Path-specific Effects Based on Information Accounts of Causality

Gong Heyang
Department of Statistics and Finance
University of Science and Technology of China
zj3712@gmail.com

Zhu Ke
Department of Statistics and Actuarial Science
University of HongKong
mazhuke@hku.hk

Abstract

Path-specific effects in mediation analysis provide a useful tool for fairness analysis, which is mostly based on nested counterfactuals. However, the dictum "no causation without manipulation" implies that path-specific effects might be induced by certain interventions. This paper proposes a new path intervention inspired by information accounts of causality, and develops the corresponding intervention diagrams and π-formula. Compared with the interventionist approach of Robins et al. (2020) based on nested counterfactuals, our proposed path intervention method explicitly describes the manipulation in structural causal model with a simple information transferring interpretation, and does not require the non-existence of recanting witness to identify path-specific effects. Hence, it could serve useful communications and theoretical focuses for mediation analysis.

1 Introduction

Analyzing the relative strength of different pathways between a decision X and an outcome Y is an interesting topic for both scientists and practitioners across many disciplines over several decades. The path-specific effects in mediation analysis have been demonstrated as a useful tool to tackle this topic (Baron & Kenny, 1986; Robins & Greenland, 1992; Pearl, 2001). This is one of the seven tasks summarized by Judea Pearl in causal inferences (Pearl, 2019b). To capture the path-specific effects, the often used approaches are based on nested counterfactuals; see Avin et al. (2005), Shpitser & Tchetgen (2016), Zhang & Bareinboim (2018), and Malinsky et al. (2019) for earlier references.

More recently, Robins et al. (2020) provided an interventionist theory of mediation to study some path-specific effects (such as natural indirect/direct effects) in single world intervention graph (SWIG). However, their interventionist framework seems lack of a formally defined concept of intervention for the general path-specific effects, and it requires the recursive assumption to well define the nested counterfactuals. Also, their path specific counterfactuals are equal to interventional counterfactuals only when no recanting witness exists. Therefore, their method are restrictive in many situations. For example, to analyze the strength of path $\pi : X \to A \to Y$ between a decision X and an outcome Y in Fig. 1(a) or (b), their method is not applicable any more.

In this paper, we propose a path intervention based on the information accounts of causality. To understand causality as information transfer, which has been formally proposed and developed in philosophy since Collier (1999), we decompose causal mechanisms into two parts: information

Figure 1: (a) A causal diagram with a recanting witness A for the path $\pi : X \to A \to Y$; (b) A cyclic causal diagram with a loop between A and Y.

transfer mechanisms and information process mechanisms. Specifically, each causal mechanism f_i in a structural causal model (**SCM**) decides how to process the collected information $e_{pa(i),i}$ from its input edges, where each $e_{j,i}$ decides the information transferred from its input node j through the edge $j \to i$. Our novel path intervention is formally defined by explicitly manipulating mechanisms in the informational decomposition of **SCM**. Henceforth, the path specific effects can be well defined by using our path intervention, even for cyclic causal diagrams such as that in Fig. 1(b),

Moreover, we develop a π-formula for identifying our interventional counterfactuals. The use of the π-formula does not need the absence of recanting witness, and hence it can be applied to identify the path-specific effects in Fig. 1(a). As a graphical tool to assist causal reasoning, we also propose the associated intervention diagram for different interventions under the informational decomposition of **SCM**. Our path-specific effects can be interpreted as simply making the decision information $X = x'$ exclusively pass through the causal path π while allowing all other information outside π to behave naturally. This interpretation is different from those in the existing graphical approaches based on nested counterfactuals.

The remaining paper is organized as follows. Section 2 gives the preliminaries and the formal definition of informational decomposition of **SCM**. Section 3 introduces our path intervention and intervention diagrams with some illustrating examples. Section 4 proposes the identification formula for path-specific effects by using our path intervention. Concluding remarks and discussions are offered in Section 5.

Notations. Our notations below are similar to those in Shpitser et al. (2020). Fix a set of indices $V \equiv \{1, ..., n\}$. For each index $i \in V$, associate a random variable X_i with state space $\mathcal{X}_i$. Given $A \subseteq V$, we will denote subsets of random variables indexed by A as $X_A \in \mathcal{X}_A \equiv \prod_{i \in A} \mathcal{X}_i$. For notational conciseness, we will sometimes use index sets to denote random variables themselves, that is, using V and A to denote X_V and X_A, respectively, and similarly using lower case x_A to denote $x_A \in \mathcal{X}_A$. Similarly, by extension, we will also use V_A to denote X_A and V_i to denote X_i. Finally, M is usually used as a mediator, and $\mathcal{M}$ is an **SCM**.

2   Preliminaries and Informational Decomposition of SCM

The centre of modern causal modeling theory lies the structural causal model (SCM) (also known as structural equation model) that makes graphical assumptions of the underlying data generating process.[1] There are many different formulations of SCM in the literature (see, e.g., Bongers et al. (2016); Pearl et al. (2009); Pearl (2019a); Schölkopf (2019)), among which we use the definition in Bareinboim et al. (2020) in this paper.

Definition 1 (SCM) *A structural causal model (SCM) $\mathcal{M}$ is a 4-tuple $(\mathbf{U}, \mathbf{V}, \mathcal{F}, P(\mathbf{U}))$, where*

- **U** *is a set of background variables, also called exogenous variables, that are determined by factors outside the model;*

- **V** *is a set $\{V_1, V_2, ..., V_n\}$ of variables, called endogenous, that are determined by other variables in the models- that is, variables in $\mathbf{U} \cup \mathbf{V}$;*

---

[1]Another popular causal modeling framework is the potential-outcome framework, and one can refer to Imbens (2020) for more details.

- $\mathcal{F}$ *is a set of functions* $\{f_1, f_2, ..., f_n\}$ *such that each* f_i *is a mapping from(the respective domains of)* $U_i \cup Pa_i$ *to* V_i*, where* $U_i \subseteq \mathbf{U}, Pa_i \subseteq \mathbf{V} \setminus V_i$*, and the entire set* $\mathcal{F}$ *forms a mapping from* $\mathbf{U}$ *to* $\mathbf{V}$*. That is, for* $i \in V = \{1, ..., n\}$*, each* $f_i \in \mathcal{F}$ *is such that*

$$v_i \leftarrow f_i(pa_i, u_i), \tag{1}$$

- $P(\mathbf{U})$ *is a probability function defined over the domain of* $\mathbf{U}$*.*

Mechanisms are activities and entities organized to produce some phenomena which are represented by structural equations in an **SCM**. A structural model is Markovian if the exogenous parent set U_i, U_j are independent whenever $i \neq j$, and the associated causal graph for a Markovian SCM $\mathcal{M} = (\mathbf{U}, \mathbf{V}, \mathcal{F}, P(\mathbf{U}))$ can be defined as follows:

Definition 2 (Causal Diagrams) $\mathcal{G}(\mathcal{M}) = (V, E)$ *is said to be a* causal diagram *(of* $\mathcal{M}$*) if constructed as follows:*

- *add a vertex for every endogenous variable in the set* V*,*
- *add an edge* $(V_j \rightarrow V_i)$ *for every* $i \in V$ *if* V_j *appears as an argument of* $f_i \in \mathcal{F}$ *in the edge set* E*.*

Various difference-making and production accounts in philosophy [2] have been taken account by Judea Pearl for developing the structural causal modeling framework. However, the information accounts of causality, one of major approaches to study causality, has been comparatively neglected. The view of causality as information transfer was first proposed in Collier (1999), further developed by Illari (2011) and Illari & Russo (2014), and recently applied in neuroscience by Feltz et al. (2019). Inspired by this view of causality, we propose the informational decomposition of **SCM** in the following:

Definition 3 (Informational Decomposition of SCM) *Consider an* **SCM** $\mathcal{M} = (\mathbf{U}, \mathbf{V}, \mathcal{F}, P(\mathbf{U}))$. *The* informational decomposition *of* $\mathcal{M}$ *is defined as, for any* $i \in V$,

$$\begin{cases} v_i \leftarrow f_i(e_{pa(i),i}, u_i), \\ e_{j,i} \leftarrow v_j, \end{cases} \tag{2}$$

where $e_{j,i}$ *represents the information on edge* (j, i) *received from its input node* V_j*.* [3]

Essentially, it can be interpreted as separating every causal mechanism f_i into two part — information process and information transfer. Mechanisms could be seen as information channels which is one centre idea of the informational accounts(Collier, 1999; Illari & Russo, 2014). Under this information transferring interpretation of **SCM**, both *do* intervention(Pearl, 1995) and *info* intervention (Gong & Zhu, 2020) are defined as following:

Definition 4 (*Do* and *info* intervention) *Given an SCM* $\mathcal{M} = (\mathbf{U}, \mathbf{V}, \mathcal{F}, P(\mathbf{U}))$ *and any* $I \subseteq V, v'_I \in \mathcal{V}_I$*, the the* do intervention $do(v'_I)$ *can be defined as a modification of structural equations to*

$$\begin{cases} v_i \leftarrow f_i(e_{pa(i),i}, u_i) \, if \, i \notin I \, else \, v'_i \\ e_{j,i} \leftarrow v_j \end{cases} \tag{3}$$

while keeping everything else constant. And info intervention $\sigma(V_I = v'_I)$ *(or, in short,* $\sigma(v'_I)$*) maps* $\mathcal{M}$ *to the info-intervention model* $\mathcal{M}^{\sigma(v'_I)}$ *with a modified mechanism for every* $i \in V$*:*

$$\begin{cases} v_i \leftarrow f_i(e_{pa(i),i}, u_i) \\ e_{j,i} \leftarrow v_j \, if \, j \notin I \, else \, v'_j \end{cases} \tag{4}$$

while keeping everything else constant.

---

[2] See more details on Illari & Russo (2014).

[3] pa_i usually stands for functional parents of node i in an **SCM**, while $pa(i)$ stands for parents in a causal diagram.

The informational decomposition of **SCM** reflect important epistemological distinctions about causality and lead to an alternative approach for mediation analysis. Comparing to the *do* intervention, the *info* intervention changes the information transferring instead of information processing mechanisms, which facilitates communication and theoretical focus as complementing causal modeling in terms of *do* operator. This motivates us to proposed a novel path intervention as modification of information transfer mechanisms. In the following section, we propose a path intervention for path-specific effect based on the above informational decomposition.

3 Path Intervention and Intervention Diagrams

Causal questions, such as what if we make something happen, can be formalized by using *do* intervention. However, the *do* intervention makes the intervention variable deterministic, by hypothetically forcing its value to a given constant. Henceforth, this causal notation is powerful but sometimes not enough for certain causal tasks, e.g., nature directed effects(NDE) defined by nested counterfactuals in mediation analysis. The dictum "no causation without manipulation" calls for an interventionist approach, recently addressed by Robins et al. (2020) based on single world intervention graphs(SWIG) with certain limitation. Here, we are giving a explicit definition of path intervention based on informational decomposition of **SCM**.

Definition 5 (Path intervention) *Consider a Makovian SCM $\mathcal{M} = (\mathbf{U}, \mathbf{V}, \mathcal{F}, P(\mathbf{U}))$. Then, the path intervention $\pi(A = a')$ (or, in short, $\pi(a')$) along a causal path π from A to Y maps $\mathcal{M}$ to the path-intervention model $\mathcal{M}^{\pi(a')}$ with a modified causal mechanism for every $i \in V$:*

$$
\begin{cases}
v_i \leftarrow f_i(e_{pa(i),i}, u_i), \\
e_{j,i} \leftarrow v_j, \\
v_i^{\pi} \leftarrow f_i(e'_{pa(i),i}, u'_i), \\
e'_{j,i} \leftarrow v_j \; if \, (j,i) \notin \pi; \; else \; if \, V_j \; is \; A \, , a'; \; else \; v_j^{\pi},
\end{cases}
\tag{5}
$$

while keep everything else constant. Here, v_i^{π} (standing for $v_i^{\pi(a')}$) and v_i are in corresponding counterfactual and factual domains, respectively, u'_i is a realization of U'_i, and U'_i is the ith entry of $\mathbf{U}'$ which is an independent and identically distributed (i.i.d.) copy of $\mathbf{U}$.

In $\mathcal{M}^{\pi(a')}$, we have a set of endogenous variables $\mathbf{V} \cup \mathbf{V}^{\pi(a')}$, where $\mathbf{V}^{\pi(a')} = \{V_1^{\pi(a')}, V_2^{\pi(a')}, ..., V_n^{\pi(a')}\}$, and all $V_i^{\pi(a')}$ (or, in short, V_i^{π}) are referred as the effect variables of this path intervention $\pi(a')$. Hence, the path intervention in Def. 4 considers both factual and counterfactual worlds to capture the cross-world nature of path-specific effects. Moreover, since the information $A = a'$ is only transferred to the descendants of A along the path π in $\mathcal{M}^{\pi(a')}$, it is easy to see that the effect variables $V_i^{\pi} \overset{d}{=} V_i$ for $(j,i) \notin \pi$. In view of this fact, we marginalize out the effect variables V_i^{π} for $(j,i) \notin \pi$ in $\mathcal{M}^{\pi(a')}$. Intuitively, our path intervention creates the counterfactual variables V^{π} to transfer the information $A = a'$ exclusively through the causal path π while allowing the information outside π to remain unchanged.

We use use the NIE(natural indirect effect) as a toy example to show how our path intervention related to existing literatures.

Example 1 (Natural indirect effect) *For an SCM with three endogenous variables, including treatment A, mediator M, Outcome Y, satisfies*

$$
\begin{cases}
a \leftarrow f_A(u_A) \\
m \leftarrow f_M(a, u_M) \\
y \leftarrow f_Y(a, m, u_Y)
\end{cases}
\tag{6}
$$

The natural indirect effect(NIE) of Y on A through the mediator M depends on $Y(a, M(a'))$ — a variable in which two different levels of a are nested within the counterfactual for Y.

4

For a causal path $\pi : A \to M \to Y$ with transferred information $A = a'$, the path intervention results in a set of modified equations

$$
\begin{cases}
a \leftarrow f_A(u_A) \\
m \leftarrow f_M(a, u_M) \\
y \leftarrow f_Y(a, m, u_Y) \\
a^\pi \leftarrow f_A(u'_A) \\
m^\pi \leftarrow f_M(e'_{AM}, u'_M); e'_{AM} \leftarrow a' \\
y^\pi \leftarrow f_Y(e'_{AY}, e'_{MY}, u'_Y); e'_{AY} \leftarrow a, e'_{MY} \leftarrow m^\pi
\end{cases}
\tag{7}
$$

Since the causal mechanisms of a and a^π coincide with the same input distribution information, the distributions of corresponding variables are the same. Henceforth, we marginalize A^π out the model, results in a causal model of variables of (A, M, Y, M^π, Y^π) with mechanisms:

$$
\begin{cases}
a \leftarrow f_A(u_A) \\
m \leftarrow f_M(a, u_M) \\
y \leftarrow f_Y(a, m, u_Y) \\
m^\pi \leftarrow f_M(e'_{AM}, u'_M); e'_{AM} \leftarrow a' \\
y^\pi \leftarrow f_Y(e'_{AY}, e'_{MY}, u'_Y); e'_{AY} \leftarrow a, e'_{MY} \leftarrow m^\pi
\end{cases}
\tag{8}
$$

We can derive a set of structural equations ruling out $e'_{..}$:

$$
\begin{cases}
a \leftarrow f_A(u_A) \\
m \leftarrow f_M(a, u_M) \\
y \leftarrow f_Y(a, m, u_Y) \\
m^\pi \leftarrow f_M(a', u'_M) \\
y^\pi \leftarrow f_Y(a, m^\pi, u'_Y)
\end{cases}
\tag{9}
$$

In the above example, the path specific effect variable Y^π are different from $Y(a, M(a'))$ in two aspects. First, NIE usually considers the contrast of outcome under two given different levels a, a' of the treatment, then mathematically $Y(a, M(a')) \neq Y(A, M(a')) = Y^{\pi(a')}$. Second, the interpretation is completely different for Y^π, which is the effect of Y had the information $A = a'$ passed through the path $\pi : A \to M \to Y$ while keeping other transferred information unchanged. Beyond this example, our path specific effects can even be well-defined for causal diagram with cycle, e.g. Fig 1(b), by the solution of an **SCM** with modified structural equations (5).

The recognition that there are mechanisms underlying the phenomena of interest, but we usually cannot determine them precisely, gives rise to the discipline of causal inference Pearl et al. (2000). Virtually every approach to causal inference works under the stringent condition that only partial knowledge of the underlying SCM is available, and usually the information of structural dependency(i.e. causal diagrams) might be sufficient for many causal effect estimation problems. Particularly, *do*-calculus is available for identification of causal queries when the causal structure is known, while its graphical criteria are placed on graphs which obtained by removing input/output edges on intervened variables from the original causal graph. Since every hypothetical intervention is creating a counterfactual world for domain variables, then we can define a causal graph corresponding to the modified mechanisms. We here introduce a novel graphical tool *intervention diagram* to further illustrate how the path intervention works.

Definition 6 (Intervention diagram) *For an intervention, which can be* do, info *or* path *intervention, the causal diagram for the modified mechanisms induced by the intervention is called the* intervention diagram (or graph).[4]

To emphasis the informational accounts of causality and differences from previous causal graphs, our intervention diagrams, which is created based on the modified mechanisms by an intervention on

---

[4]Note that intervention diagrams are not restricted to path intervention. For some cases, we may also want to include exogenous variables to represent the causal mechanisms more detailed graphically, we call the causal graph of the intervention model with both exogenous and endogenous variables the *augmented intervention diagram*.

(a) An SCM $\mathcal{M}$.

(b) $\mathcal{M}^{do(t')}$.

(c) $\mathcal{M}^{\sigma(t')}$.
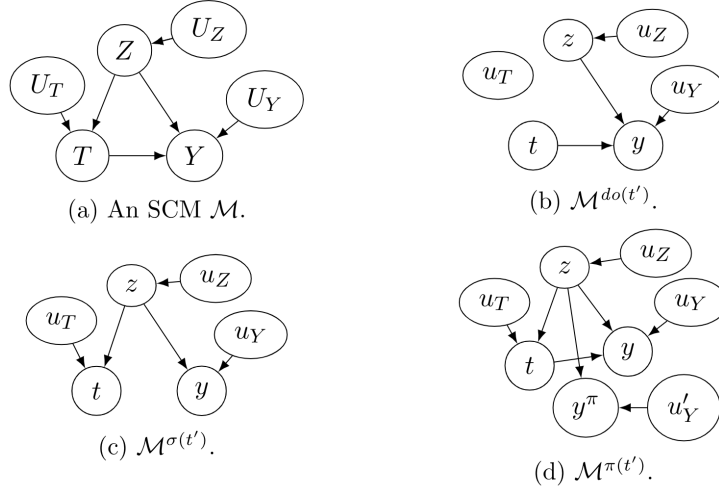
(d) $\mathcal{M}^{\pi(t')}$.

Figure 2: Causal graphs of an SCM and three augmented intervention diagrams of its intervention SCMs.

SCM, will use lowercase letter to represent variables. The following toy example is used to present the intervention diagrams for different interventions.

Example 2 *An SCM $\mathcal{M}$ with endogenous variables treatment T, outcome Y, confounder Z and theire corresponding exogenous variables U_T, U_Z, U_Y and an element t' in the domain of T, and its causal mechanisms represented by structural equations(the corresponding augmented causal graph is Fig 2(a)):*

$$\begin{cases} z \leftarrow f_Z(u_Z), \\ t \leftarrow f_T(z, u_T), \\ y \leftarrow f_Y(t, z, u_Y). \end{cases}$$

Its do-intervention SCM $\mathcal{M}^{do(t')}$ with augmented intervention diagram Fig 2(b) has the following modified structural equations:

$$\begin{cases} z \leftarrow f_Z(u_Z), \\ t \leftarrow t', \\ y \leftarrow f_Y(e_{TY}, e_{ZY}, u_Y); e_{TY} \leftarrow t, e_{ZY} \leftarrow z. \end{cases}$$

Its info-intervention SCM $\mathcal{M}^{\sigma(t')}$ with the augmented intervention diagram Fig 2(c) has the following modified structural equations:

$$\begin{cases} z \leftarrow f_Z(u_Z), \\ t \leftarrow f_T(e_{ZT}, u_T); e_{ZT} \leftarrow z, \\ y \leftarrow f_Y(e_{TY}, e_{ZY}, u_Y); e_{TY} \leftarrow t', e_{ZY} \leftarrow z. \end{cases}$$

Its path-intervention SCM $\mathcal{M}^{\pi(t')}$(for causal path $\pi : T \rightarrow Y$ with input information $T = t'$) with augmented intervention diagram Fig 2(d) [5] *has the following modified structural equations:*

$$\begin{cases} z \leftarrow f_Z(u_Z), \\ t \leftarrow f_T(e_{ZT}, u_T); e_{ZT} \leftarrow z, \\ y \leftarrow f_Y(e_{TY}, e_{ZY}, u_Y); e_{TY} \leftarrow t, e_{ZY} \leftarrow z, \\ y^\pi \leftarrow f_Y(e'_{TY}, e'_{ZY}, u'_Y); e'_{TY} \leftarrow t', e'_{ZY} \leftarrow z. \end{cases}$$

---

[5]Since T^π, Z^π are not affected by the information $T = t'$ through the path π, then we marginalize out the mechanisms for them.

For conciseness, we usually construct the (augmented) causal graph of those intervention SCMs without separating the variables in the counterfactual world created by the hypothetical intervention from the factual world, except for path intervention. In particular, the counterfactual variable $Y^{do(t')}$ created by do intervention $do(t')$ is represented with notation y instead of $y^{do(t')}$. Henceforth, we are using y rather than $y^{do(t')}$ or $y^{\sigma(t')}$ in those structural equations in the formal definition of *do* intervention and *info* intervention. However, we find that the *path* intervention is quite different from *info* or *do* intervention in that it includes cross-world structural equation between two endogenous variable represented by z and y^{π}. In fact, the cross-world counterfactuals involved in mediation analysis is exactly what makes it more complicated, and we need the counter-intuitive recanting witness criteria for identification with traditional methods. In contrast, our novel path specific effects exhibit intuitive interpretation and identification properties.

4 Identification of Path-specific Effects

Robins et al. (2020) defines the potential outcome $V_i(\pi, a, a')$ by setting A to a for path π that end in V_i, and setting A to a' for the purposes of proper causal paths from A to V_i other than π. Formally, the definition is as follows, for any $V_i \in V$:

$$V_i(\pi, a, a') \equiv a \quad \text{if} \quad V_i \text{ is } A, \tag{10}$$

$$V_i(\pi, a, a') \equiv V_i(\{V_j(\pi, a, a') \mid V_j \in pa_i^{\pi}\}, \{V_j(a') \mid V_j \in pa_i^{\overline{\pi}}\})$$

where $V_j(a') \equiv a'$ if V_j is A, and given by recursive substitution otherwise, pa_i^{π} is the set of parents of V_i along the path π, and $pa_i^{\overline{\pi}}$ is the set of all other parents of V_i. Their path specific effects rely on the recursive assumption, and no recanting witness for identification. Although our path-specific effects can be defined for non-recursive SCMs, but we will not further address that problem in this paper considering the numerous complications for cyclic causal models.

Identification for path-specific effects based on nested counterfactuals for causal graphical models is governed by the criterion known as the recanting witness criterion (Avin et al., 2005; Shpitser & Tchetgen, 2016; Shpitser et al., 2020). Here let's look into the causal diagrams Fig. 1(a), a submodel of X, M, Y can be induced through marginalization with the causal graph $X \to M \to Y$ and $M \to Y$. This is the simplest case of mediation analysis, hence the path specific effect along $\pi : X \to Y$ can be identified. However, the corresponding path $\pi : X \to A \to Y$ specific effect can not be identified because A is a recanting witness. This leads to a model with additional observational information of A to disable the identifiability of path-specific effect which conceptually contradicts the intuition that the more information the better.

In contrast, our novel path-specific effects induced by our path intervention can be identified regardless of whether it satisfies recanting witness criteria or not. In particular, for a path intervention $\pi(a')$, the joint distribution of counterfactual variables can always be identified by the π-formula. To formulate our result, we denote $desc^{\pi}(A)$ as descendants of A in the path π, i.e., all nodes in π with a parent node in causal path π.

Theorem 7 (π-formula) *In a causal diagram $\mathcal{G}$(of a Markovian SCM $\mathcal{M}$) with a factorized probability p, for a causal path $\pi : A \to \cdots \to Y$, the joint counterfactual distribution over variables $p(v^{\pi}, v)$ for path intervention $\pi(A = a')$ is identified via a equation called the π-formula:*

$$p(v_{desc^{\pi}(A)}^{\pi}, v) = \prod_{k \in desc^{\pi}(A)} p(v_k^{\pi} | e_{pa(k),k}^{\pi}) \cdot \prod_{i \in V} p(v_i | v_{pa(i)}) \tag{11}$$

where $e_{j,k}^{\pi} = v_j$ if $(j,k) \notin \pi$; else if V_j is A, a'; else v_j^{π}.

Proof By definition of path intervention, the modified causal mechanisms for $(\mathbf{V}^{\pi}, \mathbf{V})$ are, for any $i \in V$:

$$\begin{cases} v_i \leftarrow f_i(e_{pa(i),i}, u_i) \\ e_{j,i} \leftarrow v_j \\ v_i^{\pi} \leftarrow f_i(e_{pa(i),i}', u_i') \\ e_{j,i}' \leftarrow v_j \text{ if } (j,i) \notin \pi; \text{ else if } V_j \text{ is } A, a'; \text{ else } v_j^{\pi} \end{cases}$$

The intervened SCM is also acyclic with independent errors, thus factorize with the following density:

$$p(v^\pi, v) = \prod_{k \in V} p(v_k^\pi | e_{pa(k),k}^\pi) \cdot \prod_{i \in V} p(v_i | v_{pa(i)})$$

where $e_{j,k}^\pi = v_j$ if $(j,k) \notin \pi$; else if V_j is A, a'; else v_j^π. Notice that for any counterfactual variable not in $desc^\pi(A)$, it receives the information directly from factual variables and the i.i.d. copy of exogenous variable. These counterfactual variables are leafs of the intervention diagram, hence marginalizing out them directly leads to the π-formula. ∎

We then can formally identify the path specific effect for causal diagram Fig. 1(a) in the following example:

Example 3 *For a causal diagram(of a Markovian **SCM** $\mathcal{M}$) for variables X, A, M, Y, with causal relationships $X \to A \to M \to Y$ and $A \to Y$ in Fig 1(a), the joint distribution of counterfactual and factual variables in intervention diagram Fig. 3 of path intervention $\pi(x')$, where $\pi : X \to A \to Y$, can be identified the following equation:*

$$p(y^\pi, a^\pi, y, m, a, x) = p(y^\pi | a^\pi, m) p(a^\pi | x') p(y|a,m) p(m|a) p(a|x) p(x)$$

Then the density of effect variable Y^π can be obtained by summation over $\{a^\pi, y, m, a, x\}$:

$$p(y^\pi) = \sum_{a^\pi, m, a, x} p(y^\pi, a^\pi, y, m, a, x)$$

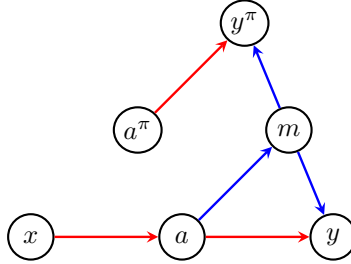$$= \sum_{m, a^\pi, a, x} p(y^\pi | a^\pi, m) p(a^\pi | x') p(m|a) p(a|x) p(x)$$

Figure 3: Intervention diagram induced by path intervention $\pi(x')$, where $\pi : X \to A \to Y$ with input information $X = x'$.

Let's see a more complicated example in which M is a recanting witness.

Example 4 *For a causal diagram(of a Markovian **SCM** $\mathcal{M}$) with a probability p that factorize, we consider the path intervention $\pi(a')$ for the causal path $\pi : A \to M \to Y$ with input information $A = a'$. Then according to intervention graph Fig. 4(b) with π-formula, we have*

$$p(y^\pi, m^\pi, z, m, a, w) = p(y^\pi | m^\pi, w, z) p(y^\pi | w, a') \cdot p(z|m,a) p(m|a,w) p(a|w) p(w)$$

Note that we use latent projection onto variables of interest. In particular, we exclude the factual variable Y in the intervention graph for conciseness, which can be marginalized over the π-formula. However you cannot marginalize out factual variable M to identify the distribution of Y^π.

This way of constructing intervention diagram is somewhat similar to twin-network (Pearl et al., 2009) for both including factual and counterfactual variables. However, there are completely different in several ways. First, the twin-SCM only have U_V which connects factual and counterfactual variables, our intervention diagram includes an i.i.d. copy of U_V for the corresponding intervened SCM. Second, there are no direct edges between factual and counterfactual variables in a twin-SCM.
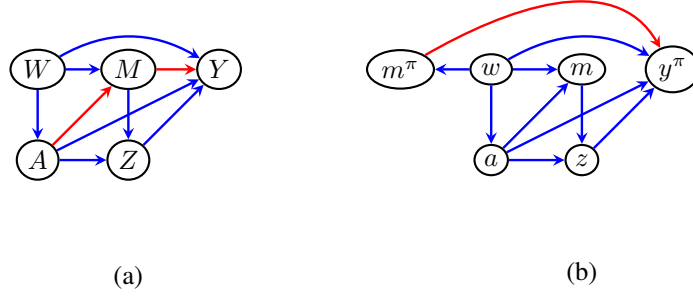
Figure 4: (a) A causal diagram with recanting witness M; (b) Intervention diagram.

In contrast, there are no direct edges between two counterfacutal variables unless there exists an edge in π connecting them, which leads to counterfactual variables usually have no output edge in the intervention diagram. In other words, parents of counterfactual variables are usually factual variables, e.g. $A, Z, W \to Y^\pi$ in Fig. 4(b). Third, Our intervention diagrams include probabilistic relations instead of deterministic relations, while the incompleteness of d-separation occurs in twin-SCM due to deterministic relations (Shpitser et al., 2020).

5 Concluding remarks and Discussion

In this paper, by relating the informational account to **SCM**, we propose a novel *path* intervention framework to formalize path-specific effects aligned the dictum "no causation without manipulation". In the *path* intervention framework, the information account of causality has suggested of separating causal mechanisms into information transfer and process. Our path intervention explicitly manipulate the mechanisms, thus path-specific effects even can be well-defined for cyclic **SCM**s. Intuitively, our path-specific effect $Y^{\pi(a')}$ is obtained by making the treatment information $A = a'$ exclusively pass through the causal path π while allowing all other information outside from the π to keep constant. Moreover, we address the identifiability of path-specific effects for causal diagrams and show that they can be identified even when recanting witness exists with the π-formula.

More discussions. Causality has been one of the basic topics of philosophy since the time of Aristotle. However during the last two decades or so, interest in causality has become very intense in the philosophy of science community, and a great variety of novel views on the subject have emerged and been developed. Among those novel views, it has been claimed that an informational account of causality, formally proposed and developed since Collier (1999) in philosophy, might be useful to the scientific problem of how we think about, and ultimately trace, causal linking, and so to causal inference and reasoning. An informational account of causality can be useful to help us reconstruct how science builds up understanding of the causal structure of the world. Broadly, we find mechanisms that help us grasp causal linking in a coarse-grained way. Then we can think in terms of causal linking in a more fine-grained way by thinking informationally. An informational account of causality may also give us the prospect of saying what causality is, in a way that is not tailored to the description of reality provided by a given discipline. And it carries the advantage over other causal metaphysics that it fares well with the applicability problem for other accounts of production (processes and mechanism)(Illari & Russo, 2014). But from an application perspective, what benefits we can gain in causal modeling? Or it is just a rhetorical flourish?

Algorithmic information theory has been used in causal modeling(see e.g. (Schölkopf, 2019)).It is also worth to mention that, broadly speaking, the information accounts of causality can also facilitate interpretation of existing widely used causal propositions. For example, regarding back-door/front-door criteria, the goal of which can be consistently considered as whether the observational information of a set of variables is enough or not to answer causal-effect estimation question, instead of conventional understanding such as controlling variables. Moreover, in general sense of information accounts, Pearl points out that questions in one layer of the causal hierarchy can only be answered when corresponding layer information are available(Pearl, 2019b; Bareinboim

et al., 2020), and Scholköpf believes causal science will enable us act and decision with information from Lorenzian imagined space (Schölkopf, 2019). The recently proposed Mini-Turing test for AI — How can machines represent causal knowledge in a way that would enable them to access the necessary information swiftly, answer questions correctly, and do it with ease, as a human can? (Pearl & Mackenzie, 2018). In summary, to build true intelligent machines, climb the ladder of data, information, knowledge and wisdom, we might need to incorporate the information accounts of causality into causal tasks.

References

Avin, C., Shpitser, I., and Pearl, J. Identifiability of path-specific effects. 2005.

Bareinboim, E., Correa, J., Ibeling, D., and Icard, T. On pearl's hierarchy and the foundations of causal inference. *ACM Special Volume in Honor of Judea Pearl (provisional title)*, 2020.

Baron, R. M. and Kenny, D. A. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, 51(6):1173, 1986.

Bongers, S., Peters, J., Schölkopf, B., and Mooij, J. M. Theoretical aspects of cyclic structural causal models. *arXiv preprint arXiv:1611.06221*, 2016.

Collier, J. D. Causation is the transfer of information. In *Causation and laws of nature*, pp. 215–245. Springer, 1999.

Feltz, B., Missal, M., and Sims, A. *Free Will, Causality, and Neuroscience*. Brill, 2019.

Gong, H. and Zhu, K. Info intervention and its generalization. *arXiv preprint arXiv:1907.11090*, 2020.

Illari, P. and Russo, F. *Causality: Philosophical theory meets scientific practice*. OUP Oxford, 2014.

Illari, P. M. Why theories of causality need production: An information transmission account. *Philosophy & Technology*, 24(2):95–114, 2011.

Imbens, G. W. Potential outcome and directed acyclic graph approaches to causality: Relevance for empirical practice in economics. *Journal of Economic Literature*, 58(4):1129–79, 2020.

Malinsky, D., Shpitser, I., and Richardson, T. A potential outcomes calculus for identifying conditional path-specific effects. *Proceedings of machine learning research*, 89:3080, 2019.

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

Pearl, J. Direct and indirect effects. *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, 2001.

Pearl, J. Causal and counterfactual inference. *The Handbook of Rationality*, pp. 1–41, 2019a.

Pearl, J. The seven tools of causal inference, with reflections on machine learning. *Communications of the ACM*, 62(3):54–60, 2019b.

Pearl, J. and Mackenzie, D. *The book of why: the new science of cause and effect*. Basic Books, 2018.

Pearl, J. et al. Models, reasoning and inference. *Cambridge, UK: Cambridge University Press*, 2000.

Pearl, J. et al. Causal inference in statistics: An overview. *Statistics Surveys*, 3:96–146, 2009.

Robins, J. M. and Greenland, S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology*, pp. 143–155, 1992.

Robins, J. M., Richardson, T. S., and Shpitser, I. An interventionist approach to mediation analysis. *arXiv preprint arXiv:2008.06019*, 2020.

Schölkopf, B. Causality for machine learning. *arXiv preprint arXiv:1911.10500*, 2019.

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

Shpitser, I., Richardson, T. S., and Robins, J. M. Multivariate counterfactual systems and causal graphical models. *arXiv preprint arXiv:2008.06017*, 2020.

Zhang, J. and Bareinboim, E. Non-parametric path analysis in structural causal models. In *Proceedings of the 34th Conference on Uncertainty in Artificial Intelligence*, 2018.