Confounder selection via iterative graph expansion

F. Richard Guo

Department of Statistics, University of Michigan

May 21, 2026
DEB Biostatistics/Bioinformatics Seminar, UCSF

Qingyuan Zhao
Cambridge

↪ F. Richard Guo and Qingyuan Zhao. "Confounder selection via iterative graph expansion" *Annals of Statistics*, 2026.

Outline

Introduction

Confounder selection

In observational studies, the single most widely used strategy to control for confounding is through covariate adjustment.

$\hookrightarrow$ I will focus on the setting that X is a (point) treatment and Y is an outcome.

☞ Time-varying confounding is NOT in this talk

Confounder selection

In observational studies, the single most widely used strategy to control for confounding is through covariate adjustment.

↪ I will focus on the setting that X is a (point) treatment and Y is an outcome.

☞ Time-varying confounding is NOT in this talk

Goal: Select a set of covariates S that can be measured such that

$$Y(x) \perp\!\!\!\perp X \mid S, \quad \text{for every treatment level } x.$$

▶ conditional exchangeability

Such a set S is called a sufficient adjustment set.

Then, under consistency + positivity assumptions,

$$p(Y(x) \mid X = x, S) = p(Y \mid X = x, S) \implies \mathbb{E}[Y(x)] = \mathbb{E}[\mathbb{E}[Y \mid X = x, S]],$$

which can be estimated with outcome regression, IPW, AIPW, etc.

Objectives in confounder selection

Objectives in confounder selection

1. **Primary**: S, consisting of covariates that are measured or can be measured, is a sufficient adjustment set.

Objectives in confounder selection

1. **Primary**: S, consisting of covariates that are measured or can be measured, is a sufficient adjustment set.
2. **Secondary**: When there are more than one sufficient adjustment sets, choose one among them to optimize some criterion, such as
   - efficiency
   - cardinality
   - cost
   - ...

Objectives in confounder selection

1. **Primary**: S, consisting of covariates that are measured or can be measured, is a sufficient adjustment set.
2. **Secondary**: When there are more than one sufficient adjustment sets, choose one among them to optimize some criterion, such as
   - efficiency in estimating the average treatment effect
   - number of covariates
   - cost
   - ...

☞ I will focus on the **primary objective** in this talk.

I will also use this opportunity to **debunk some myths** about this problem along the way. ↪ See also F. R. Guo, Lundborg, and Zhao (2022) for a recent survey.

Myth 1. This is a solved problem.

Suppose the underlying causal model can be represented by a DAG (directed acyclic graph) $\mathcal{G}$ or more generally an ADMG ($=$ DAG $+$ latent confounders), then some may argue this is a solved problem.

Back-door Criterion (Pearl, 1993) $S \subseteq \mathrm{V} \setminus \{X, Y\}$ is a sufficient adjustment set if

1. S contains no descendant of X,
2. there is no '$X \leftarrow \dots$' path between X and Y that is m-connected given S.

Myth 1. This is a solved problem.

Suppose the underlying causal model can be represented by a DAG (directed acyclic graph) $\mathcal{G}$ or more generally an ADMG (= DAG + latent confounders), then some may argue this is a solved problem.

Back-door Criterion (Pearl, 1993) $S \subseteq \mathrm{V} \setminus \{X, Y\}$ is a sufficient adjustment set if

1. S contains no descendant of X,
2. there is no '$X \leftarrow \dots$' path between X and Y that is m-connected given S.

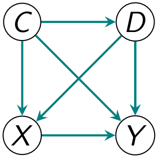
☞ If we require that

S contains no descendants of X $\Longleftarrow$ S only contains pre-treatment covariates,

then the back-door criterion is both sufficient and necessary, i.e.,

$$S \text{ is a sufficient adjustment set} \iff S \text{ meets the back-door criterion.}$$

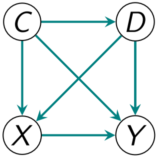
▶ It can be shown that you never have to adjust for any covariate that is post-treatment (Shpitser, VanderWeele, and Robins, 2010).



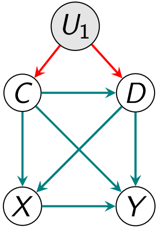
$\{C, D\}$ ✓

Back-door Criterion (Pearl, 1993) $S \subseteq \mathrm{V} \setminus \{X, Y\}$ is a sufficient adjustment set if

1. S contains no descendant of X,
2. there is no '$X \leftarrow \dots$' path between X and Y that is m-connected given S.



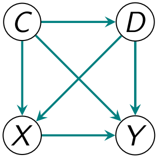
$\{C, D\}$ ✓



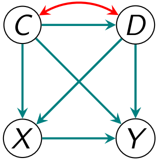
$\{C, D\}$ ✓

Back-door Criterion (Pearl, 1993) $S \subseteq \mathrm{V} \setminus \{X, Y\}$ is a sufficient adjustment set if

1. S contains no descendant of X,
2. there is no '$X \leftarrow \ldots$' path between X and Y that is m-connected given S.



$\{C, D\}$ ✓

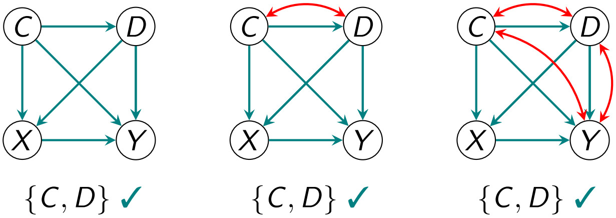


$\{C, D\}$ ✓

ADMG (acyclic directed mixed graph)

Back-door Criterion (Pearl, 1993) $S \subseteq \mathrm{V} \setminus \{X, Y\}$ is a sufficient adjustment set if

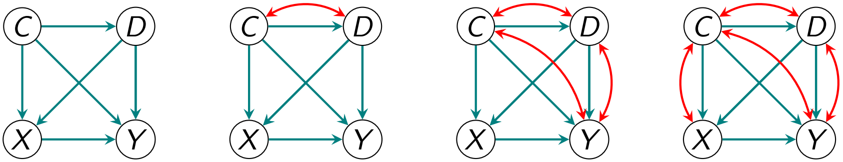
1. S contains no descendant of X,
2. there is no '$X \leftarrow \ldots$' path between X and Y that is m-connected given S.



ADMG (acyclic directed mixed graph)

Back-door Criterion (Pearl, 1993) $S \subseteq \mathrm{V} \setminus \{X, Y\}$ is a sufficient adjustment set if

1. S contains no descendant of X,
2. there is no '$X \leftarrow \dots$' path between X and Y that is m-connected given S.



$\{C, D\}$ ✓ $\{C, D\}$ ✓ $\{C, D\}$ ✓ $\{C, D\}$ ✗

ADMG (acyclic directed mixed graph) No valid adjustment set

Back-door Criterion (Pearl, 1993) $S \subseteq \mathrm{V} \setminus \{X, Y\}$ is a sufficient adjustment set if

1. S contains no descendant of X,
2. there is no '$X \leftarrow \dots$' path between X and Y that is m-connected given S.

Myth 2. Causal discovery can solve this.

Without making strong additional assumptions,

conditional exchangeability $Y(x) \perp\!\!\!\perp X \mid S$ is not testable from data.

☞ Typically, a DAG or ADMG is not uniquely identified from data. At best, one can hope to learn a set of plausible graphs instead of one single graph.

Myth 2. Causal discovery can solve this.

Without making strong additional assumptions,

conditional exchangeability $Y(x) \perp\!\!\!\perp X \mid S$ is not testable from data.

☞ Typically, a DAG or ADMG is not uniquely identified from data. At best, one can hope to learn a set of plausible graphs instead of one single graph.

Myth 3. I can take the adjustment sets from these graphs and take the union, intersection, or run a variable selection, and that will work.

$\hookrightarrow$ One can try to combine the ATEs identified from each graph in the set (Maathuis, Kalisch, and Bühlmann, 2009; Nandy, Maathuis, and Richardson, 2017; R. Guo and Perkovic, 2021). But you won't get a single point estimate. See also Ghosh and Rothenhäusler (2025) for an alternative.

Myth 2. Causal discovery can solve this.

Without making strong additional assumptions,

conditional exchangeability $Y(x) \perp\!\!\!\perp X \mid S$ is not testable from data.

☞ Typically, a DAG or ADMG is not uniquely identified from data. At best, one can hope to learn a set of plausible graphs instead of one single graph.

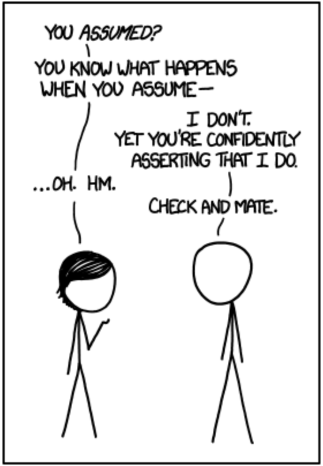
Myth 3. I can take the adjustment sets from these graphs and take the union, intersection, or run a variable selection, and that will work.

$\hookrightarrow$ One can try to combine the ATEs identified from each graph in the set (Maathuis, Kalisch, and Bühlmann, 2009; Nandy, Maathuis, and Richardson, 2017; R. Guo and Perkovic, 2021). But you won't get a single point estimate. See also Ghosh and Rothenhäusler (2025) for an alternative.

☞ Generally speaking, confounder selection requires domain knowledge about the underlying causal structure. It cannot be totally data-driven.

Myth 4. Ok, I need to draw the graph

"Suppose $\mathcal{G}$ is the causal DAG/ADMG ..."



https://xkcd.com/1339/

Myth 4. Ok, I need to draw the graph

1. **Impractical**
   - Do not know the full causal structure/mechanism
   - Even if we know it, can we readily draw it?
     - Unless the system is closed, the DAG is huge — try to think about the causes of causes, etc.
     - Reasoning locally about a pathway, mediator, etc. $\not\Rightarrow$ Equipped with global knowledge to draw the graph.
   - Help a domain expert to draw it
     - Tools and protocols (Shrier and Platt, 2008; Haber et al., 2022) are developed, but still very challenging or even impractical.

Myth 4. Ok, I need to draw the graph

1. **Impractical**
   - Do not know the full causal structure/mechanism
   - Even if we know it, can we readily draw it?
     - Unless the system is closed, the DAG is huge — try to think about the causes of causes, etc.
     - Reasoning locally about a pathway, mediator, etc. $\not\Rightarrow$ Equipped with global knowledge to draw the graph.
   - Help a domain expert to draw it
     - Tools and protocols (Shrier and Platt, 2008; Haber et al., 2022) are developed, but still very challenging or even impractical.
2. **Unnecessary**: To test the back-door criterion, it only involves partial knowledge about the graph.

   $\hookrightarrow$ Methods on representing and using the partial knowledge is still under-developed. This is the main motivation for our work.

Earlier work in this direction

Disjunctive criterion (VanderWeele and Shpitser, 2011): $\hookrightarrow$ See VanderWeele (2019) for variations.

$$S := (\mathrm{an}_{\mathcal{G}}(X) \cup \mathrm{an}_{\mathcal{G}}(Y)) \cap \{\text{observed pre-treatment covariates } \mathrm{V} \setminus \{X, Y\}\}.$$

☞ S is a sufficient adjustment set whenever $\mathrm{V} \setminus \{X, Y\}$ contains any sufficient adjustment set.

Earlier work in this direction

Disjunctive criterion (VanderWeele and Shpitser, 2011): $\hookrightarrow$ See VanderWeele (2019) for variations.

$$S := (\mathrm{an}_{\mathcal{G}}(X) \cup \mathrm{an}_{\mathcal{G}}(Y)) \cap \{\text{observed pre-treatment covariates } \mathrm{V} \setminus \{X, Y\}\}.$$

☞ S is a sufficient adjustment set whenever $\mathrm{V} \setminus \{X, Y\}$ contains any sufficient adjustment set.

This is useful when

1. it is already clear what all the observed covariates are, i.e., after the data are collected; $\hookrightarrow$ Not so useful if one wants to design a data collection plan.

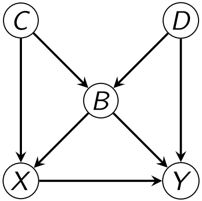
   Myth 5. I can do only confounder selection after having conceived all the observed/measurable pre-treatment covariates.

2. and the structural knowledge is rather limited.

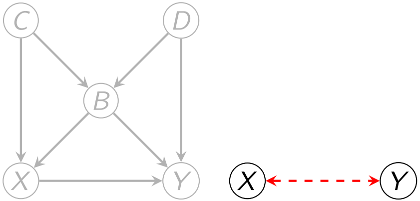
Also, this criterion does not certify the correctness of S.

Our approach in a nutshell

Our approach in a nutshell

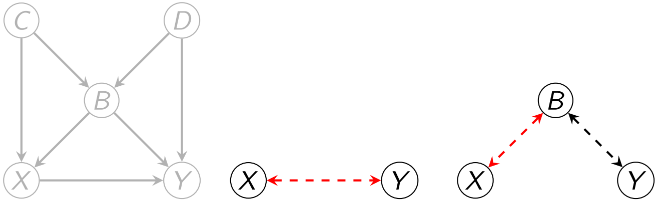


Our approach in a nutshell



1. Bottom-up approach. Start with X and Y: ⇠⇢ indicates potential confounding between them. ↪ Goal: Disconnect X and Y.

Our approach in a nutshell

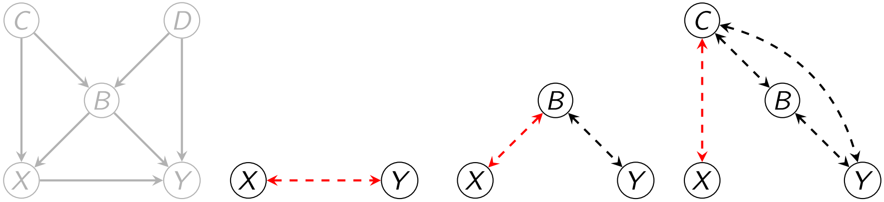


1. Bottom-up approach. Start with X and Y: ←- -→ indicates potential confounding between them.
   ↪ Goal: Disconnect X and Y.

2. Introduce B to control for confounding between X and Y. Meanwhile, potentially there is confounding between X and B, and between B and Y.

   ↪ $\{B\}$ is a primary adjustment set for X and Y.

Our approach in a nutshell



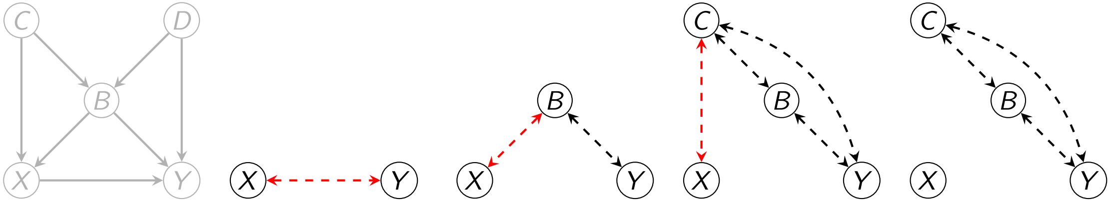
1. Bottom-up approach. Start with X and Y: $\leftarrow\!-\!\rightarrow$ indicates potential confounding between them. $\hookrightarrow$ Goal: Disconnect X and Y.

2. Introduce B to control for confounding between X and Y. Meanwhile, potentially there is confounding between X and B, and between B and Y.

   $\hookrightarrow$ $\{B\}$ is a primary adjustment set for X and Y.

3. To remove $X \leftarrow\!-\!\rightarrow B$, we introduce C.

Our approach in a nutshell



1. Bottom-up approach. Start with X and Y: $\leftarrow\!\!-\!\!\rightarrow$ indicates potential confounding between them. ↪ Goal: Disconnect X and Y.

2. Introduce B to control for confounding between X and Y. Meanwhile, potentially there is confounding between X and B, and between B and Y.

   ↪ $\{B\}$ is a primary adjustment set for X and Y.

3. To remove $X \leftarrow\!\!-\!\!\rightarrow B$, we introduce C.

4. X, Y are disconnected: $\{B, C\}$ ✓ is a sufficient adjustment set.

Myth 6. There are those things called d/m-separation and M/collider-bias. I need to be aware of them during the confounder selection process.

Myth 6. There are those things called d/m-separation and M/collider-bias. I need to be aware of them during the confounder selection process.

☞ Between two variables C and D, given all the variables on the canvas so far

$$\begin{cases} C \longleftrightarrow D & \text{(direct) confounding between } C \text{ and } D \\ C \dashleftarrow\dashrightarrow D & \text{potentially (direct) confounding between } C \text{ and } D \\ C \qquad D & \text{no (direct) confounding between } C \text{ and } D \end{cases}.$$

Here, $C \longleftrightarrow D$ if and only if there is a latent common cause U and

a pathway $C \leftarrow \cdots \leftarrow U \rightarrow \cdots \rightarrow D$ not through any variable already on canvas.

$\hookrightarrow$ There is no collider '$\rightarrow \circ \leftarrow$' on the pathway and you need not worry about it.

Myth 6. There are those things called d/m-separation and M/collider-bias. I need to be aware of them during the confounder selection process.

☞ Between two variables C and D, given all the variables on the canvas so far

$$\begin{cases} C \longleftrightarrow D & \text{(direct) confounding between } C \text{ and } D \\ C \dashleftarrow\dashrightarrow D & \text{potentially (direct) confounding between } C \text{ and } D \\ C \quad\quad D & \text{no (direct) confounding between } C \text{ and } D \end{cases}.$$

In other words, '$C \dashleftarrow\dashrightarrow D$' becomes '$C \quad\quad D$' only when

1. either U itself is added to canvas, or
2. all the $U \to \cdots \to C$ paths are blocked, or ↪ controlling mediators
3. all the $U \to \cdots \to D$ paths are blocked,

and this holds for every common cause U of C and D.

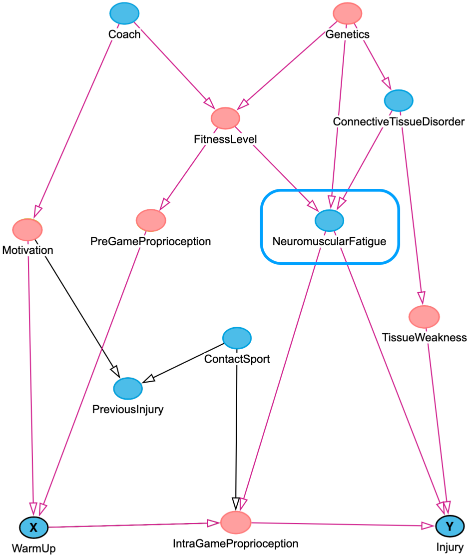
↪ You only need to reason about common causes and mediators. You do not need to know d-separation or M-bias.

Demo

https://ricguo.shinyapps.io/InteractiveConfSel/

Example

Myth 6. There are those things called d-separation and M-bias. I need to be aware of them during the confounder selection process.



Theoretical guarantees

Features of the procedure

- Through this process,

$$\text{a potential confounding arc } \circ \dashleftarrow\dashrightarrow \circ \text{ becomes } \begin{cases} \circ \dashleftarrow\dashrightarrow \circ & \text{still potential confounding} \\ \circ \longleftrightarrow \circ & \text{uncontrollable confounding} \\ \circ \qquad \circ & \text{no confounding} \end{cases}$$

Features of the procedure

- Through this process,

$$\text{a potential confounding arc } \circ \leftarrow\!\!-\!\!\rightarrow \circ \text{ becomes } \begin{cases} \circ \leftarrow\!\!-\!\!\rightarrow \circ & \text{still potential confounding} \\ \circ \longleftrightarrow \circ & \text{uncontrollable confounding} \\ \circ \qquad \circ & \text{no confounding} \end{cases}$$

- The user is interactively guided to select a $\circ \leftarrow\!\!-\!\!\rightarrow \circ$ and try to find its primary adjustment set, which is introduced onto canvas.
  - Reasoning about common causes and mediators, locally.
  - Questions are asked economically. No question is asked about the causal paths among covariates already on canvas.
    $\hookrightarrow$ Not relevant to the goal at all!

Features of the procedure

- Through this process,

$$\text{a potential confounding arc } \circ \dashleftarrow\dashrightarrow \circ \text{ becomes } \begin{cases} \circ \dashleftarrow\dashrightarrow \circ & \text{still potential confounding} \\ \circ \longleftrightarrow \circ & \text{uncontrollable confounding} \\ \circ \qquad \circ & \text{no confounding} \end{cases}$$

- The user is interactively guided to select a $\circ \dashleftarrow\dashrightarrow \circ$ and try to find its primary adjustment set, which is introduced onto canvas.
  - Reasoning about common causes and mediators, locally.
  - Questions are asked economically. No question is asked about the causal paths among covariates already on canvas.
    $\hookrightarrow$ Not relevant to the goal at all!
- The procedure terminates when
  1. X and Y are disconnected: a sufficient adjustment set (those on canvas minus $\{X, Y\}$) is found.
  2. X and Y are connected by $X \longleftrightarrow \cdots \longleftrightarrow Y$: no sufficient adjustment set exists.

Theoretical guarantees

We make the assumptions

1. The causal model is represented by an unknown DAG or ADMG;
   $\hookrightarrow$ a single-world SWIG model or Pearl's NPSEM-IE model.
2. All the covariates considered are either ancestors (i.e., direct or indirect causes) of X or ancestors of Y.

Then, the procedure must terminate in finitely many steps. Further, it is

1. **Sound:** If X and Y are disconnected, the found S is a sufficient adjustment set;
2. **Complete:** If X and Y end up being connected by $X \longleftrightarrow \cdots \longleftrightarrow Y$, then no sufficient adjustment set exists.

Some theories behind: Refined m-connections

We use notation

$$A \,\langle\text{shape}\rangle\, B \mid C \iff \exists \text{ a path of } \langle\text{shape}\rangle \text{ between } A \text{ and } B \text{ that is m-connected given } C,$$

$$A \,\langle\text{sh}\not\text{ape}\rangle\, B \mid C \iff \nexists \text{ a path of } \langle\text{shape}\rangle \text{ between } A \text{ and } B \text{ that is m-connected given } C.$$

☞ **m-connection and m-separation** (▶ '$\leftsquigarrow\!\rightsquigarrow * \leftsquigarrow\!\rightsquigarrow$' is a path of any shape)

$$A \leftsquigarrow\!\rightsquigarrow * \leftsquigarrow\!\rightsquigarrow B \mid C \iff A \not\perp\!\!\!\perp_m B \mid C \quad \text{and} \quad A \leftsquigarrow\!\rightsquigarrow \not* \leftsquigarrow\!\rightsquigarrow B \mid C \iff A \perp\!\!\!\perp_m B \mid C.$$

Some theories behind: Refined m-connections

We use notation

$$A \langle \text{shape} \rangle B \mid C \iff \exists \text{ a path of } \langle \text{shape} \rangle \text{ between } A \text{ and } B \text{ that is m-connected given } C,$$
$$A \langle \not\text{shape} \rangle B \mid C \iff \nexists \text{ a path of } \langle \text{shape} \rangle \text{ between } A \text{ and } B \text{ that is m-connected given } C.$$

☞ **m-connection and m-separation** (▶ '$\rightsquigarrow * \leftsquigarrow$' is a path of any shape)

$$A \rightsquigarrow * \leftsquigarrow B \mid C \iff A \not\perp\!\!\!\perp_m B \mid C \quad \text{and} \quad A \rightsquigarrow \not* \leftsquigarrow B \mid C \iff A \perp\!\!\!\perp_m B \mid C.$$

Refined m-connection

$$A \left\{ \begin{array}{c} \rightsquigarrow \\ \leftrightsquigarrow \\ \leftrightsquigarrow * \leftrightsquigarrow \end{array} \right\} B \mid C$$

'$\leftrightsquigarrow$': confounding arc (i.e., '$\leftarrow - \rightarrow$ / $\longleftrightarrow$' in our procedure), with no collider on the path.
'$\leftrightsquigarrow * \leftrightsquigarrow$': confounding path, where wildcard $*$ means 0,1 or more colliders on the path.

Some theories behind: Reformulated back-door criterion

Back-door criterion, reformulated Suppose $X \rightsquigarrow Y$ in $\mathcal{G}$. For any $S \subset \mathrm{V} \setminus \mathrm{de}(X)$,

$$S \text{ satisfies the back-door criterion} \iff X \leftrightsquigarrow\!\not\ast\!\leftrightsquigarrow Y \mid S.$$

Observation: confounding path is not monotone due to M-bias:

$$A \leftrightsquigarrow\!\not\ast\!\leftrightsquigarrow B \mid C \quad \not\Longrightarrow \quad A \leftrightsquigarrow\!\not\ast\!\leftrightsquigarrow B \mid C', \quad C \subset C',$$

but confounding arc is monotone

$$A \not\leftrightsquigarrow B \mid C \quad \Longrightarrow \quad A \not\leftrightsquigarrow B \mid C', \quad C \subset C'.$$

$\hookrightarrow$ Strategy: block all confounding paths by blocking confounding arcs, one at at a time.

Latent projection preserves refined m-connections

The theoretical guarantees of the procedure follow from a general theorem.

Theorem For any ADMG $\mathcal{G}$ over vertex set V and $\{A, B\} \cup C \subseteq \tilde{\mathrm{V}} \supseteq \mathrm{V}$,

$$A \left\{ \begin{array}{c} \rightsquigarrow \\ \leftrightsquigarrow \\ \leftrightsquigarrow * \leftrightsquigarrow \end{array} \right\} B \mid C \; [\mathcal{G}] \quad \Longleftrightarrow \quad A \left\{ \begin{array}{c} \rightsquigarrow \\ \leftrightsquigarrow \\ \leftrightsquigarrow * \leftrightsquigarrow \end{array} \right\} B \mid C \; [\mathcal{G}(\tilde{V})].$$

Corollary For $A, B \notin C$,

$$A \left\{ \begin{array}{c} \rightsquigarrow \\ \leftrightsquigarrow \\ \leftrightsquigarrow * \leftrightsquigarrow \end{array} \right\} B \mid C \; [\mathcal{G}] \quad \Longleftrightarrow \quad A \left\{ \begin{array}{c} \rightarrow \\ \leftrightarrow \\ \longleftrightarrow * \longleftrightarrow \end{array} \right\} B \; [\mathcal{G}(\{A, B\} \cup C)].$$

District criterion

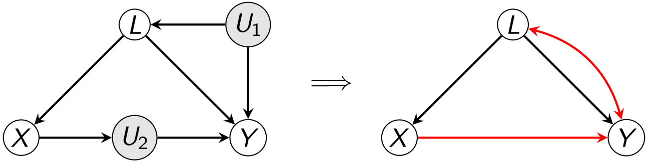
$$\underbrace{S \text{ is a sufficient adjustment set}}_{\text{block all back-door paths in } \mathcal{G}} \quad \Longleftrightarrow \quad \underbrace{X \leftrightarrow \not* \leftrightarrow Y \mid S \, [\mathcal{G}(\{X, Y\} \cup S)]}_{\text{connectivity by confounding arcs}}.$$

Latent projection

☞ Causal model represented by ADMG is closed under **latent projection**.

For $\mathrm{V} = \tilde{\mathrm{V}} \cup \mathrm{U}$ and $\tilde{\mathrm{V}} \cap \mathrm{U} = \emptyset$, let $\mathcal{G}(\tilde{\mathrm{V}})$ be the latent projection of $\mathcal{G}$ onto the margin $\tilde{\mathrm{V}}$.

$$A \left\{ \begin{matrix} \overset{\text{via } \mathrm{U}}{\rightsquigarrow} \\ \overset{\text{via } \mathrm{U}}{\leftsquigarrow} \\ \overset{\text{via } \mathrm{U}}{\leftrightsquigarrow} \end{matrix} \right\} B\ [\mathcal{G}] \quad \Longleftrightarrow \quad A \left\{ \begin{matrix} \rightarrow \\ \leftarrow \\ \leftrightarrow \end{matrix} \right\} B\ [\mathcal{G}(\tilde{\mathrm{V}})].$$



Summary

I presented an interactive procedure for confounder selection.
$\hookrightarrow$ Feel free to try the Shiny app. Please share any feedback through email/Github.

Hopefully, I debunked some common **myths** about the problem of confounder selection:

1. This is a solved problem.
2. It can be solved in a data-driven way via causal discovery.
3. Simple combinations of potential adjustment sets or running variable selection solves it.
4. I need to draw the whole graph.
5. I must have conceived all the observed variables before considering the problem.
6. I must understand d/m-separation and M-bias.

Thanks! Questions?

References I

Guo, F. Richard, Anton Rask Lundborg, and Qingyuan Zhao (2022). "Confounder selection: objectives and approaches". In: *arXiv preprint arXiv:2208.13871*.

Pearl, Judea (1993). "Comment: graphical models, causality and intervention". In: *Statistical Science* 8.3, pp. 266–269.

Shpitser, Ilya, Tyler J. VanderWeele, and James M. Robins (2010). "On the validity of covariate adjustment for estimating causal effects". In: *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*. UAI'10. Arlington, Virginia, USA: AUAI Press, pp. 527–536.

Maathuis, Marloes H, Markus Kalisch, and Peter Bühlmann (2009). "Estimating high-dimensional intervention effects from observational data". In.

Nandy, Preetam, Marloes H Maathuis, and Thomas S Richardson (2017). "Estimating the effect of joint interventions from observational data in sparse high-dimensional settings". In: *The Annals of Statistics*, pp. 647–674.

Guo, Richard and Emilija Perkovic (2021). "Minimal enumeration of all possible total effects in a Markov equivalence class". In: *International Conference on Artificial Intelligence and Statistics*. PMLR, pp. 2395–2403.

References II

- Ghosh, Aditya and Dominik Rothenhäusler (2025). "Assumption-robust causal inference". In: *arXiv preprint arXiv:2505.08729*.
- Shrier, Ian and Robert W Platt (2008). "Reducing bias through directed acyclic graphs". In: *BMC Medical Research Methodology* 8, pp. 1–15.
- Haber, Noah A et al. (2022). "DAG With Omitted Objects Displayed (DAGWOOD): a framework for revealing causal assumptions in DAGs". In: *Annals of Epidemiology* 68, pp. 64–71.
- VanderWeele, Tyler J. and Ilya Shpitser (2011). "A new criterion for confounder selection". In: *Biometrics* 67.4, pp. 1406–1413.
- VanderWeele, Tyler J. (2019). "Principles of confounder selection". In: *European Journal of Epidemiology* 34.3, pp. 211–219.
- VanderWeele, Tyler J. and Ilya Shpitser (2013). "On the definition of a confounder". In: *The Annals of Statistics* 41.1, pp. 196–220.
- Shrier, Ian (2008). "Letter to the Editor". In: *Statistics in Medicine* 27.14, pp. 2740–2741.

References III

- Rubin, Donald B. (2009). "Should observational studies be designed to allow lack of balance in covariate distributions across treatment groups?" In: *Statistics in Medicine* 28.9, pp. 1420–1423.
- Pearl, Judea (2009). "Remarks on the method of propensity score". In: *Statistics in Medicine* 28.9, pp. 1415–1416.
- Sjölander, Arvid (2009). "Propensity scores and M-structures". In: *Statistics in Medicine* 28.9, pp. 1416–1420.
- Sadeghi, Kayvan and Steffen Lauritzen (2014). "Markov properties for mixed graphs". In: *Bernoulli* 20.2, pp. 676–696.