A ADDITIONAL PRELIMINARIES

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

Probabilistic implications of d-separation. Let f be any observational density over X consistent with an MPDAG $\mathcal{G} = (V, E, U)$. Let A, Y and Z be pairwise disjoint node sets in V. If A and Y are d-separated given Z in $\mathcal{G}$, 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 independences and are therefore *Markov equivalent*.

Buckets and bucket decomposition (Perković, 2020). A node set A, $A \subseteq V$ is an *undirected connected set* in $\mathcal{G} = (V, E, U)$ if for every two distinct nodes $A_i, A_j \in A$, $A_i - \dots - A_j$ is in $\mathcal{G}$. If node set B, $B \subseteq D \subseteq V$, is a maximal undirected connected subset of D in $\mathcal{G} = (V, E, U)$, we call B a *bucket* in D. Additionally, D can be partitioned into $D = D_1 \cup \dots \cup D_K$, where each D_k, $k \in \{1, \dots, K\}$ 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, $D_1, \dots, D_K$ can be ordered in such a way that if $D_1 \to D_2$ and $D_1 \in D_i$, $D_2 \in D_j$, then $i < j$; see PCO 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 $\mathcal{D} = (V, E, \emptyset)$. Let $\mathcal{D}_{\overline{A}}$ denote the graph obtained by deleting all edges into A from $\mathcal{D}$. Similarly, let $\mathcal{D}_{\underline{A}}$ denote the graph obtained by deleting all edges out of A in $\mathcal{D}$ and let $\mathcal{D}_{\overline{A}\underline{Z}}$ denote the graph obtained by deleting all edges into A and all edges out of Z in $\mathcal{D}$.

Rule 1. If $Y \perp_{\mathcal{D}_{\overline{A}}} Z | A \cup W$, then $f(x_y | \operatorname{do}(x_a), x_w) = f(x_y | \operatorname{do}(x_a), x_z, x_w)$.

Rule 2. If $Y \perp_{\mathcal{D}_{\overline{A}\underline{Z}}} Z | A \cup W$, then $f(x_y | \operatorname{do}(x_a), \operatorname{do}(x_z), x_w) = f(x_y | \operatorname{do}(x_a), x_z, x_w)$.

Rule 3. If $Y \perp_{\mathcal{D}_{\overline{AZ(W)}}} Z | A \cup W$, then $f(x_y | \operatorname{do}(x_a), x_w) = f(x_y | \operatorname{do}(x_a), x_z, x_w)$, where $Z(W) = Z \setminus \operatorname{An}(W, \mathcal{D}_{\overline{A}})$.

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

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

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 μ_1 and covariance matrix Σ_{11} and X_2 is a $(p-q)$-dimensional multivariate Gaussian random vector with mean vector μ_2 and covariance matrix Σ_{22}. Then $\mathbb{E}[X_2 | X_1 = x_1] = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1)$.

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

---

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

Let $\mathcal{G}' = \mathcal{G}$;
while $R \neq \emptyset$ **do**
  Choose an edge $\{U \to V\}$ in R;
  $R = R \setminus \{U \to V\}$;
  if $\{U - V\}$ *or* $\{U \to V\}$ *is in* $\mathcal{G}'$ **then**
    Orient $\{U \to V\}$ in $\mathcal{G}'$;
    Iterate the rules in Figure 3 of the main text until no more can be applied;
  else
    FAIL;
  end
end
return $\mathcal{G}'$;

B PROOFS OF MAIN RESULTS

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

Since the total effect of A on Y is identified in $\mathcal{G}^*$, and because p^* is a proper path from A to Y that starts with an undirected edge in $\mathcal{G}^*$, 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 $A_1 - V_1 \leftarrow V_2$ 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, $i \in \{2, \dots, k\}$ in $\mathcal{G}$ is of the form $A_1 \rightarrow V_i$. Suppose that there is an edge between A_1 and V_i, in $\mathcal{G}$. This edge is cannot be of the form $A_1 \leftarrow V_i$, since that would imply that p is a non-causal path in $\mathcal{G}$. This edge also cannot be of the form $A_1 - V_i$, 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 $\mathcal{G}$ 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 $\mathcal{G}$ and $\mathcal{G}^*$.

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

Furthermore, p^* is a non-causal path, $A_1 - V_1$ is in $\mathcal{G}^*$, and any edge between A_1 and V_i, $i \in \{2, \dots, k\}$ 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 $\mathcal{G}^*$ 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. $\square$

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 $\mathcal{G}$, any observational density f consistent with $\mathcal{G}$ 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 $\mathcal{G}$. Let $f_1(x_Y | \mathrm{do}(x_A))$ denote the density of X_Y under the intervention $\mathrm{do}(X_A = x_A)$ computed from $f(x)$ assuming that the causal DAG belongs to $[\mathcal{G}_1]$. Analogously, let $f_2(x_Y | \mathrm{do}(x_A))$ denote the density of X_Y under the intervention $\mathrm{do}(X_A = x_A)$ 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 $\mathbb{E}_1[X_Y | \mathrm{do}(X_A = \mathbf{1})] \neq \mathbb{E}_2[X_Y | \mathrm{do}(X_A = \mathbf{1})]$, where $\mathrm{do}(X_A = \mathbf{1})$ 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 $\mathbb{E}_1[X_{Y_1} | \mathrm{do}(X_A = \mathbf{1})] \neq \mathbb{E}_2[X_{Y_1} | \mathrm{do}(X_A = \mathbf{1})]$.

The stages of this proof are as follows. First, we will establish some 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 $\mathbb{E}_1[X_{Y_1} | \mathrm{do}(X_A = \mathbf{1})] \neq \mathbb{E}_2[X_{Y_1} | \mathrm{do}(X_A = \mathbf{1})]$, which gives us the desired contradiction.

First, we establish the pertinent graphical differences between $\mathcal{G}_1$ and $\mathcal{G}_2$. For this purpose, let R_1 and R_2 be the list of edge orientations that were added to $\mathcal{G}$ to construct $\mathcal{G}_1$ and $\mathcal{G}_2$ by the `IDGraphs` algorithm. That is $\mathcal{G}_1 = \mathrm{MPDAG}(\mathcal{G}, R_1)$ and $\mathcal{G}_2 = \mathrm{MPDAG}(\mathcal{G}, R_2)$. Without loss of generality, suppose that the edge orientations in R_1 and R_2 are listed in the order that they were added by the `IDGraphs` algorithm.

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 $\mathcal{G}^* = \mathrm{MPDAG}(\mathcal{G}, R^*)$. 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 $A_1 - V_1$, $p^* = \langle A_1 = V_0, V_1, \dots, V_k = Y_1 \rangle$, $Y_1 \in Y$. Let p_1 be the path in a DAG $\mathcal{D}_1$ 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 $\mathcal{D}_2$ 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 $A_1 \rightarrow V_1 \rightarrow \dots \rightarrow Y_1$, and p_2 starts with edge $A_1 \leftarrow V_1$.

(ii) if p^* is a shielded path in $\mathcal{G}^*$, then $A_1 \rightarrow V_i$, $i \in$

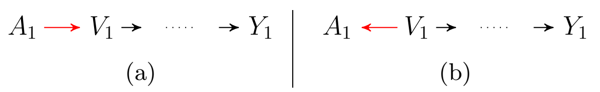

Figure B.1: DAGs (a) $\mathcal{D}_{11}$ and (b) $\mathcal{D}_{21}$ corresponding to (i) in the Proof of Theorem 3.

$\{1, \ldots, r\}$, $2 \leq r \leq k$, is in $\mathcal{G}^*$, p_2 is of the form $A_1 \leftarrow V_1 \rightarrow \cdots \rightarrow V_r \rightarrow \cdots \rightarrow Y_1$, and

(a) p_1 is of the form $A_1 \rightarrow V_1 \rightarrow \cdots \rightarrow V_r \rightarrow \cdots \rightarrow Y_1$, or

(b) p_1 is of the form $A_1 \rightarrow V_1 \leftarrow \cdots \leftarrow V_l \rightarrow \cdots \rightarrow Y_1$, $2 \leq l \leq r$.

We will now show how to choose a linear causal model consistent with $\mathcal{G}_1$ and $\mathcal{G}_2$ in each of the above cases that results in $\mathbb{E}_1[X_{Y_1} | \operatorname{do}(X_A = x_A)] \neq \mathbb{E}_2[X_{Y_1} | \operatorname{do}(X_A = x_A)]$.

(i) Consider a multivariate Gaussian density over X with mean zero, constructed using a linear causal model with Gaussian noise consistent with $\mathcal{D}_1$ and thus, also $\mathcal{G}_1$ (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 $\operatorname{var}(X_i) = 1$ for every $i \in V$.

The density $f(x)$ generated in this way is consistent with $\mathcal{D}_1$ and thus also consistent with $\mathcal{G}_1$ and $\mathcal{G}_2$ (Lauritzen et al., 1990). Moreover, $f(x)$ is consistent with DAG $\mathcal{D}_{11}$ that is obtained from $\mathcal{D}_1$ by removing all edges except for the ones on p_1; see Figure B.1(a). Additionally, $\mathcal{D}_{11}$ is Markov equivalent to DAG $\mathcal{D}_{21}$, which is obtained from $\mathcal{D}_2$ by removing all edges except for those on p_2; see Figure B.1(b). Hence, $f(x)$ is also consistent with $\mathcal{D}_{21}$.

Let $f_1(x_{Y_1} | \operatorname{do}(x_A))$ be an interventional density of X_{Y_1} under the intervention $\operatorname{do}(X_A = x_A)$ that is consistent with $\mathcal{D}_{11}$ (and $\mathcal{D}_1$). By Rules 3 and 2 of the do-calculus (Lemma A.1), we have

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

So $\mathbb{E}_1[X_{Y_1} | \operatorname{do}(X_A = \mathbf{1})] = \int x_{Y_1} f(x_Y | X_{A_1} = \mathbf{1})\, \mathrm{d}x_{Y_1} = \operatorname{cov}(X_{Y_1}, X_{A_1}) = a$ 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 $f_2(x_{Y_1} | \operatorname{do}(x_A))$ be an interventional density of X_{Y_1} consistent with $\mathcal{D}_{21}$ (and $\mathcal{D}_2$). Then $f_2(x_{Y_1} | \operatorname{do}(x_A)) = f(x_{Y_1})$ by Rule 3 of Lemma A.1. Hence, $\mathbb{E}_2[X_{Y_1} | \operatorname{do}(X_A = \mathbf{1})] = \mathbb{E}[X_{Y_1}] = 0$. Since $a \neq 0$, this completes the proof for case (i).

(ii) Consider a multivariate Gaussian density over X with mean zero, constructed using a linear causal model with Gaussian noise consistent with $\mathcal{D}_2$. 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 $\operatorname{var}(X_i) = 1$ for all $i \in V$.

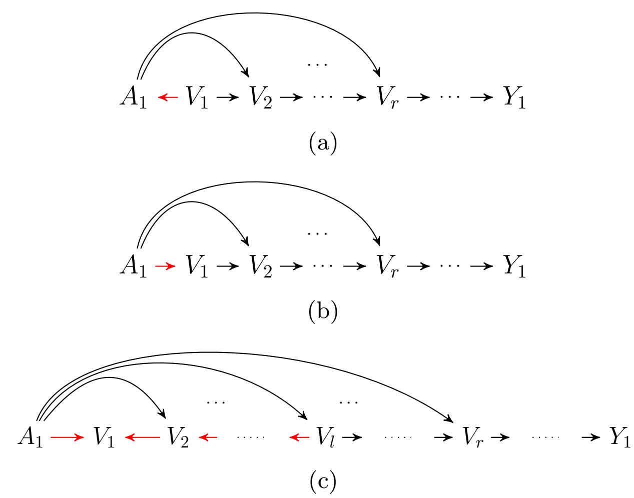

Figure B.2: DAGs (a) $\mathcal{D}_{22}$, (b) $\mathcal{D}_{11}$, and (c) $\mathcal{D}_{12}$ corresponding to (ii)a and (ii)b in the proof of Theorem 3.

The density $f(x)$ generated in this way is consistent with $\mathcal{D}_2$ and $\mathcal{G}_2$ (Lauritzen et al., 1990). Moreover, $f(x)$ is consistent with DAG $\mathcal{D}_{22}$ that is obtained from $\mathcal{D}_2$ 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 $f_2(x_{Y_1} | \operatorname{do}(x_A))$ be an interventional density of X_{Y_1} under the intervention $\operatorname{do}(X_A = x_A)$ that is consistent with $\mathcal{D}_{22}$ (and also $\mathcal{D}_2$).

We now have

$$\begin{aligned}
& f_2(x_{Y_1} | \operatorname{do}(x_A)) \\
&= f_2(x_{Y_1} | \operatorname{do}(x_{A_1})) \\
&= \int f(x_{Y_1} | \operatorname{do}(x_{A_1}), x_{V_1}) f(x_{V_1} | \operatorname{do}(x_{A_1}))\, \mathrm{d}x_{V_1} \\
&= \int f(x_{Y_1} | x_{A_1}, x_{V_1}) f(x_{V_1})\, \mathrm{d}x_{V_1}.
\end{aligned} \tag{1}$$

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 $\mathbb{E}_2[X_{Y_1} | \operatorname{do}(X_A = \mathbf{1})]$. For simplicity, we will use shorthands $\operatorname{cov}(X_{Y_1}, X_{A_1}) = a$, $\operatorname{cov}(X_{Y_1}, X_{V_1}) = b$ and $\operatorname{cov}(X_{A_1}, X_{V_1}) = c$. Now, using Lemma A.4 and Eq. (1), we have

$$
\begin{aligned}
&\mathbb{E}_2[X_{Y_1}\,|\,\mathrm{do}(X_A = 1)] \\
&= \int \mathbb{E}[X_{Y_1}|X_{A_1} = 1, X_{V_1} = x_{V_1}] f(x_{V_1})\,\mathrm{d}x_{V_1} \\
&= \int \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} 1 & c \\ c & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ x_{V_1} \end{bmatrix} f(x_{V_1})\,\mathrm{d}x_{V1} \\
&= \int \frac{1}{1-c^2} \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} 1 & -c \\ -c & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x_{V_1} \end{bmatrix} f(x_{V_1})\,\mathrm{d}x_{V1} \\
&= \frac{a-bc}{1-c^2} + \frac{-ac+b}{1-c^2}\mathbb{E}[X_{V_1}] = \frac{a-bc}{1-c^2}.
\end{aligned}
$$

Now, consider the cases (ii)a and (ii)b. Note that $f(x)$ is also consistent with $\mathcal{D}_1$, $\mathcal{G}_1$ and a DAG that is obtained from $\mathcal{D}_1$ by removing all edges except for the ones on p_1 and $A_1 \to V_i$, $i \in \{2, \dots, r\}$ (Lauritzen et al., 1990). Depending on case (ii)a or (ii)b, this will be either DAG $\mathcal{D}_{11}$ in Figure B.2(b) or DAG $\mathcal{D}_{12}$ in Figure B.2(c).

Let $f_{11}(x_{Y_1}\,|\,\mathrm{do}(x_A))$ and $f_{12}(x_{Y_1}\,|\,\mathrm{do}(x_A))$ be the interventional densities of X_{Y_1} that are consistent with $\mathcal{D}_{11}$ and $\mathcal{D}_{12}$, respectively. Note that $f_{11}(x_{Y_1}\,|\,\mathrm{do}(x_A)) = f_{12}(x_{Y_1}\,|\,\mathrm{do}(x_A))$ since

$$
\begin{aligned}
&f_{11}(x_{Y_1}\,|\,\mathrm{do}(x_A)) \\
&= f_{11}(x_{Y_1}\,|\,\mathrm{do}(x_{A_1})) \\
&= f(x_{Y_1}|x_{A_1}) \\
&= f(x_{Y_1}|x_A) = f_{12}(x_{Y_1}\,|\,\mathrm{do}(x_A)).
\end{aligned} \tag{2}
$$

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

Using Equation (2) and Lemma A.4, we have $\mathbb{E}_1[X_{Y_1}\,|\,\mathrm{do}(X_A = \mathbf{1})] = \mathbb{E}[X_{Y_1}|X_{A_1} = \mathbf{1}] = \mathrm{cov}(X_{Y_1}, X_{A_1}) = a$. To show that $\mathbb{E}_1[X_{Y_1}\,|\,\mathrm{do}(X_A = \mathbf{1})] \neq \mathbb{E}_2[X_{Y_1}\,|\,\mathrm{do}(X_A = 1)]$, 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 $c = \mathrm{cov}(X_{A_1}, X_{V_1})$ is equal the edge coefficient assigned to $A_1 \leftarrow V_1$ in $\mathcal{D}_{21}$, 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_i be the product of edge coefficients along $\langle A_1, V_i\rangle \oplus p_2(V_i, Y_1)$, $i \in \{2, \dots, r\}$. Then $a_i \in (0,1)$ for all $i \in \{1, \dots, r\}$. By Lemma A.3, we now have

$$
\begin{aligned}
a &= \mathrm{cov}(X_{Y_1}, X_{A_1}) = c \cdot a_1 + a_2 + \cdots + a_r, \\
b &= \mathrm{cov}(X_{Y_1}, X_{V_1}) = a_1 + c \cdot (a_2 + \cdots + a_r),
\end{aligned}
$$

which yields $b - ac = a_1(1-c^2) > 0$, completing the proof. □

Lemma B.1. Suppose that the total effect of A on Y is not identified given MPDAG $\mathcal{G}$. Let $p = \langle A_1 = V_0, V_1, \dots, V_k = Y_1\rangle$, $k \geq 1$, $A_1 \in A$, $Y_1 \in Y$, be a shortest proper possibly causal path from A to Y in $\mathcal{G}$. Let $\mathcal{G}_1 = \mathrm{MPDAG}(\mathcal{G}, \{A_1 \to V_1\})$ and $\mathcal{G}_2 = \mathrm{MPDAG}(\mathcal{G}, \{A_1 \leftarrow V_1\})$. Let p_1 and p_2 be the paths in $\mathcal{G}_1$ and $\mathcal{G}_2$ respectively, that consist of the same sequence of nodes as p in $\mathcal{G}$.

(i) If p is an unshielded path in $\mathcal{G}$, then
- p_1 is of the form $A_1 \to V_1 \to \cdots \to Y_1$, and
- p_2 is of the form $A_1 \leftarrow V_1 \dots Y_1$.

(ii) If p is a shielded path in $\mathcal{G}$, then
- $A_1 \to V_i$ is in $\mathcal{G}$ for all $i \in \{2, \dots, r\}, r \leq k$, $k > 1$,
- p_2 is of the form $A_1 \leftarrow V_1 \to \cdots \to Y$,
- Let $\mathcal{D}_1$ be a DAG in $[\mathcal{G}_1]$ and let p_{11} be the path in $\mathcal{D}_1$ corresponding to p_1 in $\mathcal{G}_1$ and to p in $\mathcal{G}$, then
  (a) p_{11} is of the form $A_1 \to V_1 \to \cdots \to Y_1$ in $\mathcal{D}_1$, or
  (b) p_{11} is of the form $A_1 \to V_1 \leftarrow \cdots \leftarrow V_l \to \cdots \to Y_1$, $1 < l \leq r$ in $\mathcal{D}_1$.

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 $\mathcal{G}$. 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 $\mathcal{G}$. This implies that no node V_i, $i \in \{2, \dots k-1\}$ can be a collider on either p_1 or p_2.

(i) Suppose first that p itself is unshielded. That is, no edge $\langle V_i, V_{i+2}\rangle$, $i \in \{0, k-2\}$ is in $\mathcal{G}$. Of course, since $\mathcal{G}_2$ contains edge $A_1 \leftarrow V_1$, p_2 is of the form $A_1 \leftarrow V_1 \dots Y_1$. Hence, we only need to show p_1 is a causal path in $\mathcal{G}_2$.

Since p is unshielded, p_1 is also an unshielded path. Since $A_1 \to V_1$ is in $\mathcal{G}_1$, as a consequence of iterative application of rule R1 of Meek (1995) (Fig. 3), p_1 is a causal path in $\mathcal{G}_1$.

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

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 $\mathcal{G}$. Furthermore, since p is chosen as a shortest proper possibly causal path from A to Y that starts with an undirected edge in $\mathcal{G}$, $\langle A_1, V_2\rangle$ must be of the form $A_1 \to V_2$.

If path $\langle A_1, V_2\rangle \oplus p(V_2, Y_1)$ is shielded, then by the same reasoning as above, $A_1 \to V_3$ is in $\mathcal{G}$. We can continue with the same reasoning, until we reach V_r, $r \in \{2, \dots, k\}$, so that $A_1 \to V_i$ is in $\mathcal{G}$ for $i \in \{2, \dots, r\}$ 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 $V_r \to \dots \to Y_1$. This is due to the fact that $A_1 \to V_r$ is in $\mathcal{G}$ and that $\langle A_1, V_r\rangle \oplus p(V_r, Y_1)$ is an unshielded possibly causal path in $\mathcal{G}$.

Next, we show that p_2 is of the form $A_1 \leftarrow V_1 \to \dots \to Y_1$. Since $A_1 \leftarrow V_1$ and $A_1 \to V_2$ are in $\mathcal{G}_2$ and since $\mathcal{G}_2$ is acyclic, by rule R2 of Meek (1995), the edge $\langle V_1, V_2\rangle$ is of the form $V_1 \to V_2$ in $\mathcal{G}_2$. Then since $p_2(V_1, Y_1)$ is an unshielded possibly causal path that starts with $V_1 \to V_2$, by iterative applications of rule R1 of Meek (1995), $p_2(V_1, Y_1)$ must be a causal path in $\mathcal{G}_2$.

Suppose $\mathcal{D}_1$ is a DAG in $[\mathcal{G}_1]$. From the above, we know that $A_1 \to V_1$ and if $r < k$, $V_r \to \dots Y_1$ are in $\mathcal{G}_1$ and therefore in $\mathcal{D}_1$ as well. The subpath $p(V_1, V_r)$ is a possibly causal unshielded path in $\mathcal{G}$ and hence, no node among $V_2, \dots, V_{r-1}$ 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 $\mathcal{D}_{11}$, in which case p_{11} is of the form $A_1 \to V_1 \to \dots \to Y_1$, or there is a node V_l, $1 < l \leq r$ on $p_{11}(V_{11}, V_r)$, such that $p_{11}(V_1, V_r)$ is of the form $V_1 \leftarrow \dots \leftarrow V_l \to \dots \to V_r$. $\square$

References

Lauritzen, S. L., Dawid, A. P., Larsen, B. N., and Leimer, H.-G. (1990). Independence properties of directed Markov fields. *Networks*, 20(5):491–505.

Mardia, K. V., Kent, J. T., and Bibby, J. M. (1980). *Multivariate Analysis (Probability and Mathematical Statistics)*. Academic Press London.

Meek, C. (1995). Causal inference and causal explanation with background knowledge. In *Proceedings of the 11th Annual Conference on Uncertainty in Artificial Intelligence (UAI-95)*, pages 403–410.

Pearl, J. (2009). *Causality*. Cambridge University Press, Cambridge, 2nd edition.

Perković, E. (2020). Identifying causal effects in maximally oriented partially directed acyclic graphs. In *Proceedings of the 36th Annual Conference on Uncertainty in Artificial Intelligence (UAI-20)*.

Perković, E., Kalisch, M., and Maathuis, M. H. (2017). Interpreting and using CPDAGs with background knowledge. In *Proceedings of the 33rd Annual Conference on Uncertainty in Artificial Intelligence (UAI-17)*.

Perković, E., Textor, J., Kalisch, M., and Maathuis, M. H. (2018). Complete graphical characterization and construction of adjustment sets in Markov equivalence classes of ancestral graphs. *Journal of Machine Learning Research*, 18(220):1–62.

Shpitser, I. and Pearl, J. (2006). Identification of joint interventional distributions in recursive semi-Markovian causal models. In *Proceedings of AAAI 2006*, pages 1219–1226.

Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, 20:557–585.