Causal Discovery from Interventions

Brady Neal

causalcourse.com

Structural Interventions

Single-Node Interventions

Multi-Node Interventions

Parametric Interventions

Interventional Markov Equivalence

Miscellaneous Other Settings

Structural Interventions

Single-Node Interventions

Multi-Node Interventions

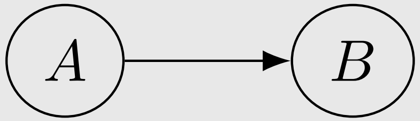
Parametric Interventions

Interventional Markov Equivalence

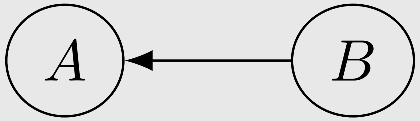
Miscellaneous Other Settings

Intervention in Two-Variable Setting

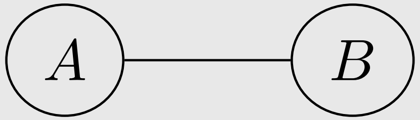
Intervention in Two-Variable Setting



or

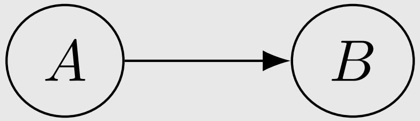


Same Markov equivalence class / essential graph

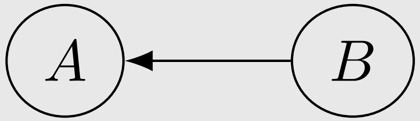


Intervention in Two-Variable Setting

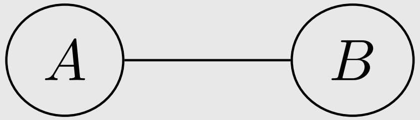
$$A := N_A$$
$$B := f(A, N_B)$$



or

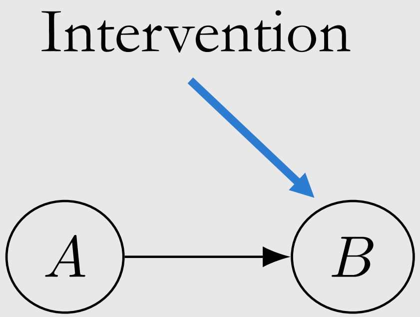


Same Markov equivalence class / essential graph

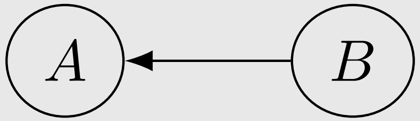


Intervention in Two-Variable Setting

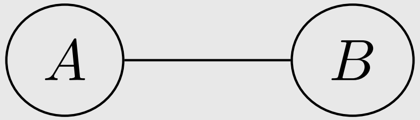
$$A := N_A$$
$$B := f(A, N_B)$$



or

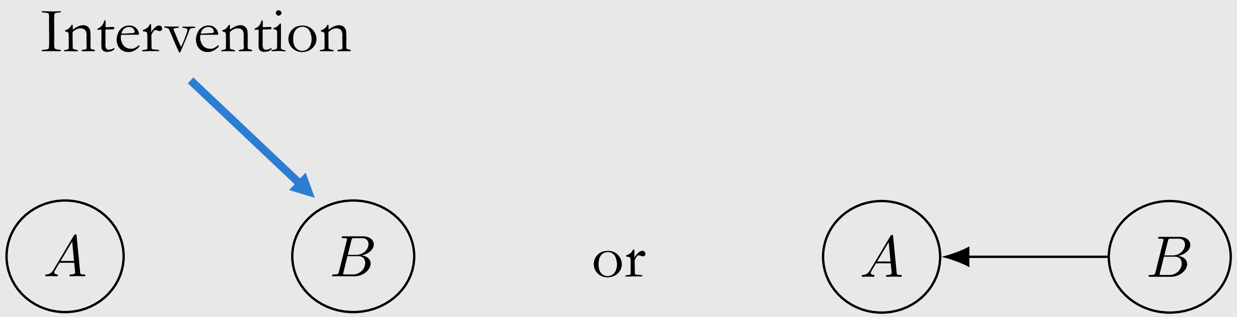


Same Markov equivalence class / essential graph

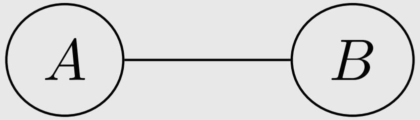


Intervention in Two-Variable Setting

$$A := N_A$$
$$B := f(A, N_B)$$

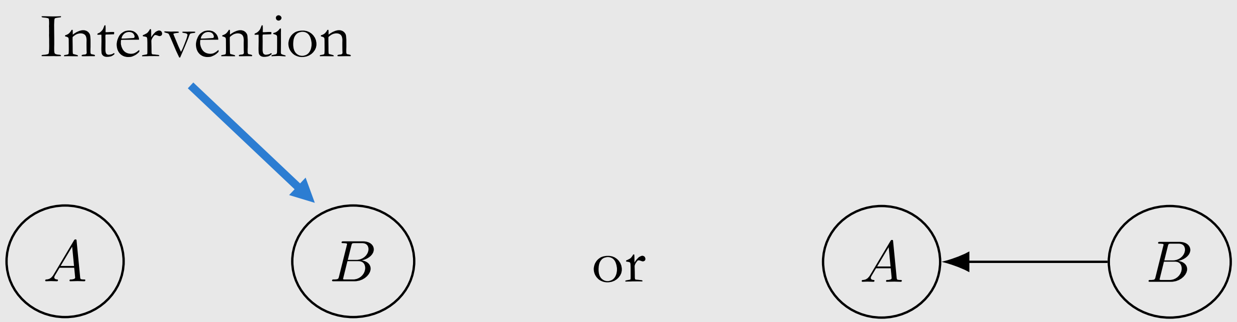


Same Markov equivalence class / essential graph



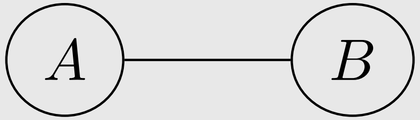
Intervention in Two-Variable Setting

$$A := N_A$$
$$B := f(A, N_B)$$



Manipulated graph

Same Markov equivalence class / essential graph



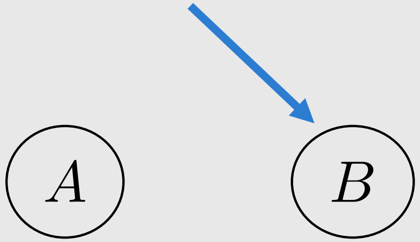
Intervention in Two-Variable Setting

$A := N_A$

~~$B := f(A, N_B)$~~

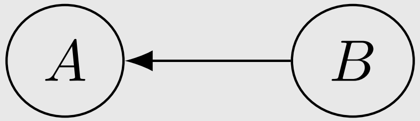
$B := N'_B$

Intervention

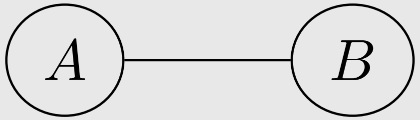


Manipulated graph

or



Same Markov equivalence class / essential graph



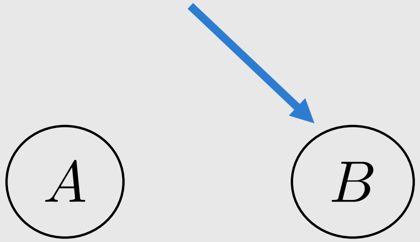
Intervention in Two-Variable Setting

$A := N_A$

~~$B := f(A, N_B)$~~

$B := N'_B$

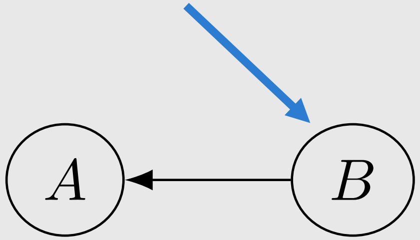
Intervention



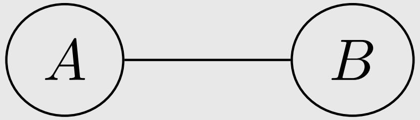
Manipulated graph

or

Intervention



Same Markov equivalence class / essential graph



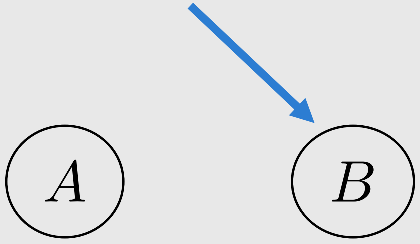
Intervention in Two-Variable Setting

$A := N_A$

~~$B := f(A, N_B)$~~

$B := N'_B$

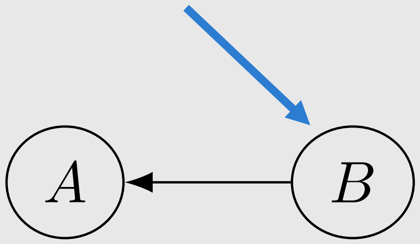
Intervention



Manipulated graph

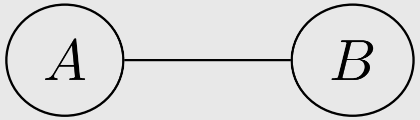
or

Intervention



Manipulated graph

Same Markov equivalence class / essential graph



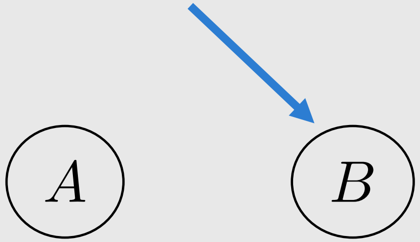
Intervention in Two-Variable Setting

$A := N_A$

~~$B := f(A, N_B)$~~

$B := N'_B$

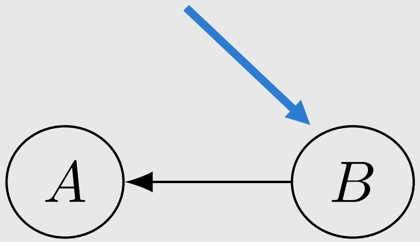
Intervention



Manipulated graph

or

Intervention

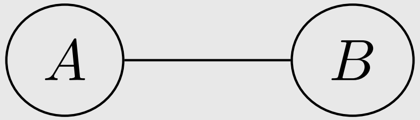


Manipulated graph

$B := N_B$

$A := f(B, N_A)$

Same Markov equivalence class / essential graph



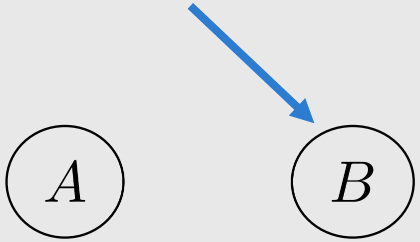
Intervention in Two-Variable Setting

$A := N_A$

~~$B := f(A, N_B)$~~

$B := N'_B$

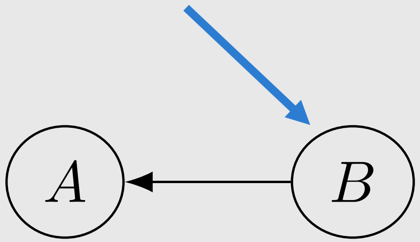
Intervention



Manipulated graph

or

Intervention



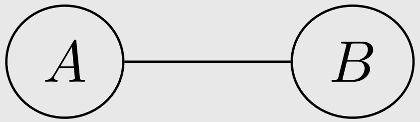
Manipulated graph

$B := N'_B$

~~$B := N_B$~~

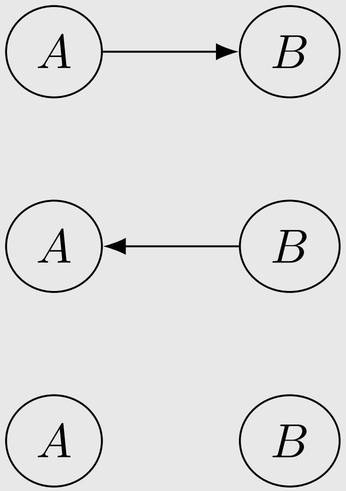
$A := f(B, N_A)$

Same Markov equivalence class / essential graph



Two Variables: Interventional Essential Graphs

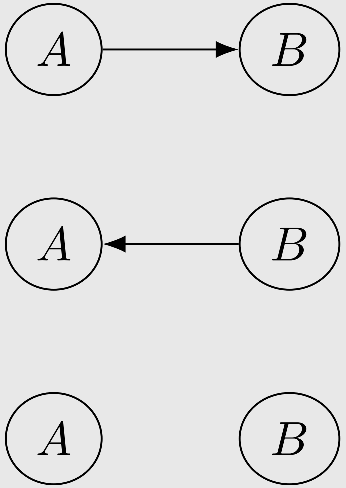
True graph



Two Variables: Interventional Essential Graphs

Interventional essential graphs

True graph

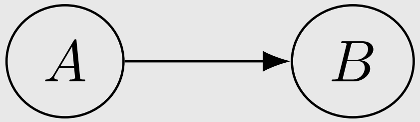


I = {A}

I = {B}

I = {}

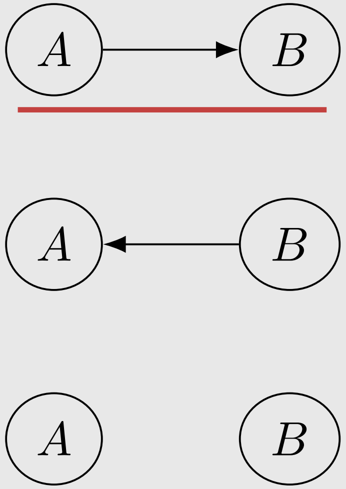
Two Variables: Interventional Essential Graphs



Interventional essential graphs

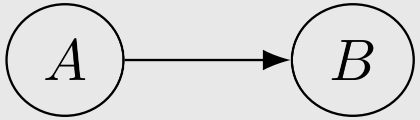
True graph

$I = \{A\}$ $I = \{B\}$ $I = \{\}$



Two Variables: Interventional Essential Graphs

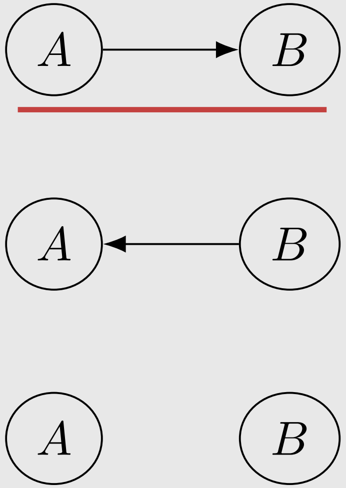
True graph under intervention I = {A}



Interventional essential graphs

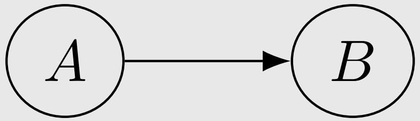
True graph

I = {A} | I = {B} | I = {}



Two Variables: Interventional Essential Graphs

True graph under intervention I = {A}



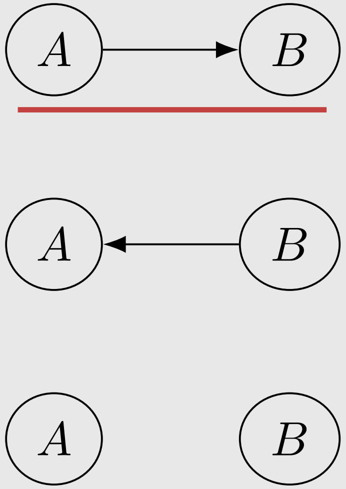
Interventional essential graphs

True graph

I = {A}

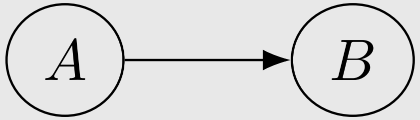
I = {B}

I = {}



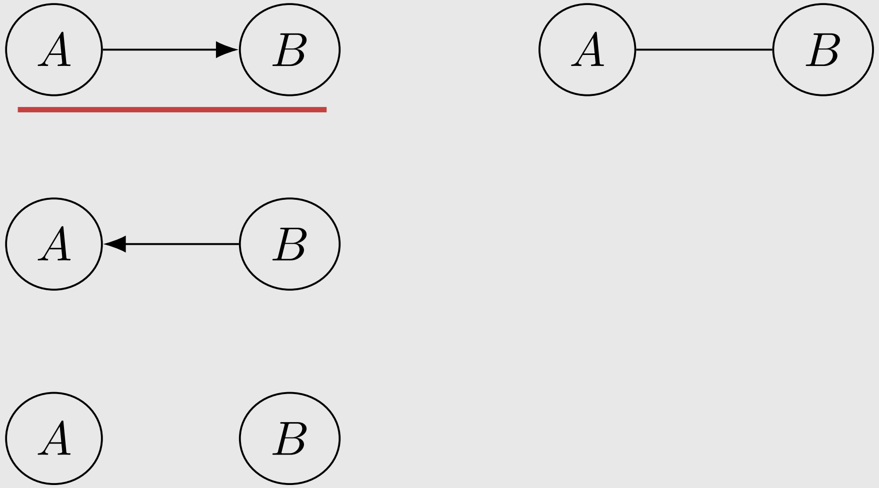
Two Variables: Interventional Essential Graphs

True graph under intervention I = {A}



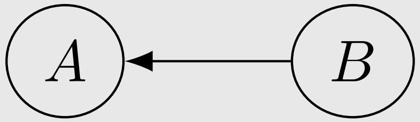
Interventional essential graphs

True graph | I = {A} | I = {B} | I = {}



Two Variables: Interventional Essential Graphs

True graph under intervention I = {A}



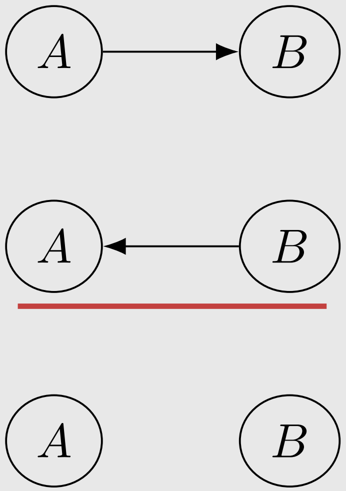
Interventional essential graphs

True graph

I = {A}

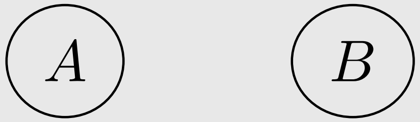
I = {B}

I = {}



Two Variables: Interventional Essential Graphs

True graph under intervention I = {A}



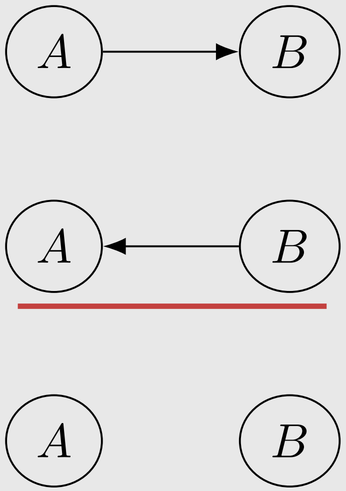
Interventional essential graphs

True graph

I = {A}

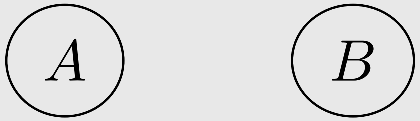
I = {B}

I = {}



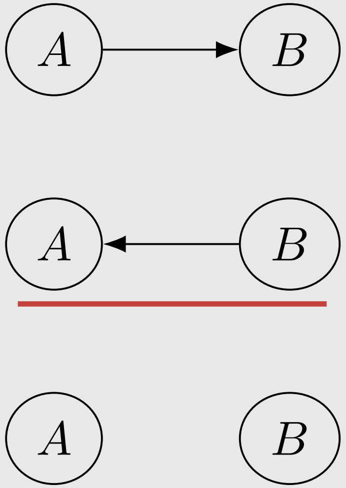
Two Variables: Interventional Essential Graphs

True graph under intervention $I = \{A\}$

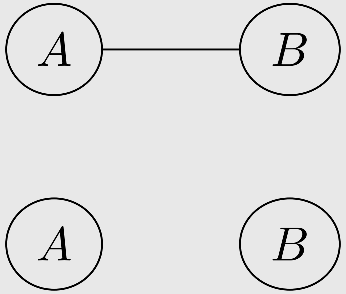


Interventional essential graphs

True graph



$I = \{A\}$

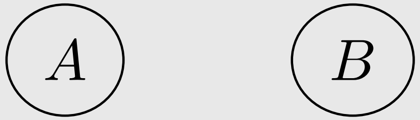


$I = \{B\}$

$I = \{\}$

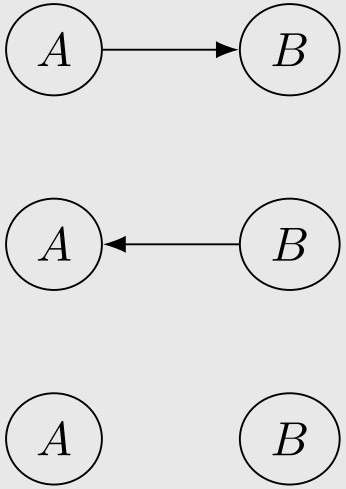
Two Variables: Interventional Essential Graphs

True graph under intervention I = {A}

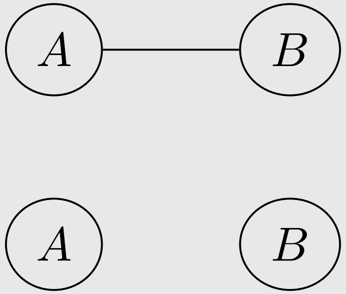


Interventional essential graphs

True graph



I = {A}

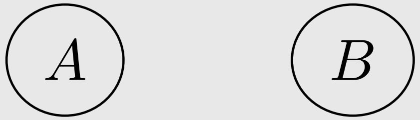


I = {B}

I = {}

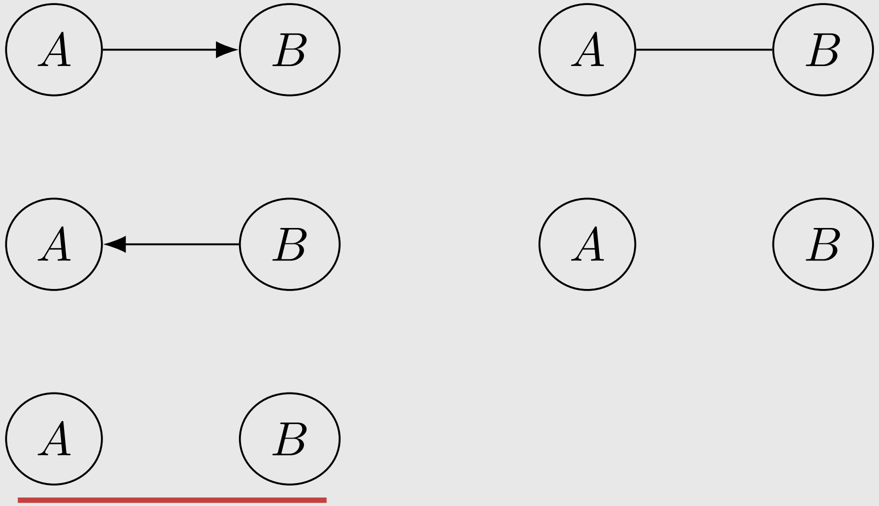
Two Variables: Interventional Essential Graphs

True graph under intervention I = {A}



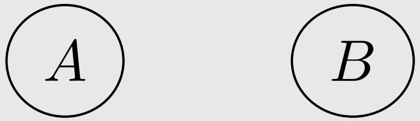
Interventional essential graphs

True graph | I = {A} | I = {B} | I = {}



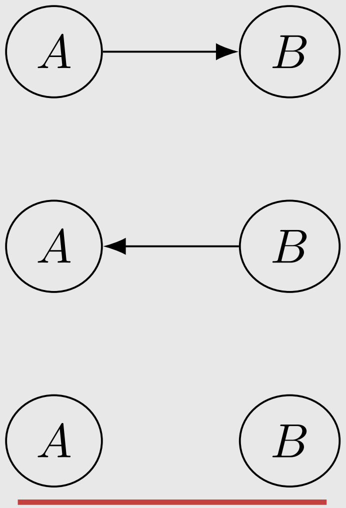
Two Variables: Interventional Essential Graphs

True graph under intervention I = {A}

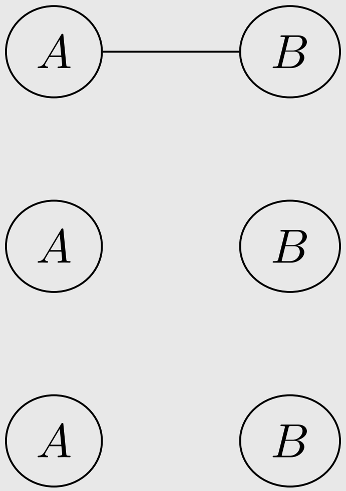


Interventional essential graphs

True graph



I = {A}

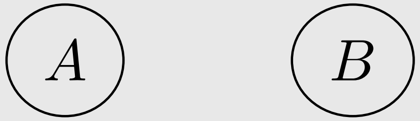


I = {B}

I = {}

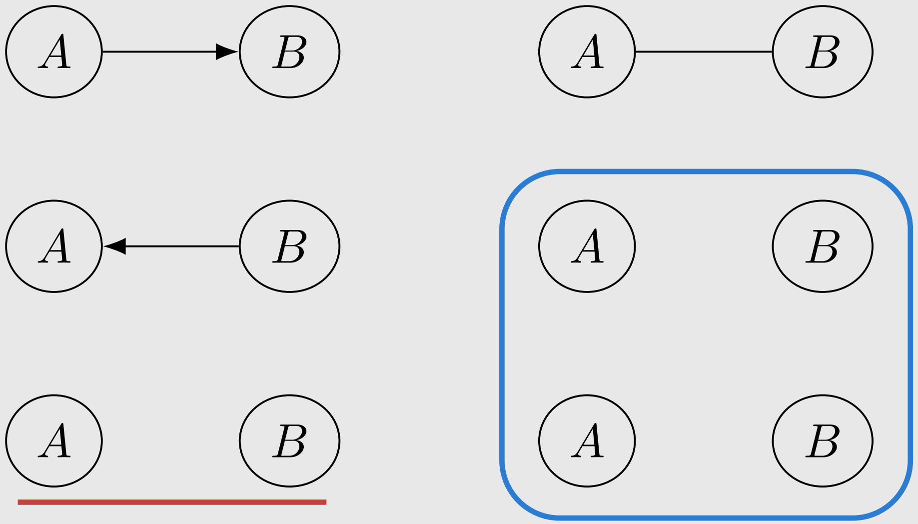
Two Variables: Interventional Essential Graphs

True graph under intervention $I = \{A\}$

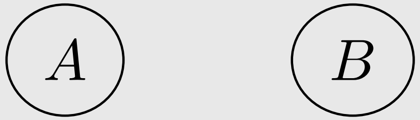


Interventional essential graphs

True graph | $I = \{A\}$ | $I = \{B\}$ | $I = \{\}$

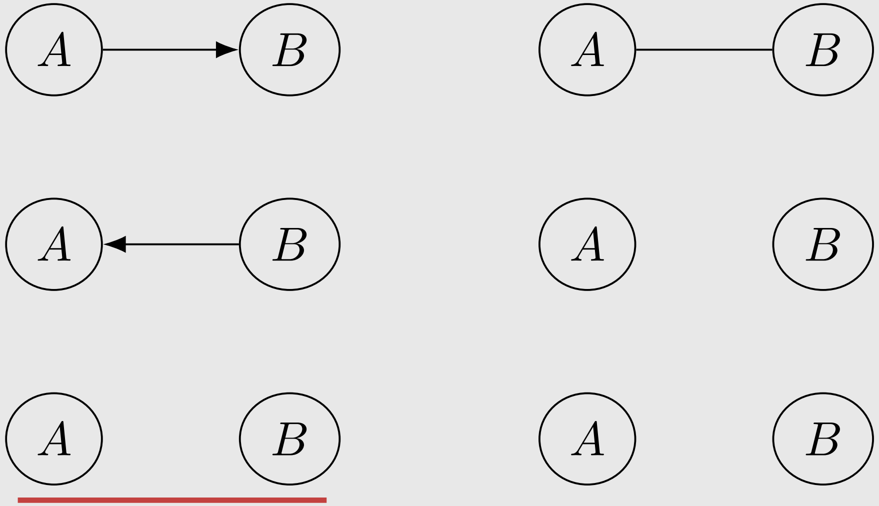


Two Variables: Interventional Essential Graphs

True graph under intervention I = {B} 

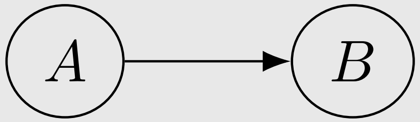
Interventional essential graphs

True graph | I = {A} | I = {B} | I = {}



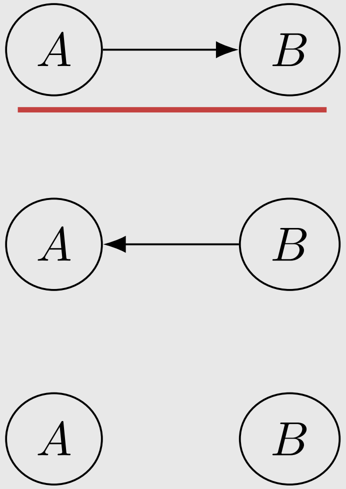
Two Variables: Interventional Essential Graphs

True graph under intervention I = {B}

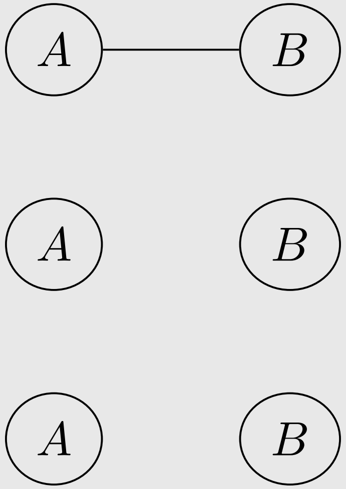


Interventional essential graphs

True graph



I = {A}

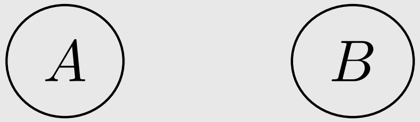


I = {B}

I = {}

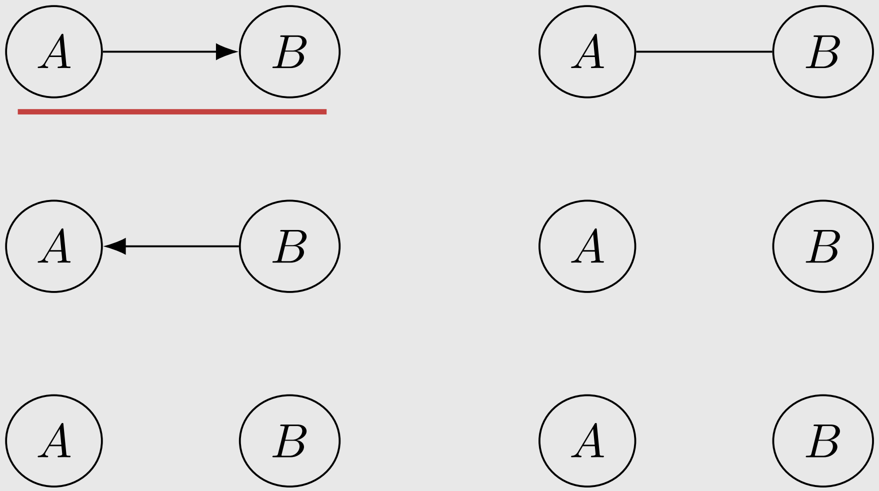
Two Variables: Interventional Essential Graphs

True graph under intervention $I = \{B\}$



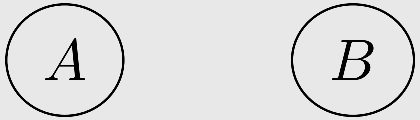
Interventional essential graphs

True graph | $I = \{A\}$ | $I = \{B\}$ | $I = \{\}$



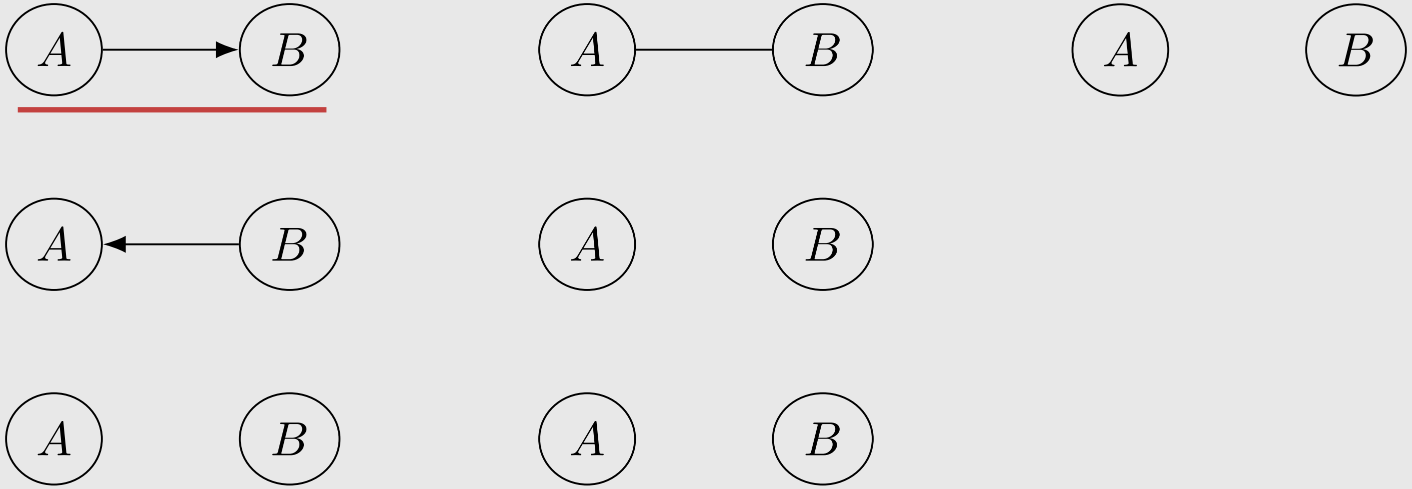
Two Variables: Interventional Essential Graphs

True graph under intervention $I = \{B\}$



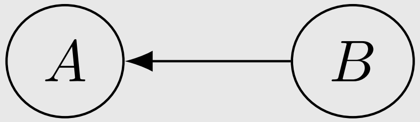
Interventional essential graphs

True graph | $I = \{A\}$ | $I = \{B\}$ | $I = \{\}$



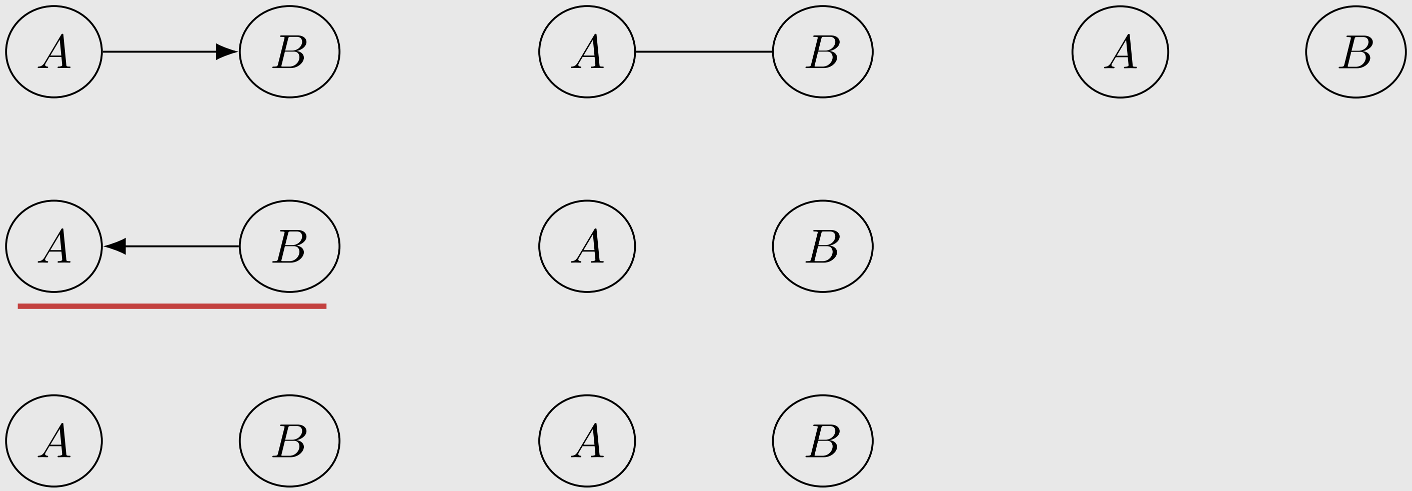
Two Variables: Interventional Essential Graphs

True graph under intervention $I = \{B\}$



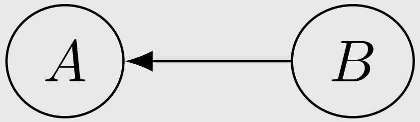
Interventional essential graphs

True graph	$I = \{A\}$	$I = \{B\}$	$I = \{\}$



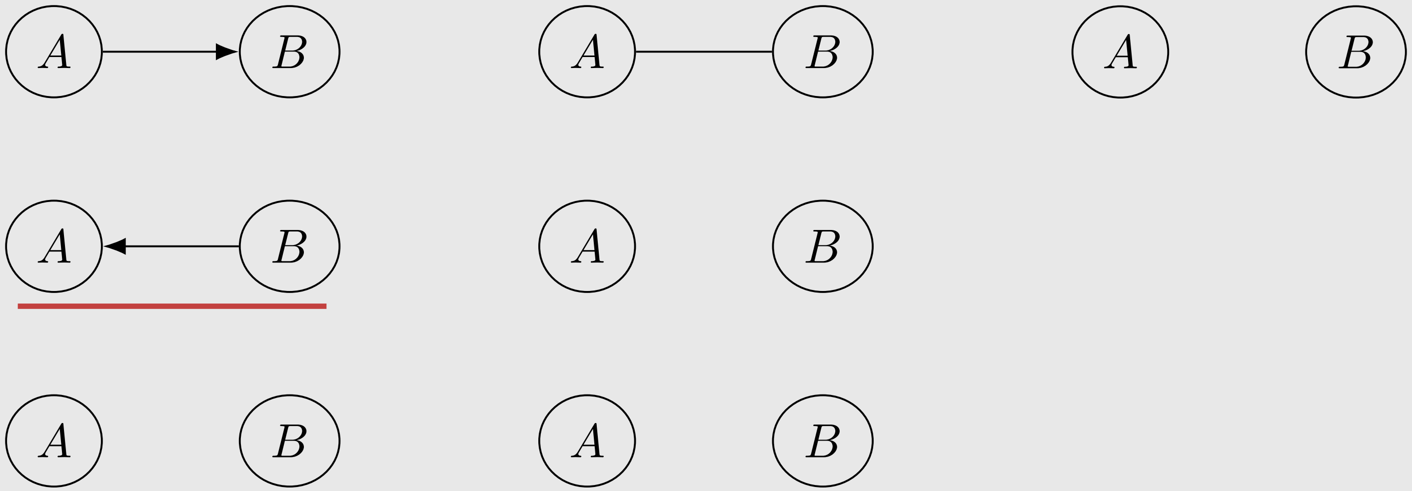
Two Variables: Interventional Essential Graphs

True graph under intervention I = {B}



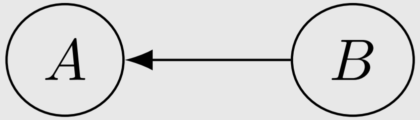
Interventional essential graphs

True graph | I = {A} | I = {B} | I = {}



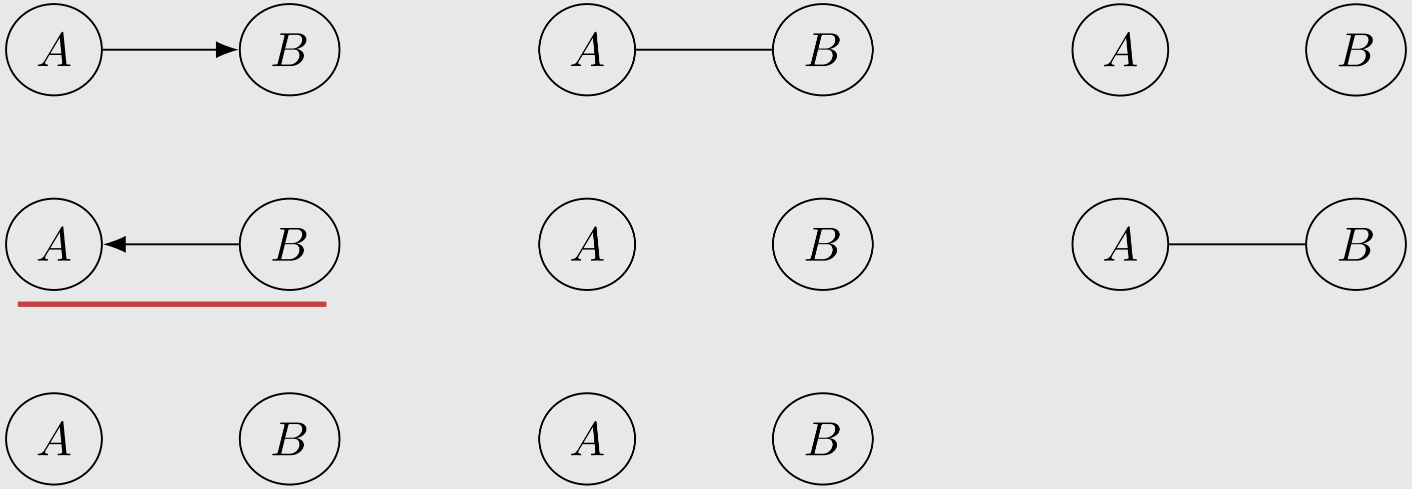
Two Variables: Interventional Essential Graphs

True graph under intervention I = {B}



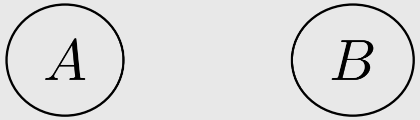
Interventional essential graphs

True graph	I = {A}	I = {B}	I = {}



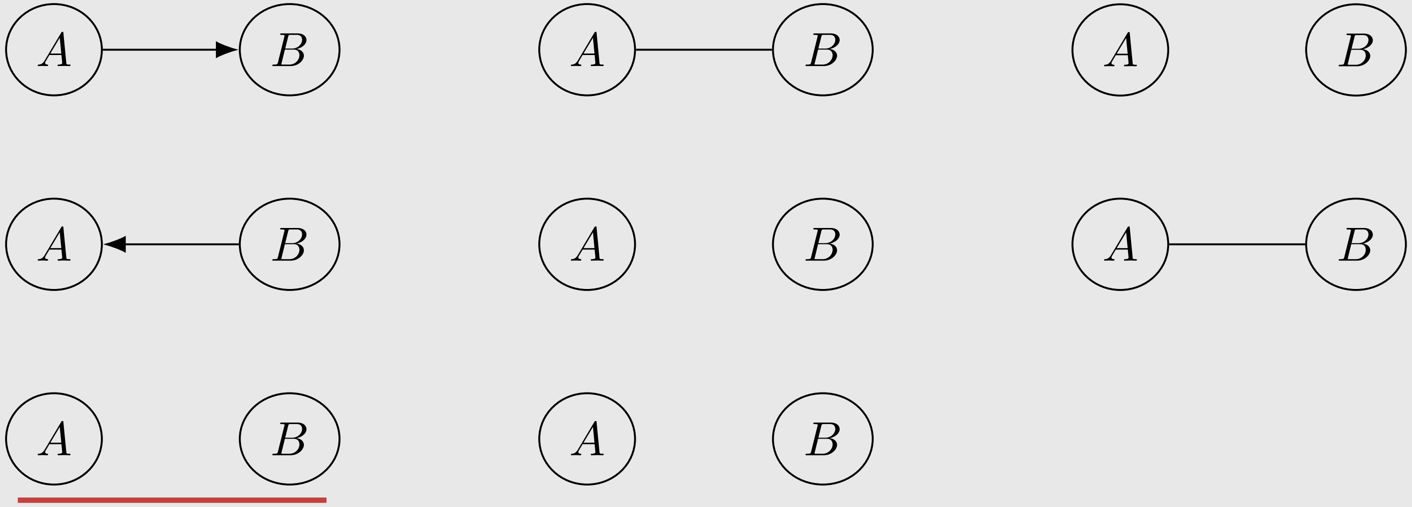
Two Variables: Interventional Essential Graphs

True graph under intervention I = {B}



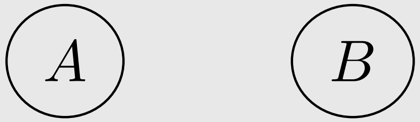
Interventional essential graphs

True graph | I = {A} | I = {B} | I = {}



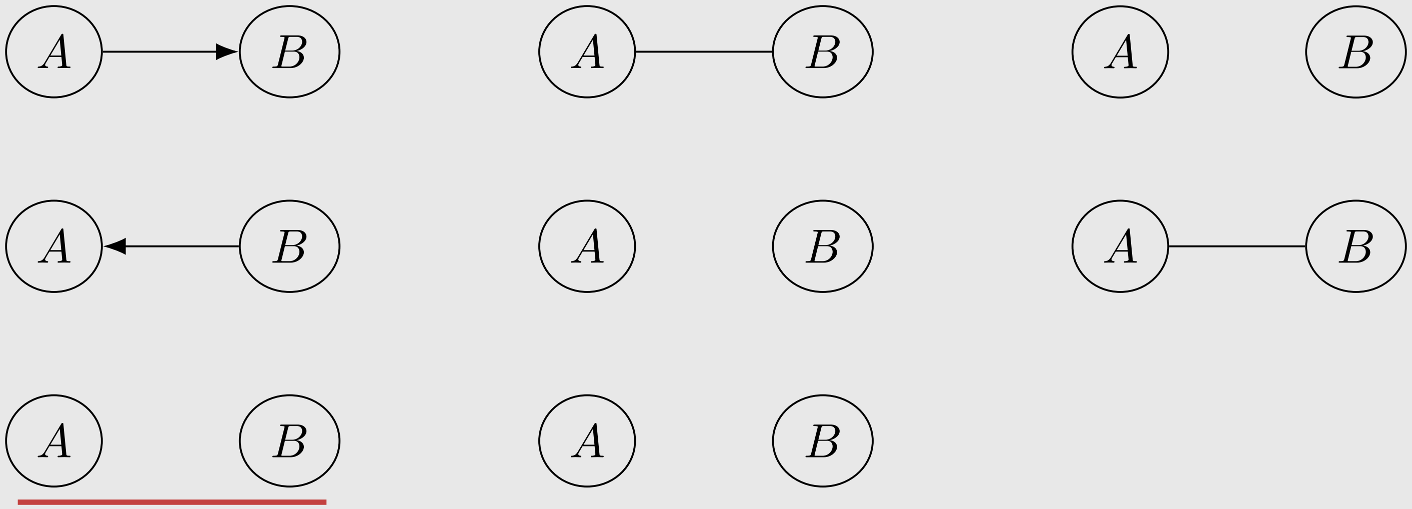
Two Variables: Interventional Essential Graphs

True graph under intervention I = {B}



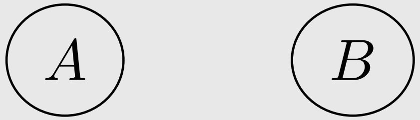
Interventional essential graphs

True graph | I = {A} | I = {B} | I = {}



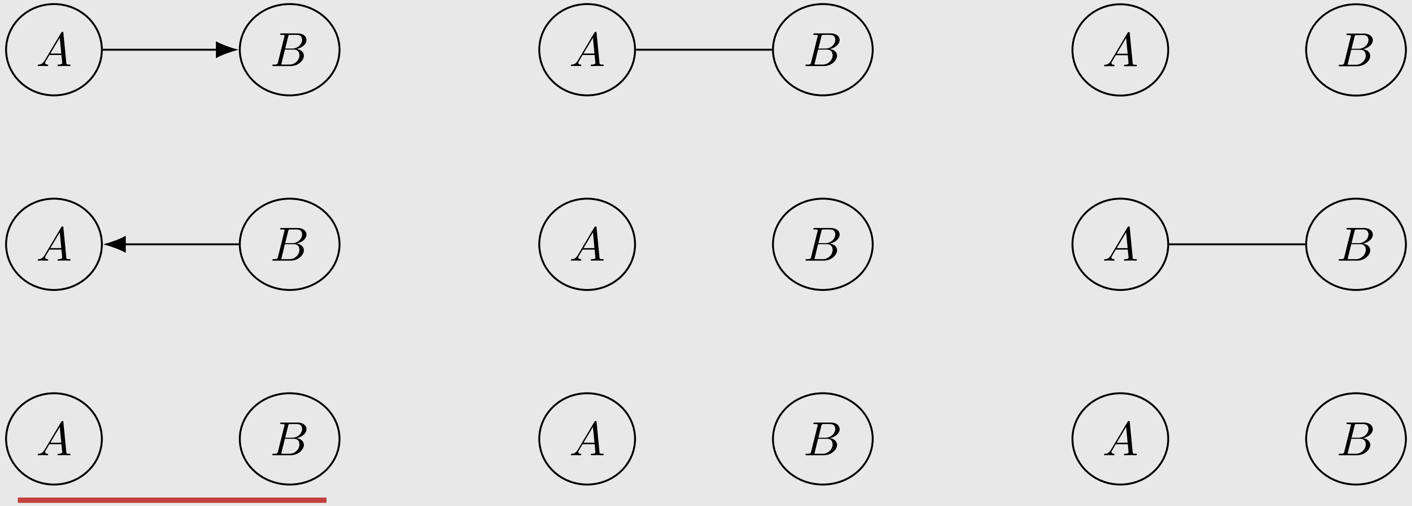
Two Variables: Interventional Essential Graphs

True graph under intervention I = {B}



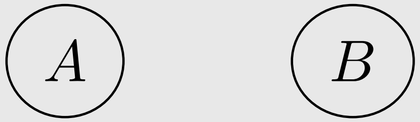
Interventional essential graphs

True graph | I = {A} | I = {B} | I = {}



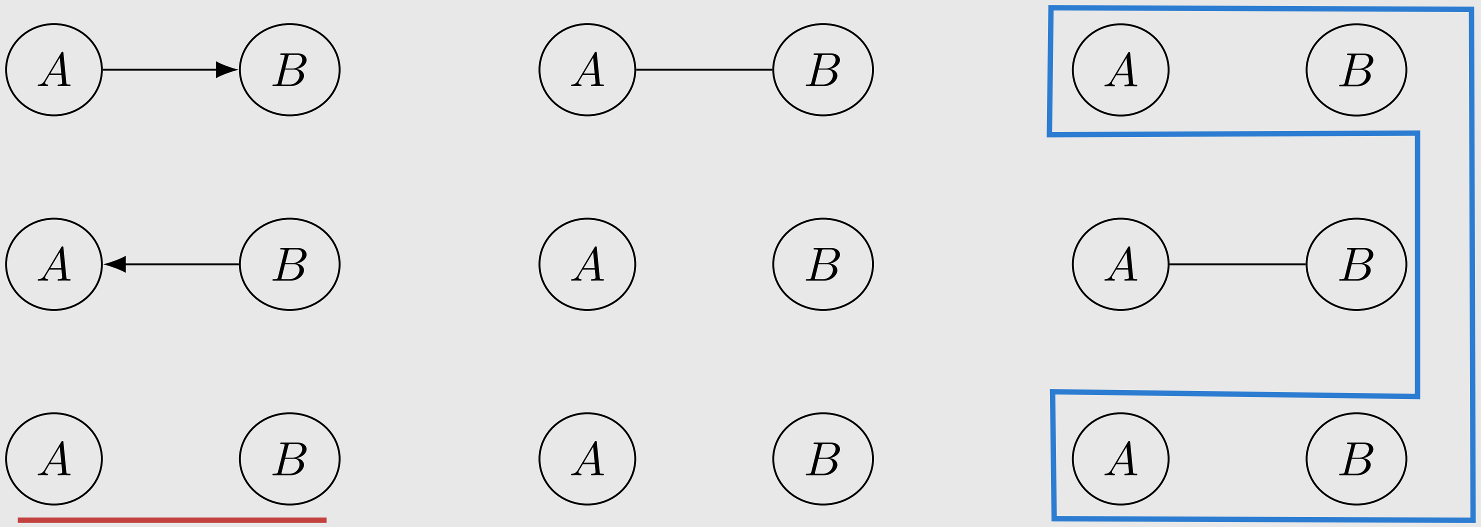
Two Variables: Interventional Essential Graphs

True graph under intervention $I = \{B\}$



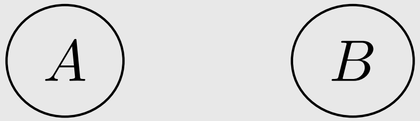
Interventional essential graphs

True graph | $I = \{A\}$ | $I = \{B\}$ | $I = \{\}$



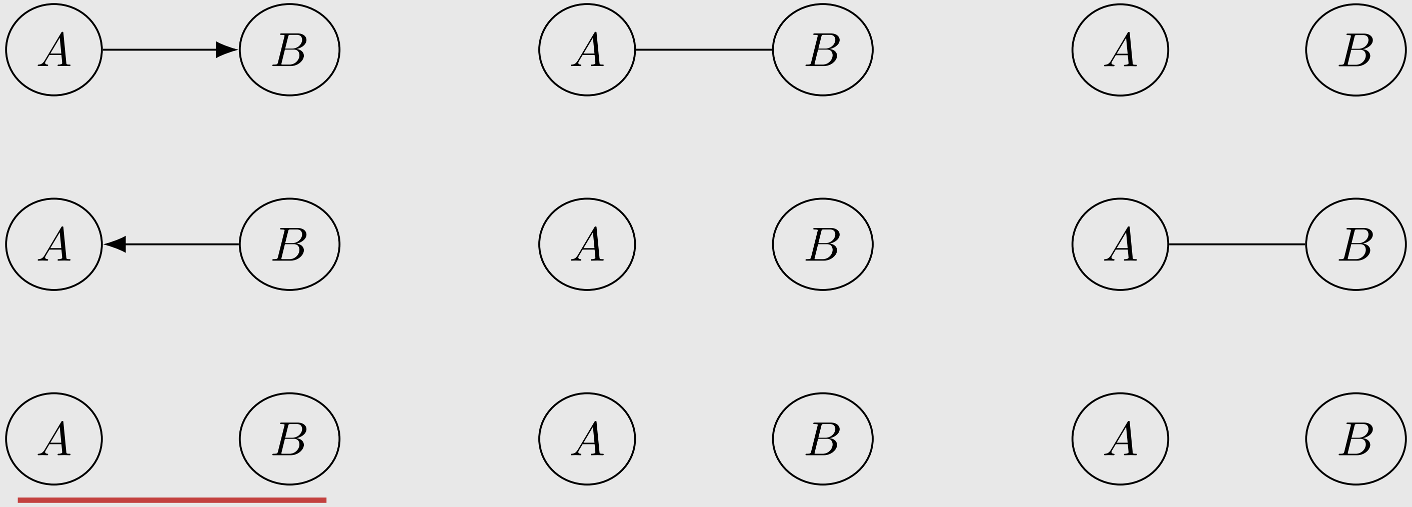
Two Variables: Interventional Essential Graphs

True graph under intervention $I = \{\}$



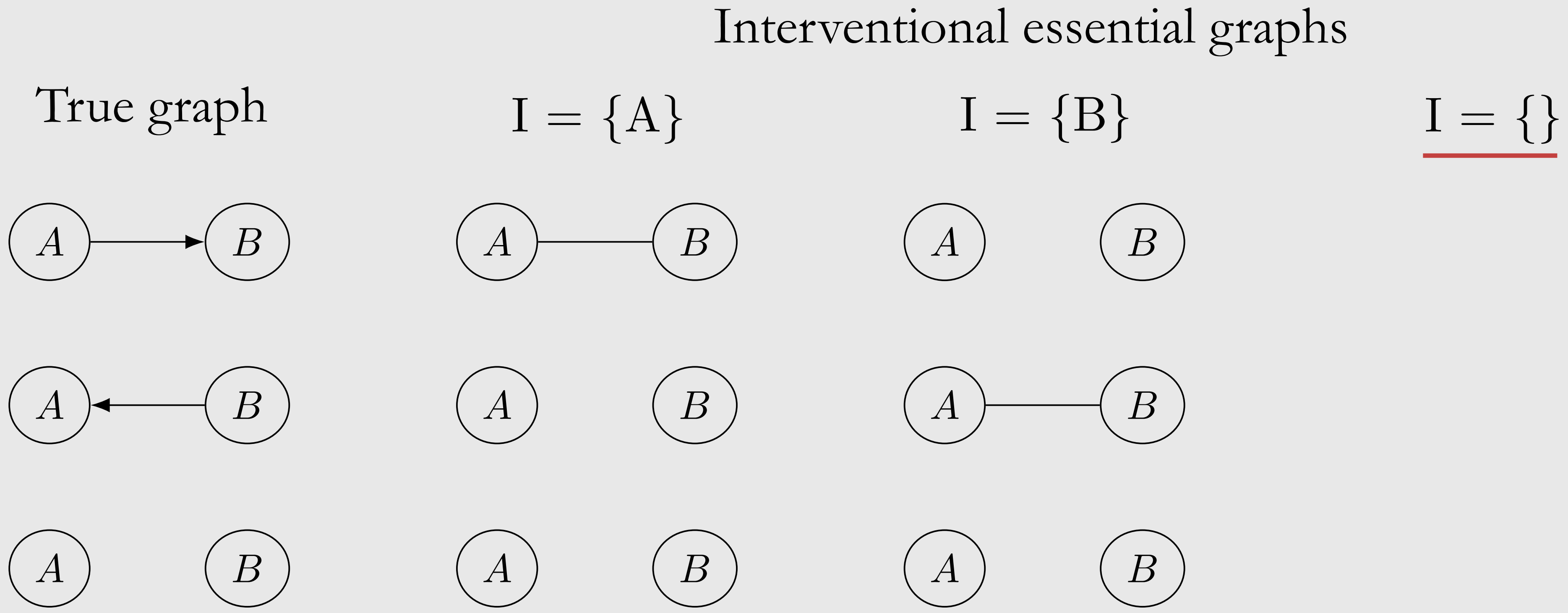
Interventional essential graphs

True graph	$I = \{A\}$	$I = \{B\}$	$I = \{\}$



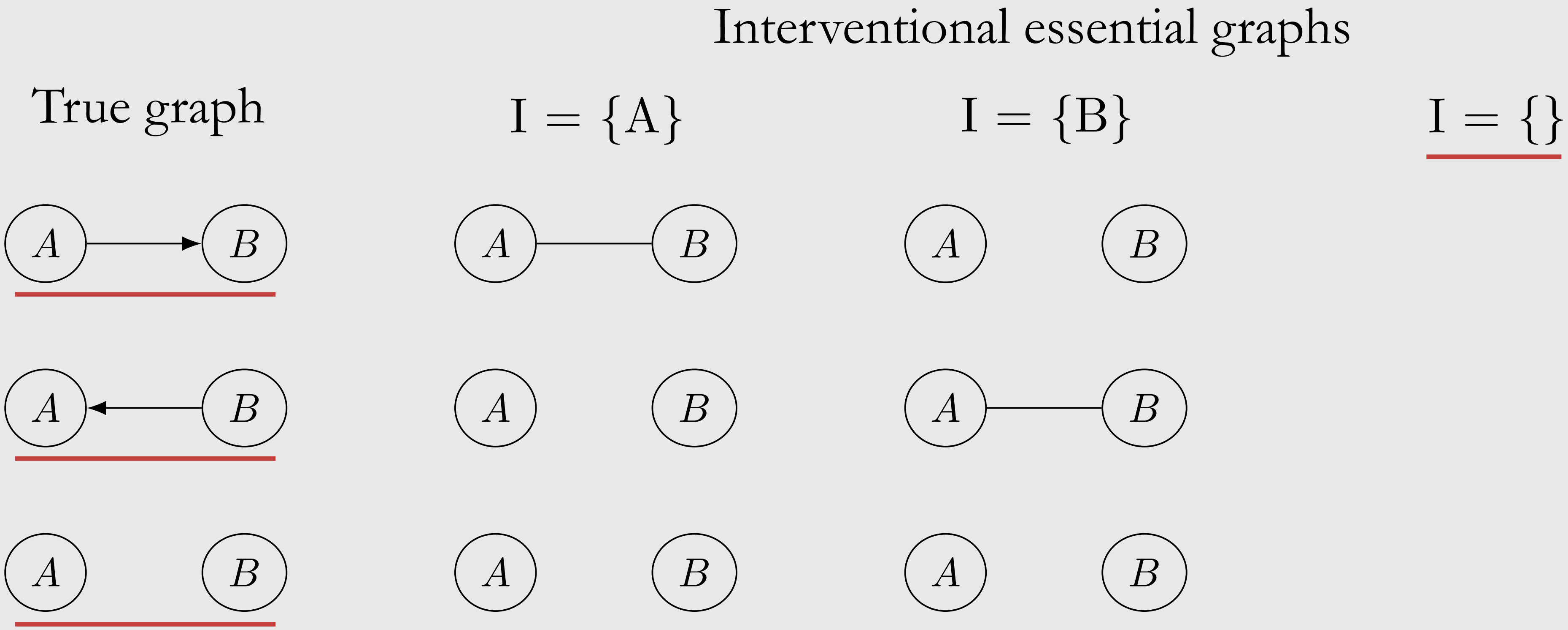
Two Variables: Interventional Essential Graphs

True graph under intervention $I = \{\}$



Two Variables: Interventional Essential Graphs

True graph under intervention $I = \{\}$

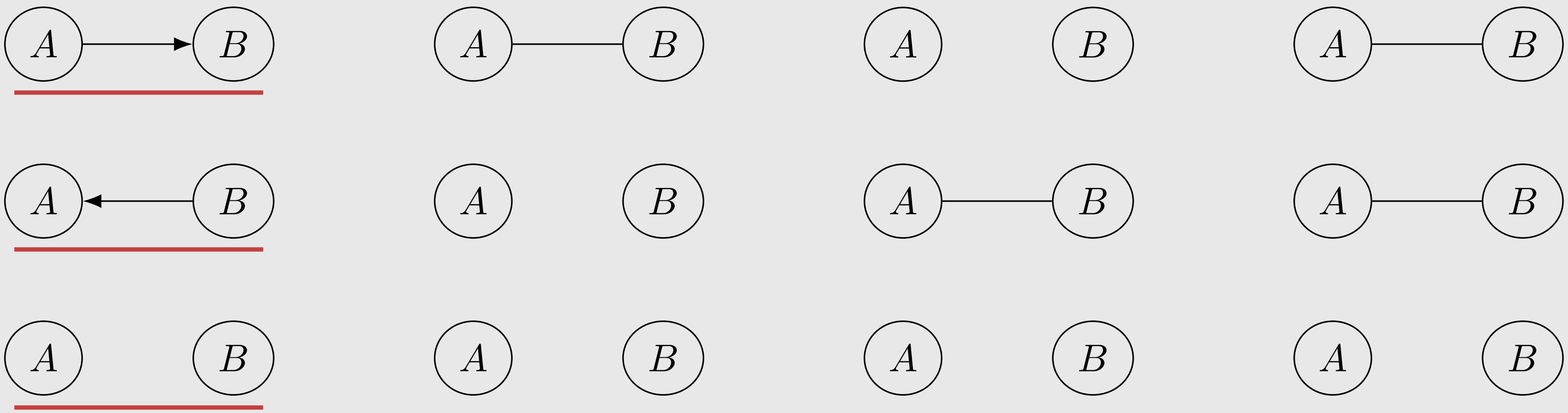


Two Variables: Interventional Essential Graphs

True graph under intervention $I = \{\}$

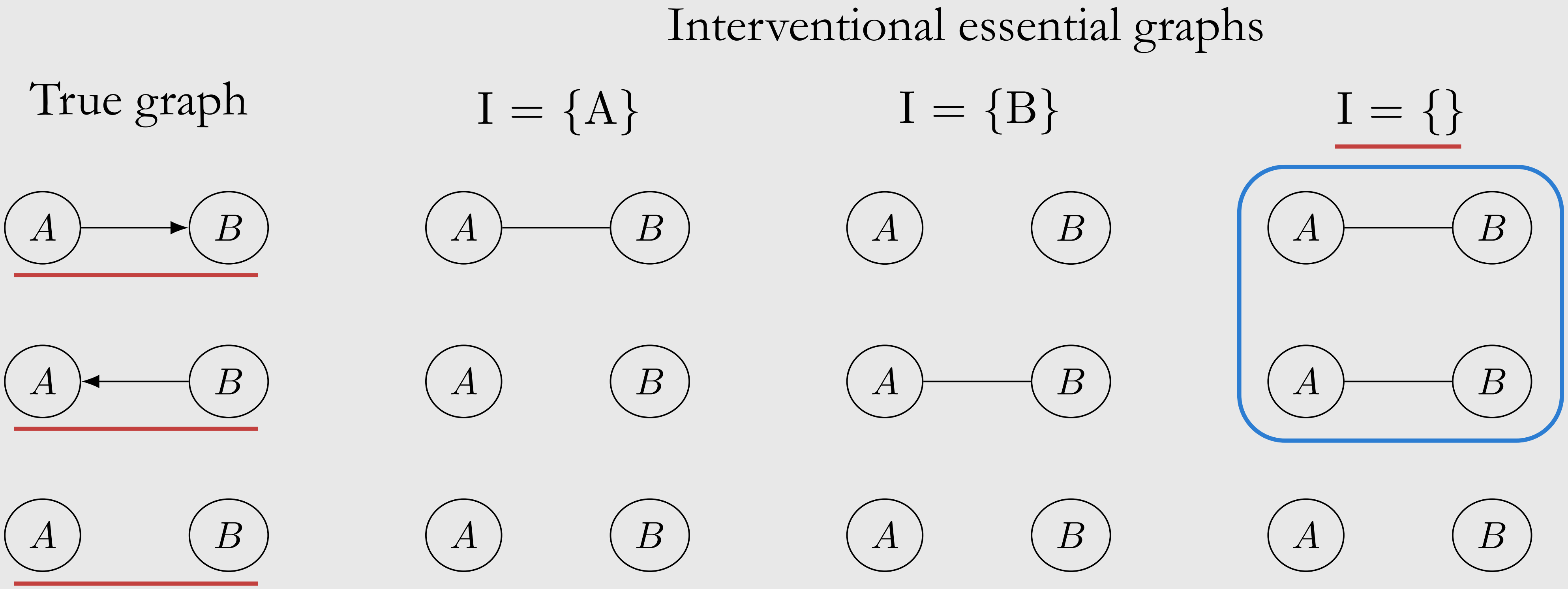
Interventional essential graphs

True graph | $I = \{A\}$ | $I = \{B\}$ | $I = \{\}$

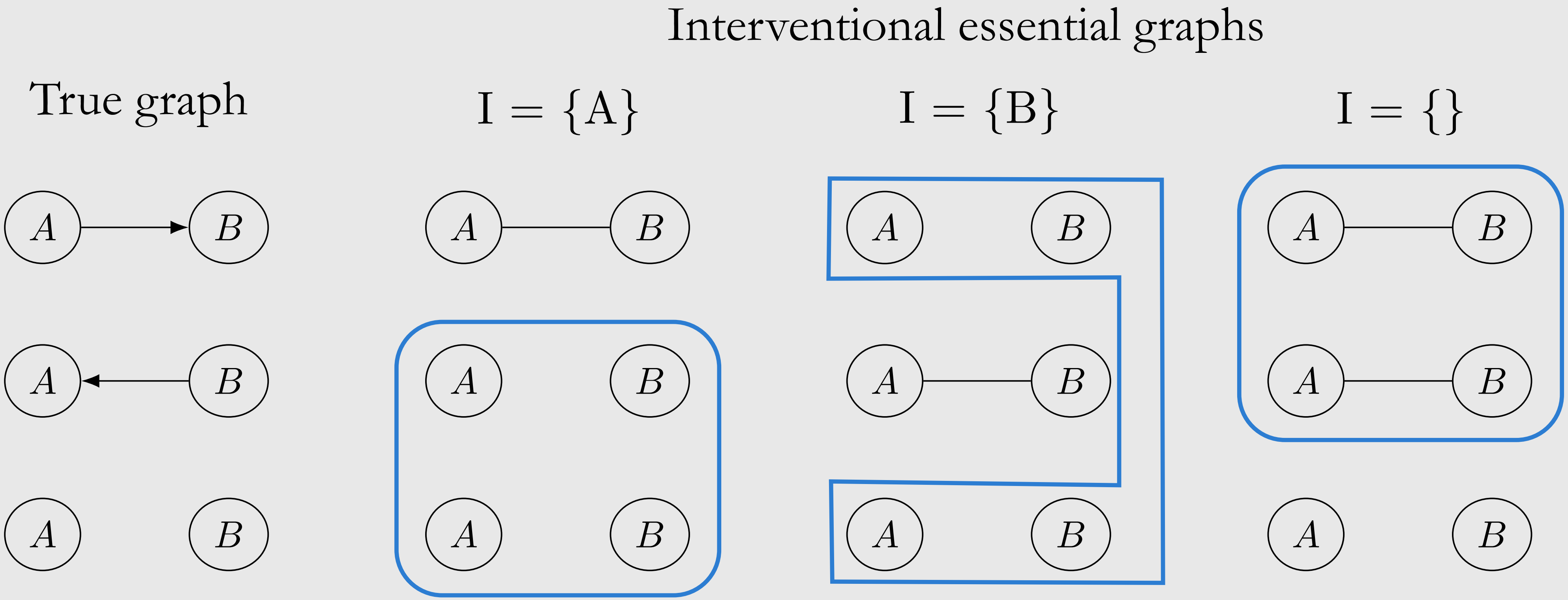


Two Variables: Interventional Essential Graphs

True graph under intervention I = {}

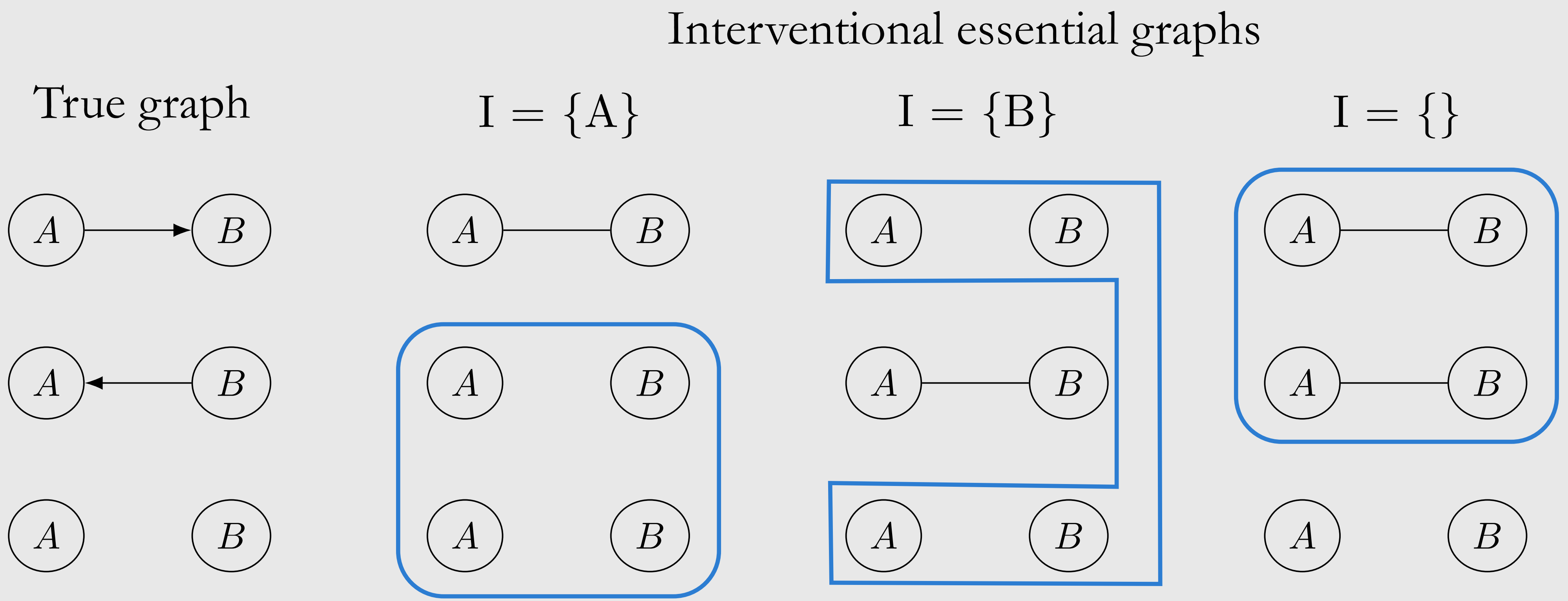


Two Variables: Interventional Essential Graphs



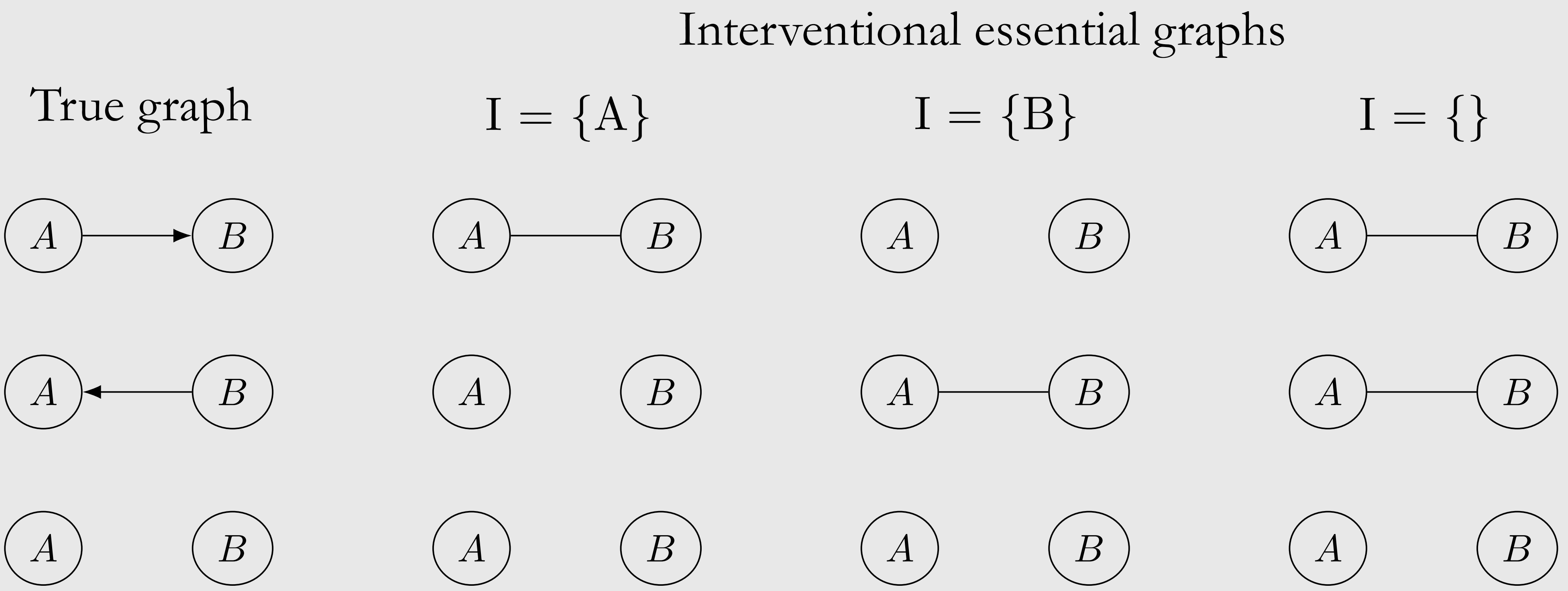
Two Variables: Interventional Essential Graphs

Need more than one intervention to identify the graph



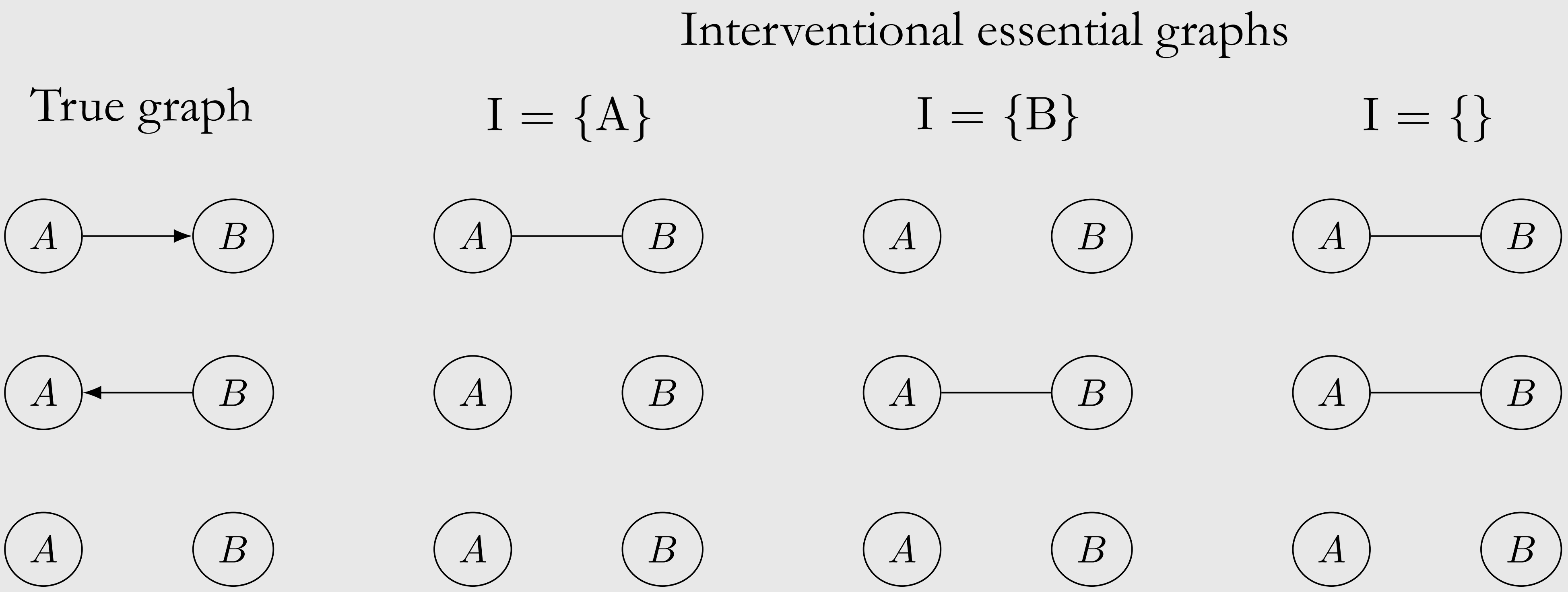
Two Variables: Two Interventions Identify the Graph

Two interventions are sufficient and necessary to identify the graph.



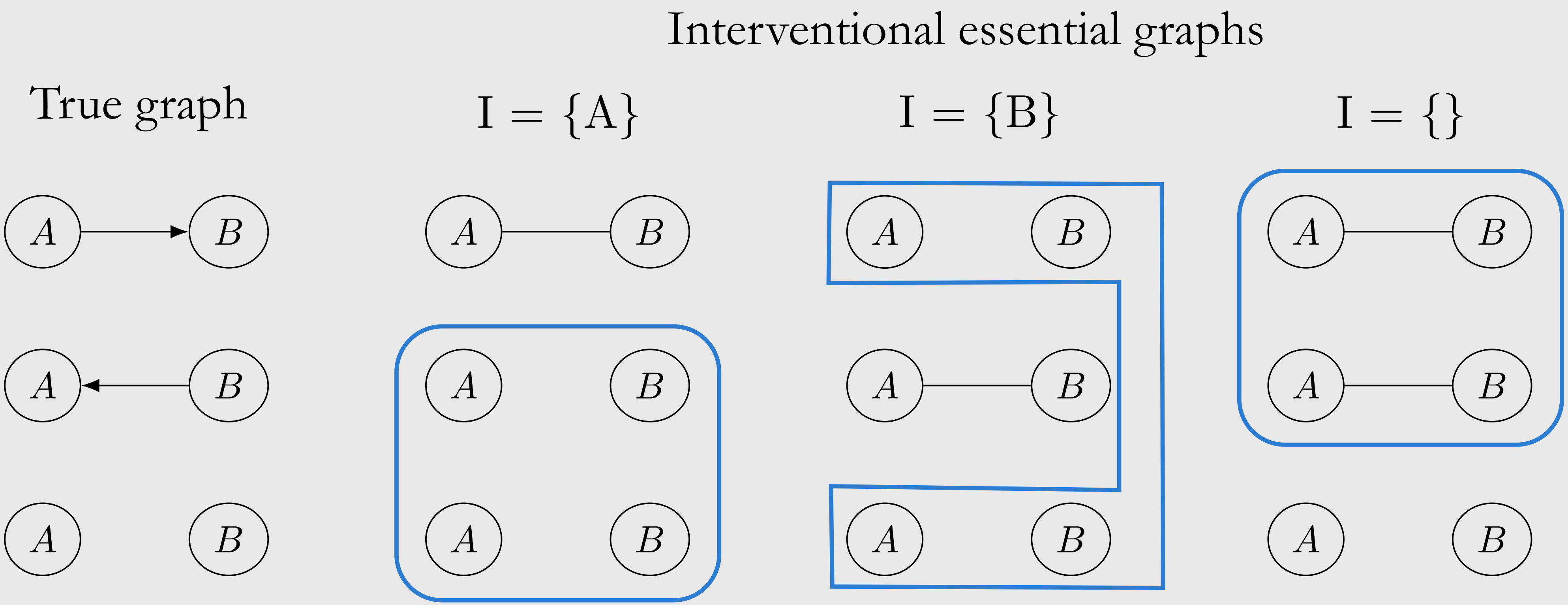
Two Variables: Two Interventions Identify the Graph

Two interventions are sufficient and necessary to identify the graph.



Two Variables: Two Interventions Identify the Graph

Two interventions are sufficient and necessary to identify the graph.



Complete Graphs Are the Worst Case

Complete Graphs Are the Worst Case

Theorem: Two graphs are Markov equivalent if and only if they have the same skeleton and same immoralities (Verma & Pearl, 1990; Frydenburg, 1990).

Complete Graphs Are the Worst Case

Theorem: Two graphs are Markov equivalent if and only if they have the same skeleton and same immoralities (Verma & Pearl, 1990; Frydenburg, 1990).

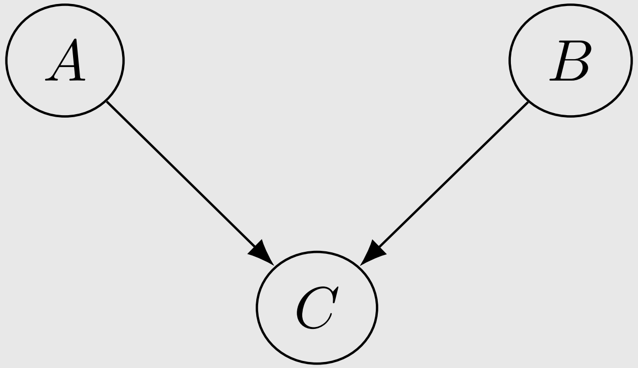
In complete graphs, there are no immoralities, so we can only get the complete skeleton graph (e.g. from PC) as the essential graph in the worse case

Complete Graphs Are the Worst Case

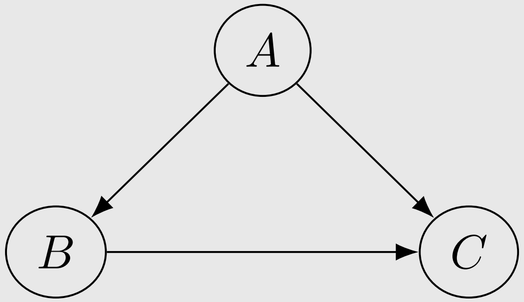
Theorem: Two graphs are Markov equivalent if and only if they have the same skeleton and same immoralities (Verma & Pearl, 1990; Frydenburg, 1990).

In complete graphs, there are no immoralities, so we can only get the complete skeleton graph (e.g. from PC) as the essential graph in the worse case

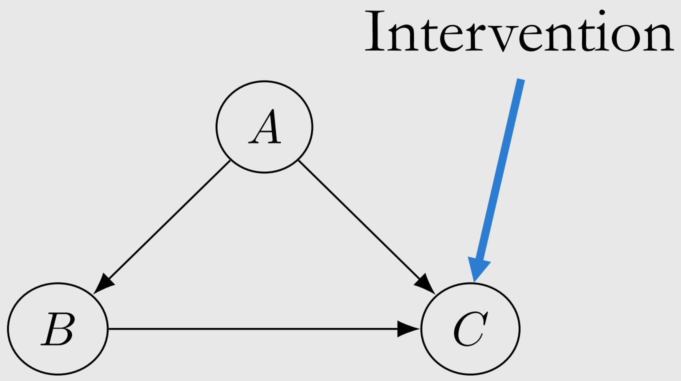
Immorality



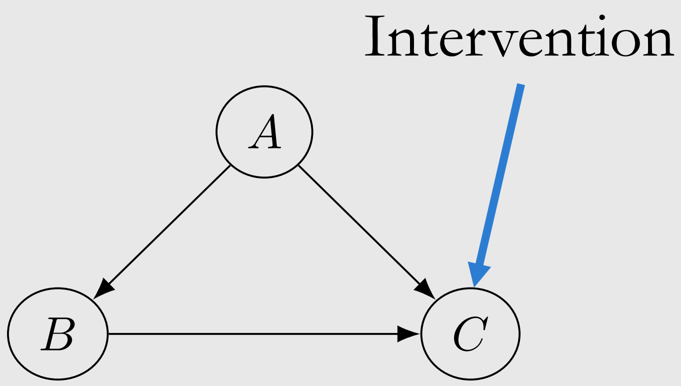
Three Variables

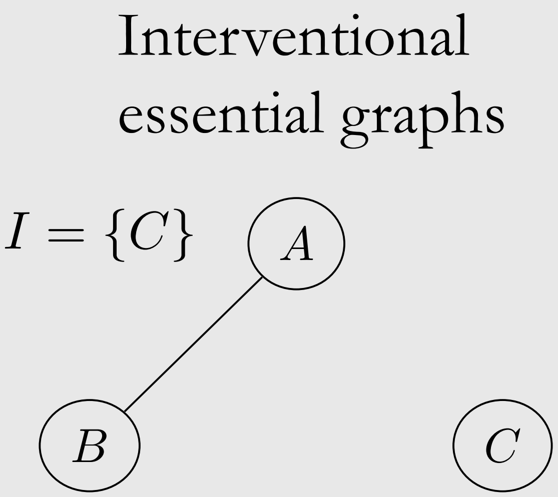


Three Variables

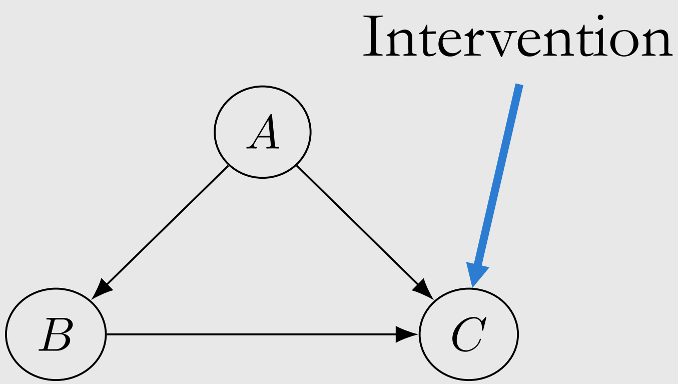


Three Variables

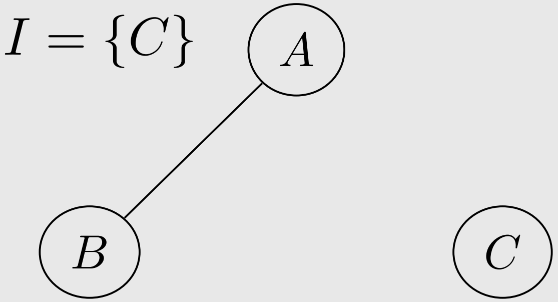




Three Variables



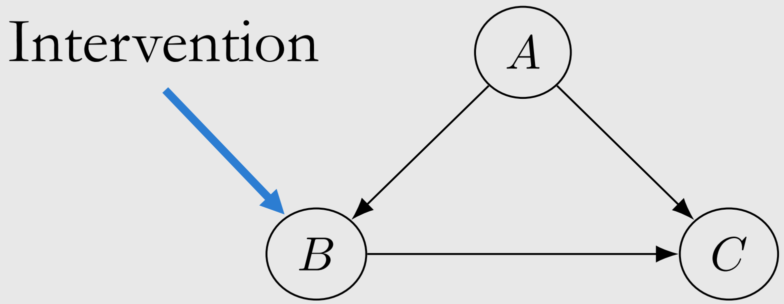
Interventional essential graphs



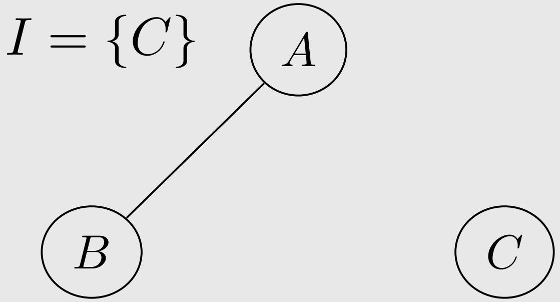
What we've learned:

- No C → A edge
- No C → B edge
- A and B connected

Three Variables



Interventional essential graphs

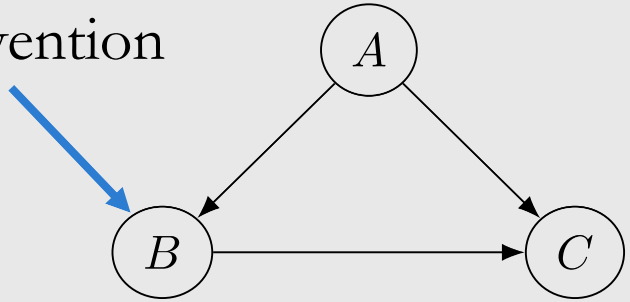


What we've learned:

- No C → A edge
- No C → B edge
- A and B connected

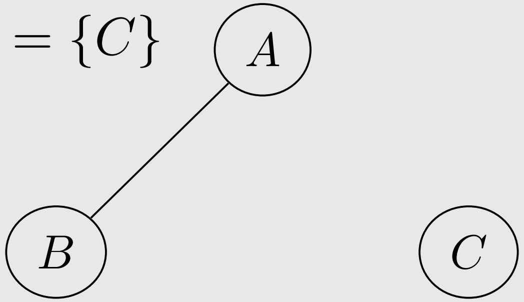
Three Variables

Intervention

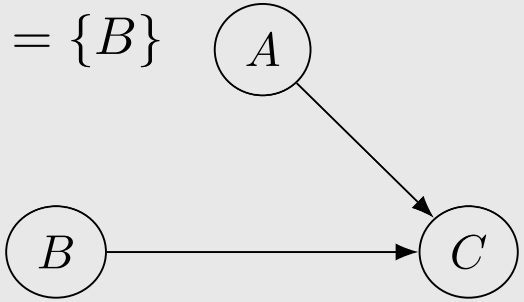


Interventional
essential graphs

$I = \{C\}$



$I = \{B\}$

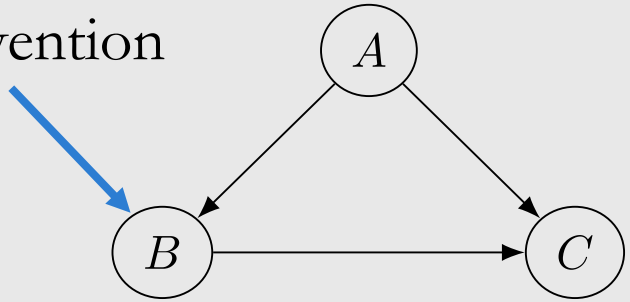


What we've learned:

- No C → A edge
- No C → B edge
- A and B connected

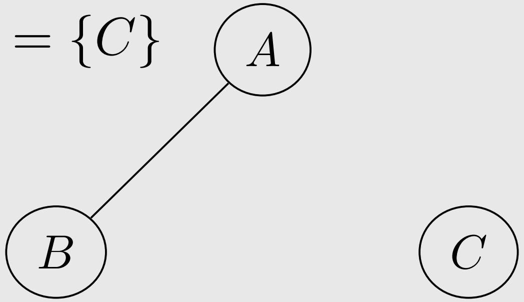
Three Variables

Intervention

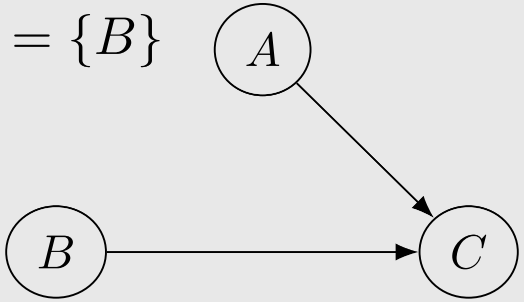


Interventional essential graphs

$I = \{C\}$



$I = \{B\}$



What we've learned:

- No C → A edge
- No C → B edge
- A and B connected
- No B → A edge
- Yes A → C edge
- Yes B → C edge

Single Variable Interventions: n – 1 Are Sufficient for n > 2 (Eberhardt et al., 2006)

Single Variable Interventions: n – 1 Are Sufficient for n > 2 (Eberhardt et al., 2006)

Intuition:

Single Variable Interventions: n – 1 Are Sufficient for n > 2 (Eberhardt et al., 2006)

Intuition:

1. First intervention identifies the adjacencies between the remaining n – 1 variables (skeleton of that subgraph).

Single Variable Interventions: n – 1 Are Sufficient for n > 2 (Eberhardt et al., 2006)

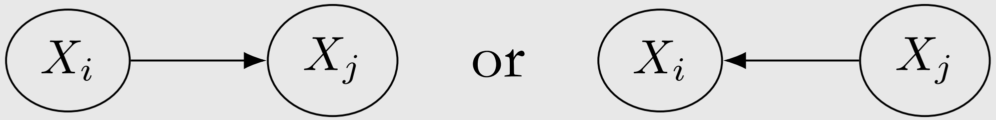
Intuition:

1. First intervention identifies the adjacencies between the remaining n – 1 variables (skeleton of that subgraph).
2. i^{th} intervention orients the edges incident on X_i

Single Variable Interventions: n – 1 Are Sufficient for n > 2 (Eberhardt et al., 2006)

Intuition:

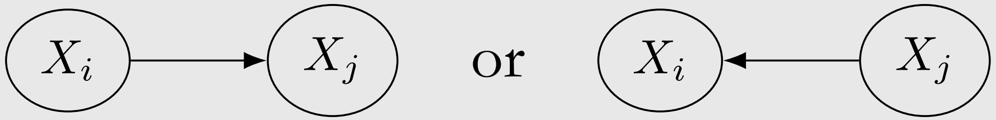
1. First intervention identifies the adjacencies between the remaining n – 1 variables (skeleton of that subgraph).
2. i^th intervention orients the edges incident on X_i



Single Variable Interventions: n – 1 Are Sufficient for n > 2 (Eberhardt et al., 2006)

Intuition:

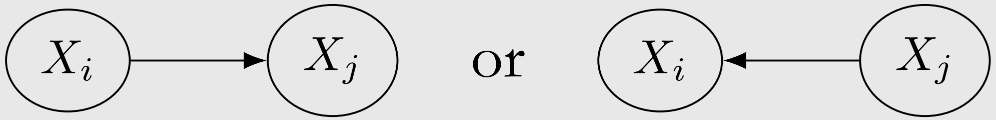
1. First intervention identifies the adjacencies between the remaining n – 1 variables (skeleton of that subgraph).
2. i^{th} intervention orients the edges incident on X_i
   1. If $X_i \perp\!\!\!\perp X_j$, then $X_i \leftarrow X_j$



Single Variable Interventions: n – 1 Are Sufficient for n > 2 (Eberhardt et al., 2006)

Intuition:

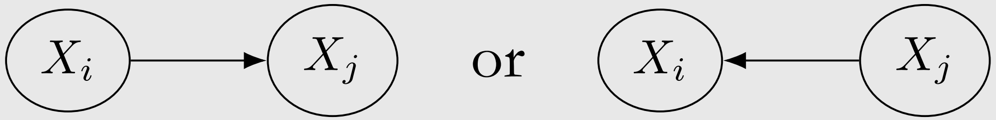
1. First intervention identifies the adjacencies between the remaining n – 1 variables (skeleton of that subgraph).
2. i^{th} intervention orients the edges incident on X_i
   1. If $X_i \perp\!\!\!\perp X_j$, then $X_i \leftarrow X_j$
   2. If $X_i \not\perp\!\!\!\perp X_j$, then $X_i \rightarrow X_j$



Single Variable Interventions: n – 1 Are Sufficient for n > 2 (Eberhardt et al., 2006)

Intuition:

1. First intervention identifies the adjacencies between the remaining n – 1 variables (skeleton of that subgraph).
2. i^{th} intervention orients the edges incident on X_i
   1. If $X_i \perp\!\!\!\perp X_j$, then $X_i \leftarrow X_j$
   2. If $X_i \not\perp\!\!\!\perp X_j$, then $X_i \rightarrow X_j$

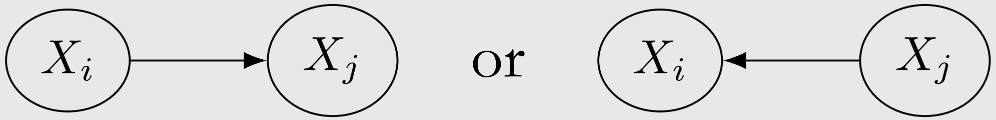


3. No intervention necessary on X_n as all those edges would have been oriented in step 2.

Single Variable Interventions: n – 1 Are Sufficient for n > 2 (Eberhardt et al., 2006)

Intuition:

1. First intervention identifies the adjacencies between the remaining n – 1 variables (skeleton of that subgraph).
2. i^{th} intervention orients the edges incident on X_i
    1. If $X_i \perp\!\!\!\perp X_j$, then $X_i \leftarrow X_j$
    2. If $X_i \not\perp\!\!\!\perp X_j$, then $X_i \rightarrow X_j$



3. No intervention necessary on X_n as all those edges would have been oriented in step 2.

Note: if you include the empty set (observational), then n are necessary in the worst case (Eberhardt et al., 2006).

Single Variable Interventions: n – 1 Are Necessary in the Worst Case (Complete Graph)

Single Variable Interventions: n – 1 Are Necessary in the Worst Case (Complete Graph)

Intuition (Eberhardt et al., 2006):

Single Variable Interventions: n – 1 Are Necessary in the Worst Case (Complete Graph)

Intuition (Eberhardt et al., 2006):

1. Choose arbitrary intervention ordering

Single Variable Interventions: n – 1 Are Necessary in the Worst Case (Complete Graph)

Intuition (Eberhardt et al., 2006):

1. Choose arbitrary intervention ordering
2. n – 2 interventions leave the $X_{n-1} - X_n$ edge undirected

Single Variable Interventions: n – 1 Are Necessary in the Worst Case (Complete Graph)

Intuition (Eberhardt et al., 2006):

1. Choose arbitrary intervention ordering
2. $n - 2$ interventions leave the $X_{n-1} - X_n$ edge undirected
3. Intervention $n - 1$ on X_{n-1} or X_n necessary to direct this final edge

Single Variable Interventions: n – 1 Are Necessary in the Worst Case (Complete Graph)

Intuition (Eberhardt et al., 2006):

1. Choose arbitrary intervention ordering
2. n – 2 interventions leave the $X_{n-1} - X_n$ edge undirected
3. Intervention n – 1 on X_{n-1} or X_n necessary to direct this final edge

Adaptivity doesn't help in the worst case (Eberhardt et al., 2006).

Questions (where the goal is to identify the graph):

1. Show that n interventions are necessary in the worse case in the three-variable setting when you use the observational data (null intervention) as one of the interventions.
2. What is the minimum number of interventions necessary in the worst case when n = 3?
3. What is the minimum number of interventions necessary in the worst case when n = 2?

Structural Interventions

Single-Node Interventions

Multi-Node Interventions

Parametric Interventions

Interventional Markov Equivalence

Miscellaneous Other Settings

Multiple-Node Interventions

Multiple-Node Interventions

Single-node interventions (Eberhardt et al., 2006):

- n – 1 are sufficient
- n – 1 are necessary in the worst case

Multiple-Node Interventions

Single-node interventions (Eberhardt et al., 2006):

- n – 1 are sufficient
- n – 1 are necessary in the worst case

Multi-node interventions with no restrictions on the number of nodes per intervention:

Multiple-Node Interventions

Single-node interventions (Eberhardt et al., 2006):

- n – 1 are sufficient
- n – 1 are necessary in the worst case

Multi-node interventions with no restrictions on the number of nodes per intervention:

- $\lfloor \log_2(n) \rfloor + 1$ interventions are sufficient (Eberhardt et al., 2005)
-

Multiple-Node Interventions

Single-node interventions (Eberhardt et al., 2006):

- n – 1 are sufficient
- n – 1 are necessary in the worst case

Multi-node interventions with no restrictions on the number of nodes per intervention:

- $\lfloor \log_2(n) \rfloor + 1$ interventions are sufficient (Eberhardt et al., 2005)
- $\lfloor \log_2(n) \rfloor + 1$ interventions are necessary in the worst case (Eberhardt et al., 2005)

Number of Interventions in the Worse Case

Eberhardt et al. (2005) found that $\lfloor \log_2(n) \rfloor + 1$ multi-node interventions are necessary for worst case graphs (complete graph), but what about for non-complete graphs?

Number of Interventions in the Worse Case

Eberhardt et al. (2005) found that $\lfloor \log_2(n) \rfloor + 1$ multi-node interventions are necessary for worst case graphs (complete graph), but what about for non-complete graphs?

Start with Markov equivalence class, then how many interventions are necessary, given that we can intervene on unlimited nodes per intervention?

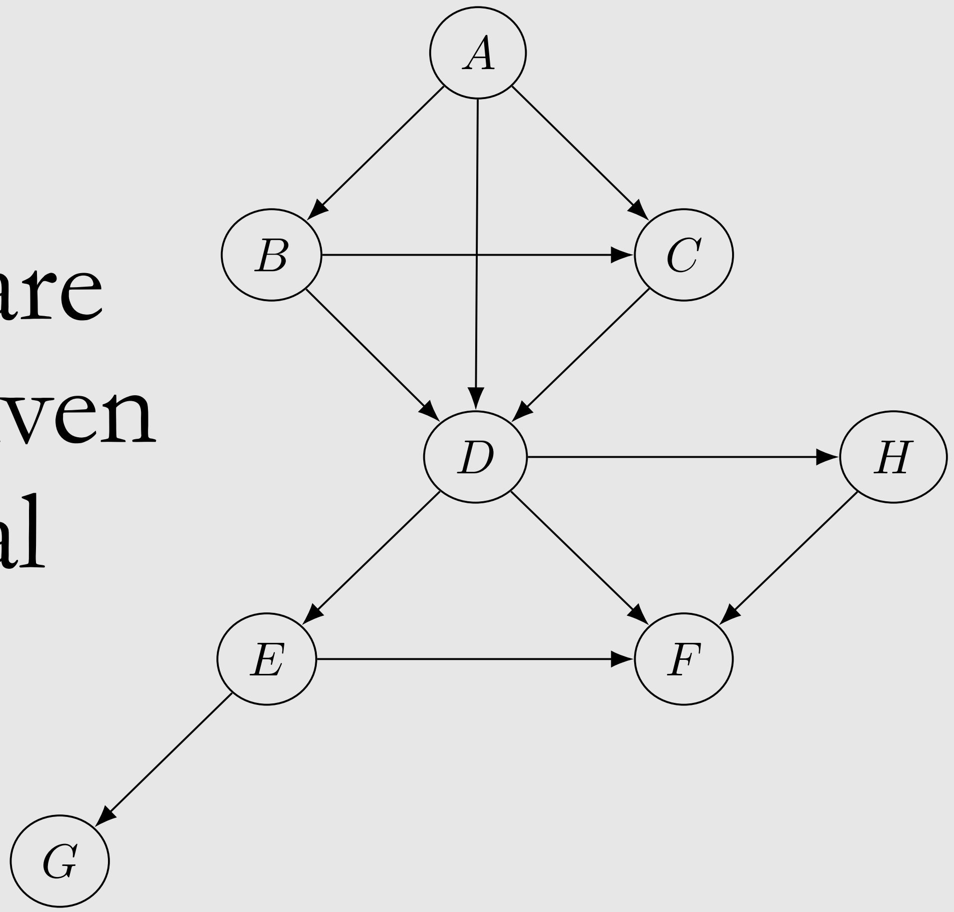
Number of Interventions in the Worse Case

Eberhardt et al. (2005) found that $\lfloor \log_2(n) \rfloor + 1$ multi-node interventions are necessary for worst case graphs (complete graph), but what about for non-complete graphs?

Start with Markov equivalence class, then how many interventions are necessary, given that we can intervene on unlimited nodes per intervention?

Theorem: $\lceil \log_2(c) \rceil$ multi-node interventions are sufficient and necessary in the worst case, where c is the size of the largest clique (conjectured by Eberhardt (2008) and proven by Hauser & Bühlmann (2014)).

Question:
In this graph, how many multi-node interventions are sufficient and necessary given that you know the essential graph?



Structural Interventions

Single-Node Interventions

Multi-Node Interventions

Parametric Interventions

Interventional Markov Equivalence

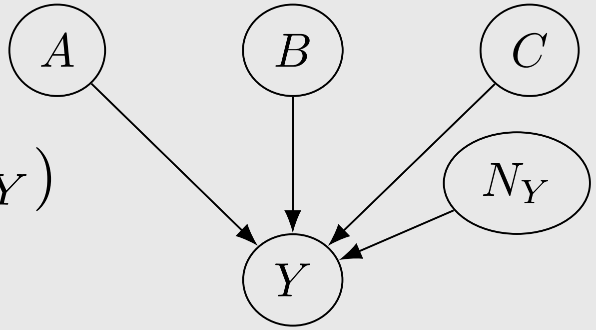
Miscellaneous Other Settings

Structural vs. Parametric Interventions

Structural vs. Parametric Interventions

Structural:

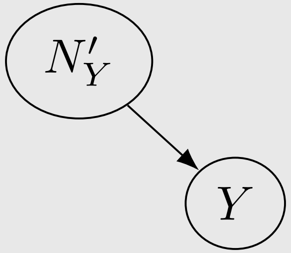
$$Y := f_\theta(A, B, C, N_Y)$$



Structural vs. Parametric Interventions

Structural:

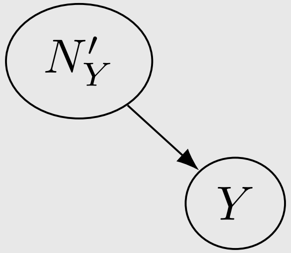
$$Y := N'_Y$$



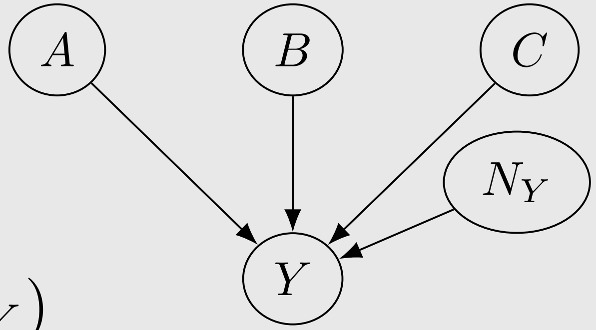
Structural vs. Parametric Interventions

Structural:

$$Y := N'_Y$$



Parametric:

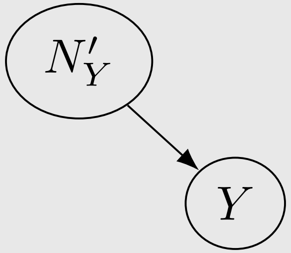


$$Y := f_\theta(A, B, C, N_Y)$$

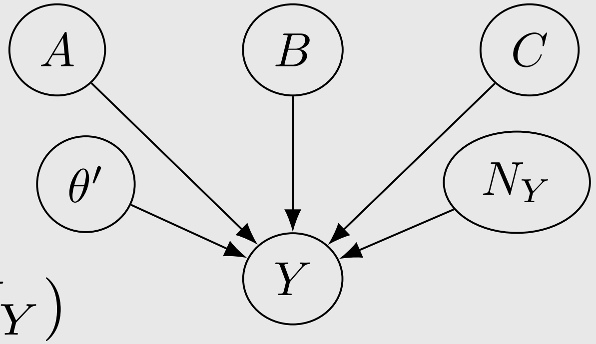
Structural vs. Parametric Interventions

Structural:

$$Y := N'_Y$$



Parametric:

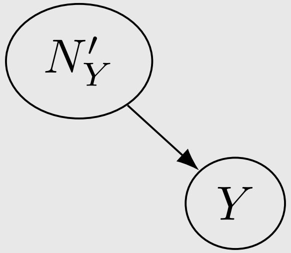


$$Y := f_{\theta'}(A, B, C, N_Y)$$

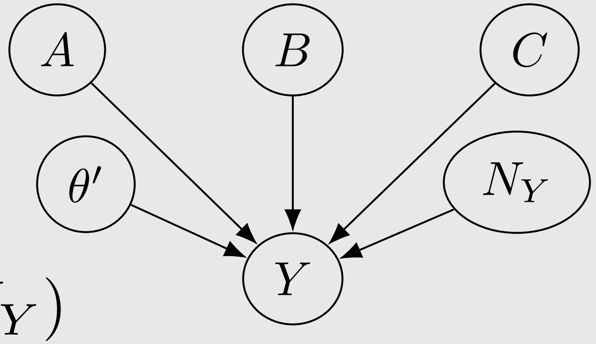
Structural vs. Parametric Interventions

Structural:

$Y := N'_Y$



Parametric:



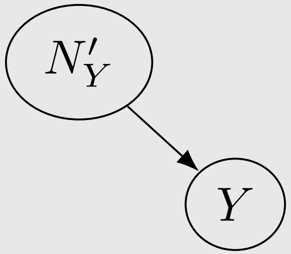
$Y := f_{\theta'}(A, B, C, N_Y)$

Other names: hard vs. soft, perfect vs. imperfect, etc.

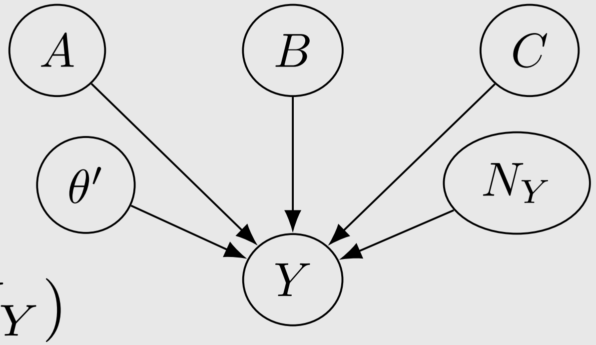
Structural vs. Parametric Interventions

Structural:

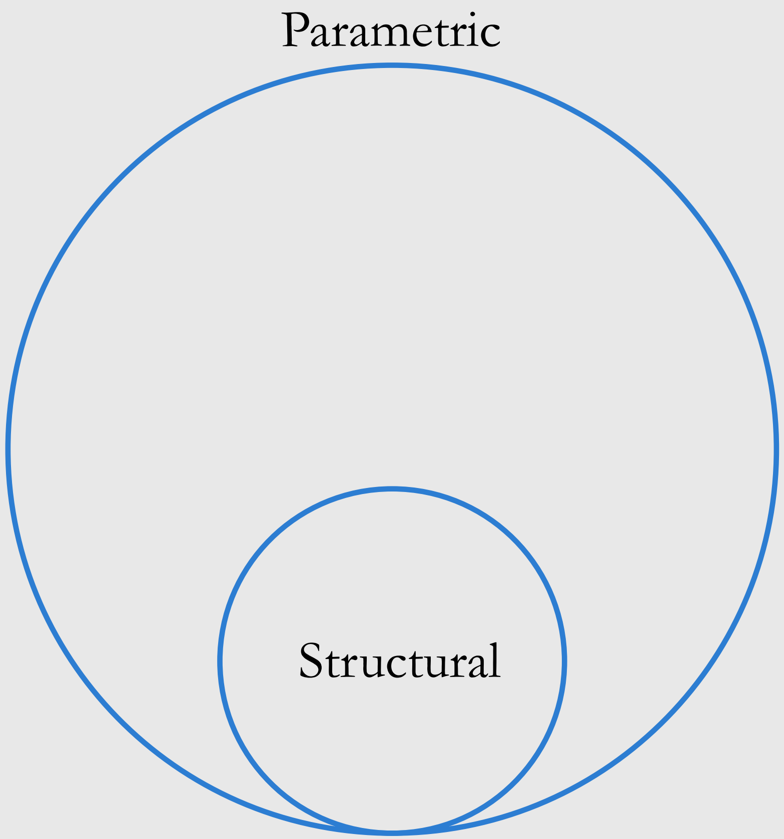
$$Y := N'_Y$$



Parametric:



$$Y := f_{\theta'}(A, B, C, N_Y)$$

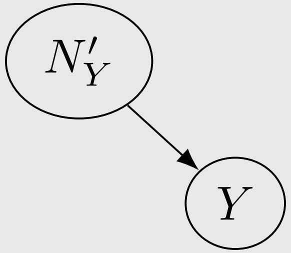


Other names: hard vs. soft, perfect vs. imperfect, etc.

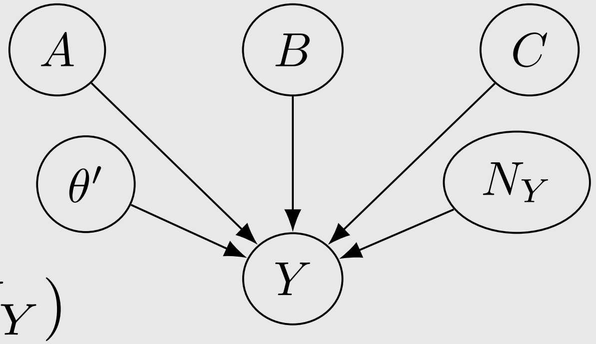
Structural vs. Parametric Interventions

Structural:

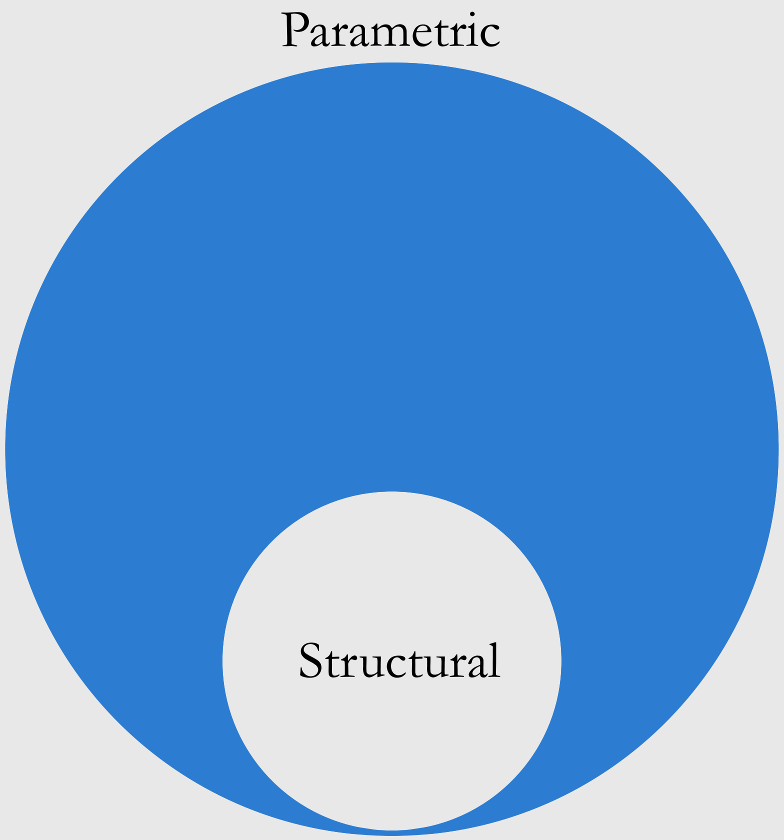
$$Y := N'_Y$$



Parametric:



$$Y := f_{\theta'}(A, B, C, N_Y)$$

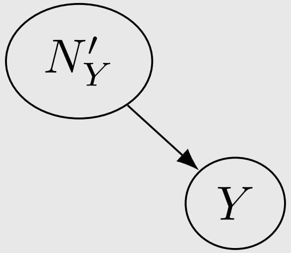


Other names: hard vs. soft, perfect vs. imperfect, etc.

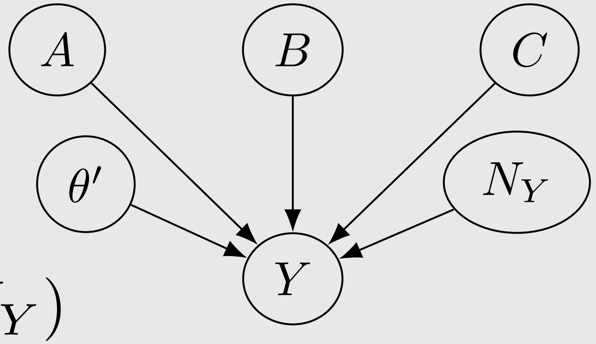
Structural vs. Parametric Interventions

Structural:

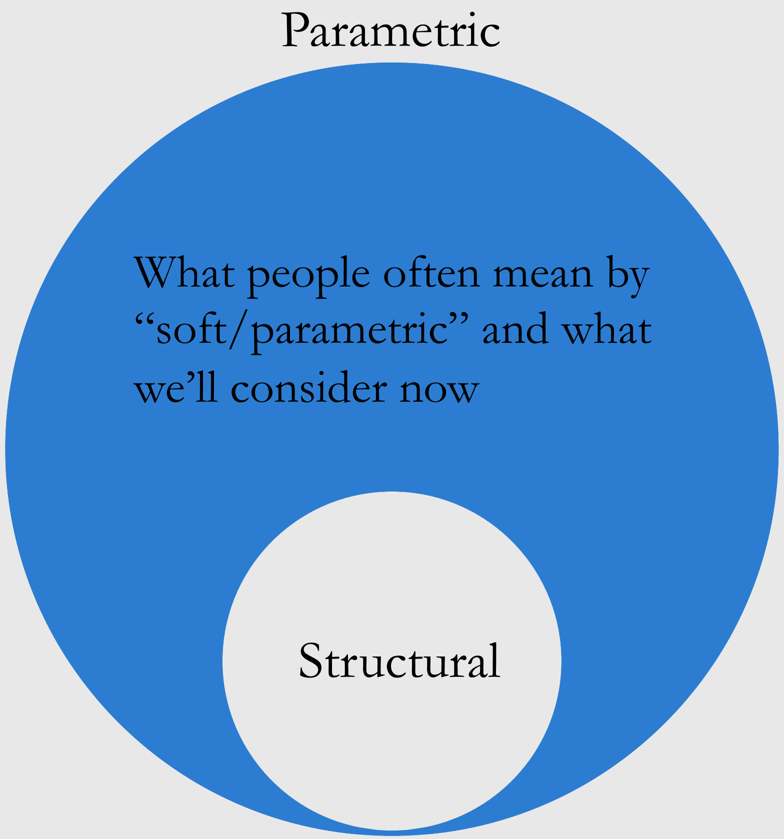
$Y := N'_Y$



Parametric:



$Y := f_{\theta'}(A, B, C, N_Y)$



Other names: hard vs. soft, perfect vs. imperfect, etc.

Number of Parametric Single-Node Interventions

Number of Parametric Single-Node Interventions

Number of interventions for identification (Eberhardt & Scheines, 2007):

Number of Parametric Single-Node Interventions

Number of interventions for identification (Eberhardt & Scheines, 2007):

- n – 1 interventions are sufficient

Number of Parametric Single-Node Interventions

Number of interventions for identification (Eberhardt & Scheines, 2007):

- n – 1 interventions are sufficient
- n – 1 interventions are necessary in the worst case

Number of Parametric Single-Node Interventions

Number of interventions for identification (Eberhardt & Scheines, 2007):

- n – 1 interventions are sufficient
- n – 1 interventions are necessary in the worst case

(same as with structural interventions)

Partial Identification with Fewer Interventions

Partial Identification with Fewer Interventions

Can identify up to the Markov equivalence classes with only observational data (no interventions) and no semi-parametric assumptions

Partial Identification with Fewer Interventions

Can identify up to the Markov equivalence classes with only observational data (no interventions) and no semi-parametric assumptions

Can identify the exact causal graphs with n – 1 single-variable interventions

Partial Identification with Fewer Interventions

Can identify up to the Markov equivalence classes with only observational data (no interventions) and no semi-parametric assumptions

Can identify the exact causal graphs with n – 1 single-variable interventions

How about the in-between? How much of the graph can we identify with fewer interventions?

How much of the graph can we identify with a given set of interventions?

Structural Interventions

Single-Node Interventions

Multi-Node Interventions

Parametric Interventions

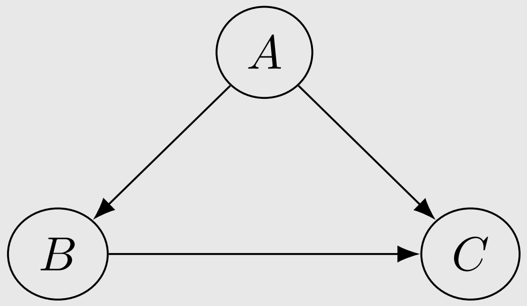
Interventional Markov Equivalence

Miscellaneous Other Settings

Interventions Introduce Immoralities: Single-Node

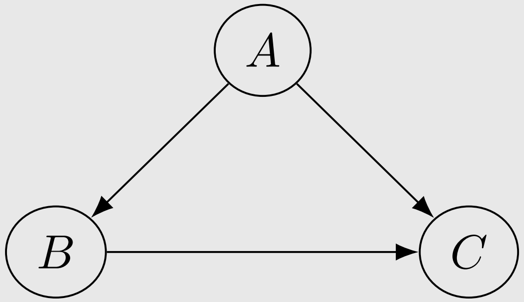
Interventions Introduce Immoralities: Single-Node

True Graph

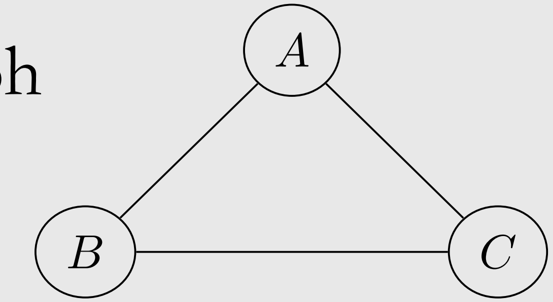


Interventions Introduce Immoralities: Single-Node

True Graph



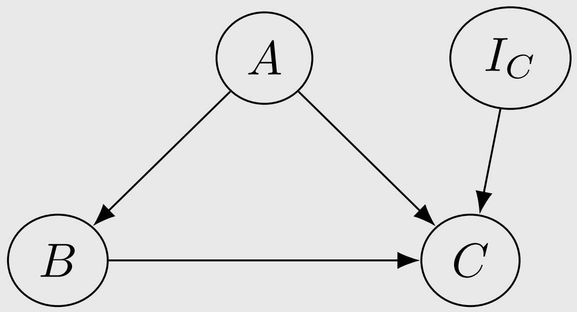
Essential Graph



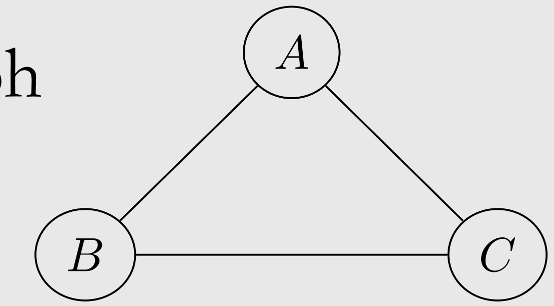
Interventions Introduce Immoralities: Single-Node

Interventional Graph

True Graph



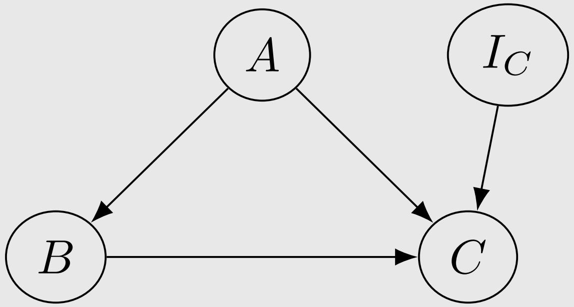
Essential Graph



Interventions Introduce Immoralities: Single-Node

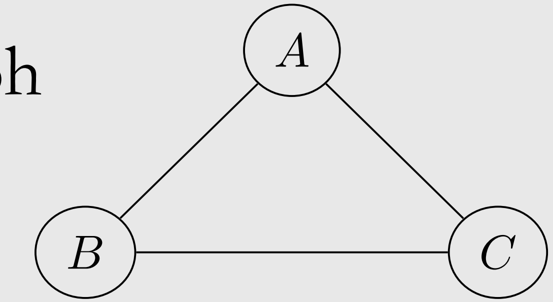
Interventional Graph

True Graph



Preserves structure, unlike structural interventions

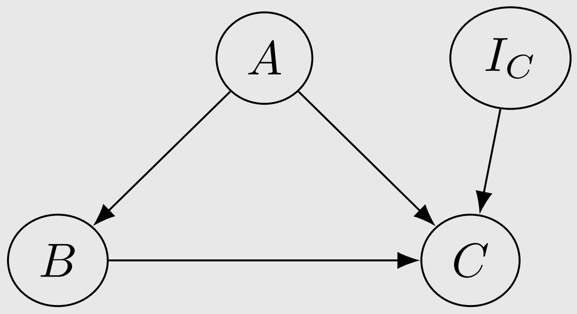
Essential Graph



Interventions Introduce Immoralities: Single-Node

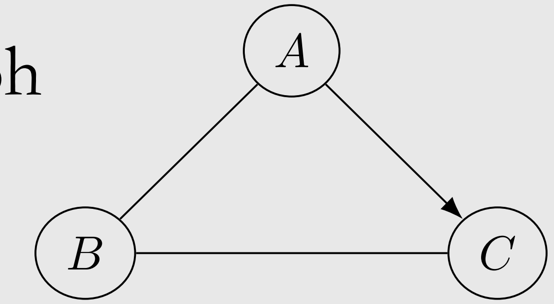
Interventional Graph

True Graph



Preserves structure, unlike structural interventions

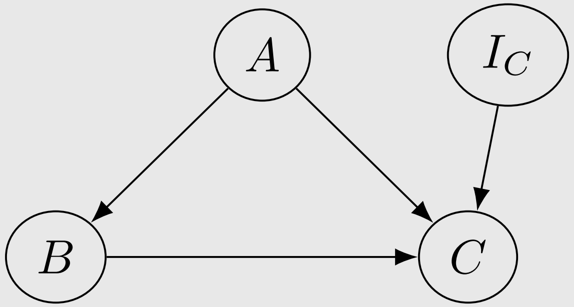
Essential Graph



Interventions Introduce Immoralities: Single-Node

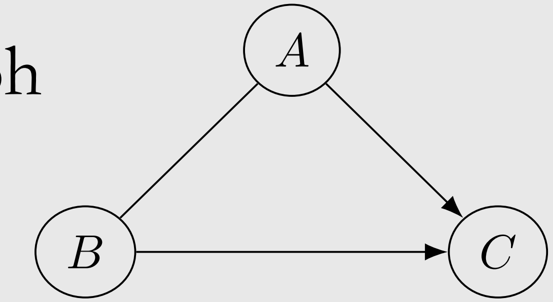
Interventional Graph

True Graph



Preserves structure, unlike structural interventions

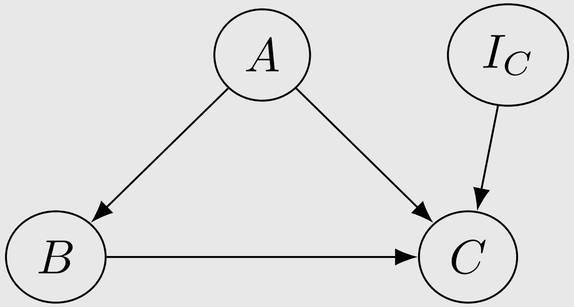
Essential Graph



Interventions Introduce Immoralities: Single-Node

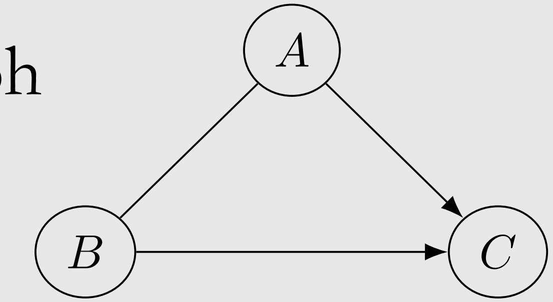
Interventional Graph

True Graph



Preserves structure, unlike structural interventions

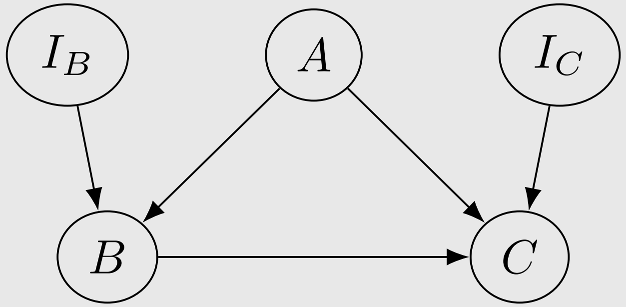
Interventional Essential Graph



Interventions Introduce Immoralities: Single-Node

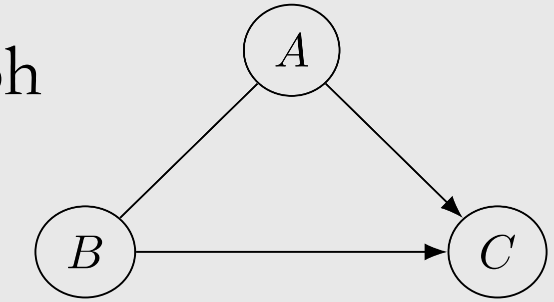
Interventional Graph

True Graph



Preserves structure, unlike structural interventions

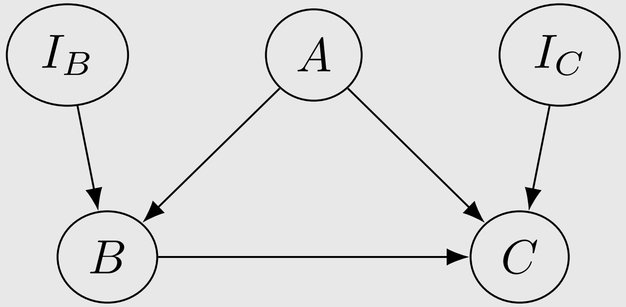
Interventional Essential Graph



Interventions Introduce Immoralities: Single-Node

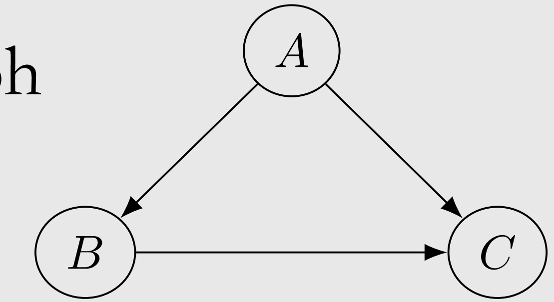
Interventional Graph

True Graph



Preserves structure, unlike structural interventions

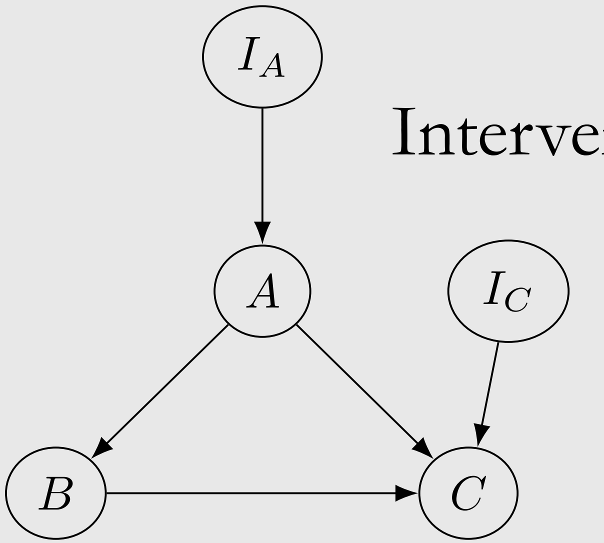
Interventional Essential Graph



Interventions Introduce Immoralities: Single-Node

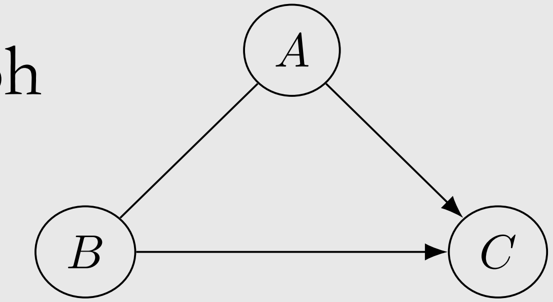
True Graph

Interventional Graph



Preserves structure, unlike structural interventions

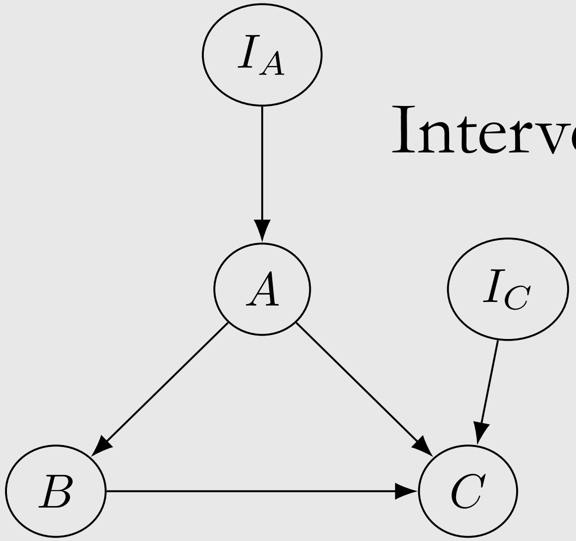
Interventional Essential Graph



Interventions Introduce Immoralities: Single-Node

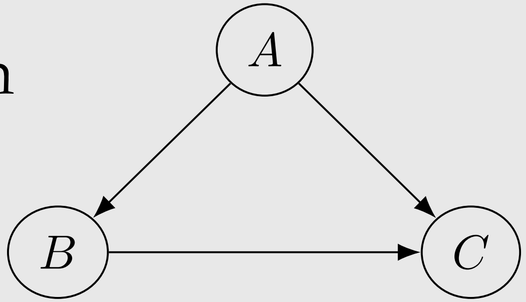
Interventional Graph

True Graph



Preserves structure, unlike structural interventions

Interventional Essential Graph



Interventional Markov Equivalence: Single-Node

Theorem: Two graphs augmented with single-node interventions are interventionally Markov equivalent if any only if they have the same skeletons and immoralities (Tian & Pearl, 2001).

Interventional Markov Equivalence: Single-Node

Theorem: Two graphs augmented with single-node interventions are interventionally Markov equivalent if any only if they have the same skeletons and immoralities (Tian & Pearl, 2001).

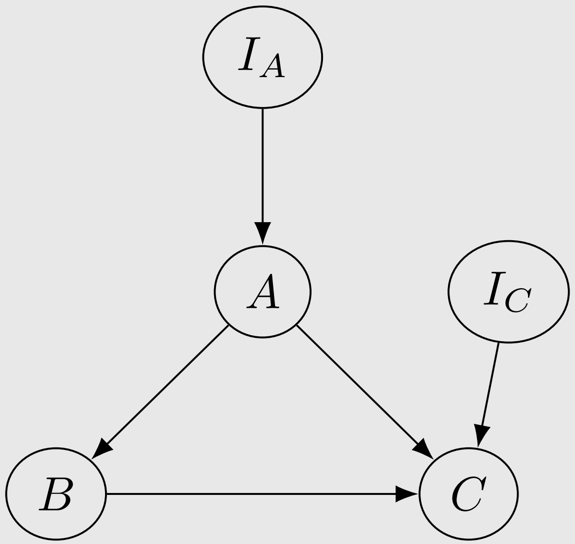
Number of interventions for identification (Eberhardt & Scheines, 2007):

- n – 1 interventions are sufficient
- n – 1 interventions are necessary in the worst case

(same as with structural interventions)

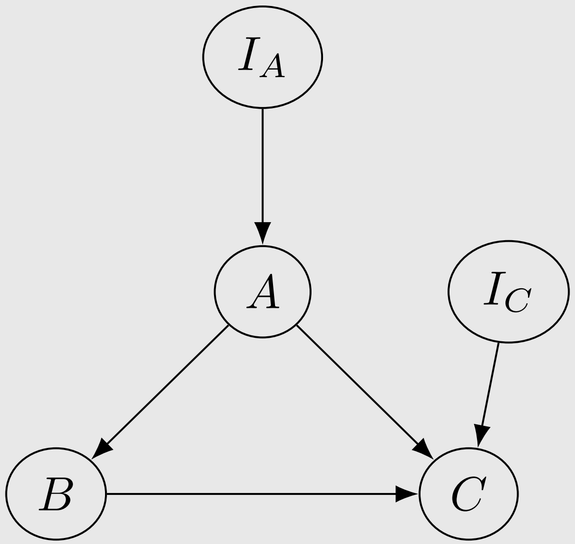
Interventional Graph: Multi-Node Interventions

Two Single-Node Interventions

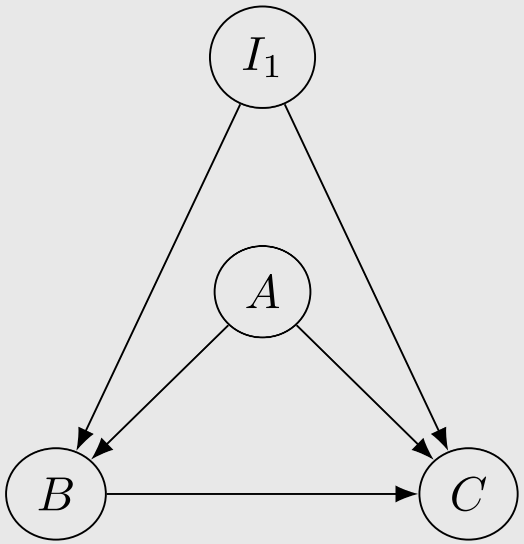


Interventional Graph: Multi-Node Interventions

Two Single-Node Interventions

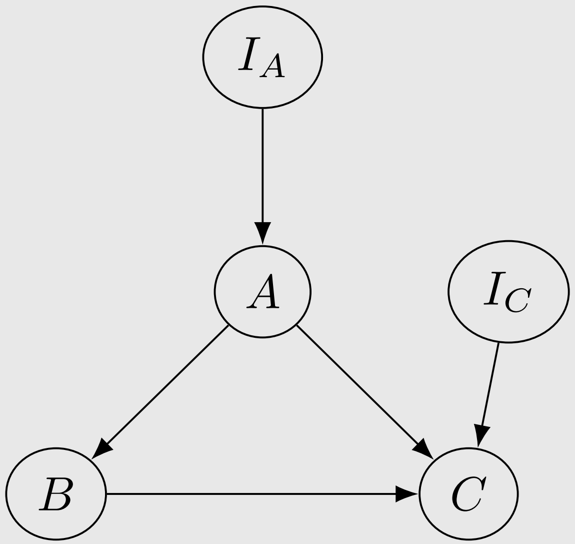


One Multi-Node Intervention

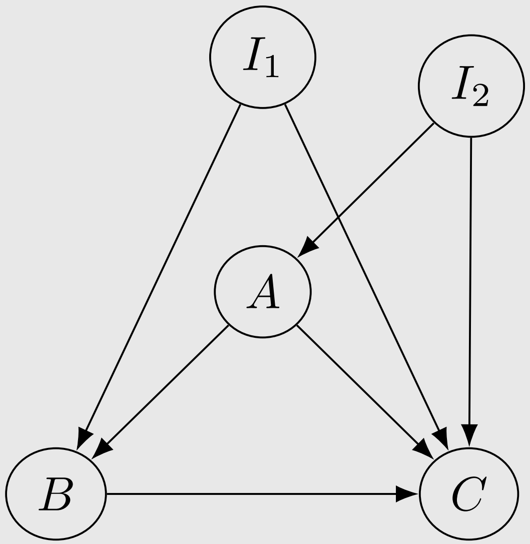


Interventional Graph: Multi-Node Interventions

Two Single-Node Interventions



Two Multi-Node Interventions

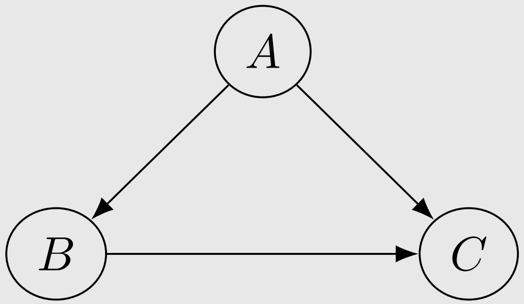


Interventional Markov Equivalence: Multi-Node

Theorem: Given the observational data, two graphs augmented with multi-node interventions are interventionally Markov equivalent if and only if they have the same skeletons and immoralities (Yang et al., 2018).

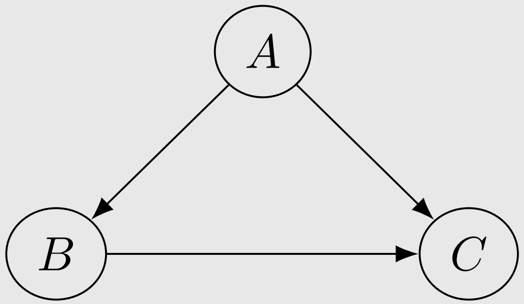
Interventional Markov Equivalence: Multi-Node

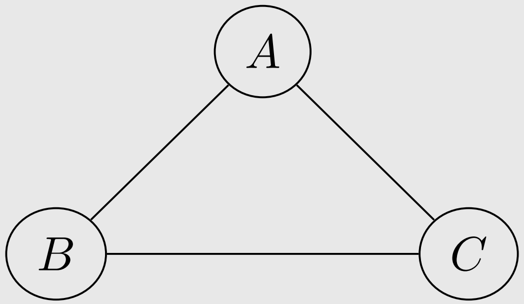
Theorem: Given the observational data, two graphs augmented with multi-node interventions are interventionally Markov equivalent if and only if they have the same skeletons and immoralities (Yang et al., 2018).



Interventional Markov Equivalence: Multi-Node

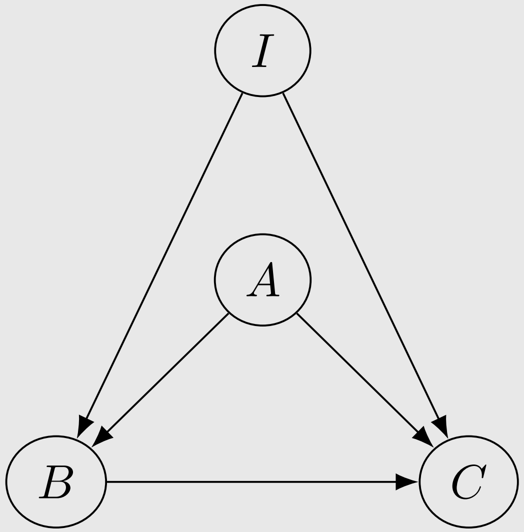
Theorem: Given the observational data, two graphs augmented with multi-node interventions are interventionally Markov equivalent if and only if they have the same skeletons and immoralities (Yang et al., 2018).

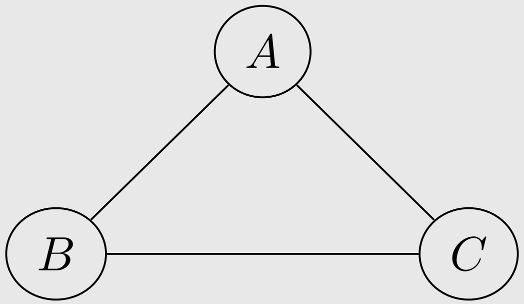




Interventional Markov Equivalence: Multi-Node

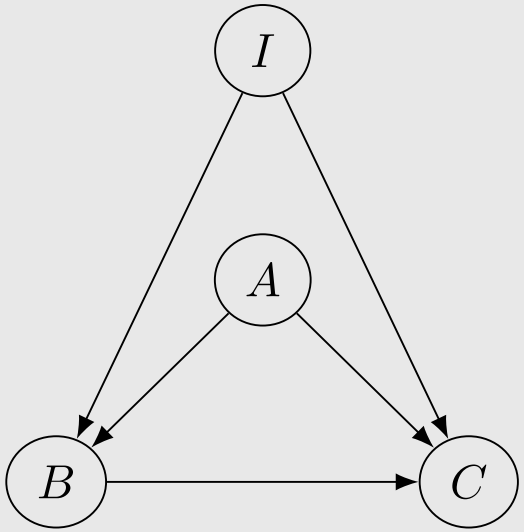
Theorem: Given the observational data, two graphs augmented with multi-node interventions are interventionally Markov equivalent if and only if they have the same skeletons and immoralities (Yang et al., 2018).

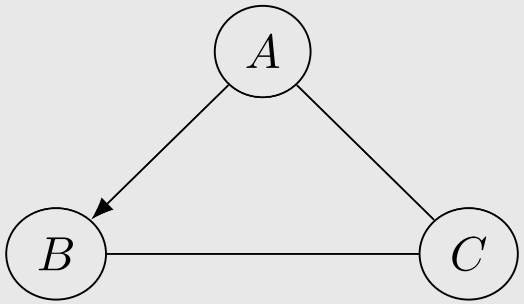




Interventional Markov Equivalence: Multi-Node

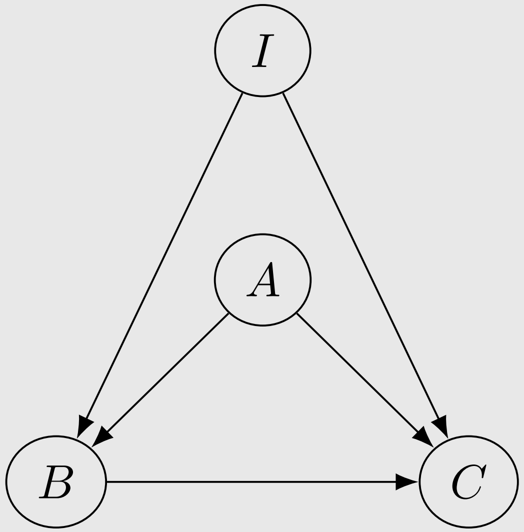
Theorem: Given the observational data, two graphs augmented with multi-node interventions are interventionally Markov equivalent if and only if they have the same skeletons and immoralities (Yang et al., 2018).

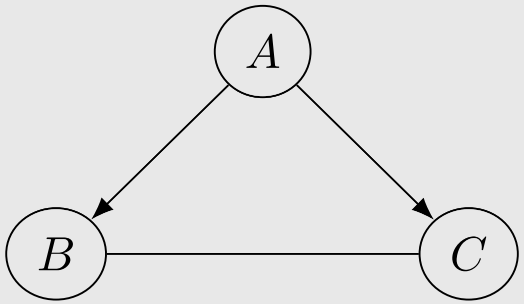




Interventional Markov Equivalence: Multi-Node

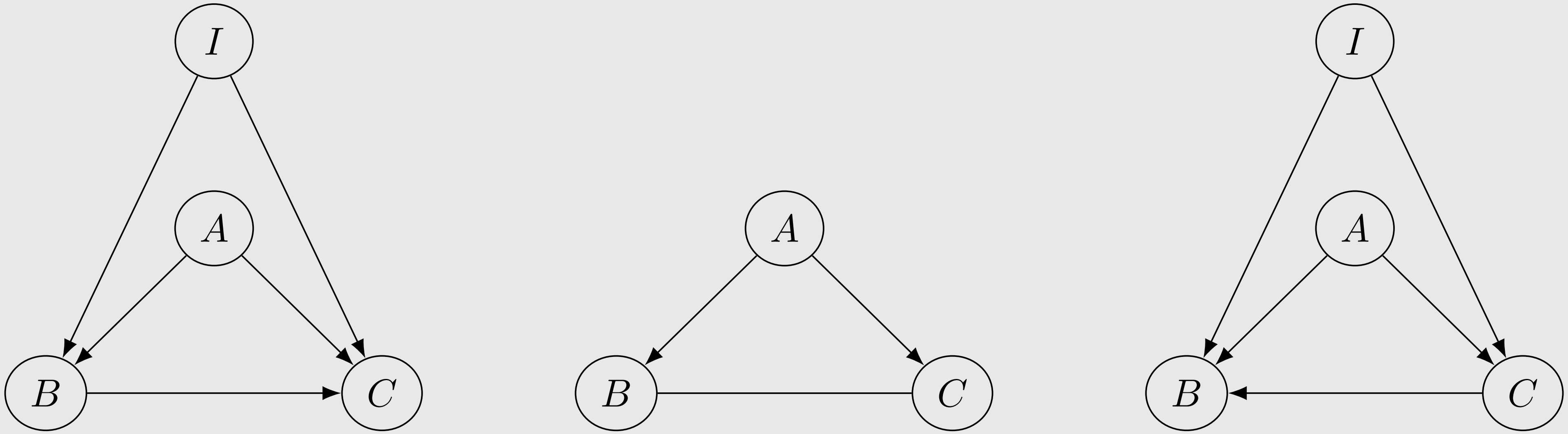
Theorem: Given the observational data, two graphs augmented with multi-node interventions are interventionally Markov equivalent if and only if they have the same skeletons and immoralities (Yang et al., 2018).





Interventional Markov Equivalence: Multi-Node

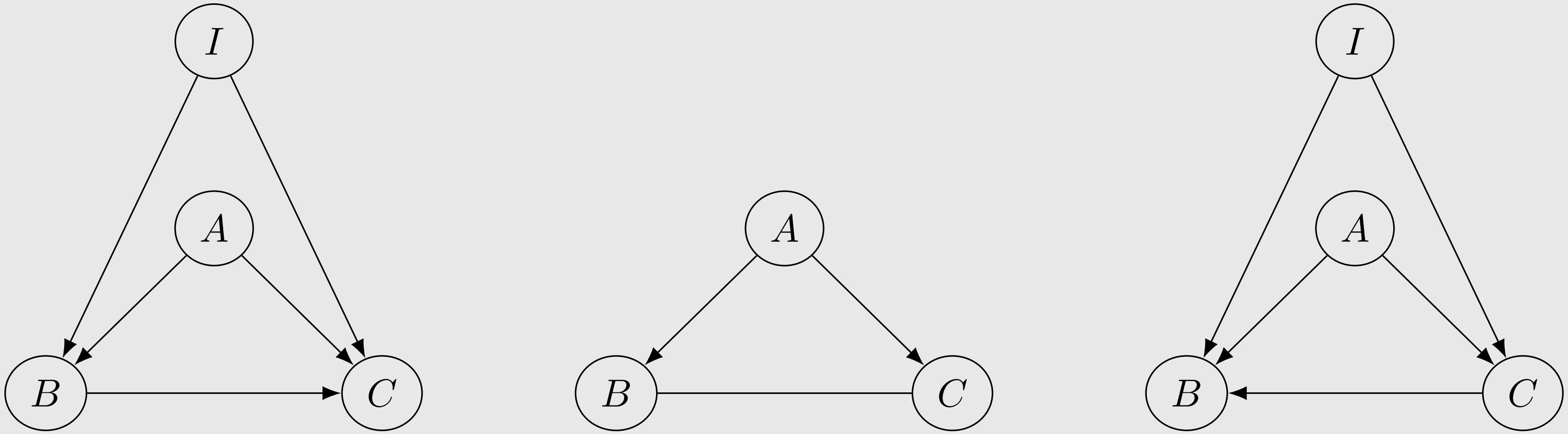
Theorem: Given the observational data, two graphs augmented with multi-node interventions are interventionally Markov equivalent if and only if they have the same skeletons and immoralities (Yang et al., 2018).



Interventional Markov Equivalence: Multi-Node

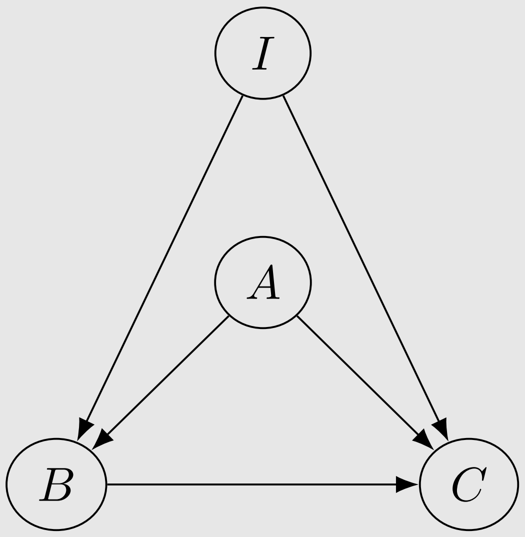
Theorem: Given the observational data, two graphs augmented with multi-node interventions are interventionally Markov equivalent if and only if they have the same skeletons and immoralities (Yang et al., 2018).

Structural/perfect analog from Hauser & Bühlmann (2012)



Questions

1. How many parametric single-node interventions are necessary and sufficient for identifying a directed acyclic graph?
2. What is the essential graph of the graph on the right (ignoring the intervention node)?
3. What is the interventional essential graph?
4. What is a graph that this graph is interventionally Markov equivalent to?



Structural Interventions

Single-Node Interventions

Multi-Node Interventions

Parametric Interventions

Interventional Markov Equivalence

Miscellaneous Other Settings

Miscellaneous Other Settings

Miscellaneous Other Settings

- Randomized algorithms

Miscellaneous Other Settings

- Randomized algorithms
  - Only need O(log log n) interventions with high probability (Hu et al., 2014)
  - Better than O(log c) from Eberhardt (2008) and Hauser & Bühlmann (2014)

Miscellaneous Other Settings

- Randomized algorithms
  - Only need O(log log n) interventions with high probability (Hu et al., 2014)
  - Better than O(log c) from Eberhardt (2008) and Hauser & Bühlmann (2014)
- Intervene on at most k variables per intervention: Shanmugam et al. (2015)

Miscellaneous Other Settings

- Randomized algorithms
  - Only need O(log log n) interventions with high probability (Hu et al., 2014)
  - Better than O(log c) from Eberhardt (2008) and Hauser & Bühlmann (2014)
- Intervene on at most k variables per intervention: Shanmugam et al. (2015)
  - Worst case: $\sim \frac{n}{2k}$

Miscellaneous Other Settings

- Randomized algorithms
  - Only need O(log log n) interventions with high probability (Hu et al., 2014)
  - Better than O(log c) from Eberhardt (2008) and Hauser & Bühlmann (2014)
- Intervene on at most k variables per intervention: Shanmugam et al. (2015)
  - Worst case: $\sim \frac{n}{2k}$
  - Randomized: $O\left(\frac{n}{k} \log \log k\right)$

Miscellaneous Other Settings

- Randomized algorithms
  - Only need O(log log n) interventions with high probability (Hu et al., 2014)
  - Better than O(log c) from Eberhardt (2008) and Hauser & Bühlmann (2014)
- Intervene on at most k variables per intervention: Shanmugam et al. (2015)
  - Worst case: $\sim \frac{n}{2k}$
  - Randomized: $O\left(\frac{n}{k} \log \log k\right)$
- Only k interventions: Ghassami et al. (2018)

Miscellaneous Other Settings

- Randomized algorithms
  - Only need O(log log n) interventions with high probability (Hu et al., 2014)
  - Better than O(log c) from Eberhardt (2008) and Hauser & Bühlmann (2014)
- Intervene on at most k variables per intervention: Shanmugam et al. (2015)
  - Worst case: $\sim \frac{n}{2k}$
  - Randomized: $O\left(\frac{n}{k}\log\log k\right)$
- Only k interventions: Ghassami et al. (2018)
- Unobserved confounding: Kocaoglu et al. (2017)