Supplement to General Identification of Dynamic Treatment Regimes Under Interference

Eli Sherman **David Arbour** **Ilya Shpitser**

1 Graphical Models Background

1.1 Statistical Graphical Models

Chain graphs and their submodels were originally conceived as *statistical* models over random variables, encoding conditional independence constraints in their factorization. For instance, a DAG $\mathcal{G}(\mathbf{V})$ represents the set of joint distributions over $\mathbf{V}$ which factorize according to:

$$p(\mathbf{V}) = \prod_{V \in \mathbf{V}} p(V \,|\, \text{pa}_{\mathcal{G}}(V)) \tag{1}$$

Similarly, an MRF $\mathcal{G}(\mathbf{V})$ represents the set of distributions over $\mathbf{V}$ which factorize according to the factorization:

$$p(\mathbf{V}) = Z^{-1} \prod_{\mathbf{C} \in \mathcal{C}(\mathcal{G}_U)} \phi_{\mathbf{C}}(\mathbf{C}),$$

where Z is a normalizing constant, $\mathcal{C}(\mathcal{G})$ denotes the set of cliques in $\mathcal{G}$, and $\phi_{\mathbf{C}}$ is an arbitrary function over $\mathbf{C}$ known as a clique potential [1].

CGs merge these notions by allowing for both directed and undirected edges. Like DAGs and MRFs, a CG $\mathcal{G}(\mathbf{V})$ represents the set of distributions over $\mathbf{V}$ that factorize according to the two-level factorization:

$$p(\mathbf{V}) = \prod_{\mathbf{B} \in \mathcal{B}(\mathcal{G})} p(\mathbf{B} | pa_{\mathcal{G}}(\mathbf{B})) \tag{2}$$

$$p(\mathbf{B} | pa_{\mathcal{G}}(\mathbf{B})) = Z(pa_{\mathcal{G}}(\mathbf{B}))^{-1} \prod_{\mathbf{C} \in \mathcal{C}^{\star}} \phi_{\mathbf{C}}(\mathbf{C}), \tag{3}$$

where $\mathcal{C}^{\star} = \{\mathbf{C} \in \mathcal{C}((\mathcal{G}_{\mathbf{B} \cup pa_{\mathcal{G}}(\mathbf{B})})^a) : \mathbf{C} \not\subseteq pa_{\mathcal{G}}(\mathbf{B})\}$, the set of cliques that intersect $\mathbf{B}$ in the augmented graph [1] on $\mathbf{B}$ and $\text{pa}_{\mathcal{G}}(\mathbf{B})$. An augmented graph $\mathcal{G}^a$ is obtained from $\mathcal{G}$ by making any edges in $\mathcal{G}$ undirected and adding undirected edges between each $V \in \text{pa}_{\mathcal{G}}(\mathbf{B})$ for all $\mathbf{B} \in \mathcal{B}(\mathcal{G})$.

Throughout the paper, we will assume that all probability distributions have full support.

1.2 Causal Graphical Models

In contrast to their statistical analogues, causal DAGs [3] and causal CGs [2] represent distributions over *counterfactual* variables. For $Y \in \mathbf{V}$ and $\mathbf{A} \subseteq \mathbf{V} \setminus V$, the counterfactual $Y(\mathbf{a})$ denotes the value of Y under the hypothetical scenario in which $\mathbf{A}$ is set to $\mathbf{a}$ via a *node* (or *atomic*) intervention [3].

In this paper, we will assume Pearl's functional model. For a DAG $\mathcal{G}(\mathbf{V})$, counterfactuals $V(\mathbf{a})$ are determined by structural equations $f_V(\mathbf{a}, \epsilon_V)$, which remain invariant under an intervention $\mathbf{a}$ with ϵ_V denoting randomness in the causal process. These *one step ahead* counterfactuals can be used to define all variables in the model via *recursive substitution*. For $\mathbf{A} \subseteq \mathbf{V} \setminus \{V\}$:

$$V(\mathbf{a}) \equiv V(\mathbf{a}_{\mathrm{pa}_{\mathcal{G}}(V)}, \{W(\mathbf{a}) : W \in \mathrm{pa}_{\mathcal{G}}(V) \setminus \mathbf{A}\}),$$

where $\mathbf{a}$ lies in the state space of $\mathbf{A}$.

A parameter is said to be *identifiable* in a model when it can be expressed as a function of observed data. In a DAG or CG $\mathcal{G}$ with all variables $\mathbf{V}_D$ or $\mathbf{V}_C$, all counterfactuals arising from node interventions are identified by the *g-formula* [5] and *chain graph g-formula* [2] respectively:

$$p(\mathbf{V}_D(\mathbf{a})) = \prod_{V \in \mathbf{V}_D \setminus \mathbf{A}} p(V | \, \mathrm{pa}_{\mathcal{G}}(V))|_{\mathbf{A}=\mathbf{a}} \tag{4}$$

$$p(\mathbf{V}_C(\mathbf{a})) = \prod_{\mathbf{B} \in \mathcal{B}(\mathcal{G})} p(\mathbf{B} \setminus \mathbf{A} | \, \mathrm{pa}_{\mathcal{G}}(\mathbf{B}), \mathbf{B} \cap \mathbf{A})|_{\mathbf{A}=\mathbf{a}} \tag{5}$$

2 Background: Identification in Causal Graphical Models

In this section, we discuss the state of latent-variable identification theory in causal graphical models. These advancements culminate with sound and complete algorithms for identification in the presence of latent variables in segregated graphs and identification of policy intervention effects in latent variable DAGs. The current work bridges these literatures.

2.1 The Nested Markov Factorization: Re-expressing the ID Algorithm

[10] gave a general condition for determining identifiability of node interventions in latent-variable DAGs. [7] re-expressed Tian's condition as a concise algorithm, known as 'ID', and proved that it is complete. Recently, [4] re-expressed the algorithm in [7] in terms of a modified nested factorization, similar to the g-formula in a type of mixed graph.

While the substance of the identification approach in [4] is identical to that in [7] and [10], the fixing operator they present enables a compact representation of existing identification theory and makes clear the connection between the ID algorithm and Robins' g-formula (Eq. 4), which is itself a modified factorization on DAGs. For these reasons, we make use of this framework for the SG policy identification extension that we present in this work. This re-formulation relies on several important concepts which we describe below.

2.1.1 Latent Projections

Rather than considering latent-variable DAGs explicitly, [4] considers a class of models known as acyclic directed mixed graphs (ADMGs). An ADMG contains directed and bi-directed edges and represents and equivalence class of latent-variable DAGs. Given a latent-variable DAG $\mathcal{G}(\mathbf{V} \cup \mathbf{H})$, where $\mathbf{V}$ is observed and $\mathbf{H}$ is latent, we can obtain the corresponding ADMG $\mathcal{G}'(\mathbf{V})$ via a latent projection operation: edges $A \to B$ in $\mathcal{G}$ are maintained in $\mathcal{G}'$; additionally, $\mathcal{G}'$ has an edge $A \to B$ for any directed path $A \to \cdots \to B$ where the intermediate nodes are all in $\mathbf{H}$, and $\mathcal{G}'$ has an edge $A \leftrightarrow B$ if there exists a path $A \leftarrow \cdots \to B$ in $\mathcal{G}$ with all intermediate nodes in $\mathbf{H}$ and no consecutive edges $\to H \leftarrow$ for $H \in \mathbf{H}$. We can also define conditional ADMGs (CADMGs) which partition nodes into sets of random variables $\mathbf{V}$ and fixed variables $\mathbf{W}$. In a CADMG $\mathcal{G}(\mathbf{V}, \mathbf{W})$, the variables $\mathbf{W}$ have no incoming edges. An ADMG $\mathcal{G}(\mathbf{V})$ is also trivially a CADMG with $\mathbf{W} = \emptyset$.

As described above, we use segregated graphs as the chain graph analogue of ADMGs. In the formulation we use, SGs represent an equivalence class of latent variable chain graphs, defined in a way that maintains their causal interpretation. A latent variable chain graph $\mathcal{G}(\mathbf{V} \cup \mathbf{H})$ is *block-safe* [6] if no $V \in \mathbf{V}$ has an incident edge from a latent variable (i.e., $H \to V$ for $H \in \mathbf{H}$ is forbidden) and no latent variable $H \in \mathbf{H}$ has an incident undirected edge. A block-safe latent variable chain graph can be represented with a segregated graph via the same latent projection operation described above.

2.1.2 Kernels and Fixing

Whereas DAGs and CGs factorize as a product of conditional probability distributions, ADMGs and SGs factorize as a product of *kernels* [1]. Again, following the notation in [6], a kernel $q_{\mathbf{V}}(\mathbf{V}|\mathbf{W})$ is a function that maps values of $\mathbf{W}$ to densities on $\mathbf{V}$ with $\sum_{\mathbf{v}\in\mathbf{V}} q_{\mathbf{V}}(\mathbf{v}|\mathbf{w}) = 1$ for each possible realization $\mathbf{w}$. As with probability distributions, for some $\mathbf{A} \subseteq \mathbf{V}$, conditioning and marginalization in kernels are defined as follows:

$$q(\mathbf{A}|\mathbf{W}) \equiv \sum_{\mathbf{V}\setminus\mathbf{A}} q(\mathbf{V}|\mathbf{W})$$

$$q(\mathbf{V}\setminus\mathbf{A}|\mathbf{A},\mathbf{W}) \equiv \frac{q(\mathbf{V}|\mathbf{W})}{q(\mathbf{A}|\mathbf{W})}.$$

The notion of fixing variables is closely tied to kernels. In a CADMG $\mathcal{G}(\mathbf{V})$, a variable V is fixable if there does not exist both a bi-directed path and a directed path to some $V' \in \mathbf{V}$, or concisely, $\mathrm{de}_{\mathcal{G}}(V) \cap \mathrm{dis}_{\mathcal{G}}(V) = \emptyset$. In a DAG $\mathcal{G}$ with a corresponding probability distribution $p(\mathbf{V})$, fixing V corresponds to applying the g-formula to obtain a new distribution $p(\mathbf{V}\setminus V)$ and a new graph $\mathcal{G}'$. For a CADMG $\mathcal{G}(\mathbf{V},\mathbf{W})$ with a corresponding kernel $q(\mathbf{V}|\mathbf{W})$, [4] defines a similar operation for fixing V in $\mathcal{G}$, denoted $\phi_V(\mathcal{G})$. This operator yields a new kernel and a new CADMG $\mathcal{G}'(\mathbf{V}\setminus\{V\},\mathbf{W}\cup\{V\})$. In this graph, all bi-directed and directed edges *into* V are removed. The operator also yields a new kernel:

$$q'(\mathbf{V}\setminus\{V\}|\mathbf{W}\cup\{V\}) \equiv \frac{q(\mathbf{V}|\mathbf{W})}{q(\mathbf{V}|\,\mathrm{nd}_{\mathcal{G}}(V),\mathbf{W})}.$$

Since the fixing operation generalizes the g-formula, it's probabilistic interpretation varies – acting as marginalization, conditioning, and sometimes neither – depending on the characteristics of the variable being fixed relative to the kernel it is being fixed in.

2.1.3 Reachability, the Nested Factorization, and ID

We can extend the notion of fixability to sets of variables $\mathbf{S} \subseteq \mathbf{V}$ in a CADMG $\mathcal{G}$. If it is possible to find a sequence $S_1, S_2, \ldots$ of the variables in $\mathbf{S}$ such that S_1 is fixable in $\mathcal{G}$, S_2 is fixable in $\phi_S(\mathcal{G})$ and so on, then $\mathbf{S}$ is fixable and $\mathbf{V}\setminus\mathbf{S}$ is said to be *reachable* in $\mathcal{G}$.

It was shown in [4] and [6] that all valid fixing sequences for $\mathbf{S}$ in a CADMG $\mathcal{G}(\mathbf{V},\mathbf{W})$ yield the same resulting CADMG $\mathcal{G}(\mathbf{V}\setminus\mathbf{S},\mathbf{W}\cup\mathbf{S})$ and analogously for the kernel obtained by fixing $\mathbf{S}$ in $q(\mathbf{V}|\mathbf{W})$. The fixing operator can therefore be defined for sets as it was for singleton variables: $\phi_{\mathbf{S}}$. A CADMG $\mathcal{G}(\mathbf{V},\mathbf{W})$ is said to satisfy the nested Markov factorization if for every fixable $\mathbf{S}$

$$\phi_{\mathbf{S}}(q(\mathbf{V}|\mathbf{W});\mathcal{G}) = \prod_{\mathbf{D}\in\mathcal{D}(\phi_{\mathbf{S}}(\mathcal{G}))} \phi_{\mathbf{V}\setminus\mathbf{D}}(q(\mathbf{V}|\mathbf{W});\mathcal{G})$$

[4] showed that $p(\mathbf{V}\cup\mathbf{H})$ satisfies the above factorization for a DAG $\mathcal{G}(\mathbf{V}\cup\mathbf{H})$ then $p(\mathbf{V})$ satisfies the factorization for the corresponding ADMG $\mathcal{G}(\mathbf{V})$. An analogous result for SGs was shown in [6], which we will discuss below.

This notation permits a reformulation of the ID algorithm as a one line formula, proven in [4] to be identical to the algorithm in [7]: Let $\mathbf{Y},\mathbf{A}$ be disjoint subsets of $\mathbf{V}$ in an ADMG $\mathcal{G}(\mathbf{V})$. Let $\mathbf{Y}^\star = \mathrm{an}_{\mathcal{G}_{\mathbf{V}\setminus\mathbf{A}}}(\mathbf{Y})$. The intervention $p(\mathbf{Y}|\mathrm{do}(\mathbf{a}))$ is identified in $\mathcal{G}$ if and only if every set (district) $\mathbf{D}\in\mathcal{D}(\mathcal{G}_{\mathbf{Y}^\star})$ is reachable and, if identification holds, then

$$p(\mathbf{Y}|\mathrm{do}(\mathbf{a})) = \sum_{\mathbf{Y}^\star\setminus\mathbf{Y}} \prod_{\mathbf{D}\in\mathcal{D}(\mathcal{G}_{\mathbf{Y}^\star})} \phi_{\mathbf{V}\setminus\mathbf{D}}(p(\mathbf{V});\mathcal{G})|_{\mathbf{A}=\mathbf{a}}. \quad (6)$$

2.2 Identification in Segregated Graphs

2.2.1 The Segregated Factorization

Building off the nested factorization for ADMGs and the chain graph factorization, we can define the segregated factorization of an SG [6]. Recall that the variables in an SG $\mathcal{G}$ can be grouped into those that lie in a non-trivial block which we denote $\mathbf{B}^\star = \cup_{\mathbf{B}\in\mathcal{B}^{nt}(\mathcal{G})}\mathbf{B}$, and those that don't, which we denote $\mathbf{D}^\star = \cup_{\mathbf{D}\in\mathcal{D}(\mathcal{G})}\mathbf{D}$.

We can factorize an SG as the product of two kernels. The first kernel corresponds to a conditional chain graph (CCG) $\mathcal{G}(\mathbf{V},\mathbf{W})$ where, as in CADMGs, $\mathbf{V}$ are random nodes and $\mathbf{W}$ are fixed. A kernel $q(\mathbf{V}|\mathbf{W})$ is said to be Markov relative to a CCG $\mathcal{G}$ if it satisfies Eq. 1 with the following modification to the outer factorization

$$q(\mathbf{V}|\mathbf{W}) = Z(\mathbf{W})^{-1}\prod_{\mathbf{B}\in\mathcal{B}(\mathcal{G})} q(\mathbf{B}|\operatorname{pa}_{\mathcal{G}}(\mathbf{B})),$$

and a similar replacement of $p(\mathbf{B}|\operatorname{pa}_{\mathcal{G}}(\mathbf{B}))$ with q in the inner factorization. We will denote the CCG obtained from a SG by $\mathcal{G}^b$ with $\mathbf{V}$ corresponding to $\mathbf{B}^\star$ and $\mathbf{W}$ to $\operatorname{pa}^s_{\mathcal{G}}(\mathbf{B}^\star)$. $\mathcal{G}^b$ contains edges between each node in $\mathbf{B}^\star$ that exists in $\mathcal{G}$ as well as those between $\operatorname{pa}^s_{\mathcal{G}}(\mathbf{B}^\star)$ in $\mathcal{G}$.

The second kernel corresponds to a CADMG which we will denote $\mathcal{G}^d$ with random nodes $\mathbf{D}^\star$ and fixed nodes $\operatorname{pa}^s_{\mathcal{G}}(\mathbf{D}^\star)$. $\mathcal{G}^d$ contains all edges between $\mathbf{D}^\star$ that are present in $\mathcal{G}$ as well as the edges between $\operatorname{pa}^s_{\mathcal{G}}(\mathbf{D}^\star)$ and $\mathbf{D}^\star$ in $\mathcal{G}$.

If each of these kernels adheres to the factorization of the respective conditional graph, then $p(\mathbf{V})$ is obeys the segregated factorization. Specifically, $p(V)$ satisfies the segregated factorization if $q(\mathbf{D}^\star|\operatorname{pa}^s_{\mathcal{G}}(\mathbf{D}^\star)$ satisfies the nested factorization and $q(\mathbf{B}^\star|\operatorname{pa}^s_{\mathcal{G}}(\mathbf{B}^\star))$ satisfies the CCG factorization.

2.2.2 The Segregated Graph ID Algorithm

Using the above extension of the nested factorization, we can now describe the extension to the ID algorithm, expressed using the fixing operator ϕ [6]: for a block safe segregated graph $\mathcal{G}(\mathbf{V})$, fix disjoint $\mathbf{Y},\mathbf{A}\subseteq\mathbf{V}$. Similar to above, let $\mathbf{Y}^\star = \operatorname{ant}_{\mathcal{G}_{\mathbf{V}\setminus\mathbf{A}}}(\mathbf{Y})$. Define $\tilde{\mathcal{G}}^d$ and $\tilde{\mathcal{G}}^b$ to be the CADMG and CCG respectively obtained from $\mathcal{G}_{\mathbf{Y}^\star}$ and

$$q(\mathbf{D}^\star) = \frac{p(\mathbf{V})}{\prod_{\mathbf{B}\in\mathcal{B}^{nt}(\mathcal{G})} p(\mathbf{B}|\operatorname{pa}_{\mathcal{G}}(\mathbf{B}))}.$$

We then have $p(\mathbf{Y}|\text{do}(\mathbf{a}))$ is identified in $\mathcal{G}$ if and only if $\mathcal{D}(\tilde{\mathcal{G}}^d)$ is reachable in $\mathcal{G}^d$ and, if it is identified, then it is equal to

$$\begin{aligned}\sum_{\mathbf{Y}^\star\setminus\mathbf{Y}}\Bigg[&\prod_{\mathbf{D}\in\mathcal{D}(\tilde{\mathcal{G}}^d)}\phi_{\mathbf{D}^\star\setminus\mathbf{D}}(q(\mathbf{D}^\star|\operatorname{pa}_{\mathcal{G}}(\mathbf{D}^\star));\mathcal{G}^d)\Bigg]\\ &\times\Bigg[\prod_{\mathbf{B}\in\mathcal{B}(\tilde{\mathcal{G}}^b)} p(\mathbf{B}\setminus\mathbf{A}|\operatorname{pa}_{\mathcal{G}_{\mathbf{Y}^\star}}(\mathbf{B}),\mathbf{B}\cap\mathbf{A})\Bigg]\Bigg|_{\mathbf{A}=\mathbf{a}}.\end{aligned} \tag{7}$$

2.3 Policy Interventions in ADMGs

Extending node interventions, [9] proposed a framework for setting an intervention node in a DAG to a policy, a function of variables preceding it in the graph. Formally, for a DAG $\mathcal{G}(\mathbf{V})$ with a topological ordering $\prec$ on $\mathbf{V}$ and an intervention set $\mathbf{A}\subseteq\mathbf{V}$, let $\mathbf{f}_{\mathbf{A}}$ be the set of policies f_A corresponding to each node A in $\mathbf{A}$. Each f_A is a function of some set $\mathbf{W}_A\subseteq\mathbf{V}_{\prec A}$ such that it maps the state space of $\mathbf{W}_A$ to the state space of A. Graphically, intervening with f_A corresponds to removing all edges *into* A in $\mathcal{G}$ and adding in edges from $\mathbf{W}_A$ to A, yielding a new graph $\mathcal{G}_{\mathbf{f}_{\mathbf{A}}}$.

For an intervention of this type, we can define a counterfactual $Y(\mathbf{f}_{\mathbf{A}})$ for $Y\in\mathbf{V}$ analogously to node interventions via recursive substitution:

$$Y(\{f_A(\mathbf{W}_{\mathbf{A}}(\mathbf{f}_{\mathbf{A}}))|A\in\operatorname{pa}_{\mathcal{G}}(Y)\cap\mathbf{A}\},\{\operatorname{pa}_{\mathcal{G}}(Y)\setminus\mathbf{A}\}(\mathbf{f}_{\mathbf{A}}))$$

This implies a policy-analogue [9] to the g-formula for $p(\{\mathbf{V}\setminus\mathbf{A}\}(\mathbf{f}_{\mathbf{A}}))$:

$$\prod_{V\in\mathbf{V}\setminus\mathbf{A}} p(V|\{f_A(\mathbf{W}_A):A\in\mathbf{A}\cap\operatorname{pa}_{\mathcal{G}}(V)\},\operatorname{pa}_{\mathcal{G}}(V)\setminus\mathbf{A})$$

An extension of these ideas to latent-variable DAGs was given in [8]. Following the Richardson re-expression of the ID algorithm, for an ADMG $\mathcal{G}$, the post-intervention graph $\mathcal{G}_{\mathbf{f_A}}$ is obtained in the same way as the fully observed case: by removing edges into $\mathbf{A}$ and adding edges from $\mathbf{W}_A$ to $\mathbf{A}$. Similarly, [8] defines $\mathbf{Y}^\star \equiv \text{an}_{\mathcal{G}_{\mathbf{f_A}}}(\mathbf{Y}) \setminus \mathbf{A}$. This leads to a policy-analogue of Eq. 6: $p(\mathbf{Y}(\mathbf{f_A}))$ is identified in $\mathcal{G}$ if and only if $p(\mathbf{Y}^\star(\mathbf{a}))$ is identified in $\mathcal{G}$; if it is identified then

$$p(\mathbf{Y}(\mathbf{f_A})) = \sum_{(\mathbf{Y}^\star \cup \mathbf{A}) \setminus \mathbf{Y}} \prod_{\mathbf{D} \in \mathcal{D}(\mathcal{G}_{\mathbf{Y}^\star})} \phi_{\mathbf{V} \setminus \mathbf{D}}(p(\mathbf{V}); \mathcal{G})|_{\tilde{\mathbf{a}}_{\text{pa}^s_{\mathcal{G}}(\mathbf{D}) \cap \mathbf{A}}}$$

where $\tilde{\mathbf{a}}_{\text{pa}^s_{\mathcal{G}}(\mathbf{D}) \cap \mathbf{A}} = \{A = f_A(\mathbf{W}_A) | A \in \text{pa}_{\mathcal{G}}(\mathbf{D}) \cap \mathbf{A}\}$ if $\text{pa}_{\mathcal{G}}(\mathbf{D}) \cap \mathbf{A} \neq \emptyset$ and $\tilde{\mathbf{a}} = \emptyset$ otherwise.

3 Proofs

Lemma 1 *Given a segregated graph $\mathcal{G}(\mathbf{V})$ and a segregation-preserving policy intervention $\mathbf{f_A}(\mathbf{Z_A})$, the post-intervention graph $\mathcal{G}_{\mathbf{f_A}}$ obtained via Procedure 2 is a segregated graph.*

Proof: In order for $\mathcal{G}_{\mathbf{f_A}}$ to be a segregated graph, it must not have a node with both an incident bi-directed and undirected edge (the 'segregation' property) and it must not have any partially directed cycles (the 'chain' property).

We first show that $\mathcal{G}_{\mathbf{f_A}}$ satisfies the segregation property. First we consider edges that appear in both $\mathcal{G}$ and $\mathcal{G}_{\mathbf{f_A}}$ (potentially with a modified functional form). Since we do not add any $\leftrightarrow$ edges when constructing $\mathcal{G}_{\mathbf{f_A}}$, and since we assumed $\mathcal{G}$ is a segregated graph, these edges are all incident to nodes that do not also have incident directed edges.

We can therefore restrict attention to undirected edges that were newly created when constructing $\mathcal{G}_{\mathbf{f_A}}$. These edges correspond to connecting two previously unconnected nodes. This requires intervening on both end points, which entails removing all incident $\leftrightarrow$ edges, as described in Procedure 2. This accounts for all possible undirected edges. In particular, we cannot convert a directed edge $X \to Y$ to an undirected edge $X - Y$: this would require intervening on X with $f_X(\mathbf{Z}_X)$ where $Y \in \mathbf{Z}_X$ which violates our construction that $\mathbf{Z} \subseteq \mathbf{V} \setminus \overline{\text{ext}}_{\mathcal{G}}(X)$.

Since no undirected edge is incident to a node that also has an incident bi-directed edge, $\mathcal{G}_{\mathbf{f_A}}$ satisfies the segregation property.

We now show that $\mathcal{G}_{\mathbf{f_A}}$ satisfies the chain property. We argue by contradiction: suppose $\mathcal{G}_{\mathbf{f_A}}$ *does* have a newly induced (relative to $\mathcal{G}$) partially directed cycle. Then, without loss of generality, one of the following sub-structures appears in $\mathcal{G}_{\mathbf{f_A}}$ but not in $\mathcal{G}$: (1) $W \to X \to Y \to W$, (2) $W \to X - Y \to W$, or (3) $W \to X - Y - W$.

Sub-structure (1) contradicts our assumption that $\mathbf{f_A}$ is segregation-preserving. Specifically, we have that $W \triangle X$ directly and $X \triangle W$ through Y, however $W \notin \mathbf{Z}_Y$.

In sub-structure (2), consider scenarios where two edges were present in $\mathcal{G}$ and we seek to add the third edge. When adding either the $W \to X$ or $Y \to W$ edge, we have that $W \triangle X$ and $X \triangle W$ (analogously for W, Y) but $X \notin \mathbf{Z}_W$ ($W \notin \mathbf{Z}_Y$) which is a contradiction. Meanwhile, adding the $X - Y$ edge requires that $Y \in \mathbf{Z}_X$, however $Y \in \overline{\text{ext}}_{\mathcal{G}}(X)$ which yields a contradiction. A similar argument involving $\triangle$ applies when only one of the three edges was present in $\mathcal{G}$ and we seek to add the other two.

In sub-structure (3) a similar argument applies. Suppose we seek to add the $W \to X$ edge with the two undirected edges present. $X \triangle W$ in the post-intervention graph but it is not the case that $W \triangle X$, yielding a contradiction. Adding the $Y - X$ edge yields a contradiction since $X \in \overline{\text{ext}}_{\mathcal{G}}(Y)$. Similarly, adding the $Y - W$ edge yields a contradiction since $Y \in \overline{\text{ext}}_{\mathcal{G}}(W)$. Again, we can make a similar argument for adding two of the three edges.

The above argument generalizes trivially to larger sub-structures in the graph (e.g., 4-cycles) and so $\mathcal{G}_{\mathbf{f_A}}$ will not have any partially directed cycles. Since $\mathcal{G}_{\mathbf{f_A}}$ satisfies both the chain property and the segregation property, it is a segregated graph. □

Theorem 1 *Let $\mathcal{G}(\mathbf{V} \cup \mathbf{H})$ be a causal LV-CG with $\mathbf{H}$ block-safe, and a topological order $\prec$. Fix disjoint $\mathbf{Y}, \mathbf{A} \subseteq \mathbf{V}$. Let $\mathbf{f_A}(\mathbf{Z_A})$ be a segregation preserving policy set. Let $\mathbf{Y}^\star \equiv \text{ant}_{\mathcal{G}_{\mathbf{f_A}}}(\mathbf{Y}) \setminus \mathbf{A}$. Let $\mathcal{G}^d, \tilde{G}^d$ be the induced CADMGs on $\mathcal{G}_{\mathbf{f_A}}$ and $\mathcal{G}_{\mathbf{Y}^\star}$, and $\tilde{G}^b$ the induced*

CCG on $\mathcal{G}_{\mathbf{Y}^\star}$. Let $q(\mathbf{D}^\star | \operatorname{pa}^s_{\mathcal{G}_{\mathbf{f_A}}}(\mathbf{D}^\star)) = \prod_{\mathbf{D}\in\mathcal{G}_{\mathbf{f_A}}} q(\mathbf{D}| \operatorname{pa}^s_{\mathcal{G}_{\mathbf{f_A}}}(\mathbf{D}))$, where $q(\mathbf{D}| \operatorname{pa}^s_{\mathcal{G}_{\mathbf{f_A}}}(\mathbf{D})) = \prod_{D\in\mathbf{D}} p(D|\mathbf{V}_{\prec D})$ if $\mathbf{D}\cap\mathbf{A} = \emptyset$ and $q = f_A(\mathbf{Z}_A)$ if $\mathbf{D}\cap\mathbf{A}\neq\emptyset$. $p(\mathbf{Y}(\mathbf{f_A}(\mathbf{Z_A})))$ is identified in $\mathcal{G}$ if and only if $p(\mathbf{Y}^\star(\mathbf{a}))$ is identified in $\mathcal{G}$ for the unrestricted class of policies. If identified, $p(\mathbf{Y}(\mathbf{f_A}(\mathbf{Z_A}))) =$

$$\begin{aligned}&\sum_{\{\mathbf{Y}^\star\cup\mathbf{A}\}\setminus\mathbf{Y}}\Bigg[\prod_{\mathbf{B}\in\mathcal{B}(\tilde{\mathcal{G}}^b)} p^\star(\mathbf{B}|\operatorname{pa}_{\mathcal{G}_{\mathbf{f_A}}}(\mathbf{B}))\Bigg]\\ &\times\Bigg[\prod_{\mathbf{D}\in\mathcal{D}(\tilde{\mathcal{G}}^d)}\phi_{\mathbf{D}^\star\setminus\mathbf{D}}(q(\mathbf{D}^\star|\operatorname{pa}^s_{\mathcal{G}_{\mathbf{f_A}}}(\mathbf{D}^\star));\mathcal{G}^d)\Bigg]\Bigg|_{\mathbf{A}=\tilde{\mathbf{a}}}\end{aligned}\tag{8}$$

where (a) $\tilde{\mathbf{a}} = \{A = f_A(\mathbf{Z}_A) : A \in \operatorname{pa}_{\mathcal{G}_{\mathbf{f_A}}}(\mathbf{D})\cap\mathbf{A}\}$ if $\operatorname{pa}_{\mathcal{G}_{\mathbf{f_A}}}(\mathbf{D})\cap\mathbf{A}\neq\emptyset$ and $\tilde{\mathbf{a}}_{\mathbf{D}} = \emptyset$ otherwise, and (b) $p^\star$ is obtained by running Procedure 1 over functions $g_{B_i}(B_{-i}, \operatorname{pa}_{\mathcal{G}_{\mathbf{f_A}}}(B_i), \epsilon_{B_i})$ where $g_{B_i}\in\mathbf{f_A}$ if $B_i\in\mathbf{A}$ and g_{B_i} is given by the observed distribution if $B_i\notin\mathbf{A}$[1].

Proof: We prove two subclaims.

Claim 1: The segregated graph policy ID formula, equation 8, is sound

We first note that each variable in $\mathbf{V}\cup\mathbf{H}$ is defined by a structural equation model. Since $\mathbf{f_A}$ is assumed to be segregation preserving, lemma 1 implies that all variables in $\mathbf{H}$ have an unchanged structural equation in $\mathbf{f_A}$. Among $\mathbf{V}$ there exist two types of variables: those that have a symmetric functional dependence with another variable (i.e., for $V_i, V_j \in \mathbf{V}$ the structural equations f_{V_i}, f_{V_j} are functions of each other), and those without symmetric dependence.

We impose an ordering on the variables in $\mathcal{G}_{\mathbf{f_A}}$ in order of their dependence on other variables in the graph: we first evaluate variables $\mathbf{V}\in(\mathbf{V}\cup\mathbf{H})$ with structural equations that don't depend on other variables ($V \sim f_V(\epsilon_V)$) and then variables that are functions of those variables and so on. Following [2], groups of variables that have symmetrically dependent structural equations are chain components corresponding to $\mathcal{B}^{nt}(\mathcal{G}_{\mathbf{f_A}})$. Variables that do not exhibit symmetric dependence are trivial chain components. Our ordering therefore implies a DAG on chain components (it is acyclic aside from in-component cycles by lemma 1).

It's clear that for trivial chain components the functions f_V immediately reach an equilibrium. We can normalize these functions, and write the margin over their corresponding variables as:

$$\prod_{V\in\mathbf{D}:\mathbf{D}\in\mathcal{D}(\mathcal{G}_{\mathbf{f_A}}(\mathbf{V}\cup\mathbf{H}))} p(V|\operatorname{pa}_{\mathcal{G}_{\mathbf{f_A}}(\mathbf{V}\cup\mathbf{H})}(V))|_{\mathbf{A}=\mathbf{f_A}}$$

Now, for each non-trivial chain component $\mathbf{B}$, the structural equations for each constituent variable treats inputs that are not in the component as known (this can be done since those variables are evaluated earlier in the ordering on the DAG of components) and evaluates each variable in the component via a Gibbs sampling process. The values obtained upon convergence can then be passed to components later in the ordering. This follows by application of proposition 6 in [2], and so we can express the DAG factorization over chain components as:

$$\begin{aligned}p(\mathbf{V}\cup\mathbf{H}(\mathbf{f_A})) = &\prod_{\mathbf{D}\in\mathcal{D}(\mathcal{G}_{\mathbf{f_A}}(\mathbf{V}\cup\mathbf{H}))} p(\mathbf{D}|\operatorname{pa}_{\mathcal{G}_{\mathbf{f_A}}(\mathbf{V}\cup\mathbf{H})}(\mathbf{D}))|_{\mathbf{A}=\mathbf{f_A}}\\ &\times\prod_{\mathbf{B}\in\mathcal{B}^{nt}(\mathcal{G}_{\mathbf{f_A}}(\mathbf{V}\cup\mathbf{H}))} p^\star(\mathbf{B}|\operatorname{pa}_{\mathcal{G}_{\mathbf{f_A}}(\mathbf{V}\cup\mathbf{H})}(\mathbf{B}))\end{aligned}$$

$\mathcal{G}_{\mathbf{f_A}}$ is a proper latent-variable chain graph.

We derive the remainder of the proof via the argument in the proof of theorem 2 in [6]. We assume without loss of generality that $\mathbf{Y}$ has no children in $\mathcal{G}(\mathbf{V})$.

[1]This distribution is identified from univariate terms but it cannot be obtained in closed-form.

Consider the chain graph factorization of $\mathcal{G}_{\mathbf{f_A}}$ derived above. Because $\mathbf{H}$ is block-safe in $\mathcal{G}$, the non-trivial blocks term can be re-written as follows:

$$\begin{aligned}\prod_{\mathbf{B}\in\mathcal{B}^{nt}(\mathcal{G}_{\mathbf{f_A}}(\mathbf{V}\cup\mathbf{H}))} p^\star(\mathbf{B}|\operatorname{pa}_{\mathcal{G}_{\mathbf{f_A}}(\mathbf{V}\cup\mathbf{H})}(\mathbf{B})) &= \prod_{\mathbf{B}\in\mathcal{B}^{nt}(\mathcal{G}_{\mathbf{f_A}}(\mathbf{V}))} p(\mathbf{B}|\operatorname{pa}_{\mathcal{G}_{\mathbf{f_A}}(\mathbf{V})}(\mathbf{B}))\\ &= \prod_{\mathbf{B}\in\mathcal{B}^{nt}(\tilde{\mathcal{G}}^b)} p^\star(\mathbf{B}|\operatorname{pa}_{\mathcal{G}_{\mathbf{f_A}}(\mathbf{V})}(\mathbf{B}))|_{\mathbf{A}=\tilde{\mathbf{a}}_\mathbf{B}}\end{aligned}$$

We are now left with the following factorization for the overall graph:

$$\begin{aligned}p(\{\mathbf{V}\cup\mathbf{H}\}(\mathbf{f_A})) = &\prod_{\mathbf{B}\in\mathcal{B}^{nt}(\tilde{\mathcal{G}}^b)} p^\star(\mathbf{B}|\operatorname{pa}_{\mathcal{G}_{\mathbf{f_A}}(\mathbf{V})}(\mathbf{B}))\\ &\times \prod_{\mathbf{D}\in(\mathbf{V}\cup\mathbf{H})\setminus\left(\bigcup_{\mathbf{B}\in\mathcal{B}^{nt}(\mathcal{G}_{\mathbf{f_A}})}\mathbf{B}\right)} \prod_{V\in\mathbf{D}\setminus\mathbf{A}} p(V|\operatorname{pa}_{\mathcal{G}_{\mathbf{f_A}}(\mathbf{V}\cup\mathbf{H})}(V)) \prod_{V\in\mathbf{D}\cap\mathbf{A}} f_V(\mathbf{Z}_V)|_{\mathbf{A}=\mathbf{f_A}}\end{aligned}$$

The factors in the second term are singleton nodes by construction and so they are defined by either observed $p(V|\operatorname{pa}_{\mathcal{G}_{\mathbf{f_A}}(\mathbf{V}\cup\mathbf{H})}(V))$ if $V\notin\mathbf{A}$ and $f_V\in\mathbf{f_A}(\mathbf{Z}_V)$ if $V\in\mathbf{A}$.

If we marginalize $\mathbf{H}$ from this second set of terms, using standard procedures [10], then the resulting expression is the kernel described in the statement of the theorem: $q(\mathbf{D}^\star|\operatorname{pa}^s_{\mathcal{G}(\mathbf{V})}(\mathbf{D}^\star)) = \prod_{\mathbf{D}\in\mathcal{D}(\mathcal{G}_{\mathbf{f_A}})} q(\mathbf{D}|\operatorname{pa}^s_{\mathcal{G}_{\mathbf{f_A}}(\mathbf{V})}(\mathbf{D}))$, where $q(\mathbf{D}|\operatorname{pa}^s_{\mathcal{G}_{\mathbf{f_A}}(\mathbf{V})}(\mathbf{D})) = \prod_{D\in\mathbf{D}} p(\mathbf{D}|\mathbf{V}_{\prec\mathbf{D}})$ if $\mathbf{D}\cap\mathbf{A}=\emptyset$ and $q(\mathbf{D}|\operatorname{pa}^s_{\mathcal{G}_{\mathbf{f_A}}(\mathbf{V})}(\mathbf{D})) = f_A(\mathbf{Z}_A)$ if $\mathbf{D}\cap\mathbf{A}\neq\emptyset$.

Since $\mathbf{Z_A}$ are all observed by assumption, we can manipulate this kernel as in the proof of soundness for theorem 2 in [6]. Whereas in [6] the authors fixed $\mathbf{A}$ to constants, here we can express setting $\mathbf{A}$ to stochastic values according to $\mathbf{f_A}$. The claim is then immediate.

Claim 2: The segregated graph policy ID formula is complete

We adapt the proof techniques in [8, 6]. At a high level, we will use the fact that $p(\mathbf{Y}^\star(\mathbf{a}))$ is not identified to demonstrate that there is a hedge in $\mathcal{G}$. We will then extend the hedge down the graph to reach $\mathbf{Y}$ via $\operatorname{ext}_{\mathcal{G}_{\mathbf{Y}^\star}}(\text{hedge})$ and $\operatorname{ant}_{\mathcal{G}_{\mathbf{Y}^\star}}(\mathbf{Y})$ to show non-identification. We do this by arguing along the partially directed paths from the hedge to $\mathbf{Y}$, which requires considering subgraphs of $\mathcal{G}_{\mathbf{Y}^\star}$. We show non-identifiability in each of an increasingly restricted submodel of $\mathcal{G}_{\mathbf{Y}^\star}$ and then show that non-identification in the submodels yields non-identification in $\mathcal{G}_{\mathbf{Y}^\star}$. More concretely, there are two complications that must be dealt with for showing completeness of policy interventions: the hedge might intersect $\mathbf{Y}$ and we must extend the hedge down to $\mathbf{Y}$ via partially directed paths. We construct a subgraph for demonstrating the latter case and then a subgraph of that for the former case. We now proceed with the proof.

Suppose $p(\mathbf{Y}^\star(\mathbf{a}))$ is not identified in $\mathcal{G}$. Then there is a district $\mathbf{D}\in\mathcal{D}(\mathcal{G}_{\mathbf{Y}^\star})$ that is not reachable in $\mathcal{G}$. Let $\mathbf{R}=\{D\in\mathbf{D}|\operatorname{ch}_\mathcal{G}(D)\cap\mathbf{D}=\emptyset\}$. Let $\mathbf{A}^\star=\mathbf{A}\cap\operatorname{pa}_\mathcal{G}(D)$. Then there exists $\mathbf{D}'\supset\mathbf{D}$, such that $\mathbf{D}$ and $\mathbf{D}'$ form a hedge for $p(\mathbf{R}|\text{do}(\mathbf{a}^\star))$ and thus $p(\mathbf{R}|\text{do}(\mathbf{a}^\star))$ is not identified by [7].

Let $\mathbf{Y}'$ be the minimal subset of $\mathbf{Y}$ such that $\mathbf{R}\subseteq\operatorname{ant}_{\mathcal{G}_{\mathbf{f_A}}}(\mathbf{Y}')$. Consider a subgraph $\mathcal{G}^\dagger$ of $\mathcal{G}_{\mathbf{f_A}}$, with vertices $\mathbf{V}'\subseteq\mathbf{V}$, consisting of all edges in $\mathcal{G}$ in the hedge on $\mathbf{D},\mathbf{D}'$ described above, and edges that lie in partially directed paths in $\mathcal{G}_{\mathbf{f_A}}$ from $\mathbf{R}$ to $\mathbf{Y}'$. We restrict attention, without loss of generality, to at most one child per node in each partially directed path such that our paths form a forest from $\mathbf{R}$ to $\mathbf{Y}'$. By Lemma 1, $\mathcal{G}^\dagger$ does not contain any directed, nor partially directed cycles. Let $\mathbf{A}^\dagger=\{\mathbf{A}^\star\cup A|A\in\mathbf{A}\text{ in }\mathcal{G}^\dagger\}$. For each $A^\dagger\in\mathbf{A}^\dagger$, we restrict attention to policies that map from $\mathbf{Z}^\dagger_{A^\dagger}$ to $A^\dagger$, where $\mathbf{Z}^\dagger_{A^\dagger}=\mathbf{Z}_{A^\dagger}\cap\mathbf{V}'$.

Now, following the proof of theorem 2 in the supplement of [6], we define an ADMG $\tilde{\mathcal{G}}^\dagger$ which has the same vertices and edges as the $\mathbf{D},\mathbf{D}'$ hedge in $\mathcal{G}^\dagger$, and has a copy of each vertex in each partially directed path from $\mathbf{R}$ to $\mathbf{Y}'$ in $\mathcal{G}^\dagger$ but replaces all the undirected edges on those partially directed paths with directed edges oriented away from $\mathbf{R}$ towards $\mathbf{Y}'$. We denote the variable copies in $\tilde{\mathcal{G}}^\dagger$ corresponding to $\mathbf{Y}'$ in $\mathcal{G}^\dagger$ by $\tilde{\mathbf{Y}}'$. This orientation is possible because each undirected edge either corresponds to a (known) policy in the intervention set, or to an observed structural equation. In either case, the observed distribution continues to argree between the two counterexamples witnessing non-identifiability. For $\mathbf{A}^\dagger$ in $\tilde{\mathcal{G}}^\dagger$, we further restrict attention to policies inducing directed edges from $\mathbf{R}$ to $\tilde{\mathbf{Y}}'$ (i.e. ignoring policies going the opposite direction that induce undirected edges). We denote these nodes by $\tilde{\mathbf{A}}^\dagger$.

We now show that $p(\tilde{\mathbf{Y}}'(\{\tilde{A}^\dagger = f_{\tilde{A}^\dagger} | \tilde{A}^\dagger \in \tilde{\mathbf{A}}^\dagger\}))$ is not identified in $\tilde{\mathcal{G}}^\dagger$ following the argument in the proof of theorem 6 in the supplement of [8]. Observe that for $\mathbf{R} \subseteq \tilde{\mathbf{Y}}'$, the subclaim is immediate by the recursive argument in the proof of theorem 4 in [8]. Otherwise, pick a node $\tilde{Y}'$ in $\tilde{\mathcal{G}}^\dagger$ such that $\mathrm{pa}_{\tilde{\mathcal{G}}^\dagger}(\tilde{Y}') \subseteq \mathbf{R}$ and $\mathrm{pa}_{\tilde{\mathcal{G}}^\dagger}(\tilde{Y}') \setminus \tilde{\mathbf{Y}}' \neq \emptyset$ (as in [8], such a vertex must exist since $\tilde{\mathcal{G}}^\dagger$ is acyclic and $\mathbf{R} \setminus \tilde{\mathbf{Y}}' \neq \emptyset$). If this $\tilde{Y}' \in \tilde{\mathbf{A}}^\dagger \setminus \mathbf{A}^\star$, the subclaim is immediate since $\tilde{Y}'$ does not intersect our hedge and we can extend down the graph using the argument in theorem 4 of [8].

If $\tilde{Y}' \in \mathbf{A}^\star$ then we can create a graph $\bar{\mathcal{G}}$ by copying the variables on the path $\tilde{Y}' \to V_1 \to \cdots \to \bar{Y} \in \tilde{\mathbf{Y}}'$ in $\tilde{\mathcal{G}}^\dagger$. We then apply the argument in theorem 4 of [8] to show that $p(\bar{Y}(\mathbf{a}^\star))$ is not identified along this path when we set $\mathbf{a}^\star$ according to the policies specified by $\mathbf{f}_{\mathbf{A}^\star}$. This follows since, by assumption, $\mathbf{f}_{\mathbf{A}^\star} \subseteq \mathbf{f}_{\mathbf{A}}$ lies in an unrestricted policy class. Now, as $p(\bar{Y}(\mathbf{a}^\star))$ is not identified in $\bar{\mathcal{G}}$, we can use the two counterexamples witnessing non-identifiability in $\bar{\mathcal{G}}$ to obtain non-identifiability for $p(\tilde{\mathbf{Y}}'(\mathbf{f}_{\tilde{\mathbf{A}}^\dagger}))$. To do so, we define new variables in $\tilde{\mathcal{G}}^\dagger$ that are the Cartesian product of variable copies created in $\bar{\mathcal{G}}$ and their corresponding variables in $\tilde{\mathcal{G}}^\dagger$. Non-identifiability follows via the standard argument in lemma 1 of [8].

Now that we have shown that $p(\tilde{\mathbf{Y}}'(\{\tilde{A}^\dagger = f_{\tilde{A}^\dagger} | \tilde{A}^\dagger \in \tilde{\mathbf{A}}^\dagger\}))$, we have two counterexamples witnessing non-identifiability in $\tilde{\mathcal{G}}^\dagger$ which agree on the observed data distribution but disagree on the counterfactual distribution. We use these counterexamples to demonstrate non-identifiability of $p(\mathbf{Y}'(\{A = f_A | A \in \mathbf{A}^\star\}))$ in $\mathcal{G}^\dagger$. To do so, we define variables along the partially directed paths from $\mathbf{R}$ to $\mathbf{Y}'$ in $\mathcal{G}^\dagger$. These variables are created by taking the Cartesian product of variable copies in $\tilde{\mathcal{G}}^\dagger$ and the corresponding variables in $\mathcal{G}^\dagger$. As before, the counterexamples continue to agree on the observed data distribution and disagree on the counterfactual distribution. Thus $p(\mathbf{Y}'(\{A = f_A | A \in \mathbf{A}^\star\}))$ is not identified in $\mathcal{G}^\dagger$. Since $\mathbf{Y}' \subseteq \mathbf{Y}^\star$, the result is immediate, subject to the remaining argument on the chain graph properties of $\mathcal{G}^\dagger$ and $\tilde{\mathcal{G}}^\dagger$ below.

Following [6], fix a block $\mathbf{B}$ in $\mathcal{G}^\dagger$. For any $B \in \mathbf{B}$, there exists a set of variables $B_1, \ldots, B_k$ in $\tilde{\mathcal{G}}^\dagger$ such that B is defined as the Cartesian product of $B_1, \ldots, B_k$. Any variable $A \in \mathrm{nb}_{\mathcal{G}^\dagger} \cup \mathrm{pa}_{\mathcal{G}^\dagger}(B)$ is similarly a Cartesian product of A variables. Then it follows that $B \perp\!\!\!\perp ((\mathrm{pa}_{\mathcal{G}^\dagger}(\mathbf{B}) \cup \mathbf{B}) \setminus (\mathrm{nb}_{\mathcal{G}^\dagger}(B) \cup \mathrm{pa}_{\mathcal{G}^\dagger}(B))) | (\mathrm{nb}_{\mathcal{G}^\dagger}(B) \cup \mathrm{pa}_{\mathcal{G}^\dagger}(B))$ by d-separation rules in the ADMG $\tilde{\mathcal{G}}^\dagger$ and that there are no colliders in $\tilde{\mathcal{G}}^\dagger$. These both follow from our vertex copy argument which separates out B from the rest of the block and eliminates the possibility of colliders by making every path from $\mathbf{R}$ to $\mathbf{Y}'$ a partially directed chain. This demonstrates that $\mathcal{G}^\dagger$ and $\tilde{\mathcal{G}}^\dagger$ (and trivially $\bar{\mathcal{G}}$) satisfy the independence constraints implied by the CG Markov property, thus proving the claim. □

4 Derivation of the Figure 2 Functional

From Fig. 2(a), we obtain $\mathcal{G}_{\mathbf{f}_\mathbf{A}}$ in Fig. 2(b) by applying the intervention detailed in Table 2. In turn, from this post-intervention graph we observe that $\mathbf{Y}^\star = \mathrm{ant}_{\mathcal{G}_{\mathbf{f}_\mathbf{A}}}(\mathbf{Y}) \setminus \mathbf{A} = \{C_2, C_3, M_3, Y_2, Y_3\}$ and obtain the induced subgraph $\mathcal{G}_{\mathbf{Y}^\star}$ in Fig. 2(c).

$\mathcal{G}_{\mathbf{Y}^\star}$ factorizes into kernels relating to district nodes and block nodes: $q_{\mathcal{D}}(C_1, A_1, M_1, Y_1, Y_2, Y_3 | C_2, M_2, M_3)$ and $q_{\mathcal{B}}(M_2, M_3, A_2, A_3, C_2, C_3 | \emptyset)$. The block nodes factorize as a product of blocks, as in the first term of Eq. 8:

$$\begin{aligned} q_{\mathcal{B}}(\mathbf{B}^\star | \mathrm{pa}_{\mathcal{G}_{\mathbf{f}_\mathbf{A}}}(\mathbf{B}^\star)) &= \prod_{\mathbf{B} \in \mathcal{B}(\tilde{\mathcal{G}}^b)} p^\star(\mathbf{B} | \mathrm{pa}_{\mathcal{G}_{\mathbf{f}_\mathbf{A}}}(\mathbf{B})) \\ &= p^\star(M_2, M_3 | A_2, A_3, C_2) p^\star(A_2, A_3 | C_2, C_3) p^\star(C_2, C_3) \end{aligned}$$

Note that $p^\star(C_2, C_3) = p(C_2, C_3)$ since the $C_2 - C_3$ block is unchanged relative to the observed data.

Separately, we must fix sets for each $\mathcal{G}_{\mathbf{Y}^\star}$ district $\{\{M_3\},\{Y_2,Y_3\}\}$ in $q_{\mathcal{D}(\mathcal{G})}$. The derivations of these pieces is as follows:

$\phi_{\mathbf{D}^\star\setminus\{M_3\}}(q(C_1,A_1,M_1,Y_1,Y_2,Y_3|C_2,M_2,M_3);\mathcal{G}^d) = \phi_{\mathbf{D}^\star}(q(C_1,A_1,M_1,Y_1,Y_2,Y_3|C_2,M_2,M_3);\mathcal{G}^d)$

This follows since M_3 is already fixed in this kernel and subgraph. Since we must fix all variables in the kernel and all variables in the kernel are fixable, this term simplifies to $p(\emptyset)=1$.

For the second kernel, we have: $\phi_{\mathbf{D}^\star\setminus\{Y_2,Y_3\}}(q(C_1,A_1,M_1,Y_1,Y_2,Y_3|C_2,M_2,M_3);\mathcal{G}^d)$

$$\begin{aligned}
&= \phi_{C_1,A_1,M_1,Y_1}(q(C_1,A_1,M_1,Y_1,Y_2,Y_3|C_2,M_2,M_3);\mathcal{G}^d)\\
&= \phi_{A_1,M_1,Y_1}(\frac{q(C_1,A_1,M_1,Y_1,Y_2,Y_3|C_2,M_2,M_3)}{p(C_1)};\phi_C(\mathcal{G}^d))\\
&= \phi_{A_1,M_1,Y_1}(q(A_1,M_1,Y_1,Y_2,Y_3|C_2,M_2,M_3,C_1);\phi_{C_1}(\mathcal{G}^d))\\
&= \phi_{M_1,Y_1}(\frac{q(A_1,M_1,Y_1,Y_2,Y_3|C_2,M_2,M_3,C_1)}{p(A_1|C_1)};\phi_{C_1,A_1}(\mathcal{G}^d))\\
&= \phi_{M_1,Y_1}(q(M_1,Y_1,Y_2,Y_3|C_2,M_2,M_3,C_1,A_1);\phi_{C_1,A_1}(\mathcal{G}^d))\\
&= \phi_{Y_1}(\frac{q(M_1,Y_1,Y_2,Y_3|C_2,M_2,M_3,C_1,A_1)}{p(M_1|A_1)};\phi_{C_1,A_1,M_1}(\mathcal{G}^d))\\
&= \phi_{Y_1}(q(Y_1,Y_2,Y_3|C_2,M_2,M_3,C_1,A_1,M_1);\phi_{C_1,A_1,M_1}(\mathcal{G}^d))\\
&= \frac{q(Y_1,Y_2,Y_3|C_2,M_2,M_3,C_1,A_1,M_1)}{p(Y_1|A_1,M_1)};\phi_{C_1,A_1,M_1,Y_1}(\mathcal{G}^d))\\
&= p(Y_2,Y_3|C_1,C_2,M_1,M_2,M_3,A_1,Y_1)
\end{aligned}$$

This yields the functional for $p(\{Y_2,Y_3\}(\mathbf{f_A}))$:

$$\sum_{\{A_1,A_2,A_3,M_2,M_3,C_2,C_3\}}\Big(p^\star(M_2,M_3|A_2,A_3,C_2)p^\star(A_2,A_3|C_2,C_3)p^\star(C_2,C_3)\\ \times p(Y_2,Y_3|C_1,C_2,M_1,M_2,M_3,A_1,Y_1)\Big)$$

5 Experimental Details and Extended Results

Each C_i, A_i, Y_i are generated according to the following densities (note that C_i is a 3-dimensional vector):

$$\begin{aligned}
C_{i,j} &\sim \text{Beta}(\alpha_j,\beta_j)\\
p(A_i=1|C_i,C_{-i}) &= \text{expit}(\sum_{j=1}^{3}\gamma_j C_{i,j} + \frac{\tau_{AC}}{|\mathcal{N}_i|}\sum_{k\in\mathcal{N}_i}\sum_{j=1}^{3}C_{k,j})\\
p(Y_i=1|A_i,A_{-i},C_i,C_{-i},Y_{-i}) &= \text{expit}\Big(\eta A_i + \sum_{j=1}^{3}\delta_j C_{i,j}\\
&\quad + \frac{1}{|\mathcal{N}_i|}\sum_{k\in\mathcal{N}_i}\Big(\tau_{YA}A_k + \tau_{YY}Y_k + \sum_{j=1}^{3}\tau_{YC}C_{k,j}\Big)\Big)
\end{aligned}$$

where $\mathcal{N}_i$ denote unit i's neighbors in $\mathcal{G}_{\mathbf{f_A}}$.

The parameters for the Beta distribution for C for both types of experiments (policy and bias) are given by:

The parameters for A_i and Y_i differ between the bias and policy experiments. For A we have:

And for Y we have:

Finally, for the policy experiment we have results similar to those in the main draft, which demonstrate the efficacy of policy interventions in selection actions that yield a more optimal outcome.

α	β
1.5	3
6	2
.8	.8

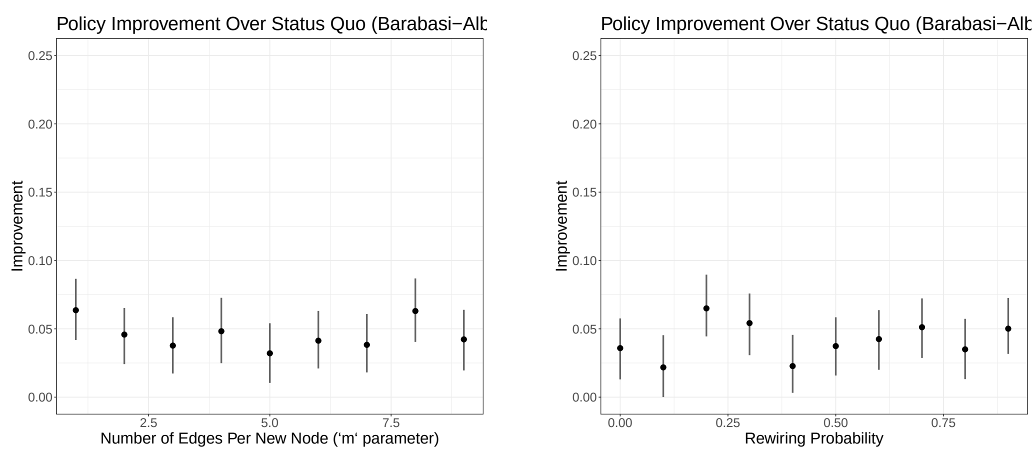
Table 1: Parameters for generating C_i

Parameter	Bias	Policy
γ_1	1	.5
γ_2	0	.2
γ_3	0	.25
τ_{AC}	0	.15

Table 2: Parameters for generating A_i

Parameter	Bias	Policy
η	-3	.6
δ_1	1	-.3
δ_2	0	.4
δ_3	0	.1
τ_{YA}	3	.2
τ_{YY}	.1	.3
τ_{YC}	0	-.2

Table 3: Parameters for generating Y_i



(a) (b)

Figure 1: Difference in expected outcomes between adopting an optimal strategy and using the status quo strategy in the Barabási-Albert model 1a and the Watts-Strogatz small world model 1b. We perform these analyses at several network densities to demonstrate the general efficacy of this approach.

References

[1] S. L. Lauritzen. *Graphical models*, volume 17. Clarendon Press, 1996.

[2] S. L. Lauritzen and T. S. Richardson. Chain graph models and their causal interpretations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(3):321–348, 2002.

[3] J. Pearl. *Causality: models, reasoning and inference*, volume 29. Springer, 2000.

[4] T. S. Richardson, R. J. Evans, J. M. Robins, and I. Shpitser. Nested markov properties for acyclic directed mixed graphs. *arXiv preprint arXiv:1701.06686*, 2017.

[5] J. Robins. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical modelling*, 7(9-12):1393–1512, 1986.

[6] E. Sherman and I. Shpitser. Identification and estimation of causal effects from dependent data. In *Advances in Neural Information Processing Systems*, pages 9424–9435, 2018.

[7] I. Shpitser and J. Pearl. Identification of joint interventional distributions in recursive semi-markovian causal models. In *Proceedings of the National Conference on Artificial Intelligence*, volume 21, page 1219. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006.

[8] I. Shpitser and E. Sherman. Identification of personalized effects associated with causal pathways. In *Proceedings of the 34th Conference on Uncertainty in Artificial Intelligence*, 2018.

[9] J. Tian. Identifying dynamic sequential plans. *Proceedings of the 24th Conference on Uncertainty in Artificial Intelligence*, 2008.

[10] J. Tian and J. Pearl. A general identification condition for causal effects. In *Proceedings of the National Conference on Artificial Intelligence*, pages 567–573, 2002.