Journal of Machine Learning Research 9 (2008) 1437-1474 Submitted 6/07; Revised 2/08; Published 7/08

Causal Reasoning with Ancestral Graphs

Jiji Zhang JIJI@HSS.CALTECH.EDU
Division of the Humanities and Social Sciences
California Institute of Technology
Pasadena, CA 91106, USA

Editor: Gregory F. Cooper

Abstract

Causal reasoning is primarily concerned with what would happen to a system under external interventions. In particular, we are often interested in predicting the probability distribution of some random variables that would result if some other variables were *forced* to take certain values. One prominent approach to tackling this problem is based on causal Bayesian networks, using directed acyclic graphs as *causal* diagrams to relate post-intervention probabilities to pre-intervention probabilities that are estimable from observational data. However, such causal diagrams are seldom fully testable given observational data. In consequence, many causal discovery algorithms based on data-mining can only output an equivalence class of causal diagrams (rather than a single one). This paper is concerned with causal reasoning given an equivalence class of causal diagrams, represented by a (partial) *ancestral graph*. We present two main results. The first result extends Pearl (1995)'s celebrated *do-calculus* to the context of ancestral graphs. In the second result, we focus on a key component of Pearl's calculus—the property of *invariance under interventions*, and give stronger graphical conditions for this property than those implied by the first result. The second result also improves the earlier, similar results due to Spirtes et al. (1993).

Keywords: ancestral graphs, causal Bayesian network, do-calculus, intervention

1. Introduction

Intellectual curiosity aside, an important reason for people to care about causality or causal explanation is the need—for example, in policy assessment or decision making—to predict consequences of actions or interventions before actually carrying them out. Sometimes we can base that prediction on similar past interventions or experiments, in which case the inference is but an instance of the classical inductive generalization. Other times, however, we do not have access to sufficient controlled experimental studies for various reasons, and can only make passive observations before interventions take place. Under the latter circumstances, we need to reason from pre-intervention or observational data to a post-intervention setting.

A prominent machinery for causal reasoning of this kind is known as *causal Bayesian network* (Spirtes et al., 1993; Pearl, 2000), which we will describe in more detail in the next section. In this framework, once the causal structure—represented by a directed acyclic graph (DAG) over a set of attributes or random variables—is fully given, every query about post-intervention probability can be answered in terms of pre-intervention probabilities. So, if every variable in the causal structure is (passively) observed, the observational data can be used to estimate the post-intervention probability of interest.

Complications come in at least two ways. First, some variables in the causal DAG may be unobserved, or worse, unobservable. So even if the causal DAG (with latent variables) is fully known, we may not be able to predict certain intervention effects because we only have data from the marginal distribution over the observed variables instead of the joint distribution over all causally relevant variables. The question is what post-intervention probability is or is not identifiable given a causal DAG with latent variables. Much of Pearl's work (Pearl, 1995, 1998, 2000), and more recently Tian and Pearl (2004) are paradigmatic attempts to address this problem.

Second, the causal structure is seldom, if ever, fully known. In the situation we are concerned with in this paper, where no substantial background knowledge or controlled study is available, we have to rely upon observational data to inform us about causal structure. The familiar curse is that very rarely can observational data determine a unique causal structure, and many causal discovery algorithms in the literature output an equivalence class of causal structures based on observational data (Spirtes et al., 1993; Meek, 1995a; Spirtes et al., 1999; Chickering, 2002).[1] Different causal structures in the class may or may not give the same answer to a query about post-intervention probability. For a simple illustration, consider two causal Bayesian networks (see Section 2 below), $X \to Y \to Z$ and $X \leftarrow Y \to Z$, over three variables X, Y and Z. The two causal structures are indistinguishable (without strong parametric assumptions) by observational data. Suppose we are interested in the post-intervention probability distribution of Y given that X is manipulated to take some fixed value x. The structure $X \to Y \to Z$ entails that the post-intervention distribution of Y is identical to the pre-intervention distribution of Y *conditional on* $X = x$, whereas the structure $X \leftarrow Y \to Z$ entails that the post-intervention distribution of Y is identical to the pre-intervention marginal distribution of Y. So the two structures give different answers to this particular query. By contrast, if we are interested in the post-intervention distribution of Z under an intervention on Y, the two structures give the same answer.

The matter becomes formidably involved when both complications are present. Suppose we observe a set of random variables $\mathbf{O}$, but for all we know, the underlying causal structure may involve extra latent variables. We will not worry about the estimation of the pre-intervention distribution of $\mathbf{O}$ in this paper, so we may well assume for simplicity that the pre-intervention distribution of $\mathbf{O}$ is known. But we are interested in queries about post-intervention probability, such as the probability of $\mathbf{Y}$ conditional on $\mathbf{Z}$ that would result under an intervention on $\mathbf{X}$ (where $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \subseteq \mathbf{O}$). The question is whether and how we can answer such queries from the given pre-intervention distribution of $\mathbf{O}$.

This problem is naturally divided into two parts. The first part is what some causal discovery algorithms attempt to achieve, namely, to learn something about the causal structure—usually features shared by all causal structures in an equivalence class—from the pre-intervention distribution of $\mathbf{O}$. The second part is to figure out, given the learned causal information, whether a post-intervention probability is identifiable in terms of pre-intervention probabilities.

This paper provides some results concerning the second part, assuming the available causal information is summarized in a (partial) *ancestral graph*. Ancestral graphical models (Richardson and Spirtes, 2002, 2003) have proved to be an elegant and useful surrogate for DAG models with latent variables (more details follow in Section 3), not the least because provably correct algorithms are available for learning an equivalence class of ancestral graphs represented by a partial ancestral graph from the pre-intervention distribution of the observed variables—in particular, from the

1. The recent work on linear non-Gaussian structural equation models (Shimizu et al., 2006) is an exception. However, we do not make parametric assumptions in this paper.

conditional independence and dependence relations implied by the distribution (Spirtes et al., 1999; Zhang, forthcoming).

We have two main results. First, we extend the *do*-calculus of Pearl (1995) to the context of ancestral graphs (Section 4), so that the resulting calculus is based on an equivalence class of causal DAGs with latent variables rather than a single one. Second, we focus on a key component of Pearl's calculus—the property of *invariance under interventions* studied by Spirtes et al. (1993), and give stronger graphical conditions for this property than those implied by the first result (Section 5). Our result improves upon the Spirtes-Glymour-Scheines conditions for invariance formulated with respect to the so-called *inducing path graphs*, whose relationship with ancestral graphs is discussed in Appendix A.

2. Causal Bayesian Network

A Bayesian network for a set of random variables $\mathbf{V}$ consists of a pair $\langle \mathcal{G}, P\rangle$, where $\mathcal{G}$ is a directed acyclic graph (DAG) with $\mathbf{V}$ as the set of vertices, and P is the joint probability function of $\mathbf{V}$, such that P factorizes according to $\mathcal{G}$ as follows:

$$P(\mathbf{V}) = \prod_{Y\in\mathbf{V}} P(Y|\mathbf{Pa}_{\mathcal{G}}(Y))$$

where $\mathbf{Pa}_{\mathcal{G}}(Y)$ denotes the set of parents of Y in $\mathcal{G}$. In a causal Bayesian network, the DAG $\mathcal{G}$ is interpreted causally, as a representation of the causal structure over $\mathbf{V}$. That is, for $X, Y \in \mathbf{V}$, an arrow from X to Y ($X \to Y$) in $\mathcal{G}$ means that X has a *direct* causal influence on Y relative to $\mathbf{V}$. We refer to a causally interpreted DAG as a **causal DAG**. The postulate that the (pre-intervention) joint distribution P factorizes according to the causal DAG $\mathcal{G}$ is known as the **causal Markov condition**.

What about interventions? For simplicity, let us focus on what Pearl (2000) calls *atomic* interventions—interventions that fix the values of the target variables—though the results in Section 5 also apply to more general types of interventions (such as interventions that confer a non-degenerate probability distribution on the target variables). In the framework of causal Bayesian network, an intervention on $\mathbf{X} \subseteq \mathbf{V}$ is supposed to be *effective* in the sense that the value of $\mathbf{X}$ is completely determined by the intervention, and *local* in the sense that the conditional distributions of other variables (variables not in $\mathbf{X}$) given their respective parents in the causal DAG are not affected by the intervention. Graphically, such an intervention amounts to erasing all arrows into $\mathbf{X}$ in the causal DAG (because variables in $\mathbf{X}$ do not depend on their original parents any more), but otherwise keeping the graph as it is. Call this modified graph the **post-intervention causal graph**.

Based on this understanding of interventions, the following postulate has been proposed by several authors in various forms (Robins, 1986; Spirtes et al., 1993; Pearl, 2000):

> **Intervention Principle** Given a causal DAG $\mathcal{G}$ over $\mathbf{V}$ and a (pre-intervention) joint distribution P that factorizes according to $\mathcal{G}$, the post-intervention distribution $P_{\mathbf{X}:=\mathbf{x}}(\mathbf{V})$—that is, the joint distribution of $\mathbf{V}$ after $\mathbf{X} \subseteq \mathbf{V}$ are manipulated to values $\mathbf{x}$ by an intervention—takes a similar, truncated form of factorization, as follows:
>
> $$P_{\mathbf{X}:=\mathbf{x}}(\mathbf{V}) = \begin{cases} \prod_{Y\in\mathbf{V}\setminus\mathbf{X}} P(Y|\mathbf{Pa}_{\mathcal{G}}(Y)) & \text{for values of } \mathbf{V} \text{ consistent with } \mathbf{X}=\mathbf{x}, \\ 0 & \text{otherwise.} \end{cases}$$

Note that in the case of a null intervention (when $\mathbf{X} = \emptyset$), the intervention principle implies the factorization of the pre-intervention distribution P according to $\mathcal{G}$, which is just the causal Markov condition. So the intervention principle generalizes the causal Markov condition: it assumes that the post-intervention distribution also satisfies the causal Markov condition with the post-intervention causal graph.

By the intervention principle, once the causal DAG is given, the post-intervention joint distribution can be calculated in terms of pre-intervention probabilities.[2] So if every variable is observed, and hence those pre-intervention probabilities can be estimated, any post-intervention probability is estimable as well.

It is time to recall the two complications mentioned in the last section. First, the intervention principle is only plausible when the given set of variables is *causally sufficient*. Here is what causal sufficiency means. Given a set of variables $\mathbf{V}$, and two variables $A, B \in \mathbf{V}$, a variable C (not necessarily included in $\mathbf{V}$) is called a *common direct cause* of A and B relative to $\mathbf{V}$ if C has a direct causal influence on A and also a direct causal influence on B relative to $\mathbf{V} \cup \{C\}$. $\mathbf{V}$ is said to be *causally sufficient* if for every pair of variables $V_1, V_2 \in \mathbf{V}$, every common direct cause of V_1 and V_2 relative to $\mathbf{V}$ is also a member of $\mathbf{V}$. It is well known that the causal Markov condition tends to fail for a causally insufficient set of variables (Spirtes et al., 1993), and even more so with the intervention principle. But in most real situations, there is no reason to assume that the set of observed variables is causally sufficient, so the causal Bayesian network may well involve latent variables.

Second, the causal DAG is not fully learnable with observational, pre-intervention data. The causal discovery algorithms in the literature—some of which are provably correct in the large sample limit assuming the causal Markov condition and its converse, causal Faithfulness condition—typically return an equivalence class of DAGs that imply the same conditional independence relations among the observed variables (according to the Markov condition), with some causal features in common that constitute the learned causal information. Given such limited causal information, a post-intervention probability may or may not be uniquely identifiable.

Taking both complications into account, the interesting question is this: what causal reasoning is warranted given the causal information learnable by algorithms that do not assume causal sufficiency for the set of observed variables, such as the FCI algorithm presented in Spirtes et al. (1999)? Before we explore the question, let us make it a little more precise with the formalism of ancestral graphs.

3. Ancestral Graphical Models

Ancestral graphical models are motivated by the need to represent data generating processes that may involve latent confounders and/or selection bias,[3] without explicitly modelling the unobserved variables (Richardson and Spirtes, 2002). We do not deal with selection bias in this paper, so we use only part of the machinery.

A (directed) *mixed graph* is a vertex-edge graph that may contain two kinds of edges: directed edges ($\rightarrow$) and bi-directed edges ($\leftrightarrow$). Between any two vertices there is at most one edge. The two ends of an edge we call *marks*. Obviously there are two kinds of marks: *arrowhead* ($>$) and *tail* ($-$). The marks of a bi-directed edge are both arrowheads, and a directed edge has one arrowhead

2. A technical issue is that some conditional probabilities may be undefined in the pre-intervention distribution. In this paper we ignore that issue by assuming that the pre-intervention distribution is strictly positive. Otherwise we just need to add the proviso “when all the conditional probabilities involved are defined” to all our results.
3. Roughly speaking, there is selection bias if the probability of a unit being sampled depends on certain properties of the unit. The kind of selection bias that is especially troublesome for causal inference is when two or more properties of interest affect the probability of being sampled, giving rise to “misleading” associations in the sample.

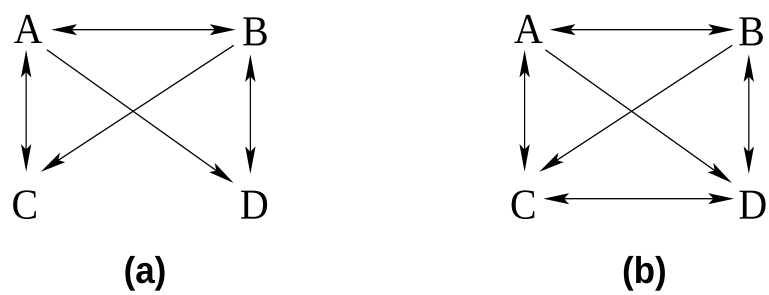

Figure 1: (a) an ancestral graph that is not maximal; (b) a maximal ancestral graph.

and one tail. We say an edge is *into* (or *out of*) a vertex if the mark of the edge at the vertex is an arrowhead (or tail).

Two vertices are said to be *adjacent* in a graph if there is an edge (of any kind) between them. Given a mixed graph $\mathcal{G}$ and two adjacent vertices X, Y therein, X is called a *parent* of Y and Y a *child* of X if $X \to Y$ is in $\mathcal{G}$; X is called a *spouse* of Y (and Y a spouse of X) if $X \leftrightarrow Y$ is in $\mathcal{G}$. A *path* in $\mathcal{G}$ is a sequence of distinct vertices $\langle V_0, \ldots, V_n \rangle$ such that for all $0 \leq i \leq n-1$, V_i and V_{i+1} are adjacent in $\mathcal{G}$. A *directed path from* V_0 *to* V_n in $\mathcal{G}$ is a sequence of distinct vertices $\langle V_0, \ldots, V_n \rangle$ such that for all $0 \leq i \leq n-1$, V_i is a parent of V_{i+1} in $\mathcal{G}$. X is called an *ancestor* of Y and Y a *descendant* of X if $X = Y$ or there is a directed path from X to Y. We use $\mathbf{Pa}_{\mathcal{G}}, \mathbf{Ch}_{\mathcal{G}}, \mathbf{Sp}_{\mathcal{G}}, \mathbf{An}_{\mathcal{G}}, \mathbf{De}_{\mathcal{G}}$ to denote the set of parents, children, spouses, ancestors, and descendants of a vertex in $\mathcal{G}$, respectively. A *directed cycle* occurs in $\mathcal{G}$ when $Y \to X$ is in $\mathcal{G}$ and $X \in \mathbf{An}_{\mathcal{G}}(Y)$. An *almost directed cycle* occurs when $Y \leftrightarrow X$ is in $\mathcal{G}$ and $X \in \mathbf{An}_{\mathcal{G}}(Y)$.[4]

Given a path $p = \langle V_0, \ldots, V_n \rangle$ with $n > 1$, V_i $(1 \leq i \leq n-1)$ is a *collider* on p if the two edges incident to V_i are both into V_i, that is, have an arrowhead at V_i; otherwise it is called a *noncollider* on p. In Figure 1(a), for example, B is a collider on the path $\langle A, B, D \rangle$, but is a non-collider on the path $\langle C, B, D \rangle$. A *collider path* is a path on which every vertex except for the endpoints is a collider. For example, in Figure 1(a), the path $\langle C, A, B, D \rangle$ is a collider path because both A and B are colliders on the path. Let $\mathbf{L}$ be any subset of vertices in the graph. An *inducing path relative to* $\mathbf{L}$ is a path on which every vertex not in $\mathbf{L}$ (except for the endpoints) is a collider on the path and every collider is an ancestor of an endpoint of the path. For example, any single-edge path is trivially an inducing path relative to any set of vertices (because the definition does not constrain the endpoints of the path). In Figure 1(a), the path $\langle C, B, D \rangle$ is an inducing path relative to $\{B\}$, but not an inducing path relative to the empty set (because B is not a collider). However, the path $\langle C, A, B, D \rangle$ is an inducing path relative to the empty set, because both A and B are colliders on the path, A is an ancestor of D, and B is an ancestor of C. To simplify terminology, we will henceforth refer to inducing paths relative to the empty set simply as inducing paths.[5]

Definition 1 (MAG) *A mixed graph is called a maximal ancestral graph (MAG) if*

i. the graph does not contain any directed or almost directed cycles (ancestral); and

4. The terminology of "almost directed cycle" is motivated by the fact that removing the arrowhead at Y on $Y \leftrightarrow X$ results in a directed cycle.
5. They are called *primitive inducing paths* by Richardson and Spirtes (2002).

ii. there is no inducing path between any two non-adjacent vertices (maximal).

The first condition is obviously an extension of the defining condition for DAGs. It follows that in an ancestral graph an arrowhead, whether on a directed edge or a bi-directed edge, implies non-ancestorship. The second condition is a technical one, but the original motivation is the familiar pairwise Markov property of DAGs: if two vertices are not adjacent, then they are d-separated by some set of other vertices. The notion of d-separation carries over to mixed graphs in a straightforward way, as we will see shortly. But in general an ancestral graph does not need to satisfy the pairwise Markov property, or what is called maximality here. A sufficient and necessary condition for maximality turns out to be precisely the second clause in the above definition, as proved by Richardson and Spirtes (2002). So although the graph in Figure 1(a) is ancestral, it is not maximal because there is an inducing path between C and D (i.e., $\langle C, A, B, D\rangle$), but C and D are not adjacent. However, each non-maximal ancestral graph has a unique supergraph that is ancestral and maximal. For example, Figure 1(b) is the unique MAG that is also a supergraph of Figure 1(a); the former has an extra bi-directed edge between C and D.

It is worth noting that both conditions in Definition 1 are obviously met by a DAG. Hence, syntactically a DAG is also a MAG, one without bi-directed edges.

An important notion in directed graphical models is that of d-separation, which captures exactly the conditional independence relations entailed by a DAG according to the Markov condition. It is straightforward to extend the notion to mixed graphs, which, following Richardson and Spirtes (2002), we call *m-separation*.

Definition 2 (m-separation) *In a mixed graph, a path p between vertices X and Y is active (or m-connecting) relative to a (possibly empty) set of vertices $\mathbf{Z}$ ($X, Y \notin \mathbf{Z}$) if*

i. every non-collider on p is not a member of $\mathbf{Z}$;

ii. every collider on p is an ancestor of some member of $\mathbf{Z}$.

X and Y are said to be m-separated by $\mathbf{Z}$ if there is no active path between X and Y relative to $\mathbf{Z}$.

Two disjoint sets of variables $\mathbf{X}$ and $\mathbf{Y}$ are m-separated by $\mathbf{Z}$ if every variable in $\mathbf{X}$ is m-separated from every variable in $\mathbf{Y}$ by $\mathbf{Z}$.

In DAGs, obviously, m-separation reduces to d-separation. The (global) Markov property of ancestral graphical models is defined by m-separation.

A nice property of MAGs is that they can represent the marginal independence models of DAGs in the following sense: given any DAG $\mathcal{G}$ over $\mathbf{V} = \mathbf{O} \cup \mathbf{L}$—where $\mathbf{O}$ denotes the set of observed variables, and $\mathbf{L}$ denotes the set of latent variables—there is a MAG over $\mathbf{O}$ alone such that for any disjoint $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \subseteq \mathbf{O}$, $\mathbf{X}$ and $\mathbf{Y}$ are d-separated by $\mathbf{Z}$ in $\mathcal{G}$ (and hence entailed by $\mathcal{G}$ to be independent conditional on $\mathbf{Z}$) if and only if they are m-separated by $\mathbf{Z}$ in the MAG (and hence entailed by the MAG to be independent conditional on $\mathbf{Z}$). The following construction gives us such a MAG:

Input: a DAG $\mathcal{G}$ over $\langle \mathbf{O}, \mathbf{L}\rangle$
Output: a MAG $\mathcal{M}_{\mathcal{G}}$ over $\mathbf{O}$

1. for each pair of variables $A, B \in \mathbf{O}$, A and B are adjacent in $\mathcal{M}_{\mathcal{G}}$ if and only if there is an inducing path between them relative to $\mathbf{L}$ in $\mathcal{G}$;

2. for each pair of adjacent variables A, B in $\mathcal{M}_{\mathcal{G}}$, orient the edge as $A \to B$ in $\mathcal{M}_{\mathcal{G}}$ if A is an ancestor of B in $\mathcal{G}$; orient it as $A \leftarrow B$ in $\mathcal{M}_{\mathcal{G}}$ if B is an ancestor of A in $\mathcal{G}$; orient it as $A \leftrightarrow B$ in $\mathcal{M}_{\mathcal{G}}$ otherwise.

It can be shown that $\mathcal{M}_{\mathcal{G}}$ is indeed a MAG and represents the marginal independence model over $\mathbf{O}$ (Richardson and Spirtes, 2002; also see Lemma 20 below). More importantly, $\mathcal{M}_{\mathcal{G}}$ also retains the ancestral relationships—and hence causal relationships under the standard interpretation—among $\mathbf{O}$. So, if $\mathcal{G}$ is the causal DAG for $\langle \mathbf{O}, \mathbf{L} \rangle$, it is fair to call $\mathcal{M}_{\mathcal{G}}$ the **causal MAG** for $\mathbf{O}$. Henceforth when we speak of a MAG over $\mathbf{O}$ *representing* a DAG over $\langle \mathbf{O}, \mathbf{L} \rangle$, we mean that the MAG is the output of the above construction procedure applied to the DAG.

Different causal DAGs may correspond to the same causal MAG. So essentially a MAG represents a set of DAGs that have the exact same d-separation structures and ancestral relationships among the observed variables. A causal MAG thus carries uncertainty about what the true causal DAG is, but also reveals features that must be satisfied by the underlying causal DAG.

There is then a natural causal interpretation of the edges in MAGs, derivative from the causal interpretation of DAGs. A directed edge from A to B in a MAG means that A is a cause of B (which is a shorthand way of saying that there is a causal pathway from A to B in the underlying DAG); a bi-directed edge between A and B means that A is not a cause of B and B is not a cause of A, which implies that there is a latent common cause of A and B (i.e., there is a latent variable L in the underlying DAG such that there is a directed path from L to A and a directed path from L to B[6]).

We borrow a simple example from Spirtes et al. (1993) to illustrate various concepts and results in this paper. Suppose we are able to observe the following variables: *Income* (I), *Parents' smoking habits* (PSH), *Smoking* (S), *Genotype* (G) and *Lung cancer* (L). The data, for all we know, are generated according to an underlying mechanism which might involve unobserved common causes. Suppose, unknown to us, the structure of the causal mechanism is the one in Figure 2, where *Profession* is an unmeasured common cause of *Income* and *Smoking*.[7]

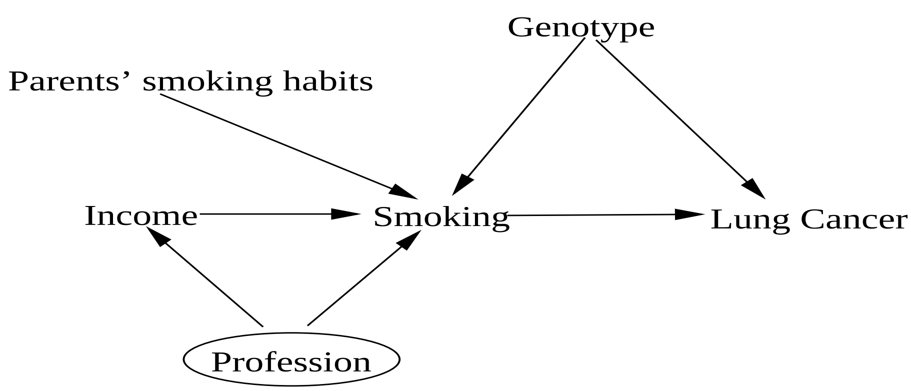

Figure 2: A causal DAG with a latent variable.

6. Note that a latent common cause is not necessarily a common *direct* cause as defined on page 4. The path from L to A, for example, may include other observed variables.
7. This example is used purely for illustrative purposes, so we will not worry why Profession is not observed but Genotype is. The exact domains of the variables do not matter either.

The causal MAG that corresponds to the causal DAG is depicted in Figure 3(a)—which *syntactically* happens to be a DAG in this case. This MAG can represent some other DAGs as well. For example, it can also represent the DAG with an extra latent common cause of *PSH* and *S*.

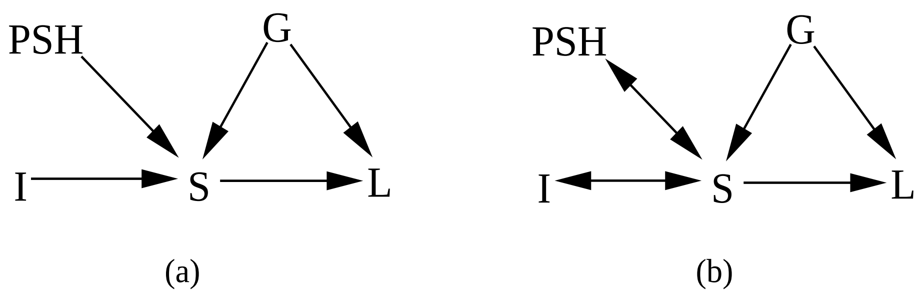

Figure 3: Two Markov Equivalent MAGs.

In general a MAG is still not fully testable with observational data. Just as different DAGs can share the exact same d-separation features and hence entail the exact same conditional independence constraints, different MAGs can entail the exact same constraints by the m-separation criterion. This is known as *Markov equivalence*. Several characterizations of the Markov equivalence between MAGs are available (Spirtes and Richardson, 1996; Ali et al., 2004; Zhang and Spirtes, 2005; Zhao et al., 2005). For the purpose of the present paper, it suffices to note that, as is the case with DAGs, all Markov equivalent MAGs have the same adjacencies and usually some common edge orientations as well. For example, the two MAGs in Figure 3 are Markov equivalent.

This motivates the following representation of equivalence classes of MAGs. Let *partial mixed graphs* denote the class of graphs that can contain four kinds of edges: $\rightarrow$, $\leftrightarrow$, $\circ\!\!-\!\!\circ$ and $\circ\!\!\rightarrow$, and hence three kinds of end marks for edges: arrowhead ($>$), tail ($-$) and circle ($\circ$).

Definition 3 (PAG) *Let $[\mathcal{M}]$ be the Markov equivalence class of an arbitrary MAG $\mathcal{M}$. The partial ancestral graph (PAG) for $[\mathcal{M}]$, $\mathcal{P}_{[\mathcal{M}]}$, is a partial mixed graph such that*

i. $\mathcal{P}_{[\mathcal{M}]}$ has the same adjacencies as $\mathcal{M}$ (and any member of $[\mathcal{M}]$) does;

ii. A mark of arrowhead is in $\mathcal{P}_{[\mathcal{M}]}$ if and only if it is shared by all MAGs in $[\mathcal{M}]$; and

iii. A mark of tail is in $\mathcal{P}_{[\mathcal{M}]}$ if and only if it is shared by all MAGs in $[\mathcal{M}]$.[8]

Basically a PAG represents an equivalence class of MAGs by displaying all common edge marks shared by all members in the class and displaying circles for those marks that are not common, much in the same way that a so-called Pattern (a.k.a. a PDAG or an essential graph) represents an equivalence class of DAGs (see, e.g., Spirtes et al., 1993, chap. 5; Chickering, 1995; Andersson et al., 1997). For instance, the PAG for our running example is drawn in Figure 4, which displays all the commonalities among MAGs that are Markov equivalent to the MAGs in Figure 3.

8. This defines what Zhang (2006, pp. 71) calls *complete* or *maximally oriented* PAGs. In this paper, we do not consider PAGs that fail to display all common edge marks in an equivalence class of MAGs (as, e.g., allowed in Spirtes et al., 1999), so we will simply use ‘PAG’ to mean ‘maximally oriented PAG’.

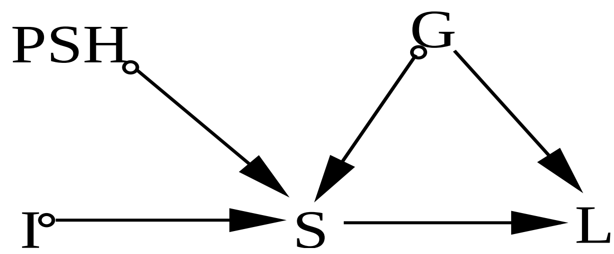

Figure 4: The PAG in our five-variable example.

Different PAGs, representing different equivalence classes of MAGs, entail different sets of conditional independence constraints. Hence a PAG is in principle fully testable by the conditional independence relations among the observed variables. Assuming the causal Markov condition and its converse, the causal Faithfulness condition,[9] there is a provably correct independence-constraint-based algorithm to learn a PAG from an oracle of conditional independence relations (Spirtes et al., 1999; Zhang, 2006, chap. 3).[10] Score-based algorithms for learning PAGs are also under investigation.

Directed paths and ancestors/descendants in a PAG are defined in the same way as in a MAG. In addition, a path between X and Y, $\langle X = V_0, \ldots, V_n = Y\rangle$, is called a *possibly directed path* from X to Y[11] if for every $0 < i \leq n$, the edge between V_{i-1} and V_i is not into V_{i-1}. Call X a *possible ancestor* of Y (and Y a *possible descendant* of X) if $X = Y$ or there is a possibly directed path from X to Y in the PAG.[12] For example, in Figure 4, the path $\langle I, S, L\rangle$ is a possibly directed path, and I is a possible ancestor of L. We use $\mathbf{PossibleAn}_{\mathcal{P}}(Y)$ to denote the set of possible ancestors of Y in $\mathcal{P}$.

In partial mixed graphs two analogues of m-connecting paths will play a role later. Let p be any path in a partial mixed graph, and W be any (non-endpoint) vertex on p. Let U and V be the two vertices adjacent to W on p. W is a *collider* on p if, as before, both the edge between U and W and the edge between V and W are into W (i.e., have an arrowhead at W, $U *\!\!\rightarrow W \leftarrow\!\!* V$). W is called a *definite non-collider* on p if the edge between U and W or the edge between V and W is out of W

9. We have introduced the causal Markov condition in its factorization form. In terms of d-separation, the causal Markov condition says that d-separation in a causal DAG implies conditional independence in the (pre-intervention) population distribution. The causal Faithfulness condition says that d-connection in a causal DAG implies conditional dependence in the (pre-intervention) population distribution. Given the exact correspondence between d-separation relations among the observed variables in the causal DAG and m-separation relations in the causal MAG, the two conditions imply that conditional independence relations among the observed variables correspond exactly to m-separation in the causal MAG, which forms the basis of constraint-based learning algorithms.
10. It is essentially the FCI algorithm (Spirtes et al., 1999), but with slight modifications (Zhang, 2006, chap. 3). The implemented FCI algorithm in the Tetrad IV package (http://www.phil.cmu.edu/projects/tetrad/tetrad4.html) is the modified version. By the way, if we also take into account the possibility of selection bias, then we need to consider a broader class of MAGs which can contain undirected edges, and the FCI algorithm needs to be augmented with additional edge inference rules (Zhang, 2006, chap. 4; forthcoming).
11. It is named a *potentially directed path* in Zhang (2006, pp. 99). The present terminology is more consistent with the names for other related notions, such as possible ancestor, possibly m-connecting path, etc.
12. The qualifier ’possible/possibly’ is used to indicate that there is some MAG represented by the PAG in which the corresponding path is directed, and X is an ancestor of Y. This is not hard to establish given the valid procedure for constructing representative MAGs from a PAG presented in Lemma 4.3.6 of Zhang (2006) or Theorem 2 of Zhang (forthcoming).

(i.e., has a tail at W, $U \leftarrow W *\!\!-\!\!* V$ or $U *\!\!-\!\!* W \rightarrow V$), or both edges have a circle mark at W and there is no edge between U and V (i.e., $U *\!\!-\!\!\circ W \circ\!\!-\!\!* V$, where U and V are not adjacent).[13] The first analogue of m-connecting path is the following:

Definition 4 (Definite m-connecting path) *In a partial mixed graph, a path p between two vertices X and Y is a definite m-connecting path relative to a (possibly empty) set of vertices $\mathbf{Z}$ ($X, Y \notin \mathbf{Z}$) if every non-endpoint vertex on p is either a definite non-collider or a collider and*

i. every definite non-collider on p is not a member of $\mathbf{Z}$;

ii. every collider on p is an ancestor of some member of $\mathbf{Z}$.

It is not hard to see that if there is a definite m-connecting path between X and Y given $\mathbf{Z}$ in a PAG, then in every MAG represented by the PAG, the corresponding path is an m-connecting path between X and Y given $\mathbf{Z}$. For example, in Figure 4 the path $\langle I, S, G\rangle$ is definitely m-connecting given L, and this path is m-connecting given L in every member of the equivalence class. A quite surprising result is that if there is an m-connecting path between X and Y given $\mathbf{Z}$ in a MAG, then there must be a definite m-connecting path (not necessarily the same path) between X and Y given $\mathbf{Z}$ in its PAG, which we will use in Section 5.

Another analogue of m-connecting path is the following:

Definition 5 (Possibly m-connecting path) *In a partial mixed graph, a path p between vertices X and Y is possibly m-connecting relative to a (possibly empty) set of vertices $\mathbf{Z}$ ($X, Y \notin \mathbf{Z}$) if*

i. every definite non-collider on p is not a member of $\mathbf{Z}$;

ii. every collider on p is a possible ancestor of some member of $\mathbf{Z}$.

Obviously a definite m-connecting path is also a possibly m-connecting path, but not necessarily vice versa. In particular, on a possibly m-connecting path it is not required that every (non-endpoint) vertex be of a "definite" status. Figure 5 provides an illustration. The graph on the right is the PAG for the equivalence class that contains the MAG on the left (in this case, unfortunately, no informative edge mark is revealed in the PAG). In the PAG, the path $\langle X, Y, Z, W\rangle$ is a possibly m-connecting path but not a definite m-connecting path relative to $\{Y, Z\}$, because Y and Z are neither colliders nor definite non-colliders on the path. Note that in the MAG, $\langle X, Y, Z, W\rangle$ is not m-connecting relative to $\{Y, Z\}$. In fact, X and W are m-separated by $\{Y, Z\}$ in the MAG. So unlike a definite m-connecting path, a mere possibly m-connecting path in a PAG does not necessarily correspond to a m-connecting path (or imply the existence of a m-connecting path) in a representative MAG in the equivalence class.[14]

As we will see, the main result in Section 4 is formulated in terms of absence of possibly m-connecting paths (what we will call, for want of a better term, definite m-separation), whereas the

13. '*' is used as wildcard that denotes any of the three possible marks: circle, arrowhead, and tail. When the graph is a PAG for some equivalence class of MAGs, the qualifier 'definite' is used to indicate that the vertex is a non-collider on the path in each and every MAG represented by the PAG, even though the circles may correspond to different marks in different MAGs. The reason why $U *\!\!-\!\!\circ W \circ\!\!-\!\!* V$ is a definite non-collider when U and V are not adjacent is because if it were a collider, it would be shared by all Markov equivalent MAGs, and hence would be manifest in the PAG.

14. This case is even more extreme in that in *every* MAG that belongs to the equivalence class, X and W are m-separated by Y and Z. So this example can be used to show that the *do*-calculus developed in Section 4 is not yet complete, though it is not clear how serious the incompleteness is.

Figure 5: Difference between possible and definite m-connecting paths: in the PAG on the right, $\langle X,Y,Z,W\rangle$ is a possibly m-connecting path relative to $\{Y,Z\}$ but *not* a definite m-connecting path relative to $\{Y,Z\}$. Also note that $\langle X,Y,Z,W\rangle$ is *not* m-connecting relative to $\{Y,Z\}$ in the MAG on the left, even though the MAG is a member of the equivalence class represented by the PAG.

main result in Section 5 is formulated in terms of absence of definite m-connecting paths. This is one important aspect in which the result in Section 5 is better than that in Section 4 (and than the analogous results presented in Spirtes et al., 1993) regarding the property of invariance under interventions. We will come back to this point after we present the PAG-based *do*-calculus.

4. Do-Calculus

Pearl (1995) developed an elegant *do*-calculus for identifying post-intervention probabilities given a single causal DAG with (or without) latent variables. To honor the name of the calculus, in this section we will use Pearl's *'do' operator* to denote post-intervention probabilities. Basically, the notation we used for the post-intervention probability function under an intervention on $\mathbf{X}$, $P_{\mathbf{X}:=\mathbf{x}}(\bullet)$, will be written as $P(\bullet|do(\mathbf{X}=\mathbf{x}))$.

The calculus contains three inference rules whose antecedents make reference to surgeries on the given causal DAG. There are two types of graph manipulations:

Definition 6 (Manipulations of DAGs) *Given a DAG $\mathcal{G}$ and a set of variables $\mathbf{X}$ therein,*

- *the* **X-lower-manipulation** *of $\mathcal{G}$ deletes all edges in $\mathcal{G}$ that are out of variables in $\mathbf{X}$, and otherwise keeps $\mathcal{G}$ as it is. The resulting graph is denoted as $\mathcal{G}_{\underline{\mathbf{X}}}$.*
- *the* **X-upper-manipulation** *of $\mathcal{G}$ deletes all edges in $\mathcal{G}$ that are into variables in $\mathbf{X}$, and otherwise keeps $\mathcal{G}$ as it is. The resulting graph is denoted as $\mathcal{G}_{\overline{\mathbf{X}}}$.*

The following proposition summarizes Pearl's *do*-calculus. (Following Pearl, we use lower case letters to denote generic value settings for the sets of variables denoted by the corresponding upper case letters. So for simplicity we write $P(\mathbf{x})$ to mean $P(\mathbf{X}=\mathbf{x})$, and $do(\mathbf{x})$ to mean $do(\mathbf{X}=\mathbf{x})$.)

Proposition 7 (Pearl) *Let $\mathcal{G}$ be the causal DAG for $\mathbf{V}$, and $\mathbf{U},\mathbf{X},\mathbf{Y},\mathbf{W}$ be disjoint subsets of $\mathbf{V}$. The following rules are sound:*

1. *if $\mathbf{Y}$ and $\mathbf{X}$ are d-separated by $\mathbf{U}\cup\mathbf{W}$ in $\mathcal{G}_{\overline{\mathbf{U}}}$, then*

$$P(\mathbf{y}|do(\mathbf{u}),\mathbf{x},\mathbf{w}) = P(\mathbf{y}|do(\mathbf{u}),\mathbf{w}).$$

2. *if* $\mathbf{Y}$ *and* $\mathbf{X}$ *are d-separated by* $\mathbf{U} \cup \mathbf{W}$ *in* $\mathcal{G}_{\underline{\mathbf{X}}\overline{\mathbf{U}}}$, *then*

$$P(\mathbf{y}|do(\mathbf{u}), do(\mathbf{x}), \mathbf{w}) = P(\mathbf{y}|do(\mathbf{u}), \mathbf{x}, \mathbf{w}).$$

3. *if* $\mathbf{Y}$ *and* $\mathbf{X}$ *are d-separated by* $\mathbf{U} \cup \mathbf{W}$ *in* $\mathcal{G}_{\overline{\mathbf{U}\mathbf{X}'}}$, *then*

$$P(\mathbf{y}|do(\mathbf{u}), do(\mathbf{x}), \mathbf{w}) = P(\mathbf{y}|do(\mathbf{u}), \mathbf{w})$$

where $\mathbf{X}' = \mathbf{X} \backslash \mathbf{An}_{\mathcal{G}_{\overline{\mathbf{U}}}}(\mathbf{W}) = \mathbf{X} \backslash (\cup_{W \in \mathbf{W}} \mathbf{An}_{\mathcal{G}_{\overline{\mathbf{U}}}}(W))$.

The proposition follows from the intervention principle (Pearl, 1995). The first rule is actually not independent—it can be derived from the other two rules (Huang and Valtorta, 2006), but it has long been an official part of the calculus. The soundness of the calculus ensures that any post-intervention probability that can be reduced via the calculus to an expression that only involves pre-intervention probabilities of observed variables is identifiable. Recently, the completeness of the calculus was also established, in the sense that any identifiable post-intervention probability can be so reduced using the calculus (Huang and Valtorta, 2006; Shpister and Pearl, 2006).

Our goal is to develop a similar calculus when the available causal information is given in a PAG. A natural idea is to formulate analogous inference rules in terms of (manipulated) PAGs, to the effect that if a certain rule is applicable given a PAG, the corresponding rule in Pearl's calculus will be applicable given the (unknown) true causal DAG. How to guarantee that? Recall that a PAG represents an equivalence class of MAGs; each MAG, in turn, represents a set of causal DAGs. The union of all these sets is the set of DAGs represented by the PAG—one of them is the true causal DAG. So a sure way to get what we want is to formulate analogous rules in terms of PAGs such that if the rule is applicable given a PAG, then for every DAG represented by the PAG, the corresponding rule in Pearl's calculus is applicable.

For this purpose, it is natural to develop the desired calculus in two steps. First, we derive an analogous *do*-calculus based on MAGs, such that if a rule is applicable given a MAG, then for every DAG represented by the MAG, the corresponding rule in Pearl's calculus is applicable. Second, we extend that to a *do*-calculus based on PAGs, such that if a rule is applicable given a PAG, then for every MAG in the equivalence class represented by the PAG, the corresponding rule in the MAG-based calculus is applicable.

Before we define appropriate analogues of graph manipulations on MAGs, it is necessary to distinguish two kinds of directed edges in a MAG, according to the following criterion.

Definition 8 (Visibility) *Given a MAG* $\mathcal{M}$, *a directed edge* $A \to B$ *in* $\mathcal{M}$ *is* **visible** *if there is a vertex C not adjacent to B, such that either there is an edge between C and A that is into A, or there is a collider path between C and A that is into A and every vertex on the path is a parent of B. Otherwise* $A \to B$ *is said to be* **invisible**.

Figure 6 gives the possible configurations that make a directed edge $A \to B$ visible. The distinction between visible and invisible directed edges is important because of the following two facts.

Lemma 9 *Let* $\mathcal{G}$ *be a DAG over* $\mathbf{O} \cup \mathbf{L}$, *and* $\mathcal{M}$ *be the MAG over* $\mathbf{O}$ *that represents the DAG. For any* $A, B \in \mathbf{O}$, *if* $A \in \mathbf{An}_{\mathcal{G}}(B)$, *and there is an inducing path relative to* $\mathbf{L}$ *between A and B that is into A in* $\mathcal{G}$, *then there is a directed edge* $A \to B$ *in* $\mathcal{M}$ *that is invisible.*

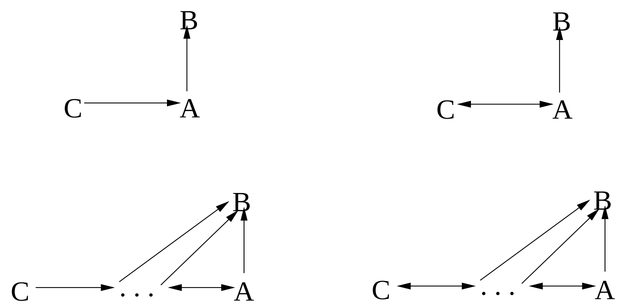

Figure 6: Possible configurations of visibility for $A \rightarrow B$.

Proof See Appendix B. ■

Taking the contrapositive of Lemma 9 gives us the fact that if $A \rightarrow B$ is visible in a MAG, then in *every* DAG represented by the MAG, there is no inducing path between A and B relative to the set of latent variables that is also into A. This implies that for every such DAG G, $G_{\underline{A}}$—the graph resulting from eliminating edges out of A in G—will not contain any inducing path between A and B relative to the set of latent variables, which means that the MAG that represents $G_{\underline{A}}$ will not contain any edge between A and B. So intuitively, deleting edges out of A in the underlying DAG corresponds to deleting visible arrows out of A in the MAG.

How about invisible arrows? Here is the relevant fact.

Lemma 10 *Let $\mathcal{M}$ be any MAG over a set of variables* $\mathbf{O}$, *and $A \rightarrow B$ be any directed edge in $\mathcal{M}$. If $A \rightarrow B$ is invisible in $\mathcal{M}$, then there is a DAG whose MAG is $\mathcal{M}$ in which A and B share a latent parent, that is, there is a latent variable L_{AB} in the DAG such that $A \leftarrow L_{AB} \rightarrow B$ is a subgraph of the DAG.*

Proof See Appendix B. ■

Obviously $A \leftarrow L_{AB} \rightarrow B$ is an inducing path between A and B relative to the set of latent variables. So if $A \rightarrow B$ in a MAG is invisible, at least for *some* DAG G represented by the MAG—and for all we know, this DAG may well be the true causal DAG—$G_{\underline{A}}$ contains $A \leftarrow L_{AB} \rightarrow B$, and hence corresponds to a MAG in which $A \leftrightarrow B$ appears.

Finally, for either $A \leftrightarrow B$ or $A \rightarrow B$ in a MAG, it is not hard to show that for *every* DAG represented by the MAG, there is no inducing path in the DAG between A and B relative to the set of latent variables that is also out of B (since otherwise B would be an ancestor of A, violating the definition of ancestral graphs). So deleting edges into B in the underlying DAG corresponds to deleting edges into B in the MAG. These considerations motivate the following definition.

Definition 11 (Manipulations of MAGs) *Given a MAG $\mathcal{M}$ and a set of variables* $\mathbf{X}$ *therein,*

- *the* **X-lower-manipulation** *of $\mathcal{M}$ deletes all those edges that are visible in $\mathcal{M}$ and are out of variables in* $\mathbf{X}$, *replaces all those edges that are out of variables in* $\mathbf{X}$ *but are invisible in $\mathcal{M}$ with bi-directed edges, and otherwise keeps $\mathcal{M}$ as it is. The resulting graph is denoted as $\mathcal{M}_{\underline{\mathbf{X}}}$.*

- *the* **X-upper-manipulation** *of $\mathcal{M}$ deletes all those edges in $\mathcal{M}$ that are into variables in* $\mathbf{X}$, *and otherwise keeps $\mathcal{M}$ as it is. The resulting graph is denoted as $\mathcal{M}_{\overline{\mathbf{X}}}$.*

We stipulate that lower-manipulation has a higher priority than upper-manipulation, so that $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ (or $\mathcal{M}_{\overline{\mathbf{X}}\underline{\mathbf{Y}}}$) denotes the graph resulting from applying the $\mathbf{X}$*-upper-manipulation to the* $\mathbf{Y}$*-lower-manipulated graph of $\mathcal{M}$.*

A couple of comments are in order. First, unlike the case of DAGs, the lower-manipulation for MAGs may introduce new edges, that is, replacing invisible directed edges with bi-directed edges. Again, the reason we do this is that an invisible directed edge from A to B allows the possibility of a latent common parent of A and B in the underlying DAG. If so, the A-lower-manipulated DAG will correspond to a MAG in which there is a bi-directed edge between A and B. Second, because of the possibility of introducing new bi-directed edges, we need the priority stipulation that lower-manipulation is to be done before upper-manipulation. The stipulation is not necessary for DAGs, because no new edges would be introduced in the lower-manipulation of DAGs, and hence the order does not matter.

Ideally, if $\mathcal{M}$ is the MAG of a DAG $\mathcal{G}$, we would like $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ to be the MAG of $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. But this is not always possible, as two DAGs represented by the same MAG before a manipulation may correspond to different MAGs after the manipulation. But we still have the following fact:

Lemma 12 *Let $\mathcal{G}$ be a DAG over $\mathbf{O} \cup \mathbf{L}$, and $\mathcal{M}$ be the MAG of $\mathcal{G}$ over $\mathbf{O}$. Let $\mathbf{X}$ and $\mathbf{Y}$ be two possibly empty subsets of $\mathbf{O}$, and $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$ be the MAG of $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. For any $A, B \in \mathbf{O}$ and $\mathbf{C} \subseteq \mathbf{O}$ that does not contain A or B, if there is an m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$, then there is an m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$.*

Proof See Appendix B. ■

Recall that a graphical model is called an *independence map* of another if any independence implied by the former is also implied by the latter (Chickering, 2002). So another way of putting Lemma 12 is that $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ is an independence map of $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$, which we write as $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}} \leq \mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. The diagram in Figure 7 visualizes what is going on.

$$
\begin{tikzcd}
\mathcal{G} \arrow[rrr, "mc"] \arrow[d, "gm"'] & & & \mathcal{M} \arrow[d, "mm"] \\
\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}} \arrow[r, "mc"'] & \mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}} & \leq & \mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}
\end{tikzcd}
$$

Figure 7: Illustration of Lemma 12: *mc* refers to MAG construction introduced in Section 3; *gm* refers to DAG manipulation; and *mm* refers to MAG manipulation.

Corollary 13 *Let $\mathcal{M}$ be a MAG over* $\mathbf{O}$*, and* $\mathbf{X}$ *and* $\mathbf{Y}$ *be two subsets of* $\mathbf{O}$*. For any $A, B \in \mathbf{O}$ and $\mathbf{C} \subseteq \mathbf{O}$ that does not contain A or B, if A and B are m-separated by $\mathbf{C}$ in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, then A and B are d-separated by $\mathbf{C}$ in $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ for every $\mathcal{G}$ represented by $\mathcal{M}$.*

Proof By Lemma 12, if A and B are m-separated by $\mathbf{C}$ in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, they are also m-separated by $\mathbf{C}$ in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$, for every $\mathcal{G}$ represented by $\mathcal{M}$, which in turn implies that A and B are d-separated by $\mathbf{C}$ in $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ for every $\mathcal{G}$ represented by $\mathcal{M}$, because d-separation relations among $\mathbf{O}$ in a DAG correspond exactly to m-separation relations in its MAG. ■

The converse of Corollary 13, however, is not true in general. To give the simplest example, consider the MAG $\mathcal{M}$ in Figure 8(a): $X \leftarrow Y \rightarrow Z$ (which happens to be a DAG syntactically). The two DAGs, $\mathcal{G}1$ in 8(b) and $\mathcal{G}2$ in 8(c), are both represented by $\mathcal{M}$. By the definition of lower-manipulation, $\mathcal{M}_{\underline{Y}}$ is the graph $X \leftrightarrow Y \leftrightarrow Z$. On the other hand, $\mathcal{G}1_{\underline{Y}}$ is $X \leftarrow L1 \rightarrow Y \quad Z$; and $\mathcal{G}2_{\underline{Y}}$ is $X \quad Y \leftarrow L2 \rightarrow Z$. Obviously, the MAG of $\mathcal{G}1_{\underline{Y}}$ is $X \leftrightarrow Y \quad Z$, and the MAG of $\mathcal{G}2_{\underline{Y}}$ is $X \quad Y \leftrightarrow Z$, both of which are *proper* subgraphs of $\mathcal{M}_{\underline{Y}}$. So an m-separation relation in $\mathcal{M}_{\underline{Y}}$—for example, X and Z are m-separated by the empty set—corresponds to a d-separation relation in both $\mathcal{G}1_{\underline{Y}}$ and $\mathcal{G}2_{\underline{Y}}$, in accord with Corollary 13.

By contrast, the converse of Corollary 13 fails for $\mathcal{M}$. It can be checked that for every $\mathcal{G}$ represented by $\mathcal{M}$, X and Z are d-separated by Y in $\mathcal{G}_{\underline{Y}}$, as evidenced by $\mathcal{G}1_{\underline{Y}}$ and $\mathcal{G}2_{\underline{Y}}$. But X and Z are not m-separated by Y in $\mathcal{M}_{\underline{Y}}$.

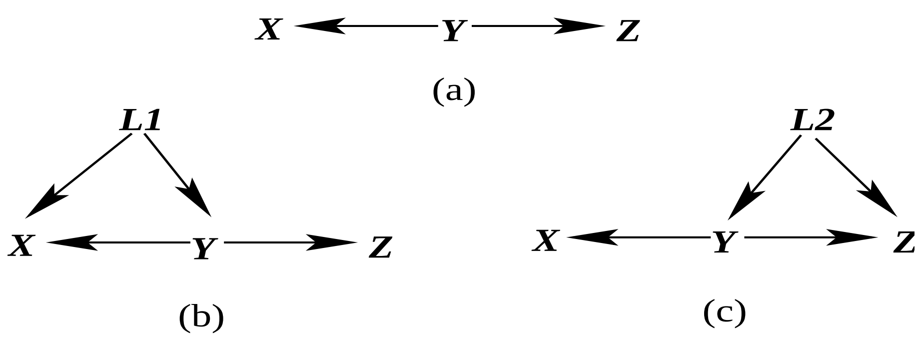

Figure 8: A counterexample to the converse of Corollary 13.

However, Definition 11 is not to be blamed for this limitation. In this simple example, one can easily enumerate all possible directed mixed graphs over X, Y, Z and see that for none of them do both Corollary 13 and its converse hold. Intuitively, this is because the MAG in Figure 8(a) implies that either $\langle X, Y \rangle$ does not have a common latent parent or $\langle Y, Z \rangle$ does not have a common latent parent in the underlying DAG. So under the Y-lower-manipulation of the underlying DAG, for all we know, either $\langle X, Y \rangle$ or $\langle Y, Z \rangle$ will become unconnected. But this disjunctive information cannot be precisely represented by a single graph.

More generally, no matter how we define $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, as long as it is a single graph, the converse of Corollary 13 will not hold in general, unless Corollary 13 itself fails. $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, as a single graph, can only aim to be a supergraph (up to Markov equivalence) of $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$ for every $\mathcal{G}$ represented by $\mathcal{M}$ (which makes Corollary 13 true). To this end, Definition 11 is 'minimal' in the following sense: two variables are adjacent in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ if and only if there exists a DAG $\mathcal{G}$ represented by $\mathcal{M}$ such that the two variables are adjacent in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$. In this regard, $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ does not have more edges than necessary. One can, for example, check this fact for the simple case in Figure 8.

We are now ready to state the intermediate theorem on MAG-based *do*-calculus.

Theorem 14 (*do*-calculus given a MAG) *Let $\mathcal{M}$ be the causal MAG over $\mathbf{O}$, and $\mathbf{U}, \mathbf{X}$, $\mathbf{Y}$, $\mathbf{W}$ be disjoint subsets of $\mathbf{O}$. The following rules are valid, in the sense that if the antecedent of the rule holds, then the consequent holds no matter which DAG represented by $\mathcal{M}$ is the true causal DAG.*

1. *if $\mathbf{Y}$ and $\mathbf{X}$ are m-separated by $\mathbf{U} \cup \mathbf{W}$ in $\mathcal{M}_{\overline{\mathbf{U}}}$, then*
$$P(\mathbf{y}|do(\mathbf{u}), \mathbf{x}, \mathbf{w}) = P(\mathbf{y}|do(\mathbf{u}), \mathbf{w}).$$
2. *if $\mathbf{Y}$ and $\mathbf{X}$ are m-separated by $\mathbf{U} \cup \mathbf{W}$ in $\mathcal{M}_{\underline{\mathbf{X}}\overline{\mathbf{U}}}$, then*
$$P(\mathbf{y}|do(\mathbf{u}), do(\mathbf{x}), \mathbf{w}) = P(\mathbf{y}|do(\mathbf{u}), \mathbf{x}, \mathbf{w}).$$
3. *if $\mathbf{Y}$ and $\mathbf{X}$ are m-separated by $\mathbf{U} \cup \mathbf{W}$ in $\mathcal{M}_{\overline{\mathbf{U}\mathbf{X}'}}$, then*
$$P(\mathbf{y}|do(\mathbf{u}), do(\mathbf{x}), \mathbf{w}) = P(\mathbf{y}|do(\mathbf{u}), \mathbf{w})$$
where $\mathbf{X}' = \mathbf{X} \backslash \mathbf{An}_{\mathcal{M}_{\overline{\mathbf{U}}}}(\mathbf{W})$.

Proof This readily follows from Proposition 7, Corollary 13, and the fact that for every $\mathcal{G}$ represented by $\mathcal{M}$, $\mathbf{An}_{\mathcal{G}_{\overline{\mathbf{U}}}}(\mathbf{W}) \cap \mathbf{O} = \mathbf{An}_{\mathcal{M}_{\overline{\mathbf{U}}}}(\mathbf{W})$. ■

As already noted, the true causal MAG is not uniquely recoverable from the pre-intervention distribution, thanks to Markov equivalence. So the main value of Theorem 14 is to facilitate the development of a PAG-based *do*-calculus. However, it is worth noting that when supplemented with some background causal knowledge, such as knowledge of the form that some variable is not a cause of another variable, it is in principle possible to determine that the true causal MAG belongs to a proper subset of the full equivalence class represented by the PAG. Depending on how strong the background knowledge is, the subset could be as big as the full equivalence class or as small as a singleton. In this sense, Theorem 14 and Theorem 17 below may be viewed as two extreme cases of a more general *do*-calculus based on a subset of Markov equivalent MAGs.

To extend the calculus to PAGs, we need to define manipulations on PAGs. They are essentially the same as the manipulations of MAGs. The definition of visibility still makes sense in PAGs, except that we will call a directed edge in a PAG *definitely visible* if it satisfies the condition for visibility in Definition 8, in order to emphasize that this edge is visible in all MAGs in the equivalence class. Despite the extreme similarity to manipulations on MAGs, let us still write down the definition of PAG manipulations for easy reference.

Definition 15 (Manipulations of PAGs) *Given a PAG $\mathcal{P}$ and a set of variables $\mathbf{X}$ therein,*

- *the* **X-lower-manipulation** *of $\mathcal{P}$ deletes all those edges that are definitely visible in $\mathcal{P}$ and are out of variables in $\mathbf{X}$, replaces all those edges that are out of variables in $\mathbf{X}$ but are not definitely visible in $\mathcal{P}$ with bi-directed edges, and otherwise keeps $\mathcal{P}$ as it is. The resulting graph is denoted as $\mathcal{P}_{\underline{\mathbf{X}}}$.*

- *the* **X-upper-manipulation** *of $\mathcal{P}$ deletes all those edges in $\mathcal{P}$ that are into variables in* $\mathbf{X}$*, and otherwise keeps $\mathcal{P}$ as it is. The resulting graph is denoted as $\mathcal{P}_{\overline{\mathbf{X}}}$.*

We stipulate that lower-manipulation has a higher priority than upper-manipulation, so that $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ (or $\mathcal{P}_{\overline{\mathbf{X}}\underline{\mathbf{Y}}}$) denotes the graph resulting from applying the $\mathbf{X}$*-upper-manipulation to the* $\mathbf{Y}$*-lower-manipulated graph of $\mathcal{P}$.*

We should emphasize that except in rare situations, $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ is not a PAG any more (i.e., not a PAG for any Markov equivalence class of MAGs). But from $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ we still gain information about m-separation in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, where $\mathcal{M}$ is a MAG that belongs to the Markov equivalence class represented by $\mathcal{P}$. Here is a simple connection. Given a MAG $\mathcal{M}$ and the PAG $\mathcal{P}$ that represents $[\mathcal{M}]$, a trivial fact is that a m-connecting path in $\mathcal{M}$ is also a possibly m-connecting path in $\mathcal{P}$. This is also true for $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ and $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$.

Lemma 16 *Let $\mathcal{M}$ be a MAG over* $\mathbf{O}$*, and $\mathcal{P}$ be the PAG for $[\mathcal{M}]$. Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be two subsets of* $\mathbf{O}$*. For any $A, B \in \mathbf{O}$ and $\mathbf{C} \subseteq \mathbf{O}$ that does not contain A or B, if a path p between A and B is m-connecting given* $\mathbf{C}$ *in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, then p, the same sequence of variables, forms a possibly m-connecting path between A and B given* $\mathbf{C}$ *in $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$.*[15]

Proof See Appendix B. ■

If there is no possibly m-connecting path between A and B given $\mathbf{C}$ in a partial mixed graph, we say that A and B are *definitely m-separated* by $\mathbf{C}$ in the graph. A *do*-calculus follows:

Theorem 17 (*do*-calculus given a PAG) *Let $\mathcal{P}$ be the causal PAG for* $\mathbf{O}$*, and $\mathbf{U}, \mathbf{X}, \mathbf{Y}, \mathbf{W}$ be disjoint subsets of* $\mathbf{O}$*. The following rules are valid:*

1. *if* $\mathbf{Y}$ *and* $\mathbf{X}$ *are definitely m-separated by $\mathbf{U} \cup \mathbf{W}$ in $\mathcal{P}_{\overline{\mathbf{U}}}$, then*

$$P(\mathbf{y}|do(\mathbf{u}), \mathbf{x}, \mathbf{w}) = P(\mathbf{y}|do(\mathbf{u}), \mathbf{w}).$$

2. *if* $\mathbf{Y}$ *and* $\mathbf{X}$ *are definitely m-separated by $\mathbf{U} \cup \mathbf{W}$ in $\mathcal{P}_{\underline{\mathbf{X}}\overline{\mathbf{U}}}$, then*

$$P(\mathbf{y}|do(\mathbf{u}), do(\mathbf{x}), \mathbf{w}) = P(\mathbf{y}|do(\mathbf{u}), \mathbf{x}, \mathbf{w}).$$

3. *if* $\mathbf{Y}$ *and* $\mathbf{X}$ *are definitely m-separated by $\mathbf{U} \cup \mathbf{W}$ in $\mathcal{P}_{\overline{\mathbf{U}\mathbf{X}'}}$, then*

$$P(\mathbf{y}|do(\mathbf{u}), do(\mathbf{x}), \mathbf{w}) = P(\mathbf{y}|do(\mathbf{u}), \mathbf{w})$$

   where $\mathbf{X}' = \mathbf{X} \backslash \mathbf{PossibleAn}_{\mathcal{P}_{\overline{\mathbf{U}}}}(\mathbf{W})$.

15. For our purpose, what we need is the obvious consequence of the lemma that if there is an m-connecting path in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, then there is a possibly m-connecting path in $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. We suspect that a stronger result might hold as well: if there is an m-connecting path in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, then there is a definite m-connecting path in $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. We can't prove or disprove the stronger result at the moment.

Proof It follows from Lemma 16 and Theorem 14. The only caveat is that in general $\mathbf{An}_{\mathcal{M}_{\overline{\mathbf{U}}}}(\mathbf{W}) \neq \mathbf{PossibleAn}_{\mathcal{P}_{\overline{\mathbf{U}}}}(\mathbf{W})$ for an arbitrary $\mathcal{M}$ represented by $\mathcal{P}$. But it is always the case that $\mathbf{An}_{\mathcal{M}_{\overline{\mathbf{U}}}}(\mathbf{W}) \subseteq \mathbf{PossibleAn}_{\mathcal{P}_{\overline{\mathbf{U}}}}(\mathbf{W})$, which means that $\mathbf{X}\backslash\mathbf{An}_{\mathcal{M}_{\overline{\mathbf{U}}}}(\mathbf{W}) \supseteq \mathbf{X}\backslash\mathbf{PossibleAn}_{\mathcal{P}_{\overline{\mathbf{U}}}}(\mathbf{W})$ for every $\mathcal{M}$ represented by $\mathcal{P}$. So it is possible that for rule (3), $\mathcal{P}_{\overline{\mathbf{UX}'}}$ leaves more edges in than necessary, but it does not affect the validity of rule (3). ■

The possibility that $\mathcal{P}_{\overline{\mathbf{UX}'}}$ leaves more edges in than necessary is one of three aspects in which our *do*-calculus may be "incomplete" in the following sense: it is possible that a rule in the PAG-based *do*-calculus is not applicable, but for every DAG compatible with the given PAG, the corresponding rule in Pearl's DAG-based calculus is applicable. The other two aspects are already noted: (1) the calculus is formulated in terms of the absence of possibly m-connecting paths (cf. Footnote 14, and more on this in the next section); and (2) the MAG-based *do*-calculus is based on Corollary 13 whose converse does not hold. Therefore, the PAG-based *do*-calculus as currently formulated may be further improved.

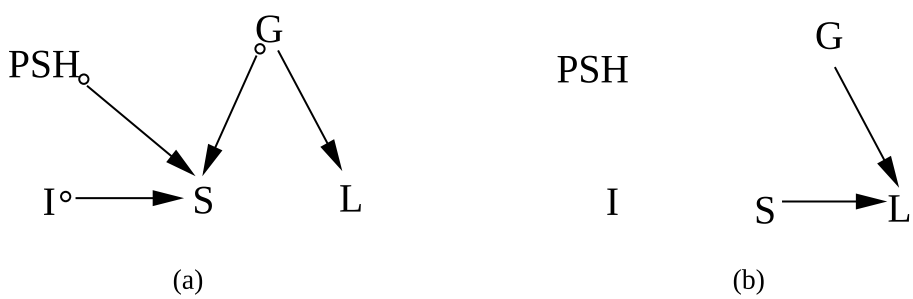

Figure 9: PAG Surgery: $\mathcal{P}_{\underline{S}}$ and $\mathcal{P}_{\overline{S}}$.

That said, let us illustrate the utility of the *do*-calculus with the simple example used in Section 3. Given the PAG in Figure 4 we can infer that $P(L|do(S),G) = P(L|S,G)$ by rule 2, because L and S are definitely m-separated by $\{G\}$ in $\mathcal{P}_{\underline{S}}$ (Figure 9(a)); and $P(G|do(S)) = P(G)$ by rule 3, because G and S are definitely m-separated in $\mathcal{P}_{\overline{S}}$ (Figure 9(b)). It follows that

$$\begin{aligned} P(L|do(S)) &= \sum_G P(L,G|do(S)) \\ &= \sum_G P(L|do(S),G)P(G|do(S)) \\ &= \sum_G P(L|S,G)P(G). \end{aligned}$$

By contrast, it is not valid in the *do*-calculus that $P(L|do(G),S) = P(L|G,S)$ because L and G are not definitely m-separated by $\{S\}$ in $\mathcal{P}_{\underline{G}}$, which is depicted in Figure 10. (Notice the bi-directed edge between L and G.)

5. Invariance Under Interventions

We now develop stronger results for a key component of *do*-calculus, the property of *invariance under interventions*, first systematically studied in Spirtes et al. (1993). The idea is simple. A

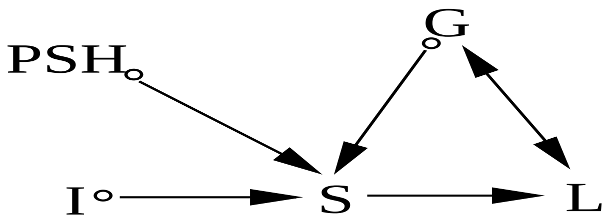

Figure 10: PAG Surgery: $\mathcal{P}_{\underline{G}}$.

conditional probability $P(\mathbf{Y}=\mathbf{y}|\mathbf{Z}=\mathbf{z})$ is said to be *invariant* under an intervention $\mathbf{X}:=\mathbf{x}$—or $do(\mathbf{X}=\mathbf{x})$—if $P_{\mathbf{X}:=\mathbf{x}}(\mathbf{y}|\mathbf{z})=P(\mathbf{y}|\mathbf{z})$.[16] This concept (under the name of 'observability') plays an important role in some interesting theoretical work on observational studies (e.g., Pratt and Schlaifer, 1988; for a good review see Winship and Morgan, 1999), and also forms the basis of the prediction algorithm presented in Spirtes et al. (1993), which seeks to identify a post-intervention probability by searching for an expression in terms of invariant probabilities.

It is also the corner stone of Pearl's *do*-calculus. To see this, let us take a closer look at the second and third rules in the *do*-calculus. The second rule of the calculus gives a graphical condition for when we can conclude

$$P(\mathbf{y}|do(\mathbf{u}),do(\mathbf{x}),\mathbf{w})=P(\mathbf{y}|do(\mathbf{u}),\mathbf{x},\mathbf{w}).$$

If we take $\mathbf{U}$ to be the empty set and write the above equation in the subscript notation, we get

$$P_{\mathbf{X}:=\mathbf{x}}(\mathbf{y}|\mathbf{w})=P(\mathbf{y}|\mathbf{x},\mathbf{w}).$$

Since $P_{\mathbf{X}:=\mathbf{x}}(\mathbf{X}=\mathbf{x})=1$, thanks to the supposed effectiveness of the intervention, we have

$$P_{\mathbf{X}:=\mathbf{x}}(\mathbf{y}|\mathbf{w})=P_{\mathbf{X}:=\mathbf{x}}(\mathbf{y}|\mathbf{x},\mathbf{w}).$$

So a special case of the second rule is a condition for $P_{\mathbf{X}:=\mathbf{x}}(\mathbf{y}|\mathbf{x},\mathbf{w})=P(\mathbf{y}|\mathbf{x},\mathbf{w})$, that is, for when $P(\mathbf{y}|\mathbf{x},\mathbf{w})$ is invariant under the intervention $\mathbf{X}:=\mathbf{x}$. In fact, the second rule is nothing but a generalization of this condition to tell when a post-intervention probability $P_{\mathbf{u}}(\mathbf{y}|\mathbf{x},\mathbf{w})$ would be invariant under a *further* intervention $\mathbf{X}:=\mathbf{x}$.

The third rule is more obviously about invariance. It is a generalization of the condition for $P_{\mathbf{X}:=\mathbf{x}}(\mathbf{y}|\mathbf{w})=P(\mathbf{y}|\mathbf{w})$, that is, for when $P(\mathbf{y}|\mathbf{w})$ is invariant under the intervention $\mathbf{X}:=\mathbf{x}$. The difference between rule 2 and rule 3 is that rule 2 is about invariance of $P(\mathbf{y}|\mathbf{z})$ under an intervention on $\mathbf{X}$ in case $\mathbf{X}\subseteq\mathbf{Z}$ $(=\mathbf{X}\cup\mathbf{W})$, whereas rule 3 is about invariance of $P(\mathbf{y}|\mathbf{z})$ under an intervention on $\mathbf{X}$ in case $\mathbf{X}$ and $\mathbf{Z}$ $(=\mathbf{W})$ are disjoint. As we mentioned earlier, the first rule is not essential, so the *do*-calculus is in effect a generalization of conditions for invariance.

We now focus on this key component of *do*-calculus, and present better graphical conditions for judging invariance given a PAG than those that are implied by the PAG-based *do*-calculus presented in the last section. The conditions for invariance implied by Pearl's (DAG-based) *do*-calculus can

16. Here we allow that $\mathbf{X}$ and $\mathbf{Z}$ have a non-empty intersection, and assume that the conditioning operation is applied to the post-intervention population (i.e., intervening comes before conditioning). As a result, when we speak of $P_{\mathbf{X}:=\mathbf{x}}(\mathbf{y}|\mathbf{z})$, we implicitly assume that $\mathbf{x}$ and $\mathbf{z}$ are consistent regarding the values for variables in $\mathbf{X}\cap\mathbf{Z}$, for otherwise the quantity is undefined.

be equivalently formulated without referring to manipulated graphs, as given in Spirtes et al. (1993, Theorem 7.1) before the *do*-calculus was invented. In this section we develop corresponding conditions in terms of PAGs. The conditions will be not only sufficient in the sense that if the conditions are satisfied, then every DAG compatible with the given PAG entails invariance, but also necessary in the sense that if the conditions fail, then there is at least one DAG compatible with the given PAG that does not entail invariance. In this aspect, the conditions are also superior to earlier results on invariance given an equivalence class of DAGs due to Spirtes et al. (1993, Theorems 7.3 and 7.4).

We first state the conditions for judging invariance given a DAG, originally presented in Spirtes et al. (1993, Theorem 7.1).

Proposition 18 (Spirtes, Glymour, Scheines) *Let $\mathcal{G}$ be the causal DAG for $\mathbf{O}\cup\mathbf{L}$, and $\mathbf{X},\mathbf{Y},\mathbf{Z}\subseteq\mathbf{O}$ be three sets of variables such that $\mathbf{X}\cap\mathbf{Y}=\mathbf{Y}\cap\mathbf{Z}=\varnothing$ (but $\mathbf{X}$ and $\mathbf{Z}$ can overlap). $P(\mathbf{y}|\mathbf{z})$ is invariant under an intervention on $\mathbf{X}$ if*

(1) for every $X\in\mathbf{X}\cap\mathbf{Z}$, there is no d-connecting path between X and any member of $\mathbf{Y}$ given $\mathbf{Z}\backslash\{X\}$ that is into X;

(2) for every $X\in\mathbf{X}\cap(\mathbf{An}_{\mathcal{G}}(\mathbf{Z})\backslash\mathbf{Z})$, there is no d-connecting path between X and any member of $\mathbf{Y}$ given $\mathbf{Z}$; and

(3) for every $X\in\mathbf{X}\backslash\mathbf{An}_{\mathcal{G}}(\mathbf{Z})$, there is no d-connecting path between X and any member of $\mathbf{Y}$ given $\mathbf{Z}$ that is out of X.[17]

Remark: Because $\mathbf{Z}\subseteq\mathbf{An}_{\mathcal{G}}(\mathbf{Z})$, $\mathbf{X}\cap\mathbf{Z},\mathbf{X}\cap(\mathbf{An}_{\mathcal{G}}(\mathbf{Z})\backslash\mathbf{Z})$ and $\mathbf{X}\backslash\mathbf{An}_{\mathcal{G}}(\mathbf{Z})$ form a partition of $\mathbf{X}$. So for each member of $\mathbf{X}$, only one of the conditions is relevant.

The proposition is an equivalent formulation of Theorem 7.1 in Spirtes et al. (1993). It is not hard to check that the proposition follows from rules 2 and 3 in the DAG-based *do*-calculus (Proposition 7); the talk of d-separation in manipulated graphs is replaced by the talk of absence of d-connecting paths of certain orientations in the original graph. Conversely, the proposition implies the special case of rules 2 and 3 where the background intervention $do(\mathbf{U})$ is empty. Specifically, clause (1) in the proposition corresponds to rule 2 in the *do*-calculus; clauses (2) and (3) correspond to rule 3 in the *do*-calculus.

Spirtes et al. (1993, pp. 164-5) argued that these conditions are also "almost necessary" for invariance. What they meant is that if the conditions are not satisfied, then the causal structure does not *entail* the invariance, although there may exist some particular distribution compatible with the causal structure such that $P(\mathbf{y}|\mathbf{z})$ is invariant under some particular intervention on $\mathbf{X}$. From now on when we speak of invariance entailed by the causal DAG, we mean that the conditions in Proposition 18 are satisfied—or equivalently, that the invariance follows from an application of rule 2 or rule 3 in the DAG-based *do*-calculus.[18] Our purpose is to demonstrate that there are corresponding graphical

17. It is not hard to see that (3) is equivalent to saying that for every $X\in\mathbf{X}\backslash\mathbf{An}_{\mathcal{G}}(\mathbf{Z})$, there is no directed path from X to any member of $\mathbf{Y}$. Lemma 23 below is an immediate corollary of this equivalent formulation.

18. This stipulation is of course not intended to be a definition of the notion of *structurally entailed invariance*. A proper definition would be to the effect that for every distribution compatible with the causal structure, $P(\mathbf{y}|\mathbf{z})$ is invariant under any intervention of $\mathbf{X}$. The argument given by Spirtes et al. (1993, pp. 164-5) for (their equivalent formulation of) Proposition 18 suggests that the conditions are sufficient and necessary for structurally entailed invariance. Their argument uses the device of what they call policy variables, extra variables introduced into the

conditions relative to a PAG that are sufficient and necessary for the conditions in Proposition 18 to hold for each and every DAG compatible with the PAG.

Once again, we develop the conditions in two steps: first to MAGs and then to PAGs. In the first step, our goal is to find sufficient and necessary conditions for invariance entailed by a MAG, as defined below:

Definition 19 (Invariance entailed by a MAG) *Let $\mathcal{M}$ be a causal MAG over* $\mathbf{O}$*, and* $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z} \subseteq \mathbf{O}$ *be three sets of variables such that* $\mathbf{X} \cap \mathbf{Y} = \mathbf{Y} \cap \mathbf{Z} = \varnothing$*,* $P(\mathbf{y}|\mathbf{z})$ *is* **entailed to be invariant under interventions on X given** $\mathcal{M}$ *if for every DAG* $\mathcal{G}(\mathbf{O}, \mathbf{L})$ *represented by* $\mathcal{M}$*,* $P(\mathbf{y}|\mathbf{z})$ *is entailed to be invariant under interventions on* $\mathbf{X}$ *given* $\mathcal{G}$ *(i.e., the conditions in Proposition 18 are satisfied).*

The question is how to judge invariance entailed by a MAG without doing the intractable job of checking the conditions in Proposition 18 for each and every compatible DAG. The next few lemmas, Lemmas 20-23, state useful connections between d-connecting paths in a DAG and m-connecting paths in the corresponding MAG. Lemma 20 records the important result due to Richardson and Spirtes (2002) that d-separation relations among observed variables in a DAG with latent variables correspond exactly to m-separation relations in its MAG.

Lemma 20 *Let $\mathcal{G}$ be any DAG over* $\mathbf{O} \cup \mathbf{L}$*, and $\mathcal{M}$ be the MAG of $\mathcal{G}$ over* $\mathbf{O}$*. For any $A, B \in$* $\mathbf{O}$ *and* $\mathbf{C} \subseteq \mathbf{O}$ *that does not contain A or B, there is a path d-connecting A and B given* $\mathbf{C}$ *in $\mathcal{G}$ if and only if there is a path m-connecting A and B given* $\mathbf{C}$ *in $\mathcal{M}$.*

Proof This is a special case of Lemma 17 and Lemma 18 in Spirtes and Richardson (1996), and also a special case of Theorem 4.18 in Richardson and Spirtes (2002). ■

Given Lemma 20, we know how to tell whether clause (2) of Proposition 18 holds in all DAGs compatible with a given MAG. For the other two conditions in Proposition 18, we need to take into account the orientations of d-connecting paths.

Lemma 21 *Let $\mathcal{G}$ be any DAG over* $\mathbf{O} \cup \mathbf{L}$*, and $\mathcal{M}$ be the MAG of $\mathcal{G}$ over* $\mathbf{O}$*. For any $A, B \in$* $\mathbf{O}$ *and* $\mathbf{C} \subseteq \mathbf{O}$ *that does not contain A or B, if there is a path d-connecting A and B given* $\mathbf{C}$ *in $\mathcal{G}$ that is into A, then there is a path m-connecting A and B given* $\mathbf{C}$ *in $\mathcal{M}$ that is either into A or contains an invisible edge out of A.*

Proof See Appendix B. ■

Lemma 22 *Let $\mathcal{M}$ be any MAG over* $\mathbf{O}$*. For any $A, B \in$* $\mathbf{O}$ *and* $\mathbf{C} \subseteq \mathbf{O}$ *that does not contain A or B, if there is a path m-connecting A and B given* $\mathbf{C}$ *in $\mathcal{M}$ that is either into A or contains an invisible edge out of A, then there exists a DAG $\mathcal{G}$ over* $\mathbf{O} \cup \mathbf{L}$ *(for some extra variables* $\mathbf{L}$*) whose MAG is $\mathcal{M}$, such that in $\mathcal{G}$ there is a path d-connecting A and B given* $\mathbf{C}$ *that is into A.*

causal DAG to represent interventions. Given the causal DAG $\mathcal{G}$, a policy variable for a variable X is an (extra) parent of X but otherwise not adjacent to any other variables in $\mathcal{G}$. Interventions can then be simulated by conditioning on the intervention variables, and invariance can be reformulated as conditional independence involving intervention variables. The conditions in Proposition 18 are equivalent to saying that the variables in $\mathbf{Y}$ are d-separated from the policy variables for $\mathbf{X}$ by $\mathbf{Z}$ (in the graph augmented by the policy variables). It thus seems plausible that these conditions are sufficient and necessary for structurally entailed invariance, given that d-separation is a sufficient and necessary condition for structurally entailed conditional independence (Geiger et al., 1990; Meek, 1995b). But Spirtes et al. did not give a rigorous proof for necessity. As an anonymous reviewer points out, the rigorous proof, if any, would need to be carefully made, and in particular, one should be careful in treating policy variables as random variables. We will not take on this task here.

Proof See Appendix B. ■

Obviously these two lemmas are related to adapting clause (1) in Proposition 18 to MAGs. The next lemma is related to clause (3).

Lemma 23 *Let $\mathcal{G}$ be any DAG over $\mathbf{O} \cup \mathbf{L}$, and $\mathcal{M}$ be the MAG of $\mathcal{G}$ over $\mathbf{O}$. For any $A, B \in \mathbf{O}$ and $\mathbf{C} \subseteq \mathbf{O}$ that does not contain B or any descendant of A in $\mathcal{G}$ (or in $\mathcal{M}$, since $\mathcal{G}$ and $\mathcal{M}$ have the same ancestral relations among variables in $\mathbf{O}$), there is a path d-connecting A and B given $\mathbf{C}$ in $\mathcal{G}$ that is out of A if and only if there is a path m-connecting A and B given $\mathbf{C}$ in $\mathcal{M}$ that is out of A.*

Proof See Appendix B. ■

Given these lemmas, the conditions in Proposition 18 are readily translated into the following conditions for invariance given a MAG.

Theorem 24 *Suppose $\mathcal{M}$ is the causal MAG over a set of variables $\mathbf{O}$. For any $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \subseteq \mathbf{O}$, $\mathbf{X} \cap \mathbf{Y} = \mathbf{Y} \cap \mathbf{Z} = \emptyset$, $P(\mathbf{y}|\mathbf{z})$ is entailed to be invariant under interventions on $\mathbf{X}$ given $\mathcal{M}$ if and only if*

(1) for every $X \in \mathbf{X} \cap \mathbf{Z}$, there is no m-connecting path between X and any member of $\mathbf{Y}$ given $\mathbf{Z} \backslash \{X\}$ that is into X or contains an invisible edge out of X;

(2) for every $X \in \mathbf{X} \cap (\mathbf{An}_{\mathcal{M}}(\mathbf{Z}) \backslash \mathbf{Z})$, there is no m-connecting path between X and any member of $\mathbf{Y}$ given $\mathbf{Z}$; and

(3) for every $X \in \mathbf{X} \backslash \mathbf{An}_{\mathcal{M}}(\mathbf{Z})$, there is no m-connecting path between X and any member of $\mathbf{Y}$ given $\mathbf{Z}$ that is out of X.

Proof Given Lemma 21, if (1) holds, then for every DAG represented by $\mathcal{M}$, the first condition in Proposition 18 holds. Given Lemma 20 and the fact that $\mathcal{M}$ and all DAGs represented by $\mathcal{M}$ have the exact same ancestral relations among $\mathbf{O}$, if (2) holds, the second condition in Proposition 18 holds for every DAG represented by $\mathcal{M}$. Moreover, given Lemma 23, if (3) holds, the third condition in Proposition 18 holds for every DAG represented by $\mathcal{M}$. So (1), (2) and (3) together imply that $P(\mathbf{y}|\mathbf{z})$ is invariant under interventions on $\mathbf{X}$ given $\mathcal{M}$.

Conversely, if (1) fails, then by Lemma 22, there is a DAG represented by $\mathcal{M}$ in which the first condition in Proposition 18 fails. Likewise with conditions (2) and (3), in light of Lemmas 20 and 23 and the fact that $\mathcal{M}$ and a DAG represented by $\mathcal{M}$ have the exact same ancestral relations among $\mathbf{O}$. So (1), (2) and (3) are also necessary for $P(\mathbf{y}|\mathbf{z})$ to be entailed to be invariant under interventions on $\mathbf{X}$ given $\mathcal{M}$. ■

For example, given the MAG in Figure 3(a), $P(L|G,S)$ is invariant under interventions on S, because the only m-connecting path between L and S given G is $\langle L, S \rangle$, which contains a visible directed edge out of L, and so the relevant clause in Theorem 24, clause (1), is satisfied. By contrast, $P(L|G,S)$ is not entailed to be invariant under interventions on G given the MAG—in the sense that there exists a causal DAG compatible with the MAG given which $P(L|G,S)$ is not entailed to be invariant under interventions on G—because clause (1) is not satisfied.

In a similar fashion, we can extend the result to invariance entailed by a PAG. Definition first:

Definition 25 (Invariance entailed by a PAG) *Let $\mathcal{P}$ be a PAG over* $\mathbf{O}$, *and* $\mathbf{X},\mathbf{Y},\mathbf{Z} \subseteq \mathbf{O}$ *be three sets of variables such that* $\mathbf{X}\cap\mathbf{Y} = \mathbf{Y}\cap\mathbf{Z} = \varnothing$, $P(\mathbf{y}|\mathbf{z})$ *is* **entailed to be invariant under interventions on X given** $\mathcal{P}$ *if for every MAG* $\mathcal{M}$ *in the Markov equivalence class represented by* $\mathcal{P}$, $P(\mathbf{y}|\mathbf{z})$ *is entailed to be invariant under interventions on* $\mathbf{X}$ *given* $\mathcal{M}$.

We need a few lemmas that state connections between m-connecting paths in a MAG and definite m-connecting paths (as opposed to mere possibly m-connecting paths) in its PAG. By the definition of definite m-connecting paths (Definition 4), definite m-connection in a PAG implies m-connection in every MAG represented by the PAG. It is not obvious, however, that m-connection in a MAG will always be revealed as definite m-connection in its PAG. Fortunately, this turns out to be true. However, the proof is highly involved, and relies on many results about the properties of PAGs and the transformation between PAGs and MAGs presented in Zhang (2006, chapters 3-4), which would take up too much space and might distract readers from the main points of the present paper. So we will simply state the fact here, and refer interested readers to Zhang (2006, chap. 5, Lemma 5.1.7) for the proof.

Lemma 26 *Let* $\mathcal{M}$ *be a MAG over* $\mathbf{O}$, *and* $\mathcal{P}$ *be the PAG that represents* $[\mathcal{M}]$. *For any* $A,B \in \mathbf{O}$ *and* $\mathbf{C} \subseteq \mathbf{O}$ *that does not contain A or B, if there is a path m-connecting A and B given* $\mathbf{C}$ *in* $\mathcal{M}$, *then there is a path definitely m-connecting A and B given* $\mathbf{C}$ *in* $\mathcal{P}$. *Furthermore, if there is an m-connecting path in* $\mathcal{M}$ *that is either into A or out of A with an invisible directed edge, then there is a definite m-connecting path in* $\mathcal{P}$ *that does not start with a definitely visible edge out of A.*

Proof See the proof of Lemma 5.1.7 in Zhang (2006, pp. 207). ■

The converse to the second part of Lemma 26 is also true.

Lemma 27 *Let* $\mathcal{P}$ *be a PAG over* $\mathbf{O}$. *For any* $A,B \in \mathbf{O}$ *and* $\mathbf{C} \subseteq \mathbf{O}$ *that does not contain A or B, if there is a path definitely m-connecting A and B given* $\mathbf{C}$ *in* $\mathcal{P}$ *that does not start with a definitely visible edge out of A, then there exists a MAG* $\mathcal{M}$ *in the equivalence class represented by* $\mathcal{P}$ *in which there is a path m-connecting A and B given* $\mathbf{C}$ *that is either into A or includes an invisible directed edge out of A.*

Proof See Appendix B. ■

Lemmas 26 and 27 are useful for establishing conditions analogous to clauses (1) and (2) in Theorem 24. For clause (3), we need two more lemmas.

Lemma 28 *Let* $\mathcal{M}$ *be a MAG over* $\mathbf{O}$, *and* $\mathcal{P}$ *be the PAG that represents* $[\mathcal{M}]$. *For any* $A,B \in \mathbf{O}$ *and* $\mathbf{C} \subseteq \mathbf{O}$ *that does not contain B or any descendant of A in* $\mathcal{M}$, *if there is a path m-connecting A and B given* $\mathbf{C}$ *in* $\mathcal{M}$ *that is out of A, then there is a path definitely m-connecting A and B given* $\mathbf{C}$ *in* $\mathcal{P}$ *that is not into A (i.e., the edge incident to A on the path is either* $A \circ\!\!-\!\!\circ$, *or* $A \circ\!\!\rightarrow$, *or* $A \rightarrow$*).*

Proof See Appendix B. ■

Lemma 29 *Let* $\mathcal{P}$ *be a PAG over* $\mathbf{O}$. *For any* $A,B \in \mathbf{O}$ *and* $\mathbf{C} \subseteq \mathbf{O}$ *that does not contain A or B, if there is a path definitely m-connecting A and B given* $\mathbf{C}$ *in* $\mathcal{P}$ *that is not into A, then there exists a MAG* $\mathcal{M}$ *represented by* $\mathcal{P}$ *in which there is a path m-connecting A and B given* $\mathbf{C}$ *that is out of A.*

Proof See Appendix B. ■

The main theorem follows.

Theorem 30 *Suppose $\mathcal{P}$ is the causal PAG over a set of variables $\mathbf{O}$. For any $\mathbf{X},\mathbf{Y},\mathbf{Z} \subseteq \mathbf{O}$ such that $\mathbf{X} \cap \mathbf{Y} = \mathbf{Y} \cap \mathbf{Z} = \varnothing$, $P(\mathbf{y}|\mathbf{z})$ is entailed to be invariant under interventions on $\mathbf{X}$ given $\mathcal{P}$ if and only if*

(1) for every $X \in \mathbf{X} \cap \mathbf{Z}$, every definite m-connecting path, if any, between X and any member of $\mathbf{Y}$ given $\mathbf{Z} \backslash \{X\}$ is out of X with a definitely visible edge;

(2) for every $X \in \mathbf{X} \cap (\mathbf{PossibleAn}_{\mathcal{P}}(\mathbf{Z}) \backslash \mathbf{Z})$, there is no definite m-connecting path between X and any member of $\mathbf{Y}$ given $\mathbf{Z}$; and

(3) for every $X \in \mathbf{X} \backslash \mathbf{PossibleAn}_{\mathcal{P}}(\mathbf{Z})$, every definite m-connecting path, if any, between X and any member of $\mathbf{Y}$ given $\mathbf{Z}$ is into X.

Proof We show that (1), (2) and (3) are sufficient and necessary for the corresponding conditions in Theorem 24 to hold for all MAGs represented by $\mathcal{P}$. It follows from Lemma 26 that if (1) holds, then the first condition in Theorem 24 holds for all MAGs represented by $\mathcal{P}$. Note moreover that for every MAG $\mathcal{M}$ represented by $\mathcal{P}$, $\mathbf{An}_{\mathcal{M}}(\mathbf{Z}) \subseteq \mathbf{PossibleAn}_{\mathcal{P}}(\mathbf{Z})$. It again follows from Lemma 26 that if (2) holds, then the second condition in Theorem 24 holds for all MAGs represented by $\mathcal{P}$. Finally, it follows from Lemma 28 (and Lemma 26) that if (3) holds, the third condition in Theorem 24 holds for all MAGs represented by $\mathcal{P}$. Hence (1), (2) and (3) are sufficient.

Conversely, if (1) fails, then by Lemma 27, there exists a MAG represented by $\mathcal{P}$ for which the first condition in Theorem 24 fails.

To show the necessity of (2), we need the fact mentioned in Footnote 11 that if X is a possible ancestor of a vertex $Z \in \mathbf{Z}$ in $\mathcal{P}$, then there exists a MAG represented by $\mathcal{P}$, in which X is an ancestor of Z. So if (2) fails, that is, there is a definite m-connecting path between a variable $X \in \mathbf{X} \cap (\mathbf{PossibleAn}_{\mathcal{P}}(\mathbf{Z}) \backslash \mathbf{Z})$ and a member of $\mathbf{Y}$ given $\mathbf{Z}$ in $\mathcal{P}$, then there exists a MAG $\mathcal{M}$ represented by $\mathcal{P}$ in which $X \in \mathbf{X} \cap (\mathbf{An}_{\mathcal{M}}(\mathbf{Z}) \backslash \mathbf{Z})$, and there is an m-connecting path between X and a member of $\mathbf{Y}$ given $\mathbf{Z}$, which violates clause (2) of Theorem 24.

Lastly, if (3) fails, that is, there is a definite m-connecting path between a variable $X \in \mathbf{X} \backslash \mathbf{PossibleAn}_{\mathcal{P}}(\mathbf{Z})$ and a member of $\mathbf{Y}$ given $\mathbf{Z}$ that is *not* into X, then it follows from Lemma 29 that there exists a MAG $\mathcal{M}$ represented by $\mathcal{P}$ in which there is an m-connecting path between X and a member of $\mathbf{Y}$ given $\mathbf{Z}$ that is out of X. Moreover, since $X \in \mathbf{X} \backslash \mathbf{PossibleAn}_{\mathcal{P}}(\mathbf{Z})$, X cannot be an ancestor of $\mathbf{Z}$ in $\mathcal{M}$, that is, $X \in \mathbf{X} \backslash \mathbf{An}_{\mathcal{M}}(\mathbf{Z})$. So $\mathcal{M}$ fails clause (3) of Theorem 24. Therefore, the conditions are also necessary. ■

For a simple illustration, consider again the PAG in Figure 4. Given the PAG, it can be inferred that $P(L|G,S)$ is invariant under interventions on I, because there is no definite m-connecting path between L and I given $\{G,S\}$, satisfying the relevant clause—clause (2)—in Theorem 30. $P(L|G,S)$ is also invariant under interventions on S because the only definitely m-connecting path between L and S given $\{G\}$ is $S \to L$ which contains a definitely visible edge out of S, satisfying the relevant clause—clause (1)—in Theorem 30.

On the other hand, for example, $P(S)$ is not entailed to be invariant under interventions on I. Note that given the MAG of Figure 3(b), $P(S)$ is indeed entailed to be invariant under interventions on I, but this invariance is not unanimously implied in the equivalence class. Given some other MAGs in the class, such as the one in Figure 3(a), $P(S)$ is not entailed to be invariant under interventions on I.

As briefly noted in the last section, the PAG-based *do*-calculus in its current form is not complete. We mentioned three issues that might be responsible for this (cf. the comments right after Theorem 17), but only one of them we are sure leads to counterexamples—examples in which a rule in the DAG-based calculus is applicable for all DAGs compatible with the given PAG, but the corresponding rule in the PAG-based calculus is not applicable. It is the fact that the calculus is formulated in terms of absence of possibly m-connecting paths. Consider the example we used to illustrate the difference between definite and possibly m-connecting paths in Section 3. Given the PAG in Figure 5, we cannot apply rule 2 of the PAG-based *do*-calculus to conclude that $P(W|do(X),Y,Z) = P(W|Y,Z)$, because there is a possibly m-connecting path between X and W relative to $\{Y,Z\}$ in the PAG (note that since $X \in \mathbf{PossibleAn}(\{Y,Z\})$, the rule does not require manipulating the graph). However, it can be shown that for every DAG compatible with the PAG, X and W are d-separated by $\{Y,Z\}$ in either the X-upper-manipulation of the DAG or in the DAG itself. So rule 2 of the DAG-based *do*-calculus is actually applicable given any DAG compatible with the PAG.

Although we suspect that such counterexamples may not be encountered often in practice, it is at least theoretically interesting to handle them. Our results in this section provide an improvement in regard to the important special case of invariance. That is, the conditions given in Theorem 30 are complete for deriving statements of invariance, in the following sense: if the conditions therein fail relative to a PAG, then there exists a DAG represented by the PAG given which the conditions in Proposition 18 do not hold. The example in Figure 5 is not a counterexample to the completeness of Theorem 30. Unlike the *do*-calculus presented in Theorem 17, Theorem 30 implies that $P(W|Y,Z)$ is entailed to be invariant under interventions on X given the PAG (and hence we can conclude that $P(W|do(X),Y,Z) = P(W|Y,Z)$), because there is no definite m-connecting path between X and W relative to $\{Y,Z\}$ in the PAG. Whether it is valid to formulate the PAG-based *do*-calculus in terms of definite m-connecting paths is an open question at this point (cf. Footnote 15).[19]

Theorem 30 is in style very similar to Theorems 7.3 and 7.4 in Spirtes et al. (1993). The latter are formulated with respect to a *partially oriented inducing path graph* (POIPG). We include in Appendix A a description of the inducing path graphs (IPGs) as well as their relationship to ancestral graphs. As shown there, syntactically the class of ancestral graphs is a proper subclass of the class of inducing path graphs. In consequence a PAG in general reveals more qualitative causal information than a POIPG. In addition, it seems MAGs are easier to parameterize than IPGs. (For a linear parametrization of MAGs, see Richardson and Spirtes, 2002.)

Apart from the advantages of working with MAGs and PAGs over IPGs and POIPGs, our Theorem 30 is superior to Spirtes et al.'s theorems in that our theorem is formulated in terms of definite m-connecting paths, whereas theirs, like the results in the last section, are formulated in terms of

19. Here is another way to view the open problem. As explained earlier, *do*-calculus is essentially a generalization of the invariance conditions. Not only does it address the question of when $(\mathbf{y}|\mathbf{z})$ is invariant under an intervention $\mathbf{X} := \mathbf{x}$, it also addresses the more general question of when a post-intervention probability $P_{\mathbf{u}}(\mathbf{y}|\mathbf{z})$ would be invariant under a *further* intervention $\mathbf{X} := \mathbf{x}$. Our results in this section do not cover the latter question. To generalize the results in terms of definite m-connecting paths to address the latter question is parallel to improving the *do*-calculus.

possibly m-connecting paths. As a result, their conditions are only sufficient but not necessary. Regarding the case in Figure 5, for example, their theorems do not imply that $P(W|Y,Z)$ is entailed to be invariant under interventions on X, due to the presence of the possibly m-connecting path in the graph (which in this case is also the POIPG). Furthermore, since definite m-connecting paths are special cases of possibly m-connecting paths, there are more possibly m-connecting paths than definite m-connecting paths to check in a PAG. This may turn out to be a computational advantage for our theorem.

6. Conclusion

Causal reasoning about consequences of interventions has received rigorous and interesting treatments in the framework of causal Bayesian networks. Much of the work assumes that the structure of the causal Bayesian network, represented by a directed acyclic graph, is fully given. In this paper we have provided some results about causal reasoning under weaker causal assumptions, represented by a maximal ancestral graph or a partial ancestral graph, the latter of which is fully testable with observational data (assuming the causal Faithfulness condition).

Theorem 17 in Section 4 gives us a *do*-calculus under testable causal assumptions, represented by a PAG. The idea is that when any rule in the calculus is applicable given the PAG, the corresponding rule in Pearl's original *do*-calculus is applicable relative to each and every DAG compatible with the PAG. The converse, however, is not true; it is not the case that whenever all DAGs compatible with the PAG sanction the application of a certain rule in the *do*-calculus, the corresponding rule in the PAG-based calculus is also applicable. An interesting project is to either improve the calculus, or to investigate more closely the extent to which the current version is not complete.

As a first step towards improvement, we examined in Section 5 an important special case of the *do*-calculus—the graphical conditions for invariance under interventions—and presented sufficient and necessary conditions for invariance given a PAG. These conditions are very similar but also superior to the analogous results proved by Spirtes et al. (1993). In the latter work, there is also an algorithm (named Prediction Algorithm) for identifying post-intervention probabilities based on the conditions for invariance. The results in this paper can certainly be used to improve that algorithm.

The search for a syntactic derivation in the *do*-calculus to identify a post-intervention probability is no minor computational task. For this reason, it is worth deriving handy graphical criteria for identifiability from the *do*-calculus. Since invariant quantities are the most basic identifiable quantities, the condition for invariance is the most basic among such graphical criteria. Other graphical criteria in the literature, including the well known "back door criterion" and "front door criterion", should be extendible to PAGs in the same way as we did for invariance. On the other hand, a novel approach to identification has been developed recently by Tian and Pearl (2004), which proves computationally attractive. To adapt that approach to ancestral graphs is probably a worthy project.

Acknowledgments

I am grateful to Clark Glymour, Thomas Richardson, and Peter Spirtes for their helpful comments on the part of my dissertation this paper is based on. Thanks also to three anonymous referees for helping improve the paper significantly. One of them, especially, made extremely detailed and helpful suggestions.

Appendix A. Inducing Path Graphs

The theory of invariance under interventions developed in this paper is largely parallel to that developed in Spirtes et al. (1993). Their theory is based on a graphical representation called inducing path graphs. This graphical object is not given an independent syntactic definition, but defined via a construction relative to a DAG (with latent variables). It is clear from the construction that this representation is closely related to MAGs. In this appendix we specify the exact relationship between them. In particular, we justify an independent syntactic definition of inducing path graphs, which makes it clear that syntactically the class of MAGs is a subclass of inducing path graphs.

An *inducing path graph (IPG)* is a directed mixed graph, defined relative to a DAG, through the following construction:

Input: a DAG $\mathcal{G}$ over $\langle \mathbf{O}, \mathbf{L} \rangle$
Output: an IPG $\mathcal{I}_\mathcal{G}$ over $\mathbf{O}$

1. for each pair of variables $A, B \in \mathbf{O}$, A and B are adjacent in $\mathcal{I}_\mathcal{G}$ if and only if there is an inducing path between them relative to $\mathbf{L}$ in $\mathcal{G}$;

2. for each pair of adjacent vertices A, B in $\mathcal{I}_\mathcal{G}$, mark the A-end of the edge as an arrowhead if there is an inducing path between A and B that is into A, otherwise mark the A-end of the edge as a tail.

It can be shown that the construction outputs a mixed graph $\mathcal{I}_\mathcal{G}$ in which the set of m-separation relations matches exactly the set of d-separation relations among $\mathbf{O}$ in the original DAG $\mathcal{G}$ (Spirtes and Verma, 1992). Furthermore, $\mathcal{I}_\mathcal{G}$ encodes information about inducing paths in the original graph, which in turn implies features of the original DAG that bear causal significance. Specifically, we have two useful facts: (i) if there is an inducing path between A and B relative to $\mathbf{L}$ that is out of A, then A is an ancestor of B in $\mathcal{G}$; (ii) if there is an inducing path between A and B relative to $\mathbf{L}$ that is into both A and B, then A and B have a common ancestor in $\mathbf{L}$ unmediated by any other observed variable.[20] So $\mathcal{I}_\mathcal{G}$, just like the MAG for $\mathcal{G}$, represents both the conditional independence relations and (features of) the causal structure among the observed variables $\mathbf{O}$. Since the above construction produces a unique graph given a DAG $\mathcal{G}$, it is fair to call $\mathcal{I}_\mathcal{G}$ the IPG for $\mathcal{G}$.

Therefore a directed mixed graph over a set of variables is an IPG if it is the IPG for some DAG. We now show that a directed mixed graph is an IPG if and only if it is maximal and does not contain a directed cycle.

Theorem 31 *For any directed mixed graph $\mathcal{I}$ over a set of variables* $\mathbf{O}$, *there exists a DAG $\mathcal{G}$ over* $\mathbf{O}$ *and possibly some extra variables* $\mathbf{L}$ *such that $\mathcal{I} = \mathcal{I}_\mathcal{G}$—that is, $\mathcal{I}$ is the IPG for $\mathcal{G}$—if and only if*

(i1) There is no directed cycle in $\mathcal{I}$; and

(i2) $\mathcal{I}$ is maximal (i.e., there is no inducing path between two non-adjacent variables).

Proof We first show that the conditions are necessary (**only if**). Suppose there exists a DAG $\mathcal{G}(\mathbf{O}, \mathbf{L})$ whose IPG is $\mathcal{I}$. In other words, $\mathcal{I}$ is the output of the IPG construction procedure given $\mathcal{G}$. If there is any directed cycle in $\mathcal{I}$, say $c = \langle O_1, \ldots, O_n, O_1 \rangle$, then between any pair of adjacent nodes in the cycle, O_i and O_{i+1} ($1 \leq i \leq n$ and $O_{n+1} = O_1$), there is an inducing path between them in $\mathcal{G}$ relative

20. For more details of the causal interpretation of IPGs, see Spirtes et al. (1993, pp. 130-138).

to $\mathbf{L}$, which, by one of the facts mentioned earlier, implies that O_i is an ancestor of O_{i+1} in $\mathcal{G}$. Thus there would be a directed cycle in $\mathcal{G}$ as well, a contradiction. Therefore there is no directed cycle in $\mathcal{I}$. To show that it is also maximal, consider any two non-adjacent nodes A and B in $\mathcal{I}$. We show that there is no inducing path in $\mathcal{I}$ between A and B. Otherwise let $p = \langle A, O_1, \ldots, O_n, B\rangle$ be an inducing path. By the construction, there is an inducing path relative to $\mathbf{L}$ in $\mathcal{G}$ between A and O_1 that is into O_1, and an inducing path relative to $\mathbf{L}$ in $\mathcal{G}$ between B and O_n that is into O_n, and for every $1 \leq i \leq i-1$, there is an inducing path relative to $\mathbf{L}$ in $\mathcal{G}$ between O_i and O_{i+1} that is into both. By Lemma 32 in Appendix B, it follows that there is an inducing path between A and B relative to $\mathbf{L}$ in $\mathcal{G}$, and so A and B should be adjacent in $\mathcal{I}$, a contradiction. Therefore $\mathcal{I}$ is also maximal.

Next we demonstrate sufficiency (**if**). If the two conditions hold, construct a DAG $\mathcal{G}$ as follows: retain all the directed edges in $\mathcal{I}$, and for each bi-directed edge $A \leftrightarrow B$ in $\mathcal{I}$, introduce a latent variable L_{AB} in $\mathcal{G}$ and replace $A \leftrightarrow B$ with $A \leftarrow L_{AB} \rightarrow B$.[21] It is easy to see that the resulting graph $\mathcal{G}$ is a DAG, as in $\mathcal{I}$ there is no directed cycle. We show that $\mathcal{I} = \mathcal{I}_{\mathcal{G}}$, the IPG for $\mathcal{G}$. For any pair of variables A and B in $\mathcal{I}$, there are four cases to consider:

Case 1: $A \rightarrow B$ is in $\mathcal{I}$. Then $A \rightarrow B$ is also in $\mathcal{G}$, so A and B are adjacent in $\mathcal{I}_{\mathcal{G}}$. In $\mathcal{I}_{\mathcal{G}}$, the edge between A and B is not $A \leftarrow B$, because otherwise B would have to be an ancestor of A in $\mathcal{G}$, a contradiction. The edge is not $A \leftrightarrow B$ either, because otherwise there would have to be a latent variable that is a parent of both A and B, which by the construction of $\mathcal{G}$ is not the case. So $A \rightarrow B$ is also in $\mathcal{I}_{\mathcal{G}}$.

Case 2: $A \leftarrow B$ is in $\mathcal{I}$. By the same argument as in *Case 1*, $A \leftarrow B$ is also in $\mathcal{I}_{\mathcal{G}}$.

Case 3: $A \leftrightarrow B$ is in $\mathcal{I}$. Then there is a L_{AB} such that $A \leftarrow L_{AB} \rightarrow B$ is in $\mathcal{G}$. Then obviously $\langle A, L_{AB}, B\rangle$ is an inducing path relative to $\mathbf{L}$ in $\mathcal{G}$ that is into both A and B, and hence $A \leftrightarrow B$ is also in $\mathcal{I}_{\mathcal{G}}$.

Case 4: A and B are not adjacent in $\mathcal{I}$. We show that they are not adjacent in $\mathcal{I}_{\mathcal{G}}$ either. For this, we only need to show that there is no inducing path between A and B relative to $\mathbf{L}$ in $\mathcal{G}$. Suppose otherwise that there is such an inducing path p between A and B in $\mathcal{G}$. Let $\langle A, O_1, \ldots, O_n, B\rangle$ be the sub-sequence of p consisting of all observed variables on p. By the definition of inducing path, all O_i's ($1 \leq i \leq n$) are colliders on p and are ancestors of either A or B. By the construction of $\mathcal{G}$, it is easy to see that O_i's are also ancestors of either A or B in $\mathcal{I}$. It is also easy to see that either $A \rightarrow O_1$ or $A \leftarrow L_{AO_1} \rightarrow O_1$ appears in $\mathcal{G}$, which implies that there is an edge between A and O_1 that is into O_1 in $\mathcal{I}$. Likewise, there is an edge between O_n and B that is into O_n in $\mathcal{I}$, and there is an edge between O_i and O_{i+1} that is into both in $\mathcal{I}$ for all $1 \leq i \leq n-1$. So $\langle A, O_1, \ldots, O_n, B\rangle$ constitutes an inducing path between A and B in $\mathcal{I}$, which contradicts the assumption that $\mathcal{I}$ is maximal. So there is no inducing path between A and B relative to $\mathbf{L}$ in $\mathcal{G}$, which means that A and B are not adjacent in $\mathcal{I}_{\mathcal{G}}$.

Therefore $\mathcal{I} = \mathcal{I}_{\mathcal{G}}$, the IPG for $\mathcal{G}$. ■

Given this theorem, it is clear that we can define IPGs in terms of $(i1)$ and $(i2)$. So a MAG is also an IPG, but an IPG is not necessarily a MAG, as the former may contain an almost directed cycle. The simplest IPG which is not a MAG is shown in Figure 11.

Spirtes et al. (1993) uses *partially oriented inducing path graphs (POIPGs)* to represent Markov equivalence classes of IPGs. The idea is exactly the same as PAGs. A (complete) POIPG displays (all) common marks in a Markov equivalence class of IPGs. An obvious fact is that given a set of conditional independence facts that admits a faithful representation by a MAG, the Markov equiva-

21. This is named *canonical DAG* in Richardson and Spirtes (2002).

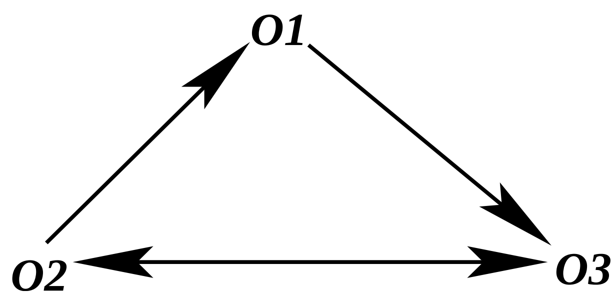

Figure 11: A simplest IPG that is not a MAG

lence class of MAGs is included in the Markov equivalence class of IPGs. It follows that the POIPG cannot contain more informative marks than the PAG, but may contain fewer. So a PAG usually reveals more qualitative causal information than a POIPG does.

Appendix B. Proofs of the Lemmas

In proving some of the lemmas, we will use the following fact, which was proved in, for example, Spirtes et al. (1999, pp. 243):

Lemma 32 *Let $\mathcal{G}(\mathbf{O},\mathbf{L})$ be a DAG, and $\langle V_0,\ldots,V_n\rangle$ be a sequence of distinct variables in $\mathbf{O}$. If (1) for all $0 \le i \le n-1$, there is an inducing path in $\mathcal{G}$ between V_i and V_{i+1} relative to $\mathbf{L}$ that is into V_i unless possibly $i = 0$ and is into V_{i+1} unless possibly $i = n-1$; and (2) for all $1 \le i \le n-1$, V_i is an ancestor of either V_0 or V_n in $\mathcal{G}$; then there is a subpath s of the concatenation of those inducing paths that is an inducing path between V_0 and V_n relative to $\mathbf{L}$ in $\mathcal{G}$. Furthermore, if the said inducing path between V_0 and V_1 is into V_0, then s is into V_0, and if the said inducing path between V_{n-1} and V_n is into V_n, then s is into V_n.*

Proof This is a special case of Lemma 10 in Spirtes et al. (1999, pp. 243). See their paper for a detailed proof. (One may think that the concatenation itself would be an inducing path between V_0 and V_n. This is almost correct, except that the concatenation may contain the same vertex multiple times. So in general it is a subsequence of the concatenation that constitutes an inducing path between V_0 and V_n.) ■

Lemma 32 gives a way to argue for the presence of an inducing path between two variables in a DAG, and hence is very useful for demonstrating that two variables are adjacent in the corresponding MAG. We will see several applications of this lemma in the subsequent proofs.

Proof of Lemma 9

Proof Since there is an inducing path between A and B relative to $\mathbf{L}$ in $\mathcal{G}$, A and B are adjacent in $\mathcal{M}$. Furthermore, since $A \in \mathbf{An}_{\mathcal{G}}(B)$, the edge between A and B in $\mathcal{M}$ is $A \to B$. We now show that it is invisible in $\mathcal{M}$. To show this, it suffices to show that for any C, if in $\mathcal{M}$ there is an edge between C and A that is into A or there is a collider path between C and A that is into A and every vertex on the path is a parent of B, then C is adjacent to B, which means that the condition for visibility cannot be met.

Let u be an inducing path between A and B relative to $\mathbf{L}$ in $\mathcal{G}$ that is into A. For any C, we consider the two possible cases separately:

Case 1: There is an edge between C and A in $\mathcal{M}$ that is into A. Then, by the way $\mathcal{M}$ is constructed from $\mathcal{G}$, there must be an inducing path u' in $\mathcal{G}$ between A and C relative to $\mathbf{L}$. Moreover, u' is into A, for otherwise A would be an ancestor of C, so that the edge between A and C in $\mathcal{M}$ would be out of A. Given u, u' and the fact that $A \in \mathbf{An}_{\mathcal{G}}(B)$, we can apply Lemma 32 to conclude that there is an inducing path between C and B relative to $\mathbf{L}$ in $\mathcal{G}$, which means C and B are adjacent in $\mathcal{M}$.

Case 2: There is a collider path p in $\mathcal{M}$ between C and A that is into A and every non-endpoint vertex on the path is a parent of B. For every pair of adjacent vertices $\langle V_i, V_{i+1}\rangle$ on p, the edge is $V_i \leftrightarrow V_{i+1}$ if $V_i \neq C$, and otherwise either $C \leftrightarrow V_{i+1}$ or $C \rightarrow V_{i+1}$. It follows that there is an inducing path in $\mathcal{G}$ between V_i and V_{i+1} relative to $\mathbf{L}$ such that the path is into V_{i+1}, and is into V_i unless possibly $V_i = C$. Given these inducing paths and the fact that every variable other than C on p is an ancestor of B, we can apply Lemma 32 to conclude that there is an inducing path between C and B relative to $\mathbf{L}$ in $\mathcal{G}$, which means C and B are adjacent in $\mathcal{M}$.

Therefore, the edge $A \rightarrow B$ is invisible in $\mathcal{M}$. ■

Proof of Lemma 10

Proof Construct a DAG from $\mathcal{M}$ as follows:

1. Leave every directed edge in $\mathcal{M}$ as it is. Introduce a latent variable L_{AB} and add $A \leftarrow L_{AB} \rightarrow B$ to the graph.

2. for every bi-directed edge $Z \leftrightarrow W$ in $\mathcal{M}$, delete the bi-directed edge. Introduce a latent variable L_{ZW} and add $Z \leftarrow L_{ZW} \rightarrow W$ to the graph.

The resulting graph we denote by $\mathcal{G}$, a DAG over $(\mathbf{O}, \mathbf{L})$, where $\mathbf{L} = \{L_{AB}\} \cup \{L_{ZW} | Z \leftrightarrow W \text{ is in } \mathcal{M}\}$. Obviously $\mathcal{G}$ is a DAG in which A and B share a latent parent. We need to show that $\mathcal{M} = \mathcal{M}_{\mathcal{G}}$, that is, $\mathcal{M}$ is the MAG of $\mathcal{G}$. For any pair of variables X and Y, there are four cases to consider:

Case 1: $X \rightarrow Y$ is in $\mathcal{M}$. Since $\mathcal{G}$ retains all directed edges in $\mathcal{M}$, $X \rightarrow Y$ is also in $\mathcal{G}$, and hence is also in $\mathcal{M}_{\mathcal{G}}$.

Case 2: $X \leftarrow Y$ is in $\mathcal{M}$. Same as *Case 1*.

Case 3: $X \leftrightarrow Y$ is in $\mathcal{M}$. Then there is a latent variable L_{XY} in $\mathcal{G}$ such that $X \leftarrow L_{XY} \rightarrow Y$ appears in $\mathcal{G}$. Since $X \leftarrow L_{XY} \rightarrow Y$ is an inducing path between X and Y relative to $\mathbf{L}$ in $\mathcal{G}$, X and Y are adjacent in $\mathcal{M}_{\mathcal{G}}$. Furthermore, it is easy to see that the construction of $\mathcal{G}$ does not create any directed path from X to Y or from Y to X. So X is not an ancestor of Y and Y is not an ancestor of X in $\mathcal{G}$. It follows that in $\mathcal{M}_{\mathcal{G}}$ the edge between X and Y is $X \leftrightarrow Y$.

Case 4: X and Y are not adjacent in $\mathcal{M}$. We show that in $\mathcal{G}$ there is no inducing path between X and Y relative to $\mathbf{L}$. Suppose otherwise that there is one. Let p be an inducing path between X and Y relative to $\mathbf{L}$ in $\mathcal{G}$ that includes a minimal number of observed variables. Let $\langle X, O_1, \ldots, O_n, Y\rangle$ be the sub-sequence of p consisting of all observed variables on p. By the definition of inducing path, all O_i's ($1 \leq i \leq n$) are colliders on p and are ancestors of either X or Y in $\mathcal{G}$. Since the construction of $\mathcal{G}$ does not create any new directed path from an observed variable to another observed variable, O_i's are also ancestors of either X or Y in $\mathcal{M}$. Since O_1 is a collider on p, either $X \rightarrow O_1$ or $X \leftarrow L_{XO_1} \rightarrow O_1$ appears in $\mathcal{G}$. Either way there is an edge between X and O_1 that is into O_1 in $\mathcal{M}$. Likewise, there is an edge between O_n and Y that is into O_n in $\mathcal{M}$.

Moreover, for all $1 \leq i \leq n-1$, the path p in $\mathcal{G}$ contains $O_i \leftarrow L_{O_iO_{i+1}} \rightarrow O_{i+1}$, because all O_i's are colliders on p. Thus in $\mathcal{M}$ there is an edge between O_i and O_{i+1}. Regarding these edges, by construction of the MAG, either all of them are bi-directed, or one of them is $A \rightarrow B$ and others are bi-directed. In the former case, $\langle X, O_1, \ldots, O_n, Y\rangle$ constitutes an inducing path between X and Y in $\mathcal{M}$, which contradicts the maximality of $\mathcal{M}$. In the latter case, without loss of generality, suppose $\langle A, B\rangle = \langle O_k, O_{k+1}\rangle$. Then $\langle X, O_1, ..., O_k = A\rangle$ is a collider path into A in $\mathcal{M}$. We now show by induction that for all $1 \leq i \leq k-1$, O_i is a parent of B in $\mathcal{M}$.

Consider O_{k-1} in the base case. O_{k-1} is adjacent to B, for otherwise $A \rightarrow B$ would be visible in $\mathcal{M}$ because there is an edge between O_{k-1} and A that is into A. The edge between O_{k-1} and B is not $O_{k-1} \leftarrow B$, for otherwise there would be $A \rightarrow B \rightarrow O_{k-1}$ and yet an edge between O_{k-1} and A that is into A in $\mathcal{M}$, which contradicts the fact that $\mathcal{M}$ is ancestral. The edge between them is not $O_{k-1} \leftrightarrow B$, for otherwise there would be an inducing path between X and Y relative to $\mathbf{L}$ in $\mathcal{G}$ that includes fewer observed variables than p does, which contradicts our choice of p. So O_{k-1} is a parent of B in $\mathcal{M}$.

In the inductive step, suppose for all $1 < m+1 \leq j \leq k-1$, O_j is a parent of B in $\mathcal{M}$, and we need to show that O_m is also a parent of B in $\mathcal{M}$. The argument is essentially the same as in the base case. Specifically, O_m and B are adjacent because otherwise it follows from the inductive hypothesis that $A \rightarrow B$ is visible. The edge is not $O_m \leftarrow B$ on pain of making $\mathcal{M}$ non-ancestral; and the edge is not $O_m \leftrightarrow B$ on pain of creating an inducing path that includes fewer observed variables than p does. So O_m is also a parent of B.

Now we have shown that for all $1 \leq i \leq k-1$, O_i is a parent of B in $\mathcal{M}$. It follows that X is adjacent to B, for otherwise $A \rightarrow B$ would be visible. Again, the edge is not $X \leftarrow B$ on pain of making $\mathcal{M}$ non-ancestral. So the edge between X and B in $\mathcal{M}$ is into B, but then there is an inducing path between X and Y relative to $\mathbf{L}$ in $\mathcal{G}$ that includes fewer observed variables than p does, which is a contradiction with our choice of p.

So our initial supposition is false. There is no inducing path between X and Y relative to $\mathbf{L}$ in $\mathcal{G}$, and hence X and Y are not adjacent in $\mathcal{M}_{\mathcal{G}}$.

Therefore $\mathcal{M} = \mathcal{M}_{\mathcal{G}}$. ∎

Proof of Lemma 12

Proof Recall the diagram in Figure 7:

$$\begin{tikzcd}
\mathcal{G} \arrow[rrr, "mc"] \arrow[d, "gm"'] & & & \mathcal{M} \arrow[d, "mm"] \\
\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}} \arrow[r, "mc"'] & \mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}} & \leq & \mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}
\end{tikzcd}$$

What we need to show is that $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ is an I-map of $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$, or in other words, whatever m-separation relation is true in the former is also true in the latter. To show this, it suffices to show that $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ is Markov equivalent to a supergraph of $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$.

For that purpose, we first establish two facts: (1) every directed edge in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$ is also in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$; and (2) for every bi-directed edge $S \leftrightarrow T$ in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$, S and T are also adjacent in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$; and the edge between S and T is either a bi-directed edge or an invisible directed edge in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$.

(1) is relatively easy to show. Note that for any $P \to Q$ in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$, $P \notin \mathbf{Y}$, for otherwise P would not be an ancestor of Q in $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, and hence would not be a parent of Q in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$; and likewise $Q \notin \mathbf{X}$, for otherwise Q would not be a descendant of P in $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, and hence would not be a child of P in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$. Furthermore, because $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ is a subgraph of $\mathcal{G}$, any inducing path between P and Q relative to $\mathbf{L}$ in $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ is also present in $\mathcal{G}$, and any directed path from P to Q in the former is also present in the latter. This entails that $P \to Q$ is also in $\mathcal{M}$, the MAG of $\mathcal{G}$. Since $P \notin \mathbf{Y}$ and $Q \notin \mathbf{X}$, $P \to Q$ is also present in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. So (1) is true.

(2) is less obvious. First of all, note that if $S \leftrightarrow T$ is in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$, then there is an inducing path between S and T relative to $\mathbf{L}$ in $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ that is into both S and T. This implies that $S, T \notin \mathbf{X}$, and moreover there is also an inducing path between S and T relative to $\mathbf{L}$ in $\mathcal{G}$ that is into both S and T. Hence there is an edge between S and T in $\mathcal{M}$, the MAG of $\mathcal{G}$. The edge in $\mathcal{M}$ is either $S \leftrightarrow T$ or, by Lemma 9, an invisible directed edge ($S \leftarrow T$ or $S \to T$).

Because $S, T \notin \mathbf{X}$, if $S \leftrightarrow T$ appears in $\mathcal{M}$, it also appears in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. If, on the other hand, the edge between S and T in $\mathcal{M}$ is directed, suppose without loss of generality that it is $S \to T$. Either $S \in \mathbf{Y}$, in which case we have $S \leftrightarrow T$ in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, because $S \to T$ is invisible in $\mathcal{M}$; or $S \notin \mathbf{Y}$, and $S \to T$ remains in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. In the latter case we need to show that $S \to T$ is still invisible in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. Suppose for the sake of contradiction that $S \to T$ is visible in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, that there is a vertex R not adjacent to T, such that either $R *\!\!\to S$ is in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ or there is a collider path c in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ between R and S that is into S on which every collider is a parent of T. We show that $S \to T$ is also visible in $\mathcal{M}$. Consider the two possible cases separately:

Case 1: $R *\!\!\to S$ is in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. If the edge is $R \to S$, it is also in $\mathcal{M}$, because manipulations of a MAG do not create new directed edges. We now show that R and T are not adjacent in $\mathcal{M}$. Suppose otherwise. The edge between R and T has to be $R \to T$ in $\mathcal{M}$. Note that $R \notin \mathbf{Y}$ for otherwise $R \to S$ would be deleted or changed into a bi-directed edge; and we have already shown that $T \notin \mathbf{X}$. It follows that $R \to T$ would be present in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ as well, a contradiction. Hence R and T are not adjacent in $\mathcal{M}$, and so the edge $S \to T$ is also visible in $\mathcal{M}$.

Suppose, on the other hand, the edge between R and S in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ is $R \leftrightarrow S$. In $\mathcal{M}$ the edge is either (i) $R \leftrightarrow S$, or (ii) $R \to S$. (It can't be $R \leftarrow S$ because then $S \in \mathbf{Y}$ and the edge $S \to T$ would not remain in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$.)

If (i) is the case, we argue that R and T are not adjacent in $\mathcal{M}$. Since $R \leftrightarrow S \to T$ is in $\mathcal{M}$, if R and T are adjacent, it has to be $R \leftrightarrow T$ or $R \to T$. In the former case, $R \leftrightarrow T$ would still be present in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ (because obviously $R, T \notin \mathbf{X}$), which is a contradiction. In the latter case, $R \to T$ is invisible in $\mathcal{M}$, for otherwise it is easy to see that $S \to T$ would also be visible. So either $R \to T$ remains in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ (if $R \notin \mathbf{Y}$), or it turns into $R \leftrightarrow T$ (if $R \in \mathbf{Y}$). In either case R and T would still be adjacent in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, a contradiction. Hence R and T are not adjacent in $\mathcal{M}$, and so the edge $S \to T$ is also visible in $\mathcal{M}$.

If (ii) is the case, then either R and T are not adjacent in $\mathcal{M}$, in which case $S \to T$ is also visible in $\mathcal{M}$; or R and T are adjacent in $\mathcal{M}$, in which case we now show that $S \to T$ is still visible. The edge between R and T in $\mathcal{M}$ has to be $R \to T$ (in view of $R \to S \to T$). Since R and T are not adjacent in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, and $R \to S$ is turned into $R \leftrightarrow S$ in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, by the definition of lower-manipulation (Definition 11), $R \to T$ is visible but $R \to S$ is invisible in $\mathcal{M}$. Because $R \to T$ is visible, by definition, there is a vertex Q not adjacent to T such that $Q *\!\!\to R$ is in $\mathcal{M}$ or there is a collider path in $\mathcal{M}$ between Q

and R that is into R on which every collider is a parent of T. But $R \rightarrow S$ is not visible, from which we can derive that $S \rightarrow T$ is visible in $\mathcal{M}$. Here is a sketch of the argument. If $Q {*\!\!\rightarrow} R$ is in $\mathcal{M}$, then Q and S must be adjacent (since otherwise $R \rightarrow S$ would be visible). It is then easy to derive that the edge between Q and S must be into S, which makes $S \rightarrow T$ visible. On the other hand, suppose there is a collider path c into R on which every collider is a parent of T. Then if there is a collider P on c such that $P \leftrightarrow S$ is in $\mathcal{M}$, we immediately get a collider path between Q and S that is into S on which every collider is a parent of T. This path makes $S \rightarrow T$ visible. Finally, if no collider on the path is a spouse of S, it is not hard to show that in order for $R \rightarrow S$ to be invisible, there has to be an edge between Q and S that is into S, which again makes $S \rightarrow T$ visible.

Case 2: There is a collider path c in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ between R and S that is into S on which every collider is a parent of T. We claim that every arrowhead on c, except possibly one at R, is also in $\mathcal{M}$. Because if an arrowhead is added at a vertex Q (which could be S) on c by the lower-manipulation, then $Q \in \mathbf{Y}$, but then the edge $Q \rightarrow T$ would not remain in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, a contradiction. So c is also a collider path in $\mathcal{M}$ that is into S. Furthermore, no new directed edges are introduced by lower-manipulation or upper-manipulation, so every collider on c is still a parent of T in $\mathcal{M}$.

It follows that if R and T are not adjacent in $\mathcal{M}$, then $S \rightarrow T$ is visible in $\mathcal{M}$. On the other hand, if R and T are adjacent in $\mathcal{M}$, it is either $R \leftrightarrow T$ or $R \rightarrow T$. Note that this edge is deleted in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. This implies that it is not $R \leftrightarrow T$ in $\mathcal{M}$: otherwise, the edge incident to R on c has to be bi-directed as well (since otherwise $\mathcal{M}$ is not ancestral), and hence if $R \leftrightarrow T$ is deleted, either the edge incident to R on c or the edge $S \rightarrow T$ should be deleted in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, which is a contradiction. So the edge is $R \rightarrow T$ in $\mathcal{M}$. Since $T \notin \mathbf{X}$ (for otherwise $S \rightarrow T$ would be deleted), $R \in \mathbf{Y}$, and $R \rightarrow T$ is visible in $\mathcal{M}$. But then it is easy to see that $S \rightarrow T$ is also visible in $\mathcal{M}$.

To summarize, we have shown that if $S \rightarrow T$ is visible in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, it is also visible in $\mathcal{M}$. Since it is not visible in $\mathcal{M}$, it is invisible in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ as well. Thus the edge between S and T is either a bi-directed edge or an invisible directed edge in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. Hence we have established (2).

The strategy to complete the proof is to show that $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ can be transformed into a supergraph of $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$ via a sequence of equivalence-preserving mark changes (Zhang and Spirtes, 2005; Tian, 2005). By (1) and (2), if $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ is not yet a supergraph of $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$, it is because some bi-directed edges in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$ correspond to directed edges in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. For any such directed edge $P \rightarrow Q$ in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ (with $P \leftrightarrow Q$ in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$), (2) implies that $P \rightarrow Q$ is invisible. It is then not hard to check that conditions in Lemma 1 of Zhang and Spirtes (2005)[22] hold for $P \rightarrow Q$ in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, and thus it can be changed into $P \leftrightarrow Q$ while preserving Markov equivalence. Furthermore, making this change will not make any other such directed edge in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ visible. It follows that $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ can be transformed into a Markov equivalent graph that is a supergraph of $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$. (We skip the details as they involve conditions for Markov equivalence we didn't have enough space to cover.)

Denote the supergraph by $\mathcal{I}$. It follows that if there is an m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$, the path is also m-connecting in $\mathcal{I}$, the supergraph of $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$. Because $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ and $\mathcal{I}$ are Markov equivalent, there is also an m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. ■

22. Here is the Lemma: Let $\mathcal{M}$ be a MAG, and $A \rightarrow B$ a directed edge in $\mathcal{M}$. Let $\mathcal{M}'$ be the graph identical to $\mathcal{M}$ except that the edge between A and B is $A \leftrightarrow B$ in $\mathcal{M}'$. (In other words, $\mathcal{M}'$ is the result of simply changing $A \rightarrow B$ into $A \leftrightarrow B$ in $\mathcal{M}$.) $\mathcal{M}'$ is a MAG and Markov equivalent to $\mathcal{M}$ if and only if
(t1) there is no directed path from A to B other than $A \rightarrow B$ in $\mathcal{M}$;
(t2)] For every $C \rightarrow A$ in $\mathcal{M}$, $C \rightarrow B$ is also in $\mathcal{M}$; and for every $D \leftrightarrow A$ in $\mathcal{M}$, either $D \rightarrow B$ or $D \leftrightarrow B$ is in $\mathcal{M}$; and
(t3) there is no discriminating path for A on which B is the endpoint adjacent to A in $\mathcal{M}$.

Proof of Lemma 16

Proof It is not hard to check that for any two variables $P, Q \in \mathbf{O}$, if P and Q are adjacent in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, then they are adjacent in $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ (though the converse is not necessarily true, because an edge not definitely visible in $\mathcal{P}$ may still be visible in $\mathcal{M}$). Furthermore, when they are adjacent in both $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ and $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, every non-circle mark on the edge in $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ is "sound" in that the mark also appears in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. The lemma obviously follows. ■

Proof of Lemma 21

Proof Spirtes and Richardson (1996), in proving their Lemma 18, gave a construction of an m-connecting path in $\mathcal{M}$ from a d-connecting path in $\mathcal{G}$. We describe the construction below.[23]

Let p be a minimal d-connecting path between A and B relative to $\mathbf{C}$ in $\mathcal{G}$ that is into A, minimal in the sense that no other d-connecting path between A and B relative to $\mathbf{C}$ that is into A is composed of fewer variables than p is.[24] Construct a sequence of variables in $\mathbf{O}$ in three steps.

Step 1: Form a sequence $\mathbf{T}$ of variables on p as follows. $\mathbf{T}[0] = A$, and $\mathbf{T}[n+1]$ is chosen to be the first vertex after $\mathbf{T}[n]$ on p that is either in $\mathbf{O}$ or a (latent) collider on p, until B is included in $\mathbf{T}$.

Step 2: Form a sequence $\mathbf{S}_0$ of variables in $\mathbf{O}$ of the same length as $\mathbf{T}$, which we assume contains m variables. For each $0 \leq n \leq m-1$, if $\mathbf{T}[n]$ is in $\mathbf{O}$, then $\mathbf{S}_0[n] = \mathbf{T}[n]$; otherwise $\mathbf{T}[n]$ is a (latent) collider on p, which, by the fact that p is d-connecting given $\mathbf{C}$, implies that there is a directed path from $\mathbf{T}[n]$ to a member of $\mathbf{C}$. So in this case, $\mathbf{S}_0[n]$ is chosen to be the first observed variable on a directed path from $\mathbf{T}[n]$ to a member of $\mathbf{C}$.

Step 3: Run the following iterative procedure:

k:=0

Repeat

If in $\mathbf{S}_k$ there is a triple of vertices $\langle \mathbf{S}_k[i-1], \mathbf{S}_k[i], \mathbf{S}_k[i+1] \rangle$ such that (1) there is an inducing path between $\mathbf{S}_k[i-1]$ and $\mathbf{S}_k[i]$ relative to $\mathbf{L}$ in $\mathcal{G}$ that is into $\mathbf{S}_k[i]$; (2) there is an inducing path between $\mathbf{S}_k[i]$ and $\mathbf{S}_k[i+1]$ relative to $\mathbf{L}$ in $\mathcal{G}$ that is into $\mathbf{S}_k[i]$; and (3) $\mathbf{S}_k[i]$ is in $\mathbf{C}$ and is an ancestor of either $\mathbf{S}_k[i-1]$ or $\mathbf{S}_k[i+1]$; then let sequence $\mathbf{S}_{k+1}$ be $\mathbf{S}_k$ with $\mathbf{S}_k[i]$ being removed;

k := k+1

Until there is no such triple of vertices in the sequence $\mathbf{S}_k$.

Let $\mathbf{S}_K$ denote the final outcome of the above three steps. Spirtes and Richardson (1996), in their Lemma 18, showed that $\mathbf{S}_K$ constitutes an m-connecting path between A and B relative to $\mathbf{C}$ in $\mathcal{M}$. We refer the reader to their paper for the detailed proof of this fact. What is left for us to show here is that the path constituted by $\mathbf{S}_K$ in $\mathcal{M}$ is either into A or out of A with an invisible edge.

In other words, we need to show that if the edge between $A = \mathbf{S}_K[0]$ and $\mathbf{S}_K[1]$ in $\mathcal{M}$ is $A \to \mathbf{S}_K[1]$, then this edge is invisible. Given Lemma 9, it suffices to show that there is an inducing path between

23. Their lemma addresses the more general case in which there may also be selection variables. The construction given here is an adaptation of theirs to fit our case.

24. In Spirtes and Richardson (1996), minimality means more than that the d-connecting path is a shortest one, but for this proof one only need to choose a shortest path.

A and $\mathbf{S}_K[1]$ relative to $\mathbf{L}$ in $\mathcal{G}$ that is into A. This is not hard to show. In fact, we can show by induction that for all $0 \leq k \leq K$, there is in $\mathcal{G}$ an inducing path between A and $\mathbf{S}_k[1]$ relative to $\mathbf{L}$ that is into A.

In the base case, notice that either (i) $\mathbf{S}_0[1]$ is an observed variable on p such that every variable between A and $\mathbf{S}_0[1]$ on p, if any, belongs to $\mathbf{L}$ and is a non-collider on p, or (ii) $\mathbf{S}_0[1]$ is the first observed variable on a directed path d starting from $\mathbf{T}[1]$ such that $\mathbf{T}[1]$ belongs to $\mathbf{L}$, lies on p and every variable between A and $\mathbf{T}[1]$ on p, if any, belongs to $\mathbf{L}$ and is a non-collider on p. In case (i), $p(A, \mathbf{S}_0[1])$ is an inducing path relative to $\mathbf{L}$, which is into A, because p is into A. In case (ii), consider $p(A, \mathbf{T}[1])$ and $d(\mathbf{T}[1], \mathbf{S}_0[1])$. Let W be the variable nearest to A on $p(A, \mathbf{T}[1])$ that is also on $d(\mathbf{T}[1], \mathbf{S}_0[1])$. ($W$ exists because $p(A, \mathbf{T}[1])$ and $d(\mathbf{T}[1], \mathbf{S}_0[1])$ at least intersect at $\mathbf{T}[1]$.) Then it is easy to see that a concatenation of $p(A,W)$ and $d(W, \mathbf{S}_0[1])$ forms an inducing path between A and $\mathbf{S}_0[1]$ relative to $\mathbf{L}$ in $\mathcal{G}$, which is into A because p is into A.

Now the inductive step. Suppose there is in $\mathcal{G}$ an inducing path between A and $\mathbf{S}_k[1]$ relative to $\mathbf{L}$ that is into A. Consider $\mathbf{S}_{k+1}[1]$. If $\mathbf{S}_{k+1}[1] = \mathbf{S}_k[1]$, it is trivial that there is an inducing path between A and $\mathbf{S}_{k+1}[1]$ that is into A. Otherwise, $\mathbf{S}_k[1]$ was removed in forming $\mathbf{S}_{k+1}$. But given the three conditions for removing $\mathbf{S}_k[1]$ in *Step 3* above, we can apply Lemma 32 (together with the inductive hypothesis) to conclude that there is an inducing path between A and $\mathbf{S}_{k+1}[1] = \mathbf{S}_k[2]$ relative to $\mathbf{L}$ in $\mathcal{G}$ that is into A. This concludes our argument. ■

Proof of Lemma 22

Proof This lemma is fairly obvious given Lemma 10. Let u be the path m-connecting A and B given $\mathbf{C}$ in $\mathcal{M}$. Let D (which could be B) be the vertex next to A on u. Construct a DAG $\mathcal{G}$ from $\mathcal{M}$ in the usual way: keep all the directed edges, replacing each bi-directed edge $X \leftrightarrow Y$ with $X \leftarrow L_{XY} \rightarrow Y$. Furthermore, if the edge between A and D is $A \rightarrow D$, it is invisible, so we can add $A \leftarrow L_{AD} \rightarrow D$ to the DAG. Then $\mathcal{G}$ is a DAG represented by $\mathcal{M}$. It is easy to check that there is a d-connecting path in $\mathcal{G}$ between A and B given $\mathbf{C}$ that is into A. ■

Proof of Lemma 23

Proof Note that because A is not an ancestor of any member of $\mathbf{C}$, if there is a path out of A d-connecting A and B given $\mathbf{C}$ in $\mathcal{G}$, the path must be a directed path from A to B. For otherwise there would be a collider on the path that is also a descendant of A, which implies that A is an ancestor of some member of $\mathbf{C}$. The sub-sequence of that path consisting of observed variables then constitutes a directed path from A to B in $\mathcal{M}$, which is of course out of A and also m-connecting given $\mathbf{C}$ in $\mathcal{M}$. The converse is as easy to show. ■

Proof of Lemma 27

Proof A path definitely m-connecting A and B given $\mathbf{C}$ in $\mathcal{P}$ is m-connecting in every MAG represented by $\mathcal{P}$, which is an immediate consequence of the definition of PAG. Let D be the vertex next to A on the definite m-connecting path in $\mathcal{P}$ between A and B given $\mathbf{C}$. All we need to show is that if the edge between A and D is not a definitely visible edge $A \rightarrow D$ in $\mathcal{P}$, then there exists a MAG represented by $\mathcal{P}$ in which the edge between A and D is not a visible edge out of A.

Obviously if the edge in $\mathcal{P}$ is not $A \rightarrow D$, there exists a MAG represented in $\mathcal{P}$ in which the edge is not $A \rightarrow D$, which follows from the fact that $\mathcal{P}$, by definition, displays all edge marks that are shared by all MAGs in the equivalence class.

So we only need to consider the case where the edge in $\mathcal{P}$ is $A \to D$, but it is not definitely visible. Now we need to use a fact proved in Lemma 3.3.4 of Zhang (2006, pp. 80): that we can turn $\mathcal{P}$ into a MAG by first changing every $\circ\!\!\to$ edge in $\mathcal{P}$ into a directed edge $\to$, and then orienting the circle component—the subgraph of $\mathcal{P}$ that consists of $\circ\!\!-\!\!\circ$ edges—into a DAG with no unshielded colliders.[25] Moreover, it is not hard to show, using the result in Meek (1995a), that we can orient the circle component—a chordal graph—into a DAG free of unshielded colliders in which every edge incident to A is oriented out of A.

Let the resulting MAG be $\mathcal{M}$. We show that $A \to D$ is invisible in $\mathcal{M}$. Suppose for the sake of contradiction that it is visible in $\mathcal{M}$. Then there exists in $\mathcal{M}$ a vertex E not adjacent to D such that either $E *\!\!\to A$ or there is a collider path between E and A that is into A and every collider on the path is a parent of D. In the former case, since $A \to D$ is not definitely visible in $\mathcal{P}$, the edge between E and A is not into A in $\mathcal{P}$, but then that edge will not be oriented as into A by our construction of $\mathcal{M}$. So the former case is impossible.

In the latter case, denote the collider path by $\langle E, E_1, ..., E_m, A\rangle$. Obviously every edge on $\langle E_1, ..., E_m, A\rangle$ is bi-directed, and so also occurs in $\mathcal{P}$ (because our construction of $\mathcal{M}$ does not introduce extra bi-directed edges). There are then two cases to consider:

Case 1: The edge between E and E_1 is also into E_1 in $\mathcal{P}$. Then the collider path appears in $\mathcal{P}$. We don't have space to go into the details here, but there is an orientation rule in constructing PAGs that makes use of a construct called "discriminating path" (e.g., Spirtes et al., 1999; Zhang, forthcoming), which would imply that if the collider path appears in $\mathcal{P}$, and every E_i $(1 \leq i \leq m)$ is a parent of D in a representative MAG $\mathcal{M}$, then every E_i is also a parent of D in $\mathcal{P}$. It follows that $A \to D$ is definitely visible in $\mathcal{P}$, a contradiction.

Case 2: The edge between E and E_1 is not into E_1 in $\mathcal{P}$, but is oriented as into E_1 in $\mathcal{M}$. This is possible only if the edge is $E \circ\!\!-\!\!\circ E_1$ in $\mathcal{P}$. But we also have $E_1 \leftrightarrow E_2$ (E_2 could be A) in $\mathcal{P}$, which, by Lemma 3.3.1 in Zhang (2006, pp. 77), implies that $E \leftrightarrow E_2$ is in $\mathcal{P}$. Then $\langle E, E_2, \ldots, A\rangle$ makes $A \to D$ definitely visible in $\mathcal{P}$, which is a contradiction. ∎

Proof of Lemma 28

Proof Note that since A does not have a descendant in $\mathbf{C}$, an m-connecting path out of A given $\mathbf{C}$ in $\mathcal{M}$ has to be a directed path from A to B such that every vertex on the path is not in $\mathbf{C}$. Then a shortest such path has to be uncovered,[26] and so will correspond to a definite m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{P}$ (on which every vertex is a definite non-collider). This path is not into A in $\mathcal{P}$ because $\mathcal{P}$ is the PAG for $\mathcal{M}$ in which the path is out of A. ∎

Proof of Lemma 29

Proof Let D be the vertex next to A on the definite m-connecting path in $\mathcal{P}$. Since the edge between A and D is not into A in $\mathcal{P}$, there exists a MAG represented by $\mathcal{P}$ in which the edge is out of A (which follows from the definition of PAG). Such a MAG obviously satisfies the lemma. ∎

25. A triple of vertices $\langle X, Y, Z\rangle$ in a graph is called an *unshielded* triple if there is an edge between X and Y, an edge between Y and Z, but no edge between X and Z. It is an *unshielded collider* if both the edge between X and Y and the edge between Z and Y are into Y.
26. A path is called *uncovered* if every consecutive triple on the path is unshielded (cf. Footnote 25).

References

R.A. Ali, T. Richardson, and P. Spirtes. Markov equivalence for ancestral graphs. Technical Report 466, Department of Statistics, University of Washington, 2004.

S. Andersson, D. Madigan, and M. Pearlman. A characterization of Markov equivalence classes of acyclic digraphs. *The Annals of Statistics* 25(2):505-541, 1997.

D.M. Chickering. A transformational characterization of equivalent Bayesian network structures. In *Proceedings of Eleventh Conference on Uncertainty in Artificial Intelligence*, pages 87-98, Morgan Kaufmann, 1995.

D.M. Chickering. Optimal structure identification with greedy search. *Journal of Machine Learning Research* 3:507-554, 2002.

D. Geiger, T. Verma, and J. Pearl. Identifying independence in Bayesian networks. *Networks 20*, pages 507-534, 1990.

Y. Huang and M. Valtorta. Pearl's calculus of intervention is complete. In *Proceedings of 22nd Conference on Uncertainty in Artificial Intelligence*, pages 217-224, AUAI Press, 2006.

C. Meek. Causal inference and causal explanation with background knowledge. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, pages 403-411, Morgan Kaufmann, 1995a.

C. Meek. Strong completeness and faithfulness in Bayesian networks, In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, pages 411-418, Morgan Kaufmann, 1995b.

J. Pearl. Causal diagrams for empirical research. *Biometrika* 82:669-710, 1995.

J. Pearl. Graphs, causality and structural equation models. *Sociological Methods and Research* 27:226-284, 1998.

J. Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, UK, 2000.

J.W. Pratt and R. Schlaifer. On the interpretation and observation of laws. *Journal of Econometrics* 39:23-52, 1988.

T. Richardson and P. Spirtes. Ancestral graph Markov models. *The Annals of Statistics* 30(4):962-1030, 2002.

T. Richardson and P. Spirtes. Causal inference via ancestral graph models. In P. Green, N. Hjort, and S. Richardson, editors, *Highly Structured Stochastic Systems*. Oxford University Press, USA, 2003.

J. Robins. A new approach to causal inference in mortality studies with sustained exposure periods—applications to control of the healthy worker survivor effect. *Mathematical Modeling* 7:1393-1512, 1986.

S. Shimizu, P.O. Hoyer, A. Hyvarinen, and A. Kerminen. A linear non-Gaussian acyclic model for causal discovery. *Journal of Machine Learning Research* 7:2003-30, 2006.

I. Shpitser and J. Pearl. Identification of conditional interventional distributions. In *Proceedings of 22nd Conference on Uncertainty in Artificial Intelligence*, pages 437-444, AUAI Press, 2006.

P. Spirtes, C. Glymour, and R. Scheines. *Causation, Prediction and Search*. Springer-Verlag., New York, 1993. (2nd ed., MIT Press, Cambridge, MA, 2000.)

P. Spirtes, C. Meek, and T. Richardson. An algorithm for causal inference in the presence of latent variables and selection bias. In C. Glymour and G.F. Cooper, editors, *Computation, Causation, and Discovery*. MIT Press, Cambridge, MA, 1999.

P. Spirtes and T. Richardson. A polynomial time algorithm for determining DAG equivalence in the presence of latent variables and selection bias. In *Proceedings of the 6th International Workshop on Artificial Intelligence and Statistics*, 1996. URL http://citeseer.ist.psu.edu/spirtes97polynomial.html.

P. Spirtes and T. Verma. Equivalence of causal models with latent variables. Technical Report Phil-36, Department of Philosophy, Carnegie Mellon University, 1992.

J. Tian and J. Pearl. On the identification of causal effects. Technical Report, Department of Computer Science, Iowa State University, 2004.

J. Tian. Generating Markov equivalent maximal ancestral graphs by single edge replacement. In *Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence*, pages 591-598, AUAI Press, 2005.

C. Winship and L.S. Morgan. The estimation of causal effects from observational data. *Annual Review of Sociology* 25:659-706, 1999.

J. Zhang and P.Spirtes. A transformational characterization of Markov equivalence for directed acyclic graphs with latent variables. In *Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence*, pages 667-674, AUAI Press, 2005.

J. Zhang. *Causal Inference and Reasoning in Causally Insufficient Systems*. PhD dissertation, Department of Philosophy, Carnegie Mellon University, 2006. URL www.hss.caltech.edu/∼jiji/dissertation.pdf.

J. Zhang. On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias. *Artificial Intelligence*, forthcoming.

H. Zhao, Z. Zheng, and B. Liu. On the Markov equivalence of maximal ancestral graphs. *Science in China (Mathematics)*, 48(4):548-562, 2005.