A ADDITIONAL PRELIMINARIES

Concatenation. We denote the concatenation of paths by $\oplus$, so that for a path $p = \langle V_1, V_2, \ldots, V_m \rangle$, $p = p(V_1, V_r) \oplus p(V_r, V_m)$, for $1 \leq r \leq m$.

Probabilistic implications of d-separation. Let $f$ be any observational density over $X$ consistent with an MPDAG $G = (V, E, U)$. Let $A, Y$ and $Z$ be pairwise disjoint node sets in $V$. If $A$ and $Y$ are d-separated given $Z$ in $G$, then $X_A$ and $X_Y$ are conditionally independent given $X_Z$ in the observational density $f$ (Lauritzen et al., 1990; Pearl, 2009). Hence, all DAGs that encode the same d-separation relationships also encode the same conditional independencies and are therefore Markov equivalent.

Buckets and bucket decomposition (Perković, 2020). A node set $A, A \subseteq V$ is an undirected connected set in $G = (V, E, U)$ if for every two distinct nodes $A_i, A_j \in A$, $A_i - \cdots - A_j$ is in $G$. If node set $B, B \subseteq D \subseteq V$, is a maximal undirected connected subset of $D$ in $G = (V, E, U)$, we call $B$ a bucket in $D$. Additionally, $D$ can be partitioned into $D = D_1 \cup \cdots \cup D_K$, where each $D_k, k \in \{1, \ldots, K\}$ is a bucket in $D$ and $D_i \cap D_j = \emptyset$ for $i \neq j$. We call the above partitioning of $D$ into buckets the bucket decomposition. Furthermore, $D_1, \ldots, D_K$ can be ordered in such a way that if $D_i \rightarrow D_2$ and $D_1 \in D_i$, $D_2 \in D_j$, then $i < j$; see PC algorithm of Perković (2020).

Lemma A.1 (Rules of the do-calculus, Pearl, 2009). Let $A, Y, Z$ and $W$ be pairwise disjoint (possibly empty) node sets in causal DAG $D = (V, E, \emptyset)$. Let $D^A$ denote the graph obtained by deleting all edges into $A$ from $D$. Similarly, let $D^A_{\overline{W}}$ denote the graph obtained by deleting all edges out of $A$ in $D$ and let $D^A_{\overline{ZW}}$ denote the graph obtained by deleting all edges into $A$ and all edges out of $Z$ in $D$.

Rule 1. If $Y \perp_{D^A} Z|A \cup W$, then $f(x_y|do(a_x), x_w) = f(x_y|do(a_x), x_z, x_w)$.

Rule 2. If $Y \perp_{D^A_{\overline{W}}} Z|A \cup W$, then $f(x_y|do(a_x), do(z), x_w) = f(x_y|do(a_x), x_z, x_w)$.

Rule 3. If $Y \perp_{D^A_{\overline{ZW}}} Z|A \cup W$, then $f(x_y|do(a_x), x_w) = f(x_y|do(a_x), x_z, x_w)$, where $Z(W) = Z \setminus \text{An}(W, D^A)$.

Lemma A.2 (Lemma 3.6 of Perković et al., 2017). Let $A$ and $Y$ be distinct nodes in a MPDAG $G$. If $p$ is a possibly causal path from $A$ to $Y$ in $G$, then a subsequence $p^*$ of $p$ forms a possibly causal unshielded path from $A$ to $Y$ in $G$.

Lemma A.3 (Wright’s rule, Wright, 1921). Let $X = AX + \varepsilon$, where $A \in \mathbb{R}^{p \times p}, X = X_Y, |V| = p,$ and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_p)^T$ is a vector of mutually independent errors with means zero and proper variance such that $\text{var}(\varepsilon_i) = 1$, for all $i \in \{1, \ldots, p\}$. Let $D = (V, E, \emptyset)$ be the corresponding DAG. For two distinct nodes $i, j \in V$, let $p_1, \ldots, p_k$ be all paths between $i$ and $j$ in $D$ that do not contain a collider. Then $\text{cov}(X_i, X_j) = \sum_{r=1}^k \pi_r$, where $\pi_r$ is the product of all edge coefficients along path $p_r$, $r \in \{1, \ldots, k\}$.

Lemma A.4. (See, e.g., Mardia et al., 1980, Theorem 3.2.4) Let $X = (X_1, X_2)$ be a $p$-dimensional multivariate Gaussian random vector with mean vector $\mu = (\mu_1, \mu_2)$ and covariance matrix $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$, so that $X_1$ is a $q$-dimensional multivariate Gaussian random vector with mean vector $\mu_1$ and covariance matrix $\Sigma_{11}$ and $X_2$ is a $(p - q)$-dimensional multivariate Gaussian random vector with mean vector $\mu_2$ and covariance matrix $\Sigma_{22}$. Then $E[X_2|X_1 = x_1] = \mu_2 + \Sigma_{21}\Sigma^{-1}_{11}(x_1 - \mu_1)$.

Algorithm 1: MPDAG (see also Meek, 1995 and Algorithm 1 of Perković et al., 2017)

input : MPDAG $G$, set of background knowledge edge orientations $R$.
output : MPDAG $G'$ or FAIL.

1 Let $G' = G$;
2 while $R \neq \emptyset$ do
3 Choose an edge $\{U \rightarrow V\}$ in $R$;
4 $R = R \setminus \{U \rightarrow V\}$;
5 if $\{U \rightarrow V\}$ or $\{U \rightarrow V\}$ is in $G'$ then
6 Orient $\{U \rightarrow V\}$ in $G'$;
7 Iterate the rules in Figure 3 of the main text until no more can be applied;
8 else
9 FAIL;
10 end
11 return $G'$;
12

B PROOFS OF MAIN RESULTS

Proof of Theorem 2. Let $p = (A_1, V_1, \ldots, V_k = Y_1), k \geq 1, A_1 \in A, Y_1 \in Y$. If $k = 1$, that is, $A_1 - Y_1$ is in $G$, the proposition clearly holds. Hence, we will assume $k > 1$. Suppose for a contradiction that there is an MPDAG $G^*$ represented by $G$ such that $A_1 - Y_1$ is in $G^*$ and that the total effect of $A$ on $Y$ is identified in $G^*$. Further, let $p^*$ be the path in $G^*$ that corresponds to path $p$ in $G$, so that $p$ and $p^*$ are both sequences of nodes $(A_1, V_1, \ldots, V_k = Y_1), k > 1$.

Since the total effect of $A$ on $Y$ is identified in $G^*$, and because $p^*$ is a proper path from $A$ to $Y$ that starts with an undirected edge in $G^*$, by Theorem 1, $p^*$ must
be a non-causal path from $A_1$ to $Y_1$ in $\mathcal{G}^*$. We show that this implies that $A_1 \rightarrow V_i \leftarrow V_2$ and $A_1 \rightarrow V_2$ are in $\mathcal{G}^*$, which contradicts that $\mathcal{G}^*$ is an MPDAG (because orientations in $\mathcal{G}^*$ are not complete with respect to R2 in Figure 3 of main text).

We first show that any existing edge between $A_1$ and $V_i$, $i \in \{2, \ldots, k\}$ in $G$ is of the form $A_1 \rightarrow V_i$. Suppose that there is an edge between $A_1$ and $V_i$ in $G$. This edge cannot be of the form $A_1 \leftarrow V_i$, since that would imply that $p$ is a non-causal path in $G$. This edge also cannot be of the form $A_1 \rightarrow V_i$, because otherwise we can concatenate $A_1 - V_i$ and $p(V_i, Y_1)$ to construct a proper possibly causal path from $A$ to $Y$ in $G$ that is shorter than $p$. Hence, any existing edge between $A_1$ and $V_i$ must be of the form $A_1 \rightarrow V_i$ in $G$ and $\mathcal{G}^*$.

Next, we show that $p^*(V_1, Y_1)$ starts with edge $V_1 \leftarrow V_2$ in $G$. Since $p$ is chosen as a shortest proper possibly causal path from $A$ to $Y$ that starts with an undirected edge in $G$, $p(V_1, Y_1)$ is a proper causal definite status path in $G$ (Lemma A.2). Then $p^+(V_2, Y_1)$ is also a path of definite status in $\mathcal{G}^*$. Additionally, since $p(V_1, Y_1)$ is a possibly causal definite status path in $G$, there cannot be any collider on $p^+(V_1, Y_1)$.

Furthermore, $p^*$ is a non-causal path, $A_1 \leftarrow V_i$ is in $\mathcal{G}^*$, and any edge between $A_1$ and $V_i$, $i \in \{2, \ldots, k\}$ is of the form $A_1 \rightarrow V_i$, so $p^+(V_1, Y_1)$ must be a non-causal path from $V_1$ to $Y$. Since $p^+(V_1, Y_1)$ is a non-causal definite status path without any colliders, it must start with an edge into $V_1$, that is $V_1 \leftarrow V_2$ is on $p^+(V_1, Y_1)$ in $\mathcal{G}^*$. Then $A_1 - V_1 \leftarrow V_2$ is in $\mathcal{G}^*$.

Now, $p^*(A_1, V_2)$ is of the form $A_1 - V_1 \leftarrow V_2$, so for $G^*$ to be an MPDAG, $\langle A_1, V_2 \rangle$ is in $\mathcal{G}^*$ (R1 in Figure 3). Then $A_1 \rightarrow V_2 \rightarrow V_1$ and $A_1 - V_1$ are in $\mathcal{G}^*$, which by R2 in Figure 3 contradicts that $\mathcal{G}^*$ is an MPDAG. □

Proof of Theorem 3. Statement (iii) directly follows from the construction of the algorithm. Statement (i) follows from the construction of the algorithm and Theorem 1.

Now we prove statement (ii). The proof follows a similar reasoning as the proof of Theorem 2 of Shpitser and Pearl (2006), proof of Theorem 57 of Perković et al. (2018) and proof of Proposition 3.2. of Perković (2020).

Suppose for a contradiction that $|L| \geq 2$ and let $\mathcal{G}_1$ and $\mathcal{G}_2$ be two different MPDAGs in $L$. Since $\mathcal{G}_1$ and $\mathcal{G}_2$ are both represented by $G$, any observational density $f$ consistent with $G$ is also consistent with $\mathcal{G}_1$ and $\mathcal{G}_2$ due to Markov equivalence.

Let $[\mathcal{G}]$ denote the set of DAGs represented by $G$. Let $f_1(x_Y|\text{do}(x_A))$ denote the density of $X_Y$ under the intervention $\text{do}(X_A = x_A)$ computed from $f(x)$ assuming that the causal DAG belongs to $[\mathcal{G}_1]$. Analogously, let $f_2(x_Y|\text{do}(x_A))$ denote the density of $X_Y$ under the intervention $\text{do}(X_A = x_A)$ computed from $f(x)$ assuming that the causal DAG belongs to $[\mathcal{G}_2]$. For the above interventional densities of $X_Y$ to differ, it suffices to show that $\mathbb{E}_1[x_Y|\text{do}(X_A = 1)] \neq \mathbb{E}_2[x_Y|\text{do}(X_A = 1)]$, where $\mathbb{E}_1(x_A = 1)$ indicates a do intervention that sets the value of every variable indexed by $A$ to 1, and $\mathbb{E}_1$ and $\mathbb{E}_2$ correspond to $f_1$ and $f_2$ respectively. Furthermore, it suffices to show that there is a node $Y_1 \in Y$ such that $\mathbb{E}_1[X_{Y_1}|\text{do}(X_A = 1)] \neq \mathbb{E}_2[X_{Y_1}|\text{do}(X_A = 1)]$, which gives us the desired contradiction.

First, we establish the pertinent graphical differences between $\mathcal{G}_1$ and $\mathcal{G}_2$ that stem from the application of Theorem 2 in the IDGraphs algorithm (Algorithm 1 in the main text). These graphical differences will be categorized as cases (i) and (ii) in Lemma B.1 below. Then, for each case, we will construct a linear causal model with Gaussian noise that imposes an observational density $f(x)$ consistent with $\mathcal{G}_1$ and $\mathcal{G}_2$ such that $\mathbb{E}_1[x_{Y_1}|\text{do}(X_A = 1)] \neq \mathbb{E}_2[x_{Y_1}|\text{do}(X_A = 1)]$, which gives us the desired contradiction.

By construction of $R_1$ and $R_2$, there is at least one edge whose orientation differs between $R_1$ and $R_2$. Without loss of generality, let $A_1 \rightarrow V_1$, $A_1 \in A$, $V_1 \in V \setminus A$ be the first edge in $R_1$ such that $A_1 \leftarrow V_1$ is in $R_2$. Also, let $R^*$ be the list of edge orientations that come before $A_1 \rightarrow V_1$ in $R_1$ and let $\mathcal{G}^* = \text{MPDAG}(G, R^*)$. Then by Theorem 2, the total effect of $A$ on $Y$ is not identified given $\mathcal{G}^*$.

Among all the shortest proper possibly causal paths from $A$ to $Y$ that start with an undirected edge in $\mathcal{G}^*$, choose $p^*$ as one that starts with $A_1 - V_i$, $p^* = \langle A_1 = V_0, V_1, \ldots, V_k = Y_1 \rangle$, $Y_1 \in Y$. Let $p_1$ be the path in a DAG $D_1$ in $[\mathcal{G}_1]$ that consists of the same sequence of nodes as $p^*$ in $\mathcal{G}^*$. Analogously, let $p_2$ be the path in a DAG $D_2$ in $[\mathcal{G}_2]$ that consists of the same sequence of nodes as $p^*$.

By Lemma B.1 we have the following cases:

(i) if $p^*$ is unshielded in $\mathcal{G}^*$, then $p_1$ is of the form $A_1 \rightarrow V_1 \rightarrow \cdots \rightarrow Y_1$, and $p_2$ starts with edge $A_1 \leftarrow V_i$.

(ii) if $p^*$ is a shielded path in $\mathcal{G}^*$, then $A_1 \rightarrow V_i$, $i \in$
Consider a multivariate Gaussian density over $X$ with mean zero, constructed using a linear causal model consistent with $D_1$ and thus, also $G_1$ (due to Markov equivalence). We define the linear causal model in such a way that all edge coefficients except for the ones on $p_1$ are 0, and all edge coefficients on $p_1$ are in $(0, 1)$ and small enough so that we can choose the error variances in such a way that $\text{var}(X_i) = 1$ for every $i \in V$.

The density $f(x)$ generated in this way is consistent with $D_1$ and thus also consistent with $G_1$ and $G_2$ (Lauritzen et al., 1990). Moreover, $f(x)$ is consistent with DAG $D_{11}$ that is obtained from $D_1$ by removing all edges except for the ones on $p_1$; see Figure B.1(a). Additionally, $D_{11}$ is Markov equivalent to DAG $D_{21}$, which is obtained from $D_2$ by removing all edges except for those on $p_2$; see Figure B.1(b). Hence, $f(x)$ is also consistent with $D_{21}$.

Let $f_1(x_Y|do(x_A))$ be an interventional density of $X_Y$ under the intervention $do(X_A = x_A)$ that is consistent with $D_{11}$ (and $D_1$). By Rules 3 and 2 of the do-calculus (Lemma A.1), we have

$$f_1(x_Y|do(x_A)) = f_1(x_Y|do(x_{A_1})) = f(x_Y|x_{A_1}).$$

So $E[11|x_Y|do(X_A = 1)] = \int f(x_Y|x_Y|X_{A_1} = 1) dx_{Y_1} = \text{cov}(X_{Y_1}, X_{A_1}) = a$ by Lemma A.4. Additionally, by Lemma A.3, $a$ is equal to the product of all edge coefficients along $p_1$ and so $a \in (0, 1)$.

Similarly, let $f_2(x_Y|do(x_A))$ be an interventional density of $X_Y$ consistent with $D_{21}$ (and $D_2$). Then $f_2(x_Y|do(x_A)) = f(x_Y)$ by Rule 3 of Lemma A.1. Hence, $E[11|x_Y|do(X_A = 1)] = E[X_Y] = 0$. Since $a \neq 0$, this completes the proof for case (i).

Consider a multivariate Gaussian density over $X$ with mean zero, constructed using a linear causal model with Gaussian noise consistent with $D_2$. We define the causal model in a way such that all edge coefficients except for the ones on $p_2$ and $A_1 \rightarrow V_i$, $i \in \{2, \ldots, r\}$ are 0, and all edge coefficients on $p_2$ and $A_1 \rightarrow V_i$ are in $(0, 1)$ and small enough so in such a way that $\text{var}(X_i) = 1$ for all $i \in V$.

The density $f(x)$ generated in this way is consistent with $D_2$ and $G_2$ (Lauritzen et al., 1990). Moreover, $f(x)$ is consistent with DAG $D_{22}$ that is obtained from $D_3$ by removing all edges except for the ones on $p_2$ and $A_1 \rightarrow V_i$, $i \in \{2, \ldots, r\}$, $2 \leq r \leq k$; see Fig. B.2(a). Let $f_2(x_Y|do(x_A))$ be an interventional density of $X_Y$ under the intervention $do(X_A = x_A)$ that is consistent with $D_{22}$ (and also $D_2$).

We now have

$$f_2(x_Y|do(x_A)) = f_2(x_Y|do(x_{A_1})) = \int f(x_{Y_1}|do(x_{A_1}), x_{V_1}) f(x_{V_1}|do(x_{A_1})) dx_{V_1} \quad (1)$$

The first line follows using Rule 3 of the do-calculus, and the third line follows from an application of Rule 2 and Rule 3; see Lemma A.1.

We now compute $E[11|x_Y|do(X_A = 1)]$. For simplicity, we will use shorthands $\text{cov}(X_{Y_1}, X_{A_1}) = a$, $\text{cov}(X_{Y_1}, X_{V_1}) = b$ and $\text{cov}(X_{A_1}, X_{V_1}) = c$. Now, us-
ing Lemma A.4 and Eq. (1), we have
\[
E_2[X_{Y_1}\mid \text{do}(X_A = 1)] = \int \mathbb{E}[X_{Y_1}\mid X_{A_i} = 1, X_{V_i} = x_{V_i}]f(x_{V_i})\,dx_{V_i}
\]
\[
= \int [a \ b] \begin{bmatrix} 1 & c \\ c & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ x_{V_1} \end{bmatrix} f(x_{V_i})\,dx_{V_i}
= \frac{1}{1 - c^2} \int [a \ b] \begin{bmatrix} 1 & -c \\ -c & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x_{V_1} \end{bmatrix} f(x_{V_i})\,dx_{V_i}
= \frac{a - bc}{1 - c^2} - \frac{ac + b}{1 - c^2}\mathbb{E}[X_{V_1}] = \frac{a - bc}{1 - c^2}.
\]

Now, consider the cases (ii)a and (ii)b. Note that \( f(x) \) is also consistent with \( D_1 \), \( G_1 \) and a DAG that is obtained from \( D_1 \) by removing all edges except for the ones on \( p_1 \) and \( A_1 \rightarrow V_1, i \in \{2, \ldots, r\} \) \citep{Lauritzen90}. Depending on case (ii)a or (ii)b, this will be either DAG \( D_{11} \) in Figure B.2(b) or DAG \( D_{12} \) in Figure B.2(c).

Let \( f_{11}(x_{Y_1}\mid \text{do}(x_A)) \) and \( f_{12}(x_{Y_1}\mid \text{do}(x_A)) \) be the interventional densities of \( X_{Y_1} \) that are consistent with \( D_{11} \) and \( D_{12} \), respectively. Note that \( f_{11}(x_{Y_1}\mid \text{do}(x_A)) = f_{12}(x_{Y_1}\mid \text{do}(x_A)) \) since
\[
\begin{align*}
f_{11}(x_{Y_1}\mid \text{do}(x_A)) &= f_{11}(x_{Y_1}\mid \text{do}(x_A)) \\
&= f(x_{Y_1}\mid x_{A_i}) \\
&= f(x_{Y_1}\mid x_{A_1}) = f_{12}(x_{Y_1}\mid \text{do}(x_A)).
\end{align*}
\]

The first two equalities above follow from Rule 3 and Rule 2 of the do-calculus, which is third and forth follow from Rule 1 and Rule 2; see again Lemma A.1. Hence, \( f_1(x_{Y_1}\mid \text{do}(x_A)) = f_{11}(x_{Y_1}\mid \text{do}(x_A)) = f_{12}(x_{Y_1}\mid \text{do}(x_A)). \)

Using Equation (2) and Lemma A.4, we have
\[
\mathbb{E}[X_{Y_1}\mid \text{do}(X_A = 1)] = \mathbb{E}[X_{Y_1}\mid X_{A_i} = 1] = \text{cov}(X_{Y_1}, X_{A_i}) = a. \]

To show that \( \mathbb{E}[X_{Y_1}\mid \text{do}(X_A = 1)] \neq \mathbb{E}[X_{Y_1}\mid \text{do}(X_A = 1)] \), we need only to show that \( a \neq (a - bc)/(1 - c^2) \).

We will show that \( b > ac \) and \( c > 0 \), which leads to \( a - bc < a - ac^2 \), that is \( (a - bc)/(1 - c^2) < a \). To show \( b > ac \) and \( c > 0 \), we need to discuss \( a, b \) and \( c \) in terms of the original linear causal model.

By Lemma A.3, we have that \( c = \text{cov}(X_{A_1}, X_{V_1}) \) is equal the edge coefficient assigned to \( A_1 \leftarrow V_1 \) in \( D_{21} \), and hence \( c \in (0,1) \). Let \( a_1 \) be the product of edge coefficients on \( p_2(V_1, Y_1) \) and let \( a_r \) be the product of edge coefficients along \( \langle A_1, V_i \rangle \oplus p_2(V_i, Y_1), i \in \{2, \ldots, r\} \). Then \( a_i \in (0,1) \) for all \( i \in \{1, \ldots, r\} \). By Lemma A.3, we now have
\[
a = \text{cov}(X_{Y_1}, X_{A_1}) = c \cdot a_1 + a_2 + \cdots + a_r,
b = \text{cov}(X_{Y_1}, X_{V_1}) = a_1 + c \cdot (a_2 + \cdots + a_r),
\]
which yields \( b - ac = a_1(1 - c^2) > 0 \), completing the proof. \( \square \)

**Lemma B.1.** Suppose that the total effect of \( A \) on \( Y \) is not identified given MPDAG \( G \). Let \( p = \langle A_1 = V_0, V_1, \ldots, V_k = Y_1 \rangle, k \geq 1 \), \( A_1 \in A, Y_1 \in Y \), be a shortest proper possibly causal path from \( A \) to \( Y \) in \( G \). Let \( G_1 = \text{MPDAG}(G, \{A_1 \rightarrow V_1\}) \) and \( G_2 = \text{MPDAG}(G, \{A_1 \leftarrow V_1\}) \). Let \( p_1 \) and \( p_2 \) be the paths in \( G_1 \) and \( G_2 \) respectively, that consist of the same sequence of nodes as \( p \) in \( G \).

(i) If \( p \) is an unshielded path in \( G \), then
\begin{itemize}
    \item \( p_1 \) is of the form \( A_1 \rightarrow V_1 \rightarrow \cdots \rightarrow Y_1 \), and
    \item \( p_2 \) is of the form \( A_1 \leftarrow V_1 \cdots Y_1 \).
\end{itemize}

(ii) If \( p \) is a shielded path in \( G \), then
\begin{itemize}
    \item \( A_1 \rightarrow V_i \) is in \( G \) for all \( i \in \{2, \ldots, r\}, r \leq k, k > 1 \),
    \item \( p_2 \) is of the form \( A_1 \leftarrow V_1 \cdots Y_1 \),
    \item \( \text{Let } D_1 \text{ be a DAG in } [G_1] \text{ and let } p_{11} \text{ be the path in } D_1 \text{ corresponding to } p_1 \text{ in } G_1 \text{ and to } p \text{ in } G \text{, then}
\end{itemize}
\begin{itemize}
    \item \( p_{11} \) is of the form \( A_1 \rightarrow V_1 \rightarrow \cdots \rightarrow Y_1 \) in \( D_{11} \), or
    \item \( p_{11} \) is of the form \( A_1 \leftarrow V_i \leftarrow \cdots \leftarrow V_r \rightarrow \cdots \rightarrow Y_1, 1 \leq i \leq r \) in \( D_{11} \).
\end{itemize}

**Proof of Lemma B.1.** Path \( p \) is chosen as a shortest proper possibly causal path from \( A \) to \( Y \) that starts with an undirected edge in \( G \). Hence, \( p(V_1, Y_1) \) must be an unshielded possibly causal path from \( V_1 \) to \( Y_1 \), otherwise we can choose a shorter path than \( p \) in \( G \). This implies that no node \( V_i, i \in \{2, \ldots, k - 1\} \) can be a collider on either \( p_1 \) or \( p_2 \).

(i) Suppose first that \( p \) itself is unshielded. That is, no edge \( (V_i, V_{i+2}), i \in \{0, k - 2\} \) is in \( G \). Of course, since \( G_2 \) contains edge \( A_1 \leftarrow V_1 \), \( p_2 \) is of the form \( A_1 \leftarrow V_1 \cdots Y_1 \). Hence, we only need to show \( p_1 \) is a causal path in \( G_2 \).

Since \( p \) is unshielded, \( p_1 \) is also an unshielded path. Since \( A_1 \rightarrow V_1 \) is in \( G_1 \), as a consequence of iterative application of rule R1 of \( \text{Meek (1995) (Fig. 3)} \), \( p_1 \) is a causal path in \( G_1 \).

(ii) Next, we suppose that \( p \) is shielded. We first show that \( A_1 \rightarrow V_i \), for all \( i \in \{2, \ldots, r\}, r \leq k \) is in \( G \).

As discussed at the beginning of this proof, \( p(V_1, Y_1) \) is unshielded. Therefore, since \( p \) is shielded and \( p(V_1, Y_1) \) is unshielded, edge \( \langle A_1, V_2 \rangle \) is in \( G \). Furthermore, since \( p \) is chosen as a shortest proper possibly causal path from \( A \) to \( Y \) that starts with an undirected edge in \( G \), \( \langle A_1, V_2 \rangle \) must be of the form \( A_1 \rightarrow V_2 \).
If path $\langle A_1, V_2 \rangle \oplus p(V_2, Y_1)$ is shielded, then by the same reasoning as above, $A_1 \rightarrow V_3$ is in $G$. We can continue with the same reasoning, until we reach $V_r$, $r \in \{2, \ldots, k\}$, so that $A_1 \rightarrow V_r$ is in $G$ for $i \in \{2, \ldots, r\}$ and $\langle A_1, V_r \rangle \oplus p(V_r, Y_1)$ is an unshielded possibly causal path.

We note that if $r < k$, $p(V_r, Y_1)$ is of the form $V_r \rightarrow \cdots \rightarrow Y_1$. This is due to the fact that $A_1 \rightarrow V_r$ is in $G$ and that $\langle A_1, V_r \rangle \oplus p(V_r, Y_1)$ is an unshielded possibly causal path in $G$.

Next, we show that $p_2$ is of the form $A_1 \leftarrow V_1 \rightarrow \cdots \rightarrow Y_1$. Since $A_1 \leftarrow V_1$ and $A_1 \rightarrow V_2$ are in $G_2$ and since $G_2$ is acyclic, by rule R2 of Meek (1995), the edge $(V_1, V_2)$ is of the form $V_1 \rightarrow V_2$ in $G_2$. Then since $p_2(V_1, Y_1)$ is an unshielded possibly causal path that starts with $V_1 \rightarrow V_2$, by iterative applications of rule R1 of Meek (1995), $p_2(V_1, Y_1)$ must be a causal path in $G_2$.

Suppose $D_1$ is a DAG in $[G_1]$. From the above, we know that $A_1 \rightarrow V_1$ and if $r < k$, $V_r \rightarrow \cdots \rightarrow Y_1$ are in $G_1$ and therefore in $D_1$ as well. The subpath $p(V_1, V_r)$ is a possibly causal unshielded path in $G$ and hence, no node among $V_2, \ldots, V_{r-1}$ is a collider on $p$, $p_1$, or $p_{11}$. It then follows that either $p_{11}(V_1, V_r)$ is a causal path in $D_{11}$, in which case $p_{11}$ is of the form $A_1 \rightarrow V_1 \rightarrow \cdots \rightarrow Y_1$, or there is a node $V_l$, $1 < l \leq r$ on $p_{11}(V_1, V_r)$, such that $p_{11}(V_1, V_r)$ is of the form $V_1 \leftarrow \cdots \leftarrow V_l \rightarrow \cdots \rightarrow V_r$. □

References


