Supplement - Learning Sparse Causal Models is not NP-hard

Abstract

This article contains detailed proofs and additional examples related to the UAI-2013 submission 'Learning Sparse Causal Models is not NP-hard'.

The supplement follows the numbering in the main submission.

1 Graphical model terminology

For reference purposes a few basic graphical model concepts, terms and definitions.

A *mixed graph* $\mathcal{G}$ is a graphical model that can contain three types of edges between pairs of nodes: directed ($\rightarrow$), bi-directed ($\leftrightarrow$), and undirected ($-$). If there is an edge $X \rightarrow Y$ in $\mathcal{G}$ then X is a *parent* of its *child* Y, if $X \leftrightarrow Y$ then X and Y are *spouses* of each other, and if $X - Y$ then they are called *neighbours*. A *path* $\pi = \langle X_1, X_2, .., X_n \rangle$ is a sequence of nodes where each successive pair (X_i, X_{i+1}) along π is adjacent (connected by an edge) in $\mathcal{G}$. A *directed path* is a path of the form $X_1 \rightarrow X_2 \rightarrow .. \rightarrow X_n$. A *directed cycle* is a directed path from a node back to itself. A *directed acyclic graph* (DAG) is graph that contains only directed edges, but has no directed cycle. A node X is an *ancestor* of Y (and Y a *descendant* of X) if there is a directed path from X to Y in $\mathcal{G}$. A vertex Z is a *collider* on a path $\pi = \langle \ldots, X, Z, Y, \ldots \rangle$ if there are arrowheads at Z on both edges from X and Y, otherwise it is a *noncollider*.

In a DAG $\mathcal{G}$, a path $\pi = \langle X, .., Y \rangle$ is said to be *unblocked* relative to a set of vertices $\mathbf{Z}$, if and only if:

- every noncollider on π is not in $\mathbf{Z}$, and
- every collider along π is an ancestor of $\mathbf{Z}$,

otherwise the path is *blocked*. If there exists an unblocked path between X and Y relative to $\mathbf{Z}$ in $\mathcal{G}$ then X and Y are said to be *d-connected* given $\mathbf{Z}$; if there is no such path then X and Y are *d-separated* by $\mathbf{Z}$.

A mixed graph $\mathcal{G}$ is *ancestral*, iff an arrowhead at X on an edge to Y implies that there is no directed path from X to Y in $\mathcal{G}$, and there are no arrowheads at nodes with undirected edges. As a result, arrowhead marks can be read as 'is not an ancestor of', and all DAGs are ancestral. When applied to an ancestral graph, d-separation is also known as m-separation. An ancestral graph is *maximal* (MAG) if for any two non-adjacent vertices there is a set that separates them.

Throughout the rest of this article, X, Y and $\mathbf{Z}$ represent disjoint (subsets of) nodes, vertices, or variables in a graph. Every MAG $\mathcal{M}$ over nodes $\mathbf{V}$ corresponds to some underlying causal DAG $\mathcal{G}$ over variables $\mathbf{V} \cup \mathbf{L} \cup \mathbf{S}$, where the (possibly empty) sets of unobserved latent variables $\mathbf{L}$ and selection nodes $\mathbf{S}$ in $\mathcal{G}$ have been marginalized out, see (Richardson and Spirtes, 2002). The set $Adj(X)$ refers to the nodes adjacent to X in $\mathcal{M}$, and $An(X)$ represents the ancestors of X in $\mathcal{M}$. Similar for sets, i.e. $Z \in Adj(\mathbf{X})$ implies $\exists X \in \mathbf{X} : Z \in Adj(X)$; idem for $An(\mathbf{X})$.

2 *D*-separating sets

This part contains the proofs for section §4.1 in the main article. We start by formalizing some terminology on D-separation:

Definition 2. In a MAG $\mathcal{M}$, two nodes X and Y are ***D*-separated** by a set of nodes $\mathbf{Z}$ if and only if:

1. $X \perp\!\!\!\perp Y \,|\, \mathbf{Z}$,
2. $\forall \mathbf{Z}' \subseteq Adj(\{X, Y\}) : X \not\perp\!\!\!\perp Y \,|\, \mathbf{Z}'$.

If $\mathbf{Z}$ D-separates X and Y, then (X, Y) is a ***D*-sep link**. A node $Z \in \mathbf{Z}$ is a ***D*-sep node** for (X, Y) if:

1. $Z \notin Adj(\{X, Y\})$,
2. $\forall \mathbf{Z}' \subseteq Adj(\{X, Y\}) : X \not\perp\!\!\!\perp Y \,|\, \mathbf{Z}_{\setminus Z} \cup \mathbf{Z}'$.

In words: X and Y are D-separated by $\mathbf{Z}$ iff they are d-separated by $\mathbf{Z}$, and all sets that *can* separate X and Y contain at least one node $Z \notin Adj(\{X,Y\})$. Such a node $Z \in \mathbf{Z}$ that cannot be made redundant by nodes adjacent to X or Y is a D-sep node, and the relation between X and Y is called a D-sep link.

2.1 Proofs - Identifying D-sep edges

We rely on the following connection between in/dependencies and (non-)ancestorship in a MAG.

Lemma 2. For disjoint (subsets of) nodes $X, Y, Z, \mathbf{Z}$ from the observed variables $\mathbf{O}$ in a causal graph $\mathcal{G}$ with selection set $\mathbf{S}$,

$$(1)\quad X \not\perp\!\!\!\perp Y \,|\, \mathbf{Z} \cup [Z] \quad \Rightarrow \quad Z \notin An_{\mathcal{G}}(\{X,Y\} \cup \mathbf{Z} \cup \mathbf{S}).$$

$$(2)\quad X \perp\!\!\!\perp Y \,|\, [\mathbf{Z} \cup Z] \quad \Rightarrow \quad Z \in An_{\mathcal{G}}(\{X,Y\} \cup \mathbf{S}),$$

where square brackets indicate a *minimal* set of nodes.

Proof. See e.g. (Spirtes et al., 1999). □

Rule (2) in Lemma 2 not only applies to the nodes in the minimal separating set, but also to all other nodes on the paths between X and Y that are blocked by $\mathbf{Z}$.

Corollary 8. If $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$, then for all nodes $\mathbf{W}$ in $\mathcal{G}$ on a path π between X and Y that is blocked by $Z \in \mathbf{Z}$ it holds that $\mathbf{W} \subset An_{\mathcal{G}}(\{X,Y\} \cup \mathbf{S})$.

Proof. Follows immediately from Lemma 2, rule (2) in combination with Lemma 3.13 in (Richardson and Spirtes, 2002), whilst noting that every node $Z \in \mathbf{Z}$ in a minimal independence $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$ is a noncollider between two paths $\langle X, .., Z\rangle + \langle Z, .., Y\rangle$ that are unblocked given $\mathbf{Z}_{\setminus Z}$.

Alternatively: A path $\pi = \langle X, .., Y\rangle$ is blocked by $Z \in \mathbf{Z}$ if Z is a noncollider between two unblocked subpaths, $\pi_{XZ} = \langle X, .., Z\rangle$ and $\pi_{ZY} = \langle Z, .., Y\rangle$, given $\mathbf{Z} \setminus Z$. Therefore all other nodes on the unblocked subpath π_{XZ} are either one of the endpoints $\{X, Z\}$ (for which it holds trivially), a collider $W \in \mathbf{W}$ along π that is in $An(\mathbf{Z} \cup \mathbf{S})$, or a noncollider $U \in \mathbf{U}$ that is not in $\mathbf{Z}$. By Lemma 2 and transitivity, all colliders $W \in \mathbf{W}$ on π_{XZ} are also in $An_{\mathcal{G}}(\{X,Y\} \cup \mathbf{S})$, either directly or via nodes in $\mathbf{Z}$. The remaining noncolliders $U \in \mathbf{U}$ are on unblocked subpaths of π_{XZ} between two nodes from $\{X, Z\} \cup \mathbf{W}$. A noncollider U implies that there is at least one tail mark at U along π_{XZ} (otherwise it would be a collider). If this tail mark at U is part of an undirected edge, then there is selection bias on U, and so $U \in An(\mathbf{S})$. If not, the tail mark is part of a directed path along π_{XZ} to a node from $\{X, Z\} \cup \mathbf{W}$: as long this is not reached the next edge along π_{XZ} cannot be a bi-directed edge (for that would imply we already reached a collider from $\mathbf{W}$), nor an undirected edge (definition of ancestral graph), and so extend a directed path until it can no longer be extended (for finite length paths) and a node from $\{X, Z\} \cup \mathbf{W}$ is reached. But that implies that U is an ancestor of one of these nodes, and so by Lemma 2 and transitivity also of $\{X,Y\} \cup \mathbf{S}$. □

This property is used (often implicitly) in several Lemmas below. To prove Lemma 3 we first derive a connection between 'not separable by adjacent nodes' and non-ancestorship:

Lemma 9. In a MAG $\mathcal{M}$ for a causal DAG $\mathcal{G}$, if $X \perp\!\!\!\perp Y \,|\, \mathbf{Z}$, but X is not independent of Y given any subset of $Adj(X)$ in $\mathcal{M}$, then $Y \notin An_{\mathcal{G}}(X \cup \mathbf{S})$.

Proof. Given the $X \perp\!\!\!\perp Y \,|\, \mathbf{Z}$, there is no edge $X - Y$ in $\mathcal{M}$, and so any unblocked path $\langle X, U, .., Y\rangle$ between them goes via some node U adjacent to X. Let $\mathbf{U} = (Adj(X) \cap An(\{X,Y\} \cup \mathbf{S}))$ be the set of all nodes in $\mathcal{M}$ that are adjacent to X and ancestor of X, Y and/or the selection set $\mathbf{S}$. There are no unblocked paths $\pi_V = \langle X, V, .., Y\rangle$ given $\mathbf{U}$ in $\mathcal{M}$ that go via a node $V \notin \mathbf{U}$: such a path would start with an edge $X \ast\!\!\rightarrow V$, but without subsequent directed path $V \longrightarrow .. \longrightarrow Y$ (otherwise $V \in \mathbf{U}$), and so π_V must contain at least one collider. Let W be the first collider along π_V when starting from X, so that $X \ast\!\!\rightarrow V \longrightarrow .. \longrightarrow W \leftarrow\!\!\ast ..Y$ (possibly $W = V$). This implies $W \in An(\mathbf{U})$, but then also $V \in An(\mathbf{U})$, and so indirect $V \in An(\{X,Y\} \cup \mathbf{S})$, contrary $V \notin \mathbf{U}$. So any (remaining) unblocked path given $\mathbf{U}$ must be a path $\pi = \langle X \ast\!\!\rightarrow U \leftarrow\!\!\ast ..Y\rangle$, with $U \in \mathbf{U}$ a collider along π, i.e. with an arrowhead on the edge to X. This signifies U is not an ancestor of X and not (ancestor of) a node in $\mathbf{S}$, see (Richardson and Spirtes, 2002), which means that this $U \in \mathbf{U}$ must be ancestor of Y. But then in turn $Y \notin An(X \cup \mathbf{S})$, otherwise $U \in An(X \cup \mathbf{S})$ through the directed path via Y, contrary the arrowhead $X \ast\!\!\rightarrow U$ in $\mathcal{M}$. □

Lemma 3. In a MAG $\mathcal{M}$ for a causal DAG $\mathcal{G}$, if two nodes X and Y are D-separated by a minimal set $\mathbf{Z}$, then

1. $X \notin An_{\mathcal{G}}(\{Y\} \cup \mathbf{Z} \cup \mathbf{S})$
2. $Y \notin An_{\mathcal{G}}(\{X\} \cup \mathbf{Z} \cup \mathbf{S})$
3. $\forall Z \in \mathbf{Z} : Z \in An_{\mathcal{G}}(\{X,Y\} \cup \mathbf{S})$

Proof. In words: in a MAG $\mathcal{M}$, D-separable nodes X and Y are not ancestors of each other, nor of any node in $\mathbf{Z}$, and are also not subject to selection bias.
1. by Lemma 9 and the definition of D-separated nodes, X is not ancestor of Y and has no selection bias, together with the observation that if X was ancestor of any node in $\mathbf{Z}$, then by 3. and acyclicity it would also be ancestor of Y and/or $\mathbf{S}$, a contradiction;

2. similar for Y; 3. follows directly from Lemma 2, rule (2), given the fact that $\mathbf{Z}$ is minimal. $\square$

Next we introduce:

Definition 3. For a set of nodes $\mathbf{X}$ in a MAG $\mathcal{M}$, the set of its *adjacent ancestors* $AA(\mathbf{X})$ is defined as $AA(\mathbf{X}) = (Adj(\mathbf{X}) \cap An(\mathbf{X})) \setminus \mathbf{X}$.

Note that $Z \in An(\mathbf{X})$ in a MAG $\mathcal{M}$ implies that $Z \in An_{\mathcal{G}}(\mathbf{X} \cup \mathbf{S})$ in the corresponding underlying causal DAG $\mathcal{G}$. As a result, for D-separable $\{X, Y\}$, adjacent nodes with selection bias are also in $AA(\{X, Y\})$. We can freely add 'adjacent ancestors' to any separating set without destroying it:

Lemma 10. In a MAG $\mathcal{M}$, if $X \perp\!\!\!\perp Y \,|\, \mathbf{Z}$, then $\forall \mathbf{W} \subseteq (\{X, Y\} \cup \mathbf{Z}) : X \perp\!\!\!\perp Y \,|\, \mathbf{Z} \cup AA(\mathbf{W})$.

Proof. Adding the nodes in $AA(\mathbf{W})$ to the separating set one by one, then by rule (1) in Lemma 2, any node that creates a dependence cannot be ancestor of any node in $(\{X, Y\} \cup \mathbf{Z})\mathbf{W}$, contrary the definition of $AA(\mathbf{W})$. So all nodes leave the independence intact, and so $X \perp\!\!\!\perp Y \,|\, \mathbf{Z} \cup AA(\mathbf{W})$. $\square$

This leads to a result to identify D-sep nodes:

Lemma 11. In a MAG $\mathcal{M}$, if two nodes X and Y are D-separated by $\mathbf{Z}$, then also $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}_{AA} \cup \mathbf{Z}_{DS}]$, with $\mathbf{Z}_{DS} \subset \mathbf{Z}$, $\mathbf{Z}_{AA} \subseteq AA(\{X, Y\})$, $\mathbf{Z}_{AA} \cap \mathbf{Z}_{DS} = \varnothing$, and where all $Z \in \mathbf{Z}_{DS}$ are D-sep nodes for $X - Y$.

Proof. We use rules (1) and (2) in Lemma 2 to construct the two sets. First we remove nodes from $\mathbf{Z}$ one-by-one until no more can be removed to obtain a minimal $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}']$, with $\mathbf{Z}' \subseteq \mathbf{Z}$. By rule (2), all nodes in $\mathbf{Z}'$ are ancestor of X, Y, and/or $\mathbf{S}$ in $\mathcal{M}$. By rule (1) we can add all nodes that are ancestor of X and/or Y without destroying the independence, and so also the subset of those that are adjacent to X or Y. This gives $X \perp\!\!\!\perp Y \,|\, AA(\{X, Y\}) \cup \mathbf{Z}''$, where $\mathbf{Z}'' = \mathbf{Z}' \setminus AA(\{X, Y\})$ now containes the subset of nodes from $\mathbf{Z}'$ that are not adjacent to X and/or Y.

We obtain $\mathbf{Z}_{DS}$ by eliminating nodes from $\mathbf{Z}''$ one by one until no more nodes can be eliminated without destroying the independence, and so $X \perp\!\!\!\perp Y \,|\, AA(\{X, Y\}) \cup [\mathbf{Z}_{DS}]$. All nodes in $\mathbf{Z}_{DS}$ satisfy the definition of D-sep node: by construction none of them are adjacent to X or Y, and if there were *some* subset of $Adj(\{X, Y\})$ that could make a node $Z \in \mathbf{Z}_{DS}$ redundant then by Lemma 2.(2) this must be a subset of $AA(\{X, Y\})$, and so by Lemma 10 the independence should also be found given $AA(\{X, Y\}) \cup \mathbf{Z}_{DS} \setminus Z$.

Finally we can obtain $\mathbf{Z}_{AA}$ by eliminating superfluous nodes from $AA(\{X, Y\})$ one by one until no more can be removed without creating a dependence. At that point the D-separating set is minimal, and so $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}_{AA} \cup \mathbf{Z}_{DS}]$. By construction the sets $\mathbf{Z}_{AA}$ and $\mathbf{Z}_{DS}$ are also disjoint.

No additional nodes from $\mathbf{Z}_{DS}$ can be eliminated during/after the process of eliminating nodes from $AA(\{X, Y\})$: if $Z \in \mathbf{Z}_{DS}$ can be eliminated only *after* some node $Z_A \in AA(\{X, Y\})$ is eliminated, then putting back Z_A after Z is removed should create a dependence, which would imply by Lemma 2, rule (1) that Z_A is not ancestor of X and/or Y (or $\mathbf{S}$), contrary the definition of $AA(\{X, Y\})$. $\square$

Note that neither $\mathbf{Z}_{AA}$ nor $\mathbf{Z}_{DS}$ need be uniquely defined for a given D-separated $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$, but may depend on the order in which nodes are removed.

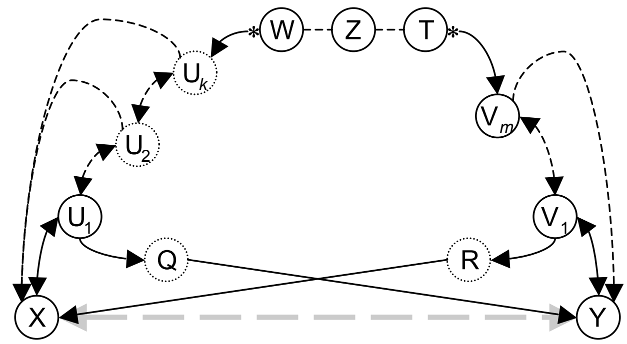

Figure 1: Path configuration for D-sep edge $X - Y$.

In the proof of Lemma 4 we rely on the fact that for each D-sep edge there is a path blocked by a D-sep node of the form depicted in Figure 1, which imposes six different identifiable dependence relations:

Lemma 12. In a MAG $\mathcal{M}$, if two nodes X and Y are D-separable, then there exists a path of nodes $X \leftarrow\!\rightarrow U_1 (.. \longleftrightarrow U_k) \leftarrow\!* W..Z..T *\!\rightarrow (V_m \longleftrightarrow ..) V_1 \longleftrightarrow Y$ (possibly with $W = Z$ and/or $Z = T$) in $\mathcal{M}$, together with minimal sets of nodes $\mathbf{Z}_{XW}, \mathbf{Z}_{YT}, \mathbf{Z}_{UV}$ such that:

1. $X \not\perp\!\!\!\perp W \,|\, \mathbf{Z}_{XW} \cup [U_1]$ and $X \not\perp\!\!\!\perp W \,|\, \mathbf{Z}_{XW} \cup [Y]$,
2. $Y \not\perp\!\!\!\perp T \,|\, \mathbf{Z}_{YT} \cup [V_1]$ and $Y \not\perp\!\!\!\perp T \,|\, \mathbf{Z}_{YT} \cup [X]$,
3. $U_1 \not\perp\!\!\!\perp V_1 \,|\, \mathbf{Z}_{UV} \cup [X]$ and $U_1 \not\perp\!\!\!\perp V_1 \,|\, \mathbf{Z}_{UV} \cup [Y]$.

Proof. By Lemma 11 we have $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}_{AA} \cup \mathbf{Z}_{DS}]$, with $\mathbf{Z}_{AA} \subseteq AA(\{X, Y\})$, and $\mathbf{Z}_{DS}$ a (sub)set of D-sep nodes not adjacent to X and/or Y.

The minimal separating set implies that D-sep node $Z \in \mathbf{Z}_{DS}$ is necessary to block an unblocked path between X and Y given $\mathbf{Z}_{AA} \cup \mathbf{Z}_{DS} \setminus Z$.

We now show this implies there is a path $\pi = \langle X \longleftrightarrow U_1 (\longleftrightarrow U_2 .. U_k) \leftarrow\!* W..Z..*\!\rightarrow V \longleftrightarrow Y \rangle$ in

$\mathcal{M}$, where all nodes U_i are adjacent to X but only U_1 has a bi-directed edge to X (similar for V at Y), and W is the first node along π starting from X that is not adjacent to X (possibly $W = Z$), see Figure 1.

Firstly, all paths between X and Y blocked by any node from $\mathbf{Z}_{AA} \cup \mathbf{Z}_{DS}$ must be *into* both X and Y: by Corollary 8 the first node U_1 encountered along π is in $An(\{X,Y\} \cup \mathbf{S})$. If the path starts with a tail from X then the same holds for X, in contradicition with Lemma 3. Idem for Y. Therefore all paths blocked by Z, including π must have $X \leftarrow\!* \,..\, *\!\rightarrow Y$.

Secondly, as the first and last step of the path blocked by Z are via nodes U_1 and V_1 adjacent to resp. X and Y that are also in the conditioning set $\mathbf{Z}_{AA} \cup \mathbf{Z}_{DS}$, it means that these must be colliders along π, leading to $X \longleftrightarrow U_1 \leftarrow\!* \,..\, *\!\rightarrow V_1 \longleftrightarrow Y$. (If other U_i also have a bidirected edge to X then redefine π with the closest such U_i to W as the new U_1; idem for V_j). As a result, the remaining U_i (if any) all have directed edges into X, as they are adjacent to X but not with a bi-directed edge, not with an edge $X \longrightarrow U_i$ for that would imply $X \in An(Y)$, and also not an undirected edge as a MAG cannot have an arrowhead on a node with an undirected edge.

Thirdly, there is always a first node W along π *not* adjacent to X, as $Z \in \pi$ is not adjacent to either X or Y, so possibly $W = Z$; idem for a node T along π when starting from Y. This means the first and last parts of π take on the configuration $X \longleftrightarrow U_1 (.. \longleftrightarrow U_k) \leftarrow\!* W$ resp. $T *\!\rightarrow (V_m \longleftrightarrow ..) V_1 \longleftrightarrow Y$ indicated in Figure 1, with noncollider Z somewhere in between.

Now for the in/dependence relations: given that X and W are not adjacent they are separated by some set $\mathbf{Z}_{XW}$ (not to be confused with $\mathbf{Z}_{AA}$ or $\mathbf{Z}_{DS}$). By construction, all $\{U_2, .., U_k\}$ are part of this set: U_k is needed to block the path $X \leftarrow U_k \leftarrow\!* W$. Conditioning on U_k unblocks the path $X \leftarrow U_{k-1} \leftrightarrow U_k \leftarrow\!* W$ so U_{k-1} is also needed, etc., all the way up to and including U_2 (but not U_1). This means there are unblocked paths *into* U_1 from both X and W given $\mathbf{Z}_{XW}$, and so also conditioning on U_1 would make X and W dependent, i.e. $X \not\perp\!\!\!\perp W \,|\, \mathbf{Z}_{XW} \cup [U_1]$. As Y is a descendant of U_1, it also implies $X \not\perp\!\!\!\perp W \,|\, \mathbf{Z}_{XW} \cup [Y]$. Idem for $Y \not\perp\!\!\!\perp T \,|\, \mathbf{Z}_{YT} \cup [V_1]$ and $Y \not\perp\!\!\!\perp T \,|\, \mathbf{Z}_{YT} \cup [X]$.

Furthermore, U_1 and V_1 cannot be adjacent in $\mathcal{M}$: they cannot be connected by a bi-directed edge, for that would make the path $\langle X, U_1, V_1, Y \rangle$ unblocked given $\mathbf{Z}_{AA} \cup \mathbf{Z}_{DS}$; they cannot be connected by an edge $U_1 \longrightarrow V_1$, for that would make U_1 ancestor of X via the path $U_1 \longrightarrow V_1 \longrightarrow .. \longrightarrow X$, contrary the bi-directed edge $X \longleftrightarrow U_1$; vice versa for an edge $U_1 \longleftarrow V_1$ (in other words: neither U_1 nor V_1 is ancestor of the other); and also not an undirected edge because of the arrowheads on other edges at both U_1 and V_1. Therefore they are conditionally independent given some minimal set $\mathbf{Z}_{UV}$. No descendant of U_1 or V_1 (including X and Y) can be part of that set, for that would imply either U_1 or V_1 was ancestor of the other, but similarly including X or Y in the conditioning set would make them dependent given that both have unblocked paths to U_1 and V_1 given $\mathbf{Z}_{UV}$. Therefore we can find both $U_1 \not\perp\!\!\!\perp V_1 \,|\, \mathbf{Z}_{UV} \cup [X]$ and $U_1 \not\perp\!\!\!\perp V_1 \,|\, \mathbf{Z}_{UV} \cup [Y]$. $\square$

By Lemma 2, rule (1), each node in Lemma 12 that destroys one of the three independencies cannot be ancestor of any node in that independence, nor of the selection set $\mathbf{S}$, and so leads to identifiable invariant edge-marks. This motivates the introduction:

Definition 4. The **Augmented Skeleton** $\mathcal{G}^+$ is obtained from a skeleton $\mathcal{G}$ by adding all invariant arrowheads that follow from single node minimal dependencies $X \not\perp\!\!\!\perp Y \,|\, \mathbf{Z} \cup [W]$ by Lemma 2, rule(1).

We assume that we have a (minimal) independence for each edge eliminated in $\mathcal{G}$ (as provided by PC), and that we can query an independence oracle for the subsequent dependencies. For D-sep edges in the augmented skeleton $\mathcal{G}^+$ this implies the following pattern:

Lemma 4. For a MAG $\mathcal{M}$, let X and Y be D-separable nodes that are adjacent in the corresponding augmented skeleton $\mathcal{G}^+$. If there are no edges in $\mathcal{G}^+$ between (other) D-separable pairs of nodes in $An(\{X,Y\})$, then $\mathcal{G}^+$ contains the following pattern: $U \leftrightarrow X \leftrightarrow Y \leftrightarrow V$, with U and V not adjacent in $\mathcal{G}^+$, and paths $V.. \rightarrow X$ and $U.. \rightarrow Y$ that do not go against an arrowhead.

Proof. We use the construction in Lemma 12 with $U = U_1$, $V = V_1$. As X and U are adjacent in $\mathcal{M}$, they are also adjacent in $\mathcal{G}^+$. Similarly for Y and V. Nodes X and Y are also (still) presumed to be adjacent in $\mathcal{G}^+$. By lemma 12, case 1. and 2., we can detect that X and Y are not ancestor of each other, giving invariant arrowheads: $X \longleftrightarrow Y$. Lemma 12 also gives $U \longleftrightarrow X$ and $V \longleftrightarrow Y$. The assumption 'no undiscovered D-sep links between $An(\{X,Y\})$' ensures that the three required independencies in Lemma Lemma 12 are already found, including the one eliminating the link between U and V in $\mathcal{M}$, and so all corresponding invariant arrowheads can be identified. As $V \in An(X)$ there has to be a path from V that can be(come) oriented as a directed path into X, which means it cannot contain an invariant arrowhead in the opposite direction; idem for $U \in An(Y)$. $\square$

Lemma 5. In a MAG $\mathcal{M}$, for two pairs of D-separable nodes $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$ and $U \perp\!\!\!\perp V \,|\, [\mathbf{W}]$, if $X \in \mathbf{W}$ and/or $Y \in \mathbf{W}$, then $U \notin \mathbf{Z}$ and/or $V \notin \mathbf{Z}$.

Proof. Suppose that $X \in \mathbf{W}$, and $\{U, V\} \subset \mathbf{Z}$. By Lemma 3, $X \in An(U)$ and/or $X \in An(V)$, as none of the D-separable nodes $\{X, Y, U, V\}$ are subject to selection bias. If $X \in An(U)$, then by acyclicity $U \in An(Y)$, which would in turn imply $X \in An(Y)$, contrary Lemma 3. Idem for the case $X \in An(V)$. Therefore U and V cannot both be in the D-separating set $\mathbf{Z}$ for X and Y. Note that if *both* $\{X, Y\} \subset \mathbf{W}$, then $U \notin \mathbf{Z}$ and $V \notin \mathbf{Z}$. □

So two D-sep links cannot be in each other's D-separating set. As a result:

Corollary 13. In a MAG $\mathcal{M}$, all D-sep links can be found by repeatedly (and exclusively) checking the augmented skeleton $\mathcal{G}^+$ for edges that appear as the middle of the bidirected triple from Lemma 4, while updating $\mathcal{G}^+$ for each D-sep link found.

Proof. Let $\mathcal{G}$ be the skeleton of a MAG $\mathcal{M}$, possibly with additional edges in $\mathcal{G}$ that all correspond to D-sep links in $\mathcal{M}$, and let $\mathcal{G}^+$ be the augmented skeleton of $\mathcal{G}$ w.r.t. the MAG $\mathcal{M}$. Then, as long as there are one or more edges in $\mathcal{G}^+$ that are not in $\mathcal{M}$, then by Lemma 5 at least one of these edges will have no unidentified D-sep links (edges in $\mathcal{G}$ that are not in $\mathcal{M}$) between its ancestors, and so by Lemma 4 this D-sep link will show up in $\mathcal{G}^+$ as the middle edge of the bidirected triple. Given a procedure to establish whether or not a candidate edge satisfying the bidirected pattern is a D-sep link (e.g. FCI's Possible-D-SEP search), then testing all candidate edges, while updating $\mathcal{G}^+$ for each D-sep link identified (remove edge and recompute arrowheads for new bidirected triples) until no more can be found, is guaranteed to find all D-sep links. This means that at the end the skeleton of $\mathcal{G}^+$ matches that of $\mathcal{M}$, and all arrowheads in $\mathcal{G}^+$ are also in $\mathcal{M}$. □

This greatly improves the practical running speed of FCI, as often no or hardly any edges need to be checked, but in itself it is not sufficient to reduce the overall complexity to polynomial time, as even a single edge may still require searching through all subsets of order N nodes. The next section shows how a different search strategy can resolve this problem.

2.2 Proofs - Capturing the D-sep nodes

All D-sep nodes for a pair (X, Y) also appear in another minimal conditional independence:

Lemma 6. In a MAG $\mathcal{M}$, if $Z \in \mathbf{Z}$ is a D-sep node in $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$, then Z is also part of a minimal separating set between two nodes from $\{X, Y\} \cup \mathbf{Z}_{\setminus Z} \cup AA(\{X, Y\})$, neither of which have selection bias.

Proof. Let $\mathbf{Z}^* = \mathbf{Z} \cup AA(\{X, Y\})$. Then, by Lemma 10 and the definition of a D-sep node, we have $X \perp\!\!\!\perp Y \,|\, \mathbf{Z}^*$ and $X \not\perp\!\!\!\perp Y \,|\, \mathbf{Z}^* \setminus Z$, which implies that Z is a noncollider on some path between X and Y that is unblocked given $\mathbf{Z}^*_{\setminus Z}$.

Consider the MAG $\mathcal{M}'$ obtained by marginalizing out all nodes not in $\{X, Y\} \cup \mathbf{Z}^*$, in accordance with the rules in (Richardson and Spirtes, 2002). As a MAG is closed under marginalization and conditioning, it follows that all ancestral and independence relations between $\{X, Y\} \cup \mathbf{Z}^*$ in $\mathcal{M}$ are also in $\mathcal{M}'$.

There are two possible cases:

(A) node Z is not adjacent to X and/or Y in $\mathcal{M}'$,

(B) node Z is adjacent to X or Y in $\mathcal{M}'$.

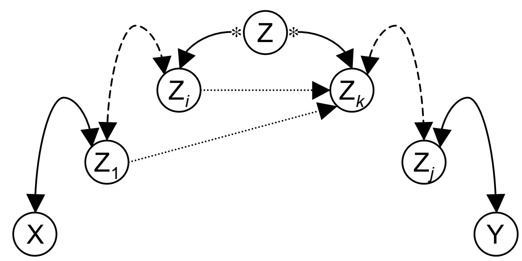

Figure 2: Canonical path in $\mathcal{M}'$ blocked by D-sep node Z.

Case (A):
If Z is not adjacent to X and/or Y in $\mathcal{M}'$ then, by the same rationale as in the proof of Lemma 12, Z blocks a path $\pi = X \leftrightarrow Z_1 \leftrightarrow .. Z_i \leftarrow\!* Z *\!\rightarrow Z_k .. \leftrightarrow Z_j \leftrightarrow Y$, but now without any other intermediate nodes that are not conditioned on, as depicted in Figure 2.
Node Z is then a noncollider between two nodes $Z_i, Z_k \in \mathbf{Z}^*$, and so if these two are not adjacent in $\mathcal{M}'$, then Z is part of some minimal independence $Z_i \perp\!\!\!\perp Z_k \,|\, [Z \cup ..]$, and the lemma is satisfied. If not, i.e. if there is an edge $Z_i *\!-\!* Z_k$ in $\mathcal{M}'$, then it cannot be an undirected edge (as that is invalid in combination with the other arrowheads at Z_i/Z_k), nor can it be a bi-directed edge $Z_i \leftrightarrow Z_k$ (for then the path $X \leftrightarrow Z_1..Z_i \leftrightarrow Z_k..Z_j \leftrightarrow Y$ would be unblocked given $\mathbf{Z}^*$). That leaves only the possibility of a directed edge, say $Z_i \rightarrow Z_k$ (indicated as dotted arc, see Figure 2). But then if the collider prior to Z_i is not adjacent to Z_k in $\mathcal{M}'$, then Z is part of minimal independence $Z_{i-1} \perp\!\!\!\perp Z_k \,|\, [\{Z_i, Z\} \cup ..]$. If not, i.e. there is

an edge $Z_{i-1} *\!-\!* Z_k$, then by the same rationale as before it cannot be an undirected edge nor a bi-directed edge. But it also cannot be an edge $Z_{i-1} \leftarrow Z_k$, for that would make Z_i ancestor of Z_{i-1} (by the directed path via Z_k), contrary the non-ancestor arrowheads at $Z_{i-1} \leftrightarrow Z_i$. Therefore it must be an arc $Z_{i-1} \rightarrow Z_k$.

We can repeat this argument all the way until we reach Z_1: if any collider is not adjacent to Z_k then Z appears in the minimal conditional independence, otherwise they all have directed arcs into Z_k, including $Z_1 \rightarrow Z_k$. But then X and Z_k cannot be adjacent in $\mathcal{M}'$, because now the edge $X \rightarrow Z_k$ is also not allowed, for that would make X ancestor of Z_k, in contradiction with Lemma 3. So then X and Z_k are separated in $\mathcal{M}'$, and all colliders $Z_1..Z_i$ are needed to separate them, and so Z as well, i.e. then $X \perp\!\!\!\perp Z_k \,|\, [\{Z_1, .., Z_i, Z\} \cup ..]$.

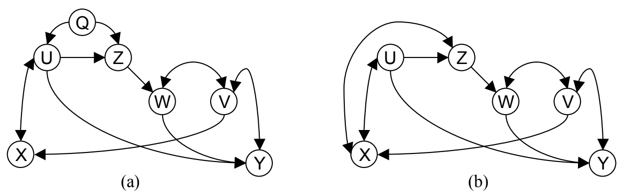

Figure 3: Example for case (B) in proof of Lemma 6. (a) MAG $\mathcal{M}$ with D-separated $X \perp\!\!\!\perp Y \,|\, [U, V, Z]$; (b) Marginal $\mathcal{M}'$ with D-sep node Z now adjacent to X.

Case (B):
If one or more nodes that are marginalized out from $\mathcal{M}'$ were needed to separate Z from X (or Y), then in $\mathcal{M}'$ node Z can be adjacent to X, as depicted in Figure 3.

First we show that this case implies that if Z is now adjacent to X, then $X \leftrightarrow Z$ in $\mathcal{M}'$, and Z is not adjacent to Y in $\mathcal{M}'$. Given that in $\mathcal{M}$ nodes X and Z can be separated by some set, but apparently not by any subset $AA(X) \subset \mathbf{Z}^*$ (as they are adjacent in $\mathcal{M}'$), it follows by Lemma 9 that $Z \notin An(X)$, which together with Lemma 3.1 implies $X \leftrightarrow Z$ in $\mathcal{M}'$. But if Z is then also adjacent to Y then the same argument applies to $Z \leftrightarrow Y$, which would mean an unblocked path $X \leftrightarrow Z \leftrightarrow Y$ when conditioning on Z, in contradiction with $X \perp\!\!\!\perp Y \,|\, \mathbf{Z}^*$.

Therefore, without loss of generalization we assume in case (B) that $X \leftrightarrow Z$ and Z not adjacent to Y in $\mathcal{M}'$. There has to be some path $\pi_{ZY} = \langle Z, W, .., Y \rangle$ in $\mathcal{M}'$ that is unblocked given $\mathbf{Z}^* \setminus Z$, otherwise Z is not needed to separate X and Y. This path must be *out of* Z (start with a tail), otherwise conditioning on Z would unblock the path $\langle X, Z \rangle + \pi_{ZY}$. There is also at least a node W (actually at least two) between Z and Y on π_{ZY}, as Z is not adjacent to Y. As all nodes except $\{X, Y\}$ in $\mathcal{M}'$ are in the conditioning set $\mathbf{Z}^*$, the path π_{ZY} must take the general form $Z \rightarrow W \leftrightarrow .. \leftrightarrow Y$. That means that W cannot be adjacent to X in $\mathcal{M}'$, because an arc $X \rightarrow W$ contradicts Lemma 3.1, an arc $X \leftarrow W$ would make Z an ancestor of X contrary the arrowheads at the presumed $X \leftrightarrow Z$, and a bi-directed edge $X \leftrightarrow W$ would leave an unblocked path $X \leftrightarrow W \leftrightarrow .. \leftrightarrow Y$ contrary $X \perp\!\!\!\perp Y \,|\, \mathbf{Z}^*$. Therefore there must be at least some set $\mathbf{Z}_{XW}$ that separates X and W, and by construction Z must be part of any such set in $\mathcal{M}'$, ergo $\exists \mathbf{Z}_{XW} \subset \mathbf{Z}^*, Z \in \mathbf{Z}_{XW} : X \perp\!\!\!\perp W \,|\, [\mathbf{Z}_{XW}]$.

For example in Figure 3(b), node Z has a bi-directed edge $X \leftrightarrow Z$ to X in $\mathcal{M}'$, and an unblocked path $Z \leftarrow W \leftrightarrow V \leftrightarrow Y$. Node W cannot be adjacent to X, and so we are indeed guaranteed to find Z from $X \perp\!\!\!\perp W \,|\, [Z, V]$.

Finally, as in both case (A) and (B) all nodes except Z along the path π have at least one incoming arrowhead on an edge in $\mathcal{M}'$ it follows that none of these other nodes can be subject to selection bias. □

Next we show that we do not need to know all minimal separating sets, but that it suffices if we have at least *one* minimal separating set per edge that is removed.

For a MAG $\mathcal{M}$, we introduce the following definitions:

Definition 5. An **independence set** $\mathcal{I} \subseteq \mathcal{I}(\mathcal{M})$ is a (sub)set of all minimal independence statements consistent with MAG $\mathcal{M}$, which contains at least one separating set for each pair of nonadjacent nodes in $\mathcal{M}$.

$Msep(\mathbf{X}, \mathcal{I})$ is the union of all nodes that appear in *some* minimal set in $\mathcal{I}$ between nodes from $\mathbf{X}$.

And a recursive definition for a set of separating nodes:

Definition 6. Let $\mathcal{I}$ be an independence set, then for a set $\mathbf{X}$ the **hierarchy** $HIE(\mathbf{X}, \mathcal{I})$ is the union of $\mathbf{X}$ and all nodes that appear in a minimal separating set in $\mathcal{I}$ between any pair of nodes in $HIE(\mathbf{X}, \mathcal{I})$.

By Lemma 2 all nodes in $HIE(\mathbf{X}, \mathcal{I})$ are ancestors of one or more nodes from $\mathbf{X}$, and we have the following straightforward relations

Lemma 14. For a given graph $\mathcal{M}$, we have that:

(a) $MSep(\{X, Y\}, \mathcal{I}(\mathcal{M})) \subseteq HIE(\{X, Y\}, \mathcal{I}(\mathcal{M}))$,
(b) $HIE(\{X, Y\}, \mathcal{I}) \subseteq HIE(\{X, Y\}, \mathcal{I}(\mathcal{M}))$, and
(c) $\exists \mathbf{Z} : X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$
$\Leftrightarrow X \perp\!\!\!\perp Y \,|\, MSep(\{X, Y\}, \mathcal{I}(\mathcal{M}))$
$\Leftrightarrow X \perp\!\!\!\perp Y \,|\, HIE(\{X, Y\}, \mathcal{I})_{\setminus \{X,Y\}}$
$\Leftrightarrow X \perp\!\!\!\perp Y \,|\, HIE(\{X, Y\}, \mathcal{I}(\mathcal{M}))_{\setminus \{X,Y\}}$.

Proof. For (a) it suffices to note that all minimal separating sets between X and Y are also present in all minimal separating sets between $An(X, Y) \supseteq \{X, Y\}$.

For (b), *any* separating set present in $\mathcal{I} \subset \mathcal{I}(\mathcal{M})$ is by definition also present in $\mathcal{I}(\mathcal{M})$.
Finally, (c) $\Rightarrow$ follows from the fact that all nodes in minimal separating sets between (ancestors) of X and Y are also ancestors of X or Y, and so adding them to any minimal separating set cannot induce a dependence, as that would imply that that nodes is *not* an ancestor of either X or Y. Therefore all such nodes can be added to any arbitrary (minimal) separating while keeping the independence. (c) $\Leftarrow$ follows from the fact that any separating set can be turned into a minimal separating set by removing nodes one-by-one while leaving the independence intact, until no more nodes can be eliminated. Note that by definition $\mathbf{Z} \subseteq HIE(\{X,Y\}, \mathcal{I}(\mathcal{M}))$. □

We now go on to show in Lemma 16 that for a MAG $\mathcal{M}$ it is sufficient to have just a single minimal independence for each separable pair of nodes (already found by PC if no undetected D-sep edges are left between nodes in $An(\{X,Y\})$, except for $X - Y$ itself) in order to find a set of nodes that is guaranteed to contain all nodes needed to separate X and Y. This may be intuitively obvious: the known sound and complete FCI algorithm also only requires a single minimal independence per separable pair of nodes, which implies that this contains sufficient information to infer all valid independence statements. Indeed, we could skip the results below and simply read the full set of separating nodes from an intermediate constructed PAG (provided there are no undiscovered D-sep links in the ancestral set), but it is instructive (and much more efficient in practice) to show that they can be added directly without recourse to an intermediate graph.

In Lemma 6, we found that the required D-sep nodes appear in some other separating set. But for a pair of separable nodes in a MAG there can be many different separating sets: some (*necessary*) nodes appear in all of them, other (*optional*) nodes only in some. This opens the possibility that we may miss certain D-sep nodes if they are optional in the separating set from Lemma 6. Fortunately, we can show that even in that case these nodes can still be found, as all optional nodes in a minimal separating set are necessary in some other minimal separating set between nodes from $An(\{X,Y\})$.

Lemma 15. In a MAG $\mathcal{M}$, if $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$ and $X \perp\!\!\!\perp Y \,|\, [\mathbf{W}]$, with $W \in (\mathbf{W} \setminus \mathbf{Z})$ an optional separating node, then W is part of a minimal conditional independence $X \perp\!\!\!\perp Z \,|\, [\mathbf{W}']$ and/or $Y \perp\!\!\!\perp Z \,|\, [\mathbf{W}']$ between another optional separating node $Z \in (\mathbf{Z} \setminus \mathbf{W})$ and at least one of the end nodes, X and/or Y, with $W \in \mathbf{W}' \subseteq \mathbf{W}$.

Proof. First note that any optional node in a minimal

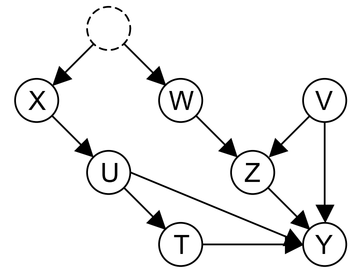

Figure 4: Two minimal separating sets: $X \perp\!\!\!\perp Y \,|\, [U, W]$ and $X \perp\!\!\!\perp Y \,|\, [U, V, Z]$, with optional node W necessary in $X \perp\!\!\!\perp Z \,|\, [W]$, and optional nodes $\{V, Z\}$ necessary in $W \perp\!\!\!\perp Y \,|\, [U, V, Z]$.

separating set cannot be adjacent to both X and Y, for if it was a non-collider between the two then it would always be needed to separate them, and if it was a collider between them then it was ancestor of neither X nor Y, and so not appear in any minimal separating set between them. Therefore each optional node in a minimal separating set is itself separable from either X and/or Y

From the given, for any optional node $W \in (\mathbf{W} \setminus \mathbf{Z})$, the set $\mathbf{Z}$ *must* block either all paths from X or from Y, otherwise $\mathbf{Z}$ would not separate X and Y because there would still be an unblocked path without W: if W were a connecting noncollider between the two unblocked paths $\langle X, .., W \rangle$ and $\langle W, .., Y \rangle$ then the path $\langle X, .., W, .., Y \rangle$ would be unblocked given $\mathbf{Z}$; if W were a collider between these paths, then by Lemma 2 it would be in $(\{X,Y\})$, as a collider cannot have selection bias, and so this directed path from W to X or Y must be blocked by some node from $\mathbf{Z}$, which in turn would unblock the path $\langle X, .., W, .., Y \rangle$ via collider $W \in (\mathbf{Z})$, again contradicting the given $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$.

Without loss of generality, assume that for a given $W_i \in (\mathbf{W} \setminus \mathbf{Z})$, the set $\mathbf{Z}$ blocks all paths between W_i and Y, so that $W_i \perp\!\!\!\perp Y \,|\, \mathbf{Z}$. For that, not all nodes may be needed, and we can eliminate nodes one by one until we obtain a subset $\mathbf{Z}_{W_iY} \subseteq \mathbf{Z}$ so that $W \perp\!\!\!\perp Y \,|\, [\mathbf{Z}_{W_iY}]$.

Furthermore, each optional $Z_j \in (\mathbf{Z} \setminus \mathbf{W})$ appears in at least one such $\mathbf{Z}_{W_iY}$ (or $\mathbf{Z}_{XW_i}$) for some optional $W_i \in (\mathbf{W} \setminus \mathbf{Z})$. By contradiction: suppose that optional Z_j does not appear in *any* minimal separating subset between one of the optional nodes in $\mathbf{W} \setminus \mathbf{Z}$ and X or Y. Then by definition $(\mathbf{W} \setminus \mathbf{Z})$ suffices to block all (remaining) unblocked paths between X and Y given $(\mathbf{Z} \cap \mathbf{W})$, and the given implies that the $(\mathbf{Z} \setminus \mathbf{W}) \setminus Z_j$ suffices to block all paths blocked by $(\mathbf{W} \setminus \mathbf{Z})$ (given $\mathbf{Z} \cap \mathbf{W}$). But that means that the union $((\mathbf{Z} \setminus \mathbf{W}) \setminus Z_j) \cup (\mathbf{Z} \cap \mathbf{W}) = (\mathbf{Z} \setminus Z_j)$ also suffices to block the same paths which implies that $\mathbf{Z}$ was not minimal, contrary the given. □

Example 1. To illustrate the proof, in Figure 4 we have $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$ and $X \perp\!\!\!\perp Y \,|\, [\mathbf{W}]$ for $\mathbf{Z} = \{U, V, Z\}$ and $\mathbf{W} = \{U, W\}$, with optional separating nodes $\{V, Z\}$ resp. W, and with necessary node $U \in (\mathbf{Z} \cap \mathbf{W})$ appearing in both. Node W blocks all remaining unblocked paths between X and $\{V, Z\}$ given U, and indeed W appears in a minimal separating set between X and optional node $Z \in \mathbf{Z}$: $X \perp\!\!\!\perp Z \,|\, [W]$. Conversely optional nodes $\{V, Z\}$ block all remaining unblocked paths between W and Y given U, and so $\{V, Z\}$ appear in the minimal separating set between Y and optional node $W \in \mathbf{W}$: $W \perp\!\!\!\perp Y \,|\, [U, V, Z]$.

So all optional nodes in two alternative minimal separating sets necessarily appear in a minimal separating set between one or more of the other optional nodes and one or both of the end nodes. If in turn this pair of nodes also has multiple minimal separating sets (optional nodes), then the same argument can be applied recursively to the set found for this one. Each step 'zooms in' further on the independence structure in the MAG, so that ultimately every optional node will appear as a necessary node in *some* minimal independence, starting from the minimal separating set originally found for X and Y.

From this it follows that:

Lemma 16. For two non-adjacent nodes, X and Y, in a MAG $\mathcal{M}$: $HIE(\{X, Y\}, \mathcal{I}) = HIE(\{X, Y\}, \mathcal{I}(\mathcal{M}))$.

Proof. $HIE(\{X, Y\}, \mathcal{I})$ and $HIE(\{X, Y\}, \mathcal{I}(\mathcal{M}))$ both recursively add all nodes that appear in the known minimal separating sets between X and Y and other nodes in the set. From Lemma 15, while adding such nodes recursively, all nodes that can appear in *some* minimal independence between X and Y will appear as necessary separators in some pair of nodes added this way, provided at least one minimal separator set is known for each separable pair. Therefore all nodes that can be added at some stage in $HIE(\{X, Y\}, \mathcal{I}(\mathcal{M}))$ (when knowing all minimal separating sets for all nonadjacent pairs) will also be found at some stage by $HIE(\{X, Y\}, \mathcal{I})$ (with just one minimal separating set per pair), and so for any $\mathcal{I}$ the two sets will be identical. □

With this the key result becomes:

Lemma 7. For a pair of D-separable nodes $\{X, Y\}$ in a MAG $\mathcal{M}$, if independence set $\mathcal{I}$ contains at least one minimal separating set for each of the nonadjacent nodes in the ancestors of X and/or Y in $\mathcal{M}$ (except for $X - Y$ itself), then $HIE(Adj(X, Y) \cap An(X, Y), \mathcal{I})_{\setminus \{X,Y\}}$ is a D-separating set for $\{X, Y\}$.

Proof. In words: if X and Y can be D-separated in $\mathcal{M}$, and we have at least one minimal separating set for each of their non-adjacent ancestors, then they are separated by the hierarchy implied by adjacent ancestors of X and/or Y. From Lemma 16, as long as at least one minimal separating set between non-adjacent ancestor nodes is in $\mathcal{I}$, then the hierarchy does not depend on which specific separating sets are found.

Furthermore, all nodes in $HIE(Adj(X, Y) \cap An(X, Y), \mathcal{I})$ are ancestor of X and/or Y, and so can be added to a D-separating set without destroying the independence, all nodes adjacent to X and/or Y that could possibly play a role in the D-separation are included in the set $Adj(X, Y) \cap An(X, Y)$. By Lemma 6, all (non-adjacent) D-sep nodes appear in some separating set between a pair of other nodes in the D-separating set that do not have selection bias, and by Lemma 12 all will be found in the hierarchy when recursively adding nodes.

As a result then $HIE(Adj(X, Y) \cap An(X, Y), \mathcal{I})$ contains a superset of a minimal D-separating set for D-sep link $X - Y$ in which no node unblocks another path, and so it is also a D-separating set. □

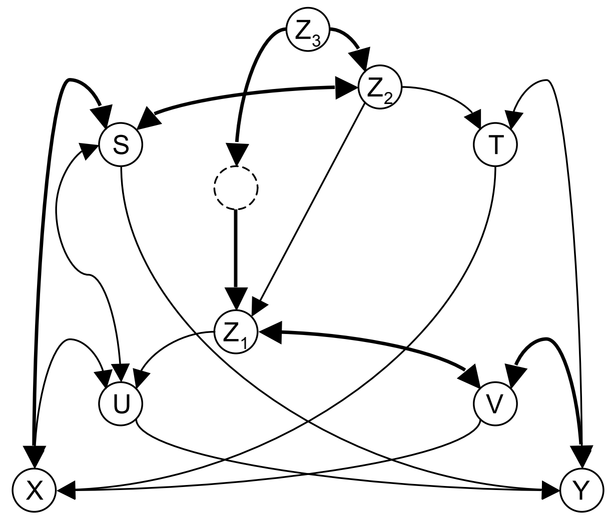

Figure 5: Illustration that it is not sufficient to only look at separating sets between nodes adjacent to $\{X, Y\}$, but that it is actually necessary to include the full recursive hierarchy: nodes X and Y are D-separated by $\{S, T, U, V, Z_1, Z_2, Z_3\}$, but Z_3 (blocking the path in bold) is *only* present (and necessary) in $S \perp\!\!\!\perp Z_1 \,|\, [Z_2.Z_3]$, with $Z_1 \notin Adj(\{X, Y\})$.

References

D. Koller and N. Friedman. *Probabilistic Graphical Models: Principles and Techniques*. The MIT Press, 2009.

J. Pearl. *Causality: models, reasoning and inference.*

Cambridge University Press, 2000.

T. Richardson and P. Spirtes. Ancestral graph Markov models. *Ann. Stat.*, 30(4):962–1030, 2002.

P. Spirtes, C. Meek, and T. Richardson. An algorithm for causal inference in the presence of latent variables and selection bias. In *Computation, Causation, and Discovery*, pages 211–252. AAAI Press, Menlo Park, CA, 1999.

P. Spirtes, C. Glymour, and R. Scheines. *Causation, Prediction, and Search.* The MIT Press, Cambridge, Massachusetts, 2nd edition, 2000.