On the Completeness of Causal Discovery in the Presence of Latent Confounding with Tiered Background Knowledge

Bryan Andrews
University of Pittsburgh

Peter Spirtes
Carnegie Mellon University

Gregory F. Cooper
University of Pittsburgh

5 SUPPLEMENT

5.1 Additional Background

Definition *inducing path*: Let X and Y be two variables in an ancestral graph and $\boldsymbol{L}$ be a set of variables not containing X or Y. A path π between X and Y is called an inducing path relative to $\boldsymbol{L}$ if every non-endpoint on π is in $\boldsymbol{L}$ or a collider on π, and every collider on π is an ancestor of X or Y. If $\boldsymbol{L} = \emptyset$, then π is called a *primitive* inducing path.

Proposition 1. (Richardson and Spirtes, 2002) An ancestral graph is maximal if and only if there are no primitive inducing paths between any two non-adjacent variables.

Definition *discriminating path*: In a MAG, a path $\pi = \langle X, \ldots, Q, B, Y \rangle$ between X and Y is a discriminating path for B if,

(i) π includes at least three edges;

(ii) B is a non-endpoint of π and adjacent to Y on π;

(iii) X is not adjacent to Y, and every variable between X and B on π is a collider on π and a parent of Y.

Definition *collider with order*: Let $\mathcal{D}_i$ $(i \geq 0)$ be the set of triple of order i in a MAG $\mathcal{G}$, defined recursively as follows:

Order 0: A triple $\langle A, B, C \rangle \in \mathcal{D}_0$ if X and Z are not adjacent.

Order i+1: A triple $\langle A, B, C \rangle \in \mathcal{D}_{i+1}$ if

(i) for all $j < i+1$, $\langle A, B, C \rangle \notin \mathcal{D}_j$, and

(ii) there is a discriminating path $\langle X, Q_1, \ldots, Q_p, B, Y \rangle$ for B with $\langle A, B, C \rangle = \langle Q_p, B, Y \rangle$ or $\langle Y, B, Q_p \rangle$ and the p colliders

$$\langle X, Q_1, Q_2 \rangle, \ldots, \langle Q_{p-1}, Q_p, B \rangle \in \bigcup_{j \leq i} \mathcal{D}_j.$$

If a triple $\langle A, B, C \rangle \in \mathcal{D}_i$ is a collider, then the triple is a collider with order i.

Proposition 2. (Ali et al., 2009) Two MAGs (over the same variables) are Markov equivalent if the share they same adjacencies and colliders with order.

Definition *partial mixed graph* A partial mixed graph is a vertex edge graph that can contain four kinds of edges: $\{\rightarrow, \leftrightarrow, \circ\!\!-\!\!\circ, \circ\!\!\rightarrow\}$.

Definition *uncovered path*: In a PMG, a path $\pi = \langle X_1, \ldots, X_k \rangle$ is said to be uncovered if for every $2 \leq i \leq k-1$, X_{i-1} and X_{i+1} are not adjacent, that is, if every consecutive triple on the path is unshielded. **Definition** *potentially directed path*: In a PMG, a path $\pi = \langle X_1, \ldots, X_k \rangle$ is said to be potentially directed from X_1 to X_k if for every $1 \leq i \leq k-1$, the edge between X_i and X_{i+1} is not into X_i or out of X_{i+1}.

5.2 Orientation Rules

FCI uses a set of orientation rules (Zhang, 2008). Rules $\mathcal{R}_5$ - $\mathcal{R}_7$ have to do with selection and are omitted, the other rules are listed below. An asterisks edge mark is used to denote that we are agnostic to the actual edge mark of an edge (it may be either a tail, arrowhead, or circle).

$\mathcal{R}_0$ For each unshielded triple $\langle \alpha, \gamma, \beta \rangle$ in $\mathcal{P}$, orient it as a collider $\alpha {*\!\!\rightarrow} \gamma {\leftarrow\!\!*} \beta$ if and only if γ is not in $sepset(\alpha, \beta)$.

$\mathcal{R}_1$ If $\alpha {*\!\!\rightarrow} \beta \circ\!\!-\!\!* \gamma$, and α and γ are not adjacent, then orient the triple as $\alpha {*\!\!\rightarrow} \beta \rightarrow \gamma$.

$\mathcal{R}_2$ If $\alpha \rightarrow \beta {*\!\!\rightarrow} \gamma$ or $\alpha {*\!\!\rightarrow} \beta \rightarrow \gamma$, and $\alpha *\!\!-\!\!\circ \gamma$, then orient $\alpha *\!\!-\!\!\circ \gamma$ as $\alpha {*\!\!\rightarrow} \gamma$.

$\mathcal{R}_3$ If $\alpha {*\!\!\rightarrow} \beta {\leftarrow\!\!*} \gamma$, $\alpha *\!\!-\!\!\circ \theta \circ\!\!-\!\!* \gamma$, α and γ are not adjacent, and $\theta *\!\!-\!\!\circ \beta$, then orient $\theta *\!\!-\!\!\circ \beta$ as $\theta {*\!\!\rightarrow} \beta$.

$\mathcal{R}_4$ If $\pi = \langle \theta, \ldots, \alpha, \beta, \gamma \rangle$ is a discriminating path between θ and γ for β, and $\beta \circ\!\!-\!\!* \gamma$; then if $\beta \in sepset(\theta, \gamma)$, orient $\beta \circ\!\!-\!\!* \gamma$ as $\beta \rightarrow \gamma$; otherwise orient the triple $\langle \alpha, \beta, \gamma \rangle$ as $\alpha \leftrightarrow \beta \leftrightarrow \gamma$.

$\mathcal{R}_8$ If $\alpha \rightarrow \beta \rightarrow \gamma$ or $\alpha -\!\!\circ \beta \rightarrow \gamma$, and $\alpha \circ\!\!\rightarrow \gamma$, orient $\alpha \circ\!\!\rightarrow \gamma$ as $\alpha \rightarrow \gamma$.

$\mathcal{R}_9$ If $\alpha \circ\!\!\rightarrow \gamma$, and $\pi = \langle \alpha, \beta, \theta, \ldots, \gamma \rangle$ is an uncovered potentially directed path from α to γ such that γ and β are not adjacent, then orient $\alpha \circ\!\!\rightarrow \gamma$ as $\alpha \rightarrow \gamma$.

$\mathcal{R}_{10}$ Suppose $\alpha \circ\!\!\rightarrow \gamma$, $\beta \rightarrow \gamma \leftarrow \theta$, π_1 is an uncovered potentially directed path from α to β, and π_2 is an uncovered potentially directed path from α to θ. Let μ be the vertex adjacent to α on π_1 (μ could be β), and ω be the vertex adjacent to α on π_2 (ω could be θ). If μ and ω are distinct and not adjacent, then orient $\alpha \circ\!\!\rightarrow \gamma$ as $\alpha \rightarrow \gamma$.

5.3 Proofs

Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG where the variables may be partitioned into three distinct subsets $\boldsymbol{O} = \boldsymbol{A} \cup \boldsymbol{B} \cup \boldsymbol{W}$. Lemma 1 proves the necessary and sufficient graphical constraint for $\mathcal{G}$ to satisfy the exogenous background knowledge $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$.

Lemma 1. Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG and $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{W}$ be three disjoint sets of variables that partition $\boldsymbol{O}$. $\mathcal{G}$ violates $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$ if and only if there exists $A \in \boldsymbol{A}$ and $B \in \boldsymbol{B}$ such that $A \leftrightarrow B$ or $B \rightarrow A$ in $\mathcal{G}$.

($\Rightarrow$) If $\mathcal{G}$ violates $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$ then there exists $A \in \boldsymbol{A}$ and $B \in \boldsymbol{B}$ such that $A \leftrightarrow B$ or $B \rightarrow A$ in $\mathcal{G}$.

If $\mathcal{G}$ violates $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$, then there exists $A \in \boldsymbol{A}$ and $B \in \boldsymbol{B}$ such that A and B are adjacent and A is not an ancestor of B. Therefore, $A \leftrightarrow B$ or $B \rightarrow A$ in $\mathcal{G}$.

($\Leftarrow$) If there exists $A \in \boldsymbol{A}$ and $B \in \boldsymbol{B}$ such that $A \leftrightarrow B$ or $B \rightarrow A$ in $\mathcal{G}$ then $\mathcal{G}$ violates $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$.

If there exists $A \in \boldsymbol{A}$ and $B \in \boldsymbol{B}$ such that $B \rightarrow A$ or $A \leftrightarrow B$ in $\mathcal{G}$ then A and B are adjacent and A is not an ancestor of B. Therefore, $\mathcal{G}$ violates $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$. □

Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG where the variables may be partitioned into three distinct subsets $\boldsymbol{O} = \boldsymbol{A} \cup \boldsymbol{B} \cup \boldsymbol{W}$. Lemma 2 proves that if $\mathcal{G}$ satisfies $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$, then the graph constructed by removing from $\mathcal{G}$ the variables in $\boldsymbol{W}$ and the edges connecting two members of $\boldsymbol{A} \cup \boldsymbol{W}$ is also a MAG satisfying $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$.

Lemma 2. Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG and $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{W}$ be three disjoint sets of variables that partition $\boldsymbol{O}$. If $\mathcal{G}$ satisfies $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$, then $\mathcal{H} = \textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})$ is a MAG satisfying $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$.

Note that $\mathcal{H}$ may be constructed from $\mathcal{G}$ by removing variables and edges.

Assume that $\mathcal{H}$ is not a mixed graph. Then $\mathcal{H}$ must contain an edge not in $\{\rightarrow, \leftrightarrow\}$ or more than one edge between a pair of adjacent variables. Since $\mathcal{H}$ may be constructed from $\mathcal{G}$ by removing variables and edges, any edge in $\mathcal{H}$ must also be in $\mathcal{G}$. But then $\mathcal{G}$ must also contain an edge not in $\{\rightarrow, \leftrightarrow\}$ or more than one edge between a pair of adjacent variables. This is a contradiction since $\mathcal{G}$ is a mixed graph. Therefore, $\mathcal{H}$ is a mixed graph.

Similarly, assume that $\mathcal{H}$ is not ancestral. Then $\mathcal{H}$ contains a directed or an almost directed cycle. Since $\mathcal{H}$ may be constructed from $\mathcal{G}$ by removing variables and edges, any path in $\mathcal{H}$ must also be in $\mathcal{G}$. But then $\mathcal{G}$ must contain a directed cycle or an almost directed cycle. This is a contradiction since $\mathcal{G}$ is ancestral. Therefore, $\mathcal{H}$ is ancestral.

Lastly, assume that $\mathcal{H}$ is not maximal. Then, by Proposition 1, a primitive inducing path π exists between two non-adjacent variables $X, Y \in \boldsymbol{O}$ in $\mathcal{H}$. Since $\mathcal{H}$ may be constructed from $\mathcal{G}$ by removing variables and edges, any path in $\mathcal{H}$ must also be in $\mathcal{G}$. However, $\mathcal{G}$ is maximal so, by Proposition 1, X and Y are adjacent in $\mathcal{G}$. Thus, the edge between X and Y must have been removed by $\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\cdot)$. Accordingly, $X, Y \in \boldsymbol{A}$ since $\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\cdot)$ removes all of the edges between any two members of $\boldsymbol{A} \cup \boldsymbol{W}$ and the variables in $\boldsymbol{W}$. Additionally, π must contain $B \in \boldsymbol{B}$ since every edge in $\mathcal{H}$ includes a member of $\boldsymbol{B}$. Therefore, by the definition of inducing path, B is an ancestor of X or Y. But then there exists $B' \in \boldsymbol{B}$ such that B' is a parent of X or Y in $\mathcal{G}$. Note that it is possible that $B' = B$, but not possible that $B' = X$ or $B' = Y$ since $X, Y \in \boldsymbol{A}$. This is a contradiction since $\mathcal{G}$ satisfies $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$. Therefore, $\mathcal{H}$ is maximal.

Since $\mathcal{G}$ satisfies $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$, by Lemma 1, there does not exist $A \in \boldsymbol{A}$ and $B \in \boldsymbol{B}$ such that $A \leftrightarrow B$ or $B \rightarrow A$ in $\mathcal{G}$. Since $\mathcal{H}$ may be constructed from $\mathcal{G}$ by removing variables and edges, there does not exist $A \in \boldsymbol{A}$ and $B \in \boldsymbol{B}$ such that $A \leftrightarrow B$ or $B \rightarrow A$ in $\mathcal{H}$. Therefore, $\mathcal{H}$ satisfies $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$. □

Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG where the variables may be partitioned into two distinct subsets $\boldsymbol{O} = \boldsymbol{A} \cup \boldsymbol{B}$. Lemma 3 proves that if $\mathcal{G}$ satisfies $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$, then the graph constructed by adding to $\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})$ the variables in $\boldsymbol{W} = \{W_1, W_2\}$ and the directed edges $W_1 \rightarrow A \leftarrow W_2$ for all $A \in \boldsymbol{A}$ is also a MAG satisfying $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$.

Lemma 3. Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG and $\boldsymbol{A}$ and $\boldsymbol{B}$ be two disjoint sets of variables that partition $\boldsymbol{O}$. If $\mathcal{G}$ satisfies $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$, then $\mathcal{H} = \textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})$ is a MAG satisfying $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$.

Let $\mathcal{H}' = \textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})$; by Lemma 2, $\mathcal{H}'$ is a MAG satisfying $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$. Note that $\mathcal{H}$ may be constructed from $\mathcal{H}'$ by inserting $\boldsymbol{W} = \{W_1, W_2\}$ and the edges $\{W \rightarrow A \mid A \in \boldsymbol{A}, W \in \boldsymbol{W}\}$.

Applying $\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\cdot)$ to $\mathcal{H}'$ inserts directed edges between non-adjacent variables (the newly inserted members of $\boldsymbol{W}$ are initially not adjacent to any variable).

Therefore, $\mathcal{H}$ is a mixed graph.

Similarly, since the members of $\boldsymbol{W}$ have no parents or spouses, any paths resulting from $\textsc{Add}^{\boldsymbol{A}}_{\boldsymbol{B}}(\cdot)$ cannot be part of a directed or almost directed cycle. Therefore, $\mathcal{H}$ is ancestral.

Lastly, assume $\mathcal{H}$ is not maximal. By Proposition 1, a primitive inducing path π exists between two non-adjacent variables $X, Y \in \boldsymbol{O}$ in $\mathcal{H}$. By construction, the set of all paths in $\mathcal{H}'$ is a subset of those in $\mathcal{H}$. Therefore, π is either a primitive inducing path in $\mathcal{H}'$ which is a contradiction because $\mathcal{H}'$ is maximal, or π is not in $\mathcal{H}'$. The latter implies that π includes an edge inserted by $\textsc{Add}^{\boldsymbol{A}}_{\boldsymbol{B}}(\cdot)$. Accordingly, $W \in \boldsymbol{W}$ is an endpoint on π and $A \in \boldsymbol{A}$ must be a collider on π. But the only path in $\mathcal{H}$ on which A is a collider is $\langle W_1 \rightarrow A \leftarrow W_2 \rangle$. However, A is not an ancestor of W_1 or W_2 which means π is not an inducing path. Therefore, $\mathcal{H}$ is maximal.

By Lemma 1, since $\mathcal{H}'$ satisfies $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$, there exists no $A \in \boldsymbol{A}$ and $B \in \boldsymbol{B}$ such that $A \leftrightarrow B$ or $B \rightarrow A$ in $\mathcal{H}'$. Since $\mathcal{H}$ may be constructed from $\mathcal{H}'$ by inserting the members of $\boldsymbol{W}$ and edges between the members of $\boldsymbol{A} \cup \boldsymbol{W}$, there exist no $A \in \boldsymbol{A}$ and $B \in \boldsymbol{B}$ such that $A \leftrightarrow B$ or $B \rightarrow A$ in $\mathcal{H}$. Therefore, $\mathcal{H}$ satisfies $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$. □

Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG where the variables may be partitioned into two distinct subsets $\boldsymbol{O} = \boldsymbol{A} \cup \boldsymbol{B}$. Lemma 4 proves that if $\mathcal{G}$ satisfies $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$, then inserting the edges in $\mathcal{G}$ connecting two members of $\boldsymbol{A}$ into a member of the Markov equivalence class constrained with $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$ for the MAG constructed by removing from $\mathcal{G}$ the edges connecting two members of $\boldsymbol{A}$ is a MAG satisfying $ebk\boldsymbol{A}_{\boldsymbol{B}}$.

Lemma 4.

Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG, $\boldsymbol{A}$ and $\boldsymbol{B}$ be two disjoint sets of variables that partition $\boldsymbol{O}$, and $\mathcal{G}' \in [\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\mathcal{G}) + ebk]$. If $\mathcal{G}$ satisfies $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$, then $\mathcal{H} = \textsc{Ins}(\mathcal{G}', \textsc{Edges}(\mathcal{G}, \boldsymbol{A}))$ is a MAG satisfying $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$.

By Lemma 2, $\mathcal{G}'$ is a MAG. Note that $\mathcal{H}$ may be constructed from $\mathcal{G}'$ by inserting the edges $\textsc{Edges}(\mathcal{G}, \boldsymbol{A})$.

Applying $\textsc{Ins}(\cdot, \textsc{Edges}(\mathcal{G}, \boldsymbol{A}))$ to the mixed graph $\mathcal{G}'$ inserts directed and bi-directed edges between non-adjacent variables (the edges between the members of $\boldsymbol{A}$ in $\mathcal{G}$ are either directed or bi-directed and the members of $\boldsymbol{A}$ in $\mathcal{G}'$ are not adjacent). Therefore, $\mathcal{H}$ is a mixed graph.

Assume $\mathcal{H}$ is not ancestral. Then there exists $X, Y \in \boldsymbol{O}$ such that $\mathcal{H}$ contains a directed path π from X to Y and $X \leftrightarrow Y$ or $X \leftarrow Y$. But π cannot be a path containing only the members of $\boldsymbol{B}$ since $\mathcal{G}'$ is ancestral, or a path containing only the members of $\boldsymbol{A}$ since $\mathcal{G}$ is ancestral. Therefore π must contain both a member of $\boldsymbol{A}$ and a member of $\boldsymbol{B}$. But $\mathcal{G}'$ satisfies $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$ so every edge between $A \in \boldsymbol{A}$ and $B \in \boldsymbol{B}$ in $\mathcal{G}'$ is directed $A \rightarrow B$. But by construction, $\mathcal{H}$ has the same edges as $\mathcal{G}'$ between a member of $\boldsymbol{A}$ and a member of $\boldsymbol{B}$. It follows π must have endpoints $X \in \boldsymbol{A}$ and $Y \in \boldsymbol{B}$. This is a contradiction since $X \leftrightarrow Y$ or $X \leftarrow Y$. Therefore, $\mathcal{H}$ is ancestral.

Similarly, assume that $\mathcal{H}$ is not maximal. Then a primitive inducing path π must exist in $\mathcal{H}$ between two non-adjacent variables $X, Y \in \boldsymbol{O}$.

Suppose π is a path containing only the members of $\boldsymbol{B}$. Then by construction, π is in $\mathcal{G}'$. But $\mathcal{G}'$ is maximal so π cannot be a primitive inducing path in $\mathcal{G}'$. Since $\mathcal{H}$ and $\mathcal{G}'$ share the same edges between the members of $\boldsymbol{B}$, X and Y are not adjacent in $\mathcal{G}'$. Therefore, π must not be an inducing path in $\mathcal{G}'$. But the ancestral relationships between the members of $\boldsymbol{B}$ in $\mathcal{H}$ are the same as they are in $\mathcal{G}'$ because no $B \in \boldsymbol{B}$ can have a member of $\boldsymbol{A}$ as a descendant. That is, $\textsc{Ins}(\cdot, \textsc{Edges}(\mathcal{G}, \boldsymbol{A}))$ cannot induce new paths between the members of $\boldsymbol{B}$. Thus, if π is not inducing in G', it cannot be inducing in $\mathcal{H}$. Therefore π cannot be a path containing only the members of $\boldsymbol{B}$.

Suppose π is a path containing only the members of $\boldsymbol{A}$. Then by construction, π is in $\mathcal{G}$. But $\mathcal{G}$ is maximal so π cannot be a primitive inducing path in $\mathcal{G}$. Since $\mathcal{H}$ and $\mathcal{G}$ share the same edges between the members of $\boldsymbol{A}$, X and Y are not adjacent in $\mathcal{G}$. Therefore, π must not be an inducing path in $\mathcal{G}$. But the ancestral relationships between the members of $\boldsymbol{A}$ in $\mathcal{H}$ are the same as they are in $\mathcal{G}$ because no $A \in \boldsymbol{A}$ can have a member of $\boldsymbol{B}$ as an ancestor. That is, $\textsc{Ins}(\cdot, \textsc{Edges}(\mathcal{G}, \boldsymbol{A}))$ cannot induce new paths between the members of $\boldsymbol{A}$. Thus, if π is not inducing in G, it cannot be inducing in $\mathcal{H}$. Therefore π cannot be a path containing only the members of $\boldsymbol{A}$.

Therefore there exists $A \in \boldsymbol{A}$ and $B \in \boldsymbol{B}$ such that the edge $A \rightarrow B$ is in π. Accordingly, π must be of the form 2b. (see Table 2). But by construction $\mathcal{G}$ and $\mathcal{G}'$ have the same adjacencies and colliders with order.

Additionally, the parent relationships between the members of π in $\mathcal{H}$ are the same as they are in $\mathcal{G}'$ because no $B \in \boldsymbol{B}$ can have a member of $\boldsymbol{A}$ as a child. It follows that $\mathcal{G}'$ and $\mathcal{H}$ have the same adjacencies and that any collider with order of the form 2b. in $\mathcal{G}'$ will be in $\mathcal{H}$. But then π is a primitive inducing path between non-adjacent variables X, Y in $\mathcal{G}'$. This is a contradiction because $\mathcal{G}'$ is maximal. Therefore, $\mathcal{G}$ is maximal.

By Lemma 1, since $\mathcal{G}'$ satisfies $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$, there exists no $A \in \boldsymbol{A}$ and $B \in \boldsymbol{B}$ such that $A \leftrightarrow B$ or $B \rightarrow A$ in $\mathcal{G}'$.

Since $\mathcal{H}$ may be constructed from $\mathcal{G}'$ by inserting edges between the members of $\boldsymbol{A}$, there exists no $A \in \boldsymbol{A}$ and $B \in \boldsymbol{B}$ such that $A \leftrightarrow B$ or $B \rightarrow A$ in $\mathcal{H}$. Therefore, $\mathcal{H}$ satisfies $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$. □

Table 1: Colliders with Order One

Case	Collider	$X {*\!\!\rightarrow} Z {\leftarrow\!\!*} Y$
1a.	$\langle A, A, A\rangle$	$A {*\!\!\rightarrow} A {\leftarrow\!\!*} A$
1b.	$\langle A, B, A\rangle$	$A \rightarrow B \leftarrow A$
1c.	$\langle A, B, B\rangle$	$A \rightarrow B {\leftarrow\!\!*} B$
1d.	$\langle B, B, B\rangle$	$B {*\!\!\rightarrow} B {\leftarrow\!\!*} B$

Table 2: Colliders with Order Greater Than One

Case	Collider	$X {*\!\!\rightarrow} Q_1 \leftrightarrow \cdots \leftrightarrow Q_p \leftrightarrow Z \leftrightarrow Y$
2a.	$\langle A, A, A\rangle$	$A {*\!\!\rightarrow} A \leftrightarrow \cdots \leftrightarrow A \leftrightarrow A \leftrightarrow A$
2b.	$\langle B, B, B\rangle$	$A \rightarrow B \leftrightarrow \cdots \leftrightarrow B \leftrightarrow B \leftrightarrow B$
2c.	$\langle B, B, B\rangle$	$B {*\!\!\rightarrow} B \leftrightarrow \cdots \leftrightarrow B \leftrightarrow B \leftrightarrow B$

Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG, $\boldsymbol{A}$ and $\boldsymbol{B}$ be two disjoint sets of variables that partition $\boldsymbol{O}$. Tables 1 and 2 illustrate the possible forms of colliders with order in $\mathcal{G}$ that satisfy $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$. More specifically, for all $A \in \boldsymbol{A}$ and $B \in \boldsymbol{B}$, any collider with order where an $A {\leftarrow\!\!*} B$ edge exists has been disregarded. Within the tables, A and B are used to denote arbitrary members of $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively. These tables are used to aid the proofs of several lemmas.

Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG where the variables may be partitioned into two distinct subsets $\boldsymbol{O} = \boldsymbol{A} \cup \boldsymbol{B}$. Lemma 5 proves that if $\mathcal{G}$ satisfies $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$, then the m.i. PAG constructed by removing from the m.i. PAG constrained with $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$ for the MAG $\mathcal{G}$ the edges connecting two members of $\boldsymbol{A}$ is equivalent to the Markov equivalence class constrained with $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$ for the MAG constructed by removing from $\mathcal{G}$ the edges connecting two members of $\boldsymbol{A}$.

Lemma 5. Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG and $\boldsymbol{A}$ and $\boldsymbol{B}$ be two disjoint sets of variables that partition $\boldsymbol{O}$. If $\mathcal{G}$ satisfies $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$, then $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}([\mathcal{G}] + ebk^{\boldsymbol{A}}_{\boldsymbol{B}}) \equiv [\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\mathcal{G})] + ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$.

($\Rightarrow$) If $\mathcal{G}$ satisfies $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$ then $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}([\mathcal{G}] + ebk^{\boldsymbol{A}}_{\boldsymbol{B}}) \subseteq [\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\mathcal{G})] + ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$.

Let $\mathcal{H}' \in ([\mathcal{G}] + ebk^{\boldsymbol{A}}_{\boldsymbol{B}})$ and $\mathcal{H} = \textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\mathcal{H}')$. By Lemma 2, $\mathcal{H}$ is a MAG satisfying $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$. Additionally, $\mathcal{H}'$ and $\mathcal{G}$ have the same adjacencies and colliders with order because they are Markov equivalent.

Applying $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\cdot)$ removes the same edges from $\mathcal{H}'$ and $\mathcal{G}$. Therefore $\mathcal{H}$ and $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\mathcal{G})$ have the same adjacencies since $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\cdot)$ removed the same edges from $\mathcal{H}'$ and $\mathcal{G}$.

Applying $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\cdot)$ removes only colliders with order of the forms 1a. and 2a. (see Tables 1 and 2) from $\mathcal{H}'$ and $\mathcal{G}$. A collider with order of the form 1b., 1c., or 1d. only depends on the edges in the collider. Thus, a collider with order of the form 1b., 1c., or 1d. is unaffected by $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\cdot)$ because the collider does not lose any edges. A collider with order of the form 2b. or 2c. is unaffected by $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\cdot)$. The path π that defines the collider does not lose any edges and the parental relationships between the endpoints and non-endpoints of π do not change because $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\cdot)$ only removes the members of $\boldsymbol{A}$ and no $B \in \boldsymbol{B}$ can be a parent of a member of $\boldsymbol{A}$. Therefore $\mathcal{H}$ and $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\mathcal{G})$ have the same collider with order since $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\cdot)$ removed the same colliders with order from $\mathcal{H}'$ and $\mathcal{G}$.

Therefore, $\mathcal{H} \in [\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\mathcal{G})] + ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$ which implies $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}([\mathcal{G}] + ebk^{\boldsymbol{A}}_{\boldsymbol{B}}) \subseteq [\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\mathcal{G})] + ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$.

($\Leftarrow$) If $\mathcal{G}$ satisfies $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$ then $[\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\mathcal{G})] + ebk^{\boldsymbol{A}}_{\boldsymbol{B}} \subseteq \textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}([\mathcal{G}] + ebk^{\boldsymbol{A}}_{\boldsymbol{B}})$.

Note that $\mathcal{G} = \textsc{Ins}(\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\mathcal{G}), \textsc{Edges}(\mathcal{G}, \boldsymbol{A}))$ because $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\cdot)$ removes the edges between the members of $\boldsymbol{A}$ which are then added back by $\textsc{Ins}(\cdot, \textsc{Edges}(\mathcal{G}, \boldsymbol{A}))$.

Let $\mathcal{H} \in ([\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\mathcal{G})] + ebk^{\boldsymbol{A}}_{\boldsymbol{B}})$ and $\mathcal{H}' = \textsc{Ins}(\mathcal{H}, \textsc{Edges}(\mathcal{G}, \boldsymbol{A}))$. By Lemma 4, $\mathcal{H}'$ is a MAG satisfying $ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$. Additionally, $\mathcal{H}$ and $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\mathcal{G})$ have the same adjacencies and colliders with order because they are Markov equivalent.

Applying $\textsc{Ins}(\cdot, \textsc{Edges}(\mathcal{G}, \boldsymbol{A}))$ inserts only colliders with order of the forms 1a. and 2a. (see Tables 1 and 2) into $\mathcal{H}$ and $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\mathcal{G})$. A collider with order of the form 1b., 1c., or 1d. only depends on the edges in the collider. Thus, a collider with order of the form 1b., 1c., or 1d. cannot be inserted by $\textsc{Ins}(\cdot, \textsc{Edges}(\mathcal{G}, \boldsymbol{A}))$ because none of the necessary edges are added. Similarly, a collider with order of the form 2b. or 2c. cannot be inserted by $\textsc{Ins}(\cdot, \textsc{Edges}(\mathcal{G}, \boldsymbol{A}))$ because none of the necessary edges are added. A collider with order of the form 2b. or 2c. cannot be induced by $\textsc{Ins}(\cdot, \textsc{Edges}(\mathcal{G}, \boldsymbol{A}))$. The parental relationships between the endpoints and non-endpoints of paths of the form 2b. or 2c. do not change because $\textsc{Ins}(\cdot, \textsc{Edges}(\mathcal{G}, \boldsymbol{A}))$ only inserts the members of $\boldsymbol{A}$ and no $B \in \boldsymbol{B}$ can be a parent of a member of $\boldsymbol{A}$. Therefore $\mathcal{H}$ and $\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\mathcal{G})$ have the same collider with order since $\textsc{Ins}(\cdot, \textsc{Edges}(\mathcal{G}, \boldsymbol{A}))$ inserts the same colliders with order into $\mathcal{H}'$ and $\mathcal{G}$.

Therefore $\mathcal{H}' \in [\mathcal{G}] + ebk^{\boldsymbol{A}}_{\boldsymbol{B}}$. Accordingly, $\mathcal{H} \in \textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}([\mathcal{G}] + ebk^{\boldsymbol{A}}_{\boldsymbol{B}})$ which implies $[\textsc{Rm}^{\boldsymbol{A}}_{\boldsymbol{B}}(\mathcal{G})] + ebk^{\boldsymbol{A}}_{\boldsymbol{B}} \subseteq$

$\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}([\mathcal{G}] + ebk_{\boldsymbol{B}}^{\boldsymbol{A}})$.

Since $\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}([\mathcal{G}] + ebk_{\boldsymbol{B}}^{\boldsymbol{A}}) \subseteq [\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})] + ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$ and $[\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})] + ebk_{\boldsymbol{B}}^{\boldsymbol{A}} \subseteq \textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}([\mathcal{G}] + ebk_{\boldsymbol{B}}^{\boldsymbol{A}})$, $\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}([\mathcal{G}] + ebk_{\boldsymbol{B}}^{\boldsymbol{A}}) \equiv [\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})] + ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$. □

Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG where the variables may be partitioned into two distinct subsets $\boldsymbol{O} = \boldsymbol{A} \cup \boldsymbol{B}$. Lemma 6 proves that if $\mathcal{G}$ satisfies $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$, then the Markov equivalence class with $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$ of the MAG constructed by removing from $\mathcal{G}$ the edges connecting two members of $\boldsymbol{A}$ is equivalent to the induced subgraph over $\boldsymbol{A} \cup \boldsymbol{B}$ of the m.i. PAG for the MAG constructed by removing from $\mathcal{G}$ the edges connecting two members of $\boldsymbol{A}$ then adding the variables in $\boldsymbol{W} = \{W_1, W_2\}$ and the directed edges $W_1 \to A \leftarrow W_2$ for all $A \in \boldsymbol{A}$.

Lemma 6. Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG and $\boldsymbol{A}$ and $\boldsymbol{B}$ be two disjoint sets of variables that partition $\boldsymbol{O}$. If $\mathcal{G}$ satisfies $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$, then $[\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})] + ebk_{\boldsymbol{B}}^{\boldsymbol{A}} \equiv [\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})](\boldsymbol{A} \cup \boldsymbol{B})$.

By Lemma 3, $\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})$ is a MAG satisfying $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$. Let $\mathcal{H} \in [\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})]$ and note that $\mathcal{H}$ has no edges connecting two members of $\boldsymbol{A}$. Therefore, $\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\cdot)$ only removes the members of $\boldsymbol{W}$. Accordingly, $\mathcal{H}(\boldsymbol{A} \cup \boldsymbol{B}) = \textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{H})$. Furthermore, note that $\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})) = \textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})$ because $\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\cdot)$ removes the variables and edges added by $\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\cdot)$. It follwos that, using Lemma 3 and 5,

$$\begin{aligned}
[\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})](\boldsymbol{A} \cup \boldsymbol{B}) &\equiv \textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}([\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})]) \\
&\equiv \textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}([\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})] + ebk_{\boldsymbol{B}}^{\boldsymbol{A}}) \\
&\equiv [\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))] + ebk_{\boldsymbol{B}}^{\boldsymbol{A}} \\
&\equiv [\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})] + ebk_{\boldsymbol{B}}^{\boldsymbol{A}}.
\end{aligned}$$

□

Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG where the variables may be partitioned into two distinct subsets $\boldsymbol{O} = \boldsymbol{A} \cup \boldsymbol{B}$. Lemma 7 proves that if $\mathcal{G}$ satisfies $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$, then the induced subgraph over $\boldsymbol{A} \cup \boldsymbol{B}$ of the m.i. PAG output by FCI run with $\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})$ as a conditional independence oracle is equivalent to the m.i. PAG output by FCI run with $\mathcal{G}$ as a conditional independence oracle and $mbk_{\boldsymbol{B}}^{\boldsymbol{A}}$.

Lemma 7. Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG and $\boldsymbol{A}$ and $\boldsymbol{B}$ be two disjoint sets of variables that partition $\boldsymbol{O}$. If $\mathcal{G}$ satisfies $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$, then $\textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))(\boldsymbol{A} \cup \boldsymbol{B}) \equiv \textsc{Fci}(\mathcal{G} + mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$.

By Lemma 3, $\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})$ is a MAG and therefore may be used a conditional independence oracle. Using $\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})$ and $\mathcal{G}$ as conditional independence oracles, FCI will return graphs with the adjacencies of $\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})$ and $\mathcal{G}$ respectively. Applying $mbk_{\boldsymbol{B}}^{\boldsymbol{A}}$ to the latter will return a graph with a subset of adjacencies of $\mathcal{G}$ where the members of $\boldsymbol{A}$ are not adjacent. Thus, $\textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))(\boldsymbol{A} \cup \boldsymbol{B})$ and $\textsc{Fci}(\mathcal{G} + mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$ will have the same adjacencies.

For orienting edges, the FCI algorithm steps through a series of rules. We will examine the application of these rules in conjunction with $mbk_{\boldsymbol{B}}^{\boldsymbol{A}}$ and note the effect on $\textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))$ and $\textsc{Fci}(\mathcal{G} + mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$. As shorthand, we use A to denote a member of $\boldsymbol{A}$ and B to denote a member of $\boldsymbol{B}$. Before the application of any rules, $mbk_{\boldsymbol{B}}^{\boldsymbol{A}}$ will orient all A ∘–∘ B adjacencies as $A \to B$ in $\textsc{Fci}(\mathcal{G} + mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$.

Consider cases of $\mathcal{R}_0$:

> $B \to B \leftarrow B$: colliders of this form will be in both $\textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))$ and $\textsc{Fci}(\mathcal{G}+mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$. Therefore $\mathcal{R}_0$ will be applied in both.
>
> $A \to B \leftarrow B$: colliders of this form will be in both $\textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))$ and $\textsc{Fci}(\mathcal{G}+mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$ Therefore $\mathcal{R}_0$ will be applied in both.
>
> $A \to B \leftarrow A$: colliders of this form will be in both $\textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))$ and $\textsc{Fci}(\mathcal{G}+mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$. Therefore $\mathcal{R}_0$ will be applied in both.

Consider cases of $\mathcal{R}_1$:

> Colliders W_1∘→ A ←∘W_2 will orient all A ∘–∘ B adjacencies as W_1∘→ $A \to B$ in $\textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))$. But orientations of this nature have been applied to $\textsc{Fci}(\mathcal{G} + mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$ through $mbk_{\boldsymbol{B}}^{\boldsymbol{A}}$. Thus, $\textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))$ and $\textsc{Fci}(\mathcal{G} + mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$ will have the same orientations among the members of $\boldsymbol{A} \cup \boldsymbol{B}$.

Consider cases of $\mathcal{R}_4$:

> $A{*}\!\!\to \cdots \leftrightarrow A_\alpha \leftrightarrow A_\beta \to B_\gamma$: paths of this form are impossible because the members of $\boldsymbol{A}$ are not adjacent in either $\textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))$ or $\textsc{Fci}(\mathcal{G} + mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$.
>
> $A/B{*}\!\!\to \cdots \leftrightarrow B_\alpha \leftrightarrow B_\beta \to B_\gamma$: the $A \leftrightarrow B$ case is impossible, but otherwise paths of this form will be in both $\textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))$ and $\textsc{Fci}(\mathcal{G} + mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$. Therefore the B_β will be in the sepset of B_α and B_γ in both $\textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))$ and $\textsc{Fci}(\mathcal{G} + mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$. Accordingly, R_4 will orient $B_\beta \to B_\gamma$ in both.
>
> $A/B{*}\!\!\to \cdots \leftrightarrow B_\alpha \leftrightarrow B_\beta \leftrightarrow B_\gamma$: the $A \leftrightarrow B$ case is impossible, but otherwise paths of this form will be in both $\textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))$ and $\textsc{Fci}(\mathcal{G} + mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$. Therefore the B_β will not be in the sepset of B_α and B_γ in either $\textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))$ or $\textsc{Fci}(\mathcal{G} + mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$. Accordingly, R_4 will orient $B_\alpha \leftrightarrow B_\gamma \leftrightarrow B_\gamma$ in both.

All the remaining rules propagate preexisting orientations throughout the graph. Therefore $\textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))$

and $\textsc{Fci}(\mathcal{G} + mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$ will have the same adjacencies and orientations among the members of $\boldsymbol{A} \cup \boldsymbol{B}$ and $\textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))(\boldsymbol{A} \cup \boldsymbol{B}) \equiv \textsc{Fci}(\mathcal{G} + mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$ □

Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG where the variables may be partitioned into two distinct subsets $\boldsymbol{O} = \boldsymbol{A} \cup \boldsymbol{B}$. Lemma 8 proves that if $\mathcal{G}$ satisfies $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$, then running FCI using $\mathcal{G}$ as a conditional independence oracle and incorporating modified background knowledge $mbk_{\boldsymbol{B}}^{\boldsymbol{A}}$ recovers the sound and complete set of edges that connect two members of $\boldsymbol{B}$.

Lemma 8. Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG and $\boldsymbol{A}$ and $\boldsymbol{B}$ be two disjoint sets of variables that partition $\boldsymbol{O}$. If $\mathcal{G}$ satisfies $ebk_{\boldsymbol{B}}^{\boldsymbol{A}}$, then $\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}([\mathcal{G}] + ebk_{\boldsymbol{B}}^{\boldsymbol{A}}) \equiv \textsc{Fci}(\mathcal{G} + mbk_{\boldsymbol{B}}^{\boldsymbol{A}})$.

FCI is sound and complete in the sense that, given a conditional independence oracle, FCI will return the m.i. PAG (Spirtes et al., 2000; Zhang, 2008). Thus, for a MAG $\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})$, we have $[\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})] \equiv \textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))$. It follows that, using Lemmas 5, 6, and 7,

$$\begin{aligned}\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}([\mathcal{G}] + ebk_{\boldsymbol{B}}^{\boldsymbol{A}}) &\equiv [\textsc{Rm}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})] + ebk_{\boldsymbol{B}}^{\boldsymbol{A}} \\ &\equiv [\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G})](\boldsymbol{A} \cup \boldsymbol{B}) \\ &\equiv \textsc{Fci}(\textsc{Add}_{\boldsymbol{B}}^{\boldsymbol{A}}(\mathcal{G}))(\boldsymbol{A} \cup \boldsymbol{B}) \\ &\equiv \textsc{Fci}(\mathcal{G} + mbk_{\boldsymbol{B}}^{\boldsymbol{A}}).\end{aligned}$$

□

Let $\mathcal{G}$ be a MAG where the variables may be partitioned into $n > 1$ disjoint subsets $\boldsymbol{T} = \{\boldsymbol{T}_1, \ldots, \boldsymbol{T}_n\}$. Let $\boldsymbol{A}_i = \bigcup_{j=1}^{i-1} \boldsymbol{T}_j$, $\boldsymbol{B}_i = \boldsymbol{T}_i$, and $\boldsymbol{O}_i = \boldsymbol{A}_i \cup \boldsymbol{B}_i$. Lemma 9 proves that if $\mathcal{G}$ satisfies $tbk^{\boldsymbol{T}}$, then the induced subgraph over $\boldsymbol{O}_i$ of $\mathcal{G}$ is a MAG satisfying $tbk^{\boldsymbol{T}}$ for all $1 \leq i \leq n$.

Lemma 9. Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG and $\boldsymbol{T} = \{\boldsymbol{T}_1, \ldots, \boldsymbol{T}_n\}$ be a partitioning of $\boldsymbol{O}$. Let $\boldsymbol{A}_i = \bigcup_{j=1}^{i-1} \boldsymbol{T}_j$, $\boldsymbol{B}_i = \boldsymbol{T}_i$, and $\boldsymbol{O}_i = \boldsymbol{A}_i \cup \boldsymbol{B}_i$. If $\mathcal{G}$ satisfies $tbk^{\boldsymbol{T}}$, then $\mathcal{H} = \mathcal{G}(\boldsymbol{O}_i)$ is a MAG satisfying $tbk^{\boldsymbol{T}}$ over the variables $\boldsymbol{O}_i$ for all $1 \leq i \leq n$.

Since $\mathcal{H}$ is a subgraph of $\mathcal{G}$, $\mathcal{H}$ contains a subset of the variables and edges in $\mathcal{G}$. Therefore $\mathcal{H}$ is a mixed graph.

Similarly, the set of paths in $\mathcal{H}$ is a subset of the paths in $\mathcal{G}$. Since $\mathcal{G}$ is ancestral, $\mathcal{G}$ does not contain any directed or almost directed cycles. Therefore $\mathcal{H}$ does not contain any direct or almost direct cycles; $\mathcal{H}$ is ancestral.

Let $\boldsymbol{L}_i = \boldsymbol{O} \setminus \boldsymbol{O}_i$ denote the variables removed by the subgraph operation. Assume that $\mathcal{H}$ is not maximal. Since $\mathcal{G}$ is maximal, the subgraph operation made an inducing path π between two members of $\boldsymbol{O}_i$ a primitive inducing path in $\mathcal{H}$. That is, π is inducing for $\boldsymbol{L}_i$ in $\mathcal{G}$, but not for $\emptyset$. Therefore, there exists a non-collider $L \in \boldsymbol{L}_i$ because if no such L exists, then π is inducing for $\emptyset$. But if L is a non-collider on π, then it is an ancestor an the endpoint of π. This is a contradiction because the endpoints of π are members of $\boldsymbol{O}_i$ and, by $tbk^{\boldsymbol{T}}$, L cannot be an ancestor of a member of $\boldsymbol{O}_i$. Therefore, $\mathcal{H}$ is maximal.

Since $\mathcal{G}$ satisfies $tbk^{\boldsymbol{T}}$, there exist no $A \in \boldsymbol{A}_i$ and $B \in \boldsymbol{B}_i$ such that $A \leftrightarrow B$ or $B \rightarrow A$ in $\mathcal{G}(\boldsymbol{O}_i)$ for all $1 \leq i \leq n$. Therefore, the subgraph $\mathcal{G}(\boldsymbol{O}_i)$ satisfies $tbk^{\boldsymbol{T}}$ over the variables $\boldsymbol{O}_i$. □

Table 3: Colliders with Order One

Case	Collider	$X{*}\!\!\rightarrow Z \leftarrow\!\!{*}Y$
3a.	$\langle O_i, O_i, O_i \rangle$	$O_i{*}\!\!\rightarrow O_i \leftarrow\!\!{*}O_i$
3b.	$\langle O_i, L_i, O_i \rangle$	$O_i \rightarrow L_i \leftarrow O_i$
3c.	$\langle O_i, L_i, L_i \rangle$	$O_i \rightarrow L_i \leftarrow\!\!{*}L_i$
3d.	$\langle L_i, L_i, L_i \rangle$	$L_i{*}\!\!\rightarrow L_i \leftarrow\!\!{*}L_i$

Table 4: Colliders with Order Greater Than One

Case	Collider	$X{*}\!\!\rightarrow Q_1 \leftrightarrow \cdots \leftrightarrow Q_p \leftrightarrow Z \leftrightarrow Y$
4a.	$\langle O_i, O_i, O_i \rangle$	$O_i{*}\!\!\rightarrow O_i \leftrightarrow \cdots \leftrightarrow O_i \leftrightarrow O_i \leftrightarrow O_i$
4b.	$\langle L_i, L_i, L_i \rangle$	$O_i \rightarrow L_i \leftrightarrow \cdots \leftrightarrow L_i \leftrightarrow L_i \leftrightarrow L_i$
4c.	$\langle L_i, L_i, L_i \rangle$	$L_i{*}\!\!\rightarrow L_i \leftrightarrow \cdots \leftrightarrow L_i \leftrightarrow L_i \leftrightarrow L_i$

Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG and $\boldsymbol{T} = \{\boldsymbol{T}_1, \ldots, \boldsymbol{T}_n\}$ be a partitioning of $\boldsymbol{O}$. Let $\boldsymbol{A}_i = \bigcup_{j=1}^{i-1} \boldsymbol{T}_j$, $\boldsymbol{B}_i = \boldsymbol{T}_i$, $\boldsymbol{O}_i = \boldsymbol{A}_i \cup \boldsymbol{B}_i$, and $\boldsymbol{L}_i = \boldsymbol{O} \setminus \boldsymbol{O}_i$. Tables 3 and 4 illustrate the possible forms of colliders with order in $\mathcal{G}$ that satisfy $tbk^{\boldsymbol{T}}$. More specifically, for all $O_i \in \boldsymbol{O}_i$ and $L_i \in \boldsymbol{L}_i$, any collider with order where an $O_i \leftarrow\!\!{*}L_i$ edge exists has been disregarded. Within the tables, O_i and L_i are used to denote arbitrary members of $\boldsymbol{O}_i$ and $\boldsymbol{L}_i$, respectively. These tables are used to aid the proofs of several lemmas.

Let $\mathcal{G}$ be a MAG where the variables may be partitioned into $n > 1$ disjoint subsets $\boldsymbol{T} = \{\boldsymbol{T}_1, \ldots, \boldsymbol{T}_n\}$. Let $\boldsymbol{A}_i = \bigcup_{j=1}^{i-1} \boldsymbol{T}_j$, $\boldsymbol{B}_i = \boldsymbol{T}_i$, and $\boldsymbol{O}_i = \boldsymbol{A}_i \cup \boldsymbol{B}_i$. Lemma 10 proves that if $\mathcal{G}$ satisfies $tbk^{\boldsymbol{T}}$, then the induced subgraph over $\boldsymbol{O}_i$ of the m.i. PAG constrained with $tbk^{\boldsymbol{T}}$ for $\mathcal{G}$ is equivalent to the Markov equivalence class constrained by $tbk^{\boldsymbol{T}}$ for the induced subgraph over $\boldsymbol{O}_i$ of $\mathcal{G}$ for all $1 \leq i \leq n$.

Lemma 10. Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG and $\boldsymbol{T} = \{\boldsymbol{T}_1, \ldots, \boldsymbol{T}_n\}$ be a partitioning of $\boldsymbol{O}$. Let $\boldsymbol{A}_i = \bigcup_{j=1}^{i-1} \boldsymbol{T}_j$, $\boldsymbol{B}_i = \boldsymbol{T}_i$, and $\boldsymbol{O}_i = \boldsymbol{A}_i \cup \boldsymbol{B}_i$. If $\mathcal{G}$ satisfies $tbk^{\boldsymbol{T}}$, then $([\mathcal{G}] + tbk^{\boldsymbol{T}})(\boldsymbol{O}_i) \equiv [\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}}$ for all $1 \leq i \leq n$.

($\Rightarrow$) If $\mathcal{G}$ satisfies $tbk^{\boldsymbol{T}}$, then $([\mathcal{G}] + tbk^{\boldsymbol{T}})(\boldsymbol{O}_i) \subseteq [\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}}$ for all $1 \leq i \leq n$.

Without loss of generality, pick an arbitrary $1 \leq i \leq n$. Let $\mathcal{H} \in [\mathcal{G}] + tbk^{\boldsymbol{T}}$ and note that $\mathcal{H}$ and $\mathcal{G}$ have the same adjacencies and colliders with order because they are Markov equivalent. Let $\mathcal{H}' = \mathcal{H}(\boldsymbol{O}_i)$ and note that by Lemma 9, $\mathcal{H}'$ is a MAG satisfying $tbk^{\boldsymbol{T}}$ over $\boldsymbol{O}_i$.

Applying the subgraph operation with respect to $\boldsymbol{O}_i$ removes colliders with order of the forms 3b., 3c., 3d., 4b., and 4c. (see Table 3 and 4) from $\mathcal{G}$. Consider the remaining colliders.

A collider with order of the form 3a. only depends on the edges in the collider. Thus, a collider with order of the form 3a. is unaffected by subgraph operation with respect to $\boldsymbol{O}_i$ because the collider does not lose any edges. A collider with order of the form 4a is unaffected by the subgraph operation. The path π that defines the collider does not lose any edges and the parental relationships between the endpoints and non-endpoints of π do not change because the subgraph operation with respect to $\boldsymbol{O}_i$ only removes the members of $\boldsymbol{L}_i$ and no $O_i \in \boldsymbol{O}_i$ can be a child of a member of $\boldsymbol{L}_i$.

Thus, $\mathcal{H}'$ and $[\mathcal{G}(\boldsymbol{O}_i)]$ have the same adjacencies and colliders with order. Therefore, $\mathcal{H}' \in [\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}}$ which implies $([\mathcal{G}] + tbk^{\boldsymbol{T}})(\boldsymbol{O}_i) \subseteq [\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}}$ for all $1 \leq i \leq n$.

($\Leftarrow$) If $\mathcal{G}$ satisfies $tbk^{\boldsymbol{T}}$, then $[\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}} \subseteq ([\mathcal{G}] + tbk^{\boldsymbol{T}})(\boldsymbol{O}_i)$ for all $1 \leq i \leq n$.

Without loss of generality, pick an arbitrary $1 \leq i \leq n$. Let $\mathcal{H} \in [\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}}$, $\boldsymbol{L}_i = \boldsymbol{O} \setminus \boldsymbol{O}_i$, and $\mathcal{H}' = \text{INS}(\text{RM}^{\boldsymbol{O}_i}_{\boldsymbol{L}_i}(\mathcal{G}), \text{EDGES}(\mathcal{H}, \boldsymbol{O}_i))$.

$\mathcal{H}$ and $\mathcal{G}$ have the same adjacencies and colliders with order because they are Markov equivalent. By construction, $\mathcal{H}'$ and $\mathcal{H}$ have the same adjacencies over the members of $\boldsymbol{O}_i$. Similarly, $\mathcal{H}'$ and $\mathcal{H}$ share any paths π that define a colliders with order in $\mathcal{H}$ that contains only members of $\boldsymbol{O}_i$. $\mathcal{H}'$ and $\mathcal{H}$ share the same parental relationships between the endpoints and non-endpoints of π because no $O_i \in \boldsymbol{O}_i$ can be a parent of a member of $\boldsymbol{L}_i$ in $\mathcal{H}'$ or $\mathcal{H}$ due to $\mathcal{H}$ and $\mathcal{G}$ satisfying $tbk^{\boldsymbol{T}}$. Therefore, $\mathcal{H}'$ and $\mathcal{H}$ have the same adjacencies and colliders with order over the $\boldsymbol{O}_i$ variables. Accordingly, $\mathcal{H}'$ and $\mathcal{G}$ have the same adjacencies and colliders with order over the $\boldsymbol{O}_i$ variables.

By construction, $\mathcal{H}'$ and $\mathcal{G}$ have the same adjacencies over the $\boldsymbol{L}_i$ variables. Similarly, $\mathcal{H}'$ and $\mathcal{G}$ share any paths π that define a colliders with order in $\mathcal{G}$ that contains only members of $\boldsymbol{L}_i$. $\mathcal{H}'$ and $\mathcal{G}$ share the same parental relationships between the endpoints and non-endpoints of π because no $L_i \in \boldsymbol{L}_i$ can have a member of $\boldsymbol{O}_i$ as a child in $\mathcal{H}'$ or $\mathcal{H}$ due to $\mathcal{H}$ and $\mathcal{G}$ satisfying $tbk^{\boldsymbol{T}}$. Therefore, $\mathcal{H}'$ and $\mathcal{G}$ have the same adjacencies and colliders with order over the $\boldsymbol{L}_i$ variables.

Using the same logic, any adjacencies between a member of $\boldsymbol{O}_i$ and a member of $\boldsymbol{L}_i$ will be in both $\mathcal{H}'$ and $\mathcal{G}$. Furthermore, $\mathcal{H}'$ and $\mathcal{G}$ share any paths π that define a colliders with order in $\mathcal{G}$ between a member of $\boldsymbol{O}_i$ and a member of $\boldsymbol{L}_i$ because π must be of form 3b., 3c., or 4b. (see Tables 3 and 4). $\mathcal{H}'$ and $\mathcal{G}$ share the same parental relationships between the endpoints and non-endpoints of π because no $L_i \in \boldsymbol{L}_i$ can have a member of $\boldsymbol{L}_i$ as a child in $\mathcal{H}'$ or $\mathcal{H}$ due to $\mathcal{H}$ and $\mathcal{G}$ satisfying $tbk^{\boldsymbol{T}}$.

Finally, note that the edges between the members of $\boldsymbol{O}_i$ are the same as $\mathcal{G}$, thus, since $\mathcal{G}$ satisfies $tbk^{\boldsymbol{T}}$, there are not violations of $tbk^{\boldsymbol{T}}$ within the members of $\boldsymbol{O}_i$. Additionally, the edges between a member of $\boldsymbol{O}_i$ and a member of $\boldsymbol{L}_i$ satisfy $tbk^{\boldsymbol{T}}$. Therefore $\mathcal{H}'$ satisfies $tbk^{\boldsymbol{T}}$.

Therefore, $\mathcal{H}' \in [\mathcal{G}] + tbk^{\boldsymbol{T}}$ and, by construction, $\mathcal{H}'(\boldsymbol{O}_i) = \mathcal{H}$ which implies $\mathcal{H} \in ([\mathcal{G}] + tbk^{\boldsymbol{T}})(\boldsymbol{O}_i)$. Therefore $[\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}} \subseteq ([\mathcal{G}] + tbk^{\boldsymbol{T}})(\boldsymbol{O}_i)$ for all $1 \leq i \leq n$.

Since $([\mathcal{G}] + tbk^{\boldsymbol{T}})(\boldsymbol{O}_i) \subseteq [\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}}$ for all $1 \leq i \leq n$ and $[\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}} \subseteq ([\mathcal{G}] + tbk^{\boldsymbol{T}})(\boldsymbol{O}_i)$ for all $1 \leq i \leq n$, $([\mathcal{G}] + tbk^{\boldsymbol{T}})(\boldsymbol{O}_i) \equiv [\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}}$ for all $1 \leq i \leq n$. □

Let $\mathcal{G}$ be a MAG where the variables may be partitioned into $n > 1$ disjoint subsets $\boldsymbol{T} = \{\boldsymbol{T}_1, \ldots, \boldsymbol{T}_n\}$. Let $\boldsymbol{A}_i = \bigcup_{j=1}^{i-1} \boldsymbol{T}_j$, $\boldsymbol{B}_i = \boldsymbol{T}_i$, and $\boldsymbol{O}_i = \boldsymbol{A}_i \cup \boldsymbol{B}_i$. Lemma 11 proves that if $\mathcal{G}$ satisfies $tbk^{\boldsymbol{T}}$, then the m.i. PAG constructed by removing from the m.i. PAG constrained with $tbk^{\boldsymbol{T}}$ for the induced subgraph over $\boldsymbol{O}_i$ of $\mathcal{G}$ the edges connecting two members of $\boldsymbol{A}_i$ is equivalent to the m.i. PAG constructed by removing from the m.i. PAG constrained with $ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}$ for the induced subgraph over $\boldsymbol{O}_i$ of $\mathcal{G}$ the edges connecting two members of $\boldsymbol{A}_i$ is equivalent to for all $1 \leq i \leq n$.

Lemma 11. Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG and $\boldsymbol{T} = \{\boldsymbol{T}_1, \ldots, \boldsymbol{T}_n\}$ be a partitioning of $\boldsymbol{O}$. Let $\boldsymbol{A}_i = \bigcup_{j=1}^{i-1} \boldsymbol{T}_j$, $\boldsymbol{B}_i = \boldsymbol{T}_i$, and $\boldsymbol{O}_i = \boldsymbol{A}_i \cup \boldsymbol{B}_i$. If $\mathcal{G}$ satisfies $tbk^{\boldsymbol{T}}$, then $\text{RM}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}}) \equiv \text{RM}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i})$ for all $1 \leq i \leq n$.

($\Rightarrow$) If $\mathcal{G}$ satisfies $tbk^{\boldsymbol{T}}$, then $\text{RM}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}}) \subseteq \text{RM}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i})$ for all $1 \leq i \leq n$.

Without loss of generality, pick an arbitrary i such that $1 \leq i \leq n$. Let $\mathcal{H} \in \text{RM}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}})$ and note that any MAG that satisfies $tbk^{\boldsymbol{T}}$ satisfies $ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}$

since $ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}$ is contained within $tbk^{\boldsymbol{T}}$. Therefore, $\mathcal{H} \in \textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i})$ which implies $\textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}}) \subseteq \textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i})$ for all $1 \leq i \leq n$.

($\Leftarrow$) If $\mathcal{G}$ satisfies $tbk^{\boldsymbol{T}}$, then $\textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}) \subseteq \textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}})$ for all $1 \leq i \leq n$

Without loss of generality, pick an arbitrary i such that $1 \leq i \leq n$. Let $\mathcal{H} \in \textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i})$ and note that, by Lemmas 2 and 9, $\mathcal{H}$ is a MAG satisfying $ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}$. Let $\mathcal{H}' \in \textsc{Ins}(\mathcal{H}, \textsc{Edges}(\mathcal{G}(\boldsymbol{O}_i), \boldsymbol{A}_i))$ and note that, by Lemma 4, $\mathcal{H}'$ is a MAG satisfying $ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}$. By construction, $\mathcal{H}'$ and $\mathcal{G}(\boldsymbol{O}_i)$ have the same adjacencies.

Additionally, $\mathcal{H}'$ and $\mathcal{G}(\boldsymbol{O}_i)$ share any path π that defines a collider with order in $\mathcal{G}(\boldsymbol{O}_i)$ that contains only members of $\boldsymbol{A}_i$ since they both contain the same edges over $\boldsymbol{A}_i$. $\mathcal{H}'$ and $\mathcal{G}(\boldsymbol{O}_i)$ share the same parental relationships between the endpoints and non-endpoints of π because no $A_i \in \boldsymbol{A}_i$ can have a member of $\boldsymbol{B}_i$ as a parent in $\mathcal{H}'$ or $\mathcal{G}(\boldsymbol{O}_i)$ due to $\mathcal{H}'$ and $\mathcal{G}(\boldsymbol{O}_i)$ satisfying $ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}$. Therefore, $\mathcal{H}'$ and $\mathcal{G}(\boldsymbol{O}_i)$ have the same adjacencies and colliders with order over the $\boldsymbol{A}_i$ variables.

Similarly, $\mathcal{H}'$ and $\mathcal{G}(\boldsymbol{O}_i)$ share any path π that defines a colliders with order in $\mathcal{G}(\boldsymbol{O}_i)$ that contains only members of $\boldsymbol{B}_i$ since π has no ambiguity in $[\mathcal{G}(\boldsymbol{O}_i)]$ (by virtue of being a collider or discriminating path). $\mathcal{H}'$ and $\mathcal{G}(\boldsymbol{O}_i)$ share the same parental relationships between the endpoints and non-endpoints of π because no $B_i \in \boldsymbol{B}_i$ can have a member of $\boldsymbol{A}_i$ as a child in $\mathcal{H}'$ or $\mathcal{G}(\boldsymbol{O}_i)$ due to $\mathcal{H}'$ and $\mathcal{G}(\boldsymbol{O}_i)$ satisfying $ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}$. Therefore, $\mathcal{H}'$ and $\mathcal{G}(\boldsymbol{O}_i)$ have the same adjacencies and colliders with order over the $\boldsymbol{B}_i$ variables.

Finally, $\mathcal{H}'$ and $\mathcal{G}(\boldsymbol{O}_i)$ share any path π that defines a collider with order in $\mathcal{G}(\boldsymbol{O}_i)$ that contains a member of $\boldsymbol{A}_i$ and a member of $\boldsymbol{B}_i$. π must be of the form 1b., 1c., or 2b. (see Table 1 and 2)—the edges involving the member of A are oriented due $\mathcal{H}'$ $ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}$ and the remaining edges have no ambiguity in $[\mathcal{G}(\boldsymbol{O}_i)]$ (by virtue of being a collider or discriminating path). $\mathcal{H}'$ and $\mathcal{G}(\boldsymbol{O}_i)$ share the same parental relationships between the endpoints and non-endpoints of π because no $A_i \in \boldsymbol{A}_i$ can have a member of $\boldsymbol{B}_i$ as a parent in $\mathcal{H}'$ or $\mathcal{G}(\boldsymbol{O}_i)$ due to $\mathcal{H}'$ and $\mathcal{G}(\boldsymbol{O}_i)$ satisfying $ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}$. Therefore $\mathcal{H}'$ and $\mathcal{G}(\boldsymbol{O}_i)$ share the same colliders with order.

Note that the edges in $\mathcal{H}'$ between the members of $\boldsymbol{A}_i$ are identical to $\mathcal{G}$, thus, since $\mathcal{G}$ satisfies $tbk^{\boldsymbol{T}}$, there are no violations of $tbk^{\boldsymbol{T}}$ within the members of $\boldsymbol{A}_i$. Additionally, the edges connecting a member of $\boldsymbol{A}_i$ and a member of $\boldsymbol{B}_i$ in $\mathcal{H}'$ satisfy $ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}$. Therefore $\mathcal{H}'$ satisfies $tbk^{\boldsymbol{T}}$ over the variables $\boldsymbol{O}_i$.

Therefore $\mathcal{H}' \in [\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}}$, $\mathcal{H} \in \textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}})$ which implies $\textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}) \subseteq \textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}})$ for all $1 \leq i \leq n$.

Since $\textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}}) \subseteq \textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i})$ for all $1 \leq i \leq n$ and $\textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}) \subseteq \textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}})$ for all $1 \leq i \leq n$, $\textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}}) \equiv \textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i})$ for all $1 \leq i \leq n$. $\square$

Let $\mathcal{G}$ be a MAG where the variables may be partitioned into $n > 1$ disjoint subsets $\boldsymbol{T} = \{\boldsymbol{T}_1, \ldots, \boldsymbol{T}_n\}$. Let $\boldsymbol{A}_i = \bigcup_{j=1}^{i-1} \boldsymbol{T}_j$, $\boldsymbol{B}_i = \boldsymbol{T}_i$, and $\boldsymbol{O}_i = \boldsymbol{A}_i \cup \boldsymbol{B}_i$. Lemma 12 extends the results of Lemma 8 by showing that if $\mathcal{G}$ satisfies $tbk^{\boldsymbol{T}}$, then for all $1 \leq i \leq n$, running FCI on $\boldsymbol{O}_i$ using $\mathcal{G}(\boldsymbol{O}_i)$ as a conditional independence oracle and incorporating modified background knowledge $mbk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}$ recovers the sound and complete set of edges that connect two members of $\boldsymbol{B}_i$.

Lemma 12. Let $\mathcal{G} = (\boldsymbol{O}, \boldsymbol{E})$ be a MAG and $\boldsymbol{T} = \{\boldsymbol{T}_1, \ldots, \boldsymbol{T}_n\}$ be a partitioning of $\boldsymbol{O}$. Let $\boldsymbol{A}_i = \bigcup_{j=1}^{i-1} \boldsymbol{T}_j$, $\boldsymbol{B}_i = \boldsymbol{T}_i$, and $\boldsymbol{O}_i = \boldsymbol{A}_i \cup \boldsymbol{B}_i$. If $\mathcal{G}$ satisfies $tbk^{\boldsymbol{T}}$, then $\textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}(([\mathcal{G}] + tbk^{\boldsymbol{T}})(\boldsymbol{O}_i)) \equiv \textsc{Fci}(\mathcal{G}(\boldsymbol{O}_i) + mbk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i})$ for all $1 \leq i \leq n$.

Using Lemmas 8, 10, and 11,

$$\begin{aligned} \textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}(([\mathcal{G}] + tbk^{\boldsymbol{T}})(\boldsymbol{O}_i)) &\equiv \textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + tbk^{\boldsymbol{T}}) \\ &\equiv \textsc{Rm}^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}([\mathcal{G}(\boldsymbol{O}_i)] + ebk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}) \\ &\equiv \textsc{Fci}(\mathcal{G}(\boldsymbol{O}_i) + mbk^{\boldsymbol{A}_i}_{\boldsymbol{B}_i}). \end{aligned}$$

$\square$

References

Ali, R. A., Richardson, T. S., and Spirtes, P. (2009). Markov equivalence for ancestral graphs. *The Annals of Statistics*, 37:2808–2837.

Richardson, T. S. and Spirtes, P. (2002). Ancestral graph Markov models. *The Annals of Statistics*, 30:962–1030.

Spirtes, P., Glymour, C., and Scheines, R. (2000). *Causation, Prediction, and Search*. MIT Press.

Zhang, J. (2008). On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias. *Artificial Intelligence*, 172:1873–1896.