Comparing Causal Frameworks: Potential Outcomes, Structural Models, Graphs, and Abstractions (Supplement)

These appendices contain demonstrations of the results in the main text as well as additional technical notes.

A Modeling (§1)

A.1 Preliminaries

Proposition A.1. Uniform RCMs are dense (in the weak subspace topology on counterfactual distributions[1]): for any $\mathcal{R}$, $\varepsilon > 0$, there is a $\mathcal{R}'$ whose distribution is uniform that ε-approximates $\mathcal{R}$ (in, e.g., the standard metric on counterfactual distributions considered as points in Euclidean space, or the Lévy-Prohorov metric).

Proof. Follows by density of $\mathbb{Q} \subset \mathbb{R}$ since any RCM with a rational distribution is representable also by a uniform distribution on units. □

Proof of Prop. 1. Suppose $\mathcal{R} = \langle \mathcal{U}, \mathbf{V}, \mathfrak{O}, \{f_{Y_\mathbf{x}}\}_{Y_\mathbf{x} \in \mathfrak{O}}, P \rangle$ is representable. Then we have an SCM $\mathcal{M} = \langle \mathbf{U}, \mathbf{V}, \{f_V\}_V, P' \rangle \in \mathfrak{M}_{\text{uniq}}$ inducing its counterfactual distribution on $\text{Val}(\mathfrak{O})$. For each $u \in \mathcal{U}$, there must be a $\mathbf{u}_u \in \text{Val}(\mathbf{U})$ such that $P'(\mathbf{u}_u) > 0$ and $\mathcal{M}_\mathbf{x}, \mathbf{u}_u \models Y = f_{Y_\mathbf{x}}(u)$ for each $Y_\mathbf{x} \in \mathfrak{O}$. Consider the RCM $\mathcal{R}' = \langle \mathcal{U}, \mathbf{V}, \{Y_\mathbf{x}\}_{\text{all } Y_\mathbf{x}}, \{f'_{Y_\mathbf{x}}\}_{Y_\mathbf{x}}, P \rangle$ where we define $f'_{Y_\mathbf{x}}(u) = f_{Y_\mathbf{x}}(u)$ for each $Y_\mathbf{x} \in \mathfrak{O}$, and $f'_{Y_\mathbf{x}} = \pi_Y(\mathbf{v})$ for each $Y_\mathbf{x} \notin \mathfrak{O}$ where $\mathbf{v}$ is the unique solution such that $\mathcal{M}_\mathbf{x}, \mathbf{u}_u \models \mathbf{v}$. Note that $\mathcal{R}'$ extends $\mathcal{R}$ by construction, has no proper extension itself, and satisfies composition and reversibility by the soundness direction of [2, Thm. 3.2], as these principles must apply to $\mathcal{M}$.

Conversely suppose we have the extension $\mathcal{R}'$. Then by the completeness direction of the proof of [2, Thm. 3.2] we can construct an SCM $\mathcal{M} \in \mathfrak{M}_{\text{uniq}}$ representing $\mathcal{R}'$ and thus also $\mathcal{R}$. Specifically, we can derive a unique maximal consistent set Γ from the outcomes defined in $\mathcal{R}'$ (unique maximality is guaranteed since $\mathcal{R}'$ has no proper extension and consistency since $\mathcal{R}'$ meets composition and reversibility), and from there define the equations of $\mathcal{M}$. □

A.2 Causal Abstraction

In this section we will use the following technical result to decompose constructive partial translations.

Lemma A.1. Suppose τ is a constructive translation of $\mathbf{V}_\text{L}$ to $\mathbf{V}_\text{H}$ and let $\mathbf{x}_\text{L}, \mathbf{x}_\text{H}$ be partial settings of $\mathbf{X}_\text{L} \subset \mathbf{V}_\text{L}, \mathbf{X}_\text{H} \subset \mathbf{V}_\text{H}$ respectively. Then $\tau(\mathbf{x}_\text{L}) = \mathbf{x}_\text{H}$ iff:

1. For each $X_\text{H} \in \mathbf{X}_\text{H}$, we have that $\tau_{X_\text{H}}(\theta_{X_\text{H}}) = \pi_{X_\text{H}}(\mathbf{x}_\text{H})$ where we define $\theta_{V_\text{H}} = \{\mathbf{x} \in \text{Val}(\Pi_{V_\text{H}}) : \pi_{\mathbf{X}_\text{L} \cap \Pi_{V_\text{H}}}(\mathbf{x}) = \pi_{\mathbf{X}_\text{L} \cap \Pi_{V_\text{H}}}(\mathbf{x}_\text{L})\}$ for each $V_\text{H} \in \mathbf{V}_\text{H}$.

2. For each $V_\text{H} \notin \mathbf{X}_\text{H}$, we have that $\tau_{V_\text{H}}(\theta_{V_\text{H}}) = \text{Val}(V_\text{H})$.

[1] See [4, 1] for more detail on this topology and inducing metrics.

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

Proof. First we show the "if" direction. It follows easily from 1 that $\tau(\mathbf{v}_\mathrm{L}) \in \pi^{-1}_{\mathbf{X}_\mathrm{H}}(\mathbf{x}_\mathrm{H})$ for any $\mathbf{v}_\mathrm{L} \in \pi^{-1}_{\mathbf{X}_\mathrm{L}}(\mathbf{x}_\mathrm{L})$, and it follows from 2 and 1 that for any $\mathbf{v}_\mathrm{H} \in \pi^{-1}_{\mathbf{X}_\mathrm{H}}(\mathbf{x}_\mathrm{H})$ there is a $\mathbf{v}_\mathrm{L} \in \pi^{-1}_{\mathbf{X}_\mathrm{L}}(\mathbf{x}_\mathrm{L})$ with $\tau(\mathbf{v}_\mathrm{L}) = \mathbf{v}_\mathrm{H}$ (specifically, any extension of $\mathbf{x}_\mathrm{L}$ works).

Now we show the "only if" direction. To show 1: suppose not. Then since θ_{V_H} is always nonempty there is some $\mathbf{x} \in \mathrm{Val}(\Pi_{X_\mathrm{H}})$ such that $\pi_{\mathbf{X}_\mathrm{L} \cap \Pi_{X_\mathrm{H}}}(\mathbf{x}) = \pi_{\mathbf{X}_\mathrm{L} \cap \Pi_{X_\mathrm{H}}}(\mathbf{x}_\mathrm{L})$ and $\tau_{X_\mathrm{H}}(\mathbf{x}) \neq \pi_{X_\mathrm{H}}(\mathbf{x}_\mathrm{H})$. But then we cannot have $\tau(\mathbf{v}_\mathrm{L}) \in \pi^{-1}_{\mathbf{X}_\mathrm{H}}(\mathbf{x}_\mathrm{H})$ for any $\mathbf{v}_\mathrm{L} \in \pi^{-1}_{\mathbf{X}_\mathrm{L}}(\mathbf{x}_\mathrm{L}) \cap \pi^{-1}_{\Pi_{X_\mathrm{H}}}(\mathbf{x})$, which is nonempty. To show 2: suppose not. Then there is some $v_\mathrm{H} \in \mathrm{Val}(V_\mathrm{H})$ such that there is no $\mathbf{x} \in \mathrm{Val}(\Pi_{V_\mathrm{H}})$ with $\pi_{\mathbf{X}_\mathrm{L} \cap \Pi_{V_\mathrm{H}}}(\mathbf{x}) = \pi_{\mathbf{X}_\mathrm{L} \cap \Pi_{V_\mathrm{H}}}(\mathbf{x}_\mathrm{L})$ and $\tau_{V_\mathrm{H}}(\mathbf{x}) = v_\mathrm{H}$. But then for any $\mathbf{v}_\mathrm{H} \in \pi^{-1}_{\mathbf{X}_\mathrm{H}}(\mathbf{x}_\mathrm{H}) \cap \pi^{-1}_{V_\mathrm{H}}(v_\mathrm{H})$, which is nonempty since $V_\mathrm{H} \notin \mathbf{X}_\mathrm{H}$, there cannot be any $\mathbf{v}_\mathrm{L} \in \tau^{-1}(\mathbf{v}_\mathrm{H}) \cap \pi^{-1}_{\mathbf{X}_\mathrm{L}}(\mathbf{x}_\mathrm{L})$, contradicting that $\tau(\mathbf{x}_\mathrm{L}) = \mathbf{x}_\mathrm{H}$. □

Proof of Prop. 2. Suppose $\mathcal{H}$ is not effective so there is some $\{(\mathbf{y}^j_\mathrm{H})_{\mathbf{x}^j_\mathrm{H}}\}_{1 \le j \le n}$ in its counterfactual support such that for some k and Y we have that $\pi_Y(\mathbf{y}^k_\mathrm{H}) \neq \pi_Y(\mathbf{x}^k_\mathrm{H})$. This means there is some $(\mathbf{y}^l_\mathrm{L})_{\mathbf{x}^l_\mathrm{L}}$ in a counterfactual in the support of $\mathcal{L}$ with $\tau(\mathbf{y}^l_\mathrm{L}) = \mathbf{y}^k_\mathrm{H}$ and $\tau(\mathbf{x}^l_\mathrm{L}) = \mathbf{x}^k_\mathrm{H}$. By Lem. A.1, 1, there is no $\mathbf{x} \in \mathrm{Val}(\Pi_Y)$ such that $\pi_{\mathbf{Y}^l_\mathrm{L} \cap \Pi_Y}(\mathbf{x}) = \pi_{\mathbf{Y}^l_\mathrm{L} \cap \Pi_Y}(\mathbf{y}^l_\mathrm{L})$ and $\pi_{\mathbf{X}^l_\mathrm{L} \cap \Pi_Y}(\mathbf{x}) = \pi_{\mathbf{X}^l_\mathrm{L} \cap \Pi_Y}(\mathbf{x}^l_\mathrm{L})$. This implies that there is some $Y' \in \mathbf{X}^l_\mathrm{L} \cap \mathbf{Y}^l_\mathrm{L} \cap \Pi_Y$ such that $\pi_{Y'}(\mathbf{x}^l_\mathrm{L}) \neq \pi_{Y'}(\mathbf{y}^l_\mathrm{L})$, contradicting the effectiveness of $\mathcal{L}$.

As for the second claim, by induction it suffices to show this for the case where $\mathcal{H}'$ is identical to $\mathcal{H}$ except that it lacks a single potential outcome $Y_\mathbf{x}$. Consider the model $\mathcal{L}'$ formed applying the mapping $\big\{(\mathbf{y}^i_\mathrm{L})_{\mathbf{x}^i_\mathrm{L}}\big\}_{1 \le i \le m} \mapsto \big\{(\mathbf{z}^i_\mathrm{L})_{\mathbf{x}^i_\mathrm{L}}\big\}_{1 \le i \le m}$ of counterfactuals to $\mathcal{L}$, where $\mathbf{z}^i_\mathrm{L} = \pi_{\mathbf{Y}^i_\mathrm{L} \setminus \Pi_Y}(\mathbf{y}^i_\mathrm{L})$ if $\tau(\mathbf{x}^i_\mathrm{L}) = \mathbf{x}$ and $\mathbf{z}^i_\mathrm{L} = \mathbf{y}^i_\mathrm{L}$ otherwise. It follows from Lem. A.1 that $\mathcal{H}' \prec_\tau \mathcal{L}'$. □

Proof of Thm. 1. Let $\mathcal{R}$ be as in Def. 1 and index its potential outcomes $\mathfrak{O}$ as $\big\{(\mathbf{Y}^i)_{\mathbf{x}^i}\big\}_{1 \le i \le n} \cup \{(\mathbf{Y}^{n+1})_{\mathbf{x}^{n+1}}\}$ where $\mathbf{x}^i \neq \varnothing$ for any $1 \le j \le n$ and $\mathbf{x}^{n+1} = \varnothing$; here $\varnothing$ represents the empty intervention on $\mathbf{V}$. We assume that $n > 0$, since otherwise $\mathcal{R}$ is trivially representable. The potential response functions of $\mathcal{R}$ are $\{f_{Y_\mathbf{x}}\}_{Y_\mathbf{x}}$.

Define a low-level set of variables $\mathbf{V}_\mathrm{L} = \{(V, j) : {}^{V \in \mathbf{V}}_{1 \le j \le n+1}\}$ with $\mathrm{Val}(V, j) = \mathrm{Val}(V) \cup \{\star\}$ where $\star \notin \mathrm{Val}(V)$ for each $(V, j) \in \mathbf{V}_\mathrm{L}$. Define τ as a constructive translation with a partition Π, defined by $\Pi_V = \big\{(V, j)\big\}_{1 \le j \le n+1}$ for each $V \in \mathbf{V}$, and $\tau_V(\mathbf{p}_\mathrm{L}) = p_j$ iff $\mathbf{p}_\mathrm{L} \in \mathrm{Val}(\Pi_V)$ is such that there is exactly one j such that $\pi_{(V,j)}(\mathbf{p}_\mathrm{L}) \neq \star$, with p_j this value. Let

$$\mathfrak{O}_\mathrm{L} = \big\{(Y, i)_{\mathbf{x}^i_\mathrm{L}} : {}^{1 \le i \le n}_{Y \in \mathbf{Y}^i}\big\} \cup \{(Y, n+1)_{\varnothing_\mathrm{L}} : Y \in \mathbf{Y}^{n+1}\} \cup \left\{(Y, i)_{\mathbf{x}^j_\mathrm{L}} : \substack{1 \le i, j \le n \\ i \neq j \\ Y \in \mathbf{Y}^i}\right\}$$
$$\cup \big\{(Y, n+1)_{\mathbf{x}^i_\mathrm{L}} : {}^{1 \le i \le n}_{Y \in \mathbf{Y}^{n+1}}\big\} \cup \big\{(Y, i)_{\varnothing_\mathrm{L}} : {}^{1 \le i \le n}_{Y \in \mathbf{Y}^i}\big\} \tag{A.1}$$

where for each $1 \le i \le n$, we let $\mathbf{X}^i_\mathrm{L} = \{(X, i) : X \in \mathbf{X}^i\}$ with $\pi_{(X,i)}(\mathbf{x}^i_\mathrm{L}) = \pi_X(\mathbf{x}^i)$ for every $X \in \mathbf{X}^i$, $\varnothing_\mathrm{L}$ is an empty intervention on $\mathbf{V}_\mathrm{L}$, and define a set of potential responses $\mathfrak{F}_\mathrm{L} = \{f^\mathrm{L}_{Z_\mathbf{x}}\}_{Z_\mathbf{x}}$ for each of these outcomes via

$$f^\mathrm{L}_{Z_\mathbf{x}} = \begin{cases} f_{Y_{\mathbf{x}^i}}, Z_\mathbf{x} = (Y, i)_{\mathbf{x}^i_\mathrm{L}} \\ f_{Y_{\mathbf{x}^{n+1}}}, Z_\mathbf{x} = (Y, n+1)_{\varnothing_\mathrm{L}} \\ f_\star, \text{otherwise} \end{cases} \tag{A.2}$$

where $f_\star$ is a constant function with $f_\star(u) = \star$ for any $u \in \mathcal{U}$.

We claim that $\mathcal{R}_\mathrm{L} = \langle \mathcal{U}, \mathbf{V}_\mathrm{L}, \mathfrak{O}_\mathrm{L}, \{f^\mathrm{L}_{Z_\mathbf{x}}\}_{Z_\mathbf{x}}, P \rangle$, where $\mathcal{U}$ and P are the same as those in $\mathcal{R}$, abstracts to $\mathcal{R}$ and is representable. To show the former, we employ the following result:

Lemma A.2. Let $\mathbf{x}_\mathrm{L} \in \mathrm{Val}(\mathbf{X}_\mathrm{L})$ for some $\mathbf{X}_\mathrm{L} \subset \mathbf{V}_\mathrm{L}$ and suppose that there is some $\mathbf{X} \subset \mathbf{V}$ and $j \in \{1, \dots, n+1\}$ such that $\pi_{(V,i)}(\mathbf{x}_\mathrm{L}) \neq \star$ iff $V \in \mathbf{X}$ and $i = j$; suppose further that for each V, there is at least one $i \in \{1, \dots, n+1\}$ such that $(V, i) \notin \mathbf{X}_\mathrm{L}$ or $\pi_{(V,i)}(\mathbf{x}_\mathrm{L}) \neq \star$. Then $\tau(\mathbf{x}_\mathrm{L}) = \mathbf{x}_\mathrm{H}$ where $\mathbf{X}_\mathrm{H} = \mathbf{X}$ and $\pi_X(\mathbf{x}_\mathrm{H}) = \pi_{(X,i)}(\mathbf{x}_\mathrm{L})$ for each $X \in \mathbf{X}$.

Proof. This follows directly from the construction of τ, in light of Lem. A.1. □

Lem. A.2 implies that $\tau(\varnothing_{\mathrm{L}}) = \varnothing$, $\tau(\mathbf{x}^i_{\mathrm{L}}) = \mathbf{x}^i$ for each i. Further, for each $u \in \mathcal{U}$ and i it implies that $\tau(\mathbf{y}^{\mathrm{L}}_i(u)) = \mathbf{y}_i(u)$ where $\mathbf{y}^{\mathrm{L}}_i(u)$ is the partial setting of $\mathbf{V}_{\mathrm{L}}$ induced by $\mathcal{R}_{\mathrm{L}}$ for potential outcomes under $\mathbf{x}^i_{\mathrm{L}}$ for unit u, viz., by (A.1) and (A.2), that with the projections $\pi_{(Y,i)}(\mathbf{y}^{\mathrm{L}}_i(u)) = f_{Y_{\mathbf{x}^i}}(u)$ for each $Y \in \mathbf{Y}^i$, $\pi_{(Y,j)}(\mathbf{y}^{\mathrm{L}}_i(u)) = \star$ for each $j \neq i$, $Y \in \mathbf{Y}^j$; and $\mathbf{y}_i(u)$ is the analogue in $\mathcal{R}$, with $\pi_Y(\mathbf{y}_i(u)) = f_{Y_{\mathbf{x}^i}}(u)$ for $Y \in \mathbf{Y}^i$. These facts establish that $P^{\mathcal{R}}_{\mathrm{cf}} = \tau_*(P^{\mathcal{R}_{\mathrm{L}}}_{\mathrm{cf}})$ so that $\mathcal{R} \prec_\tau \mathcal{R}_{\mathrm{L}}$.

Next we demonstrate that $\mathcal{R}_{\mathrm{L}}$ is representable by constructing an explicit SCM representation $\mathcal{M}_{\mathrm{L}} = \langle \{U\}, \mathbf{V}_{\mathrm{L}}, \{e_{V_{\mathrm{L}}}\}_{V_{\mathrm{L}} \in \mathbf{V}_{\mathrm{L}}}, P \rangle$, where U is an exogenous variable with $\mathrm{Val}(U) = \mathcal{U}$, while P is the same distribution as in $\mathcal{R}$ and $\mathcal{R}_{\mathrm{L}}$. It will be convenient to set up $\mathcal{M}_{\mathrm{L}}$ so as to be *recursive*. We say that an $\mathcal{M}$ over endogenous variables $\mathbf{V}$ is recursive if there is a total order $<$ on $\mathbf{V}$ such that $\mathbf{Pa}_V \subset \{V' \in \mathbf{V} : V' < V\}$ for every $V \in \mathbf{V}$. Recursiveness guarantees uniqueness and existence of solutions under every unit and intervention, regardless of the particular structural functions $\{f_V\}_V$ and merely by virtue of their signatures:

Lemma A.3. If $\mathcal{M}$ is recursive then there is a unique $\mathbf{v} \in \mathrm{Val}(\mathbf{V})$ such that $\mathcal{M}_{\mathbf{x}}, \mathbf{u} \vDash \mathbf{v}$ for any $\mathbf{x} \in \bigcup_{\mathbf{X} \subset \mathbf{V}} \mathrm{Val}(\mathbf{X})$, unit $\mathbf{u}$.

Proof. Assume without loss $\mathbf{V} = \{V_1, \dots, V_m\}$ with $V_1 < \cdots < V_m$ and prove by induction, the inductive hypothesis being that there is a unique $\mathbf{v}$ over $\{V_1, \dots, V_j\}$ for $j \leq m$, which extends as $\pi_{V_{j+1}}(\mathbf{v}) = f_{V_{j+1}}\big(\mathbf{u}, \pi_{\{V_1,\dots,V_j\}}(\mathbf{v})\big)$ if $V_{j+1} \notin \mathbf{X}$ and $\pi_{V_{j+1}}(\mathbf{x})$ otherwise. □

For the recursive order $<$, pick any $<$ such that $(V,h) < (V',h')$ for any $V, V' \in \mathbf{V}$ and $1 \leq h, h' \leq n+1$ whenever $h < h'$, for any $1 \leq j \leq n$, we have that $(X,j) < (Y,j)$ for any $X \in \mathbf{X}^j \setminus \mathbf{Y}^j$, $Y \in \mathbf{Y}^j$, and $(Y_l, n+1) < (Y_{l'}, n+1)$ whenever $l < l'$, where we fix an arbitrary indexing $\{Y_l : 1 \leq l \leq |\mathbf{Y}^{n+1}|\}$ of $\mathbf{Y}^{n+1}$. Let $\mathbf{Pa}_{V_{\mathrm{L}}} = \{V' \in V_{\mathrm{L}} : V' < V_{\mathrm{L}}\}$, $\mathbf{U}_{V_{\mathrm{L}}} = \mathbf{U}$ for each V_{L} and define the structural function $e_{(V,h)}$ by

$$e_{(V,h)}(u,\mathbf{p}) = \begin{cases} f_{V_\varnothing}(u), & h = n+1, V \in \mathbf{Y}^{n+1}, \mathbf{p} \text{ is such that } \pi_{(Y_{l'},n+1)}(\mathbf{p}) = f_{V_\varnothing}(u) \\ & \qquad \text{for any } l' \text{ such that } n+1 \leq l' < l \text{ and} \\ & \quad \pi_{(V',h')}(\mathbf{p}) = \star \text{ for any other } (V',h') \\ f_{V_{\mathbf{x}^h}}(u), & 1 \leq h \leq n, V \in \mathbf{Y}^h, \left|\mathbf{X}^h \setminus \mathbf{Y}^h\right| > 0, \mathbf{p} \text{ is such that} \\ & \quad \pi_{(X,h)}(\mathbf{p}) = \pi_X(\mathbf{x}_h) \text{ for any} \\ & \quad X \in \mathbf{X}^h \setminus \mathbf{Y}^h \text{ and } \pi_{(V',h')}(\mathbf{p}) = \star \text{ for any other } (V',h') \\ \star, & \text{otherwise} \end{cases} . \tag{A.3}$$

We can find the unique solution of $\mathcal{M}_{\mathrm{L}}$ under any of our interventions of interest by following these equations in order:

Lemma A.4 (Solutions of $\mathcal{M}_{\mathrm{L}}$)**.** Let $\mathbf{x} \in \{\mathbf{x}^j_{\mathrm{L}}\}_{1 \leq j \leq n} \cup \{\varnothing_{\mathrm{L}}\}$ and $i \in \mathcal{U}$. Then $(\mathcal{M}_{\mathrm{L}})_{\mathbf{x}}, i \vDash \mathbf{v} \in \mathrm{Val}(\mathbf{V}_{\mathrm{L}})$ where

$$\pi_{(V,h)}(\mathbf{v}) = \begin{cases} \pi_V(\mathbf{x}_j) & \mathbf{x} = \mathbf{x}^j_{\mathrm{L}} \text{ for some } j, V \in \mathbf{X}^j \setminus \mathbf{Y}^j, h = j \\ f_{V_{\mathbf{x}^j}}(i) & \mathbf{x} = \mathbf{x}^j_{\mathrm{L}} \text{ for some } j, V \in \mathbf{Y}^j, h = j \\ f_{V_\varnothing}(i) & \mathbf{x} = \varnothing_{\mathrm{L}}, h = n+1, V \in \mathbf{Y}^{n+1} \\ \star & \text{otherwise} \end{cases} . \tag{A.4}$$

Proof. Suppose $\mathbf{x} = \mathbf{x}^j_{\mathrm{L}}$ for some $j \leq n$. We claim that $\pi_{(V,k)}(\mathbf{v}) = \star$ for any V and $k < j$. The only way it could be otherwise is in the second case of (A.3), but since $k \neq j$ and $\left|\mathbf{X}^k \setminus \mathbf{Y}^k\right| > 0$ there is here at least one $(X,k) < (Y^k, k)$ with $\pi_{(X,k)}(\mathbf{v}) = \star$, so that this case in fact cannot apply. Now for $k = j$, it is clear by construction of $\mathbf{x}^j_{\mathrm{L}}$ that $\pi_{(V,j)}(\mathbf{v}) = \pi_V(\mathbf{x}^j)$ for each $V \in \mathbf{X}^j \setminus \mathbf{Y}^j$. For $V \in \mathbf{Y}^j$, if $V \in \mathbf{X}^j$ then by effectiveness and the construction of $\mathbf{x}^j_{\mathrm{L}}$, we have that $\pi_{(V,j)}(\mathbf{v}) = \pi_V(\mathbf{x}^j) = f_{V_{\mathbf{x}^j}}(i)$. Otherwise, since $\mathbf{X}^j \neq \varnothing$ and given what we have already found about $\mathbf{v}$ we fall into the second case in (A.3). Thus $\pi_{(V,j)}(\mathbf{v}) = f_{V_{\mathbf{x}^j}}(i)$. For $k > j$, if $k \leq n$ then the second case of (A.3) cannot apply so $\mathbf{v}$ extends only by $\star$'s, and if $k = n+1$ then since $\mathbf{x}^j \neq \varnothing$ the first case of (A.3) likewise cannot apply, and we get only $\star$'s. This gives us precisely the $\mathbf{v}$ specified by the first, second, and fourth cases in (A.4).

Next suppose $\mathbf{x} = \varnothing_{\mathrm{L}}$. Then $\pi_{(V,h)}(\mathbf{v}) = \star$ for any V and $h < n+1$: it could only be otherwise in the second case of (A.3) where since $\left|\mathbf{X}^h \setminus \mathbf{Y}^h\right| > 0$ there is here at least one $(X, h) < (V, h)$ with $\pi_{(X,h)}(\mathbf{v}) = \star$ by the third case of (A.3). Now if $h = n+1$ then we have $\pi_{(Y_1,n+1)}(\mathbf{v}) = f_{V_\varnothing}(i)$ by the first case of (A.3) and (inductively) likewise for $Y_2, \dots, Y_{|\mathbf{Y}^{n+1}|}$. For any $V \notin \mathbf{Y}^{n+1}$, we have $\pi_{(V,n+1)}(\mathbf{v}) = \star$. We thus obtain exactly the last two cases of (A.4). □

To show that $p_{\mathrm{cf}}^{\mathcal{M}_{\mathrm{L}}}$ marginalizes to $p_{\mathrm{cf}}^{\mathcal{R}_{\mathrm{L}}}$, because $\mathcal{M}_{\mathrm{L}}$ and $\mathcal{R}_{\mathrm{L}}$ are constructed so as to share the same set of units and distribution over them, it suffices to show that for each $i \in \mathcal{U}$ we have that $(\mathcal{M}_{\mathrm{L}})_{\mathbf{x}}, i \vDash (V, h) = f^{\mathrm{L}}_{(V,h)_{\mathbf{x}}}(i)$ for every $(V, h)_{\mathbf{x}} \in \mathfrak{D}_{\mathrm{L}}$. This is easily seen by inspecting and comparing (A.4), (A.1), and (A.2). □

SUTVA

The following remarks are in the context of the example from the discussion after Def. 9 (see also that after Thm. 1).

Remark A.1. We show how our framework can model unit-wise treatment assignment and the first part of SUTVA as making a particular abstraction viable.

Consider a model $\langle \{s\}, \{U, Y, T\}, \mathfrak{D}, \mathfrak{F}, \cdot \rangle$ where s is some singleton element and U is an endogenized unit variable (so that the exogenous unit set $\mathcal{U}$ is trivial in this model), Y encodes survival with $\mathrm{Val}(Y) = \{0, 1\}$, and T is a treatment variable that encodes whether *each* unit was assigned to treatment, with $\mathrm{Val}(T) = \bigtimes_{u \in \mathrm{Val}(U)} \mathrm{Val}(Y)$. The list $\mathfrak{D}$ contains the $2^{|\mathrm{Val}(U)|}|\mathrm{Val}(U)|$ potential responses $Y_{\mathbf{x}}$ for each possible $\mathbf{x}$ where $\pi_T(\mathbf{x}) = \mathbf{u} \in \bigtimes_{u \in \mathrm{Val}(U)} \{0, 1\}$ and $\pi_U(\mathbf{x}) = u \in \mathrm{Val}(U)$, and $\mathfrak{F}$ maps this to the survival outcome for patient u when the vector $\mathbf{u}$ encodes treatment assignments for each patient. Under the above SUTVA assumption, this outcome depends only on the component $\pi_u(\mathbf{u})$ of $\pi_T(\mathbf{x})$ corresponding to $U = u$. This means that a constructive abstraction τ to high-level variables $\{T', Y\}$ where $\mathrm{Val}(T') = \mathrm{Val}(U) \times \{0, 1\}$ with partition Π is viable: $\Pi_{T'} = \{T, U\}$, $\Pi_Y = \{Y\}$, and $\tau(\mathbf{x}) = \big(\pi_U(\mathbf{x}), \pi_{\pi_U(\mathbf{x})}(\pi_T(\mathbf{x}))\big)$. In turn, given a distribution on units, we can exogenize them out of T', showing that this model is equivalent to the familiar one with binary treatment and potential outcomes $Y_t(u)$.

Remark A.2. We show how to model several alternative ways mentioned by [5] of making the second part of SUTVA hold.

The first way is for each unit to receive only one treatment level; then the possibility of inconsistency in abstraction is excluded. Our framework requires that the potential responses are total functions (with domain the set of units) and therefore that the outcome for each unit and treatment level is defined; nevertheless, it is possible to redo the analysis from the beginning up to causal abstraction, modifying Def. 1 to admit partial potential response functions. One would then find that this version of SUTVA guarantees the viability of the abstraction collapsing all treatment levels.

The second way is to admit stochastic treatments. Suppose we have two treatment levels tr^i, tr^j and some $u \in \mathcal{U}$, occurring with probability $P(u)$, for which $Y_{\mathrm{tr}^i}(u) \neq Y_{\mathrm{tr}^j}(u)$. Then for any $p^i, p^j \geq 0$ with $p^i + p^j = 1$ we can form a new model with tr^i, tr^j merged into a single treatment tr, but u split into two new units u^i, u^j with respective probabilities $p^i P(u)$, $p^j P(u)$. In this model we define $Y_{\mathrm{tr}}(u^i) = Y_{\mathrm{tr}^i}(u)$, $Y_{\mathrm{tr}}(u^j) = Y_{\mathrm{tr}^j}(u)$, and the violation to the condition in question is removed since we end up with only a single treatment level tr. Further, the expectation of Y_{tr} conditioned on drawing u^i or u^j is equal to the weighted average of $Y_{\mathrm{tr}^i}(u)$ and $Y_{\mathrm{tr}^j}(u)$. This is really a different definition in which we mix incompatible outcomes rather than consider them to invalidate the entire abstraction—study of such a notion may prove fruitful, but is outside the scope of this article.

A final way to make this condition hold is to coarsen the outcome, by, e.g. blurring multiple health statuses into the binary distinction between surviving and not. This can in fact already be seen in our framework. Let τ be the abstraction from the running example coarsening treatment levels (values of T), τ' be that coarsening the outcome (values of Y), and $\mathcal{L}$ be our lowest-level model. Then it is possible that we have some $\mathcal{H}', \mathcal{H}$ for which $\mathcal{H} \prec_\tau \mathcal{H}' \prec_{\tau'} \mathcal{L}$, but no $\mathcal{H}$ for which $\mathcal{H} \prec_\tau \mathcal{L}$. In this sense (by moving to the already coarser model $\mathcal{H}'$) the second part of SUTVA can be made to hold.

B Inference (§2)

Proposition B.1. $(6), (7), (8) \vDash \text{LATE} = \text{ITT}_1/\text{ITT}_2$.

Proof. Regarding the encoding of expected values: for any function $f : \text{Val}_{\text{cf}} \to \mathbb{R}$ we can define an expected value $\mathbf{E}_{\mathbf{u}}(f) = \sum_{\mathbf{u} \in \mathcal{U}} P(\mathbf{u}) f(\mathfrak{O}(\mathbf{u}))$, abbreviated $\mathbf{E}(f)$. If f is integer-valued then $\mathbf{E}(f)$ is expressible within our formal language. Thus LATE is equal to a ratio[2] of our terms,

$$\text{LATE} = \mathbf{E}(Y_{x^+} - Y_{x^-} | x^+_{z^+} \wedge x^-_{z^-}) = \frac{\mathbf{P}(y^+_{x^+} \wedge y^-_{x^-} \wedge x^+_{z^+} \wedge x^-_{z^-}) - \mathbf{P}(y^-_{x^+} \wedge y^+_{x^-} \wedge x^+_{z^+} \wedge x^-_{z^-})}{\mathbf{P}(x^+_{z^+} \wedge x^-_{z^-})}.$$

We show $\text{LATE} = \text{ITT}_1/\text{ITT}_2$; (6) shows the second term of ITT_2 vanishes, while the first term is the denominator of LATE, so it suffices to show equality of the numerator and ITT_1. All of the following arithmetic moves rely only on AX. Note that

$$\mathbf{P}(y^+_{z^+, X_{z^+}} \wedge y^-_{z^-, X_{z^-}}) = \sum_{x, x'} \mathbf{P}\big(y^+_{z^+, x} \wedge x_{z^+} \wedge y^-_{z^-, x'} \wedge x'_{z^-}\big) = \sum_{x, x'} \mathbf{P}(y^+_x \wedge x_{z^+} \wedge y^-_{x'} \wedge x'_{z^-})$$

the last equality following from (8) and (7). Any terms for $x = x'$ vanish by (7) (using also that $\mathbf{P}(\epsilon) = 0 \to \mathbf{P}(\epsilon \wedge \zeta) = 0$ is derivable from basic probabilistic logic), the term for $x = x^-$ and $x' = x^+$ vanishes by (6), and we are left with exactly the first term of the numerator. We can analogously derive equality of the expressions with negative coefficients. □

Proposition B.2. Under representability, $\text{ITT}_1 = \mathbf{E}(Y_{z^+} - Y_{z^-})$.

Proof. We claim that $\text{ITT}_1 = \mathbf{E}(Y_{z^+, X_{z^+}} - Y_{z^-, X_{z^-}}) = \mathbf{E}(Y_{z^+} - Y_{z^-})$ and thus that

$$\mathbf{P}\big(y^+_{z^+, X_{z^+}} \wedge y^-_{z^-, X_{z^-}}\big) - \mathbf{P}\big(y^-_{z^+, X_{z^+}} \wedge y^+_{z^-, X_{z^-}}\big) = \mathbf{P}(y^+_{z^+} \wedge y^-_{z^-}) - \mathbf{P}(y^-_{z^+} \wedge y^+_{z^-})$$

under representability. It suffices to show Rep. $\vDash y_{z, X_z} \leftrightarrow y_z$ where Rep. stands for representability. Note that Rep. $\vDash y_{z, X_z} \leftrightarrow \bigvee_{x, y'} y_{zx} \wedge x_z \wedge y'_z \leftrightarrow \bigvee_x y_{xz} \wedge x_z \wedge y_z$, the last step since $(2) \vDash x_z \wedge y'_z \to y'_{xz}$ contradicting y_{xz} when $y' \neq y$. Similarly by setting $y' = y$, we have Rep. $\vDash \bigvee_x y_{xz} \wedge x_z \wedge y_z \leftrightarrow \bigvee_x x_z \wedge y_z \leftrightarrow y_z$. Note that no additional outcomes are used here except Y_z, which is necessary to state the result. □

Proof of Thm. 2. Soundness is straightforward. For completeness, show that any consistent formula φ is satisfiable. By the completeness proof for AX [3, Thm. 6] we get a distribution P satisfying φ over joint valuations $\text{Val}(\mathfrak{O}_\varphi)$, where $\mathfrak{O}_\varphi$ is the set of potential outcomes appearing in φ. To construct a satisfying model $\mathcal{R}$, add a unit u that witnesses any $\mathbf{o} \in \text{Val}(\mathfrak{O}_\varphi)$ for which $P(\mathbf{o}) > 0$. It is clear that $\mathcal{R} \in \mathfrak{R}(S)$ since φ is consistent with $\text{T}(S)$. □

Proof of Cor. 1. Observe that $\mathfrak{R}_{\text{eff}} = \mathfrak{R}(1)$. □

Proof of Thm. 3. Soundness has been shown in prior works [7], and in any case is straightforward ($\mathsf{ER}^{\mathcal{G}}$ since each variable can only be a function of its parents in $\mathfrak{M}(\mathcal{G})$ and $\mathsf{cf\text{-}sep}^{\mathcal{G}}$ since the functional mechanisms determining unconfounded nodes must be independent; see also Cor. B.1 below for a derivation from d-separation).

Completeness: our proof is a modification of the proof for AX_3 in [3, Thm. 6]. We take the same steps as in that proof, the strategy being to show that any consistent φ is satisfiable, up to and including the normal form in Lem. 8, except that we use the full set $\mathbf{V}$ in place of $\mathbf{V}_\varphi$ (this is possible since $\mathbf{V}$ is finite in our case).

Let $\mathcal{C}$ be the set of connected components under the edge relation $\dashleftarrow\!\!\dashrightarrow$ of $\mathcal{G}$. Each $\mathbf{C} \in \mathcal{C}$ is a subset $\mathbf{C} \subset \mathbf{V}$ that is a complete graph under $\dashleftarrow\!\!\dashrightarrow$, by assumption. For each $\mathbf{C}$ let

$$\Delta_{\mathbf{C}} = \Bigg\{ \bigwedge_{\substack{C \in \mathbf{C} \\ \mathbf{p}_C \in \text{Val}(\mathbf{Pa}^{\mathcal{G}}_C)}} C_{\mathbf{p}_C} = c_{C, \mathbf{p}_C} : \substack{c_{C, \mathbf{p}_C} \in \text{Val}(C) \text{ for each } C \in \mathbf{C}, \\ \mathbf{p}_C \in \text{Val}(\mathbf{Pa}^{\mathcal{G}}_C)} \Bigg\}$$

[2] The ratio $\mathbf{t}_1 = {}^{\mathbf{t}_2}/_{\mathbf{t}_3}$ is an abbreviation for $\mathbf{t}_1 \cdot \mathbf{t}_3 = \mathbf{t}_2 \in \mathcal{L}$. Likewise in Ex. 4.

be the collection of counterfactuals describing complete functional mechanisms for each $C \in \mathbf{C}$. The notation $C_{\mathbf{p}_C} = c_{C,\mathbf{p}_C}$ represents the outcome that $C = c_{C,\mathbf{p}_C}$ under the intervention $\mathbf{p}_C$; we use such notation in order to be explicit about the choice of variable C. The below is our analogue of [3, Lem. 8].

Lemma B.1. Let φ be a conjunction of literals. Then there are polynomial terms $\{\mathbf{t}_i, \mathbf{t}'_{i'}\}_{i,i'}$ in the variables $\left\{\mathbf{P}(\delta_{\mathbf{C}}) : {}^{\delta_{\mathbf{C}} \in \Delta_{\mathbf{C}}}_{\mathbf{C} \in \mathcal{C}}\right\}$ such that

$$\mathrm{T}(\mathcal{G}) \vdash \varphi \leftrightarrow \bigwedge_{\substack{\delta_{\mathbf{C}} \in \Delta_{\mathbf{C}} \\ \mathbf{C} \in \mathcal{C}}} \mathbf{P}(\delta_{\mathbf{C}}) \geq 0 \wedge \bigwedge_{\mathbf{C} \in \mathcal{C}} \Big[\sum_{\delta_{\mathbf{C}} \in \Delta_{\mathbf{C}}} \mathbf{P}\left(\delta_{\mathbf{C}}\right) = 1 \Big] \wedge \bigwedge_i \mathbf{t}_i \geq 0 \wedge \bigwedge_{i'} \mathbf{t}'_{i'} > 0. \tag{B.1}$$

Proof. This will follow from the fact that we have the below (B.2) for any $\delta \in \Delta_{\mathrm{sat}}$. Here Δ_{sat} is the set defined in [3] as the satisfiable subset of complete interventional state descriptions $\Delta = \left\{ \bigwedge_{\substack{\mathbf{x} \in \mathrm{Val}(\mathbf{X}) \\ \mathbf{X} \subset \mathbf{V}}} \bigwedge_{V \in \mathbf{V}} V_{\mathbf{x}} = \pi_V(\mathbf{v}_{\mathbf{x}}) : {}^{\mathbf{v}_{\mathbf{x}} \in \mathrm{Val}(\mathbf{V})}_{\text{for each } \mathbf{x}} \right\}$.

$$\mathrm{T}(\mathcal{G}) \vdash \mathbf{P}(\delta) = \prod_{\mathbf{C} \in \mathcal{C}} \mathbf{P}\Bigg(\bigwedge_{\substack{C \in \mathbf{C} \\ \mathbf{p}_C \in \mathrm{Val}(\mathbf{Pa}^{\mathcal{G}}_C)}} C_{\mathbf{p}_C} = c_{C,\mathbf{p}_C} \Bigg) \tag{B.2}$$

where $c_{C,\mathbf{p}_C} \in \mathrm{Val}(C)$ for each C is the outcome of C in δ in the outer conjunct corresponding to $\mathbf{x} = \mathbf{p}_C$ and the inner conjunct corresponding to $V = C$. To see (B.2), note that for each $V \in \mathbf{V}$ we can remove all conjuncts in a δ for any $\mathbf{x}$ that do not correspond to a setting $\mathbf{p}_C$ of the parents $\mathbf{Pa}^{\mathcal{G}}_C$ by $\mathsf{ER}^{\mathcal{G}}$ and composition, which we can use since SCM is part of $\mathrm{T}(\mathcal{G})$. We thereby obtain $\mathbf{P}(\delta) = \mathbf{P}\big(\bigwedge_{\substack{V \in \mathbf{V} \\ \mathbf{p} \in \mathbf{Pa}^{\mathcal{G}}_V}} V_{\mathbf{p}} = \pi_V(\mathbf{v}_{\mathbf{p}}) \big)$ and this gives us the $c_{V,\mathbf{p}} = \pi_V(\mathbf{v}_{\mathbf{p}})$. Next, we apply $\mathsf{cf\text{-}sep}^{\mathcal{G}}$ to this expression, giving us the final product in (B.2) factorizing over all $\mathbf{C} \in \mathcal{C}$. □

Now, as in the proof of [3, Thm. 6], the Positivstellensatz shows that consistency of φ implies there is a real solution to (B.1) for $\left\{\mathbf{P}(\delta_{\mathbf{C}})\right\}_{\substack{\delta_{\mathbf{C}} \in \Delta_{\mathbf{C}} \\ \mathbf{C} \in \mathcal{C}}}$. We construct a model $\mathcal{M}$ with diagram $\mathcal{G}$ inducing these probabilities and thereby satisfying φ. Note that any $\delta_{\mathbf{C}} \in \Delta_{\mathbf{C}}$ defines a function $f^{\delta_{\mathbf{C}}}_C : \mathrm{Val}(\mathbf{Pa}^{\mathcal{G}}_C) \to \mathrm{Val}(C)$ for each $C \in \mathbf{C}$ via setting $f^{\delta_{\mathbf{C}}}_C(\mathbf{p}_C)$ to the value C is set to under $\mathbf{p}_C$ in $\delta_{\mathbf{C}}$. We let $\mathcal{M} = \langle \{U_{\mathbf{C}}\}_{\mathbf{C} \in \mathcal{C}}, \mathbf{V}, \{f_V\}_V, P \rangle$ where P factors as $P\big(\{u_{\mathbf{C}}\}_{\mathbf{C} \in \mathcal{C}}\big) = \prod_{\mathbf{C} \in \mathcal{C}} P(u_{\mathbf{C}})$ and for any variable $C \in \mathbf{V}$ with $C \in \mathbf{C}$, we define parent sets $\mathbf{Pa}_C = \mathbf{Pa}^{\mathcal{G}}_C$ and $\mathbf{U}_C = \{U_{\mathbf{C}}\}$; this implies that $\mathcal{M}$ has diagram $\mathcal{G}$ (there is a bidirected edge between any C, C' in the same $\mathbf{C}$, and the statement assumes $\mathcal{G}$'s bidirected arcs form a disjoint union of complete graphs). Finally let $\mathrm{Val}(U_{\mathbf{C}}) = \Delta_{\mathbf{C}}$ for each $\mathbf{C}$, and let $P(\delta_{\mathbf{C}})$ be the probability $\mathbf{P}(\delta_{\mathbf{C}})$ from our solution to (B.1); define the structural functions by $f_C(\mathbf{p}_C, \delta_{\mathbf{C}}) = f^{\delta_{\mathbf{C}}}_C(\mathbf{p}_C)$ for each $C \in \mathbf{C} \in \mathcal{C}$. By construction, it is clear that $\mathcal{M}$ induces exactly the probabilities from our solution to (B.1). □

The following definition of d-separation and its extension to mixed diagrams are standard (e.g., [6, 8]). In a directed graph $\mathcal{D}$ over nodes $\mathbf{V}$, let $\mathbf{De}^{\mathcal{D}}_V = \{V' \in \mathbf{V} : V \to \cdots \to V'\}$ be the set of *descendants* of $V \in \mathbf{V}$, the transitive closure of parenthood, with $V \in \mathbf{De}^{\mathcal{D}}_V$.

Definition B.1 (Directed d-separation)**.** Given directed graph $\mathcal{D}$, the nodes X and Y are *d-separated* given a set of nodes $\mathbf{Z}$ if every path (consisting of arrows of either direction) from X to Y either has a *collider* node (to be defined shortly) on the path M such that $\mathbf{De}^{\mathcal{D}}_M \cap \mathbf{Z} = \varnothing$, or a non-collider node M' on the path such that $M' \in \mathbf{Z}$. A collider is a vertex M on the path such that the path is of the form $X \rightleftharpoons \cdots \to M \leftarrow \cdots \rightleftharpoons Y$ (the symbol $\rightleftharpoons$ denoting an arrow of either direction). The subsets $\mathbf{X}, \mathbf{Y}$ of nodes are d-separated given $\mathbf{Z}$ if every $X \in \mathbf{X}$ and $Y \in \mathbf{Y}$ are d-separated given $\mathbf{Z}$.

Definition B.2 (Mixed d-separation)**.** Given mixed causal diagram $\mathcal{G}$ over $\mathbf{V}$, let $\mathcal{C}$ be the set of maximal cliques in $\mathcal{G}$ under its edge relation $\dashleftarrow\!\!\dashrightarrow$. Thus every $\mathbf{C} \in \mathcal{C}$ is a subset $\mathbf{C} \subset \mathbf{V}$ of vertices such that $C \dashleftarrow\!\!\dashrightarrow C'$ for every $C, C' \in \mathbf{C}$, and there is no $\mathbf{C}' \supsetneq \mathbf{C}$ with this property. Then form the directed graph $\mathcal{D}^{\mathcal{G}}$ with nodes $\mathbf{V} \cup \mathcal{C}$ and edges

$$\{V \to V' : V, V' \in \mathbf{V}, V \to V' \text{ in } \mathcal{G}\} \cup \{\mathbf{C} \to V : \mathbf{C} \in \mathcal{C}, V \in \mathbf{V}, V \in \mathbf{C}\}.$$

Finally, given $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \subset \mathbf{V}$, we say that $\mathbf{X}$ and $\mathbf{Y}$ are d-separated in $\mathcal{G}$ given $\mathbf{Z}$ if the same holds in $\mathcal{D}^{\mathcal{G}}$ (according to Def. B.1).

Corollary B.1. Let $\mathsf{d\text{-}sep}^{\mathcal{G}}$ be an encoding of all the conditional independences implied by the graphical d-separation criterion over the diagram $\mathcal{G}$ (including over counterfactuals, where these are determined by parallel networks). Then the system $\mathrm{T}(\mathsf{ER}^{\mathcal{G}}) + \mathsf{d\text{-}sep}^{\mathcal{G}} + \mathsf{SCM}$ is sound and complete under the same assumptions as Thm. 3.

Proof. For soundness, see [9] or any standard discussion of d-separation, e.g., that in [7] or [8]. For completeness, in light of Thm. 3 it suffices to show that any instance of $\mathsf{cf\text{-}sep}^{\mathcal{G}}$, as in Def. 13, is an instance of $\mathsf{d\text{-}sep}^{\mathcal{G}}$. Consider the parallel network $\mathcal{G}'$ with $n + n'$ copies of $\mathcal{G}$ and the two sets of nodes $\mathbf{Y} = \{(Y_i)_{\mathbf{p}_i}\}_{1\leq i\leq n}$, $\mathbf{Y}' = \{(Y_i')_{\mathbf{p}_i'}\}_{1\leq i\leq n'}$. Since there is no $Y \in \mathbf{Y}$, $Y' \in \mathbf{Y}'$ for which $Y = Y'$ or $Y \dashleftarrow\!\dashrightarrow Y'$ in $\mathcal{G}$ and all ingoing edges to Y and Y' have been severed in $\mathcal{G}'$, the only possibility for a (direction-agnostic) path between some such Y, Y' is a bidirected path of length ≥ 2 or a bidirected path between proper descendants of Y, Y'. All such paths contain colliders, so that d-separation delivers the (unconditional) independence (9). □

References

[1] P. Billingsley. *Convergence of Probability Measures*. Wiley, 2nd edition, 1999.

[2] J. Y. Halpern. Axiomatizing causal reasoning. *Journal of AI Research*, 12:317–337, 2000.

[3] D. Ibeling and T. Icard. Probabilistic reasoning across the causal hierarchy. In *Proceedings of AAAI*, 2020.

[4] D. Ibeling and T. Icard. A topological perspective on causal inference. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 5608–5619. Curran Associates, Inc., 2021.

[5] G. W. Imbens and D. B. Rubin. *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press, 2015.

[6] J. Pearl. *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufmann, San Francisco, CA, 1988.

[7] J. Pearl. *Causality*. Cambridge University Press, 2009.

[8] I. Shpitser and J. Pearl. Complete identification methods for the causal hierarchy. *Journal of Machine Learning Research*, 9:1941–1979, 2008.

[9] T. Verma and J. Pearl. Causal networks: Semantics and expressiveness. In *Proceedings of the Fourth Annual Conference on Uncertainty in Artificial Intelligence*, UAI ’88, USA, 1988. Elsevier Science Inc.