TECHNICAL REPORT
R-55
Sep, 2019

A Calculus for Stochastic Interventions: Causal Effect Identification and Surrogate Experiments

Juan D. Correa and **Elias Bareinboim**
Causal Artificial Intelligence Laboratory
Department of Computer Science
Columbia University
{jdcorrea,eb}@cs.columbia.edu

Abstract

Some of the most prominent results in causal inference have been developed in the context of atomic interventions, following the semantics of the *do*-operator and the inferential power of *do*-calculus. In practice, however, many real-world settings call for more complex types of treatments that cannot be represented by a simple atomic intervention. In this paper, we investigate a general class of interventions that covers some non-trivial types of policies (including conditional and stochastic), going beyond the atomic case. Our goal is to develop a general understanding and formal machinery to reason about the effects of those policies, similar to the robust theory developed to handle the atomic case. Specifically, we introduce a new set of inference rules (akin to *do*-calculus) that can be used to derive claims about general interventions, which we call σ-calculus. We develop a graph-based, efficient procedure for finding estimands of the effect of general policies as a function of the available observational and experimental distributions. We then prove that our algorithm and σ-calculus are both sound for the tasks of identification (Pearl, 1995) and z-identification (Bareinboim and Pearl, 2012) under this class of interventions.

1 Introduction

Causal relations are considered highly valuable and desirable throughout the data-driven sciences due to their inherent interpretability and robustness to changing conditions. In machine learning, for example, they play a key role due to their amenability to extrapolation to new, unforeseen situations, and also their capability to support robust decision-making. Making sense of the world and constructing coherent and transparent explanations about it, almost invariably, hinge on our ability to learn and reason with cause and effect relationships (Pearl, 2000; Spirtes et al., 2001; Bareinboim and Pearl, 2016; Pearl and Mackenzie, 2018).

One of the most common ways of learning about causal relations is through controlled experimentation. In practice, however, performing experiments is not always feasible due to its potentially harmful side effects, financial constraints, and ethical considerations. This impossibility leads to one of the fundamental challenges in causal inference, namely, to determine whether the effect of an intervention can be computed without directly experimenting on the system, which is known as the problem of *identification of causal effects* (Pearl, 2000, Def. 3.2.4). Among many types of interventions, the simplest and best understood is called *atomic*. In modern causal inference, atomic interventions are usually modeled through the *do*-operator (Pearl, 1995), which is denoted by $do(X{=}x)$. Formally, $do(X{=}x)$ represents the symbolic operation of replacing the underlying causal mechanism that naturally dictates the behavior of a variable X with a constant value x[1].

The identification task relies on assumptions about the underlying causal system, which is usually encoded in the form of a causal graphical model. For concreteness, consider the causal diagram in Fig. 1(a), where X represents the choice to *smoke*, W *age*, Z a set of risk factors leading to *tendency to smoke* (e.g., peer pressure, education, SES, psychological age), and Y the development or not of *lung cancer*. The goal of the task is to compute the average effect of X on Y based on the observational (i.e., non-interventional) distribution $P(W, Z, X, Y)$. Using the *do*-operator, this quantity can be formally written as $P(Y|do(X = x))$, which describes the behavior of Y when X is fixed to x (smoking or not) *regardless of* Z or any other confounding factors. Furthermore, the difference between two specific do-distributions, $P(Y|do(X = 1)) - P(Y|do(X = 0))$, measures the causal changes of Y that are due to the deliberate variations of X.

Each specific interventional distribution induces a new hypothetical regime that can be represented through a causal diagram reflecting the corresponding change in mechanism. For $do(X = x)$, this corresponds to a diagram where the arrows incoming to X are removed (Fig. 1(b)). We will annotate these diagrams with an explicit *regime* node σ_X to

[1]This primitive has appeared at different times and contexts through causality's history. It was introduced in econometrics by (Haavelmo, 1943; Strotz and Wold, 1960). In statistics, potential outcomes were introduced in the context of randomized experiments by (Neyman, 1923), and then connected with observational studies by (Rubin, 1974). In mathematical logic, counterfactuals were discussed by (Lewis, 1973) with a possible worlds semantics. They were then given a general and algorithmic treatment through graphical models in AI (Pearl, 1993b; Pearl, 1995).

indicate that the causal mechanism of X has changed. This is a critical construct, which will be discussed in Sec. 3).

There exists a growing body of literature concerned with identification of do-interventions from data collected under observational and experimental regimes, including celebrated results such as *do-calculus* (Pearl, 1993a; Pearl, 1995; Pearl, 2000), and complete graphical and algorithmic conditions (Tian and Pearl, 2002a; Tian, 2004; Shpitser and Pearl, 2006a; Shpitser and Pearl, 2006b; Huang and Valtorta, 2006; Bareinboim and Pearl, 2012; Lee et al., 2019).

While the intervention $do(X = 0)$ describes with mathematical precision a counterfactual world where smoking is banned from society, it is unlikely, in practice, that a policy could be implemented such that cigarettes would be completely wiped out from the streets. In other words, we could eventually predict the effect of this new, idealized policy, however unlikely to be implemented in reality. The tension between the result of the formal analysis and the practical, realizable result of an implemented policy has been a point of intense debate in causal circles (Woodward, 2003; Heckman, 2005; Cartwright, 2007; Pearl, 2010).

In this paper, we offer a mathematical solution to address this decade-old debate. Going back to our example, for concreteness, policy-makers contemplate a more strict regulation on underage smoking and higher taxes on cigarettes sales that could be set in place. A sensible question in this context could be – what is the effect of a policy that inhibits smoking in people under 21 years of age, by 90%? Such intervention is certainly non-atomic (which would entail that a 100% decrease in smoking should be enforced for this group), and in this case, the underlying mechanism for X is replaced with a softer mechanism; these interventions are sometimes called *soft* or *stochastic* interventions.

Even though deciding the identifiability of complex interventions has been studied in the literature, there is still work to be done (Pearl, 2000, Ch. 4). For instance, (Pearl and Robins, 1995) studied the effect of interventions in longitudinal settings where the decision in each time step is dependant on the previous ones, which was called *conditional plans*. Further, other works investigated the effect of *stochastic* interventions, where the original causal mechanism of the treatment variable is replaced with a new known function (Dawid, 2002; Didelez et al., 2006; Tian, 2008; Shpitser and Sherman, 2018). For the case when the new function is unknown, the problem has been studied under the rubrics of transportability (Bareinboim and Pearl, 2014; Bareinboim and Pearl, 2016; Correa and Bareinboim, 2019).

Despite the high level of sophistication and generality achieved for reasoning with atomic interventions, we highlight the glaring difference with the non-atomic case. For instance, there exist no counterpart for *do*-calculus in the non-atomic case nor general results on identifiability from experimental distributions produced by soft interventions. In this paper, we develop a general, symbolic, and algorithmic treatment for identifiability of arbitrary non-atomic interventions from both observational and experimental distributions. More specifically, our contributions are as follows:

1. **Symbolic characterization.** We introduce a set of inference rules, called σ-calculus, to reason about the effect of general types of intervention. Further, we provide a syntactical method for deriving and verifying claims about such interventions given a causal graph.
2. **Algorithmic solution.** We develop an efficient procedure to determine the identifiability of the (conditional) effect of non-atomic interventions from observational and experimental distributions given a causal diagram.

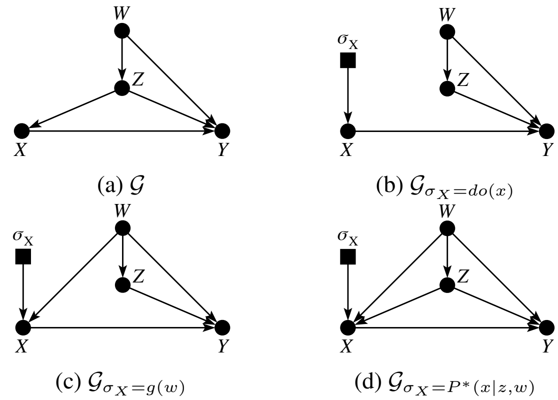


Figure 1: (a) original causal diagram $\mathcal{G}$. (b), (c), and (d) show the causal diagrams after an atomic, conditional, and stochastic intervention, respectively. See discussion in the introduction and examples 1 and 2 for details.

2 Preliminaries

The basic semantic framework our work rests on is the Structural Causal Models paradigm (Pearl, 2000, Ch. 7), which allows one to represent the data-generation process and different types of interventions:

Definition 1 (SCM)**.** A Structural Causal Model M is a 4-tuple $\langle \mathbf{U}, \mathbf{V}, \mathcal{F}, P(\mathbf{u})\rangle$, where $\mathbf{U}$ is a set of exogenous (latent) variables; $\mathbf{V}$ is a set of endogenous (observed) variables; $\mathcal{F}$ is a collection of functions such that each variable $V_i \in \mathbf{V}$ is determined by a function $f_i \in \mathcal{F}$. Each f_i is a mapping from a set of exogenous variables $U_i \subseteq \mathbf{U}$ and a set of endogenous variables $Pa_i \subseteq \mathbf{V} \setminus \{V_i\}$ to the domain of V_i. The uncertainty is encoded through a probability distribution over the exogenous variables, $P(\mathbf{U})$.

Note that this definition allows for latent confounders, so the model is also known as *Semi-Markovian*. Each SCM M is associated with a *causal diagram* where every $V_i \in \mathbf{V}$ is a vertex, there is a directed edge $(V_j \rightarrow V_i)$ for every $V_i \in \mathbf{V}$ and $V_j \in Pa_i$, and a bidirected edge $(V_i \leftarrow\dashrightarrow V_j)$ for every pair $V_i, V_j \in \mathbf{V}$ such that $U_i \cap U_j \neq \emptyset$ (V_i and V_j have a common latent confounder).

We assume that the underlying model is recursive, that is, there are no cyclic dependencies among the variables. Equivalently, the causal diagram corresponding to the SCM is acyclic. The observable distribution is derived from M as

$$P(\mathbf{v}) = \sum_{\mathbf{u}} \prod_{\{i|V_i \in \mathbf{V}\}} P(v_i \mid pa_i, u_i) P(\mathbf{u}), \quad (1)$$

where every term $P(v_i|pa_i, u_i)$ is governed by the corresponding function $f_i \in \mathcal{F}$ that represents an autonomous mechanisms affecting only V_i, locally (Aldrich, 1989).

A $do(\mathbf{X}{=}\mathbf{x})$ intervention results in a new structural causal model $M_{\mathbf{x}}$, which represents the state of the system after the hypothetical intervention takes place. As for M, assumptions about the causal structure of $M_{\mathbf{x}}$ can be seen as the corresponding causal diagram $\mathcal{G}_{\overline{\mathbf{X}}}$, which is the same as $\mathcal{G}$ but for the absence of all edges incoming towards $\mathbf{X}$. Moreover, $M_{\mathbf{x}}$ induces a probability distribution $P(\mathbf{V}|do(\mathbf{x}))$ that can be established using Eq. (1) in the context of $M_{\mathbf{x}}$, i.e.:

$$P(\mathbf{v}|do(\mathbf{x})){=}\sum_{\mathbf{u}}\prod_{\{i|V_i\in\mathbf{V}\}}P(v_i|pa_i,u_i,do(\mathbf{x}))P(\mathbf{u}|do(\mathbf{x})). \quad (2)$$

The key observation here is that for every $V_i \in \mathbf{V}, V_i \notin \mathbf{X}$, $P(v_i \mid pa_i, u_i, do(\mathbf{x})) = P(v_i \mid pa_i, u_i)$, because the functions f_i in M and $M_{\mathbf{x}}$ are the same. Similarly, $P(\mathbf{u} \mid do(\mathbf{x})) = P(\mathbf{u})$ since exogenous variables are not affected by the do-operation. Moreover, for $V_i \in \mathbf{X}$, the function f_i in $M_{\mathbf{x}}$ is independent of $\mathbf{U} \cup (\mathbf{V} \setminus \{V_i\})$, hence $Pa_i = \emptyset, U_i = \emptyset$, and the corresponding term $P(V_i = v_i \mid do(\mathbf{x})) = 1$, if v_i is consistent with $\mathbf{x}$; and 0, otherwise. Then, $P(\mathbf{v}|do(\mathbf{x}))$ in Eq. (2) is also equal to:

$$\begin{cases}\sum_{\mathbf{u}}\prod_{\{i|V_i\in\mathbf{V}\setminus\mathbf{X}\}}P(v_i|pa_i,u_i)P(\mathbf{u}) & \mathbf{v}\text{ consistent with }\mathbf{x}\\ 0 & \mathbf{v}\text{ inconsistent with }\mathbf{x}\end{cases}. \quad (3)$$

In the special case of Markovian models, where every unobservable variable in $\mathbf{U}$ affects at most one observable, Eq. (3) is called the "truncated factorization product" (Pearl, 1993a; Pearl, 2000; Bareinboim et al., 2012), which yields a mapping from the pre ($P(\mathbf{V})$) to the post-interventional distribution ($P(\mathbf{V}|do(\mathbf{x}))$). In Fig. 1(a), for example, the effect $P(y|do(x)){=}\sum_{w,z}P(y|x,z,w)P(z|w)P(w)$ by Eq. (3).

It is unlikely that one could observe all variables in the system in most practical applications. Consequently, realistic causal diagrams usually account for latent (unobserved, exogenous) variables that affect more than one observable, which are represented through bidirected edges. In the following sections, we will address the problem of identifying the effect of stochastic interventions in such class of models.

We follow standard notation in the field. Random variables are denoted with uppercase letters (e.g, C) while their instantiations to particular values are written in lowercase (e.g, c). Similarly, letters in bold (e.g, $\mathbf{C}$) represent sets of variables, and lowercase-bold letters (e.g., $\mathbf{c}$) a particular value assignment for them. Further, we denote by $\mathcal{G}_{\overline{\mathbf{W}}\underline{\mathbf{X}}}$ the graph that is the same as $\mathcal{G}$ except that the edges incoming to variables in $\mathbf{W}$ and the edges going out from variables in $\mathbf{X}$ are removed. Let $\mathcal{G}_{[\mathbf{C}]}$ be the subgraph of $\mathcal{G}$ made only of nodes in $\mathbf{C}{\subset}\mathbf{V}$ and the edges between them. We define $Pa(\mathbf{C})$ and $An(\mathbf{C})$, as the union of $\mathbf{C} \subset \mathbf{V}$ with its parents and ancestors, respectively. Also, the expression $(\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z})_{\mathcal{G}}$ denotes that the variables in $\mathbf{X}$ are separated from the variables in $\mathbf{Y}$ conditioned on $\mathbf{Z}$ according to the d-separation criterion in the graph $\mathcal{G}$ (Pearl, 2000).

The proofs are provided in the Appendix.

3 Moving Beyond Atomic Interventions

In general, the result of an intervention encompasses a new *regime* where the data-generating process differs from that of the natural system only in the mechanisms associated with the variables that have been intervened (Pearl, 1994; Dawid, 2002; Dawid, 2015). From this point of view, we use *regime indicators* as discussed in (Pearl, 2000, Sec. 3.2.2) and (Dawid, 2002) to represent different types of interventions. The regime indicator for interventions on a variable X is denoted by σ_X, and encodes the fact that the function f_x in M has been replaced by a new function f_x^*. This operation results in a new model M_{σ_X}, with causal diagram $\mathcal{G}_{\sigma_X}$, and inducing a distribution $P(\mathbf{V};\sigma_X)$. See Fig. 1(b)-(d) for a few examples of post-interventional diagrams.

In particular, depending on the intervention, the function f_x^* could receive as inputs the values of variables other than the original parents Pa_x and U_x. Accordingly, we will denote as Pa_x^* and U_x^* the set of observable and unobservable parents of X in M_{σ_X}, as dictated by f_x^*. To avoid clutter, when a regime indicator σ_X is present in a probability expression, such as $P(x|pa_x, u_x;\sigma_X)$, Pa_x and U_x correspond to Pa_x^* and U_x^*, respectively. Naturally, this means that $\mathcal{G}_{\sigma_X}$ may not be a subgraph of $\mathcal{G}$, as it occurs with *do*-interventions. One important assumption used throughout the paper is that the hypothetical model M_{σ_X} resulting from the intervention σ_X does not contain cycles. Following the convention in (Dawid, 2002), we augment $\mathcal{G}_{\sigma_{\mathbf{X}}}$ with a node σ_{X_i} for every $X_i \in \mathbf{X}$ that graphically denote the targets of intervention, together with the edge $(\sigma_{X_i} \to X_i)$.

Representing Different Types of Interventions

Qualitatively different types of interventions can be modeled by assigning different *strategies* to the indicator σ_X using the construct discussed above. We list in Table 1 general types of interventions that will be used in the remaining of the paper. The *idle* intervention represents the natural state of the system; *atomic* or *do* interventions replace the function f_X with a constant, while *conditional* ones replace it with a deterministic function of some observables pa_x^*. The *stochastic* type sets the new f_X^* such that the variable X will follow a pre-specified distribution $P^*(X|pa_x^*)$. To simplify notation, whenever the strategy assigned to σ_X is clear from the context, we will omit it in the probability expressions. Also, we may just write $P(\mathbf{V})$ whenever $P(\mathbf{V};\sigma_X{=}\emptyset)$. For a set $\mathbf{X}{\subset}\mathbf{V}$, let $\sigma_{\mathbf{X}}{=}\{\sigma_{X_1},\dots\}$ represent an intervention affecting the functions f_{x_i} of every $X_i{\in}\mathbf{X}$.

Example 1 (Conditional Intervention)**.** In the context of a tutoring program, suppose that in Fig. 1(a) W represents previous GPA of a student, Z student's motivation, X after hours tutoring (or not), and Y the GPA at the end of the term. Currently, students seek tutoring voluntarily, which depends on their motivation. Given the limited amount of resources, the school is considering to make after hours tutoring mandatory for students with low GPAs, and offering this service only to them. The proposed intervention can be encoded as $\sigma_X = g(w)$, where $g(w) = 1$ if W is low GPA, and 0 otherwise. Graphically, this change in policy is represented by the diagram in Fig. 1(c), where X now depends on W, not on Z. Still, we highlight that X was dependent on Z in the observational regime and its corresponding dataset.

Type	Strategy	$P(x \mid pa_x, u_x; \sigma_X)$	
Idle	$\emptyset$	(unaltered)	
Atomic/*do*	$do(X = x')$	$\delta(x, x')$	(4)
Conditional	$do(X = g(pa_x^*))$	$\delta(x, g(pa_x^*))$	(5)
Stochastic/Random	$P^*(X \mid pa_x^*)$	$P^*(x \mid pa_x^*)$	(6)

Table 1: (1st column) Different types of interventions. (2nd) The corresponding strategies that can be assigned to the indicator variable σ_X. (3rd) Distributions that X will display after the intervention is implemented. $\delta(a, b) = 1$ if $a = b$, and 0 otherwise.

Example 2 (Stochastic Intervention)**.** Recall the discussion about the new smoking policy in the introduction. One could estimate the effect of reducing by 90% smoking on people under 21 years old by reasoning about a stochastic intervention $P^*(x|w,z)$, depicted in Fig.1(d), such that $P^*(X{=}1|W{<}21, z){=}(0.1){\times}P(X{=}1|z)$, for every z.

Interestingly, the randomization procedure used in a controlled experiment (Fisher, 1951) – represented by the *do*-operator – can be seen as the implementation of the stochastic intervention $\sigma_X{=}P^*(X)$, with $P^*(x){=}1/2$, for $x{=}\{0, 1\}$. This procedure induces the distribution $P(\mathbf{v}; \sigma_X{=}P^*(X))$. Evidently, Fisher's randomization is physical, while the inferences studied here are about how to determine a causal effect without actually performing the intervention in the real world. To understand this connection more precisely, we first condition the post-interventional distribution, $P(\mathbf{v}; \sigma_X{=}P^*(X))$, on X, which leads to $P(\mathbf{v}|X{=}x; \sigma_X{=}P^*(x))$. Now notice that each individual for which $X{=}x$ under σ_X is assigned treatment completely at random (i.e., without the influence of any other factor), which is the very definition of $do(x)$, hence $P(\mathbf{v}|x; \sigma_X{=}P^*(x)) = P(\mathbf{v} \mid do(x))$.

Effect of General Interventions

Regardless of the particular type of intervention, we can reason about the distribution that (the hypothetical) $M_{\sigma_\mathbf{X}}$ induces. Let $\mathbf{U}^*$ be the set of all unobservable variables in $M_{\sigma_\mathbf{X}}$, then using Eq. (1) we have:

$$P(\mathbf{v}; \sigma_\mathbf{X}) = \sum_{\mathbf{u}^*} \prod_{\{i|V_i \in \mathbf{V}\}} P(v_i|pa_i, u_i; \sigma_X)P(\mathbf{u}^*; \sigma_X). \quad (7)$$

Every $V_i \in \mathbf{V} \setminus \mathbf{X}$ is governed by the same function in M and M_{σ_X}, by definition, hence $P(v_i|pa_i, u_i; \sigma_\mathbf{X}) = P(v_i|pa_i, u_i)$. For the exogenous, the variables in the set $\mathbf{U}^* \setminus \mathbf{U}$ were introduced due to $\sigma_\mathbf{X}$ and were not originally in M (e.g., the randomness for a stochastic intervention). Since $\mathbf{U}$ is not affected by $\sigma_\mathbf{X}$, it follows $P(\mathbf{u}; \sigma_\mathbf{X}){=}P(\mathbf{u})$, and

$$P(\mathbf{v}; \sigma_\mathbf{X}) = \sum_{\mathbf{u}^*} \prod_{\{i|V_i \in \mathbf{X}\}} P(v_i|pa_i, u_i; \sigma_\mathbf{X})P(\mathbf{u}^* \setminus \mathbf{u}; \sigma_\mathbf{X}) \prod_{\{i|V_i \in \mathbf{V} \setminus \mathbf{X}\}} P(v_i|pa_i, u_i)P(\mathbf{u}). \quad (8)$$

While Eq. (8) holds in general, the distribution $P(\mathbf{U})$ is not observed. The challenge is then to find a function of the observed distribution $P(\mathbf{V})$ that is guaranteed to be equal to the probability query of interest in the intervened model $M_{\sigma_\mathbf{X}}$, for any M inducing $\mathcal{G}$. Formally,

Definition 2 (Effect Identifiability)**.** Let $\mathbf{Y}, \mathbf{X}, \mathbf{W} \subset \mathbf{V}$ with $\mathbf{W} \cap \mathbf{Y} = \emptyset$. The (conditional) effect of an intervention specified by $\sigma_\mathbf{X} = \{\sigma_{X_1}, \ldots, \sigma_{X_n}\}$ on a set of outcome variables $\mathbf{Y}$, conditional on $\mathbf{W}$, $P(\mathbf{y}|\mathbf{w}; \sigma_\mathbf{X})$, is said to be identifiable in $\mathcal{G}$, if it is uniquely computable from the joint distribution $P(\mathbf{V})$, for every assignment $(\mathbf{y}, \mathbf{w})$, in every model that induces $\mathcal{G}$ and $P(\mathbf{V})$.

Remark 1. An important distinction between atomic and more general interventions is that the former implicitly conditions on the intervened variable $\mathbf{X}$, more formally,

$$P(\mathbf{y} \mid do(\mathbf{x})) = P(\mathbf{y}; \sigma_\mathbf{X} = do(\mathbf{X} = \mathbf{x})) \quad (9)$$
$$= P(\mathbf{y} \mid \mathbf{x}; \sigma_\mathbf{X} = do(\mathbf{X} = \mathbf{x})). \quad (10)$$

Eq. (9) follows by definition, and Eq. (10) is immediate since under the intervention $\sigma_X{=}do(\mathbf{X}{=}\mathbf{x})$, the probability of $\mathbf{X}$ being different than the constant $\mathbf{x}$ is zero. In general, $P(\mathbf{y}; \sigma_\mathbf{X})$ and $P(\mathbf{y}|\mathbf{x}; \sigma_\mathbf{X})$ need not match one another.

Interestingly, while atomic interventions always reduce the model structure, a policy-maker could envision a new policy taking into account a wide range of covariates, not matching the observational regime and previous policies (as with Examples 1, 2).

4 A Calculus for General Interventions

In this section, we introduce a set of inference rules, in the spirit of *do*-calculus (Pearl, 2000, Sec. 3.4), capable of handling both atomic and non-atomic interventions, which we call σ-calculus.

Theorem 1. *[Inference Rules – σ-calculus] Let $\mathcal{G}$ be a causal diagram compatible with a structural causal model M, with endogenous variables $\mathbf{V}$. For any disjoint subsets $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \subseteq \mathbf{V}$, two disjoint subsets $\mathbf{T}, \mathbf{W} \subseteq \mathbf{V} \setminus (\mathbf{Z} \cup \mathbf{Y})$ (i.e., possibly including $\mathbf{X}$), the following rules are valid for any intervention strategies $\sigma_\mathbf{X}, \sigma_\mathbf{Z}$, and $\sigma'_\mathbf{Z}$:*

Rule 1 *(Insertion/Deletion of observations):*

$$P(\mathbf{y} \mid \mathbf{w}, \mathbf{t}; \sigma_\mathbf{X}) = P(\mathbf{y} \mid \mathbf{w}; \sigma_\mathbf{X}) \quad \text{if } (\mathbf{Y} \perp\!\!\!\perp \mathbf{T} \mid \mathbf{W}) \text{ in } \mathcal{G}_{\sigma_\mathbf{X}}. \quad (11)$$

Rule 2 *(Change of regimes under observation):*

$$P(\mathbf{y} \mid \mathbf{z}, \mathbf{w}; \sigma_\mathbf{x}, \sigma_\mathbf{z}) = P(\mathbf{y} \mid \mathbf{z}, \mathbf{w}; \sigma_\mathbf{x}, \sigma'_\mathbf{z}) \quad \text{if } (\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{W}) \text{ in } \mathcal{G}_{\sigma_\mathbf{X}\sigma_\mathbf{Z}\underline{\mathbf{Z}}} \text{ and } \mathcal{G}_{\sigma_\mathbf{X}\sigma'_\mathbf{Z}\underline{\mathbf{Z}}}. \quad (12)$$

Rule 3 *(Change of regimes without observation):*

$$P(\mathbf{y} \mid \mathbf{w}; \sigma_\mathbf{x}, \sigma_\mathbf{z}) = P(\mathbf{y} \mid \mathbf{w}; \sigma_\mathbf{x}, \sigma'_\mathbf{z}) \quad \text{if } (\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{W}) \text{ in } \mathcal{G}_{\sigma_\mathbf{X}\sigma_\mathbf{Z}\overline{\mathbf{Z}(\mathbf{W})}} \text{ and } \mathcal{G}_{\sigma_\mathbf{X}\sigma'_\mathbf{Z}\overline{\mathbf{Z}(\mathbf{W})}}, \quad (13)$$

where $\mathbf{Z}(\mathbf{W}) \subseteq \mathbf{Z}$ *is the set of elements in* $\mathbf{Z}$ *that are not ancestors of* $\mathbf{W}$ *in* $\mathcal{G}_{\sigma_{\mathbf{X}}}$.

The rules above follow from the semantics of $\sigma_{\mathbf{X}} = \{\sigma_{X_1}, \ldots, \sigma_{X_k}\}$ as indicator of the change in causal mechanism of each variable in $\mathbf{X}$ according to a specified strategy. Rule 1 ascertains the validity of the *d*-separation criterion for reading conditional independence constraints in the post-interventional distribution $P(\mathbf{V}; \sigma_{\mathbf{X}})$ using the interventional graph $\mathcal{G}_{\sigma_{\mathbf{X}}}$. Rule 2 establishes a condition that guarantees that the corresponding probability distribution is the same under interventions $\sigma'_{\mathbf{Z}}$ and $\sigma_{\mathbf{Z}}$ while $\mathbf{Z} = \mathbf{z}$ is observed. Rule 3 establishes a condition for changing the regime indicator from $\sigma'_{\mathbf{Z}}$ to $\sigma_{\mathbf{Z}}$ without affecting the associated probability. This rule differs from rule 2 since it is only applicable when $\mathbf{Z}$ is not observed.

In particular, these rules can be applied with $\sigma'_{\mathbf{Z}}$ having $\sigma_{Z_i}{=}\emptyset$, to make one or more regime indicators for $Z_i \in \mathbf{Z}$ idle. When all indicators are idle the expression is estimable from observational data. Differently than in the case of atomic interventions and *do*-calculus, causal diagrams induced by intervened models in this context are not necessarily subgraphs of the original diagram, hence σ-calculus needs to verify separation conditions in the corresponding two models. In Appendix. B, we revisit a classical example from (Pearl and Robins, 1995) that misses this subtlety and reaches an incorrect conclusion. The same appendix also provides a more detailed comparison of both calculi.

Comparison between σ-calculus and *do*-calculus

Independences in *do*-calculus rules usually include conditioning on $\mathbf{X}$. Notice that in our rules $\mathbf{W}$ could include variables in $\mathbf{X}$, accounting for situations when the expression has conditioning on $\mathbf{X}$ or part of it, but not necessarily the whole set every time.

The new rule 2 allows one to change across regimes when the variable under intervention is being observed. This is consistent with the traditional rule 2 and remark 1 about the $do(.)$ operator having an implicit conditioning on the intervened variable. Consider the back-door graph in Fig. 1(a) and an intervention $\sigma_X = do(g(z))$, which is associated with $\mathcal{G}_{\sigma_X} = \mathcal{G}$ (same argument in the observational and new interventional regime). Using the new rule 2, we have:

$$P(y \mid x, z; \sigma_X) = P(y \mid x, z),$$

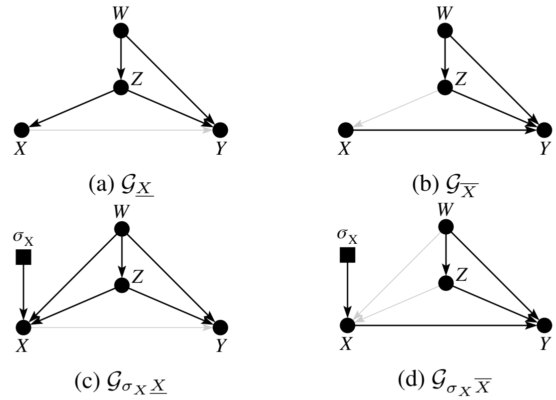
since $(X \perp\!\!\!\perp Y \mid Z)$ holds in both $\mathcal{G}_{\underline{X}}$ and $\mathcal{G}_{\sigma_X \underline{X}}$ (same graph in this case, Fig. 2(a)). One may be tempted to apply the new rule 2 as its *do*-calculus counterpart, trying to claim that

$$P(y \mid z; \sigma_X) = P(y \mid x, z).$$

However, this is not the case for many models compatible with the graph (see appendix for details).

Rule 3 licenses the addition or removal of a regime altogether. This rule is not the exact counterpart of the same rule in *do*-calculus. Consider again Fig. 1(a) with $\sigma_Z{=}P^*(z|w)$ and its effect on W conditioned on Z. Traditional rule 3 tests for $(W \perp\!\!\!\perp Z)$ in $\mathcal{G}_{\overline{Z}}$ which leads to

$$P(w \mid do(z)) = P(w \mid z; \sigma_Z) = P(w).$$



(a) $\mathcal{G}_{\underline{X}}$ (b) $\mathcal{G}_{\overline{X}}$ (c) $\mathcal{G}_{\sigma_X \underline{X}}$ (d) $\mathcal{G}_{\sigma_X \overline{X}}$

Figure 2: Graphs used to test the conditions required by rules 2 and 3 of σ-calculus in the derivation of the query in Example 2 where $\sigma_X = P^*(x|z, w)$. Arrows shown in gray indicate they have been cut.

In contrast, we consider the σ_Z in σ-calculus:

$$P(w \mid z; \sigma_Z) = P(w)P(z \mid w; \sigma_Z)/P(z; \sigma_Z),$$

which is almost always different than $P(w)$. The important distinction to make at this point is that for soft intervention on Z, we are not necessarily conditioning on it. Without conditioning, rule 3 of σ-calculus and independence $(W \perp\!\!\!\perp Z)$ in $\mathcal{G}_{\overline{Z}}$ (and $\mathcal{G}_{\sigma_Z \overline{Z}}$) yield $P(w; \sigma_Z){=}P(w)$. In contrast

$$P(w|z; \sigma_Z = do(z)) = P(w; \sigma_Z = do(z)) = P(w)$$

can be obtained by applying first rule 1 with $(W \perp\!\!\!\perp Z)$ in $\mathcal{G}_{\sigma_Z = do(z)}$ and then rule 3 (more discussion in appendix).

Examples of Symbolic Derivations

We illustrate the use of σ-calculus rules by solving the question in Example 2. Recall that our goal is to identify $P(y; \sigma_X)$ with $\sigma_X = P^*(x|w, z)$. We start by conditioning on the set $\{X, Z, W\}$,

$$P(y; \sigma_X) = \sum_{x,z,w} P(y|x, z, w; \sigma_X)P(x|z, w; \sigma_X)P(z, w; \sigma_X), \quad (14)$$

Note that Rule 2 can be applied with $\sigma'_X{=}\emptyset$ to infer $P(y|x, z, w; \sigma_X) = P(y|x, z, w)$ following the independence $(Y \perp\!\!\!\perp X \mid Z, W)$ in the $\mathcal{G}_{\underline{X}}$ and $\mathcal{G}_{\sigma_X \underline{X}}$ (see Figs. 2(a) and (c), respectively). Also, Rule 3 ($\sigma'_X{=}\emptyset$) leads to $P(z, w; \sigma_X) = P(z, w)$, licensed by $(Z, W \perp\!\!\!\perp X)$ in $\mathcal{G}_{\overline{X}}$ and $\mathcal{G}_{\sigma_X \overline{X}}$ (Figs. 2(b), (d)). Next, we replace $P(x|z, w; \sigma_X)$ using Eq. (6) by virtue of $\sigma_X{=}P^*(x|z, w)$:

$$P(y; \sigma_X) = \sum_{x,z,w} P(y|x, z, w)P^*(x|z, w)P(z, w). \quad (15)$$

Notice that all terms in the right hand side of Eq. (15) are either obtainable from $P(\mathbf{V})$ or defined by the new intervention itself, which means the target effect is identifiable (see appendix B for a more detailed example).

A natural albeit important consequence of Thm. 1 is described in the following corollary:

Corollary 1. *Considering only atomic (and idle) interventions, σ-calculus reduces to do-calculus.*

Identifying Effects with (Atomic and Non-atomic) Surrogate Experiments

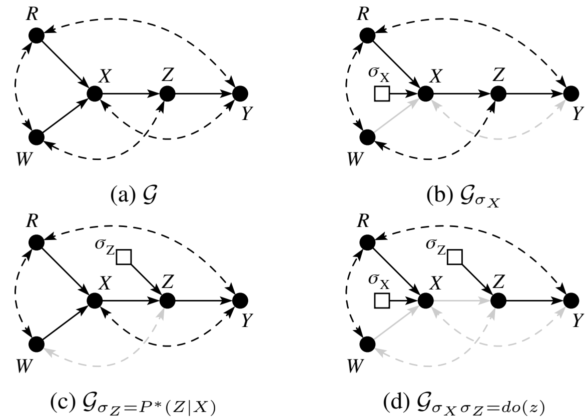
It's not uncommon that the effect of a certain intervention is not identifiable from observational data alone whenever unobserved confounders are present. It may be the case that experiments over surrogate variables may be available for use, which has been called in the literature the problem of *z-identification* (Bareinboim and Pearl, 2012; Lee et al., 2019). For instance, experiments over a set of surrogate variables $\mathbf{Z}$ may be more accessible to manipulation than the target effect $\sigma_{\mathbf{X}}$. In this case, still, experiments are assumed to be the product of controlled trials, that is, of atomic interventions. In this section, we leverage data from surrogate experiments obtained from general interventions.

Example 3. To illustrate this setting, consider the causal diagram in Fig. 3(a) and the effect $P(y|r, z; \sigma_X = P^*(X|R))$, which is not identifiable from $P(\mathbf{V})$. However, suppose a distribution $P(\mathbf{V}; \sigma_Z = P^*(Z|X))$ is given as an additional input. We can then write the target effect as

$$
\begin{aligned}
&P(y|r, z; \sigma_X) \\
&= P(y|r, z; \sigma_X, \sigma_Z{=}do(z)) &(16)\\
&= P(y|r, z; \sigma_Z{=}do(z)) &(17)\\
&= \sum_{x'} P(y|r, x', z; \sigma_Z{=}do(z)) P(x'|r, z; \sigma_z{=}do(z)). &(18)
\end{aligned}
$$

Eq. (16) follows from Rule 2 and the independence $(Y \perp\!\!\!\perp Z \mid R)$ in $\mathcal{G}_{\sigma_X \sigma_Z=\emptyset \underline{Z}}$ and $\mathcal{G}_{\sigma_X \sigma_Z=do(z)\underline{Z}}$; Eq. (17) by Rule 3 with $(Y \perp\!\!\!\perp X \mid R, Z)$ in $\mathcal{G}_{\sigma_X \sigma_Z=do(z)\overline{X}}$ and $\mathcal{G}_{\sigma_X=\emptyset \sigma_Z=do(z)\overline{X}}$. At this point, note that if the given experiment was randomized (i.e., $do(z)$), the target effect would be identifiable. However, the given distribution came from policy $\sigma_Z{=}P^*(Z|X)$. Nevertheless, we can condition on X and obtain Eq. (18). We can then apply Rule 2 to change the strategy of σ_Z from $do(Z)$ to $P^*(Z|X)$ due to $(Y \perp\!\!\!\perp Z \mid R, X)$ in $\mathcal{G}_{\sigma_Z=do(z)\underline{Z}}$ and $\mathcal{G}_{\sigma_Z=P^*(Z|X)\underline{Z}}$ that license $P(y|r, x', z; \sigma_Z{=}do(z)) = P(y|r, x', z; \sigma_Z{=}P^*(Z|X))$. Finally, the second factor in Eq. (18) can be obtained from the observational data by applying Rule 1 with $(X \perp\!\!\!\perp Z)$ in $\mathcal{G}_{\sigma_Z=do(z)}$ to remove the observation on z, followed by Rule 3 and $(X \perp\!\!\!\perp Z \mid R)$ in $\mathcal{G}_{\sigma_Z \overline{Z}}$ and $\mathcal{G}_{\sigma_Z=\emptyset \overline{Z}}$ to change σ_Z to the idle regime. Putting the pieces together, we obtain the following expression:

$$
P(y|r, z; \sigma_X) = \sum_{x'} P(y|r, x', z; \sigma_Z{=}P^*(Z|X))\, P(x'|r). \quad (19)
$$



(a) $\mathcal{G}$ (b) $\mathcal{G}_{\sigma_X}$ (c) $\mathcal{G}_{\sigma_Z=P^*(Z|X)}$ (d) $\mathcal{G}_{\sigma_X \sigma_Z=do(z)}$

Figure 3: (a) is the original diagram for which we want to identify the effect $P(y|r, z; \sigma_X{=}P^*(X|R))$ corresponding to the diagram in (b). Experimental data is given in the form of $P(\mathbf{V}; \sigma_Z{=}P^*(Z|X))$ corresponding to (c). Diagram in (d) is intermediate in the derivation of the target effect (see text for details).

5 Identifying the Effect of General Interventions Systematically

Even though σ-calculus is a great tool for understanding and reasoning about the logical implications of general interventions, searching for a derivation in moderately-sized causal models can be a very challenging task given the combinatorial nature of the problem. Also, the solution of realistic applications involving models with thousands of variables requires the use of computers. In this section, we develop an algorithmic solution for identifying the (conditional) effect of general interventions (Table 1) from observational and experimental data, based on a given causal diagram $\mathcal{G}$.

Consider a query of interest $P(\mathbf{y}, \mathbf{w}; \sigma_{\mathbf{X}})$ and let $\mathbf{D} = An(\mathbf{Y} \cup \mathbf{W})_{\mathcal{G}_{\sigma_{\mathbf{X}}}}$, then from Eq. (7) we can sum out variables that are not ancestors of $(\mathbf{Y} \cup \mathbf{W})$ and obtain

$$
\begin{aligned}
P(\mathbf{y}, \mathbf{w}; \sigma_{\mathbf{X}}) &= \sum_{\mathbf{v} \setminus (\mathbf{y} \cup \mathbf{w})} P(\mathbf{v}; \sigma_{\mathbf{X}}) \\
&= \sum_{\mathbf{d} \setminus (\mathbf{y} \cup \mathbf{w})} \sum_{\mathbf{u}^*} \prod_{\{i | V_i \in \mathbf{D}\}} P(v_i | pa_i, u_i; \sigma_{\mathbf{X}}) P(\mathbf{u}^*; \sigma_{\mathbf{X}}). \quad (20)
\end{aligned}
$$

For convenience, and following (Tian and Pearl, 2002a), we define for any $\mathbf{C} \subseteq \mathbf{V}$ the quantity $Q[\mathbf{C}](\mathbf{v})$, called *c-factor*, to denote the following function

$$
\begin{aligned}
&Q[\mathbf{C}; \sigma_{\mathbf{X}}](\mathbf{v}) = \\
&\quad \sum_{u(\mathbf{C})} \prod_{\{i | V_i \in \mathbf{C}\}} P(v_i | pa_i, u_i; \sigma_{\mathbf{X}}) P(u(\mathbf{C}); \sigma_{\mathbf{X}}), \quad (21)
\end{aligned}
$$

where $U(\mathbf{C}) = \bigcup_{V_i \in \mathbf{C}} U_i$. In particular, note that $Q[\mathbf{V}; \sigma_{\mathbf{X}}](\mathbf{v}) {=} P(\mathbf{v}; \sigma_{\mathbf{X}})$ and when $\sigma_{\mathbf{X}} = \emptyset$, $Q[\mathbf{C}; \sigma_{\mathbf{X}}] = Q[\mathbf{C}]$. For convenience, we will often write $Q[\mathbf{C}](\mathbf{v})$ as $Q[\mathbf{C}]$, and whenever $\mathbf{C} = \{V_i\}$ we will write $Q[V_i]$ instead of $Q[\{V_i\}]$. Using c-factors, Eq. (20) translates into $P(\mathbf{y}, \mathbf{w}; \sigma_{\mathbf{X}}) = \sum_{\mathbf{d} \setminus (\mathbf{y} \cup \mathbf{w})} Q[\mathbf{D}; \sigma_{\mathbf{X}}]$. Now consider the query $P(\mathbf{y} | \mathbf{w}; \sigma_{\mathbf{X}})$, we can write

$$
P(\mathbf{y}|\mathbf{w}; \sigma_{\mathbf{X}}) = \frac{P(\mathbf{y}, \mathbf{w}; \sigma_{\mathbf{X}})}{\sum_{\mathbf{y}} P(\mathbf{y}, \mathbf{w}; \sigma_{\mathbf{X}})} = \frac{\sum_{\mathbf{d} \setminus (\mathbf{y} \cup \mathbf{w})} Q[\mathbf{D}; \sigma_{\mathbf{X}}]}{\sum_{\mathbf{d} \setminus \mathbf{w}} Q[\mathbf{D}; \sigma_{\mathbf{X}}]}. \quad (22)
$$

Moreover, Eq. (22) can be further simplified as stated in the following

Lemma 1. *Let $\mathbf{Y}, \mathbf{X}, \mathbf{W} \subset \mathbf{V}$ with $\mathbf{W} \cap \mathbf{Y} = \emptyset$ and let $\mathcal{G}$ be a causal diagram over the variables $\mathbf{V}$. The effect $P(\mathbf{y} \mid \mathbf{w}; \sigma_{\mathbf{X}})$ is given by:*

$$
P(\mathbf{y}|\mathbf{w}; \sigma_{\mathbf{X}}) = \sum\nolimits_{\mathbf{a} \setminus (\mathbf{y} \cup \mathbf{w})} Q[\mathbf{A}; \sigma_{\mathbf{X}}] \Big/ \sum\nolimits_{\mathbf{a} \setminus \mathbf{w}} Q[\mathbf{A}; \sigma_{\mathbf{X}}], \quad (23)
$$

where $\mathbf{A}$ is the set of all variables connected to $\mathbf{Y}$ (including $\mathbf{Y}$) by any path (regardless of the directionality) in $\mathcal{G}_{\sigma_{\mathbf{X}}[\mathbf{D}]\underline{\mathbf{W}}}$, with $\mathbf{D} = An(\mathbf{Y} \cup \mathbf{W})_{\mathcal{G}_{\sigma_{\mathbf{X}}}}$.

The problem we need to solve now is to determine if, and how, the c-factor $Q[\mathbf{A}; \sigma_{\mathbf{X}}]$ can be computed from the observed data (i.e., $P(\mathbf{V})$=$Q[\mathbf{V}]$). To do so, we will leverage the machinery developed by (Tian and Pearl, 2002a; Huang and Valtorta, 2006) that deals with the identification of c-factors from other (larger) c-factors. First, note that the set of observable variables present in a causal diagram $\mathcal{G}$ can be partitioned into sets called *c-components* (Tian and Pearl, 2002a). Two variables are in the same c-component set if and only if they are connected by a path composed entirely of bidirected edges in $\mathcal{G}$. Using this notion we state the following results, which will be key for our algorithm:

Lemma 2. *Let $\mathbf{A}$ be defined relative to $\mathbf{X}, \mathbf{Y}, \mathbf{W}$ as in lemma 1, then:*

$$Q[\mathbf{A}; \sigma_{\mathbf{X}}] = Q\left[\mathbf{A}^{\mathbf{X}}; \sigma_{\mathbf{X}}\right] Q\left[\mathbf{A} \setminus \mathbf{A}^{\mathbf{X}}\right], \qquad (24)$$

where $\mathbf{A}^{\mathbf{X}}$ is the union of the c-components of $\mathcal{G}_{\sigma_{\mathbf{X}}[\mathbf{A}]}$ containing variables in $\mathbf{X}$.

Note that the factor $Q[\mathbf{A} \setminus \mathbf{A}^{\mathbf{X}}]$ in Eq. (24) corresponds to the idle regime. Hence, to asses if such c-factor is computable from $P(\mathbf{V})$ or a given $P(\mathbf{V}; \sigma_{\mathbf{Z}})$, we can use the algorithm IDENTIFY from (Tian and Pearl, 2002a) and the following lemma.

Lemma 3. *Let $\sigma_{\mathbf{Z}}$ indicate any intervention on $\mathbf{Z}$ and let $\mathbf{C} \subseteq \mathbf{V}$. Then, $Q[\mathbf{C}] = Q[\mathbf{C}; \sigma_{\mathbf{Z}}]$ if $\mathbf{C} \cap \mathbf{Z} = \emptyset$.*

What is left is to reason about the c-factor $Q[\mathbf{A}^{\mathbf{X}}; \sigma_{\mathbf{X}}]$. In the case of atomic, conditional and stochastic interventions; $\mathbf{A}^{\mathbf{X}}$ is simply $\mathbf{X}$ because for those interventions, variables in $\mathbf{X}$ do not share unobservable parents with any other variable under intervention. Therefore, $Q[\mathbf{A}^{\mathbf{X}}; \sigma_{\mathbf{X}}] = \prod_{X \in \mathbf{X}} Q[X; \sigma_{\mathbf{X}}]$ where each $Q[X; \sigma_{\mathbf{X}}] = Q[X; \sigma_X]$ is obtained by replacing it with the corresponding equation among (4), (5) or (6).

Following the discussion in this section, we propose the algorithm σ-IDENTIFY (Alg. 1). This procedure takes as input the variables defining a query, the specification of $\sigma_{\mathbf{X}}$ (i.e., what type of intervention is being applied and its arguments), a set of available distributions ($\mathbb{Z} = \{\sigma_\emptyset\}$ when only $P(\mathbf{V})$ is known.) and the causal diagram. The subroutine 'REPLACE' handles factors of intervened variables, replacing them according to the type of intervention. σ-IDENTIFY runs in $O(n^4 z)$ time, where n is the number of nodes in $\mathcal{G}$ and $z = |\mathbb{Z}|$ (see Appendix C).

For an illustration, we run σ-IDENTIFY to identify $P(y|r; \sigma_X{=}P^*(X|R))$ in Fig. 3(a), where $\mathbf{Y} = \{Y\}$, $\mathbf{W} = \{R\}$ and $\sigma_{\mathbf{X}} = \{\sigma_X{=}P^*(X|R)\}$; from observation and experimental data $\mathbb{Z} = \{\sigma_\emptyset, \sigma_Z{=}P^*(Z|X)\}$. Here, $\mathbf{A} = \{R, X, Z, Y\}$ and the c-components of $\mathcal{G}_{\sigma_X[\mathbf{A}]}$ are $\mathbf{A}_1 = \{R, Y\}$, $\mathbf{A}_2 = \{Z\}$ and $\mathbf{A}_3 = \{X\}$. The loop in line 2 will pick up $\mathbf{A}_1$ with $\sigma_{\mathbf{Z}} = \sigma_Z = P^*(Z|X)$, where $\mathbf{B}_i = \{R, X, Y, W\}$ is the c-component of $\mathcal{G}_{\sigma_{\mathbf{Z}}}$ (Fig. 3(c)) containing $\mathbf{A}_1$, for which IDENTIFY will return $\sum_{x'} P(y|r, x', z; \sigma_Z) P(x', r; \sigma_Z)$. Next, in the same loop, $\mathbf{A}_2$ matches with $\sigma_{\mathbf{Z}} = \sigma_\emptyset$ and IDENTIFY returns $\sum_{r', w} P(z|r', w, x) P(r', w)$. Line 9 handles $\mathbf{A}_3$ and replaces it with $P^*(x|r)$ according to the intervention $\sigma_{\mathbf{X}}$. Finally, the return expression is

$$P(y|r; \sigma_X) = Q[\mathbf{A}; \sigma_{\mathbf{X}}] \Big/ \sum\nolimits_y Q[\mathbf{A}; \sigma_{\mathbf{X}}], \qquad (25)$$

where

$$Q[\mathbf{A}; \sigma_{\mathbf{X}}] = \sum_{x,z} P^*(x|r) \Big(\sum_{x'} P(y|r, x', z; \sigma_Z) P(r, x'; \sigma_Z)\Big) \Big(\sum_{r', w} P(z|r', w, x) P(r', w)\Big) \qquad (26)$$

Algorithm 1 σ-IDENTIFY$(\mathbf{Y}, \mathbf{W}, \sigma_{\mathbf{X}}, \mathbb{Z}, \mathcal{G})$

Input: $\mathcal{G}$ causal diagram over a set of variables $\mathbf{V}$, $\mathbf{Y}, \mathbf{W} \subseteq \mathbf{V}$ disjoint subsets of variables, an intervention strategy $\sigma_{\mathbf{X}}$ defined over a set $\mathbf{X} \subseteq \mathbf{V}$, and a set $\mathbb{Z} = \{\sigma_{\mathbf{Z}_i}\}_{i=1}^n$ of known (interventional) distributions.
Output: $P(\mathbf{y}|\mathbf{w}; \sigma_{\mathbf{X}})$ in terms of available distributions or FAIL.

1: let $\mathbf{A}$ be defined as in lemma 1, and let $\mathbf{A}_1, \ldots, \mathbf{A}_n$ be the set of c-components of $\mathcal{G}_{\sigma_{\mathbf{X}}[\mathbf{A}]}$.
2: **for each** $\mathbf{A}_i$ containing no variable in $\mathbf{X}$ and every $\sigma_{\mathbf{Z}} \in \mathbb{Z}$ such that $\mathbf{A}_i \cap \mathbf{Z} = \emptyset$ **do**
3:     let $\mathbf{B}_i$ be the c-component of $\mathcal{G}_{\sigma_{\mathbf{Z}}}$ such that $\mathbf{A}_i \subseteq \mathbf{B}_i$.
4:     **if** IDENTIFY$(\mathbf{A}_i, \mathbf{B}_i, Q[\mathbf{B}_i], \mathcal{G}_{\sigma_{\mathbf{Z}}})$ does not FAIL **then**
5:         $Q[\mathbf{A}_i; \sigma_{\mathbf{X}}] =$ IDENTIFY$(\mathbf{A}_i, \mathbf{B}_i, Q[\mathbf{B}_i], \mathcal{G}_{\sigma_{\mathbf{Z}}})$.
6:         move to the next $\mathbf{A}_i$.
7:     **end if**
8: **end for**
9: **for each** $\mathbf{A}_i$ containing variables in $\mathbf{X}$ let $Q[\mathbf{A}_i; \sigma_{\mathbf{X}}] =$ REPLACE$(\mathbf{A}_i, \sigma_{\mathbf{X}})$.
10: **if** any $Q[\mathbf{A}_i]$ was not assigned **then return** FAIL.
11: let $Q[\mathbf{A}; \sigma_{\mathbf{X}}] = \prod_i Q[\mathbf{A}_i; \sigma_{\mathbf{X}}]$.
12: **return** $\sum_{\mathbf{a} \setminus (\mathbf{y} \cup \mathbf{w})} Q[\mathbf{A}; \sigma_{\mathbf{X}}] \Big/ \sum_{\mathbf{a} \setminus \mathbf{w}} Q[\mathbf{A}; \sigma_{\mathbf{X}}]$.

Theorem 2. *The effect $P(\mathbf{y} \mid \mathbf{w}; \sigma_{\mathbf{X}})$ is identifiable if σ-IDENTIFY (Alg. 1) does not fail. Moreover, the expression returned is a valid estimand for the effect.*

6 Conclusions

In this paper, we introduced a set of inference rules for reasoning about the effect of general interventions (Thm. 1), which has been called σ-calculus. The σ-calculus allows one to discover and verify from the causal graph, logical statements about general interventions generated by an arbitrary SCM. We showed how these rules can be used to identify the effect of interventions from a combination of observational and experimental data. Finally, we developed an algorithm (Alg. 1) that decides in an automated fashion whether a reduction of the effect of interest to the set of observed quantities (observational and experimental) exists; if so, it also returns the corresponding mapping. The algorithm and σ-calculus were proven sound and efficient for the task of identification of general interventions (Thm. 2), subsuming previous treatment for atomic interventions by *do*-calculus.

Acknowledgements

This research was supported by grants from IBM Research, Adobe Research, NSF IIS-1704352, and IIS-1750807 (CAREER).

References

Aldrich, J. (1989). Autonomy. *Oxford Economic Papers*, 41:15–34.

Bareinboim, E., Brito, C., and Pearl, J. (2012). Local Characterizations of Causal Bayesian Networks. In Croitoru, M., Rudolph, S., Wilson, N., Howse, J., and Corby, O., editors, *GKR 2011, LNAI 7205*, pages 1–17.

Bareinboim, E. and Pearl, J. (2012). Causal Inference by Surrogate Experiments: z-Identifiability. In *Proceedings of the Twenty-Eighth Conference on Uncertainty in Artificial Intelligence*, pages 113–120. AUAI Press.

Bareinboim, E. and Pearl, J. (2014). Transportability from Multiple Environments with Limited Experiments: Completeness Results. In *Advances in Neural Information Processing Systems 27*, pages 280–288.

Bareinboim, E. and Pearl, J. (2016). Causal inference and the data-fusion problem. *Proceedings of the National Academy of Sciences*, 113(27):7345–7352.

Cartwright, N. (2007). *Hunting Causes and Using Them: Approaches in Philosophy and Economics*. Cambridge University Press, New York, NY.

Correa, J. D. and Bareinboim, E. (2019). From Statistical Transportability to Estimating the Effects of Stochastic Interventions. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*.

Dawid, A. P. (2002). Influence diagrams for causal modelling and inference. *International Statistical Review*, 70:161–189.

Dawid, A. P. (2015). Statistical Causality from a Decision-Theoretic Perspective. *Annual Review of Statistics and Its Application*, 2(1):273–303.

Didelez, V., Dawid, A. P., and Geneletti, S. (2006). Direct and indirect effects of sequential treatments. In *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence*, pages 138–146. AUAI Press.

Fisher, R. A. (1951). *The Design of Experiments*. Oliver and Boyd, Edinburgh, 6th edition.

Haavelmo, T. (1943). The Statistical Implications of a System of Simultaneous Equations. *Econometrica*, 11(1):1–12.

Heckman, J. J. (2005). The Scientific Model of Causality. *Sociological Methodology*, 35:1–97.

Huang, Y. and Valtorta, M. (2006). Identifiability in Causal Bayesian Networks: A Sound and Complete Algorithm. In *Proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI 2006)*, pages 1149–1156. AAAI Press.

Huang, Y. and Valtorta, M. (2008). On the completeness of an identifiability algorithm for semi-Markovian models. *Annals of Mathematics and Artificial Intelligence*, 54(4):363–408.

Lee, S., Correa, J. D., and Bareinboim, E. (2019). General Identifiability with Arbitrary Surrogate Experiments. In *Proceedings of the Thirty-Fifth Conference Annual Conference on Uncertainty in Artificial Intelligence*, Corvallis, OR. AUAI Press.

Lewis, D. (1973). *Counterfactuals*. Harvard University Press, Cambridge, MA.

Neyman, J. (1923). Sur les applications de la thar des probabilities aux experiences Agaricales: Essay des principle.

Pearl, J. (1993a). Aspects of graphical models connected with causality. *Proceedings of the 49th Session of the International Statistical Institute*, 1(August):399–401.

Pearl, J. (1993b). Graphical Models, Causality, and Intervention. *Statistical Science*, 8(3):266–273.

Pearl, J. (1994). A probabilistic calculus of actions. In *Uncertainty in Artificial Intelligence 10*, pages 454–462. Morgan Kaufmann, San Mateo, CA.

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, 82(4):669–688.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press, New York, NY, USA, 2nd edition.

Pearl, J. (2010). Review of N. Cartwright 'Hunting Causes and Using Them'. *Economics and Philosophy*, 26:69–77.

Pearl, J. and Mackenzie, D. (2018). *The Book of Why*. Basic Books, New York.

Pearl, J. and Robins, J. M. (1995). Probabilistic evaluation of sequential plans from causal models with hidden variables. In Besnard, P. and Hanks, S., editors, *Uncertainty in Artificial Intelligence 11*, pages 444–453. Morgan Kaufmann, San Francisco.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688–701.

Shpitser, I. and Pearl, J. (2006a). Identification of Conditional Interventional Distributions. In *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence*, pages 437–444, Corvallis, OR. AUAI Press.

Shpitser, I. and Pearl, J. (2006b). Identification of Joint Interventional Distributions in Recursive semi-Markovian Causal Models. In *Proceedings of the Twenty-First AAAI Conference on Artificial Intelligence*, volume 2, pages 1219–1226.

Shpitser, I. and Sherman, E. (2018). Identification of Personalized Effects Associated With Causal Pathways. In *Proceedings of the 34th Conference on Uncertainty in Artificial Intelligence*, pages 530–539.

Spirtes, P., Glymour, C. N., and Scheines, R. (2001). *Causation, Prediction, and Search*. MIT Press, 2nd edition.

Strotz, R. H. and Wold, H. O. A. (1960). Recursive versus nonrecursive systems: An attempt at synthesis. *Econometrica*, 28:417–427.

Tian, J. (2004). Identifying Conditional Causal Effects. In *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*, UAI '04, pages 561–568, Arlington, Virginia, United States. AUAI Press.

Tian, J. (2008). Identifying Dynamic Sequential Plans. In *In Proceedings of the Twenty-Fourth Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-08)*, pages 554–561, Corvallis, Oregon. AUAI Press.

Tian, J. and Pearl, J. (2002a). A General Identification Condition for Causal Effects. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI 2002)*, pages 567–573. AAAI Press/The MIT Press.

Tian, J. and Pearl, J. (2002b). On the Testable Implications of Causal Models with Hidden Variables. *Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence (UAI-02)*, pages 519–527.

Woodward, J. (2003). *Making Things Happen*. Oxford University Press, New York, NY.

A Calculus for Stochastic Interventions: Causal Effect Identification and Surrogate Experiments - Appendix

A Proofs for Section 3

Proposition 1. *Let $\sigma_{\mathbf{X}}$ be an intervention where every variable in $\mathbf{U}^* \setminus \mathbf{U}$ affects at most one $X \in \mathbf{X}$. Then for any Markovian model the effect $P(\mathbf{y};\sigma_{\mathbf{X}})$ is given by*

$$\sum\nolimits_{\mathbf{v}\setminus\mathbf{y}} \sum\nolimits_{u(\mathbf{X})} \prod\nolimits_{\{i|V_i\in\mathbf{X}\}} P(v_i|pa_i;\sigma_{\mathbf{X}}) \prod\nolimits_{\{i|V_i\in\mathbf{V}\setminus\mathbf{X}\}} P(v_i|pa_i). \tag{27}$$

Proof. In this kind of models the set $\mathbf{U}$ can partitioned into disjoint sets, with U_i affecting only V_i. Also, it is the case that $(U_i \perp\!\!\!\perp Pa_i)$ hence

$$P(\mathbf{u}) = \prod_i P(u_i) = \prod_i P(u_i|pa_i).$$

From Eq. 8 we can distribute each $P(u_i|pa_i)$ so that $P(v_i|pa_i,u_i)P(u_i) = P(v_i|pa_i,u_i)P(u_i|pa_i) = P(v_i,u_i|pa_i)$.

$$P(\mathbf{v};\sigma_{\mathbf{X}}) = \sum_{\mathbf{u}^*} \prod_{\{i|V_i\in\mathbf{X}\}} P(v_i|pa_i,u_i;\sigma_{\mathbf{X}})P(\mathbf{u}^*\setminus\mathbf{u};\sigma_{\mathbf{X}}) \prod_{\{i|V_i\in\mathbf{V}\setminus\mathbf{X}\}} P(v_i,u_i \mid pa_i). \tag{28}$$

The sum over $\mathbf{U}$ can be broken into sums over $\mathbf{U} \setminus \mathbf{U}$ and for each U_i. Pulling the U_i's sums into the product we have:

$$P(\mathbf{v};\sigma_{\mathbf{X}}) = \sum_{\mathbf{u}^*\setminus\mathbf{u}} \prod_{\{i|V_i\in\mathbf{X}\}} P(v_i|pa_i,u_i;\sigma_{\mathbf{X}})P(\mathbf{u}^*\setminus\mathbf{u};\sigma_{\mathbf{X}}) \prod_{\{i|V_i\in\mathbf{V}\setminus\mathbf{X}\}} \sum_{u_i} P(v_i,u_i \mid pa_i), \tag{29}$$

which simplifies to

$$P(\mathbf{v};\sigma_{\mathbf{X}}) = \sum_{\mathbf{u}^*\setminus\mathbf{u}} \prod_{\{i|V_i\in\mathbf{X}\}} P(v_i|pa_i,u_i;\sigma_{\mathbf{X}})P(\mathbf{u}^*\setminus\mathbf{u};\sigma_{\mathbf{X}}) \prod_{\{i|V_i\in\mathbf{V}\setminus\mathbf{X}\}} P(v_i \mid pa_i). \tag{30}$$

If the randomness of the interventions in $\sigma_{\mathbf{X}}$ is not shared among them (i.e. each $U_i \in \mathbf{U}^* \setminus \mathbf{U}$ only affects a particular $V_i \in \mathbf{X}$), then we can follow a similar reasoning to remove the sum over $\mathbf{U}^* \setminus \mathbf{U}$ and get

$$P(\mathbf{v};\sigma_{\mathbf{X}}) = \prod_{\{i|V_i\in\mathbf{X}\}} P(v_i|pa_i;\sigma_{\mathbf{X}}) \prod_{\{i|V_i\in\mathbf{V}\setminus\mathbf{X}\}} P(v_i|pa_i). \tag{31}$$

Finally, we can sum over $\mathbf{V} \setminus \mathbf{Y}$ in both sides to obtain Eq. (27). □

B Proofs and Examples for Section 4

Theorem 1. *[Inference Rules – σ-calculus] Let $\mathcal{G}$ be a causal diagram compatible with a structural causal model M, with endogenous variables $\mathbf{V}$. For any disjoint subsets $\mathbf{X},\mathbf{Y},\mathbf{Z} \subseteq \mathbf{V}$, two disjoint subsets $\mathbf{T},\mathbf{W} \subseteq \mathbf{V} \setminus (\mathbf{Z}\cup\mathbf{Y})$ (i.e., possibly including $\mathbf{X}$), the following rules are valid for any intervention strategies $\sigma_{\mathbf{X}},\sigma_{\mathbf{Z}}$, and $\sigma'_{\mathbf{Z}}$:*

Rule 1 *(Insertion/Deletion of observations):*

$$P(\mathbf{y} \mid \mathbf{w},\mathbf{t};\sigma_{\mathbf{X}}) = P(\mathbf{y} \mid \mathbf{w};\sigma_{\mathbf{X}}) \quad \textit{if } (\mathbf{Y} \perp\!\!\!\perp \mathbf{T} \mid \mathbf{W}) \textit{ in } \mathcal{G}_{\sigma_{\mathbf{X}}}. \tag{11}$$

Rule 2 *(Change of regimes under observation):*

$$P(\mathbf{y} \mid \mathbf{z},\mathbf{w};\sigma_{\mathbf{x}},\sigma_{\mathbf{z}}) = P(\mathbf{y} \mid \mathbf{z},\mathbf{w};\sigma_{\mathbf{x}},\sigma'_{\mathbf{z}}) \quad \textit{if } (\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{W}) \textit{ in } \mathcal{G}_{\sigma_{\mathbf{X}}\sigma_{\mathbf{Z}}\underline{\mathbf{Z}}} \textit{ and } \mathcal{G}_{\sigma_{\mathbf{X}}\sigma'_{\mathbf{Z}}\underline{\mathbf{Z}}}. \tag{12}$$

Rule 3 *(Change of regimes without observation):*

$$P(\mathbf{y} \mid \mathbf{w};\sigma_{\mathbf{x}},\sigma_{\mathbf{z}}) = P(\mathbf{y} \mid \mathbf{w};\sigma_{\mathbf{x}},\sigma'_{\mathbf{z}}) \quad \textit{if } (\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{W}) \textit{ in } \mathcal{G}_{\sigma_{\mathbf{X}}\sigma_{\mathbf{Z}}\overline{\mathbf{Z}(\mathbf{W})}} \textit{ and } \mathcal{G}_{\sigma_{\mathbf{X}}\sigma'_{\mathbf{Z}}\overline{\mathbf{Z}(\mathbf{W})}}, \tag{13}$$

where $\mathbf{Z}(\mathbf{W}) \subseteq \mathbf{Z}$ is the set of elements in $\mathbf{Z}$ that are not ancestors of $\mathbf{W}$ in $\mathcal{G}_{\sigma_{\mathbf{X}}}$.

Proof. **Rule 1** Both sides of the expression refer to the same model M_{σ_X} and the corresponding causal diagram $\mathcal{G}_{\sigma_X}$. So the condition licenses the equality by application of the d-separation criterion in the context of this pair.

To prove the next two rules, we will consider a new SCM M^* with observable variables $\mathbf{V}^* = \mathbf{V} \cup \sigma_{\mathbf{Z}}$, $\sigma_{\mathbf{Z}} = \{\sigma_Z\}_{Z\in\mathbf{Z}}$, that is a new node for each variable affected by interventions on variables in $\mathbf{Z}$. M^* has a set of unobservables $\mathbf{U}^* \supseteq \mathbf{U}$, and the distribution $P(\mathbf{u})$ is the same. Further, M^* has a set of functions $\mathcal{F}^*$ such that $f^*_{v_i} = f_{v_i}$ for $V_i \in \mathbf{V} \setminus (\mathbf{X}\cup\mathbf{Z})$, for $X \in \mathbf{X}$ let $f^*_x = f_{\sigma_X}$ (the function for X in the model $M_{\sigma_{\mathbf{X}}}$). Finally, for $f^*_z, Z \in \mathbf{Z}$ let

$$f^*_z = \begin{cases} f_{\sigma'_Z} & \text{if } \sigma_Z = 0 \\ f_{\sigma_Z} & \text{if } \sigma_Z = 1 \end{cases}, \tag{32}$$

where f_{σ_Z} is the function of Z in the model $M_{\sigma_{\mathbf{Z}}}$ and $f_{\sigma'_Z}$ the same function in $M_{\sigma'_{\mathbf{Z}}}$.

The model M^* induces a graph $\mathcal{G}^*$ where pa_i for any $V_i \in \mathbf{V}\setminus\mathbf{Z}$ is the same in as in $\mathcal{G}_{\sigma_{\mathbf{X}}}$, while $pa_i, V_i \in \mathbf{Z}$ is the union of the parents of Z in $\mathcal{G}_{\sigma'_{\mathbf{Z}}}$ and $\mathcal{G}_{\sigma_{\mathbf{Z}}}$.

Let P^* denote the probability distribution induced by M^*. We have that $P^*(\mathbf{v}|\sigma_{\mathbf{Z}} = 1)$ is exactly the same as $P(\mathbf{v};\sigma_{\mathbf{X}},\sigma_{\mathbf{Z}})$ while $P^*(\mathbf{v}|\sigma_{\mathbf{Z}} = 0)$ behaves as $P(\mathbf{v};\sigma_{\mathbf{X}},\sigma'_{\mathbf{Z}})$. It follows that for any pair of disjoint sets $\mathbf{A},\mathbf{B} \subset \mathbf{V}$:

$$P^*(\mathbf{a} \mid \mathbf{b},\sigma_{\mathbf{Z}} = 1) = P(\mathbf{a} \mid \mathbf{b};\sigma_{\mathbf{X}},\sigma_{\mathbf{Z}}), \text{ and} \tag{33}$$

$$P^*(\mathbf{a} \mid \mathbf{b},\sigma_{\mathbf{Z}} = 0) = P(\mathbf{a} \mid \mathbf{b};\sigma_{\mathbf{X}},\sigma'_{\mathbf{Z}}). \tag{34}$$

Rule 2 If $(\sigma_{\mathbf{Z}} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{W},\mathbf{Z})$ in $\mathcal{G}^*$ it follows

$$P(\mathbf{y} \mid \mathbf{z},\mathbf{w};\sigma_{\mathbf{X}},\sigma_{\mathbf{Z}}) = P^*(\mathbf{y} \mid \mathbf{z},\mathbf{w},\sigma_{\mathbf{Z}} = 1) \tag{35}$$

$$= P^*(\mathbf{y} \mid \mathbf{z},\mathbf{w},\sigma_{\mathbf{Z}} = 0) \tag{36}$$

$$= P(\mathbf{y} \mid \mathbf{z},\mathbf{w};\sigma_{\mathbf{X}},\sigma'_{\mathbf{Z}}). \tag{37}$$

If this independence does not hold, there exists a path from σ_Z to some $Y \in \mathbf{Y}$ in $\mathcal{G}^*$ that is d-connected given $\mathbf{W} \cup \mathbf{Z}$. Let $\overline{p}$ (without loss of generality) be such a path with no node in $\mathbf{Z}$ other than Z in it. The path $\overline{p}$ must start with $\sigma_Z \to Z \leftarrow A$, for some variable A, else it is blocked by conditioning on Z. The edge $(Z \leftarrow A)$ is either present in $\mathcal{G}_{\sigma_\mathbf{X}\sigma_\mathbf{Z}\underline{\mathbf{Z}}}$ or in $\mathcal{G}_{\sigma_\mathbf{X}\sigma'_\mathbf{Z}\underline{\mathbf{Z}}}$, which implies the portion of $\overline{p}$ from Z to Y is present in one of those graphs and d-connected given $\mathbf{W}$, which leads to a contradiction to the conditions in the rule.

Rule 3 If $(\sigma_\mathbf{Z} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{W})$ in $\mathcal{G}^*$ it follows

$$P(\mathbf{y} \mid \mathbf{w}; \sigma_\mathbf{X}, \sigma_\mathbf{Z}) = P^*(\mathbf{y} \mid \mathbf{w}, \sigma_\mathbf{Z} = 1) \tag{38}$$
$$= P^*(\mathbf{y} \mid \mathbf{w}, \sigma_\mathbf{Z} = 0) \tag{39}$$
$$= P(\mathbf{y} \mid \mathbf{w}; \sigma_\mathbf{X}, \sigma'_\mathbf{Z}). \tag{40}$$

If this independence does not hold, there exists a path from σ_Z to some $Y \in \mathbf{Y}$ in $\mathcal{G}^*$ that is d-connected given $\mathbf{W}$. Let $\overline{p}$ (without loss of generality) be such a path with no node in $\mathbf{Z}$ other than Z in it. If $\overline{p}$ starts with $\sigma_Z \to Z \leftarrow A$, Z must have a descendant in $\mathbf{W}$ which implies $Z \notin \mathbf{Z}(\mathbf{W})$. Hence the edge $(Z \leftarrow A)$ is in $\mathcal{G}_{\sigma_\mathbf{X}\sigma_\mathbf{Z}\overline{\mathbf{Z}(\mathbf{W})}}$ or $\mathcal{G}_{\sigma_\mathbf{X}\sigma'_\mathbf{Z}\overline{\mathbf{Z}(\mathbf{W})}}$. If $\overline{p}$ starts with $\sigma_Z \to Z \to A$, $(Z \to A)$ is also in one of those graphs.
Then, the portion of $\overline{p}$ from Z to Y exists either in $\mathcal{G}_{\sigma_\mathbf{X}\sigma_\mathbf{Z}\overline{\mathbf{Z}(\mathbf{W})}}$ or $\mathcal{G}_{\sigma_\mathbf{X}\sigma'_\mathbf{Z}\overline{\mathbf{Z}(\mathbf{W})}}$ and is d-connected given $\mathbf{W}$, a contradiction to at least one of the independences in the rule.

□

The next result shows that the rules of *do*-calculus follow from σ-calculus when only atomic (and idle) interventions are considered. Before presenting the proof we recall the rules of *do*-calculus:

Rule 1

$$P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{w}, \mathbf{t}) = P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{w}) \quad \text{if } (\mathbf{Y} \perp\!\!\!\perp \mathbf{T} \mid \mathbf{X}, \mathbf{W}) \text{ in } \mathcal{G}_{\overline{\mathbf{X}}}. \tag{41}$$

Rule 2

$$P(\mathbf{y} \mid do(\mathbf{x}), do(\mathbf{z}), \mathbf{w}) = P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{z}, \mathbf{w}) \quad \text{if } (\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{X}, \mathbf{W}) \text{ in } \mathcal{G}_{\overline{\mathbf{X}}\underline{\mathbf{Z}}}. \tag{42}$$

Rule 3

$$P(\mathbf{y} \mid do(\mathbf{x}), do(\mathbf{z}), \mathbf{w}) = P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{w}) \quad \text{if } (\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{X}, \mathbf{W}) \text{ in } \mathcal{G}_{\overline{\mathbf{X}\mathbf{Z}(\mathbf{W})}}, \tag{43}$$

where $\mathbf{Z}(\mathbf{W}) \subseteq \mathbf{Z}$ are not ancestors of $\mathbf{W}$ in $\mathcal{G}_{\overline{\mathbf{X}}}$.

Corollary 1. *Considering only atomic (and idle) interventions, σ-calculus reduces to do-calculus.*

Proof. Consider the rules of σ-calculus with $\sigma_\mathbf{X} = do(\mathbf{x})$, $\sigma_\mathbf{Z} = do(\mathbf{z})$ and $\sigma'_\mathbf{Z} = \emptyset$. Accordingly the graphs associated with the conditions become

$$\mathcal{G}_{\sigma_\mathbf{X}} = \mathcal{G}_{\overline{\mathbf{X}}}, \tag{44}$$
$$\mathcal{G}_{\sigma_\mathbf{Z}} = \mathcal{G}_{\overline{\mathbf{Z}}}, \tag{45}$$
$$\mathcal{G}_{\sigma'_\mathbf{Z}} = \mathcal{G}. \tag{46}$$

Also let $\mathbf{W}' = \mathbf{X} \cup \mathbf{W}$.

Rule 1

Consider rule 1 of σ-calculus, if $(\mathbf{Y} \perp\!\!\!\perp \mathbf{T} \mid \mathbf{W}') \equiv (\mathbf{Y} \perp\!\!\!\perp \mathbf{T} \mid \mathbf{X}, \mathbf{W})$ in $\mathcal{G}_{\overline{\mathbf{X}}}$ we have

$$P(\mathbf{y} \mid \mathbf{w}', \mathbf{t}; \sigma_\mathbf{X}) = P(\mathbf{y} \mid \mathbf{w}'; \sigma_\mathbf{X}), \tag{47}$$

replacing $\mathbf{W}'$ leads to

$$P(\mathbf{y} \mid \mathbf{x}, \mathbf{w}, \mathbf{t}; \sigma_\mathbf{X}) = P(\mathbf{y} \mid \mathbf{x}, \mathbf{w}; \sigma_\mathbf{X}). \tag{48}$$

Then by Eq. (10) we have

$$P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{w}, \mathbf{t}) = P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{w}), \tag{49}$$

which matches rule 1 in *do*-calculus.

Rule 2

Consider rule 2 of σ-calculus, that needs to be tested in $\mathcal{G}_{\overline{\mathbf{X}\mathbf{Z}}\underline{\mathbf{Z}}}$ and $\mathcal{G}_{\overline{\mathbf{X}}\underline{\mathbf{Z}}}$. It is trivial to see that any conditional independence satisfied by the latter is also satisfied by the former. Then if $(\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{W}') \equiv (\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{X}, \mathbf{W})$ in $\mathcal{G}_{\overline{\mathbf{X}}\underline{\mathbf{Z}}}$:

$$P(\mathbf{y} \mid \mathbf{z}, \mathbf{w}'; \sigma_\mathbf{X}, \sigma_\mathbf{Z}) = P(\mathbf{y} \mid \mathbf{z}, \mathbf{w}'; \sigma_\mathbf{X}), \tag{50}$$

replacing $\mathbf{W}'$ leads to

$$P(\mathbf{y} \mid \mathbf{z}, \mathbf{x}, \mathbf{w}, ; \sigma_\mathbf{X}, \sigma_\mathbf{Z}) = P(\mathbf{y} \mid \mathbf{z}, \mathbf{x}, \mathbf{w}; \sigma_\mathbf{X}). \tag{51}$$

Then by Eq. (10) we have

$$P(\mathbf{y} \mid do(\mathbf{x}), do(\mathbf{z}), \mathbf{w}) = P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{z}, \mathbf{w}), \tag{52}$$

matching rule 2 of *do*-calculus.

Rule 3

Consider rule 3 of σ-calculus, that needs to be tested in $\mathcal{G}_{\overline{\mathbf{X}\mathbf{Z}}\overline{\mathbf{Z}(\mathbf{W}')}}$ and $\mathcal{G}_{\overline{\mathbf{X}\mathbf{Z}(\mathbf{W}')}}$. As before, we only need to test the latter, because it satisfies a subset of the independencies of the former. Moreover, in that graph $\mathbf{X}$ has no ancestors (because of the $\overline{\mathbf{X}}$), hence no node in $\mathbf{Z}$ has descendants in $\mathbf{X}$ and $\mathbf{Z}(\mathbf{W} \cup \mathbf{X}) = \mathbf{Z}(\mathbf{W})$. Then if $(\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{W}') \equiv (\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{X}, \mathbf{W})$ in $\mathcal{G}_{\overline{\mathbf{X}\mathbf{Z}(\mathbf{W})}}$, we have

$$P(\mathbf{y} \mid \mathbf{w}'; \sigma_\mathbf{X}, \sigma_\mathbf{Z}) = P(\mathbf{y} \mid \mathbf{w}'; \sigma_\mathbf{X}), \tag{53}$$

replacing $\mathbf{W}'$ leads to

$$P(\mathbf{y} \mid \mathbf{x}, \mathbf{w}; \sigma_\mathbf{X}, \sigma_\mathbf{Z}) = P(\mathbf{y} \mid \mathbf{x}, \mathbf{w}; \sigma_\mathbf{X}). \tag{54}$$

Note that $(\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{X}, \mathbf{W})$ in $\mathcal{G}_{\overline{\mathbf{X}\mathbf{Z}(\mathbf{W})}}$ implies the same independence in $\mathcal{G}_{\overline{\mathbf{X}\mathbf{Z}}}$ which could only have less edges. Then, we can apply rule 1 of σ-calculus to introduce an observation on $\mathbf{z}$ to the expression in the left hand side:

$$P(\mathbf{y} \mid \mathbf{z}, \mathbf{x}, \mathbf{w}; \sigma_\mathbf{X}, \sigma_\mathbf{Z}) = P(\mathbf{y} \mid \mathbf{x}, \mathbf{w}; \sigma_\mathbf{X}); \tag{55}$$

finally by Eq. (10) we have

$$P(\mathbf{y} \mid do(\mathbf{x}), do(\mathbf{z}), \mathbf{w}) = P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{w}), \tag{56}$$

matching rule 3 of *do*-calculus.

□

Comparison between σ-calculus and *do*-calculus

Independences in *do*-calculus rules usually include conditioning on $\mathbf{X}$. Notice that in our rules $\mathbf{W}$ could include variables in $\mathbf{X}$, accounting for situations when the expression has conditioning on $\mathbf{X}$ or part of it, but not necessarily the whole set every time.

The new rule 2 allows one to change across regimes when the variable under intervention is being observed. This is consistent with the traditional rule 2 and remark 1 about the $do(.)$ operator having an implicit conditioning on the intervened variable. Consider the back-door graph in Fig. 1(a) and an intervention $\sigma_X = do(g(z))$, which is associated with $\mathcal{G}_{\sigma_X} = \mathcal{G}$ (same argument in the observational and new interventional regime). Using the new rule 2, we have:

$$P(y \mid x, z; \sigma_X) = P(y \mid x, z), \tag{57}$$

since $(X \perp\!\!\!\perp Y \mid Z)$ holds in both $\mathcal{G}_{\underline{X}}$ and $\mathcal{G}_{\sigma_X \underline{X}}$ (same graph in this case, Fig. 2(a)). One may be tempted to apply the new rule 2 as its *do*-calculus counterpart, trying to claim that

$$P(y \mid z; \sigma_X) = P(y \mid x, z). \tag{58}$$

However, this is not the case, to see why, first condition on X in the l.h.s.,

$$P(y \mid z; \sigma_X) = \sum_x P(y \mid x, z; \sigma_X) P(x \mid z; \sigma_X) \tag{59}$$

$$= \sum_x P(y \mid x, z) P(x \mid z; \sigma_X), \tag{60}$$

where the last equality follows from Eq. (57). Note that, in general, Eq. (60) will not be equal to $P(y \mid x, z)$.

Now consider the intervention $\sigma_X = P^*(X|W)$. Using rule 2 with $(Y \perp\!\!\!\perp X \mid Z, W)$ in $\mathcal{G}_{\underline{X}}$ and $\mathcal{G}_{\sigma_X \underline{X}}$ yield

$$P(y \mid x, z, w; \sigma_X) = P(y \mid x, z, w). \tag{61}$$

One might surmise that the effect of X on Y conditioned on Z, no longer on W, could be computed, i.e.,

$$P(y \mid x, z; \sigma_X) = P(y \mid x, z). \tag{62}$$

We can manipulate both sides of the equality to see clearly the differences. First we massage the expressions such that we have the probability of each variable given its parents, for the left hand side:

$$P(y \mid x, z; \sigma_X)$$

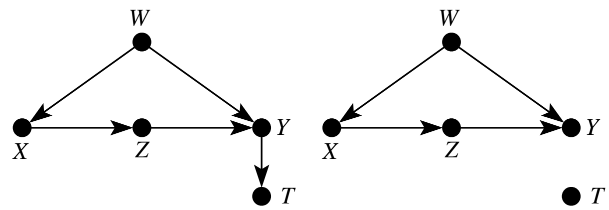
$$= \frac{P(y, x \mid z; \sigma_X)}{P(x \mid z; \sigma_X)} \tag{63}$$

$$= \frac{\sum_w P(y, x, \mid z, w; \sigma_X) P(w; \sigma_X)}{\sum_w P(x \mid z, w; \sigma_X) P(w; \sigma_X)} \tag{64}$$

$$= \frac{\sum_w P(y \mid x, z, w; \sigma_X) P(x \mid z, w; \sigma_X) P(w; \sigma_X)}{\sum_w P(x \mid z, w; \sigma_X) P(w; \sigma_X)} \tag{65}$$

From eq. (61) and rule 3 with $(W \perp\!\!\!\perp X)$ in $\mathcal{G}_{\sigma_X \overline{X}}$ and $\mathcal{G}_{\overline{X}}$ we have $P(w; \sigma_X) = P(w)$, it follows

$$P(y|x, z; \sigma_X) = \frac{\sum_w P(y|x, z, w) P(x|z, w; \sigma_X) P(w)}{\sum_w P(x|z, w; \sigma_X) P(w)} \tag{66}$$



(a) $\mathcal{G}$ (b) $\mathcal{G}_{\overline{T}}$

Figure 4: Model used to exemplify the use of rule 3 (see text).

For the right hand side,

$$P(y \mid x, z) = \sum_w P(y \mid x, z, w) P(w \mid x, z) \tag{67}$$

$$= \sum_w P(y \mid x, z, w) P(w) \tag{68}$$

If $P(x|z, w; \sigma_X)$ is independent of W then eqs. (66) and (68) become the same. It is easy to find a model not satisfying this constraint and equality. This example shows that even though there is no open confounding path passing through W in the observational case, W will be needed due to its role in the new policy affecting X - i.e., Eq. (61) holds while Eq. (62) does not.

Rule 3 licenses the addition or removal of a regime altogether. This rule is not the exact counterpart of the same rule in *do*-calculus. As mentioned before, having the $do(.)$ operator not only specifies a change in regime but a conditioning on the intervened variables. To illustrate this point, consider the causal diagram in Fig. 4(a) and the intervention $\sigma_T = do(t = P^*(t \mid y))$ and its effect on Y. Traditional rule 3 tests for $(Y \perp\!\!\!\perp T)$ in $\mathcal{G}_{\overline{T}}$ (Fig. 4(b)) which leads to

$$P(y \mid do(t)) = P(y \mid t; \sigma_T) = P(y). \tag{69}$$

In contrast, we consider the σ_t in σ-calculus:

$$P(y \mid t; \sigma_T) = \frac{P(y) P(t \mid y; \sigma_T)}{P(t; \sigma_T)}, \tag{70}$$

which is almost always different than $P(y)$. The important distinction to make at this point is that for soft intervention on T, we are not necessarily conditioning on it. By rule 3 of σ-calculus and the independence $(Y \perp\!\!\!\perp T)$ in $\mathcal{G}_{\overline{T}}$ (and $\mathcal{G}_{\sigma_T \overline{T}}$) we have

$$P(y; \sigma_T) = P(y). \tag{71}$$

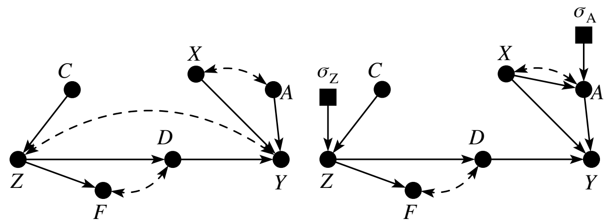
If the intervention was $\sigma_T = do(t)$, to reach the same conclusion as in (69), we can use rule 1 with $(Y \perp\!\!\!\perp T)$ in $\mathcal{G}_{\sigma_T} = \mathcal{G}_{\overline{T}}$ and rule 3 with $(Y \perp\!\!\!\perp T)$ in $\mathcal{G}_{\overline{T}}$ (note that the independences collapsed to the same one), i.e.,

$$P(y \mid do(t)) = P(y \mid t; \sigma_T) \tag{72}$$

$$= P(y; \sigma_T) \quad \text{(Rule 1)} \tag{73}$$

$$= P(y). \quad \text{(Rule 3)} \tag{74}$$

Consider the same causal diagram with the intervention $\sigma_X = g(w)$. The independece $(Y \perp\!\!\!\perp X \mid Z, W)$ in



(a) $\mathcal{G}$ (b) $\mathcal{G}_{\sigma_X\sigma_A}$

Figure 5: Pair of models associated with Example 4.

$\mathcal{G}_{\sigma_X\overline{X(Z,W)}} = \mathcal{G}_{\sigma_X} = \mathcal{G}$ licenses the application of rule 1 and 3, which implies, respectively,

$$P(y \mid x, z, w; \sigma_X) = P(y \mid z, w; \sigma_X) = P(y \mid z, w). \tag{75}$$

More elaborate examples

Example 4 (A derivation in σ-calculus)**.** Consider the causal diagram in Fig. 5(a) and the target effect $P(y \mid x, z; \sigma_Z, \sigma_A)$ with $\sigma_Z = P^*(z|c)$ and $\sigma_A = P^*(a|x)$. We can use the rules of σ-calculus to derive the effect as follows:

$$P(y \mid x, z; \sigma_Z, \sigma_A) \tag{76}$$

condition on D, A

$$= \sum_{d,a} P(y \mid x, z, d, a; \sigma_Z, \sigma_A) P(d, a \mid x, z; \sigma_Z, \sigma_A) \tag{77}$$

Rule 1: $(Y \perp\!\!\!\perp Z \mid X, D, A)$ in $\mathcal{G}_{\sigma_Z\sigma_A}$. Also since $(D, Z \perp\!\!\!\perp A, X)$ in $\mathcal{G}_{\sigma_Z\sigma_A}$ we can factorize the second term:

$$= \sum_{d,a} P(y \mid x, d, a; \sigma_Z, \sigma_A) \\ P(d \mid z; \sigma_Z, \sigma_A) P(a \mid x; \sigma_Z, \sigma_A) \tag{78}$$

Rule 2: $(Y \perp\!\!\!\perp D \mid X, A)$ in $\mathcal{G}_{\sigma_Z\sigma_A\underline{D}}, \mathcal{G}_{\sigma_Z\sigma_A\sigma_D=do(d)\underline{D}}$.

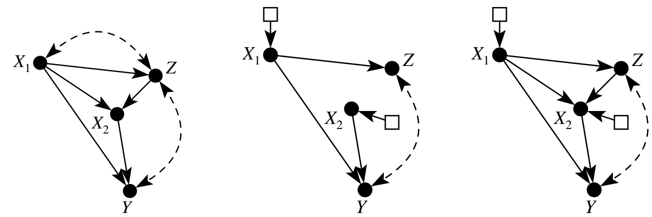
$$= \sum_{d,a} P(y \mid x, d, a; \sigma_Z, \sigma_A, \sigma_D) \\ P(d \mid z; \sigma_Z, \sigma_A) P(a \mid x; \sigma_Z, \sigma_A) \tag{79}$$

Rule 3: $(D \perp\!\!\!\perp A \mid Z)$ in $\mathcal{G}_{\sigma_Z\sigma_A\overline{A}}$ and $\mathcal{G}_{\sigma_Z\overline{A}}$, also rule 3: $(A \perp\!\!\!\perp Z \mid X)$ in $\mathcal{G}_{\sigma_Z\sigma_A\overline{Z}}$ and $\mathcal{G}_{\sigma_A\overline{Z}}$

$$= \sum_{d,a} P(y|x, d, a; \sigma_Z, \sigma_A, \sigma_D) P(d|z; \sigma_Z) P(a|x; \sigma_A) \tag{80}$$

Rule 3: $(Y \perp\!\!\!\perp Z \mid X, D, A)$ in $\mathcal{G}_{\sigma_Z\sigma_A\sigma_D\overline{Z(X,D,A)}} = \mathcal{G}_{\sigma_Z\sigma_A\sigma_D}$

$$= \sum_{d,a} P(y \mid x, d, a; \sigma_A, \sigma_D) P(d \mid z; \sigma_Z) P(a \mid x; \sigma_A) \tag{81}$$



(a) $\mathcal{G}$ (b) $\mathcal{G}_{\sigma_\mathbf{X}}$ (C1) (c) $\mathcal{G}_{\sigma_\mathbf{X}}$ (C2)

Figure 6: Causal diagrams associated with Example 5.

Rule 2: $(D \perp\!\!\!\perp Z)$ in $\mathcal{G}_{\sigma_Z\underline{Z}}, \mathcal{G}_{\underline{D}}$.

$$= \sum_{d,a} P(y \mid x, d, a; \sigma_A, \sigma_D) P(d \mid z) P(a \mid x; \sigma_A) \tag{82}$$

Reordering the terms and conditioning on Z (as Z' to avoid confusion with the argument Z in the query):

$$= \sum_{d,a} P(d \mid z) P(a \mid x; \sigma_A) \\ \sum_{z'} P(y \mid x, z', d, a; \sigma_A, \sigma_D) P(z' \mid x, d, a; \sigma_A, \sigma_D) \tag{83}$$

Rule 1: $(Z \perp\!\!\!\perp X, D, A)$ in $\mathcal{G}_{\sigma_A\sigma_D}$.

$$= \sum_{d,a} P(d \mid z) P(a \mid x; \sigma_A) \\ \sum_{z'} P(y \mid x, z', d, a; \sigma_D) P(z'; \sigma_A, \sigma_D) \tag{84}$$

Rule 3: $(Z \perp\!\!\!\perp A, D)$ in $\mathcal{G}_{\sigma_A\sigma_D\overline{A,D}}$ and $\mathcal{G}_{\overline{A,D}}$.

$$= \sum_{d,a} P(d \mid z) P(a \mid x; \sigma_A) \sum_{z'} P(y \mid x, z', d, a) P(z') \tag{85}$$

Motivation from a Classical Example

Example 5. Consider the the dynamic plan problem first studied in (Pearl and Robins, 1995) and shown in Fig. 6(a). The goal is to assess the distribution $P(\mathbf{y}; \sigma_\mathbf{X})$ in two hypothetical environments where:

C1 the values of (X_1, X_2) have been fixed to (x_1, x_2), in standard do-form, i.e., $\sigma_\mathbf{X} = \{\sigma_{X_1} = do(x_1), \sigma_{X_2} = do(x_2)\}$. The corresponding $\mathcal{G}_{\sigma_\mathbf{X}}$ is shown in Fig. 6(b).

C2 the value of X_1 is fixed to x_1 and X_2 is set conditionally on X_1, Z based on a function $g(x_1, z)$. That is, σ_{X_1} is as before and $\sigma_{X_2} = \delta(x_2, g(x_1, z))$. The corresponding $\mathcal{G}_{\sigma_\mathbf{X}}$ is shown in Fig. 6(c).

In (Pearl and Robins, 1995), the impact of the plan on the outcome variable $\mathbf{Y}$ is denoted as $P(\mathbf{y} \mid \hat{x_1}, \ldots, \hat{x_n})$, here it is written as $P(\mathbf{y}; \sigma_\mathbf{x})$. While the first scenario (C1) is identifiable by the rules of *do*-calculus (and σ-calculus), contrary to previous beliefs (Pearl, 2000, pp.120), and as hinted by (Tian, 2008), the plan (C2) is not identifiable from $P(\mathbf{v})$. The extra edge $Z \to X_2$ in Fig. 6(c) (C2) makes X_1 and Y dependent conditional on X_2, hence the same derivation strategy does not work.

We formally show that the effect of $\sigma_{\mathbf{X}}$ (C2) is not identifiable. $P(y;\sigma_{\mathbf{X}})$ can be written as

$$P(y;\sigma_{\mathbf{X}}) = \sum_{x_1,x_2,z} P(\mathbf{x};\sigma_{\mathbf{X}})Q[Z,Y]. \tag{86}$$

From (Huang and Valtorta, 2008) we have that $Q[Z,Y]$ is not identifiable from $P(x_1,x_2,z,y) = Q[\mathbf{V}]$. This means that there exists two models M_1 and M_2, compatible with $\mathcal{G}$, that induce the same $P(\mathbf{v})$ but for some $\mathbf{v}' = (x_1', x_2', z', y')$ we have $Q^1[Z,Y](\mathbf{v}') = a$, $Q^2[Z,Y](\mathbf{v}') = b$ with $a \neq b$. Assume, without loss of generality that $a > b$. Then consider the intervention $do(X_1 = x_1'), do(X_2 = g(x_1, z))$ with

$$g(x_1,z) = \begin{cases} x_2' & \text{, if } (x_1,z) = (x_1',z') \\ \text{other than } x_2' & \text{, if } (x_1,z) \neq (x_1',z') \end{cases}. \tag{87}$$

We will extend a strategy used by (Huang and Valtorta, 2008) to construct two models M_1' and M_2' where the domain of Y is $\mathcal{D}_Y \times \{0,1\}$, where $\mathcal{D}_Y$ is the domain of Y in M_1, M_2. Let $F(x_2)$ be a probability function from $\mathcal{D}_{X_2}$ to $\{0,1\}$, such that $P(F(x_2) = i) > 0, i = 0,1$ and $P(F(x_2) = 0) = 1 - P(F(x_2) = 1)$. In $M_i', i = 1,2$ we define:

$$P_i^{M_i'}((y,k)|x_2) = P^{M_i}(y|x_2)P(F(x_2) = k). \tag{88}$$

And for $V_j \in \{X_1, X_2, Z\}$ let $P^{M_i'}(v_j|pa_j) = P^{M_i}(v_j|pa_j)$. We can verify that

$$\begin{aligned} &P^{M_1'}(\mathbf{v}\setminus y,(y,k)) \\ &\quad = Q^{M_1'}[\mathbf{V}\setminus\{Y\},(Y,K)](\mathbf{v}\setminus y,(y,k)) && (89)\\ &\quad = Q^{M_1}[\mathbf{V}\setminus\{Y\},(Y,K)](\mathbf{v})P(F(x_2)=k) && (90)\\ &\quad = Q^{M_2}[\mathbf{V}\setminus\{Y\},(Y,K)](\mathbf{v})P(F(x_2)=k) && (91)\\ &\quad = Q^{M_2'}[\mathbf{V}\setminus\{Y\},(Y,K)](\mathbf{v}\setminus y,(y,k)) && (92)\\ &\quad = P^{M_2'}(\mathbf{v}\setminus y,(y,k)). && (93) \end{aligned}$$

Under intervention, we have that $P^{M_i^*}(y'), i{=}1{,}2$ is given by

$$P^{M_i^*}((y',0)) = \sum_{x_1,x_2,z} Q^*[X_1]Q^*[X_2]Q^{M_i'}[Z,Y,K]. \tag{94}$$

Note that $Q^*[X_1] = 1$ for $X_1 = x_1'$ and 0 otherwise. Similarly, $Q^*[X_2] = 1$ for $X_2 = g(x_1,z)$ and 0. Hence

$$\begin{aligned} &P^{M_i^*}((y',0)) \\ &= \sum_{x_2,z} Q^*[X_2](x_1',x_2,z)Q^{M_i'}[Z,Y,K](z,x_2,(y',0)). \end{aligned} \tag{95}$$

By construction we have

$$\begin{aligned} &Q^{M_i'}[Z,Y,K](z,x_2,(y',0)) \\ &\quad = Q^{M_i}[Z,Y,K](z,x_2,y')P(F(x_2)=0). \end{aligned} \tag{96}$$

Let $P(F(x_2')=0) = 1/2$ and $P(F(x_2)=0) = (a-b)/4$, for $x_2 \neq x_2'$. It yields:

$$\begin{aligned} P^{M_i^*}((y',0)) = &(\tfrac{1}{2})Q^{M_i'}[Z,Y](z',x_2',y') + \\ &\left(\tfrac{a-b}{4}\right) \sum_{\substack{x_2\neq x_2',\\ x_2=g(x_1',z)}} Q^{M_i'}[Z,Y](z,x_2,y'). \end{aligned} \tag{97}$$

For M_1^*:

$$\begin{aligned} P^{M_1^*}((y',0)) &= \tfrac{1}{2}a \left(\tfrac{a-b}{4}\right) \sum_{\substack{x_2\neq x_2',\\ x_2=g(x_1',z)}} Q^{M_i'}[Z,Y](z,x_2,y'). \\ &> \tfrac{1}{2}a. \end{aligned} \tag{98}$$

As for M_2^*:

$$\begin{aligned} P^{M_1^*}((y',0)) &= \tfrac{1}{2}b + \left(\tfrac{a-b}{4}\right) \sum_{\substack{x_2\neq x_2',\\ x_2=g(x_1',z)}} Q^{M_i'}[Z,Y](z,x_2,y') \\ &< \tfrac{1}{2}b + \tfrac{a-b}{4} < \tfrac{1}{2}a. \end{aligned} \tag{99}$$

Then, M_1', M_2' and M_1^*, M_2^* are compatible with $\mathcal{G}$ and $\mathcal{G}^*$, match in $P(\mathbf{v})$ and provide different plan effects.

This example suggests that the rules of *do*-calculus are not suitable to reason about interventions where the graphical structure of the intervened diagram is contingent to the type of intervention.

C Proofs for Section 5

Lemma 1. *Let $\mathbf{Y},\mathbf{X},\mathbf{W} \subset \mathbf{V}$ with $\mathbf{W}\cap\mathbf{Y} = \emptyset$ and let $\mathcal{G}$ be a causal diagram over the variables $\mathbf{V}$. The effect $P(\mathbf{y} \mid \mathbf{w};\sigma_{\mathbf{X}})$ is given by:*

$$P(\mathbf{y}|\mathbf{w};\sigma_{\mathbf{X}}) = \sum\nolimits_{\mathbf{a}\setminus(\mathbf{y}\cup\mathbf{w})} Q[\mathbf{A};\sigma_{\mathbf{X}}] \Big/ \sum\nolimits_{\mathbf{a}\setminus\mathbf{w}} Q[\mathbf{A};\sigma_{\mathbf{X}}]\ , \tag{23}$$

where $\mathbf{A}$ is the set of all variables connected to $\mathbf{Y}$ (including $\mathbf{Y}$) by any path (regardless of the directionality) in $\mathcal{G}_{\sigma_{\mathbf{X}}[\mathbf{D}]\underline{\mathbf{W}}}$, with $\mathbf{D} = An(\mathbf{Y}\cup\mathbf{W})_{\mathcal{G}_{\sigma_{\mathbf{X}}}}$.

Proof. As implied by eq. (22) the query only depends on variables in $\mathbf{D}$ hence we can focus on the subgraph $\mathcal{G}_{\sigma_{\mathbf{X}}[\mathbf{D}]}$. The set $\mathbf{A}$ is defined in terms of $\mathcal{G}_{\sigma_{\mathbf{X}}[\mathbf{D}]\underline{\mathbf{W}}}$ which is equal to $\mathcal{G}_{\sigma_{\mathbf{X}}[\mathbf{D}]}$ except for the absence of edges outgoing from $\mathbf{W}$, but have the same set of bidirected edges. Let $\mathbf{A}' = \mathbf{D}\setminus\mathbf{A}$, then the definition of $\mathbf{A}$ and implies that there are no bidirected arrows crossing between $\mathbf{A}$ and $\mathbf{A}'$ in $\mathcal{G}_{\sigma_{\mathbf{X}}}$. To witness suppose there exists $A \in \mathbf{A}$, $A' \in \mathbf{A}'$ with a bidirected edge between them. We have that some $Y \in \mathbf{Y}$ is connected by some path to A, hence also connected to $\mathbf{A}'$, but then A' has to be in $\mathbf{A}$, a contradiction.

This mean that $\mathbf{A}$ and $\mathbf{A}'$ share no c-component and by (Tian and Pearl, 2002b, Lemma 2) we have that $Q[\mathbf{D};\sigma_{\mathbf{X}}] = Q[\mathbf{A};\sigma_{\mathbf{X}}]Q[\mathbf{A}';\sigma_{\mathbf{X}}]$. From Eq. (21) we can see that $Q[\mathbf{C};\sigma_{\mathbf{X}}]$ is a function of $Pa(\mathbf{C})_{\mathcal{G}_{\sigma_{\mathbf{X}}}}$, which is defined as the set of variables in $\mathbf{C}$ and their observable parents in $\mathcal{G}_{\sigma_{\mathbf{X}}}$ (all unobservable parents affecting variables in $\mathbf{C}$ are being summed out). We want to claim that by definition of $\mathbf{A}$ the sets $Pa(\mathbf{A})_{\mathcal{G}_{\sigma_{\mathbf{X}}}}$ and $Pa(\mathbf{A}')_{\mathcal{G}_{\sigma_{\mathbf{X}}}}$ may intersect only on $\mathbf{W}$. Suppose this is not the case, then there exists variables $A \in \mathbf{A}$ and $A' \in \mathbf{A}'$ sharing a parent $B \notin \mathbf{W}$, but this implies that some $Y \in \mathbf{Y}$ that is connected to A by definition is also connected to A' by extension $A \leftarrow B \rightarrow A'$ in $\mathcal{G}_{\sigma_{\mathbf{X}}\underline{\mathbf{W}}}$, a contradiction. Then let $\mathbf{H} = Pa(\mathbf{A})_{\mathcal{G}_{\sigma_{\mathbf{X}}}} \setminus \mathbf{W}$ and $\mathbf{H}' = Pa(\mathbf{A}')_{\mathcal{G}_{\sigma_{\mathbf{X}}}} \setminus \mathbf{W}$ and note that $\mathbf{H} \cup \mathbf{H}'$ must be equal to $\mathbf{D} \setminus \mathbf{W}$.

Consider the denominator in Eq. (22):

$$\sum_{\mathbf{d}\setminus\mathbf{w}} Q[\mathbf{D};\sigma_{\mathbf{X}}] = \sum_{\mathbf{h},\mathbf{h}'} Q[\mathbf{A};\sigma_{\mathbf{X}}]Q[\mathbf{A}';\sigma_{\mathbf{X}}] \tag{100}$$

$$= \left(\sum_{\mathbf{h}'} Q[\mathbf{A}';\sigma_{\mathbf{X}}]\right)\left(\sum_{\mathbf{h}} Q[\mathbf{A};\sigma_{\mathbf{X}}]\right), \tag{101}$$

where the last equality follows from the fact that $Q[\mathbf{A}';\sigma_{\mathbf{X}}]$ is not a function of any variable in $\mathbf{H}$ while the same is true for $Q[\mathbf{A};\sigma_{\mathbf{X}}]$ and $\mathbf{H}'$. Since $\mathbf{Y}$ is contained in $\mathbf{A}$ by definition, it follows that

$$\sum_{\mathbf{d}\setminus(\mathbf{y}\cup\mathbf{w})} Q[\mathbf{D};\sigma_{\mathbf{X}}] = \left(\sum_{\mathbf{h}'} Q[\mathbf{A}';\sigma_{\mathbf{X}}]\right)\left(\sum_{\mathbf{h}\setminus\mathbf{y}} Q[\mathbf{A};\sigma_{\mathbf{X}}]\right). \tag{102}$$

Dividing Eq. (102) by (101) we get an expression equivalent to Eq. (22) where the factors $(\sum_{\mathbf{h}'} Q[\mathbf{A}';\sigma_{\mathbf{X}}])$ appearing in numerator and denominator cancel out, and we are left with

$$P(\mathbf{y} \mid \mathbf{w};\sigma_{\mathbf{X}}) = \frac{\sum_{\mathbf{h}\setminus\mathbf{y}} Q[\mathbf{A};\sigma_{\mathbf{X}}]}{\sum_{\mathbf{h}} Q[\mathbf{A};\sigma_{\mathbf{X}}]}. \tag{103}$$

Next, we argue that $\mathbf{H} = \mathbf{A} \setminus \mathbf{W}$. First it is easy to see that $(\mathbf{A} \setminus \mathbf{W}) \subseteq \mathbf{H}$, for the other side suppose for the sake of contradiction that there exists a variable $B \in \mathbf{H} \setminus (\mathbf{A} \setminus \mathbf{W})$. Then B is a parent of some variable in $A \in \mathbf{A}$ and is not in $\mathbf{W}$, but this mean that B is connected to A in $\mathcal{G}_{\sigma_{\mathbf{X}}\underline{\mathbf{W}}}$ and has to be in $\mathbf{A}$, a contradiction. Therefore, we can further rewrite Eq. (103) as

$$P(\mathbf{y} \mid \mathbf{w};\sigma_{\mathbf{X}}) = \frac{\sum_{\mathbf{a}\setminus(\mathbf{y}\cup\mathbf{w})} Q[\mathbf{A};\sigma_{\mathbf{X}}]}{\sum_{\mathbf{a}\setminus\mathbf{w}} Q[\mathbf{A};\sigma_{\mathbf{X}}]}. \tag{104}$$

□

Lemma 2. *Let $\mathbf{A}$ be defined relative to $\mathbf{X}, \mathbf{Y}, \mathbf{W}$ as in lemma 1, then:*

$$Q[\mathbf{A};\sigma_{\mathbf{X}}] = Q\left[\mathbf{A}^{\mathbf{X}};\sigma_{\mathbf{X}}\right] Q\left[\mathbf{A} \setminus \mathbf{A}^{\mathbf{X}}\right], \tag{24}$$

where $\mathbf{A}^{\mathbf{X}}$ is the union of the c-components of $\mathcal{G}_{\sigma_{\mathbf{X}}[\mathbf{A}]}$ containing variables in $\mathbf{X}$.

Proof. This follows immediately from (Tian and Pearl, 2002b, Lemma 2) since by definition $\mathbf{A}^{\mathbf{X}}$ do not share any c-component with $\mathbf{A}$. □

Lemma 3. *Let $\sigma_{\mathbf{Z}}$ indicate any intervention on $\mathbf{Z}$ and let $\mathbf{C} \subseteq \mathbf{V}$. Then, $Q[\mathbf{C}] = Q[\mathbf{C};\sigma_{\mathbf{Z}}]$ if $\mathbf{C} \cap \mathbf{Z} = \emptyset$.*

Proof. By definition of $Q[\mathbf{C};\sigma_{\mathbf{Z}}]$ (Eq. (21)) we have

$$Q[\mathbf{C};\sigma_{\mathbf{Z}}](\mathbf{v}) = \sum\nolimits_{u(\mathbf{C})} \prod\nolimits_{\{i|V_i\in\mathbf{C}\}} P(v_i \mid pa_i, u_i;\sigma_{\mathbf{Z}})P(u(\mathbf{C});\sigma_{\mathbf{Z}}), \tag{105}$$

since no variable in $\mathbf{C}$ is in $\mathbf{Z}$ every term $P(v_i \mid pa_i, u_i;\sigma_{\mathbf{Z}}) = P(v_i \mid pa_i, u_i)$ and interventions cannot affect the distribution of variables in $\mathbf{U}$, hence $P(u(\mathbf{C});\sigma_{\mathbf{Z}}) = P(u(\mathbf{C}))$. Replacing those terms, we obtain exactly the definition of $Q[\mathbf{C}](\mathbf{v})$. □

Lemma 4. *Suppose $Q[\mathbf{A};\sigma_{\mathbf{X}}]$ is not identifiable from a set of available distributions in a causal diagram $\mathcal{G}$. Let $A_1, A_2 \in \mathbf{A}$ such that there exists an edge $A_1 \rightarrow A_2$ in $\mathcal{G}$. Then $\sum_{a_1} Q[\mathbf{A}]$ is not identifiable from the same input either.*

Proof. Let M_1 and M_2 be the two models witnessing the non-identifiability of $Q[\mathbf{A};\sigma_{\mathbf{X}}]$, they agree on available distributions, but for some value-assignment $\mathbf{v}'$ we have $Q^1[\mathbf{A};\sigma_{\mathbf{X}}](\mathbf{v}') = \alpha$, $Q^2[\mathbf{A};\sigma_{\mathbf{X}}](\mathbf{v}') = \beta$ with $\alpha \neq \beta$. Assume, without loss of generality that $\alpha > \beta$. We will extend a strategy used by (Huang and Valtorta, 2008) to construct two models M_1' and M_2' where the domain of A_2 is $\mathcal{D}_{A_2} \times \{0, 1\}$, where $\mathcal{D}_{A_2}$ is the domain of A_2 in M_1, M_2. Let $F(A_1)$ be a probability function from $\mathcal{D}_{A_1}$ to $\{0, 1\}$, such that $P(F(a_1) = i) > 0, i = 0, 1$ and $P(F(a_1) = 0) = 1 - P(F(a_1) = 1)$. In $M_i', i = 1, 2$ we define:

$$P_i^{M_i'}((a_2, k)|pa_{a_2}, u_{a_2}) = P^{M_i}(a_2|pa_{a_2}, u_{a_2})P(F(a_1)=k). \tag{106}$$

And for $V_j \in \mathbf{V} \setminus \{A_2\}$ let $P^{M_i'}(v_j|pa_j, u_j) = P^{M_i}(v_j|pa_j, u_j)$. We can verify that for any $\sigma_{\mathbf{Z}_j} \in \mathbb{Z}$

$$P^{M_1'}(\mathbf{v} \setminus a_2, (a_2, k);\sigma_{\mathbf{Z}_j})$$

$$= Q^{M_1'}[\mathbf{V} \setminus \{A_2\}, (A_2, K);\sigma_{\mathbf{Z}_j}](\mathbf{v} \setminus a_2, (a_2, k)) \tag{107}$$

$$= Q^{M_1}[\mathbf{V} \setminus \{A_2\}, (A_2, K);\sigma_{\mathbf{Z}_j}](\mathbf{v})P(F(a_1) = k) \tag{108}$$

$$= Q^{M_2}[\mathbf{V} \setminus \{A_2\}, (A_2, K);\sigma_{\mathbf{Z}_j}](\mathbf{v})P(F(a_1) = k) \tag{109}$$

$$= Q^{M_2'}[\mathbf{V} \setminus \{A_2\}, (A_2, K);\sigma_{\mathbf{Z}_j}](\mathbf{v} \setminus a_2, (a_2, k)) \tag{110}$$

$$= P^{M_2'}(\mathbf{v} \setminus a_2, (a_2, k);\sigma_{\mathbf{Z}_j}). \tag{111}$$

Consider the assignment $\mathbf{v}' \setminus \{a_2\}, (a_2', 0)$, by construction we have

$$Q^{M_i'}[\mathbf{A};\sigma_{\mathbf{X}}](\mathbf{v}' \setminus \{a_2\}, (a_2', 0)) = Q^{M_i}[\mathbf{A};\sigma_{\mathbf{X}}](\mathbf{v}')P(F(a_1') = 0). \tag{112}$$

Let $P(F(a'_1) = 0) = 1/2$ and $P(F(a_1) = 0) = (\alpha - \beta)/4$, for $a_1 \neq a'_1$. It yields:

$$\sum_{a_1} Q^{M'_i}[\mathbf{A}; \sigma_\mathbf{X}](\mathbf{v}' \setminus \{a_1, a_2\}, (a'_2, 0))$$
$$= \sum_{a_1} Q^{M_i}[\mathbf{A}; \sigma_\mathbf{X}](\mathbf{v}' \setminus \{a_1, a_2\}) P(F(a_1) = 0) \quad (113)$$

For M'_1 this means

$$\sum_{a_1} Q^{M'_1}[\mathbf{A}; \sigma_\mathbf{X}](\mathbf{v}' \setminus \{a_1, a_2\}, (a'_2, 0))$$
$$= \tfrac{1}{2}\alpha + \left(\tfrac{\alpha-\beta}{4}\right) \sum_{a_1 \neq a'_1} Q^{M_i}[\mathbf{A}; \sigma_\mathbf{X}](\mathbf{v}' \setminus \{a_1, a_2\}) \quad (114)$$
$$> \tfrac{1}{2}\alpha \quad (115)$$

As for M'_2:

$$\sum_{a_1} Q^{M'_1}[\mathbf{A}; \sigma_\mathbf{X}](\mathbf{v}' \setminus \{a_1, a_2\}, (a'_2, 0))$$
$$= \tfrac{1}{2}b + \left(\tfrac{a-b}{4}\right) \sum_{a_1 \neq a'_1} Q^{M_i}[\mathbf{A}; \sigma_\mathbf{X}](\mathbf{v}' \setminus \{a_1, a_2\}) \quad (116)$$
$$< \tfrac{1}{2}\beta + \tfrac{\alpha-\beta}{4} \quad (117)$$
$$< \tfrac{1}{2}\alpha. \quad (118)$$

Then, M'_1 and M'_2 are compatible with $\mathcal{G}$, match in the available distributions and yield different $\sum_{a_1} Q[\mathbf{A}; \sigma_\mathbf{X}]$. □

Theorem 2. *The effect $P(\mathbf{y} \mid \mathbf{w}; \sigma_\mathbf{X})$ is identifiable if σ-IDENTIFY (Alg. 1) does not fail. Moreover, the expression returned is a valid estimand for the effect.*

Proof. (if) σ-IDENTIFY starts by computing the set $\mathbf{A}$ as defined in lemma 1, which define the c-factor $Q[\mathbf{A}; \sigma_\mathbf{X}]$ that determines the target query.

Next, it factorizes $Q[\mathbf{A}; \sigma_\mathbf{X}]$ according to the c-components of $\mathcal{G}_{\sigma_\mathbf{X}[\mathbf{A}]}$ and tries to identify each one of them individually from any of the available distributions whenever $\mathbf{A}_i \cap \mathbf{X} = \emptyset$. Given $P(\mathbf{V}; \sigma_\mathbf{Z})$, if $\mathbf{A}_i \cap \mathbf{Z} = \emptyset$, then lemma 3 guarantees that $Q[\mathbf{A}_i; \sigma_\mathbf{Z}] = Q[\mathbf{A}_i]$ and by the same lemma $Q[\mathbf{A}_i] = Q[\mathbf{A}_i; \sigma_\mathbf{X}]$. To do this the algorithm runs IDENTIFY with the graph and distribution that correspond to $\sigma_\mathbf{Z}$ and stores the result when it succeeds.

For those $\mathbf{A}_i$ that contain variables in $\mathbf{X}$, the algorithm obtains them by using the replacement strategy corresponding to the type of intervention (see discussion in the main text).

Finally, by lemma 2 and lemma 2 in (Tian and Pearl, 2002b), it follows that as long as every $Q[\mathbf{A}_i]$ was identified, their product is equal to $Q[\mathbf{A}; \sigma_\mathbf{X}]$. Finally, the algorithm employs equation 23 to return a correct estimand for $P(\mathbf{y} \mid \mathbf{w}; \sigma_\mathbf{X})$.

□

Complexity Analysis of σ-calculus

Let $n = |\mathbf{V}|$ and $z = |\mathbb{Z}|$. Operations in σ-IDENTIFY such as compute the set of ancestors and find the set of C-components in a graph can be done in $O(n^2)$ time. The number of C-components is at most n, hence the total number of times the for-loops in the algorithm could execute and call *Identify* is nz. IDENTIFY (see (Tian and Pearl, 2002a; Huang and Valtorta, 2006)) recursively reduces the input c-factor at least by a variables each time, and the operations used can be performed in $O(n^2)$; overall it takes $O(n^3)$ to return an expression of FAIL. Consequently, σ-IDENTIFY runs in $O(n^4 z)$.