Intervening on Network Ties

Eli Sherman
Department of Computer Science
Johns Hopkins University
Baltimore, MD

Ilya Shpitser
Department of Computer Science
Johns Hopkins University
Baltimore, MD

Abstract

A foundational tool for making causal inferences is the emulation of randomized control trials via variable interventions. This approach has been applied to a wide variety of contexts, from health to economics [3, 6]. Variable interventions have long been studied in independent and identically distributed (iid) data contexts, but recently non-iid settings, such as networks with interacting agents [8, 17, 28] have attracted interest. In this paper, we propose a type of *structural intervention* [12] relevant in network contexts: the *network intervention*. Rather than estimating the effect of changing variables, we consider changes to social network structure resulting from creation or severance of ties between agents. We define the *individual participant* and *average bystander effects* for these interventions and describe identification criteria. We then prove a series of theoretical results that show existing identification theory obtains minimally KL-divergent distributions corresponding to network interventions. Finally, we demonstrate estimation of effects of network interventions via a simulation study.

1 INTRODUCTION

Practitioners in applied fields such as medicine, epidemiology, and economics often seek *causal* understanding of the processes they observe. In turn, an in-depth causal understanding can inform decisions and improve policy. The gold standard for obtaining this understanding is the randomized control trial (RCT). Unfortunately, RCTs are often infeasible due to expense or ethical concerns. The field of causal inference provides a framework for emulating RCTs in such situations, using observational data.

A fundamental element of the causal framework is the notion of interventions. Researchers select one or more 'treatment' variables and outcomes of interest. The value of the outcome of interest is estimated under the hypothetical scenario in which the value of the treatment variable is changed to a specific, researcher-chosen value.

A number of assumptions are required for such estimates to have a valid causal interpretation. A common assumption is that data samples are independent and identically distributed (iid), which permits the application of conventional statistical methods. Nevertheless, it is easy to think of cases where this assumption does not hold. Here we focus on domains pertaining to networks of interacting study subjects such as infectious disease spread and social networks. Recently, several papers have proposed methods for obtaining inferences from dependent data [28, 24, 17, 16]. As in the iid setting, these papers emulate RCTs by intervening on variables and estimating the effects on downstream outcomes.

Unfortunately, existing methods for network inference are ill-suited to consider more general changes to the network. For instance, in urban development economics authors have proposed housing vouchers as a 'treatment' to incent families to move to neighborhoods with greater opportunity for upward social mobility [3]. Evaluating the effect of extracting a family from one neighborhood and placing them in a new neighborhood with new social connections isn't possible by considering changes to values of variables alone: the network itself changes.

In this paper we extend the classical causal inference framework to consider changes to social network structure. First, we review different network representations in the causal inference literature as well as notions of interventions that have departed from conventional variable interventions. Next, we give a motivating example based on the global political economy. Extending [12], we propose *network interventions*; interventions on the *structure* of a network where ties between units are formed or

broken. We define the *individual participant* and *average bystander* effects of these interventions, analogous to the network effects described in [8] and discuss identification. We then demonstrate that post-severance distributions satisfy independence constraints for the severed units while remaining minimally KL-divergent from the pre-intervention distributions. Finally, we demonstrate estimation of network intervention effects from observational data via a simulation study.

2 REVIEW

Causal Networks

Since networks are of interest to a variety of fields, there are numerous representations, each with their own advantages and limitations. These representations were developed as a means of studying interference – the phenomenon that arises when neighbors' treatments causally affect each other's outcomes. While the present work doesn't focus explicitly on interference, we discuss it here since our work is complementary to that literature.

A widely used approach, characterized in [17], represents networks with directed acyclic graphs (DAGs), where network connections appear as directed edges from one individual's variables to another's. This approach lends a natural causal interpretation that follows from a rich literature on causal DAGs. Importantly, the relationships between individuals are encoded in the functional relationships represented by edges connecting different units; when two individuals are not friends, edges will be absent.

Recent work [28, 24, 15] advocates representing networks with Lauritzen-Wermuth-Freydenburg (LWF) chain graphs (CGs), which were given a causal interpretation in [11]. CGs extend DAGs by permitting representation of symmetric relationships (i.e. stable-state equilibria) via undirected edges. [15] argued that CGs under the LWF interpretation can approximate feedback processes when those processes are slow. While CGs provide a more general representation, their interpretation in the context of the present work is somewhat complicated.

Beyond these notions, there is a substantial literature on probabilistic relational models [9, 5]. These models generalize conventional graphical models by employing first-order logic to describe the nature of relationships between entities. These models have been extended to causal inference in network settings [1], however, similar to chain graphs, their use in public health contexts like those considered here is not yet well established. For these reasons, we will restrict attention to graphical models.

Aside from graphical representations, a large subset of the interference literature formalizes inter-unit relationships algebraically as in [8]. Many of these formulations could be reformulated using graphical models.

Structural Interventions

The majority of the causal inference literature has focused on hypothetical experiments wherein interventions are made upon *variables* (e.g. smoking status) and the effects of interventions are considered with respect to some outcome (e.g. lung cancer). The two dominating frameworks, the Neyman-Rubin potential outcomes framework [14, 23] and Pearl's graph-based framework [18] differ primarily in their philosophical approach, and recently researchers have begun to use their terminology and mechanics interchangeably. See, for example [20]. The causal interpretation of variable interventions under these (and other) frameworks is the subject of literature at the intersection of applied fields and the philosophy of science. These discussions are broad and have a lengthy history. For an incomplete survey, we encourage the interested reader to consult the works of Halpern, Tian, and Pearl [30, 7], and Woodward [31].

In the past two decades, there has been a movement towards defining more general notions of intervention. Korb proposed several generalizations, including interventions that are stochastic with respect to the treatment variable [10]. Eberhardt and Scheines discussed similar ideas, contrasting 'hard' and 'soft' interventions, corresponding to changing causal structure (e.g. removing edges) and parametric form respectively [4]. They proposed using this continuum of interventions to aid in causal discovery efforts (see also [30]). Malinsky proposed a framework for considering the effects of changes to the structure of a causal model on a 'macro level' [12]. In this framework, one modifies structural equations or manipulates parameters in order to evaluate counterfactuals pertaining to the world in which macro level features are different. Finally, unrelated to philosophy, [16] proposes a type of edge intervention in social networks as a means of understanding changes in network ties. In addition, interventions on paths and edges were considered in the context of mediation analysis in [26]. In the current work, we build on these ideas to evaluate general interventions on network ties, enabling us to envision the counterfactual world in which two units are severed or connected.

3 MOTIVATING EXAMPLE: THE POLITICAL ECONOMY

In this section we give a motivating example: a model of trade relations between countries, in which network interventions provide a means of understanding counterfactual changes to network structure.
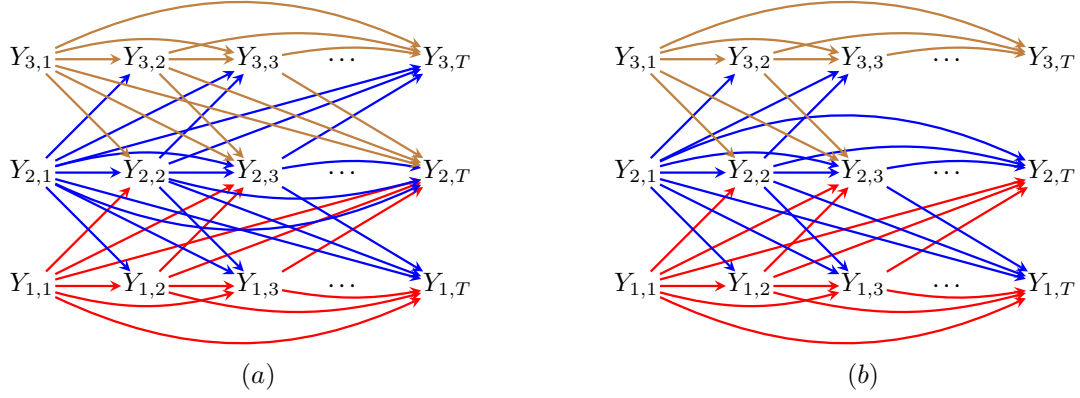
Figure 1: (a) A DAG representing time series cross sectional data on three countries where country 2 has a trade agreement with countries 1 and 3; (b) the DAG in (a) after an intervention is performed, severing the alliance between countries 2 and 3 at $t = 3$.

In global economics, policies made by one country, such as treaties, trade deals, and tariffs, have a direct impact on the nations geographically and diplomatically connected to the policymaker. In light of current events, we refer to the interventions represented in Fig. 1 as the 'Brexit' scenario (for severing two or more countries) and the 'Turkey joins the EU' scenario (for connecting two or more countries). These types of temporal DAG models, also known as dynamic Bayes nets [13], correspond to time series cross sectional data from the political economy literature [2]. Each country i is represented by temporally sequential observations $Y_{i,1}, Y_{i,2}, \ldots Y_{i,T}$ where each Y is a vector of economic variables (GDP, unemployment rate, open-market funds rate, etc.).

As a generalization of Fig. 1, we can imagine the network having several countries, each with multiple neighbors. We can then consider the hypothetical effect of a 'clean break' at time t between one country and some or all of its neighbors (introducing non-stationarity [22]). This is represented by moving from Fig. 1 (a) to Fig. 1 (b) by severing the connection between countries 2 and 3 at $t = 3$. We can also consider the reverse intervention, where two previously unconnected countries are connected, corresponding to the signing of a trade agreement. Graphically, this corresponds to moving from Fig. 1 (b) to (a). Using the framework we propose in this paper, a decision-maker could evaluate these hypothetical policies *prior to* implementation and ensure that they have the intended effect.

4 REPRESENTING NETWORKS WITH CAUSAL DAGs

Throughout this paper, we will consider performing causal inference in social networks represented by DAGs. In

this section we formalize the notation needed to define network interventions and their associated effects.

Causal DAG Prerequisites

We follow standard notation from the probabilistic graphical models literature with vertices and random variables used interchangeably. We will represent random variables and their realizations via capital letters (V) and lowercase letters (v) respectively. Sets of variables and their realizations will be denoted in boldface ($\mathbf{V}$ and $\mathbf{v}$). We also define shorthand notation for standard graphical notions: parents, $\mathrm{pa}_{\mathcal{G}}(V) \equiv \{W | W \to V\}$; ancestors, $\mathrm{an}_{\mathcal{G}}(V) \equiv \{W | W \to \cdots \to V\}$; children, $\mathrm{ch}_{\mathcal{G}}(V) \equiv \{W | V \to W\}$; descendants $\mathrm{de}_{\mathcal{G}}(V) \equiv \{W | V \to \cdots \to W\}$. Each of these can be generalized disjunctively to sets: $\mathrm{pa}_{\mathcal{G}}(\mathbf{S}) \equiv \cup_{V \in \mathbf{S}} \mathrm{pa}_{\mathcal{G}}(V)$. We also define non-descendants $\mathrm{nd}_{\mathcal{G}}(\mathbf{S}) = \mathbf{V} \setminus \mathrm{de}_{\mathcal{G}}(\mathbf{S})$. Finally $\mathfrak{X}_V$ denotes the state space of the variable V.

A statistical DAG model $\mathcal{G}$ with vertices $\mathbf{V}$ is associated with a set of probability distributions on random variables in $\mathbf{V}$ that satisfy the factorization: $p(\mathbf{V}) = \prod_{V \in \mathbf{V}} p(V | \mathrm{pa}_{\mathcal{G}}(V))$.

In such models, the absences of edges between variables, relative to a complete DAG, encode independences. These correspond to the local Markov property of DAGs: $\mathbf{X} \perp\!\!\!\perp \mathrm{nd}_{\mathcal{G}}(\mathbf{X}) \setminus \mathrm{pa}_{\mathcal{G}}(\mathbf{X}) | \mathrm{pa}_{\mathcal{G}}(\mathbf{X})$.

Extending statistical DAGs, a *causal* DAG model $\mathcal{G}$ with vertices $\mathbf{V}$ is associated with a set of distributions on *counterfactual* variables in $\mathbf{V}$. For $Y \in \mathbf{V}$ and $\mathbf{A} \subset \mathbf{V} \setminus Y$, a counterfactual $Y(\mathbf{a})$ describes the value of Y under the hypothetical scenario in which $\mathbf{A}$ is set to $\mathbf{a}$ via an intervention [18]. We will describe generalizations to this convention in the next section.

In this paper, we assume Pearl's *functional model* of a DAG $\mathcal{G}(\mathbf{V})$. Under this model, if $\mathbf{a}$ are the values of the parents of $V \in \mathbf{V}$, then $V(\mathbf{a})$ is determined by a structural equation $f(\mathbf{a}, \epsilon_V)$ where f is invariant to changes to the values of $\mathbf{a}$ and ϵ_V is an error term. We will further assume that there are no hidden variables in the models discussed in this paper. Relaxing this assumption for network interventions is the subject of future work.

The above counterfactuals, often referred to as *one step ahead* counterfactuals, permit us to describe all variables in the model via *recursive substitution*:

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

where $\mathbf{A} \in \mathbf{V}$ and $\mathbf{a} \in \mathfrak{X}_{\mathbf{A}}$.

A parameter in a model is said to be *identified* if it is expressable as a function of the observed data. In causal DAGs with no unobserved variables, all counterfactual distributions $p(\mathbf{V}(\mathbf{a}))$ are identified by the *g-formula* [21]:

$$p(\{W(\mathbf{a}) : W \in \mathbf{V} \setminus \mathbf{A}\}) = \prod_{W \in \mathbf{V} \setminus \mathbf{A}} p(W | \mathrm{pa}_{\mathcal{G}}(W))|_{\mathbf{A}=\mathbf{a}}$$

As an example, consider a single-unit version of Fig. 2. If we are interested in the effect of setting $A = a$, the interventional distribution $p(\mathbf{V}(a))$ is given by $p(Y|A = a, C)p(C)$.

While classical interventions set a variable to a value, we are often interested in how an intervention affects an outcome along multiple pathways, such as the separate effects of smoking, smoke inhalation and nicotine, on a patient's risk of lung cancer. In these cases, it is natural to think of interventions in which we intervene on the treatment node with different values for each edge out of the node. For instance, we might consider setting smoking status to 0 for the sake of the smoke inhalation edge and to a reference value for the nicotine exposure edge, corresponding to having the patient smoke e-cigarettes.

Formally, for a set of treatment variables $\mathbf{A}$ the set of edges out of $\mathbf{A}$ is denoted by α. Interventions are performed with a multiset $\mathfrak{a}_\alpha$ which maps edges to constant values for A or to the natural value of A for each $A \in \mathbf{A}$. As with node interventions, for $\mathbf{A}_\alpha = \{A | (AB)_\rightarrow \in \alpha\}$, where $(AB)_\rightarrow \in \alpha$ signifies that an edge $A \rightarrow B$ is in α, edge interventions given by $p(\{W(\mathfrak{a}_\alpha) : W \in \mathbf{V} \setminus \mathbf{A}_\alpha\})$ are identified by the edge g-formula [27] with $\mathrm{pa}_{\mathcal{G}}^{\bar{\alpha}}(V) = \{W | (WV)_\rightarrow \notin \alpha\}$:

$$\prod_{W \in \mathbf{V} \setminus \mathbf{A}_\alpha} p(W | \mathfrak{a}_{(ZW)_\rightarrow \in \alpha}, \mathrm{pa}_{\mathcal{G}}^{\bar{\alpha}}(W)). \tag{1}$$

If we again consider a single-unit version of Fig. 2, when we intervene with $\mathfrak{a}_\alpha = \{(CA)_\rightarrow = c, (CY)_\rightarrow = c'\}$, the distribution $p(\{W(\mathfrak{a}_\alpha) : W \in \mathbf{V} \setminus \mathbf{A}_\alpha\})$ is given by $p(Y|A, C = c')p(A|C = c)$.

As an alternative generalization to classical interventions, we might be interested in customizing treatments according to unit-specific characteristics. For instance, we might want to choose a cancer patient's chemotherapy regimen according to the specific characteristics of their tumor. Rather than setting treatments to fixed values, we set them to analyst-specified functions of pre-treatment covariates. This type of policy intervention is the subject of the dynamic treatment regime (DTR) literature [29, 25].

Formally, for a set of treatment variables $\mathbf{A}$, the set of pre-treatment covariates we wish to use to set each $A \in \mathbf{A}$ is denoted $\mathbf{C}_A$. Policy interventions entail setting $\mathbf{A}$ to the set of functions $\mathbf{f}_{\mathbf{A}}$, where $f_A \in \mathbf{f}_{\mathbf{A}}$ maps $\mathfrak{X}_{\mathbf{C}_A} \rightarrow \mathfrak{X}_A$. Responses to policy interventions, $p(\{W(\mathbf{f}_{\mathbf{A}}) : W \in \mathbf{V} \setminus \mathbf{A}\})$, are identified by the policy g-formula [25]:

$$\prod_{W \in \mathbf{V} \setminus \mathbf{A}} p(W | \{f_A(\mathbf{C}_A) : A \in \mathbf{A} \cap \mathrm{pa}_{\mathcal{G}}(W)\}, \mathrm{pa}_{\mathcal{G}}(W) \setminus \mathbf{A}) \tag{2}$$

Continuing with our single-unit example for Fig. 2, suppose we are interesting in setting A to a policy that is a function of C: $f_A(C)$. Then the counterfactual distribution $p(\mathbf{V}(f_A(C)))$ is given by $p(Y|A = f_A(C), C)p(C)$.

DAG Representation of Network Data

In this paper we will represent networks of interacting agents with DAGs following [17]. We will assume each network $\mathcal{G}$ is associated with a probability distribution $p(\mathbf{V})$ and that $\mathcal{G}$ has a causal interpretation as described above. Denote the set of agents ('units' or 'subjects') in $\mathcal{G}$ by $\mathcal{A}$. $\mathcal{G}$ can be partitioned into sub-graphs $\mathcal{G}_i$ with variables $\mathbf{V}_i \subset \mathbf{V}$ for each agent $i \in \mathcal{A}$. The marginal distribution for agent i is therefore denoted $p(\mathbf{V}_i)$. The notation $-i$ will refer to $\mathcal{A} \setminus i$. Analogously, $\mathcal{G}_{-i}$ denotes the subgraph of $\mathcal{G}$ where $\mathbf{V}_i$ and its associated edges have been removed.

We define the notion of *unit homogeneity*. This assumption has two parts: a) if there exists a unit $i \in \mathcal{A}$ with variable $V_i \in \mathbf{V}_i$, then there is a corresponding $V_k \in \mathbf{V}_k$ for all $k \in \mathcal{A}$ with an analogous interpretation; and b) if there exists a unit $i \in \mathcal{A}$ with variables $V_i, U_i \in \mathbf{V}_i$ such that $V_i \in \mathrm{pa}_{\mathcal{G}_i}(U_i)$, then $V_k \in \mathrm{pa}_{\mathcal{G}_k}(U_k)$ for all $k \in \mathcal{A}$. The first part ensures that units are all of the same 'type' (e.g. all agents have the same demographic variables, and the same outcome variable). The second part ensures that the existence of a relationship between one unit's variables implies the same relationship exists for all other units.

For an example of these definitions, consider Fig. 2. Each unit has a variable of each 'type' (e.g. C, A, Y) and the connections between variables are the same for each unit (e.g. $C_i \rightarrow A_i$ in all units).
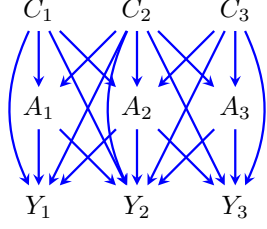
Figure 2: A simple social network represented by a DAG. The network exhibits unit homogeneity, symmetric connections, and homogeneous connections.

On the network level, we define the notions of connectedness, symmetry of connections, and homogeneity of connections. Two units $i, j \in \mathcal{A}$, with $i \neq j$, are said to be *connected* if for some $V_i \in \mathbf{V}_i$ and some $U_j \in \mathbf{V}_j$ it is the case that $V_i \in \mathrm{pa}_{\mathcal{G}}(U_j)$. The connection between i and j is said to be *symmetric* if the vice-versa relationship holds. That is, if i and j are connected and the connection is symmetric then for all $V_i \in \mathrm{pa}_{\mathcal{G}}(U_j)$, we have $V_j \in \mathrm{pa}_{\mathcal{G}}(U_i)$, where V_i is analogous to V_j and U_i is analogous to U_j. The set of units connected to unit i in $\mathcal{G}$, also referred to as i's *neighbors*, will be denoted $\mathcal{N}_{\mathcal{G}}(i)$. Finally, we define homogeneity of connections, which ensures that the relationships across the network are similar. If i and j are connected and there is an edge between some $V_i \in \mathbf{V}_i$ and some $V_j \in \mathbf{V}_j$ then network connections are *homogeneous* if for *all* connected units k, l in the network, an edge is present between the analogous $V_k \in \mathbf{V}_k$ and $V_l \in \mathbf{V}_l$.

We further define homogeneity of functional form which strengthens the notion of homogeneity for connections by imposing that, for any pair of connected nodes, the marginal distribution with respect to those two nodes is the same as the marginal distribution for any other pair of connected nodes (e.g. $p(\mathbf{V}_i, \mathbf{V}_j) = p(\mathbf{V}_k, \mathbf{V}_l)$ for all $i \neq j$ and $k \neq l$)). Under this assumption, pairwise relationships between units are the same, regardless of the type of unit. This assumption is reasonable in certain applied contexts, such as infectious disease spread, which is governed by a process that operates in the same way for any unit in the population.

For an example of these definitions, once again consider Fig. 2. Connections are symmetric (e.g. $C_1 \rightarrow A_2$ and $C_2 \rightarrow A_1$) and homogeneous (e.g. $C_1 \rightarrow A_2$ and likewise $C_3 \rightarrow A_2$).

5 NETWORK INTERVENTIONS

In this section we introduce the notion of *network interventions* where we intervene on the *structure* of a network by adding or removing edges, changing relationships be-

tween units. We define effects of these interventions and give identification criteria in §6, describe appealing properties of certain network interventions with respect to KL-divergence in §7, and discuss estimation in §8.

Severance Interventions

We will call interventions in which we sever two individuals in a network 'severance interventions'. For a graph $\mathcal{G}$ with pre-intervention distribution $p(\mathbf{V})$, where $\mathbf{V}$ is partitioned by $\{\mathbf{V}_i | i \in \mathcal{A}\}$, we denote the intervention severing units i and j by $i \bowtie j$. Graphically, this corresponds to removing *all* edges between $\mathbf{V}_i$ and $\mathbf{V}_j$, yielding the graph $\mathcal{G}_{i\bowtie j}$. We will define responses to severances with respect to individual units (e.g. $p(V_i(i \bowtie j))$). The joint response is simply the joint distribution over these counterfactuals.

We propose two different types of severance intervention. Each formulation has a corresponding causal interpretation and one could use either formulation depending on the application.

The first formulation, which we will call 'value-based' severance and is closely tied to classical mediation analysis, generalizes edge interventions [27] to networks. We intervene on variables in an edge-specific manner, replacing cross-unit edges into a unit, say i, with synthetic edges into i that represent fixed relationships no longer dependent on variables in the previously connected unit. For $\mathbf{V}_i \subset \mathbf{V}$, let $\mathbf{A}_{\mathbf{V}_i} = \mathrm{pa}_{\mathcal{G}_j}(\mathbf{V}_i)$, the parents of $\mathbf{V}_i$ in $\mathbf{V}_j$. We consider setting $\mathbf{A}_{\mathbf{V}_i}$ to $\mathbf{a}_{\mathbf{V}_i}$ for the sake of edges from $\mathbf{A}_{\mathbf{V}_i}$ to $\mathbf{V}_i$. All other edges out of $\mathbf{A}_{V_i}$ maintain the observed values of their source node so that for all $V_j \in \mathbf{V} \setminus \mathbf{V}_i$, V_j's pre- and post-intervention distributions are the same. Since the intervention values are constant, i and j are no longer connected. Returning to our diplomacy example, one might choose $\mathbf{a}_{\mathbf{V}_i}$ to be a reference value in the network, such as network averages of economic variables. Formally, $p(V_i(i \bowtie j; \mathbf{a}_{\mathbf{V}_i})) =$

$$p\big(V_i(\mathbf{A}_{\mathbf{V}_i} = \mathbf{a}_{\mathbf{V}_i}, \{V_j(\mathbf{a}_{\mathbf{V}_i}) : V_j \in \mathrm{pa}_{\mathcal{G}_{-j}}(V_i)\})\big)$$

The second formulation, which we call 'stochastic' severance, entails marginalizing out the parents from the severed unit. We phrase these as policy interventions.

Consider $V_i \in \mathbf{V}_i$ and let $\mathbf{A} = \{A \in \mathbf{V}_i | \mathrm{pa}_{\mathcal{G}_j}(A) \neq \emptyset\}$ (i.e. $\mathbf{A}$ is the set of unit-i variables with parents in unit j). The counterfactual $p(V_i(i \bowtie_{\mathbf{f}_{\mathbf{A}}} j))$ corresponds to selecting a set of stochastic policies $\mathbf{f}_{\mathbf{A}}$ where each f_A is *unit-structure preserving* (see below). The counterfactual is given by the recursive formula:

$$p(V_i(i \bowtie_{\mathbf{f}_{\mathbf{A}}} j)) = f_{V_i}(\{C(\mathbf{f}_{\mathbf{A}}) : C \in \mathrm{pa}_{\mathcal{G}}(A) \setminus \mathrm{pa}_{\mathcal{G}_j}(A)\})$$

For a policy f_{V_i} to be unit-structure preserving, $p(V_i | \mathrm{pa}_{\mathcal{G}_{-j}}(V_i))$ must be the same in the pre- and post-intervention distributions. This ensures that unit i's causal

structure is maintained. Formally, $f_{V_i}(\{W : W \in \mathrm{pa}_{\mathcal{G}_{-j}}(V_i)\}) =$

$$\int_{\mathrm{pa}_j(V_i)} p(V_i | \mathrm{pa}_{\mathcal{G}}(V_i)) d(\mathrm{pa}_j(V_i)),$$

where $\mathrm{pa}_j(V_i) = \mathrm{pa}_{\mathcal{G}}(V_i) \cap \mathbf{V}_j$.

We will argue in §7 that post-severance distributions are minimally KL-divergent from $p(\mathbf{V})$ among the class of distributions corresponding to the DAG with reduced edge set. Specifically, this holds for value-based severances if, instead of fixing values, we allow the source nodes of edge interventions to vary and average over those nodes. Likewise, for stochastic severances, the KL result holds if we pick f_{V_i} such that V_i and it's remaining parents $\mathrm{pa}_{\mathcal{G}_{i \bowtie j}}(V_i)$ have a particular relationship.

Connection Interventions

We will call interventions in which we adjoin two previously unconnected individuals in a network *connection interventions*. We will denote the intervention where units i and j are joined by $i \diamond j$. Graphically, this corresponds to inserting one or more edges from $\mathcal{G}_i$ to $\mathcal{G}_j$ or vice-versa, yielding $\mathcal{G}_{i \diamond j}$. As before, we will define responses to connection interventions with respect to individual units (e.g. $p(V_i(i \diamond j))$). The joint response $p(\mathbf{V}(i \diamond j))$ is simply the joint distribution over these counterfactual variables. We describe three separate and increasingly general formulations of connection interventions.

Interventions Under Functional Form Homogeneity

If we assume that the functional forms of network ties are homogeneous, and further assume that each structural equation in the network aggregates arbitrarily many inputs, then the new structural equation for each variable is determined by the equations for the analogous variables in the network.

We might be interested in counterfactual situations that are not present in the observed data, such as the case when connecting two units results in one unit having more neighbors than any unit in the observed data. Because we assume homogeneity of functional form, we can only allow for classes of policies that can flexibly handle an arbitrary number of neighbor nodes.

For the intervention to be well-defined, we must have $f_V \in \mathcal{F}$, where $\mathcal{F}$ is a class of aggregator functions of the form $f(h_U(U_1, U_2, \dots), h_W(W_1, W_2, \dots), \dots)$. Each h_Z maps $\mathcal{Z} \to \mathbb{R}$ where $\mathcal{Z}$ is an arbitrary-sized multiset of Z-type variables. In turn, f maps $\mathcal{H} \to \mathbb{R}$ where $\mathcal{H}$ is the arbitrary-sized multiset of outputs from the h functions. For instance, if V_i has parents of types $\mathbf{U}, \mathbf{W} \subset \mathbf{V}$, we might select h_U to output the mean of

the U's, h_W to output the median of the W's, and f to output the sum of those two values.

Suppose i and m, and i and k are connected in $\mathcal{G}$. Then under functional form homogeneity, the relationships between $V_i \in \mathbf{V}_i$ and $U_m \in \mathrm{pa}_{\mathcal{G}_m}(V_i)$ and V_i and the analogous $U_k \in \mathrm{pa}_{\mathcal{G}_k}(V_i)$ are governed by a function f_V. In the post-intervention distribution, $p(V_i(i \diamond_{f_V} j))$, where units i and j are connected, the relationship between V_i and U_j is also governed by f_V. The associated counterfactual is given by:

$$p(V_i(i \diamond_{f_V} j)) = p(V_i = f_V(\{V(i \diamond j) : V \in \mathrm{pa}_{\mathcal{G}_{i \diamond j}}(V_i)\}))$$

Intervening With Known Policies

We can relax the assumption of homogeneous network ties by intervening with a known functional form. As with the previous formulation, the analyst is interested in understanding the effect of inducing a specific relationship. Continuing our diplomacy example from §3, consider Turkey as a candidate for EU membership. Since Turkey has a large, robust economy, it may be able to negotiate a more favorable entrance with specific parameters, similar to Switzerland's non-member bilateral treaties. This formulation represents the inverse operation of function form-based severance interventions.

We wish to evaluate the effect of connecting units i and j with a known induced relationship. In the pre-intervention distribution, $V_i \in \mathbf{V}_i$ is determined by $f_{V_i}(\mathrm{pa}_{\mathcal{G}}(V_i), \epsilon_{V_i}) \in \mathcal{F}_{V_i}$. For the intervention to be valid, the analyst must specify $f'_{V_i} \in \mathcal{F}'_{V_i}$ where $\mathcal{F}'_{V_i}$ is a family of unit-structure preserving functions. The counterfactual is defined as:

$$p(V_i(i \diamond_{f'_{V_i}} j)) = p(V_i = f'_{V_i}(\{V(i \diamond j) : V \in \mathrm{pa}_{\mathcal{G}_{i \diamond j}}(V_i)\}))$$

In this context, the notion of a unit-structure preserving policy is the same as before, however for notational clarity we define $\mathbf{S} = \mathrm{pa}_{\mathcal{G}_{i \diamond j}}(V_i) \setminus \mathrm{pa}_{\mathcal{G}}(V_i)$, V_i's new parents in the post-intervention graph, and rephrase the definition as:

$$p(V_i | \mathrm{pa}_{\mathcal{G}}(V_i)) = \int_{\mathbf{S}} f'_{V_i}(\mathrm{pa}_{\mathcal{G}_{i \diamond j}}(V_i)) p(\mathbf{S}) d\mathbf{S} \qquad (3)$$

Intervening with Unknown Policies

In the most general formulation, we do not assume the analyst knows the interventional policy in advance. Instead, we formalize a procedure for picking an optimal policy to govern the relationship between connected units subject to some known constraints. In the example where we consider Turkey joining the EU, this corresponds to the EU and Turkey negotiating a treaty that jointly optimizes their outcomes (e.g. mean per-capita GDP).

Building on the preceding subsection, we can simply express this type of intervention as an optimization on some

jointly defined criterion, such as utility, within a class of policies. Let $\mathcal{F}'_{V_i}$ and $\mathcal{F}'_{V_j}$ be families of unit-structure preserving candidate policies for V_i and V_j. Let $\mathcal{C}$ be a known set of constraints that the solution must satisfy (e.g. Turkey cannot trade away more natural resources than it has). Let $g((V_i, V_j)(f_{V_i}, f_{V_j}))$ be a known function that captures the joint outcome for units i and j under a given pair of f's. Then the optimal f's are given by:

$$\underset{f_{V_i} \in \mathcal{F}'_{V_i}, f_{V_j} \in \mathcal{F}'_{V_j}}{\arg\max} E[g(V_i, V_j)(f_{V_i}, f_{V_j})] \text{ subject to } \mathcal{C}$$

Solving this optimization corresponds to evaluating $p(\mathbf{V}(i \diamond j)(f_{V_i}, f_{V_j}))$ for each pair of candidate f's that satisfy $\mathcal{C}$ and picking the best pair.

6 EFFECTS AND IDENTIFICATION OF NETWORK INTERVENTIONS

Hudgens and Halloran [8] defined the direct, spillover, and network average effects for interference settings. Respectively, these correspond to the effect on unit i's outcome when i's treatment is modified, the effect on i's outcome when i's neighbor's treatment is modified, and the average effect on all units' when someone's treatment is modified (i.e. the sum of the direct and spillover effects). Since these effects are defined for a particular type of node intervention, it is necessary to define analogous effects for network interventions.

We define two new effects: the individual participant effect (IPE), and the average bystander effect (ABE). The IPE is defined for units i and j when they are the subjects of a network intervention. IPE_i is the contrast between i's observed and interventional outcomes. For severances (with connections defined analogously), this contrast is given by $IPE_i(i \bowtie j) = Y_i - E[Y_i(i \bowtie j)]$. We can also define the average participant effect (APE) as the mean of IPE_i and IPE_j.

The ABE captures the contrast for units not directly involved in a network intervention. By the Markov property of DAGs, for a network intervention on i and j, the ABE is non-trivial for i and j's pre-intervention neighbors $\mathcal{N}_\mathcal{G}(i) \cup \mathcal{N}_\mathcal{G}(j) \setminus \{i, j\}$ (e.g. the other countries i and j have treaties with). For severances, $ABE(i \bowtie j) =$

$$\frac{1}{|(\mathcal{N}_i \cup \mathcal{N}_j) \setminus \{i, j\}|} \sum_{k \in (\mathcal{N}_i \cup \mathcal{N}_j) \setminus \{i, j\}} Y_k - E[Y_k(i \bowtie j)]$$

Connections are defined analogously. Following [8], the average effect on the network (e.g. the effect on the 'global' economy) is the sum of APE and ABE.

Identification

For a given intervention type, if the IPE is identified then the ABE is also identified and vice versa. We therefore

focus on the criteria for identification of each type of intervention we've discussed.

Under our setup, value-based severance interventions are the network analogue of edge interventions in mediation settings. For a severance of units i and j, let α be the set of edges out of $\text{pa}_\mathcal{G}(\mathbf{V}_i) \cap \mathbf{V}_j$. If $\mathfrak{a}_\alpha$ specifies a constant value for each edge $\mathbf{V}_j \to \mathbf{V}_i$ and that the source nodes for all other edges in α are random, then $p(\mathbf{V}_i(i \bowtie j; \mathfrak{a}_\alpha))$ is identified by the edge g-formula (Eq. 1).

For instance, in Fig. 2, if we are interested in the effect on Y_2 of severing units 2 and 3 by setting $\mathbf{A}_{V_2} = \mathbf{a}_{V_2} = \{C_3 = c_3, A_3 = a_3\}$ for the sake of the edges $(C_3 Y_2)_\to$ and $(A_3 Y_2)_\to$, then:

$$\begin{aligned} p(V_2(2 \bowtie 3); \mathbf{a}_{\mathbf{V}_3}) = \\ p(Y_2 | A_1, A_2, C_1, C_2, A_3 = a_3, C_3 = c_3) \\ \times p(A_1 | C_1, C_2) p(A_2 | C_1, C_2, C_3 = c_3) p(C_1) p(C_2) \end{aligned}$$

The other interventions we define entail a change in the functional form of the variables of interest. Suppose we wish to join units i and j with $\mathbf{A} = \{V \in \mathbf{V}_i | \text{pa}_{\mathcal{G}_{i \diamond j}}(V) \cap \mathbf{V}_j \neq \emptyset\}$ and $\mathbf{C}_A = \{V \in \mathbf{V}_j | A \in \text{ch}_{\mathcal{G}_{i \diamond j}}(V)\}$. Then, if $\mathbf{f}_\mathbf{A}$ are all functions that either satisfying the aggregator properties for the homogeneous case, or are unit-structure preserving for the non-homogeneous case, the counterfactual $p(\mathbf{V}_i(i \diamond_{\mathbf{f}_\mathbf{A}} j))$ is identified by the policy g-formula (Eq. 2). Stochastic severances (e.g. of units i and j) are also identified under our setup, with $\mathbf{A} = \{A \in \mathbf{V}_i | \text{pa}_{\mathcal{G}_j}(A) \neq \emptyset\}$ and $\mathbf{C}_A = \text{pa}_\mathcal{G}(A) \setminus \text{pa}_{\mathcal{G}_j}(A)$ for each $A \in \mathbf{A}$ and $\mathbf{f}_\mathbf{A}$ satisfying unit-structure preservation for each f_A. For homogeneous connections, we must also estimate the parameters of each aggregator function $(h_V, h_W, \text{etc.})$ from observed data. These are identified by maximum likelihood from $\mathcal{G}$

As an example, if we are interested in performing a stochastic severance on units 2 and 3 in Fig. 2, suppose we set $f_{V_i}(\text{pa}_{\mathcal{G}_{2 \bowtie 3}}(V_i)) = p(V_i | \text{pa}_\mathcal{G}(V_i) \setminus \text{pa}_{\mathcal{G}_3}(V_i))$ for each $V_i \in \mathbf{V}_i$. Then the identifying functional for the effect on $\mathbf{V}_2$ is given by:

$$\begin{aligned} p(\mathbf{V}_2(2 \bowtie_{f'_{V_i}} 3)) = p(Y_2 | A_1, A_2, C_1, C_2) \\ \times p(A_1 | C_1, C_2) p(A_2 | C_1, C_2) p(C_1) p(C_2) \end{aligned}$$

Latent-Variable Network Interventions

Throughout this work, we have assumed that data is representable by a DAG where all variables are observed. We can relax this assumption to allow for models in which some variables are latent. In these cases, the interpretation of the proposed interventions remains the same, however, identification conditions will be modified slightly.

Consider a latent-variable DAG $\mathcal{G}(\mathbf{V} \cup \mathbf{H})$ with $\mathbf{V}$ observed and $\mathbf{H}$ hidden. From $\mathcal{G}$, we can obtain a acyclic

directed mixed graph (ADMG) $\mathcal{G}'(\mathbf{V})$ via a latent projection operation [19]. $\mathcal{G}'$ represents an equivalence class of graphs that share the same observed variables and set of independence constraints [19].

Identification of network interventions in an ADMG $\mathcal{G}'$ relies on the assumptions described in the previous section, existing non-parametric identification theory for ADMGs, and the requirement that the network intervention operates only on edges that are present in both $\mathcal{G}$ and $\mathcal{G}'$. As pointed out previously, value-based severances in DAGs can be identified by the edge g-formula. Under the relaxation allowing for latent variables, value-based severances are instead identifiable according to a version of the ID algorithm adapted to edge interventions, proven sound and complete in [25]. Likewise, for stochastic severances and for connection interventions, if the identification conditions described in the previous sub-section hold, then the respective interventions are identifiable according to a version of the ID algorithm adapted to policy interventions, proven sound and complete in [25]).

7 OPTIMAL CHOICE OF POST-SEVERANCE DISTRIBUTION

In this section we prove a series of results regarding the KL-divergence from a distribution $p(\mathbf{V})$, corresponding to a known DAG $\mathcal{G}$, to another distribution $\tilde{p}(\mathbf{V})$, corresponding to a DAG in which edges have been removed. The results demonstrate that the KL-divergence from p to $\tilde{p}$ is minimized when $\tilde{p}$ takes on a form similar to the g-formula [21]. These *probabilistic* results help justify the g-formula and edge g-formula as intuitive tools for analyzing causal queries in DAGs. Moreover, these results motivate the manner in which we perform severances.

The first result demonstrates that when removing edges between a node A and its parents, a simple modification to the factorization of $\mathcal{G}$, removing A's parents from the term for A yields the KL-minimal distribution satisfying the independence constraints implied by the severance.

Theorem 1 *Let $\mathbf{V}$ be a set of random variables with $p(\mathbf{V})$ corresponding to a DAG $\mathcal{G}$. Let $A \in \mathbf{V}$. Let $\mathcal{P}(\mathbf{V})$ be the set of probability distributions that factorize according to $\mathcal{G}$. Then*

$$p(A) \prod_{V \in \mathbf{V} \setminus A} p(V | \operatorname{pa}_{\mathcal{G}}(V)) = \operatorname*{arg\,min}_{\tilde{p} \in \mathcal{P}(\mathbf{V})} D_{KL}(p || \tilde{p})$$
$$s.t. \ A \perp\!\!\!\perp \operatorname{pa}_{\mathcal{G}}(A)$$

The second result generalizes the first by allowing for edge removal between A and a *subset* of its parents.

Theorem 2 *Let $\mathbf{V}$ be a set of random variables with $p(\mathbf{V})$ corresponding to a DAG $\mathcal{G}$. Let $A \in \mathbf{V}$ and $\mathbf{B} \subseteq \mathbf{V}$*

such that $\mathbf{B} \subseteq \operatorname{pa}_{\mathcal{G}}(A)$. Let $\mathcal{P}(\mathbf{V})$ be the set of probability distributions that factorize according to $\mathcal{G}$. Then

$$p(A | \operatorname{pa}_{\mathcal{G}}(A) \setminus \mathbf{B}) \prod_{V \in \mathbf{V} \setminus A} p(V | \operatorname{pa}_{\mathcal{G}}(V))$$
$$= \operatorname*{arg\,min}_{\tilde{p} \in \mathcal{P}(\mathbf{V})} D_{KL}(p || \tilde{p}) \ s.t. \ A \perp\!\!\!\perp \mathbf{B} | \operatorname{pa}_{\mathcal{G}}(A) \setminus \mathbf{B}$$

The following result generalizes the previous theorem to allow for removal of *any* set of edges in $\mathcal{G}$. This result corresponds to directly to severance interventions. If we remove the dependence of each variable on the parents for which we remove edges, and otherwise keep the variable functionally consistent with its original structural equation, the result is the minimally KL-divergent distribution from the original distribution that reflects the severance.

Theorem 3 *Let $\mathbf{V}$ be a set of random variables with $p(\mathbf{V})$ corresponding to a DAG $\mathcal{G}$. Let $\mathbf{A} \in \mathbf{V}$ and for each $A \in \mathbf{A}$ define $\operatorname{In}(A) \subseteq \operatorname{pa}_{\mathcal{G}}(A)$, the set of parents of A whose edges into A we wish to remove. Let $\mathcal{P}(\mathbf{V})$ be the set of probability distributions that factorize according to $\mathcal{G}$. Then*

$$\prod_{A \in \mathbf{A}} p(A | \operatorname{pa}_{\mathcal{G}}(A) \setminus \operatorname{In}(A)) \prod_{V \in \mathbf{V} \setminus \mathbf{A}} p(V | \operatorname{pa}_{\mathcal{G}}(V))$$
$$= \operatorname*{arg\,min}_{\tilde{p} \in \mathcal{P}(\mathbf{V})} D_{KL}(p || \tilde{p})$$
$$s.t. \ A \perp\!\!\!\perp \operatorname{In}(A) | \operatorname{pa}_{\mathcal{G}}(A) \setminus \operatorname{In}(A) \ \forall A \in \mathbf{A}$$

The final two results are corollaries of Thm. 3 and are closely related to classical causal inference. The first corresponds to variable interventions where we fix some $\mathbf{A} \subseteq \mathbf{V}$ to a value $\mathbf{a}$. The KL-closest distribution to $p(\mathbf{V})$ is given by the g-formula, where terms for each $A \in \mathbf{A}$ are removed and variables with parents in $\mathbf{A}$ are evaluated with those parents set to $\mathbf{a}$.

Theorem 4 *Let $\mathbf{V}$ be a set of random variables with $p(\mathbf{V})$ corresponding to a DAG $\mathcal{G}$. Let $\mathbf{A} \subseteq \mathbf{V}$ and assume that for some $\mathbf{a}$ we have $p(\mathbf{A} = \mathbf{a}) > 0$. Let $\mathcal{P}(\mathbf{V})$ be the set of probability distributions that factorize according to $\mathcal{G}$. Then*

$$\prod_{V \in \mathbf{V} \setminus \mathbf{A}} p(V | \operatorname{pa}_{\mathcal{G}}(V)) |_{\mathbf{A} = \mathbf{a}} = \operatorname*{arg\,min}_{\tilde{p} \in \mathcal{P}(\mathbf{V})} D_{KL}(p || \tilde{p}) \ s.t.$$
$$\tilde{p}(A_i | \operatorname{nd}_{\mathcal{G}}(A_i)) = I(A_i = a_i) \ for \ i = \{1, \dots, |\mathbf{A}|\}$$

The final result, which can be found in the appendix, generalizes the above theorem to edge interventions [27]. This result corresponds to the value-based formulation of severances. When we fix a set of edges to constant values, the resulting distribution is given by the edge g-formula and is the KL-closest distribution to the pre-intervention distribution that reflects the fact that those edges have been fixed.

95% Confidence Intervals of Bias			
Intervention	*Erdős-Rényi*	*Barabasi-Albert*	*Watts-Strogatz*
Homogeneous Connection	(-.0049, .0020)	(-.0021, .0006)	(-.0024, .0010)
Known Connection	(-.0014, .0010)	(-.0004, .0016)	(-.0018, .0020)
Unknown Connection	(-.0035, .0025)	(-.0134, .0124)	(-.0280, .0093)
Stochastic Severance	(-.0015, .0043)	(-.0096, .0066)	(-.0032, .0020)
Value Severance	(-.0088, .0112)	(-.0010, .0020)	(-.0048, .0016)

Table 1: 95% confidence intervals for the bias of estimates of each type of network intervention.

8 EXPERIMENTS

We now describe a set of simulation studies which demonstrate the feasibility of obtaining unbiased estimates of the effects of network interventions. In these experiments we assume *partial interference*: we observe M samples of a network, each with N units. While we do not consider full interference scenarios, in which the analyst has access to only a single sample of the network, similar results could be obtained in that setting using the auto-g-computation algorithm [28]. We also assume that all pre-intervention networks satisfy symmetry of connections, and homogeneity of units, connections, and functional form.

We consider a social network graph resembling Fig. 2 where all variables C, A, and Y are binary. In four separate experiments we demonstrate estimation across varying social network generators, varied attachment probabilities for the Erdős-Rényi generator, varied network sizes, and varied sample sizes. For the latter three experiements we restrict attention to the stochastic severance intervention. For each unit i we generate values for $\mathbf{V}_i$ according to log-linear models with parameters τ_C, τ_A, τ_Y. For the detailed setup, please see the appendix.

For each experiment we estimate the average IPE by separately applying the intervention to each unit in the network. For severances we remove the connection between the unit of interest and it's highest degree neighbor while for connections we connect the unit to it's highest degree *non*-neighbor.

8.1 ESTIMATION AND EVALUATION

For each experiment we first fit models for each variable type given it's parents via MLE where features for neighbor variables are sums of those variables. We estimate values of endogenous nodes using Monte Carlo sampling using these fit models and exogenous nodes via the empirical distribution. We estimate values in the pre- and post-intervention worlds and report the mean difference between these estimates across all units and all samples of the network. For specific details on the mechanics of each intervention type, please see the appendix.

To evaluate the performance of this estimation technique, we generate 'ground truth' graphs for each intervention and generate values for the Y_i's of interest. For each simulated network we generated 1000 bootstrap samples. We compare the intervention effects to the ground truth effects and obtain the bias of our approach. As presented in Table 1 above, and Tables 2 - 4 in the appendix, the 95% confidence interval for each experiment covers the ground truth bias and thus shows that the effects of network interventions can be consistently estimated.

9 DISCUSSION

In this paper we proposed a framework for intervening on the *structure* of a social network graph by severing or creating connections between subjects. We defined effects that extend the network effects defined in [8]. We then proved that for severances, and causal interventions generally, the g-formula and edge g-formula obtain distributions that are minimally KL-divergent from the pre-intervention distribution subject to the independence constraint imposed by the intervention. Finally, we demonstrated that these effects can be estimated from observational data via a simulation study.

In the future, this framework could be generalized to chain graph models to allow for more flexibility of network representation.

References

[1] D. Arbour, D. Garant, and D. Jensen. Inferring network effects from observational data. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 715–724. ACM, 2016.

[2] N. Beck and J. N. Katz. Modeling dynamics in time-series–cross-section political economy data. *Annual Review of Political Science*, 14:331–352, 2011.

[3] R. Chetty, N. Hendren, and L. F. Katz. The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, 106(4):855–902, 2016.

[4] F. Eberhardt and R. Scheines. Interventions and causal inference. *Philosophy of Science*, 74(5):981–995, 2007.

[5] N. Friedman, L. Getoor, D. Koller, and A. Pfeffer. Learning probabilistic relational models. In *IJCAI*, volume 99, pages 1300–1309, 1999.

[6] T. A. Glass, S. N. Goodman, M. A. Hernán, and J. M. Samet. Causal inference in public health. *Annual review of public health*, 34:61–75, 2013.

[7] J. Y. Halpern and J. Pearl. Causes and explanations: A structural-model approach. part i: Causes. *The British journal for the philosophy of science*, 56(4):843–887, 2005.

[8] M. G. Hudgens and M. E. Halloran. Toward causal inference with interference. *Journal of the American Statistical Association*, 103(482):832–842, 2008.

[9] D. Koller, N. Friedman, S. Džeroski, C. Sutton, A. McCallum, A. Pfeffer, P. Abbeel, M.-F. Wong, D. Heckerman, C. Meek, et al. *Introduction to statistical relational learning*. MIT press, 2007.

[10] K. B. Korb, L. R. Hope, A. E. Nicholson, and K. Axnick. Varieties of causal intervention. In *Pacific Rim International Conference on Artificial Intelligence*, pages 322–331. Springer, 2004.

[11] 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.

[12] D. Malinsky. Intervening on structure. *Synthese*, 195(5):2295–2312, May 2018.

[13] K. P. Murphy and S. Russell. Dynamic bayesian networks: representation, inference and learning. 2002.

[14] J. Neyman. Sur les applications de la théorie des probabilités aux experiences agricoles: Essai des principes. *Roczniki Nauk Rolniczych*, 10:1–51, 1923.

[15] E. L. Ogburn, I. Shpitser, and Y. Lee. Causal inference, social networks, and chain graphs. *arXiv preprint arXiv:1812.04990*, 2018.

[16] E. L. Ogburn, O. Sofrygin, I. Diaz, and M. J. van der Laan. Causal inference for social network data. *arXiv preprint arXiv:1705.08527*, 2017.

[17] E. L. Ogburn, T. J. VanderWeele, and others. Causal diagrams for interference. *Statistical science*, 29(4):559–578, 2014.

[18] J. Pearl. *Causality*. Cambridge university press, 2009.

[19] 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.

[20] T. S. Richardson and J. M. Robins. Single world intervention graphs (swigs): A unification of the counterfactual and graphical approaches to causality. *Center for the Statistics and the Social Sciences, University of Washington Series. Working Paper*, 128(30):2013, 2013.

[21] 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.

[22] J. W. Robinson and A. J. Hartemink. Non-stationary dynamic bayesian networks. In *Advances in neural information processing systems*, pages 1369–1376, 2009.

[23] D. B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688, 1974.

[24] E. Sherman and I. Shpitser. Identification and estimation of causal effects from dependent data. In *Advances in Neural Information Processing Systems*, pages 9445–9456, 2018.

[25] 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.

[26] I. Shpitser and E. T. Tchetgen. Causal inference with a graphical hierarchy of interventions. *Annals of statistics*, 44(6):2433, 2016.

[27] I. Shpitser and E. T. Tchetgen. Causal inference with a graphical hierarchy of interventions. *Annals of statistics*, 44(6):2433, 2016.

[28] E. J. T. Tchetgen, I. Fulcher, and I. Shpitser. Auto-G-Computation of Causal Effects on a Network. *arXiv preprint arXiv:1709.01577*, 2017.

[29] J. Tian. Identifying dynamic sequential plans. *arXiv preprint arXiv:1206.3292*, 2012.

[30] J. Tian and J. Pearl. Causal discovery from changes. In *Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence*, pages 512–521. Morgan Kaufmann Publishers Inc., 2001.

[31] J. Woodward. Causation and manipulability. 2001.