Supplement - Learning Sparse Causal Models is not NP-hard

Tom Claassen[*], **Joris M. Mooij**[‡], **and Tom Heskes**[*]
[*] Institute for Computer and Information Science, Radboud University Nijmegen
[‡] Informatics Institute, University of Amsterdam
The Netherlands

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 Preliminaries

For reference purposes a few basic graphical model concepts, terms and definitions. For details, see e.g., Richardson and Spirtes (2002).

1.1 Graphical model terminology

A *mixed graph* $\mathcal{G}$ is a graphical model that can contain three types of edges between pairs of nodes: directed ($\leftarrow$, $\rightarrow$), bidirected ($\leftrightarrow$), and undirected ($-$). In this paper we only consider graphs with at most one edge between each pair of nodes, and with no node with an edge to itself. 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* $\pi = \langle X_1, X_2, \ldots, X_n \rangle$ is an ordered sequence of distinct 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 $X_1 \rightarrow X_2 \rightarrow \ldots \rightarrow X_n$. A *directed cycle* is a directed path from X_1 to X_n in combination with a directed edge $X_n \rightarrow X_1$. A *directed acyclic graph* (DAG) is a graph that contains only directed edges, but has no directed cycle. The *skeleton* $\mathcal{S}$ of a graph $\mathcal{G}$ is the undirected graph corresponding to the structure of $\mathcal{M}$, so that for each edge in $\mathcal{G}$ there is a undirected edge in $\mathcal{S}$. 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 $\mathcal{G}$, or $X = Y$. 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, i.e., if $X *\!\!\rightarrow Z \leftarrow\!\!* Y$ (where the symbol $*$ stands for either an arrowhead mark or a tail mark), otherwise it is a *noncollider*. A *trek* is a path without colliders.

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

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

otherwise the path is *blocked*. We say that a path $\pi = \langle X, \ldots, Y \rangle$ is *blocked by node* $Z \in \mathbf{Z}$ iff π is blocked given $\mathbf{Z}$, but unblocked relative to $\mathbf{Z}_{\setminus Z}$. 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 $\mathcal{M}$ is an *ancestral graph* (AG) iff an arrowhead at X on an edge to Y implies that there is no directed path from X to Y in $\mathcal{M}$, 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. In an ancestral graph $\mathcal{M}$ a node X is said to be *anterior* to a node Y if there is a so-called *anterior path* from X to Y in $\mathcal{M}$ of the form $X - \ldots - (Z) \rightarrow \ldots \rightarrow Y$, possibly with $Z = X$ (no undirected part) or with $Z = Y$ (no directed part), or if $X = Y$. Arrowhead marks in an ancestral graph can therefore also be read as 'is not anterior to'. When applied to an ancestral graph d-separation is also known as m-separation.[1] An ancestral graph is *maximal* (MAG) if for any two nonadjacent vertices there is a set that separates them. A path π between two nodes (X, Y) in an ancestral graph is *inducing with respect to a set of nodes* $\mathbf{Z}$ iff every collider on π is ancestor of X or Y, and every noncollider is in $\mathbf{Z}$. Inducing paths w.r.t $\mathbf{Z} = \varnothing$ are called *primitive*.

Throughout the rest of this article, X, Y and $\mathbf{Z}$ rep-

[1] Sometimes m-connected is defined using 'anterior' instead of 'ancestor' in condition (2) of 'unblocked', but as colliders have no undirected edges these two are equivalent.

resent disjoint (subsets of) nodes (vertices, variables) in a graph, with sets denoted in boldface. The set $Adj(X)$ refers to the nodes adjacent to X in an AG $\mathcal{M}$, $An(X)$ represents the ancestors of X in $\mathcal{M}$, and $Ant(X)$ the nodes anterior to X in $\mathcal{M}$. Similar for sets, i.e., $X \in Adj(\mathbf{Z})$ implies $\exists Z \in \mathbf{Z} : X \in Adj(Z)$; idem for $An(\mathbf{Z})$ and $Ant(\mathbf{Z})$.

Every (maximal) ancestral graph $\mathcal{M}$ 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 and conditioned out, see (Richardson and Spirtes, 2002). We denote the ancestors of $\mathbf{W} \subseteq \mathbf{V} \cup \mathbf{L} \cup \mathbf{S}$ in $\mathcal{G}$ as $An_{\mathcal{G}}(\mathbf{W})$, where the subscript $\mathcal{G}$ highlights that the ancestorship relation is with respect to the underlying DAG $\mathcal{G}$ instead of $\mathcal{M}$. Pairs of nodes in $\mathbf{V}$ that share a common ancestor in the subgraph of $\mathcal{G}$ over $\mathbf{L}$ are said to be *confounded*. Nodes in $An_{\mathcal{G}}(\mathbf{S})$ are said to be subject to *selection bias*. Both confounding and selection bias can give rise to links between nodes in a MAG, where confounding is associated with bidirected edges and selection bias with undirected edges.

The following helpful properties for reading ancestral information from a (M)AG $\mathcal{M}$ corresponding to an underlying causal DAG $\mathcal{G}$ are shown in (Richardson and Spirtes, 2002):

$$
\begin{array}{lll}
(1)\ X \multimap Y \in \mathcal{M} & \Rightarrow & X \in An_{\mathcal{G}}(Y \cup \mathbf{S}), \\
(2)\ X \leftarrow\!\!\ast\, Y \in \mathcal{M} & \Rightarrow & X \notin An_{\mathcal{G}}(Y \cup \mathbf{S}), \\
(3)\ X \leftrightarrow Y \in \mathcal{M} & \Rightarrow & \{X, Y\} \subset De_{\mathcal{G}}(\mathbf{L}), \\
(4)\ X - Y \in \mathcal{M} & \Rightarrow & \{X, Y\} \subset An_{\mathcal{G}}(\mathbf{S}).
\end{array}
$$

Conversely, a node subject to selection bias in $\mathcal{G}$ has no arrowhead in $\mathcal{M}$, and a node not subject to selection bias is not part of an undirected edge in $\mathcal{M}$. Finally, note that $X \in An_{\mathcal{G}}(\mathbf{Y} \cup \mathbf{S}) \iff X \in Ant(\mathbf{Y})$ for all $X \in \mathcal{M}$, $\mathbf{Y} \subseteq \mathcal{M}$.

The following definition is a special case of the definition in section 4.2.1 of (Richardson and Spirtes, 2002). Given an ancestral graph $\mathcal{M}$ over nodes $\mathbf{V} \cup \mathbf{L}$, the *marginal MAG* $\mathcal{M}'$ over nodes $\mathbf{V}$ has the following edges: $X, Y \in \mathbf{V}$ are adjacent in $\mathcal{M}'$ if there does not exist a set $\mathbf{Z} \subseteq \mathbf{V} \setminus \{X, Y\}$ that m-separates X, Y in $\mathcal{M}$, and in that case the edge in $\mathcal{M}'$ has an arrowhead at X if and only if $X \notin Ant_{\mathcal{M}}(Y)$, and has an arrowhead at Y if and only if $Y \notin Ant_{\mathcal{M}}(X)$.

1.2 Ancestral graph properties

We rely on the following connection between in/dependences and (non-)ancestorship in an ancestral graph.

Lemma 2. For disjoint (subsets of) nodes $X, Y, Z, \mathbf{Z}$ in an ancestral graph $\mathcal{M}$,

$$
\begin{array}{lll}
(1)\ X \not\perp\!\!\!\perp Y \,|\, \mathbf{Z} \cup [Z] & \Rightarrow & Z \notin Ant(\{X, Y\} \cup \mathbf{Z}). \\
(2)\ X \perp\!\!\!\perp Y \,|\, \mathbf{Z} \cup [Z] & \Rightarrow & Z \in Ant(\{X, Y\} \cup \mathbf{Z}). \\
(3)\ X \perp\!\!\!\perp Y \,|\, [\mathbf{Z} \cup Z] & \Rightarrow & Z \in Ant(\{X, Y\}),
\end{array}
$$

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

Proof. See e.g., Corollary to Lemma 14 in (Spirtes et al., 1999), and Lemma 2 in (Claassen and Heskes, 2011). □

We use the following result on anteriorship for nodes on unblocked paths:

Lemma 3.13 In an ancestral graph $\mathcal{M}$, if π is a path m-connecting X and Y given $\mathbf{Z}$, then every vertex on π is in $Ant(\{X, Y\} \cup \mathbf{Z})$.

Proof. See 3.13 in (Richardson and Spirtes, 2002). □

As a result, rule (3) in Lemma 2 not only applies to the nodes in the minimal separating set, but also to *all other nodes* on the paths in $\mathcal{M}$ between X and Y that become unblocked given only a subset $\mathbf{Z}' \subset \mathbf{Z}$.

Corollary 8. Let $\mathcal{M}$ be an ancestral graph, and suppose that $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$. If a path π in $\mathcal{M}$ between X and Y is unblocked given some subset $\mathbf{Z}' \subset \mathbf{Z}$, then all nodes on π are in $Ant(\{X, Y\})$.

Proof. Follows from Lemma 3.13, given that Lemma 2 rule (3) ensures that $\mathbf{Z}' \subset \mathbf{Z} \subseteq Ant(\{X, Y\})$. □

Similarly, we can freely add anterior nodes to any separating set without introducing a dependence:

Corollary 9. In an ancestral graph $\mathcal{M}$, if $X \perp\!\!\!\perp Y \,|\, \mathbf{Z}$, then $\forall \mathbf{W} \subseteq Ant(\{X, Y\} \cup \mathbf{Z})_{\setminus \{X,Y\}} : X \perp\!\!\!\perp Y \,|\, \mathbf{Z} \cup \mathbf{W}$.

Proof. Adding the nodes in $\mathbf{W}$ to the separating set one by one, then by rule (1) in Lemma 2, any node that creates a dependence cannot be anterior to any node in $\{X, Y\} \cup \mathbf{Z}$, contrary the assumed. So all added nodes leave the original independence intact, and therefore $X \perp\!\!\!\perp Y \,|\, \mathbf{Z} \cup \mathbf{W}$. □

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 $\mathbf{Z}$ iff:

1. $X \perp\!\!\!\perp Y \,|\, \mathbf{Z}$,

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

If $\mathbf{Z}$ D-separates X and Y, then (X,Y) is called a ***D*-sep link**, and a node $Z \in \mathbf{Z}$ is called 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 Identifying *D*-sep links

To prove Lemma 3 from the main article we first derive a connection between 'not separable by adjacent nodes' and non-anteriorship:

Lemma 10. In an ancestral graph $\mathcal{M}$, if $X \perp\!\!\!\perp Y \,|\, \mathbf{Z}$, but X is not independent of Y given any subset of $Adj(X)$ in $\mathcal{M}$, then $Y \notin Ant(X)$ and Y is not part of an undirected edge.

Proof. From $X \perp\!\!\!\perp Y \,|\, \mathbf{Z}$ there is no edge between X and Y in $\mathcal{M}$. Let $\mathbf{W} = Adj(X) \cap Ant(\{X,Y\})$ be the set of all nodes in $\mathcal{M}$ that are adjacent to X and have an anterior path to X and/or Y. According to the assumed then $X \not\perp\!\!\!\perp Y \,|\, \mathbf{W}$, and so there are one or more unblocked paths of the form $\langle X, U, \ldots, Y\rangle$ relative to $\mathbf{W}$ in $\mathcal{M}$ (as there is no direct edge). By Lemma 3.13 we know that implies $U \in Ant(\{X,Y\} \cup \mathbf{W})$. From $\mathbf{W} \subset Ant(\{X,Y\})$ and transitivity of 'anteriorship' then follows $U \in Ant(\{X,Y\}$, which combined with the fact that U is adjacent to X implies $U \in \mathbf{W}$.

But given that path $\pi = \langle X, U, \ldots, Y\rangle$ is unblocked relative to $\mathbf{W}$, node $U \in \mathbf{W}$ must be a collider along this path with arrowhead $X *\!\!\rightarrow U$ in $\mathcal{M}$. This means $U \notin Ant(X)$, which leaves $U \in Ant(Y)$. But then also $Y \notin Ant(X)$, otherwise (again by transitivity) U would still be anterior to X in $\mathcal{M}$. From the fact that U is collider along π we know that it is not part of an undirected edge, and so Y as descendant of U also cannot be part of an undirected edge in $\mathcal{M}$. □

This also applies directly to D-sep links.

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

(1) $X \notin Ant(Y \cup \mathbf{Z})$,
(2) $Y \notin Ant(X \cup \mathbf{Z})$,
(3) $\forall Z \in \mathbf{Z} : Z \in Ant(\{X,Y\})$,
(4) X and Y are not part of an undirected edge.

Proof. (1) from the definition of D-separated nodes and Lemma 10 follows that X is not anterior to Y; but X also cannot be anterior to any node in $\mathbf{Z}$, otherwise by (3) and transitivity/acyclicity it would either still be anterior to Y, or it would by anterior to itself which would imply a directed cycle (given that (4) implies there cannot be an undirected edge at X); therefore $X \notin Ant(Y \cup \mathbf{Z})$;

(2) idem for Y;

(3) Lemma 2 rule (3), given that $\mathbf{Z}$ is minimal.

(4) follows directly from Lemma 10. □

Note that it is possible that one or more nodes in $\mathbf{Z}$ (including D-sep nodes) are part of an undirected edge in $\mathcal{M}$.

Next we introduce:

Definition 3. For a set of nodes $\mathbf{X}$ in an ancestral graph $\mathcal{M}$, the set $AA(\mathbf{X})$ (*adjacent anteriors*) is defined as $AA(\mathbf{X}) = (Adj(\mathbf{X}) \cap Ant(\mathbf{X})) \setminus \mathbf{X}$.

In the context of D-sep links (X,Y) we usually refer to $AA(\{X,Y\})$ as the set of *adjacent ancestors*, as then $Ant(\{X,Y\}) = An(\{X,Y\})$, by Lemma 3-(4).

With this we can bring D-separation in standard form:

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 nodes in $\mathbf{Z}_{DS}$ (possibly empty) are D-sep nodes for (X,Y).

Proof. We use rules (1)-(3) 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 (3), all nodes in $\mathbf{Z}'$ are anterior to X and/or Y. By Corollary 9 we obtain $X \perp\!\!\!\perp Y \,|\, AA(\{X,Y\}) \cup \mathbf{Z}''$, where $\mathbf{Z}'' = \mathbf{Z}' \setminus AA(\{X,Y\})$ contains 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 then $X \perp\!\!\!\perp Y \,|\, AA(\{X,Y\}) \cup [\mathbf{Z}_{DS}]$. If (X,Y) is not a D-sep link, then $\mathbf{Z}_{DS} = \varnothing$. 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.

By construction the sets $\mathbf{Z}_{AA}$ and $\mathbf{Z}_{DS}$ are 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, in contradiction with Corollary 9. Therefore, at that point the D-separating set is minimal, i.e., $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}_{AA} \cup \mathbf{Z}_{DS}]$.

All nodes in $\mathbf{Z}_{DS}$ (if nonempty) satisfy the definition of D-sep node: by construction none of them are adjacent to X or Y, and if there were some subset $\mathbf{W} \subseteq Adj(\{X, Y\})$ that could make a node $Z \in \mathbf{Z}_{DS}$ redundant then, by Lemma 2.(2), that subset $\mathbf{W}$ must be a subset of $AA(\{X, Y\})$, and so by Corollary 9 the independence should also be found given $AA(\{X, Y\}) \cup \mathbf{Z}'''$ with $\mathbf{Z}''' \subseteq \mathbf{Z}_{DS} \setminus \{Z\}$: a contradiction. □

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.

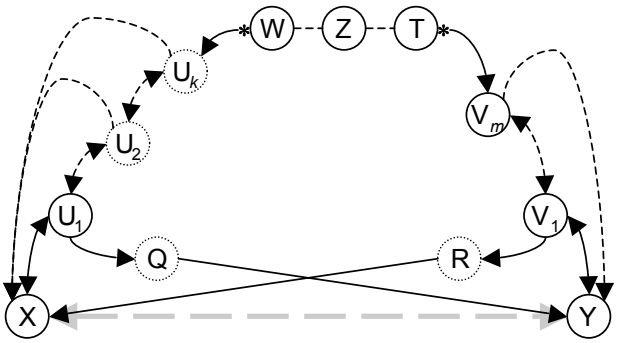

Figure 1: Path configuration for D-sep link $X - Y$.

In the proof of Lemma 4 we rely on the fact that for each D-sep link there is a path blocked by a D-sep node of the form depicted in Figure 1, which imposes six *identifiable* minimal dependence relations in (4)-(6), below:

Lemma 12. In a MAG $\mathcal{M}$, if nodes X and Y are D-separable, then there are nodes $\{U, V, W, T\} \in An(\{X, Y\})$ such that:

(1) $X \longleftrightarrow U$ and $V \longleftrightarrow Y$ in $\mathcal{M}$,

(2) $U \in An(Y)$ and $V \in An(X)$,

(3) $U \notin An(V)$ and $V \notin An(U)$,

(4) $W \notin Adj(X)$ and $\forall \mathbf{Z}_{XW}$ with $X \perp\!\!\!\perp W \,|\, [\mathbf{Z}_{XW}]$: $X \not\perp\!\!\!\perp W \,|\, \mathbf{Z}_{XW} \cup [U]$ and $X \not\perp\!\!\!\perp W \,|\, \mathbf{Z}_{XW} \cup [Y]$,

(5) $T \notin Adj(Y)$ and $\forall \mathbf{Z}_{YT}$ with $Y \perp\!\!\!\perp T \,|\, [\mathbf{Z}_{YT}]$: $Y \not\perp\!\!\!\perp T \,|\, \mathbf{Z}_{YT} \cup [V]$ and $Y \not\perp\!\!\!\perp T \,|\, \mathbf{Z}_{YT} \cup [X]$,

(6) $U \notin Adj(V)$ in $\mathcal{M}$, and $\forall \mathbf{Z}_{UV}, U \perp\!\!\!\perp V \,|\, [\mathbf{Z}_{UV}]$: $U \not\perp\!\!\!\perp V \,|\, \mathbf{Z}_{UV} \cup [X]$ and $U \not\perp\!\!\!\perp V \,|\, \mathbf{Z}_{UV} \cup [Y]$.

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. Let $Z \in \mathbf{Z}_{DS}$ and define $\mathbf{Z} := AA(\{X, Y\}) \cup \mathbf{Z}_{DS}$. Then, $X \perp\!\!\!\perp Y \,|\, \mathbf{Z}$ but $X \not\perp\!\!\!\perp Y \,|\, \mathbf{Z}_{\setminus Z}$, and so there must be a path π that is (only) blocked by noncollider Z (relative to the other $\mathbf{Z}_{\setminus Z}$).

We now show that we can take this path to be of the form $\pi = X \leftrightarrow U_1 (\leftrightarrow U_2 \leftrightarrow .. \leftrightarrow U_k) \leftarrow * W \cdots Z \cdots T * \rightarrow (V_m \leftrightarrow .. \leftrightarrow V_2 \leftrightarrow) V_1 \leftrightarrow Y$ in $\mathcal{M}$, where all nodes U_i are colliders along π and are adjacent to X, but only U_1 has a bidirected edge to X (similar for V_i at Y), and W is the first node along π starting from X that is not adjacent to X (possibly $W = Z$), and similarly for T. See also Figure 1.

Firstly, all paths between X and Y blocked by a node $Z \in \mathbf{Z}_{DS}$ must be *into* both X and Y: given that there are no undirected edges to X and/or Y in $\mathcal{M}$ (Lemma 3), then by Corollary 8 the first node U encountered along any such path must be in $An(\{X, Y\})$. But if this path starts with a tail from X then necessarily $X \longrightarrow U$, so that $U \in An(Y)$, which in turn implies $X \in An(Y)$, in contradiction with Lemma 3. Idem for Y. Therefore all paths blocked by node Z, including π, must have $X \leftarrow * \ldots * \rightarrow Y$.

Secondly, all paths between X and Y blocked by a node $Z \in \mathbf{Z}_{DS}$ must go via at least two other nodes $U_1 \in Adj(X)$ resp. $V_1 \in Adj(Y)$, as Z is presumed to be not adjacent to X and/or Y. As both these nodes satisfy the criteria for $AA(\{X, Y\})$ they are part of the conditioning set $\mathbf{Z}$, and so they must be colliders along π (otherwise Z was not needed to block it). The same holds for all subsequent nodes up to U_k and V_m along π that are adjacent to X and/or Y. Therefore the path π blocked by Z must have the general form $\pi = X \leftrightarrow U (\leftrightarrow \ldots \leftrightarrow U') \leftarrow * \ldots * \rightarrow (V' \leftrightarrow \ldots \leftrightarrow) V \leftrightarrow Y$.

Next, starting from X, at some point along π the first node W must be encountered that is *not* adjacent to X (possibly $W = Z$). Take U_1 as the first node encountered along π with a bidirected edge to X when starting from W in the direction of X. Then all other, up to U_k nodes between U_1 and W are colliders along π with a directed edge into X (again by Lemma 3 and Corollary 8). Similar for some nodes T and V_1 for Y. Therefore there exists a path blocked by Z of the form $\pi = X \leftrightarrow U_1 (\leftrightarrow U_2 \leftrightarrow \ldots \leftrightarrow U_k) \leftarrow * W \cdots Z \cdots T * \rightarrow (V_m \leftrightarrow \ldots \leftrightarrow V_2 \leftrightarrow) V_1 \leftrightarrow Y$ in $\mathcal{M}$, as indicated in Figure 1, with Z as noncollider along the path. Note that $W \in An(\{X, Y\})$ (and similarly for T): if $W \longrightarrow U_k$ on π this is immediate, and if $W \longleftrightarrow U_k$ on π, then W is not part of an undirected edge, which in combination with Corollary 8 leads to $W \in An(\{X, Y\})$.

With generic path π we can prove statements (1)-(6), equating U_1 with U and V_1 with V:

(1) By construction, we have $X \leftrightarrow U_1$ and $V_1 \leftrightarrow Y$ along π.

(2) By Corollary 8, all nodes along π are in $Ant(\{X, Y\})$. As both U_1 and V_1 are colliders along π this reduces to $\{U_1, V_1\} \subset An(\{X, Y\})$. The bidirected edge $X \leftrightarrow U_1$ implies $U_1 \notin An(X)$, and so: $U_1 \in An(Y)$; vice versa for $V_1 \in An(X)$.

(3) If $U_1 \in An(V_1)$, then $V_1 \in An(X)$ and transitivity would imply $U_1 \in An(X)$, contrary the bidirected edge $X \leftrightarrow U_1$, and so $U_1 \notin An(V_1)$; idem for V_1 and Y.

(4) For the in/dependence relations on π: given that X and W are not adjacent, they are separated by some minimal set $\mathbf{Z}_{XW}$ (not to be confused with $\mathbf{Z}_{AA}$ or $\mathbf{Z}_{DS}$). By construction, all $\{U_2, \ldots, U_k\}$ 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). As this holds for *any* (minimal) set $\mathbf{Z}_{XW}$ that can separate X and W, it means there are unblocked paths *into* U_1 from both X and W given $\mathbf{Z}_{XW}$, and so then conditioning on U_1 will 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 $X \not\perp\!\!\!\perp W \,|\, \mathbf{Z}_{XW} \cup [Y]$.

(5) Idem $Y \not\perp\!\!\!\perp T \,|\, \mathbf{Z}_{YT} \cup [V_1]$ and $Y \not\perp\!\!\!\perp T \,|\, \mathbf{Z}_{YT} \cup [X]$.

(6) Finally, U_1 and V_1 cannot be adjacent in $\mathcal{M}$: they cannot be connected by a bidirected edge, for that would make the path $\langle X, U_1, V_1, Y\rangle$ unblocked given $\mathbf{Z}$; by (3) they cannot be connected by an edge $U_1 \rightarrow V_1$ or $U_1 \leftarrow V_1$; and they cannot be connected by an undirected edge because they are both colliders along π. Therefore U_1 and V_1 are conditionally independent given some minimal set $\mathbf{Z}_{UV}$. For any such minimal separating set $\mathbf{Z}_{UV}$, no descendant of U_1 or V_1 (including X and Y) can be part of it, for that would imply either U_1 or V_1 was ancestor of the other. Including X or Y in the conditioning set would make them dependent given that both X and Y 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 $U_1 \not\perp\!\!\!\perp V_1 \,|\, \mathbf{Z}_{UV} \cup [Y]$. □

By Lemma 2, rule (1), each node in Lemma 12 that destroys one of the three independences cannot be anterior to any node in that independence, and so leads to identifiable invariant edge-marks (arrowheads).

To make this more precise we first introduce the following definitions:

Definition 4. A **minimal independence set** $\mathcal{I}(\mathcal{M})$ is a set of minimal independencies consistent with a MAG $\mathcal{M}$. It is called a **minimal independence model** if it contains at least one separating set for each pair of nonadjacent nodes in the MAG $\mathcal{M}$.

The skeleton $\mathcal{S}$ implied by a minimal independence set $\mathcal{I}(\mathcal{M})$ corresponds to the undirected graph with no edges between any $(X, Y) : X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}] \in \mathcal{I}(\mathcal{M})$. Note that a minimal independence *model* $\mathcal{I}(\mathcal{M})$ uniquely identifies the Markov equivalence class of $\mathcal{M}$.

Definition 5. Let $\mathcal{S}$ be the skeleton implied by a minimal independence set $\mathcal{I}(\mathcal{M})$. Then the **Augmented Skeleton** $\mathcal{S}^+$ is obtained by adding invariant arrowheads at all nodes W on edges to $\{X, Y\} \cup \mathbf{Z}$ in $\mathcal{S}$ that create a single node minimal dependence $X \not\perp\!\!\!\perp Y \,|\, \mathbf{Z} \cup [W]$, for all $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}] \in \mathcal{I}(\mathcal{M})$.

Augmentation boils down to repeated application of Lemma 2, rule (1).

From now on we assume that $\mathcal{I}(\mathcal{S})$ represents a minimal independence set as output by the PC algorithm with possible addition of one or more D-separating sets, consistent with a MAG $\mathcal{M}$. We also assume that we can query an independence oracle for the subsequent dependencies. This implies that the corresponding skeleton $\mathcal{S}$ matches the skeleton of $\mathcal{M}$, except that it may contain zero, one, or more additional (undirected) edges that all correspond to D-sep links in $\mathcal{M}$. For D-sep links in the corresponding augmented skeleton $\mathcal{S}^+$ this leads to the following pattern:

Lemma 4. Let $\mathcal{S}^+$ be the augmented skeleton obtained from a minimal independence set $\mathcal{I}(\mathcal{S})$ consistent with a MAG $\mathcal{M}$, such that the only additional edges in $\mathcal{S}^+$ that do not correspond with an edge in $\mathcal{M}$ are D-sep links. Let (X, Y) be an edge in $\mathcal{S}^+$ corresponding to a D-sep link in the MAG $\mathcal{M}$. If there are no (additional) edges in $\mathcal{S}^+$ between other D-separable pairs of nodes in $An(\{X, Y\})$, then $\mathcal{S}^+$ contains the following pattern:

(1) $U \leftrightarrow X \leftrightarrow Y \leftrightarrow V$ in $\mathcal{S}^+$,

(2) U and V not adjacent in $\mathcal{S}^+$,

(3) paths $V \cdots \rightarrow X$ and $U \cdots \rightarrow Y$ that do not contain arrowheads in the direction of V, resp. U.

Proof. Follows from Lemma 12.

(1) As X and U are adjacent in $\mathcal{M}$, they are also adjacent in $\mathcal{S}^+$. Similarly for Y and V. Nodes X and Y are also (still) presumed to be adjacent in $\mathcal{S}^+$. The assumption 'no edges in $\mathcal{S}^+$ between other D-sep links in $An(\{X, Y\})$' ensures that the three non-adjacencies (4)-(6) in Lemma 12 are present in $\mathcal{I}(\mathcal{M})$; the six subsequent dependences in Lemma 12 are found by the augmentation procedure, each time adding arrowheads

to the corresponding edge. Ultimately this means that $\mathcal{S}^+$ contains the invariant pattern: $U \leftrightarrow X \leftrightarrow Y \leftrightarrow V$.

(2) In particular (6) in Lemma 12 ensures that U and V are not adjacent in $\mathcal{M}$. The assumption 'no edges in $\mathcal{S}^+$ between other D-sep links in $An(\{X,Y\})$' ensures that (U,V) are not adjacent in $\mathcal{S}^+$ either.

(3) As $V \in An(X)$ there has to be a path from V in $\mathcal{S}^+$ that can be(come) oriented as a directed path into X. This means the augmentation procedure cannot add an invariant arrowhead in the opposite direction; idem for $U \in An(Y)$. □

The following result generalizes Lemma 5 in the original article as a minimal separating set $X \perp\!\!\!\perp Y \mid [\mathbf{Z}]$ automatically implies $\mathbf{Z} \subset Ant(\{X,Y\})_{\setminus\{X,Y\}}$.

Lemma 5. Let (X,Y) and (U,V) be two possibly overlapping but nonidentical pairs of D-separable nodes in a MAG $\mathcal{M}$. If $\{X,Y\} \subset An(\{U,V\})$, then $\{U,V\} \nsubseteq An(\{X,Y\})$.

Proof. Suppose $U \in An(X)$. If $U \neq X$ then by the given and acylicity $X \in An(V)$, which by transitivity implies $U \in An(V)$, contrary Lemma 3 rule (1). Idem for $U \in An(Y)$. So either $U \in \{X,Y\}$ or $U \notin An(\{X,Y\})$. Idem for V. But if both $U \in \{X,Y\}$ and $V \in \{X,Y\}$, with $U \neq V$ in a D-sep link, then the two D-separable pairs would be identical. Therefore at least one is not ancestor of $\{X,Y\}$, and so $\{U,V\} \nsubseteq An(\{X,Y\})$. □

So two D-separable node pairs cannot both be present in each others D-separating set. In fact, the ancestor relation induces a partial order over the D-sep links:

Lemma 13. Let $\Phi = \big\{\{U_1,V_1\},..,\{U_n,V_n\}\big\}$ be a set of distinct (but not necessarily disjoint) D-sep links in a MAG $\mathcal{M}$. Then the relation $\{U_i,V_i\} \preceq \{U_j,V_j\} \iff \{U_i,V_i\} \subseteq An(\{U_j,V_j\})$ defines a partial order over Φ.

Proof. For all $\{U_i,V_i\},\{U_j,V_j\},\{U_k,V_k\} \in \Phi$:

1. Reflexivity: $(\{U_i,V_i\} \preceq \{U_i,V_i\})$ is trivial.
2. Antisymmetry: (if $\{U_i,V_i\} \preceq \{U_j,V_j\}$ and $\{U_j,V_j\} \preceq \{U_i,V_i\}$ then $\{U_i,V_i\} = \{U_j,V_j\}$) follows from Lemma 5.
3. Transitivity: (if $\{U_i,V_i\} \preceq \{U_j,V_j\}$ and $\{U_j,V_j\} \preceq \{U_k,V_k\}$, then $\{U_i,V_i\} \preceq \{U_k,V_k\}$) follows from transitivity of the ancestor relationship of nodes in a MAG.

This implies the relation $\preceq$ satisfies the conditions of a partial order over the elements in Φ. □

As a result, in every non-empty (sub)set of D-separable node pairs there is at least one pair that does not have both nodes of any of the other pairs in its ancestors:

Lemma 14. If $\Phi = \big\{\{U_1,V_1\},..,\{U_n,V_n\}\big\}$ is a non-empty set of distinct (but not necessarily disjoint) D-sep links in a MAG $\mathcal{M}$, then there is a $\{U_i,V_i\} \in \Phi$ such that $\forall j \neq i : \{U_j,V_j\} \nsubseteq Ant(\{U_i,V_i\})$.

Proof. By Lemma 3 rule (4), D-sep nodes are not part of an undirected edge, so the statement reduces to $\forall j \neq i : \{U_j,V_j\} \nsubseteq An(\{U_i,V_i\})$. In terms of the partial order defined in Lemma 13 this is equivalent to stating that there exists a minimal element with respect to $\preceq$, i.e., an element $\{U_j,V_j\} \in \Phi$ such that there is no other element $\{U_i,V_i\} \in \Phi$ (with $i \neq j$) that precedes it, i.e., such that $\{U_i,V_i\} \preceq \{U_j,V_j\}$. As any finite partially ordered set has at least one minimal element, this proves the lemma. □

This means that if there are still one or more unidentified D-sep links in the augmented skeleton $\mathcal{S}^+$, then at least one of these has no unidentified D-sep links between any two of its ancestors, and so for that D-sep link the bidirected edge pattern of Lemma 4 is guaranteed to appear in $\mathcal{S}^+$. Therefore we can employ the following search strategy to check for D-sep links.

Lemma 15. In a MAG $\mathcal{M}$, all D-sep links can be found by repeatedly (and exclusively) checking an augmented skeleton $\mathcal{S}^+$ for edges that appear as the middle link of the bidirected triple from Lemma 4, while updating $\mathcal{S}^+$ for each D-sep link found.

Proof. Let $\mathcal{S}$ be the skeleton of $\mathcal{M}$, possibly with additional edges in $\mathcal{S}$ that all correspond to D-sep links in $\mathcal{M}$, (e.g. as obtained from the PC-search stage in the FCI algorithm). Let $\mathcal{S}^+$ be the augmented skeleton of $\mathcal{S}$ w.r.t. minimal (in)dependencies implied by $\mathcal{M}$. Then, as long as there are one or more edges in $\mathcal{S}^+$ that are not in $\mathcal{M}$, then by Lemma 14 at least one of these edges will have no unidentified D-sep links (edges in $\mathcal{S}$ that are not in $\mathcal{M}$) between its ancestors, and so by Lemma 4 this D-sep link will show up in $\mathcal{S}^+$ 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 $\mathcal{S}^+$ for each D-sep link identified (remove edge and compute 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 $\mathcal{S}^+$ matches that of the MAG $\mathcal{M}$, and all arrowheads in $\mathcal{S}^+$ 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 (after the augmented skeleton has been constructed), but in itself it is not sufficient to guarantee a reduction of 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

In proving some of the Lemmas below we often consider marginal MAGs, i.e. MAGs $\mathcal{M}'$ obtained by marginalizing out one or more nodes from a base MAG $\mathcal{M}$ in accordance with the rules in (Richardson and Spirtes, 2002) (see also section 1.1 for a definition).

First some properties of unblocked paths in an ancestral graph relative to the adjacent ancestors of D-sep link $\{X, Y\}$, used in the proof of Lemma 6.

All paths ultimately blocked by one or more of the D-sep nodes are unblocked relative to $AA(\{X, Y\})$.

Lemma 16. In a MAG $\mathcal{M}$, if $X \perp\!\!\!\perp Y \,|\, AA(\{X, Y\}) \cup \mathbf{Z}$ with $\mathbf{Z} \subset Ant(\{X, Y\})$, and π is a path between X and Y that is unblocked relative to $AA(\{X, Y\}) \cup \mathbf{Z}'$ for $\mathbf{Z}' \subset \mathbf{Z}$, then:

(1) all colliders on π are in $An(AA(\{X, Y\}))$;

(2) π is unblocked given $AA(\{X, Y\}) \cup \mathbf{Z}^*$, $\forall \mathbf{Z}^* \subseteq \mathbf{Z}'$.

Proof. (1) A path π is unblocked relative to a set $\mathbf{Z}$ iff every noncollider along π is not in $\mathbf{Z}$, and every collider on π is ancestor of some node in $\mathbf{Z}$. If every noncollider along π is not present in $AA(\{X, Y\}) \cup \mathbf{Z}'$ then they are also not present for a subset $\mathbf{Z}^* \subseteq \mathbf{Z}'$. Furthermore, every node $Z \in \mathbf{Z}'$ that is a descendant of some collider along π is in $An(\{X, Y\})$ (given $\mathbf{Z} \subset Ant(\{X, Y\})$ and the arrowhead at Z as descendant of the collider), and so has a directed path to $\{X, Y\}$. This directed path goes via penultimate node $U \in AA(\{X, Y\})$, and so it follows that Z, and so by transitivity the colliders along π as well, are ancestor of a node in $AA(\{X, Y\})$.

(2) Therefore π remains unblocked relative to $AA(\{X, Y\})$ in combination with any subset $\mathbf{Z}^* \subset \mathbf{Z}' \subset Ant(\{X, Y\})$, including $\mathbf{Z}^* = \varnothing$. □

Also, paths blocked by D-sep nodes correspond to sequences of bidirected edges in marginal MAGs.

Lemma 17. In a MAG $\mathcal{M}$ with D-separable X and Y, if $X \perp\!\!\!\perp Y \,|\, AA(\{X, Y\}) \cup \mathbf{Z} \cup [Z]$ with $\mathbf{Z} \subset Ant(\{X, Y\})$, then a path π between X and Y in $\mathcal{M}$ that is unblocked relative to $AA(\{X, Y\}) \cup \mathbf{Z}$ corresponds to a sequence of three or more bidirected edges connecting X and Y in all marginal MAGs $\mathcal{M}'$ over $\{X, Y\} \cup AA(\{X, Y\}) \cup \mathbf{Z}'$, with $\mathbf{Z}' \subseteq \mathbf{Z}$.

Proof. Below we first construct a sequence of unblocked treks in the MAG $\mathcal{M}$ between nodes in $\{X, Y\} \cup AA(\{X, Y\}) \cup \mathbf{Z}'$ that connects X and Y (*steps 1-3*). Then we map this sequence to the bidirected edge path in the marginal MAG $\mathcal{M}'$ (*steps 4-6*). Let π be a path in $\mathcal{M}$ between D-separable (X, Y) that is unblocked relative to $AA(\{X, Y\}) \cup \mathbf{Z}$.

Step 1: map π to sequence σ_U of unblocked treks in $\mathcal{M}$.
Let $U_1, \ldots, U_m$ be the colliders in $\mathcal{M}$ along the path π blocked by Z. By Lemma 16-(1), all colliders $U_i \in An(AA(\{X, Y\}))$. Using similar reasoning as in the beginning of the proof of Lemma 12, the path π blocked by Z (which is nonadjacent to X and Y) must be of the form $X \leftrightarrow U_1 \leftarrow\!* \ldots *\!\rightarrow U_m \leftrightarrow Y$. Each successive pair of colliders (U_i, U_{i+1}) along unblocked π must be connected by a trek (possibly a single edge $\leftrightarrow$) that does not contain any node in $AA(\{X, Y\}) \cup \mathbf{Z}$, and so $\sigma_U = [X, U_1, \ldots, U_m, Y]$ corresponds to a sequence of treks connecting X and Y in $\mathcal{M}$ that are unblocked relative to any subset of $AA(\{X, Y\}) \cup \mathbf{Z}$.

Step 2: map σ_U to sequence σ_V of unblocked treks between nodes in $\{X, Y\} \cup AA(\{X, Y\}) \cup \mathbf{Z}'$.
By Lemma 16 the path π in $\mathcal{M}$ is also unblocked given $AA(\{X, Y\}) \cup \mathbf{Z}'$, for any subset $\mathbf{Z}' \subseteq \mathbf{Z}$, and each collider U_i along π is in $An(AA(\{X, Y\}))$. For each U_i let V_i be the first descendant of U_i in $\mathcal{M}$ that is in $AA(\{X, Y\}) \cup \mathbf{Z}'$ (possibly $U_i = V_i$; in particular, $U_1 = V_1$ and $U_m = V_m$). If there are two or more such descendants (along different paths) then simply pick one of these at random. As a result, in $\mathcal{M}$ there are treks between V_i and V_{i+1}, and each such trek is again unblocked given any subset of $AA(\{X, Y\}) \cup \mathbf{Z}'$. Note that the concatenation of the three treks $V_i \longleftarrow \ldots \longleftarrow U_i$, $U_i \leftarrow\!* \ldots *\!\rightarrow U_{i+1}$, $U_{i+1} \longrightarrow \ldots \longrightarrow V_{i+1}$ is not necessarily a trek, as a node may occur more than once. This can be remedied by taking a "shortcut" via that node. Note that this node cannot become a collider, as at least one of the occurrences of that node must be on one of the directed paths $V_i \longleftarrow \ldots \longleftarrow U_i$ or $U_{i+1} \longrightarrow \ldots \longrightarrow V_{i+1}$ (because $U_i \leftarrow\!* \ldots *\!\rightarrow U_{i+1}$ is a trek), and that means that at least one of the edges at that node will have a tail. The result is a trek in $\mathcal{M}$ between V_i and V_{i+1} that is unblocked given any subset of $AA(\{X, Y\}) \cup \mathbf{Z}'$. Therefore $\sigma_V = [X, V_1, \ldots, V_m, Y]$ corresponds to a sequence of unblocked treks in $\mathcal{M}$ between nodes in $\{X, Y\} \cup AA(\{X, Y\}) \cup \mathbf{Z}'$.

Step 3: map σ_V to sequence σ_Z of unblocked treks between distinct nodes in $\{X, Y\} \cup AA(\{X, Y\}) \cup \mathbf{Z}'$ in $\mathcal{M}$.
It is possible that there are duplicates in σ_V, (i.e. $V_i = V_j$ with $i \neq j$), e.g. in case a descendant in $AA(\{X, Y\}) \cup \mathbf{Z}'$ is shared by multiple U_i. In

that case we can remove all nodes $[V_{i+1}, \ldots, V_j]$ from the sequence σ_V while still keeping a contiguous sequence of unblocked treks between X and Y. Assume we repeatedly merge such doublets (removing all intermediate nodes) until we are left with a sequence $\sigma_Z = [X, Z_1, \ldots, Z_k, Y]$ of distinct nodes $Z_i \in AA(\{X,Y\}) \cup \mathbf{Z}'$, with $Z_i \neq Z_j$ for $i \neq j$, and where each (Z_i, Z_{i+1}) is connected by a trek (with arrows into Z_i and Z_{i+1}) in $\mathcal{M}$ that is unblocked given any subset $\{X,Y\} \cup AA(\{X,Y\}) \cup \mathbf{Z}'$.
Note that the mapping $\sigma_V \rightarrow \sigma_Z$ is not necessarily unique, e.g. if $\sigma_V = [1,3,4,3,5,4,6,2,6,7]$, then either $\sigma_Z = [1,3,5,4,6,7]$ or $\sigma_Z = [1,3,4,6,7]$ will do.

Step 4: match sequence σ_Z to path π' in $\mathcal{M}'$.
As each pair (Z_i, Z_{i+1}) in the sequence σ_Z is connected by a trek in $\mathcal{M}$ that does not contain any noncolliders that are in $\{X,Y\} \cup AA(\{X,Y\}) \cup \mathbf{Z}'$, it follows that there is an unblocked path between each such pair given any subset of $\{X,Y\} \cup AA(\{X,Y\}) \cup \mathbf{Z}'$, and so each pair (Z_i, Z_{i+1}) must be adjacent in the marginal MAG $\mathcal{M}'$ over $\{X,Y\} \cup AA(\{X,Y\}) \cup \mathbf{Z}'$. That means the sequence of nodes σ_Z corresponds to a path $\pi' = \langle X, Z_1, \ldots, Z_k, Y \rangle$ in $\mathcal{M}'$. Note that this specific path π' is *not* necessarily the unblocked path in $\mathcal{M}'$ relative to $AA(\{X,Y\}) \cup \mathbf{Z}'$ we are looking for (we construct this next). For example, in Figure 2, the sequence $\sigma_Z = [X, Z_1, Z_i, Z_k, Z_j, Y]$ does indeed correspond to a path $X \leftrightarrow Z_1 \leftrightarrow Z_i \rightarrow Z_k \leftrightarrow Z_j \leftrightarrow Y$ in $\mathcal{M}'$, but this path is not unblocked relative to all other nodes in $\mathcal{M}'$ except $\{X,Y\}$, as Z_i is a noncollider along π'. However, this induces a bidirected edge $X \leftrightarrow Z_k$ in $\mathcal{M}'$, so that the path $\pi^* = X \leftrightarrow Z_k \leftrightarrow Z_j \leftrightarrow Y$ over a subset of nodes along π' does correspond to an unblocked path in $\mathcal{M}'$ relative to all other nodes.

Step 5: find bidirected edges that span nodes along π'.
Even though the trek in $\mathcal{M}$ between each (Z_i, Z_{i+1}) is *into* both nodes, they are not necessarily connected by a bidirected edge along π' in $\mathcal{M}'$, as one may be ancestor of the other in $\mathcal{M}$ (and so also in $\mathcal{M}'$). Here we show that this induces bidirected edges in the marginal MAG $\mathcal{M}'$ between other nodes along π', such that a bidirected edge path π^* over a subset of nodes along π' connecting X and Y in $\mathcal{M}'$ remains. (Note that nodes on π' are never connected by an undirected edge, as they all have arrowhead marks in $\mathcal{M}$, being (descendants of) colliders along the original π.)

We now identify *(maximal) inducing groups* of successive nodes $[Z_i, .., Z_j]$ along π' that are all ancestor (in $\mathcal{M}$) of the first or the last node ('sink') in the same group, and where two successive inducing groups along π' overlap on the last, resp. first sink node, so that $[X, .., Z_i], [Z_i, .., Z_j], .., [Z_k, .., Z_m], [Z_m, .., Y]$ along π. Each group $[Z_p, .., Z_r]$ is constructed as follows: starting from the last sink node of the previous group Z_p (or X for the first group), add successive nodes along π' to the group until the first node Z_q is encountered that is *not* in $An(Z_p)$. Then, starting from Y back along π', find the first node Z_r such that all nodes from Z_q up to Z_r along π' are in $An(Z_r)$ in $\mathcal{M}$ (possibly $Z_r = Z_q$). Then $[Z_p, .., Z_r]$ is the next maximal inducing group.

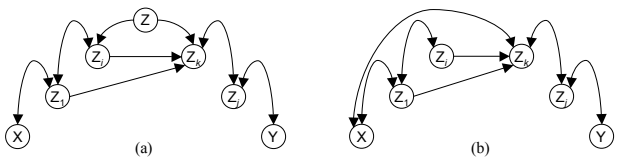

Figure 2: (a) Path in $\mathcal{M}$ blocked by D-sep node Z, (b) Idem in marginal MAG $\mathcal{M}'$ (without Z) with corresponding inducing groups $[X, Z_1, Z_i, Z_k], [Z_k, Z_j], [Z_j, Y]$ and induced edge $X \leftrightarrow Z_k$.

The two sink nodes of a maximal inducing group are connected by an *inducing path* in $\mathcal{M}$ (hence the name, see also §1) with respect to nodes not in $\{X,Y\} \cup AA(\{X,Y\}) \cup \mathbf{Z}'$:

1. by construction all nodes $\{Z_p, .., Z_r\}$ in the group are in $An(\{Z_p, Z_r\})$,
2. Z_p and Z_r are connected by a path in $\mathcal{M}$ on which all colliders are in $An(\{Z_p, Z_r\})$.

The inducing path between Z_p and Z_r can be constructed as follows. First, one concatenates all the treks $Z_p \leftarrow * \ldots * \rightarrow Z_{p+1}, Z_{p+1} \leftarrow * \ldots * \rightarrow Z_{p+2}, \ldots, Z_{r-1} \leftarrow * \ldots * \rightarrow Z_r$. In case a node Q occurs more than once, one takes the shortcut by deleting all intermediate nodes between the left-most and right-most occurrence of Q. Such Q can occur at most once on the original path π, and the other occurrence(s) must be on a directed path to one of the nodes $Z_p, \ldots, Z_r$. Therefore, if the remaining node Q becomes a collider when making the shortcut, then $Q \in An(\{Z_p, \ldots, Z_r\}) = An(\{Z_p, Z_r\})$.

As a result, by Theorem 4.2-(i) in Richardson and Spirtes (2002) the two sink nodes of a maximal inducing group are connected by an edge in $\mathcal{M}'$. Furthermore, this edge must be a bidirected edge, as sink nodes in a maximal inducing group cannot be ancestor of each other:

1. $Z_r \notin An(Z_p)$, otherwise by transitivity also $Z_q \in An(Z_p)$, contrary the given,
2. $Z_p \notin An(Z_r)$, because for $Z_p = X$ (first group) as part of D-sep link (X,Y), the given $AA(\{X,Y\}) \cup \mathbf{Z} \subset Ant(\{X,Y\})$ and acyclicity implies $X \notin An(Z_r) \subset An(\mathbf{Z})$; and if Z_p is the sink node

of the previous group $[Z_m, .., Z_p]$, then all nodes $[Z_p, .., Z_r]$ would also satisfy the conditions for inclusion in that group, and we would have obtained $[Z_m, .., Z_r]$, contrary the given.

Step 6: obtain bidirected edge path π^* in $\mathcal{M}'$.
Therefore in $\mathcal{M}'$ there is a path π^* of bidirected edges connecting X to Y via the sink nodes of neighbouring inducing groups along π'. The connection between the path π in $\mathcal{M}$ and the paths π' and π^* in $\mathcal{M}'$ is illustrated in Figure 3.

For every π' there are at least three distinct inducing groups: the first and last edge along π' in $\mathcal{M}'$ (corresponding to $X \leftrightarrow U_1$ resp. $U_m \leftrightarrow Y$ along π in $\mathcal{M}$, as both U_1 and U_m as subset of $AA(\{X,Y\})$ are also in $\mathcal{M}'$) are part of disjoint groups, as by Lemma 12-(3) $U_1 \notin An(U_m)$, and vice versa. That means there must be at least one other group to bridge the gap between U_1 and U_m from which follows that there are at least *three* such bidirected edges on the path π^* connecting X and Y in $\mathcal{M}'$. This proves the Lemma. □

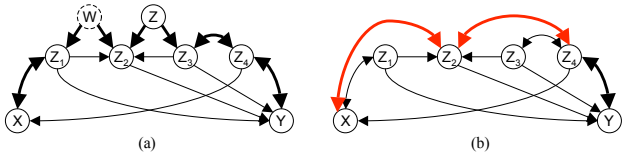

Figure 3: Induced bidirected edge path.
(a) Path $\pi = \langle X, Z_1, W, Z_2, Z, Z_3, Z_4, Y\rangle$ in $\mathcal{M}$ blocked by D-sep node Z, (b) Corresponding path $\pi' = \langle X, Z_1, Z_2, Z_3, Z_4, Y\rangle$ in marginal MAG $\mathcal{M}'$ over $\{X,Y\} \cup AA(\{X,Y\})(= \{Z_1, Z_2, Z_3, Z_4\})$, with maximal inducing groups $[X, Z_1, Z_2], [Z_2, Z_3, Z_4], [Z_4, Y]$, leading to induced edges $X \leftrightarrow Z_2$ and $Z_2 \leftrightarrow Z_4$ (red), to give the bidirected edge path $\pi^* = \langle X, Z_2, Z_4, Y\rangle$ in $\mathcal{M}'$ (bold).

With this we can prove Lemma 6 and Lemma 18 on being able to find all required D-sep nodes sequentially as part of a separating set between nodes already found.

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 another pair of nodes from $\{X,Y\} \cup \mathbf{Z}_{\setminus Z} \cup AA(\{X,Y\})$, neither of which are part of an undirected edge in $\mathcal{M}$.

Proof. By Lemma 17, in the marginal MAG $\mathcal{M}'$ over $\{X,Y\} \cup \mathbf{Z}_{\setminus Z} \cup AA(\{X,Y\})$ there is a sequence of at least 3 bidirected edges connecting X and Y. If this sequence of edges still exists in the MAG $\mathcal{M}^*$ over $\{X,Y\} \cup \mathbf{Z} \cup AA(\{X,Y\})$, then X and Y are not separated given $\mathbf{Z} \cup AA(\{X,Y\})$, contrary the given. Note that any adjacency in $\mathcal{M}^*$ that is also present in $\mathcal{M}'$ must have identical arrow/tail-marks in $\mathcal{M}^*$ and $\mathcal{M}'$, as both edges must express the same anterior relations in $\mathcal{M}$. Therefore at least one of these edges is eliminated in $\mathcal{M}^*$, which implies the existence of a set that can separate these two nodes from $\{X,Y\} \cup \mathbf{Z}_{\setminus Z} \cup AA(\{X,Y\})$. D-sep node Z is necessarily part of that set, otherwise the edge would already be eliminated in $\mathcal{M}'$. Given that both separated nodes have arrowhead marks in $\mathcal{M}'$ (as part of a bidirected edge path) it also follows that neither can be part of undirected edge as well. □

Lemma 18. In a MAG $\mathcal{M}$ with D-sep link (X,Y), all D-sep nodes $\mathbf{Z}_{DS}$ in $X \perp\!\!\!\perp Y \,|\, AA(\{X,Y\}) \cup [\mathbf{Z}_{DS}]$ can be found sequentially as part of a minimal separating set between a pair of nodes already found, starting from $\{X,Y\} \cup AA(\{X,Y\})$.

Proof. Starting from the marginal MAG $\mathcal{M}_0$ over $\{X,Y\} \cup AA(\{X,Y\})$ we infer from Lemma 17 that there is a sequence of bidirected edges in $\mathcal{M}_0$ connecting X and Y. Analogous to the rationale in Lemma 6: if this sequence of edges still exists in $\mathcal{M}^*$ over $\{X,Y\} \cup AA(\{X,Y\}) \cup \mathbf{Z}_{DS}$, then X and Y are not separated, contradicting the assumptions. Therefore at least one of these edges must be eliminated in $\mathcal{M}^*$, which implies the existence of a minimal separating set containing one or more of the nodes from $\mathbf{Z}_{DS}$. If only a subset $\mathbf{Z}_1 \subset \mathbf{Z}$ of nodes are needed in this separating set, then we can apply the same argument again to the marginal MAG $\mathcal{M}_1$ over $\{X,Y\} \cup AA(\{X,Y\}) \cup \mathbf{Z}_1$. Each time we find new nodes from $\mathbf{Z}$ that are part of a separating set between some pair of nodes we already found, until all have been added. □

2.3 Proofs - Building the hierarchy

Lemma 18 describes a procedure to find all required D-sep nodes for a given D-sep link (X,Y), starting from the adjacent ancestors of X and Y. However, there is a possible snag in the sense that the lemma ensures that each D-sep node can be found as part of *some* minimal separating set between a pair of nodes already found, but standard search strategies like PC only look for a *single* minimal sepset between each separable pair of nodes. As a result, if there are multiple possible minimal sepsets it is not a priori guaranteed that the D-sep nodes we are looking for are indeed contained in the sepset returned by the PC-search stage. However, it turns out that they can still be found as part of an 'ancestral superset' that is built recursively from results already obtained.

For that we introduce the following recursive definition for a set of separating nodes:

Definition 6. Let $\mathcal{I}$ be a minimal 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 $HIE(\mathbf{X}, \mathcal{I})$.

The recursion as formula: let $\mathbf{Q}_0 = \mathbf{X}$, and $\mathbf{Q}_{i+1} = \mathbf{Q}_i \cup \left(\bigcup_j \mathbf{W}_j : U_j, V_j \in \mathbf{Q}_i, U_j \perp\!\!\!\perp V_j \,|\, [\mathbf{W}_j] \in \mathcal{I}\right)$. Then, if n is the lowest index for which $\mathbf{Q}_{n+1} = \mathbf{Q}_n$, $HIE(\mathbf{X}, \mathcal{I}) = \mathbf{Q}_n$, with $n <$ nr. of variables in $\mathcal{I}$. Note that $\mathbf{X} \subseteq HIE(\mathbf{X}, \mathcal{I}) \subseteq Ant(\mathbf{X})$ for any $\mathbf{X}$ and any minimal independence set $\mathcal{I}(\mathcal{M})$.

The crucial result is now that for an arbitrary minimal independence set the hierarchy of a D-sep link and its adjacent ancestors is guaranteed to be a separating superset that contains all required D-sep nodes.

Lemma 7. In a MAG $\mathcal{M}$ with minimal independence model $\mathcal{I}(\mathcal{M})$, suppose that X and Y are non-adjacent. Let $\mathcal{I}(\mathcal{S}^+) \subset \mathcal{I}(\mathcal{M})$ be a subset that contains the same separating sets between ancestors of X and/or Y in $\mathcal{M}$, except that it does not contain a separating set for $\{X, Y\}$. If $\mathbf{Q} = HIE(\{X, Y\} \cup AA(\{X, Y\}), \mathcal{I}(\mathcal{S}^+))_{\setminus\{X,Y\}}$, then $\mathbf{Q}$ is a separating set between X and Y, i.e. $X \perp\!\!\!\perp Y \,|\, \mathbf{Q}$.

Proof. As X and Y are non-adjacent in $\mathcal{M}$, there is a minimal separating set $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$. By Lemma 11 and Corollary 9, $X \perp\!\!\!\perp Y \,|\, [\mathbf{Z}]$ implies that $X \perp\!\!\!\perp Y \,|\, \mathbf{Q} \cup [\mathbf{Z}^*]$, with $\mathbf{Q} \supset AA(\{X, Y\})$ as defined above and where all $\mathbf{Z}^* \subset \mathbf{Z}$ are D-sep nodes for (X, Y).

If (X, Y) is not a D-sep link then $\mathbf{Z}^* = \varnothing$, and the previous immediately reduces to $X \perp\!\!\!\perp Y \,|\, \mathbf{Q}$.

Suppose $\mathbf{Z}^* \neq \varnothing$, i.e. $\mathbf{Q}$ is not a D-separating set for (X, Y). Then we can write $X \perp\!\!\!\perp Y \,|\, \mathbf{Q} \cup \mathbf{Z}^*_{\setminus Z} \cup [Z]$, for some $Z \in \mathbf{Z}^*$, and so there is an unblocked path π in $\mathcal{M}$ relative to $\mathbf{Q} \cup \mathbf{Z}^*_{\setminus Z}$.

By Lemma 17 any path π unblocked without Z corresponds to a sequence of three or more bidirected edges in any marginal MAG over $\{X, Y\} \cup AA(\{X, Y\}) \cup \mathbf{Q}'$ with $\mathbf{Q}' \subseteq (\mathbf{Q} \cup \mathbf{Z}^*_{\setminus Z})$, including the marginal MAG $\mathcal{M}'$ over $\{X, Y\} \cup \mathbf{Q}$.

Therefore there is an unblocked path π' between X and Y relative to $\mathbf{Q}$ in this $\mathcal{M}'$ that consists of a sequence of three or more bidirected edges. If these edges are still present in $\mathcal{M}$, then the path is still unblocked relative to any set that includes $\mathbf{Q}$, counter to $X \perp\!\!\!\perp Y \,|\, \mathbf{Q} \cup [\mathbf{Z}^*]$. If one of the edges from π' is no longer present in $\mathcal{M}$, then that implies the existence of a minimal separating set in $\mathcal{I}(\mathcal{M})$ between two nodes from $\{X, Y\} \cup \mathbf{Q}$. Any such separating set would satisfy the conditions for inclusion in the hierarchy $\mathbf{Q}$ (definition 6), and so already be part of $\mathbf{Q}$. But then the corresponding edge should be absent from the marginal MAG $\mathcal{M}'$ over $\{X, Y\} \cup \mathbf{Q}$, contrary the assumption that it was part of the path π'.

Therefore the assumption that $\mathbf{Z}^* \neq \varnothing$ leads to a contradiction, and so $\mathbf{Q}$ must already contain all required nodes to form a separating set for (X, Y), i.e. $X \perp\!\!\!\perp Y \,|\, \mathbf{Q}$. □

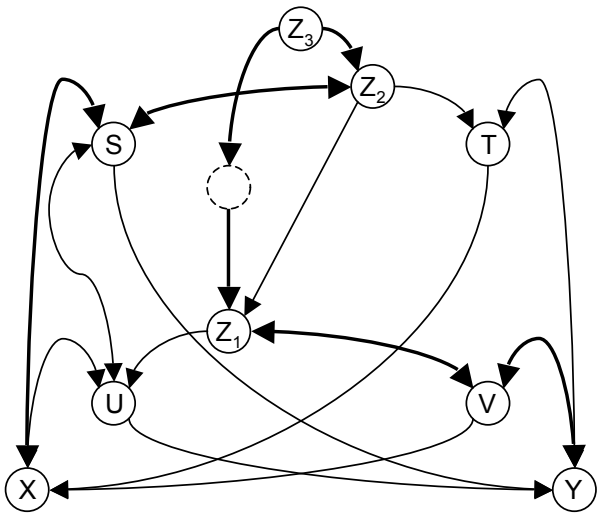

Figure 4: Illustration that it is not sufficient to look at separating sets between nodes adjacent to $\{X, Y\}$, but that it is necessary to include the full recursive hierarchy: nodes X and Y are D-separated by $\{S, T, U, V, Z_1, Z_2, Z_3\}$, but Z_3 (blocking the path in bold) is *only* present (and necessary) in $S \perp\!\!\!\perp Z_1 \,|\, [Z_2, Z_3]$, with $Z_1 \notin Adj(\{X, Y\})$.

The resulting D-sep set can be converted into a minimal D-separating set in at most N^2 additional independence tests, by removing redundant nodes one-by-one until no more can be found, see (Tian et al., 1998).

While searching for D-sep links in the augmented skeleton $\mathcal{S}^+$ we may not yet know the true adjacent ancestors of a D-sep candidate pair $\{X, Y\}$ in the underlying MAG $\mathcal{M}$. For a node with degree bounded by k, it has to be a combination of at most k from N nodes, which for a pair implies one from worst case order $N^k \times N^k = N^{2k}$ different sets. If there is a D-separating set for candidate $\{X, Y\}$, then it is guaranteed to appear in the hierarchy implied by one of these sets. Therefore, in a sparse graph, we need at most a polynomial number of tests (in N) to find a D-sep set.

We can now prove the main claim of the article:

Theorem 1. Let $\mathcal{M}$ be a MAG over N observed nodes corresponding to a distribution that is faithful to some underlying causal DAG $\mathcal{G}$, such that the node degree in $\mathcal{M}$ is bounded by some constant k, then the sound and complete equivalence class PAG $\mathcal{P}$ can be obtained

from worst case polynomial order $N^{2(k+2)}$ independence tests, even when latent variables and selection bias may be present.

Proof. Follows from a combination of (known) complexity results for the three main stages required to obtain the PAG $\mathcal{P}$:

1. find PC-skeleton graph $\mathcal{S}$ from adjacency search,
2. eliminate D-sep links to obtain the skeleton of $\mathcal{M}$,
3. orient invariant edge marks to obtain PAG $\mathcal{P}$.

The first stage is known to require worst case order N^{k+2} independence tests, as it searches for subsets $\leq k$ from $N-2$ variables in order to eliminate at most $N(N-1)$ edges (Spirtes et al., 2000).

This article has shown that the second stage can be completed in a number of independence tests that is also worst case polynomial order in N. First it suffices to find/update the augmented skeleton $\mathcal{S}^+$ (Definition 4) in order to obtain up to N^2 possible candidate D-sep links (Lemma 4). Augmenting the PC-graph may take order N^3 tests as it needs to check for at most $N-2$ nodes for $N(N-1)$ edges eliminated so far. Lemma 16 ensures that as long as not all D-sep links have been found then at least one of these has no unidentified D-sep links between its ancestors and so can be identified as a bidirected edge triple in $\mathcal{S}^+$ (Lemma 4). For each possible D-sep link $\{X, Y\}$ we need to find the set of possible adjacent ancestors $AA(\{X, Y\})$ (Definition 3) in $\mathcal{M}$ to compute the corresponding hierarchy (Lemma 7). For both candidates $\{X, Y\}$ this implies searching for at most k nodes from $N-2$ possible neighbours, which boils down to worst case $N^k \times N^k = N^{2k}$ independence tests. On finding a D-sep set we may need order N^2 tests to convert it into a minimal separating set (Tian et al., 1998), update the augmented skeleton (N^3), and possibly recheck up to N^2 previously tried-but-failed D-sep links. This leads to an overall complexity of the second stage of worst case order $N^2 \times N^{2k} \times N^2 = N^{2(k+2)}$, where augmenting the graph and conversion into minimal D-sep sets do not contribute to the leading order terms.

The third stage does not require additional independence tests at all. Complexity of the orientation rules lies mainly in checking for each edge mark for existence of certain paths in $\mathcal{P}$, which can be done in order N^2k by a generic 'reachability' algorithm, with $Nk \sim$ nr. of edges. As there are N^2 possible edge marks, this overall complexity is worst case order N^4k.

As the first and third stage do not contribute to the leading order terms from the second stage, it implies that the overall complexity of finding the sound and complete PAG takes at most order $N^{2(k+2)}$ independence tests. □

Note that in practice the typical performance is much better than this worst case result suggests: relatively few D-sep candidates with even fewer D-sep links. This limits the nr. of adjacent nodes per candidate pair in the second stage to $\sim k$, which reduces the most expensive term from N^{2k} down to constant order 2^{2k} (twice all subsets from k nodes), independent of N.

A description of the FCI+ algorithm that implements this result can be found in the main article.

Acknowledgements

This research was supported by the NWO (Netherlands Organization for Scientic Research), grant nr. 612.001.202. JM was supported by NWO VENI grant nr. 639.031.036.

References

T. Claassen and T. Heskes. A logical characterization of constraint-based causal discovery. In *Proc. of the 27th Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 135–144, 2011.

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.

J. Tian, A. Paz, and J. Pearl. Finding minimal d-separators. Technical Report R-254, UCLA Cognitive Systems Laboratory, 1998.