

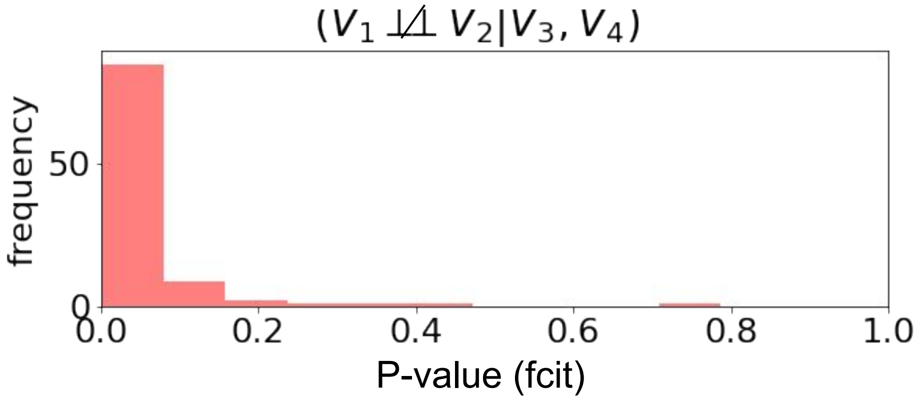
- Monte Carlo: sample from $P(\mathbf{U})$ and apply $V_i(\mathbf{U})$

- Test $(V_1 \overset{?}{\perp\!\!\!\perp} V_2 | V_3, V_4)_P$ numerically…

  Conditional independence + continuous domain? That's hard…

- …using FCIT

(a fast independence test which tries to predict V_1 from V_2 after V_3 and V_4 are provided)



Low p-values $\longrightarrow$ conditional dependence

[FCIT: Chalupka et al. 2018]

Numerics: Directed Global Markov Property

$$V_1(\mathbf{U}) = U_1$$
$$V_2(\mathbf{U}) = U_2$$
$$V_3(\mathbf{U}) = \frac{U_1U_4+U_3}{1-U_1U_2}$$
$$V_4(\mathbf{U}) = \frac{U_1U_4+U_3}{1-U_1U_2}U_2 + U_4$$

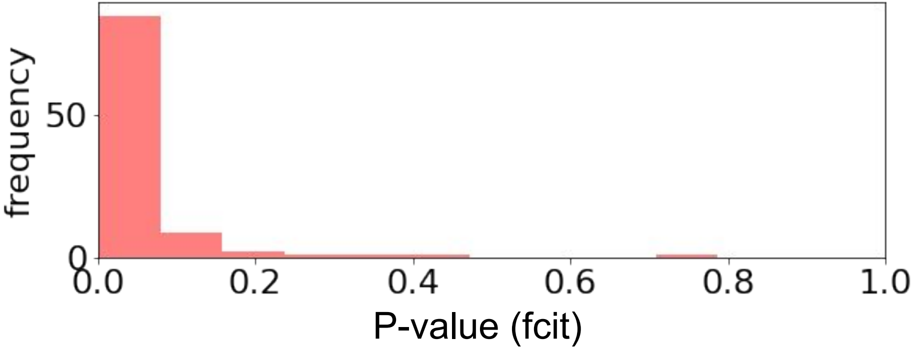
$$P(\mathbf{U}) \xrightarrow{V_i(\mathbf{U})} P(\mathbf{V})$$

- ➢ Monte Carlo: sample from $P(\mathbf{U})$ and apply $V_i(\mathbf{U})$

- Test $(V_1 \overset{?}{\perp\!\!\!\perp} V_2 | V_3, V_4)_P$ numerically…
  - 🤔 Conditional independence + continuous domain? That's hard…
- …using FCIT

(a fast independence test which tries to predict V_1 from V_2 after V_3 and V_4 are provided)

$(V_1 \not\perp\!\!\!\perp V_2 | V_3, V_4)$



Low p-values ⟶ conditional dependence

$$(V_1 \not\perp\!\!\!\perp V_2 | V_3, V_4)_P$$

$$(V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_G \not\Rightarrow (V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_P$$

Example: Asymptotics

$$V_1(\mathbf{U}) = U_1$$
$$V_2(\mathbf{U}) = U_2$$
$$V_3(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2}$$
$$V_4(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2} U_2 + U_4$$

$$P(\mathbf{U}) \xrightarrow{V_i(\mathbf{U})} P(\mathbf{V})$$

Example: Asymptotics

$$V_1(\mathbf{U}) = U_1$$
$$V_2(\mathbf{U}) = U_2$$
$$V_3(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2}$$
$$V_4(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2} U_2 + U_4$$

$$P(\mathbf{U}) \xrightarrow{V_i(\mathbf{U})} P(\mathbf{V})$$

🤔 $V_i(\mathbf{U})$ is found by fixing $\mathbf{u}$ and solving for the equilibrium of the system

Example: Asymptotics

$$V_1(\mathbf{U}) = U_1$$
$$V_2(\mathbf{U}) = U_2$$
$$V_3(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2}$$
$$V_4(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2} U_2 + U_4$$

$$P(\mathbf{U}) \xrightarrow{V_i(\mathbf{U})} P(\mathbf{V})$$

- 🤔 $V_i(\mathbf{U})$ is found by fixing $\mathbf{u}$ and solving for the equilibrium of the system
- 🤔 For linear SCMs*, these $V_i(\mathbf{U})$ at equilibrium preserve the conditional independencies of $\boldsymbol{G}$ (hence, d-separation is valid)

*If $P(\mathbf{V}) > 0$ everywhere, and there's no trivial dependencies

Example: Asymptotics

$$V_1(\mathbf{U}) = U_1$$
$$V_2(\mathbf{U}) = U_2$$
$$V_3(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2}$$
$$V_4(\mathbf{U}) = \frac{U_1 U_4 + U_3}{1 - U_1 U_2} U_2 + U_4$$

$$P(\mathbf{U}) \xrightarrow{V_i(\mathbf{U})} P(\mathbf{V})$$

- 🤔 $V_i(\mathbf{U})$ is found by fixing $\mathbf{u}$ and solving for the equilibrium of the system
- 🤔 For linear SCMs*, these $V_i(\mathbf{U})$ at equilibrium preserve the conditional independencies of $\boldsymbol{G}$ (hence, d-separation is valid)

Perhaps nonlinear SCMs whose *asymptomatic dynamics* are bounded by a linear SCM will also inherit these nice properties of $V_i(\mathbf{U})$

*If $P(\mathbf{V}) > 0$ everywhere, and there's no trivial dependencies

Example: Asymptotics

$$V_1(\mathbf{U}) = U_1$$
$$V_2(\mathbf{U}) = U_2$$
$$V_3(\mathbf{U}) = \frac{U_1U_4+U_3}{1-U_1U_2}$$
$$V_4(\mathbf{U}) = \frac{U_1U_4+U_3}{1-U_1U_2}U_2 + U_4$$

$$P(\mathbf{U}) \xrightarrow{V_i(\mathbf{U})} P(\mathbf{V})$$

- 🤔 $V_i(\mathbf{U})$ is found by fixing $\mathbf{u}$ and solving for the equilibrium of the system
- 🤔 For linear SCMs*, these $V_i(\mathbf{U})$ at equilibrium preserve the conditional independencies of $\boldsymbol{G}$ (hence, d-separation is valid)

💡 Perhaps nonlinear SCMs whose *asymptomatic dynamics* are bounded by a linear SCM will also inherit these nice properties of $V_i(\mathbf{U})$

- What happens if we try to bound this example??

*If $P(\mathbf{V}) > 0$ everywhere, and there's no trivial dependencies

Example: Linear Bounding

Bound each f_i by $\quad g_i = (\sum_j a_{ij} V_j) + U_i \quad a_{ij} = \max_{\mathbf{v}} |\frac{\partial f_i}{\partial V_j}(\mathbf{v}, \mathbf{u})|$

Example: Linear Bounding

Bound each f_i by $g_i = (\sum_j a_{ij} V_j) + U_i$ $a_{ij} = \max_{\mathbf{v}} |\frac{\partial f_i}{\partial V_j}(\mathbf{v}, \mathbf{u})|$

$$f_1 : V_1 \leftarrow U_1$$
$$f_2 : V_2 \leftarrow U_2$$
$$f_3 : V_3 \leftarrow V_1 V_4 + U_3$$
$$f_4 : V_4 \leftarrow V_2 V_3 + U_4$$

Example: Linear Bounding

Bound each f_i by $g_i = \left(\sum_j a_{ij} V_j\right) + U_i$ $\quad a_{ij} = \max_{\mathbf{v}} \left|\frac{\partial f_i}{\partial V_j}(\mathbf{v}, \mathbf{u})\right|$

$$
\begin{aligned}
f_1 &: V_1 \leftarrow U_1 \\
f_2 &: V_2 \leftarrow U_2 \\
f_3 &: V_3 \leftarrow V_1 V_4 + U_3 \\
f_4 &: V_4 \leftarrow V_2 V_3 + U_4
\end{aligned}
$$

$$
\begin{aligned}
g_1 &: V_1 \leftarrow U_1 \\
g_2 &: V_2 \leftarrow U_2 \\
g_3 &: V_3 \leftarrow \boxed{a_{31}} V_1 + \boxed{a_{34}} V_4 + U_3 \\
g_4 &: V_4 \leftarrow \boxed{a_{42}} V_2 + \boxed{a_{43}} V_3 + U_4
\end{aligned}
$$

Example: Linear Bounding

Bound each f_i by $g_i = \left(\sum_j a_{ij} V_j\right) + U_i$ $\quad a_{ij} = \max_{\mathbf{v}} \left|\frac{\partial f_i}{\partial V_j}(\mathbf{v}, \mathbf{u})\right|$

$f_1 : V_1 \leftarrow U_1 \longrightarrow a_{1j} = 0 \longrightarrow g_1 : V_1 \leftarrow U_1$

$f_2 : V_2 \leftarrow U_2 \longrightarrow a_{2j} = 0 \longrightarrow g_2 : V_2 \leftarrow U_2$

$f_3 : V_3 \leftarrow V_1 V_4 + U_3 \longrightarrow g_3 : V_3 \leftarrow \boxed{a_{31}} V_1 + \boxed{a_{34}} V_4 + U_3$

$f_4 : V_4 \leftarrow V_2 V_3 + U_4 \longrightarrow g_4 : V_4 \leftarrow \boxed{a_{42}} V_2 + \boxed{a_{43}} V_3 + U_4$

$a_{31} = \max_{\mathbf{v}} |V_4| \quad a_{34} = \max_{\mathbf{v}} |V_1| \quad a_{42} = \max_{\mathbf{v}} |V_3| \quad a_{43} = \max_{\mathbf{v}} |V_2|$

Example: Linear Bounding

Bound each f_i by $g_i = \left(\sum_j a_{ij} V_j\right) + U_i$ $\qquad a_{ij} = \max_{\mathbf{v}} \left|\frac{\partial f_i}{\partial V_j}(\mathbf{v}, \mathbf{u})\right|$

$$f_1 : V_1 \leftarrow U_1 \longrightarrow a_{1j} = 0 \longrightarrow g_1 : V_1 \leftarrow U_1$$

$$f_2 : V_2 \leftarrow U_2 \longrightarrow a_{2j} = 0 \longrightarrow g_2 : V_2 \leftarrow U_2$$

$$f_3 : V_3 \leftarrow V_1 V_4 + U_3 \longrightarrow g_3 : V_3 \leftarrow \boxed{a_{31}} V_1 + \boxed{a_{34}} V_4 + U_3$$

$$f_4 : V_4 \leftarrow V_2 V_3 + U_4 \longrightarrow g_4 : V_4 \leftarrow \boxed{a_{42}} V_2 + \boxed{a_{43}} V_3 + U_4$$

$$a_{31} = \max_{\mathbf{v}} |V_4| \quad a_{34} = \max_{\mathbf{v}} |V_1| \quad a_{42} = \max_{\mathbf{v}} |V_3| \quad a_{43} = \max_{\mathbf{v}} |V_2|$$

- Some terms go to infinity, so we can't bound this example!
  - (these correspond directly to unboundedness of the potential response function)

Example: Linear Bounding

Bound each f_i by $g_i = (\sum_j a_{ij} V_j) + U_i$ $\qquad a_{ij} = \max_{\mathbf{v}} |\frac{\partial f_i}{\partial V_j}(\mathbf{v}, \mathbf{u})|$

$$f_1 : V_1 \leftarrow U_1 \longrightarrow a_{1j} = 0 \longrightarrow g_1 : V_1 \leftarrow U_1$$
$$f_2 : V_2 \leftarrow U_2 \longrightarrow a_{2j} = 0 \longrightarrow g_2 : V_2 \leftarrow U_2$$
$$f_3 : V_3 \leftarrow V_1 V_4 + U_3 \longrightarrow g_3 : V_3 \leftarrow \boxed{a_{31}} V_1 + \boxed{a_{34}} V_4 + U_3$$
$$f_4 : V_4 \leftarrow V_2 V_3 + U_4 \longrightarrow g_4 : V_4 \leftarrow \boxed{a_{42}} V_2 + \boxed{a_{43}} V_3 + U_4$$

$$a_{31} = \max_{\mathbf{v}} |V_4| \quad a_{34} = \max_{\mathbf{v}} |V_1| \quad a_{42} = \max_{\mathbf{v}} |V_3| \quad a_{43} = \max_{\mathbf{v}} |V_2|$$

- Some terms go to infinity, so we can’t bound this example!
  - (these correspond directly to unboundedness of the potential response function)
- What if we tweak the example?

Example: Intrinsic SCM

$$M \begin{cases} \mathbf{u}, \mathbf{v} \in [-0.5, 0.5] \\ U_i \sim \mathcal{N}(0,1) \cap [-0.5, 0.5] \quad iid. \\ f_1 : V_1 \leftarrow U_1 \\ f_2 : V_2 \leftarrow U_2 \\ f_3 : V_3 \leftarrow V_1 V_4 + U_3 \\ f_4 : V_4 \leftarrow V_2 V_3 + U_4 \end{cases}$$

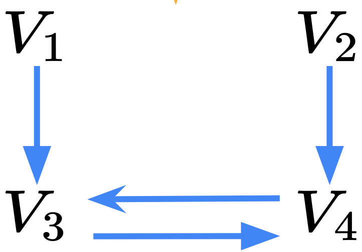
Example: Intrinsic SCM

$$M\left\{\begin{array}{l}\mathbf{u}, \mathbf{v} \in [-0.5, 0.5] \\ U_i \sim \mathcal{N}(0,1) \cap [-0.5, 0.5] \quad iid. \\ \\ f_1 : V_1 \leftarrow U_1 \\ f_2 : V_2 \leftarrow U_2 \\ f_3 : V_3 \leftarrow V_1 V_4 + U_3 \\ f_4 : V_4 \leftarrow V_2 V_3 + U_4 \end{array}\right.$$

$$\begin{tikzcd} V_1 \arrow[d] & V_2 \arrow[d] \\ V_3 \arrow[r, shift right] & V_4 \arrow[l, shift right] \end{tikzcd}$$

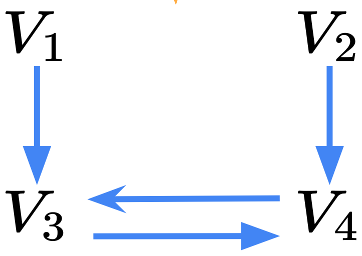
Example: Intrinsic SCM

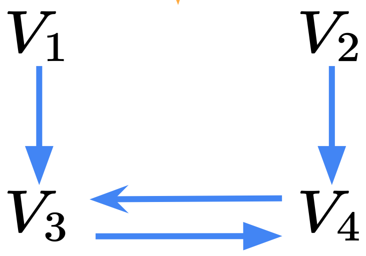
$$
M \begin{cases}
\mathbf{u}, \mathbf{v} \in [-0.5, 0.5] \\
U_i \sim \mathcal{N}(0,1) \cap [-0.5, 0.5] \quad iid. \\
f_1 : V_1 \leftarrow U_1 \\
f_2 : V_2 \leftarrow U_2 \\
f_3 : V_3 \leftarrow V_1 V_4 + U_3 \\
f_4 : V_4 \leftarrow V_2 V_3 + U_4
\end{cases}
\qquad
\mathcal{L}(M) \begin{cases}
\mathbf{u}, \mathbf{v} \in [-0.5, 0.5] \\
U_i \sim \mathcal{N}(0,1) \cap [-0.5, 0.5] \quad iid. \\
g_1 : V_1 \leftarrow U_1 \\
g_2 : V_2 \leftarrow U_2 \\
g_3 : V_3 \leftarrow a_{31} V_1 + a_{34} V_4 + U_3 \\
g_4 : V_4 \leftarrow a_{42} V_2 + a_{43} V_3 + U_4
\end{cases}
$$



Example: Intrinsic SCM

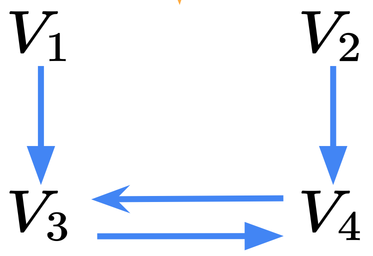
$$M \left\{ \begin{array}{l} \color{red}{\mathbf{u}, \mathbf{v} \in [-0.5, 0.5]} \\ \color{red}{U_i \sim \mathcal{N}(0,1) \cap [-0.5, 0.5] \quad iid.} \\ f_1 : V_1 \leftarrow U_1 \\ f_2 : V_2 \leftarrow U_2 \\ f_3 : V_3 \leftarrow V_1 V_4 + U_3 \\ f_4 : V_4 \leftarrow V_2 V_3 + U_4 \end{array} \right. \longrightarrow \quad \mathcal{L}(M) \left\{ \begin{array}{l} \color{red}{\mathbf{u}, \mathbf{v} \in [-0.5, 0.5]} \\ \color{red}{U_i \sim \mathcal{N}(0,1) \cap [-0.5, 0.5] \quad iid.} \\ g_1 : V_1 \leftarrow U_1 \\ g_2 : V_2 \leftarrow U_2 \\ g_3 : V_3 \leftarrow a_{31} V_1 + a_{34} V_4 + U_3 \\ g_4 : V_4 \leftarrow a_{42} V_2 + a_{43} V_3 + U_4 \end{array} \right.$$



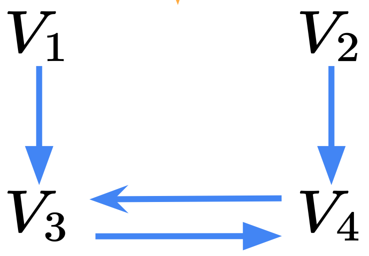


Example: Intrinsic SCM

$$M \begin{cases} \mathbf{u}, \mathbf{v} \in [-0.5, 0.5] \\ U_i \sim \mathcal{N}(0,1) \cap [-0.5, 0.5] \quad iid. \\ f_1 : V_1 \leftarrow U_1 \\ f_2 : V_2 \leftarrow U_2 \\ f_3 : V_3 \leftarrow V_1 V_4 + U_3 \\ f_4 : V_4 \leftarrow V_2 V_3 + U_4 \end{cases} \longrightarrow \mathcal{L}(M) \begin{cases} \mathbf{u}, \mathbf{v} \in [-0.5, 0.5] \\ U_i \sim \mathcal{N}(0,1) \cap [-0.5, 0.5] \quad iid. \\ g_1 : V_1 \leftarrow U_1 \\ g_2 : V_2 \leftarrow U_2 \\ g_3 : V_3 \leftarrow a_{31} V_1 + a_{34} V_4 + U_3 \\ g_4 : V_4 \leftarrow a_{42} V_2 + a_{43} V_3 + U_4 \end{cases}$$



$(V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_G$ ✓



Example: Intrinsic SCM

$$M \left\{ \begin{array}{l} \mathbf{u}, \mathbf{v} \in [-0.5, 0.5] \\ U_i \sim \mathcal{N}(0,1) \cap [-0.5, 0.5] \quad iid. \\ f_1 : V_1 \leftarrow U_1 \\ f_2 : V_2 \leftarrow U_2 \\ f_3 : V_3 \leftarrow V_1 V_4 + U_3 \\ f_4 : V_4 \leftarrow V_2 V_3 + U_4 \end{array} \right. \longrightarrow \mathcal{L}(M) \left\{ \begin{array}{l} \mathbf{u}, \mathbf{v} \in [-0.5, 0.5] \\ U_i \sim \mathcal{N}(0,1) \cap [-0.5, 0.5] \quad iid. \\ g_1 : V_1 \leftarrow U_1 \\ g_2 : V_2 \leftarrow U_2 \\ g_3 : V_3 \leftarrow a_{31} V_1 + a_{34} V_4 + U_3 \\ g_4 : V_4 \leftarrow a_{42} V_2 + a_{43} V_3 + U_4 \end{array} \right.$$

$$\begin{tikzcd} V_1 \arrow[d] & V_2 \arrow[d] \\ V_3 \arrow[r, shift right] & V_4 \arrow[l, shift right] \end{tikzcd} \longleftarrow (V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_G \checkmark \longrightarrow \begin{tikzcd} V_1 \arrow[d] & V_2 \arrow[d] \\ V_3 \arrow[r, shift right] & V_4 \arrow[l, shift right] \end{tikzcd}$$

$$(V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_P \checkmark$$

Example: Intrinsic SCM

$$M\begin{cases} \mathbf{u},\mathbf{v}\in[-0.5,0.5] \\ U_i\sim\mathcal{N}(0,1)\cap[-0.5,0.5] \quad iid. \\ f_1 : V_1 \leftarrow U_1 \\ f_2 : V_2 \leftarrow U_2 \\ f_3 : V_3 \leftarrow V_1 V_4 + U_3 \\ f_4 : V_4 \leftarrow V_2 V_3 + U_4 \end{cases} \longrightarrow \mathcal{L}(M)\begin{cases} \mathbf{u},\mathbf{v}\in[-0.5,0.5] \\ U_i\sim\mathcal{N}(0,1)\cap[-0.5,0.5] \quad iid. \\ g_1 : V_1 \leftarrow U_1 \\ g_2 : V_2 \leftarrow U_2 \\ g_3 : V_3 \leftarrow a_{31}V_1 + a_{34}V_4 + U_3 \\ g_4 : V_4 \leftarrow a_{42}V_2 + a_{43}V_3 + U_4 \end{cases}$$

$$\begin{tikzcd} V_1 \arrow[d] & V_2 \arrow[d] \\ V_3 \arrow[r, shift right] & V_4 \arrow[l, shift right] \end{tikzcd} \longleftarrow (V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_G \checkmark \longrightarrow \begin{tikzcd} V_1 \arrow[d] & V_2 \arrow[d] \\ V_3 \arrow[r, shift right] & V_4 \arrow[l, shift right] \end{tikzcd}$$

$(V_1 \overset{?}{\perp\!\!\!\perp} V_2 | V_3, V_4)_P$?

$(V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_P \checkmark$

Numerics: Intrinsic SCM

$U_i \sim \mathcal{N}(0,1) \cap [-0.5, 0.5] \quad iid.$

$$M \quad \boxed{\begin{aligned} V_1(\mathbf{U}) &= U_1 \\ V_2(\mathbf{U}) &= U_2 \\ V_3(\mathbf{U}) &= \tfrac{U_1U_4+U_3}{1-U_1U_2} \\ V_4(\mathbf{U}) &= \tfrac{U_1U_4+U_3}{1-U_1U_2}U_2 + U_4 \end{aligned}}$$

Numerics: Intrinsic SCM

$U_i \sim \mathcal{N}(0,1) \cap [-0.5, 0.5] \quad iid.$

$$M \quad \boxed{\begin{aligned} V_1(\mathbf{U}) &= U_1 \\ V_2(\mathbf{U}) &= U_2 \\ V_3(\mathbf{U}) &= \tfrac{U_1 U_4 + U_3}{1 - U_1 U_2} \\ V_4(\mathbf{U}) &= \tfrac{U_1 U_4 + U_3}{1 - U_1 U_2} U_2 + U_4 \end{aligned}}$$

Test $(V_1 \overset{?}{\perp\!\!\!\perp} V_2 | V_3, V_4)_P$ again, numerically

- Using $P(\mathbf{U}) \xrightarrow{V_i(\mathbf{U})} P(\mathbf{V})$ but with restricted domain

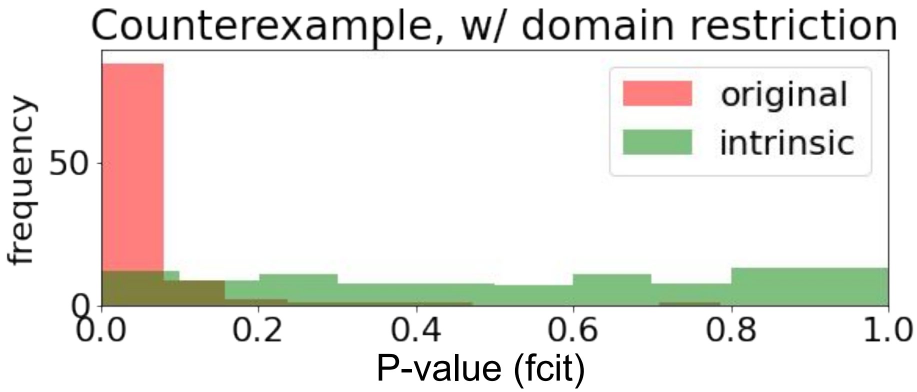
Numerics: Intrinsic SCM

$U_i \sim \mathcal{N}(0,1) \cap [-0.5, 0.5] \quad iid.$

$$M \quad \boxed{\begin{aligned} V_1(\mathbf{U}) &= U_1 \\ V_2(\mathbf{U}) &= U_2 \\ V_3(\mathbf{U}) &= \frac{U_1U_4+U_3}{1-U_1U_2} \\ V_4(\mathbf{U}) &= \frac{U_1U_4+U_3}{1-U_1U_2}U_2 + U_4 \end{aligned}}$$

Test $(V_1 \overset{?}{\perp\!\!\!\perp} V_2 | V_3, V_4)_P$ again, numerically

- Using $P(\mathbf{U}) \xrightarrow{V_i(\mathbf{U})} P(\mathbf{V})$ but with restricted domain



Uniform p-values $\longrightarrow$ conditional independence

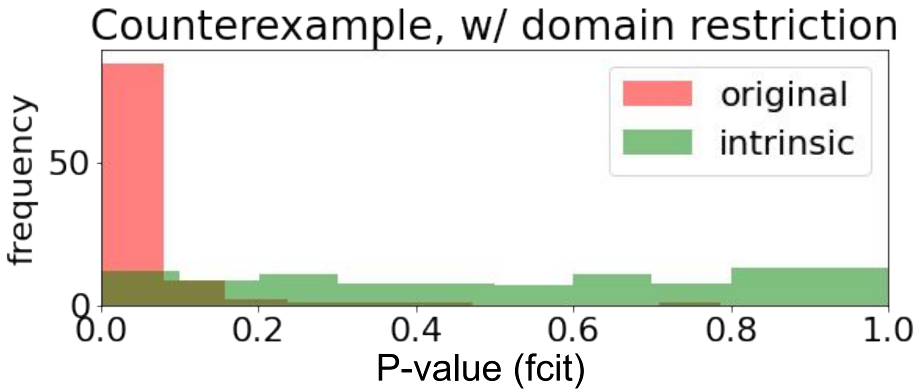
Numerics: Intrinsic SCM

$U_i \sim \mathcal{N}(0,1) \cap [-0.5, 0.5] \quad iid.$

$$M \quad \boxed{\begin{aligned} V_1(\mathbf{U}) &= U_1 \\ V_2(\mathbf{U}) &= U_2 \\ V_3(\mathbf{U}) &= \frac{U_1U_4+U_3}{1-U_1U_2} \\ V_4(\mathbf{U}) &= \frac{U_1U_4+U_3}{1-U_1U_2}U_2 + U_4 \end{aligned}}$$

Test $(V_1 \overset{?}{\perp\!\!\!\perp} V_2 | V_3, V_4)_P$ again, numerically

- Using $P(\mathbf{U}) \xrightarrow{V_i(\mathbf{U})} P(\mathbf{V})$ but with restricted domain



Uniform p-values $\longrightarrow$ conditional independence

$$(V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_P$$

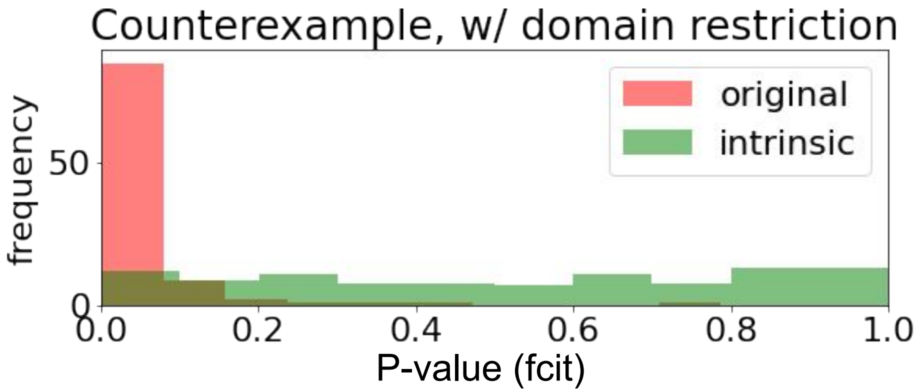
Numerics: Intrinsic SCM

$U_i \sim \mathcal{N}(0,1) \cap [-0.5, 0.5] \quad iid.$

$$M \quad \begin{aligned} V_1(\mathbf{U}) &= U_1 \\ V_2(\mathbf{U}) &= U_2 \\ V_3(\mathbf{U}) &= \frac{U_1 U_4 + U_3}{1 - U_1 U_2} \\ V_4(\mathbf{U}) &= \frac{U_1 U_4 + U_3}{1 - U_1 U_2} U_2 + U_4 \end{aligned}$$

Test $(V_1 \overset{?}{\perp\!\!\!\perp} V_2 | V_3, V_4)_P$ again, numerically

- Using $P(\mathbf{U}) \xrightarrow{V_i(\mathbf{U})} P(\mathbf{V})$ but with restricted domain



Uniform p-values $\longrightarrow$ conditional independence

$$(V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_P$$

$$(V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_G \implies (V_1 \perp\!\!\!\perp V_2 | V_3, V_4)_P$$

was inherited by the linear bound!

Next week’s directions

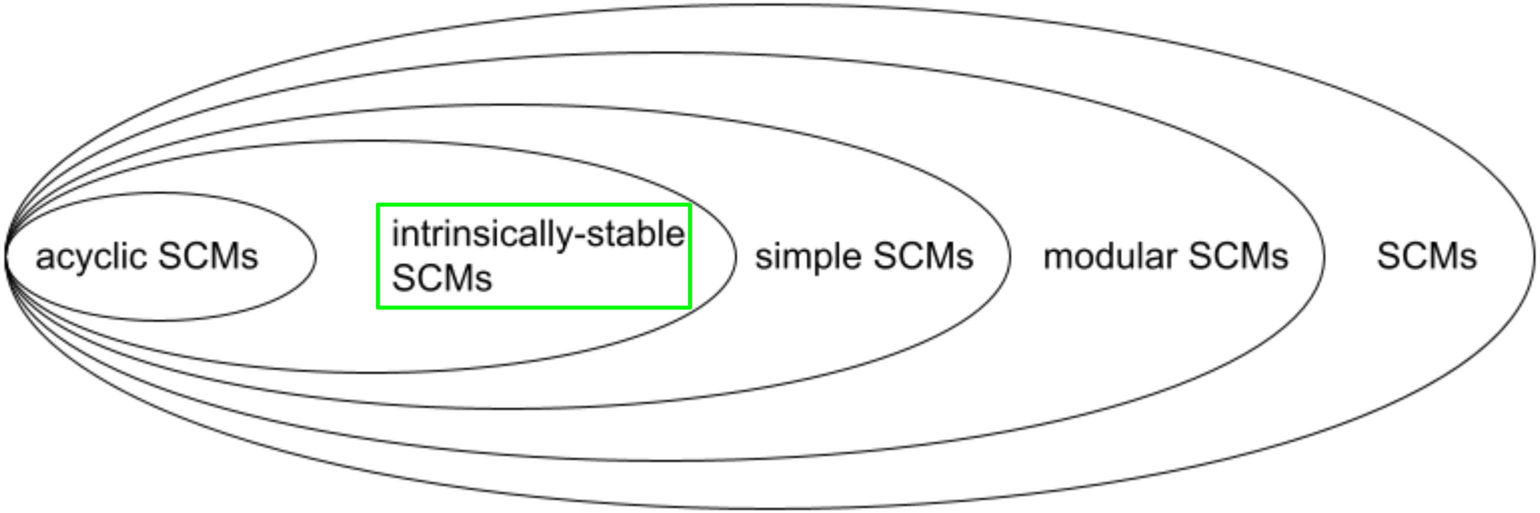
Observational Case: Proof Sketch

Strategy:

- Relate M to a linear SCM L(M) which satisfies the directed global Markov property.
- Show that M and L(M):
  - respect the same graph
  - are observationally equivalent.

$$
\begin{tikzcd}
 & & P_M(V) \arrow[d, equal] & \\
 & & P_{\mathscr{L}(M)}(V) \arrow[r] & (X \perp\!\!\!\perp Y \mid Z)_{P_{\mathscr{L}(M)}} \\
M \arrow[r] \arrow[uurr] \arrow[ddrr] & \mathscr{L}(M) \arrow[ur] \arrow[dr] & & \\
 & & G_{\mathscr{L}}(M) \arrow[r] \arrow[d, equal] & (X \perp\!\!\!\perp Y \mid Z)_{G_{\mathscr{L}(M)}} \arrow[uu, Rightarrow] \\
 & & G_M &
\end{tikzcd}
$$

Conjectured relation to other spaces of cyclic SCMs



Continuing Discussions: Game Theory + Causality

- Lewis Hammond (Oxford, Computer Science: Dphil student)
  - Extending Pearl's causal hierarchy to the game-theoretic domain

- Christian Kroer (Columbia, IOER department: Assistant professor)
  - Research: Game Theory, scalable optimization methods (for finding the potential responses)

- How relates:
  - Game theory is a primary motivation for cyclic causality
  - Local intrinsic stability

Next To-do's

- Prove Observational Equivalence
  - Get examples for each lemma
- Example of finding proper P(U) for linear bound
  - show numerically as well
  - Also an example of when it fails
- Prove (or disprove):
  - acyclic SCMs ⊂ intrinsic SCMs ⊂ simple SCMs
  - Example of each:
    - Linear bound for acyclic
    - Demonstrating unique solvability for each subset of intrinsic example
    - Example of simple SCM that's not intrinsic
- Meet with Christian Kroer, see if worth collaborating

References

- Krzysztof Chalupka, Pietro Perona, and Frederick Eberhardt. Fast conditional independence test for vector variables with large sample sizes, 2018.
- Patrick Forr’e and Joris M. Mooij. Markov properties for graphical models with cycles and latent variables. arXiv:Statistics Theory, 2017.
- Peter Spirtes. Directed cyclic graphical representations of feedback models. CoRR, abs/1302.4982, 2013.
- Stephan Bongers, Patrick Forré, Jonas Peters, and Joris M. Mooij. Foundations of structural causal models with cycles and latent variables. The Annals of Statistics, 49(5):2885 – 2915, 2021.