Using background knowledge for the estimation of total causal effects

Interpreting and using CPDAGs with background knowledge

Emilija Perković, ETH Zurich

Joint work with Markus Kalisch and Marloes Maathuis

Causal effects

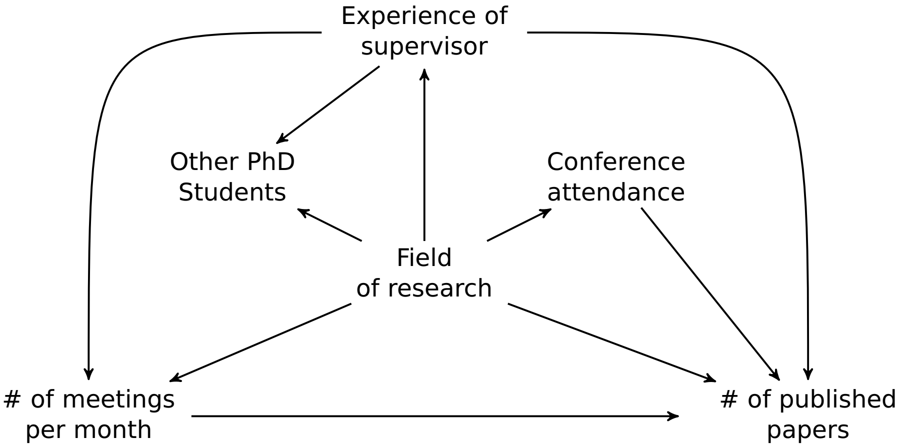


Figure: DAG $\mathcal{D}$.

Causal effects

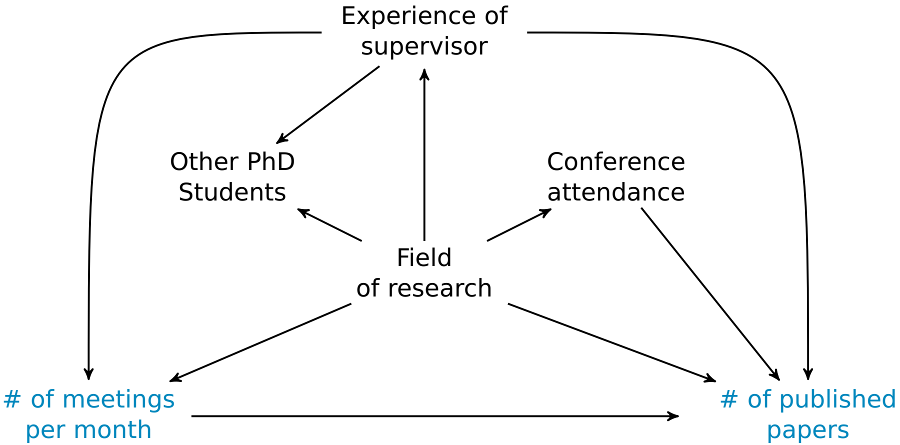


Figure: DAG $\mathcal{D}$.

Goal

- Estimate the total causal effect of **X** on **Y**

- *do*(**x**): an outside intervention that sets variables **X** to **x**.

Observational data

Randomized control studies

Goal

- Estimate the total causal effect of $\mathbf{X}$ on $\mathbf{Y}$
  - the average change in $\mathbf{Y}$ due to $do(\mathbf{x})$ -

- $do(\mathbf{x})$: an outside intervention that sets variables $\mathbf{X}$ to $\mathbf{x}$.

Observational data

Randomized control studies

Goal

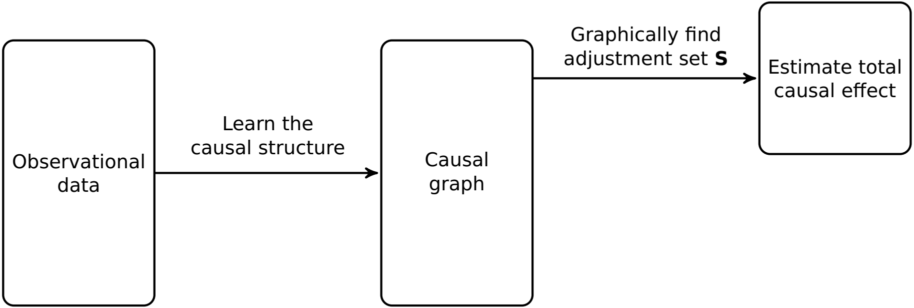
- Estimate the total causal effect of $\mathbf{X}$ on $\mathbf{Y}$
  - the average change in $\mathbf{Y}$ due to $do(\mathbf{x})$ -
  from observational data.

- $do(\mathbf{x})$: an outside intervention that sets variables $\mathbf{X}$ to $\mathbf{x}$.

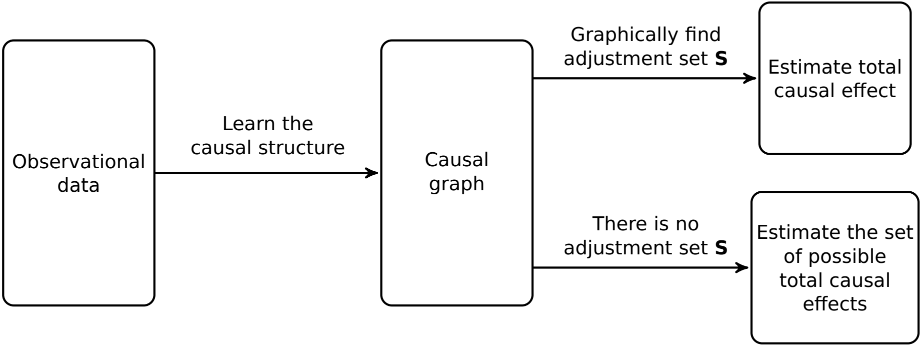
Observational data

Randomized control studies

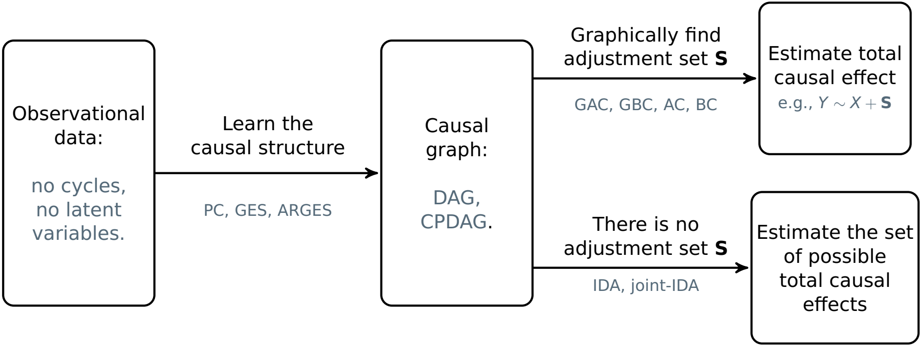
Framework


Observational data
Learn the causal structure
Causal graph
Graphically find adjustment set **S**
Estimate total causal effect

Framework

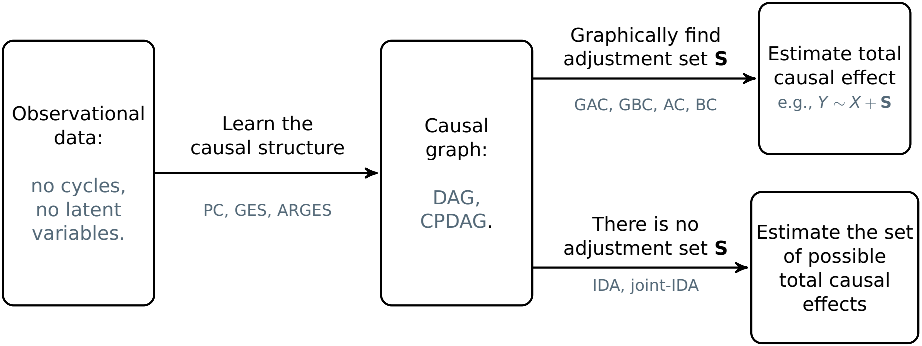

Observational
data
Learn the
causal structure
Causal
graph
Graphically find
adjustment set **S**
Estimate total
causal effect
There is no
adjustment set **S**
Estimate the set
of possible
total causal
effects

Framework



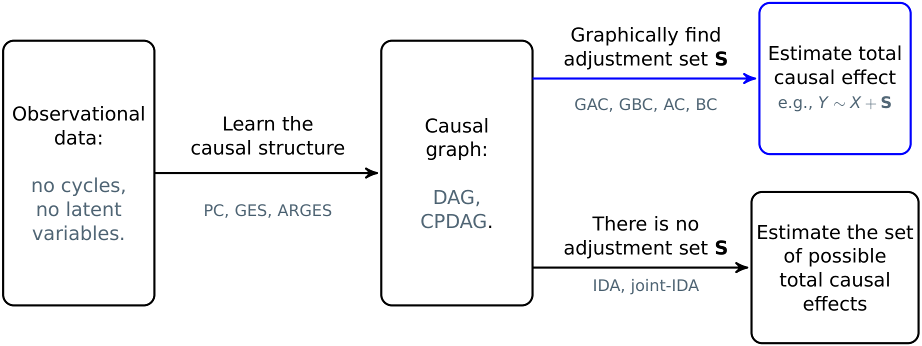
- PC (Spirtes et al, 1993), GES (Chickering, 2002), ARGES (Nandy et al, 2016).

Framework



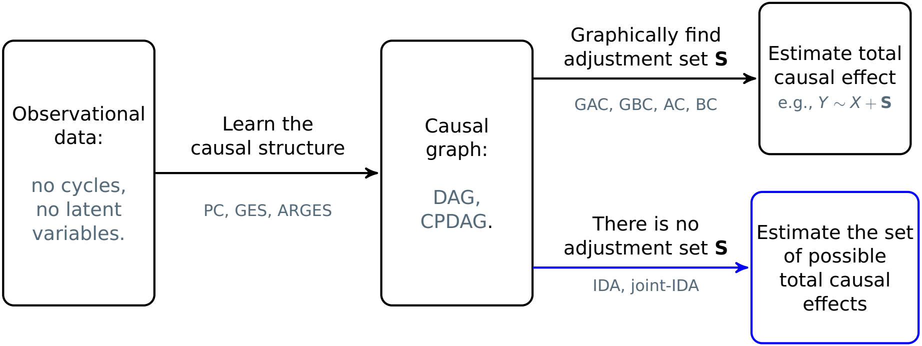
- BC (Pearl, 1993), AC (Shpitser et al. 2012; van der Zander et al. 2014), GAC (Perkovic et al, 2015,2017a), GBC (Maathuis and Colombo, 2016).
- IDA (Maathuis et al., 2009), joint-IDA (Nandy et al, 2017).

Framework

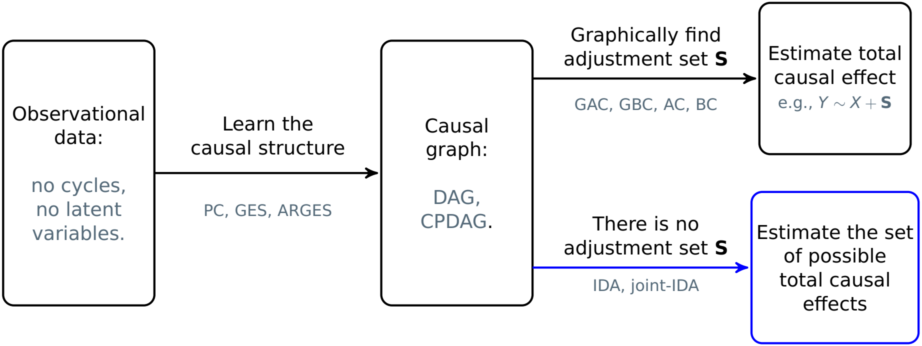


- Causal effects are often estimated by adjusted regression.
- Adjustment sets depend on the causal structure, which can be represented by a graph.

Framework

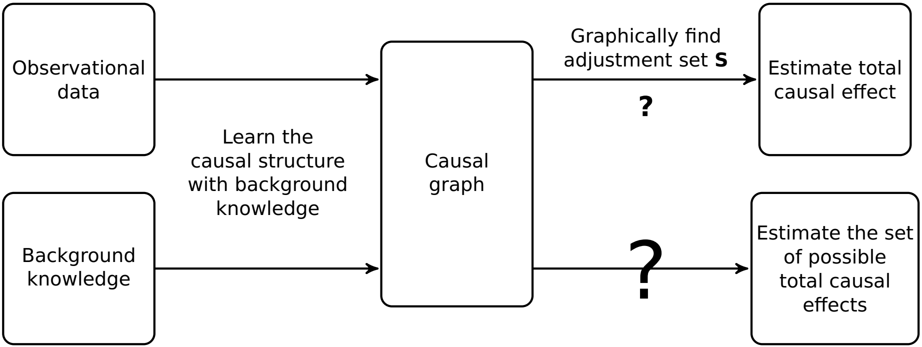

Observational data:
no cycles, no latent variables.
Learn the causal structure
PC, GES, ARGES
Causal graph:
DAG, CPDAG.
Graphically find adjustment set **S**
GAC, GBC, AC, BC
Estimate total causal effect
e.g., $Y \sim X + \mathbf{S}$
There is no adjustment set **S**
IDA, joint-IDA
Estimate the set of possible total causal effects

Framework



- If the total causal effect is different between DAGs in the equivalence class, then no adjustment set.
- Then **often** the set of possible total causal effects will contain zero.

Framework



- How much does background knowledge help to identify total causal effects?

Sources of background knowledge

- Applications - Expert knowledge of some causal relations, previous studies etc.
- Using a mix of observational and interventional data
- Model restrictions

Sources of background knowledge

- Applications - Expert knowledge of some causal relations, previous studies etc. (Meek, 1995; Scheines et al., 1998)
- Using a mix of observational and interventional data (Hauser and Bühlmann, 2012; Wang et al., 2017)
- Model restrictions (Hoyer et al., 2008; Ernest et al. 2016; Eigenmann et al., 2017)

Causal effects

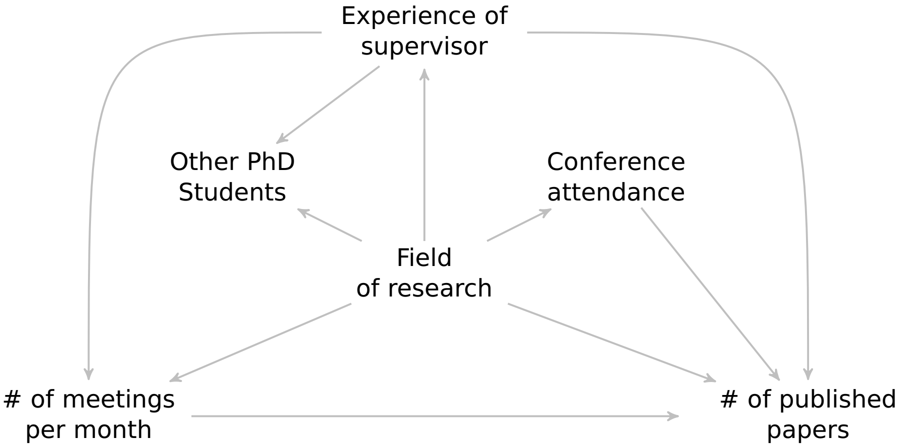


Figure: DAG $\mathcal{D}$.

Causal effects

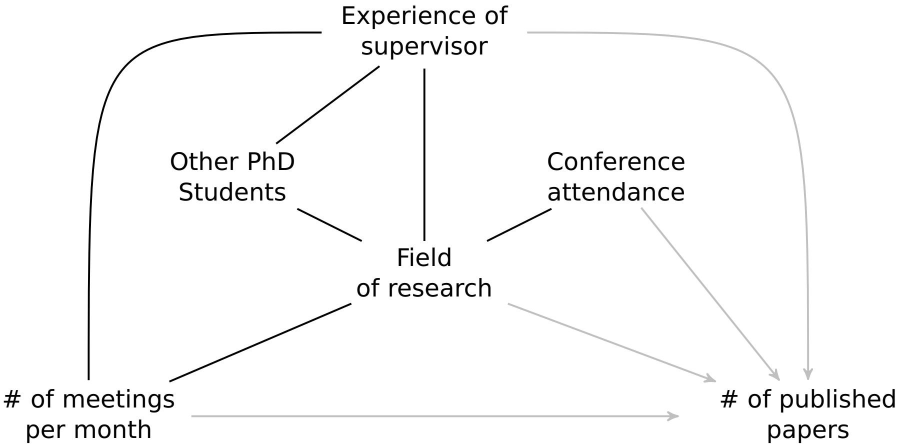


Figure: CPDAG $\mathcal{C}$ of DAG $\mathcal{D}$.

Causal effects

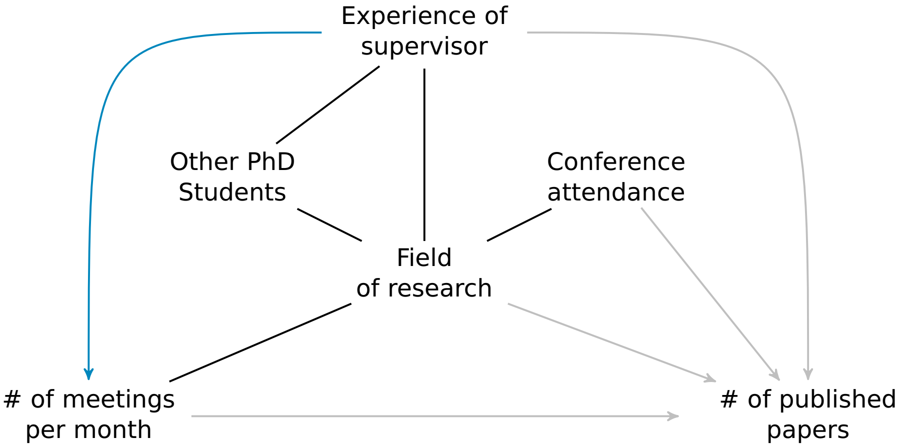


Figure: CPDAG $\mathcal{C}$ with background knowledge.

Causal effects

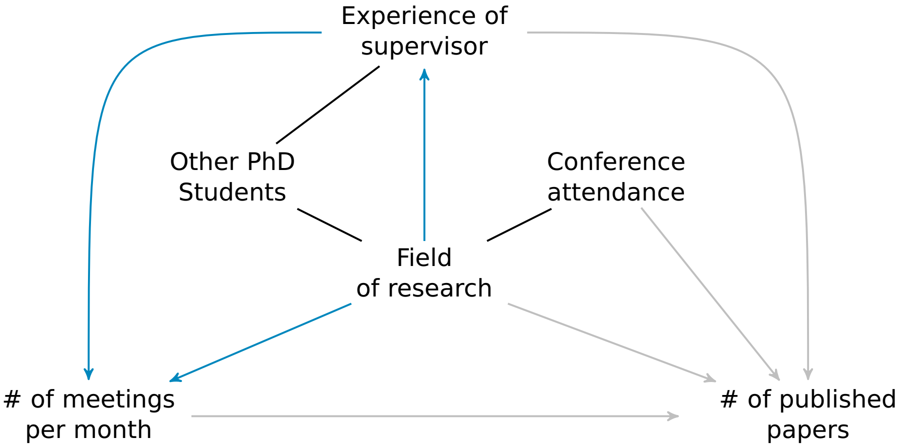
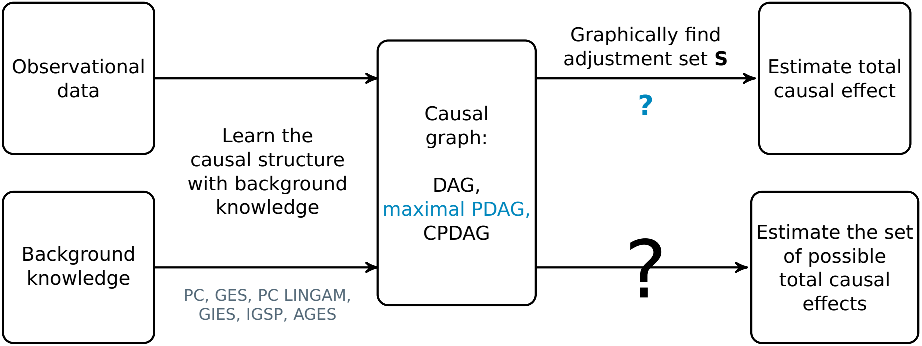


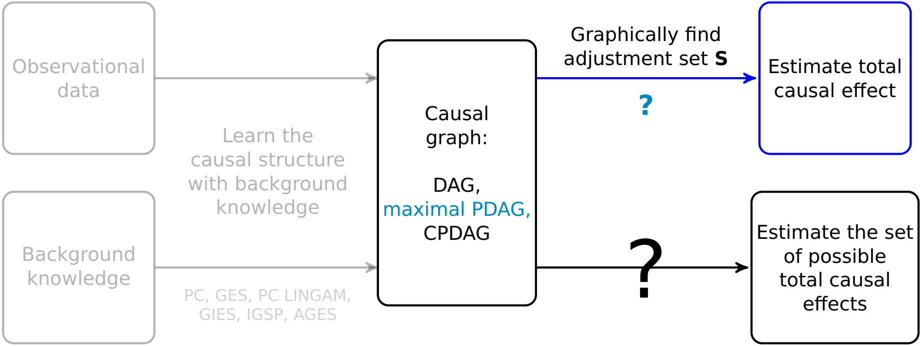
Figure: CPDAG $\mathcal{C}$ with background knowledge.

Framework



- PC (Spirtes et al, 1993) or GES (Chickering, 2002) + background knowledge (Meek, 1995; TETRAD, Scheines et al., 1998)
- PC LINGAM (Hoyer et al., 2008), GIES (Hauser and Bühlmann, 2012), AGES (Eigenmann et al., 2017), IGSP (Wang et al., 2017).

Framework



Observational data
Background knowledge
Learn the causal structure with background knowledge
PC, GES, PC LINGAM, GIES, IGSP, AGES
Causal graph:
DAG,
maximal PDAG,
CPDAG
Graphically find adjustment set **S**
?
Estimate total causal effect
?
Estimate the set of possible total causal effects

Adjustment criterion for maximal PDAGs

Theorem (Perković et al, 2017b):
S is an adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ and $\mathcal{G}$ if and only if:

Amenability	$\mathcal{G}$ is **b-amenable** relative to $(\mathbf{X}, \mathbf{Y})$.
Forbidden Set	**S** does not contain nodes in **b-Forbidden**$(\mathbf{X}, \mathbf{Y}, \mathcal{G})$.
Blocking	**S** blocks all **proper b-non-causal definite status** paths from **X** to **Y** in $\mathcal{G}$.

Adjustment criterion for maximal PDAGs

Theorem (Perković et al, 2017b):
S is an adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ and $\mathcal{G}$ if and only if:

- Amenability: $\mathcal{G}$ is **b-amenable** relative to $(\mathbf{X}, \mathbf{Y})$.
- Forbidden Set: **S** does not contain nodes in **b-Forbidden**$(\mathbf{X}, \mathbf{Y}, \mathcal{G})$.
- Blocking: **S** blocks all **proper b-non-causal definite status** paths from **X** to **Y** in $\mathcal{G}$.

Maximal PDAGs can contain partially directed cycles and do not have a chordal undirected component.

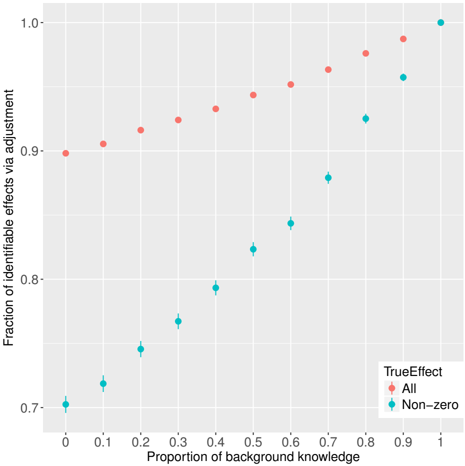
Adjustment criterion for maximal PDAGs

Theorem (Perković et al, 2017b):
$\mathbf{S}$ is an adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ and $\mathcal{G}$ if and only if:

Amenability	$\mathcal{G}$ is **b-amenable** relative to $(\mathbf{X}, \mathbf{Y})$.
Forbidden Set	$\mathbf{S}$ does not contain nodes in **b-Forbidden**$(\mathbf{X}, \mathbf{Y}, \mathcal{G})$.
Blocking	$\mathbf{S}$ blocks all **proper b-non-causal definite status** paths from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$.

In a linear setting the total causal effect of X on Y is then the linear regression coefficient of X in the regression $Y \sim X + \mathbf{S}$.

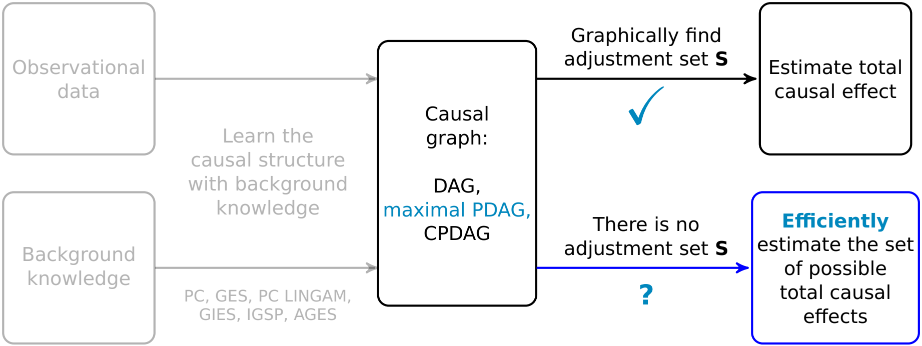
Does an adjustment set always exist?



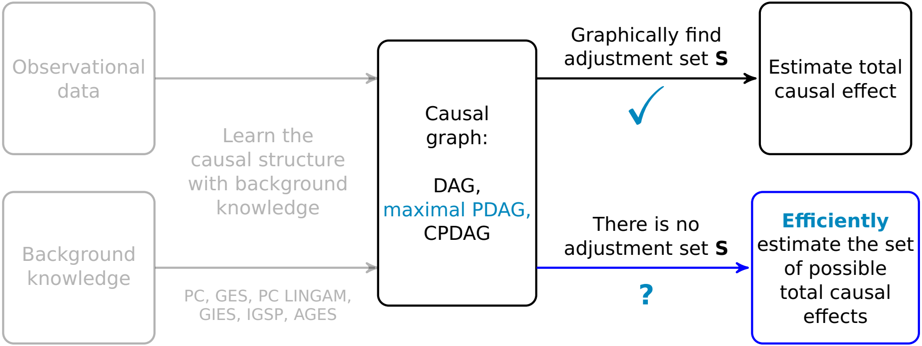
Randomly sampled
20 000 DAGs:
$p \in \{20, 30, \ldots, 100\}$,
$E[N] \in \{3, 4, \ldots, 10\}$.

X - randomly chosen,
Y - connected to X
and $Y \not\to X$.

Framework


Observational data
Background knowledge
Learn the causal structure with background knowledge
PC, GES, PC LINGAM, GIES, IGSP, AGES
Causal graph:
DAG, maximal PDAG, CPDAG
Graphically find adjustment set **S**
✓
Estimate total causal effect
There is no adjustment set **S**
?
Efficiently estimate the set of possible total causal effects

Framework



- Modify IDA and joint-IDA framework.
- Assume a linear Gaussian generating mechanism.

IDA and joint-IDA in maximal PDAGs

- Use sets of direct causes (parents) of **X** to estimate all possible total causal effect of **X** on **Y**.

IDA and joint-IDA in maximal PDAGs

- Use sets of direct causes (parents) of $\mathbf{X}$ to estimate all possible total causal effect of $\mathbf{X}$ on $\mathbf{Y}$.

- Find all sets of parents of $\mathbf{X}$ in $\mathcal{G}$

IDA and joint-IDA in maximal PDAGs

- Use sets of direct causes (parents) of $\mathbf{X}$ to estimate all possible total causal effect of $\mathbf{X}$ on $\mathbf{Y}$.
- Find all sets of parents of $\mathbf{X}$ in $\mathcal{G}$ *in an efficient way*.

IDA and joint-IDA in maximal PDAGs

- Use sets of direct causes (parents) of **X** to estimate all possible total causal effect of **X** on **Y**.

- Find all sets of parents of **X** in $\mathcal{G}$ *in an efficient way*.

The same local algorithm cannot be used with added background knowledge, due to partially directed cycles.

IDA and joint-IDA in maximal PDAGs

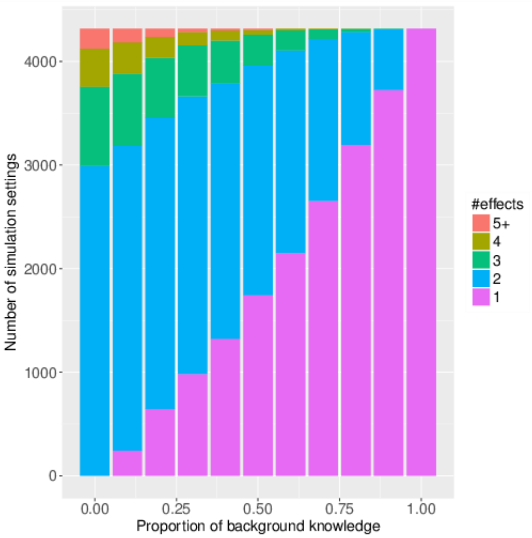
- Use sets of direct causes (parents) of **X** to estimate all possible total causal effect of **X** on **Y**.

- Find all sets of parents of **X** in $\mathcal{G}$ *in an efficient way*.

The same local algorithm cannot be used with added background knowledge, due to partially directed cycles.

Runtime	Median	Mean	Max
IDA w/o bg	0.003	0.003	0.009
IDA w bg	0.003	0.016	4.881

Identifiability gain with background knowledge.



4315 out of 20 000 sampled DAGs:
$p \in \{20, 30, \ldots, 100\}$,
$E[N] \in \{3, 4, \ldots, 10\}$,
$n = 200$.

X - randomly chosen,
Y - connected to X
and $Y \not\to X$.

Implementation

Algorithms implemented in R package pcalg on CRAN:

Implementation

Algorithms implemented in `R` package `pcalg` on CRAN:

- `addBgKnowledge(`$\mathcal{G}$, $\mathbf{X}$, $\mathbf{Y}$`)`

Implementation

Algorithms implemented in R package `pcalg` on CRAN:

- `addBgKnowledge(`$\mathcal{G}, \mathbf{X}, \mathbf{Y}$`)`
- `isValidGraph(`$\mathcal{G}$`,graph.type)`

Implementation

Algorithms implemented in `R` package `pcalg` on CRAN:

- `addBgKnowledge(`$\mathcal{G}$, $\mathbf{X}$, $\mathbf{Y}$`)`

- `isValidGraph(`$\mathcal{G}$`,graph.type)`

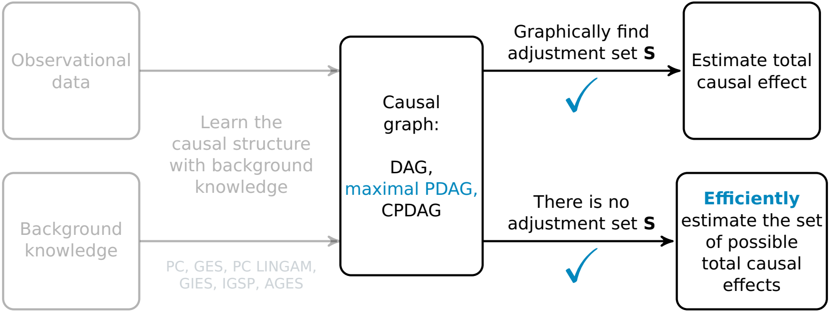
- `gac(`$\mathcal{G}$`,`$\mathbf{X}$`,`$\mathbf{Y}$`,`$\mathbf{S}$`,graph.type)`
- `adjustment(`$\mathcal{G}$`,graph.type,`$\mathbf{X}$`,`$\mathbf{Y}$`,set.type)`

Implementation

Algorithms implemented in R package `pcalg` on CRAN:

- `addBgKnowledge(`$\mathcal{G}, \mathbf{X}, \mathbf{Y}$`)`
- `isValidGraph(`$\mathcal{G}$`,graph.type)`
- `gac(`$\mathcal{G}$`,`$\mathbf{X}$`,`$\mathbf{Y}$`,`$\mathbf{S}$`,graph.type)`
- `adjustment(`$\mathcal{G}$`,graph.type,`$\mathbf{X}$`,`$\mathbf{Y}$`,set.type)`
- `ida(`X, Y`,cov.mat,`$\mathcal{G}$`,graph.type)`
- `jointIda(`$\mathbf{X}, \mathbf{Y}$`,cov.mat,technique,` $\mathcal{G}$ `, graph.type)`

Our contribution



- A necessary and sufficient graphical adjustment criterion for maximal PDAGs.
- Efficient IDA and joint-IDA algorithms for maximal PDAGs.
- Implemented new algorithms and extended the existing algorithms in `pcalg`.

Thanks! See you at the poster session!

Joint work with
Marloes Maathuis, Markus Kalisch

References:

- Perković, Kalisch and Maathuis (2017b). Interpreting and using CPDAGs with background knowledge *UAI 2017*.
- Perković, Textor, Kalisch and Maathuis (2015). A complete generalized adjustment criterion. *UAI 2015*.
- Perković, Textor, Kalisch and Maathuis (2017a). Complete graphical characterization and construction of adjustment sets in Markov equivalence classes of ancestral graphs. *arXiv:1606.06903, to appear in JMLR*.