Total causal effects in MPDAGs: identification and minimal enumeration

Emilija Perković

Department of Statistics, University of Washington

Joint with F. Richard Guo

Goal

- Estimate the total causal effect of A on Y

Observational data

Randomized control studies

Goal

- Estimate the total causal effect of A on Y
  - the change in Y due to $do(a)$-
  from observational data.

- $do(a)$: an intervention that sets variables A to a.

Observational data

Randomized control studies

Goal

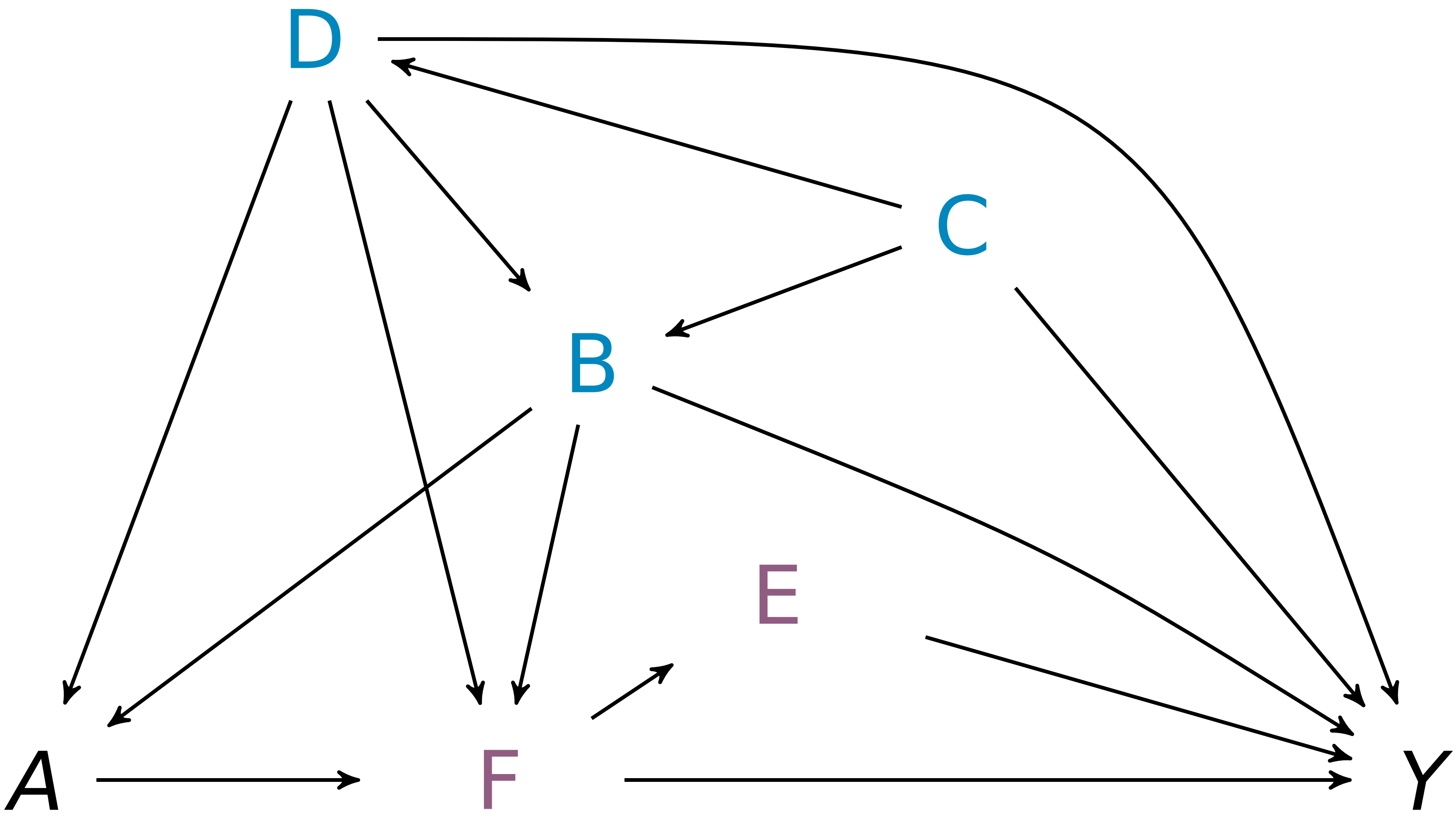
- Estimate the total causal effect of A on Y
  - the change in Y due to $do(a)$-
  from observational data.

- $do(a)$: an intervention that sets variables A to a.
  $f(y|do(a)) \neq f(y|a)$.

Observational data

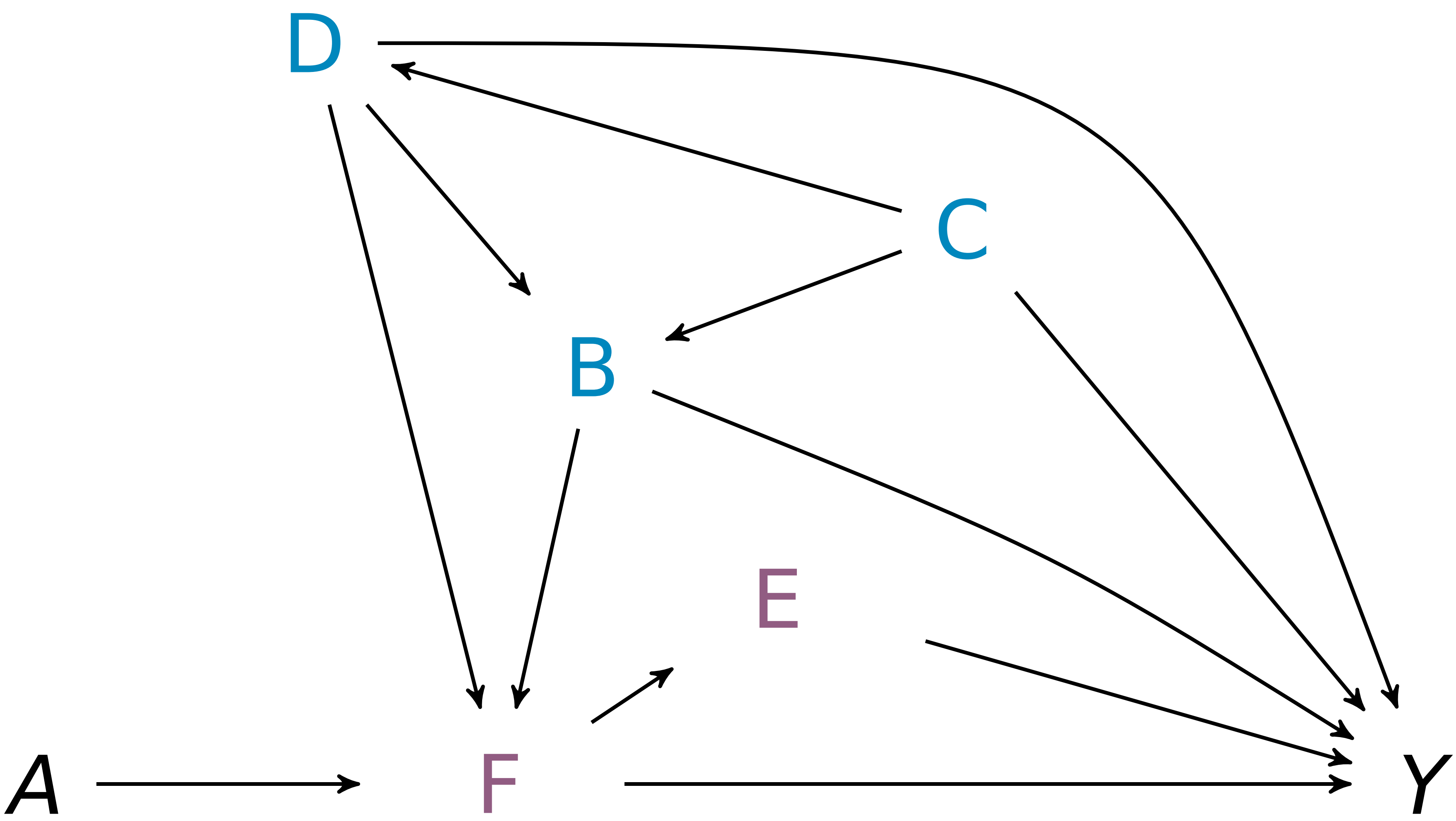
Randomized control studies

Observational Causal DAG



Causal Directed Acyclic Graph (DAG) $\mathcal{D}$.

Interventional Causal DAG



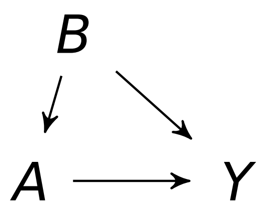
Causal DAG $\mathcal{D}$ **after a "do"-intervention on** A**.**

DAGs and distributions

- $do(\mathbf{a})$: an intervention that sets variables $\mathbf{A}$ to $\mathbf{a}$.
- Observational density $f(\mathbf{v})$, Interventional density $f(\mathbf{v}|do(\mathbf{a}))$.

- A DAG $\mathcal{D}$ is causal if for all observational and interventional densities:

$$f(\mathbf{v}) = \prod_{V_j \in \mathbf{V}} f(v_j | pa(v_j, \mathcal{D})) \quad \text{and} \quad f(\mathbf{v}|do(\mathbf{a})) = \prod_{V_j \in \mathbf{V} \setminus \mathbf{A}} f(v_j | pa(v_j, \mathcal{D})).$$



$$f(b, a, y) = f(y|b, a) f(a|b) f(b)$$

$$f(b, y|do(a)) = f(y|b, a) f(b)$$

How to define a causal effect?

Total causal effect

- Total causal effect - $\tau_{\mathbf{ay}}$ - is *some functional* of $f(\mathbf{y}|do(\mathbf{a}))$, $P(\mathbf{Y}|do(\mathbf{a}))$.
- Examples: $E[Y|do(A=a+1)] - E[Y|do(A=a)]$, $\frac{\partial}{\partial a}E(Y|do(a))$, OR, RR...

Identifiability

- A causal effect is identifiable from observational data if

$$f(\mathbf{y}|do(\mathbf{a})) \text{ is computable from } f(\mathbf{v}).$$

How to define a causal effect?

Total causal effect

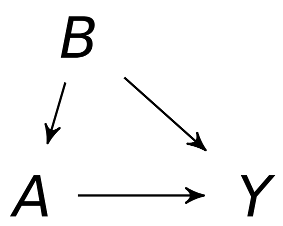
- Total causal effect - $\tau_{\mathbf{ay}}$ - is *some functional* of $f(\mathbf{y}|do(\mathbf{a}))$, $P(\mathbf{Y}|do(\mathbf{a}))$.
- Examples: $E[Y|do(A = a + 1)] - E[Y|do(A = a)]$, $\frac{\partial}{\partial a}E(Y|do(a))$, OR, RR...

Identifiability

- A causal effect is identifiable from observational data if

$$f(\mathbf{y}|do(\mathbf{a})) \text{ is computable from } f(\mathbf{v}).$$

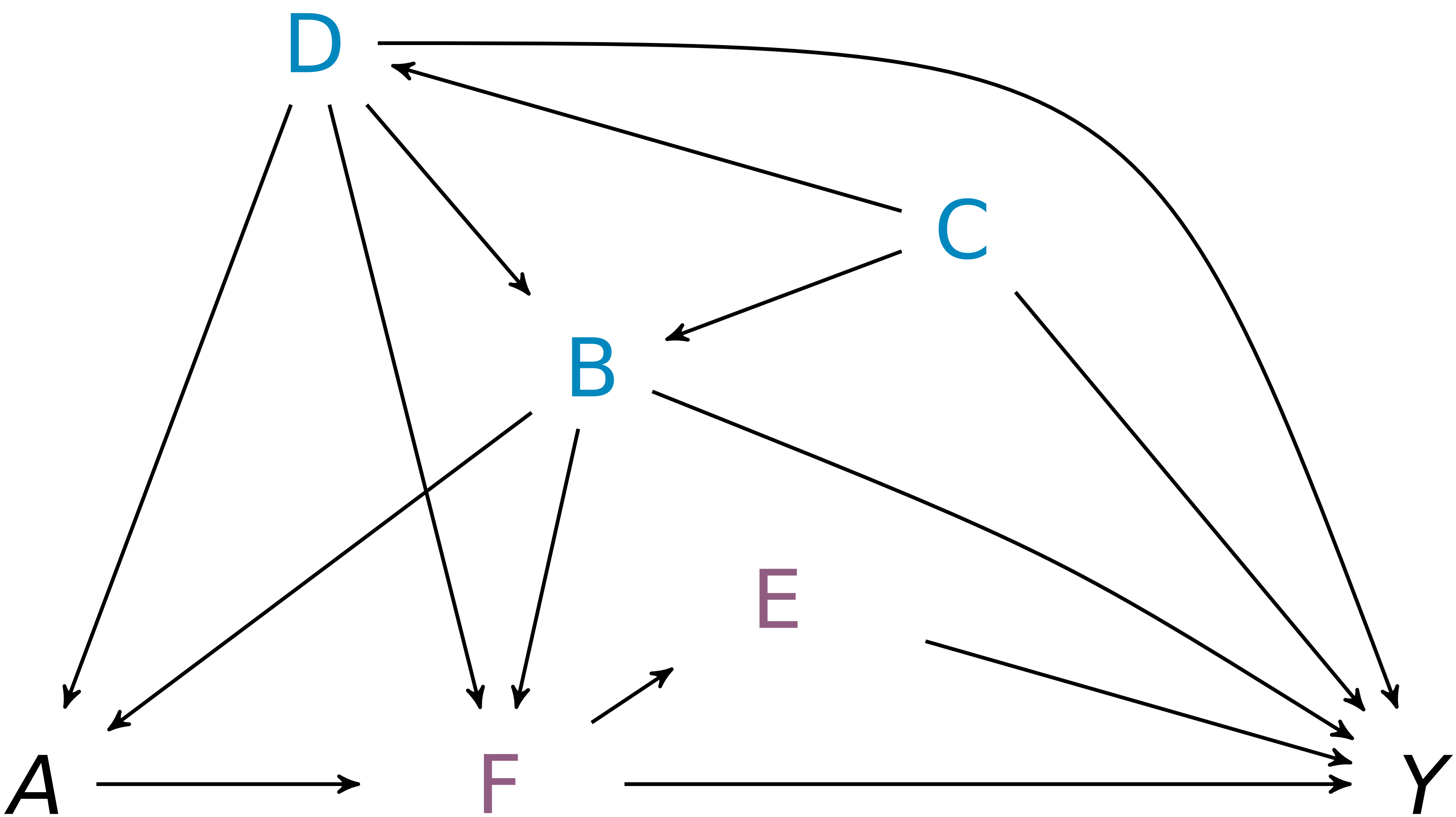
- Given the causal DAG, every total causal effect is identifiable.



$$f(y|do(a)) = \int f(b, y|do(a))db$$
$$= \int f(y|b, a)f(b)db.$$

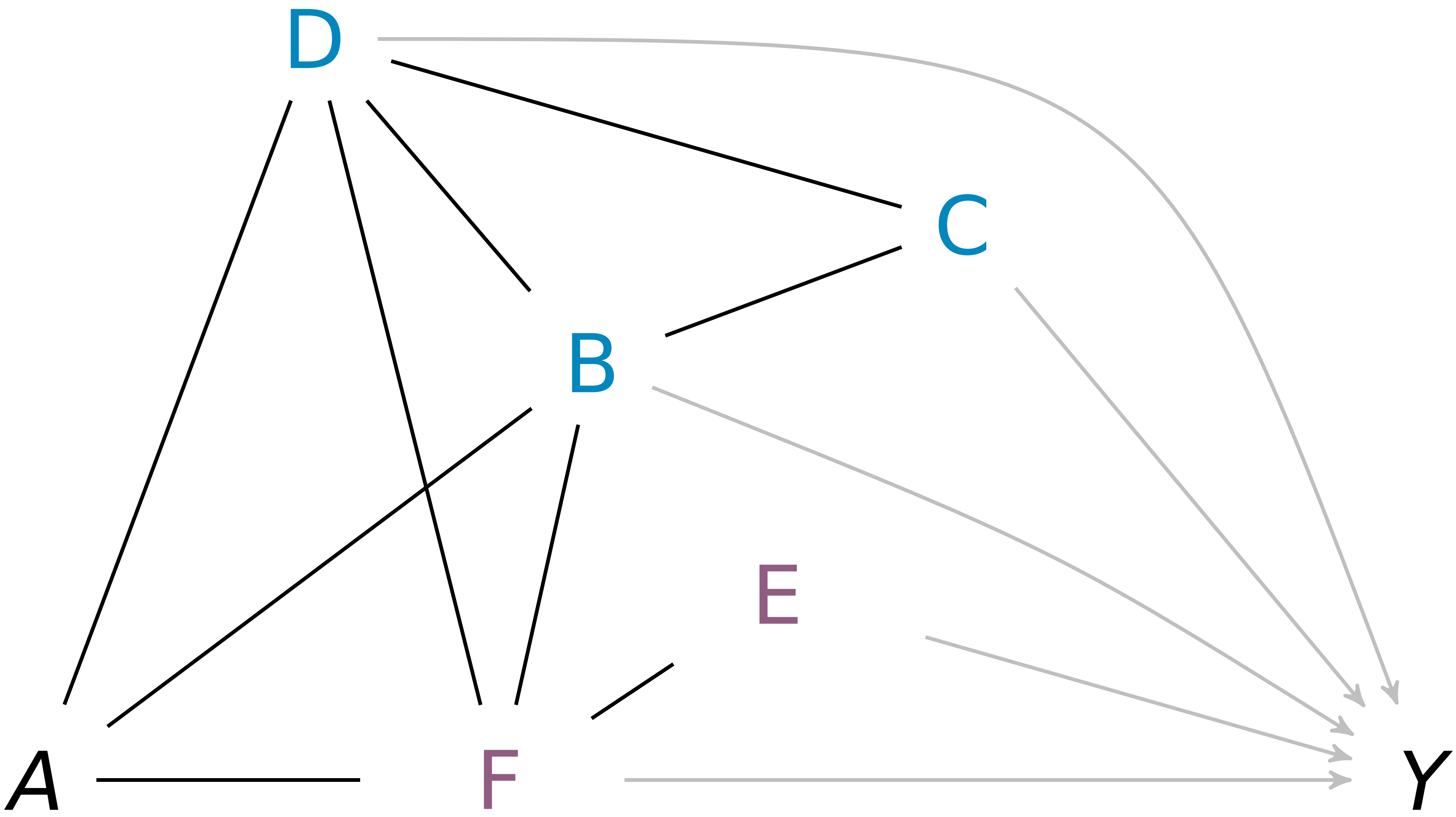
(Generalized) G-formula

Problem solved?



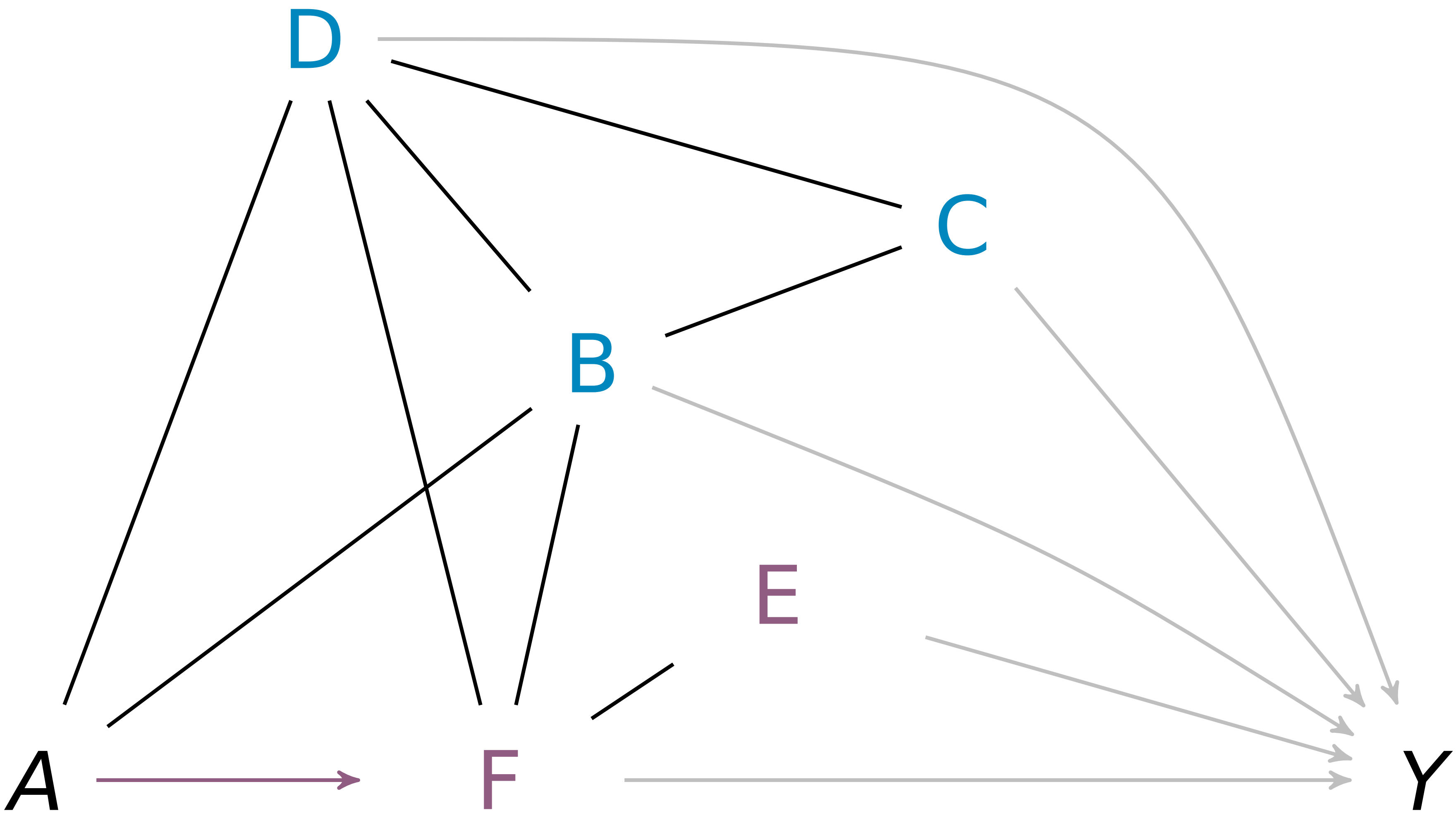
DAG $\mathcal{D}$.

Problem solved?



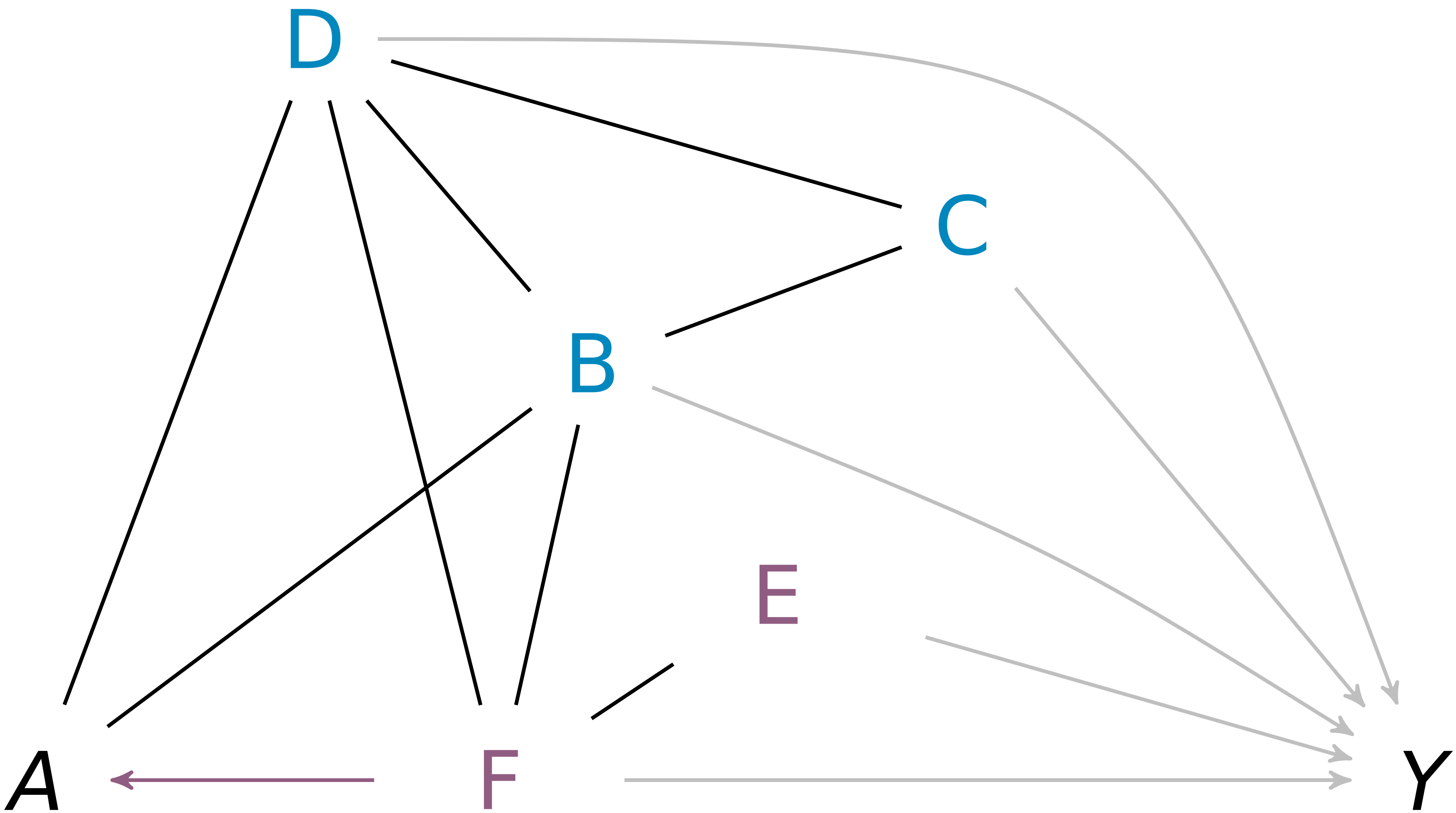
Completed Partially Directed Acyclic Graph (CPDAG) $\mathcal{C}$.

Problem solved?



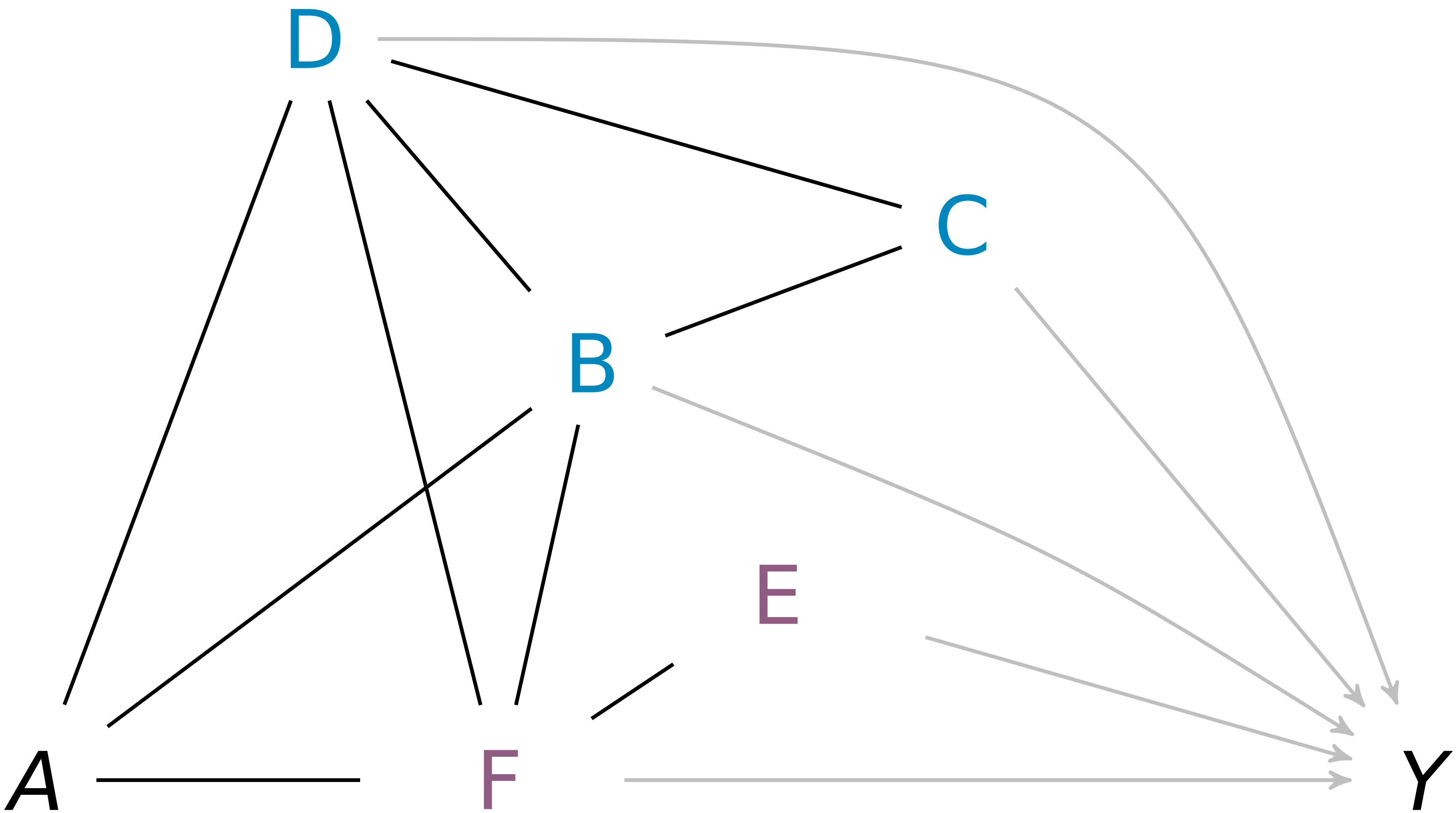
Completed Partially Directed Acyclic Graph (CPDAG) $\mathcal{C}$.

Problem solved?



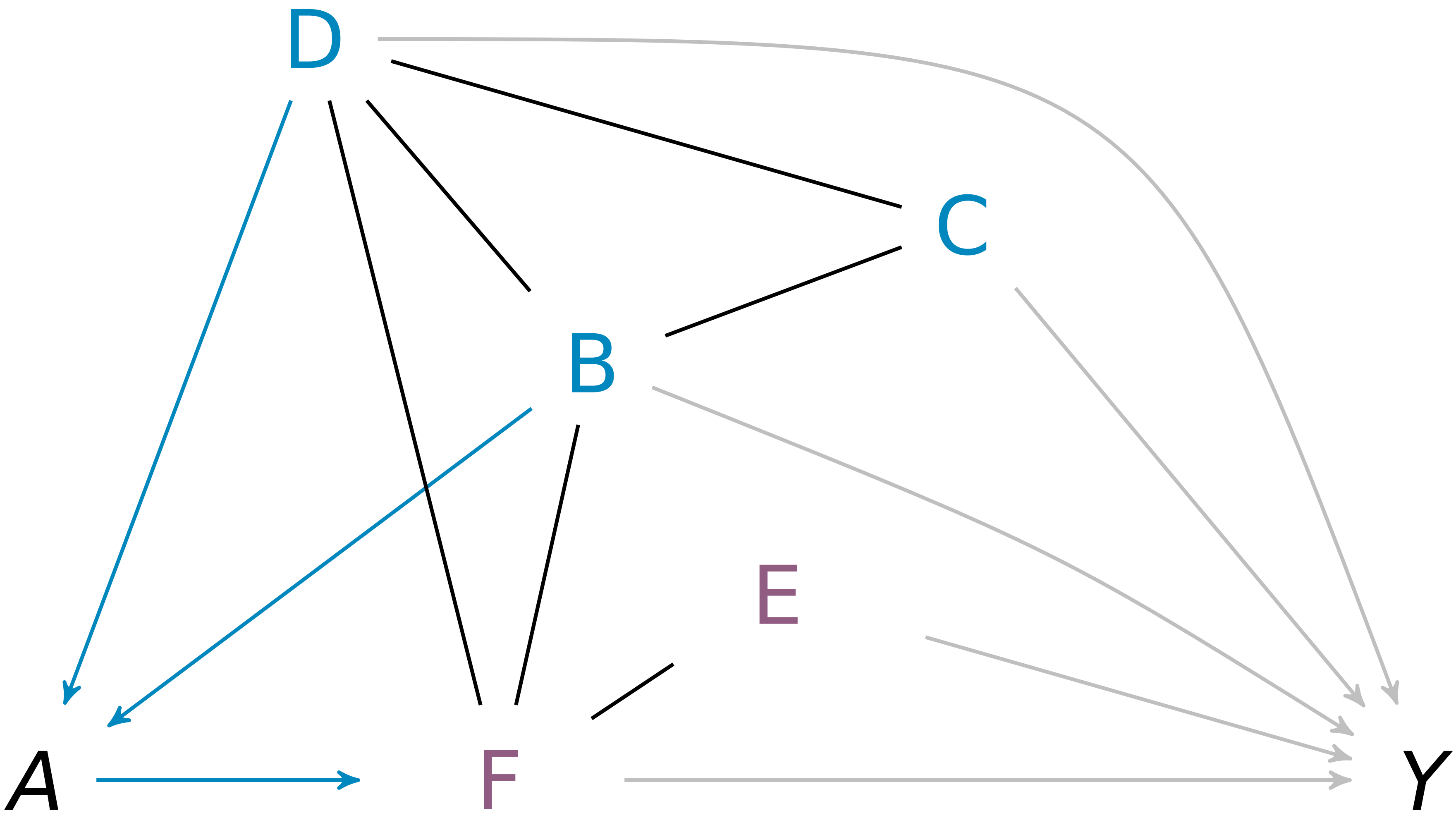
Completed Partially Directed Acyclic Graph (CPDAG) $\mathcal{C}$.

Problem solved?



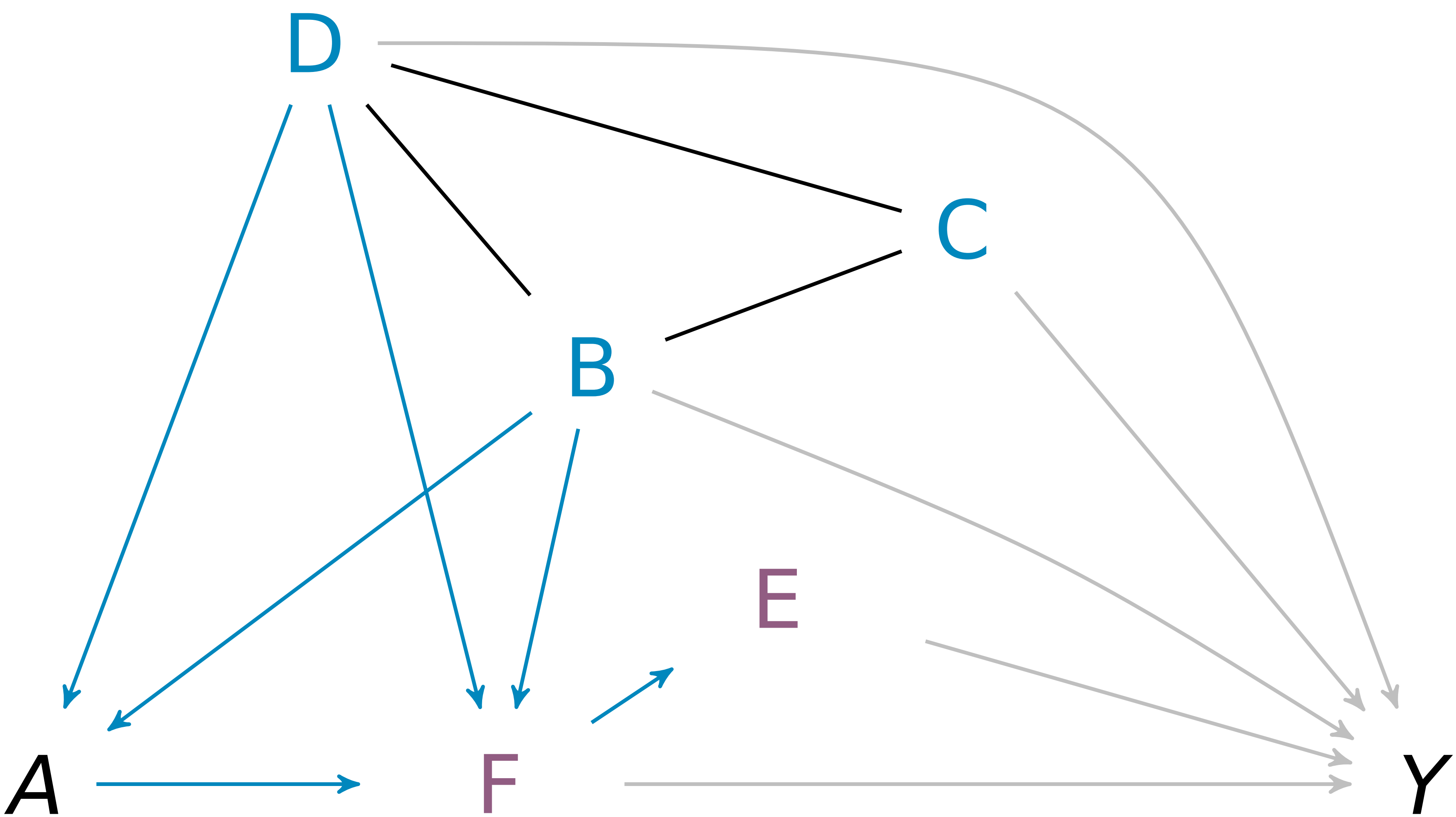
Completed Partially Directed Acyclic Graph (CPDAG) $\mathcal{C}$.

Problem solved?



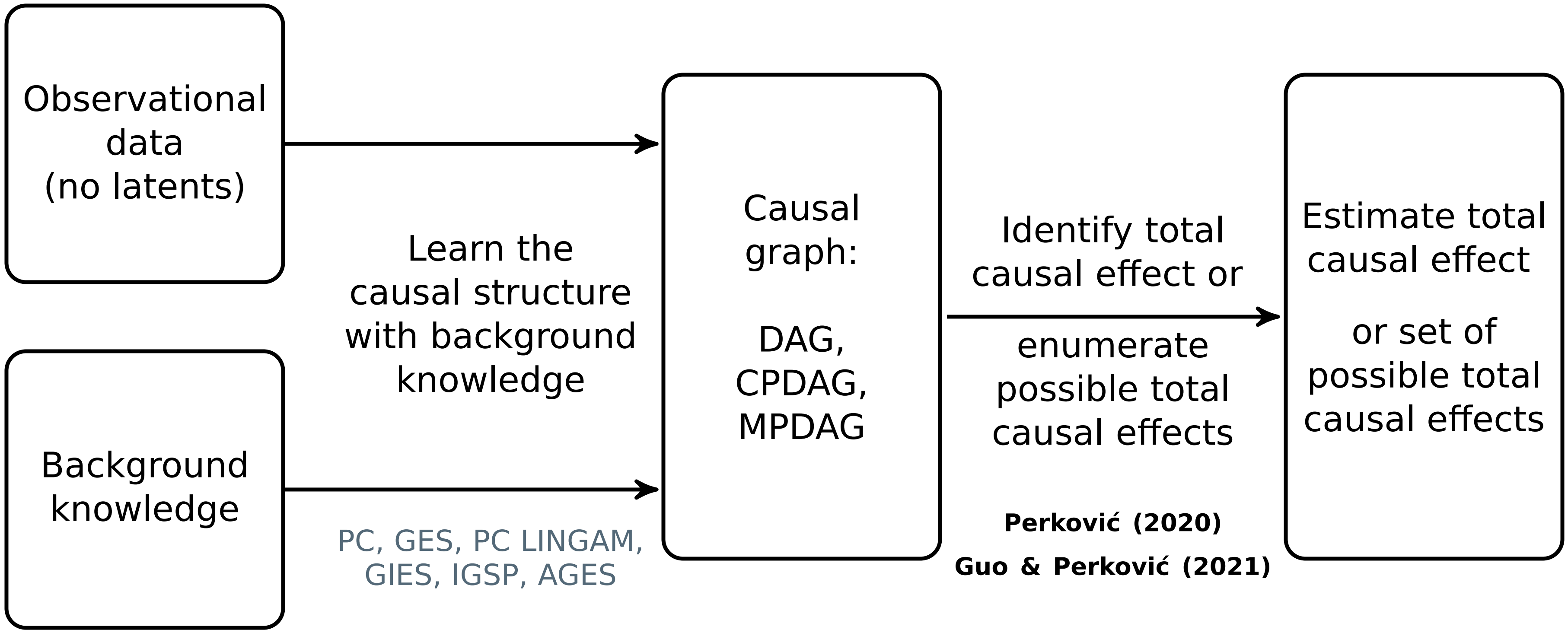
PDAG with Bg. Knowledge.

Problem solved?



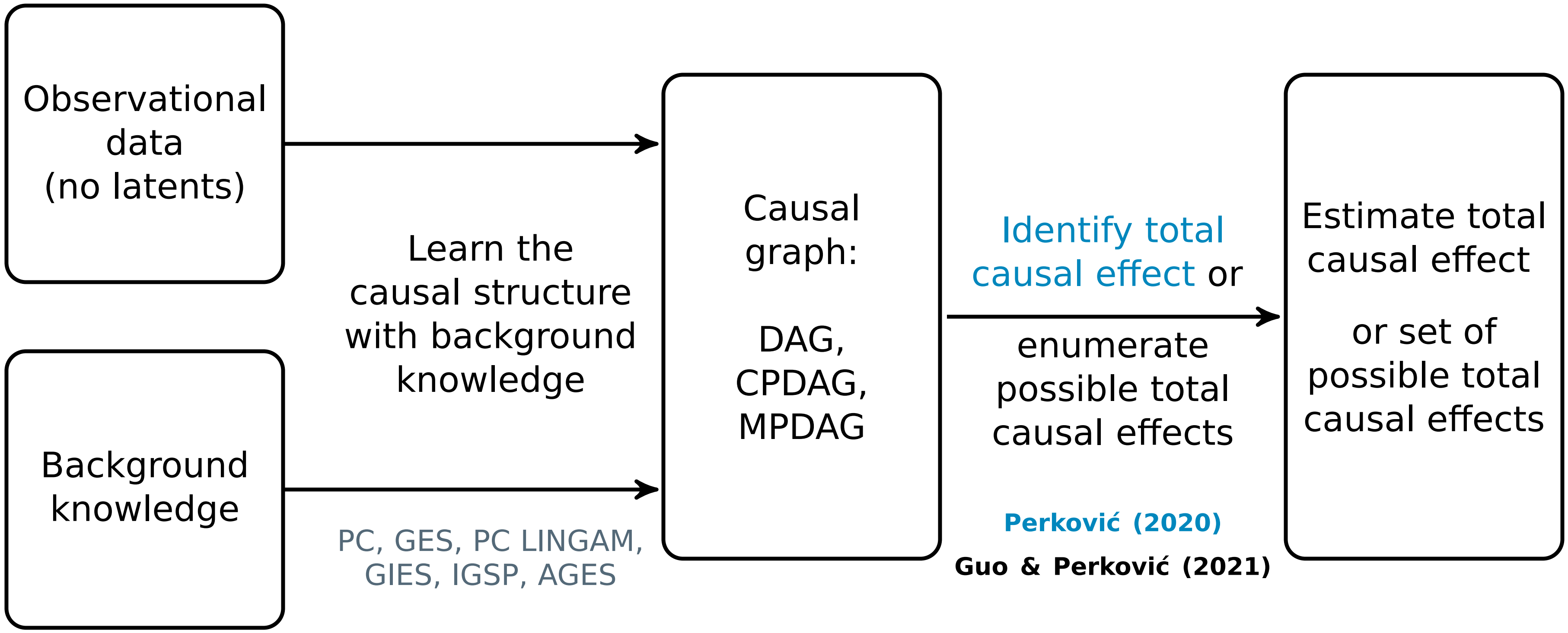
Maximally oriented Partially Directed Acyclic Graph (MPDAG) $\mathcal{G}$.

Framework



- PC (Spirtes et al, 1993), GES (Chickering, 2002) + Adding background knowledge (Meek, 1995; TETRAD, Scheines et al., 1998), PC LINGAM (Hoyer et al., 2008), GIES (Hauser and Bühlmann, 2012), IGSP (Wang et al., 2017), etc.
- Other framing: start with a DAG and remove some directional information while keeping the orientations closed under Meek orientation rules (Meek, 1995).

Framework



- PC (Spirtes et al, 1993), GES (Chickering, 2002) + Adding background knowledge (Meek, 1995; TETRAD, Scheines et al., 1998), PC LINGAM (Hoyer et al., 2008), GIES (Hauser and Bühlmann, 2012), IGSP (Wang et al., 2017), etc.
- Other framing: start with a DAG and remove some directional information while keeping the orientations closed under Meek orientation rules (Meek, 1995).

Overview of graphical criteria for identification

Graphical criterion	DAG	CPDAG	MPDAG
Adjustment (Pearl ’93, Shpitser et al ’10, Perković et al ’15, ’17, ’18)	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$
G-formula, Truncated Factorization (Robins ’86, Pearl ’93, Spirtes ’93)	$\Leftrightarrow$		

$\Rightarrow$ - sufficient for identification,
$\Leftrightarrow$ - necessary and sufficient for identification

Overview of graphical criteria for identification

Graphical criterion	DAG	CPDAG	MPDAG
Adjustment (Pearl '93, Shpitser et al '10, Perković et al '15, '17, '18)	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$
G-formula, Truncated Factorization (Robins '86, Pearl '93, Spirtes '93)	$\Leftrightarrow$		

$\Rightarrow$ - sufficient for identification,
$\Leftrightarrow$ - necessary and sufficient for identification

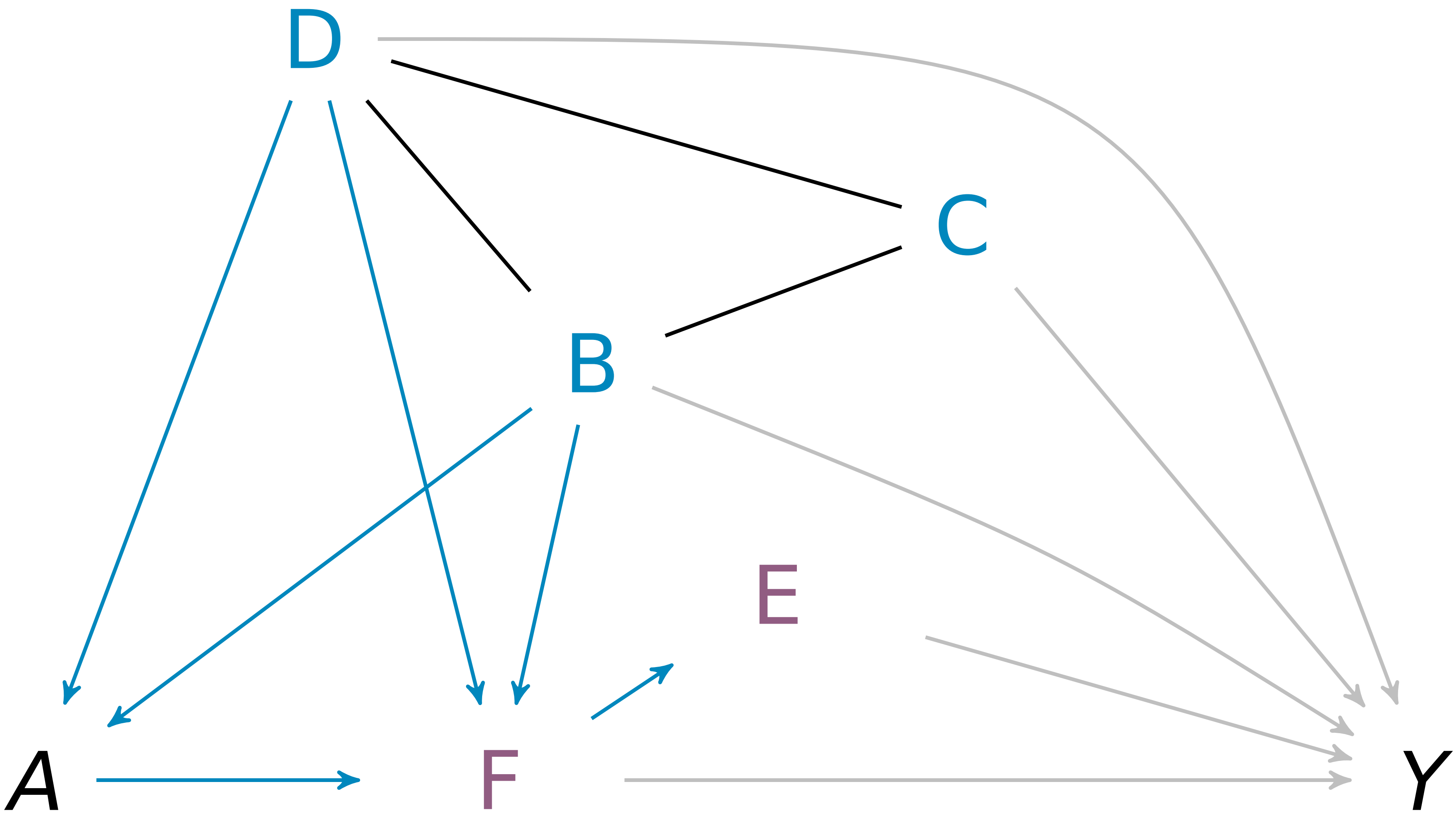
Adjustment: $\mathbf{Z}$ is an adjustment set if

$$f(\mathbf{y}|do(\mathbf{a})) = \int f(\mathbf{y}|\mathbf{a}, \mathbf{z}) f(\mathbf{z}) d\mathbf{z}$$

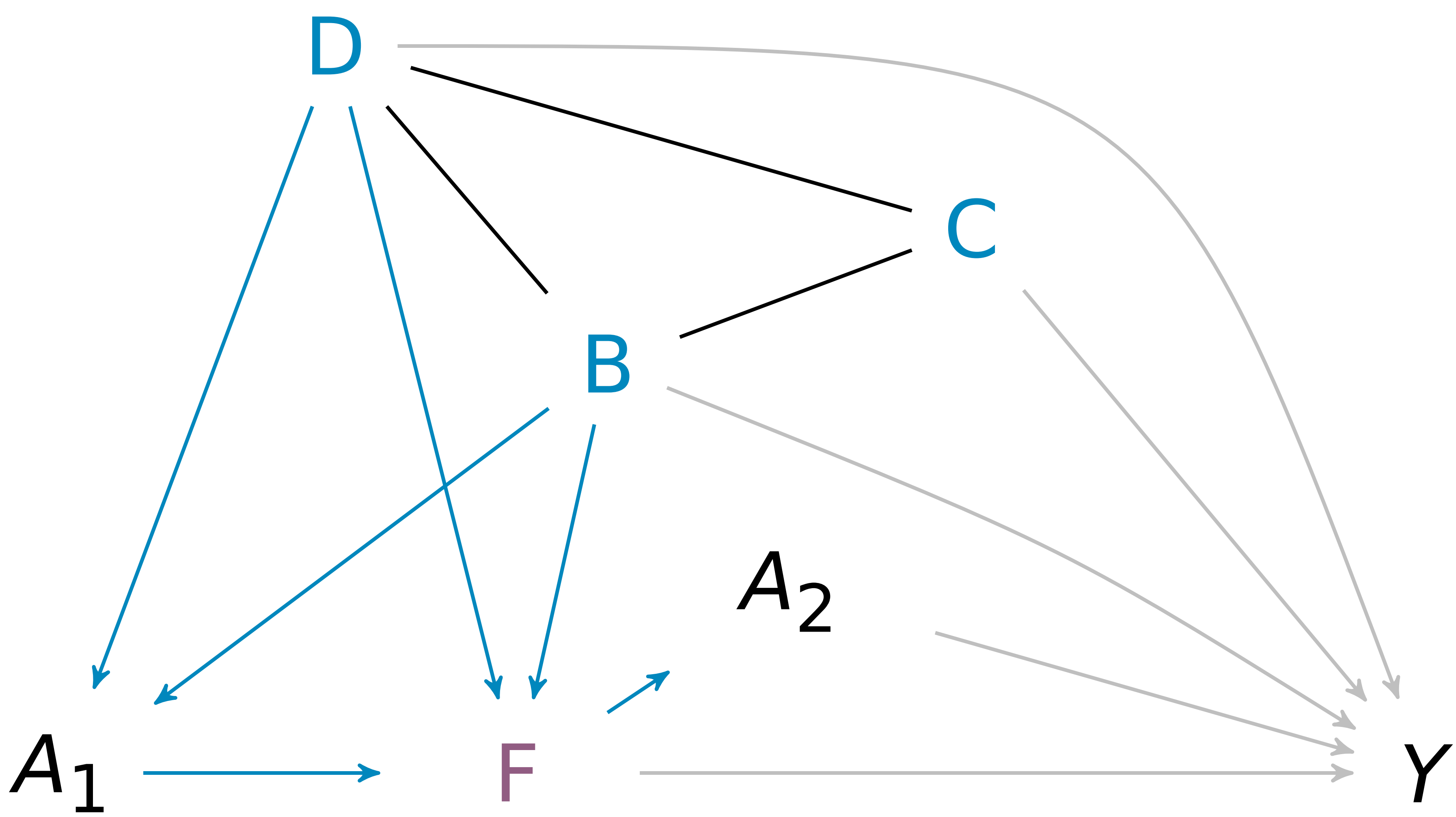
Truncated Factorization: Let $\mathbf{V}' = \mathbf{V} \setminus \{\mathbf{A} \cup \mathbf{Y}\}$, then

$$f(\mathbf{y}|do(\mathbf{a})) = \int \prod_{V_i \in \mathbf{V} \setminus \mathbf{A}} f(v_i | pa(v_i, \mathcal{D})) d\mathbf{v}'.$$

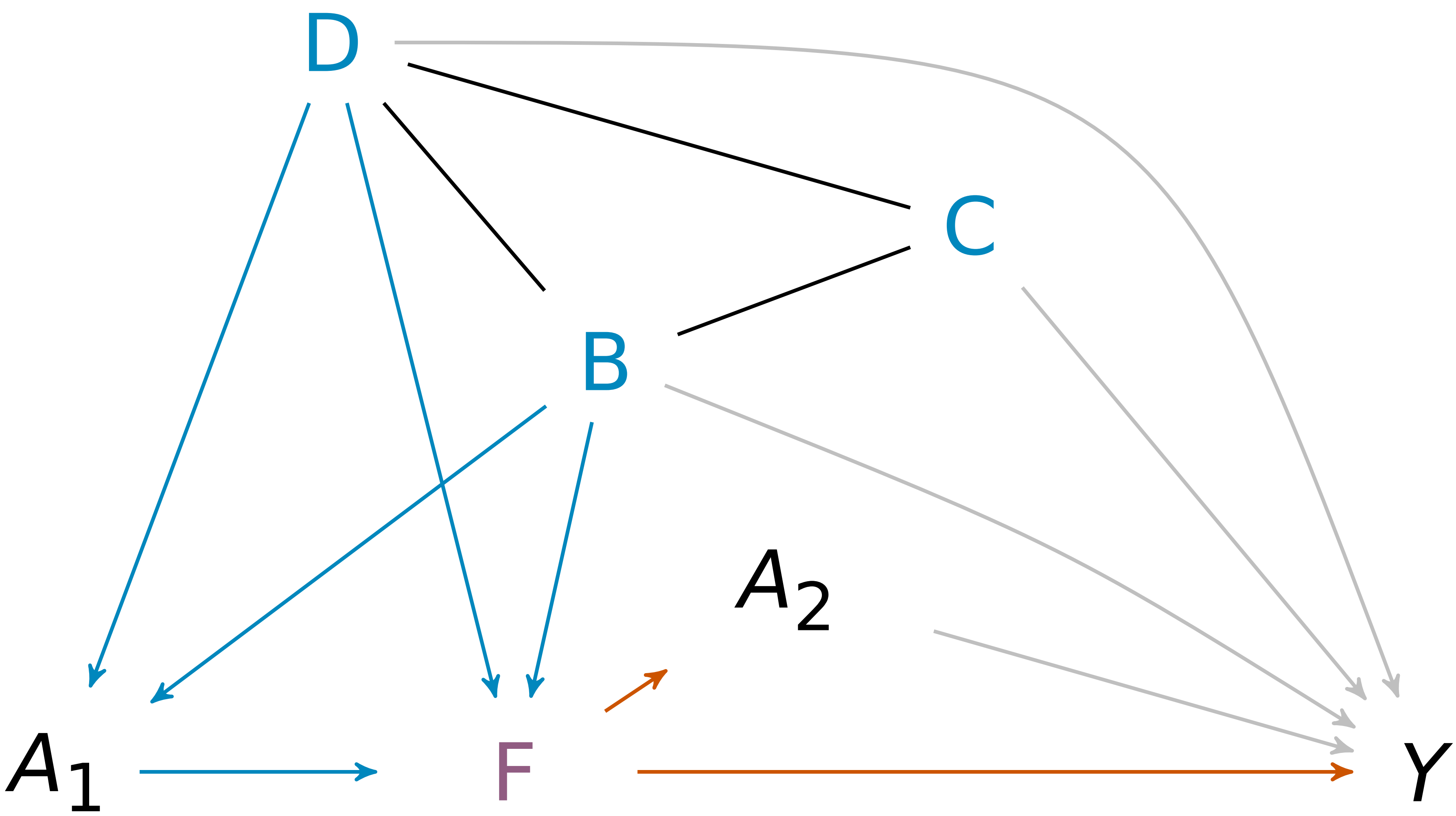
Joint Effects



Does and adjustment set always exist?



Does and adjustment set always exist?



Overview of graphical criteria for identification

Graphical criterion	DAG	CPDAG	MPDAG
Adjustment (Pearl '93, Shpitser et al '10, Perković et al '15, '17, '18)	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$
G-formula, Truncated Factorization (Robins '86, Pearl '93)	$\Leftrightarrow$		

$\Rightarrow$ - sufficient for identification,
$\Leftrightarrow$ - necessary and sufficient for identification

Overview of graphical criteria for identification

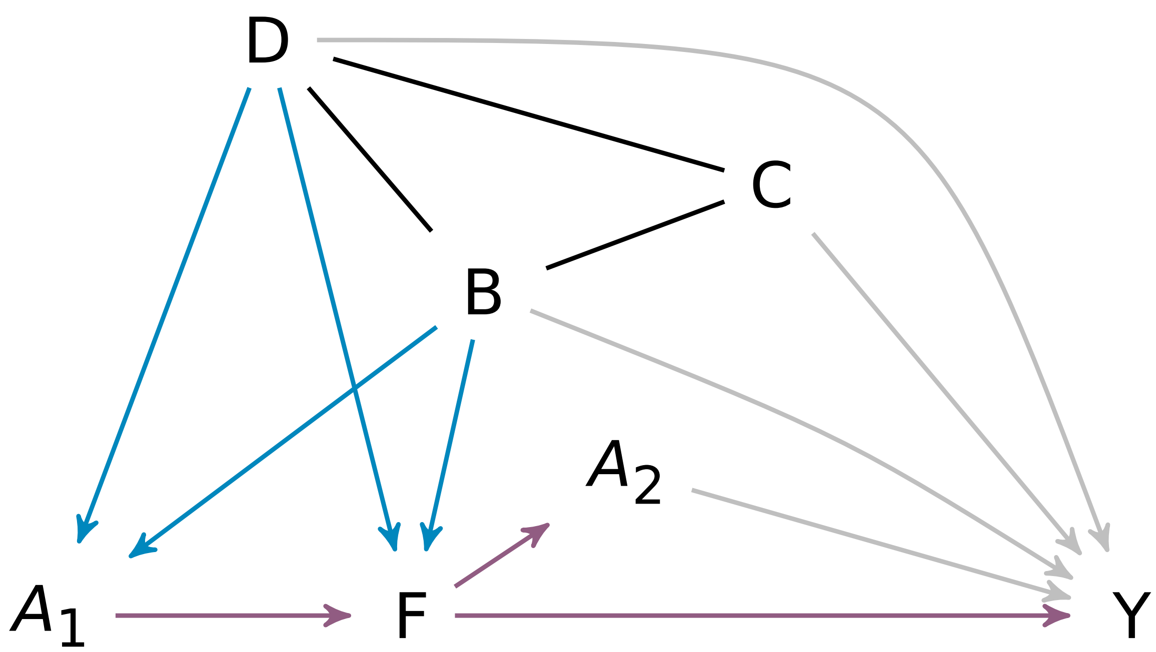
Graphical criterion	DAG	CPDAG	MPDAG
Adjustment (Pearl '93, Shpitser et al '10, Perković et al '15, '17, '18)	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$
G-formula, Truncated Factorization (Robins '86, Pearl '93)	$\Leftrightarrow$		
Causal identification formula (Perković '20)	$\Leftrightarrow$	$\Leftrightarrow$	$\Leftrightarrow$

$\Rightarrow$ - sufficient for identification,
$\Leftrightarrow$ - necessary and sufficient for identification

Necessary and sufficient condition

Theorem (Perković, 2020)

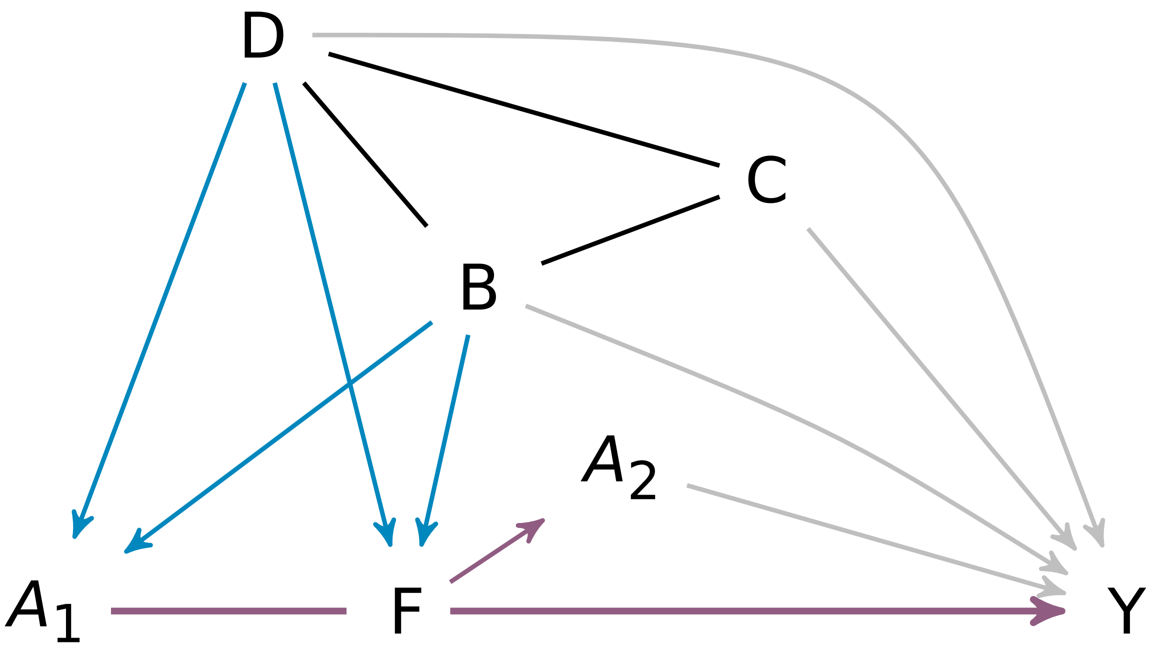
The total causal effect of **A** on **Y** is identifiable in MPDAG $\mathcal{G}$ if and only if **all proper possibly causal paths** from **A** to **Y** start with a directed edge in $\mathcal{G}$.



Necessary and sufficient condition

Theorem (Perković, 2020)

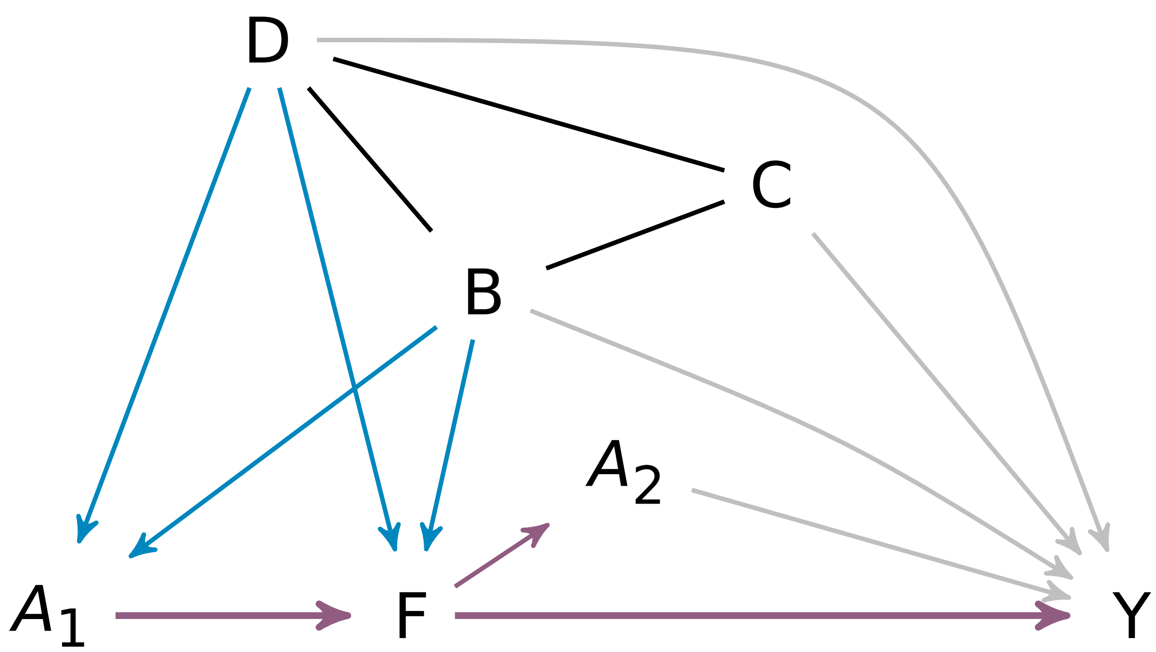
The total causal effect of **A** on **Y** is identifiable in MPDAG $\mathcal{G}$ if and only if **all proper possibly causal paths** from **A** to **Y** start with a directed edge in $\mathcal{G}$.



Necessary and sufficient condition

Theorem (Perković, 2020)

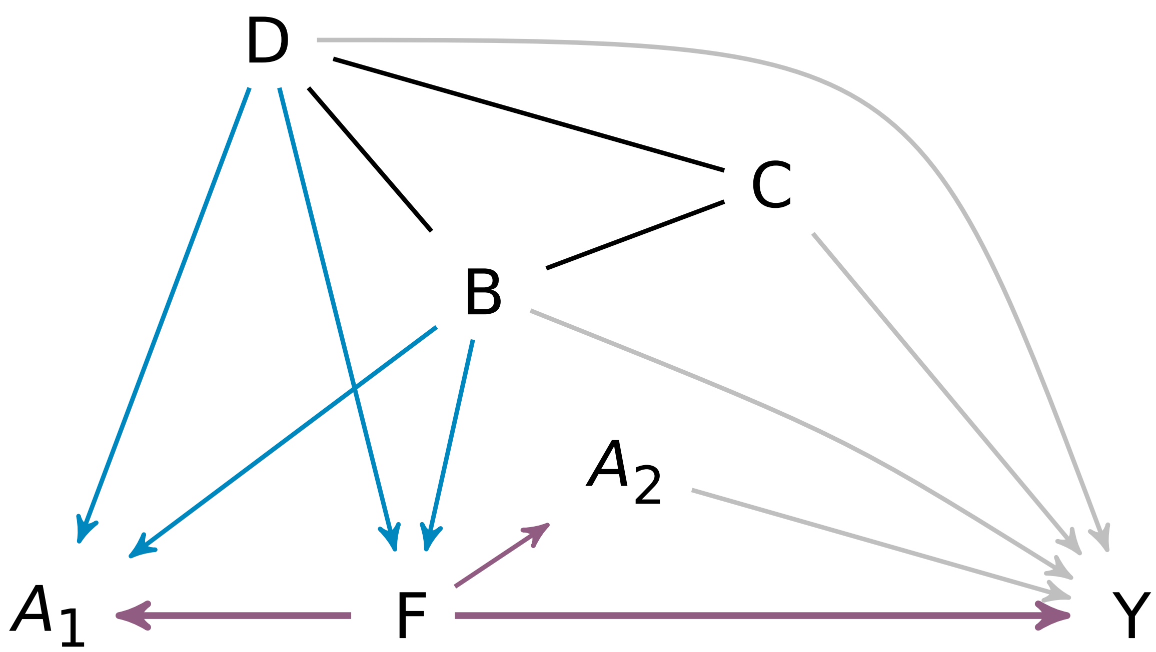
The total causal effect of **A** on **Y** is identifiable in MPDAG $\mathcal{G}$ if and only if **all proper possibly causal paths** from **A** to **Y** start with a directed edge in $\mathcal{G}$.



Necessary and sufficient condition

Theorem (Perković, 2020)

The total causal effect of **A** on **Y** is identifiable in MPDAG $\mathcal{G}$ if and only if **all proper possibly causal paths** from **A** to **Y** start with a directed edge in $\mathcal{G}$.



Causal identification formula

Theorem (Perković, 2020)

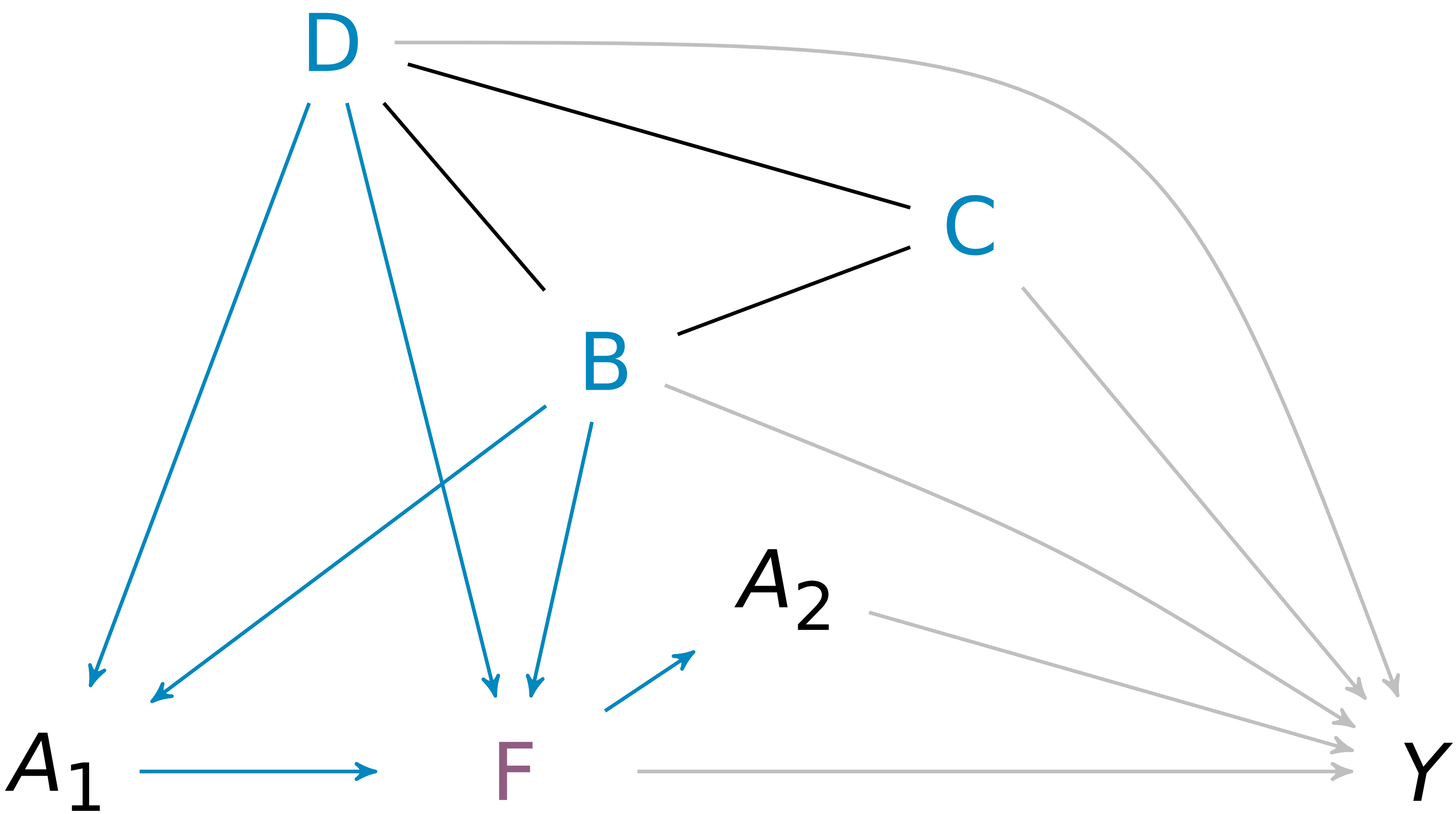
If **all proper possibly causal paths** from $\mathbf{A}$ to $\mathbf{Y}$ start with a directed edge in $\mathcal{G}$, then

$$f(\mathbf{y}|do(\mathbf{a})) = \int \prod_{i=1}^{k} f(\mathbf{s_i}|pa(\mathbf{s_i}, \mathcal{G}))d\mathbf{s},$$

where $\mathbf{S} = an(\mathbf{Y}, \mathcal{G}_{\mathbf{V} \setminus \mathbf{A}}) \setminus \mathbf{Y}$,
and $(\mathbf{S_1}, \dots, \mathbf{S_k})$ is a partition of $\mathbf{S} \cup \mathbf{Y}$ into undirected connected sets in $\mathcal{G}$.

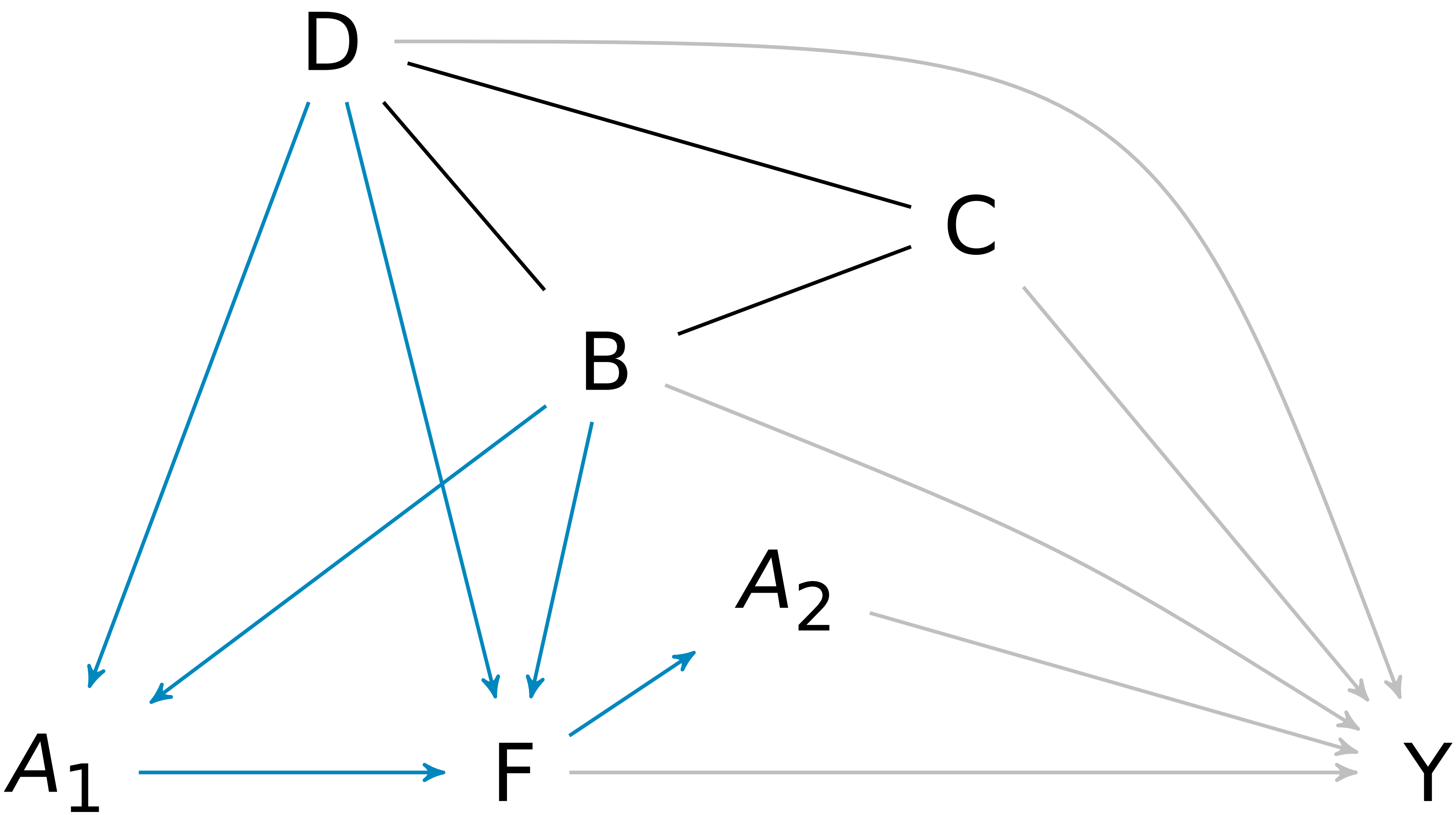
- $\mathbf{S} \cup \mathbf{Y} = an(\mathbf{Y}, \mathcal{G}_{\mathbf{V} \setminus \mathbf{A}})$ - nodes that have a causal path to $\mathbf{Y}$ that is not through $\mathbf{A}$.
- $(\mathbf{S_1}, \dots, \mathbf{S_k})$ - maximal connected components of $\mathbf{S} \cup \mathbf{Y}$ in the induced undirected subgraph of $\mathcal{G}$.

How to use the causal identification formula?



$$f(y|do(a_1,a_2)) = \int f(y|f,b,c,d,a_2)f(f|b,d,a_1)f(b,c,d)df\,db\,dc\,dd$$

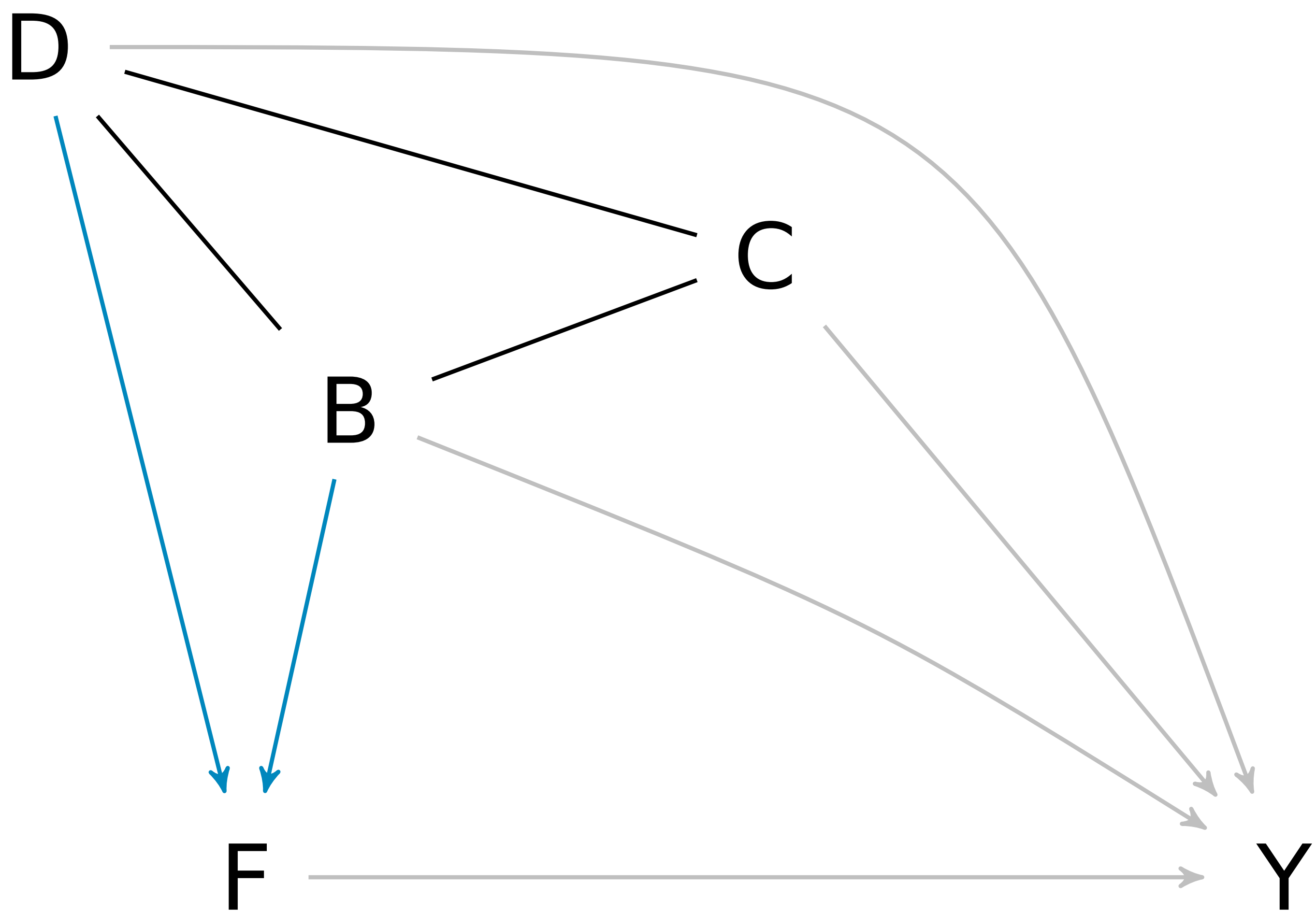
$f(y|do(a_1, a_2)) = ?$



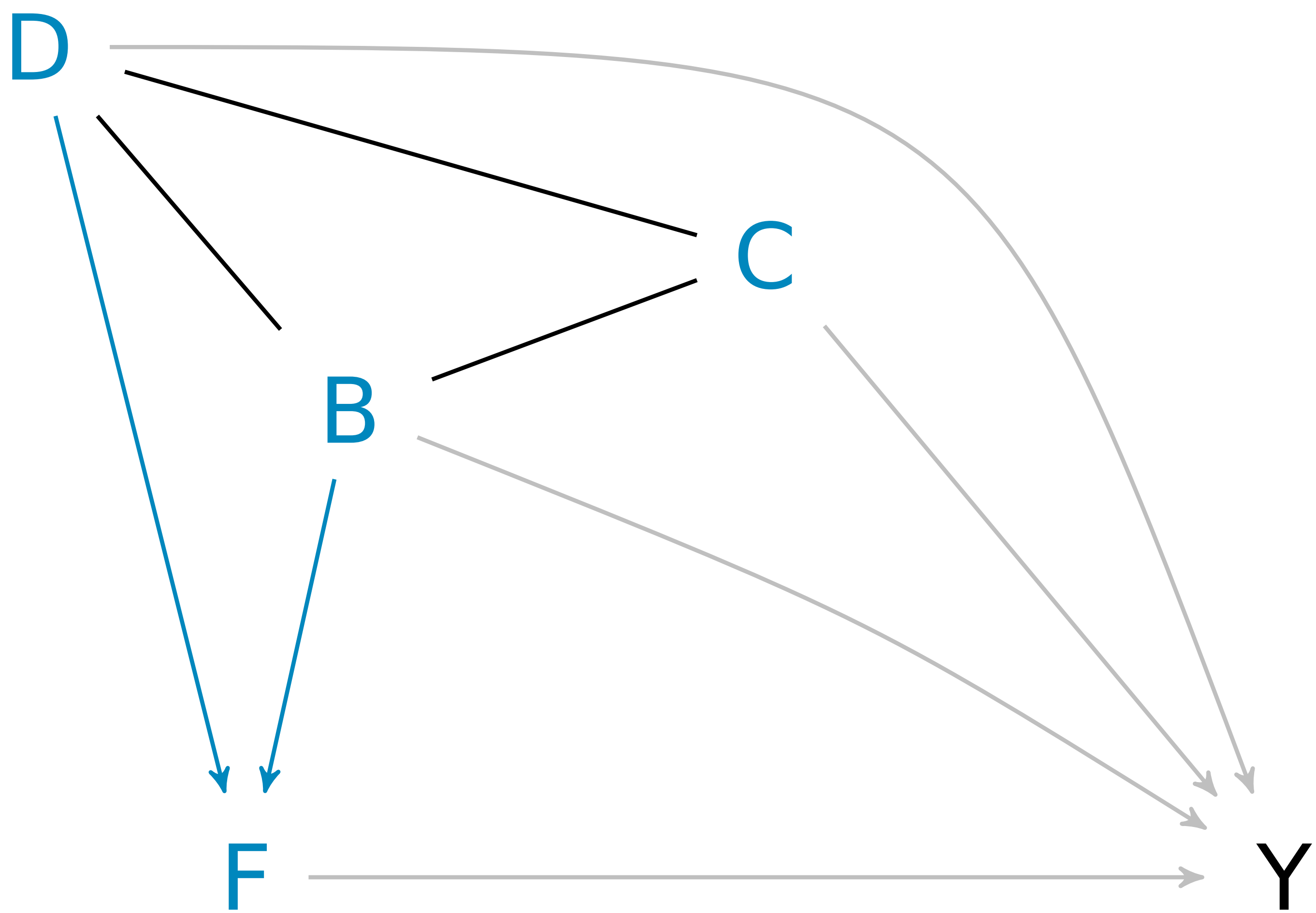
- $\mathbf{S} = an(Y, \mathcal{G}_{\mathbf{V} \setminus \mathbf{A}}) \setminus \{Y\} =$

Partition of $\mathbf{S} \cup \{Y\} =$

$f(y|do(a_1, a_2)) = ?$

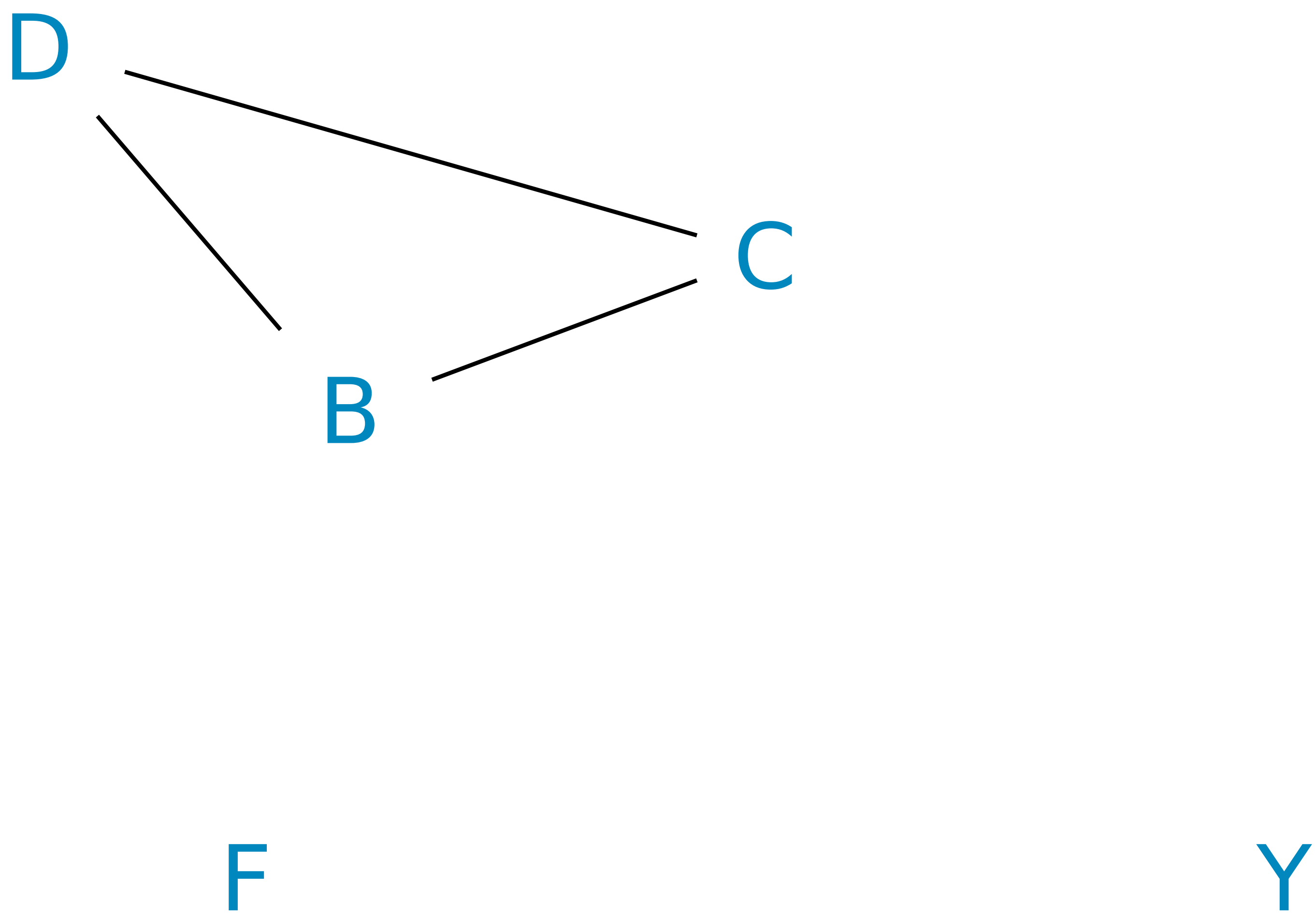


- $\mathbf{S} = an(Y, \mathcal{G}_{\mathbf{V} \setminus \mathbf{A}}) \setminus \{Y\} =$ Partition of $\mathbf{S} \cup \{Y\} =$



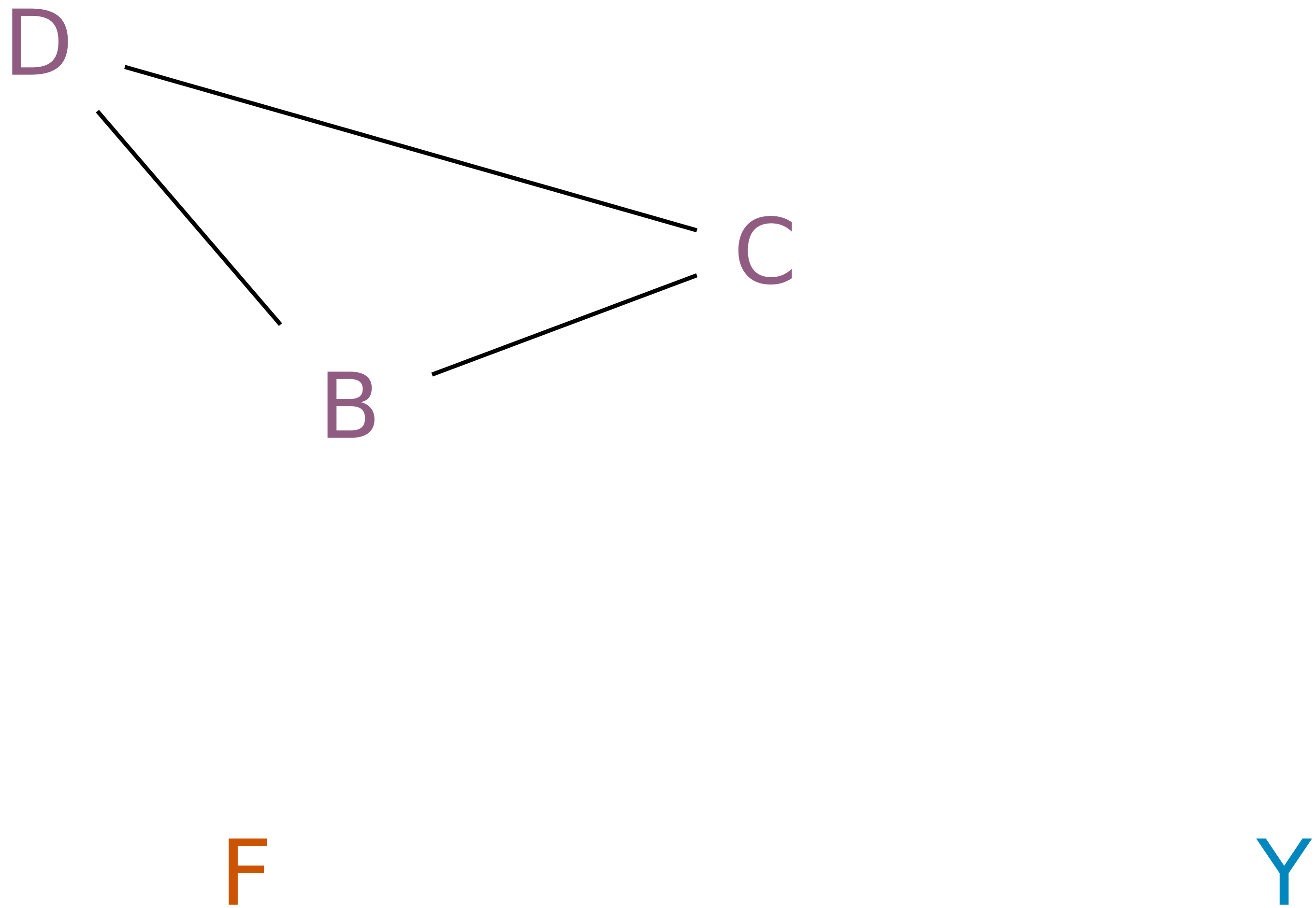
- $\mathbf{S} = an(Y, \mathcal{G}_{\mathbf{V}\setminus\mathbf{A}}) \setminus \{Y\} = \{F, B, C, D\}$, Partition of $\mathbf{S} \cup \{Y\} =$

$f(y|do(a_1, a_2)) = ?$



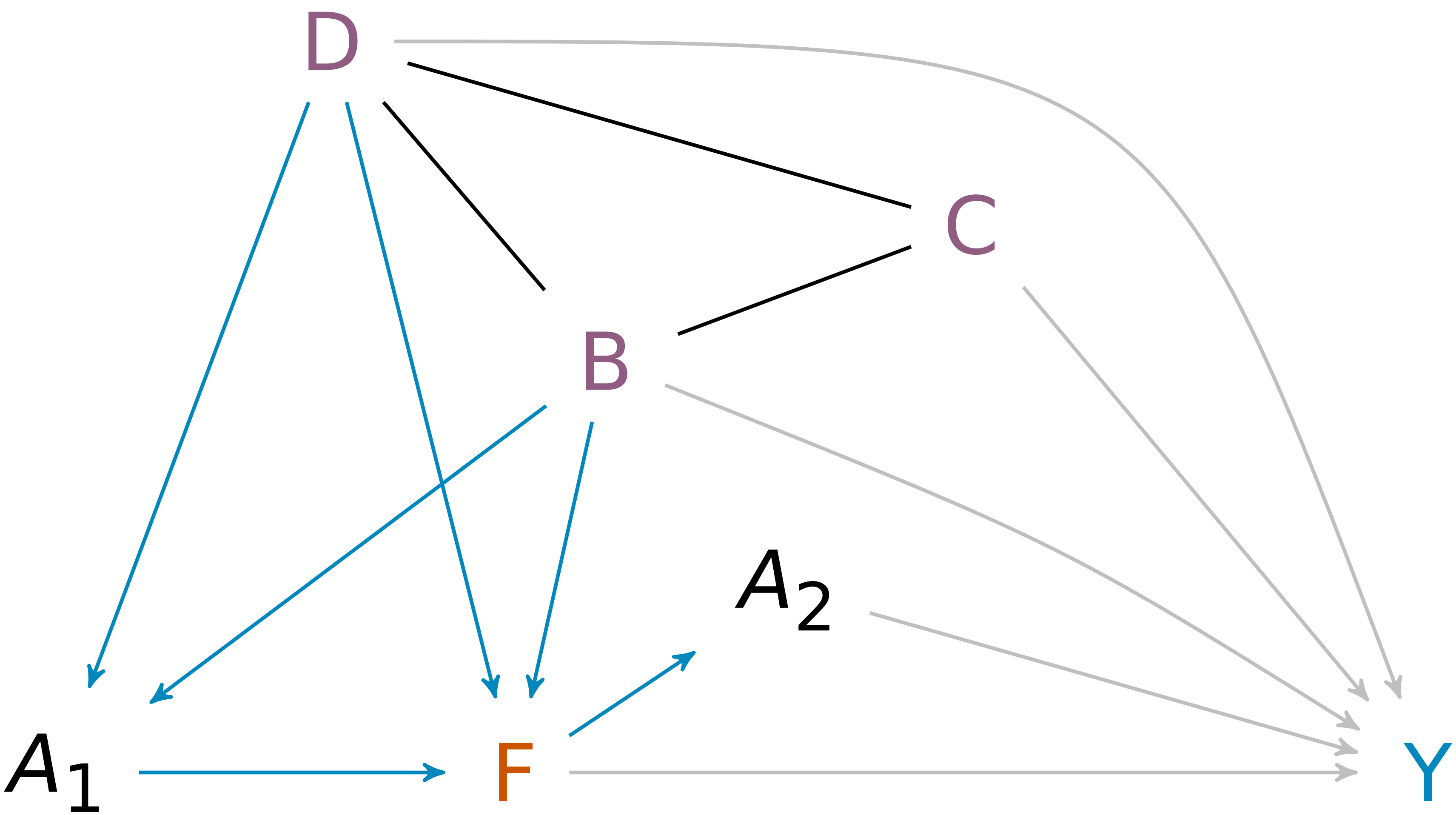
- $\mathbf{S} = an(Y, \mathcal{G}_{\mathbf{V}\setminus\mathbf{A}}) \setminus \{Y\} = \{F, B, C, D\}$, Partition of $\mathbf{S} \cup \{Y\} =$

$f(y|do(a_1, a_2)) = ?$



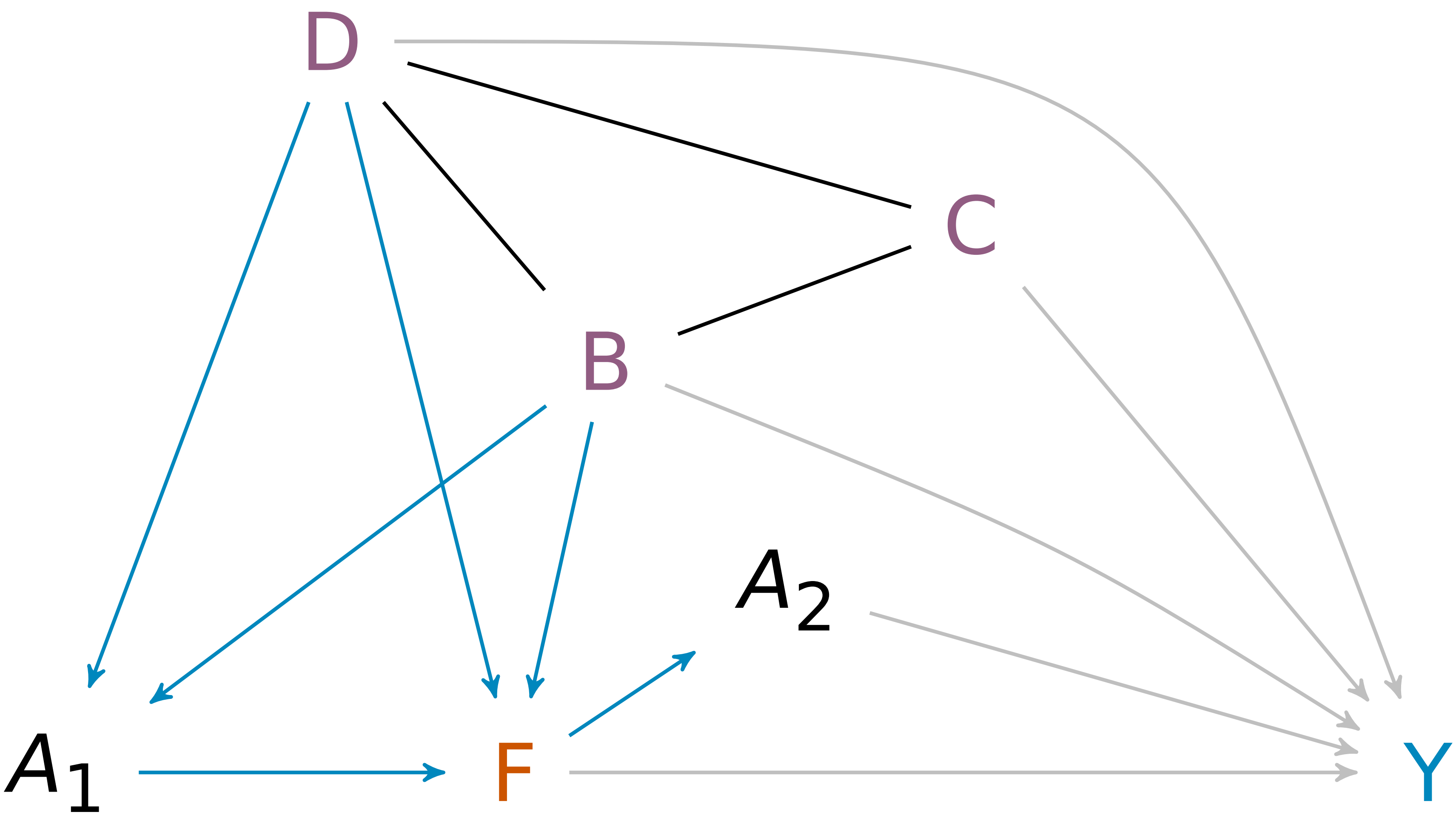
- $\mathbf{S} = an(Y, \mathcal{G}_{\mathbf{V} \setminus \mathbf{A}}) \setminus \{Y\} = \{F, B, C, D\}$, Partition of $\mathbf{S} \cup \{Y\} = (\{B, C, D\}, \{F\}, \{Y\})$.

$f(y|do(a_1, a_2)) = ?$



- $\mathbf{S} = an(Y, \mathcal{G}_{\mathbf{V}\setminus\mathbf{A}}) \setminus \{Y\} = \{B, C, D, F\}$, Partition of $\mathbf{S} \cup \{Y\} = (\{B, C, D\}, \{F\}, \{Y\})$.

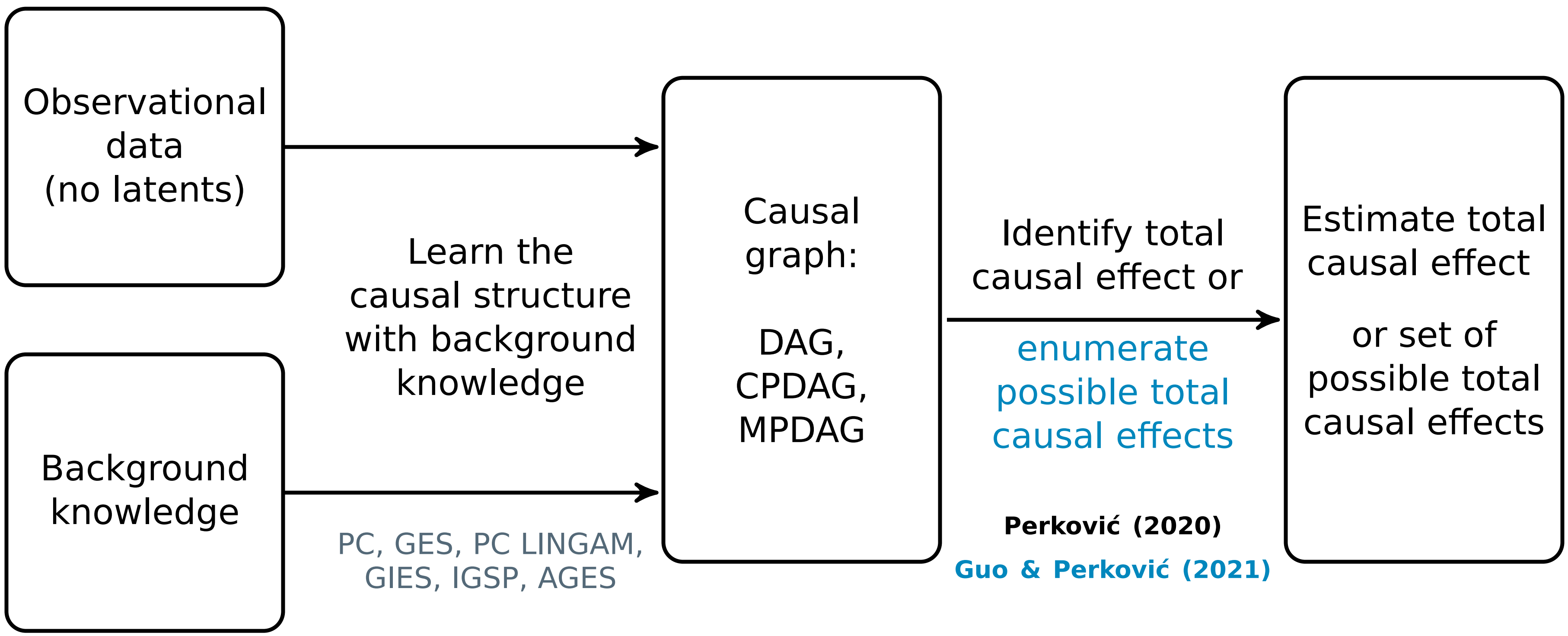
$f(y|do(a_1, a_2)) = ?$



- $\mathbf{S} = an(Y, \mathcal{G}_{\mathbf{V} \setminus \mathbf{A}}) \setminus \{Y\} = \{B, C, D, F\}$, Partition of $\mathbf{S} \cup \{Y\} = (\{B, C, D\}, \{F\}, \{Y\})$.

$$f(y|do(a_1, a_2)) = \int f(y, b, c, d, f, |do(a_1, a_2)) d\mathbf{s} = \int f(y|b, c, d, f, a_2) f(f|b, a_1) f(b, c, d) d\mathbf{s}.$$

Framework



- What if the causal effect is not identifiable?

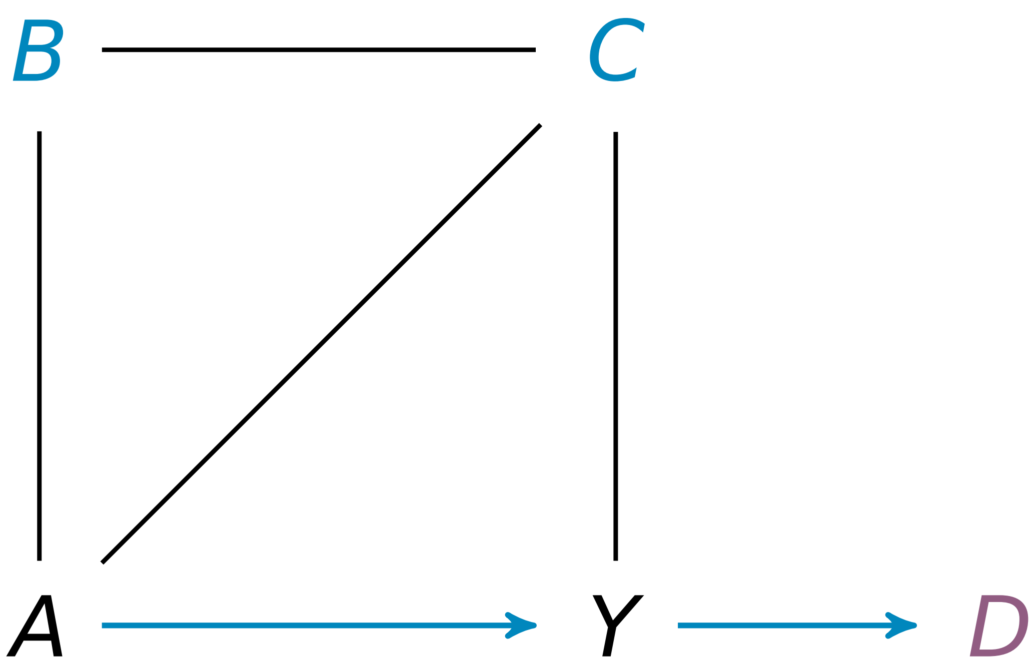
Theorem (Perković, 2020)

The total causal effect of **A** on **Y** is identifiable in MPDAG $\mathcal{G}$ if and only if **all proper possibly causal paths** from **A** to **Y** start with a directed edge in $\mathcal{G}$.

Causal effect is not identifiable

Theorem (Perković, 2020)

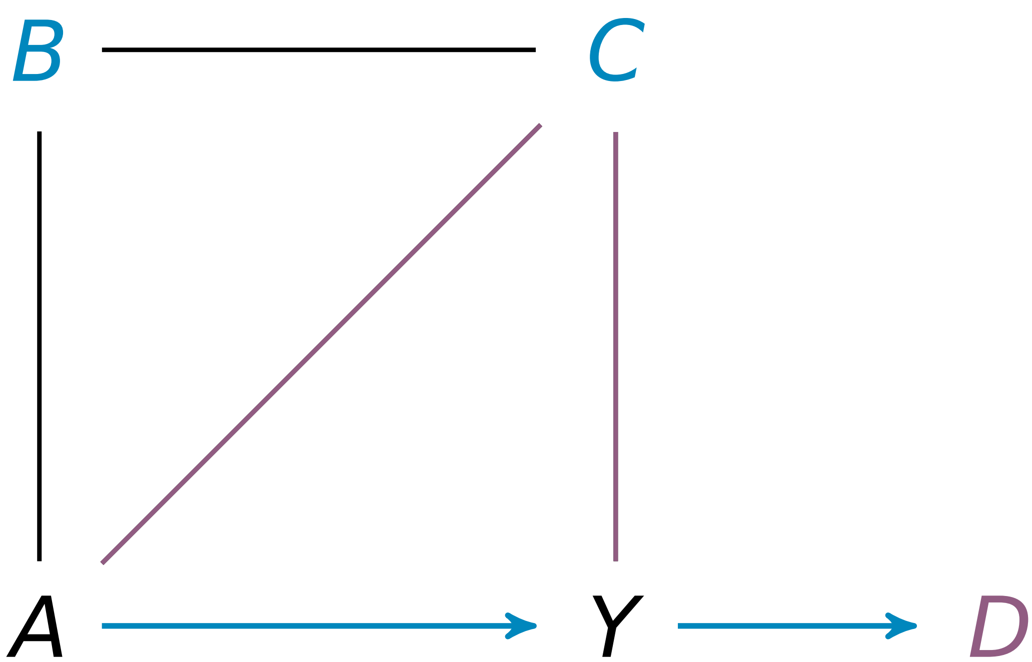
The total causal effect of $\mathbf{A}$ on $\mathbf{Y}$ is identifiable in MPDAG $\mathcal{G}$ if and only if **all proper possibly causal paths** from $\mathbf{A}$ to $\mathbf{Y}$ start with a directed edge in $\mathcal{G}$.



Causal effect is not identifiable

Theorem (Perković, 2020)

The total causal effect of $\mathbf{A}$ on $\mathbf{Y}$ is identifiable in MPDAG $\mathcal{G}$ if and only if **all proper possibly causal paths** from $\mathbf{A}$ to $\mathbf{Y}$ start with a directed edge in $\mathcal{G}$.



- How to enumerate all possible total causal effects?

IDA Enumeration Types

Theorem (Perković, 2020)

The total causal effect of **A** on **Y** is identifiable in MPDAG $\mathcal{G}$ if and only if **all proper possibly causal paths** from **A** to **Y** start with a directed edge in $\mathcal{G}$.

For an MPDAG $\mathcal{G}$, we look for sub-MPDAGs $\mathcal{G}_1, \dots, \mathcal{G}_m$ such that

1. **complete**: $[\mathcal{G}] = [\mathcal{G}_1] \dot{\cup} [\mathcal{G}_2] \dot{\cup} \cdots \dot{\cup} [\mathcal{G}_m]$
2. $f(\mathbf{y}|\mathrm{do}(\mathbf{a}))$ is identifiable under each $\mathcal{G}_i$

IDA Enumeration Types

Theorem (Perković, 2020)

The total causal effect of **A** on **Y** is identifiable in MPDAG $\mathcal{G}$ if and only if **all proper possibly causal paths** from **A** to **Y** start with a directed edge in $\mathcal{G}$.

For an MPDAG $\mathcal{G}$, we look for sub-MPDAGs $\mathcal{G}_1, \dots, \mathcal{G}_m$ such that

1. **complete**: $[\mathcal{G}] = [\mathcal{G}_1]\dot{\cup}[\mathcal{G}_2]\dot{\cup}\cdots\dot{\cup}[\mathcal{G}_m]$
2. $f(\mathbf{y}|do(\mathbf{a}))$ is identifiable under each $\mathcal{G}_i$

   We could enumerate over

   - all DAGs (Maathuis et al, '09)
   - the valid parent sets of A (Maathuis et al, '09, Nandy et al, '17, Perković et al, '17, Witte et al, '20, Fang and He, '20)
   - orientation of $A-$ on possibly causal paths to Y (Liu et al, '20)

IDA Enumeration Types

Theorem (Perković, 2020)

The total causal effect of **A** on **Y** is identifiable in MPDAG $\mathcal{G}$ if and only if **all proper possibly causal paths** from **A** to **Y** start with a directed edge in $\mathcal{G}$.

For an MPDAG $\mathcal{G}$, we look for sub-MPDAGs $\mathcal{G}_1, \ldots, \mathcal{G}_m$ such that

1. **complete**: $[\mathcal{G}] = [\mathcal{G}_1]\dot{\cup}[\mathcal{G}_2]\dot{\cup}\cdots\dot{\cup}[\mathcal{G}_m]$
2. $f(\mathbf{y}|\mathrm{do}(\mathbf{a}))$ is identifiable under each $\mathcal{G}_i$

   We could enumerate over

   - all DAGs (Maathuis et al, '09)
   - the valid parent sets of A (Maathuis et al, '09, Nandy et al, '17, Perković et al, '17, Witte et al, '20, Fang and He, '20)
   - orientation of $A-$ on possibly causal paths to Y (Liu et al, '20)
3. **minimal**: maps $f \mapsto f(\mathbf{y}|do(\mathbf{a}))$ are distinct under each $\mathcal{G}_i \Rightarrow$ possible causal effects $f \mapsto \frac{\partial}{\partial a_i}\mathbb{E}(\mathbf{Y}|\mathrm{do}(\mathbf{A}) = \mathbf{a})$ are distinct functionals!

   - None of the above approaches are minimal!

Optimal enumeration

Theorem (Perković, 2020)

The total causal effect of $\mathbf{A}$ on $\mathbf{Y}$ is identifiable in MPDAG $\mathcal{G}$ if and only if **all proper possibly causal paths** from $\mathbf{A}$ to $\mathbf{Y}$ start with a directed edge in $\mathcal{G}$.

Input: MPDAG $\mathcal{G}$, $\mathbf{Y} \subset \mathbf{V}$ and $\mathbf{A} \subset \mathbf{V} \setminus \mathbf{Y}$.

Algorithm FirstTry

Optimal enumeration

Theorem (Perković, 2020)

The total causal effect of $\mathbf{A}$ on $\mathbf{Y}$ is identifiable in MPDAG $\mathcal{G}$ if and only if **all proper possibly causal paths** from $\mathbf{A}$ to $\mathbf{Y}$ start with a directed edge in $\mathcal{G}$.

Input: MPDAG $\mathcal{G}$, $\mathbf{Y} \subset \mathbf{V}$ and $\mathbf{A} \subset \mathbf{V} \setminus \mathbf{Y}$.

Algorithm FirstTry

1. Pick $A_1 - V_1$ such that $A_1 \in \mathbf{A}$ and there is a proper possibly causal path $A_1, V_1, \dots, Y_1$, $Y_1 \in \mathbf{Y}$.

Optimal enumeration

Theorem (Perković, 2020)

The total causal effect of $\mathbf{A}$ on $\mathbf{Y}$ is identifiable in MPDAG $\mathcal{G}$ if and only if **all proper possibly causal paths** from $\mathbf{A}$ to $\mathbf{Y}$ start with a directed edge in $\mathcal{G}$.

Input: MPDAG $\mathcal{G}$, $\mathbf{Y} \subset \mathbf{V}$ and $\mathbf{A} \subset \mathbf{V} \setminus \mathbf{Y}$.

Algorithm FirstTry

1. Pick $A_1 - V_1$ such that $A_1 \in \mathbf{A}$ and there is a proper possibly causal path $A_1, V_1, \dots, Y_1$, $Y_1 \in \mathbf{Y}$.
2. $\mathcal{G}_1 \leftarrow$ MPDAG($\mathcal{G}, A_1 \rightarrow V_1$), $\mathcal{G}_2 \leftarrow$ MPDAG($\mathcal{G}, A_1 \leftarrow V_1$)

Optimal enumeration

Theorem (Perković, 2020)

The total causal effect of $\mathbf{A}$ on $\mathbf{Y}$ is identifiable in MPDAG $\mathcal{G}$ if and only if **all proper possibly causal paths** from $\mathbf{A}$ to $\mathbf{Y}$ start with a directed edge in $\mathcal{G}$.

Input: MPDAG $\mathcal{G}$, $\mathbf{Y} \subset \mathbf{V}$ and $\mathbf{A} \subset \mathbf{V} \setminus \mathbf{Y}$.

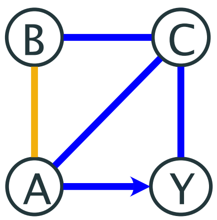
Algorithm FirstTry

1. Pick $A_1 - V_1$ such that $A_1 \in \mathbf{A}$ and there is a proper possibly causal path $A_1, V_1, \dots, Y_1$, $Y_1 \in \mathbf{Y}$.
2. $\mathcal{G}_1 \leftarrow \text{MPDAG}(\mathcal{G}, A_1 \rightarrow V_1)$, $\mathcal{G}_2 \leftarrow \text{MPDAG}(\mathcal{G}, A_1 \leftarrow V_1)$
3. Recurse on $\mathcal{G}_1$ and $\mathcal{G}_2$ until $f(\mathbf{y}|\text{do}(\mathbf{a}))$ is identified

MPDAG$(\mathcal{G}, R)$ adds orientations R to $\mathcal{G}$ and completes Meek orientation rules.

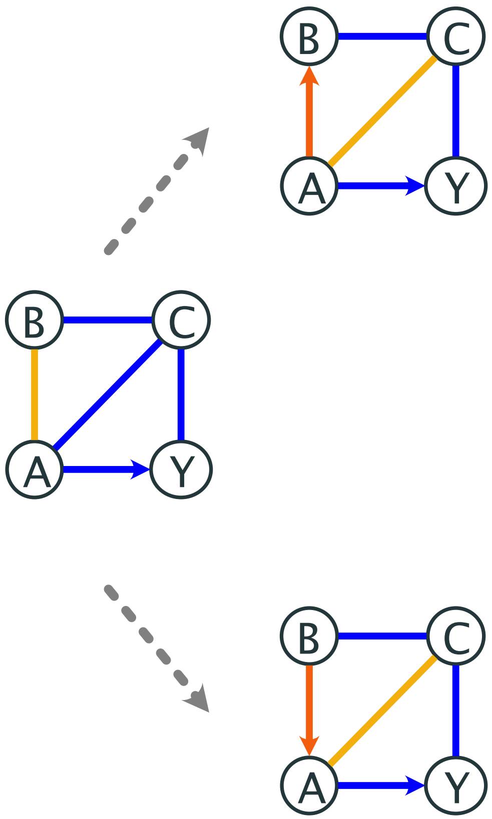
Optimal enumeration

Orienting $A - B$ first ...



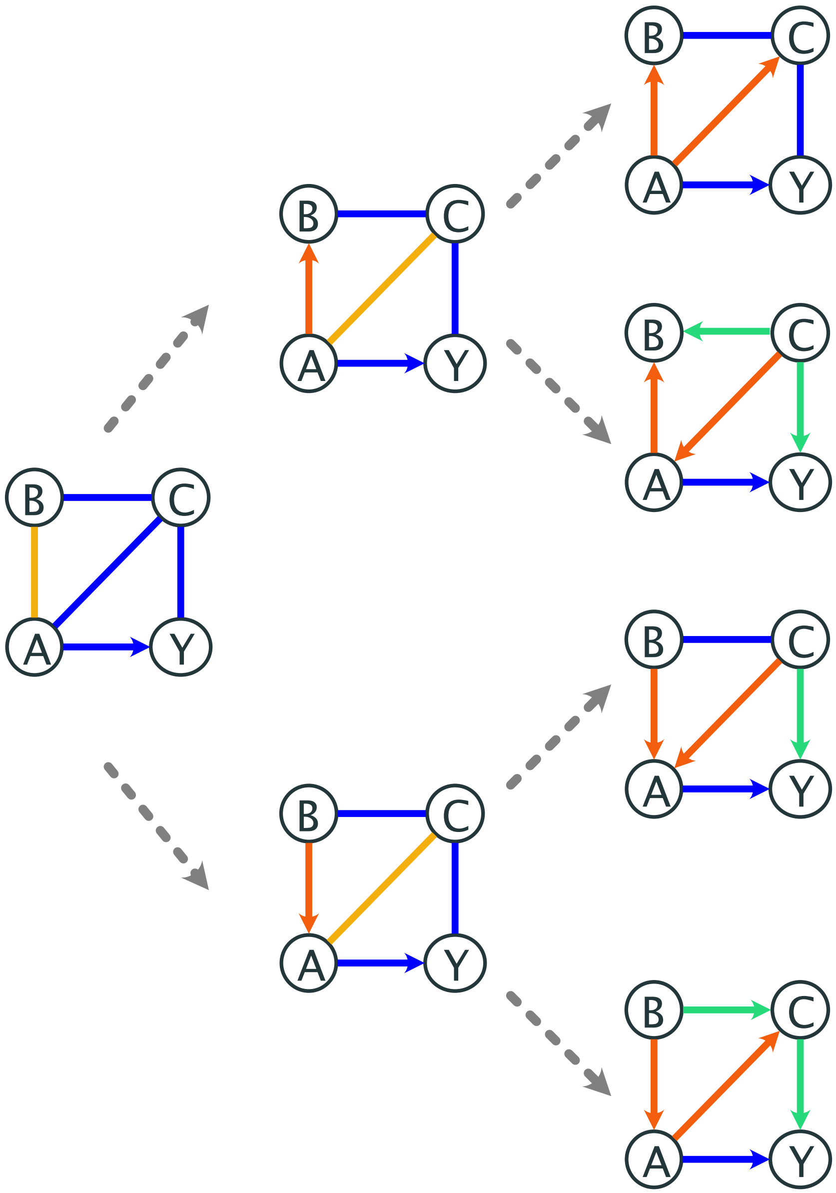
Optimal enumeration

Orienting $A - B$ first ...



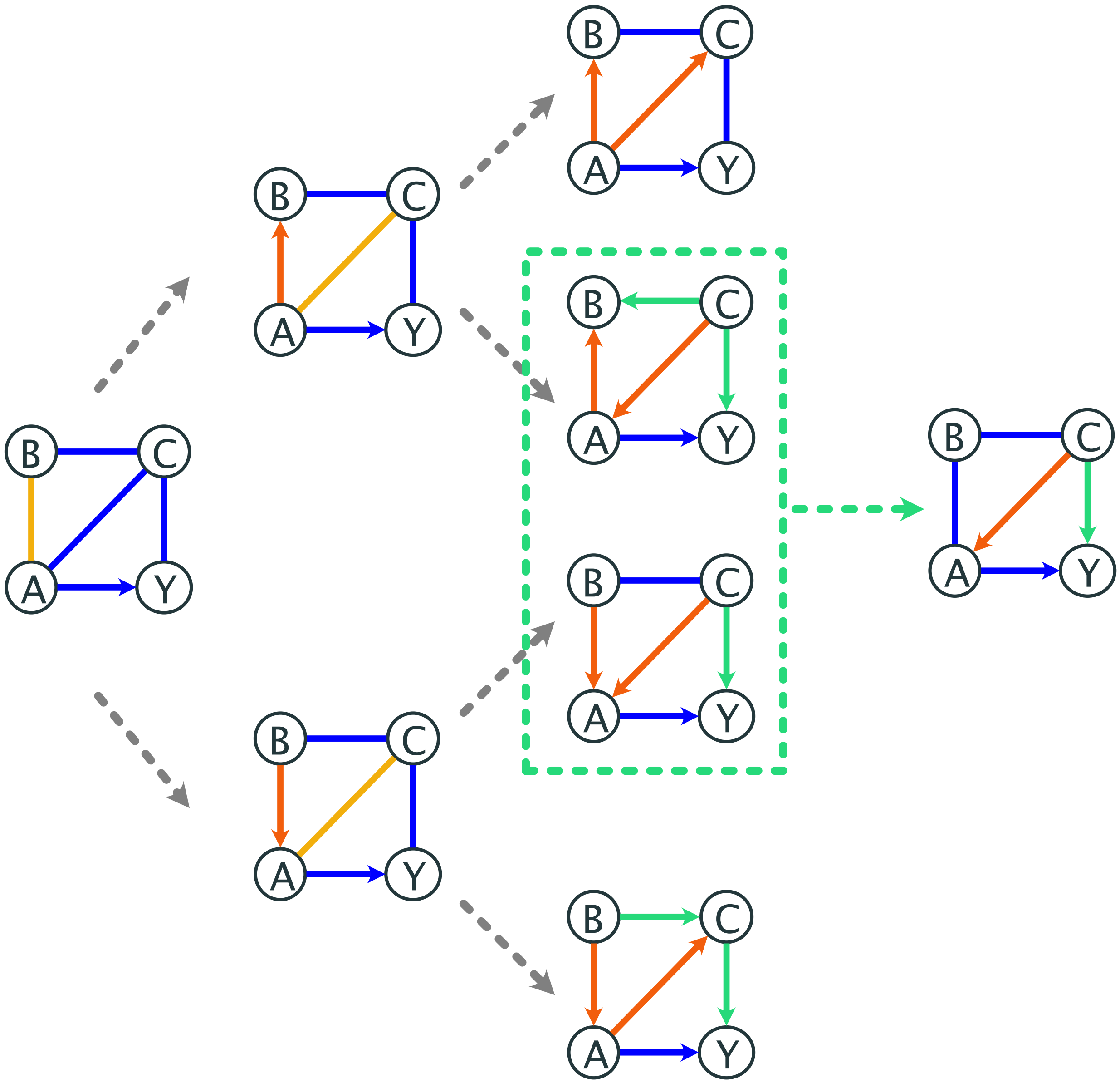
Optimal enumeration

Orienting $A - B$ first ...



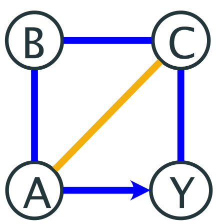
Optimal enumeration

Orienting $A - B$ first ...



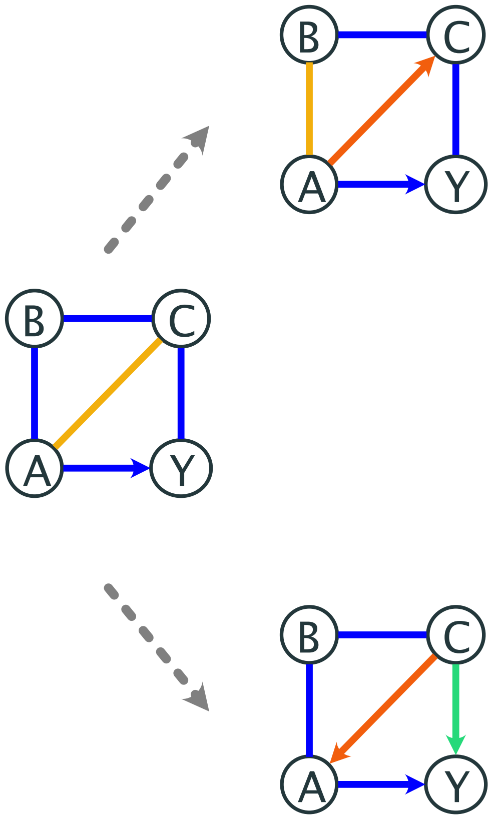
Optimal enumeration

Orienting $A - C$ first ...



Optimal enumeration

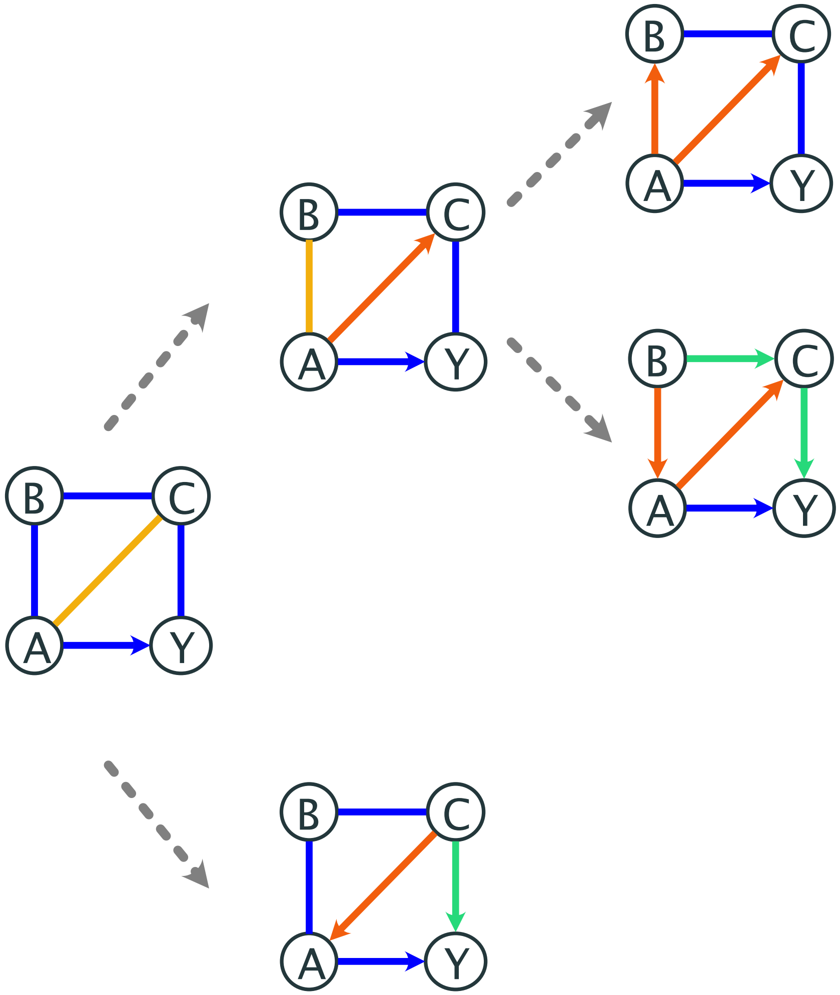
Orienting $A - C$ first ...



Optimal enumeration

Orienting $A - C$ first ...

- $A - C$ should be oriented first because the *status* of $A - B - C - Y$ depends on $A - C - Y$.



Optimal enumeration

Algorithm IDGraphs, (Guo & Perković, 2021)

1. Pick $A_1 - V_1$ such that $A_1 \in \mathbf{A}$ and $A_1, V_1, \ldots, Y_1$, $Y_1 \in \mathbf{Y}$ is a shortest proper possibly causal path from $\mathbf{A}$ to $\mathbf{Y}$.
2. $\mathcal{G}_1 \leftarrow$ MPDAG$(\mathcal{G}, A_1 \to V_1)$, $\mathcal{G}_2 \leftarrow$ MPDAG$(\mathcal{G}, A_1 \leftarrow V_1)$
3. Recurse on $\mathcal{G}_1$ and $\mathcal{G}_2$ until identified

Optimal enumeration

Algorithm IDGraphs, (Guo & Perković, 2021)

1. Pick $A_1 - V_1$ such that $A_1 \in \mathbf{A}$ and $A_1, V_1, \ldots, Y_1$, $Y_1 \in \mathbf{Y}$ is a shortest proper possibly causal path from **A** to **Y**.
2. $\mathcal{G}_1 \leftarrow$ MPDAG$(\mathcal{G}, A_1 \rightarrow V_1)$, $\mathcal{G}_2 \leftarrow$ MPDAG$(\mathcal{G}, A_1 \leftarrow V_1)$
3. Recurse on $\mathcal{G}_1$ and $\mathcal{G}_2$ until identified

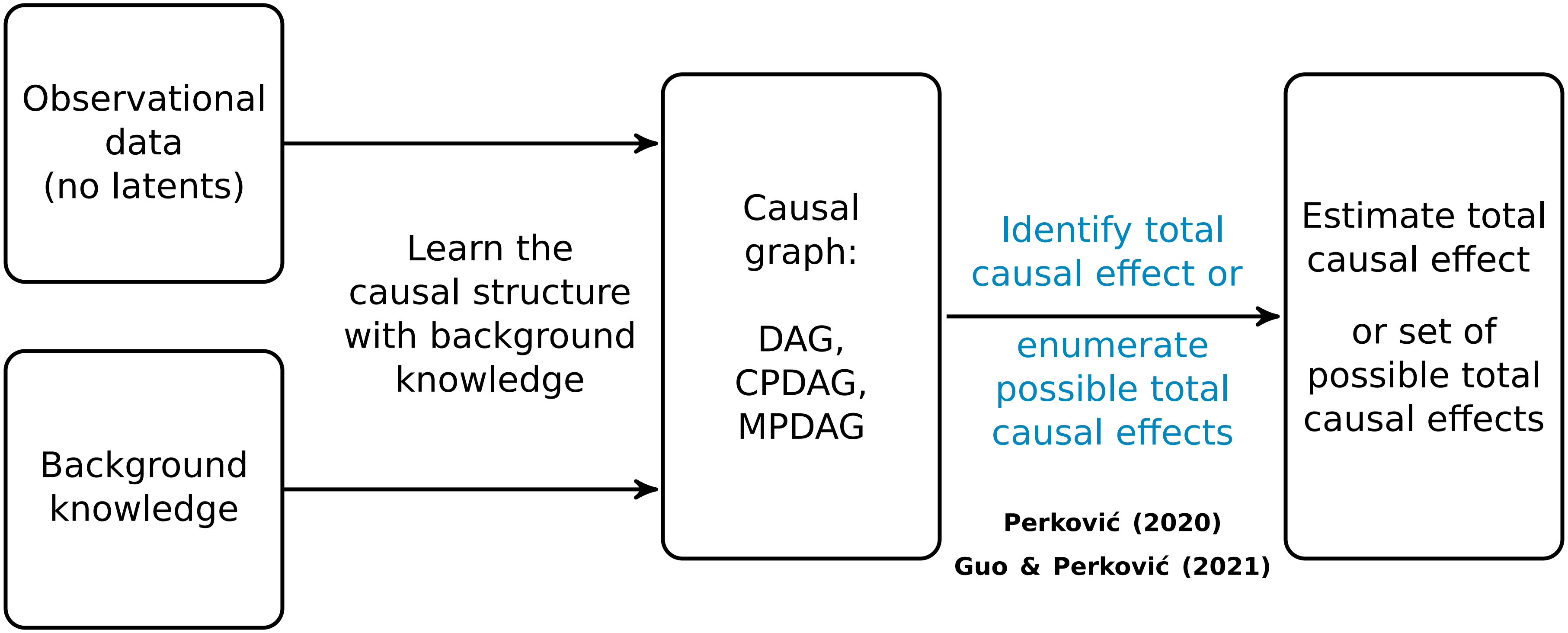
Theorem (Guo & Perković, 2021)

$(\mathcal{G}_1, \ldots, \mathcal{G}_m)$ output by the algorithm is **complete** and **minimal**.

Hence, each $\mathcal{G}_i$ represents the minimal set of additional orientations required for a particular interventional distribution/possible effect!

In contrast, the existing algorithms will output 4 effects for this example, but two of them are different estimates of the same possible effect!

Framework



- **R package** `eff2`: github.com/richardkwo/eff2

Thanks!