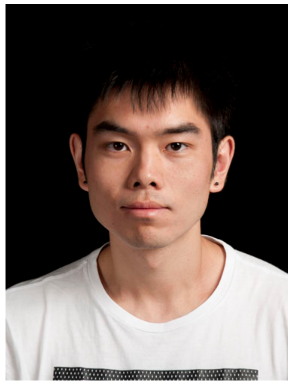
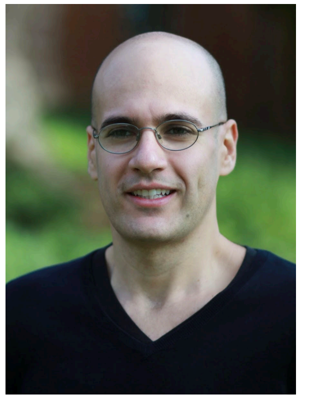
Partial Counterfactual Identification from Observational and Experimental Data

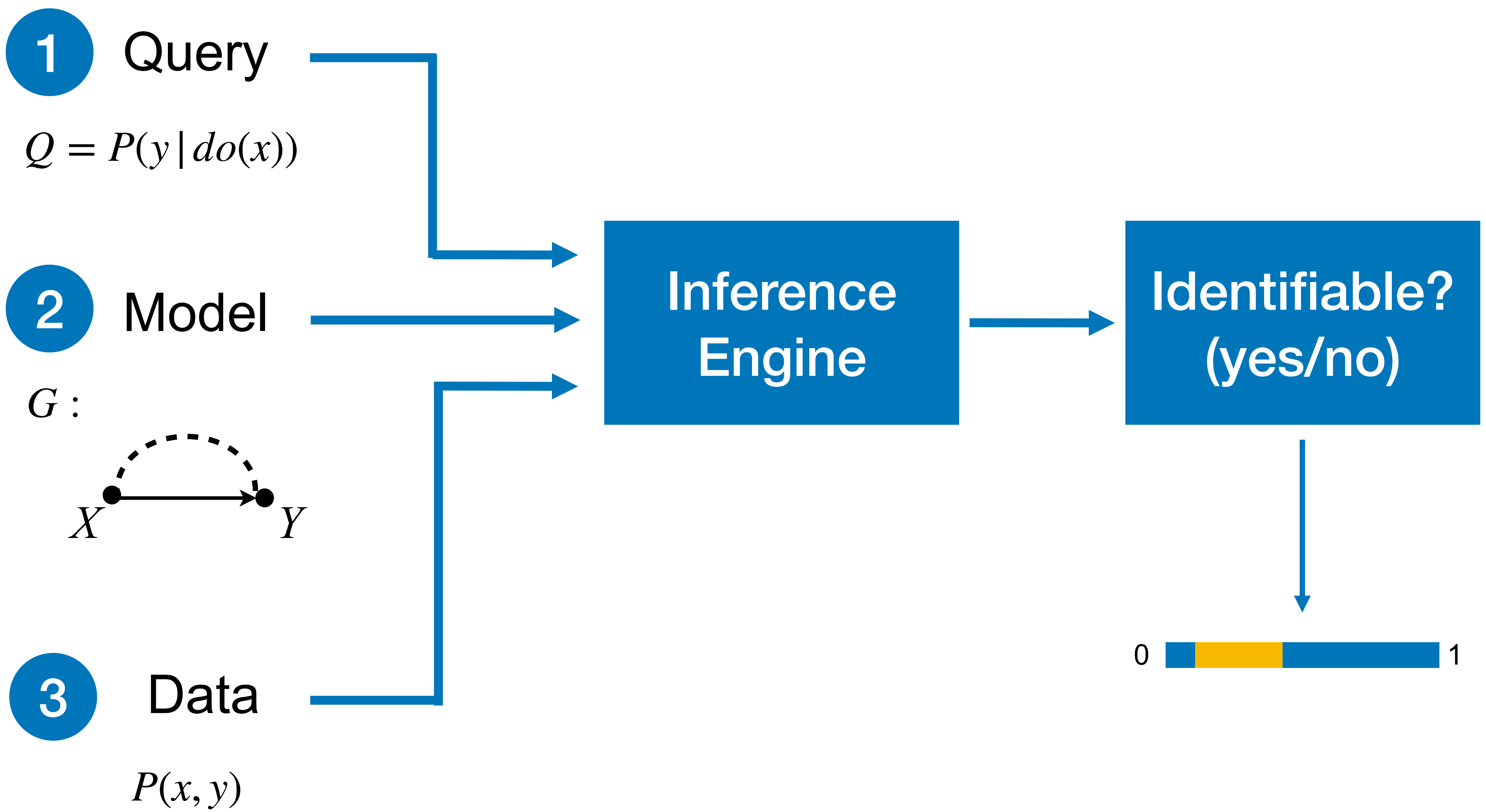




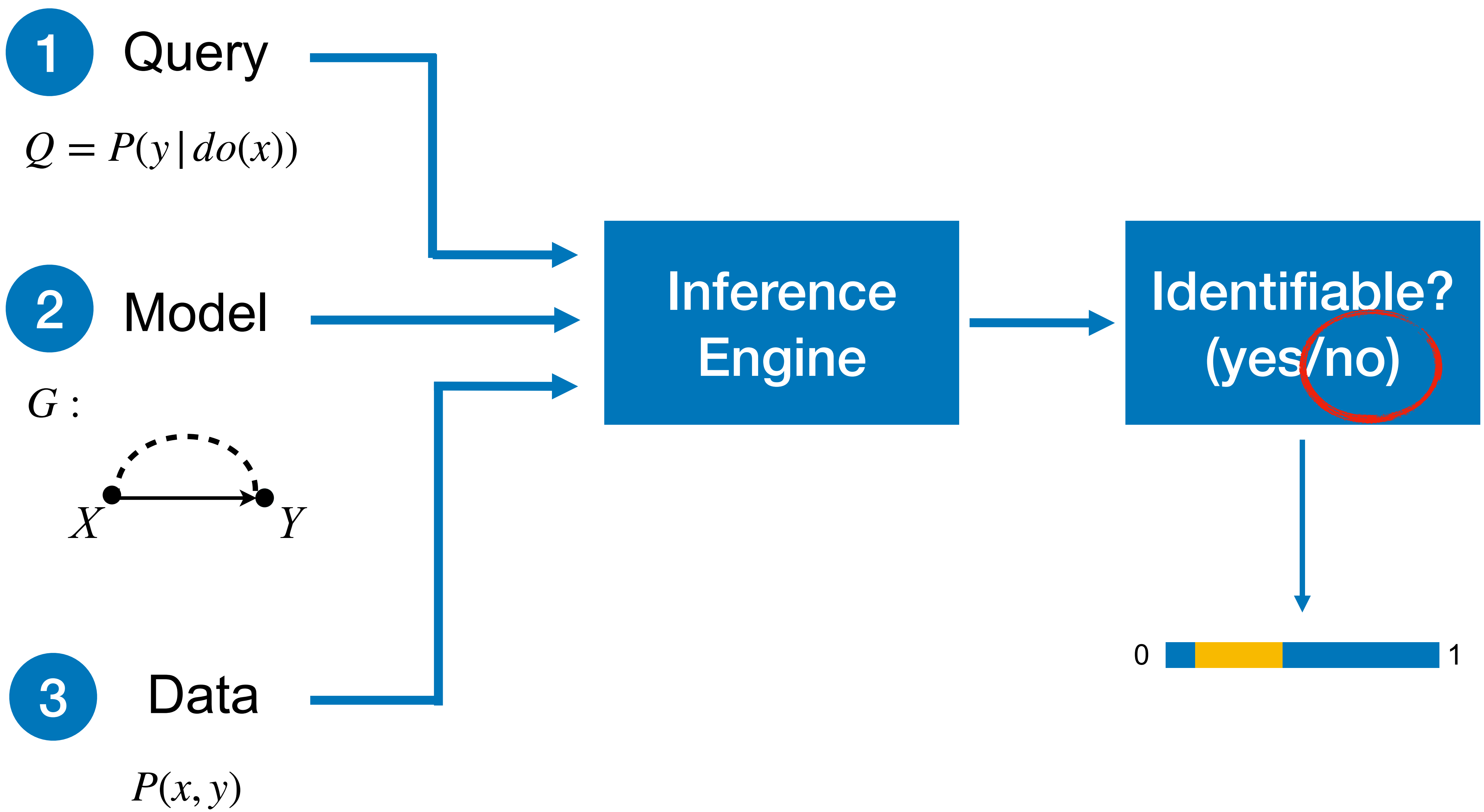
Junzhe Zhang[1], Jin Tian[2], Elias Bareinboim[1]
[1]Columbia University
[2]Iowa State University

Thirty-ninth International Conference on Machine Learning, 2022

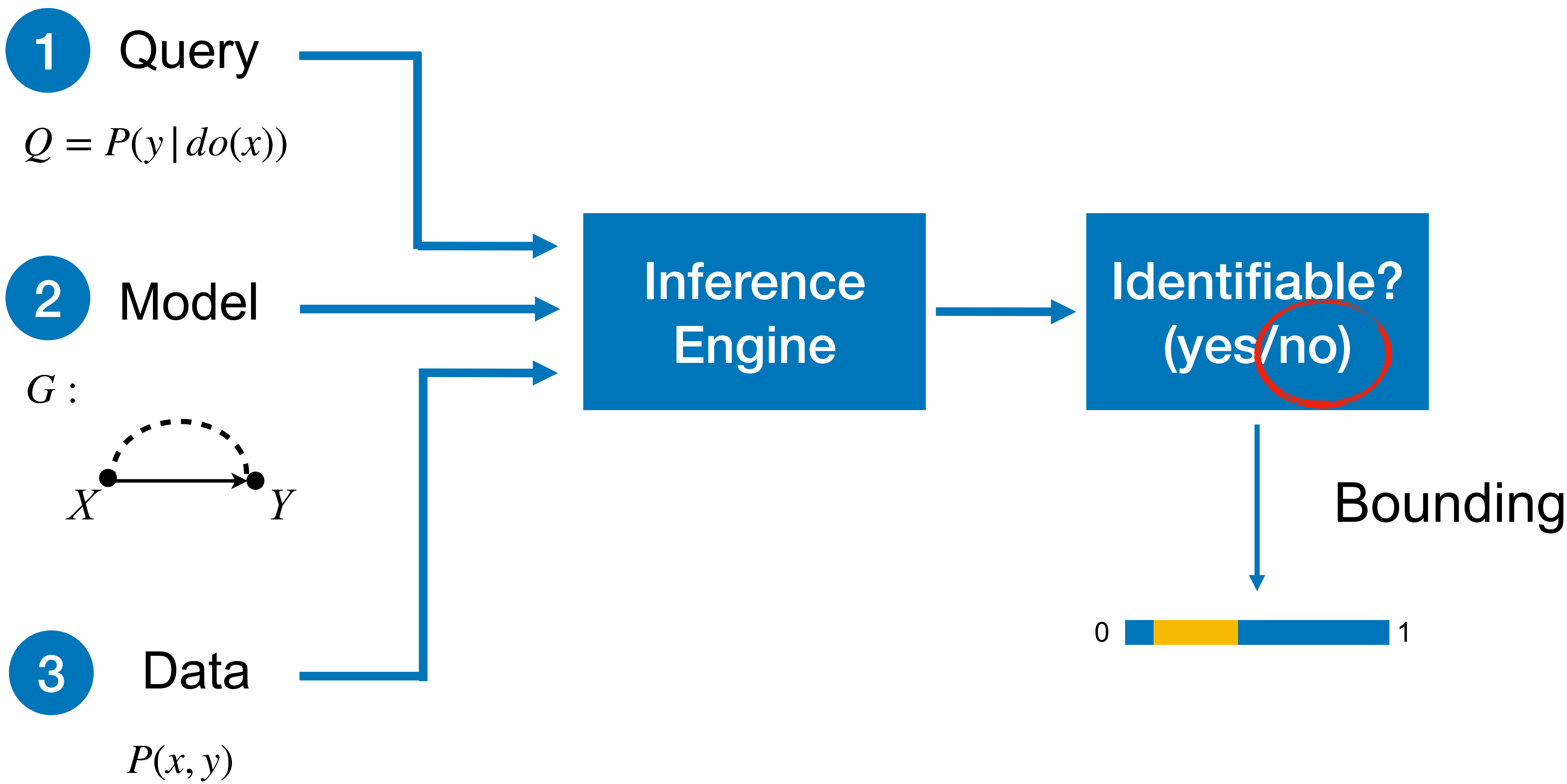
The Partial Identification Problem



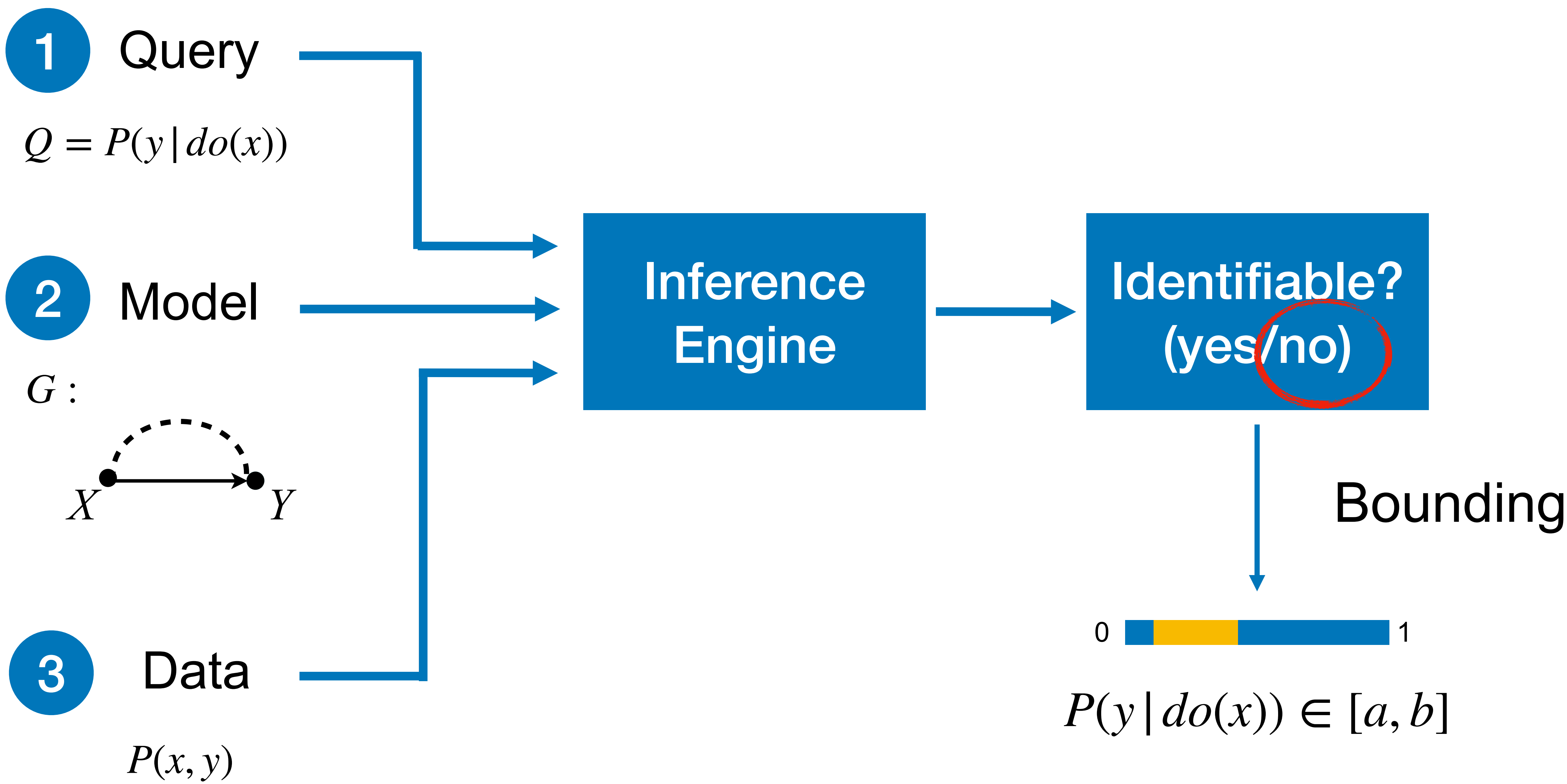
The Partial Identification Problem



The Partial Identification Problem

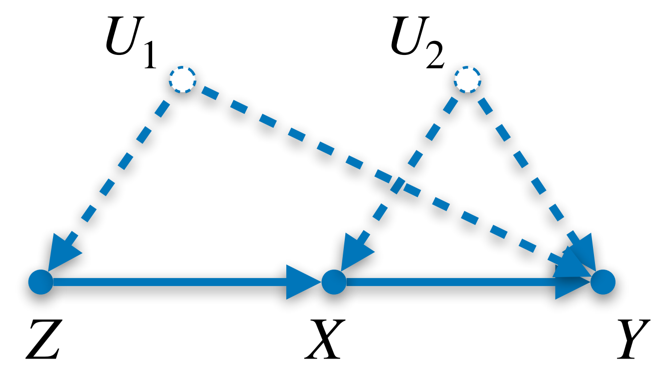


The Partial Identification Problem



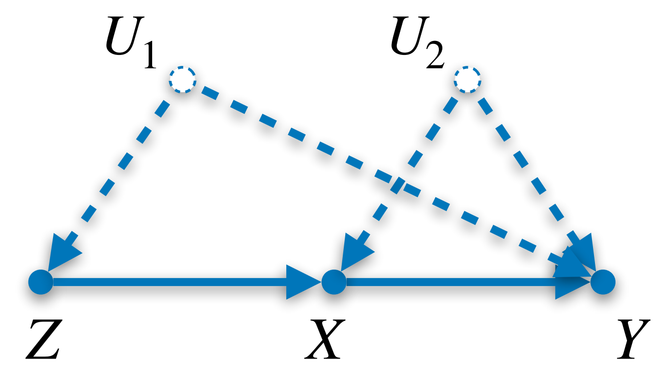
Partial Identification of Causal Effects

Task. Given the observational distribution $P(\mathbf{v})$ in an arbitrary causal diagram G, bound $P(\mathbf{y} \mid do(\mathbf{x}))$ for any $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$.



Partial Identification of Causal Effects

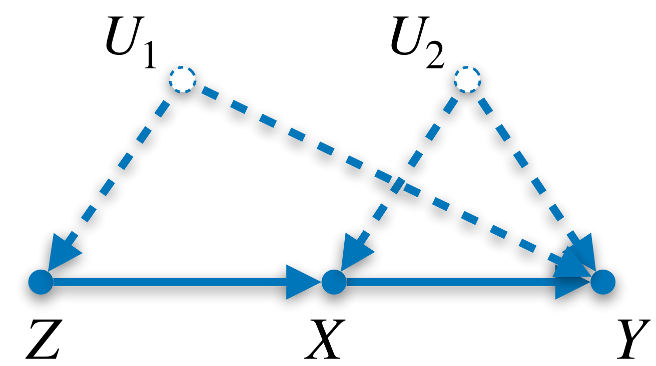
Task. Given the observational distribution $P(\mathbf{v})$ in an arbitrary causal diagram G, bound $P(\mathbf{y} \mid do(\mathbf{x}))$ for any $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$.



- We assume that the domain of $\mathbf{V}$ is discrete and finite.

Partial Identification of Causal Effects

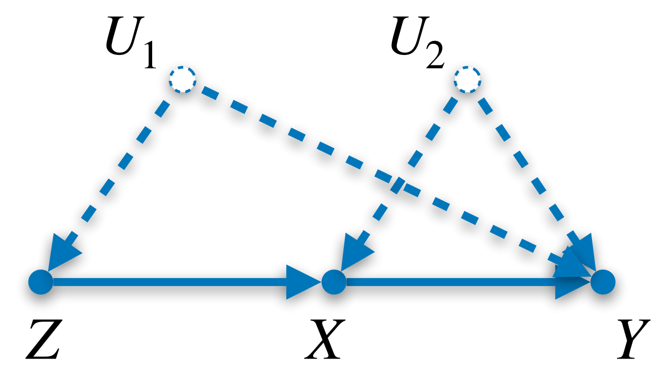
Task. Given the observational distribution $P(\mathbf{v})$ in an arbitrary causal diagram G, bound $P(\mathbf{y} \mid do(\mathbf{x}))$ for any $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$.



- We assume that the domain of $\mathbf{V}$ is discrete and finite.
- Let $\mathscr{M}$ denote the set of all possible SCMs compatible with G.

Partial Identification of Causal Effects

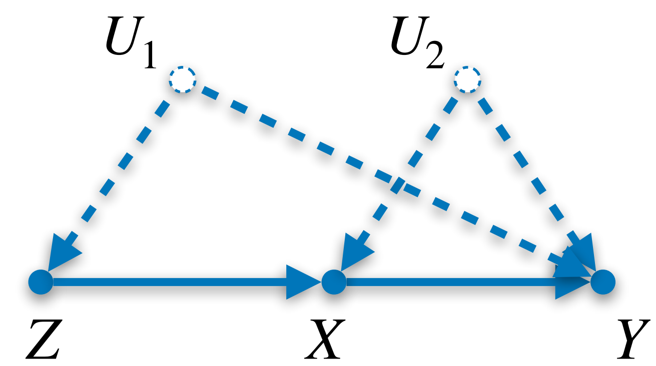
Task. Given the observational distribution $P(\mathbf{v})$ in an arbitrary causal diagram G, bound $P(\mathbf{y} \mid do(\mathbf{x}))$ for any $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$.



- We assume that the domain of $\mathbf{V}$ is discrete and finite.
- Let $\mathscr{M}$ denote the set of all possible SCMs compatible with G.
- Given $P(\mathbf{v})$, $P(\mathbf{y} \mid do(\mathbf{x}))$ is bounded in $[a, b]$ where:

Partial Identification of Causal Effects

Task. Given the observational distribution $P(\mathbf{v})$ in an arbitrary causal diagram G, bound $P(\mathbf{y} \mid do(\mathbf{x}))$ for any $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$.

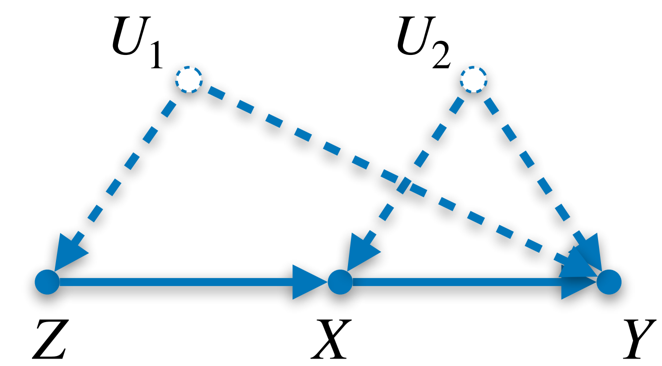


- We assume that the domain of $\mathbf{V}$ is discrete and finite.
- Let $\mathscr{M}$ denote the set of all possible SCMs compatible with G.
- Given $P(\mathbf{v})$, $P(\mathbf{y} \mid do(\mathbf{x}))$ is bounded in $[a, b]$ where:

$$a = \min P_M(\mathbf{y} \mid do(\mathbf{x})), \qquad \forall M \in \mathscr{M},$$
$$b = \max P_M(\mathbf{y} \mid do(\mathbf{x})). \quad \textbf{s.t.} \quad P_M(\mathbf{v}) = P(\mathbf{v}).$$

Partial Identification of Causal Effects

Task. Given the observational distribution $P(\mathbf{v})$ in an arbitrary causal diagram G, bound $P(\mathbf{y} \mid do(\mathbf{x}))$ for any $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$.

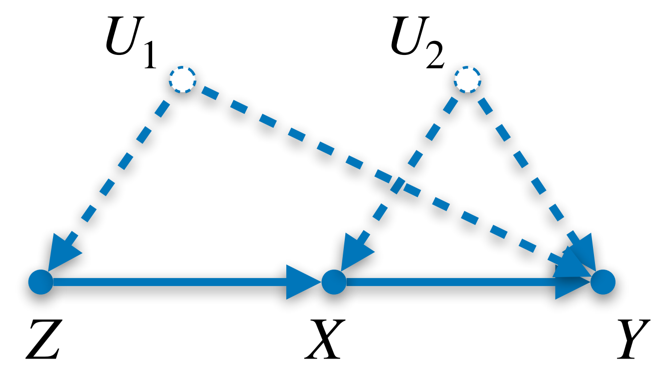


- We assume that the domain of $\mathbf{V}$ is discrete and finite.
- Let $\mathscr{M}$ denote the set of all possible SCMs compatible with G.
- Given $P(\mathbf{v})$, $P(\mathbf{y} \mid do(\mathbf{x}))$ is bounded in $[a, b]$ where:

$$a = \min P_M(\mathbf{y} \mid do(\mathbf{x})), \quad b = \max P_M(\mathbf{y} \mid do(\mathbf{x}))\,. \quad \text{s.t.} \quad \forall M \in \mathscr{M}, \quad P_M(\mathbf{v}) = P(\mathbf{v})\,.$$

Partial Identification of Causal Effects

Task. Given the observational distribution $P(\mathbf{v})$ in an arbitrary causal diagram G, bound $P(\mathbf{y} \mid do(\mathbf{x}))$ for any $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$.



- We assume that the domain of $\mathbf{V}$ is discrete and finite.
- Let $\mathscr{M}$ denote the set of all possible SCMs compatible with G.
- Given $P(\mathbf{v})$, $P(\mathbf{y} \mid do(\mathbf{x}))$ is bounded in $[a, b]$ where:

$$a = \min P_M(\mathbf{y} \mid do(\mathbf{x})), \quad \text{s.t.} \quad \forall M \in \mathscr{M},$$
$$b = \max P_M(\mathbf{y} \mid do(\mathbf{x})). \quad \quad P_M(\mathbf{v}) = P(\mathbf{v}).$$

Solving this optimization is difficult since parametric form of $\mathscr{F}, P(U)$ are not provided.

Canonical Causal Models

Canonical Causal Models

Definition. A canonical SCM is a SCM $M = \langle \mathbf{V}, \mathbf{U}, \mathscr{F}, P(\mathbf{U}) \rangle$ where

Canonical Causal Models

Definition. A canonical SCM is a SCM $M = \langle \mathbf{V}, \mathbf{U}, \mathscr{F}, P(\mathbf{U}) \rangle$ where

- Every $V \in \mathbf{V}$ is decided by a function $v \leftarrow f_V(\mathrm{pa}_V, u_V)$ taking values in a discrete and finite domain Ω_V.

Canonical Causal Models

Definition. A canonical SCM is a SCM $M = \langle \mathbf{V}, \mathbf{U}, \mathscr{F}, P(\mathbf{U}) \rangle$ where

- Every $V \in \mathbf{V}$ is decided by a function $v \leftarrow f_V(\mathrm{pa}_V, u_V)$ taking values in a discrete and finite domain Ω_V.

- Every $U \in \mathbf{U}$ are drawn from a discrete domain Ω_U with cardinality

Canonical Causal Models

Definition. A canonical SCM is a SCM $M = \langle \mathbf{V}, \mathbf{U}, \mathscr{F}, P(\mathbf{U}) \rangle$ where

- Every $V \in \mathbf{V}$ is decided by a function $v \leftarrow f_V(\mathrm{pa}_V, u_V)$ taking values in a discrete and finite domain Ω_V.
- Every $U \in \mathbf{U}$ are drawn from a discrete domain Ω_U with cardinality

$$|\Omega_U| = \prod_{V \in \mathbf{C}(U)} |\Omega_{\mathrm{Pa}_V}| \times |\Omega_V|$$

Canonical Causal Models

Definition. A canonical SCM is a SCM $M = \langle \mathbf{V}, \mathbf{U}, \mathscr{F}, P(\mathbf{U}) \rangle$ where

- Every $V \in \mathbf{V}$ is decided by a function $v \leftarrow f_V(\mathrm{pa}_V, u_V)$ taking values in a discrete and finite domain Ω_V.

- Every $U \in \mathbf{U}$ are drawn from a discrete domain Ω_U with cardinality

$$|\Omega_U| = \prod_{V \in \mathbf{C}(U)} |\Omega_{\mathrm{Pa}_V}| \times |\Omega_V|$$

where $\mathbf{C}(U)$ is the c-component in G that covers U.

Canonical Causal Models

Definition. A canonical SCM is a SCM $M = \langle \mathbf{V}, \mathbf{U}, \mathscr{F}, P(\mathbf{U}) \rangle$ where

- Every $V \in \mathbf{V}$ is decided by a function $v \leftarrow f_V(\mathrm{pa}_V, u_V)$ taking values in a discrete and finite domain Ω_V.

- Every $U \in \mathbf{U}$ are drawn from a discrete domain Ω_U with cardinality

$$|\Omega_U| = \prod_{V \in \mathbf{C}(U)} |\Omega_{\mathrm{Pa}_V}| \times |\Omega_V|$$

where $\mathbf{C}(U)$ is the c-component in G that covers U.

Two endogenous variables are in the same c-component if and only if they are connected by a bi-directed path.

Canonical SCMs

Canonical SCMs

Theorem. For any SCM M, there exists a canonical SCM N s.t.

Canonical SCMs

Theorem. For any SCM M, there exists a canonical SCM N s.t.

1. M and N are compatible with the same causal diagram G;

Canonical SCMs

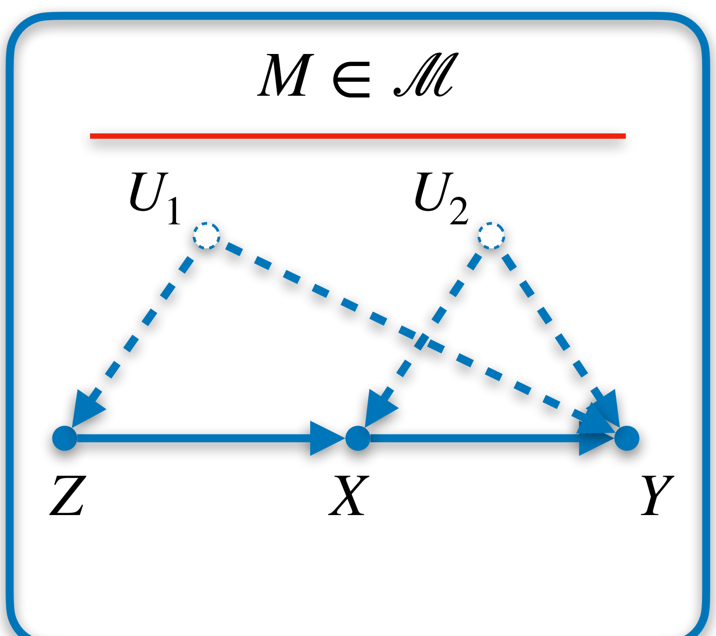
Theorem. For any SCM M, there exists a canonical SCM N s.t.

1. M and N are compatible with the same causal diagram G;
2. For any subsets $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$, $P_M(\mathbf{y} \mid do(\mathbf{x})) = P_N(\mathbf{y} \mid do(\mathbf{x}))$.

Canonical SCMs

Theorem. For any SCM M, there exists a canonical SCM N s.t.

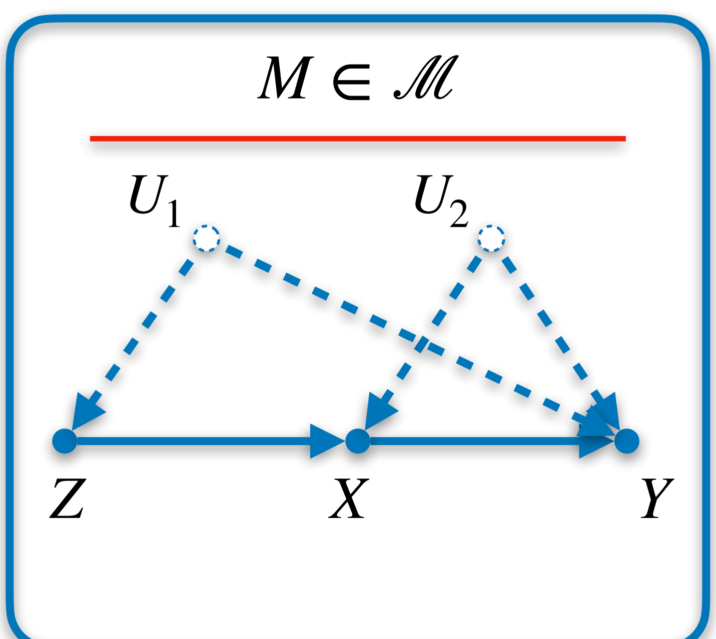
1. M and N are compatible with the same causal diagram G;
2. For any subsets $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$, $P_M(\mathbf{y} \mid do(\mathbf{x})) = P_N(\mathbf{y} \mid do(\mathbf{x}))$.

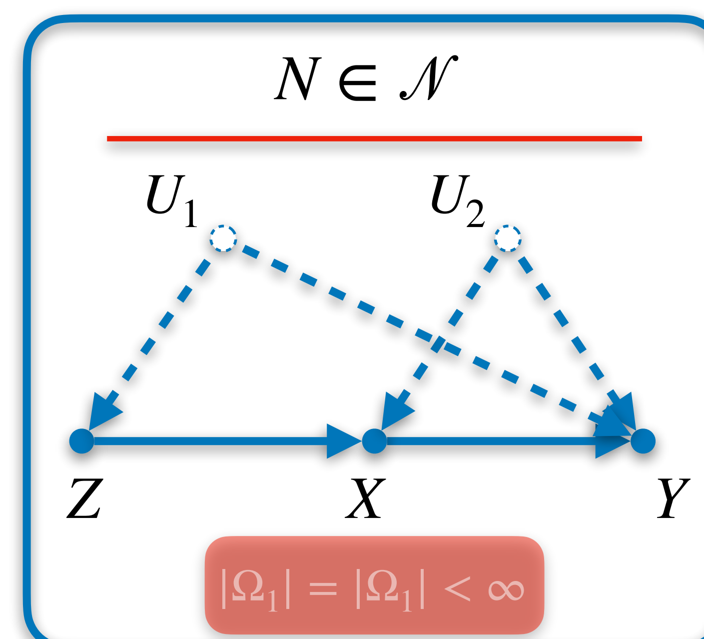


Canonical SCMs

Theorem. For any SCM M, there exists a canonical SCM N s.t.

1. M and N are compatible with the same causal diagram G;
2. For any subsets $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$, $P_M(\mathbf{y} \mid do(\mathbf{x})) = P_N(\mathbf{y} \mid do(\mathbf{x}))$.

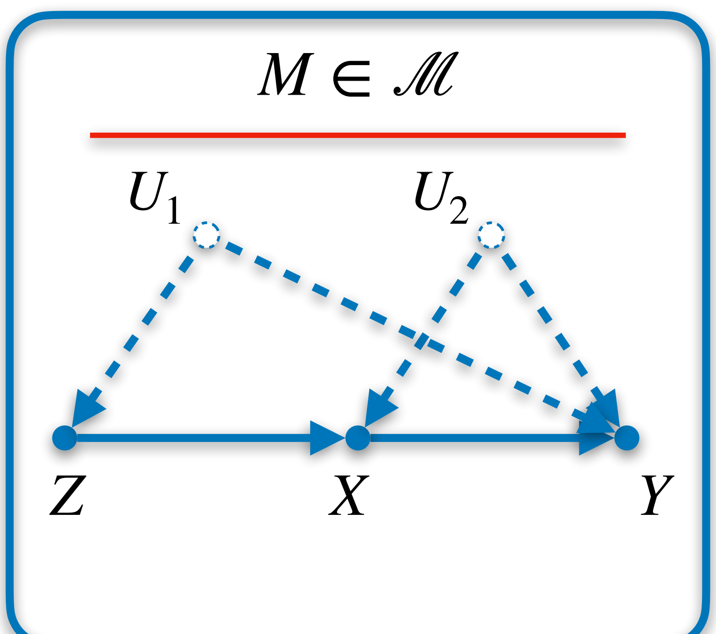




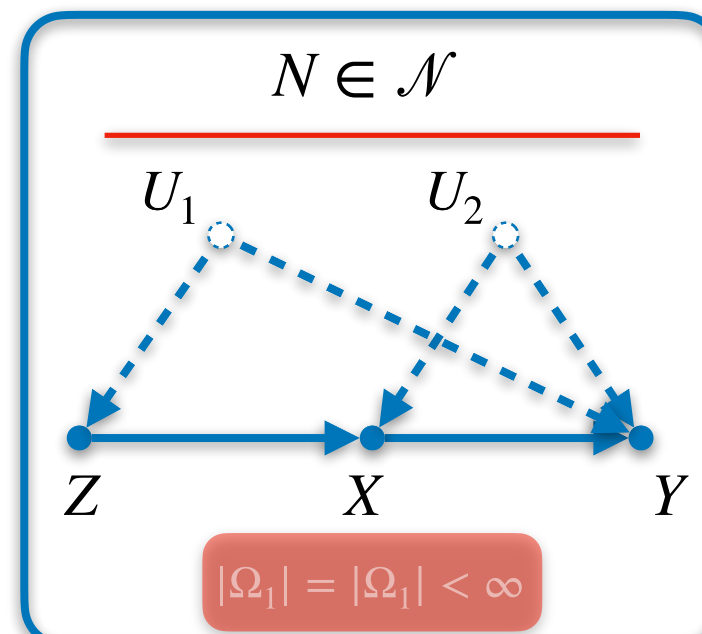
Canonical SCMs

Theorem. For any SCM M, there exists a canonical SCM N s.t.

1. M and N are compatible with the same causal diagram G;
2. For any subsets $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$, $P_M(\mathbf{y} \mid do(\mathbf{x})) = P_N(\mathbf{y} \mid do(\mathbf{x}))$.



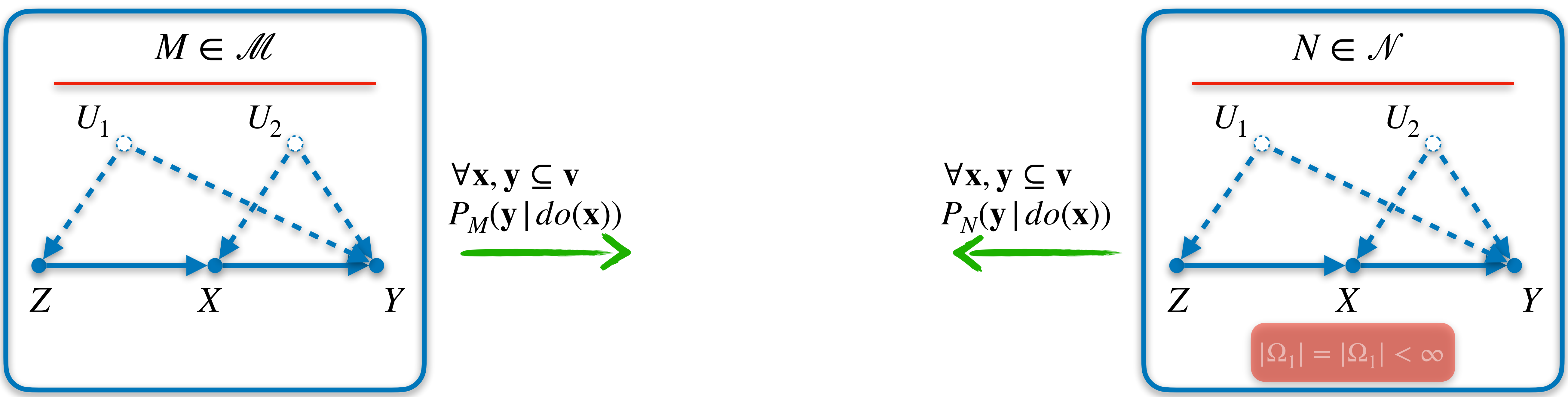
$\forall \mathbf{x}, \mathbf{y} \subseteq \mathbf{v}$
$P_M(\mathbf{y} \mid do(\mathbf{x}))$



Canonical SCMs

Theorem. For any SCM M, there exists a canonical SCM N s.t.

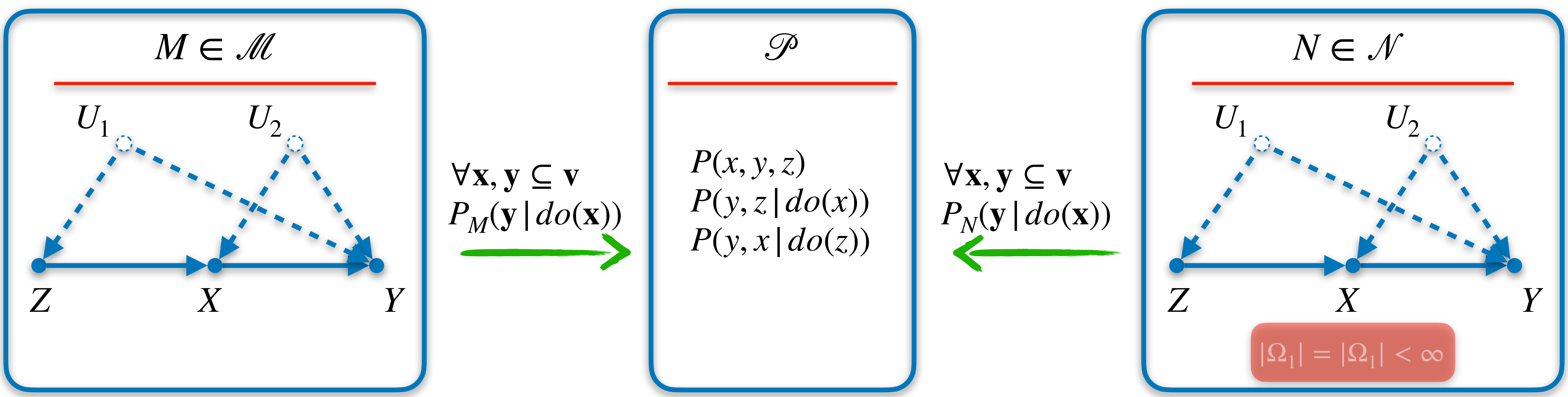
1. M and N are compatible with the same causal diagram G;
2. For any subsets $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$, $P_M(\mathbf{y} \mid do(\mathbf{x})) = P_N(\mathbf{y} \mid do(\mathbf{x}))$.



Canonical SCMs

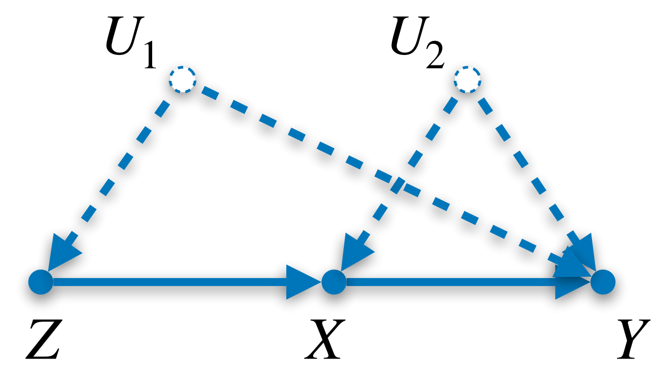
Theorem. For any SCM M, there exists a canonical SCM N s.t.

1. M and N are compatible with the same causal diagram G;
2. For any subsets $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$, $P_M(\mathbf{y} \mid do(\mathbf{x})) = P_N(\mathbf{y} \mid do(\mathbf{x}))$.



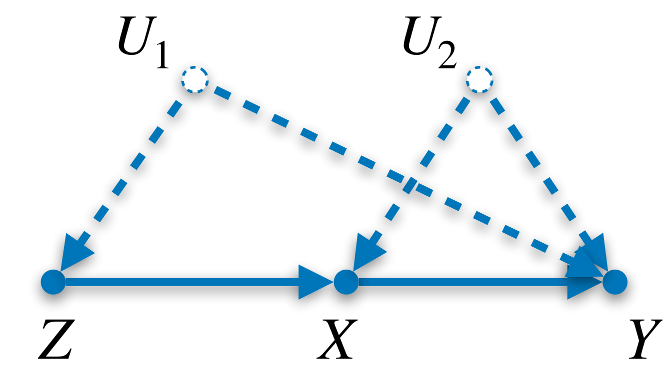
Partial Identification of Causal Effects: Revisit

Task. Given the observational distribution $P(\mathbf{v})$ in an arbitrary causal diagram G, bound $P(\mathbf{y} \mid do(\mathbf{x}))$ for any $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$.



Partial Identification of Causal Effects: Revisit

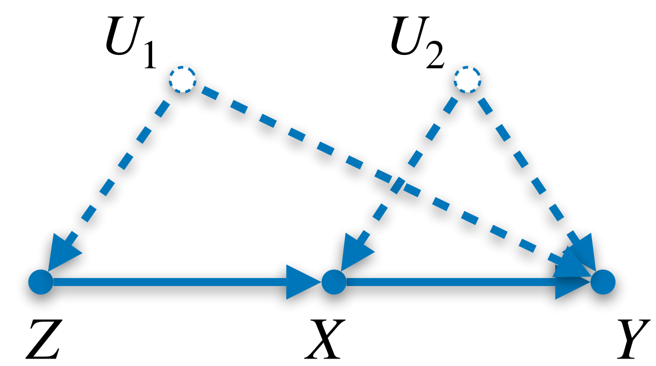
Task. Given the observational distribution $P(\mathbf{v})$ in an arbitrary causal diagram G, bound $P(\mathbf{y} \mid do(\mathbf{x}))$ for any $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$.



- We assume that the domain of $\mathbf{V}$ is discrete and finite.

Partial Identification of Causal Effects: Revisit

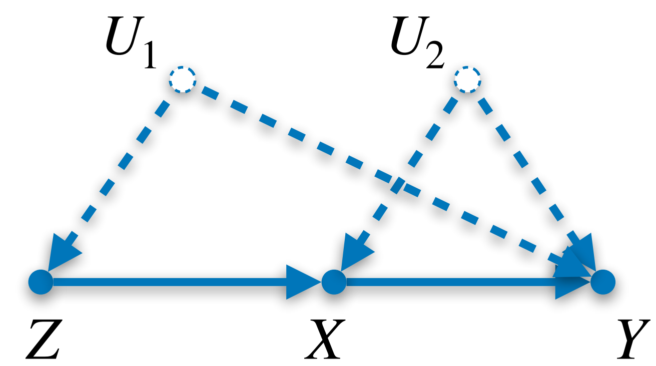
Task. Given the observational distribution $P(\mathbf{v})$ in an arbitrary causal diagram G, bound $P(\mathbf{y} \mid do(\mathbf{x}))$ for any $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$.



- We assume that the domain of $\mathbf{V}$ is discrete and finite.
- Let $\mathcal{N}$ denote the set of all canonical SCMs compatible with G.

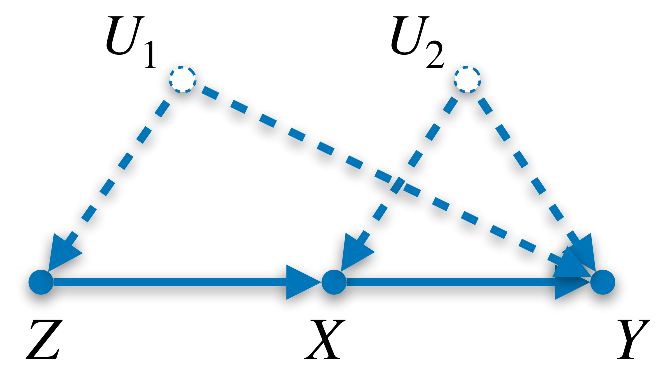
Partial Identification of Causal Effects: Revisit

Task. Given the observational distribution $P(\mathbf{v})$ in an arbitrary causal diagram G, bound $P(\mathbf{y} \mid do(\mathbf{x}))$ for any $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$.



- We assume that the domain of $\mathbf{V}$ is discrete and finite.
- Let $\mathcal{N}$ denote the set of all canonical SCMs compatible with G.
- Given $P(\mathbf{v})$, $P(\mathbf{y} \mid do(\mathbf{x}))$ is bounded in $[a, b]$ where:

Partial Identification of Causal Effects: Revisit



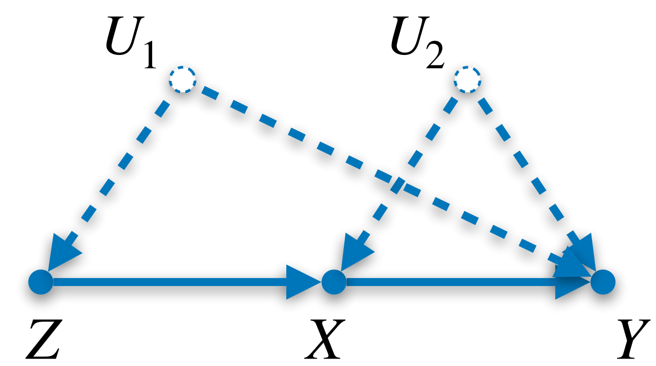
Task. Given the observational distribution $P(\mathbf{v})$ in an arbitrary causal diagram G, bound $P(\mathbf{y} \mid do(\mathbf{x}))$ for any $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$.

- We assume that the domain of $\mathbf{V}$ is discrete and finite.
- Let $\mathcal{N}$ denote the set of all canonical SCMs compatible with G.
- Given $P(\mathbf{v})$, $P(\mathbf{y} \mid do(\mathbf{x}))$ is bounded in $[a, b]$ where:

$$a = \min P_N(\mathbf{y} \mid do(\mathbf{x})), \quad \text{s.t.} \quad \forall N \in \mathcal{N},$$
$$b = \max P_N(\mathbf{y} \mid do(\mathbf{x}))\,. \qquad P_N(\mathbf{v}) = P(\mathbf{v})\,.$$

Partial Identification of Causal Effects: Revisit

Task. Given the observational distribution $P(\mathbf{v})$ in an arbitrary causal diagram G, bound $P(\mathbf{y} \mid do(\mathbf{x}))$ for any $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$.

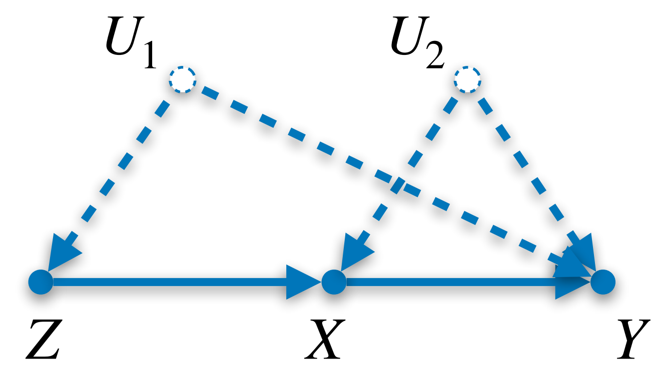


- We assume that the domain of $\mathbf{V}$ is discrete and finite.
- Let $\mathcal{N}$ denote the set of all canonical SCMs compatible with G.
- Given $P(\mathbf{v})$, $P(\mathbf{y} \mid do(\mathbf{x}))$ is bounded in $[a, b]$ where:

$$
\begin{aligned}
a &= \min P_N(\mathbf{y} \mid do(\mathbf{x})), & \quad \text{s.t.} \quad & \forall N \in \mathcal{N}, \\
b &= \max P_N(\mathbf{y} \mid do(\mathbf{x}))\,. & & P_N(\mathbf{v}) = P(\mathbf{v})\,.
\end{aligned}
$$

Partial Identification of Causal Effects: Revisit

Task. Given the observational distribution $P(\mathbf{v})$ in an arbitrary causal diagram G, bound $P(\mathbf{y} \mid do(\mathbf{x}))$ for any $\mathbf{X}, \mathbf{Y} \subseteq \mathbf{V}$.

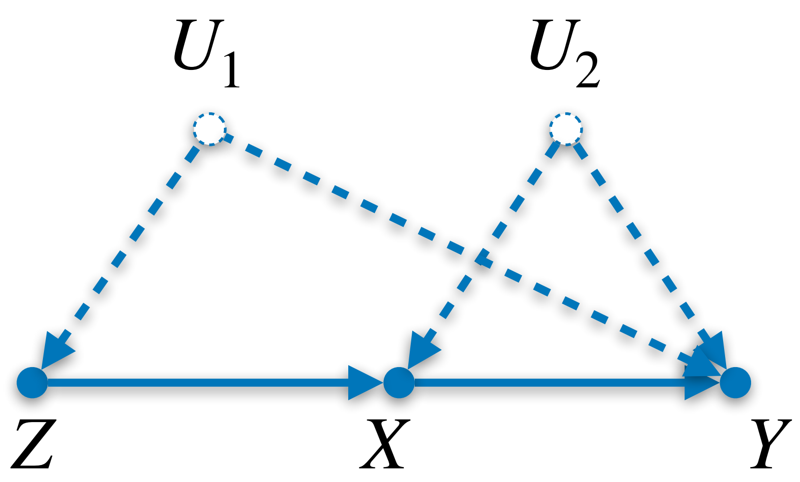


- We assume that the domain of $\mathbf{V}$ is discrete and finite.
- Let $\mathcal{N}$ denote the set of all canonical SCMs compatible with G.
- Given $P(\mathbf{v})$, $P(\mathbf{y} \mid do(\mathbf{x}))$ is bounded in $[a, b]$ where:

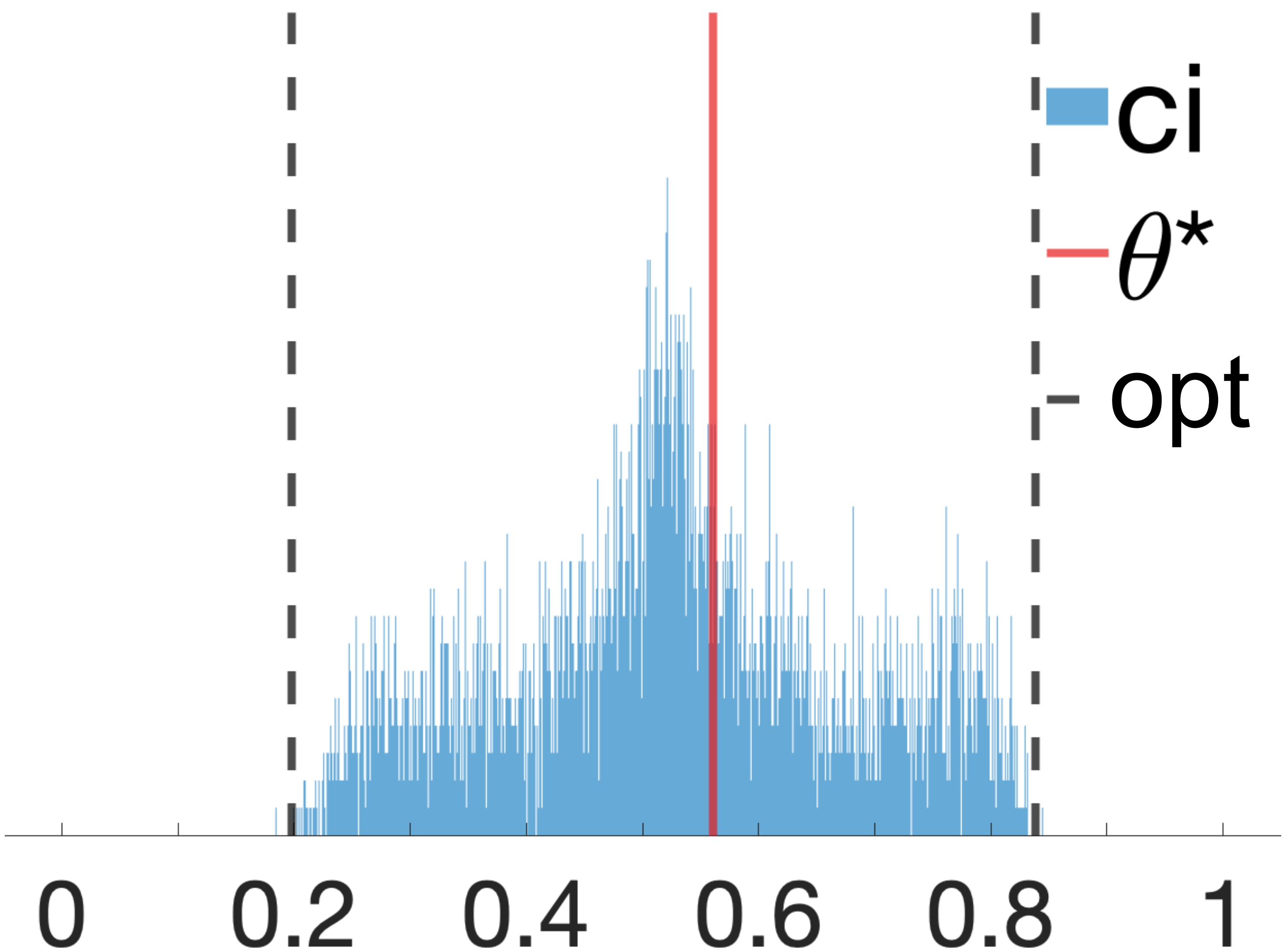
$$a = \min P_N(\mathbf{y} \mid do(\mathbf{x})), \quad b = \max P_N(\mathbf{y} \mid do(\mathbf{x})) \,. \quad \text{s.t.} \quad \forall N \in \mathcal{N}, \; P_N(\mathbf{v}) = P(\mathbf{v}) \,.$$

This problem is reducible to an equivalent polynomial optimization program

Example: Non-IV



- $X, Y, Z \in \{0,1\}$
- $U_1, U_2 \in \mathbb{R}$
- Data - $P(x, y, z)$
- Query - $P(y \mid do(x))$



$N = 1000$

Conclusions

Conclusions

- We introduce canonical causal models that could represent all interventional distributions in an arbitrary causal diagram.

Conclusions

- We introduce canonical causal models that could represent all interventional distributions in an arbitrary causal diagram.
- It reduces partial causal identification to equivalent polynomial programs.

Conclusions

- We introduce canonical causal models that could represent all interventional distributions in an arbitrary causal diagram.
- It reduces partial causal identification to equivalent polynomial programs.
- What is in the paper (Contributions):

Conclusions

- We introduce canonical causal models that could represent all interventional distributions in an arbitrary causal diagram.
- It reduces partial causal identification to equivalent polynomial programs.
- What is in the paper (Contributions):
    - Generalized canonical SCMs that could represent all counterfactual distributions in a causal digram.

Conclusions

- We introduce canonical causal models that could represent all interventional distributions in an arbitrary causal diagram.
- It reduces partial causal identification to equivalent polynomial programs.
- What is in the paper (Contributions):
  - Generalized canonical SCMs that could represent all counterfactual distributions in a causal digram.
  - Effective posterior sampling methods to approximate optimal bounds over unknown counterfactual probabilities from observational and experimental data.