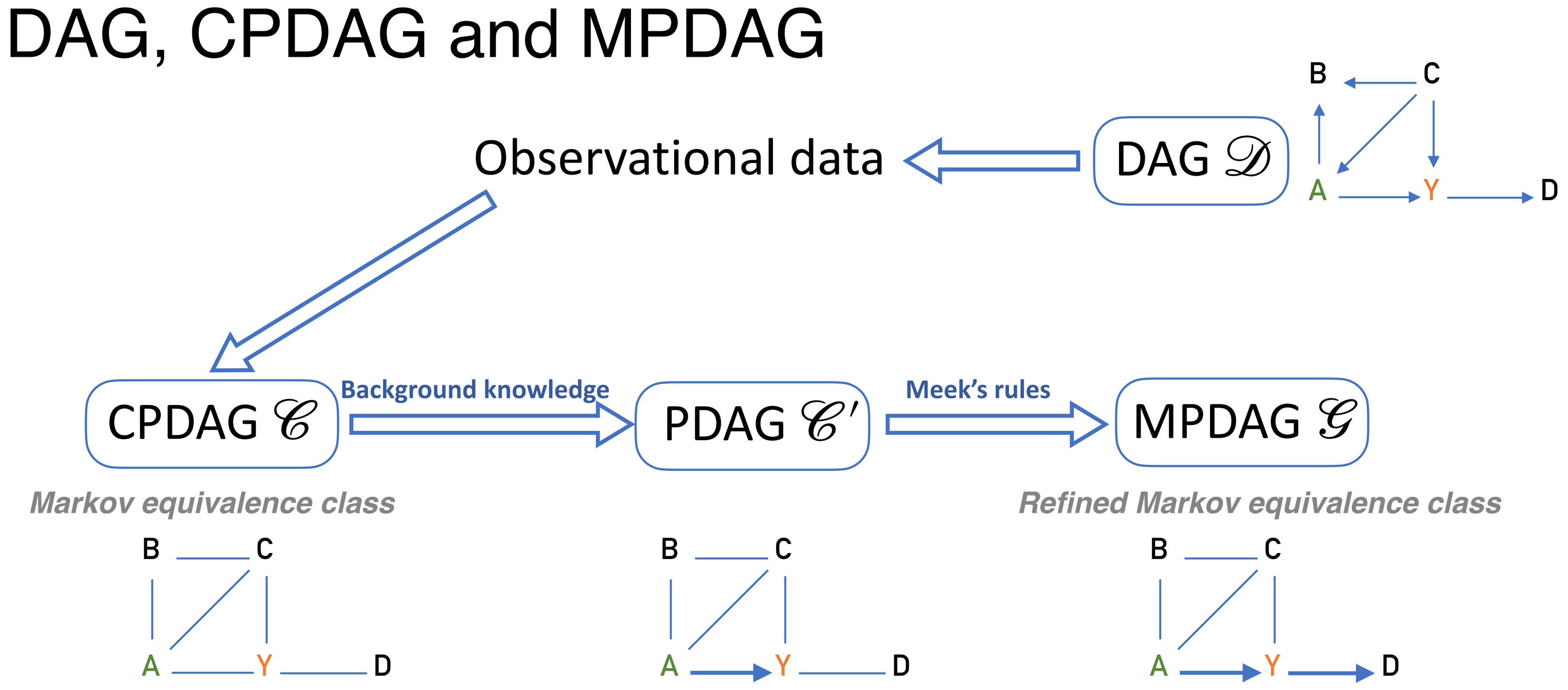
Minimal Enumeration of All Possible Total Effects in a Markov Equivalence Class

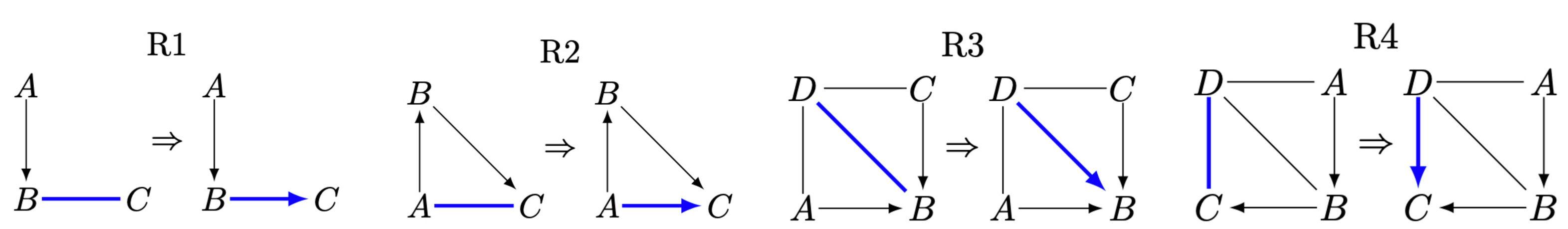
F. Richard Guo and Emilija Perković
Department of Statistics
University of Washington, Seattle

- Causal inference from observational data
- Underlying causal DAG $\mathscr{D}$ is known up to a refined Markov equivalence class, represented by MPDAG $\mathscr{G}$
- Effect of A on Y is unidentified from $\mathscr{G}$
- Goal: Report the set of possible effects of A on Y

DAG, CPDAG and MPDAG



Meek's rules (1995)

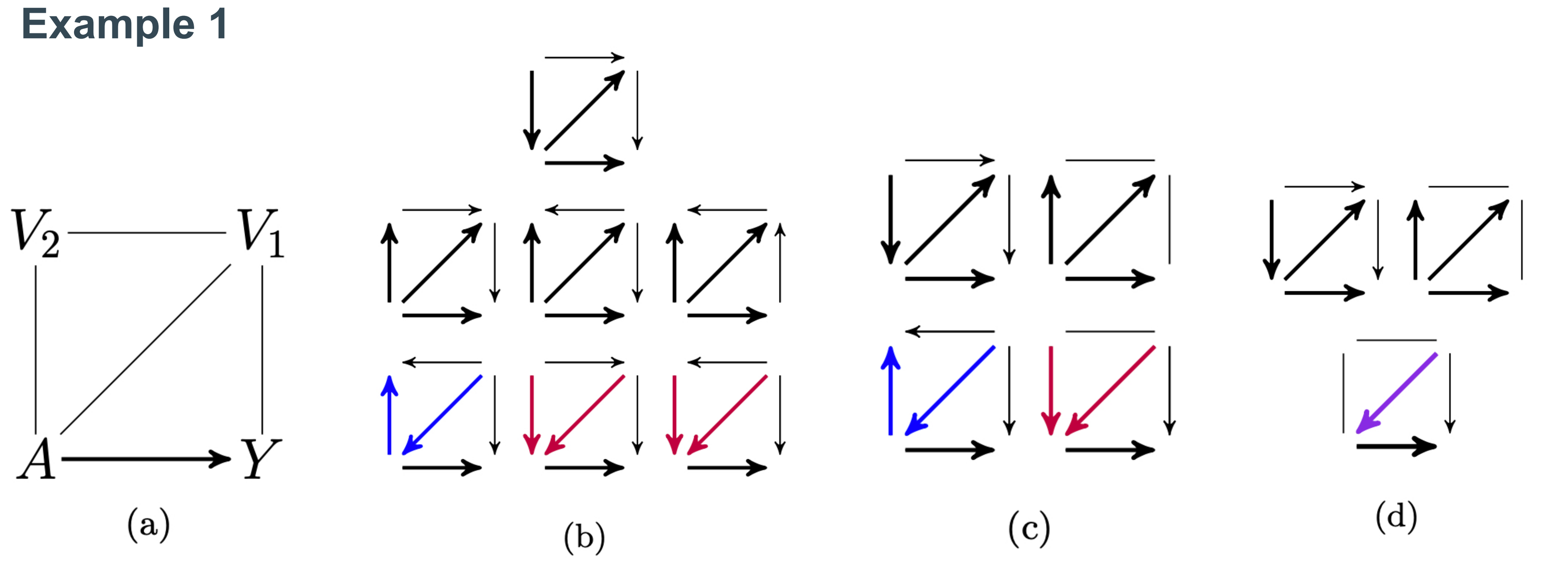


Identification condition

A total effect is identified given an equivalence class of DAGs if it can be expressed as a functional of the observed distribution, which is the same for all DAGs in the equivalence class

Theorem 1 (Perković 2010) The total effect of A on Y is identified in MPDAG $\mathscr{G}$ if and only if every proper possibly causal path from A to Y starts with a directed edge in $\mathscr{G}$.

Example 1



(a) MPDAG $\mathscr{G}$, (b) all DAGs represented by $\mathscr{G}$, (c) all MPDAGs with distinct parent sets of A represented by $\mathscr{G}$, (d) all MPDAGs with distinct functionals $P(Y \mid \mathrm{do}(A))$ represented by $\mathscr{G}$.

Main results

Theorem 2 Let $\mathscr{G}$ be a causal MPDAG. Let A and Y be disjoint node sets in $\mathscr{G}$ such that the total effect of A on Y is not identified given $\mathscr{G}$. Suppose $p = \langle A_1, V_1, \ldots, Y_1 \rangle$ for $A_1 \in A$, $Y_1 \in Y$ is a **shortest** proper possibly causal path from A to Y such that $A_1 - V_1$. Then the total effect of A on Y is not identified in any MPDAG $\mathscr{G}^*$ that is represented by $\mathscr{G}$ and contains the undirected edge $A_1 - V_1$.

IDGraphs Algorithm

Input: *MPDAG $\mathscr{G}$, disjoint node sets A and Y*

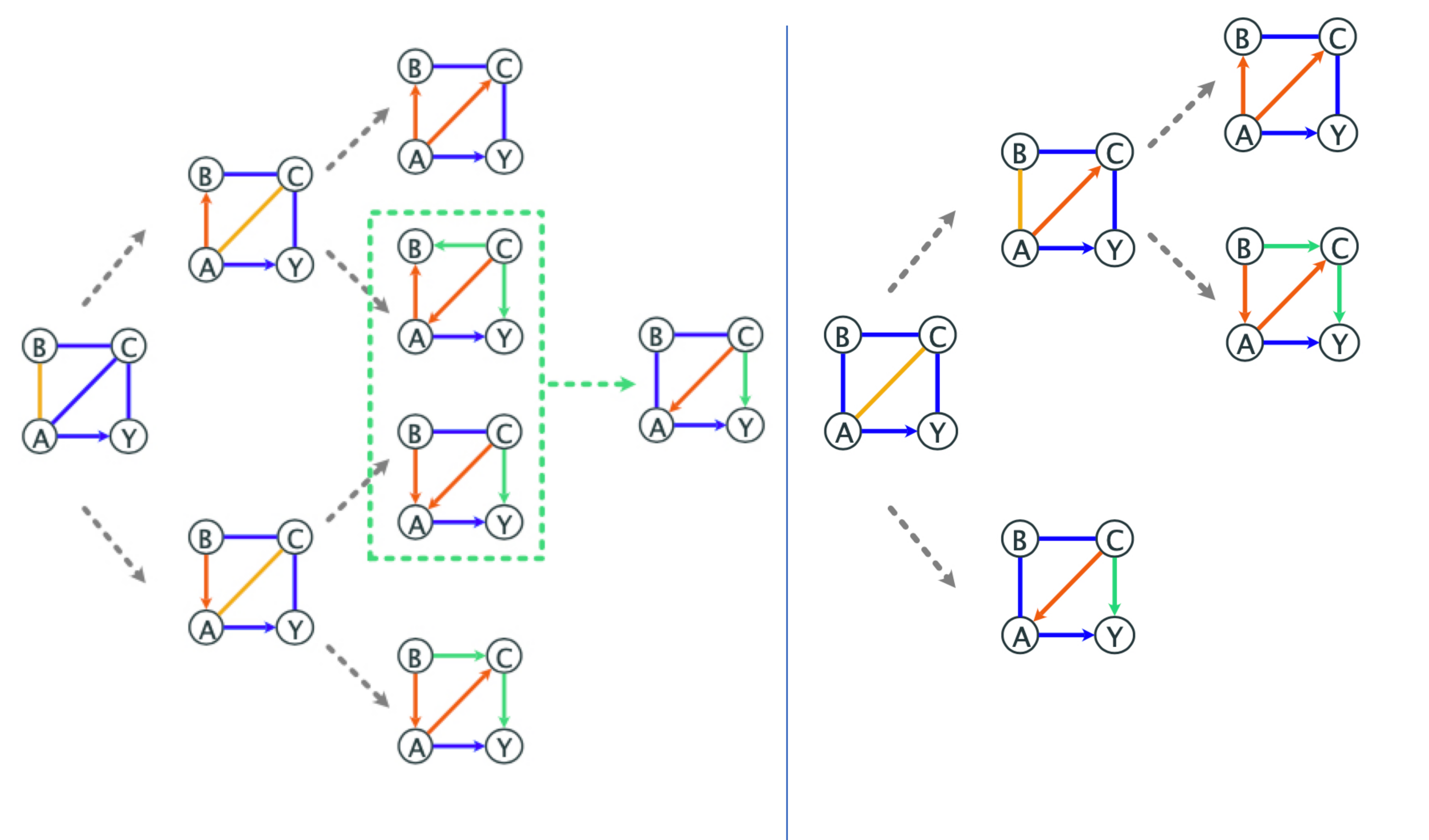
Output: *the minimal set of MPDAGs with identified effects that partition $\mathscr{G}$*

1. Pick $A_1 - V_1$ such that $A_1 \in A$ and $A_1, V_1, \ldots, Y$ is a **shortest** proper possibly causal path from A to Y.
2. Let $\mathscr{G}_1 \leftarrow$ MPDAG$(\mathscr{G}, A_1 \rightarrow V_1)$ and $\mathscr{G}_2 \leftarrow$ MPDAG$(\mathscr{G}, A_1 \leftarrow V_1)$
3. Recurse on $\mathscr{G}_1$ and $\mathscr{G}_2$ until identified.

Theorem 3 Suppose $\mathscr{G}$ is a causal MPDAG and A, Y are two disjoint node sets in $\mathscr{G}$. Let $L = \{\mathscr{G}_1, \ldots, \mathscr{G}_n\}$ be the output of **IDGraphs**$(A, Y, \mathscr{G})$. Then the following statements hold.

1. The total effect of A on Y is identified in each $\mathscr{G}_i$.
2. For any $i \neq j$, there exists an observational density f that is consistent with $\mathscr{G}$ such that the effect identified from f in $\mathscr{G}_i$ is different from the effect identified from f in $\mathscr{G}_j$.
3. L is a partition of $\mathscr{G}$ in terms of DAGs represented.

Orienting the shortest path first



Non-minimal: A - B is oriented first. Minimal: A - C is oriented first.

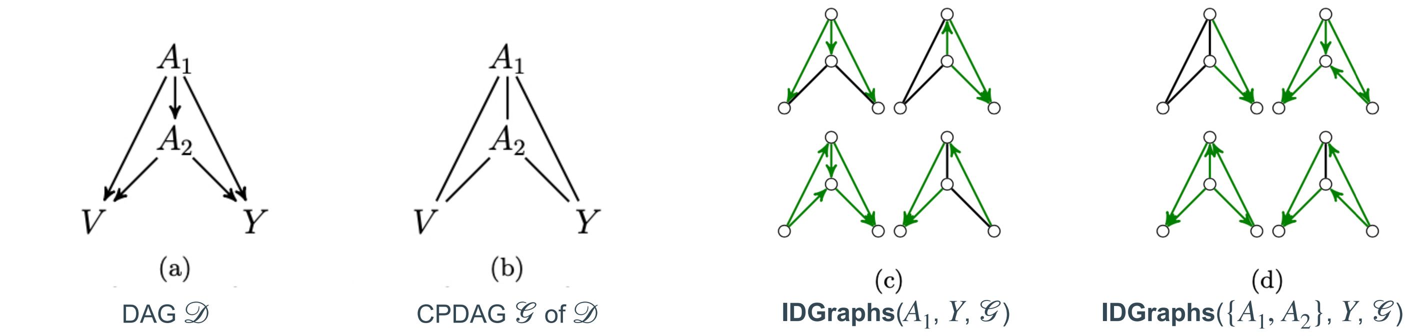
Comparison

	Comp. Cost	$\lvert A\rvert = 1$	$\lvert A\rvert > 1$	Minimal
Naive - Enumerate all DAGs:				
global IDA (Maathuis et al, 2009)	$\mathcal{O}(\lvert V\rvert !)$	✓	-	No
global joint IDA (Nandy et al, 2017)	$\mathcal{O}(\lvert V\rvert !)$	✓	✓	No
Enumerate valid parent sets of A:				
local IDA (Maathuis et al, 2009, Fang & He, 2020)	$\mathcal{O}(2^{l(\mathscr{G})})$	✓	-	No
semi-local IDA, joint IDA (Nandy et al, 2017)	$\mathcal{O}(2^{l(\mathscr{G})} poly(\lvert V\rvert))$	✓	✓	No
optimal IDA (Witte et al, 2020)	$\mathcal{O}(2^{l(\mathscr{G})} poly(\lvert V\rvert))$	✓	~	Yes
Enum. $A-$ on poss. causal paths to Y:				
collapsible IDA (Liu et. al, 2020)	$\mathcal{O}((\lvert V\rvert + \lvert E\rvert) 2^{r(\mathscr{G})})$	✓	-	No
Recursively enum. shortest path first				
`IDGraphs` (Guo & Perković)	$\mathcal{O}(2^{m(\mathscr{G})} poly(\lvert V\rvert))$	✓	✓	Yes

Note: $m(\mathscr{G}) \leq r(\mathscr{G}) \leq l(\mathscr{G})$.

Simulations

Example 2 Data is simulated from linear causal model with Gaussian error.

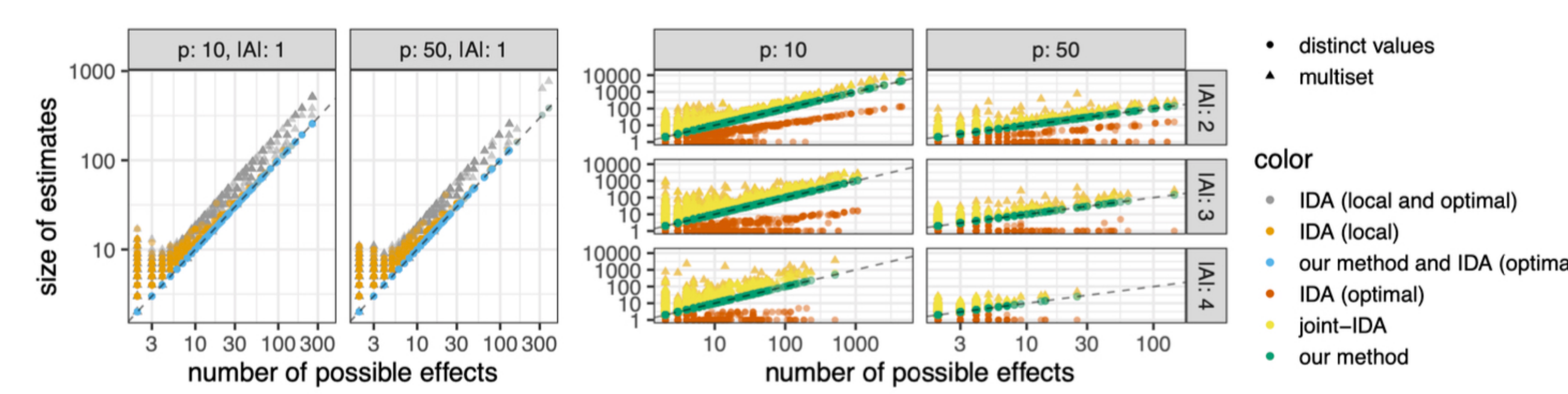


DAG $\mathscr{D}$ — CPDAG $\mathscr{G}$ of $\mathscr{D}$ — **IDGraphs**$(A_1, Y, \mathscr{G})$ — **IDGraphs**$(\{A_1, A_2\}, Y, \mathscr{G})$

	A_1 on Y (See (c))	A_1, A_2 on Y (See (d))
true effect	3	(2,1)
true possible effects	{3, 2, 1.8, 0}	$\{(2,1), (3,0), (0,2), (0,0)\}$
our method	{2.9, 2.1, 1.9, 0}	$\{(2.1,0.9), (2.9,0), (0,1.9), (0,0)\}$
IDA (optimal) (Witte et al., 2020)	$\{2.9, (2.1)^2, 1.9, 0\}$	$\{(2.1,0.9)^6, (0,0)^2, (\mathtt{NA},\mathtt{NA})^2\}$
IDA (local) (Maathuis et al., 2009)	{2.9, 2.1, 2.2, 1.9, 0}	—
joint-IDA (Nandy et al., 2017)	—	$\{(2.1,0.9)^2, (2.2,0.9), (1.9,1.1), (2.2,1.1)^2, (0,1.9), (2.9,0), (0,0)^2\}$

Random instances $\mathscr{D}$ is drawn from an Erdős–Rényi ensemble.

- We consider graphs $p = 10$ and $p = 50$, where the average degree k is drawn from $\{2,\ldots,8\}$ for the former and $\{2,\ldots,45\}$ for the latter.
- We take $\mathscr{G}$ to be the CPDAG of $\mathscr{D}$.
- Treatment variables A and outcome Y are randomly selected such that the total effect is unidentified in $\mathscr{G}$. The size of A varies from 1 to 4.
- For each instance, the possible effects of A on Y are estimated given $\mathscr{G}$ and 500 independent samples generated by a corresponding linear causal model.



For more details: https://bit.ly/20USlGL