On Identifiability of Conditional Causal Effects (Supplementary Material)

Yaroslav Kivva[1] **Jalal Etesami**[1, 3] **Negar Kiyavash**[1,2]

[1]School of Computer and Communication Sciences, EPFL, Lausanne, Switzerland
[2]College of Management of Technology, EPFL, Lausanne, Switzerland
[3]TUM School of Computation, Information and Technology, Technical University of Munich

1 TECHNICAL PROOFS

1.1 NON C-GID CAUSAL EFFECTS

For proving Lemma 1, Lemma 4 and Lemma 7, it suffices to introduce two models that agree on the known distributions but disagree:

- on the causal effect $Q[\mathbf{L}'|\mathbf{L}'']$ (for Lemma 1),
- on the causal effect $P_{\mathbf{x}'}(\mathbf{d}|\mathbf{s}\backslash\mathbf{d})$ (for Lemma 4),
- on the causal effect $P_{\mathbf{x}'}(\widetilde{\mathbf{d}}|\mathbf{s}\backslash\widetilde{\mathbf{d}})$ (for Lemma 6).

To do so, we require a result from [Kivva et al., 2022] and couple of definitions and notations which we present in the next section.

1.1.1 Baseline Models

In this section, we present two models which we use as our baseline models for proving the non-identifiability parts.

Theorem 1. *Theorem 1 Kivva et al. [2022] Suppose $\breve{\mathbf{S}} \subseteq \mathbf{V}$ is a single c-component. $Q[\breve{\mathbf{S}}]$ is gID from $(\mathbb{A}, \mathcal{G})$ if and only if there exists $\mathbf{A} \in \mathbb{A}$ such that $\breve{\mathbf{S}} \subseteq \mathbf{A}$ and $Q[\breve{\mathbf{S}}]$ is ID in $\mathcal{G}[\mathbf{A}]$.*

To introduce the baseline models, we use the models from the proof of Theorem 1 in [Kivva et al., 2022]. Note that in the proof of Lemma 4 and 7, we use $\mathbf{S}_1$ and $\breve{\mathbf{S}}$ interchangeably, i.e., $\breve{\mathbf{S}} = \mathbf{S}_1$.

Suppose that $Q[\breve{\mathbf{S}}]$ is not gID from $(\mathbb{A}, \mathcal{G})$ and there exists $i \in [0, m]$, such that $\breve{\mathbf{S}} \subset \mathbf{A}_i$. Without loss of generality, let $\breve{\mathbf{S}} \subset \mathbf{A}_i$ for $i \in [0, \breve{k}]$ and $\breve{\mathbf{S}} \not\subseteq \mathbf{A}_i$ for $i \in [\breve{k}+1, m]$. This allows us to define a particular graph which we use throughout our proof. More precisely, under these assumptions, Lemma 2 and the above theorem guarantee that for each $i \in [0, \breve{k}]$, there exists a $\breve{\mathbf{S}}$-rooted c-forest $\mathcal{F}_i$ over a subset of observed variables $\mathbf{B}_i$ ($\breve{\mathbf{S}} \subsetneq \mathbf{B}_i \subseteq \mathbf{A}_i$) such that $\mathcal{F}_0[\breve{\mathbf{S}}] = \mathcal{F}_j[\breve{\mathbf{S}}]$ for $j \in [1, \breve{k}]$. In words, induced subgraphs of $\mathcal{F}_i$s over the set $\breve{\mathbf{S}}$ are the same. We define graph $\mathcal{G}'$ as the union of all the subgraphs in $\{\mathcal{F}_i\}_{i=0}^{\breve{k}}$ with the observed variables $\breve{\mathbf{V}} := \bigcup_{i=0}^{\breve{k}} \mathbf{B}_i$ and the unobserved variables which we denoted by $\breve{\mathbf{U}}$.

To properly define a SEM $\mathcal{M}$ over a causal graph $\mathcal{G}$, it suffices to define the domain set of each node X in $\mathcal{G}$ with its associated conditional distribution $P(X|Pa_{\mathcal{G}}(X))$. Note that if for some variable X in $\mathcal{G}$, its domain $\mathfrak{X}(X)$ or $P(X|Pa_{\mathcal{G}}(X))$ are not specified, then by default, we assume $\mathfrak{X}(X) := \{0\}$ and $P(X = 0|Pa_{\mathcal{G}}(X)) = 1$.

Let U_0 be an unobserved variable from subgraph $\mathcal{F}_0$ that has one child in $\breve{\mathbf{S}}$ and one child in $\breve{\mathbf{T}} := \breve{\mathbf{V}}\backslash\breve{\mathbf{S}}$. In high-level, our

Accepted for the 39th Conference on Uncertainty in Artificial Intelligence (UAI 2023).

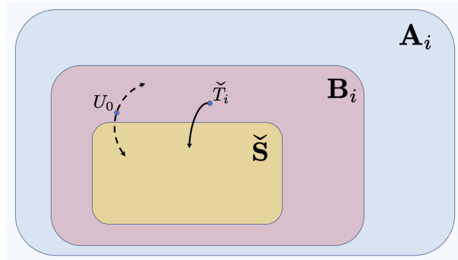

Figure 1: An illustration of the definitions $\mathbf{B}_i, U_0, \mathbf{A}_i, \check{\mathbf{S}}$, and $\check{T}_i$.

baseline models $\mathcal{M}_1$ and $\mathcal{M}_2$ have the same distributions over all variables in graph $\mathcal{G}$ except the variable U_0. Especially,

$$P^{M_1}(V|Pa_{\mathcal{G}}(V)) = P^{\mathcal{M}_2}(V|Pa_{\mathcal{G}}(V)), \quad V \in \mathbf{V}, \forall \tag{1}$$

$$P^{\mathcal{M}_1}(U) = P^{\mathcal{M}_2}(U) = \frac{1}{|\mathfrak{X}(U)|}, \quad \forall U \in \mathbf{U} \backslash \{U_0\}, \tag{2}$$

where $|\cdot|$ denotes the cardinality of a given set. For the sake of brevity, we drop the superscripts $\mathcal{M}_1$ and $\mathcal{M}_2$ for the distributions in Equations (1) and (2). We denote the domain of variable U_0 to be $\mathfrak{X}(U_0) := \{\gamma_1, \ldots, \gamma_d\}$, where γ_js are vectors and d is an integer number to be defined later. In model $\mathcal{M}_1$, we define U_0 to have uniform distribution over $\mathfrak{X}(U_0)$, i.e., $P^{\mathcal{M}_1}(U_0 = \gamma_j) = 1/d$. In model $\mathcal{M}_2$, we define $P^{\mathcal{M}_2}(U_0 = \gamma_j) := p_j$, where $j \in [1:d]$ and

$$\sum_{j=1}^{d} p_j = 1,$$
$$p_j > 0 \quad \forall j \in [1:d].$$

For $i \in [0:m]$, $j \in [1:d]$, $u_0 = \gamma_j$ and any $\mathbf{v} \in \mathfrak{X}(\mathbf{V})$, we define:

$$\begin{aligned} \theta_{i,j}(\mathbf{v}) &:= \sum_{\mathbf{U} \backslash \{U_0\}} \prod_{X \in \mathbf{A}_i} P(x \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U} \backslash \{U_0\}} P(u), \\ \eta_j(\mathbf{v}) &:= \sum_{\mathbf{U} \backslash \{U_0\}} \prod_{X \in \check{\mathbf{S}}} P(x \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U} \backslash \{U_0\}} P(u). \end{aligned} \tag{3}$$

Note that in the above equations, $u_0 = \gamma_j$ may appear as a parent of an observed variable. Using the above definitions, we can re-write the Q-notation in Equation (3) as follows

$$Q^{\mathcal{M}_1}[\mathbf{A}_i](\mathbf{v}) = \sum_{j=1}^{d} \frac{1}{d} \theta_{i,j}(\mathbf{v}), \tag{4}$$

$$Q^{\mathcal{M}_2}[\mathbf{A}_i](\mathbf{v}) = \sum_{j=1}^{d} p_j \theta_{i,j}(\mathbf{v}), \tag{5}$$

$$Q^{\mathcal{M}_1}[\check{\mathbf{S}}](\mathbf{v}) = \sum_{j=1}^{d} \frac{1}{d} \eta_j(\mathbf{v}), \tag{6}$$

$$Q^{\mathcal{M}_2}[\check{\mathbf{S}}](\mathbf{v}) = \sum_{j=1}^{d} p_j \eta_j(\mathbf{v}). \tag{7}$$

Denote the set of unobserved variables in $\mathcal{G}'[\check{\mathbf{S}}]$ by $\check{\mathbf{U}}^{\check{\mathbf{S}}}$ and its complement set in $\check{\mathbf{U}} \backslash (\check{\mathbf{U}}^{\check{\mathbf{S}}} \cup \{U_0\})$ by $\check{\mathbf{U}}^{\check{\mathbf{T}}}$. For each $i \in [0:\check{k}]$, let $\check{T}_i$ be a node in $\mathbf{B}_i \backslash \check{\mathbf{S}}$ such that $Ch_{\mathcal{F}_i}(\check{T}_i) \neq \varnothing$. Node $\check{T}_i$ exists because $\mathcal{F}_i$ is a $\check{\mathbf{S}}$-rooted c-forest. Figure 1 illustrates an example of the above definitions.

We define the domains of $X \in \check{\mathbf{V}} \cup \check{\mathbf{U}}$ as follows. Note that $\check{\mathbf{V}} = \check{\mathbf{S}} \cup \check{\mathbf{T}}$ and $\check{\mathbf{U}} = \check{\mathbf{U}}^{\check{\mathbf{S}}} \cup \check{\mathbf{U}}^{\check{\mathbf{T}}} \cup \{U_0\}$.

$$\begin{aligned}
&\mathfrak{X}(X) := [0:\kappa], \quad \forall X \in \check{\mathbf{S}},\\
&\mathfrak{X}(X) := [0:\kappa], \quad \forall X \in \check{\mathbf{U}}^{\check{\mathbf{S}}},\\
&\mathfrak{X}(X) := \{0,1\}^{\alpha(T)}, \quad \forall X \in \check{\mathbf{T}},\\
&\mathfrak{X}(X) := \{0,1\}^{\alpha(U)}, \quad \forall X \in \mathbf{U}^{\check{\mathbf{T}}},\\
&\mathfrak{X}(X) := [0:\kappa] \times \{0,1\}^{\alpha(U_0)-1}, \quad X = U_0,
\end{aligned}$$

where κ is an odd number greater than 4. Function $\alpha(X)$ is only defined for $X \in \check{\mathbf{T}} \cup \check{\mathbf{U}}^{\check{\mathbf{T}}} \cup \{U_0\}$ and it denotes the number of subgraphs in $\{\mathcal{F}_j\}_{j=0}^{k}$ that contains X. From the above definition, it is clear that d, the domain size of U_0 is equal to $(\kappa+1)2^{\alpha(U_0)-1}$.

Suppose that $X \in \check{\mathbf{T}} \cup \check{\mathbf{U}}^{\check{\mathbf{T}}} \cup \{U_0\}$ and X belongs to $\mathcal{F}_{i_1}, \mathcal{F}_{i_2}, \cdots, \mathcal{F}_{i_{\alpha(X)}}$, where $i_1 < i_2 < \cdots < i_{\alpha(X)}$. We use $(X[i_1], X[i_2], \ldots, X[i_{\alpha X}])$ to represent X. Note that depending on where X belongs to, its vector size is different. If $X \in \check{\mathbf{U}}^{\check{\mathbf{T}}} \cup \{U_0\}$, both its distribution and its domain are specified above. If $X \in \check{\mathbf{T}}$, we define the entries of its corresponding vector as

$$X[i_j] \equiv \left(\sum_{Y \in Pa_{\mathcal{F}_{i_j}}(X)} Y[i_j] \right) \pmod 2,$$

where $j \in [1:\alpha(X)]$. This specifies the distribution of $P(X|Pa_{\mathcal{G}}(X))$ for $X \in \check{\mathbf{T}}$. What is left to specify is the domains and the distributions of variables in $\check{\mathbf{S}}$.

Recall that U_0 has one child in $\check{\mathbf{S}}$ and one child in $\check{\mathbf{T}}$. We denote the child in $\check{\mathbf{S}}$ by S_0. For each $S \in \check{\mathbf{S}} \backslash \{S_0\}$ and any realization of $Pa_{\mathcal{G}'}(S)$, we define $\mathbb{I}(S)$ to be one if there exists $i \in [0:k]$ such that

1. $\check{T}_i \in Pa_{\mathcal{G}'}(S)$ and $\check{T}_i[i] = 0$, or
2. there exists $X \in Pa_{\mathcal{G}'}(S) \backslash (\check{\mathbf{U}}^{\check{\mathbf{S}}} \cup \{\check{T}_i\})$ such that $\mathcal{F}_i$ contains X and $X[i] = 1$,

and zero, otherwise. It is noteworthy that according to the definition of $\check{T}_i$, it belongs to $\mathbf{B}_i \backslash \check{\mathbf{S}} \subseteq \check{\mathbf{T}}$ which means $\check{T}_i[i]$ is well-defined according to the above definition. Note that the above definition holds for all $S \in \check{\mathbf{S}} \backslash \{S_0\}$. When $S = S_0$, we define $\mathbb{I}(S_0)$ to be one if there exists $i \in [0:k]$, such that

1. $\check{T}_i \in Pa_{\mathcal{G}'}(S)$ and $\check{T}_i[i] = 0$, or
2. $i \neq 0$, $\mathcal{F}_i$ contains U_0, and $U_0[i] = 1$, or
3. there exists $X \in Pa_{\mathcal{G}'}(S) \backslash (\check{\mathbf{U}}^{\check{\mathbf{S}}} \cup \{\check{T}_i, U_0\})$ such that $\mathcal{F}_i$ contains X and $X[i] = 1$.

For each $S \in \check{\mathbf{S}}$, we define $\mathfrak{X}(S) := \{0, ..., \kappa\}$ and for $s \in \mathfrak{X}(S)$,

$$P(S = s \mid Pa_{\mathcal{G}}(S)) := \begin{cases} \frac{1}{\kappa+1} & \text{if } \mathbb{I}(S) = 1 \\ 1 - \kappa\epsilon & \text{if } \mathbb{I}(S) = 0 \text{ and } s \equiv M(S) \pmod{\kappa+1}, \\ \epsilon & \text{if } \mathbb{I}(S) = 0 \text{ and } s \not\equiv M(S) \pmod{\kappa+1}, \end{cases} \tag{8}$$

where $0 < \epsilon < \frac{1}{\kappa}$ and

$$M(S) := \begin{cases} \sum_{x \in Pa_{\mathcal{G}'[\check{\mathbf{S}}]}(S)} x & \text{, if } S \in \check{\mathbf{S}} \backslash \{S_0\}, \\ u_0[0] + \sum_{x \in Pa_{\mathcal{G}'[\check{\mathbf{S}}]}(S)} x & \text{, if } S = S_0 \, . \end{cases} \tag{9}$$

Note that $M(S)$ is an integer number because $Pa_{\mathcal{G}'[\check{\mathbf{S}}]}(S) \subseteq \check{\mathbf{U}}^{\check{\mathbf{S}}}$ and thus all terms in the summations in (9) belong to $[0:\kappa]$. With this, we finish defining the models and now we are ready to present some of their properties.

Let $\mathbf{\Gamma}$ denote a subset of $\mathfrak{X}(U_0) = \{\gamma_1, ..., \gamma_d\}$ with $\frac{\kappa+1}{2}$ elements that is given by

$$\mathbf{\Gamma} := \Big\{ (2x, 0, \cdots, 0) : \; x \in [0 : \frac{\kappa - 1}{2}] \Big\}.$$

Recall that for $\mathbf{v} \in \mathfrak{X}(\mathbf{V})$ and $i \in [0:m]$, $\theta_i(\mathbf{v})$ and $\eta(\mathbf{v})$ are two vectors in $\mathbb{R}^d$ with j-th entry corresponding to $U_0 = \gamma_j$. Suppose that $\mathbf{\Gamma} = \{\gamma_{j_1}, ..., \gamma_{j_{\frac{\kappa+1}{2}}}\}$. Next result shows that in the constructed models, all entries of $\theta_i(\mathbf{v})$ with indices in $\{j_1, ..., j_{\frac{\kappa+1}{2}}\}$ are equal.

Lemma 1 (Kivva et al. [2022]). *For any $\mathbf{v} \in \mathfrak{X}(\mathbf{V})$, $i \in [0:m]$, and both models, we have*

$$\theta_{i,j_1}(\mathbf{v}) = \theta_{i,j_2}(\mathbf{v}) = \cdots = \theta_{i,j_{\frac{\kappa+1}{2}}}(\mathbf{v}).$$

The next two lemmas are used to prove the existence of parameters ϵ and $\{p_j\}_{j=1}^d$ such that the constructed models $\mathcal{M}_1$ and $\mathcal{M}_2$ agree on the known distributions but disagree on the target causal effect.

Lemma 2 (Kivva et al. [2022]). *There exists $0 < \epsilon < \frac{1}{\kappa}$ such that there exists $\mathbf{v}_0 \in \mathfrak{X}(\mathbf{V})$ and $1 \leqslant r < t \leqslant \frac{\kappa+1}{2}$ such that*

$$\eta_{j_r}(\mathbf{v}_0) \neq \eta_{j_t}(\mathbf{v}_0).$$

Lemma 3 (Kivva et al. [2022]). *Consider a set of vectors $\{c_i\}_{i=1}^n$, where $c_i \in \mathbb{R}^d$. Assume $c \in \mathbb{R}^d$ is a vector that is linearly independent of $\{c_i\}_{i=1}^n$, then there is a vector $b \in \mathbb{R}^d$ such that*

$$\begin{aligned} &\langle c_i, b\rangle = 0, \quad \forall i \in [1:n], \\ &\langle c, b\rangle \neq 0. \end{aligned}$$

1.2 PROOF OF LEMMA 1

Herein, we present the proof of our first lemma. But first, we need the following technical lemmas. Assume that $\breve{\mathbf{S}}'$ and $\breve{\mathbf{S}}''$ are two disjoint non-empty subsets of $\breve{\mathbf{S}}$, such that $\breve{\mathbf{S}} = \breve{\mathbf{S}}' \cup \breve{\mathbf{S}}''$.

Let $\breve{\mathbf{S}}^\dagger := \mathbf{V}\backslash\breve{\mathbf{S}}'$ and $\breve{\mathbf{s}}^\dagger \in \mathfrak{X}(\breve{\mathbf{S}}^\dagger)$. For $u_0 = \gamma_j$, where $j \in [1:d]$, we define

$$\phi_j(\breve{\mathbf{s}}^\dagger) := \sum_{\breve{\mathbf{S}}'} \sum_{\breve{\mathbf{U}}'\backslash\{U_0\}} \prod_{X \in \breve{\mathbf{S}}} P(x|Pa_{\mathcal{G}}(X)) \prod_{U \in \breve{\mathbf{U}}'\backslash\{U_0\}} P(u). \tag{10}$$

Note that U_0 may appear as a parent of some observed variables in the above equation. Recall that $\mathbf{\Gamma} = \{\gamma_{j_1}, ..., \gamma_{j_{\frac{\kappa+1}{2}}}\}$.

Lemma 4. *For any $\breve{\mathbf{s}}^\dagger \in \mathfrak{X}(\breve{\mathbf{S}}^\dagger)$ and both models, we have*

$$\phi_{j_1}(\breve{\mathbf{s}}^\dagger) = \phi_{j_2}(\breve{\mathbf{s}}^\dagger) = \cdots = \phi_{j_{\frac{\kappa+1}{2}}}(\breve{\mathbf{s}}^\dagger).$$

Proof. We fix a realization $\breve{\mathbf{s}}^\dagger$ of $\breve{\mathbf{S}}^\dagger$. Suppose that l_1 and l_2 are two integers, such that

$$\begin{aligned} \gamma_{l_1} &= (2x, 0, \ldots, 0), \\ \gamma_{l_2} &= (2x + 2 \pmod{\kappa+1}, 0, \ldots, 0), \end{aligned}$$

and x is an integer in $[0 : \frac{\kappa-1}{2}]$. Recall that

$$\phi_j(\breve{\mathbf{s}}^\dagger) := \sum_{\breve{\mathbf{S}}'} \sum_{\breve{\mathbf{U}}\backslash\{U_0\}} \prod_{X \in \breve{\mathbf{S}}} P(x|Pa_{\mathcal{G}}(X)) \prod_{U \in \breve{\mathbf{U}}\backslash\{U_0\}} P(u).$$

We consider two cases:

1. Suppose that there exists a variable $S \in \breve{\mathbf{S}}$ such that $\mathbb{I}(S) = 1$. Then, there is a sequence of variables $U_0, \hat{S}_1, \hat{U}_1, \hat{S}_2, \hat{U}_2, \ldots, \hat{U}_l, S$, such that U_0 is a parent of $\hat{S}_1$, $S \in \breve{\mathbf{S}}$ is a children of $\hat{U}_l \in \breve{\mathbf{U}}^{\breve{\mathbf{S}}}$ and $\hat{U}_j \in \breve{\mathbf{U}}^{\breve{\mathbf{S}}}$ is a parent of variables $\hat{S}_j$ and $\hat{S}_{j+1}$ for $j \in [1:l-1]$. Let $\hat{\mathbf{U}} := \{\hat{U}_1, \ldots, \hat{U}_l\}$. For a given realization $\mathbf{u}_1$ of $\breve{\mathbf{U}}^{\breve{\mathbf{S}}}$, we define $\mathbf{u}_2 \in \mathfrak{X}(\breve{\mathbf{U}}^{\breve{\mathbf{S}}})$ by

$$\begin{aligned} &\mathbf{u}_2[\hat{U}_j] := \mathbf{u}_1[\hat{U}_j] + 2(-1)^j \pmod{\kappa+1}, \quad j \in [1:l], \\ &\mathbf{u}_2[U] := \mathbf{u}_1[U], \quad \forall U \in \breve{\mathbf{U}}^{\breve{\mathbf{S}}}\backslash\hat{\mathbf{U}}. \end{aligned} \tag{11}$$

This implies

$$P(\hat{s}|Pa_{\mathcal{G}}(\hat{S}))\Big|_{(\breve{\mathbf{U}}^{\mathbf{S}},U_0)=(\mathbf{u}_1,\gamma_{l_1})} = P(\hat{s}|Pa_{\mathcal{G}}(\hat{S}))\Big|_{(\breve{\mathbf{U}}^{\mathbf{S}},U_0)=(\mathbf{u}_2,\gamma_{l_2})},$$

for any $\hat{S} \in \breve{\mathbf{S}}$ and consequently, $\phi_{l_1}(\breve{\mathbf{s}}^\dagger) = \phi_{l_2}(\breve{\mathbf{s}}^\dagger)$.

2. Suppose that there is no variable in $\breve{\mathbf{S}}$ with $\mathbb{I}(\cdot) = 1$. Denote by S a node in $\breve{\mathbf{S}}'$ with the shortest path to the node U_0 by bidirected edges. Suppose $\hat{s}$ is a realization of S and the shortest path is $U_0, \hat{S}_1, \hat{U}_1, \hat{S}_2, \hat{U}_2, \ldots, \hat{U}_l, S$, so that U_0 is a parent of $\hat{S}_1$, S is a child of $\hat{U}_l \in \breve{\mathbf{U}}^{\breve{\mathbf{S}}}$ and $\hat{U}_j \in \breve{\mathbf{U}}^{\breve{\mathbf{S}}}$ is a parent of variables $\hat{S}_j$ and $\hat{S}_{j+1}$ for $j \in [1 : l-1]$. Let $\hat{\mathbf{U}} := \{\hat{U}_1, \ldots, \hat{U}_l\}$. For a given realization $\mathbf{u}_1$ of $\breve{\mathbf{U}}^S$, we define $\mathbf{u}_2 \in \mathfrak{X}(\breve{\mathbf{U}}^{\breve{\mathbf{S}}})$ by

$$\begin{aligned}&\mathbf{u}_2[\hat{U}_j] := \mathbf{u}_1[\hat{U}_j] + 2(-1)^j \pmod{\kappa+1}, \quad j \in [1:l],\\ &\mathbf{u}_2[U] := \mathbf{u}_1[U], \quad \forall U \in \breve{\mathbf{U}}^{\breve{\mathbf{S}}}\backslash\hat{\mathbf{U}},\end{aligned} \tag{12}$$

For a given realization $\breve{\mathbf{s}}_1$ of $\breve{\mathbf{S}}'$, we define $\breve{\mathbf{s}}_2 \in \mathfrak{X}(\breve{\mathbf{S}}')$ as follows

$$\begin{aligned}&\breve{\mathbf{s}}_2[S''] := \breve{\mathbf{s}}_1[S''], \quad \forall S'' \in \breve{\mathbf{S}}'\backslash\{S\},\\ &\breve{\mathbf{s}}_2[S] := \breve{\mathbf{s}}_1[S] + 2(-1)^l \pmod{\kappa+1},\end{aligned} \tag{13}$$

Note that with the above modifications for any $\widetilde{S} \in \breve{\mathbf{S}}$, we get

$$\widetilde{s}_1 - M(\widetilde{S}) \equiv \widetilde{s}_2 - M(\widetilde{S}) \pmod{\kappa+1},$$

where $\widetilde{s}_1$ is a realization of $\widetilde{S}\Big|_{(\breve{\mathbf{S}}^\dagger,\breve{\mathbf{S}}')=(\breve{\mathbf{s}}^\dagger,\breve{\mathbf{s}}_1)}$, $\widetilde{s}_2$ is a realization of $\widetilde{S}\Big|_{(\breve{\mathbf{S}}^\dagger,\breve{\mathbf{S}}')=(\breve{\mathbf{s}}^\dagger,\breve{\mathbf{s}}_2)}$, and $M(\cdot)$ is given by Equation (9). Therefore:

$$P(\widetilde{s}|Pa_{\mathcal{G}}(\widetilde{S}))\Big|_{(\breve{\mathbf{U}}^{\mathbf{S}},U_0,S)=(\mathbf{u}_1,\gamma_{l_1},s_1)} = P(\widetilde{s}|Pa_{\mathcal{G}}(\widetilde{S}))\Big|_{(\breve{\mathbf{U}}^{\mathbf{S}},U_0,S)=(\mathbf{u}_2,\gamma_{l_2},s_2)},$$

for any $\widetilde{S} \in \breve{\mathbf{S}}$ and thus $\phi_{l_1}(\breve{\mathbf{s}}^\dagger) = \phi_{l_2}(\breve{\mathbf{s}}^\dagger)$.

To summarize, we proved that $\phi_{l_1}(\breve{\mathbf{s}}^\dagger) = \phi_{l_2}(\breve{\mathbf{s}}^\dagger)$. By varying x within $[0 : \frac{\kappa-1}{2}]$ in the definition of γ_{l_1} and γ_{l_2}, we conclude the lemma. □

In order to have consistent notations in the appendix, we restate Lemma 1 using $\breve{\mathbf{S}}, \breve{\mathbf{S}}', \breve{\mathbf{S}}''$ instead of $\mathbf{L}, \mathbf{L}', \mathbf{L}''$ respectively.

Lemma 1. *Suppose $\breve{\mathbf{S}} \subseteq \breve{\mathbf{V}}$ is a single c-component, such that $\breve{\mathbf{S}} = \breve{\mathbf{S}}' \cup \breve{\mathbf{S}}''$ for some disjoint sets $\breve{\mathbf{S}}'$ and $\breve{\mathbf{S}}''$. $Q[\breve{\mathbf{S}}'|\breve{\mathbf{S}}'']$ is c-gID from $(\mathbb{A}, \mathcal{G})$ if and only if $Q[\breve{\mathbf{S}}' \cup \breve{\mathbf{S}}'']$ is gID from $(\mathbb{A}, \mathcal{G})$.*

Proof.
Sufficiency: We use Assume that $Q[\breve{\mathbf{S}}' \cup \breve{\mathbf{S}}'']$ is gID from $(\mathbb{A}, \mathcal{G})$, then $Q[\breve{\mathbf{S}}'|\breve{\mathbf{S}}'']$ is c-gID from $(\mathbb{A}, \mathcal{G})$. This is an immediate result of applying Equation (4), i.e.,

$$Q[\breve{\mathbf{S}}'|\breve{\mathbf{S}}''](\mathbf{v}) = \frac{Q[\breve{\mathbf{S}}](\mathbf{v})}{\sum_{\breve{\mathbf{S}}'} Q[\breve{\mathbf{S}}](\mathbf{v})}$$

Necessity: We prove this by contradiction. Assume that $Q[\breve{\mathbf{S}}' \cup \breve{\mathbf{S}}'']$ is not gID from $(\mathbb{A}, \mathcal{G})$. We will show that $Q[\breve{\mathbf{S}}'|\breve{\mathbf{S}}'']$ is not c-gID from $(\mathbb{A}, \mathcal{G})$. To this end, we will construct two models $\mathcal{M}_1$ and $\mathcal{M}_2$ such that for each $i \in [0 : m]$ and any $\mathbf{v} \in \mathbf{V}$:

$$Q^{\mathcal{M}_1}[\mathbf{A}_i](\mathbf{v}) = Q^{\mathcal{M}_2}[\mathbf{A}_i](\mathbf{v}), \tag{14}$$

$$\sum_{\breve{\mathbf{S}}'} Q^{\mathcal{M}_1}[\breve{\mathbf{S}}](\mathbf{v}') = \sum_{\breve{\mathbf{S}}'} Q^{\mathcal{M}_2}[\breve{\mathbf{S}}](\mathbf{v}'), \tag{15}$$

but there exists $\mathbf{v}_0 \in \mathfrak{X}(\mathbf{V})$ such that:

$$Q^{\mathcal{M}_1}[\breve{\mathbf{S}}](\mathbf{v}_0) \neq Q^{\mathcal{M}_2}[\breve{\mathbf{S}}](\mathbf{v}_0). \tag{16}$$

Equations (15)-(16) yield

$$Q[\breve{\mathbf{S}}'|\breve{\mathbf{S}}'']^{\mathcal{M}_1}(\mathbf{v}_0) \neq Q[\breve{\mathbf{S}}'|\breve{\mathbf{S}}'']^{\mathcal{M}_2}(\mathbf{v}_0).$$

This means that $Q[\breve{\mathbf{S}}'|\breve{\mathbf{S}}'']$ is not c-gID from $(\mathbb{A}, \mathcal{G})$.

We consider two cases.

First case: Suppose that there exists $i \in [0, m]$, such that $\check{\mathbf{S}} \subset \mathbf{A}_i$. For this, we consider the models constructed in the section 1.1.1:

$$\sum_{\check{\mathbf{S}}'} Q[\check{\mathbf{S}}]^{\mathcal{M}_1}(\mathbf{v}) = \sum_{j=1}^{d} \frac{1}{d}\phi_j(\mathbf{v}[\check{\mathbf{S}}^{\dagger}]), \tag{17}$$

$$\sum_{\check{\mathbf{S}}'} Q[\check{\mathbf{S}}]^{\mathcal{M}_2}(\mathbf{v}) = \sum_{j=1}^{d} p_j\phi_j(\mathbf{v}[\check{\mathbf{S}}^{\dagger}]). \tag{18}$$

and according to the Equations (3) and (10), we have

$$Q^{\mathcal{M}_2}[\mathbf{A}_i](\mathbf{v}) - Q^{\mathcal{M}_1}[\mathbf{A}_i](\mathbf{v}) = \sum_{j=1}^{d}(p_j - \frac{1}{d})\theta_{i,j}(\mathbf{v}) \tag{19}$$

$$\sum_{\check{\mathbf{S}}'} Q[\check{\mathbf{S}}]^{\mathcal{M}_2}(\mathbf{v}) - \sum_{\check{\mathbf{S}}'} Q[\check{\mathbf{S}}]^{\mathcal{M}_1}(\mathbf{v}) = \sum_{j=1}^{d}(p_j - \frac{1}{d})\phi_j(\mathbf{v}[\check{\mathbf{S}}^{\dagger}]) \tag{20}$$

$$Q^{\mathcal{M}_2}[\check{\mathbf{S}}](\mathbf{v}_0) - Q^{\mathcal{M}_1}[\check{\mathbf{S}}](\mathbf{v}_0) = \sum_{j=1}^{d}(p_j - \frac{1}{d})\eta_j(\mathbf{v}_0) \tag{21}$$

$$\sum_{j=1}^{d} p_j - 1 = \sum_{j=1}^{d}(p_j - \frac{1}{d}) \tag{22}$$

Therefore, it suffices to solve a system of linear equations over parameters $\{p_j\}_{j=1}^d$ and show that it admits a solution.

$$\sum_{j=1}^{d}(p_j - \frac{1}{d})\theta_{i,j}(\mathbf{v}) = 0, \quad \forall \mathbf{v} \in \mathfrak{X}(\mathbf{V}), i \in [0, m], \tag{23}$$

$$\sum_{j=1}^{d}(p_j - \frac{1}{d})\phi_j(\check{\mathbf{s}}^{\dagger}) = 0, \quad \forall \check{\mathbf{s}}^{\dagger} \in \mathfrak{X}(\check{\mathbf{S}}^{\dagger}), i \in [0, m], \tag{24}$$

$$\sum_{j=1}^{d}(p_j - \frac{1}{d})\eta_j(\mathbf{v}_0) \neq 0, \quad \exists \mathbf{v}_0 \in \mathfrak{X}(\check{\mathbf{V}}), \tag{25}$$

$$(p_j - \frac{1}{d}) = 0, \tag{26}$$

$$0 < p_j < 1, \quad \forall j \in [1 : d]. \tag{27}$$

However, the system of linear equations (23)-(27) admits a solution with respect to $\{p_j\}_{j=1}^d$ if and only if the following system of equations has a solution with respect to parameters $\{\beta_j\}_{j=1}^d$:

$$\sum_{j=1}^{d} \beta_j\theta_{i,j}(\mathbf{v}) = 0, \quad \forall \mathbf{v} \in \mathfrak{X}(\mathbf{V}), i \in [0 : m] \tag{28}$$

$$\sum_{j=1}^{d} \beta_j\phi_j(\check{\mathbf{s}}^{\dagger}) = 0, \quad \forall \check{\mathbf{s}}^{\dagger} \in \mathfrak{X}(\check{\mathbf{S}}^{\dagger}), i \in [0 : m] \tag{29}$$

$$\sum_{j=1}^{d} \beta_j\eta_j(\mathbf{v}_0) \neq 0, \quad \exists \mathbf{v}_0 \in \mathfrak{X}(\mathbf{V}) \tag{30}$$

$$\sum_{j=1}^{d} \beta_j = 0. \tag{31}$$

Clearly, if $\{\beta_j^*\}$ is a solution for system (28)-(31), then

$$p_j^* := \frac{1}{d} + \frac{\beta_j^*}{2hd}, \tag{32}$$

is a solution for (23)-(27), where $h := \max_{j\in[1:d]} |\beta_j^*|$.

According to Lemma 1 and Lemma 4, for any $i \in [0:m]$, $\mathbf{v} \in \mathfrak{X}(\mathbf{V})$ and $\breve{\mathbf{s}}^\dagger \in \mathfrak{X}(\breve{\mathbf{S}}^\dagger)$, we have

$$\theta_{i,j_1}(\mathbf{v}) = \theta_{i,j_2}(\mathbf{v}) = \cdots = \theta_{i,j_{\frac{\kappa+1}{2}}}(\mathbf{v}),$$
$$\phi_{j_1}(\breve{\mathbf{s}}^\dagger) = \phi_{j_2}(\breve{\mathbf{s}}^\dagger) = \cdots = \phi_{j_{\frac{k+1}{2}}}(\breve{\mathbf{s}}^\dagger),$$

and by Lemma 2, we know that there exists $\mathbf{v}_0 \in \mathfrak{X}(\mathbf{V})$ and $1 \leqslant r < t \leqslant \frac{\kappa+1}{2}$ such that

$$\eta_{j_r}(\mathbf{v}_0) \neq \eta_{j_t}(\mathbf{v}_0).$$

The latter means that the vector $\big(\eta_1(\mathbf{v}_0), \eta_2(\mathbf{v}_0), \ldots, \eta_d(\mathbf{v}_0)\big)$ is linearly independent from vectors:

$$\big(1, 1, \ldots, 1\big), \tag{33}$$
$$\big(\theta_{i,1}(\mathbf{v}), \theta_{i,2}(\mathbf{v}), \ldots, \theta_{i,d}(\mathbf{v})\big), \quad \forall \mathbf{v} \in \mathfrak{X}(\mathbf{V}), \quad \forall i \in [0:m], \tag{34}$$
$$\big(\phi_1(\breve{\mathbf{s}}^\dagger), \phi_2(\breve{\mathbf{s}}^\dagger), \ldots, \phi_d(\breve{\mathbf{s}}^\dagger)\big), \quad \forall \breve{\mathbf{s}}^\dagger \in \mathfrak{X}(\breve{\mathbf{S}}^\dagger). \tag{35}$$
$$\tag{36}$$

Combining the last result with Lemma 3 imply the existence of a solution $\{\beta_j^*\}$ and subsequently the existence of two models $\mathcal{M}_1$ and $\mathcal{M}_2$ that satisfy equation (14), (15) and (16).

Second case: Suppose that there is no $i \in [0, m]$, such that $\breve{\mathbf{S}} \subset \mathbf{A}_i$. Suppose $S^* \in \breve{\mathbf{S}}$ and denote by $\mathcal{G}^*$ the graph obtained from graph $\mathcal{G}$ through the following procedure:

1. Add nodes T_0^* and U_0^* to graph $\mathcal{G}$.
2. Draw a direct edge from T_0^* to S^*.
3. Draw direct edges from U_0^* to S^* and T_0^*.

We define $\mathbf{A}_{m+1} := \breve{\mathbf{S}} \cup \{T_0^*\}$ and $\mathbb{A}^* := \mathbb{A} \cup \{\mathbf{A}_{m+1}\}$. To summarize, we have

- $\mathbf{V}$ is a set of all observed variables in graph $\mathcal{G}$;
- $\mathbf{U}$ is a set of all unobserved variables in graph $\mathcal{G}$;
- $\mathbf{V}^* = \mathbf{V} \cup \{T_0^*\}$;
- $\mathbf{U}^* = \mathbf{U} \cup \{U_0^*\}$.

Note that $Q[\breve{\mathbf{S}}]$ is not identifiable in $\mathcal{G}^*[\mathbf{A}_{m+1}]$ and therefore $Q[\breve{\mathbf{S}}]$ remains not gID from $(\mathbb{A}^*, G^*)$. Since $\breve{\mathbf{S}} \in \mathbf{A}_{m+1}$, according to the **First case**, we can construct models $\mathcal{M}_1^*$ and $\mathcal{M}_2^*$ for the graph $\mathcal{G}^*$ and set $\mathbb{A}^*$. These two models satisfy the following properties

- $\mathfrak{X}(U_0^*) = [0:\kappa]$ and $d = \kappa + 1$.
- $\mathfrak{X}(T_0^*) = \{0, 1\}$.

For the graph $\mathcal{G}^*$, we define

$$\theta_{i,j}(\mathbf{v}, T_0^* = t_0) := \sum_{\mathbf{U}} \prod_{X\in\mathbf{A}_i} P(x \mid Pa_{\mathcal{G}*}(X)) \prod_{U\in\mathbf{U}} P(u), \quad i \in [0:m],\ j \in [1:d],$$
$$\theta_{m+1,j}(\mathbf{v}, T_0^* = t_0) := \sum_{\mathbf{U}} P(T_0^* = t_0) \prod_{X\in\mathbf{A}_i} P(x \mid Pa_{\mathcal{G}*}(X)) \prod_{U\in\mathbf{U}} P(u), \quad j \in [1:d],$$
$$\phi_j(\breve{\mathbf{s}}^\dagger) := \sum_{\breve{\mathbf{S}}'} \sum_{\breve{\mathbf{U}}} \prod_{X\in\breve{\mathbf{S}}} P(x|Pa_{\mathcal{G}*}(X)) \prod_{U\in\breve{\mathbf{U}}} P(u), \quad j \in [1:d]$$
$$\eta_j(\mathbf{v}, T_0^* = t_0) := \sum_{\mathbf{U}} \prod_{X\in\breve{\mathbf{S}}} P(x \mid Pa_{\mathcal{G}*}(X)) \prod_{U\in\mathbf{U}} P(u).$$

Now, we are ready to construct two models $\mathcal{M}_1$ and $\mathcal{M}_2$ for $\mathcal{G}$.

- For all $S \in \breve{\mathbf{S}} \backslash \{S^*\}$, we define

$$P^{\mathcal{M}_i}(S|Pa_{\mathcal{G}}(S)) := P(S|Pa_{\mathcal{G}^*}(S)), \quad i \in \{1,2\}.$$

- For $S = S^*$, we define

$$P^{\mathcal{M}_1}(S|Pa_{\mathcal{G}}(S)) := P(S|Pa_{\mathcal{G}}(S), T_0^* = 1, U_0^* = 0),$$
$$P^{\mathcal{M}_2}(S|Pa_{\mathcal{G}}(S)) := P(S|Pa_{\mathcal{G}}(S), T_0^* = 1, U_0^* = 2).$$

Suppose that $\gamma_{r_1} = 0$ and $\gamma_{r_2} = 2$, then

- In model $\mathcal{M}_1$:

$$Q[\mathbf{A}_i]^{\mathcal{M}_1}(\mathbf{v}) = \theta_{r_1,j}(\mathbf{v}, T_0^* = 1), \quad i \in [0,m],$$
$$\sum_{\breve{\mathbf{S}}'} Q^{\mathcal{M}_1}[\breve{\mathbf{S}}](\mathbf{v}) = \phi_{r_1}(\mathbf{v}[\breve{\mathbf{S}}^\dagger]),$$
$$Q^{\mathcal{M}_1}[\breve{\mathbf{S}}](\mathbf{v}) = \eta_{r_1}(\mathbf{v}).$$

- In model $\mathcal{M}_2$:

$$Q[\mathbf{A}_i]^{\mathcal{M}_2}(\mathbf{v}) = \theta_{r_2,j}(\mathbf{v}, T_0^* = 1), \quad i \in [0,m],$$
$$\sum_{\breve{\mathbf{S}}'} Q^{\mathcal{M}_2}[\breve{\mathbf{S}}](\mathbf{v}) = \phi_{r_2}(\mathbf{v}[\breve{\mathbf{S}}^\dagger]),$$
$$Q^{\mathcal{M}_2}[\breve{\mathbf{S}}](\mathbf{v}) = \eta_{r_2}(\mathbf{v}).$$

According to the Lemmas 1 and 4 for any $\mathbf{v} \in \mathfrak{X}(\mathbf{V})$

$$Q^{\mathcal{M}_1}[\mathbf{A}_i](\mathbf{v}) = Q^{\mathcal{M}_2}[\mathbf{A}_i](\mathbf{v}),$$
$$\sum_{\breve{\mathbf{S}}'} Q^{\mathcal{M}_1}[\breve{\mathbf{S}}](\mathbf{v}') = \sum_{\breve{\mathbf{S}}'} Q^{\mathcal{M}_2}[\breve{\mathbf{S}}](\mathbf{v}),$$

however, using Lemma 2 and for $\mathbf{v}_0 = (0, ..., 0)$, we get

$$Q^{\mathcal{M}_1}[\breve{\mathbf{S}}](\mathbf{v}_0) \neq Q^{\mathcal{M}_2}[\breve{\mathbf{S}}](\mathbf{v}_0).$$

□

1.3 PROOF OF THE PROPERTIES IN SECTION 4.2.2 & SECTION 4.2.3

Recall that in Sections 4.2.2 and 4.2.3, we present two sets of properties which we prove them here. We only present the formal proof of the set of properties in Sections 4.2.2 since the other set of properties in Section 4.2.3 can be shown similarly.

1. If path p contains a chain $W' \to W \to W''$ or a fork $W' \leftarrow W \to W''$, then node W does not belong to any of the sets $\mathbf{X}'$, $\mathbf{Z}'$ or $\mathbf{Y}'$.
2. If path p contains a collider $W' \to W \leftarrow W''$, then there is a directed path p_W from W to a node in $\mathbf{Z}'$. Moreover, none of the intermediate nodes in the path p_W belong to the set $\mathbf{X}' \cup \mathbf{Z}' \cup \mathbf{Y}'$.
3. Path p does not contain any node from the set $\mathbf{X}'$.

Proof.
1. The first property is obvious since path p is not blocked by the set $\mathbf{X}' \cup (\mathbf{Z}' \backslash \{Z'\}) \cup (\mathbf{Z}' \backslash \{Y'\})$.

2. Suppose W is a collider as defined and let assume that R is the closest descendant of the variable W that unblocks path p. Note that $R \notin \mathbf{X}'$ since it unblocks p in the graph $\mathcal{G}_{\overline{\mathbf{X}'}, \underline{\{Z'\}}}$, i.e. no incoming edges in $\mathbf{X}'$.

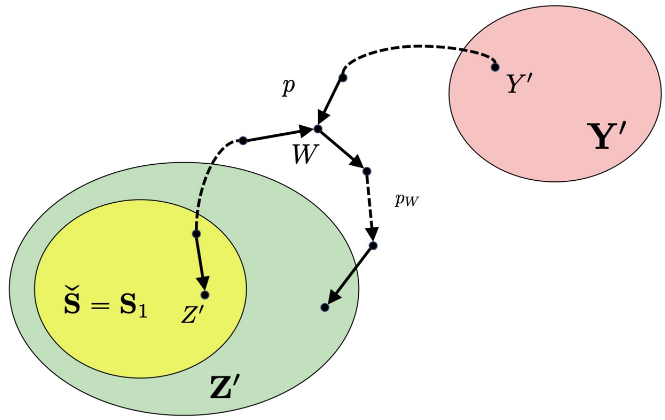

Figure 2: An illustration of the path p, collider W and its corresponding path p_W.

All variables except R in the shortest directed path from W to R do not belong to the set $\mathbf{X}' \cup \mathbf{Y}' \cup \mathbf{Z}'$. Assume that $R \in \mathbf{Y}'$ and p' is a path obtained by combining two paths: one from Z' to W in p and the other one from W to R (defined above). It is easy to see that p' is also unblocked, but it contains less number of colliders than p. This is impossible according to the definition of the path p. Thus, R must be in the set $\mathbf{Z}'$. This concludes the proof of the second property.

3. We prove this by contradiction. Suppose that there is a variable $R \in \mathbf{X}'$ on the path p. Since p is unblocked, then X is a collider or a descendant of a collider. This is impossible due to property 2. □

1.4 PROOF OF LEMMA 4

Recall that $\mathbf{S} = Anc_{\mathcal{G}[\breve{\mathbf{V}} \setminus \mathbf{X}']}(\mathbf{Y}' \cup \mathbf{Z}')$ and it is assumed that is not gID from $(\mathbb{A}, \mathcal{G})$. $\mathbf{S}$ consists of $\mathbf{S}_1, ..., \mathbf{S}_n$ as its single c-components where $\mathbf{S}_1$ is not gID. Let $\breve{\mathbf{S}} = \mathbf{S}_1$. Clearly, we can add $\{\mathbf{S}_i\}_{i=2}^n$ to the known distributions and $\breve{\mathbf{S}}$ remains not gID, i.e., $\breve{\mathbf{S}}$ is not gID from $(\mathbb{A}', \mathcal{G})$, where $\mathbb{A}' := \mathbb{A} \cup \{\mathbf{S}_i\}_{i=2}^n$. For simplicity, we denote $\mathbb{A}' = \{\mathbf{A}'\}_{i=0}^{m'}$. Hence, using the method in Section 1.1.1, we can construct two models $\mathcal{M}_1$ and $\mathcal{M}_2$ that are the same over the known distributions and different over $Q[\mathbf{S}_1]$. These models disagree on the distribution $Q[\mathbf{S}]$ as well, because $Q[\mathbf{S}] = \prod_{i=1}^n Q[\mathbf{S}_i]$. Below, we use these two models to introduce two new models to prove Lemma 4.

1.4.1 New models for Lemma 4

Recall that $\mathcal{P}$ is a collection of paths $\{p\} \cup \{p_W | W \in \mathbf{F}\}$, where $\mathbf{F}$ is a set of all colliders on the path p. Moreover, $\mathbf{D}$ is a set of all observed nodes on the paths in $\mathcal{P}$ excluding the ones in $\mathbf{Z}'$. Figure 2 demonstrates some variables used in this proof and their relations for clarity.

Herein, we define new models $\mathcal{M}_1'$ and $\mathcal{M}_2'$ based on the models $\mathcal{M}_1$ and $\mathcal{M}_2$. Let $\mathbf{D}_{\mathcal{P}}$ be the set of all variables (observed and unobserved) on the paths in $\mathcal{P}$. We say that a variable D is a **starting node** of path $\hat{p} \in \mathcal{P}$ if

- $D = Z'$ and $\hat{p} = p$ or
- $D \in \mathbf{F}$, i.e., it is a collider on path p and $\hat{p} = p_D$.

Note that D can be a starting node of only one path. According to the definition of a starting node, if D is a starting node for some path then either D is a collider on the path p or D is Z'.

For $R \in \mathbf{V} \cup \mathbf{U}$, let $\alpha_p(R)$ be the number of paths in $\mathcal{P}$ that contains R. Furthermore, we use $\mathfrak{X}(R)'$ and $\mathfrak{X}(R)$ to denote its domain in $\mathcal{M}_1'$ or $\mathcal{M}_2'$ (variables in different models have the same domains) and in $\mathcal{M}_1$ or $\mathcal{M}_2$, respectively. We define $\mathfrak{X}(R)'$ as follows:

- If R is a starting node for a path in $\mathcal{P}$:

$$\mathfrak{X}(R)' := \mathfrak{X}(R) \times [0 : \kappa]^{\alpha_p(R)-1}.$$

- If R is not a starting node for any of the paths in $\mathcal{P}$, then:

$$\mathfrak{X}(R)^{'} := \mathfrak{X}(R) \times [0:\kappa]^{\alpha_p(R)}.$$

Consequently, if R does not belong to any of the paths in $\mathcal{P}$, then $\mathfrak{X}(R)^{'} = \mathfrak{X}(R)$.

Consider $R \in \mathbf{V} \cup \mathbf{U}$. According to the domains definitions above, R is a vector that is a concatenation of the vector coming from $\mathfrak{X}(R)$ in model $\mathcal{M}_1$ (or $\mathcal{M}_2$) and some additional coordinates. These additional coordinates are defined based on $\alpha_p(R)$. More precisely, if R is not a starting node of a path $\hat{p} \in \mathcal{P}$, then there is a coordinate assigned to this path, denoted by $R[\hat{p}]$, otherwise, if R is a starting node of $\hat{p} \in \mathcal{P}$, then there is no coordinate assigned this path.

Let $\mathbf{O} \subseteq \mathbf{V} \cup \mathbf{U}$. For any realization $\mathbf{o} \in \mathfrak{X}(\mathbf{O})^{'}$ of $\mathbf{O}$, we denote by $\mathbf{o}^{\mathcal{M}} \in \mathfrak{X}(\mathbf{O})$, a realization of $\mathbf{O}$ that is consistent with $\mathbf{o}$. With a slight abuse of notation, we use $\mathbf{O}$ and $\mathbf{O}^{\mathcal{M}}$ to denote realizations of $\mathbf{O}$ in models $\mathcal{M}_i'$ and $\mathcal{M}_i$, respectively. $\mathbf{O}^{\mathcal{M}}$ means realizations in $\mathfrak{X}(\mathbf{O})$ from model $\mathcal{M}_i$ that are consistent with realizations in $\mathfrak{X}(\mathbf{O})^{'}$ from model $\mathcal{M}_i'$.

Recall that $\mathbf{D}_{\mathcal{P}}$ is a set of all variables on the paths in $\mathcal{P}$. Let $D \in \mathbf{D}_{\mathcal{P}}$. We denote by $\mathcal{P}_D$, the set of all paths $\hat{p}$, such that $\hat{p} \in \mathcal{P}$, D belongs to path $\hat{p}$, and D is not a starting node of path $\hat{p}$. We are now ready to define the probabilities of $P^{\mathcal{M}_i'}(D|Pa_{\mathcal{G}}(D))$ for any $D \in \mathbf{V} \cup \mathbf{U}$ and $i \in \{1,2\}$.

- If D does not belong to the set $\mathbf{D}_{\mathcal{P}}$, we define

$$P^{\mathcal{M}_i'}(D|Pa_{\mathcal{G}}(D)) := P^{\mathcal{M}_i}(D^{\mathcal{M}}|Pa_{\mathcal{G}}(D)).$$

- If D belongs to the set $\mathbf{D}_{\mathcal{P}} \backslash \{Z'\}$, we define

$$P^{\mathcal{M}_i'}(D|Pa_{\mathcal{G}}(D)) := P^{\mathcal{M}_i}(D^{\mathcal{M}}|Pa_{\mathcal{G}}(D)) \prod_{\hat{p} \in \mathcal{P}_D} f_{\hat{p}}(D[\hat{p}]|Pa_{\hat{p}}(D)),$$

  where $Pa_{\hat{p}}(D)$ denotes the parents of D on path $\hat{p}$ and $f_{\hat{p}}(D|Pa_{\hat{p}}(D))$ is given below.

  Definition of function $f_{\hat{p}}(D[\hat{p}]|Pa_{\hat{p}}(D))$:

  - When there exists a variable $W \in \mathbf{F}$, such that $\hat{p} = p_W$ and D is a child of W on path p_W (i.e., $W \in Pa_{p_W}(D)$), then we define

$$f_{p_W}(D[p_W]|Pa_{p_W}(D)) := \begin{cases} 1-\kappa\epsilon & \text{if } D[p_W] \equiv W[p] \pmod{\kappa+1}, \\ \epsilon & \text{if } D[p_W] \not\equiv W[p] \pmod{\kappa+1}. \end{cases}$$

  - When $Pa_{\hat{p}}(D) = \varnothing$,

$$f_{\hat{p}}(D[\hat{p}]) := \frac{1}{\kappa+1}.$$

  - Otherwise,

$$f_{\hat{p}}(D[\hat{p}]|Pa_{\hat{p}}(D)) := \begin{cases} 1-\kappa\epsilon & \text{if } D[\hat{p}] \equiv \sum_{D' \in Pa_{\hat{p}}(D)} D'[\hat{p}] \pmod{\kappa+1} \\ \epsilon & \text{if } D[\hat{p}] \not\equiv \sum_{D' \in Pa_{\hat{p}}(D)} D'[\hat{p}] \pmod{\kappa+1}, \end{cases} \tag{37}$$

    Note that $P^{\mathcal{M}_i'}(D|Pa_{\mathcal{G}}(D))$ is a probability distribution since for different paths $\hat{p}_1$ and $\hat{p}_2$, $D[\hat{p}_1]$ and $D[\hat{p}_2]$ are different and also

$$\sum_{D[\hat{p}] \in \mathfrak{X}(D[\hat{p}])} f_{\hat{p}}(D[\hat{p}]|Pa_{\hat{p}}(D)) = 1$$

- If $D = Z'$ and W is a parent of Z' in path p. Note that such W exists because p is an unblocked backdoor path in graph $\mathcal{G}_{\overline{\mathbf{X}'},\underline{\{Z'\}}}$. Recall that Z' is a variable from the set $\check{\mathbf{S}}$. In this case, we define

$$P^{\mathcal{M}_i'}(Z'|Pa_{\mathcal{G}}(Z')) := P'(Z'^{\mathcal{M}}|Pa_{\mathcal{G}}(Z')) \prod_{\hat{p} \in \mathcal{P}_{Z'}} f_{\hat{p}}(Z'[\hat{p}]|Pa_{\hat{p}}(Z')), \tag{38}$$

  where $P'(\cdot|\cdot)$ is given by

$$P'(Z'^{\mathcal{M}} = z' \mid Pa_{\mathcal{G}}(Z')) := \begin{cases} \frac{1}{\kappa+1} & \text{if } \mathbb{I}(Z') = 1, \\ 1-\kappa\epsilon & \text{if } \mathbb{I}(Z') = 0 \text{ and } z' \equiv M'(Z') \pmod{\kappa+1}, \\ \epsilon & \text{if } \mathbb{I}(Z') = 0 \text{ and } z' \not\equiv M'(Z') \pmod{\kappa+1}, \end{cases}$$

and $M'(\cdot)$ is defined similar to (9) and is given by

$$M'(Z') := \begin{cases} W[p] + \sum_{x \in Pa_{\mathcal{G}'[\breve{\mathbf{S}}]}(Z')} x^{\mathcal{M}} & \text{, if } Z' \in \breve{\mathbf{S}} \backslash \{S_0\}, \\ W[p] + u_0^{\mathcal{M}}[0] + \sum_{x \in Pa_{\mathcal{G}'[\breve{\mathbf{S}}]}(Z')} x^{\mathcal{M}} & \text{, if } Z' = S_0 \,. \end{cases} \tag{39}$$

Note that for any $W \in (\mathbf{V} \cup \mathbf{U}) \backslash \{U_0\}$, we have

$$P^{\mathcal{M}'_1}(W|Pa_{\mathcal{G}}(W)) = P^{\mathcal{M}'_2}(W|Pa_{\mathcal{G}}(W)).$$

Therefore, we will use $P^{\mathcal{M}'}(W|Pa_{\mathcal{G}}(W))$ instead of $P^{\mathcal{M}'_1}(W|Pa_{\mathcal{G}}(W))$ or $P^{\mathcal{M}'_2}(W|Pa_{\mathcal{G}}(W))$ for $W \in (\mathbf{V} \cup \mathbf{U}) \backslash \{U_0\}$.

We also have

$$\begin{aligned} P^{\mathcal{M}'_1}(U_0) &= \frac{1}{d} \prod_{\hat{p} \in \mathcal{P}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]), \\ P^{\mathcal{M}'_2}(U_0) &= P^{\mathcal{M}_2}(U_0^{\mathcal{M}}) \prod_{\hat{p} \in \mathcal{P}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]). \end{aligned} \tag{40}$$

Recall that $\mathbf{S} = Anc_{\mathcal{G}[\mathbf{V} \backslash \mathbf{X}']}(\mathbf{Y}', Z')$. Let $\mathbf{D}' := \mathbf{S} \backslash \mathbf{D}$ and $\mathbf{D}^{\dagger} := \mathbf{V} \backslash \mathbf{D}$. For $i \in [0 : m']$, $j \in [1 : d]$, $\mathbf{v} \in \mathfrak{X}(\mathbf{V})'$ and $\mathbf{d}^{\dagger} \in \mathfrak{X}(\mathbf{D}^{\dagger})'$, we define $\theta'_{i,j}(\mathbf{v})$, $\phi'_j(\mathbf{d}^{\dagger})$ and $\eta'_j(\mathbf{v})$ as follows:

$$\theta'_{i,j}(\mathbf{v}) := \sum_{U_0[\mathcal{P}]} \prod_{\hat{p} \in \mathcal{P}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \sum_{\mathbf{U} \backslash \{U_0\}} \prod_{X \in \mathbf{A}'_i} P^{\mathcal{M}'}(x \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U} \backslash \{U_0\}} P^{\mathcal{M}'}(u^{\mathcal{M}}), \tag{41}$$

$$\phi'_j(\mathbf{d}^{\dagger}) := \sum_{U_0[\mathcal{P}]} \prod_{\hat{p} \in \mathcal{P}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \sum_{\mathbf{D}} \sum_{\mathbf{U} \backslash \{U_0\}} \prod_{X \in \mathbf{S}} P^{\mathcal{M}'}(x \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U} \backslash \{U_0\}} P^{\mathcal{M}'}(u^{\mathcal{M}}), \tag{42}$$

$$\eta'_j(\mathbf{v}) := \sum_{U_0[\mathcal{P}]} \prod_{\hat{p} \in \mathcal{P}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \sum_{\mathbf{U} \backslash \{U_0\}} \prod_{X \in \mathbf{S}} P^{\mathcal{M}'}(x \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U} \backslash \{U_0\}} P^{\mathcal{M}'}(u^{\mathcal{M}}), \tag{43}$$

where $\sum_{U_0[\mathcal{P}]}$ is a summation over all realizations of the random variables $\{U_0[\hat{p}] | \, \hat{p} \in \mathcal{P}_{U_0}\}$.

Next, we prove three lemmas similar to Lemmas 1, 2, and 4 for the new models $\mathcal{M}'_1$ and $\mathcal{M}'_2$.

Lemma 6. *For any $\mathbf{v} \in \mathfrak{X}(\mathbf{V})'$ and $i \in [0 : m']$, we have*

$$\theta'_{i,j_1}(\mathbf{v}) = \theta'_{i,j_2}(\mathbf{v}) = \cdots = \theta'_{i,j_{\frac{\kappa+1}{2}}}(\mathbf{v}).$$

Proof. By substituting $P^{\mathcal{M}'}$ from the above into Equation (41) and rearranging the terms, we obtain

$$\begin{aligned} \theta'_{i,j}(\mathbf{v}) = &\sum_{\mathbf{U}[\mathcal{P}]} \prod_{\hat{U}[\hat{p}] \in \mathbf{U}[\mathcal{P}]} f_{\hat{p}}(\hat{U}) \prod_{\hat{X}[\hat{p}] \in \mathbf{A}'_i[\mathcal{P}]} f_{\hat{p}}(\hat{X}[\hat{p}] | Pa_{\hat{p}}(\hat{X})) \\ &\times \Big(\sum_{\underline{\mathbf{U}}^{\mathcal{M}}} P'(Z'|Pa_{\mathcal{G}}(Z')) \prod_{X \in \mathbf{A}'_i \backslash \{Z'\}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U} \backslash \{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}}) \Big), \end{aligned}$$

where $\mathbf{U}[\mathcal{P}] := \bigcup_{U \in \mathbf{U}} \{U[\hat{p}] | \, \hat{p} \in \mathcal{P}_U\}$, $\underline{\mathbf{U}} := \mathbf{U} \backslash \{U_0\}$, and by definition $\underline{\mathbf{U}}^{\mathcal{M}}$ is all realizations of elements in set $\underline{\mathbf{U}}$ in $\mathfrak{X}(\underline{\mathbf{U}})$ that are consistent with realizations in $\mathfrak{X}(\underline{\mathbf{U}})'$. Suppose variable W belongs to the path p and Z' is a child of W in that path. By the construction of $P(Z'|Pa_{\mathcal{G}}(Z'))$, we have

$$P'(Z' + W[p] | Pa_{\mathcal{G}}(Z')) = P^{\mathcal{M}}(Z'|Pa_{\mathcal{G}}(Z')). \tag{44}$$

This is because $M'(Z' + W[p]) = M(Z')$. Let $\mathbf{v}' \in \mathfrak{X}(\mathbf{V})$ be a realization that is consistent with $\mathbf{v}^{\mathcal{M}}[\mathbf{V} \backslash \{Z'\}]$ and

$$\mathbf{v}'[Z'] = \mathbf{v}^{\mathcal{M}}[Z'] - \mathbf{v}[W[p]].$$

In this case, using (44), we have

$$\begin{aligned} \theta'_{i,j}(\mathbf{v}) = &\sum_{\mathbf{U}[\mathcal{P}]} \prod_{\hat{U}[\hat{p}] \in \mathbf{U}[\mathcal{P}]} f_{\hat{p}}(\hat{U}) \prod_{\hat{X}[\hat{p}] \in \mathbf{A}'_i[\mathcal{P}]} f_{\hat{p}}(\hat{X}[\hat{p}] | Pa_{\hat{p}}(\hat{X})) \\ &\times \Big(\sum_{\underline{\mathbf{U}}^{\mathcal{M}}} P^{\mathcal{M}}(\mathbf{v}'^{\mathcal{M}}[Z'] | Pa_{\mathcal{G}}(Z')) \prod_{X \in \mathbf{A}'_i \backslash \{Z'\}} P^{\mathcal{M}}(x \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U} \backslash \{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}}) \Big). \end{aligned}$$

Note that the terms inside the big parenthesis is equal to $\theta_{i,j}(\mathbf{v}')$ given in (3), i.e.,

$$\theta'_{i,j}(\mathbf{v}) = \sum_{\mathbf{U}[\mathcal{P}]} \prod_{\hat{U}[\hat{p}]\in\mathbf{U}[\mathcal{P}]} f_{\hat{p}}(\hat{U}) \prod_{\hat{X}[\hat{p}]\in\mathbf{A}'_i[\mathcal{P}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X}))\theta_{i,j}(\mathbf{v}').$$

In the last equation, all terms on the right hand side except $\theta_{i,j}(\mathbf{v}')$ are independent of the realization of $\{U_0\}^{\mathcal{M}}$, i.e., independent of index j. For $j \in \{j_1, ..., j_{\frac{\kappa+1}{2}}\}$ and using the result of Lemma 1 that says $\theta_{i,j_1}(\mathbf{v}') = ... = \theta_{i,j_{\frac{\kappa+1}{2}}}(\mathbf{v}')$, we can conclude the result. □

Lemma 7. *For any $\mathbf{d}^\dagger \in \mathfrak{X}(\mathbf{D}^\dagger)'$, we have*

$$\phi_{j_1}(\mathbf{d}^\dagger) = \phi_{j_2}(\mathbf{d}^\dagger) = \cdots = \phi_{j_{\frac{\kappa+1}{2}}}(\mathbf{d}^\dagger).$$

Proof. Similar to the previous lemma, by substituting $P^{\mathcal{M}'}$ from their definitions into Equation (42) and rearranging the terms, we obtain

$$\phi'_j(\mathbf{d}^\dagger) := \sum_{\mathbf{U}[\mathcal{P}]} \sum_{\underline{\mathbf{U}}^{\mathcal{M}}} \sum_{\mathbf{D}} \prod_{\hat{p}\in\mathcal{P}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \prod_{\hat{X}[\hat{p}]\in\mathbf{S}[\mathcal{P}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X})) \\ \times\Big(P'(Z'|Pa_{\mathcal{G}}(Z')) \prod_{X\in\mathbf{S}\setminus\{Z'\}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U\in\mathbf{U}\setminus\{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}})\Big), \tag{45}$$

where $\mathbf{U}[\mathcal{P}] := \bigcup_{U\in\mathbf{U}}\{U[\hat{p}]|\, \hat{p} \in \mathcal{P}_U\}$, $\underline{\mathbf{U}} := \mathbf{U}\setminus\{U_0\}$. Suppose that l_1 and l_2 are two integers such that

$$\gamma_{l_1} := (2x, 0, \ldots, 0)), \\ \gamma_{l_2} := (2x + 2 \pmod{\kappa + 1}, 0, \ldots, 0),$$

and x is an integer in $[0 : \frac{\kappa-1}{2}]$. We will prove that $\phi_{l_1}(\mathbf{d}^\dagger) = \phi_{l_2}(\mathbf{d}^\dagger)$.

Suppose that path p is the sequence of variables: Z', D'_1, $D'_2 \ldots D'_{k'_1}$, $D'_{k'_1+1} := Y'$. Note that there is a direct edge between any consecutive nodes in this path and furthermore, the direct edge between Z' and D'_1 is pointing toward Z', i.e., $Z' \leftarrow D'_1$.

On the other hand, since Z' and U_0 are both in $\breve{\mathbf{S}}$ ($\breve{\mathbf{S}} = \mathbf{S}_1$ by construction), then there exists a shortest path $U_0, \hat{S}'_1, \hat{U}'_1, \hat{S}'_2, \hat{U}'_2, \ldots, \hat{U}'_{l'}, Z'$, such that U_0 is a parent of $\hat{S}'_1 \in \breve{\mathbf{S}}$, Z' is a child of $\hat{U}'_l \in \breve{\mathbf{U}}^{\breve{\mathbf{S}}}$, and $\hat{U}'_j \in \breve{\mathbf{U}}^{\breve{\mathbf{S}}}$ is a parent of variables $\hat{S}'_j \in \breve{\mathbf{S}}$ and $\hat{S}'_{j+1} \in \breve{\mathbf{S}}$ for $j \in [1 : l' - 1]$. Let $\hat{\mathbf{U}}' := \{\hat{U}'_1, \ldots, \hat{U}'_l\}$, i.e., unobserved nodes in this shortest path except U_0. For a given realization $\mathbf{o}_1$ of $\mathbf{U} \cup \mathbf{D}$, we define $\mathbf{o}_2 \in \mathfrak{X}(\mathbf{U} \cup \mathbf{D})'$ as follows

$$\mathbf{o}_2^{\mathcal{M}}[\hat{U}'_j] := \mathbf{o}_1^{\mathcal{M}}[\hat{U}'_j] + 2(-1)^j \pmod{\kappa + 1}, \quad j \in [1 : l'], \tag{46}$$

For D'_1, we have

$$\mathbf{o}_2[D'_1[p]] = \mathbf{o}_1[D'_1[p]] - 2(-1)^{l'} \pmod{\kappa + 1}. \tag{47}$$

Note that with these modifications, for any $\widetilde{S} \in \breve{\mathbf{S}}\setminus\{Z'\}$, we have

$$\widetilde{s}_2 - M(\widetilde{S}) \equiv \widetilde{s}_1 - M(\widetilde{S}) \pmod{\kappa + 1},$$

where $\widetilde{s}_1$ is a realization of $\widetilde{S}\Big|_{(\mathbf{U}\cup\mathbf{D},\mathbf{D}^\dagger,U_0^{\mathcal{M}})=(\mathbf{o}_1,\mathbf{d}^\dagger,\gamma_{l_1})}$, $\widetilde{s}_2$ is a realization of $\widetilde{S}\Big|_{(\mathbf{U}\cup\mathbf{D},\mathbf{D}^\dagger,U_0^{\mathcal{M}})=(\mathbf{o_2},\mathbf{d}^\dagger,\gamma_{l_2})}$, and $M(\cdot)$ is given by Equation (9). Additionally,

$$\mathbf{o}_2^{\mathcal{M}}[Z'] - M'(Z') \equiv \mathbf{o}_1^{\mathcal{M}}[Z'] - M'(Z') \pmod{\kappa + 1},$$

where $M'(\cdot)$ is defined in Equation (39). This implies that for any $\widetilde{S} \in \mathbf{S}$, we have

$$P^{\mathcal{M}}(\widetilde{s}|Pa_{\mathcal{G}}(\widetilde{S}))\Big|_{(\mathbf{U}\cup\mathbf{D},\mathbf{D}^\dagger,U_0^{\mathcal{M}})=(\mathbf{o}_1,\mathbf{d}^\dagger,\gamma_{l_1})} = P^{\mathcal{M}}(\widetilde{s}|Pa_{\mathcal{G}}(\widetilde{S}))\Big|_{(\mathbf{U}\cup\mathbf{D},\mathbf{D}^\dagger,U_0^{\mathcal{M}})=(\mathbf{o_2},\mathbf{d}^\dagger,\gamma_{l_2})}.$$

Let $c := -2(-1)^{l'}$, then Equation (47) becomes

$$\mathbf{o}_2^{\mathcal{M}}[D'_1[p]] = \mathbf{o}_1^{\mathcal{M}}[D'_1[p]] + c \pmod{\kappa + 1}. \tag{48}$$

Suppose that D'_j is not a collider on the path p and $j \in [2 : k'_1 + 1]$. We define $\mu(D'_j)$ to be the number of colliders on a part of the path p from D'_1 to D'_{j-1}. Thus, for those $j \in [2 : k'_1 + 1]$ that D'_j is not a collider, we define

$$\mathbf{o}_2^{\mathcal{M}}[D'_j[p]] := \mathbf{o}_1^{\mathcal{M}}[D'_j[p]] + c(-1)^{\mu(D'_j)}. \tag{49}$$

Note that the modifications in (49) might only affect the function $f_p(\cdot|\cdot)$. Next, we show that after these modifications, function $f_p(\cdot|\cdot)$ remains unchanged. To do so, for $j \in [1 : k'_1 + 1]$, we consider four different cases:

1. If D'_j has no parents, then it is obvious that

$$f_p(D'_j[p])\Big|_{(\mathbf{U}\cup\mathbf{D},U_0^{\mathcal{M}})=(\mathbf{o}_1,\gamma_{l_1})} = f_p(D'_j[p])\Big|_{(\mathbf{U}\cup\mathbf{D},U_0^{\mathcal{M}})=(\mathbf{o}_2,\gamma_{l_2})}.$$

2. D'_j is a collider, then $\mu(D'_{j+1}) = \mu(D'_{j-1}) + 1$ and

$$\mathbf{o}_1[D'_{j+1}[p]] + \mathbf{o}_1[D'_{j-1}[p]] = \mathbf{o}_2[D'_{j+1}[p]] + \mathbf{o}_2[D'_{j-1}[p]],$$

   and hence, according to Equation (37), we have

$$f_p(D'_j[p]|Pa_p(D'_j))\Big|_{(\mathbf{U}\cup\mathbf{D},U_0^{\mathcal{M}})=(\mathbf{o}_1,\gamma_{l_1})} = f_p(D'_j[p]|Pa_p(D'_j))\Big|_{(\mathbf{U}\cup\mathbf{D},U_0^{\mathcal{M}})=(\mathbf{o}_2,\gamma_{l_2})}.$$

3. D'_j is a child of D'_{j+1}, then $\mu(D'_j) = \mu(D'_{j+1})$ and

$$\mathbf{o}_1[D'_j[p]] - \mathbf{o}_1[D'_{j+1}[p]] = \mathbf{o}_2[D'_j[p]] - \mathbf{o}_2[D'_{j+1}[p]].$$

   According to Equation (37), we imply that

$$f_p(D'_j[p]|Pa_p(D'_j))\Big|_{(\mathbf{U}\cup\mathbf{D},U_0^{\mathcal{M}})=(\mathbf{o}_1,\gamma_{l_1})} = f_p(D'_j[p]|Pa_p(D'_j))\Big|_{(\mathbf{U}\cup\mathbf{D},U_0^{\mathcal{M}})=(\mathbf{o}_2,\gamma_{l_2})}.$$

4. D'_j is a child of D'_{j-1}, then $\mu(D'_j) = \mu(D'_{j-1})$ and

$$\mathbf{o}_1[D'_j[p]] - \mathbf{o}_1[D'_{j-1}[p]] = \mathbf{o}_2[D'_j[p]] - \mathbf{o}_2[D'_{j-1}[p]].$$

   Similarly, according to Equation (37), we get

$$f_p(D'_j[p]|Pa_p(D'_j))\Big|_{(\mathbf{U}\cup\mathbf{S},U_0)=(\mathbf{o}_1,\gamma_{l_1})} = f_p(D'_j[p]|Pa_p(D'_j))\Big|_{(\mathbf{U}\cup\mathbf{S},U_0)=(\mathbf{o}_2,\gamma_{l_2})}.$$

This concludes that for any $j \in [1 : k'_1 + 1]$,

$$f_p(D'_j[p]|Pa_p(D'_j))\Big|_{(\mathbf{U}\cup\mathbf{D},U_0^{\mathcal{M}})=(\mathbf{o}_1,\gamma_{l_1})} = f_p(D'_j[p]|Pa_p(D'_j))\Big|_{(\mathbf{U}\cup\mathbf{D},U_0^{\mathcal{M}})=(\mathbf{o}_2,\gamma_{l_2})}.$$

Note that the aforementioned transformation of $\mathbf{o}_1$ affects only those realizations of variables that are used for the marginalization in the Equation (45). Putting the above results together implies that the terms in (45) remain unchanged, i.e.,

$$\begin{aligned}
&\prod_{\hat{p}\in\mathcal{P}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \prod_{\hat{X}[\hat{p}]\in\mathbf{S}[\mathcal{P}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X}))\\
&\times \Big(P'(Z'|Pa_{\mathcal{G}}(Z')) \prod_{X\in\mathbf{S}\backslash\{Z'\}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U\in\mathbf{U}\backslash\{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}})\Big)\Big|_{(\mathbf{U}\cup\mathbf{D},\mathbf{D}^\dagger,U_0^{\mathcal{M}})=(\mathbf{o}_1,\mathbf{d}^\dagger,\gamma_{l_1})}\\
&= \prod_{\hat{p}\in\mathcal{P}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \prod_{\hat{X}[\hat{p}]\in\mathbf{S}[\mathcal{P}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X}))\\
&\times \Big(P'(Z'|Pa_{\mathcal{G}}(Z')) \prod_{X\in\mathbf{S}\backslash\{Z'\}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U\in\mathbf{U}\backslash\{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}})\Big)\Big|_{(\mathbf{U}\cup\mathbf{D},\mathbf{D}^\dagger,U_0^{\mathcal{M}})=(\mathbf{o}_2,\mathbf{d}^\dagger,\gamma_{l_2})}
\end{aligned}$$

This implies that $\phi_{l_1}(\mathbf{d}^\dagger) = \phi_{l_2}(\mathbf{d}^\dagger)$. By varying x within $[0 : \frac{\kappa-1}{2}]$ in the definition of γ_{l_1} and γ_{l_2}, we obtain the result. $\square$

Lemma 8. *There exists $0 < \epsilon < \frac{1}{\kappa}$, such that there exists $\mathbf{v}_0 \in \mathfrak{X}(\mathbf{V})'$ and $1 \leqslant r < t \leqslant \frac{\kappa+1}{2}$ such that*

$$\eta'_{j_r}(\mathbf{v}_0) \neq \eta'_{j_t}(\mathbf{v}_0).$$

Proof. By substituting $P^{\mathcal{M}'}$ from their definitions into Equation (43) and rearranging the terms, we obtain

$$\begin{aligned} \eta_j^{'}(\mathbf{v}_0) = &\sum_{\mathbf{U}^{\mathcal{M}}\setminus\{U_0\}^{\mathcal{M}}} \sum_{\mathbf{U}[\mathcal{P}]} \prod_{\hat{U}[\hat{p}]\in\mathbf{U}[\mathcal{P}]} f_{\hat{p}}(\hat{U}) \prod_{\hat{X}[\hat{p}]\in\mathbf{S}[\mathcal{P}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X})) \\ &\times\Big(P'(Z'|Pa_{\mathcal{G}}(Z')) \prod_{X\in\mathbf{S}\setminus\{Z'\}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U\in\mathbf{U}\setminus\{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}})\Big), \end{aligned} \tag{50}$$

Next, we define $\mathbf{v}_0 \in \mathfrak{X}(\mathbf{V})'$ such that the conditions in the lemma hold.

- For any path $\hat{p} \in \mathcal{P}$ and any node W on the path $\hat{p}$ that is not a starting node for path $\hat{p}$, we define

$$\mathbf{v}_0[W[\hat{p}]] := 0.$$

- For any variable $S \in \check{\mathbf{S}}$, we define

$$\mathbf{v}_0^{\mathcal{M}}[S] := 0.$$

- For the remaining part of $\mathbf{v}_0$, we choose a realization such that for the selected $\mathbf{v}_0$, there exists a realization for the unobserved variables $\mathbf{U}$ that ensures $\mathbb{I}(S) = 0$ for all $S \in \check{\mathbf{S}}$. This is clearly possible due to the definition of $\mathbb{I}(S)$.

Assume r and t are such that $\gamma_{j_r} := (0, 0, \ldots, 0)$ and $\gamma_{j_t} := (2, 0, \ldots, 0)$. To finish the proof of the lemma, it is enough to show that $\eta'_{j_r}(\mathbf{v}_0)$ and $\eta'_{j_t}(\mathbf{v}_0)$ are two different polynomial functions of parameter ϵ. We prove that these two polynomials are different by showing that $\eta'_{j_r}(\mathbf{v}_0) \neq \eta'_{j_t}(\mathbf{v}_0)$ for $\epsilon = 0$.

We only need to consider the non-zero terms in Equation (50). From (50), we have

$$\begin{aligned} &\prod_{\hat{U}[\hat{p}]\in\mathbf{U}[\mathcal{P}]} f_{\hat{p}}(\hat{U}) \prod_{\hat{X}[\hat{p}]\in\mathbf{S}[\mathcal{P}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X})) \\ &\times\Big(P'(Z'|Pa_{\mathcal{G}}(Z')) \prod_{X\in\mathbf{S}\setminus\{Z'\}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U\in\mathbf{U}\setminus\{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}})\Big). \end{aligned} \tag{51}$$

Note that $f_{\hat{p}}(\hat{U}) = \frac{1}{\kappa+1}$ and $f_{\hat{p}}(\hat{X}|Pa_{\hat{p}'}(\hat{X}))$ is non-zero only

- when there exists a variable $W \in \mathbf{F}$ such that $\hat{p}' = p_W$, $\hat{X}$ is a child of W in path p_W, and

$$\hat{X}[\hat{p}'] \equiv W[p] \pmod{\kappa+1}.$$

- when the following holds

$$\hat{X}[\hat{p}'] \equiv \sum_{\hat{X}'\in Pa_{\hat{p}'}(\hat{X})\setminus\{W\}} \hat{X}'[\hat{p}'] \pmod{\kappa+1}.$$

Similarly, $P^{\mathcal{M}}(X|Pa_{\mathcal{G}}(X))$ is non-zero

- if $\mathbb{I}(X) = 1$ (i.e. $P^{\mathcal{M}}(X|Pa_{\mathcal{G}}(X)) = \frac{1}{\kappa+1}$), or
- if $X \equiv M(X) \pmod{\kappa+1}$ for $P^{\mathcal{M}}(X|Pa_{\mathcal{G}}(X))$,

$P'(Z'|Pa_{\mathcal{G}}(Z'))$ is non-zero

- if $\mathbb{I}(Z') = 1$ (i.e. $P^{\mathcal{M}}(Z'|Pa_{\mathcal{G}}(Z')) = \frac{1}{\kappa+1}$), or
- $Z' \equiv M'(Z') \pmod{\kappa+1}$ for $P'(Z'|Pa_{\mathcal{G}}(Z'))$.

Let fix a realization $\mathbf{u} \in \mathfrak{X}(\mathbf{U}\backslash\{U_0^{\mathcal{M}}\})'$. We consider two scenarios:

I) Assume that for this realization, there is a variable $S \in \breve{\mathbf{S}}$, such that $\mathbb{I}(S) = 1$ and S is the closest variable to U_0 considering only paths with bidirected edges in $\mathcal{G}'[\breve{\mathbf{S}}]$. The value of $S^{\mathcal{M}}$ does not depend on its parents because of $\mathbb{I}(S) = 1$ and Equation (8). Additionally in the graph $\mathcal{G}'[\breve{\mathbf{S}}]$, there exists a path $U_0, \hat{S}_1', \hat{U}_1', \hat{S}_2', \hat{U}_2', \dots, \hat{U}_{l'}', S$, such that U_0 is a parent of $\hat{S}_1' \in \breve{\mathbf{S}}$, S is a child of $\hat{U}_l' \in \breve{\mathbf{U}}^{\breve{\mathbf{S}}}$, and $\hat{U}_j' \in \breve{\mathbf{U}}^{\breve{\mathbf{S}}}$ is a parent of variables $\hat{S}_j' \in \breve{\mathbf{S}}$ and $\hat{S}_{j+1}' \in \breve{\mathbf{S}}$ for $j \in [1 : l' - 1]$. Let $\hat{\mathbf{U}}' := \{\hat{U}_1', \dots, \hat{U}_l'\}$. We define $\mathbf{u}' \in \mathfrak{X}(\mathbf{U}\backslash\{U_0^{\mathcal{M}}\})'$ that is consistent with $\mathbf{u}$ except the variables in $\hat{\mathbf{U}}'$. For these variables, we define

$$\mathbf{u}'^{\mathcal{M}}[\hat{U}_j] := \mathbf{u}^{\mathcal{M}}[\hat{U}_j] + 2(-1)^j \pmod{\kappa + 1}, \quad j \in [1 : l'], \tag{52}$$

With this modification for any $\widetilde{S} \in \mathbf{S}$, we have

$$P^{\mathcal{M}}(\widetilde{s}|Pa_{\mathcal{G}}(\widetilde{S}))\Big|_{(\mathbf{U})=(\mathbf{u},\gamma_{j_r})} = P(\widetilde{s}|Pa_{\mathcal{G}}(\widetilde{S}))\Big|_{(\mathbf{U})=(\mathbf{u}',\gamma_{j_t})}.$$

Therefore for all such realizations of $\mathbf{u}$, the summation of the following terms for both $\eta_{j_r}'(\mathbf{v}_0)$ and $\eta_{j_t}'(\mathbf{v}_0)$ will be the same,

$$\begin{aligned}&\prod_{\hat{U}[\hat{p}] \in \mathbf{U}[\mathcal{P}]} f_{\hat{p}}(\hat{U}) \prod_{\hat{X}[\hat{p}] \in \mathbf{S}[\mathcal{P}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X})) \\ &\times \Big(P'(Z'|Pa_{\mathcal{G}}(Z')) \prod_{X \in \mathbf{S}\backslash\{Z'\}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U}\backslash\{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}})\Big).\end{aligned} \tag{53}$$

II) Assume that for all $S \in \breve{\mathbf{S}}$, we have $\mathbb{I}(S) = 0$. We consider a realization $U_0^{\mathcal{M}} = \gamma_{j_r}$ and $\mathbf{u}$ such that:

- $\mathbf{u}[\breve{\mathbf{U}}^{\breve{\mathbf{S}}}] = \mathbf{0}$, and
- for all $U \in \mathbf{U}$ and any path $\hat{p} \in \mathcal{P}$ which contains U, $\mathbf{u}[U[\hat{p}]] = 0$.

We claim that for such $\mathbf{u}$,

$$\begin{aligned}&\prod_{\hat{U}[\hat{p}] \in \mathbf{U}[\mathcal{P}]} f_{\hat{p}}(\hat{U}) \prod_{\hat{X}[\hat{p}] \in \mathbf{S}[\mathcal{P}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X})) \\ &\times \Big(P'(Z'|Pa_{\mathcal{G}}(Z')) \prod_{X \in \mathbf{S}\backslash\{Z'\}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U}\backslash\{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}})\Big).\end{aligned}$$

is non-zero. To prove this claim, we consider four cases:

- assume that $\hat{p} \in \mathcal{P}$ and exists a variable W such that $\hat{p} = p_W$. Let $\hat{X}$ be a child of W in path p_W. From the definitions of $\mathbf{u}$ and $\mathbf{v}_0$, we get
$$\hat{X}[\hat{p}] \equiv W[\hat{p}] \pmod{\kappa + 1},$$
and therefore $f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X})) = 1$. The above holds because $\hat{X}[\hat{p}] = 0 = W[\hat{p}] \pmod{\kappa + 1}$.
- assume that $\hat{p} \in \mathcal{P}$ and $\hat{X}$ is a variable on this path such that it is neither a starting node on $\hat{p}$ nor a child of a starting node on path $\hat{p}$. Then, from the definitions of $\mathbf{v}_0$ and $\mathbf{u}$ we get
$$\hat{X}[\hat{p}] \equiv \sum_{\hat{X} \in Pa_{\hat{p}}(\hat{X})} \hat{X}[\hat{p}] \pmod{\kappa + 1},$$
and therefore $f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X})) = 1$. The above holds because all the variables in the above equation are zero.
- assume $X \in \breve{\mathbf{S}}\backslash\{Z'\}$. From the definitions of $\mathbf{v}_0$ and $\mathbf{u}$, we get
$$X^{\mathcal{M}} \equiv M(X) \pmod{\kappa + 1},$$
and therefore $P^{\mathcal{M}}(x^{\mathcal{M}}|Pa_{\mathcal{G}}(X)) = 1$. Again, the above holds because all the terms are zero.

- assume $X = Z'$, then
$$Z' \equiv M'(Z') \pmod{\kappa + 1},$$
and consequently $P'(Z'|Pa_{\mathcal{G}}(Z')) = 1$.

Now, we consider the case when $U_0^{\mathcal{M}} = \gamma_{j_t}$. Assume that $W \in \mathbf{F}$ and W' is the last descendant of W on the path p_W. From the properties which we proved in Section 1.3, we have $W' \in \mathbf{Z}'$ and by the definition of $\mathbf{v}_0$, we have $W'[p_w] = 0$. Assume W'' is a parent of W' on the path p_W. Note that $f_{p_W}(W'[p_W]|Pa_{p_W}(W')) \neq 0$ if and only if $W''[p_W] = 0$. Repeating the above reasoning for variables from W' to W, we conclude that $W[p]$ must be equal to 0, otherwise, there would be a term in Equation (53) that is zero and this contradicts with the fact that Equation (53) is non-zero.

Assume that $Z', W_1', W_2', \dots, W_{k'}', W_{k'+1}' := Y'$ are the nodes on the path p. Next, we prove by backward induction that $W_i' = 0$ for all $i \in [1 : k'+1]$. By definition of $\mathbf{v_0}$, we know that $Y'[p] = 0$. If $W_i'[p] = 0$ for all $i \in [k''+1 : k']$, we will prove that $W_{k''}''[p] = 0$ as well. To do so, we consider the following three cases:

- $W_{k''}'$ is a collider on a path p. Then the fact that $W_{k''}''[p] = 0$ follows immediately from the aforementioned reasoning and the fact that Equation (53) is non-zero.
- $W_{k''}'$ is a child of $W_{k''+1}'$ and it is not a collider. Then, $f_p(W_{k''}'|Pa_p(W_{k''}'')) \neq 0$ if and only if $W_{k''}'[p] = W_{k''+1}'[p] = 0$.
- $W_{k''}'$ is a parent of $W_{k''+1}'$. Then, $f_p(W_{k''+1}'|Pa_p(W_{k''}'')) \neq 0$ if and only if $0 = W_{k''+1}'[p] = W_{k''}'[p]$.

This implies that $W_1' = 0$. Therefore, $P'(Z'|Pa_{\mathcal{G}}(Z')) = P^{\mathcal{M}}(Z'|Pa_{\mathcal{G}}(Z'))$ because $M'(Z') = M(Z')$. Furthermore, the above arguments show that all $f_{\hat{p}}(\cdot|\cdot)$ terms in Equation (53) are equal to one. This simplifies the Equation (53) to

$$\prod_{X \in \mathbf{S}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U} \setminus \{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}}).$$

However by the proof of Lemma 6 Kivva et al. [2022], we know that there is no realization of $\mathbf{U}^{\mathcal{M}}$ consistent with $U_0 = \gamma_{j_t}$ such that:

- $\mathbb{I}(S) = 0$ for all $S \in \check{\mathbf{S}}$, and
- $x^{\mathcal{M}} \equiv M(X) \pmod{\kappa+1}$ for all $X \in \check{\mathbf{S}}$. The latter is a necessary condition for $P^{\mathcal{M}}(x|Pa_{\mathcal{G}}(X))$ being non-zero.

To summarize, we showed that for $U_0^{\mathcal{M}} = \gamma_{j_r}$, Equation (53) is non-zero while it is zero for $U_0^{\mathcal{M}} = \gamma_{j_t}$. This implies that $\eta_{j_r}'(\mathbf{v}_0) \neq \eta_{j_t}'(\mathbf{v}_0)$ for $\epsilon = 0$. □

1.4.2 Proof of Lemma 4

Lemma 4. *Let* $\mathbf{S} := Anc_{\mathcal{G}[\mathbf{V} \setminus \mathbf{X}']}(\mathbf{Y}', \mathbf{Z}')$ *and* $\mathbf{D}$ *is a set of all nodes on the paths in* $\mathcal{P}$ *excluding* $\mathbf{Z}'$. *Then,*

$$P_{\mathbf{x}'}(\mathbf{d}|\mathbf{s} \setminus \mathbf{d}) = \frac{Q[\mathbf{S}]}{\sum_{\mathbf{D}} Q[\mathbf{S}]} = Q[\mathbf{D}|\mathbf{S} \setminus \mathbf{D}] \tag{54}$$

is not c-gID from $(\mathbb{A}, \mathcal{G})$.

Proof. We will show that $Q[\mathbf{D}|\mathbf{S} \setminus \mathbf{D}]$ is not c-gID from $(\mathbb{A}', \mathcal{G})$, where $\mathbb{A}' := \mathbb{A} \cup \{\mathbf{S}_i\}_{i=2}^n$. To this end, we will construct two models $\mathcal{M}_1$ and $\mathcal{M}_2$ such that for each $i \in [0 : m']$ and any $\mathbf{v} \in \mathbf{V}$:

$$Q^{\mathcal{M}_1}[\mathbf{A}_i'](\mathbf{v}) = Q^{\mathcal{M}_2}[\mathbf{A}_i'](\mathbf{v}), \tag{55}$$
$$\sum_{\mathbf{D}} Q^{\mathcal{M}_1}[\mathbf{S}](\mathbf{v}) = \sum_{\mathbf{D}} Q^{\mathcal{M}_2}[\mathbf{S}](\mathbf{v}'), \tag{56}$$

but there exists $\mathbf{v}_0 \in \mathfrak{X}(\mathbf{V})'$ such that:

$$Q^{\mathcal{M}_1}[\mathbf{S}](\mathbf{v}_0) \neq Q^{\mathcal{M}_2}[\mathbf{S}](\mathbf{v}_0). \tag{57}$$

Note that Equations (56)-(57) yield

$$Q[\mathbf{D}|\mathbf{S} \setminus \mathbf{D}]^{\mathcal{M}_1}(\mathbf{v}_0) \neq Q[\mathbf{D}|\mathbf{S} \setminus \mathbf{D}]^{\mathcal{M}_2}(\mathbf{v}_0).$$

This means that $Q[\mathbf{D}|\mathbf{S}\backslash\mathbf{D}]$ is not c-gID from $(\mathbb{A}', \mathcal{G})$.

To this end, we consider two cases.

First case:
Suppose that there exists $i \in [0, m]$, such that $\check{\mathbf{S}} \subset \mathbf{A}_i$. Then, consider the models $\mathcal{M}'_1$ and $\mathcal{M}'_2$ constructed in the section 1.4.1. According to the definitions of models $\mathcal{M}'_1$ and $\mathcal{M}'_2$ for any $\mathbf{v} \in \mathfrak{X}(\mathbf{V})'$, and any $i \in [0 : m']$, and any $g \in \{1, 2\}$, we have

$$\begin{aligned}
Q[\mathbf{A}'_i]^{\mathcal{M}'_g}(\mathbf{v}) &:= \sum_{U_0^{\mathcal{M}}} P^{\mathcal{M}_g}(u_0^{\mathcal{M}}) \sum_{U_0[\mathcal{P}]} \prod_{\hat{p} \in \mathcal{P}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \sum_{\mathbf{U}\backslash\{U_0\}} \prod_{X \in \mathbf{A}'_i} P^{\mathcal{M}'}(x \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U}\backslash\{U_0\}} P^{\mathcal{M}'}(u^{\mathcal{M}}),\\
\sum_{\mathbf{D}} Q[\mathbf{S}]^{\mathcal{M}'_g}(\mathbf{v}) &:= \sum_{U_0^{\mathcal{M}}} P^{\mathcal{M}_g}(u_0^{\mathcal{M}}) \sum_{U_0[\mathcal{P}]} \prod_{\hat{p} \in \mathcal{P}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \sum_{\mathbf{D}} \sum_{\mathbf{U}\backslash\{U_0\}} \prod_{X \in \mathbf{S}} P^{\mathcal{M}'}(x \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U}\backslash\{U_0\}} P^{\mathcal{M}'}(u^{\mathcal{M}}),\\
Q[\mathbf{S}]^{\mathcal{M}'_g}(\mathbf{v}) &:= \sum_{U_0^{\mathcal{M}}} P^{\mathcal{M}_g}(u_0^{\mathcal{M}}) \sum_{U_0[\mathcal{P}]} \prod_{\hat{p} \in \mathcal{P}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \sum_{\mathbf{U}\backslash\{U_0\}} \prod_{X \in \mathbf{S}} P^{\mathcal{M}'}(x \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U}\backslash\{U_0\}} P^{\mathcal{M}'}(u^{\mathcal{M}}).
\end{aligned}$$

We can re-writing the above equations using the notations of $\theta'_{i,j}$, ϕ'_j, and η'_j,

$$\begin{aligned}
Q[\mathbf{A}'_i]^{\mathcal{M}_1}(\mathbf{v}) &= \sum_{j=1}^{d} \frac{1}{d}\theta'_{i,j}(\mathbf{v}),\\
Q[\mathbf{A}'_i]^{\mathcal{M}_2}(\mathbf{v}) &= \sum_{j=1}^{d} p_j\theta'_{i,j}(\mathbf{v}),\\
\sum_{\mathbf{D}} Q[\mathbf{S}]^{\mathcal{M}_1}(\mathbf{v}) &= \sum_{j=1}^{d} \frac{1}{d}\phi'_j(\mathbf{v}[\mathbf{D}^\dagger]),\\
\sum_{\mathbf{D}} Q[\mathbf{S}]^{\mathcal{M}_2}(\mathbf{v}) &= \sum_{j=1}^{d} p_j\phi'_j(\mathbf{v}[\mathbf{D}^\dagger]),\\
Q[\mathbf{S}]^{\mathcal{M}_1}(\mathbf{v}) &= \sum_{j=1}^{d} \frac{1}{d}\eta'_j(\mathbf{v}),\\
Q[\mathbf{S}]^{\mathcal{M}_2}(\mathbf{v}) &= \sum_{j=1}^{d} p_j\eta'_j(\mathbf{v}).
\end{aligned}$$

The above equations imply the following equations.

$$\begin{aligned}
Q^{\mathcal{M}_2}[\mathbf{A}'_i](\mathbf{v}) - Q^{\mathcal{M}_1}[\mathbf{A}'_i](\mathbf{v}) &= \sum_{j=1}^{d}(p_j - \frac{1}{d})\theta'_{i,j}(\mathbf{v})\\
\sum_{\mathbf{D}} Q[\mathbf{S}]^{\mathcal{M}_2}(\mathbf{v}) - \sum_{\mathbf{D}} Q[\mathbf{S}]^{\mathcal{M}_1}(\mathbf{v}) &= \sum_{j=1}^{d}(p_j - \frac{1}{d})\phi'_j(\mathbf{v}[\mathbf{D}^\dagger])\\
Q^{\mathcal{M}_2}[\check{\mathbf{S}}](\mathbf{v}_0) - Q^{\mathcal{M}_1}[\check{\mathbf{S}}](\mathbf{v}_0) &= \sum_{j=1}^{d}(p_j - \frac{1}{d})\eta'_j(\mathbf{v}_0)\\
\sum_{j=1}^{d} p_j - 1 = \sum_{j=1}^{d}(p_j - \frac{1}{d}).
\end{aligned}$$

To prove the statement of the lemma it suffices to solve the following system of linear equations over parameters $\{p_j\}_{j=1}^{d}$

and show that it admits a solution.

$$
\begin{aligned}
&\sum_{j=1}^{d}(p_j-\frac{1}{d})\theta'_{i,j}(\mathbf{v})=0,\quad \forall \mathbf{v}\in\mathfrak{X}(\mathbf{V})', i\in[0:m'],\\
&\sum_{j=1}^{d}(p_j-\frac{1}{d})\phi'_j(\mathbf{d}^\dagger)=0,\quad \forall \mathbf{d}^\dagger\in\mathfrak{X}(\mathbf{D}^\dagger), i\in[0:m'],\\
&\sum_{j=1}^{d}(p_j-\frac{1}{d})\eta'_j(\mathbf{v}_0)\neq 0,\quad \exists \mathbf{v}_0\in\mathfrak{X}(\mathbf{V})',\\
&(p_j-\frac{1}{d})=0,\\
&0<p_j<1,\quad \forall j\in[1:d].
\end{aligned}
$$

Analogous to the proof of Lemma 1, we use Lemmas 6, 7, and 8 instead of Lemmas 1, 4 and 2 respectively and conclude the result.

Second case:
Suppose that there is no $i\in[0,m]$, such that $\breve{\mathbf{S}}\subset\mathbf{A}_i$. This case is identical to the **Second case** of the Lemma 1. □

1.5 PROOF OF LEMMA 7

Recall that $\mathbf{S}=Anc_{\mathcal{G}[\breve{\mathbf{V}}\setminus\mathbf{X}']}(\mathbf{Y}'\cup\mathbf{Z}')$ and it is assumed that is not gID from $(\mathbb{A},\mathcal{G})$. $\mathbf{S}$ consists of $\mathbf{S}_1,...,\mathbf{S}_n$ as its single c-components where $\mathbf{S}_1$ is not gID. Let $\breve{\mathbf{S}}=\mathbf{S}_1$. Clearly, we can add $\{\mathbf{S}_i\}_{i=2}^n$ to the known distributions and $\breve{\mathbf{S}}$ remains not gID, i.e., $\breve{\mathbf{S}}$ is not gID from $(\mathbb{A}',\mathcal{G})$, where $\mathbb{A}':=\mathbb{A}\cup\{\mathbf{S}_i\}_{i=2}^n$. For simplicity, we denote $\mathbb{A}'=\{\mathbf{A}'\}_{i=0}^{m'}$. Hence, using the method in Section 1.1.1, we can construct two models $\mathcal{M}_1$ and $\mathcal{M}_2$ that are the same over the known distributions and different over $Q[\mathbf{S}_1]$. These models disagree on the distribution $Q[\mathbf{S}]$ as well, because $Q[\mathbf{S}]=\prod_{i=1}^n Q[\mathbf{S}_i]$. Below, we use these two models to introduce two new models to prove Lemma 7.

1.5.1 New models for Lemma 7

Recall that T is a node in $\breve{\mathbf{S}}\setminus(\mathbf{Z}'\cup\mathbf{Y}')$, p_T is a shortest directed path from node T to the node Z', $\widetilde{\mathbf{F}}$ is a set of all colliders on the path $\widetilde{p}$, $\widetilde{\mathcal{P}}:=\{\widetilde{p}\}\cup\{p_T\}\cup\{\widetilde{p}_W|W\in\widetilde{\mathbf{F}}\}$ and $\widetilde{\mathbf{D}}$ is a set of all nodes on the paths from $\widetilde{\mathcal{P}}$ excluding the nodes in $\mathbf{Z}'$. Let $\widetilde{\mathbf{D}}_{\mathcal{P}}$ be a set of all variables that belong to at least one path in $\mathcal{P}$.

Similar to the Section 1.4.1 further we define new models $\widetilde{\mathcal{M}}_1$ and $\widetilde{\mathcal{M}}_2$ based on the models $\mathcal{M}_1$ and $\mathcal{M}_2$ defined in Section 1.1.1. We say that a variable D is a **starting node** of the path $\hat{p}\in\widetilde{\mathcal{P}}$ if

- $D=T$ and $\hat{p}=p_T$, or
- $D=Z'$ and $\hat{p}=\widetilde{p}$, or
- $D\in\widetilde{\mathbf{F}}$, i.e., it is a collider on path $\widetilde{p}$, and $\hat{p}=\widetilde{p}_D$.

Note that D can be a starting node of only one path.

For $R\in\mathbf{V}\cup\mathbf{U}$, let $\widetilde{\alpha}_p(R)$ be the number of paths in $\widetilde{\mathcal{P}}$ that contains R. Furthermore, we use $\widetilde{\mathfrak{X}}(R)$ and $\mathfrak{X}(R)$ to denote its domain in $\widetilde{\mathcal{M}}_1$ or $\widetilde{\mathcal{M}}_2$ (variables in different models have the same domains) and in $\mathcal{M}_1$ or $\mathcal{M}_2$ respectively. We define $\widetilde{\mathfrak{X}}(R)$ as follows:

- If R is a starting node for one of the paths in $\mathcal{P}$

$$\widetilde{\mathfrak{X}}(R):=\mathfrak{X}(R)\times[0:\kappa]^{\widetilde{\alpha}_p(R)-1}.$$

- If R is not a starting node for any of the paths in $\mathcal{P}$, then:

$$\widetilde{\mathfrak{X}}(R):=\mathfrak{X}(R)\times[0:\kappa]^{\widetilde{\alpha}_p(R)}.$$

Consequently, if R does not belong to any of the paths in $\widetilde{\mathcal{P}}$, then $\widetilde{\mathfrak{X}}(R) := \mathfrak{X}(R)$.

Consider $R \in \mathbf{V} \cup \mathbf{U}$. According to the domain's definitions above, R is a vector that is a concatenation of the vector coming from $\mathfrak{X}(R)$ in model $\mathcal{M}_1$ (or $\mathcal{M}_2$) and some additional coordinates. These additional coordinates are defined based on $\widetilde{\alpha}_p(R)$. More precisely, if R is not a starting node of a path $\hat{p} \in \widetilde{\mathcal{P}}$, then there is a coordinate assigned to this path, denoted by $R[\hat{p}]$, otherwise, if R is a starting node of $\hat{p} \in \widetilde{\mathcal{P}}$, then there is no coordinate assigned this path.

Let $\mathbf{O} \in \mathbf{V} \cup \mathbf{U}$. For any realization $\mathbf{o} \in \widetilde{\mathfrak{X}}(\mathbf{O})$ of $\mathbf{O}$ we denote by $\mathbf{o}^{\mathcal{M}} \in \mathfrak{X}(O)$ a realization of $\mathbf{O}$ that is consistent with $\mathbf{o}$. With slight abuse of notation, we use $\mathbf{O}$ and $\mathbf{O}^{\mathcal{M}}$ to denote realizations of $\mathbf{O}$ in models $\widetilde{\mathcal{M}}_i$ and $\mathcal{M}_i$, respectively. $\mathbf{O}^{\mathcal{M}}$ means realizations in $\mathfrak{X}(O)$ that are consistent with realizations in $\widetilde{\mathfrak{X}}(\mathbf{O})$.

Recall that $\mathbf{D}_{\widetilde{\mathcal{P}}}$ is a set of all variables on the paths in $\widetilde{\mathcal{P}}$. Let $D \in \widetilde{\mathbf{D}}_{\mathcal{P}}$. We denote by $\widetilde{\mathcal{P}}_D$ the set of all paths $\hat{p}$, such that $\hat{p} \in \widetilde{\mathcal{P}}$, D belongs to the path $\hat{p}$, and D is not a starting node of path $\hat{p}$. We are ready now to define the probabilities of $P^{\widetilde{\mathcal{M}}_i}(D|Pa_{\mathcal{G}}(D))$ for any $D \in \mathbf{V} \cup \mathbf{U}$ and $i \in \{1, 2\}$.

- If D does not belong to the set $\widetilde{\mathbf{D}}_{\mathcal{P}}$, we define

$$P^{\widetilde{\mathcal{M}}_i}(D|Pa_{\mathcal{G}}(D)) := P^{\mathcal{M}_i}(D|Pa_{\mathcal{G}}(D)).$$

- If D belongs to the set $\widetilde{\mathbf{D}}_{\mathcal{P}} \backslash \{Z'\}$, we define

$$P^{\widetilde{\mathcal{M}}_i}(D|Pa_{\mathcal{G}}(D)) := P^{\mathcal{M}_i}(D^{\mathcal{M}}|Pa_{\mathcal{G}}(D)) \prod_{\hat{p} \in \widetilde{\mathcal{P}}_D} f_{\hat{p}}(D[\hat{p}]|Pa_{\hat{p}}(D)), \tag{58}$$

  where

  - if $D \neq Z'$ or $\hat{p} \neq p_T$ then $Pa_{\hat{p}}(D)$ is a parents of D in a path $\hat{p}$;
  - if $D = Z'$ and $\hat{p} = p_T$ then $Pa_{\hat{p}}(Z')$ is a parents of Z' on the paths $\hat{p}$ and p_T;
  - $f_{\hat{p}}(D|Pa_{\hat{p}}(D))$ is given below.

  Definition of function $f_{\hat{p}}(D[\hat{p}]|Pa_{\hat{p}}(D))$:

  - When there exists a variable $W \in \widetilde{\mathbf{F}}$ such that $\hat{p} = \widetilde{p}_W$ and D is a child of W on path p_W,

$$f_{\hat{p}}(D[\hat{p}]|Pa_{\hat{p}}(D)) := \begin{cases} 1 - \kappa\epsilon & \text{if } D[\hat{p}] \equiv W[\widetilde{p}] \pmod{\kappa + 1} \\ \epsilon & \text{if } D[\hat{p}] \not\equiv W[\widetilde{p}] \pmod{\kappa + 1}. \end{cases}$$

  - When $\hat{p} = p_T$ and D is a child of T on path $\hat{p}$,

$$f_{\hat{p}}(D[\hat{p}]|Pa_{\hat{p}}(D)) := \begin{cases} 1 - \kappa\epsilon & \text{if } D[\hat{p}] \equiv T^{\mathcal{M}} \pmod{\kappa + 1} \\ \epsilon & \text{if } D[\hat{p}] \not\equiv T^{\mathcal{M}} \pmod{\kappa + 1}. \end{cases}$$

  - When $\hat{p} = p_T$ and $D = Z'$. Suppose Z' is a child of W' on a path p_T and is a child of W'' on a path $\widetilde{p}$,

$$f_{\hat{p}}(Z'[p_T]|Pa_{\hat{p}}(D)) := \begin{cases} 1 - \kappa\epsilon & \text{if } Z'[p_T] \equiv W'[p_T] + W''[\widetilde{p}] \pmod{\kappa + 1} \\ \epsilon & \text{if } Z'[p_T] \not\equiv W'[p_T] + W''[\widetilde{p}] \pmod{\kappa + 1}. \end{cases}$$

  - When $Pa_{\hat{p}}(D) = \varnothing$,

$$f_{\hat{p}}(D[\hat{p}]) := \frac{1}{\kappa + 1}.$$

  - Otherwise,

$$f_{\hat{p}}(D[\hat{p}]|Pa_{\hat{p}}(D)) := \begin{cases} 1 - \kappa\epsilon & \text{if } D[\hat{p}] \equiv \sum_{D' \in Pa_{\hat{p}}(D)} D'[\hat{p}] \pmod{\kappa + 1} \\ \epsilon & \text{if } D[\hat{p}] \not\equiv \sum_{D' \in Pa_{\hat{p}}(D)} D'[\hat{p}] \pmod{\kappa + 1}, \end{cases} \tag{59}$$

  Note that $P^{\widetilde{\mathcal{M}}_i}(D|Pa_{\mathcal{G}}(D))$ is a probability distribution since for different paths $\hat{p}_1$ and $\hat{p}_2$, $D[\hat{p}_1]$ and $D[\hat{p}_2]$ are different and also

$$\sum_{D[\hat{p}] \in \widetilde{\mathfrak{X}}(D[\hat{p}])} f_{\hat{p}}(D[\hat{p}]|Pa_{\hat{p}}(D)) = 1$$

Note that for any $W \in (\mathbf{V} \cup \mathbf{U})\backslash\{U_0\}$, we have

$$P^{\widetilde{\mathcal{M}}_1}(W|Pa_{\mathcal{G}}(W)) = P^{\widetilde{\mathcal{M}}_2}(W|Pa_{\mathcal{G}}(W)).$$

Therefore, we will use $P^{\widetilde{\mathcal{M}}}(W|Pa_{\mathcal{G}}(W))$ instead of $P^{\widetilde{\mathcal{M}}_1}(W|Pa_{\mathcal{G}}(W))$ or $P^{\widetilde{\mathcal{M}}_2}(W|Pa_{\mathcal{G}}(W))$ for $W \in (\mathbf{V} \cup \mathbf{U})\backslash\{U_0\}$. We also have

$$\begin{aligned} P^{\widetilde{\mathcal{M}}_1}(U_0) &= \frac{1}{d} \prod_{\hat{p} \in \widetilde{\mathcal{P}}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]), \\ P^{\widetilde{\mathcal{M}}_2}(U_0) &= P^{\mathcal{M}_2}(U_0^{\mathcal{M}}) \prod_{\hat{p} \in \widetilde{\mathcal{P}}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]). \end{aligned} \tag{60}$$

Recall that $\mathbf{S} = Anc_{\mathcal{G}[\mathbf{V}\backslash\mathbf{X}']}(\mathbf{Y}' \cup Z')$. Let $\widetilde{\mathbf{D}}' := \mathbf{S}\backslash\widetilde{\mathbf{D}}$ and $\widetilde{\mathbf{D}}^{\dagger} := \mathbf{V}\backslash\widetilde{\mathbf{D}}$. For $i \in [0 : m']$, $j \in [1 : d]$, $\mathbf{v} \in \widetilde{\mathfrak{X}}(\mathbf{V})$ and $\mathbf{d}^{\dagger} \in \widetilde{\mathfrak{X}}(\widetilde{\mathbf{D}}^{\dagger})$, we define $\theta_{i,j}^{'}(\mathbf{v})$, $\phi_j^{'}(\mathbf{d}^{\dagger})$ and $\eta_j^{'}(\mathbf{v})$ as follows:

$$\widetilde{\theta}_{i,j}(\mathbf{v}) := \sum_{U_0[\widetilde{\mathcal{P}}]} \prod_{\hat{p} \in \widetilde{\mathcal{P}}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \sum_{\mathbf{U}\backslash\{U_0\}} \prod_{X \in \mathbf{A}_i'} P^{\widetilde{\mathcal{M}}}(x \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U}\backslash\{U_0\}} P^{\widetilde{\mathcal{M}}}(u^{\mathcal{M}}), \tag{61}$$

$$\widetilde{\phi}_j(\mathbf{d}^{\dagger}) := \sum_{U_0[\widetilde{\mathcal{P}}]} \prod_{\hat{p} \in \widetilde{\mathcal{P}}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \sum_{\widetilde{\mathbf{D}}} \sum_{\mathbf{U}\backslash\{U_0\}} \prod_{X \in \mathbf{S}} P^{\widetilde{\mathcal{M}}}(x \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U}\backslash\{U_0\}} P^{\widetilde{\mathcal{M}}}(u^{\mathcal{M}}), \tag{62}$$

$$\widetilde{\eta}_j(\mathbf{v}) := \sum_{U_0[\widetilde{\mathcal{P}}]} \prod_{\hat{p} \in \widetilde{\mathcal{P}}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \sum_{\mathbf{U}\backslash\{U_0\}} \prod_{X \in \mathbf{S}} P^{\widetilde{\mathcal{M}}}(x \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U}\backslash\{U_0\}} P^{\widetilde{\mathcal{M}}}(u^{\mathcal{M}}), \tag{63}$$

where $\sum_{U_0[\widetilde{\mathcal{P}}]}$ is a summation over realizations of the random variables $\{U_0[\hat{p}] | \, \hat{p} \in \widetilde{\mathcal{P}}_{U_0}\}$.

Next, we prove three lemmas similar to Lemmas 6, 7, and 8 for the new models $\widetilde{\mathcal{M}}_1$ and $\widetilde{\mathcal{M}}_2$.

Lemma 10. *For any* $\mathbf{v} \in \widetilde{\mathfrak{X}}(\mathbf{V})$ *and* $i \in [0 : m']$*,*

$$\widetilde{\theta}_{i,j_1}(\mathbf{v}) = \widetilde{\theta}_{i,j_2}(\mathbf{v}) = \cdots = \widetilde{\theta}_{i,j_{\frac{\kappa+1}{2}}}(\mathbf{v}).$$

Proof. By substituting $P^{\widetilde{\mathcal{M}}}$ from the above into Equation (41) and rearranging the terms, we obtain

$$\begin{aligned} \widetilde{\theta}_{i,j}(\mathbf{v}) = &\sum_{\mathbf{U}[\widetilde{\mathcal{P}}]} \prod_{\hat{U}[\hat{p}] \in \mathbf{U}[\widetilde{\mathcal{P}}]} f_{\hat{p}}(\hat{U}) \prod_{\hat{X}[\hat{p}] \in \mathbf{A}_i'[\widetilde{\mathcal{P}}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X})) \times \\ &\times \Big(\sum_{\mathbf{U}^{\mathcal{M}}\backslash\{U_0\}^{\mathcal{M}}} \prod_{X \in \mathbf{A}_i'} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U}\backslash\{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}}) \Big) \end{aligned}$$

Note that the terms inside the big parenthesis of the above equation is equal to $\theta_{i,j}$ given by 3, i.e.,

$$\widetilde{\theta}_{i,j}(\mathbf{v}) = \sum_{\mathbf{U}[\widetilde{\mathcal{P}}]} \prod_{\hat{U}[\hat{p}] \in \mathbf{U}[\widetilde{\mathcal{P}}]} f_{\hat{p}}(\hat{U}) \prod_{\hat{X}[\hat{p}] \in \mathbf{A}_i'[\widetilde{\mathcal{P}}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X}))\theta_{i,j}(\mathbf{v}^{\mathcal{M}}).$$

In the last equation, all terms on the right-hand side except $\theta_{i,j}(\mathbf{v}^{\mathcal{M}})$ are independent of the realization of $\{U_0\}^{\mathcal{M}}$, i.e., independent of index j. For $j \in \{j_1, j_2, \ldots, j_{\frac{\kappa+1}{2}}\}$ and using the result of Lemma 1 that says $\theta_{i,j_1}(\mathbf{v}) = \theta_{i,j_2}(\mathbf{v}) = \cdots = \theta_{i,j_{\frac{\kappa+1}{2}}}(\mathbf{v})$, we can conclude the result. ∎

Lemma 11. *For any* $\mathbf{d}^{\dagger} \in \widetilde{\mathfrak{X}}(\widetilde{\mathbf{D}}^{\dagger})$*:*

$$\widetilde{\phi}_{j_1}(\mathbf{d}^{\dagger}) = \widetilde{\phi}_{j_2}(\mathbf{d}^{\dagger}) = \cdots = \widetilde{\phi}_{j_{\frac{k+1}{2}}}(\mathbf{d}^{\dagger}).$$

Proof. Similar to the previous lemma, by substituting $P^{\widetilde{\mathcal{M}}}$ from their definitions into Equation (62) and rearranging the terms, we obtain:

$$\begin{aligned}\widetilde{\phi}_j(\mathbf{d}^\dagger) := &\sum_{U_0[\widetilde{\mathcal{P}}]} \sum_{\mathbf{U}^{\mathcal{M}}\setminus\{U_0\}^{\mathcal{M}}} \sum_{\widetilde{\mathbf{D}}} \prod_{\hat{p}\in\widetilde{\mathcal{P}}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \prod_{\hat{X}[\hat{p}]\in\mathbf{S}[\widetilde{\mathcal{P}}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X}))\times \\ &\times\Big(\prod_{X\in\mathbf{S}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U\in\mathbf{U}\setminus\{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}})\Big)\end{aligned} \tag{64}$$

Suppose that l_1 and l_2 are two integers such that

$$\begin{aligned}\gamma_{l_1} &= (2x, 0, \dots, 0)),\\ \gamma_{l_2} &= (2x+2 \pmod{\kappa+1}, 0, \dots, 0),\end{aligned}$$

and x is an integer in $[0 : \frac{\kappa-1}{2}]$. We will prove that $\widetilde{\phi}_{l_1}(\mathbf{d}^\dagger) = \widetilde{\phi}_{l_2}(\mathbf{d}^\dagger)$.

Suppose that path p is the sequence of variables: Z', D'_1, $D'_2 \dots$, $D'_{k'_1}$, $D'_{k'_1+1} := Y'$ and path p_T is a sequence of variables: $T_0 := T, T_1, \dots, T_{k'_2}, T_{k'_2+1} := Z'$. Note that direct edge between Z' and D'_1 is pointing toward Z', i.e., $Z' \leftarrow D'_1$ and for all $i \in [0, k'_2]$ variable T_i is a parent of T_{i+1} on the path p_T.

On the other hand, since T and U_0 are both in $\breve{\mathbf{S}}$ ($\breve{\mathbf{S}} = \mathbf{S}_1$ by construction), then there exists a shortest path $U_0, \hat{S}'_1, \hat{U}'_1, \hat{S}'_2, \hat{U}'_2, \dots, \hat{U}'_{l'}, T$, such that U_0 is a parent of $\hat{S}'_1 \in \breve{\mathbf{S}}$, T is a child of $\hat{U}'_l \in \breve{\mathbf{U}}^{\breve{\mathbf{S}}}$, and $\hat{U}'_j \in \breve{\mathbf{U}}^{\breve{\mathbf{S}}}$ is a parent of variables $\hat{S}'_j \in \breve{\mathbf{S}}$ and $\hat{S}'_{j+1} \in \breve{\mathbf{S}}$ for any $j \in [1 : l'-1]$. Let $\hat{\mathbf{U}}' := \{\hat{U}'_1, \dots, \hat{U}'_l\}$, i.e., unobserved nodes in this shortest path except U_0. For any given realization $\mathbf{o}_1 \in \widetilde{\mathfrak{X}}(\mathbf{U} \cup \widetilde{\mathbf{D}})$, we define $\mathbf{o}_2 \in \widetilde{\mathfrak{X}}(\mathbf{U} \cup \widetilde{\mathbf{D}})$ as follows,

$$\begin{aligned}\mathbf{o}_2^{\mathcal{M}}[\hat{U}'_j] &:= \mathbf{o}_1^{\mathcal{M}}[\hat{U}'_j] + 2(-1)^j \pmod{\kappa+1}, \quad \forall j \in [0 : l'],\\ \mathbf{o}_2^{\mathcal{M}}[T] &:= \mathbf{o}_1^{\mathcal{M}}[\hat{U}'_j] + 2(-1)^{l'} \pmod{\kappa+1}.\end{aligned} \tag{65}$$

Note that if $\mathbf{o}_1^{\mathcal{M}}[U_0] = \gamma_{l_1}$ then $\mathbf{o}_2^{\mathcal{M}}[U_0] = \gamma_{l_2}$. With these modifications for any $S \in \breve{\mathbf{S}}$, we obtain

$$s_1 - M(S) \equiv s_2 - M(S) \pmod{\kappa+1},$$

where s_1 is a realization of $S\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}},\widetilde{\mathbf{D}}^\dagger)=(\mathbf{o}_1,\mathbf{d}^\dagger)}$, s_2 is a realization of $S\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}},\widetilde{\mathbf{D}}^\dagger)=(\mathbf{o}_2,\mathbf{d}^\dagger)}$, $M(\cdot)$ is given by Equation (9). This implies for any $S \in \mathbf{S}$, we have

$$P^{\mathcal{M}}(s|Pa_{\mathcal{G}}(S))\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}},\widetilde{\mathbf{D}}^\dagger)=(\mathbf{o}_1,\mathbf{d}^\dagger)} = P(s|Pa_{\mathcal{G}}(S))\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}},\widetilde{\mathbf{D}}^\dagger)=(\mathbf{o}_2,\mathbf{d}^\dagger)}.$$

Let $c = -2(-1)^{l'}$ and we define

$$\begin{aligned}\mathbf{o}_2^{\mathcal{M}}[T_j[p_T]] &:= \mathbf{o}_1^{\mathcal{M}}[T_j[p_T]] - c \pmod{\kappa+1}, \quad \forall j \in [1 : k'_2],\\ \mathbf{o}_2^{\mathcal{M}}[D'_1[\widetilde{p}]] &:= \mathbf{o}_1^{\mathcal{M}}[D'_1[\widetilde{p}]] + c \pmod{\kappa+1}.\end{aligned}$$

This implies that for all $j \in [1 : k'_2+1]$ we have

$$f_{p_T}(T_j|Pa_{p_T}(T_j))\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}},\widetilde{\mathbf{D}}^\dagger)=(\mathbf{o}_1,\mathbf{d}^\dagger)} = f_{p_T}(T_j|Pa_{p_T}(T_j))\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}},\widetilde{\mathbf{D}}^\dagger)=(\mathbf{o}_2,\mathbf{d}^\dagger)}.$$

Assume that D'_j is not a collider on the path $\widetilde{p}$ and $j \in [2 : k'_1+1]$. We define $\mu(D'_j)$ to be the number of colliders on a part of the path $\widetilde{p}$ from D'_1 to D'_{j-1}. Thus, for those $j \in [2 : k'_1+1]$ that D'_j is not a collider, we define

$$\mathbf{o}_2^{\mathcal{M}}[D'_j[\widetilde{p}]] := \mathbf{o}_1^{\mathcal{M}}[D'_j[\widetilde{p}]] + c(-1)^{\mu(D'_j)}. \tag{66}$$

Note that the modifications in (66) might only affect the function $f_{\widetilde{p}}(\cdot|\cdot)$. Next, we show that after these modifications, function $f_{\widetilde{p}}(\cdot|\cdot)$ remains unchanged. To do so, for $j \in [1 : k'_1+1]$ we consider four different cases:

1. If D'_j has no parents, then it is obvious that

$$f_{\widetilde{p}}(D'_j[\widetilde{p}])\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}})=(\mathbf{o}_1)} = f_{\widetilde{p}}(D'_j[\widetilde{p}])\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}})=(\mathbf{o}_2)}.$$

2. If D'_j is a collider, then $\mu(D'_{j+1}) = \mu(D'_{j-1}) + 1$ and

$$\mathbf{o}_1[D'_{j+1}[\widetilde{p}]] + \mathbf{o}_1[D'_{j-1}[\widetilde{p}]] = \mathbf{o}_2[D'_{j+1}[\widetilde{p}]] + \mathbf{o}_2[D'_{j-1}[\widetilde{p}]],$$

   and hence, according to the Equation (59), we have

$$f_{\widetilde{p}}(D'_j[\widetilde{p}]|Pa_{\widetilde{p}}(D'_j))\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}})=(\mathbf{o}_1)} = f_{\widetilde{p}}(D'_j[\widetilde{p}]|Pa_{\widetilde{p}}(D'_j))\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}})=(\mathbf{o}_2)}.$$

3. If D'_j is a child of D'_{j+1}, then $\mu(D'_j) = \mu(D'_{j+1})$ and

$$\mathbf{o}_1[D'_j[\widetilde{p}]] - \mathbf{o}_1[D'_{j+1}[\widetilde{p}]] = \mathbf{o}_2[D'_j[\widetilde{p}]] - \mathbf{o}_2[D'_{j+1}[\widetilde{p}]].$$

   According to Equation (59), we imply that

$$f_{\widetilde{p}}(D'_j[\widetilde{p}]|Pa_{\widetilde{p}}(D'_j))\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}})=(\mathbf{o}_1)} = f_{\widetilde{p}}(D'_j[\widetilde{p}]|Pa_{\widetilde{p}}(D'_j))\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}})=(\mathbf{o}_2)}.$$

4. If D'_j is a child of D'_{j-1}, then $\mu(D'_j) = \mu(D'_{j-1})$ and

$$\mathbf{o}_1[D'_j[\widetilde{p}]] - \mathbf{o}_1[D'_{j-1}[\widetilde{p}]] = \mathbf{o}_2[D'_j[\widetilde{p}]] - \mathbf{o}_2[D'_{j-1}[\widetilde{p}]].$$

   Similarly, according to Equation (59), we get

$$f_{\widetilde{p}}(D'_j[\widetilde{p}]|Pa_{\widetilde{p}}(D'_j))\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}})=(\mathbf{o}_1)} = f_{\widetilde{p}}(D'_j[\widetilde{p}]|Pa_{\widetilde{p}}(D'_j))\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}})=(\mathbf{o}_2)}.$$

This concludes that for any $j \in [1 : k'_1 + 1]$,

$$f_{\widetilde{p}}(D'_j[\widetilde{p}]|Pa_{\widetilde{p}}(D'_j))\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}})=(\mathbf{o}_1)} = f_{\widetilde{p}}(D'_j[\widetilde{p}]|Pa_{\widetilde{p}}(D'_j))\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}})=(\mathbf{o}_2)}.$$

Note that the aforementioned transformation of $\mathbf{o}_1$ affects only those realizations of variables that are used for marginalization in the Equation (64). Putting the above results together implies that the terms in Equation (64) remain unchanged, i.e.,

$$\begin{aligned}
&\prod_{\hat{p}\in\widetilde{\mathcal{P}}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \prod_{\hat{X}[\hat{p}]\in\mathbf{S}[\widetilde{\mathcal{P}}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X})) \\
&\times \Big(\prod_{X\in\mathbf{S}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U\in\mathbf{U}} P^{\mathcal{M}}(u^{\mathcal{M}})\Big)\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}})=(\mathbf{o}_1)} = \\
&= \prod_{\hat{p}\in\widetilde{\mathcal{P}}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \prod_{\hat{X}[\hat{p}]\in\mathbf{S}[\widetilde{\mathcal{P}}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X})) \\
&\times \Big(\prod_{X\in\mathbf{S}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U\in\mathbf{U}\setminus\{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}})\Big)\Big|_{(\mathbf{U}\cup\widetilde{\mathbf{D}})=(\mathbf{o}_2)}
\end{aligned}$$

This implies that $\widetilde{\phi}_{l_1}(\mathbf{d}^\dagger) = \widetilde{\phi}_{l_2}(\mathbf{d}^\dagger)$. By varying x within $[0 : \frac{\kappa-1}{2}]$ in the definition of γ_{l_1} and γ_{l_2}, we obtain the result. □

Lemma 12. *There exists* $0 < \epsilon < \frac{1}{\kappa}$*, such that there exists* $\mathbf{v}_0 \in \widetilde{\mathfrak{X}}(\mathbf{V})$ *and* $1 \leqslant r < t \leqslant \frac{\kappa+1}{2}$ *such that*

$$\widetilde{\eta}_{j_r}(\mathbf{v}_0) \neq \widetilde{\eta}_{j_t}(\mathbf{v}_0).$$

Proof. By substituting $P^{\widetilde{\mathcal{M}}}$ from their definitions into Equation (63) and rearranging the terms, we obtain

$$\begin{aligned}
\widetilde{\eta}_j(\mathbf{v}_0) = &\sum_{\mathbf{U}^{\mathcal{M}}\setminus\{U_0\}^{\mathcal{M}}} \sum_{\mathbf{U}[\widetilde{\mathcal{P}}]} \prod_{\hat{U}[\hat{p}]\in\mathbf{U}[\widetilde{\mathcal{P}}]} f_{\hat{p}}(\hat{U}) \prod_{\hat{X}[\hat{p}]\in\mathbf{S}[\widetilde{\mathcal{P}}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X}))\times \\
&\times\Big(\prod_{X\in\mathbf{S}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U\in\mathbf{U}\setminus\{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}})\Big),
\end{aligned} \tag{67}$$

Next, we define $\mathbf{v}_0 \in \widetilde{\mathfrak{X}}(\mathbf{V})$ such that the conditions in the lemma hold.

- For any path $\hat{p} \in \widetilde{\mathcal{P}}$ and any node W on the path $\hat{p}$ that is not a starting node for path $\hat{p}$ we define:

$$\mathbf{v}_0[W[\hat{p}]] := 0;$$

- For any variable $S \in \breve{\mathbf{S}}$, we define

$$\mathbf{v}_0^{\mathcal{M}}[S] := 0;$$

- For the remaining part of $\mathbf{v}_0$, we choose a realization such that for the selected $\mathbf{v}_0$, there exists a realization for the unobserved variables $\mathbf{U}$ that ensures $\mathbb{I}(I)(S) = 0$ for all $S \in \breve{\mathbf{S}}$. This is clearly possible due to the definition of $\mathbb{I}(S)$.

Assume r and t are such that $\gamma_{j_r} = (0, 0, \ldots, 0)$ and $\gamma_{j_t} = (2, 0, \ldots, 0)$. To finish the proof of the lemma, it is enough to show that $\widetilde{\eta}_{j_r}(\mathbf{v}_0)$ and $\widetilde{\eta}_{j_t}(\mathbf{v}_0)$ are two different polynomial functions of parameter ϵ. We prove that those two polynomials are different by showing that $\widetilde{\eta}_{j_r}(\mathbf{v}_0) \neq \widetilde{\eta}_{j_t}(\mathbf{v}_0)$ for $\epsilon = 0$.

We only need to consider the non-zero terms in Equation (67). From (67), we have

$$\begin{aligned} &\prod_{\hat{U}[\hat{p}] \in \mathbf{U}[\widetilde{\mathcal{P}}]} f_{\hat{p}}(\hat{U}) \prod_{\hat{X}[\hat{p}] \in \mathbf{S}[\widetilde{\mathcal{P}}]} f_{\hat{p}}(\hat{X}[\hat{p}] | Pa_{\hat{p}}(\hat{X})) \\ &\times \Big(\prod_{X \in \mathbf{S}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U \in \mathbf{U} \setminus \{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}}) \Big). \end{aligned} \tag{68}$$

Note that $f_{\hat{p}}(\hat{U}) = \frac{1}{\kappa+1}$ and $f_{\hat{p}}(\hat{X}|Pa_{\hat{p}}(\hat{X}))$ is non-zero only:

- when $\hat{p} = p_T$, $\hat{X}$ is a child of T on the path p_T, and

$$\hat{X}[\hat{p}] \equiv T^{\mathcal{M}} \pmod{\kappa + 1}.$$

- when $\hat{p} = p_T$, $\hat{X} = Z'$, and

$$Z'[\hat{p}] \equiv W'[\hat{p}] + W''[\widetilde{p}] \pmod{\kappa + 1},$$

  where W' is a parent of Z' on the path p_T and W'' is a parent of Z' on the path $\widetilde{p}$.
- when there exists a variable $W \in \mathbf{F}$ such that $\hat{p}' = \widetilde{p}_W$, $\hat{X}$ is a child of W in path $\widetilde{p}_W$, and

$$\hat{X}[\hat{p}] \equiv W[\widetilde{p}] \pmod{\kappa + 1}.$$

- when the following holds

$$\hat{X}[\hat{p}] \equiv \sum_{\hat{X}' \in Pa_{\hat{p}}(\hat{X}) \setminus \{W\}} \hat{X}'[\hat{p}] \pmod{\kappa + 1}.$$

Similarly, $P^{\mathcal{M}}(X|Pa_{\mathcal{G}}(X))$ is non-zero

- if $\mathbb{I}(X) = 1$ (i.e. $P^{\mathcal{M}}(X^{\mathcal{M}}|Pa_{\mathcal{G}}(X)) = \frac{1}{\kappa+1}$), or
- if $X^{\mathcal{M}} \equiv M(X) \pmod{\kappa + 1}$ for $P^{\mathcal{M}}(X^{\mathcal{M}}|Pa_{\mathcal{G}}(X))$.

Let fix a realization $\mathbf{u} \in \widetilde{\mathfrak{X}}(\mathbf{U} \setminus \{U_0^M\})$. We consider two scenarios:

I) Assume that for this realization, there is a variable $S \in \breve{\mathbf{S}}$, such that $\mathbb{I}(S) = 1$ and S is the closest variable to U_0 considering only paths with bidirected edges in $\mathcal{G}'[\breve{\mathbf{S}}]$. The value of $S^{\mathcal{M}}$ does not depend on its parents because of $\mathbb{I}(S) = 1$ and Equation (8). Additionally in the graph $\mathcal{G}'[\breve{\mathbf{S}}]$ there exists a path $U_0, \hat{S}_1', \hat{U}_1', \hat{S}_2', \hat{U}_2', \ldots, \hat{U}_{l'}', S$, such that U_0 is a parent of $\hat{S}_1' \in \breve{\mathbf{S}}$, S is a child of $\hat{U}_l' \in \breve{\mathbf{U}}^{\breve{\mathbf{S}}}$, and $\hat{U}_j' \in \breve{\mathbf{U}}^{\breve{\mathbf{S}}}$ is a parent of variables $\hat{S}_j' \in \breve{\mathbf{S}}$ and $\hat{S}_{j+1}' \in \breve{\mathbf{S}}$ for $j \in [1 : l' - 1]$. Let $\hat{\mathbf{U}}' := \{\hat{U}_1', \ldots, \hat{U}_l'\}$. We define $\mathbf{u}' \in \widetilde{\mathfrak{X}}(\mathbf{U} \setminus \{U_0^{\mathcal{M}}\})$ that is consistent with $\mathbf{u}$ except the variables in $\breve{\mathbf{U}}$. For these variables, we define

$$\mathbf{u}'^{\mathcal{M}}[\hat{U}_j] := \mathbf{u}^{\mathcal{M}}[\hat{U}_j] + 2(-1)^j \pmod{\kappa + 1}, \quad j \in [1 : l'], \tag{69}$$

With this modification for any $\widetilde{S} \in \mathbf{S}$, we have

$$P^{\mathcal{M}}(\widetilde{s}|Pa_{\mathcal{G}}(\widetilde{S}))\Big|_{(\mathbf{U})=(\mathbf{u},\gamma_{l_1})} = P(\widetilde{s}|Pa_{\mathcal{G}}(\widetilde{S}))\Big|_{(\mathbf{U})=(\mathbf{u}',\gamma_{l_2})}.$$

Therefore for all such realizations of $\mathbf{u}$ the summation of the following terms for both $\widetilde{\eta}_{j_r}(\mathbf{v}_0)$ and $\widetilde{\eta}_{j_t}(\mathbf{v}_0)$ will be the same,

$$\prod_{\hat{U}[\hat{p}]\in\mathbf{U}[\widetilde{\mathcal{P}}]} f_{\hat{p}}(\hat{U}) \prod_{\hat{X}[\hat{p}]\in\mathbf{S}[\widetilde{\mathcal{P}}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X}))\times \\ \times\Big(\prod_{X\in\mathbf{S}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U\in\mathbf{U}\setminus\{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}})\Big). \tag{70}$$

II) Assume that for all $S \in \check{\mathbf{S}}$, we have $\mathbb{I}(S) = 0$. We consider a realization $U_0^{\mathcal{M}} = \gamma_{j_r}$ and $\mathbf{u}$ such that:

- $\mathbf{u}[\check{\mathbf{U}}^{\check{\mathbf{S}}}] = \mathbf{0}$.
- for all $U \in \mathbf{U}$ and any path $\hat{p} \in \mathcal{P}$ which contains U, $\mathbf{u}[U[\hat{p}]] = 0$.

We claim that for such $\mathbf{u}$,

$$\prod_{\hat{U}[\hat{p}]\in\mathbf{U}[\widetilde{\mathcal{P}}]} f_{\hat{p}}(\hat{U}) \prod_{\hat{X}[\hat{p}]\in\mathbf{S}[\widetilde{\mathcal{P}}]} f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X}))\times \\ \times\Big(\prod_{X\in\mathbf{S}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U\in\mathbf{U}\setminus\{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}})\Big).$$

is non-zero. To prove this claim we consider 5 cases:

- assume that $\hat{p} = p_T$ and $\hat{X} = Z'$. Denote by W' parent of Z' on the path p_T and by W'' parent of Z' on the path $\widetilde{p}$. From the definition of $\mathbf{u}$ and $\mathbf{v_0}$, we get
$$\hat{X}[\hat{p}] \equiv W'[\hat{p}] + W''[\widetilde{p}],$$
and therefore $f_{\hat{p}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X}))} = 1$. The latter is true because $\hat{X}[\hat{p}] \equiv W'[\hat{p}] \equiv W''[\widetilde{p}] \equiv 0 \pmod{\kappa+1}$.
- assume that $\hat{p} = p_T$ and $\hat{X}$ is a child of T. From the definition of $\mathbf{u}$ and $\mathbf{v}_0$ we get
$$\hat{X} \equiv T^{\mathcal{M}} \pmod{\kappa+1},$$
and therefore $f_{\hat{p}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X}))} = 1$. The above holds because all the variables in the above equation are zero.
- assume that $\hat{p} \in \widetilde{\mathcal{P}}$ and exists a variable W such that $\hat{p} = \widetilde{p}_W$. Let $\hat{X}$ is a child of W in a path $\widetilde{p}_W$. From the definitions of $\mathbf{u}$ and $\mathbf{v}_0$, we get
$$\hat{X}[\hat{p}] \equiv W[\hat{p}] \pmod{\kappa+1},$$
and therefore $f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X})) = 1$. Again, the above holds because all the terms are zero.
- assume that $\hat{p} \in \widetilde{\mathcal{P}}$ and $\hat{X}$ is a variable on this path such that it is neither a starting node of the path $\hat{p}$ nor a child of a starting node on the path $\hat{p}$. Then, from the definitions of $\mathbf{v}_0$ and $\mathbf{u}$, we get
$$\hat{X}[\hat{p}] \equiv \sum_{\hat{X}\in Pa_{\hat{p}}(\hat{X})\setminus\{W\}} \hat{X}[\hat{p}] \pmod{\kappa+1},$$
and therefore $f_{\hat{p}}(\hat{X}[\hat{p}]|Pa_{\hat{p}}(\hat{X})) = 1$. Again, the above holds because all the terms are zero.
- assume $X \in \check{\mathbf{S}}$. Then, from the definitions of $\mathbf{v}_0$ and $\mathbf{u}$, we get
$$X^{\mathcal{M}} \equiv M(X) \pmod{\kappa+1},$$
and consequently $P^{\mathcal{M}}(x^{\mathcal{M}}|Pa_{\mathcal{G}}(X)) = 1$.

Now we consider the case when $U_0^{\mathcal{M}} = \gamma_{j_t}$.

Note that the following term depends only the realization of $\mathbf{U}^{\mathcal{M}}$ and $\mathbf{v}_0^{\mathcal{M}}$.

$$\prod_{X\in\mathbf{S}} P^{\mathcal{M}}(x^{\mathcal{M}} \mid Pa_{\mathcal{G}}(X)) \prod_{U\in\mathbf{U}\setminus\{U_0\}} P^{\mathcal{M}}(u^{\mathcal{M}}).$$

However by the proof of Lemma 6 Kivva et al. [2022] we know that there is no realization of $\mathbf{U}^{\mathcal{M}}$ such that:

- $\mathbb{I}(S) = 0$ for all $S \in \check{\mathbf{S}}$, and
- $U_0^{\mathcal{M}} = \gamma_{j_t}$, and
- $x^{\mathcal{M}} \equiv M(X)(\kappa + 1)$ for all $X \in \check{\mathbf{S}}$. The latter is a necessary condition for $P^{\mathcal{M}}(x|Pa_{\mathcal{G}}(X))$ being non-zero.

To summarize, we showed that for $U_0^{\mathcal{M}} = \gamma_{j_r}$, Equation (70) is non-zero while it is zero for $U_0^{\mathcal{M}} = \gamma_{j_t}$. This implies that $\widetilde{\eta}_{j_r}(\mathbf{v}_0) \neq \widetilde{\eta}_{j_t}(\mathbf{v}_0)$ for $\epsilon = 0$. □

1.5.2 Proof of Lemma 7

Lemma 7. *Let* $\mathbf{S} := Anc_{\mathcal{G}[\mathbf{V}\setminus\mathbf{X}']}(\mathbf{Y}', \mathbf{Z}')$ *and* $\widetilde{\mathbf{D}}$ *is a set of all nodes on the paths in* $\mathcal{P}$ *excluding* $\mathbf{Z}'$*. Then,*

$$P_{\mathbf{x}'}(\widetilde{\mathbf{d}}|\mathbf{s}\setminus\widetilde{\mathbf{d}}) = \frac{Q[\mathbf{S}]}{\sum_{\widetilde{\mathbf{D}}} Q[\mathbf{S}]} = Q[\widetilde{\mathbf{D}}|\mathbf{S}\setminus\widetilde{\mathbf{D}}] \tag{71}$$

is not c-gID from $(\mathbb{A}, \mathcal{G})$*.*

Proof. We will show that $Q[\widetilde{\mathbf{D}}|\mathbf{S}\setminus\widetilde{\mathbf{D}}]$ is not c-gID from $(\mathbb{A}', \mathcal{G})$, where $\mathbb{A}' := \mathbb{A} \cup \{\mathbf{S}_i\}_{i=2}^n$. To this end, we will specify two models $\mathcal{M}_1$ and $\mathcal{M}_2$ such that for each $i \in [0 : m']$ and any $\mathbf{v} \in \widetilde{\mathfrak{X}}(\mathbf{V})$:

$$Q^{\mathcal{M}_1}[\mathbf{A}'_i](\mathbf{v}) = Q^{\mathcal{M}_2}[\mathbf{A}'_i](\mathbf{v}), \tag{72}$$

$$\sum_{\widetilde{\mathbf{D}}} Q^{\mathcal{M}_1}[\mathbf{S}](\mathbf{v}') = \sum_{\widetilde{\mathbf{D}}} Q^{\mathcal{M}_2}[\mathbf{S}](\mathbf{v}'), \tag{73}$$

but there exists $\mathbf{v}_0 \in \widetilde{\mathfrak{X}}(\mathbf{V})$ such that:

$$Q^{\mathcal{M}_1}[\mathbf{S}](\mathbf{v}_0) \neq Q^{\mathcal{M}_2}[\mathbf{S}](\mathbf{v}_0). \tag{74}$$

Note that using Equations (73)-(57) yield

$$Q[\widetilde{\mathbf{D}}|\mathbf{S}\setminus\widetilde{\mathbf{D}}]^{\mathcal{M}_1}(\mathbf{v}_0) \neq Q[\widetilde{\mathbf{D}}|\mathbf{S}\setminus\widetilde{\mathbf{D}}]^{\mathcal{M}_2}(\mathbf{v}_0).$$

This means that $Q[\widetilde{\mathbf{D}}|\mathbf{S}\setminus\widetilde{\mathbf{D}}]$ is not c-gID from $(\mathbb{A}', \mathcal{G})$.

Two this end, we consider two cases.

First case:
Suppose that there exists $i \in [0, m]$, such that $\check{\mathbf{S}} \subset \mathbf{A}_i$. Further we consider models $\widetilde{\mathcal{M}}_1$ and $\widetilde{\mathcal{M}}_2$ constructed in Section 1.5.1. According to the definitions of models $\widetilde{\mathcal{M}}_1$ and $\widetilde{\mathcal{M}}_2$ for any $\mathbf{v} \in \widetilde{\mathfrak{X}}(\mathbf{V})$, and any $i \in [0 : m']$, and any $g \in \{1, 2\}$:

$$Q[\mathbf{A}'_i]^{\widetilde{\mathcal{M}}_g}(\mathbf{v}) := \sum_{U_0^{\mathcal{M}}} P^{\mathcal{M}_g}(u_0^{\mathcal{M}}) \sum_{U_0[\widetilde{\mathcal{P}}]} \prod_{\hat{p}\in\widetilde{\mathcal{P}}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \sum_{\mathbf{U}\setminus\{U_0\}} \prod_{X\in\mathbf{A}'_i} P^{\widetilde{\mathcal{M}}}(x \mid Pa_{\mathcal{G}}(X)) \prod_{U\in\mathbf{U}\setminus\{U_0\}} P^{\widetilde{\mathcal{M}}}(u),$$

$$\sum_{\widetilde{\mathbf{D}}} Q^{\widetilde{\mathcal{M}}_g}[S](\mathbf{v}) := \sum_{U_0^{\mathcal{M}}} P^{\mathcal{M}_g}(u_0^{\mathcal{M}}) \sum_{U_0[\widetilde{\mathcal{P}}]} \prod_{\hat{p}\in\widetilde{\mathcal{P}}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \sum_{\widetilde{\mathbf{D}}} \sum_{\mathbf{U}\setminus\{U_0\}} \prod_{X\in\mathbf{S}} P^{\widetilde{\mathcal{M}}}(x \mid Pa_{\mathcal{G}}(X)) \prod_{U\in\mathbf{U}\setminus\{U_0\}} P^{\widetilde{\mathcal{M}}}(u),$$

$$Q^{\widetilde{\mathcal{M}}_g}[\mathbf{S}](\mathbf{v}) := \sum_{U_0^{\mathcal{M}}} P^{\mathcal{M}_g}(u_0^{\mathcal{M}}) \sum_{U_0[\widetilde{\mathcal{P}}]} \prod_{\hat{p}\in\widetilde{\mathcal{P}}_{U_0}} f_{\hat{p}}(U_0[\hat{p}]) \sum_{\mathbf{U}\setminus\{U_0\}} \prod_{X\in\mathbf{S}} P^{\widetilde{\mathcal{M}}}(x \mid Pa_{\mathcal{G}}(X)) \prod_{U\in\mathbf{U}\setminus\{U_0\}} P^{\widetilde{\mathcal{M}}}(u).$$

We can re-writing the above equations using the notations of $\widetilde{\theta}_{i,j}$, $\widetilde{\phi}_j$, and $\widetilde{\eta}_j$,

$$Q[\mathbf{A}'_i]^{\widetilde{\mathcal{M}}_1}(\mathbf{v}) = \sum_{j=1}^{d} \frac{1}{d}\widetilde{\theta}_{i,j}(\mathbf{v}),$$

$$Q[\mathbf{A}'_i]^{\widetilde{\mathcal{M}}_2}(\mathbf{v}) = \sum_{j=1}^{d} p_j\widetilde{\theta}_{i,j}(\mathbf{v}),$$

$$\sum_{\tilde{\mathbf{D}}} Q[\mathbf{S}]^{\widetilde{\mathcal{M}}_1}(\mathbf{v}) = \sum_{j=1}^{d} \frac{1}{d}\widetilde{\phi}_j(\mathbf{v}[\tilde{\mathbf{D}}^\dagger]),$$

$$\sum_{\tilde{\mathbf{D}}} Q[\mathbf{S}]^{\widetilde{\mathcal{M}}_2}(\mathbf{v}) = \sum_{j=1}^{d} p_j\widetilde{\phi}_j(\mathbf{v}[\tilde{\mathbf{D}}^\dagger]),$$

$$Q[\mathbf{S}]^{\widetilde{\mathcal{M}}_1}(\mathbf{v}) = \sum_{j=1}^{d} \frac{1}{d}\widetilde{\eta}_j(\mathbf{v}),$$

$$Q[\mathbf{S}]^{\widetilde{\mathcal{M}}_2}(\mathbf{v}) = \sum_{j=1}^{d} p_j\widetilde{\eta}_j(\mathbf{v}).$$

The above equations imply the following equations.

$$Q^{\widetilde{M}_2}[\mathbf{A}'_i](\mathbf{v}) - Q^{\widetilde{\mathcal{M}}_1}[\mathbf{A}'_i](\mathbf{v}) = \sum_{j=1}^{d}(p_j - \frac{1}{d})\widetilde{\theta}_{i,j}(\mathbf{v})$$

$$\sum_{\tilde{\mathbf{D}}} Q[\mathbf{S}]^{\widetilde{\mathcal{M}}_2}(\mathbf{v}) - \sum_{\tilde{\mathbf{D}}} Q[\mathbf{S}]^{\widetilde{\mathcal{M}}_1}(\mathbf{v}) = \sum_{j=1}^{d}(p_j - \frac{1}{d})\widetilde{\phi}_j(\mathbf{v}[\tilde{\mathbf{D}}^\dagger])$$

$$Q^{\widetilde{\mathcal{M}}_2}[\mathbf{S}](\mathbf{v}_0) - Q^{\widetilde{\mathcal{M}}_1}[\mathbf{S}](\mathbf{v}_0) = \sum_{j=1}^{d}(p_j - \frac{1}{d})\widetilde{\eta}_j(\mathbf{v}_0)$$

$$\sum_{j=1}^{d} p_j - 1 = \sum_{j=1}^{d}(p_j - \frac{1}{d}).$$

To prove the statement of the lemma it suffices to solve a following system of linear equations over parameters $\{p_j\}_{j=1}^d$ and show that it admits a solution.

$$\sum_{j=1}^{d}(p_j - \frac{1}{d})\widetilde{\theta}_{i,j}(\mathbf{v}) = 0, \quad \forall \mathbf{v} \in \widetilde{\mathfrak{X}}(\mathbf{V}), i \in [0:m'],$$

$$\sum_{j=1}^{d}(p_j - \frac{1}{d})\widetilde{\phi}_j(\mathbf{d}^\dagger) = 0, \quad \forall \mathbf{d}^\dagger \in \widetilde{\mathfrak{X}}(\tilde{\mathbf{D}}^\dagger), i \in [0:m'],$$

$$\sum_{j=1}^{d}(p_j - \frac{1}{d})\widetilde{\eta}_j(\mathbf{v}_0) \neq 0, \quad \exists \mathbf{v}_0 \in \widetilde{\mathfrak{X}}(\mathbf{V}),$$

$$(p_j - \frac{1}{d}) = 0,$$

$$0 < p_j < 1, \quad \forall j \in [1:d].$$

Analogous to the proof of Lemma 1, we use Lemmas 10, 11, and 12 instead of Lemmas 1, 4 and 2 respectively and conclude the result.

Second case:
Suppose that there is no $i \in [0, m]$, such that $\check{\mathbf{S}} \subset \mathbf{A}_i$. This case we solve exactly the same as the **Second case** of the Lemma 1. □

1.6 PROOF OF LEMMA 3

Lemma 3. *Suppose that* $\mathbf{X}$*,* $\mathbf{Y}$ *and* $\mathbf{Z}$ *are disjoint subsets of* $\mathbf{V}$ *in graph* $\mathcal{G}$ *and variables* $Z_1 \in \mathbf{Z}$*,* $Z_2 \in \mathbf{Y} \cup \mathbf{Z}$*, such that there is a directed edge from* Z_1 *to* Z_2 *in* $\mathcal{G}$*. If the causal effect* $P_{\mathbf{x}}(\mathbf{y}|\mathbf{z})$ *is not c-gID from* $(\mathbb{A}, \mathcal{G})$*, then the causal effect* $P_{\mathbf{x}}(\mathbf{y}|\mathbf{z}\backslash\{z_1\})$ *is also not c-gID from* $(\mathbb{A}, \mathcal{G})$*.*

Proof. By the basic probabilistic manipulations, we get

$$P_{\mathbf{x}}(\mathbf{y}|\mathbf{z}) = \frac{P_x(\mathbf{y}, \mathbf{z})}{P_x(\mathbf{z})},$$
$$P_{\mathbf{x}}(\mathbf{y}|\mathbf{z}\backslash\{z_1\}) = \frac{P_x(\mathbf{y}, \mathbf{z}\backslash\{z_1\})}{P_x(\mathbf{z}\backslash\{z_1\})}.$$

Using Markov factorization property in graph $\mathcal{G}$, we have

$$P_{\mathbf{x}}(\mathbf{y}, \mathbf{z}) = \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Y}\cup\mathbf{Z})} \sum_{\mathbf{U}} \prod_{W \in \mathbf{V}\backslash\mathbf{X}} P(w \mid Pa_{\mathcal{G}}(W)) \prod_{U \in \mathbf{U}} P(u),$$
$$P_{\mathbf{x}}(\mathbf{z}) = \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Z})} \sum_{\mathbf{U}} \prod_{W \in \mathbf{V}\backslash\mathbf{X}} P(w \mid Pa_{\mathcal{G}}(W)) \prod_{U \in \mathbf{U}} P(u).$$

And similarly, we have

$$\begin{aligned} P_{\mathbf{x}}(\mathbf{y}, \mathbf{z}\backslash\{Z_1\}) &= \sum_{Z_1} \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Y}\cup\mathbf{Z})} \sum_{\mathbf{U}} \prod_{W \in \mathbf{V}\backslash\mathbf{X}} P(w \mid Pa_{\mathcal{G}}(W)) \prod_{U \in \mathbf{U}} P(u), \\ P_{\mathbf{x}}(\mathbf{z}\backslash\{Z_1\}) &= \sum_{Z_1} \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Z})} \sum_{\mathbf{U}} \prod_{W \in \mathbf{V}\backslash\mathbf{X}} P(w \mid Pa_{\mathcal{G}}(W)) \prod_{U \in \mathbf{U}} P(u). \end{aligned} \tag{75}$$

Since $P_{\mathbf{x}}(\mathbf{y}|\mathbf{z})$ is not c-gID from $(\mathbb{A}, \mathcal{G})$, there exists $\mathcal{M}_1$ and $\mathcal{M}_2$ such that

$$Q^{\mathcal{M}_1}[\mathbf{A}_i](\mathbf{v}) = Q^{\mathcal{M}_2}[\mathbf{A}_i](\mathbf{v}), \ \forall \mathbf{v} \in \mathfrak{X}(\mathbf{V}), \ \forall i \in [0:m],$$

$$P_{\mathbf{x}}^{\mathcal{M}_1}(\mathbf{y}|\mathbf{z}) \neq P_{\mathbf{x}}^{\mathcal{M}_2}(\mathbf{y}|\mathbf{z}), \ \exists \mathbf{x} \in \mathfrak{X}(\mathbf{X}), \ \exists \mathbf{y} \in \mathfrak{X}(\mathbf{Y}).$$

Using $\mathcal{M}_1$ and $\mathcal{M}_2$, we construct two models $\mathcal{M}'_1$ and $\mathcal{M}'_2$. To do so, we first take any surjective function $F\colon \mathfrak{X}(Z_1) \to \{0,1\}$ and define a function $\Psi\colon \{0,1\} \times \mathfrak{X}(Z_1) \to (0,1)$ that satisfies $\Psi(0, z_1) + \Psi(1, z_1) = 1$ for any $z_1 \in \mathfrak{X}(Z_1)$.

For any node S that either belongs to the set of unobserved variables or belongs to $\mathbf{V}\backslash(\{Z_2\} \cup Ch_{\mathcal{G}}(Z_2))$, we define

$$P^{\mathcal{M}'_i}(s|Pa_{\mathcal{G}}(S)) := P^{\mathcal{M}_i}(s|Pa_{\mathcal{G}}(S)), \quad i \in \{1,2\}.$$

The domain of Z_2 in $\mathcal{M}'_i$ is defined as $\mathfrak{X}(Z_2)^{\mathcal{M}} \times \{0,1\}$, where $\mathfrak{X}(Z_2)^{\mathcal{M}}$ is the domain of Z_2 in $\mathcal{M}$ (either $\mathcal{M}_1$ or $\mathcal{M}_2$). For $z_2 \in \mathfrak{X}(Z_2)^{\mathcal{M}}$, $i \in \{1,2\}$, and $k \in \{0,1\}$, we define

$$P^{\mathcal{M}'_i}((z_2, k) \mid Pa_{\mathcal{G}}(Z_2)\backslash\{Z_1\}, z_1) := P^{\mathcal{M}_i}(z_2 \mid Pa_{\mathcal{G}}(Z_2)) \times \Psi(F(z_1) \oplus k, z_1).$$

Due to the property of function Ψ, the above definitions are valid probabilities, i.e., for any realizations $(Pa_{\mathcal{G}}(Z_2), z_1)$, the following holds

$$\sum_{k \in \{0,1\}} \sum_{z_2 \in \mathfrak{X}(Z_2)^{\mathcal{M}}} P^{\mathcal{M}'_i}((z_2, k)|pa(Z_2), z_1) = 1.$$

For each $S \in Ch_{\mathcal{G}}(Z_2)$, we define:

$$P^{\mathcal{M}'_i}(s \mid Pa_{\mathcal{G}}(S)\backslash\{Z_2\}, (z_2, k)) := P^{\mathcal{M}_i}(s \mid Pa_{\mathcal{G}}(S)\backslash\{Z_2\}, z_2), \quad i \in \{1,2\}, k \in \{0,1\}.$$

Next, we show that $Q^{\mathcal{M}'_1}[\mathbf{A}_i](\mathbf{v}) = Q^{\mathcal{M}'_2}[\mathbf{A}_i](\mathbf{v})$ for each $\mathbf{v} \in \mathfrak{X}(\mathbf{V})$ and $i \in [0:m]$. Suppose $\mathbf{v}$ is a realization of $\mathbf{V}$ in $\mathcal{M}'_1$ with realizations z_1 and (z_2, k) for Z_1 and Z_2, respectively. Consider two cases:

- $Z_2 \notin \mathbf{A}_i$: In this case, we have

$$\begin{aligned}
Q^{\mathcal{M}_1'}[\mathbf{A}_i](\mathbf{v}) &= \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{\mathcal{M}_1'}(a \mid Pa_{\mathcal{G}}(A)) \prod_{U \in \mathbf{U}} P^{\mathcal{M}_1'}(u) \\
&= \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{\mathcal{M}_1}(a \mid Pa_{\mathcal{G}}(A)) \prod_{U \in \mathbf{U}} P^{\mathcal{M}_1}(u) = Q^{\mathcal{M}_1}[\mathbf{A}_i](\mathbf{v}) = Q^{\mathcal{M}_2}[\mathbf{A}_i](\mathbf{v}) \\
&= \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{\mathcal{M}_2}(a \mid Pa_{\mathcal{G}}(A)) \prod_{U \in \mathbf{U}} P^{\mathcal{M}_2}(u) \\
&= \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{\mathcal{M}_2'}(a \mid Pa_{\mathcal{G}}(A)) \prod_{U \in \mathbf{U}} P^{\mathcal{M}_2'}(u) \\
&= Q^{\mathcal{M}_2'}[\mathbf{A}_i](\mathbf{v}).
\end{aligned}$$

- $Z_2 \in \mathbf{A}_i$: In this case, we have

$$\begin{aligned}
Q^{\mathcal{M}_1'}[\mathbf{A}_i](\mathbf{v}) &= \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{\mathcal{M}_1'}(a \mid Pa_{\mathcal{G}}(A)) \prod_{U \in \mathbf{U}} P^{\mathcal{M}_1'}(u) \\
&= \Psi\left(F(z_1) \oplus k, z_1\right) \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{\mathcal{M}_1}(a \mid Pa_{\mathcal{G}}(A)) \prod_{U \in \mathbf{U}} P^{\mathcal{M}_1}(u) \\
&= \Psi(F(z_1) \oplus k, z_1) Q^{\mathcal{M}_1}[\mathbf{A}_i](\mathbf{v}) = \Psi(F(z_1) \oplus k, z_1) Q^{\mathcal{M}_2}[\mathbf{A}_i](\mathbf{v}) \\
&= \Psi(F(z_1) \oplus k, z_1) \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{M_2}(a \mid Pa_{\mathcal{G}}(A)) \prod_{U \in \mathbf{U}} P^{M_2}(u) \\
&= \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{\mathcal{M}_2'}(a \mid Pa_{\mathcal{G}}(A))) \prod_{U \in \mathbf{U}} P^{\mathcal{M}_2'}(u) \\
&= Q^{\mathcal{M}_2'}[\mathbf{A}_i](\mathbf{v}).
\end{aligned}$$

Therefore, $Q^{\mathcal{M}_1'}[\mathbf{A}_i](\mathbf{v}) = Q^{\mathcal{M}_2'}[\mathbf{A}_i](\mathbf{v})$ for each $\mathbf{v} \in \mathfrak{X}(\mathbf{V})$ and $i \in [0:m]$.

On the other hand, we know that there exists $\hat{\mathbf{x}} \in \mathfrak{X}(\mathbf{X})^{\mathcal{M}}$, $\hat{\mathbf{y}} \in \mathfrak{X}(\mathbf{Y})^{\mathcal{M}}$ and $\hat{\mathbf{z}} \in \mathfrak{X}(\mathbf{Z})^{\mathcal{M}}$ such that $P_{\hat{\mathbf{x}}}^{\mathcal{M}_1}(\hat{\mathbf{y}}|\hat{\mathbf{z}}) \neq P_{\hat{\mathbf{x}}}^{\mathcal{M}_2}(\hat{\mathbf{y}}|\hat{\mathbf{z}})$. According to Equations (75), we have

$$\begin{aligned}
&P_{\mathbf{x}}^{\mathcal{M}_i'}(\mathbf{y}, \mathbf{z}\backslash\{Z_1\}) = \sum_{z_1 \in \mathfrak{X}(Z_1)} \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Y}\cup\mathbf{Z})} \sum_{\mathbf{U}} P^{\mathcal{M}_i'}((z_2, k)|Pa_{\mathcal{G}}(Z_2)) \prod_{W \in \mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i'}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U \in \mathbf{U}} P(u) \\
&= \sum_{z_1 \in \mathfrak{X}(Z_1)} \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Y}\cup\mathbf{Z})} \sum_{\mathbf{U}} \Psi(F(z_1) \oplus k, z_1) P^{\mathcal{M}_i}(z_2|Pa_{\mathcal{G}}(Z_2)) \prod_{W \in \mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U \in \mathbf{U}} P(u) \\
&= \sum_{z_1 \in \mathfrak{X}(Z_1)} \Psi(F(z_1) \oplus k, z_1) \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Y}\cup\mathbf{Z})} \sum_{\mathbf{U}} P^{\mathcal{M}_i}(z_2|Pa_{\mathcal{G}}(Z_2)) \prod_{W \in \mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U \in \mathbf{U}} P(u) \\
&= \sum_{z_1 \in \mathfrak{X}(Z_1)} \Psi(F(z_1) \oplus k, z_1) P_{\mathbf{x}}^{\mathcal{M}_i}(\mathbf{y}, \mathbf{z}).
\end{aligned}$$

Let us denote $\mathfrak{X}(Z_1) = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$. For $z_1 = \alpha_j$ and $j \in [1:n]$, we also denote

$$\begin{aligned}
&\psi_j := \Psi(F(\alpha_j) \oplus 0, \alpha_j), \\
&\beta_j^{\mathcal{M}_i} := P_{\hat{\mathbf{x}}}^{\mathcal{M}_i}(\hat{\mathbf{y}}, \hat{\mathbf{z}}[\mathbf{Z}\backslash\{Z_1\}], \alpha_j).
\end{aligned}$$

This leads to

$$P_{\mathbf{x}}^{\mathcal{M}_i}(\mathbf{y}, \mathbf{z}\backslash\{Z_1\}) = \sum_{j=1}^{n} \psi_j \beta_j^{\mathcal{M}_i},$$

for realizations $\mathbf{y}$ consistent with $\hat{\mathbf{y}}$, realization $\mathbf{x}$ consistent with $\hat{\mathbf{x}}$, $\mathbf{z}$ consistent with $\hat{\mathbf{z}}$, and $Z_2 = (z_2, k)$ consistent with $\hat{\mathbf{y}} \cup \hat{\mathbf{z}}$ and $k = 0$. Recall that ψ_j is a real number from the interval $(0, 1)$. Note that ψ_j is independent from any other ψ_l for $l \neq j$.

Next, we consider two cases:

- Assume that $Z_2 \in \mathbf{Z}$. In this case, we have

$$
\begin{aligned}
P_{\mathbf{x}}^{\mathcal{M}_i'}(\mathbf{z}\backslash\{Z_1\}) &= \sum_{z_1 \in \mathfrak{X}(Z_1)} \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Z})} \sum_{\mathbf{U}} P^{\mathcal{M}_i'}((z_2,k)|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i'}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \sum_{z_1 \in \mathfrak{X}(Z_1)} \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Z})} \sum_{\mathbf{U}} \Psi(F(z_1)\oplus k, z_1) P^{\mathcal{M}_i}(z_2|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \sum_{z_1 \in \mathfrak{X}(Z_1)} \Psi(F(z_1)\oplus k, z_1) \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Z})} \sum_{\mathbf{U}} P^{\mathcal{M}_1}(z_2|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \sum_{z_1 \in \mathfrak{X}(Z_1)} \Psi(F(z_1)\oplus k, z_1) P^{\mathcal{M}_i}(\mathbf{z}).
\end{aligned}
$$

For $j \in [1:n]$ and $z_1 = \alpha_j$, we denote

$$
\gamma_j^{\mathcal{M}_i} := P_{\hat{\mathbf{x}}}^{\mathcal{M}_i}(\hat{\mathbf{z}}[\mathbf{Z}\backslash\{Z_1\}], \alpha_j),
$$

which leads to

$$
P_{\mathbf{x}}^{\mathcal{M}_i}(\mathbf{z}\backslash\{Z_1\}) = \sum_{j=1}^{n} \psi_j \gamma_j^{\mathcal{M}_i},
$$

for realizations $\mathbf{y}$ consistent with $\hat{\mathbf{y}}$, realization $\mathbf{x}$ consistent with $\hat{\mathbf{x}}$, $\mathbf{z}$ consistent with $\hat{\mathbf{z}}$, and $Z_2 = (z_2, k)$ consistent with $\hat{\mathbf{y}} \cup \hat{\mathbf{z}}$ and $k = 0$. Thus, for such realizations, we have

$$
P_{\mathbf{x}}^{\mathcal{M}_i}(\hat{\mathbf{y}}|\hat{\mathbf{z}}\backslash\{Z_1\}) = \frac{\sum_{j=1}^{n} \psi_j \beta_j^{\mathcal{M}_i}}{\sum_{j=1}^{n} \psi_j \gamma_j^{\mathcal{M}_i}}.
$$

By the assumption of the lemma, there exists $j \in [1:n]$ such that

$$
\frac{\beta_j^{\mathcal{M}_1}}{\gamma_j^{\mathcal{M}_1}} \neq \frac{\beta_j^{\mathcal{M}_2}}{\gamma_j^{\mathcal{M}_2}},
$$

or equivalently,

$$
\beta_j^{\mathcal{M}_1}\gamma_j^{\mathcal{M}_2} \neq \beta_j^{\mathcal{M}_2}\gamma_j^{\mathcal{M}_1}.
$$

Without loss of generality, we assume that the aforementioned inequality holds for $j = 1$. Next, we prove that there exists a parameters $\{\psi_j\}_{j=1}^{n}$ such that

$$
\frac{\sum_{j=1}^{n} \psi_j \beta_j^{\mathcal{M}_1}}{\sum_{j=1}^{n} \psi_j \gamma_j^{\mathcal{M}_1}} \neq \frac{\sum_{j=1}^{n} \psi_j \beta_j^{\mathcal{M}_2}}{\sum_{j=1}^{n} \psi_j \gamma_j^{\mathcal{M}_2}},
$$

or equivalently,

$$
\sum_{j=1}^{n} \psi_j \beta_j^{\mathcal{M}_1} \sum_{j=1}^{n} \psi_j \gamma_j^{\mathcal{M}_2} - \sum_{j=1}^{n} \psi_j \beta_j^{\mathcal{M}_2} \sum_{j=1}^{n} \psi_j \gamma_j^{\mathcal{M}_1} \neq 0.
$$

Note that the left hand side is a quadratic equation with respect to parameter ψ_1, e.g.,

$$
(\beta_1^{\mathcal{M}_1}\gamma_1^{\mathcal{M}_2} - \beta_1^{\mathcal{M}_2}\gamma_1^{\mathcal{M}_1})\psi_1^2
$$

Since $\beta_1^{\mathcal{M}_1}\gamma_1^{\mathcal{M}_2} - \beta_1^{\mathcal{M}_2}\gamma_1^{\mathcal{M}_1} \neq 0$, then we can find $\{\psi_j\}_{j=1}^{n}$, such that

$$
\sum_{j=1}^{n} \psi_j \beta_j^{\mathcal{M}_1} \sum_{j=1}^{n} \psi_j \gamma_j^{\mathcal{M}_2} - \sum_{j=1}^{n} \psi_j \beta_j^{\mathcal{M}_2} \sum_{j=1}^{n} \psi_j \gamma_j^{\mathcal{M}_1} \neq 0.
$$

This is possible because $\psi_i \in (0, 1)$. This concludes the proof of the lemma for this case.

- Assume that $Z_2 \in \mathbf{Y}$. Suppose that $P^{\mathcal{M}_1}_{\mathbf{x}}(\mathbf{y}, z_1|\mathbf{z}\backslash\{Z_1\}) = P^{\mathcal{M}_2}_{\mathbf{x}}(\mathbf{y}, z_1|\mathbf{z}\backslash\{Z_1\})$ for all $\mathbf{x} \in \mathfrak{X}(\mathbf{X})$, $\mathbf{y} \in \mathfrak{X}(\mathbf{Y})$ and $\mathbf{z} \in \mathfrak{X}(\mathbf{Z})$. Then,

$$P^{\mathcal{M}_1}_{\mathbf{x}}(\mathbf{y}|\mathbf{z}) = \frac{P^{\mathcal{M}_1}(\mathbf{y}, z_1|\mathbf{z}\backslash\{Z_1\})}{P^{\mathcal{M}_1}_{\mathbf{x}}(z_1|\mathbf{z}\backslash\{Z_1\})} = \frac{P^{\mathcal{M}_1}(\mathbf{y}, z_1|\mathbf{z}\backslash\{Z_1\})}{P^{\mathcal{M}_2}_{\mathbf{x}}(z_1|\mathbf{z}\backslash\{Z_1\})} = P^{\mathcal{M}_2}_{\mathbf{x}}(\mathbf{y}|\mathbf{z}).$$

This is impossible as $P^{\mathcal{M}_1}_{\hat{\mathbf{x}}}(\hat{\mathbf{y}}|\hat{\mathbf{z}}) \neq P^{\mathcal{M}_2}_{\hat{\mathbf{x}}}(\hat{\mathbf{y}}|\hat{\mathbf{z}})$. Thus, there exist $\hat{x}' \in \mathfrak{X}(\mathbf{X})$, $\hat{y}' \in \mathfrak{X}(\mathbf{Y})$, and $\hat{z}' \in \mathfrak{X}(\mathbf{Z})$, such that

$$P^{\mathcal{M}_1}_{\hat{\mathbf{x}}'}(\hat{\mathbf{y}}', \hat{z}_1'|\hat{\mathbf{z}}'\backslash\{Z_1\}) \neq P^{\mathcal{M}_2}_{\hat{\mathbf{x}}'}(\hat{\mathbf{y}}', \hat{z}_1'|\hat{\mathbf{z}}'\backslash\{Z_1\}).$$

On the other hand, we have

$$\begin{aligned}
P^{\mathcal{M}_i'}_{\mathbf{x}}(\mathbf{y}, \mathbf{z}) &= \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Y}\cup\mathbf{Z})} \sum_{\mathbf{U}} P^{\mathcal{M}_i'}((z_2, k)|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i'}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Y}\cup\mathbf{Z})} \sum_{\mathbf{U}} \Psi(F(z_1) \oplus k, z_1) P^{\mathcal{M}_i}(z_2|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \Psi(F(z_1) \oplus k, z_1) \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Y}\cup\mathbf{Z})} \sum_{\mathbf{U}} P^{\mathcal{M}_i}(z_2|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \Psi(F(z_1) \oplus k, z_1) P^{\mathcal{M}_i}_{\mathbf{x}}(\mathbf{y}, \mathbf{z}).
\end{aligned}$$

For $j \in [1:n]$, we define

$$\beta'^{\mathcal{M}_i}_j := P^{\mathcal{M}_i}_{\hat{\mathbf{x}}}(\hat{\mathbf{y}}', \hat{\mathbf{z}}'[\mathbf{Z}\backslash\{Z_1\}], \alpha_j).$$

Suppose $m \in [1:n]$, where $z_1 = \alpha_m$ and it is consistent with $\hat{\mathbf{z}}'$. We assign $k = 0$ and denote

$$\beta'^{\ \mathcal{M}_i}_m := P^{\mathcal{M}_i}_{\hat{x}'}(\hat{\mathbf{y}}', \hat{\mathbf{z}}').$$

This results in

$$P^{\mathcal{M}_i'}_{\hat{\mathbf{x}}'}(\hat{\mathbf{y}}', \hat{\mathbf{z}}') = \beta'^{\ \mathcal{M}_i}_m \psi_m.$$

We also have

$$\begin{aligned}
P^{\mathcal{M}_i'}_{\mathbf{x}}(\mathbf{z}\backslash\{Z_1\}) &= \sum_{z_1\in\mathfrak{X}(Z_1)} \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Z})} \sum_{\mathbf{U}} P^{\mathcal{M}_i'}((z_2, k)|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i'}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \sum_{z_1\in\mathfrak{X}(Z_1)} \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Z})} \sum_{\mathbf{U}} \Psi(F(z_1) \oplus k, z_1) P^{\mathcal{M}_i}(z_2|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \sum_{z_1\in\mathfrak{X}(Z_1)} \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Z})} \sum_{\mathbf{U}} \sum_{k\in\{0,1\}} \Psi(F(z_1) \oplus k, z_1) P^{\mathcal{M}_i}(z_2|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \sum_{z_1\in\mathfrak{X}(Z_1)} P^{\mathcal{M}_i}(\mathbf{z}).
\end{aligned}$$

For $j \in [1:n]$ and $z_1 = \alpha_j$, we denote

$$\gamma'^{\mathcal{M}_i}_j := P^{\mathcal{M}_i}_{\hat{\mathbf{x}}}(\hat{\mathbf{z}}'[\mathbf{Z}\backslash\{Z_1\}]),$$

and from the above equation, we get

$$P^{\mathcal{M}_i}_{\mathbf{x}}(\mathbf{z}\backslash\{Z_1\}) = \sum_{j=1}^{n} \gamma^{\mathcal{M}_i}_j,$$

for realizations $\mathbf{y}$ consistent with $\hat{\mathbf{y}}'$, realization $\mathbf{x}$ consistent with $\hat{\mathbf{x}}'$, $\mathbf{z}$ consistent with $\hat{\mathbf{z}}'$, and $Z_2 = (z_2, k)$ consistent with $\hat{\mathbf{y}}' \cup \hat{\mathbf{z}}'$ and $k = 0$. We have

$$P^{\mathcal{M}_i}_{\mathbf{x}}(\hat{\mathbf{y}}|\hat{\mathbf{z}}\backslash\{Z_1\}) = \frac{\sum_{j=1}^{n} \psi_j \beta'^{\mathcal{M}_i}_j}{\sum_{j=1}^{n} \gamma^{\mathcal{M}_i}_j}.$$

By the assumption of the lemma, we have

$$\frac{{\beta'_m}^{\mathcal{M}_1}}{\sum_{j=1}^{n}\gamma_j^{\mathcal{M}_1}} \neq \frac{{\beta'_m}^{\mathcal{M}_2}}{\sum_{j=1}^{n}\gamma_j^{\mathcal{M}_2}}.$$

Next, we prove that there exists a set of parameters $\{\psi_j\}_{j=1}^{n}$, such that

$$\frac{\sum_{j=1}^{n}\psi_j {\beta'_j}^{\mathcal{M}_1}}{\sum_{j=1}^{n}\gamma_j^{\mathcal{M}_1}} \neq \frac{\sum_{j=1}^{n}\psi_j {\beta'_j}^{\mathcal{M}_2}}{\sum_{j=1}^{n}\gamma_j^{\mathcal{M}_2}}$$

or equivalently,

$$\frac{\sum_{j=1}^{n}\psi_j {\beta'_j}^{\mathcal{M}_1}}{\sum_{j=1}^{n}\gamma_j^{\mathcal{M}_1}} - \frac{\sum_{j=1}^{n}\psi_j {\beta'_j}^{\mathcal{M}_2}}{\sum_{j=1}^{n}\gamma_j^{\mathcal{M}_2}} \neq 0.$$

Note that left hand side of the above equation is linear with respect to parameter ψ_m with the following coefficient,

$$\frac{{\beta'_m}^{\mathcal{M}_1}}{\sum_{j=1}^{n}\gamma_j^{\mathcal{M}_1}} - \frac{{\beta'_m}^{\mathcal{M}_2}}{\sum_{j=1}^{n}\gamma_j^{\mathcal{M}_2}} \neq 0.$$

This ensures that we can can find a realization of $\{\psi_j\}_{j=1}^{n}$, such that

$$\frac{\sum_{j=1}^{n}\psi_j {\beta'_j}^{\mathcal{M}_1}}{\sum_{j=1}^{n}\gamma_j^{\mathcal{M}_1}} - \frac{\sum_{j=1}^{n}\psi_j {\beta'_j}^{\mathcal{M}_2}}{\sum_{j=1}^{n}\gamma_j^{\mathcal{M}_2}} \neq 0.$$

This concludes the proof of the lemma for second case.

□

1.7 PROOF OF LEMMA 6

Lemma 6. *Suppose that $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ are disjoint subsets of $\mathbf{V}$ in graph $\mathcal{G}$ and variables $Z_1 \in \mathbf{Y}$, $Z_2 \in \mathbf{Y} \cup \mathbf{Z}$, such that there is a directed edge from Z_1 to Z_2 in $\mathcal{G}$. If the causal effect $P_{\mathbf{x}}(\mathbf{y}|\mathbf{z})$ is not c-gID from $(\mathbb{A}, \mathcal{G})$, then the causal effect $P_{\mathbf{x}}(\mathbf{y}\backslash\{z_1\}|\mathbf{z})$ is also not c-gID from $(\mathbb{A}, \mathcal{G})$.*

Proof. By the basic probabilistic manipulations, we get

$$P_{\mathbf{x}}(\mathbf{y}|\mathbf{z}) = \frac{P_x(\mathbf{y},\mathbf{z})}{P_x(\mathbf{z})},$$
$$P_{\mathbf{x}}(\mathbf{y}\backslash\{z_1\}|\mathbf{z}) = \frac{P_x(\mathbf{y}\backslash\{z_1\},\mathbf{z})}{P_x(\mathbf{z})}.$$

Using Markov factorization property in graph $\mathcal{G}$, $P_{\mathbf{x}}(\mathbf{y})$ will be

$$P_{\mathbf{x}}(\mathbf{y},\mathbf{z}) = \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Y}\cup\mathbf{Z})}\sum_{\mathbf{U}}\prod_{W\in\mathbf{V}\backslash\mathbf{X}} P(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u),$$
$$P_{\mathbf{x}}(\mathbf{z}) = \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Z})}\sum_{\mathbf{U}}\prod_{W\in\mathbf{V}\backslash\mathbf{X}} P(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u).$$

And similarly, we have

$$\begin{aligned} P_{\mathbf{x}}(\mathbf{y}\backslash\{Z_1\},\mathbf{z}) &= \sum_{Z_1}\sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Y}\cup\mathbf{Z})}\sum_{\mathbf{U}}\prod_{W\in\mathbf{V}\backslash\mathbf{X}} P(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u), \\ P_{\mathbf{x}}(\mathbf{z}) &= \sum_{Z_1}\sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Z})}\sum_{\mathbf{U}}\prod_{W\in\mathbf{V}\backslash\mathbf{X}} P(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u). \end{aligned} \tag{76}$$

Since $P_{\mathbf{x}}(\mathbf{y}|\mathbf{z})$ is not gID from $(\mathbb{A}, \mathcal{G})$, there exists $\mathcal{M}_1$ and $\mathcal{M}_2$ such that

$$Q^{\mathcal{M}_1}[\mathbf{A}_i](\mathbf{v}) = Q^{\mathcal{M}_2}[\mathbf{A}_i](\mathbf{v}),\ \forall \mathbf{v} \in \mathfrak{X}(\mathbf{V}),\ \forall i \in [0:m],$$

$$P^{\mathcal{M}_1}_{\mathbf{x}}(\mathbf{y}|\mathbf{z}) \neq P^{\mathcal{M}_2}_{\mathbf{x}}(\mathbf{y}|\mathbf{z}),\ \exists \mathbf{x} \in \mathfrak{X}(\mathbf{X}),\ \exists \mathbf{y} \in \mathfrak{X}(\mathbf{Y}).$$

Using $\mathcal{M}_1$ and $\mathcal{M}_2$, we construct two models $\mathcal{M}'_1$ and $\mathcal{M}'_2$. Define a surjective function $F\colon\ \mathfrak{X}(Z_1) \to \{0,1\}$ and a function $\Psi\colon\ \{0,1\} \times \mathfrak{X}(Z_1) \to (0,1)$ such that $\Psi(0, z_1) + \Psi(1, z_1) = 1$ for each $z_1 \in \mathfrak{X}(Z_1)$.

For any node S which is either unobserved or in $\mathbf{V}\backslash(\{Z_2\} \cup Ch_{\mathcal{G}}(Z_2))$, we define

$$P^{\mathcal{M}'_i}(s|Pa_{\mathcal{G}}(S)) = P^{\mathcal{M}_i}(s|Pa_{\mathcal{G}}(S)),$$

where $i \in \{1,2\}$. The domain of Z_2 in $\mathcal{M}'_i$ is defined as $\mathfrak{X}(Z_2)^{\mathcal{M}} \times \{0,1\}$, where $\mathfrak{X}(Z_2)^{\mathcal{M}}$ is the domain of Z_2 in $\mathcal{M}$ (either $\mathcal{M}_1$ or $\mathcal{M}_2$). For $z_2 \in \mathfrak{X}(Z_2)^{\mathcal{M}}$, $i \in \{0,1\}$, and $k \in \{0,1\}$, we define

$$P^{\mathcal{M}'_i}((z_2,k) \mid Pa_{\mathcal{G}}(Z_2)\backslash\{Z_1\}, z_1) = P^{\mathcal{M}_i}(z_2 \mid Pa_{\mathcal{G}}(Z_2))\Psi(F(z_1) \oplus k, z_1).$$

Moreover, for a fixed realization $(Pa_{\mathcal{G}}(Z_2), z_1)$, we have

$$\sum_{k \in \{0,1\}} \sum_{z_2 \in \mathfrak{X}(Z_2)^{\mathcal{M}}} P^{\mathcal{M}'_i}((z_2,k)|pa(Z_2), z_1) = 1.$$

For each $S \in Ch_{\mathcal{G}}(Z_2)$, we define:

$$P^{\mathcal{M}'_i}(s \mid Pa_{\mathcal{G}}(S)\backslash\{Z_2\}, (z_2,k)) = P^{\mathcal{M}_i}(s \mid Pa_{\mathcal{G}}(S)\backslash\{Z_2\}, z_2).$$

Next, we show that $Q^{\mathcal{M}'_1}[\mathbf{A}_i](\mathbf{v}) = Q^{\mathcal{M}'_2}[\mathbf{A}_i](\mathbf{v})$ for each $\mathbf{v} \in \mathfrak{X}(\mathbf{V})$ and $i \in [0:m]$. Suppose $\mathbf{v}$ is a realization of $\mathbf{V}$ in $\mathcal{M}'_1$ with realizations z_1 and (z_2,k) for Z_1 and Z_2, respectively. Consider two cases:

- $Z_2 \notin \mathbf{A}_i$: In this case, we have

$$\begin{aligned}
Q^{\mathcal{M}'_1}[\mathbf{A}_i](\mathbf{v}) &= \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{\mathcal{M}'_1}(a \mid Pa_{\mathcal{G}}(A)) \prod_{U \in \mathbf{U}} P^{\mathcal{M}'_1}(u) \\
&= \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{\mathcal{M}_1}(a \mid Pa_{\mathcal{G}}(A)) \prod_{U \in \mathbf{U}} P^{\mathcal{M}_1}(u) = Q^{\mathcal{M}_1}[\mathbf{A}_i](\mathbf{v}) = Q^{\mathcal{M}_2}[\mathbf{A}_i](\mathbf{v}) \\
&= \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{\mathcal{M}_2}(a \mid Pa_{\mathcal{G}}(A)) \prod_{U \in \mathbf{U}} P^{\mathcal{M}_2}(u) \\
&= \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{\mathcal{M}'_2}(a \mid Pa_{\mathcal{G}}(A)) \prod_{U \in \mathbf{U}} P^{\mathcal{M}'_2}(u) \\
&= Q^{\mathcal{M}'_2}[\mathbf{A}_i](\mathbf{v}).
\end{aligned}$$

- $Z_2 \in \mathbf{A}_i$: In this case, we have

$$\begin{aligned}
Q^{\mathcal{M}'_1}[\mathbf{A}_i](\mathbf{v}) &= \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{\mathcal{M}'_1}(a \mid Pa_{\mathcal{G}}(A)) \prod_{U \in \mathbf{U}} P^{\mathcal{M}'_1}(u) \\
&= \Psi\left(F(z_1) \oplus k, z_1\right) \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{\mathcal{M}_1}(a \mid Pa_{\mathcal{G}}(A)) \prod_{U \in \mathbf{U}} P^{\mathcal{M}_1}(u) \\
&= \Psi(F(z_1) \oplus k, z_1) Q^{\mathcal{M}_1}[\mathbf{A}_i](\mathbf{v}) = \Psi(F(z_1) \oplus k, z_1) Q^{\mathcal{M}_2}[\mathbf{A}_i](\mathbf{v}) \\
&= \Psi(F(z_1) \oplus k, z_1) \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{M_2}(a \mid Pa_{\mathcal{G}}(A)) \prod_{U \in \mathbf{U}} P^{M_2}(u) \\
&= \sum_{\mathbf{U}} \prod_{A \in \mathbf{A}_i} P^{\mathcal{M}'_2}(a \mid Pa_{\mathcal{G}}(A))) \prod_{U \in \mathbf{U}} P^{\mathcal{M}'_2}(u) \\
&= Q^{\mathcal{M}'_2}[\mathbf{A}_i](\mathbf{v}).
\end{aligned}$$

Therefore, $Q^{\mathcal{M}'_1}[\mathbf{A}_i](\mathbf{v}) = Q^{\mathcal{M}'_2}[\mathbf{A}_i](\mathbf{v})$ for each $\mathbf{v} \in \mathfrak{X}(\mathbf{V})$ and $i \in [0:m]$.

On the other hand, we know that there exists $\hat{\mathbf{x}} \in \mathfrak{X}(\mathbf{X})^{\mathcal{M}}$, $\hat{\mathbf{y}} \in \mathfrak{X}(\mathbf{Y})^{\mathcal{M}}$ and $\hat{\mathbf{z}} \in \mathfrak{X}(\mathbf{Z})^{\mathcal{M}}$ such that $P^{\mathcal{M}_1}_{\hat{\mathbf{x}}}(\hat{\mathbf{y}}|\hat{\mathbf{z}}) \neq P^{\mathcal{M}_2}_{\hat{\mathbf{x}}}(\hat{\mathbf{y}}|\hat{\mathbf{z}})$.

According to Equations (76), we have

$$\begin{aligned}
P^{\mathcal{M}'_i}_{\mathbf{x}}(\mathbf{y}\backslash\{Z_1\},\mathbf{z}) &= \sum_{z_1\in\mathfrak{X}(Z_1)}\sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Y}\cup\mathbf{Z})}\sum_{\mathbf{U}} P^{\mathcal{M}'_i}((z_2,k)|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}'_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \sum_{z_1\in\mathfrak{X}(Z_1)}\sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Y}\cup\mathbf{Z})}\sum_{\mathbf{U}} \Psi(F(z_1)\oplus k, z_1) P^{\mathcal{M}_i}(z_2|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \sum_{z_1\in\mathfrak{X}(Z_1)} \Psi(F(z_1)\oplus k, z_1) \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Y}\cup\mathbf{Z})}\sum_{\mathbf{U}} P^{\mathcal{M}_i}(z_2|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \sum_{z_1\in\mathfrak{X}(Z_1)} \Psi(F(z_1)\oplus k, z_1) P^{\mathcal{M}_i}_{\mathbf{x}}(\mathbf{y},\mathbf{z}).
\end{aligned}$$

Let us denote $\mathfrak{X}(Z_1) = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$. For $z_1 = \alpha_j$ and $j \in [1:n]$, we also denote

$$\begin{aligned}
&\psi_j = \psi(F(\alpha_j)\oplus 0, \alpha_j), \\
&P^{\mathcal{M}_i}_{\hat{\mathbf{x}}}(\hat{\mathbf{y}}[\mathbf{Y}\backslash\{Z_1\}], \hat{\mathbf{z}}, \alpha_j) = \beta^{\mathcal{M}_i}_j.
\end{aligned}$$

For $Z_2 = (\hat{z}[Z_2], 0)$ we have:

$$P^{\mathcal{M}_i}_{\hat{\mathbf{x}}}(\hat{\mathbf{y}}\backslash\{Z_1\}, \hat{\mathbf{z}}) = \sum_{j=1}^{n} \psi_j \beta^{\mathcal{M}_i}_j.$$

Recall that ψ_j is a real number from the interval $(0,1)$. Note that ψ_j is independent from any other ψ_l for $l \neq j$.

Next, we consider two cases:

- Assume that $Z_2 \in \mathbf{Z}$. In this case, we have

$$\begin{aligned}
P^{\mathcal{M}'_i}_{\mathbf{x}}(\mathbf{z}) &= \sum_{z_1\in\mathfrak{X}(Z_1)}\sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Z})}\sum_{\mathbf{U}} P^{\mathcal{M}'_i}((z_2,k)|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}'_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \sum_{z_1\in\mathfrak{X}(Z_1)}\sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Z})}\sum_{\mathbf{U}} \Psi(F(z_1)\oplus k, z_1) P^{\mathcal{M}_i}(z_2|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \sum_{z_1\in\mathfrak{X}(Z_1)} \Psi(F(z_1)\oplus k, z_1) \sum_{\mathbf{V}\backslash(\mathbf{X}\cup\mathbf{Z})}\sum_{\mathbf{U}} P^{\mathcal{M}_1}(z_2|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\backslash(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \sum_{z_1\in\mathfrak{X}(Z_1)} \Psi(F(z_1)\oplus k, z_1) P^{\mathcal{M}_i}(\mathbf{z}).
\end{aligned}$$

  We denote

$$P^{\mathcal{M}_i}_{\hat{\mathbf{x}}}(\hat{\mathbf{z}}) = \gamma^{\mathcal{M}_i},$$

  which leads to

$$P^{\mathcal{M}_i}_{\mathbf{x}}(\mathbf{z}) = \sum_{j=1}^{n} \psi_j \gamma^{\mathcal{M}_i},$$

  for $Z_2 = (\hat{\mathbf{z}}[Z_2], 0)$. Thus,

$$P^{\mathcal{M}_i}_{\mathbf{x}}(\hat{\mathbf{y}}\backslash\{Z_1\}|\hat{\mathbf{z}}) = \frac{\sum_{j=1}^{n}\psi_j\beta^{\mathcal{M}_i}_j}{\sum_{j=1}^{n}\psi_j\gamma^{\mathcal{M}_i}}.$$

  By the assumption of the lemma, there exists $j \in [1:n]$ such that

$$\frac{\beta^{\mathcal{M}_1}_j}{\gamma^{\mathcal{M}_1}} \neq \frac{\beta^{\mathcal{M}_2}_j}{\gamma^{\mathcal{M}_2}},$$

or equivalently,

$$\beta_j^{\mathcal{M}_1}\gamma^{\mathcal{M}_2} \neq \beta_j^{\mathcal{M}_2}\gamma^{\mathcal{M}_1}.$$

Without loss of generality, we assume that the aforementioned inequality holds for $j = 1$. Next, we prove that there exists a parameters $\{\psi_j\}_{j=1}^n$ such that

$$\frac{\sum_{j=1}^n \psi_j \beta_j^{\mathcal{M}_1}}{\sum_{j=1}^n \psi_j \gamma^{\mathcal{M}_1}} \neq \frac{\sum_{j=1}^n \psi_j \beta_j^{\mathcal{M}_2}}{\sum_{j=1}^n \psi_j \gamma^{\mathcal{M}_2}},$$

or equivalently,

$$\sum_{j=1}^n \psi_j \beta_j^{\mathcal{M}_1} \sum_{j=1}^n \psi_j \gamma^{\mathcal{M}_2} - \sum_{j=1}^n \psi_j \beta_j^{\mathcal{M}_2} \sum_{j=1}^n \psi_j \gamma^{\mathcal{M}_1} \neq 0.$$

Note that the left hand side is a quadratic equation with parameter ψ_1 that contains the following term

$$(\beta_1^{\mathcal{M}_1}\gamma^{\mathcal{M}_2} - \beta_1^{\mathcal{M}_2}\gamma^{\mathcal{M}_1})\psi_1^2$$

Since $\beta_1^{\mathcal{M}_1}\gamma^{\mathcal{M}_2} - \beta_1^{\mathcal{M}_2}\gamma^{\mathcal{M}_1} \neq 0$, then we can can find realization of $\{\psi_j\}_{j=1}^n$, such that

$$\sum_{j=1}^n \psi_j \beta_j^{\mathcal{M}_1} \sum_{j=1}^n \psi_j \gamma^{\mathcal{M}_2} - \sum_{j=1}^n \psi_j \beta_j^{\mathcal{M}_2} \sum_{j=1}^n \psi_j \gamma^{\mathcal{M}_1} \neq 0.$$

which concludes the proof of the lemma for this case.

- Assume that $Z_2 \in \mathbf{Y}$. In this case we have:

$$\begin{aligned}
P_{\mathbf{x}}^{\mathcal{M}'_i}(\mathbf{z}) &= \sum_{\mathbf{V}\setminus(\mathbf{X}\cup\mathbf{Z})} \sum_{\mathbf{U}} P^{\mathcal{M}'_i}((z_2,k)|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\setminus(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}'_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \sum_{\mathbf{V}\setminus(\mathbf{X}\cup\mathbf{Z})} \sum_{\mathbf{U}} \Psi(F(z_1)\oplus k, z_1) P^{\mathcal{M}_i}(z_2|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\setminus(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= \sum_{\mathbf{V}\setminus(\mathbf{X}\cup\mathbf{Z})} \sum_{\mathbf{U}} \sum_{k\in\{0,1\}} \Psi(F(z_1)\oplus k, z_1) P^{\mathcal{M}_i}(z_2|Pa_{\mathcal{G}}(Z_2)) \prod_{W\in\mathbf{V}\setminus(\mathbf{X}\cup\{Z_2\})} P^{\mathcal{M}_i}(w \mid Pa_{\mathcal{G}}(W)) \prod_{U\in\mathbf{U}} P(u) \\
&= P^{\mathcal{M}_i}(\mathbf{z}).
\end{aligned}$$

  We denote

$$P_{\hat{\mathbf{x}}}^{\mathcal{M}_i}(\hat{\mathbf{z}}) = \gamma^{\mathcal{M}_i}.$$

  Thus,

$$P_{\mathbf{x}}^{\mathcal{M}_i}(\hat{\mathbf{y}}\setminus\{Z_1\}|\hat{\mathbf{z}}) = \frac{\sum_{j=1}^n \psi_j {\beta_j}^{\mathcal{M}_i}}{\gamma^{\mathcal{M}_i}}.$$

  By the assumption of the lemma exist $m \in [1:n]$ such that

$$\frac{{\beta_m}^{\mathcal{M}_1}}{\gamma_j^{\mathcal{M}_1}} \neq \frac{{\beta_m}^{\mathcal{M}_2}}{\gamma^{\mathcal{M}_2}}.$$

  Next, we prove that there exists a set of parameters $\{\psi_j\}_{j=1}^n$, such that

$$\frac{\sum_{j=1}^n \psi_j {\beta_j}^{\mathcal{M}_1}}{\gamma^{\mathcal{M}_1}} \neq \frac{\sum_{j=1}^n \psi_j {\beta_j}^{\mathcal{M}_2}}{\gamma^{\mathcal{M}_2}}$$

  or equivalently,

$$\frac{\sum_{j=1}^n \psi_j {\beta'_j}^{\mathcal{M}_1}}{\gamma^{\mathcal{M}_1}} - \frac{\sum_{j=1}^n \psi_j {\beta'_j}^{\mathcal{M}_2}}{\gamma^{\mathcal{M}_2}} \neq 0.$$

  Note that left hand side of the above equation is linear with respect to parameter ψ_m with the following coefficient,

$$\frac{{\beta'_m}^{\mathcal{M}_1}}{\gamma^{\mathcal{M}_1}} - \frac{{\beta'_m}^{\mathcal{M}_2}}{\gamma^{\mathcal{M}_2}} \neq 0.$$

This ensures that we can can find a realization of $\{\psi_j\}_{j=1}^n$, such that

$$\frac{\sum_{j=1}^n \psi_j {\beta'_j}^{\mathcal{M}_1}}{\sum_{j=1}^n \gamma^{\mathcal{M}_1}} - \frac{\sum_{j=1}^n \psi_j {\beta'_j}^{\mathcal{M}_2}}{\sum_{j=1}^n \gamma^{\mathcal{M}_2}} \neq 0.$$

This concludes the proof of the lemma for the second case.

□

2 ON THE POSITIVITY ASSUMPTION IN THE LITERATURE

As it was pointed out in Kivva et al. [2022], positivity assumption is crucial for proving the completeness part. More precisely, the completeness of an algorithm means that if the algorithm does not compute a given conditional causal effect, then it cannot be computed uniquely by any other algorithms. To prove the completeness, two models $\mathcal{M}_1$ and $\mathcal{M}_2$ are constructed such that they are both positive and induce the same set of distributions as the ones given in the problem statement, i.e., $Q[\mathbf{A}_0]$, $Q[\mathbf{A}_1]$, ..., $Q[\mathbf{A}_m]$, but they result in different values for the conditional causal effect of interest, i.e., $P_{\mathbf{x}}^{\mathcal{M}_1}(\mathbf{y} \mid \mathbf{z}) \neq P_{\mathbf{x}}^{\mathcal{M}_2}(\mathbf{y} \mid \mathbf{z})$. Hence, $P_{\mathbf{x}}(\mathbf{y} \mid \mathbf{z})$ cannot be uniquely computed .

In Lee et al. [2019, 2020] and Correal et al. [2021] for the completeness part authors constructed such models $\mathcal{M}_1$ and $\mathcal{M}_2$, but the models violate the positivity assumption. That is, it is possible to have examples in which a given causal effect is identifiable under the positivity assumption while it is possible to construct two non-positive models that show the causal effect is not identifiable (Kivva et al. [2022]). Violation of positivity assumption renders some distributions ill-defined (conditioning on zero-probability events). That is why computing a causal effect in the classical setting with do-calculus implicitly contains steps in which we can cancel out a distribution (e.g., Q) that appears on both side of an equality, i.e., $P_1 \cdot Q = P_2 \cdot Q \Rightarrow P_1 = P_2$. Clearly, this is only possible when $Q > 0$. If positivity is violated, then such steps in computing a causal effect cannot be used.

2.1 GENERAL TRANSPORTABILITY

The work of Lee et al. [2020] proves the completeness part of the c-gID problem by constructing two models that agree on the observable distributions and disagree on the target causal effect. Those models does not satisfy the positivity assumption by the construction. A similar flaw existed in the proof of Lee et al. [2019], which was specified in details later by Kivva et al. [2022]. Given that Lee et al. [2020] does not discuss whether their models can be transformed into positive ones.

For further details, we refer to the technical report of Lee et al. [2020], which contains the proofs.

Parametrizations for an s-Thicket: According to the appendix of Lee et al. [2020], the models in Lemma 3 which is one of the main Lemmas for proving the completeness result are based on the ones in Lee et al. [2019]. These models violate the positivity assumption according to Kivva et al. [2022] and should be substituted with a fixed ones.

Parametrization for an Extended s-Thicket: According to Eq. (5) and (6) in Lee et al. [2020], it is easy to observe that several observed variables are deterministic functions of other observed variables. This implies that there exists a realization of observed variables such that the conditional probability of one observed variable given the rest is zero. This is against the positivity assumption.

Parametrization for an Extended s-Thicket with a Path-Witnessing Subgraph: In Eq. (7) of Lee et al. [2020] if $v_{\mathcal{P}}$ is an observed variable with only observed parents on a backdoor path $\mathcal{P}$, again, $v_{\mathcal{P}}$ will be a deterministic function of only observed variables. This again does not satisfy the positivity. In general, such $v_{\mathcal{P}}$ would always exist.

Please note that the errata for Lee et al. [2019] can potentially fix the issue for s-Thicket, but not for extended s-Thicket or Extended s-Thicket with a Path-Witnessing Subgraph (the last two cases).

2.2 COUNTERFACTUAL IDENTIFICATION

Here, we refer to the technical report of Correa et al. [2021] and construct a simple example that demonstrates our main concerns about the proof of the completness part of the c-gID problem.

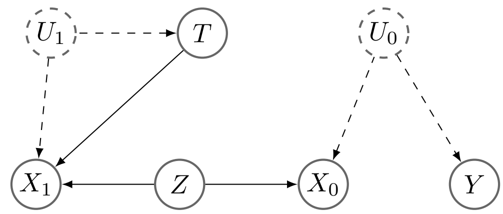

Figure 3: DAG $\mathcal{G}$ with $\mathbf{X} = \{X_0, X_1\}$ and $\mathbf{Y} = \{Y\}$.

Recall that a causal effect $P_{\mathbf{T}}(\mathbf{Y}|\mathbf{X})$ can be written as a counterfactual $P(\mathbf{Y}_*, \mathbf{X}_*)$, where $\mathbf{Y}_* \cup \mathbf{X}_* = \{W_{[\mathbf{T}]} | W \in \mathbf{V}(\mathbf{Y}_* \cup \mathbf{X}_*)\}$ and $[\mathbf{T}]$ denotes an intervention under which the counterfactual value is observed. Now, consider the graph $\mathcal{G}$ in Fig. 3. Suppose that the known distribution is $P(\mathbf{V})$ and the target conditional causal effect is

$$P_T(Y|X_0, X_1) = P(\mathbf{X}_*, \mathbf{Y}_*),$$

where $\mathbf{X}_* = \{X_{0[T]}, X_{1[T]}\}$, $\mathbf{Y}_* = \{Y_{[T]}\}$, $\mathbf{X} = \{X_0, X_1\}$ and $\mathbf{Y} = \{Y\}$. Note that for both X_0 and X_1, there exists an active backdoor path to Y, thus, we cannot use the second rule of do-calculus to simplify $P_T(Y|X_0, X_1)$. Please note that in this graph, X_0, X_1, Z, Y belong to the same ancestral component (Def. 7 in [2]) induced by $\mathbf{X}_* \cup \mathbf{Y}_*$ given $\mathbf{X}_*$. This is because $Z \in An(X_{0[T]})_{\mathcal{G}_{\underline{X_0}}} \cap An(Z_{[T]})_{\mathcal{G}} \cap An(X_{1[T]})_{\mathcal{G}_{\underline{X_1}}}$ and there is a bidirected arrow between X_0 and Y. This ancestral component contains Y and based on the definition of $\mathbf{D}_*$ (after Eq. (69) in [2]), we have $\mathbf{D}_* = \{X_{0[T]}, X_{1[T]}, Z_{[T]}, Y_{[T]}\}$. Furthermore, according to Equation (70) in Correa et al. [2021] is

$$\rho(\mathbf{x}, \mathbf{y}) := \sum_{\mathbf{d}_* \setminus (\mathbf{y}_* \cup \mathbf{x}_*)} P(\bigwedge_{D_t \in \mathbf{D}_*} D_{\mathbf{pa}_d} = d)$$

and in our example, it is equivalent to

$$P_T(Y = \mathbf{y}[Y], X_0 = \mathbf{x}[X_0], X_1 = \mathbf{x}[X_1]).$$

In part of the proof, they encounter a setting in which $\rho(\mathbf{x}, \mathbf{y})$ and $\rho(\mathbf{x})$ are not g-ID and they need to show that

$$\rho(\mathbf{y}|\mathbf{x}) = \rho(\mathbf{x}, \mathbf{y}) / \rho(\mathbf{x})$$

is not c-gID. To do so, they consider two models $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$ that shows $\rho(\mathbf{x}, \mathbf{y})$ is not g-ID and transform them into two new models to prove the non-c-gID of $\rho(\mathbf{y}|\mathbf{x})$. According to Correa et al. [2021], realizations $\mathbf{x}', \mathbf{y}'$ are such that for models $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$:

$$\rho^{(1)}(\mathbf{y}', \mathbf{x}') \neq \rho^{(2)}(\mathbf{y}', \mathbf{x}').$$

Models $\mathcal{M}^{(1)'}$ and $\mathcal{M}^{(2)'}$ obtained from models $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$ as follows:

1. Append an extra bit U_p to the node U_0.
2. X_p and Y_p binary unobserved variables defined for variables X_0 and Y, respectively.
3. Rename X_0 and Y as $\widetilde{X}_0$ and $\widetilde{Y}$ and make them unobserved, then X_0 and Y are defined in models $\mathcal{M}^{(1)'}$ and $\mathcal{M}^{(2)'}$ as $X_0 := \mathbf{x}'[X_0]$ if $X_p = 1$ and $\widetilde{X}_0$, otherwise. Similarly, they defined Y using Y_p and $\widetilde{Y}$.

According to the definitions of $\rho(\mathbf{x}, \mathbf{y})$, $\widetilde{Y}$, and $\widetilde{X}$, and using the law of total probability, we have

$$\rho'(\mathbf{x}', \mathbf{y}') = \sum_{X_p, Y_p, \widetilde{X}, \widetilde{Y}} \rho'(\mathbf{x}', \mathbf{y}', \widetilde{X}_0, \widetilde{Y} | X_p, Y_p) P(X_p, Y_p)$$

and therefore

$$\begin{aligned} \rho'(\mathbf{x}', \mathbf{y}') = &P(X_p = 0, Y_p = 0)\rho(\mathbf{x}', \mathbf{y}') + \\ &P(X_p = 1, Y_p = 0)\rho(\mathbf{x}'[X_1], \mathbf{y}') + \\ &P(X_p = 0, Y_p = 1)\rho(\mathbf{x}') + \\ &P(X_p = 1, Y_p = 1)\rho(\mathbf{x}'[X_1]). \end{aligned} \tag{77}$$

Clearly, $\rho(\mathbf{x}'[X_1]) \neq 1$ otherwise, the positivity assumption does not hold. On the other hand, based on Eq. (78)-(81) Correa et al. [2021], $\rho'(\mathbf{x}', \mathbf{y}')$ is computed by

$$\begin{aligned} &P(X_p = 0, Y_p = 0)\rho(\mathbf{x}', \mathbf{y}')+ \\ &P(X_p = 1, Y_p = 0)\rho(\mathbf{y}')+ \\ &P(X_p = 0, Y_p = 1)\rho(\mathbf{x}')+ \\ &P(X_p = 1, Y_p = 1). \end{aligned} \tag{78}$$

In general, (77) and (78) are not equal unless for example, $\mathbf{X} = \{X\}$ and $\mathbf{Y} = \{Y\}$.

Moreover the rest of the proof in Correa et al. [2021], i.e., Eq. (83)-(92) heavily relies on Eq. (78), therefore without corresponding fix the whole proof for the completeness part in c-gID problem falls apart.