Causal Inference and Reasoning in Causally Insufficient Systems

Jiji Zhang

Department of Philosophy

Carnegie Mellon University

July, 2006

Abstract

The big question that motivates this dissertation is the following: under what conditions and to what extent can passive observations inform us of the structure of causal connections among a set of variables and of the potential outcome of an active intervention on some of the variables? The particular concern here revolves around the common kind of situations where the variables of interest, though measurable themselves, may suffer from confounding due to unobserved common causes.

Relying on a graphical representation of causally insufficient systems called maximal ancestral graphs, and two well-known principles widely discussed in the literature, the causal Markov and Faithfulness conditions, we show that the FCI algorithm, a sound inference procedure in the literature for inferring features of the unknown causal structure from facts of probabilistic independence and dependence, is, with some extra sound inference rules, also *complete* in the sense that any feature of the causal structure left undecided by the inference procedure is indeed underdetermined by facts of probabilistic independence and dependence.

In addition, we consider the issue of quantitative reasoning about effects of local interventions with the FCI-learnable features of the unknown causal structure. We improve and generalize two important pieces of work in the literature about identifying intervention effects. We also provide some preliminary study of the testability of the Causal Faithfulness Condition.

Acknowledgements

There is no way I can express my gratitude to Peter Spirtes properly in a few sentences — I simply owe him too much in all these years at CMU. Peter's enormous influence of this dissertation is evident to anyone familiar with his own work, not to mention his direct contributions that will be acknowledged, only too briefly, in a few footnotes later. I feel very fortunate, in particular, to benefit from Peter's incredible intellectual respect for students, which, in my case, encouraged promising but initially unsophisticated ideas.

I am also very grateful to the other committee members. I thank Clark Glymour for his rich ideas, sharp insights, challenging questions, and his encouragements from time to time that inspired much self-confidence I needed. I greatly appreciate his academic and career advices all through these years. I thank Thomas Richardson for being an exact reader of my dissertation, without whose careful and insightful comments the dissertation would not have taken its present shape. I thank Richard Scheines for supervising my work despite his busy schedule, and for his great confidence in my ability.

I am glad that I chose the philosophy department at CMU to pursue my graduate study six years ago, despite advices to the contrary at that time. I feel lucky and proud to have been part of this open-minded and thought-provoking environment. I wish to thank especially Teddy Seidenfeld and Kevin Kelly for many, many long and rewarding conservations concerning statistics, decision theory, formal epistemology, and philosophy of science in general, which helped me to see great things beyond causal graphs. I would also like to thank Cristina Bicchieri for stimulating my interest in the fascinating field of experimental economics, and for her trust and confidence in me.

There are too many others to list here that have taught, helped or entertained me in the past six years, including friends in the United States who made me less nostalgic and friends back in China who made me more so. Thanks to all of them!

Finally, I want to thank my parents for their perpetual love, support and understanding, and my fiancee, Jing Zhan, for her constant encouragement and love that keep me optimistic about the future.

Contents

List of Figures

Chapter 1

Introduction

Probably few people today would agree with Hume that "all reasonings concerning matter of fact seem to be founded on the relation of Cause and Effect." Many predictions we make about yet unobserved features, e.g., tomorrow's weather or a woman's age, are based on information of statistical correlations among various features of the world, and oftentimes rightly so. In these cases we reason about matter of fact without having to rely on knowledge of cause and effect[1]. In other cases, however, the basis of reasoning does involve causal knowledge. This is so whenever the objects of inquiry concern consequences of actions, effects of policies or outcomes of interventions in general. As long as such questions are of interest, as they have been all along, causal inference and reasoning are unavoidable.

Unavoidable as they are, inferences about causal relations have been the cynosure of skeptical arguments ever since Hume. Skepticism in general embraces two ways that evidence could leave reality undetermined: global underdetermination – alternative, contradictory hypotheses may be compatible with all of the evidence there could

[1]This does not necessarily contradict what Hume had in mind. For example, Hume could be merely asserting that all empirical relations that reasonings are founded upon, including statistical/probablistic associations, are derivative (ontologically?) of the relation of cause and effect.

possibly be; and local underdetermination – alternative, contradictory hypotheses may be compatible with all of the available evidence at any time. Hume's critique of our causal knowledge invokes both kinds: if "causation" is meant to be some sort of "necessary connection" that transcends the observable regularity, then it is underdetermined in the first, global sense; alternatively, even if "causation" is defined as nothing more than the "constant conjunction of events", there is still the issue of inferring genuine regularities from the temporarily observed patterns, which suffers from underdetermination of the second sort.

This two-fold picture is largely inherited in the modern statistical inference of causality, which can be viewed as involving two general steps: inferring (objective) probabilities from sample statistics; and inferring causality from probabilities (cf. Papineau 1994). The first step – with primarily underdetermination of the second sort – has of course been a central subject of statistics. A typical antidote statisticians provide to the skeptical worry, for example, is to demonstrate long term properties of statistical inference – in particular, to prove asymptotic convergence to the true hypothesis (or at least asymptotic avoidance of false conclusions), and more powerfully, to calculate rates of the convergence or bounds of the errors. The second step – the one that motivates this dissertation as well as much work on probabilistic theories of causation, however, is sabotaged by underdetermination of the more radical sort, unless further assumptions are adopted that in some way restrict alternative causal hypotheses and their relations to possible data. Making further assumptions about causality is thus commonplace in treatises on causal inference methodologies. Mill, for example, made a principle of Hume's hope that similar causes produce similar effects (Mill 1843); Fisher implicitly assumed that potential causes could be randomized to eliminate confounding (Fisher 1935); and both of them were committed to some version of the well known principle of the common cause[2], which is of central import in the recent philosophical literature.

There are two types of projects associated with any candidate principles proposed to support causal inference. One may be called the "Humean" project and the other the "Millian" project. The former is concerned with the nature and status of the assumptions themselves, and seeks to provide a justification of the assumptions or, when no rational basis is found, a psychological explanation of why people believe in those assumptions. The latter, by contrast, is concerned with the epistemic consequences of the assumptions, and seeks to develop methodologies of causal inference justifiable *given* the assumptions. These two projects are obviously complementary but also largely orthogonal to each other, which makes it possible to pursue both in parallel. Philosophers tend to emphasize the first type of inquiry much more than the second, but the present dissertation, if only to counterbalance the tendency, will focus on the second type of inquiry.

Specifically, this dissertation aims to contribute to the methodologies of causal inference in observational, non-experimental settings based on graphical representations and methods, a field that has been drawing increasing attention, in the past three decades or so, from researchers in various disciplines such as philosophy, statistics, computer science and artificial intelligence. Before I synopsize what the contributions are, it is helpful to introduce some relevant background of the research program.

[2]Randomization is intended to do more than the elimination of confounding. For Fisher, it is an important device to create a well-defined sampling distribution (under the null hypothesis of no causal effect) so as to facilitate the calculation of significance levels. But to calculate the significance level, for example in the classical tea-tasting case, it is necessary to interpret what the null hypothesis of "no effect" means in terms of the relationship between the lady's responses and the schedules of adding milk, where it is assumed that "no effect" implies "no covariation" (or independence). Mill's commitment to the principle of the common cause, on the other hand, is evident in his fifth canon of induction.

1.1 Variable Causation and Directed Acyclic Graphs

A large body of philosophical work on causation is devoted to metaphysical issues. One problem, for example, is about relata of the causal relation: what is causation a relation *of*? There is yet nothing close to a consensus on this issue, or any reason to expect one is forthcoming. Indeed, if we look at the actual use of the word in daily discourse, we find a variety of bearers of causal relations. A cause can be an event, an object (or a feature thereof), a property, a fact, a state of affairs, or an agent, as far as linguistic evidence goes. Causation, to name another widely discussed topic, is also classified into token-level or singular causation on the one hand, and type-level or general causation on the other. There is the controversy about which level is more fundamental and whether one level is definable in terms of the other. The literature on this topic is again equivocal and does not exhibit any sign of convergence of opinions. Fortunately, the subject that concerns me in this dissertation does not depend crucially on settlement of these metaphysical issues.

My interest lies in one manifestation of causal relations that is probably of the most practical import among conceptions of cause and effect. No matter how vague and multifaceted is the meaning of the word "cause", a key intuition most people share is that causes make effects happen but not vice versa, and causes can be exploited to produce changes in their effects. In the simplest case, for example, we can imagine "taking away" a cause (be it an event, an object or a fact) that was present, and as such producing a change from an effect being present to that effect being absent (or from the effect being very probably present to being less probably present). Indeed a major practical reason for acquiring knowledge of causal relations is to predict effects of interventions, a central task in many decision and policy making scenarios. For this purpose it is convenient to conceive causal influence as operating among *variables*, and to equate the presence or absence of causal influence with the potential variance or invariance upon manipulation.

Some philosophers maintain that causal relations are relations between contrast pairs or transitional events (Hitchcock 1996, Belnap 2005). Under such views, variables are a suitable surrogate for the fundamental causal relata, as a contrast or a transition is naturally represented by a variable taking different values. On the other hand, it is fairly straightforward to define variable causation in terms of event causation or object causation or what have you, as long as the fundamental relata can be represented by variables being instantiated to specific values. For example, we can generically define a variable A being a cause of another variable B in terms of *some value* of A being a cause of *some value* of B. If the appropriate causal relata are states of affairs, then different values of A or B correspond to different states of affairs; if the appropriate causal relata are events, then different values of A or B correspond to different events (or non-happening of events), and so forth. In this way, the controversy over causal relata translates into a controversy about the ontological category of the range of the "right" variables for causal analysis, but does not affect the intelligibility of the notion of variable causation. So, to make sense of variable causation does not have to wait for a resolution of the debate over causal relata.

Variable causation is often viewed as type-level causation in the causal modelling literature, rightly so because the purpose is usually to model a type of systems or a population of individuals that admits sampling and statistical analysis. Conceptually, however, the notion of variable causation seems to cut through both type-level and token-level. We speak of both generic variables, such as brain weight in general, and specific variables, such as Albert Einstein's brain weight in particular, just as we speak of both generic and concrete events (Strevens 2003). Accordingly we may talk about general variable causation as well as singular variable causation. Again which level is ontologically basic is controversial, but that controversy does not matter from an epistemic point of view.

Variable causation is amenable to formal analysis in that it naturally admits mathematical representations. The mathematical representations that will play a major role in this dissertation are graphs, consisting of vertices that represent variables[3], and various types of edges between pairs of vertices that represent relationships between variables. The tradition of representing causal relations graphically can be traced back at least to Sewell Wright's pioneering work in genetics (e.g., Wright 1921, 1934). In recent years graphical models have also received a lot of attention in statistics (Whittaker 1990, Lauritzen 1996) and computer science (Pearl 1988, Jordan 1998, Neapolitan 2004).

The most widely used graphs in causal modelling are directed acyclic graphs (DAGs). A *directed graph (DG)* is a mathematical object consisting of a pair $\langle \mathbf{V}, \mathbf{E} \rangle$, where $\mathbf{V}$ is a set of vertices and $\mathbf{E}$ is a set of arrows. An arrow is an ordered pair of nodes, say $\langle A, B \rangle$, represented visually by $A \rightarrow B$. Given a graph $\mathcal{G}(\mathbf{V}, \mathbf{E})$, if $\langle A, B \rangle \in \mathbf{E}$, then A and B are said to be *adjacent*, and A is called a *parent* of B and B a *child* of A. A *path* in $\mathcal{G}$ is a sequence of distinct vertices $\langle V_0, ..., V_n \rangle$ such that for $0 \leq i \leq n-1$, V_i and V_{i+1} are adjacent in $\mathcal{G}$. A *directed path* in $\mathcal{G}$ from A to B is a sequence of distinct vertices $\langle V_0, ..., V_n \rangle$ such that $V_0 = A$, $V_n = B$ and for $0 \leq i \leq n-1$, V_i is a parent of V_{i+1} in $\mathcal{G}$, i.e., all arrows on the path point in the same direction. A is called an *ancestor* of B and B a *descendant* of A if $A = B$ or there is a directed path from A to B. DAGs are simply DGs in which there are no directed cycles, or in other words, there are no two distinct vertices in the graph that are ancestors of each other.

[3]Throughout I will talk about vertices and variables interchangeably.

DAGs are a great tool in statistical modelling in virtue of their standard probabilistic semantics: a DAG over a set of random variables encodes a set of conditional independence constraints on the joint probability distribution by its Markov property. The (local) Markov property specifies that every variable in the DAG is independent of its non-descendants (i.e., variables that are not its descendants) conditional on its parents. These conditional independence constraints entail still others, all of which can be read off the DAG by a graphical criterion known as d-separation (Pearl 1988), which will be defined later in Chapter 2.

More importantly for our purpose, DAGs have a natural causal semantics: vertices represent variables, and arrows represent direct causal relationship between pairs of variables.[4] By "direct causal relationship" it is meant that there is no variable that mediates the relationship, or intuitively, that there exists a manipulated change in the variable cause that will be followed by a change in the variable effect, while holding all other variables fixed.[5] The last clause makes it clear that it only makes sense to talk about "direct causes" relative to a given set of variables. Arrows in a DAG, we stipulate, represent "direct causation" relative to the variables in the DAG. An obvious limitation of acyclic graphs is that they cannot represent feedback or non-recursive mechanisms. To represent and infer feedback via cyclic graphs is a difficult subject that deserves a whole other dissertation (see Richardson 1996), and will not be considered in this one.

[4]Strictly speaking, I would also include the manipulation principle introduced in Chapter 5 to be an important component of the causal interpretation of DAGs.

[5]As emphasized by almost everyone working in this field, this account of "direct cause" in terms of manipulations does not make a definition in the usual reductive sense, because manipulation is not a non-causal term. However, the elucidation of one causal notion by way of another causal notion does not need to be fruitlessly circular, as Woodward (2003) vigorously argued.

1.2 Causal Sufficiency, Markov and Faithfulness Conditions

Hume is commonly regarded as identifying cause and effect with merely constant conjunction of events or objects (plus spacial contiguity and time order). This prototype of the so-called regularity theory of causation encountered various objections from the very beginning (e.g., Reid 1788), but the motivating insight is seldom denied, i.e., the only aspect of "causation" that can have direct empirical basis is constant conjunction, or from our point of view, statistical associations in general (plus perhaps a time order). By contrast, causation beyond constant conjunction does not lend itself to direct observation.

The recent advance in causal modelling does not refute this point, but starts by positing bridge principles that connect causal structures to correlation patterns. As I said, this dissertation is for the most part a sort of "Millian" effort, an effort to explore the epistemic and methodological consequences of certain principles. The principles I will rely upon are precisely two bridge principles that are becoming influential: the *Causal Markov Condition (CMC)* and the *Causal Faithfulness Condition (CFC)*. In particular, I aim to continue and improve the existing study of the consequences of these two principles.

The CMC is closely related to the concept of causal sufficiency. Given a set of variables $\mathbf{V}$, also referred to as a causal system in this dissertation, 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 is a direct cause of A and also a direct cause of 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 $\mathbf{V}$. Only for causally sufficient systems do we have reason to assume the CMC. (A number of counterexamples to the condition in causally insufficient systems can be found in Spirtes et al. 1993/2000, pp. 33-35, page number referring to the 2nd edition.)

Given a causally sufficient set of variables, the CMC states that every variable in the set is probabilistically independent of its non-effects (i.e., variables that are not its effects) conditional on its direct causes. This condition obviously brings together the probabilistic semantics and the causal semantics of DAGs, so that the same DAG can represent a set of variables both causally and probabilistically. The CFC is the converse principle, which, to put it simply, states that no conditional independence other than those entailed by the CMC holds. I will explain the content in more detail in the next chapter. For now, note that the two conditions together provide a link between causality and probability, which, if valid, opens up the possibility of inferring causal information from probabilistic information.

1.3 A Representation of Causally Insufficient Systems: Maximal Ancestral Graphs

A major worry among statisticians towards inferring causation from correlation is that there are unobserved or latent variables that contribute to the observed correlation pattern among observed variables – variables of which we can and do measure the values. In such cases the set of observed variables may be causally insufficient.

For a causally insufficient system $\mathbf{O}$, DAGs over $\mathbf{O}$ do not provide a satisfactory representation, because the CMC typically fails of $\mathbf{O}$.[6] The causal DAG over $\mathbf{O}$ – i.e.,

[6]By this I mean that some variable in $\mathbf{O}$ is not probabilistically independent of its non-effects conditional on its direct causes *in the set*.

the DAG in which there is an arrow from A to B, $A \to B$, if and only if A is a direct cause of B relative to $\mathbf{O}$ – does not represent the joint probability over $\mathbf{O}$ properly in the sense of entailing the right conditional independence constraints.[7] On the other hand, a DAG that represents the joint probability correctly cannot be causally accurate. So no DAG over just the variables in $\mathbf{O}$ can do the dual representational job.

Suppose, as I will do throughout the dissertation, that a causally insufficient system can nonetheless be extended to a causally sufficient system by adding a finite number of extra variables. Then in theory we can represent a causally insufficient set of observed variables $\mathbf{O}$ both causally and probabilistically by a DAG over $\mathbf{V} = \mathbf{O} \cup \mathbf{L}$, where $\mathbf{L}$ is a set of unobserved or latent variables added to make the whole set $\mathbf{V}$ causally sufficient.[8] In this dissertation, however, I will not directly work with DAGs with latent variables. The reason is two-fold. First, there are well-known undesirable statistical properties of DAG models with latent variables. For example, DAG models with latent variables are not always identified (Bollen 1989); DAG models with latent variables are usually not curved exponential families (Geiger et al. 2001), and hence such commonly used score as BIC may not be a consistent model selection criterion (Haughton 1988); Moreover, it is in general difficult to even calculate such penalized likelihood scores for DAG models with latent variables, because they in general do not have a well-defined dimension, being unions of curved exponential families of different dimensions (Geiger et al. 2001); and so forth.

[7] Another more important complication is that the causal DAG of a causally insufficient system does not enable us to identify all effects of interventions on the system, even if the CMC happens to hold of the system. In other words, the manipulation principle to be discussed in Chapter 5 fails. This, however, is an inherent problem with causally insufficient systems that will not be resolved by adopting the ancestral graphical representation that I will rely upon. We will come back to this issue in Chapter 5.

[8] More generally there can also be what we call selection variables, which will be considered in Chapters 3 and 4.

Second, and more importantly, occasions often arise in which there is very little information about unmeasured variables. Sometimes researchers fail to measure a variable because it is not easy or practically impossible to measure, but other times researchers fail to measure a variable because they have no idea about the existence of the variable or the relevance of the variable never comes into mind. In the latter situations, except for the suspicion that the set of observed variables is insufficient, there is no reason to posit anything specific about latent variables, such as the number of the relevant latent variables and their locations in the causal network. We then face the following dilemma: one horn is to put no constraint on latent variables, which implies, among other difficulties, that there are literally infinitely many DAGs with latent variables to consider as candidate causal models; the other horn is to posit groundless constraints on latent variables, which is undesirable because inference with latent variable DAG models is sensitive to the assumptions made about latent variables.

I will thus rely on an alternative graphical representation of causally insufficient systems developed in Richardson and Spirtes (2002), which does not explicitly include latent variables and represents them only implicitly. Again, this choice means that I target at those situations where a set of observed variables is not known or assumed to be causally sufficient, but little is known about the extra relevant variables, if any. Moreover the query itself only concerns the observed variables, such as what would happen to an observed variable Y if another observed variable X is intervened to take some value, and there is no interest in latent variables *per se*, except that they may be relevant to the inference about those observed variables. As will become clear, however, even when latent variables themselves are of interest, the methodology based on the representation to be introduced is not useless, as the method, among other things, may indicate the locations of latent variables in the causal network.

The alternative representation is a generalization of DAGs. A *directed mixed graph*[9] is just like a directed graph except that it can contain, besides directed edges or arrows ($\rightarrow$), also bi-directed edges or double-headed arrows ($\leftrightarrow$). All graphical notions introduced earlier, *adjacency, parent/child, ancestor/descendant, path* and *directed path*, obviously remain meaningful. In addition, if there is a bi-directed edge $A \leftrightarrow B$ in a directed mixed graph G, then A is called a *spouse* of B and B a spouse of A. An *almost directed cycle* occurs if there are two variables A and B such that A is both an ancestor and a spouse of B.

Given a path $u = \langle V_0, ..., V_n \rangle$ with $n > 1$, V_i $(1 \leq i \leq n-1)$ is a *collider* on u if the two edges incident to V_i are both into V_i, i.e., have an arrowhead at V_i; otherwise it is a *noncollider* on u. A path is called a *collider path* if every vertex on it (except for the endpoints) is a collider along the path. Let $\mathbf{L}$ be any subset of vertices in $\mathcal{G}$, 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. When $\mathbf{L}$ is empty we simply call the path an *inducing path.*[10]

The representation to be used extensively later is called a maximal ancestral graph (MAG). A mixed graph is *ancestral* if it does not contain any directed or almost directed cycle. It is *maximal* if no inducing path is present between any two non-adjacent vertices in the graph. A *MAG* is a mixed graph that is both ancestral and maximal. Note that syntactically a DAG is a special case of MAG, simply a MAG without bi-directed edges.

A nice feature of MAGs is that they can represent the marginal independence models of DAGs in the following sense: given any DAG $\mathcal{G}$ over $\mathbf{V} = \mathbf{O} \cup \mathbf{L}$, there is

[9]Richardson and Spirtes (2002) actually considered more general mixed graphs that can also include undirected edges (—). Chapters 3 and 4 will consider this general case.

[10]It is called a *primitive inducing path* in Richardson and Spirtes (2002).

a MAG over $\mathbf{O}$ alone such that for any three disjoint sets of variables $\mathbf{A}, \mathbf{B}, \mathbf{C} \subseteq \mathbf{O}$, if $\mathbf{A}$ and $\mathbf{B}$ are entailed to be independent conditional on $\mathbf{C}$ by $\mathcal{G}$ (according to the Markov property of DAGs) if and only if the conditional independence is entailed by the MAG, by the Markov property of MAGs to be formally defined in Chapter 2. 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 a inducing path between them relative to $\mathbf{L}$ in $\mathcal{G}$;

2. for each pair of adjacent vertices A, B in $\mathcal{M}_{\mathcal{G}}$, orient the edge as $A \rightarrow B$ in $\mathcal{M}_{\mathcal{G}}$ if $A \in \mathbf{An}_{\mathcal{G}}(B)$; orient it as $A \leftarrow B$ in $\mathcal{M}_{\mathcal{G}}$ if $B \in \mathbf{An}_{\mathcal{G}}(A)$; 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). More importantly, notice that $\mathcal{M}_{\mathcal{G}}$ also retains the ancestral relationships — and hence causal relationships under the standard interpretation — among $\mathbf{O}$ in $\mathcal{G}$. 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 $\mathbf{O}$. It is thus hopeful to infer causal information from probabilistic information via MAGs. There are of course also limitations of working with MAGs. First, it is not the case that the causal MAG retains all causal information needed to predict all effects of interventions among the observed variables (see Chapter 5), as different DAGs with latent variables can entail different intervention effects but correspond to the same MAG, even though a MAG does retain all qualitative information regarding whether an observed variable X has a causal influence on another observed variable Y.

Second, MAGs do not give any clue about the causal structure among latent variables, except for indicating where the relevant latent variable might be. I will not be concerned with queries about causal structure among the unobserved in the present work, and I refer interested readers to Silva et al. (2006) for a novel treatment of the issue.

1.4 Overview of the Dissertation

The foregoing background description probably falls short of rigor, most of which will be formally introduced again later. But it should be enough for the current overview. My primary interest in the dissertation is to study the consequence of the CMC and the CFC for causal inference, so I will not attempt to dig into the literature that debates about the status of these principles themselves. Still, in Chapter 2, I will try to shed some light on the testability of the Causal Faithfulness Condition. In particular, I will present a simple finding that, surprisingly enough, has not been noticed or emphasized before. The simple fact is that the CFC can be dissected into at least two components, one of which is in principle testable given the other. This observation suggests a simple twist to a familiar causal discovery procedure, known as the PC algorithm (Spirtes et al. 1993/2000). The resulting algorithm, called Conservative PC (CPC), is provably correct under a weaker-than-standard Faithfulness condition. I shall also argue, in a preliminary fashion, that the CPC algorithm is in a sense robust against limited violations of the CMC, given that the CFC (or for that matter, the weaker Faithfulness condition) holds. In addition to presenting these substantive results, this chapter will also serve the purpose of describing relevant background of causal discovery under the assumption of causal sufficiency, which is helpful for understanding later chapters on causal discovery without the assumption of causal sufficiency.

The rest of the dissertation will take the two principles, CMC and CFC, as given, and investigate the possibility and limit of inferring causal information from *facts of probabilistic independence and dependence* given these two principles.[11] First of all, the issue is to infer qualitative causal structures, represented by causal graphs, from patterns of probabilistic independence and dependence. It is still true under the two assumptions that correlation does not imply causation. In general, the true causal graph is underdetermined by a pattern of correlations, or in other words, there are multiple causal graphs that satisfy the CMC and CFC with the given pattern of conditional independence constraints – such graphs are what we call *Markov equivalent*. Nonetheless, these Markov equivalent causal structures can and do share some common features that are hence uniquely determined. An important project is then to fully characterize the extent of underdetermination due to Markov equivalence, and to develop a feasible calculus that can derive all valid qualitative causal information entailed by a pattern of conditional independence constraints, assuming the CMC

[11]Let me note a couple of limitations right away. First, the qualitative information about conditional independence or dependence is presumably not the only type of probabilistic information that one can employ to infer causal information. Given suitable parametric assumptions, for example, there has been an important thread of work that focuses on inferring causal structure of linear models from a type of probabilistic constraint known as the *tetrad* constraints (Glymour, et al. 1987, Spirtes, et al. 1993/2000, Silva, et al. 2003, 2006). It is also known that some DAGs with latent variables entail non-parametric constraints on the marginal probability over observed variables that do not take the form of conditional independence. There has been some detailed study of how to derive these constraints from a given DAG (e.g., Tian and Pearl 2002), but the reserve direction, i.e., how these non-parametric constraints may be employed to recover the unknown causal graph, is not yet well understood. So the current dissertation does not aim to reveal the full implication of the CMC and CFC for causal inference, but rather the implication of the two principles for causal inference based on conditional independence and dependence facts. In some sense, the implication of the latter sort will be fully given by this work.

Second, I will bypass the step of statistical inference to the conditional independence and dependence facts. It is undoubtedly an extremely important issue that will engage statisticians for a long time. In the subsequent chapters I will often talk about a *perfect oracle of conditional independence*, which is only available in the large sample limit provided that there are consistent statistical tests of conditional independence.

and CFC.

If the set of observed variables can be assumed to be causally sufficient, the target causal graph is a DAG. For this case, there are sound and complete algorithms for extracting causal information out of an oracle of probabilistic independence, an oracle that is assumed to decide queries about conditional independence correctly (e.g. Meek 1995). The output of such algorithms, a graphical object that represents an equivalence class of causal structures, displays all and only those common features shared by all causal structures that satisfy the CMC and CFC with the oracle.

However, it is seldom (or never, according to some epidemiologists) unproblematic to assume that the observed causal system is causally sufficient. Without the assumption of causal sufficiency, the target causal graph is in general not a DAG over the observed variables, but rather, as we briefly explained, a MAG over the observed variables (since we rely on conditional independence and dependence facts only). In this regard, there is a sound algorithm, known as the FCI algorithm, for extracting causal information out of an oracle of conditional independence (Spirtes et al. 1999), but it is unknown whether the algorithm is complete. To establish completeness amounts to fully characterizing commonalities in an equivalence classes of MAGs, which turns out to be substantially more difficult than the analogous task for DAGs.

I will tackle this completeness problem in Chapters 3 and 4. Chapter 3 will prove that the FCI algorithm is actually "arrowhead complete", complete regarding common arrowheads among Markov equivalent MAGs. In Chapter 4, I point out that the FCI algorithm is not "tail complete", i.e., not complete regarding arrow tails shared by all Markov equivalent MAGs. I then present an augmented FCI algorithm which proves to be complete. This part on completeness is probably the most difficult and significant portion of the dissertation, which, in addition to the theoretical significance in its own right, have potential applications in further research.

For example, in Chapter 4 I will also present a result which is to a large extent a by-product of the arguments used there to prove completeness. The result is a transformational characterization of Markov equivalent MAGs, MAGs that are indistinguishable given an oracle of conditional independence. As I shall explain in due time, I expect this result to be useful for several purposes, given that the analogous result for DAGs established by Chickering (1995) found several interesting applications.

I begin this introduction by indicating that we care about acquiring causal information not because it constitutes the basis for all actual and possible empirical reasonings, but because it enables reasonings about effects of certain changes. In Chapter 5, I address the issue of quantitative causal reasoning that intends to make use of the qualitative causal information inferred from the FCI algorithm. In a general form, the inference problem of interest is this: given a set of observed random variables, whose joint probability distribution can be consistently inferred from (non-experimental) data, we want to figure out, for three subsets of variables $\mathbf{X}, \mathbf{Y}$ and (possibly empty) $\mathbf{Z}$, whether or not we can infer the probability of $\mathbf{Y}$ conditional on $\mathbf{Z}$) given that $\mathbf{X}$ were intervened to follow certain probability distribution (Spirtes et al. 1993/2000). We will specify the formal details of the intervention in question in Chapter 5, but essentially the inference is about the outcome of certain "changes" imposed from outside the given system based on information about the probability distribution *before* the change.[12]

[12]Put this way, the problem looks much harder than what Hume was worried about. The kind of induction Hume wrestled with would be warranted by a principle of the uniformity of nature, as Hume apparently thought. There can be, of course, many versions of the uniformity principle, but Hume's version – that similar causes are followed by similar effects – is obviously too vague and simple to even indicate how an inference of intervention effects can be carried out. The kind of uniformity or invariance principle needed for this inference will be discussed in Chapter 5.

The attempted solution to this problem proceeds in two stages: (1) extract qualitative information of the causal structure – represented by a causal graph – from the joint probability distribution, say, from the conditional independence and dependence relations implies by the joint distribution; and (2) predict the effect of certain interventions based on the acquired causal information and the joint distribution over the observed variables before interventions.[13] The causal information obtainable in (1) is not always enough for identifying a given causal quantity, so an important subtask in (2) is to determine whether the quantity of interest is predictable at all given the information about causal structure gained in (1).

Stage (1) is tackled in Chapters 3 and 4 (with the limitation that only conditional independence and dependence facts are exploited). Chapter 5 thus deals exclusively with stage (2). An output from the (augmented) FCI algorithm is assumed to be given, and the question is how much quantitative reasoning is warranted. My contribution here will be to improve or generalize two existing methodologies, one due to Spirtes et al. (1993/2000) and the other due to Pearl (2000). The former has a theory of invariance based on a graphical representation called "inducing path graphs", which is a less elegant representation than ancestral graphs. I will thus develop a similar, and in an important aspect better, theory of invariance based on MAGs. The latter is known for its seminal work on causal reasoning based on a (single) causal DAG with latent variables. But, as one would naturally worry, the true causal DAG is seldom fully discoverable. I will hence adapt Pearl's celebrated *do*-calculus to cases where the available causal information is summarized by an output of the FCI algorithm.

I conclude the dissertation by pointing out a few open problems in Chapter 6.

---

[13]In practice, the pre-intervention distribution needs to be inferred from data, and hence there is usually also a step of estimating parameters.

Chapter 2

Decomposition and Testability of the Causal Faithfulness Condition

Much of the recent advance in causal modelling starts by explicitly acknowledging the Causal Markov Condition (CMC) and the Causal Faithfulness Condition (CFC) as axioms. Any principle that receives such a treatment will almost certainly be subject to philosophical scrutiny. Since my purpose is not to defend these two principles, but rather to explore their implications on the epistemology and methodology of causal inference, I will leave aside the literature that attempts to justify or criticize them. A large part of that literature is devoted to the CMC, which, for the most part, centers around a few problem cases, such as EPR-like funny business (e.g., Artzenius 1992, Steel 2003), apparent correlations between causally unconnected time series (e.g., Yule 1926, Sober 1987, 2001, Hoover 2003), and examples of the sort of Wesley Salmon's interactive fork (e.g., Salmon 1980, Cartwright 1999, Hausman and Woodward 2004). Fewer reflections are cast on the CFC, perhaps because it looks more like a standard methodological assumption of simplicity than a substantive posit about the world. It is clear, however, that standard asymptotical justifications of causal discovery procedures regard the CFC as substantive (Spirtes et al. 1993/2000). As a substantive principle, the CFC incurs criticisms that cite versions of Simpson's paradox or more generally cases where multiple causal pathways exactly cancel each other, and cases where causal transitivity fails. Relevant discussions have thus focused on how often or rarely cancellation or failure of transitivity could occur, or in what domains would the CFC be particularly shaky or safe. In this chapter I shall adopt a different perspective. I will consider the issue of detecting violations of the CFC and the possibility of relaxing the assumption of CFC within a standard approach to causal discovery.

This chapter is distinctive in this dissertation in that it studies the epistemic status of the CFC itself, albeit from a different angle than the normal one in the literature, whereas later chapters will take the CMC and CFC as axioms and study what follow from them. The results to be presented are preliminary in a couple of respects. Most notably they are established in the context of causal discovery with the assumption of causal sufficiency, unlike the rest of the dissertation that deals with causally insufficiency systems. However, it should become clear later that the work can be extended to the context of causally insufficient systems, which I list as an open but relatively straightforward project in the concluding chapter. For now, this apparent limitation also gives us an opportunity to go through some details of causal discovery for causally sufficient systems, which will be helpful for understanding later chapters as well.

In the main I will focus on a decomposition of the CFC into two components that suggests a simple test of one component assuming the other. I report a piece of joint work with Joseph Ramsey and Peter Spirtes on constructing a conservative version of the well-known PC procedure for inferring causal structure from independence/dependence facts assuming causal sufficiency.[1] The resulting algorithm proves to be superior in both theoretical properties and empirical performance.

2.1 A Simple Note on Testing the CFC

The general issue is whether and how the CFC can be tested, assuming the CMC holds. Let us first recall what the CFC says. The CFC essentially states the converse of the CMC. The CMC says, for a given set of variables $\mathbf{V}$, that every variable is probabilistically independent of its non-effects conditional on its direct causes relative to $\mathbf{V}$. This condition (as thus formulated) typically fails if the set of variables is not causally sufficient or there is causal feedback among the given variables. Note that, as explained in Chapter 1, either of these scenarios will render DAGs (over the given set of variables alone) an "improper" representation of the causal structure. In this chapter, we assume neither of these scenarios obtain, and will express this assumption by simply saying that the causal structure of the given set of variables can be represented by a DAG. Let us reformulate the CMC in terms of the causal DAG.

[1] Much of the ensuing discussion is also included in Ramsey, Zhang and Spirtes (2006), which focuses on evaluating the empirical performance of the resulting algorithm. This collaboration reflects an interesting convergence of and interplay between theoretical work and empirical work. Peter Spirtes and I, while discussing issues related to my master thesis on uniform consistency (Zhang 2002) two years ago, realized that different components of the CFC serves very different purposes, and the full CFC need not be assumed if computational complexity is not a concern. But at that time we thought the algorithm theoretically correct under a weaker assumption would not be computationally feasible. From a totally different motivation, Joe Ramsey came up with virtually the same proposal of modifying PC due to his frustration with some undesirable features of the typical output of the PC algorithm. To our surprise and joy, Joe's simulation work clearly shows that the conservative version runs almost as fast as the standard version but improves accuracy significantly. Most interestingly, one possible explanation we can think of for why the conservative version has better performance at modest sample sizes seems to link well, in a way, to the issue of uniform consistency that initially concerned Peter and me.

> **Causal Markov Condition**: Given a set of variables whose structure can be represented by a DAG, every variable is probabilistically independent of its non-descendants in the DAG conditional on its parents in the DAG.

The CFC is accordingly formulated as such:

> **Causal Faithfulness Condition**: Given a set of variables whose causal structure can be represented by a DAG, no conditional independence holds among the variables unless entailed by the causal Markov condition.

Evidently the CFC describes a relationship between the causal DAG over a set of variables and the probability distribution of that set of variables. It dictates a list of probabilistic dependence relations that are required by a given causal DAG. Thus to test whether the CFC actually holds requires in general information about the true causal structure and the true probability distribution (see Spanos 2006 for an example of testing the CFC with assumptions about causal structure). But if the purpose of testing the CFC is ultimately for the sake of inferring causal structure, it is obviously not very useful to have a test that requires information about the causal structure in the first place. So can we detect failure of the CFC without knowing the causal structure? Not always, but there is a distinction to draw between violations of the CFC that are not detectable and violations of the CFC that are in principle detectable with access to the probabilistic information alone.

The idea is simple. Assuming the CMC holds, if a probability distribution does not satisfy the CFC with any possible causal structure it is Markov to, then no matter what the true causal structure is, the CFC is violated. By contrast, if the true probability distribution is faithful to some causal structure but not to the true one, it would be a violation of the CFC that cannot be recognized without information about the true causal structure.

To fully characterize detectable violations of the CFC in this sense is an ongoing project. In what follows we will only make a step towards that end. In particular, we explore a natural decomposition of the CFC suggested by familiar procedures of learning (equivalence classes of) causal DAGs. We show that assuming one component of the decomposition holds, any failure of the other component is in principle detectable. Moreover, the relevant test is readily incorporated into the familiar causal learning procedure, and clearly improves the empiricial performance.

2.2 Causal Inference with Causal Sufficiency

A quick review of some familiar methods for inferring causal DAGs is in order. For one thing, the decomposition to be introduced arises naturally from a standard routine for inferring causal DAGs (when the set of observed variables is known or assumed to be causally sufficient) from an oracle of conditional independence constraints. Moreover, a central topic in later chapters is inference of causal MAGs (when the set of observed variables is not assumed to be causally sufficient), which is in many ways parallel to, though significantly harder than, inference of causal DAGs. So an understanding of inference of causal DAGs will be very helpful for reading later chapters.

Suppose we have a set of observed variables $\mathbf{V}$, which is known or assumed to be causally sufficient. The (unknown) causal structure can be represented by a DAG, as it is assumed throughout this dissertation that causation is acyclic. Furthermore we assume the CMC holds, so all conditional independence constraints entailed by this causal DAG via its Markov property hold of the true joint probability distribution over $\mathbf{V}$. To better understand those entailed constraints, it is time to formally define the d-separation criterion.

Given a path u in a DAG, a non-endpoint vertex V on u is called a *collider* if the two edges incident to V on u are both into V ($\rightarrow V \leftarrow$), otherwise V is called a *non-collider*.

Definition 2.2.1 (d-separation). *In a DAG, a path u between vertices A and B is* **active** *(***d-connecting***) relative to a set of vertices $\mathbf{Z}$ ($A, B \notin \mathbf{Z}$) if*

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

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

A and B are said to be **d-separated** *by $\mathbf{Z}$ if there is no active path between A and B relative to $\mathbf{Z}$.*

The importance of the d-separation criterion lies in the fact that it captures exactly the conditional independence constraints entailed by the local Markov property of a DAG (Pearl 1988). The d-separation criterion is also referred to as the *global Markov property*, and the useful fact is that the local and the global Markov properties are equivalent for DAGs.

Given this fact, the CMC can be rephrased as saying that for any three disjoint subsets of variables $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$, if $\mathbf{A}$ and $\mathbf{B}$ are d-separated by $\mathbf{C}$ in the causal DAG, then $\mathbf{A}$ and $\mathbf{B}$ are independent conditional on $\mathbf{C}$.[2] The CFC hence says the following:

> **Causal Faithfulness Condition**: Given a set of variables whose causal structure can be represented by a DAG, for any three disjoint subsets of variables $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$, if $\mathbf{A}$ and $\mathbf{B}$ are *not* d-separated by $\mathbf{C}$ in the causal DAG, then $\mathbf{A}$ and $\mathbf{B}$ are *not* independent conditional on $\mathbf{C}$.

[2] A note for interested readers: there are good reasons to assume this formulation of the CMC in linear feedback systems represented by directed cyclic graphs (Spirtes 1995, Koster 1996) as well as non-recursive structural equations among discrete variables with equilibrium solutions (Pearl and Dechter 1996), though the earlier formulation fails.

The CMC and CFC together set up a perfect correspondence between conditional independence constraints and d-separation features of the causal DAG. In other words, an oracle of conditional independence constraints over the given set of observed variables translates into an oracle of d-separation features in the causal DAG. One prominent approach to causal discovery, the independence-constraint-based approach, attempts precisely to recover the causal structure from the d-separation features.

An important limitation is that d-separation features usually do not uniquely pick out a DAG. Multiple DAGs over the same set of variables can have the exact same d-separation features, and hence entail the exact same conditional independence relations among the variables. These DAGs are referred to as *Markov equivalent* and are indistinguishable by an oracle of conditional independence facts. Hence the output of independence-constraint-based procedures is a Markov equivalence class of DAGs, represented by a graphical object, known as a *Pattern* or *PDAG* or *essential graph*, that displays common features shared by all DAGs in the class.

Two simple facts about d-separation are particularly relevant to constraint-based causal discovery procedures (see e.g. Neapolitan 2004, pp. 89 for proofs):

Lemma 2.2.1. *Two variables are adjacent in a DAG if and only if they are not d-separated by any subset of other variables in the DAG.*

Call a triple of variables $\langle X, Y, Z \rangle$ in a DAG an *unshielded triple* if X and Z are both adjacent to Y but are not adjacent to each other. It is an *unshielded collider* if $X \rightarrow Y \leftarrow Z$, i.e., the edge between X and Y and the one between Z and Y are both into Y; otherwise it is an *unshielded non-collider*.

Lemma 2.2.2. *In a DAG, any unshielded triple $\langle X, Y, Z \rangle$ is a collider if and only if all sets that d-separate X from Z do not contain Y; it is a non-collider if and only if all sets that d-separate X from Z contain Y.*

Lemma 2.2.1 implies that all Markov equivalent DAGs have the same adjacencies. Lemma 2.2.2 implies that all Markov equivalent DAGs have the same unshielded colliders (and unshielded non-colliders). In fact, the converse is also true, i.e., two DAGs with the same adjacencies and the same unshielded colliders are Markov equivalent (Verma and Pearl 1991).

Constraint-based causal discovery proceeds in two stages. The first stage infers the common adjacencies shared by all DAGs Markov equivalent to the true causal graph, and the second stage infers some (preferably all) arrow orientations shared by all DAGs Markov equivalent to the true causal graph. A well-known representative of constraint-based algorithms is the PC algorithm (Spirtes et al. 2000). The algorithm is reproduced below, in which $ADJ(G, X)$ denotes the set of nodes adjacent to X in a graph G:

PC Algorithm

S1 Form the complete undirected graph U on the set of variables $\mathbf{V}$;

S2 $n = 0$

repeat

For each pair of variables X and Y that are adjacent in (the current) U such that $ADJ(U, X)\backslash\{Y\}$ or $ADJ(U, Y)\backslash\{X\}$ has at least n elements, check through the subsets of $ADJ(U, X)\backslash\{Y\}$ and the subsets of $ADJ(U, Y)\backslash\{X\}$ that have exactly n variables. If a subset S is found conditional on which X and Y are independent, remove the edge between X and Y in U, and record S as $Sepset(X, Y)$;

$n = n + 1$;

until for each ordered pair of adjacent variables X and Y, $ADJ(U,X)\backslash\{Y\}$ has less than n elements.

S3 Let P be the graph resulting from step S2. For each unshielded triple $\langle A,B,C\rangle$ in P, orient it as $A \rightarrow B \leftarrow C$ iff. B is not in $Sepset(A,C)$.

S4 Execute the following orientation rules until none of them applies:

a If $A \rightarrow B$ — C, A and C are not adjacent, orient as $B \rightarrow C$.

b If $A \rightarrow B \rightarrow C$ and A — C, orient as $A \rightarrow C$.

c If $A \rightarrow B \leftarrow C$, A — D — C, D — B, and A and C are not adjacent, orient D — B as $D \rightarrow B$.

In the PC algorithm, S2 constitutes the adjacency stage; S3 and S4 constitute the orientation stage. In S2, the PC algorithm essentially searches for a conditioning set for each pair of variables that renders them independent, which we henceforth call a *screen-off* conditioning set. Why it does so is obvious given Lemma 2.2.1. What distinguishes the PC algorithm from other constraint-based algorithms is the way it performs search. Two tricks are employed: (1) it starts with the conditioning set of size 0 (i.e., the empty set) and gradually increases the size of the conditioning set; and (2) it confines the search of a screen-off conditioning set for two variables within the potential parents – i.e., the currently adjacent nodes – of the two variables, and thus systematically narrows down the space of possible screen-off sets as the search goes on. These two tricks increase both computational and statistical efficiency in most real cases.

In S3, the PC algorithm uses a very simple criterion to identify unshielded colliders or non-colliders. For any unshielded triple $\langle X,Y,Z\rangle$, it simply checks whether or not Y is contained in the screen-off set for X and Z found in the adjacency stage. The connection between this rule and Lemma 2.2.2 should be obvious. S4 consists of orientation propagation rules based on information about non-colliders obtained in S3 and the assumption of acyclicity. These rules are shown to be both sound and complete in Meek (1995). In plain terms, they pick out *all* remaining orientations that are shared by all DAGs Markov equivalent to the true causal DAG. The major task in Chapters 3 and 4 is to prove the analogous completeness theorem for an algorithm that infers causal MAGs.

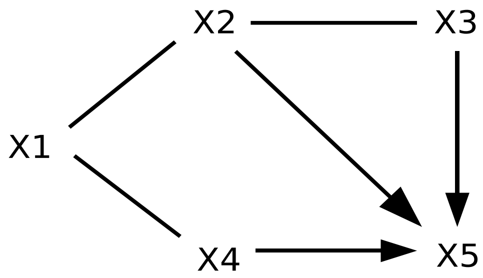

Figure 2.1: A sample output from the PC algorithm.

A typical output of the PC algorithm is shown in Figure 2.1. It is a graphical object containing both directed edges and undirected edges. Although the true causal graph is not fully known, this output reveals quite some causal information, for example, that $X2, X3, X4$ are direct causes of $X5$. For people who have resources to perform controlled experiments, the output suggests what experiments are needed in order to fully discover the true causal graph (Murphy 2001, Frederick et al. 2005).

2.3 The CFC Decomposed

The two stages of constraint-based causal discovery algorithms suggest a natural decomposition of the CFC into two parts, which will be referred to as *Adjacency-Faithfulness* and *Orientation-Faithfulness*, respectively.

> **Adjacency-Faithfulness**: Given a set of variables $\mathbf{V}$ whose causal structure can be represented by a DAG G, if two variables X, Y are adjacent in G, then they are dependent conditional on any subset of $\mathbf{V} \backslash \{X, Y\}$.

That this condition follows from the CFC should be clear in light of Lemma 2.2.1. The Adjacency-Faithfulness condition is the part of the CFC that is used to justify the stage of recovering adjacencies in the causal graph in constraint-based algorithms. Again, as we saw in the PC algorithm, this stage proceeds by searching for screen-off sets for pairs of variables, and by the causal Markov and Adjacency-Faithfulness conditions, two variables are not adjacent if and only if a screen-off set for them is found.

> **Orientation-Faithfulness**: Given a set of variables $\mathbf{V}$ whose causal structure can be represented by a DAG G, let $\langle X, Y, Z \rangle$ be any unshielded triple in G.
>
> (O1) if $X \rightarrow Y \leftarrow Z$, then X and Z are dependent given any subset of $\mathbf{V} \backslash \{X, Z\}$ that contains Y;
>
> (O2) otherwise, X and Z are dependent conditional on any subset of $\mathbf{V} \backslash \{X, Z\}$ that does not contain Y.

Orientation-Faithfulness is entailed by the CFC in light of Lemma 2.2.2. It is called Orientation-Faithfulness for the obvious reason that it serves to justify the step of identifying *unshielded colliders* (and *unshielded non-colliders*) in constraint-based algorithms. In particular, for any unshielded triple $\langle X, Z, Y \rangle$ resulting from the adjacency stage, a screen-off set for X and Y must have been found. The Orientation-Faithfulness condition then implies that the triple is an unshielded collider if and only if the screen-off set does not contain Z, which is exactly what the PC algorithm checks.

We should note that the Adjacency-Faithfulness and the Orientation-Faithfulness do not constitute an exhaustive decomposition of the CFC. Both of them are consequences of the CFC, but they together do not imply the CFC. Consider, for example, a causal graph consisting of a simple chain $X \rightarrow Y \rightarrow Z \rightarrow W$. We can easily cook up a case/parameterization for this causal structure where transitivity of causation fails (see McDermott's example below) and as a result X is independent of W, which violates the CFC. But it does not have to violate the Adjacency-Faithfulness or the Orientation-Faithfulness. For example, we can make a case where the only conditional independence relations that hold are $X \perp\!\!\!\perp W$, $X \perp\!\!\!\perp W | Y$, $X \perp\!\!\!\perp W | Z$ and $X \perp\!\!\!\perp W | Y, Z$.[3] It is easy to check that Adjacency-Faithfulness and Orientation-Faithfulness are both satisfied, whereas the CFC is violated due to the independence between X and W.

It is worth noting, however, that the correctness of the PC algorithm only depends on the truth of Adjacency-Faithfulness and Orientation-Faithfulness. As long as these two components of the CFC holds, the PC will not err given the right oracle of conditional independence. In our example, for instance, the PC algorithm outputs $X \text{—} Y \text{—} Z \text{—} W$, with $\langle X, Y, Z \rangle$ and $\langle Y, Z, W \rangle$ being unshielded non-colliders, which is obviously correct. So the parts of the CFC that get actually used in constraint-based causal discovery are the Adjacency-Faithfulness and the Orientation-Faithfulness.

[3] $\perp\!\!\!\perp$ is a symbol that denotes probabilistic independence introduced by Dawid (1979). The vertical bar | denotes conditioning.

The above example shows that in general there exist cases where Adjacency-Faithfulness and Orientation-Faithfulness are both satisfied but the full Faithfulness condition is violated. It is of course equally obvious that in general there exist cases where the Adjacency-Faithfulness condition holds but the Orientation-Faithfulness condition fails. Again, this can happen with a simple chain $A \rightarrow B \rightarrow C$ where causal transitivity fails along this chain. McDermott (1995) gave an interesting example of the sort. The story goes roughly as this: a right-handed terrorist is about to press a detonation button to explode a building when a dog bites his right hand, so he uses his left hand instead to press the button and triggers the explosion. Intuitively, the dog-bite causes the terrorist pressing the button with his left hand, which in turn causes the explosion, but the dog-bite does not cause the explosion. Transitivity of causation apparently fails. (Other putative cases of failure of causal transitivity can be found in e.g. Hitchcock 2001.)

This example is supposedly a case token-level event causation, but there is no difficulty turning it into a case of type-level variable causation. Let A be the variable that takes two values: 'yes' if dog bites, and 'no' otherwise; B be the variable that takes three values: 'right' if the terrorist presses the button with his right hand, 'left' if he does it with his left hand, and 'none' if he does not press the button at all; and C be the variable that takes two values: 'yes' if explosion occurs, and 'no' otherwise. If we assign probabilities in line with McDermott's intended story, we find that $A \perp\!\!\!\perp C$ and $A \perp\!\!\!\perp C|B$.

Note that for restricted class of causal structures and family of probability distributions, adjacency-faithfulness may imply orientation-faithfulness. In other words, there may not be any probability from the given family that is adjacent-faithful but not orientation-faithful to a causal structure in the given class. For example, in the above case of a simple chain $A \to B \to C$, if we restrict to binary variables or Gaussian variables that bear linear relationships, there do not exist distributions that are adjacency-faithful but not orientation-faithful to $A \to B \to C$. (The example given in the previous paragraph involves a variable with three categories.) More generally there are known results (e.g., Becker et al. 2000) that imply that in binary tree-like networks adjacency-faithfulness implies orientation-faithfulness. This result can be generalized to Gaussian tree-like networks as well. [4] However, if we do not restrict to tree-like causal structures and consider general DAGs, both binary and linear Gaussian networks admit failure of orientation-faithfulness but not adjacency-faithfulness. The simplest example is cancellation of two causal pathways $A \to B \to D$ and $A \to B \to D$. It follows that the Adjacency-Faithfulness condition is indeed weaker than the CFC.

Now the main point. If we assume the causal Markov and Adjacency-Faithfulness conditions are true, we can in principle test whether Orientation-Faithfulness fails of a particular unshielded triple. Suppose we have a perfect oracle of conditional independence relations, which is in principle available in the large sample limit. Since the CMC and the Adjacency-Faithfulness conditions are by assumption true, out of the oracle one can construct correct adjacencies and non-adjacencies, and thus correct unshielded triples in the causal graph. For such an unshielded triple, say, $\langle X, Y, Z \rangle$, if there is a subset of $\mathbf{V} \backslash \{X, Z\}$ containing Y that renders X and Z independent and a subset not containing Y that renders X and Z independent, then Orientation-Faithfulness definitely fails on this triple. This failing condition can of course be verified by the oracle.

Note that this simple test of Orientation-Faithfulness does not rely on knowing what the true causal graph is, even though Orientation-Faithfulness is a relation be-

[4] Thanks to Thomas Richardson for pointing this out.

tween a probability distribution and a graph. The reason why this test works is already hinted earlier, namely, that it actually checks whether the distribution (or the conditional independence oracle) is Orientation-Faithful to any DAG. A distribution that satisfies the causal Markov and Adjacency-Faithfulness conditions with the true DAG but fails the above test is not Orientation-Faithful to any DAG, and in particular, not Orientation-Faithful to the true DAG. This is why we do not need to know what the true graph is in order to detect a violation of Orientation-Faithfulness.

This suggests that theoretically we can relax the standard CFC and still have provably correct and informative causal discovery procedures. In fact, a main result to be established is that the PC algorithm, though incorrect under the weaker, Adjacency-Faithfulness condition, can be revised such that the modified version – that we call CPC (conservative PC) – is (1) correct given the Adjacency-Faithfulness condition; and (2) as informative as the standard PC algorithm if the CFC actually obtains.

2.4 The Conservative PC (CPC) Algorithm

Before we present a modified PC algorithm, it is helpful to explain how the PC algorithm can make mistakes under the causal Markov and Adjacency-Faithfulness conditions. Basically the causal Markov and Adjacency-Faithfulness conditions guarantee that adjacencies (and non-adjacencies) resulting from the adjacency stage of the PC algorithm are asymptotically correct. However, these two conditions do not imply the truth of Orientation-Faithfulness, and when the latter fails, the PC algorithm will err even in the large sample limit.

Consider again the adapted McDermott's case where $A \rightarrow B \rightarrow C$ is the true causal structure and the true probability distribution implies that $A \perp\!\!\!\perp C$ and $A \perp\!\!\!\perp C | B$.

The causal Markov and Adjacency-Faithfulness conditions are both satisfied, but Orientation-Faithfulness is not true of the triple $\langle A, B, C\rangle$. Now, given the correct conditional independence oracle, the PC algorithm would remove the edge between A and C in S2 because $A \perp\!\!\!\perp C$, and later in S3 orient the triple as $A \rightarrow B \leftarrow C$ because B is not in the screen-off set found in S2, i.e., the empty set. Simple as it is, the example suffices to establish that the PC algorithm is not asymptotically correct under the causal Markov and Adjacency-Faithfulness assumptions. We can of course construct any number of examples in which the PC algorithm makes any number of mistakes in the large sample limit.

It is not hard, however, to modify the PC algorithm to retain correctness under the weaker assumption. Indeed a predecessor of the PC algorithm, called the SGS algorithm (Spirtes et al. 1993/2000), is almost correct. The SGS algorithm decides whether an unshielded triple $\langle X, Y, Z\rangle$ is a collider or a non-collider by literally checking whether ($O1$) or ($O2$) in the statement of Orientation-Faithfulness is true. Theoretically all it lacks is a clause that acknowledges the failure of Orientation-Faithfulness when neither ($O1$) nor ($O2$) passes the check. Practically, however, the SGS algorithm is a terribly inefficient algorithm. In terms of computation, it is best case exponential because it has to check dependence between X and Z conditional on every subset of $\mathbf{V}\backslash\{X, Z\}$. Moreover, the oracle of conditional independence is in practice provided by statistical tests. When too many statistical tests of conditional independence have to be done, it is exceedingly likely that some of them will err, and we suspect that almost every unshielded triple will be marked as unfaithful if we run the SGS algorithms on more than a few variables.

Fortunately, the leading idea in the adjacency stage of the PC algorithm can be exploited here. In principle in order to test whether $\langle X, Y, Z\rangle$ is a collider, a non-collider, or an unfaithful triple, we only need to check subsets of the variables that are potential parents of X and Z. The fact that this is sufficient to make a procedure asymptotically correct under the CMC and Adjacency-Faithfulness condition is demonstrated in Theorem 2.4.1 below. This trick, as we shall see, makes the modified algorithm almost as fast as the PC algorithm in simulations.[5]

The modified PC algorithm, called CPC (Conservative PC), replaces S3 in PC with the following S3', and otherwise remains the same.

S3' Let P be the graph resulting from step 1. For each unshielded triple $\langle A, B, C\rangle$, check all subsets of A's potential parents and of C's potential parents:

(a) If B is NOT in any such set conditional on which A and C are independent, orient $A - B - C$ as $A \rightarrow B \leftarrow C$;

(b) if B is in all such sets conditional on which A and C are independent, leave $A - B - C$ as it is, i.e., a non-collider;

(c) otherwise, mark the triple as "unfaithful" or "ambiguous" by an underline.

(Note: of course a triple marked "unfaithful" does not count as a non-collider in S4(a) and S4(c).)

It should be clear why the modified PC algorithm is labelled "conservative": it is more cautious than the PC algorithm in drawing unambiguous conclusions about causal orientations. A typical output of the CPC algorithm is shown in Figure 2.2, where the underlinings (which are lines crossing pairs of edges in the figure) denote marks of "unfaithful".

The conservativeness is what is needed to make the algorithm correct under the causal Markov and Adjacency-Faithfulness assumptions.

[5]This idea was initially suggested by Joseph Ramsey, who tried to improve the practical performance of the PC algorithm by imposing extra checks on edge orientations.

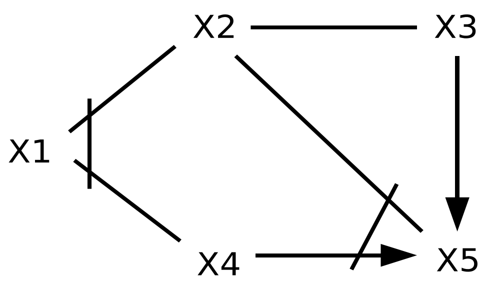

Figure 2.2: A sample output from the CPC algorithm.

Theorem 2.4.1 (Correctness of CPC). *Under the causal Markov and Adjacency-Faithfulness assumptions, the CPC algorithm is asymptotically correct in the sense that given a perfect conditional independence oracle, the algorithm returns a graphical object such that (1) it has the same adjacencies as the true causal graph does; and (2) all arrowheads and unshielded non-colliders in it are also in the true graph.*

Proof. Suppose the true causal graph is G, and all conditional independence judgments are correct. The Markov and Adjacency-Faithfulness assumptions imply that the undirected graph P resulting from step $S2$ has the same adjacencies as G does (Spirtes et al. 1993/2000). Now consider step $S3'$. If $S3'(a)$ obtains, then $A \rightarrow B \leftarrow C$ must be a subgraph of G, because otherwise by the Markov assumption, either A's parents or C's parents d-separate A and C, which means that there is a subset $\mathbf{S}$ of either A's potential parents or C's potential parents containing B such that $A \perp\!\!\!\perp C | \mathbf{S}$, contradicting the antecedent in $S3'(a)$. If $S3'(b)$ obtains, then $A \rightarrow B \leftarrow C$ cannot be a subgraph of G (and hence the triple must be an unshielded non-collider), because otherwise by the Markov assumption, there must be a subset $\mathbf{S}$ of either A's potential parents or C's potential parents not containing B such that $A \perp\!\!\!\perp C | \mathbf{S}$, contradicting the antecedent in $S3'(b)$. So neither $S3'(a)$ nor $S3'(b)$ will introduce an orientation error. Trivially $S3'(c)$ does not produce an orientation error, and it has been proven (in e.g., Meek 1995) that $S4$ will not produce any, which completes the proof. $\square$

Note that the two edges that figure in a triple marked "unfaithful" may still be oriented via other unshielded triples or by some propogation rules. For example, a triple $\langle A, B, C\rangle$ is marked as unfaithful, but there may be an unshielded collider $A \rightarrow B \leftarrow D$ and an unshielded collider $C \rightarrow B \leftarrow D$ in the output of the CPC, in which case we get $\underline{A \rightarrow B \leftarrow C}$. In such cases as $\underline{A \rightarrow B \leftarrow C}$, $\underline{A \rightarrow B \rightarrow C}$, $\underline{A \leftarrow B \rightarrow C}$ and $\underline{A - B \rightarrow C}$, the underlining serves no real purpose and can be removed. The remaining triples marked unfaithful by the CPC algorithm in the large sample limit are truly ambiguous in that either a collider or a non-collider is compatible with the conditional independence judgments. We conjecture but cannot yet prove that the CPC algorithm is complete in the sense that for every undirected edge in the output, there is a DAG that orients the edge in one way and a DAG that orients the edge in the other way such that both DAGs satisfy the causal Markov and Adjacency-Faithfulness assumptions with the given oracle of conditional independence.

The following theorem is obvious, of which we omit the proof.

Theorem 2.4.2. *Asymptotically the CPC algorithm and the PC algorithm produce the same output under the causal Markov and Faithfulness assumptions.*

Therefore, the CPC algorithm in principle does not sacrifice informativeness if the standard causal Faithfulness condition actually obtains. This theoretical result does not imply that the CPC algorithm will not be significantly less informative than the PC algorithm in realistic sample sizes. The latter needs to be studied empirically. In the next section I report some simulation work done by Joe Ramsey to show that it still pays to be conservative when the CFC holds. In particular, the CPC algorithm runs almost as fast as the PC algorithm, does not noticeably sacrifice information, but drastically decreases the amount of *misinformation* provided by PC with moderate sample sizes.[6]

2.5 Some Simulation Results

There are two immediate worries about the CPC algorithm. First, the extra check the CPC does might be a computational overhead that significantly slows down the procedure. Second, the theoretical superiority of the CPC algorithm over the PC algorithm may not necessarily cash out in practice if the situations where the Adjacency-Faithfulness but not the Orientation-Faithfulness holds do not arise often. (We will not try to make an argument to the contrary here, which does not mean that we endorse the claim). When the CFC actually holds, wouldn't the CPC algorithm be unnecessarily conservative? The following simulation results will address these worries by showing that the CPC algorithm in practice performs better than the PC algorithm, regardless of whether Orientation-Faithfulness holds or not. Even when the data are generated from a distribution Markov and Faithful to the true causal graph, it pays to be conservative on realistic sample sizes. In such cases, of course triples marked by the CPC algorithm should not be interpreted as "unfaithful". They are triples deemed ambiguous in the sense that the sample does not provide strong evidence to favor the collider structure over the non-collider structure or vice versa. In practice this is the interpretation we should give to the marked triples.

The simulations illustrate that the extra independence checks invoked in the CPC algorithms do not render CPC significantly slower than PC and that CPC is significantly more accurate than PC in terms of arrow orientations. The simulations were

[6] I thank Joe for generously permitting me to report the results here, and thank Clark Glymour for suggesting to include the simulation study for the readers' sake. The following section is a slightly extended version of what is reported in Ramsey et al. (2006).

performed on linear Gaussian models, with variations for sparser and denser graphs. The number of variables in the generated graphs ranges from 5 to 100 variables. For the sparser case, for each d from 5 to 100 in increments of 5, five random graphs with d random variables were selected uniformly from the space of DAGs with at most d edges and with a maximum degree of 10. For each such graph, a random linear structural equation model with independent Gaussian errors was constructed. For each such model, a random data set of 1000 cases was generated, to which PC and CPC were applied with significance level $\alpha = 0.05$ for each hypothesis test of conditional independence.[7] For denser models, the only difference is that the generated DAGs can have up to $2 \times d$ edges instead of just d edges.

The output in each instance was compared to the Pattern for the true DAG in that instance, the true Pattern. Performance statistics were recorded, including elapsed time and false positive and negative counts for arrowheads, unshielded non-colliders, and adjacencies, etc. For each number of variables, each performance statistic was averaged over the five random models constructed with that many variables, for PC and for CPC, respectively.

Figure 2.3 shows that for both sparser and denser models, the extra times CPC spend (recorded in seconds) are negligible, even in denser models with many variables. In all figures in this section, the performance statistics for the PC algorithm are represented by triangles and those for the CPC algorithm are represented by circles; sparser models use filled symbols, and denser models used unfilled symbols. The horizontal axis is the number of variables in the true DAG.

Since the CPC algorithm uses the exact same adjacency search routine as the PC algorithm, the number of adjacency errors do not differ. So our focus was on orientation errors. However, in our counting orientation errors, we decided to be

[7] The tests based on Fisher's Z transformation of correlation and partial correlation were used.

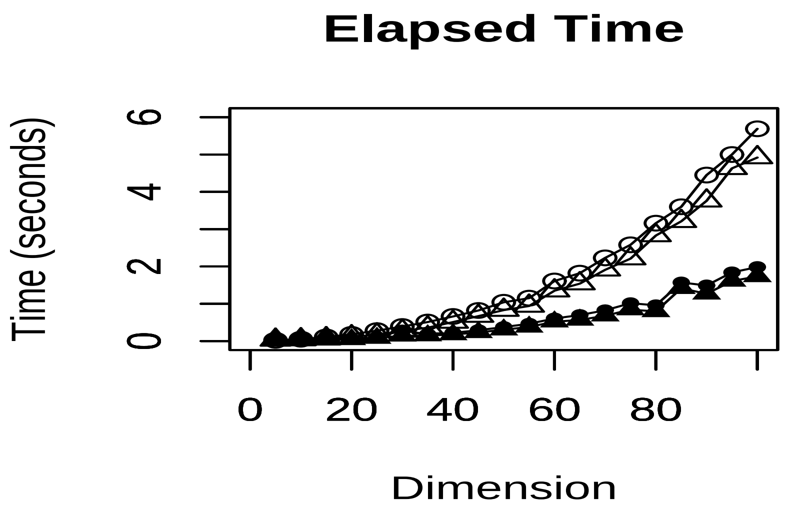

Figure 2.3: Average Elapsed Time

strict and incorporate adjacency errors in a way. Specifically, in our scoring scheme, an arrowhead removal error (false negative) occurs when the true pattern P_1 contains $A \rightarrow B$, but the output P_2 either does not contain an edge between A and B or does contain an edge between A and B but there is no arrowhead on this edge at B. Analogously, an arrowhead addition error (false positive) occurs when the output P_2 contains $A \rightarrow B$, but the true Pattern P_1 either does not contain the edge at all or has an edge with no arrowhead at B.

Obviously this scoring renders some but not all of the adjacency errors matter for evaluating orientation accuracy. For example, if A and B are not adjacent in P_1, but $A \text{---} B$ is in P_2, this is counted as an adjacency addition error, but not an arrowhead addition or removal error. In contrast, if $A \rightarrow B$ is in P_2, this is counted as an adjacency addition error and an arrowhead addition error, because of the arrowhead at B. We chose this scoring rule to differentiate between adjacency errors that lead to further orientations and adjacency errors that stand idle.

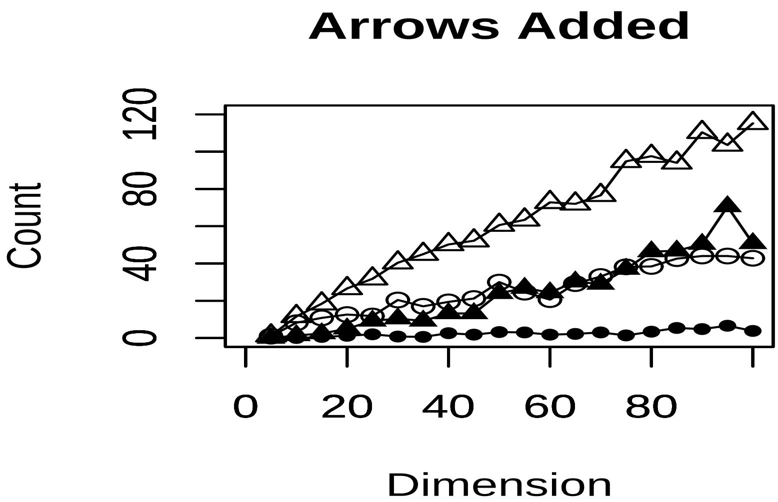

Figure 2.4: Average Counts of Arrow False Positives

Figure 2.4 shows that for both sparser and denser models, the number of extra arrowheads introduced is far better controlled by CPC than by PC. In other words, the conservativeness of the CPC algorithm avoids many false arrowheads output by the PC algorithm. For sparser models, the error is particularly well-controlled by the CPC algorithm, almost perfectly.

This would not be impressive if the CPC algorithm simply avoided mistakes at the cost of information. One may easily avoid mistakes by refusing to make judgments all the time. Our simulations suggest, however, that this is not the case with CPC. Figure 2.5 shows that for both sparser and denser models, the number of arrowhead removal errors committed by CPC is almost indistinguishable from the number of arrowhead removal errors committed by PC. So CPC does not sacrifice the true arrowheads inferred by PC.

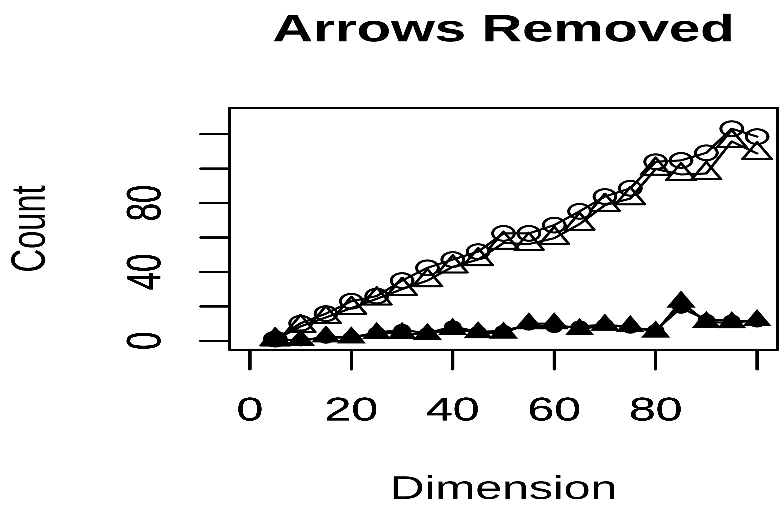

Figure 2.5: Average Counts of Arrow False Negatives

Besides false positive and false negative of arrowheads, we also recorded false positive and false negative for unshielded non-colliders, because unshielded non-colliders also contain important causal information, which, though disjunctive in nature, can easily lead to arrowhead orientations given some background knowledge. The counting method is similar. There is an unshielded non-collider addition error for the triple $\langle X, Y, Z\rangle$ if they form an unshielded non-collider in P_2, but in P_1 they either do not form an unshielded triple or form an unshielded collider. An ambiguous triple in G_2 does not count as an unshielded non-collider addition error, regardless of what is in G_1. Unshielded non-collider removal errors are calculated in an analogous fashion.

The performance of CPC and that of PC regarding false positive and false negative unshielded non-colliders are almost indistinguishable, as shown in Figures 2.6 and 2.7.

We also see that both PC and CPC do well regarding non-collider errors for sparser

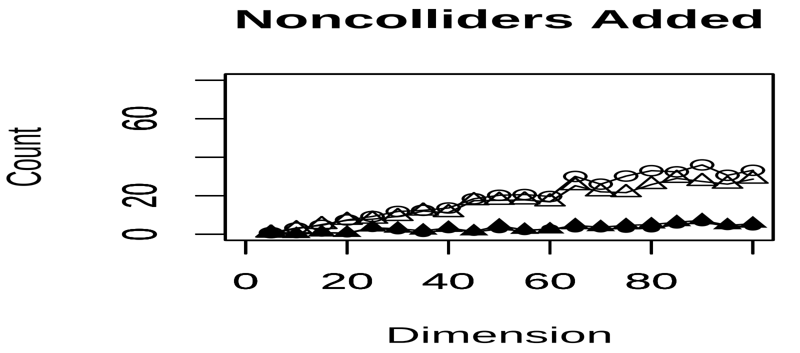

Figure 2.6: Average Counts of Non-Collider False Positive

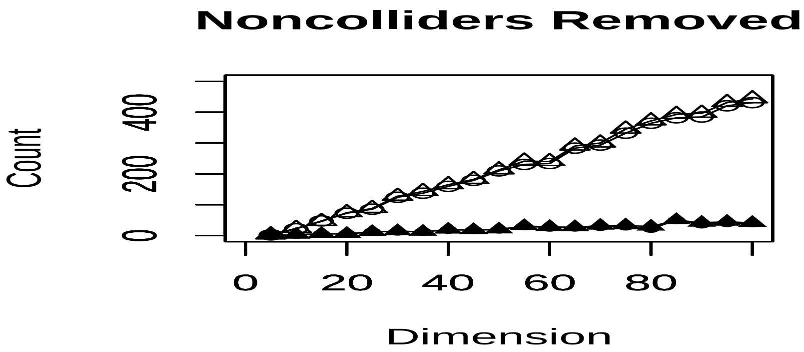

Figure 2.7: Average counts of Non-Collider False Negative

graphs, but are not satisfactory for denser graphs. This is a well-known limitation for constraint-based causal discovery algorithms, and is probably due to the considerable number of adjacency errors encountered in denser graphs.

In a word, CPC does no worse than PC in any aspect, but drastically decreases the number of false arrowheads inferred by the PC algorithm — in this sense a "Pareto improvement". The simulations suggest that the PC algorithm too often infers that unshielded triples are colliders, and the CPC algorithm provides the right antidote to this by means of the extra checks it performs. Similar simulations were carried out parameterizing random graphs using discrete variables with 2 to 4 categories, but otherwise with identical setup to the sparser continuous simulations above. Results are very similar.

A consequence of CPC's well control of false arrowheads is that it outputs way fewer "illegitimate" bi-directed edges. One complaint about constraint-based causal discovery procedures, and the PC algorithm in particular, is that at moderate sample sizes their outputs are usually outside the class of objects the algorithms aim to produce. Bi-directed edges should not occur in DAGs, and hence should not occur in Patterns that represent Markov equivalence classes of DAGs. From Figure 2.8 we can see that the number of bi-directed edges output by the CPC is almost zero in sparser graphs, and is in any case much smaller than the number of bi-directed edges resulting from the PC algorithm.

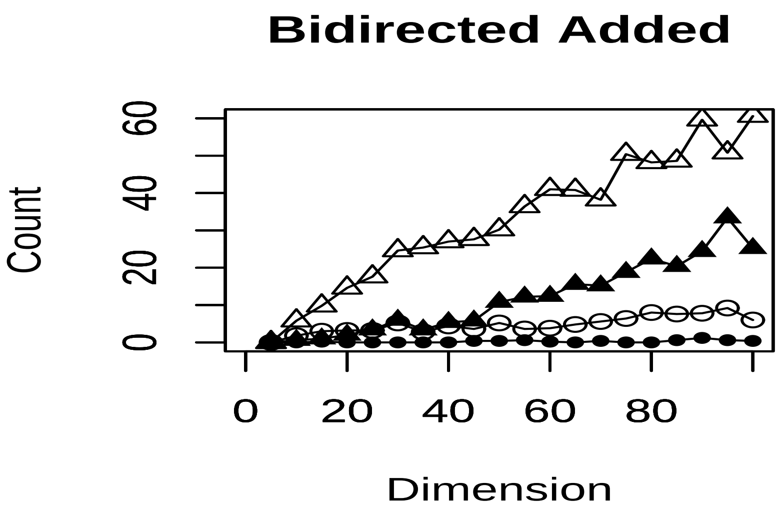

Figure 2.8: Average Counts of Bi-directed Edges

We have established in theory that the CPC inference procedure would be more accurate than the PC inference procedure when the Orientation-Faithfulness fails — indeed in the large sample limit the former is correct whereas the latter is often not. But why is CPC more accurate than PC at moderate sample sizes when Orientation-Faithfulness does hold? We think the reason is actually analogous. At

Although PC is correct in the large sample limit if Orientation-Faithfulness is not violated, it is very liable to error on realistic sample sizes if Orientation-Unfaithfulness is almost violated. By "almost violations" of Orientation-Faithfulness we mean the kind of situations where two variables, though entailed to be dependent conditional on some variables by the Orientation-Faithfulness condition, are close to be conditionally independent. How to quantify the "closeness" and just how close is close enough to cause trouble depend on distributional assumptions and sample sizes, and will not be formally pursued here. But intuitively if a distribution, though orientation-faithful to the true causal graph, is very similar to a distribution that is not orientation-faithful, then it cannot be well distinguished from the unfaithful distribution at a certain sample size. Eventually the faithfulness of the distribution will be revealed given more and more data, but at a certain sample size it may well be regarded as unfaithful, because it is close enough to being unfaithful relative to that sample size.

Almost-violations of Orientation-Faithfulness can arise in several ways – for example, when a triple chain is almost non-transitive, or more generally, when one of the edges in an unshielded triple is very weak — and are likely to arise especially relative to small sample sizes. When they happen, the CPC procedure tends to mark the relevant triples as ambiguous and avoids jumping into judgments of colliders as the PC procedure tends to do. That, we think, is a major reason for the improvement gained by CPC. In this regard, CPC seems to provide a partial solution to handling close-to-unfaithfulness, a situation pointed out by several authors as a major obstacle to reliable causal inference (Meek 1996, Robins et al. 2003, Zhang and Spirtes 2003).

2.6 Adjacency Error and the CPC algorithm

Earlier we drew a distinction between "detectable" and "undetectable" failure of the CFC. The testability of Orientation-Faithfulness given the CMC and the Adjacency-Faithfulness is certainly not the whole story. For example, assuming the CMC, some violations of the Adjacency-Faithfulness may be detectable. A simple instance of this is the following case: suppose there are two fair coins and a bell. The bell rings if and only if independent flips of the two coins respectively turn out the same (both heads or both tails). In this case, we have three binary variables and the true causal graph is a unshielded collider: $coin1 \rightarrow bell \leftarrow coin2$. The distribution violates the Adjacency-Faithfulness because, as can be easily calculated, $coin1 \perp\!\!\!\perp bell$ and $coin2 \perp\!\!\!\perp bell$. However, the violation is detectable because the distribution does not satisfy the CMC and the CFC with any causal DAG. So to fully characterize detectable violations of the CFC remains an open problem.

Still, the testability result about Orientation-Faithfulness is particularly nice in that the test is local and can be easily incorporated into constraint-based algorithms. Moreover, precisely due to the locality, some causal information can still be obtained even when violations of Orientation-Faithfulness are detected. Other detectable unfaithfulness may call for more global tests.

We should keep in mind that the testability result is based on the CMC and the Adjacency-Faithfulness condition, which ensure that the adjacencies can be correctly inferred from a perfect oracle of conditional independence. A simple Duhemian point is that a detected "unfaithful" triple indicates either an adjacency error or a real violation of Orientation-Faithfulness. Adjacency errors could be due to failure of the Adjacency-Faithfulness, or failure of the CMC, or in practical cases due to statistical errors. So a mark of "unfaithful" in the CPC algorithm suggests that extra checks on the adjacencies in the triple should be performed, if possible. It is important and complementary to the work presented in this chapter to explore ways that can increase the accuracy of the estimated adjacencies. Steck and Tresp (1996, also see Steck 2001), for example, have made interesting contributions related to this issue, which result in the NPC algorithm implemented in the Hugin package.

In this section we will present a very preliminary result intended to suggest that the CPC algorithm may be more robust than the PC algorithm against adjacency errors. The result is about whether adjacency errors will further lead to orientation errors. An adjacency error is usually not as consequential as an ensuing orientation error, as the former by itself does not imply an unambiguous causal claim, but an ensuing orientation error typically implies a substantial error in causal judgments. So it would be a virtue to have some measure against propagating adjacency errors into further orientation errors.

For the moment we only have a very simple and limited theorem on offer, but we expect to have a more interesting story to tell regarding this issue.

Theorem 2.6.1. *Suppose A — B is the only adjacency error made in the adjacency stage of the CPC algorithm, and the CMC holds of all other pairs of variables. Then given a perfect oracle of conditional independence, the CPC algorithm does not produce any orientation error except possibly an orientation of A — B. Moreover, if A is not an ancestor of B in the true causal DAG, the CPC algorithm does not orient A — B into $A \to B$.*

Proof. Consider S3' first. Since A — B is the only adjacency error, applications of

S3' can produce errors only on unshielded triples that include A — B. Consider any unshielded triple that includes A — B, say, without loss of generality, $\langle A, B, C\rangle$. We show that S3' will not mis-orient the edge between B and C, and will not orient A — B as $A \to B$ unless A is an ancestor of B in the true causal DAG.

Note that B and C are truly adjacent by assumption, so there are two cases to consider:

Case 1: $B \to C$ appears in the true causal DAG. We argue that there is a set of variables (adjacent to A or C) containing B that d-separates A and C. Note that A and C are truly non-adjacent. If A is not a descendant of C, then A is d-separated from C by the set of C's parents which contains B. If A is a descendant of C, we claim that A is d-separated from C by A's parents plus B, $\mathbf{Pa}(A) \cup \{B\}$. Suppose otherwise, that there is a d-connecting path u between A and C given $\mathbf{Pa}(A) \cup \{B\}$. Then the arrow on u incident to A must be out of A. It is then easy to derive that A is an ancestor of either C or a member of $\mathbf{Pa}(A) \cup \{B\}$, which contradicts acyclicity. So there exists a set containing B that d-separates A and C. Since the CMC holds (of the pair A and C), there is a screen-off set for A and C containing B. Then it is clear that S3' will not orient the triple $\langle A, B, C\rangle$ as a collider, and thus will not produce an orientation error.

Case 2: $B \leftarrow C$ appears in the true causal DAG. Since A and C are d-separated by either A's parents or C's parents, there is a screen-off set for A and C not containing B. So S3' will not judge the triple to be a non-collider. If it marks the triple as ambiguous, it does not commit an orientation error. If it orients the triple as $A \to B \leftarrow C$, it correctly orients $B \leftarrow C$, and we show that A is an ancestor of B in the true causal graph. Suppose otherwise, that A is not an ancestor of B. Then A and C are d-separated by $\mathbf{Pa}(A) \cup \{B\}$. Because any d-connecting between A and

C relative to $\mathbf{Pa}(A) \cup \{B\}$ must be out of A, which implies that A is an ancestor of either C or a member of $\mathbf{Pa}(A) \cup \{B\}$, which is impossible. It follows that S3' would not have oriented the triple as a collider.

Next consider S4. We argue by induction. Assume orientations so far satisfy the theorem. Consider S4(a) first. Again we only need to worry about S4(a) being applied to unshielded non-colliders that involve the edge between A and B. Note that the proof for S3' above already establishes that for any vertex C such that $\langle A, B, C \rangle$ forms an unshielded the triple, it will be judged to be a non-collider only if $B \rightarrow C$ appears in the true causal DAG (*Case 1* above). Moreover, by the inductive hypothesis, every orientation so far satisfies the theorem, so the only possible application of S4(a) on this triple is to orient the edge between B and C as $B \rightarrow C$, which is correct.

For S4(b), if S4(b) orients $A — B$ as $A \rightarrow B$, obviously A has to be an ancestor of B. And if S4(b) orients some other edge based on $A \rightarrow B$, obviously that orientation is correct given that A is an ancestor of B.

For S4(c), if S4(c) orients $A — B$ as $A \rightarrow B$, again it is easy to see that A has to be an ancestor of B in the true causal graph. And if S4(b) orients some other edge based on $A \rightarrow B$, that orientation will be correct given that A is an ancestor of B. □

So, given the conditions in Theorem 2.6.1, the only orientation error that could result is on the edge that should not have been there at all. In other words, the adjacency error does not lead to false orientations of other edges. The possible orientation error on the falsely added edge is also mitigated by the fact that it tracks the ancestral relationship in the true causal DAG. By contrast, the PC algorithm does not have this property. A single adjacency error can easily lead to orientation errors on other edges. As an simplest example, suppose the true causal DAG is $A \quad B \rightarrow C$, but there is an adjacency error that results in $A — B — C$ as the output of the adjacency stage of the PC (and the CPC) algorithm. If the conditions in Theorem 2.6.1 hold, then $A \perp\!\!\!\perp C$, and accordingly the PC algorithm will falsely judge the triple to be a collider $A \rightarrow B \leftarrow C$.

A false adjacency can arise due to failure of the CMC. Even though, for example, A and B are not adjacent in the true causal DAG, a typical failure of CMC on the pair A and B will render it the case that no set of other variables can screen off A from B. In that case, A and B will be falsely judged to be adjacent. Theorem 2.6.1, *prima facie*, suggests that the CPC algorithm is in a sense robust against a failure of the CMC on a single pair of variables. But it is not clear whether there are interesting cases where the CMC only fails on two variables against which the CPC but not the PC algorithm is robust.[8] In the example we gave in the end of the last paragraph, for instance, if the CMC fails of A and B, one would expect that it also fails of A and C so that A is not independent of C. We suspect that there are more interesting results along the line of Theorem 2.6.1 that await further research.

2.7 Related Issues

The CPC algorithm proposed in this chapter is provably correct under the causal Markov assumption plus a weaker-than-standard Faithfulness assumption, the Adjacency-Faithfulness assumption. It is a conservative generalization of the PC algorithm in that it theoretically gives the same answer as the PC does under the standard assumptions.

[8] Moreover, if the failure is a superficial one due to causal insufficiency, then the fact that the orientation rules used by the PC algorithm are also used by the FCI algorithm to be discussed in Chapter 3 implies that the PC algorithm is robust in the sense of Theorem 2.6.1. Thanks to Thomas Richardson for this point.

The other prominent approach to causal discovery and to learning graphical models in general is known as the score-based search. A natural question here is how a score-based algorithm would perform when the true distribution is only Adjacency-Faithful to the true causal graph. We will not undertake an extended discussion here, but it seems clear that such algorithms could err even in the large sample limit given typical scores such as BIC. For example, we constructed a case where two binary variables A, C and a quaternary variable form a causal graph $A \rightarrow B \leftarrow C$. We parameterized the graph producing a distribution such that $A \perp\!\!\!\perp C$ and $A \perp\!\!\!\perp C|B$, which violates the Orientation-Faithfulness.[9] The GES algorithm (Meek 1996, Chickering 2002), for example, when fed sufficient data from this distribution, outputs A—B—C, an unshielded non-collider. This is to be expected from any algorithm using a consistent score such as BIC or the BDe score, as a consistent score would (eventually) prefer the model with fewer parameters – in this case, the non-collider model – if both models contain the true distribution.

The foregoing analysis of decomposing and testing the CFC is confined to the context of inferring causal structure of causally sufficient systems. We see no principal reason that this confinement is necessary. Indeed some preliminary work is underway that intends to prove parallel results and make parallel improvements to causal inference algorithms without assuming that the set of observed variables is causally sufficient. We will report some basic ideas only in the final concluding chapter, as

[9] A causal story for how this could arise is this: imagine B is actually a compound variable $B = (B_1, B_2)$, where B_1 and B_2 are both binary. A and C are probabilistically independent. A is a cause of B_1 and C is a cause of B_2 such that the two causal mechanisms are autonomous with no interaction. So when we consider the three variables A, B and C, the causal DAG is indeed $A \rightarrow B \leftarrow C$, but it is not hard to see that $A \perp\!\!\!\perp C|B$. One may object that this example is based on an "improper" choice of a compound variable. But it is not clear what the criterion is for defining variables. At least this example does not look more artificial than the failure of causal transitivity along a simple chain $A \rightarrow B \rightarrow C$. The generalized dog-bit case, for example, is also based on a particular choice of variables.

those ideas are rooted in standard procedures of inferring causal structure of causally insufficient systems assuming the CMC and the CFC. To this central topic we now turn.

Chapter 3

Inference of 'Non-Cause' without Causal Sufficiency: Arrowhead Completeness

This chapter and the next will tackle an open problem that has been out there for over ten years (Spirtes et al. 1993/2000). As we briefly explained in Chapter 1, a major subtlety of causally insufficient systems is that DAGs over just the observed variables typically do not provide a proper representation: any such DAG either misrepresents the causal structure or misrepresents the probability distribution. If we do not want to explicitly introduce latent variables in the representation (which are indeed desirable to avoid if possible), we need to use graphs with a richer expressive machinery. One of such kinds is the class of *inducing path graphs* (Spirtes and Verma 1992, Spirtes et al. 1993/2000). Another kind is the class of MAGs roughly introduced in Chapter 1. A representative constraint-based causal discovery algorithm for causally insufficient systems is known as the FCI algorithm. The algorithm has at least two versions, one targeted at learning causal inducing path graphs (Spirtes et al. 1993/2000), and the other targeted at learning causal MAGs (Spirtes et al. 1999), but neither of them was proved to be complete in the sense that the output of the algorithm contains all valid features, features shared by all causal structures compatible with the given oracle of conditional independence assuming the CMC and CFC.

In fact, neither of them is complete. An open problem is thus to augment the FCI algorithm with additional inference rules and to prove its completeness, which is the one to be solved in the current and next chapters. Recent studies of MAGs (esp. Richardson and Spirtes 2002, 2003) show that MAGs are superior than inducing path graphs in several aspects — for example, that MAGs are much easier to parameterize. The Appendix also shows that syntactically MAGs form a subclass of inducing path graphs, which means that a Markov equivalence class of MAGs can contain more commonalities than a Markov equivalence class of inducing path graphs. So we will use MAGs in this dissertation. Since MAGs are not only a proper representation for causal inference, but also a useful tool in statistical modelling, the work presented here is hopefully also a contribution to the general statistical literature on graphical models.

We break the completeness result into two parts. In this chapter, we prove that the FCI algorithm is actually complete for inferring arrowheads that are common among all possible causal MAGs compatible with a given oracle of conditional independence. As shall be explained, arrowheads represent non-causes. They report information of the form: some variable is *not* a cause of another variable.[1] In the next chapter, we provide additional inference rules such that the resulting inference procedure is proved to be also sound and complete for inferring valid tails, which represent positive

[1]Another way of saying this may be more informative but less concise. An arrowhead at B on an edge between A and B means that either A is a cause of B or there is a latent common cause of A and B or both. Thanks to Clark Glymour for emphasizing this point.

statements of cause and effect.[2]

The class of ancestral graphical models as introduced by Richardson and Spirtes (2002) is more general than the one needed for representing causally insufficient systems as roughly introduced in Chapter 1. For full generality, we will consider inferring general ancestral graphs from conditional independence facts in these two chapters, and will highlight the special case of inferring causal structures for causally insufficient systems. The rest of the chapter is organized as follows. We introduce some details of ancestral graphs and their probabilistic and causal semantics in section 3.1. Then in section 3.2 we introduce the main components of the FCI algorithm as presented in Spirtes et al. (1999), and reproduce the proof of its soundness. Lastly, in section 3.3, the proof for the arrowhead completeness of the FCI algorithm is given.

3.1 Ancestral Graphs and Their Interpretations

The following example attributed to Chris Meek in Richardson (1998) illustrates nicely the major motivation of ancestral graphs:

> "The graph [Figure 3.1] represents a randomized trial of an ineffective drug with unpleasant side-effects. Patients are randomly assigned to the treatment or control group (A). Those in the treatment group suffer unpleasant side-effects, the severity of which is influenced by the patient's general level of health (H), with sicker patients suffering worse side-effects. Those patients who suffer sufficiently severe side-effects are likely to drop out of the study. The selection variable (Sel) records whether or not a patient remains in the study, thus for

[2]Dividing the completeness result into two chapters is largely a stylistic matter in view of the length of the proof, but also intends to highlight the modularity of the inference rules. The arrowhead completeness result in this chapter is joint work with Peter Spirtes, who, among other things, proves the key lemma, Lemma 3.3.1. The arrowhead completeness result is also obtained independently by Ayesha Ali and Thomas Richardson in a slightly different framework (Ali et al. 2005).

all those remaining in the study $Sel = StayIn$. Since unhealthy patients who are taking the drug are more likely to drop out, those patients in the treatment group who remain in the study tend to be healthier than those in the control group. Finally health status (H) influences how rapidly the patient covers (R)." (Richardson 1998, pp. 234)

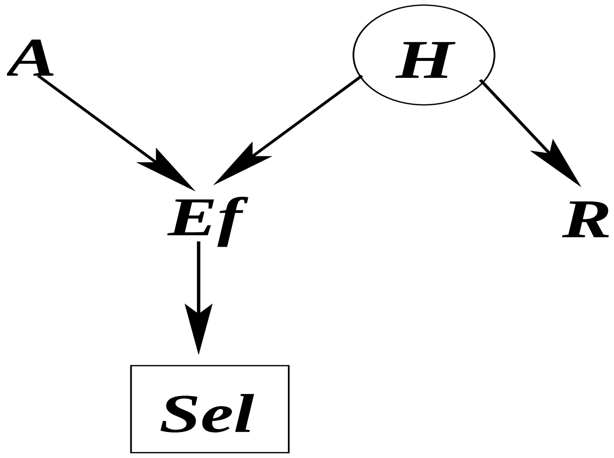

Figure 3.1: A Causal Mechanism with Latent and Selection Variables

Simple as this case is, it shows how the presence of latent confounders and/or selection variables matters. The two variables of primary interest, A and R, will be observed to be positively correlated, even though the supposed causal mechanism represented in Figure 3.1 seems to entail independence between them. Note that this correlation is not due to sample variation, but rather corresponds to genuine probabilistic association induced by design, the design that only the subjects that eventually stay in the study are considered. The observed correlation is in effect a correlation conditional on the variable Sel (taking the value $StayIn$), which should indeed be non-zero given the causal structure. This case thus provides a canonical example of what is called "selection effect" or "selection bias", in that a subject is sampled in virtue of the value of certain variable or variables – which are called *selection variables* – that are causally influenced by some other variables in the system. The point is that in those situations where selection variables matter, any probabilistic relationship inferrable from data is conditional upon (certain values of) the selection variables.

In Figure 3.1, H is supposed to be unobserved or latent, and hence the set of observed variables $\{A, Ef, R\}$ is causally insufficient. Note that if H were measured, we would in principle be able to find out that A and R are independent conditional on H (and, implicitly, on Sel), which would naturally suggest caution against drawing a causal conclusion about A and R. Without observing H, no screen-off set can be found between A and R, which makes a relatively strong appearance of causal relationship between A and R. As noted above, the current and the next chapters take into account both the possibility of causal insufficiency and the possibility of selection effect, but we will make comments now and then about situations where only causal insufficiency is possibly an issue.

As explained in Chapter 1, it is not possible to represent a causally insufficient system properly with a DAG over the observed variables, let alone a situation with both latent and selection variables. Oftentimes, *a fortiori*, there is no DAG over the observed variables that can even represent the (marginal) probability perfectly (i.e., that entails all and only those conditional independence relations implied by the marginal probability over the observed variables). The primary motivation of ancestral graphs is precisely the need to represent the presence of latent common causes and selection variables in the causal process that generates the data. Besides directed edges ($\rightarrow$), an ancestral graph can also contain bi-directed edges ($\leftrightarrow$, associated with the presence of latent common causes), and undirected edges (—, associated with the presence of selection variables).

3.1.1 Syntax of Ancestral Graphs

A **mixed graph** is a graph consisting of vertices and edges that may contain any of the three kinds of edges (directed, bi-directed and undirected) and at most one edge between any two vertices. The two ends of an edge we call **marks** or **orientations**. Obviously two kinds of marks can appear in a mixed graph: **arrowhead** ($>$) or **tail** ($-$). Specifically, the marks of an undirected edge are both tails; the marks of a bi-directed edge are both arrowheads; and a directed edge has one arrowhead and one tail. Sometimes 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 A, B therein, A is a **parent** of B and B is a **child** of A if $A \to B$ is in $\mathcal{G}$; A is called a **spouse** of B (and B a spouse of A) if $A \leftrightarrow B$ is in $\mathcal{G}$; A is called a **neighbor** of B (and B a neighbor of A) if $A — B$ is in $\mathcal{G}$. A **path** in $\mathcal{G}$ is a sequence of distinct vertices $\langle V_0, ..., V_n \rangle$ such that for $0 \leq i \leq n-1$, 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 $\langle V_0, ..., V_n \rangle$ such that for $0 \leq i \leq n-1$, V_i is a parent of V_{i+1} in $\mathcal{G}$. A is called an **ancestor** of B and B a **descendant** of A if $A = B$ or there is a directed path from A to B. We use $\mathbf{Pa}_{\mathcal{G}}, \mathbf{Ch}_{\mathcal{G}}, \mathbf{Sp}_{\mathcal{G}}, \mathbf{Ne}_{\mathcal{G}}, \mathbf{An}_{\mathcal{G}}, \mathbf{De}_{\mathcal{G}}$ to denote the set of parents, children, spouses, neighbors, ancestors, and descendants of a vertex in $\mathcal{G}$, respectively. A **directed cycle** occurs in $\mathcal{G}$ when $B \to A$ is in $\mathcal{G}$ and $A \in \mathbf{An}_{\mathcal{G}}(B)$. An **almost directed cycle** occurs when $B \leftrightarrow A$ is in $\mathcal{G}$ and $A \in \mathbf{An}_{\mathcal{G}}(B)$.

Definition 3.1.1. *A mixed graph is* **ancestral** *if the following three conditions hold:*

(a1) there is no directed cycle;

(a2) there is no almost directed cycle;

(a3) if there is an undirected edge between V_1 and V_2, i.e., $V_1 — V_2$, then V_1 and V_2 have no parents or spouses.

Obviously DAGs and undirected graphs (UGs) – graphs in which all edges are undirected – meet the definition, and hence are special cases of ancestral graphs. The first condition in Definition 3.1.1 is just the familiar one for DAGs. Together with the second condition, they define a nice connotation of arrowheads — that is, an arrowhead implies non-ancestorship, which induces a natural causal interpretation of ancestral graphs. The third condition requires that there is no edge into any vertex in the *undirected* component of an ancestral graph. This property simplifies parameterization and fitting of ancestral graphs (Richardson and Spirtes 2002, Drton and Richardson 2003), and it still allows selection effect to be properly represented.

3.1.2 Probabilistic Interpretation of Ancestral Graphs

Ancestral graphs are interpreted as encoding conditional independence relations by a graphical criterion that generalizes d-separation (Definition 2.2.1) for DAGs, called *m-separation*. The definition of d-separation essentially carries over to m-separation except that colliders and non-colliders admit more edge configurations in ancestral graphs than they do in DAGs. Given a path u in a graph, a non-endpoint vertex V on u is called a **collider** if the two edges incident to V on u are both into V, otherwise V is called a **non-collider**. The following definition is virtually the same as Definition 2.2.1.

Definition 3.1.2 (m-separation)**.** *In an ancestral graph, a path u between vertices A and B is* **active** *(***m-connecting***) relative to a set of vertices* $\mathbf{Z}$ *($A, B \notin \mathbf{Z}$) if*

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

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

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

Let $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ be three disjoint sets of vertices. $\mathbf{X}$ and $\mathbf{Y}$ are said to be m-separated by $\mathbf{Z}$ if $\mathbf{Z}$ m-separates every member of $\mathbf{X}$ from every member of $\mathbf{Y}$.

For DAGs, this probabilistic interpretation reduces to d-separation[3]. The following property is true of DAGs: if two vertices are not adjacent in a DAG, then there is a subset of other vertices that m-separates (d-separates) the two. This, however, is not always true of ancestral graphs. For example, the graph (a) in Figure 3.2 is an ancestral graph that fails this condition: C and D are not adjacent, but no subset of $\{A, B\}$ m-separates them. This motivates the following definition:

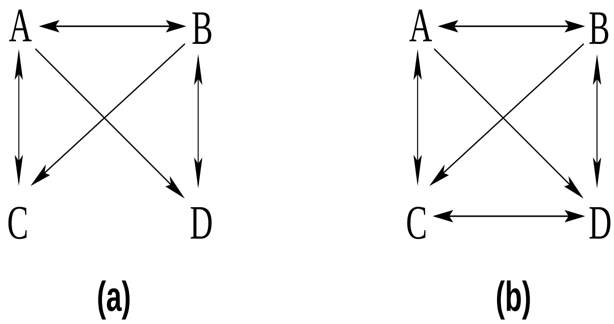

Figure 3.2: (a) an ancestral graph that is not maximal; (b) a maximal ancestral graph

[3]When there are only undirected edges in an ancestral graph, m-separation also reduces to the probabilistic interpretation of undirected graphs, the simple separation criterion.

Definition 3.1.3 (maximality). *An ancestral graph is said to be* **maximal** *if for any two non-adjacent vertices, there is a set of vertices that m-separates them.*

As already noted, DAGs are all maximal. In fact, maximality corresponds to the property known as the *pairwise Markov property*, i.e., every missing edge corresponds to a conditional independence relation, which, recall, is particularly relevant to inference of adjacencies in the PC algorithm. It is shown in Richardson and Spirtes (2002) that every non-maximal ancestral graph has a unique supergraph that is ancestral and maximal, and furthermore, every non-maximal ancestral graph can be transformed into the maximal supergraph by a series of additions of bi-directed edges. For example, in Figure 3.2, (b) is the unique maximal supergraph of (a), which has an extra bi-directed edge between C and D. From now on, we focus on maximal ancestral graphs (MAGs).

Maximality is closely related to the notion of inducing path, defined below:

Definition 3.1.4 (inducing path). *In an ancestral graph, let A, B be any two variables and $\mathbf{L}$, $\mathbf{S}$ be two disjoint sets of variables not containing A, B. A path u between A and B is called an* **inducing path relative to L and S** *if every non-endpoint vertex on u is either in $\mathbf{L}$ or a collider, and every collider on u is an ancestor of either A, B, or a member of $\mathbf{S}$.*

When $\mathbf{L}$ and $\mathbf{S}$ are both empty, the path u is simply called an **inducing path**[4] *between A and B.*

According to this definition, if A and B are adjacent, then the edge between them is trivially an inducing path. An important fact is that given an ancestral graph over $\mathbf{V}$, the presence of an inducing path relative to $\mathbf{L}$ and $\mathbf{S}$ is necessary and sufficient for two vertices, say A and B, not to be m-separated by any $\mathbf{C} \cup \mathbf{S}$,

[4]It is named *primitive* inducing path in Richardson and Spirtes (2002).

where $\mathbf{C} \subseteq \mathbf{V} \backslash (\mathbf{L} \cup \{A, B\})$. This property will play an important role in linking a causal DAG with latent variables to a causal MAG (see 3.1.3). In the special case of $\mathbf{L}$ and $\mathbf{S}$ being empty, the presence of an inducing path is necessary and sufficient for two vertices not to be m-separated by any set of other variables, which is obviously connected to maximality. (Recall that in Chapter 1 we actually give a definition of maximality in terms of inducing paths). We write the connection down as a proposition for later reference, a proof of which can be found in Richardson and Spirtes (2002).

Proposition 3.1.1. *An ancestral graph is maximal if and only if there is no inducing path between any two non-adjacent vertices in the graph.*

Probabilistically, a MAG represents the set of joint distributions over its vertices that satisfy its *global Markov property*, i.e., the set of distributions of which the conditional independence relations as implied by the m-separation relations in the MAG hold. Hence, if two MAGs share the same m-separation features, then they represent the same set of distributions. In this case, we call them **Markov equivalent**.

Definition 3.1.5 (Markov equivalence)**.** *Two MAGs $\mathcal{G}_1, \mathcal{G}_2$ (with the same set of vertices) are* **Markov equivalent** *if for any three disjoint sets of vertices $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$, $\mathbf{X}$ and $\mathbf{Y}$ are m-separated by $\mathbf{Z}$ in $\mathcal{G}_1$ if and only if $\mathbf{X}$ and $\mathbf{Y}$ are m-separated by $\mathbf{Z}$ in $\mathcal{G}_2$.*

There are sufficient and necessary conditions for Markov equivalence of MAGs that can be checked in polynomial time (Spirtes and Richardson 1996, Ali et al. 2004). To present a version of the conditions, the following notions are needed.

Definition 3.1.6 (unshielded collider)**.** *In a MAG, a triple of vertices $\langle A, B, C \rangle$ forms an* **unshielded collider** *if A and C are not adjacent, and there is an edge between A and B and one between B and C such that both edges are into B.*

It is well known that two DAGs are Markov equivalent if and only if they have the same adjacencies and the same unshielded colliders (Verma and Pearl 1990). These conditions are still necessary for Markov equivalence between MAGs, but are not sufficient. For two MAGs to be Markov equivalent, some shielded colliders have to be present in both or neither of the graphs. The next definition is related to this.

Definition 3.1.7 (discriminating path)**.** *In a MAG, a path between D and C, $u = \langle D, \cdots, A, B, C \rangle$, is a* **discriminating path** *for B if*

i. u includes at least three edges;

ii. B is a non-endpoint vertex on u, and is adjacent to C on u; and

iii. D is not adjacent to C, and every vertex between D and B is a collider on u and is a parent of C.

A canonical pictorial illustration of an discriminating path is included in Figure 3.3. Note that we write a discriminating path in such a form $u = \langle D, \cdots, A, B, C \rangle$, that is, we specify the endpoints and the vertices adjacent to B, the vertex being *discriminated*. The ellipsis therein designates any number (possibly zero) of other vertices. More generally, we adopt it as a convention for depicting a path: the vertices specified in the sequence are understood as distinct, and the ellipsis could be any number (possibly zero) of vertices.

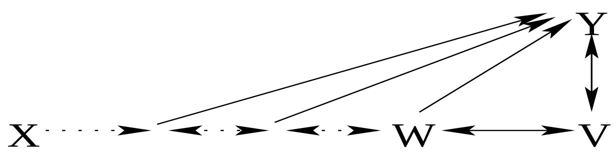

Figure 3.3: A discriminating path between X and Y for V

Discriminating paths behave similarly to unshielded triples in the following way (cf. Lemma 2.2.2): if a path between D and C is discriminating for B, then B is a collider on the path if and only if every set that m-separates D and C excludes B; and B is a non-collider on the path if and only if every set that m-separates D and C contains B. Thus we have the following proposition, proved in Spirtes and Richardson (1996):

Proposition 3.1.2. *Two MAGs over the same set of vertices are Markov equivalent if and only if*

(e1) They have the same adjacencies;

(e2) They have the same unshielded colliders;

(e3) If a path u is a discriminating path for a vertex B in both graphs, then B is a collider on the path in one graph if and only if it is a collider on the path in the other.

Given an arbitrary MAG $\mathcal{G}$, we denote its Markov equivalence class, the set of MAGs Markov equivalent to $\mathcal{G}$, by $[\mathcal{G}]$. A mark in $\mathcal{G}$ is said to be **invariant** if the mark is the same in all members of $[\mathcal{G}]$. According to Proposition 3.1.2, all members of $[\mathcal{G}]$ have the same adjacencies. But between two adjacent vertices, the edge, and hence one or both of the marks on the edge, may be variant across $[\mathcal{G}]$. An important task is then to fully characterize the invariant marks in $\mathcal{G}$, or in other words, the common marks shared by every member of $[\mathcal{G}]$. This will become crucial for the sake of causal inference, once we interpret MAGs causally.

3.1.3 Causal Interpretation of Maximal Ancestral Graphs

Consider again the simple motivating example in Figure 3.1. No conditional independence relation among A, Ef, R holds (conditionally on Sel implicitly). The non-vanishing correlations among them, however, hardly warrant any causal conclusion. In fact, according to the general definition to be presented shortly, the MAG over A, Ef, R that represents the situation is $R \leftarrow A — Ef \rightarrow R$. The causal interpretation of this MAG has to derive from the underlying causal DAG it represents.

More generally, the pattern of association and independence among a set of observed variables can be "misleading" about causal structure for at least two reasons. First, the set of observed variables may be causally insufficient (in the usual sense noted above), and some or all associations are due to unobserved common causes (or confounders as they are usually called). Second, the population that samples are drawn from may be just a subpopulation of the population of interest. The subpopulation, in particular, is characterized by a set of unobserved selections or conditioning variables such that units in the subpopulation share the values of the selection variables. If so, the pattern of association and independence among the observed variables is really a pattern conditional upon the selection variables.

Formally it is natural to represent such a situation by a causal DAG over the union of three disjoint sets of variables, $\mathbf{V} = \mathbf{O} \cup \mathbf{L} \cup \mathbf{S}$, where $\mathbf{O}$ denotes a set of observed variables, $\mathbf{L}$ denotes a set of latent or unobserved variables, and $\mathbf{S}$ denotes a set of unobserved selection variables to be conditioned upon. The DAG entails a set of conditional independence constraints among $\mathbf{V}$. Among these constraints, what are in principle testable are ones of the form $\mathbf{A} \perp\!\!\!\perp \mathbf{B} | \mathbf{C} \cup \mathbf{S}$[5], where $\mathbf{A}, \mathbf{B}, \mathbf{C} \subseteq \mathbf{O}$.

[5]Strictly speaking, we are conditioning on a specific value or vector of values of $\mathbf{S}$, so it is more accurate to write $\mathbf{A} \perp\!\!\!\perp \mathbf{B} | \mathbf{C} \cup \mathbf{S} = s$. This note applies to every occasion we write a conditional independence relation with selection variables in the conditioning set. Thanks to Thomas Richardson for emphasizing this.

A distinctive property of MAGs is that they can represent such in-principle-testable constraints without explicitly introducing $\mathbf{L}$ and $\mathbf{S}$. Given any DAG $\mathcal{G}$ over $\mathbf{V} = \mathbf{O} \cup \mathbf{L} \cup \mathbf{S}$, there exists a MAG over $\mathbf{O}$ alone such that for any three disjoint sets of variables $\mathbf{A}, \mathbf{B}, \mathbf{C} \subseteq \mathbf{O}$, if $\mathbf{A}$ and $\mathbf{B}$ are entailed to be independent conditional on $\mathbf{C} \cup \mathbf{S}$ by $\mathcal{G}$ if and only if $\mathbf{A}$ and $\mathbf{B}$ are entailed to be independent conditional on $\mathbf{C}$ by the MAG. (This is obviously a generalization of the fact introduced in section 1.3.) When this is the case, we say the MAG *probabilistically represents* the DAG. The following construction gives us such a MAG:

Input: a DAG $\mathcal{G}$ over $\langle \mathbf{O}, \mathbf{L}, \mathbf{S} \rangle$

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 relative to $\mathbf{L}$ and $\mathbf{S}$ between them in $\mathcal{G}$;

2. for each pair of adjacent vertices A, B in $\mathcal{M}_\mathcal{G}$, orient the edge between them as follows:

   (a) orient it as $A \rightarrow B$ in $\mathcal{M}_\mathcal{G}$ if $A \in \mathbf{An}_\mathcal{G}(B \cup \mathbf{S})$ and $B \notin \mathbf{An}_\mathcal{G}(A \cup \mathbf{S})$;

   (b) orient it as $A \leftarrow B$ in $\mathcal{M}_\mathcal{G}$ if $B \in \mathbf{An}_\mathcal{G}(A \cup \mathbf{S})$ and $A \notin \mathbf{An}_\mathcal{G}(B \cup \mathbf{S})$;

   (c) orient it as $A \leftrightarrow B$ in $\mathcal{M}_\mathcal{G}$ if $A \notin \mathbf{An}_\mathcal{G}(B \cup \mathbf{S})$ and $B \notin \mathbf{An}_\mathcal{G}(A \cup \mathbf{S})$;

   (d) orient it as $A — B$ in $\mathcal{M}_\mathcal{G}$ if $A \in \mathbf{An}_\mathcal{G}(B \cup \mathbf{S})$ and $B \in \mathbf{An}_\mathcal{G}(A \cup \mathbf{S})$.

It can be shown that $\mathcal{M}_\mathcal{G}$ is indeed a MAG and probabilistically represents $\mathcal{G}$ (Richardson and Spirtes 2002). Moreover, it is easy to see that $\mathcal{M}_\mathcal{G}$ also encodes ancestral relationships in $\mathcal{G}$. So, if $\mathcal{G}$ is the causal DAG for $\langle \mathbf{O}, \mathbf{L}, \mathbf{S} \rangle$, a causal reading of $\mathcal{M}_\mathcal{G}$ readily follows. Notice that $\mathcal{M}_\mathcal{G}$ is a unique outcome of the above construction, so it is fair to call it *the* causal MAG over $\mathbf{O}$. Let us put down the now obvious causal interpretation of MAGs:

Causal Interpretation of MAGs

Edges in a MAG $\mathcal{G}$ are to be interpreted as follows:

1. $A \rightarrow B$ means that A is a cause of B or some (unobserved) selection variable, but B is not a cause of A or any (unobserved) selection variable.[6][7]

2. $A \leftrightarrow B$ means that A is not a cause of B or any (unobserved) selection variable, and B is not a cause of A or any (unobserved) selection variable.[8]

3. $A \text{---} B$ means that A and B are both ancestors of some selection variable. (The relationship between A and B is not clear.)

In short, arrowheads in a MAG are interpreted as "non-cause" (hence the title of this chapter), and tails are interpreted as "cause" (of either an observed variable or a selection variable). Note that in a situation where no selection effect is present (i.e., $\mathbf{S} = \emptyset$), the causal MAG will not contain any undirected edges, as implied, for example, by the third clause above. Therefore, regarding our main subject, causal insufficiency, alone, it suffices to consider *directed MAGs* as briefly described in Chapter 1. Furthermore, if there is no selection effect, directed edges (or tails) in a MAG get a more elegant and informative interpretation: $A \rightarrow B$ means that A is a cause of B. The interpretation of a directed path (or more generally, a partial directed path –

[6]When we say A is a cause of B, it means that there is a causal pathway from A to B in the true causal structure (with possibly latent variables). The presence of a causal pathway may not be sufficient for attributing cause and effect – for one thing, whether causation is transitive is controversial – and it is probably more appropriate to use "candidate cause" or "*prima facie* cause" in this regard. However, since we are assuming the CFC, causal transitivity is also assumed.

[7]This disjunctive interpretation may sound hardly useful, but it is a reflection the fact that the presence of selection effects seriously limits the possibility of inferring useful causal information from observations. Moreover, this disjunctive information may be combined with other information to deduce more useful information. For example, if there is also an arrowhead at A, then it can be deduced that A is not a cause of any selection variable, but a cause of B.

[8]Since the fact that there is an edge between A and B at all implies that A and B are not probabilistically independent conditional on any subset of other observed variables, the above causal interpretation of $A \leftrightarrow B$ also implies that there is a latent common cause of A and B.

a path in which arrowheads, if any, point to the same direction) from A to B readily follows.

3.2 Inferring Causal Structure in the Presence of Latent Confounders and Selection Variables – the FCI Algorithm

Since a causal MAG involves observed variables only, it is hopeful to infer from data (features of) the causal structure via the MAG representation in the presence of latent common causes and/or selection selection variables, given suitable assumptions. Given an arbitrary set of observed variables $\mathbf{O}$, we again skip over statistical inference and assume a reliable oracle of conditional independence is available. Since there may be selection effects, a peculiarity of this oracle is that there exists a (possibly empty) set of variables $\mathbf{S}$ such that whatever conditional independence query, say $A \perp\!\!\!\perp B | \mathbf{C}$, passed to the oracle will be translated into a query with $\mathbf{S}$ being added to the conditioning set, $A \perp\!\!\!\perp B | \mathbf{C} \cup \mathbf{S}$ (for some particular value of $\mathbf{S}$).

The two main assumptions to be relied upon are the CMC and the CFC. In particular, although $\mathbf{O}$ may be causally insufficient, we assume that there always exists a (possibly empty) set of variables $\mathbf{L}$ such that the causal structure over $\mathbf{V} = \mathbf{O} \cup \mathbf{L} \cup \mathbf{S}$ is properly represented by a DAG, i.e., the CMC holds of the causal DAG and the true joint distribution over $\mathbf{V}$. Assume furthermore that the CFC also holds between them. These two assumptions imply that the given oracle satisfies the CMC and CFC with the true causal MAG.[9] The question is how much information about the MAG

[9] Of course the assumptions we will rely upon henceforth are simply that the (marginal) distribution of the observed variables $\mathbf{O}$ satisfy the CMC and CFC with the true causal MAG over $\mathbf{O}$, whether or not there is an underlying causal DAG over a superset $\mathbf{V}$ of $\mathbf{O}$ such that the joint

can be inferred from the oracle.

By Definition 3.1.5, if the oracle is Markov and Faithful to the true causal MAG $\mathcal{G}$, it is also Markov and Faithful to all (and only) causal MAGs Markov equivalent to $\mathcal{G}$. Hence the oracle, given the CMC and the CFC, only determines up to $[\mathcal{G}]$, the Markov equivalence class of the true causal MAG. The aim is thus to infer all invariant features of $\mathcal{G}$, features that are common across $[\mathcal{G}]$.

Richardson (1996) introduced a class of graphical objects called *partial ancestral graphs* (PAGs) to represent the output of his causal inference algorithm in linear feedback systems. It turns out PAGs can also be used to represent Markov equivalence classes of MAGs. Spirtes et al. (1999) used PAGs to represent the output of the FCI (Fast Causal Inference) algorithm, an algorithm for causal inference in the presence of latent and selection variables. The basic idea of the FCI algorithm is not unlike the PC algorithm described in the previous chapter. Specifically, the FCI algorithm also consists of two stages: the adjacency stage and the orientation stage. In the adjacency stage, for every pair of variables, the algorithm searches for a screen-off set and leaves an edge (whose orientation is yet undecided) between the two if and only if no screen-off set is found. Just as in the PC algorithm, the adjacency search in the FCI algorithm also exploits some sort of online pruning of the space of possible screen-off sets to improve computational (and statistical) efficiency. Part of the trick is indeed exactly the same as one exploited in the PC algorithm. However, the PC adjacency search alone is not enough for MAGs, and the FCI algorithm has to do extra work with a fairly complicated rationale. Since we are not interested in improving the adjacency search of the FCI algorithm, which is obviously correct (and complete[10]) given

distribution of **V** satisfies the CMC and CFC with the causal DAG. But there does not seem to be a good reason for assuming the former without the latter.

[10]The correctness is regarding both adjacencies and non-adjacencies, which automatically implies completeness.

a reliable oracle, we will ignore the details and take the correctness of the adjacency stage for granted. Below is (an equivalent version of) the FCI algorithm from Spirtes et al. (1999), with the details of the adjacency stage F2 omitted. (Three kinds of edge marks are used in the algorithm: arrowhead ($>$), tail ($-$) and circle ($\circ$). In the presentation below, a meta-symbol, star ($*$), is also used as a wildcard that denotes any of the three marks.[11] More specifically, if "$*$" appears in an antecedent of an orientation rule, that means it does not matter whether the mark at that place is an arrowhead, or a tail, or a circle. If "$*$" appears in the consequence of a rule, that means the mark at that place remains what it was before the firing of the rule. Also, in writing orientation rules, we use Greek letters to denote generic variables.)

FCI Algorithm

F1 Form the complete graph U on the set of variables where between every pair of variables there is an edge $\circ\!\!-\!\!\circ$;

F2 For every pair of variables A and B, search *in some clever way* for a screen-off set. If such as set S is found, remove the edge between A and B in U, and record S as $Sepset(A, B)$;

F3 Let P be the graph resulting from step F2. Execute the orientation rule:

  $\mathcal{R}0$ For each unshielded triple $\langle \alpha, \gamma, \beta \rangle$ in P, orient it as a collider $\alpha {*\!\!\rightarrow}\, \gamma \leftarrow\!\!* \beta$ iff. γ is not in $Sepset(\alpha, \beta)$.

F4 Execute the following orientation rules until none of them applies:

[11] By this we mean the rule in question applies no matter which of the three marks actually appears in the position of $*$. It does not mean that all three marks *can* appear in that position.

$\mathcal{R}1$ If $\alpha{*\!\!\to}\, \beta \circ\!\!-\!\!-\!\!* \gamma$, and α and γ are not adjacent, then orient the triple as $\alpha{*\!\!\to}\, \beta \to \gamma$.

$\mathcal{R}2$ If $\alpha \to \beta{*\!\!\to}\, \gamma$ or $\alpha{*\!\!\to}\, \beta \to \gamma$, and $\alpha *\!\!-\!\!\circ\, \gamma$, then orient $\alpha *\!\!-\!\!\circ\, \gamma$ as $\alpha{*\!\!\to}\, \gamma$.

$\mathcal{R}3$ If $\alpha{*\!\!\to}\, \beta \leftarrow\!\!* \gamma$, $\alpha *\!\!-\!\!\circ\, \theta \circ\!\!-\!\!* \gamma$, α and γ are not adjacent, and $\theta *\!\!-\!\!\circ\, \beta$, then orient $\theta *\!\!-\!\!\circ\, \beta$ as $\theta{*\!\!\to}\, \beta$.

$\mathcal{R}4$ If $u = \langle \theta, ..., \alpha, \beta, \gamma \rangle$ is a discriminating path between θ and γ for β, and $\beta \circ\!\!-\!\!* \gamma$; then if $\beta \in Sepset(\theta, \gamma)$, orient $\beta \circ\!\!-\!\!* \gamma$ as $\beta \to \gamma$; otherwise orient the triple $\langle \alpha, \beta, \gamma \rangle$ as $\alpha \leftrightarrow \beta \leftrightarrow \gamma$.[12]

The output of the FCI algorithm is what is called a PAG. It is clear from the algorithm that a PAG can contain three kinds of marks: tail ($-$), arrowhead ($>$) and circle ($\circ$). So, by simple combinatorics, there could at most be six types of edges: $-$, $\to$, $\leftrightarrow$, $\circ\!\!-$, $\circ\!\!-\!\!\circ$, $\circ\!\!\to$. The circle is obviously intended to be an uninformative or ambiguous mark, which indicates that the corresponding mark may be either an arrowhead or a tail in the true causal MAG.

The question is whether all circles in the FCI output should really be ambiguous in the sense that some MAG Markov equivalent to the true causal MAG has an arrowhead there and some other MAG Markov equivalent to the true causal MAG has a tail there. In other words, the question is whether the FCI output is a complete PAG for the Markov equivalence class of the true causal MAG, as defined below (recall that a mark in $\mathcal{G}$ is said to be **invariant** if the mark is the same in all MAGs Markov equivalent to $\mathcal{G}$):

Definition 3.2.1 (CPAG). *Let* $[\mathcal{G}]$ *be the Markov equivalence class of an arbitrary MAG* $\mathcal{G}$. *The* **complete (or maximally oriented) partial ancestral graph**

[12]See Ali et al. (2005) for an alternative and perhaps more efficient formulation of the rules that takes on a special kind of discriminating paths.

(CPAG) *for* $[\mathcal{G}]$[13], $\mathcal{P}_{\mathcal{G}}$, *is a graph with (possibly) three kinds of marks (and hence six kinds of edges:* —, →, ↔, ∘—, ∘–∘, ∘→*), such that*

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

ii. A mark of arrowhead is in $\mathcal{P}_{\mathcal{G}}$ *if and only if it is invariant in* $[\mathcal{G}]$*; and*

iii. A mark of tail is in $\mathcal{P}_{\mathcal{G}}$ *if and only if it is invariant in* $[\mathcal{G}]$.

The difference between a CPAG and a PAG as previously employed in the literature[14] is of course that the latter is not alleged to contain all invariant arrowheads or tails. Another representation of Markov equivalence classes of MAGs is introduced by Ali (2002), called *joined graphs*, which aims only to represent all invariant arrowheads and hence do not distinguish invariant tails from variant marks. Clearly the most complete representation of a Markov equivalence class of MAGs regarding the common arrowheads and tails is the CPAG.

It will become clear that the FCI algorithm does not output the CPAG for the true causal MAG, even given a reliable oracle of conditional independence. However, with a reliable oracle, the FCI output does satisfy (i) and (ii) in Definition 3.2.1, as well as half of (iii). The fact that the output satisfies (i) is fairly obvious, and we refer readers to Spirtes et al. (1999) for a rigorous demonstration.

That it also satisfies (ii), however, amounts to saying that all arrowheads contained in the FCI output are valid in that they appear in every MAG that satisfies the CMC and CFC with the given oracle, and also that none of the other marks in the output could be a valid arrowhead. By 'half of (iii)' we mean that all tails in the FCI output are valid, but they do not necessarily exhaust invariant tails.

[13]Sometimes we will also refer to it as "the CPAG for $\mathcal{G}$".

[14]It is worth noting that a graphical object named *partially oriented inducing path graph (POIPG)* is studied in Spirtes et al. (1993/2000), which, however, can be shown to be just a PAG that is not complete in the sense of Definition 3.2.1. See Appendix.

All this is equivalent to saying that the orientation inference rules $\mathcal{R}0$ and $\mathcal{R}1$–$\mathcal{R}4$, the latter of which are illustrated in Figure 3.4, are sound and together complete with respect to invariant arrowheads.[15] Soundness is not that hard to prove, which was already shown in Spirtes et al. (1999). But the proof of completeness, as is usually the case, is fairly complicated.

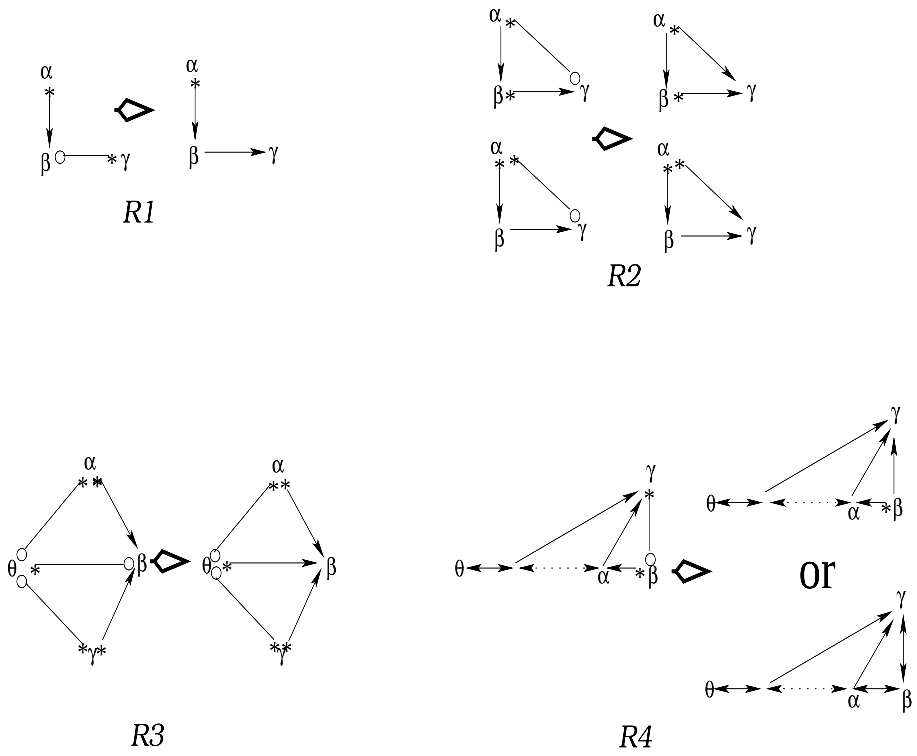

Figure 3.4: Graphical illustrations of $\mathcal{R}1 - \mathcal{R}4$

In order to prove soundness, we need a couple of simple facts (see, e.g., Spirtes and Richardson (1996) for proofs):

Lemma 3.2.1. *In a MAG, any unshielded triple $\langle X, Y, Z\rangle$ is a collider if and only if all sets that m-separate X from Z do not contain Y; it is a non-collider if and only if all sets that m-separate X from Z contain Y.*

[15] They are also independent in the sense that none of them can be derived from other rules.

Lemma 3.2.2. *In a MAG, if $u = \langle W, \cdots, X, Y, Z\rangle$ is a path between W and Z discriminating for Y, then the triple $\langle X, Y, Z\rangle$ is a collider if and only if all sets that m-separate W from Z do not contain Y; and it is a non-collider if and only if all sets that m-separate W from Z contain Y.*

Lemma 3.2.1 is obviously a generalization of Lemma 2.2.2. Lemma 3.2.2 reveals that discriminating paths bear a similar property as unshielded triples. We now prove the soundness of $\mathcal{R}0 - \mathcal{R}4$.

Theorem 3.2.1. *$\mathcal{R}0 - \mathcal{R}4$ are sound, and hence every non-circle mark in the FCI output is an invariant mark in the true causal MAG.*

Proof. Denote the (unknown) true causal MAG as $\mathcal{G}_T$. The soundness of $\mathcal{R}0$ readily follows from Lemma 3.2.1. For the other rules, it suffices to show that for each rule, if a mixed graph satisfies the antecedent of the rule but contains a mark different than what the rule requires, the graph is either not a MAG or not Markov equivalent to $\mathcal{G}_T$, and hence not a member of $[\mathcal{G}_T]$. (Then an inductive argument goes through.) In every case that follows, we assume the antecedent of the rule holds in the graph under consideration.

$\mathcal{R}1$: Suppose a mixed graph, contrary to what the rule requires, has an arrowhead at β. Then it contains an unshielded collider $\langle \alpha, \beta, \gamma\rangle$ which is not in $\mathcal{G}_T$, for otherwise it would have been picked up by $\mathcal{R}0$. Hence it is not Markov equivalent to $\mathcal{G}_T$ by Proposition 3.1.2. Furthermore, if the mark at γ is a tail, then $\alpha {*\!\!\rightarrow} \beta -\!\!\!-\gamma$ appears, which means the graph is not ancestral by Definition 3.1.1.

$\mathcal{R}2$: Suppose a mixed graph, contrary to what the rule requires, has a tail at γ. If it is $\alpha -\!\!\!- \gamma$, the graph is not ancestral because the edge between β and γ is into γ. If it is $\alpha \leftarrow \gamma$, then either γ is an ancestor of β and $\beta {*\!\!\rightarrow} \gamma$, or β is an ancestor of α and $\alpha {*\!\!\rightarrow} \beta$. In either case the graph is not ancestral.

$\mathcal{R}3$: Suppose a mixed graph, contrary to what the rule requires, has a tail at β. If it is $\theta —\beta$, the graph is not ancestral because the edge between α and β is into β. Suppose it is $\theta \leftarrow \beta$ that appears in the graph. Notice that if the triple $\langle \alpha, \theta, \gamma \rangle$ is an unshielded collider in the graph, then the graph is not Markov equivalent to $\mathcal{G}$. On the other hand, if it is not a collider, then at least one of the two edges is out of θ. Note that neither the edge between α and θ, nor the edge between γ and θ can be undirected, for otherwise the graph is not ancestral due to the presence of $\theta \leftarrow \beta$. So either $\theta \rightarrow \alpha$, or $\theta \rightarrow \gamma$; that is, β is either an ancestor of α or an ancestor of γ. In either case, the graph is not ancestral because $\alpha *\!\!\rightarrow \beta \leftarrow\!\!* \gamma$ is present.

$\mathcal{R}4$: There are two cases to consider.

Case 1: $\beta \in \mathit{Sepset}(\theta, \gamma)$. By the CFC and Lemma 3.2.2, the triple $\langle \alpha, \beta, \gamma \rangle$ has to be a non-collider in $\mathcal{G}_T$, and also, by Proposition 3.1.2, in all members of $[\mathcal{G}_T]$. Suppose a mixed graph contains the triple as a non-collider, but contrary to what the rule requires, has an arrowhead at β on the edge between β and γ. Then the edge between β and α is out of β. If it is $\beta —\alpha$, then the graph is not ancestral because there is an arrowhead at β; if it is $\beta \rightarrow \alpha$ that appears in the graph, recall that by the definition of discriminating path (Definition 3.1.7), α is a parent of γ. So β is an ancestor of γ, which, together with our supposition, makes the graph not ancestral. Furthermore, if the edge between β and γ is $\beta —\gamma$, then the graph is not ancestral because α is a parent of γ. Therefore, any MAG equivalent to $\mathcal{G}_T$ has to contain $\beta \rightarrow \gamma$, as the rule requires.

Case 2: $\beta \in \mathit{Sepset}(\theta, \gamma)$. By the CFC and Lemma 3.2.2, the triple $\langle \alpha, \beta, \gamma \rangle$ has to be a collider in $\mathcal{G}_T$, and also, by Proposition 3.1.2, in all members of $[\mathcal{G}_T]$. Also, by the definition of discriminating path, α is a collider on the path, which means $\alpha \leftrightarrow \beta \leftarrow\!\!* \gamma$ is in any MAG equivalent to $\mathcal{G}_T$. Furthermore, because $\alpha \rightarrow \gamma$ is present, $\alpha \leftrightarrow \beta \leftarrow \gamma$ will make the graph not ancestral. Therefore, $\alpha \leftrightarrow \beta \leftrightarrow \gamma$ is in any MAG equivalent to $\mathcal{G}_T$. $\square$

The real challenge for this chapter, however, is to demonstrate the arrowhead completeness of these rules, which deserves (and requires) a whole new section.

3.3 Arrowhead Completeness of the FCI Algorithm

The soundness result tells us that given a perfect oracle of conditional independence, the FCI procedure outputs a PAG whose informative (i.e., non-circle) marks are all valid, assuming the CMC and CFC. Henceforth we use $\mathcal{P}_{FCI}$ to denote the PAG output by FCI. We aim to show that $\mathcal{P}_{FCI}$ is also arrowhead complete. For this to be true, it has to be the case that every circle in $\mathcal{P}_{FCI}$ can be oriented as a tail. What this means is that for every circle in $\mathcal{P}_{FCI}$, there exists a MAG in the Markov equivalent class of the true causal MAG such that the corresponding mark is a tail.

To the end of showing this we need to establish a few properties of $\mathcal{P}_{FCI}$. Certain properties are already evident given soundness. For example, the defining properties of ancestral graphs, i.e., (a1)-(a3) in Definition 3.1.1 all hold of $\mathcal{P}_{FCI}$, and there is no inducing path between two non-adjacent vertices. Other lemmas are not as obvious, and some of them require quite non-trivial demonstrations. We will present the main argument in section 3.3.1, postponing some long proofs to 3.3.3. In section 3.3.2 we discuss the significance of the arrowhead completeness result.

3.3.1 The Main Argument

The first lemma establishes a property crucial for the subsequent argument, which is analogous to the one established by Meek (1995) in the context of inferring DAGs.

We refer to the property as **CP**1.

Lemma 3.3.1. *In $\mathcal{P}_{FCI}$, the following property holds:*

CP1 *for any three vertices A, B, C, if $A{*}\!\!\rightarrow B \circ\!\!-\!\!* C$, then there is an edge between A and C with an arrowhead at C, namely, $A{*}\!\!\rightarrow C$. Furthermore, if the edge between A and B is $A \rightarrow B$, then the edge between A and C is either $A \rightarrow C$ or $A \circ\!\!\rightarrow C$ (i.e., it is not $A \leftrightarrow C$).*

Proof. See section 3.3.3. □

As we shall see, the property **CP**1 is a defining feature of CPAGs in general, and is of central importance in the ensuing argument. We first derive an easy corollary, which concerns *circle paths*, which are defined to be paths on which every mark is a circle, or equivalently, every edge is of the type $\circ\!\!-\!\!\circ$.

Lemma 3.3.2. *In $\mathcal{P}_{FCI}$, for any two vertices A and B, if there is a circle path, i.e., a path consisting of $\circ\!\!-\!\!\circ$ edges, between A and B, then:*

(i) if there is an edge between A and B, the edge is not into A or B.

(ii) for any other vertex C, $C{*}\!\!\rightarrow A$ if and only if $C{*}\!\!\rightarrow B$. Furthermore, $C \leftrightarrow A$ if and only if $C \leftrightarrow B$.

Proof. We do induction on the length of the circle path. For (i), the base case is trivial. In the inductive step, suppose the proposition holds when there is a circle path consisting of n $\circ\!\!-\!\!\circ$ edges between any two vertices. Consider the case in which the circle path between A and B has $n+1$ edges. Let D be the vertex adjacent to B on the circle path. By the inductive hypothesis, the edge between A and D, if any, is not into D. This implies that the edge between A and B, if any, is not into B, otherwise **CP**1 does not hold of the triple $A{*}\!\!\rightarrow B \circ\!\!-\!\!\circ D$. By symmetry, the edge between A and B, if any, is not into A either. Hence (i) is true.

For (ii), notice that it is a direct consequence of **CP**1 that if $A \circ\!\!-\!\!\circ B$, then for any other vertex C, $C{*}\!\!\rightarrow A$ iff $C{*}\!\!\rightarrow B$. Furthermore, if $C \leftrightarrow A$, then the edge between C and B can't be $C \circ\!\!\rightarrow B$, for then the triple $A \leftrightarrow C \circ\!\!\rightarrow B$ would violate **CP**1. Neither can it be $C \rightarrow B$, for then $C \rightarrow B \circ\!\!-\!\!\circ A$ would violate **CP**1 (because $C \leftrightarrow A$ is present). So it has to be $C \leftrightarrow B$. By symmetry, $C \leftrightarrow B$ also implies $C \leftrightarrow A$. Thus the base case holds. The inductive step is similar to that in (i). □

Since $\mathcal{P}_{FCI}$ is sound, i.e., every non-circle mark therein is invariant in $[\mathcal{G}_T]$ (where $\mathcal{G}_T$ denotes the true causal MAG), any MAG equivalent to $\mathcal{G}_T$ should contain the non-circle marks in $\mathcal{P}_{FCI}$. Therefore, every MAG equivalent to $\mathcal{G}$ is a further orientation of $\mathcal{P}_{FCI}$ in the sense of changing circles into arrowheads or tails. Again, to prove arrowhead completeness, we need to show that every circle can be oriented into a tail in some MAG orientation of $\mathcal{P}_{FCI}$. We now define a general orientation operation on general *partial mixed graphs*, graphs that can contain the three kinds of marks.

Definition 3.3.1 (Tail Augmentation)**.** *Let $\mathcal{H}$ be any partial mixed graph.* **Tail augmentation** *of $\mathcal{H}$ is defined as the following set of operations on $\mathcal{H}$:*

- *change all $\circ\!\!\rightarrow$ edges into directed edges $\rightarrow$;*
- *change all $\circ\!\!-$ edges into undirected edges $-$;*
- *for any $A \circ\!\!-\!\!\circ B$, if there is no arrowhead into A or B, then change the edge into an undirected edge $A - B$.*[16]

[16] A more radical tail augmentation that changes every circle into a tail would give us a Joined Graph as defined by Ali (2002), who also provided a simple extension of the m-separation criterion that can be applied to the augmented graph.

The resulting graph is called the **tail augmented graph (TAG)** *of $\mathcal{H}$, denoted by $\mathcal{H}^{tag}$.*

Obviously the tail augmentation changes some circles in a partial mixed graph into tails, but does not introduce any new arrowhead or affect any non-circle mark already in the graph. Furthermore, after the tail augmentation, all remaining circles in the graph, if any, belong to $\circ\!\!-\!\!\circ$ edges.

Now consider the TAG of $\mathcal{P}_{FCI}$, $\mathcal{P}_{FCI}^{tag}$. The next lemma establishes some important properties of $\mathcal{P}_{FCI}^{tag}$.

Lemma 3.3.3. *Let $\mathcal{P}_{FCI}^{tag}$ be the TAG of $\mathcal{P}_{FCI}$. In $\mathcal{P}_{FCI}^{tag}$,*

(i) (a1)-(a3) (in Definition 3.1.1) and **CP**1 *hold;*

(ii) there is no inducing path between two non-adjacent vertices;

(iii) there is no such triple as $A - B \circ\!\!-\!\!\circ C$; and

(iv) there is no chordless cycle consisting of $\circ\!\!-\!\!\circ$ edges, i.e., there is no cycle of length 4 or more consisting of $\circ\!\!-\!\!\circ$ edges without an edge (chord) linking two non-consecutive vertices on the cycle.

Proof. See section 3.3.3. □

Let the **circle component** of any partial mixed graph be the induced subgraph that consists of all $\circ\!\!-\!\!\circ$ edges in the graph. We denote the circle component of $\mathcal{P}_{FCI}^{tag}$ by $\mathbf{C}(\mathcal{P}_{FCI}^{tag})$. A key implication of **CP**1 is that no matter how we orient the remaining $\circ\!\!-\!\!\circ$ edges (i.e., $\mathbf{C}(\mathcal{P}_{FCI}^{tag})$), no new unshielded colliders or directed cycles or almost directed cycles would be created that involve the arrowheads already present in $\mathcal{P}_{FCI}^{tag}$. This implication will be explored in the next lemma, which shows that if we orient

$\mathbf{C}(\mathcal{P}^{tag}_{FCI})$ into a directed acyclic graph with no unshielded colliders, then the resulting graph is a maximal ancestral graph and is Markov equivalent to $\mathcal{G}_T$.

Lemma 3.3.4. *Let $\mathcal{P}^{tag}_{FCI}$ be the TAG of $\mathcal{P}_{FCI}$. If we further orient $\mathbf{C}(\mathcal{P}^{tag}_{FCI})$, the circle component of $\mathcal{P}^{tag}_{FCI}$, into a DAG with no unshielded colliders, the resulting graph is a MAG and is Markov equivalent to $\mathcal{G}$.*

Proof. See section 3.3.3. □

We are almost there. To orient $\mathbf{C}(\mathcal{P}^{tag}_{FCI})$ into a DAG is trivial: an arbitrary ordering over the vertices in $\mathbf{C}(\mathcal{P}^{tag}_{FCI})$ would do. But that does not in general yield a DAG with no unshielded colliders. In fact, as is well known, an undirected graph can be oriented into a DAG with no unshielded colliders if and only if it is chordal (see, e.g., Meek 1995). A graph is **chordal** (a.k.a. **triangular**) if there is no cycle of length 4 or more without an edge (chord) linking two non-consecutive vertices on the cycle (cf. (iv) in Lemma 3.3.3). As we expect, $\mathbf{C}(\mathcal{P}^{tag}_{FCI})$ is indeed chordal.

Lemma 3.3.5. *The circle component of $\mathcal{P}^{tag}_{FCI}$, $\mathbf{C}(\mathcal{P}^{tag}_{FCI})$, is chordal.*

Proof. Suppose for contradiction that there is a cycle $\langle V_0, V_1, \cdots, V_{n-1}, V_n, V_0\rangle$ in $\mathbf{C}(\mathcal{P}^{tag}_{FCI})$ such that no non-consecutive vertices on the cycle are adjacent. We argue that the cycle is also chordless in $\mathcal{P}^{tag}_{FCI}$, which contradicts proposition (iv) of Lemma 3.3.3. Suppose on the contrary that in $\mathcal{P}^{tag}_{FCI}$ there is an edge linking two nonadjacent vertices on the cycle, say, V_i and V_j. The edge is either $V_i - V_j$ or is into at least one of them. By (iii) of Lemma 3.3.3, there is no such pattern as $- \circ\!\!-\!\!\circ$ in $\mathcal{P}^{tag}_{FCI}$, so the former case is impossible. By Lemma 3.3.2, since there is a circle path between V_i and V_j, the edge between V_i and V_j, if any, is not into V_i or V_j in $\mathcal{P}_{FCI}$, and hence is not into V_i or V_j in $\mathcal{P}^{tag}_{FCI}$. So the latter case is also impossible. Hence the cycle is also chordless in $\mathcal{P}^{tag}_{FCI}$, a contradiction to (iv) of Lemma 3.3.3. □

Therefore, $\mathbf{C}(\mathcal{P}_{FCI}^{tag})$ can be oriented into a DAG with no unshielded colliders. By Lemma 3.3.4, there is a MAG in $[\mathcal{G}_T]$ — the Markov equivalence class to which $\mathcal{G}_T$ belongs — in which all circles except possibly ones in $\mathbf{C}(\mathcal{P}_{FCI}^{tag})$ are marked as tails. So the circles outside $\mathbf{C}(\mathcal{P}_{FCI}^{tag})$ do not hide invariant arrowheads. The next lemma due to Meek (1995) entails that the circles in $\mathbf{C}(\mathcal{P}_{FCI}^{tag})$ do not hide invariant arrowheads either.

Lemma 3.3.6 **(Meek 1995).** *Let X and Y be any two vertices adjacent in a chordal graph. That graph can be oriented into a directed acyclic graph with no unshielded colliders in which the edge between X and Y is oriented as $X \to Y$.*

The main theorem of this chapter is now evident.

Theorem 3.3.1. *Assuming the CMC and CFC, if a perfect oracle of conditional independence is given as input, the FCI algorithm outputs a PAG that is both sound and complete with respect to arrowheads.*

Proof. It follows readily from Lemma 3.3.4, Lemma 3.3.5 and Lemma 3.3.6 that for every circle in $\mathcal{P}_{FCI}$, there is a member of $[\mathcal{G}_T]$ in which the circle is oriented as a tail. □

3.3.2 Significance of the Arrowhead Completeness Result

The significance of Theorem 3.3.1 is not simply what it literally says, that every invariant arrowhead in the true causal MAG is contained in the FCI output. The further implication lies in the fact that from the sound and arrowhead complete PAG output by the FCI algorithm, we can actually read off all valid negative causal sentences, sentences of the form that "variable A is *not* a cause of variable B" or more accurately, "there is no causal pathway from variable A to variable B in the true causal structure." The way to read them off is by checking what we call *potentially anterior path*. In a PAG, a path between X and Y is a **potentially anterior path** from X to Y if there is no arrowhead on the path pointing towards X.[17] (Similarly, a path between X and Y in a MAG is a *anterior path* from X to Y if there is no arrowhead on the path pointing towards X.) If there is no potentially anterior path in $\mathcal{P}_{FCI}$ from a variable A to another variable B, there is no anterior or partially directed path in any member of $[\mathcal{G}_T]$ from A to B, and hence there cannot be any causal pathway from A to B in the (unknown) true causal DAG. Conversely, if there is a potentially anterior path from A to B in $\mathcal{P}_{FCI}$, our argument in this section makes it easy to derive that in some member of $[\mathcal{G}_T]$ there is an anterior or partially directed path from A to B, and hence the sentence "variable A is *not* a cause of variable B" is not valid[18] in the sense of being true in every model compatible with the given oracle assuming the CMC and CFC. Therefore, the FCI algorithm, as a causal inference system, is actually complete with respect to deriving valid negative causal sentences from a perfect oracle of conditional independence and the two axioms.

What about positive causal sentences? We turn to this question in the next chapter.

3.3.3 Omitted Proofs

Proof of Lemma 3.3.1

[17] This notion is closely related to the notion of *potentially directed path* defined in the next chapter. In $\mathcal{P}_{FCI}$, they amount to the same thing as there are no such edges as $\circ\!\!-$ or $-$ in $\mathcal{P}_{FCI}$. In graphs where $\circ\!\!-$ edges or $-$ edges are present, however, a potentially directed path is also a potentially anterior path but not necessarily vice versa.

[18] Strictly speaking, if we consider the possibility of selection effect, the fact that there is a directed path in a MAG from A to B means only that there is a causal pathway (in the underlying true DAG) from A to either B or to some selection variable. However, it is easy to show that for such a MAG, there is always a DAG (with possibly latent and selection variables) compatible with this MAG such that a causal pathway from A to B is present in the DAG, though this is not necessarily true in all DAGs compatible with the MAG.

Proof. Let $\mathbf{M} = \{Y | \exists X, Z$ such that $X{*}\!\!\rightarrow Y \circ\!\!-\!\!* Z$ but not $X{*}\!\!\rightarrow Z$ is in $\mathcal{P}_{FCI}\}$. We need to show that $\mathbf{M}$ is empty. Suppose for the sake of contradiction that $\mathbf{M}$ is not empty. Let Y_0 be a vertex in $\mathbf{M}$ such that no *proper ancestor*[19] of Y_0 in $\mathcal{P}_{FCI}$ is in $\mathbf{M}$. (This specification is legitimate because there is no directed cycle in $\mathcal{P}_{FCI}$.) Let $\mathbf{F}_{Y_0} = \{X | \exists Z$ such that $X{*}\!\!\rightarrow Y_0 \circ\!\!-\!\!* Z$ but not $X{*}\!\!\rightarrow Z$ is in $\mathcal{P}_{FCI}\}$. Since $Y_0 \in \mathbf{M}$, $\mathbf{F}_{Y_0}$ is not empty. Choose X_0 in $\mathbf{F}_{Y_0}$ such that no proper *descendant* of X_0 in $\mathcal{P}_{FCI}$ is in $\mathbf{F}_{Y_0}$. Finally, choose any Z_0 such that $X_0{*}\!\!\rightarrow Y_0 \circ\!\!-\!\!* Z_0$ but not $X_0{*}\!\!\rightarrow Z_0$ is in $\mathcal{P}_{FCI}$. We will manage to derive a contradiction out of this.

Note that X_0 and Z_0 must be adjacent, otherwise the circle at Y_0 on $Y_0 \circ\!\!-\!\!* Z_0$ would have been oriented by $\mathcal{R}0$ or $\mathcal{R}1$. Furthermore, the edge between X_0 and Z_0 is not out of Z_0, i.e., the mark at Z_0 on the edge is not a tail. The reason is this: it is evident that no — or ∘— could result from applications of $\mathcal{R}0 - \mathcal{R}4$, and hence none is present in $\mathcal{P}_{FCI}$. So if the edge between X_0 and Z_0 is out of Z_0, then it must be $X_0 \leftarrow Z_0$. But then $Y_0 \circ\!\!-\!\!* Z_0$ would have been oriented as $Y_0 \leftarrow\!\!* Z_0$ by $\mathcal{R}2$. This is a contradiction. Hence the edge between X_0 and Z_0 is not out of Z_0. Since by our supposition, the edge is not into Z_0 either, the mark at Z_0 on the edge between X_0 and Z_0 has to be a circle, namely $X_0 *\!\!-\!\!\circ Z_0$.

Below we enumerate the ways in which the arrowhead at Y_0 on $X_0{*}\!\!\rightarrow Y_0$ could have been oriented, and derive a contradiction in each case.

Case 1: $X_0{*}\!\!\rightarrow Y_0$ is oriented by $\mathcal{R}0$. That means there is a vertex W (distinct from Y_0 or Z_0) such that W is not adjacent to X_0, and $X_0{*}\!\!\rightarrow Y_0 \leftarrow\!\!* W$ appears in $\mathcal{P}_{FCI}$. This implies that Z_0 and W are adjacent, for otherwise the circle at Y_0 on $Y_0 \circ\!\!-\!\!* Z_0$ would have been oriented by either $\mathcal{R}0$ or $\mathcal{R}1$. Furthermore, because $X_0 *\!\!-\!\!\circ Z_0$, it is not the case that $Z_0 \leftarrow\!\!* W$, otherwise the circle at Z_0 would have been oriented by $\mathcal{R}0$ or $\mathcal{R}1$. It follows that either $Z_0 \circ\!\!-\!\!* W$ or $Z_0 \rightarrow W$ (again, because

[19] A **proper ancestor** of a vertex is an ancestor distinct from the vertex itself.

no — or —∘ is present). In the former case, $X_0 *\!\!-\!\!\circ Z_0 \circ\!\!-\!\!* W$ and $X_0 *\!\!\rightarrow Y_0 \leftarrow\!\!* W$, and hence $Y_0 \circ\!\!-\!\!* Z_0$ should have been oriented as $Y_0 \leftarrow\!\!* Z_0$ by $\mathcal{R}3$; in the latter case, $Z_0 \rightarrow W *\!\!\rightarrow Y_0$, and hence $Y_0 \circ\!\!-\!\!* Z_0$ should have been oriented as $Y_0 \leftarrow\!\!* Z_0$ by $\mathcal{R}2$. So in either case it is a contradiction.

Case 2: $X_0 *\!\!\rightarrow Y_0$ is oriented by $\mathcal{R}1$, which means that there is a vertex W (distinct from Y_0) not adjacent to Y_0 such that $W *\!\!\rightarrow X_0 \rightarrow Y_0$ is in $\mathcal{P}_{FCI}$. It is not the case that $X_0 \leftarrow\!\!\circ Z_0$, otherwise $Y_0 \circ\!\!-\!\!* Z_0$ would have been oriented by $\mathcal{R}2$ to be $Y_0 \leftarrow\!\!* Z_0$. So $X_0 \circ\!\!-\!\!\circ Z_0$ is in $\mathcal{P}_{FCI}$. It follows that W and Z_0 are adjacent, otherwise the circle at X_0 on $X_0 \circ\!\!-\!\!\circ Z_0$ would be oriented by $\mathcal{R}0$ or $\mathcal{R}1$. Now the unshielded triple $Y_0 \circ\!\!-\!\!* Z_0 *\!\!-\!\!* W$ cannot be a collider, for otherwise $X_0 \rightarrow Y_0 \circ\!\!\rightarrow Z_0$, and $X_0 \circ\!\!-\!\!\circ Z_0$ would be oriented as $X_0 \circ\!\!\rightarrow Z_0$ by $\mathcal{R}2$. Since it is a non-collider, it cannot be that $W *\!\!\rightarrow Z_0$, otherwise $Y_0 \circ\!\!-\!\!* Z_0$ would be oriented as $Y_0 \leftarrow Z_0$. Now we have $W *\!\!\rightarrow X_0 \circ\!\!-\!\!\circ Z_0$ but not $W *\!\!\rightarrow Z_0$ in $\mathcal{P}_2$. So X_0 is in $\mathbf{M}$ and is a parent of Y_0 in $\mathcal{P}_{FCI}$, which contradicts our choice of Y_0.

Case 3: $X_0 *\!\!\rightarrow Y_0$ is oriented by $\mathcal{R}2$. There are two sub-cases to consider.

Case 3.1: There is a vertex W (distinct from Z_0) such that $X_0 \rightarrow W *\!\!\rightarrow Y_0$ appears in $\mathcal{P}_{FCI}$. Then W and Z_0 must be adjacent, for otherwise the circle at Y_0 on $Y_0 \circ\!\!-\!\!* Z_0$ would be oriented by either $\mathcal{R}0$ or $\mathcal{R}1$. Furthermore, it is not the case that $W *\!\!\rightarrow Z_0$, otherwise by $\mathcal{R}2$, $X_0 *\!\!\rightarrow Z_0$, a contradiction. Now we have $W *\!\!\rightarrow Y_0 \circ\!\!-\!\!* Z_0$ but not $W *\!\!\rightarrow Z_0$. So W is in $\mathbf{F}_{Y_0}$ and is a child of X_0, which contradicts our choice of X_0.

Case 3.2: There is a vertex W (distinct from Z_0) such that $X_0 *\!\!\rightarrow W \rightarrow Y_0$ appears in $\mathcal{P}_{FCI}$. Again, W and Z_0 must be adjacent, for the same reason as in *3.1*. Furthermore, it must be the case that $W \circ\!\!-\!\!* Z_0$. If not, either $W \rightarrow Z_0$ or $W \leftarrow\!\!* Z_0$. In the former case, $\mathcal{R}2$ would dictate that $X_0 *\!\!\rightarrow Z_0$, which contradicts our assumption; in the latter case, $\mathcal{R}2$ would dictate that $Y_0 \leftarrow\!\!* Z_0$, which also contradicts our assumption. Now we have $X_0 *\!\!\rightarrow W \circ\!\!-\!\!* Z_0$ but not $X_0 *\!\!\rightarrow Z_0$. So W is in $\mathbf{M}$ and is a parent of Y_0, which contradicts our choice of Y_0.

Case 4: $X_0 *\!\!\rightarrow Y_0$ is oriented by $\mathcal{R}3$. That means there are two non-adjacent vertices U and V (both distinct from Z_0) such that $U *\!\!-\!\!* X_0 *\!\!-\!\!* V$ is a non-collider (which, at the time $X_0 *\!\!\rightarrow Y_0$ gets oriented, is $U *\!\!-\!\!\circ X_0 \circ\!\!-\!\!* V$ as required by the antecedent of $\mathcal{R}3$), and $U *\!\!\rightarrow Y_0 \leftarrow\!\!* V$ is a collider in $\mathcal{P}_{FCI}$. U and V must be adjacent to Z_0, otherwise the circle at Y_0 on $Y_0 \circ\!\!-\!\!* Z_0$ would be oriented by either $\mathcal{R}0$ or $\mathcal{R}1$. Furthermore, since $U *\!\!-\!\!* X_0 *\!\!-\!\!* V$ is a non-collider, either $U *\!\!-\!\!\circ X_0 \circ\!\!-\!\!* V$, or $U \leftarrow X_0$, or $X_0 \rightarrow V$ appears in $\mathcal{P}_{FCI}$. It follows that the triple $\langle U, Z_0, V\rangle$ is not a collider, otherwise $X_0 *\!\!-\!\!\circ Z_0$ should be oriented as $X_0 *\!\!\rightarrow Z_0$ by $\mathcal{R}3$ or $\mathcal{R}2$, contrary to our assumption. Also, neither $Z_0 \rightarrow U$ nor $Z_0 \rightarrow V$ is the case, otherwise $Y_0 \circ\!\!-\!\!* Z_0$ should be oriented as $Y_0 \leftarrow\!\!* Z_0$ by $\mathcal{R}2$, contrary to our assumption. Then it must be the case that $U *\!\!-\!\!\circ Z_0 \circ\!\!-\!\!* V$. Again, by $\mathcal{R}3$, $Y_0 \circ\!\!-\!\!* Z_0$ should be oriented as $Y_0 \leftarrow\!\!* Z_0$, a contradiction.

Case 5: $X_0 *\!\!\rightarrow Y_0$ is oriented by $\mathcal{R}4$. There are three sub-cases to consider.

Case 5.1: There is a discriminating path $u = \langle U, ..., W, X_0, Y_0\rangle$ for X_0 in $\mathcal{P}_{FCI}$ (and $X_0 \in \mathit{Sepset}(U, Y_0)$), which orients the edge as $X_0 \rightarrow Y_0$. By the definition of discriminating path, $W \leftarrow\!\!* X_0$ and $W \rightarrow Y_0$ (and $W \neq Z_0$). So W is adjacent to Z_0, otherwise the circle at Y_0 on $Y_0 \circ\!\!-\!\!-\!\!* Z_0$ would have been oriented by either $\mathcal{R}0$ or $\mathcal{R}1$. It is not the case that $W \rightarrow Z_0$, for otherwise $X_0 *\!\!-\!\!\circ Z_0$ would be oriented as $X_0 *\!\!\rightarrow Z_0$ by $\mathcal{R}2$. It is not the case that $W \leftarrow\!\!* Z_0$, for otherwise $Y_0 \circ\!\!-\!\!* Z_0$ would be oriented as $Y_0 \leftarrow\!\!* Z_0$ by $\mathcal{R}2$, contrary to our assumption. So it has to be that $W \circ\!\!-\!\!* Z_0$. Now we have $X_0 *\!\!\rightarrow W \circ\!\!-\!\!* Z_0$ but not $X_0 *\!\!\rightarrow Z_0$. So W is in $\mathbf{M}$ and is a parent of Y_0 in $\mathcal{P}_{FCI}$, which contradicts our choice of Y_0.

Case 5.2: There is a discriminating path $u = \langle U, ..., X_0, Y_0, W\rangle$ for Y_0 in $\mathcal{P}_{FCI}$ (and $Y_0 \notin Sepset(U, W)$), which orients the triple as $X_0 \leftrightarrow Y_0 \leftrightarrow W$. It follows that $W \neq Z_0$. Moreover, W is adjacent to Z_0; if not, the circle at Y_0 on $Y_0 \circ\!\!-\!\!* Z_0$ would have been oriented by either $\mathcal{R}0$ or $\mathcal{R}1$. By the definition of discriminating path, X_0 is a parent of W, i.e., $X_0 \to W$. Hence it is not the case that $W *\!\!\to Z_0$, otherwise $X_0 *\!\!\to Z_0$ by $\mathcal{R}2$, contrary to what we established at the beginning. So we have $W \leftrightarrow Y_0 \circ\!\!-\!\!* Z_0$ but not $W *\!\!\to Z_0$, which means W is in $\mathbf{F}_{Y_0}$. But W is a child of X_0, which contradicts our choice of X_0.

Case 5.3: There is a discriminating path $u = \langle U, ..., W, Y_0, X_0\rangle$ for Y_0 in $\mathcal{P}_{FCI}$ (and $Y_0 \notin Sepset(U, X_0)$), which orients the triple as $W \leftrightarrow Y_0 \leftrightarrow X_0$. It follows that $W \neq Z_0$. The contradiction in this case is the least obvious, and needs several non-trivial steps to be revealed.

Note that Z_0 is not on u because it is not the case that $Z_0 \to X_0$ as we showed at the beginning. As a first step, we show that for every vertex Q on u between U and W (including W but not U), it is not the case that $Q \leftarrow\!\!* Z_0$. Otherwise, for any such Q, $u(U, Q) \oplus Q \leftarrow\!\!* Z_0 \oplus Z_0 \circ\!\!-\!\!* X_0$ is a discriminating path for Z_0. (We use $\oplus$ to denote the concatenation operation of paths). So the circle at Z_0 on $Z_0 \circ\!\!-\!\!* X_0$ would be oriented by $\mathcal{R}4$, a contradiction.

Next, we establish that every vertex on u between U and W (including U and W) is adjacent to Z_0. Suppose not, let V be the closest vertex to Y_0 on $u(U, W)$ that is not adjacent to Z_0. If $V = W$, the circle at Y_0 on $Y_0 \circ\!\!-\!\!* Z_0$ would have been oriented by either $\mathcal{R}0$ or $\mathcal{R}1$. If $V \neq W$, let T be the first vertex after V on $u(V, W)$, which is adjacent to Z_0 (because of our choice of V). Because u is a discriminating path, the edge between V and T is $V *\!\!\to T$. Since $\langle V, T, Z_0\rangle$ is an unshielded triple, the edge between T and Z_0 is either $T \to Z_0$ or $T \leftarrow\!\!* Z_0$ (by either $\mathcal{R}1$ or $\mathcal{R}0$. Hence it is

$T \rightarrow Z_0$, as the latter case has been ruled out in the previous step. Then we can show that every vertex on $u(T, W)$ (including W) is a parent of Z_0. Otherwise, let R be the closest vertex to T on $u(T, W)$ that is not a parent of Z_0. Then $u(V, R) \oplus R \ast\!\!-\!\!\ast Z_0$ is a discriminating path for R (because every vertex between T and R is a parent of Z_0, by our choice of R). Since it is not the case that $R \rightarrow Z_0$, the edge between R and Z_0 must be oriented as $R \leftrightarrow Z_0$, which, however, has been shown impossible. Hence every vertex on $u(T, W)$ (including W) is a parent of Z_0. Then $u(V, Y_0) \oplus Y_0 \circ\!-\!\ast Z_0$ is a discriminating path for Y_0, which means the circle at Y_0 on $Y_0 \circ\!-\!\ast Z_0$ would have been oriented by $\mathcal{R}4$, a contradiction. So every vertex on $u(U, W)$ (including U and W) is adjacent to Z_0.

The contradiction we are about to carry out is on the adjacency between U and Y_0. We first argue that U is not adjacent to Y_0. Suppose for contradiction that they are adjacent. By the definition of discriminating path, U is not adjacent to X_0, so $U \ast\!\!-\!\!\ast Y_0 \leftrightarrow X_0$ is an unshielded triple. It follows that either $\mathcal{R}0$ or $\mathcal{R}1$ could apply here, and it is either $U \leftarrow Y_0 \leftrightarrow X_0$ or $U \ast\!\!\rightarrow Y_0 \leftrightarrow X_0$. The former case is impossible, because that would make u an inducing path between U and X_0, two non-adjacent vertices, which contradicts the maximality of $\mathcal{G}$. In the latter case, we claim that $U \ast\!\!-\!\!\circ Z_0$ is present. Otherwise, either $U \ast\!\!\rightarrow Z_0$ or $U \leftarrow Z_0$. In the former case the circle at Z_0 on $Z_0 \circ\!\!-\!\!\ast X_0$ would be oriented by either $\mathcal{R}0$ or $\mathcal{R}1$; in the latter case, let S be the vertex next to U on u. Then by $\mathcal{R}2$, the edge between S and Z_0 would be oriented as $S \leftarrow\!\!\ast Z_0$, which we have shown to be impossible. So $U \ast\!\!-\!\!\circ Z_0 \circ\!\!-\!\!\ast X_0$ is present. Now, by $\mathcal{R}3$, the edge between Y_0 and Z_0 would be oriented as $Y_0 \leftarrow\!\!\ast Z_0$, a contradiction. Hence U and Y_0 are not adjacent.

An immediate corollary of the above argument is that it is not the case that $Y_0 \circ\!\!\rightarrow Z_0$. Otherwise, the edge between U and Z_0 would be oriented either as $U \ast\!\!\rightarrow Z_0$

or as $U \leftarrow Z_0$ (because U and Y_0 are not adjacent). But neither of the two cases can be true, as shown in the above argument. It follows that $Y_0 \circ\!\!-\!\!\circ Z_0$ is present.

Now we are ready to complete the argument. We show (by induction) that every vertex on u between W and U, and in particular U, is adjacent to Y_0, which yields a contradiction. Obviously W is adjacent to Y_0. In the inductive step, we show that if a vertex S_1 between W and U on u is adjacent to Y_0, then the next vertex S_2 (the one further from W), if any (i.e., when $S_1 \neq U$), is also adjacent to Y_0. Suppose otherwise, that S_1 is adjacent to Y_0 but S_2 is not. Because $S_2 *\!\!-\!\!* Z_0 \circ\!\!-\!\!\circ Y_0$ is an unshielded triple, it is not the case that $S_2 *\!\!\rightarrow Z_0$. Note that this further rules out $S_1 \rightarrow Z_0$, as the latter implies the former by $\mathcal{R}2$. $S_2 \leftarrow Z_0$ is also impossible, for in that case we have $S_1 \leftarrow\!\!* Z_0$ by $\mathcal{R}2$, which we have ruled out. Hence the only possible case is $S_2 *\!\!-\!\!\circ Z_0 \circ\!\!-\!\!\circ Y_0$. Now let us focus on the triple $S_2 *\!\!\rightarrow S_1 *\!\!-\!\!* Y_0$. It is an unshielded triple, which implies either $S_1 \leftarrow\!\!* Y_0$ or $S_1 \rightarrow Y_0$. In the former case, we can apply $\mathcal{R}3$ to orient the edge between S_1 and Z_0 as $S_1 \leftarrow\!\!* Z_0$, which we have shown to be impossible; in the latter case, since neither $S_1 \rightarrow Z_0$ nor $S_1 \leftarrow\!\!* Z_0$ can be true, it must be $S_1 \circ - * Z_0$. Thus we have $S_2 *\!\!\rightarrow S_1 \circ - * Z_0$ but not $S_2 *\!\!\rightarrow Z_0$. Hence S_1 is in $\mathbf{M}$ and is a parent of Y_0. This contradicts our choice of Y_0. So S_2 is also adjacent to Y_0, and by induction U is also adjacent to Y_0. Hence a contradiction, which concludes *Case 5.3*.

This enumeration exhausts the possible ways that the arrowhead at Y_0 on $X_0 *\!\!\rightarrow Y_0$ was oriented. Hence, the initial supposition that $\mathbf{M}$ is non-empty leads to contradiction. Furthermore, for any $A \rightarrow B \circ\!\!\rightarrow\!\!* C$ in $\mathcal{P}_{FCI}$, it is not the case that $A \leftrightarrow C$, for otherwise the circle at B on $B \circ\!\!-\!\!* C$ could be oriented as an arrowhead by $\mathcal{R}2$. Since we have shown that $A *\!\!\rightarrow C$ appears, it is either $A \rightarrow C$ or $A \circ\!\!\rightarrow C$. Therefore, $\mathbf{CP}1$ holds of $\mathcal{P}_{FCI}$. $\square$

Proof of Lemma 3.3.3

Proof. First we prove (i). For (a1), suppose for contradiction that there is a directed cycle. Let $c = \langle V_0, \cdots, V_n, V_0 \rangle$ be a shortest directed cycle in $\mathcal{P}_{FCI}^{tag}$, that is, no other directed cycle has fewer edges than c does. Since no directed cycle is present in $\mathcal{P}_{FCI}$, the corresponding cycle in $\mathcal{P}_{FCI}$ must contain a $\circ\!\!\rightarrow$ edge. That is, there exists i such that $V_{i-1}\ast\!\!\rightarrow V_i\circ\!\!\rightarrow V_{i+1}$ is in $\mathcal{P}_{FCI}$. Because **CP**1 holds of $\mathcal{P}_{FCI}$, there is an edge $V_{i-1}\ast\!\!\rightarrow V_{i+1}$ in $\mathcal{P}_{FCI}$. The edge can't be $V_{i-1} \leftrightarrow V_{i+1}$ for the following reason: the edge between V_{i-1} and V_i is either $V_{i-1}\circ\!\!\rightarrow V_i$ or $V_{i-1} \rightarrow V_i$. In the former case, the triple $V_{i+1} \leftrightarrow V_{i-1}\circ\!\!\rightarrow V_i$ would violate **CP**1; in the latter case, the circle at V_i on $V_i\circ\!-\!\ast\, V_{i+1}$ should have been oriented to an arrowhead by $\mathcal{R}2$. So either $V_{i-1}\circ\!\!\rightarrow V_{i+1}$ or $V_{i-1} \rightarrow V_{i+1}$ is in $\mathcal{P}_{FCI}$, which means $V_{i-1} \rightarrow V_{i+1}$ is in $\mathcal{P}_{FCI}^{tag}$. But then $\langle V_0, \cdots, V_{i-1}, V_{i+1}, \cdots, V_n, V_0 \rangle$ is a shorter cycle than c is, hence a contradiction. So there is no directed cycle in $\mathcal{P}_{FCI}^{tag}$.

For (a2), suppose for contradiction that there is an almost directed cycle in $\mathcal{P}_{FCI}^{tag}$. Let $c = \langle V_0, \cdots, V_n, V_0 \rangle$ be a shortest one. Without loss of generality, suppose the bi-directed edge in the cycle is $V_0 \leftrightarrow V_n$, and $\langle V_0, V_1, \cdots, V_n \rangle$ is a directed path from V_0 to V_n. It is obvious that $V_0 \leftrightarrow V_n$ is also in $\mathcal{P}_{FCI}$, because no extra arrowheads are introduced in $\mathcal{P}_{FCI}^{tag}$. Since no almost directed cycle is present in $\mathcal{P}_{FCI}$, the corresponding path between V_0 and V_n in $\mathcal{P}_{FCI}$ contains a $\circ\!\!\rightarrow$ edge. If the edge between V_0 and V_1 is not $\circ\!\!\rightarrow$, then there must exist $1 \leq i \leq n-1$ such that $V_{i-1}\ast\!\!\rightarrow V_i\circ\!\!\rightarrow V_{i+1}$ is in $\mathcal{P}_{FCI}$. By the same argument we went through in proving (a1), there is a shorter directed path from V_0 to V_n and hence a shorter almost directed cycle. So it is $V_0\circ\!\!\rightarrow V_1$ that appears in $\mathcal{P}_{FCI}$. Then by **CP**1, $V_n\ast\!\!\rightarrow V_1$ is in $\mathcal{P}_{FCI}$, which means that either $V_n \rightarrow V_1$ or $V_n \leftrightarrow V_1$ is in $\mathcal{P}_{FCI}^{tag}$. In the former case, there is a directed cycle in $\mathcal{P}_{FCI}^{tag}$, which we have shown to be impossible; in the latter case, there is a shorter almost directed cycle, a contradiction.

For (a3), note that any $X \text{—} Y$ in $\mathcal{P}_{FCI}^{tag}$ corresponds to either $X \text{—} Y$ or $X \circ\!\!-\!\!\circ Y$ in $\mathcal{P}_{FCI}$ (remember there is no $\circ\!\!-$ edge in $\mathcal{P}_{FCI}$). In the former case, there is no edge into X or Y due to the soundness of $\mathcal{P}_{FCI}$; in the latter case, the definition of tail augmentation guarantees that there is no edge into X or Y.

For **CP**1, note that no extra arrowheads are introduced in $\mathcal{P}_{FCI}^{tag}$ and hence any pattern of $*\!\!\rightarrow \circ\!\!-\!\!*$ in $\mathcal{P}_{FCI}^{tag}$ is also in $\mathcal{P}_{FCI}$. Since **CP**1 holds of $\mathcal{P}_{FCI}$, it also holds of $\mathcal{P}_{FCI}^{tag}$.

Next, we prove (ii). It is convenient to define the *rank* of an inducing path. By definition (Definition 3.1.4), an inducing path is one on which every vertex (except the endpoints) is a collider and is an ancestor of one of the endpoints. In other words, from each interior vertex on the path there is a directed path to one of the endpoints. Let the *rank* of each interior vertex on the path be the length of a shortest directed path from that vertex to one of the endpoints. We define the *rank* of an inducing path as the length of the path plus the sum of the ranks of the interior vertices.

Suppose for contradiction that in $\mathcal{P}_{FCI}^{tag}$ there is an inducing path between two non-adjacent vertices X and Y. Let $p = \langle X = V_0, V_1, \cdots, V_{n-1}, Y = V_n \rangle$ be the one of the lowest rank. By definition, V_i's ($1 \leq i \leq n-1$) are colliders on the path and are ancestors of either X or Y. This implies that V_1 is an ancestor of Y and V_{n-1} is an ancestor of X, otherwise there would be a directed or almost directed cycle in $\mathcal{P}_{FCI}^{tag}$, which we have shown to be absent. For the same reason, the edge between X and V_1 is $X \leftrightarrow V_1$, and the edge between V_{n-1} and Y is $V_{n-1} \leftrightarrow Y$. So every edge on p is bi-directed. Since no extra arrowheads are introduced in $\mathcal{P}_{FCI}^{tag}$, these bi-directed edges on p are also in $\mathcal{P}_{FCI}$.

Furthermore, note that $\mathcal{P}_{FCI}$ is sound and hence should not contain any inducing path between X and Y. It follows that not every interior vertex on p is an ancestor of X or Y in $\mathcal{P}_{FCI}$. Let V_j $(1 \leq j \leq n-1)$ be such a vertex, that is, V_j is not an ancestor of X or Y in $\mathcal{P}_{FCI}$. Without loss of generality, suppose in $\mathcal{P}_{FCI}^{tag}$, V_j is an ancestor of Y. Let d be a shortest directed path from V_j to Y in $\mathcal{P}_{FCI}^{tag}$. Since d is not a directed path in $\mathcal{P}_{FCI}$, d must contain a $\circ\!\!\rightarrow$ edge in $\mathcal{P}_{FCI}$. Since d is a shortest one, **CP**1 implies that $\circ\!\!\rightarrow$ can only appear as the first edge on the directed path (by the argument we have used several times above). That is, let V_{j1} be the vertex adjacent to V_j on d, then $V_j \circ\!\!\rightarrow V_{j1}$ is in $\mathcal{P}_{FCI}$.

Now we argue that if V_{j1} is not on $p(V_j, Y)$, then there is a V_k $(j+1 \leq k \leq n-1)$ such that $V_k \leftrightarrow V_{j1}$ is in $\mathcal{P}_{FCI}$. Suppose not; we prove by induction that for every $j+1 \leq i \leq n$, either $V_i \circ\!\!\rightarrow V_{j1}$ or $V_i \rightarrow V_{j1}$ is present in $\mathcal{P}_{FCI}$. The base case is easy. Since $V_{j+1} \leftrightarrow V_j \circ\!\!\rightarrow V_{j1}$ is in $\mathcal{P}_{FCI}$, by **CP**1, we have $V_{j+1} *\!\!\rightarrow V_{j1}$. Since it is not bi-directed by the supposition, it is either $V_{j+1} \circ\!\!\rightarrow V_{j1}$ or $V_{j+1} \rightarrow V_{j1}$. In the inductive step, suppose $V_{j+1}, ..., V_m$ all satisfy the claim, we argue that V_{m+1} also satisfies the claim. V_{m+1} must be adjacent to V_{j1}, otherwise either, for some k, $V_k \circ\!\!\rightarrow V_{j1}$ will be oriented as $V_k \leftrightarrow V_{j1}$ $(j+1 \leq k \leq m)$ by $\mathcal{R}4$, or all $V_k \circ\!\!\rightarrow V_{j1}$ will be oriented into $V_k \rightarrow V_{j1}$, and hence $V_j \circ\!\!\rightarrow V_{j1}$ will be oriented by $\mathcal{R}4$. Furthermore, the edge between V_{m+1} and V_{j1} is $V_{m+1} *\!\!\rightarrow V_{j1}$. This is because either $V_m \circ\!\!\rightarrow V_{j1}$ or $V_m \rightarrow V_{j1}$ appears. In the former case, $V_{m+1} *\!\!\rightarrow V_{j1}$ by **CP**1; in the latter case, $V_{m+1} *\!\!\rightarrow V_{j1}$ by $\mathcal{R}2$. Lastly, since the edge between V_{m+1} and V_{j1} is not bi-directed (in the case of $V_{m+1} = Y$, it is not $Y \leftrightarrow V_{j1}$ by (a2) because V_{j1} is an ancestor of Y in $\mathcal{P}_{FCI}^{tag}$), it is either $V_{m+1} \circ\!\!\rightarrow V_{j1}$ or $V_{m+1} \rightarrow V_{j1}$. This completes the induction. But then either $Y \circ\!\!\rightarrow V_{j1}$ or $Y \rightarrow V_{j1}$, which contradicts the fact that V_{j1} is an ancestor of Y in $\mathcal{P}_{FCI}^{tag}$.

So there is a V_k $(j+1 \leq k \leq n-1)$ such that $V_k \leftrightarrow V_{j1}$ is in $\mathcal{P}_{FCI}$.

By essentially the same argument, we can show that if V_{j1} is not on $p(X, V_j)$, then there is a V_h $(0 \leq h \leq j-1)$ such that $V_h \leftrightarrow V_{j1}$ is in $\mathcal{P}_{FCI}$. (The only difference is that we rule out the case $X \circ\!\!\rightarrow V_{j1}$ and the case $X \rightarrow V_{j1}$ not because V_{j1} is an ancestor of X, but because X cannot be an ancestor of Y in $\mathcal{P}^{tag}_{FCI}$, for otherwise an almost directed cycle would be present.) Therefore, if V_{j1} is not on p, then the path $\langle V_0 = X, \cdots, V_h, V_{j1}, V_k, \cdots, V_n = Y\rangle$ (V_h could be V_0) is an inducing path between X and Y but is of a lower rank than p, a contradiction. On the other hand, if V_{j1} is on p, without loss of generality, suppose it is on $p(X, V_j)$, then $\langle V_0 = X, \cdots, V_{j1}, V_k, \cdots, V_n = Y\rangle$ is an inducing path between X and Y but is of a lower rank than p, a contradiction. Hence there is no inducing path between two non-adjacent vertices in $\mathcal{P}^{tag}_{FCI}$.

It is easy to demonstrate (iii). Suppose for contradiction that there is such a triple $X - Y \circ\!\!-\!\!\circ Z$ in $\mathcal{P}^{tag}_{FCI}$. By the definition of tail augmentation and the fact that there is no $\circ\!\!-$ edge in $\mathcal{P}_{FCI}$, in $\mathcal{P}_{FCI}$ the edge between X and Y is either $X - Y$ or $X \circ\!\!-\!\!\circ Y$. In the first case, obviously there is no edge into Y or Z in $\mathcal{P}_{FCI}$, for otherwise an arrowhead would meet an undirected edge, which contradicts the soundness of $\mathcal{P}_{FCI}$. In the second case, since $X \circ\!\!-\!\!\circ Y$ is changed to $X - Y$ in the tail augmentation, there is no edge into Y in $\mathcal{P}_{FCI}$, which implies, by Lemma 3.3.2, that there is no edge into Z either. Either way, $Y \circ\!\!-\!\!\circ Z$ should be changed to $Y - Z$ in the tail augmentation.

Lastly we prove (iv). (iv) is actually obvious given a rule to be presented in the next chapter. Just notice that if there is any chordless cycle consisting of $\circ\!\!-\!\!\circ$ edges in $\mathcal{P}^{tag}_{FCI}$, which would of course also be present in $\mathcal{P}_{FCI}$, then every edge on that cycle should be undirected edges in the true causal MAG, as otherwise a directed cycle or an almost directed cycle should result in the MAG in light of $\mathcal{R}1$, which is a contradiction. Since $\mathcal{P}_{FCI}$ is sound, it follows that no edge is into any vertex on that cycle in $\mathcal{P}_{FCI}$. So the tail augmentation should have changed every $\circ\!\!-\!\!\circ$ on that cycle into undirected edges. $\square$

Proof of Lemma 3.3.4

Proof. Let $\mathcal{H}$ denote the resulting graph. We first show that $\mathcal{H}$ is a MAG. Since $\mathcal{H}$ is obviously a mixed graph, we only need to check that $(a1) - (a3)$ in Definition 3.1.1 hold, and that there is no inducing path between two non-adjacent vertices. The argument is very similar to the one we saw in the previous lemma, so we will only highlight the strategy.

There is no directed cycle in $\mathcal{H}$. Otherwise let c be a shortest one. The corresponding cycle in $\mathcal{P}^{tag}_{FCI}$ must contain $\rightarrow \circ\!\!-\!\!\circ$, because there is no directed cycle in $\mathcal{P}^{tag}_{FCI}$ (Lemma 3.3.3) and by assumption $\mathbf{C}(\mathcal{P}^{tag}_{FCI})$ is oriented into a DAG. Then $\mathbf{CP}1$ of $\mathcal{P}^{tag}_{FCI}$ implies that there is a even shorter directed cycle in $\mathcal{H}$, a contradiction.

For almost the same reason, there is no almost directed cycle in $\mathcal{H}$.

For any $X - Y$ in $\mathcal{H}$, it is also in $\mathcal{P}^{tag}_{FCI}$, because no new undirected edge is created in $\mathcal{H}$. We have shown in Lemma 3.3.3 that there is no edge into X or Y in $\mathcal{P}^{tag}_{FCI}$, and that there is no $\circ\!\!-\!\!\circ$ edge incident to X or Y in $\mathcal{P}^{tag}_{FCI}$. It obviously follows that there is no edge into X or Y in $\mathcal{H}$.

To show that $\mathcal{H}$ is maximal, we will again use the rank of an inducing path as defined in the proof of the previous lemma. Suppose for contradiction that there is an inducing path between two non-adjacent vertices X and Y in $\mathcal{H}$. Consider one that is of the lowest rank, $p = \langle V_0 = X, V_1, \cdots, V_{n-1}, V_n = Y\rangle$. As we have shown in the proof of the previous lemma, every edge on p is bi-directed, which is also in $\mathcal{P}^{tag}_{FCI}$, because no new bi-directed edge is created in $\mathcal{H}$. We shall argue that in $\mathcal{P}^{tag}_{FCI}$ it is also the case that every V_i ($1 \leq i \leq n-1$) is an ancestor of either X or Y, and hence p is also an inducing path in $\mathcal{P}^{tag}_{FCI}$, which contradicts (ii) of Lemma 3.3.3. Here is the argument. For an arbitrary V_i ($1 \leq i \leq n-1$), by supposition, it is an ancestor of X or Y in $\mathcal{H}$. Without loss of generality, suppose it is an ancestor of Y. Let d be a shortest directed path from V_i to Y. Then d must also be a directed path in $\mathcal{P}^{tag}_{FCI}$. Suppose not, then it contains a $\circ\!\!-\!\!\circ$ in $\mathcal{P}^{tag}_{FCI}$. Furthermore, the first edge must be $V_i \circ\!\!-\!\!\circ V_{i1}$, for otherwise $\rightarrow \circ\!\!-\!\!\circ$ would appear on the path and **CP**1 implies there is a shorter directed path in $\mathcal{H}$. Now by Lemma 3.3.2, we have $V_{i-1} \leftrightarrow V_{i1}$ and $V_{i+1} \leftrightarrow V_{i1}$, which means we can replace V_i with V_{i1} and create an inducing path with a lower rank (because the directed path from V_{i1} to Y is shorter than the one from V_i to Y). Contradiction. So d is also a directed path in $\mathcal{P}^{tag}_{FCI}$, which means V_i is also an ancestor of Y in $\mathcal{P}^{tag}_{FCI}$. That is, p is also an inducing path between X and Y in $\mathcal{P}^{tag}_{FCI}$, which contradicts the previous lemma.

Therefore $\mathcal{H}$ is a maximal ancestral graph.

Now that we have shown $\mathcal{H}$ to be a MAG, we only need to check the conditions in Proposition 3.1.2 to demonstrate its Markov equivalence to $\mathcal{G}_T$. Obviously they have the same adjacencies (because they both have the same adjacencies as $\mathcal{P}_{FCI}$), so ($e1$) holds. For ($e2$), notice that every unshielded collider in $\mathcal{G}_T$ is also in $\mathcal{P}_{FCI}$ – which is guaranteed by $\mathcal{R}0$ – and hence is also in $\mathcal{H}$. Conversely, for any unshielded collider in $\mathcal{H}$, in $\mathcal{P}^{tag}_{FCI}$ the triple is either $*\!\!\rightarrow\leftarrow\!\!*$, or $*\!\!\rightarrow \circ\!\!-\!\!\circ$, or $\circ\!\!-\!\!\circ \circ\!\!-\!\!\circ$. The latter two cases are impossible, because by **CP**1 $*\!\!\rightarrow \circ\!\!-\!\!\circ$ implies that the triple is shielded; and by assumption, the circle component is oriented into a DAG with no unshielded colliders. So it must be the first case. Then the unshielded collider is also in $\mathcal{P}_{FCI}$ (because no arrowhead is introduced in tail augmentation), and hence also in $\mathcal{G}_T$.

Thus if $\mathcal{H}$ and $\mathcal{G}_T$ are not Markov equivalent, it is due to a violation of $(e3)$. That is, there is a path $u = \langle W, ..., X, Y, Z\rangle$ that is discriminating for Y in both graphs, but the triple $\langle X, Y, Z\rangle$ is a collider in one of the graphs but a non-collider in the other. Note that if the triple is a collider in $\mathcal{H}$, then it is easy to deduce from the definition of discriminating path that $X \leftrightarrow Y \leftrightarrow Z$ is in $\mathcal{H}$. But every bi-directed edge in $\mathcal{H}$ is also in $\mathcal{P}_{FCI}$ (because neither tail augmentation nor the further orientation of $\mathcal{P}_{FCI}^{tag}$ creates any new bi-directed edge), so $A \leftrightarrow B \leftrightarrow C$ is also in $\mathcal{G}_T$. Therefore, it can only be the case that $\langle A, B, C\rangle$ is a collider in $\mathcal{G}_T$ and a non-collider in $\mathcal{H}$. We will derive a contradiction from this.

First of all, we argue that if every collider on $u(W, Y) = \langle W, \cdots, X, Y\rangle$ is present in $\mathcal{P}_{FCI}$, then every vertex between W and X (including A) is a parent of Z in $\mathcal{P}_{FCI}$. The argument goes by induction. Let U be the vertex next to W on u. $\langle W, U, Z\rangle$ is an unshielded triple (because by the definition of discriminating path, W and Z are not adjacent). Since by assumption U is a collider on the path, $W {*\!\!\rightarrow} U$ is in $\mathcal{P}_{FCI}$; so the edge between U and Z is either oriented as $U \leftarrow\!\!* Z$ by $\mathcal{R}0$ or $U \rightarrow Z$ by $\mathcal{R}1$. It cannot be the former case, because in $\mathcal{G}_T$ (and in $\mathcal{H}$) we have $U \rightarrow Z$. Hence $U \rightarrow Z$ is in $\mathcal{P}_{FCI}$. Now suppose the first n vertices after W on u are all parents of Z in $\mathcal{P}_{FCI}$. Then the edge between the $n + 1^{st}$ vertex and Z can be oriented by $\mathcal{R}4$. Because it is a parent of Z in $\mathcal{G}$, the edge will be oriented as a directed edge into Z in $\mathcal{P}_{FCI}$. End of induction. Therefore, if every collider on $u(W, Y)$ is present in $\mathcal{P}_{FCI}$, u is also a discriminating path in $\mathcal{P}_{FCI}$, which means the triple $\langle X, Y, Z\rangle$ would be oriented as a collider by $\mathcal{R}4$, as is the case in $\mathcal{G}_T$. Then it would be a collider in $\mathcal{H}$, too, a contradiction.

Thus some collider on $u(W, Y)$ is not present in $\mathcal{P}_{FCI}$. In other words, some arrowheads on the path correspond to circles in $\mathcal{P}_{FCI}$. Note also that only the first and/or the last collider on the path can be absent from $\mathcal{P}_{FCI}$, because bi-directed edges in $\mathcal{H}$, if any, are all in $\mathcal{P}_{FCI}$ as well. Below we consider three cases separately.

Case 1: $u(W, Y)$ only has three vertices, $\langle W, X, Y\rangle$. So in $\mathcal{P}_{FCI}$, it is either (a) $W \circ\!\!-\!\!\circ X \leftarrow\!\!* Y$, or (b) $W *\!\!\rightarrow X \circ\!\!-\!\!\circ Y$, or (c) $W \circ\!\!-\!\!\circ X \circ\!\!-\!\!\circ Y$ (because extra arrowheads are only introduced in the orientation of the circle component of $\mathcal{P}^{tag}_{FCI}$). In (a) and (b), by **CP**1, W and Y are adjacent. In (c), because $\langle W, X, Y\rangle$ is oriented as a collider in $\mathcal{H}$, and by assumption no unshielded collider is introduced in the orientation of $\mathbf{C}(\mathcal{P}^{tag}_{FCI})$, W and Y must also be adjacent ($W \circ\!\!-\!\!\circ Y$). So $\langle W, Y, Z\rangle$ is an unshielded triple, in any case. Since $\langle X, Y, Z\rangle$ is a non-collider in $\mathcal{H}$, it must be that $Y \rightarrow Z$, which can be easily deduced from the definition of discriminating path. Thus $\langle W, Y, Z\rangle$ is an unshielded non-collider in $\mathcal{H}$. We already showed that $\mathcal{G}_T$ and $\mathcal{H}$ have the same unshielded colliders, so $\langle W, Y, Z\rangle$ is also an unshielded non-collider in $\mathcal{G}_T$. Furthermore, since $\langle X, Y, Z\rangle$ is a collider in $\mathcal{G}_T$, we have $X \leftrightarrow Y \leftrightarrow Z$ in $\mathcal{G}_T$, and hence the edge between W and Y must be $W \leftarrow Y$ (to avoid collider). But then the path u is an inducing path between W and Z in $\mathcal{G}_T$, which contradicts the fact that $\mathcal{G}$ is maximal.

Case 2: $u(W, Y)$ has four vertices, $\langle W, U, X, Y\rangle$. So in $\mathcal{P}_{FCI}$, it is either (a) $W \circ\!\!-\!\!\circ U \leftrightarrow X \leftarrow\!\!* Y$, or (b) $W *\!\!\rightarrow U \leftrightarrow X \circ\!\!-\!\!\circ Y$, or (c) $W \circ\!\!-\!\!\circ U \leftrightarrow X \circ\!\!-\!\!\circ Y$. In (c), it is easy to deduce from Lemma 3.3.2 that $W \leftrightarrow Y$ is in $\mathcal{P}_{FCI}$, and hence is in both $\mathcal{G}_T$ and $\mathcal{H}$. But then the triple $\langle W, Y, Z\rangle$ is an unshielded collider in $\mathcal{G}_T$ but not in $\mathcal{H}$, contrary to what we already showed. In (a), Lemma 3.3.2 implies that $W \leftrightarrow X$. Hence $X \rightarrow Z$ is also present in $\mathcal{P}_{FCI}$. So the path $\langle W, X, Y, Z\rangle$ is a discriminating path for Y in $\mathcal{P}_{FCI}$, which means, by the argument in *Case 1*, the edge between Y

and Z should be the same in $\mathcal{G}_T$ and $\mathcal{H}$, contrary to the assumption. In (b), Lemma 3.3.2 implies that $U \leftrightarrow Y$. For the same reason as in (a), the path $\langle W, U, Y, Z \rangle$ is a discriminating path for Y in $\mathcal{P}_{FCI}$, hence, by the argument in *Case 1*, the edge between Y and Z should be the same in $\mathcal{G}_T$ and $\mathcal{H}$, contrary to the assumption.

Case 3: $u(W, Y)$ has more than four vertices, $\langle W, U, V_1, ..., V_2, X, Y \rangle$ (V_1 and V_2 could be the same vertex). Again there are three cases: (a) $W \circ\!\!-\!\!\circ U \leftrightarrow V_1 \cdots V_2 \leftrightarrow X \leftarrow\!\!* Y$, or (b) $W *\!\!\rightarrow U \leftrightarrow V_1 \cdots V_2 \leftrightarrow X \circ\!\!-\!\!\circ Y$, or (c) $W \circ\!\!-\!\!\circ U \leftrightarrow V_1 \cdots V_2 \leftrightarrow X \circ\!\!-\!\!\circ Y$. In any of the three cases, by essentially the same argument as we saw in *Case 2* (or more rigorously, an inductive argument with *Case 2* as the base case), there would be a discriminating path in $\mathcal{P}_{FCI}$ for Y that ends at Z, so the edge between Y and Z should be the same in $\mathcal{G}_T$ and $\mathcal{H}$, contrary to the assumption.

Therefore, the initial supposition of non-equivalence is false. $\mathcal{H}$ and $\mathcal{G}_T$ are Markov equivalent. $\square$

Chapter 4

Inference of 'Cause': Tail Completeness

The FCI algorithm is not only complete with respect to inferring invariant arrowheads, it also produces a PAG that represents a maximal set of conditional independence relations. That is, even if some circles hide invariant tails, replacing any number of them in the resulting PAG with tails will not introduce more m-separation relations than there already are. This observation implies that no more use can we make of facts about conditional independence and dependence. Still, the algorithm is not yet complete with respect to tails. Just as is the case with $\mathcal{R}1 - \mathcal{R}3$, there are more tail inference rules that follow from acyclicity and ancestral-ness combined with the orientation information already in the PAG. Extra tails resulting from such rules may give us more qualitative causal information about what causes what. In this chapter we present additional orientation inference rules to make the inference system fully complete.

The rest of the chapter is organized as follows. Section 4.1 introduces and explains a few more orientation rules that introduce tails. We demonstrate the soundness of these rules in section 4.2, and then present the long and convoluted argument of completeness in 4.3. In section 4.4, we switch gears a little bit and exploit some results established in 4.3 to prove a transformational relation between Markov equivalent directed MAGs. The transformational result is analogous to the one established by Chickering (1995) for DAGs, which has several interesting applications.

4.1 More Tail Inference Rules

To introduce the tail inference rules, we need a few more graphical notions. Recall that we call any graph that may contain the three kinds of marks a **partial mixed graph (PMG)**.

Definition 4.1.1 (uncovered path)**.** *In a PMG, a path $u = \langle V_0, \cdots, V_n \rangle$ is said to be* **uncovered** *if for every $1 \leq i \leq n-1$, V_{i-1} and V_{i+1} are not adjacent, i.e., every consecutive triple on the path is unshielded.*

Definition 4.1.2 (potentially directed path)**.** *In a PMG, a path $u = \langle V_0, \cdots, V_n \rangle$ is said to be* **potentially directed** *(abbreviated as* **p.d.***) from V_0 to V_n if for every $0 \leq i \leq n-1$, the edge between V_i and V_{i+1} is not into V_i, nor is it out of V_{i+1}.*

Potentially directed paths are already mentioned in the end of last chapter. Intuitively, a p.d. path is one that could be oriented into a directed path by changing the circles on the path into appropriate tails or arrowheads. As we shall see, uncovered p.d. paths play an important role in locating invariant tails. A special case of a p.d. path is also mentioned earlier, where every edge is of the form $\circ\!\!-\!\!\circ$. Recall that we call such a path, a path that consists solely of $\circ\!\!-\!\!\circ$ edges, a **circle path**.

We will present the rules in two blocks to highlight certain modularity. Here is the first block:

$\mathcal{R}5$ For every (remaining) $\alpha \circ\!\!-\!\!\circ \beta$, if there is a path $u = \langle \alpha, \gamma, \cdots, \theta, \beta \rangle$ that is an uncovered circle path between α and β s.t. α, θ are not adjacent and β, γ are not adjacent, then orient $\alpha \circ\!\!-\!\!\circ \beta$ and every edge on u as undirected (—).

$\mathcal{R}6$ If $\alpha - \beta \circ\!\!-\!\!* \gamma$ (α and γ may or may not be adjacent), then orient $\beta \circ\!\!-\!\!* \gamma$ as $\beta -\!\!* \gamma$.

$\mathcal{R}7$ If $\alpha -\!\!\circ \beta \circ\!\!-\!\!* \gamma$, and α, γ are not adjacent, then orient $\beta \circ\!\!-\!\!* \gamma$ as $\beta -\!\!* \gamma$.

The pictorial illustrations of $\mathcal{R}5$–$\mathcal{R}7$ are given in Figure 4.1. These rules are obviously related to undirected edges. $\mathcal{R}5$ explicitly lead to undirected edges, and $\mathcal{R}6$ explicitly depend on undirected edges. So if it is known beforehand that there is no undirected edge in the true causal MAG, the two rules are not necessary. In that case, moreover, $\mathcal{R}7$ will not get triggered at all, because neither $\mathcal{R}0 - \mathcal{R}5$ introduced earlier nor $\mathcal{R}8 - \mathcal{R}10$ to be introduced shortly can lead to $-\!\!\circ$ edges, which are in the antecedent of $\mathcal{R}7$.

The quantifier in $\mathcal{R}5$ indicates that it will be executed once and for all (before $\mathcal{R}6$ and $\mathcal{R}7$), as is the case with $\mathcal{R}0$. It will become clear from the proof of soundness that $\mathcal{R}5 - \mathcal{R}7$ are primarily motivated by the third condition in the definition of ancestral graphs ((a3) in Definition 3.1.1), namely, the restriction on the endpoints of undirected edges.

Therefore, only when the presence of selection effects is an issue do we need to include $\mathcal{R}5 - \mathcal{R}7$. For dealing with causal insufficiency alone or learning directed MAGs, we need only $\mathcal{R}0 - \mathcal{R}4$ plus the following rules:

$\mathcal{R}8$ If $\alpha \to \beta \to \gamma$ or $\alpha -\!\!\circ \beta \to \gamma$, and $\alpha \circ\!\!\to \gamma$, orient $\alpha \circ\!\!\to \gamma$ as $\alpha \to \gamma$.

$\mathcal{R}9$ If $\alpha \circ\!\!\to \gamma$, and $u = \langle \alpha, \beta, \theta, \cdots, \gamma \rangle$ is an uncovered p.d. path from α to γ such that γ and β are not adjacent, then orient $\alpha \circ\!\!\to \gamma$ as $\alpha \to \gamma$.

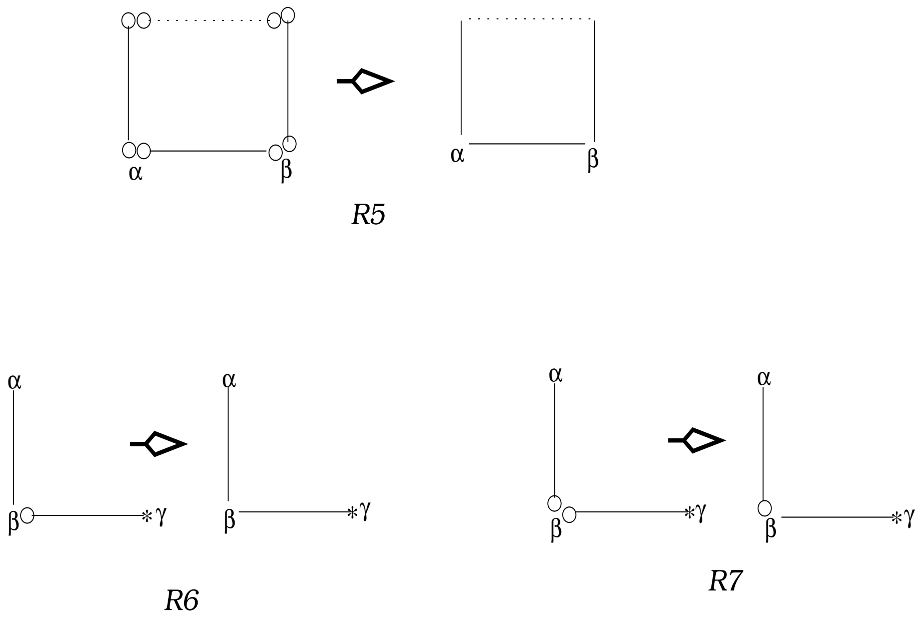

Figure 4.1: Graphical illustrations of $\mathcal{R}5 - \mathcal{R}7$

$\mathcal{R}10$ Suppose $\alpha\circ\!\!\rightarrow \gamma$, $\beta \rightarrow \gamma \leftarrow \theta$, u_1 is an uncovered p.d. path from α to β, and u_2 is an uncovered p.d. path from α to θ. Let μ be the vertex adjacent to α on u_1 (μ could be β), and ω be the vertex adjacent to α on u_2 (ω could be θ). If μ and ω are distinct, and are not adjacent, then orient $\alpha\circ\!\!\rightarrow \gamma$ as $\alpha \rightarrow \gamma$.

These rules are visualized in Figure 4.2. All of them are about turning partially directed edges $\circ\!\!\rightarrow$ into directed ones $\rightarrow$. Such additional orientations are particularly informative when there is no selection effect, as in that situation a directed edge, say, $A \rightarrow B$ means that A is a cause of B whereas $A\circ\!\!\rightarrow B$ only says that B is not a cause of A.

There are obviously cases where some of the additional rules are applicable, and so the FCI algorithm in its present form is not complete with respect to tails. It is not hard to show, for every rule except for $\mathcal{R}8$, that there are cases where that rule alone is applicable. So all these rules except possibly $\mathcal{R}8$ are independent of each other. It

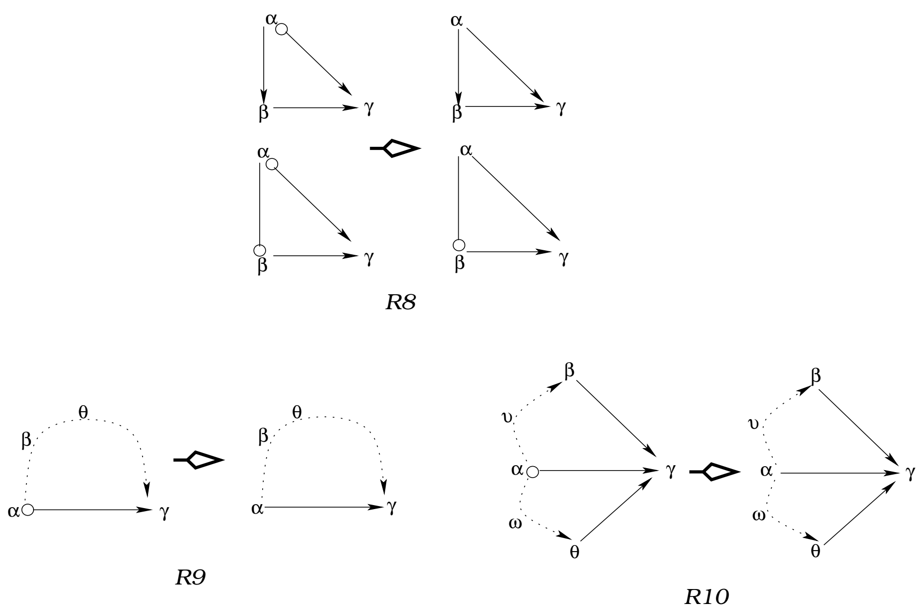

Figure 4.2: Graphical illustrations of $\mathcal{R}8 - \mathcal{R}10$

is not yet clear whether $\mathcal{R}8$ can be derived from other rules.[1] It is noticeable that $\mathcal{R}5$, $\mathcal{R}9$ and $\mathcal{R}10$, just as $\mathcal{R}4$ in the FCI algorithm, involve checking special paths, and as such are computationally more expensive than other rules.[2]

We present $\mathcal{R}8-\mathcal{R}10$ after $\mathcal{R}5-\mathcal{R}7$ because it is clear that any firing of $\mathcal{R}8-\mathcal{R}10$ will not trigger any extra firing of $\mathcal{R}5 - \mathcal{R}7$. Moreover, as will become clear later, $\mathcal{R}5 - \mathcal{R}7$ are the necessary steps in transforming a PAG into a MAG in which all bi-directed edges and undirected edges are invariant. In other words, if we would like to turn the FCI output PAG into a representative MAG with a minimum number of

[1] One thing we do know is that in principle (i.e., given a perfect oracle of conditional independence) $\mathcal{R}8$ is not needed if only $\mathcal{R}0-\mathcal{R}4$ have been fired. In other words, given a perfect oracle of conditional independence, just applying $\mathcal{R}0 - \mathcal{R}4$ will not create an occasion where $\mathcal{R}8$ alone is applicable. So although $\mathcal{R}8$ is actually included in some version of the FCI algorithm discussed in the previous chapter, in theory it is not needed.

[2] The implementation details shall not concern us in this dissertation, so we simply note that the antecedent of each rule that involves (uncovered) paths, in the worst case, can be checked in $O(mn)$, with m being the number of edges and n being the number of vertices in the graph. More efficient implementation seems possible given a further elaboration of the properties of uncovered p.d. paths.

bi-directed and undirected edges, perhaps for the purpose of fitting and scoring, as in Spirtes et al. (1997), then we need to apply $\mathcal{R}5 - \mathcal{R}7$ but do not need to apply $\mathcal{R}8 - \mathcal{R}10$. On the other hand, since in principle $\mathcal{R}5 - \mathcal{R}7$ are relevant only when undirected edges may be present, they in principle will not be invoked if there is no selection bias.[3]

We call the inference procedure resulting from adding $\mathcal{R}5 - \mathcal{R}10$ to the step F4 of the FCI algorithm the *Augmented FCI (AFCI)* algorithm. Our aim in this chapter is to prove that AFCI is both sound and complete, assuming the CMC and CFC.

4.2 Soundness of the Additional Rules

Soundness, as usual, is not difficult to demonstrate.

Theorem 4.2.1. *The tail inference rules, $\mathcal{R}5 - \mathcal{R}10$, are sound.*

Proof. Denote the (unknown) true causal MAG by $\mathcal{G}_T$. Again, for each of the rules, we show that any mixed graph that violates the rule does not belong to $[\mathcal{G}_T]$.

$\mathcal{R}5$: Note that the antecedent of this rule implies that $\langle \alpha, \gamma, \cdots, \theta, \beta, \alpha \rangle$ forms an uncovered cycle that consists of $\circ\!\!-\!\!\circ$ edges. Suppose a mixed graph, contrary to what the rule requires, has an arrowhead on this cycle. By our argument for the soundness of $\mathcal{R}1$, it should be clear that the cycle must be oriented as a directed cycle to avoid unshielded colliders that are not in $\mathcal{G}_T$. But then the graph is not ancestral.

$\mathcal{R}6$: It is clear that if any graph, contrary to what the rule requires, contains $\alpha \,—\, \beta \leftarrow\!\!* \gamma$, the graph is not ancestral.

[3]We add "in principle" here to caution that this is only true with a prefect conditional independence oracle, which in practice may be approximated well by a sufficiently large sample. If the sample size is small, however, there may be occasions where the firing conditions of $\mathcal{R}5$ and $\mathcal{R}7$ are satisfied even though in theory they should never be invoked.

$\mathcal{R}7$: Suppose a mixed graph, contrary to what the rule requires, has an arrowhead at β on the edge between β and γ. Then either α —$\beta \leftarrow\!\!* \gamma$ is present, in which case the graph is not ancestral; or $\alpha \rightarrow \beta \leftarrow\!\!* \gamma$ is present, in which case the graph contains an unshielded collider that is not in $\mathcal{G}_T$.

$\mathcal{R}8$: This rule is analogous to $\mathcal{R}2$. Obviously if a mixed graph, contrary to what the rule requires, contains $\alpha \leftrightarrow \gamma$, then either an almost directed cycle is present or there is an arrowhead into an undirected edge, and hence the graph is not ancestral.

$\mathcal{R}9$: The same argument for the soundness of $\mathcal{R}5$ applies here. If a mixed graph, contrary to what the rule requires, contains $\alpha \leftrightarrow \gamma$, then the uncovered path u must be a directed path (from α to γ) in a graph, to avoid unshielded colliders which are not present in $\mathcal{G}_T$. But then the graph is not ancestral.

$\mathcal{R}10$: This rule is analogous to $\mathcal{R}3$. The antecedent of the rule implies that the triple $\langle \mu, \alpha, \omega \rangle$ is not a collider in $\mathcal{G}_T$, which means at least one of the two edges involved in the triple is out of α in any MAG equivalent to $\mathcal{G}_T$. Now, suppose a mixed graph, contrary to what the rule requires, contains $\alpha \leftrightarrow \gamma$. Then the edge(s) out of α must be a directed edge for the graph to be ancestral. It follows that either u_1 or u_2 is a directed path in the graph to avoid unshielded colliders which are not in $\mathcal{G}_T$. In either case, α is an ancestor of γ, and hence the graph is not ancestral. □

So, every tail introduced by $\mathcal{R}5 - \mathcal{R}10$ is indeed valid given what are already in the PAG. We now turn to the most difficult part of this dissertation — to show that the output of the AFCI algorithm is indeed the CPAG for the true causal MAG.

4.3 Completeness of the Augmented FCI Algorithm

Let us get straight on what we need to show. Let $\mathcal{P}_{AFCI}$ be the output of the AFCI algorithm (given a perfect oracle of conditional independence). The only difference between $\mathcal{P}_{FCI}$ and $\mathcal{P}_{AFCI}$ is that the latter reveals possibly more invariant tails, but no more arrowheads. It trivially follows, for example, that property **CP**1 also holds of $\mathcal{P}_{AFCI}$. Since $\mathcal{P}_{FCI}$ is already arrowhead complete, we know that for every remaining circle in $\mathcal{P}_{AFCI}$, there is a MAG Markov equivalent to $\mathcal{G}_T$ in which the circle is turned into a tail. What is left to show is that for every remaining circle in $\mathcal{P}_{AFCI}$, there is also a MAG Markov equivalent to $\mathcal{G}_T$ in which the circle is turned into an arrowhead.

Our proof of this fact, unfortunately, will be even more involved than the proof of arrowhead completeness. The difficulty lies roughly in the following subtlety concerning $\circ\!\!\rightarrow$ edges. Recall that in showing arrowhead completeness, we can actually turn all circles on $\circ\!\!\rightarrow$ edges simultaneously into tails, via the operation of tail augmentation. Doing this does not take us outside the Markov equivalence class. By contrast, it is in general not possible to turn all such circles simultaneously into arrowheads without breaking Markov equivalence. This fact constitutes the major obstacle to showing tail completeness, and will be dealt with in a fairly roundabout fashion.

Indeed showing that any circle on $\circ\!\!-$ or $\circ\!\!-\!\!\circ$ edges can be oriented as an arrowhead in some Markov equivalent MAG is of comparable complexity as showing arrowhead completeness. The argument for this part is presented in section 4.3.1. We take a break in 4.3.2, in which we present an important corollary that follows from the results in 4.3.1. We then pick up the really difficult task in 4.3.3 — to show that any circle on $\circ\!\!\rightarrow$ edges can be oriented as an arrowhead. The significance of the completeness result is discussed in section 4.3.4. Proofs of some lemmas are to be found in 4.3.5.

4.3.1 Circles on $\circ\!\!-\!\!-$ and $\circ\!\!-\!\!\circ$ Edges

First, quite naturally, we need an operation analogous to tail augmentation.

Definition 4.3.1 (Arrowhead Augmentation)**.** *Let $\mathcal{H}$ be any partial mixed graph.* **Arrowhead augmentation** *of $\mathcal{H}$ is defined as the following set of operations on $\mathcal{H}$:*

- *change all $\circ\!\!\rightarrow$ edges into directed edges $\rightarrow$;*
- *change all $-\!\!\circ$ edges into directed edges $\rightarrow$.*

The resulting graph is called the **arrowhead augmented graph (AAG)** *of $\mathcal{H}$, denoted by $\mathcal{H}^{aag}$.*

Note that the arrowhead augmentation and the tail augmentation are common in their treatments of $\circ\!\!\rightarrow$ edges. They are distinguished by their treatments of $-\!\!\circ$ edges: the tail augmentation turns the circles into tails, whereas the arrowhead augmentation turns the circles into arrowheads. Furthermore, unlike the tail augmentation, the arrowhead augmentation does not affect any $\circ\!\!-\!\!\circ$ edge.

Let $\mathcal{P}^{aag}_{AFCI}$ be the AAG of $\mathcal{P}_{AFCI}$. We will prove a lemma about $\mathcal{P}^{aag}_{AFCI}$ analogous to Lemma 3.3.3. For that purpose, we need to establish some properties of $\mathcal{P}_{AFCI}$ concerning $-\!\!\circ$ edges. The next lemma establishes a property we call **CP**2.

Lemma 4.3.1. *In $\mathcal{P}_{AFCI}$, the following property holds:*

CP2 *For any two vertices A, B, if $A -\!\!\circ B$, then there is no edge into A or B.*

Proof. Since **CP**1 holds of $\mathcal{P}_{AFCI}$, for any $A -\!\!\circ B$ in $\mathcal{P}_{AFCI}$, if $C{*}\!\!\rightarrow B$ is present, then $C{*}\!\!\rightarrow A$ is also present. So it suffices to prove that for any $A -\!\!\circ B$, there is no edge into A. Let $\mathbf{E} = \{X -\!\!\circ Y \text{ in } \mathcal{P}_{AFCI} | \exists Z \text{ s.t. } Z{*}\!\!\rightarrow X \text{ is in } \mathcal{P}_{AFCI}\}$. We need to show that $\mathbf{E}$ is empty. Suppose for contradiction that it is not empty. Let

$X_0 \multimap Y_0 \in \mathbf{E}$ be the member of $\mathbf{E}$ that gets oriented first in the orientation process, that is, the tail marks on other edges in $\mathbf{E}$, if any, get oriented after $X_0 \circ\!\!-\!\!\circ Y_0$ is oriented as $X_0 \multimap Y_0$. Choose any Z_0 such that $Z_0 *\!\!\rightarrow X_0$ is in $\mathcal{P}_{AFCI}$. Since $X_0 \circ\!\!-\!\!\circ Y_0$ is oriented as $X_0 \multimap Y_0$ either by $\mathcal{R}6$ or $\mathcal{R}7$, we consider the two cases one by one:

Case 1: It is oriented by $\mathcal{R}6$. That means there is a vertex W such that $W - X_0$ is in $\mathcal{P}_{AFCI}$. But then $Z_0 *\!\!\rightarrow X_0 - W$ violates (a3) in the definition of ancestral graphs, which contradicts the soundness of $\mathcal{P}_{AFCI}$.

Case 2: It is oriented by $\mathcal{R}7$. That means, at the time of the orientation, there is a vertex W such that W, Y_0 are not adjacent, and there is an edge $W \multimap X_0$ between them. This implies that either $W \multimap X_0$ or $W - X_0$ appears in $\mathcal{P}_{AFCI}$ (as no arrowhead is added by any of $\mathcal{R}5 - \mathcal{R}10$). The latter case is again ruled out by (a3) in the definition of ancestral graphs. In the former case, since $Z_0 *\!\!\rightarrow X_0$ is in $\mathcal{P}_{AFCI}$, by **CP**1, $Z_0 *\!\!\rightarrow W$ is in $\mathcal{P}_{AFCI}$, too. But then $W \multimap X_0$ is in $\mathbf{E}$ and gets oriented before $X_0 \multimap Y_0$ does, which contradicts our choice of $X_0 \multimap Y_0$.

Hence the supposition that $\mathbf{E}$ is not empty is false. **CP**2 holds of $\mathcal{P}_{AFCI}$. $\square$

A path $\langle V_0, \cdots, V_n \rangle$ is called a **tail-circle path** from V_0 to V_n if for every i $(0 \leq i \leq n-1)$, the edge between V_i and V_{i+1} is $V_i \multimap V_{i+1}$.

Lemma 4.3.2. *In $\mathcal{P}_{AFCI}$, the following hold:*

(i) For any $A \multimap B$, there is an uncovered tail-circle path from an endpoint of an undirected edge to B that ends with the edge $A \multimap B$.

(ii) If u is an uncovered tail-circle path, then any two non-consecutive vertices on u are not adjacent.

Proof. Let **TC** be the set of $\multimap$ edges in $\mathcal{P}_{AFCI}$. We order the members of **TC** by their order of occurrence in the orientation process. (i) can be proved by induction.

Let $X \multimap Y$ be the "first" edge in **TC** – that is, it gets oriented as such before any other member of **TC** does (i.e., the others were still $\circ\!\!-\!\!\circ$ edges). Among all the orientation rules, only $\mathcal{R}6$ and $\mathcal{R}7$ could yield $\multimap$ edges. If $X \multimap Y$ is oriented by $\mathcal{R}6$, then obviously X is an endpoint of an undirected edge; if $X \multimap Y$ is oriented by $\mathcal{R}7$, which means there is a vertex Z such that Z, Y are not adjacent, and $Z \multimap X \circ\!\!-\!\!\circ Y$ is the configuration at the point of orienting $X \circ\!\!-\!\!\circ Y$. If $Z \multimap X$ remains in $\mathcal{P}_{AFCI}$, then it belongs to **TC**, and it occurs earlier than $X \multimap Y$ does, which contradicts our choice of $X \multimap Y$. So in $\mathcal{P}_{AFCI}$ it must be $Z - X$ (because no orientation rule will orient $\multimap$ into $\rightarrow$). Hence in either case X is an endpoint of an undirected edge. Then $X \multimap Y$ is an uncovered tail-circle path from an endpoint of an undirected edge to Y.

Now we show the inductive step. Suppose the first n edges in **TC** satisfy (i); consider the $n+1^{st}$ edge, $U \multimap W$, in **TC**. Again, it is oriented by $\mathcal{R}6$ or $\mathcal{R}7$. If it is oriented by $\mathcal{R}6$, then U is an endpoint of an undirected edge, and $U \multimap W$ constitutes an uncovered tail-circle path from U to W; if it is oriented by $\mathcal{R}7$, then there is a vertex V such that V, W are not adjacent, and $V \multimap U \circ\!\!-\!\!\circ W$ is the configuration at the point of orienting $X \circ\!\!-\!\!\circ Y$. If $V \multimap U$ remains in $\mathcal{P}_{AFCI}$, then it is one of the first n edges in **TC**. By the inductive hypothesis, there is an uncovered tail-circle path, T, from an endpoint of an undirected edge to U that includes the edge $V \multimap U$. Since V, W are not adjacent, T appended to $U \multimap W$ constitutes an uncovered tail-circle path from an endpoint of an undirected edge to W. If, on the other hand, $V \multimap U$ is not in $\mathcal{P}_{AFCI}$, then it must be $V - U$, which makes U an endpoint of an undirected edge, and $U \multimap W$ the desired path. Therefore, for every edge in **TC**, the property stated in (i) holds.

Next we prove (ii). If u has only one edge, the proposition trivially holds, because there is no pair of non-consecutive vertices; if u has two edges, the proposition also trivially holds, because u is uncovered, and the only pair of non-consecutive vertices on u are by definition non-adjacent.

Now suppose the proposition holds for those uncovered circle-tail paths that have fewer than n edges. Consider an uncovered circle-tail path with n edges: $V_0 \multimap V_1 \cdots V_{n-1} \multimap V_n$. By the inductive hypothesis, the only pair of non-consecutive vertices that could be adjacent is V_0 and V_n. By **CP**2 (Lemma 4.3.1), the edge between V_0 and V_n is not into V_0 or V_n. It is not an undirected edge either, for otherwise the circle at V_n on $V_{n-1} \multimap V_n$ should have been oriented by $\mathcal{R}6$. However, $\langle V_0, V_1, \cdots, V_{n-1}, V_n, V_0 \rangle$ forms an uncovered cycle, so at least one of the $\circ\!\!-\!\!\circ$ edges on the cycle should have been oriented as $-$ by $\mathcal{R}5$ before any $\multimap$ edge appears, which contradicts the fact that there is no $-$ edge on the cycle. So V_0 and V_n are not adjacent. $\square$

The main use we make of Lemma 4.3.2 is to establish two properties of $\mathcal{P}_{AFCI}$ we call **CP**3 and **CP**4, respectively.

Lemma 4.3.3. *In $\mathcal{P}_{AFCI}$, the following property holds:*

CP3 *For any three vertices A, B, C, if $A \multimap B \circ\!\!-\!\!* C$, then A and C are adjacent. Furthermore, if $A \multimap B \circ\!\!-\!\!\circ C$, then $A \multimap C$; if $A \multimap B \circ\!\!\rightarrow C$, then $A \rightarrow C$ or $A \circ\!\!\rightarrow C$.*

Proof. The first claim is obvious. If $A \multimap B \circ\!\!-\!\!* C$, but A, C are not adjacent, then the circle at B on $B \circ\!\!-\!\!* C$ should have been oriented by $\mathcal{R}7$.

Suppose, more specifically, that $A \multimap B \circ\!\!-\!\!\circ C$. Consider the edge between A and C. Lemma 3.3.1 implies that it is not into C. Lemma 4.3.1 implies that it is not into A. It is not undirected either, for otherwise the circle at C on $B \circ\!\!-\!\!\circ C$ could be oriented by $\mathcal{R}6$. Hence it is either (1) $A \circ\!\!- C$; or (2) $A \circ\!\!-\!\!\circ C$; or (3) $A -\!\!\circ C$. We now show that (1) and (2) are impossible.

Suppose for contradiction that (1) or (2) is the case. By (i) in Lemma 4.3.2, there is an uncovered tail-circle path u from E, an endpoint of an undirected edge, to B that includes the edge $A -\!\!\circ B$. We claim that for every vertex V on u, either $V \circ\!\!-\!\!\circ C$ or $V \circ\!\!- C$ is present. The argument goes by induction. Obviously B and A satisfy the claim. Suppose, starting from B, the n'th vertex on u, V_n, satisfies the claim. Consider the $n+1$'st vertex on u, V_{n+1}. Since u is a tail-circle path, we have $V_{n+1} -\!\!\circ V_n$. By the inductive hypothesis, $V_n \circ\!\!-\!\!\circ C$ or $V_n \circ\!\!- C$. So, as we have established, V_{n+1} and C must be adjacent. Again, Lemma 3.3.1 implies that the edge between them is not into C. Lemma 4.3.1 implies that the edge between them is not into V_{n+1}. The edge is not undirected either, for otherwise the circle at C on $B \circ\!\!-\!\!\circ C$ could be oriented by $\mathcal{R}6$. Furthermore, by (ii) in Lemma 4.3.2, V_{n+1} and B are not adjacent. So the edge between V_{n+1} and C can't be $V_{n+1} -\!\!\circ C$, for otherwise the circle at C on $C \circ\!\!-\!\!\circ B$ could be oriented by $\mathcal{R}7$. It follows that either $V_{n+1} \circ\!\!-\!\!\circ C$ or $V_{n+1} \circ\!\!- C$. Therefore, every vertex on u, in particular the endpoint E satisfies the claim, i.e., that either $E \circ\!\!-\!\!\circ C$ or $E \circ\!\!- C$ occurs. But E is an endpoint of an undirected edge, and hence the circle at E on $E \circ\!\!-\!\!\circ C$ or $E \circ\!\!- C$ could be oriented. Contradiction.

So neither (1) nor (2) is the case, which means $A -\!\!\circ C$ occurs in $\mathcal{P}_{AFCI}$.

On the other hand, if it is $A -\!\!\circ B \circ\!\!\rightarrow C$ that occurs in $\mathcal{P}_{AFCI}$, then Lemma 4.3.1 implies that the edge between A and C has an arrowhead at C (due to the arrowhead on $B \circ\!\!\rightarrow C$), and that there is no arrowhead at A (due to the presence of $A -\!\!\circ B$). So it is either $A \rightarrow C$ or $A \circ\!\!\rightarrow C$. $\square$

Lemma 4.3.4. *In $\mathcal{P}_{AFCI}$, the following property holds:*

CP4 *For any $A \multimap B$, there is no tail-circle path from B to A. That is, there is no such cycle as $A \multimap B \multimap C \multimap \cdots \multimap A$.*

Proof. We first argue that if there is any such cycle in $\mathcal{P}_{AFCI}$, then there is a cycle with only three edges, i.e., $A \multimap B \multimap C \multimap A$. To show this, note that for any such cycle $c = \langle V_0, V_1, V_2, \cdots, V_n, V_0 \rangle$ with more than three edges, c can't be uncovered, otherwise every edge on c would have been oriented as — by $\mathcal{R}5$. That means there is a consecutive triple on c which is shielded. Without loss of generality, suppose $\langle V_0, V_1, V_2 \rangle$ is shielded, i.e., V_0 and V_2 are adjacent. The edge between V_0 and V_2 can't contain an arrowhead, as Lemma 4.3.1 shows; it can't be undirected, for otherwise some circle on c should been oriented by $\mathcal{R}6$; it can't be $\circ\!\!-\!\!\circ$, as implied by Lemma 4.3.3 (because $V_0 \multimap V_1 \multimap V_2$ is present). So it is either $V_0 \multimap V_2$ or $V_2 \multimap V_0$. In either case, there is a shorter cycle than c that consists of $\multimap$ edges. Hence we have established that for any such cycle with more than three edges, there is a shorter one. It follows that if there is such a cycle at all, there must be one with only three edges.

So, to prove **CP**4, it suffices to show that $A \multimap B \multimap C \multimap A$ is impossible. Suppose for contradiction that $A \multimap B \multimap C \multimap A$ appears in $\mathcal{P}_{AFCI}$. By (i) in Lemma 4.3.2, there is an uncovered tail-circle path u from E, an endpoint of an undirected edge, to B that includes the edge $A \multimap B$. We claim that for every vertex V on u between A and E (including A and E), $C \multimap V$ is present in $\mathcal{P}_{AFCI}$. The argument is by induction. The vertex A, by supposition, satisfies the claim. Suppose, starting from A, the n'th vertex on u, V_n, satisfies the claim. Consider the $n+1^{st}$ vertex on u, V_{n+1}. Since u is a tail-circle path, we have $V_{n+1} \multimap V_n$. By the inductive hypothesis, $C \multimap V_n$. So by Lemma 4.3.3, V_{n+1} and C are adjacent. Lemma 4.3.1 implies that the edge between them is not into either vertex. The edge is not undirected either, for otherwise the circle at C on B —∘ C could be oriented by $\mathcal{R}6$. Furthermore, by (ii) in Lemma 4.3.2, V_{n+1} and B are not adjacent. Since B —∘ C, the edge between V_{n+1} and C must be oriented as C —∘ V_{n+1}. Therefore, every vertex between A and E, in particular the endpoint E, satisfies the claim. But E is an endpoint of an undirected edge, and hence the circle at E on C —∘ E could be oriented. This is a contradiction. □

We are now ready to prove a lemma about the arrowhead augmented graph of $\mathcal{P}_{AFCI}$, $\mathcal{P}^{aag}_{AFCI}$, which is analogous to Lemma 3.3.3, the lemma about the tail augmented graph of $\mathcal{P}_{FCI}$. One obvious fact is that $\mathcal{P}^{tag}_{AFCI} = \mathcal{P}^{tag}_{FCI}$ – i.e., the tail augmented graph of $\mathcal{P}_{AFCI}$ is the same as the tail augmented graph of $\mathcal{P}_{FCI}$, which we will use freely below.

Lemma 4.3.5. *Let $\mathcal{P}^{aag}_{AFCI}$ be the arrowhead augmented graph of $\mathcal{P}_{AFCI}$. In $\mathcal{P}^{aag}_{AFCI}$*

(i) (a1)-(a3) (in Definition 3.1.1) and **CP**1 *hold;*

(ii) there is no inducing path between two non-adjacent vertices;

(iii) there is no such triple as A — B ∘—∘C; and

(iv) every unshielded collider in $\mathcal{P}^{aag}_{AFCI}$ is also in $\mathcal{P}_{AFCI}$, i.e., arrowhead augmentation does not create any new unshielded colliders.

Proof. We first demonstrate (i). For (a1), suppose for contradiction that there is a directed cycle in $\mathcal{P}^{aag}_{AFCI}$. Since there is no directed cycle in $\mathcal{P}^{tag}_{FCI}$ or $\mathcal{P}^{tag}_{AFCI}$, as proved in Lemma 3.3.3, at least one edge in the cycle must correspond to a —∘ edge in $\mathcal{P}_{AFCI}$ (because the treatment of ∘→ edges is the same in both tail augmentation and arrowhead augmentation). On the other hand, not all edges in the cycle correspond to —∘ edges in $\mathcal{P}_{AFCI}$, as implied by **CP**4 (Lemma 4.3.4). This means that at least one arrowhead in the cycle is already present in $\mathcal{P}_{AFCI}$. It follows that there will be an arrowhead meeting a $\multimap$ edge in $\mathcal{P}_{AFCI}$, which contradicts **CP**2 (Lemma 4.3.1). So there is no directed cycle in $\mathcal{P}_{AFCI}^{aag}$.

For (a2), suppose for contradiction that there is an almost directed cycle in $\mathcal{P}_{AFCI}^{aag}$. Again, one edge therein must correspond to a $\multimap$ edge in $\mathcal{P}_{AFCI}$, since we already showed that there is no almost directed cycle in $\mathcal{P}_{FCI}^{tag}$ or $\mathcal{P}_{AFCI}^{tag}$ (Lemma 3.3.3). Also, because no new bi-directed edge is introduced by the arrowhead augmentation, the bi-directed edge in the cycle is also in $\mathcal{P}_{AFCI}$. Then it is easy to see that there must be an arrowhead meeting a $\multimap$ edge in $\mathcal{P}_{AFCI}$, which contradicts **CP**2. So there is no almost directed cycle in $\mathcal{P}_{AFCI}^{aag}$.

For (a3), note that no new undirected edge is introduced in the arrowhead augmentation, and new arrowheads are introduced only by way of changing $\multimap$ into $\rightarrow$. Since $\mathcal{P}_{AFCI}$ satisfies (a3) by soundness, and no such pattern as $\multimap$ — appears in $\mathcal{P}_{AFCI}$ (for otherwise the circle could be oriented by $\mathcal{R}6$), it obviously follows that (a3) holds of $\mathcal{P}_{AFCI}^{aag}$.

The fact that **CP**1 also holds of $\mathcal{P}_{AFCI}^{aag}$ follows directly from **CP**3 of $\mathcal{P}_{AFCI}$ (Lemma 4.3.3) and the fact that **CP**1 holds of $\mathcal{P}_{FCI}$.

To see (ii) is true, it suffices to note the following: if there is an inducing path in $\mathcal{P}_{AFCI}^{aag}$ between two non-adjacent vertices, the path must consist of bi-directed edges (which follows from (a1) and (a2), as we have gone through in proving Lemma 3.3.3). Every bi-directed edge in $\mathcal{P}_{AFCI}^{aag}$ is also in $\mathcal{P}_{AFCI}$, so every vertex on the inducing path would have an arrowhead into it in $\mathcal{P}_{AFCI}$. It follows that no edge on any directed path from a vertex on the path to one of the endpoints corresponds to a $\multimap$ edge in $\mathcal{P}_{AFCI}$, for otherwise **CP**2 would be violated. So if there is any inducing path in

$\mathcal{P}^{aag}_{AFCI}$ between two non-adjacent vertices, it would also be present in $\mathcal{P}^{tag}_{AFCI}$, which we have shown to be impossible. Therefore (ii) is true in $\mathcal{P}^{aag}_{AFCI}$.

(iii) is obvious given $\mathcal{R}6$ because no new undirected edge is introduced in the arrowhead augmentation.

(iv) follows from **CP**2 and **CP**3 of $\mathcal{P}_{AFCI}$. Specifically, **CP**2 implies that the extra colliders produced by the arrowhead augmentation can only come from such patterns as $-\!\!\circ\ \circ\!\!-$ in $\mathcal{P}_{AFCI}$, but **CP**3 implies that they are shielded. $\square$

Lemma 4.3.5 immediately leads to the following result, analogous to Lemma 3.3.4.

Lemma 4.3.6. *Let $\mathcal{P}^{aag}_{AFCI}$ be the AAG of $\mathcal{P}_{AFCI}$. If we further orient $\mathbf{C}(\mathcal{P}^{aag}_{AFCI})$, the circle component of $\mathcal{P}^{aag}_{AFCI}$, into a DAG with no unshielded colliders, the resulting graph is a MAG and is Markov equivalent to $\mathcal{G}_T$.*

Proof. Let $\mathcal{H}$ be the resulting MAG. Given Lemma 4.3.5, the exact same argument as in Lemma 3.3.4 can be used to to argue that $\mathcal{H}$ is a MAG.

The argument for the Markov equivalence between $\mathcal{H}$ and $\mathcal{G}_T$ is also similar. (iv) in Lemma 4.3.5 (plus the argument in Lemma 3.3.4) ensure that they have the same unshielded colliders. So if they are not Markov equivalent, it must be that there is a path $u = \langle W, ..., X, Y, Z\rangle$ that is discriminating for Y in both graphs, but the triple $\langle X, Y, Z\rangle$ is a collider in one of the graphs but a non-collider in the other. By the same argument as in Lemma 3.3.4, we can show that it must be a collider in $\mathcal{G}_T$ and a non-collider in $\mathcal{H}$.

If none of the edges on $u(W, Y)$ corresponds to a $-\!\!\circ$ edge in $\mathcal{P}_{AFCI}$, obviously the same argument as in 3.3.4 can be applied to derive a contradiction. Suppose some edge on u corresponds to a $-\!\!\circ$ edge in $\mathcal{P}_{AFCI}$. This is either the first or the last edge on $u(W, Y)$, as every other edge in between, if any, is a bi-directed edge in $\mathcal{H}$ (and in $\mathcal{G}_T$) and hence also a bi-directed in $\mathcal{P}_{AFCI}$. But by **CP**2 (Lemma 4.3.1), there is no arrowhead into either endpoint of a $—\!\circ$ edge. It follows that there is no vertex between W and X on u (for otherwise there exists a bi-directed edge on u, which is already in $\mathcal{P}_{AFCI}$, and thus $\mathcal{CP}2$ would be violated). Hence either $W —\!\circ X \circ\!— Y$, or $W \circ\!—\!\circ X \circ\!— Y$, or $W —\!\circ X \circ\!—\!\circ Y$ occurs in $\mathcal{P}_{AFCI}$. In the first two cases, $\langle X, Y, Z\rangle$ are non-colliders in both graphs. In the last case, by **CP**3 (Lemma 4.3.3), we have $W —\!\circ Y$ in $\mathcal{P}_{AFCI}$. The corresponding edge in $\mathcal{G}_T$ must be $W \rightarrow Y$, because $\langle X, Y, Z\rangle$ is a collider in $\mathcal{G}_T$ (and hence it can't be $W — Y$). We also know that $Y \rightarrow Z$ is in $\mathcal{H}$, since $\langle X, Y, Z\rangle$ is a (discriminated) non-collider in $\mathcal{H}$. But then $\langle W, Y, Z\rangle$ is an unshielded collider in $\mathcal{G}_T$ but not in $\mathcal{H}$, contrary to what we already established.

Therefore, $\mathcal{H}$ and $\mathcal{G}_T$ are Markov equivalent. □

Lemma 4.3.7. *The circle component of $\mathcal{P}^{aag}_{AFCI}$, $\mathbf{C}(\mathcal{P}^{aag}_{AFCI})$, is chordal.*

Proof. This is guaranteed by $\mathcal{R}5$. □

Lemma 3.3.6, Lemma 4.3.6, Lemma 4.3.7 together imply that after $\mathcal{R}5 - \mathcal{R}7$ are done, the circles on the $\circ\!—\!\circ$ and $—\!\circ$ edges do not hide any invariant tails. In other words, for any circle on $\circ\!—\!\circ$ or $—\!\circ$, there is a MAG belonging to $[\mathcal{G}_T]$ in which the circle is marked as an arrowhead. So what is left to show is that $\mathcal{R}8 - \mathcal{R}10$ are sufficient to pick up all the invariant tails hidden in the $\circ\!\!\rightarrow$ edges.

4.3.2 An Important Corollary

This last step, as we hinted, is the least obvious, and our proof for it is going to be quite convoluted. Before we delve into that complicated demonstration, we note an important corollary that follows from the foregoing arguments. A major part of the corollary says that every Markov equivalence class of MAGs has a representative with the minimum number of bi-directed edges and undirected edges, or put it differently, a representative whose bi-directed edges and undirected edges are all invariant (and hence appear in every member of the class).

Corollary 4.3.8. *For every MAG $\mathcal{G}$, there is a MAG $\mathcal{H}$ Markov equivalent to $\mathcal{G}$ such that all bi-directed and undirected edges in $\mathcal{H}$ are invariant, and every directed edge in $\mathcal{G}$ is also in $\mathcal{H}$.*

Proof. $\mathcal{G}$ can serve as the input oracle of conditional independence to the AFCI algorithm. Let the output be $\mathcal{P}_G$. If follows from Lemma 4.3.6 that as long as we orient $\mathbf{C}(\mathcal{P}_G^{aag})$, the circle component of $\mathcal{P}_G^{aag}$, into a DAG with no unshielded colliders, we get a MAG Markov equivalent to $\mathcal{G}$ such that all bi-directed and undirected edges therein are invariant, because no additional bi-directed edges or undirected edges are created in the arrowhead augmentation of $\mathcal{P}_G$.

Let $\mathcal{G}^*$ be the subgraph of $\mathcal{G}$ that corresponds to $\mathbf{C}(\mathcal{P}_G^{aag})$. It is easy to see that all directed edges that are in $\mathcal{G}$ but not in $\mathcal{G}^*$ are already contained in $\mathcal{P}_G^{aag}$. Hence, to prove $\mathcal{H}$ as described exists, it suffices to show that $\mathbf{C}(\mathcal{P}_G^{aag})$ can be oriented into a DAG with no unshielded colliders that retains all the directed edges of $\mathcal{G}^*$.

This is not hard to show. Let $\mathcal{G}_u^*$ be the undirected component of $\mathcal{G}^*$. It is chordal, otherwise it would have been oriented by $\mathcal{R}5$ as undirected in $\mathcal{P}_G$, and hence would not be part of $\mathbf{C}(\mathcal{P}_G^{aag})$. So the part of $\mathbf{C}(\mathcal{P}_G^{aag})$ that corresponds to $\mathcal{G}_u^*$ can be oriented into a DAG with no unshielded colliders. Orient it into any such DAG, $\mathcal{D}_1$.

The rest of $\mathbf{C}(\mathcal{P}_G^{aag})$ will be oriented as follows. The ancestor relationship in $\mathcal{G}^*$ naturally induces a partial order over the vertices therein. Since $\mathcal{G}^*$ is ancestral (as it is a subgraph of an ancestral graph), no edge is into the vertices of $\mathcal{G}_u^*$, which implies that no vertex precedes any vertex of $\mathcal{G}_u^*$ in the partial order. Thus we can extend this partial order to a total order such that every vertex of $\mathcal{G}_u^*$ precedes every vertex not in $\mathcal{G}_u^*$. Orient the rest of $\mathbf{C}(\mathcal{P}_G^{aag})$ according to this total order, and we get a DAG $\mathcal{D}_2$. $\mathcal{D}_2$ obviously retains all the directed edges of $\mathcal{G}^*$, as it respects the partial order induced by $\mathcal{G}^*$. So every arrowhead in $\mathcal{D}_2$ is also in $\mathcal{G}^*$, which implies that $\mathcal{D}_2$ does not contain any unshielded collider (for otherwise $\mathcal{G}^*$ would contain unshielded colliders too, which contradicts the fact that it is a counterpart of $\mathbf{C}(\mathcal{P}_G^{aag})$).

Let $\mathcal{D}$ denote the resulting DAG orientation of $\mathbf{C}(\mathcal{P}_G^{aag})$, i.e., the union of $\mathcal{D}_1$ and $\mathcal{D}_2$. This union will not create any unshielded collider, because every edge between a vertex in $\mathcal{D}_1$ and a vertex not in $\mathcal{D}_1$ is out of the former, by our construction of $\mathcal{D}_2$. So $\mathcal{D}$ is the desired DAG orientation of $\mathbf{C}(\mathcal{P}_G^{aag})$ that has no unshielded colliders and retains all the directed edges of $\mathcal{G}^*$. □

This corollary has several potential applications. One is probably related to fitting and scoring MAGs. In fact, the statistical significance of a special case of this Corollary is already explored by Drton and Richardson (2004) in the context of fitting bi-directed Gaussian graphical models. That special case is also useful in proving a transformational property of directed MAGs, analogous to the one for DAGs established by Chickering (1995). We shall present that result in the last section of this chapter. It turns out that a key lemma for proving the transformational result is also useful for our current purpose of proving tail completeness. We thus present it here, which gives sufficient and necessary conditions under which changing a directed edge ($\rightarrow$) into a bi-directed edge ($\leftrightarrow$) in a MAG preserves Markov equivalence.

Lemma 4.3.9. *Let $\mathcal{G}$ be an arbitrary MAG, and $A \rightarrow B$ an arbitrary directed edge in $\mathcal{G}$. Let $\mathcal{G}'$ be the graph identical to $\mathcal{G}$ except that the edge between A and B is $A \leftrightarrow B$. (In other words, $\mathcal{G}'$ is the result of simply changing the mark at A on $A \rightarrow B$ from a tail into an arrowhead.) $\mathcal{G}'$ is a MAG and Markov equivalent to $\mathcal{G}$ if and only if* *(t1) A is not an endpoint of an undirected edge;*

(t2) there is no directed path from A to B other than $A \to B$;

(t3) For any $C \to A$ in $\mathcal{G}$, $C \to B$ is also in $\mathcal{G}$; and for any $D \leftrightarrow A$ in $\mathcal{G}$, either $D \to B$ or $D \leftrightarrow B$ is in $\mathcal{G}$;

(t4) There is no discriminating path for A on which B is the endpoint adjacent to A.

Proof. See section 4.3.5. □

4.3.3 Circles on $\circ\!\!\rightarrow$ Edges

We now turn to the difficult task of showing that for every $\circ\!\!\rightarrow$ edge in $\mathcal{P}_{AFCI}$, there is a MAG equivalent to $\mathcal{G}_T$ in which the edge is oriented as $\leftrightarrow$. Our argument is going to be roundabout, with two major steps. Let $J \circ\!\!\rightarrow K$ be an arbitrary $\circ\!\!\rightarrow$ edge in $\mathcal{P}_{AFCI}$. In the first step, we show that we can orient $\mathbf{C}(\mathcal{P}_{AFCI})$ — the circle component of $\mathcal{P}_{AFCI}$, which is the same as $\mathbf{C}(\mathcal{P}^{aag}_{AFCI})$, the circle component of the AAG of $\mathcal{P}_{AFCI}$ — into a DAG with no unshielded colliders that satisfies certain conditions relative to $J \circ\!\!\rightarrow K$. By Lemma 4.3.6, the arrowhead augmentation together with this DAG orientation of $\mathbf{C}(\mathcal{P}_{AFCI})$ yield a MAG equivalent to $\mathcal{G}_T$. In the second step, we argue that this particular MAG can be transformed into a MAG containing $J \leftrightarrow K$ through a sequence of equivalence-preserving changes of $\to$ to $\leftrightarrow$. It then follows that the resulting MAG with $J \leftrightarrow K$ is also equivalent to $\mathcal{G}_T$, which gives us what we need.

The following definitions specify the conditions we want a DAG orientation of $\mathbf{C}(\mathcal{P}_{AFCI})$ to satisfy.

Definition 4.3.2 (Relevance)**.** *Let $J \circ\!\!\rightarrow K$ be an arbitrary $\circ\!\!\rightarrow$ edge in $\mathcal{P}_{AFCI}$. For any $A \circ\!\!\rightarrow B$ in $\mathcal{P}_{AFCI}$, it is said to be* **relevant** *to $J \circ\!\!\rightarrow K$ if* *(i)* $A = J$ *or there is a p.d. path from* J *to* A *in* $\mathcal{P}_{AFCI}$ *such that no vertex on the path (including the endpoints) is a parent of* K*; and*

(ii) $B = K$ *or* B *is a parent of* K *(namely* $B \to K$*) in* $\mathcal{P}_{AFCI}$*.*

If $A \circ\!\!\rightarrow B$ *is relevant to* $J \circ\!\!\rightarrow K$*, we say that* A *is* **circle-relevant** *to* $J \circ\!\!\rightarrow K$*, and* B *is* **arrowhead-relevant** *to* $J \circ\!\!\rightarrow K$*.*

Informally, relevant edges are those that may have to be changed to bi-directed edges ($\leftrightarrow$) before the edge between J and K can be so oriented. The rationale behind the formal definition above will be revealed by the proof of Lemma 4.3.32. We use $\mathbf{REL}(J \circ\!\!\rightarrow K)$ to denote the set of $\circ\!\!\rightarrow$ edges relevant to $J \circ\!\!\rightarrow K$ in $\mathcal{P}_{AFCI}$. Notice that $J \circ\!\!\rightarrow K$ itself belongs to this set. It will also be convenient to denote the set of circle-relevant vertices by $\mathbf{CR}(J \circ\!\!\rightarrow K)$, and the set of arrow-relevant vertices by $\mathbf{AR}(J \circ\!\!\rightarrow K)$.

Definition 4.3.3. *A DAG orientation of* $\mathbf{C}(\mathcal{P}_{AFCI})$ *– the circle component of* $\mathcal{P}_{AFCI}$ *– is said to be* **agreeable** *to* $J \circ\!\!\rightarrow K$ *if the following three conditions hold:*

$\mathbf{C_1}$ *For any* $A \circ\!\!\rightarrow B \in \mathbf{REL}(J \circ\!\!\rightarrow K)$ *and* $B \circ\!\!-\!\!\circ C$ *in* $\mathcal{P}_{AFCI}$*, if* $C \notin \mathbf{AR}(J \circ\!\!\rightarrow K)$*, then* $B \circ\!\!-\!\!\circ C$ *is oriented as* $B \to C$ *in the DAG;*

$\mathbf{C_2}$ *For any* $A \circ\!\!\rightarrow B \in \mathbf{REL}(J \circ\!\!\rightarrow K)$ *and* $A \circ\!\!-\!\!\circ C$ *in* $\mathcal{P}_{AFCI}$*, if* C *is a parent of* B *(namely* $C \to B$*) in* $\mathcal{P}_{AFCI}$*, then* $A \circ\!\!-\!\!\circ C$ *is oriented as* $A \leftarrow C$ *in the DAG;*

$\mathbf{C_3}$ *For any* $A \circ\!\!\rightarrow B \in \mathbf{REL}(J \circ\!\!\rightarrow K)$ *and* $A \circ\!\!-\!\!\circ C$ *in* $\mathcal{P}_{AFCI}$*, if* C *is not adjacent to* B *in* $\mathcal{P}_{AFCI}$*, then* $A \circ\!\!-\!\!\circ C$ *is oriented as* $A \to C$ *in the DAG.*

Since we will henceforth refer to $\mathbf{C_1} - \mathbf{C_3}$ very frequently, some further explanation of them is in order. Roughly speaking, they are all motivated as necessary for a $\circ\!\!\rightarrow$ edge (relevant to $J \circ\!\!\rightarrow K$) to meet the conditions in Lemma 4.3.9. This is especially clear in $\mathbf{C_2}$ and $\mathbf{C_3}$. Regarding a relevant edge $A\circ\!\!\rightarrow B$ (which will be $A \rightarrow B$ after arrowhead augmentation), violation of $\mathbf{C_2}$ will fail condition (t2), and violation of $\mathbf{C_3}$ will fail condition (t3) in Lemma 4.3.9 for changing $A \rightarrow B$ into $A \leftrightarrow B$. For $\mathbf{C_1}$, notice that if the antecedent holds, we have either $A \rightarrow C$ or $A\circ\!\!\rightarrow C$ in $\mathcal{P}_{AFCI}$, by property **CP**1 (Lemma 3.3.1). In either case, $A \rightarrow C$ will appear in $\mathcal{P}^{aag}_{AFCI}$. So if $\mathbf{C_1}$ is violated, i.e., if $B \circ\!\!-\!\!\circ C$ is oriented as $B \leftarrow C$, then (t2) in Lemma 4.3.9 fails. (It will not matter, however, if $C \in \mathbf{AR}(J\circ\!\!\rightarrow K)$; because in that case, as will become clear later, $A\circ\!\!\rightarrow C$ is in $\mathcal{P}_{AFCI}$. Then $A\circ\!\!\rightarrow C \in \mathbf{REL}(J\circ\!\!\rightarrow K)$, which can be dealt with before $A\circ\!\!\rightarrow B$.)

It is far less obvious, however, that $\mathbf{C_1} - \mathbf{C_3}$ suffice to ensure the existence of a sequence of equivalence-preserving changes that can eventually turn $J \rightarrow K$ into $J \leftrightarrow K$. The demonstration of this fact will be postponed until Lemma 4.3.32. Before that, we need to establish the even less obvious fact that $\mathbf{C}(\mathcal{P}_{AFCI})$ can be oriented into a DAG with no unshielded colliders that satisfies $\mathbf{C_1} - \mathbf{C_3}$ relative to $J\circ\!\!\rightarrow K$.

One way to orient a chordal graph into a DAG free of unshielded colliders is via the Meek orientation rules (Meek 1995):

Meek's Algorithm

Input: a chordal unoriented graph $\mathcal{U}$

Output: a DAG orientation of $\mathcal{U}$ (with no unshielded colliders)

Repeat

1. choose a yet unoriented edge $A \circ\!\!-\!\!\circ B$ in $\mathcal{U}$;

2. orient the edge into $A \rightarrow B$ and close orientations under the following rules:[4]

[4]There is another rule in Meek (1995), which corresponds to S4(c) in the PC algorithm presented in Chapter 2. However, the antecedent will never be met in orienting a chordal graph into a DAG *with no unshielded colliders*. So we need not include that one here.

UR$_1$ If $A \to B \circ\!\!-\!\!\circ C$, A and C are not adjacent, orient as $B \to C$.

UR$_2$ If $A \to B \to C$ and $A \circ\!\!-\!\!\circ C$, orient as $A \to C$.

UR$_3$ If $A \to B \to C$, $A \circ\!\!-\!\!\circ D \circ\!\!-\!\!\circ C$, $B \circ\!\!-\!\!\circ D$, and A and C are not adjacent, orient $D \circ\!\!-\!\!\circ C$ as $D \to C$.

Until every edge is oriented in $\mathcal{H}$.

We now adapt the algorithm to fit our purpose. Given an arbitrary edge $J \circ\!\!\rightarrow K$ in $\mathcal{P}_{AFCI}$, let $\mathbf{E_n}, n = 1, 2, 3$ denote the set of $\circ\!\!-\!\!\circ$ edges whose orientations are required by condition $\mathbf{C_n}$ in Definition 4.3.3. (Note that they are not necessarily disjoint.)

Orientation Algorithm for The Circle Component of $\mathcal{P}_{AFCI}$

Input: $\mathbf{C}(\mathcal{P}_{AFCI})$, $\mathcal{P}_{AFCI}$, and an edge $J \circ\!\!\rightarrow K$ therein

Output: a DAG orientation of $\mathbf{C}(\mathcal{P}_{AFCI})$ with no unshielded colliders

Let $\mathcal{D} = \mathbf{C}(\mathcal{P}_{AFCI})$

Repeat

If some edge in $\mathbf{E_1}$ is yet unoriented in $\mathcal{D}$

(a) choose such an edge $A \circ\!\!-\!\!\circ B \in \mathbf{E_1}$, and orient it as condition $\mathbf{C_1}$ requires;

(b) close orientations under $\mathbf{UR}_1, \mathbf{UR}_2, \mathbf{UR}_3$.

Else If some edge in $\mathbf{E_2}$ is yet unoriented in $\mathcal{D}$;

(a) choose such an edge $A \circ\!\!-\!\!\circ B \in \mathbf{E_2}$, and orient it as condition $\mathbf{C_2}$ requires;

(b) close orientations under $\mathbf{UR}_1, \mathbf{UR}_2, \mathbf{UR}_3$.

Else If some edge in $\mathbf{E_3}$ is yet unoriented in $\mathcal{D}$;

(a) choose such an edge $A \circ\!\!-\!\!\circ B \in \mathbf{E_3}$, and orient it as condition $\mathbf{C_3}$ requires;

(b) close orientations under $\mathbf{UR}_1, \mathbf{UR}_2, \mathbf{UR}_3$.

Else

(a) choose a yet unoriented edge $A \circ\!\!-\!\!\circ B$ in $\mathcal{D}$;

(b) orient the edge into $A \rightarrow B$ and close orientations under $\mathbf{UR}_1, \mathbf{UR}_2, \mathbf{UR}_3$.

Until every edge is oriented in $\mathcal{D}$

Return $\mathcal{D}$

Given the correctness of Meek's algorithm, this Orientation Algorithm obviously returns a DAG orientation of $\mathbf{C}(\mathcal{P}_{AFCI})$ with no unshielded colliders. The question is whether it is also agreeable to $J \circ\!\!\rightarrow K$. We answer this question affirmatively in Corollary 4.3.29, to which we now proceed.

We begin by noting some facts about (uncovered) p.d. paths (see Definition 4.1.2) in $\mathcal{P}_{AFCI}$.

Lemma 4.3.10. *If $u = \langle A, \cdots, B \rangle$ is a p.d. path from A to B in $\mathcal{P}_{AFCI}$, then some subsequence of u forms an uncovered p.d. path from A to B in $\mathcal{P}_{AFCI}$.*

Proof. We prove it by induction on the length of u. If there is only one edge on u, then it is trivially an uncovered p.d. path from A to B. If there are two edges on u, namely $u = \langle A, C, B \rangle$, either it is already uncovered, or it is covered so that A and B are adjacent. In the latter case, we show that the edge between A and B constitutes an uncovered p.d. path from A to B, or in other words, the edge between A and B is not into A or out of B.

We first argue that it is not into A. Suppose for contradiction that the mark at A on the edge between A and B is an arrowhead. Then the edge between A and C can't have a circle mark at A, for otherwise by **CP**1 (Lemma 3.3.1), the edge between C and B has an arrowhead at C, which contradicts the fact that u is potentially directed. It follows that the edge between A and C must have a tail at A in $\mathcal{P}_{AFCI}$. Since the edge between A and B is into A, it follows from **CP**2 (Lemma 4.3.1) that the edge between A and C is $A \rightarrow C$. Then the mark at C on the edge between C and B must be an arrowhead, as implied by $\mathcal{R}2$, a contradiction. So the edge between A and B is not into A.

Next we argue that it is not out of B either. Suppose for contradiction that the mark at B on the edge between A and B is a tail. Then it is either $A - B$ or $A \circ\!\!- B$. The former implies that the edge between C and B has a tail at B by $\mathcal{R}6$, which contradicts the fact that u is potentially directed. So it must be $A \circ\!\!- B$. It obviously follows, by **CP**1 and **CP**2, that there can't be any arrowhead on u, so it is either $A \circ\!\!-\!\!\circ C \circ\!\!-\!\!\circ B$, or $A \circ\!\!-\!\!\circ C -\!\!\circ B$, or $A -\!\!\circ C \circ\!\!-\!\!\circ B$ or $A -\!\!\circ C -\!\!\circ B$. The first three cases contradict **CP**3 (Lemma 4.3.3), and the last case contradicts **CP**4 (Lemma 4.3.4).

The inductive step is very easy. Suppose the proposition holds when the length of u is $n-1$ ($n \geq 3$). Consider the case where u has n edges. Either u is already uncovered, or there is a triple $\langle X, Y, Z \rangle$ on the path which is shielded. In the latter case, by the foregoing argument, the edge between X and Z is not into X or out of Z. So if we replace $\langle X, Y, Z \rangle$ with the edge between X and Z on u, we get a subsequence of u which is a p.d. path from A to B with length $n-1$. By the inductive hypothesis, a subsequence of the new path, which is also a subsequence of u, forms an uncovered p.d. path from A to B. This concludes our argument. $\square$

Lemma 4.3.11. *If u is an uncovered p.d. path from A to B in $\mathcal{P}_{AFCI}$, then*

(i) if there is an $\circ\!\!\rightarrow$ or $-\!\!\circ$ edge on u, then any $\circ\!\!-\!\!\circ$ edge on u is before that edge,

and any $\rightarrow$ edge on u is after that edge;

(ii) u does not include both a $\circ\!\!\rightarrow$ edge and a $—\!\!\circ$ edge; and

(iii) there is at most one $\circ\!\!\rightarrow$ edge on u.

Proof. To see (i) is true, notice that since u is uncovered and potentially directed, any edge after a $\circ\!\!\rightarrow$ edge or a $\rightarrow$ edge on u must be oriented as $\rightarrow$ by $\mathcal{R}1$. So no $\circ\!\!—\!\!\circ$ can appear after a $\circ\!\!\rightarrow$ edge on u, and no $\rightarrow$ can appear before a $\circ\!\!\rightarrow$ edge on u. The same is true with a $—\!\!\circ$ edge. Since u is uncovered, any edge on u after $—\!\!\circ$ will be oriented as $—\!\!\circ$ or $\rightarrow$ by either $\mathcal{R}7$ or $\mathcal{R}1$.

(ii) and (iii) are evident given the argument for (i). For (iii), just note that any edge after a $\circ\!\!\rightarrow$ edge on u must be oriented as a $\rightarrow$ edge. For (ii), suppose for contradiction that u contains both a $\circ\!\!\rightarrow$ edge and a $—\!\!\circ$ edge. Then the $—\!\!\circ$ edge does not appear after the $\circ\!\!\rightarrow$ edge on u, because any edge after $\circ\!\!\rightarrow$ on u must be oriented as $\rightarrow$ by $\mathcal{R}1$. On the other hand, the $\circ\!\!\rightarrow$ does not appear after the $—\!\!\circ$ edge on u, because any edge after $—\!\!\circ$ on u is either $—\!\!\circ$ or $\rightarrow$. This is a contradiction. $\square$

Lemma 4.3.12. *In $\mathcal{P}_{AFCI}$, if there is a p.d. path from A to B, then the edge between A and B, if any, is not into A.*

Proof. By Lemma 4.3.10, there is an uncovered p.d. path u from A to B. Suppose for contradiction that there is an edge between A and B which is into A, namely $A \leftarrow\!\!* B$ is in $\mathcal{P}_{AFCI}$. There can't be a $—\!\!\circ$ edge on u for the following reason: the first $—\!\!\circ$ edge, if any, is either incident to A or is connected to A by a circle path, according to Lemma 4.3.11. In either case, by Lemma 3.3.2, there is an edge into the tail endpoint of the $—\!\!\circ$ edge, which contradicts **CP**2 (Lemma 4.3.1).

So, by Lemma 4.3.11, u is of the form: $\circ\!\!—\!\!\circ \cdots \circ\!\!\rightarrow\rightarrow \cdots \rightarrow$. It takes little effort to see that Lemma 3.3.2 entails that there is an edge between B and an ancestor of

B which is into that ancestor. This contradicts the soundness of $\mathcal{P}_{AFCI}$. $\square$

Lemma 4.3.13. *In $\mathcal{P}_{AFCI}$, if there is a p.d. path from A to B that is into B, then every uncovered p.d. path from A to B is into B.*

Proof. Suppose for contradiction that an uncovered p.d. path from A to B is not into B. That is, the last edge on the path is not $\circ\!\!\rightarrow$ or $\rightarrow$. The last edge can't be $-\!\!\circ$ either, because there is a p.d. path into B. So the last edge must be $\circ\!\!-\!\!\circ$, and hence by Lemma 4.3.11, the path must be a circle path. Let C be the vertex adjacent to B on the p.d. path into B, which means $C{*}\!\!\rightarrow B$. Since there is a circle path between A and B, it follows from Lemma 3.3.2 that $C{*}\!\!\rightarrow A$. But there is a p.d. path from A to C, which contradicts Lemma 4.3.12. $\square$

Corollary 4.3.14. *In $\mathcal{P}_{AFCI}$, if A, B are adjacent, and there is a p.d. path from A to B that is into B, then the edge between A and B is either $A \circ\!\!\rightarrow B$ or $A \rightarrow B$.*

Proof. By Lemma 4.3.12, the edge between A and B is not into A. It follows that it is not out of B, because there is a path into B, which rules out the possibility of $A - B$ or $A \circ\!\!- B$ by Lemma 4.3.1. Hence the edge between A and B is an uncovered p.d. path from A to B. By Lemma 4.3.13, it is into B, which means it is either $A \circ\!\!\rightarrow B$ or $A \rightarrow B$. $\square$

Lemma 4.3.15. *If there is a circle path between two adjacent vertices in $\mathcal{P}_{AFCI}$, then the edge between the two vertices is $\circ\!\!-\!\!\circ$.*

Proof. By Lemma 3.3.2, there is no arrowhead on the edge between the two vertices. The edge obviously can't be $-$. If it is $-\!\!\circ$, then it is easy to derive a contradiction from **CP**3 (Lemma 4.3.3). So the edge between the two vertices must be $\circ\!\!-\!\!\circ$ in $\mathcal{P}_{AFCI}$. $\square$

Lemma 4.3.16. *Let u be an uncovered circle path in $\mathcal{P}_{AFCI}$. If A and B are two non-consecutive vertices on u, then A and B are not adjacent in $\mathcal{P}_{AFCI}$.*

Proof. It follows from Lemma 4.3.15 and the fact that $\mathbf{C}(\mathcal{P}_{AFCI})$ is chordal. $\square$

The next couple of lemmas establish two useful facts for the endpoints of the edges in $\mathbf{REL}(J \circ\!\!\rightarrow K)$.

Lemma 4.3.17. *For any $A \circ\!\!\rightarrow B \in \mathbf{REL}(J \circ\!\!\rightarrow K)$, there is an uncovered p.d. path u from J to B in $\mathcal{P}_{AFCI}$ such that for every vertex V on u other than B, there is an edge $V \circ\!\!\rightarrow K$.*

Proof. Obviously if $A = J$ or $B = K$, the lemma is trivial. Suppose $A \neq j$ and $B \neq K$. By Definition 4.3.2, there is a p.d. path from J to A in $\mathcal{P}_{AFCI}$ such that no vertex on the path (including the endpoints) is a parent of K. B is not on this p.d. path, for otherwise by Lemma 4.3.12 the edge between B and A should be into A. This p.d. path concatenated with $A \circ\!\!\rightarrow B$ constitutes a p.d. path from J to B which is into B. Lemma 4.3.10 implies that there is an uncovered p.d. path u from J to B such that every vertex on u other than B is not a parent of K. This path, by Lemma 4.3.13, is into B. We now argue that for every vertex V on u other than B, there is an edge $V \circ\!\!\rightarrow K$ in $\mathcal{P}_{AFCI}$.

The base case is trivial, as $J \circ\!\!\rightarrow K$ is by assumption in $\mathcal{P}_{AFCI}$. Suppose for the n'th vertex on u, V_n, there is an edge $V_n \circ\!\!\rightarrow K$ in $\mathcal{P}_{AFCI}$. We show that for the $n+1^{st}$ vertex on u, V_{n+1}, if $V_{n+1} \neq B$, then there is an edge $V_{n+1} \circ\!\!\rightarrow K$ in $\mathcal{P}_{AFCI}$.

Suppose $V_{n+1} \neq B$. Note that since $B \neq K$, by Definition 4.3.2, B is a parent of K. Hence the path $u(V_{n+1}, B) \oplus B \rightarrow K$ is an uncovered p.d. path from V_{n+1} to K which is into K (note that K is not on u in view of the edge $B \rightarrow K$ and Lemma 4.3.12). Thus by Corollary 4.3.14 and by the fact that no vertex on u other than

B is a parent of K, we can conclude that if V_{n+1} is adjacent to K, then the edge is $V_{n+1}\circ\!\!\rightarrow K$. Suppose for contradiction that there is no such an edge $V_{n+1}\circ\!\!\rightarrow K$ in $\mathcal{P}_{AFCI}$. If follows that V_{n+1} and K are not adjacent. Let W be the closest vertex to V_{n+1} on $u(V_{n+1}, B)$ that is adjacent to K. (W exists because B is adjacent to K.) Again, by Corollary 4.3.14, the edge between W and K is either $W\circ\!\!\rightarrow K$ or $W \rightarrow K$. So the path $u(V_n, W) \oplus \langle W, K\rangle$ is an uncovered p.d. path from V_n to K (it is uncovered because by our choice of W no vertex between V_n and W is adjacent to K). Together with the fact that V_{n+1} is not adjacent to K, this implies that the circle at V_n on $V_n\circ\!\!\rightarrow K$ could be oriented by $\mathcal{R}9$, which is a contradiction. Therefore there is an edge $V_{n+1}\circ\!\!\rightarrow K$ in $\mathcal{P}_{AFCI}$. $\square$

Lemma 4.3.18. *If $A\circ\!\!\rightarrow B \in$ **REL**$(J\circ\!\!\rightarrow K)$, then there is an edge $A\circ\!\!\rightarrow K$ in $\mathcal{P}_{AFCI}$.*

Proof. If $A = J$ or $B = K$, there is obviously an edge $A\circ\!\!\rightarrow K$ in $\mathcal{P}_{AFCI}$. Suppose $A \neq J$ and $B \neq K$. Since $A\circ\!\!\rightarrow B \in$ **REL**$(J\circ\!\!\rightarrow K)$, by Lemma 4.3.17, there is an uncovered p.d. path u from J to B in $\mathcal{P}_{AFCI}$ such that for every vertex V on u other than B, there is an edge $V\circ\!\!\rightarrow K$. By Lemma 4.3.13, we know that u is also into B. Let X be the vertex adjacent to B on u. We have $X\circ\!\!\rightarrow K$ in $\mathcal{P}_{AFCI}$. Also, because $B \neq K$, $B \rightarrow K$ is in $\mathcal{P}_{AFCI}$, so the edge between X and B can't be $X \rightarrow B$, for otherwise $X\circ\!\!\rightarrow K$ could be oriented by $\mathcal{R}8$. It follows that $X\circ\!\!\rightarrow B$ is in $\mathcal{P}_{AFCI}$, because u is into B.

Now, suppose for contradiction that A is not adjacent to K. Then the path $\langle A, B, X, K\rangle$ is a discriminating path for X (Definition 3.1.7). Hence the circle on $X\circ\!\!\rightarrow K$ could have been oriented by $\mathcal{R}4$, a contradiction. So A is adjacent to K. By Corollary 4.3.14, the edge between A and K is either $A \rightarrow K$ or $A\circ\!\!\rightarrow K$. But by definition (Definition 4.3.2), A is not a parent of K, so it must be $A\circ\!\!\rightarrow K$ in

$\mathcal{P}_{AFCI}$. $\square$

We are now ready to make important steps towards showing that in the course of the *Orientation Algorithm*, no violation of $\mathbf{C_1} - \mathbf{C_3}$ (Definition 4.3.3) would occur, and hence the output DAG orientation of $\mathbf{C}(\mathcal{P}_{AFCI})$ is agreeable to $J \circ\!\!\rightarrow K$. For this purpose, we assume, without loss of generality, that $\mathbf{UR}_1$ has priority over $\mathbf{UR}_2$ and $\mathbf{UR}_3$ in the sense that whenever two or more different rules can be fired, $\mathbf{UR}_1$ will always be applied first, if applicable. The following series of lemmas will amount to showing that if we choose a $\circ\!\!-\!\!\circ$ edge to orient away from violation of $\mathbf{C_1} - \mathbf{C_3}$ (as the *Orientation Algorithm* does), that orientation will not trigger any violation of $\mathbf{C_1} - \mathbf{C_3}$ by applications of $\mathbf{UR}_1$ alone. Notice that the stereotype of a chain of $\mathbf{UR}_1$ firings is that the first edge on an uncovered circle path $\circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ$ is oriented out of the first vertex, which triggers repeated applications of $\mathbf{UR}_1$ that orient the whole circle path. That is why most of the next block of lemmas are concerned with an uncovered circle path.

Lemma 4.3.19. *For any two vertices $B, C \in \mathbf{AR}(J \circ\!\!\rightarrow K)$, there is no uncovered circle path between B and C consisting of more than one edge in $\mathcal{P}_{AFCI}$.*

Proof. If one of B and C is K, it is manifest in the definition of relevance that there is a directed edge between them, and hence there is no circle path between them, as implied by Lemma 3.3.2. So we only need to consider the case where neither of them is K, that is, both of them are parents of K. Suppose for contradiction that in $\mathcal{P}_{AFCI}$ there is an uncovered circle path u between B and C that includes two or more $\circ\!\!-\!\!\circ$ edges. It follows, by Lemma 4.3.16, that B and C are not adjacent. Let A be such a vertex that $A \circ\!\!\rightarrow B \in \mathbf{REL}(J \circ\!\!\rightarrow K)$. It follows from Lemma 3.3.2 that either $A \circ\!\!\rightarrow C$ or $A \rightarrow C$ is in $\mathcal{P}_{AFCI}$. Furthermore, because A is not a parent of K, it must be $A \circ\!\!\rightarrow C$. Now consider the edge $A \circ\!\!\rightarrow K$, which is shown to be present by Lemma

4.3.18. It could be oriented by $\mathcal{R}10$, because $A \circ\!\!\rightarrow B$ is an uncovered p.d. path from A to B, a parent of K; $A \circ\!\!\rightarrow C$ is an uncovered p.d. path from A to C, a parent of K; B and C are not adjacent. Hence a contradiction. $\square$

Lemma 4.3.20. *Suppose $A \circ\!\!\rightarrow B \in \mathbf{REL}(J \circ\!\!\rightarrow K)$. If $A \circ\!\!-\!\!\circ C$ appears in $\mathcal{P}_{AFCI}$ and C is a parent of B in $\mathcal{P}_{AFCI}$ (i.e. the edge $A \circ\!\!-\!\!\circ C$ is required by condition $\mathbf{C_2}$ to be oriented as $A \leftarrow C$), then C is a parent of K in $\mathcal{P}_{AFCI}$.*

Proof. If $B = K$, it is trivial that C is a parent of K. Suppose $B \neq K$. Since $A \circ\!\!\rightarrow B \in \mathbf{REL}(J \circ\!\!\rightarrow K)$, B is a parent of K. By Lemma 4.3.18, $A \circ\!\!\rightarrow K$ is present in $\mathcal{P}_{AFCI}$. It follows that C is adjacent to K, for otherwise $\langle C, B, A, K\rangle$ would constitute a discriminating path for A in $\mathcal{P}_{AFCI}$, and the circle at A on $A \circ\!\!\rightarrow K$ could be oriented by $\mathcal{R}4$. Furthermore, the edge between C and K must be $C \rightarrow K$, as required by $\mathcal{R}8$. Hence C is a parent of K. $\square$

Lemma 4.3.21. *Suppose $A \circ\!\!\rightarrow B \in \mathbf{REL}(J \circ\!\!\rightarrow K)$, $A \circ\!\!-\!\!\circ C$ and C is a parent of B in $\mathcal{P}_{AFCI}$ (i.e. the edge $A \circ\!\!-\!\!\circ C$ is required by condition $\mathbf{C_2}$ to be oriented as $A \leftarrow C$). Then*

(1) if for some $D \in \mathbf{AR}(J \circ\!\!\rightarrow K)$, $C \circ\!\!-\!\!\circ D$ is in $\mathcal{P}_{AFCI}$, then $C \in \mathbf{AR}(J \circ\!\!\rightarrow K)$ (so that the edge $C \circ\!\!-\!\!\circ D$ is not subject to $\mathbf{C_1}$);

(2) If $u = \langle C, A, \ldots\rangle$ is an uncovered circle path, no vertex (except possibly C) on u is in $\mathbf{AR}(J \circ\!\!\rightarrow K)$.

Proof. To show (1), note that if $D \in \mathbf{AR}(J \circ\!\!\rightarrow K)$, then there is some vertex X such that $X \circ\!\!\rightarrow D \in \mathbf{REL}(J \circ\!\!\rightarrow K)$. By $\mathbf{CP}1$ (Lemma 3.3.1), $X \circ\!\!\rightarrow C$ or $X \rightarrow C$ is in $\mathcal{P}_{AFCI}$. By Lemma 4.3.20, C is a parent of K. So it is not $X \rightarrow C$ in $\mathcal{P}_{AFCI}$, otherwise $X \circ\!\!\rightarrow K$, which is shown to be present by Lemma 4.3.18, could be oriented as $X \to K$ by $\mathcal{R}8$. So it must be $X\circ\!\!\rightarrow C$ in $\mathcal{P}_{AFCI}$. Since $X\circ\!\!\rightarrow D \in \mathbf{REL}(J\circ\!\!\rightarrow K)$ and C is a parent of K, $X\circ\!\!\rightarrow C$ obviously satisfies Definition 4.3.2, which means $C \in \mathbf{AR}(J\circ\!\!\rightarrow K)$.

To prove (2), suppose for contradiction that some vertex $E \neq C$ on u is in $\mathbf{AR}(J\circ\!\!\rightarrow K)$. Obviously $E \neq K$, otherwise $A\circ\!\!\rightarrow E$ would be present in $\mathcal{P}_{AFCI}$ by Lemma 4.3.18, which contradicts Lemma 3.3.2. So E is a parent of K. Now consider the edge $A\circ\!\!\rightarrow K$, which is implied to exist by Lemma 4.3.18. $A \circ\!\!-\!\!\circ C$ constitutes an uncovered p.d. path from A to C, a parent of K, as implied by Lemma 4.3.20; $u(A, E)$ is an uncovered p.d. path from A to E, a parent of K. Since u is uncovered, $A\circ\!\!\rightarrow K$ could be oriented as $A \to K$ by $\mathcal{R}10$, a contradiction. □

Lemma 4.3.22. *For any uncovered circle path $u = \langle A, \cdots E\rangle$ in $\mathcal{P}_{AFCI}$, either the edge incident to A is not required by $\mathbf{C_2}$ to be oriented out of A, or the edge incident to E is not required by $\mathbf{C_2}$ to be oriented out of E.*

Proof. Suppose for contradiction that the contrary is true. By Lemma 4.3.20, both A and E are parents of K. Let B be the vertex adjacent to A on u. By our supposition, $A \circ\!\!-\!\!\circ B$ is required by $\mathbf{C_2}$ to be oriented as $A \to B$. This means, by Definition 4.3.3, that there is a vertex C such that $B\circ\!\!\rightarrow C \in \mathbf{REL}(J\circ\!\!\rightarrow K)$ (and A is a parent of C). Consider $B\circ\!\!\rightarrow K$, which is shown to be present by Lemma 4.3.18. $B \circ\!\!-\!\!\circ A$ constitutes an uncovered p.d. path from B to A, a parent of K; $u(B, E)$ constitutes an uncovered p.d. path from B to E, a parent of K. Thus it is easy to see that $B\circ\!\!\rightarrow K$ could be oriented as $B \to K$ by $\mathcal{R}10$, a contradiction. □

Lemma 4.3.23. *If $A\circ\!\!\rightarrow B \in \mathbf{REL}(J\circ\!\!\rightarrow K)$, and $u = \langle A, C, \cdots\rangle$ is an uncovered circle path such that C is not adjacent to B in $\mathcal{P}_{AFCI}$ (so that the edge between A and C is required by $\mathbf{C_3}$ to be oriented as $A \to C$), then no vertex on u is a parent of K in $\mathcal{P}_{AFCI}$.*

Proof. Since $A \circ\!\!\rightarrow B \in \mathbf{REL}(J \circ\!\!\rightarrow K)$, by Lemma 4.3.18, $A \circ\!\!\rightarrow K$ is present in $\mathcal{P}_{AFCI}$. Suppose for contradiction that a vertex D (which could be C) on u is a parent of K. By definition (Definition 4.3.2), either $B = K$ or B is a parent of K. We consider the two cases separately and derive a contradiction in each.

Case 1: $B = K$, and hence K and C are not adjacent (which means D can't be C in this case). So $u(A, D) \oplus D \rightarrow K$ is a p.d. path from A to K such that the vertex adjacent to A on the path, namely C, is not adjacent to K. Let E be the first vertex after C on the path which is adjacent to K (there must be one, because D is adjacent to K). The edge between E and K, by Corollary 4.3.14, is either $E \circ\!\!\rightarrow K$ or $E \rightarrow K$. It follows that $\langle A, C, \cdots, E, K\rangle$ forms an uncovered p.d. path from A to K such that C and K are not adjacent. Hence $A \circ\!\!\rightarrow K$ could be oriented as $A \rightarrow K$ by $\mathcal{R}9$, a contradiction.

Case 2: $B \rightarrow K$ is in $\mathcal{P}_{AFCI}$. Then $u(A, D)$ is an uncovered p.d. path from A to D, a parent of K, and $A \circ\!\!\rightarrow B$ is an uncovered p.d. path from A to B, a parent of K. Since C and B are not adjacent, the edge $A \circ\!\!\rightarrow K$ could be oriented as $A \rightarrow K$ by $\mathcal{R}10$, a contradiction. □

Lemma 4.3.24. *Suppose $A \circ\!\!\rightarrow B, C \circ\!\!\rightarrow D \in \mathbf{REL}(J \circ\!\!\rightarrow K)$, $A \neq C$ and $u = \langle A, \cdots, C\rangle$ is an uncovered circle path in $\mathcal{P}_{AFCI}$. Either the vertex next to A on u is adjacent to B (so that $\mathbf{C_3}$ does not require orienting the edge out of A), or the vertex next to C on u is adjacent to D (so that $\mathbf{C_3}$ does not require orienting the edge out of C).*

Proof. Suppose for contradiction that the vertex next to A on u (which could be C) is not adjacent to B, and the vertex next to C on u (which could be A) is not adjacent to D. We consider three cases separately and derive a contradiction in each.

Case 1: $B = D$. In this case, since D is not adjacent to the vertex next to C on

u, $u \oplus C\circ\!\!\rightarrow B$ is an uncovered p.d. path from A to B such that the vertex adjacent to A on the path is not adjacent to B. Hence $A\circ\!\!\rightarrow B$ could be oriented by $\mathcal{R}9$ as $A \rightarrow B$, a contradiction.

Case 2: $B \neq D$ and one of them is K. Without loss of generality, suppose $B = K$. Since $C\circ\!\!\rightarrow D \in \mathbf{REL}(J\circ\!\!\rightarrow K)$, and $D \neq K$, by definition (Definition 4.3.2), D is a parent of K (B). Then $u \oplus C\circ\!\!\rightarrow D$ constitutes an uncovered p.d. path from A to D such that the vertex adjacent to A on the path is not adjacent to B. This is exactly the same situation as *Case 1* in the proof of Lemma 4.3.23, which implies that $A\circ\!\!\rightarrow B$ could be oriented as $A \rightarrow B$ by $\mathcal{R}9$, a contradiction.

Case 3: $B \neq D$ and neither of them is K. By definition (Definition 4.3.2), both B and D are parents of K. Consider the edge $A\circ\!\!\rightarrow K$, which is shown to be present by Lemma 4.3.18. Since $A\circ\!\!\rightarrow B$ is an uncovered p.d. path from A to B, a parent of K, $u \oplus C\circ\!\!\rightarrow D$ is an uncovered p.d. path from A to D, a parent of K, and that the vertex next to A on u is not adjacent to B, the edge $A\circ\!\!\rightarrow K$ could be oriented as $A \rightarrow K$ by $\mathcal{R}10$, a contradiction. □

In our *Orientation Algorithm*, some $\circ\!\!-\!\!\circ$ edges are explicitly oriented to satisfy one of $\mathbf{C_1} - \mathbf{C_3}$. Lemmas 4.3.19, 4.3.21, 4.3.22, 4.3.23, 4.3.24 ensure that such orientations will not at the same time violate $\mathbf{C_1} - \mathbf{C_3}$ (and hence $\mathbf{C_1}$, $\mathbf{C_2}$ and $\mathbf{C_3}$ are consistent in themselves and with each other). Furthermore, these lemmas imply that if we propagate such orientations with $\mathbf{UR}_1$ alone, no violation of $\mathbf{C_1} - \mathbf{C_3}$ would occur. (These claims will be formally demonstrated in Lemma 4.3.25 and Lemma 4.3.28.)

However, it is not yet clear whether propagations with $\mathbf{UR}_1 - \mathbf{UR}_3$ together will create violations of $\mathbf{C_1} - \mathbf{C_3}$. We resolve this worry in Lemma 4.3.25 and Lemma 4.3.28, with which we establish the key fact that no violation of $\mathbf{C_1} - \mathbf{C_3}$ would occur in the *Orientation Algorithm.* Note that if any violation were to occur, it could only occur by the end of the third stage of the *Orientation Algorithm*, namely before all $\circ\!\!-\!\!\circ$ edges in $\mathbf{E_1} \cup \mathbf{E_2} \cup \mathbf{E_3}$ get oriented. Let $\mathcal{D}^*$ be the resulting graph at the end of the third stage of the *Orientation Algorithm*. Clearly the $\circ\!\!-\!\!\circ$ edges left in $\mathcal{D}^*$, if any, do not belong to $\mathbf{E_1} \cup \mathbf{E_2} \cup \mathbf{E_3}$, and hence are not relevant to $\mathbf{C_1} - \mathbf{C_3}$. The next lemma states two important properties of $\mathcal{D}^*$. (We assume, without loss of generality, that $\mathbf{UR}_1$ has priority over $\mathbf{UR}_2$ and $\mathbf{UR}_3$.)

Lemma 4.3.25. *Let $\mathcal{D}^*$ be the resulting graph at the end of the third stage of the Orientation Algorithm.*

(i) for any vertex $W \in \mathbf{AR}(J \circ\!\!\rightarrow K)$, there is no edge directed into W in $\mathcal{D}^*$;

(ii) for any three vertices X, Y, Z, if $X \circ\!\!\rightarrow Y \in \mathbf{REL}(J \circ\!\!\rightarrow K)$, $X \circ\!\!-\!\!\circ Z$ and Z is a parent of Y in $\mathcal{P}_{AFCI}$, then there is no edge directed into Z in $\mathcal{D}^*$.

Proof. See section 4.3.5. $\square$

Corollary 4.3.26. *In the course of the Orientation Algorithm, no violation of* $\mathbf{C_1}$ *occurs.*

Proof. This follows trivially from (i) in Lemma 4.3.25. $\square$

Corollary 4.3.27. *In the course of the Orientation Algorithm, no violation of* $\mathbf{C_2}$ *occurs.*

Proof. This follows trivially from (ii) in Lemma 4.3.25. $\square$

Lemma 4.3.28. *In the course of the Orientation Algorithm, no violation of* $\mathbf{C_3}$ *occurs.*

Proof. See section 4.3.5. $\square$

Let $\mathcal{D}_{J\circ\to K}$ be the DAG output of the **Orientation Algorithm**. We have thus proved the following proposition:

Corollary 4.3.29. *$\mathcal{D}_{J\circ\to K}$ is a DAG orientation of $\mathbf{C}(\mathcal{P}_{AFCI})$ free of unshielded colliders and agreeable to $J\circ\!\!\to K$.*

Proof. It follows from Corollary 4.3.26, Corollary 4.3.27, Lemma 4.3.28 and the correctness of Meek's algorithm. □

Let $\mathcal{H}_{J\circ\to K}$ be the graph resulting from orienting $\mathbf{C}(\mathcal{P}^{aag}_{AFCI})$ — the circle component of the arrowhead augmented graph of $\mathcal{P}_{AFCI}$, which is the same as the circle component of $\mathcal{P}_{AFCI}$ — into $\mathcal{D}_{J\circ\to K}$. By Lemma 4.3.6 and Corollary 4.3.29, $\mathcal{H}_{J\circ\to K}$ is a MAG equivalent to $\mathcal{G}$. Note that in $\mathcal{H}_{J\circ\to K}$ there is the edge $J \to K$. As planned earlier, what is left to show is that $\mathcal{H}_{J\circ\to K}$ can be transformed into a MAG where $J \leftrightarrow K$ appears through a sequence of equivalence-preserving changes of $\to$ into $\leftrightarrow$ (recall Lemma 4.3.9).

First we mention two simple facts about $\mathcal{P}_{AFCI}$.

Lemma 4.3.30. *For any $A\circ\!\!\to B$ in $\mathcal{P}_{AFCI}$, if there is a p.d. path u other than $A\circ\!\!\to B$ from A to B, then some vertex on u is adjacent to both A and B.*

Proof. The argument is an induction on the length of u. If u consists of two edges, the interior vertex on u (i.e., other than A or B) is obviously adjacent to both A and B. Suppose u consists of three edges. If it is covered, then obviously one of the two interior vertices is adjacent to both A and B. If it is uncovered, then the vertex adjacent to A on u must also be adjacent to B, for otherwise $A\circ\!\!\to B$ could be oriented by $\mathcal{R}9$. In the inductive step, suppose the proposition holds if u consists of less than n edges. Consider the case where u consists of n edges. If it is covered, then a subsequence of u constitutes a p.d. path from A to B with less than n edges, and hence by the inductive hypothesis, a vertex on u is adjacent to both A and B. If u is uncovered, then the vertex adjacent to A on u must also be adjacent to B, for otherwise $A \circ\!\!\rightarrow B$ could be oriented by $\mathcal{R}9$. $\square$

Lemma 4.3.31. *Suppose $C \leftarrow\!\!\circ A \circ\!\!\rightarrow B$ is in $\mathcal{P}_{AFCI}$. If C and B are not adjacent, then $A \circ\!\!\rightarrow B \notin \mathbf{REL}(J \circ\!\!\rightarrow K)$ or $A \circ\!\!\rightarrow C \notin \mathbf{REL}(J \circ\!\!\rightarrow K)$.*

Proof. Suppose for contradiction that $A \circ\!\!\rightarrow B \in \mathbf{REL}(J \circ\!\!\rightarrow K)$ and $A \circ\!\!\rightarrow C \in \mathbf{REL}(J \circ\!\!\rightarrow K)$. By Lemma 4.3.18, $A \circ\!\!\rightarrow K$ is in $\mathcal{P}_{AFCI}$. It also follows that $B \neq K$ and $C \neq K$, for otherwise B and C would be adjacent. Then, by Definition 4.3.2, both B and C are parents of K, which implies that $A \circ\!\!\rightarrow K$ could be oriented by $\mathcal{R}10$ because C and B are not adjacent, a contradiction. $\square$

Now we present the key lemma, of which we provide an informal explanation here. Note that in $\mathcal{H}_{J \circ\rightarrow K}$ all $\circ\!\!\rightarrow$ edges in $\mathcal{P}_{AFCI}$ are oriented as $\rightarrow$. So in particular, all edges in $\mathbf{REL}(J \circ\!\!\rightarrow K)$ are oriented as $\rightarrow$. Let $\mathcal{M}$ be any MAG identical to $\mathcal{H}_{J \circ\rightarrow K}$ except possibly that some $\circ\!\!\rightarrow$ edges in $\mathbf{REL}(J \circ\!\!\rightarrow K)$ are oriented as $\leftrightarrow$ (instead of $\rightarrow$). The lemma below shows that if not all edges in $\mathbf{REL}(J \circ\!\!\rightarrow K)$ are oriented as $\leftrightarrow$ in $\mathcal{M}$, then some $\rightarrow$ in $\mathcal{M}$ corresponding to some $\circ\!\!\rightarrow$ edge in $\mathbf{REL}(J \circ\!\!\rightarrow K)$ satisfies the conditions in Lemma 4.3.9, and hence can be changed into $\leftrightarrow$ while preserving equivalence. As a special case, for example, in $\mathcal{H}_{J \circ\rightarrow K}$ some $\rightarrow$ edge corresponding to a $\circ\!\!\rightarrow$ edge in $\mathbf{REL}(J \circ\!\!\rightarrow K)$ can be changed into $\leftrightarrow$. After this change, some remaining $\rightarrow$ corresponding to a $\circ\!\!\rightarrow$ edge, if any, can be further changed to $\leftrightarrow$ while preserving equivalence. This process can be continued until every edge in $\mathbf{REL}(J \circ\!\!\rightarrow K)$, and in particular $J \circ\!\!\rightarrow K$, can be oriented as $J \leftrightarrow K$ while preserving the Markov equivalence with $\mathcal{G}$.

Lemma 4.3.32. *Let $\mathcal{M}$ be any MAG identical to $\mathcal{H}_{J \circ\rightarrow K}$ except possibly that some $\circ\!\!\rightarrow$ edges in $\mathbf{REL}(J \circ\!\!\rightarrow K)$ are oriented as $\leftrightarrow$ (instead of $\rightarrow$). Let*

$$\mathbf{RREL} = \{A \to B \ in\ \mathcal{M} | A{\circ\!\!\to} B \ is\ in\ \mathcal{P}_{AFCI} \ and\ A{\circ\!\!\to} B \in \mathbf{REL}(J{\circ\!\!\to} K)\}$$

If **RREL** *is not empty, then some edge therein can be changed to* $\leftrightarrow$ *while preserving Markov equivalence with* $\mathcal{M}$*.*

Proof. See section 4.3.5. □

Corollary 4.3.33. $\mathcal{G}_{J\circ\!\to K}$ *can be transformed via a series of equivalence-preserving changes into a MAG where* $J \leftrightarrow K$ *appears.*

Proof. Using Lemma 4.3.32, a simple induction on the number of edges in $\mathbf{REL}(J{\circ\!\!\to} K)$ suffices. □

We thus arrive at a major result of this dissertation.

Theorem 4.3.1. *Assuming the CMC and the CFC, the Augmented FCI algorithm returns the CPAG for the true causal MAG given a perfect oracle of conditional independence. In other words, for every circle in* $\mathcal{P}_{AFCI}$*, the output of the augmented FCI algorithm, there is a MAG Markov equivalent to the true causal MAG in which the circle is oriented as an arrowhead, and there is a MAG Markov equivalent to the true causal MAG in which the circle is marked as a tail.*

Proof. It follows readily from Lemma 4.3.6, Corollary 4.3.29 and Corollary 4.3.33, and Theorem 3.3.1. □

4.3.4 Significance of the Completeness Result

The CPAG constructed from a perfect oracle of conditional independence relations is useful not only because it displays all commonalities among the class of Markov equivalent MAGs that includes the true causal MAG, and hence clearly reveals what features of the unknown true causal MAG are or are not underdetermined by the oracle of conditional independence relations, but also because, given the causal interpretation of MAGs, we can easily read off the CPAG valid sentences about causal relations as entailed by the CMC and CFC together with the given conditional independence facts, in particular, sentences of the form "variable A is not a cause of variable B or any selection variable", and sentences of the form "variable A is a cause of variable B or some selection variable".

The primary concern of this dissertation is about causal insufficiency, not about selection effects. Suppose it is known that there are no selection effects, but the set of observed variables may be causally insufficient. In this case, as we said, we do not need $\mathcal{R}5 - \mathcal{R}7$. Moreover, the CPAG constructed by the AFCI procedure will not contain any — edge or ∘— edge. As mentioned in the previous chapter, this CPAG gives us all valid negative causal sentences by a simple criterion: the sentence "variable A is not a cause of variable B" is valid if and only if there is no potentially directed path from A to B in the CPAG. In other words, if there is a potentially directed path from A to B in the CPAG, then the sentence is false in some causal models (DAGs with latent variables) that satisfy the axioms; if there is no potentially directed path from A to B in the CPAG, the sentence is true in all causal models that satisfy the axioms.

The validity of positive causal sentences is more subtle. A sufficient condition for the sentence "variable A is a cause of variable B" to be valid[5] is obviously that there is a directed path from A to B in the CPAG. It is also easy to see that the sentence "variable A is a cause of B" is not valid if there is an edge between A and B but the edge is not $A \to B$. Another sufficient condition for the sentence to be invalid is that there is no directed path and at most one partially directed path from A to B. However, a sufficient and necessary condition for validity is not obvious. The presence

[5]Again, causal transitivity is assumed because the CFC is assumed.

of a directed path in the CPAG is not necessary. For example, consider the CPAG for four variables consisting of the following edges: $C \circ\!\!-\!\!\circ A \circ\!\!-\!\!\circ D$, $C \rightarrow B \leftarrow D$. Although there is no directed path from A to B in this CPAG, it is true that in every MAG that belongs to the Markov equivalence class represented by the CPAG, A is either a parent of C or a parent of D, and hence in every such MAG, A is an ancestor of B. This means that the sentence "A is a cause of B" is valid. We thus conjecture the following sufficient and necessary condition: the sentence "variable A is a cause of B" is valid if and only if there is a directed path from A to B in the CPAG or there exist two uncovered partially directed paths from A to B in the CPAG such that the vertices adjacent to A on the two paths respectively are distinct and are not adjacent.[6]

Apart from causal inference, the completeness result leads to a syntactic characterization of CPAGs by the orientation rules. The characterization, unlike the definition given in 3.2.1, does not need to mention a MAG. To borrow a term from Andersson et al. (1997), we say a non-circle mark in a partial mixed graph is *protected* by an orientation rule if it could be introduced by that orientation rule, given all other marks in the graph. The next theorem gives the necessary and sufficient conditions for a partial mixed graph — that is, a graph consisting of $\rightarrow$, $\leftrightarrow$, $--$, $\circ\!\!\rightarrow$, $\circ\!\!-\!\!\circ$, $\circ\!\!-$ — to be a CPAG for some Markov equivalence class of MAGs.

Theorem 4.3.2. *A partial mixed graph is a CPAG for a Markov equivalence class of MAGs if and only if*

(i) (a1)-(a3) (in Definition 3.1.1) and $\mathbf{CP}1 - \mathbf{CP}4$ *hold; and there is no inducing path between two non-adjacent vertices;*

(ii) the circle component is chordal;

[6] This is obviously related to $\mathcal{R}10$.

(iii) it is closed under $\mathcal{R}8 - \mathcal{R}10$[7]*; and*

(iv) every non-circle mark is protected by one of $\mathcal{R}0 - \mathcal{R}10$.

Note that the characterization of essential graphs — graphs that represent Markov equivalence classes of DAGs — given in Andersson et al. (1997) is essentially of the same sort. A proof of the theorem is not hard to construct given what we have shown in the last section, in particular, Lemma 4.3.6 (or Lemma 3.3.4). In particular, we can construct a MAG by arrowhead augmenting (or tail augmenting) the given partial mixed graph and orienting the circle component as a DAG with no unshielded colliders. The resulting MAG is a member of the Markov equivalence class of which the given graph is the CPAG.

4.3.5 Omitted Proofs

Proof of Lemma 4.3.9

Proof. We first show that each of the conditions is necessary (**only if**). Obviously if (t1) or (t2) fails, $\mathcal{G}'$ will not be ancestral. The failure of (t3) could be due to one of the following two cases:

Case 1: there is a vertex C which is a parent of A but not a parent of B. If B and C are not adjacent, then there is an unshielded collider in $\mathcal{G}'$ but not in $\mathcal{G}$, and hence the two graphs are not Markov equivalent. If B and C are adjacent, then $\mathcal{G}$ can't be ancestral (unless we have $C \to B$).

Case 2: there is a vertex C which is a spouse of A but not a parent or spouse of B. Again, if B and C are not adjacent, the two graphs can't be Markov equivalent because there is an unshielded collider in $\mathcal{G}$ but not in $\mathcal{G}'$. If B and C are adjacent,

[7] Other rules are taken care of by $\mathbf{CP}1 - \mathbf{CP}4$.

the edge between them must be $B \to C$ by the supposition. But then there is an almost directed cycle in $\mathcal{G}$.

Suppose for contradiction that (t4) fails, i.e., there is a discriminating path $u = \langle U, \cdots, V, A, B \rangle$ for A. If the edge between V and A is into A, then $\mathcal{G}$ and $\mathcal{G}'$ are not Markov equivalent, because (e3) in Proposition 3.1.2 is violated. If, on the other hand, the edge between V and A is not into A, then it must be $A \to V$. By the definition of discriminating path (Definition 3.1.7), V is a parent of B. So we have $A \to V \to B \leftrightarrow A$ in $\mathcal{G}'$, an almost directed cycle.

Next, we demonstrate the sufficiency of the conditions (**if**). Suppose (t1)-(t4) are met. We first verify that $\mathcal{G}'$ is a MAG, i.e., it is both ancestral and maximal. Suppose for contradiction that $\mathcal{G}'$ is not ancestral. Since $\mathcal{G}$ is ancestral, and $\mathcal{G}'$ differs from $\mathcal{G}$ only regarding the edge between A and B, in $\mathcal{G}'$ the violation of the definition of ancestral graphs (Definition 3.1.1) must involve the edge between A and B. So it can't be a violation of (a1), because a directed cycle would not involve $A \leftrightarrow B$. If it is a violation of (a2), i.e., there is an almost directed cycle in $\mathcal{G}'$, then that cycle includes $A \leftrightarrow B$, which means either A is an ancestor of B or B is an ancestor of A in $\mathcal{G}'$. The former case contradicts (t2), and the latter case yields a directed cycle in $\mathcal{G}$. So there can't be any violation of (a2) in $\mathcal{G}'$. Lastly, if there is a violation of (a3) in $\mathcal{G}'$, it must be that there is an undirected edge incident to A, which contradicts (t1). Hence $\mathcal{G}'$ must be ancestral.

To show that $\mathcal{G}'$ is maximal, suppose for the sake of contradiction that there is an inducing path u in $\mathcal{G}'$ between two non-adjacent vertices, D and E. Then u must include $A \leftrightarrow B$, otherwise u is also an inducing path in $\mathcal{G}$. Furthermore, A is not an endpoint of u, otherwise u is still an inducing path in $\mathcal{G}$ (in fact, there will be an almost directed path in $\mathcal{G}$ in that case). Suppose, without loss of generality, that D is the endpoint closer to A on u than it is to B. We show that some vertex on $u(D, A)$ other than A is B's spouse. Suppose not; we argue by induction that every vertex on $u(A, D)$, and in particular D, is a parent of B. By (t3), the vertex adjacent to A on $u(D, A)$ is either a parent or a spouse of B, but it is not a spouse by supposition, so it is a parent. In the inductive step, suppose the first n vertices next to A on $u(D, A)$ are B's parents, then the $n+1^{st}$ vertex V must be adjacent to B, otherwise the sub-path of u between this vertex and B forms a discriminating path for A which contradicts (t4). The edge between V and B obviously can't be undirected. Furthermore, by supposition, V is not a spouse of B, i.e., it is not $V \leftrightarrow B$. It can't be $V \leftarrow B$ either, because in that case there would be an almost directed cycle in $\mathcal{G}'$ (as the vertex before V, by the inductive hypothesis, is a parent of B), which we have shown to be impossible. So V must be a parent of B. Thus we have shown that every vertex on $u(A, D)$, and in particular D, is a parent of B. Then B must be an ancestor of E, because by the definition of inducing path (Definition 3.1.4), B is an ancestor of either D or E. So D is an ancestor of E, and it is obvious that the vertex adjacent to E on u must be an ancestor of D, which implies that there is an almost directed cycle in $\mathcal{G}'$. which we have shown to be absent. Hence a contradiction. So some vertex on $u(D, A)$ other than A is a spouse of B. Let C be such a vertex on $u(D, A)$. Replacing $u(C, B)$ on u with $C \leftrightarrow B$ yields an inducing path between D and E in $\mathcal{G}$, which contradicts the fact that $\mathcal{G}$ is maximal.

Having shown that $\mathcal{G}'$ is a MAG, we now verify that $\mathcal{G}$ and $\mathcal{G}'$ satisfy the conditions for Markov equivalence in Proposition 3.1.2. Obviously they have the same adjacencies, and share the same colliders except possibly A. But A will not be a collider in an unshielded triple, for condition (t3) requires that any vertex that is incident to an edge into A is also adjacent to B. So the only worry is that a triple $\langle C, A, B\rangle$ might be discriminated by a path, but (t4) guarantees that there is no such path. Therefore, $\mathcal{G}'$ is Markov equivalent to $\mathcal{G}$. $\square$

Proof of Lemma 4.3.25

Proof. We first demonstrate (i). Suppose for contradiction that some $\circ\!\!-\!\!\circ$ edge is oriented into a vertex in $\mathbf{AR}(J\circ\!\!\rightarrow K)$ by the end of the third stage of the *Orientation Algorithm*. Let the first occurrence of such an orientation be $A \circ\!\!-\!\!\circ B$ being oriented as $A \rightarrow B$, where $B \in \mathbf{AR}(J\circ\!\!\rightarrow K)$. We consider all the possible ways in which this orientation could occur and derive a contradiction in each.

Case 1: $A \circ\!\!-\!\!\circ B$ is oriented as $A \rightarrow B$ to satisfy one of $\mathbf{C_1} - \mathbf{C_3}$. Since B is in $\mathbf{AR}(J\circ\!\!\rightarrow K)$, $\mathbf{C_1}$ does not dictate this orientation. It can't be $\mathbf{C_2}$, as entailed by (2) in Lemma 4.3.21. So it must be $\mathbf{C_3}$, which means there is a vertex E such that $A\circ\!\!\rightarrow E \in \mathbf{REL}(J\circ\!\!\rightarrow K)$ and E, B are not adjacent. Then Lemma 4.3.23 implies that B is not a parent of K. Furthermore, by Lemma 4.3.18, $A\circ\!\!\rightarrow K$ is present in $\mathcal{P}_{AFCI}$, which implies that $B \neq K$ (because the edge between A and B is $A \circ\!\!-\!\!\circ B$ in $\mathcal{P}_{AFCI}$). It follows that B is not in $\mathbf{AR}(J\circ\!\!\rightarrow K)$, which is a contradiction.

Case 2: $A \circ\!\!-\!\!\circ B$ is oriented as $A \rightarrow B$ by an application of $\mathbf{UR}_2$. That is, there is a vertex C such that $A \circ\!\!-\!\!\circ C \circ\!\!-\!\!\circ B$ is in $\mathbf{C}(\mathcal{P}_{AFCI})$, and is oriented as $A \rightarrow C \rightarrow B$ before $A \circ\!\!-\!\!\circ B$ is oriented. Then $C \circ\!\!-\!\!\circ B$ being oriented as $C \rightarrow B$ would be an earlier occurrence of orientation into B. This contradicts our choice of $A \circ\!\!-\!\!\circ B$.

Case 3: $A \circ\!\!-\!\!\circ B$ is oriented as $A \rightarrow B$ by an application of $\mathbf{UR}_3$. Again, it is easy to see that this contradicts the assumption that $A \rightarrow B$ is the first orientation into B.

Case 4: $A \circ\!\!-\!\!\circ B$ is oriented as $A \rightarrow B$ by an application of $\mathbf{UR}_1$. Then there is a chain of applications of $\mathbf{UR}_1$ (which could consist of just one application) that leads to $A \to B$ where the first edge on the chain is *not* oriented by $\mathbf{UR}_1$. So there are three subcases to consider:

Case 4.1: the first edge is oriented to satisfy one of $\mathbf{C_1} - \mathbf{C_3}$. If it is $\mathbf{C_1}$, then in $\mathbf{C}(\mathcal{P}_{AFCI})$ there is an uncovered circle path with more than one edge between two vertices in $\mathbf{AR}(J \circ\!\!\to K)$, which contradicts Lemma 4.3.19. It can't be $\mathbf{C_2}$, as entailed by (2) in Lemma 4.3.21. So it must be $\mathbf{C_3}$, but in that case Lemma 4.3.23 implies that B is not a parent of K and Lemma 4.3.18 implies that $B \neq K$, which contradict the membership of B in $\mathbf{AR}(J \circ\!\!\to K)$.

Case 4.2: the first edge is oriented by $\mathbf{UR}_2$. That is, there are three vertices X, Y and Z (Z could be A) such that $X \circ\!\!-\!\!\circ Y \circ\!\!-\!\!\circ Z$ is in $\mathbf{C}(\mathcal{P}_{AFCI})$, and is oriented as $X \to Y \to Z$, which in turn orients the edge $X \circ\!\!-\!\!\circ Z$ as $X \to Z$. And $X \to Z$ initiates a chain of $\mathbf{UR}_1$ applications on an uncovered circle path $u = \langle X, Z, \cdots, B \rangle$ that eventually leads to the orientation of $A \to B$. Now we argue that for every vertex V on u between Z and B, there is an edge between Y and V already oriented as $Y \to V$ before $X \to Z$ is thus oriented. The argument is by induction. Let V_1 be the first vertex next to Z on u (V_1 is B if Z is A). Y and V_1 must be adjacent in $\mathbf{C}(\mathcal{P}_{AFCI})$, for otherwise $Z \circ\!\!-\!\!\circ V_1$ would be oriented as $Z \to V_1$ by $\mathbf{UR}_1$ *before* $X \circ\!\!-\!\!\circ Z$ is oriented by $\mathbf{UR}_2$, according to our convention of the priority of $\mathbf{UR}_1$. Since X and V_1 are not adjacent (because u is uncovered), $Y \circ\!\!-\!\!\circ V_1$ should be oriented as $Y \to V_1$ by $\mathbf{UR}_1$ before $X \to Z$ is thus oriented. In the inductive step, suppose $Y \to V_n$ is oriented as such before $X \to Z$ is thus oriented, where V_n is the n'the vertex after Z on u. Consider the $n + 1^{st}$ vertex V_{n+1}. Again, it must be adjacent to Y, otherwise the edge $V_n \circ\!\!-\!\!\circ V_{n+1}$ should be oriented by $\mathbf{UR}_1$ before $X \circ\!\!-\!\!\circ Z$ is oriented by our convention of the priority of $\mathbf{UR}_1$. Furthermore, by Lemma 4.3.16, X and V_{n+1} are not adjacent, so the edge $Y \circ\!\!-\!\!\circ V_{n+1}$ should be oriented as $Y \to V_{n+1}$ by $\mathbf{UR}_1$ before $X \circ\!\!-\!\!\circ Z$ gets oriented. Hence, in particular, $Y \to B$ is already present before $X \circ\!\!-\!\!\circ Z$ gets oriented, and hence before $A \circ\!\!-\!\!\circ B$ gets oriented. This contradicts our choice of $A \circ\!\!-\!\!\circ B$.

Case 4.3: the first edge is oriented by $\mathbf{UR}_3$. That is, there are four vertices X, Y, Z, W (Z could be A) such that $W \circ\!\!-\!\!\circ Y \circ\!\!-\!\!\circ Z$, $W \circ\!\!-\!\!\circ X \circ\!\!-\!\!\circ Z$, $X \circ\!\!-\!\!\circ Y$ are in $\mathbf{C}(\mathcal{P}_{AFCI})$, and that W, Z are not adjacent. Furthermore, $W \circ\!\!-\!\!\circ Y \circ\!\!-\!\!\circ Z$ is oriented as $W \to Y \to Z$, which in turn orients the edge $X \circ\!\!-\!\!\circ Z$ as $X \to Z$. This then initiates a chain of $\mathbf{UR}_1$ applications on an uncovered circle path $u = \langle X, Z, \cdots, B\rangle$ that eventually leads to the orientation of $A \to B$. Notice that W, Z are not adjacent, so $\langle W, X, Z, \cdots, B\rangle$ is also an uncovered path. By the exact same argument as in *Case 4.2*, we can show that for every vertex V between Z and B on u, there is an edge between Y and V already oriented as $Y \to V$ before $X \to Z$ is thus oriented. So in particular, $Y \to B$ is already present before $X \circ\!\!-\!\!\circ Z$ gets oriented, and hence before $A \circ\!\!-\!\!\circ B$ gets oriented. This contradicts our choice of $A \circ\!\!-\!\!\circ B$.

Thus we have established (i).

The proof of (ii) is completely parallel to the proof of (i). The only notable difference is that in the counterparts of *Case 1* and *Case 4.1*, we need to cite some different lemmas. Take *Case 1* for example. We need to argue that an orientation to satisfy one of $\mathbf{C_1} - \mathbf{C_3}$ will not be an orientation that violates (ii). For $\mathbf{C_1}$, it suffices to cite (1) in Lemma 4.3.21; for $\mathbf{C_2}$, we need to cite Lemma 4.3.22; for $\mathbf{C_3}$, we need Lemma 4.3.23. Other details are virtually the same as the arguments for (i). $\square$

Proof of Lemma 4.3.28

Proof. Suppose for contradiction that a violation of $\mathbf{C_3}$ occurs. Let the first occurrence be $A \circ\!\!-\!\!\circ C$ oriented as $A \leftarrow C$. This means there is a vertex B such that $A \circ\!\!\rightarrow B \in \mathbf{REL}(J \circ\!\!\rightarrow K)$, but C, B are not adjacent. Again, this orientation must occur by the end of the third stage of the *Orientation Algorithm*, so the following are all the possible ways in which this orientation could occur. We derive a contradiction in each.

Case 1: $A \circ\!\!-\!\!\circ C$ is oriented as $A \leftarrow C$ to satisfy one of $\mathbf{C_1} - \mathbf{C_3}$. Lemma 4.3.23 implies that it is not $\mathbf{C_1}$ or $\mathbf{C_2}$. So it must be $\mathbf{C_3}$, which, however, contradicts Lemma 4.3.24 (note that $A \circ\!\!-\!\!\circ C$ is the circle path relevant to the conditions of 4.3.24).

Case 2: $A \circ\!\!-\!\!\circ C$ is oriented as $A \leftarrow C$ by an application of $\mathbf{UR}_2$, which means there is a D such that $C \circ\!\!-\!\!\circ D \circ\!\!-\!\!\circ A$ is in $\mathbf{C}(\mathcal{P}_{AFCI})$ and is already oriented as $C \rightarrow D \rightarrow A$ (before $A \leftarrow C$ is thus oriented). Then D must be adjacent to B, otherwise $A \leftarrow D$ would be an earlier violation of $\mathbf{C_3}$. Furthermore, because $D \circ\!\!-\!\!\circ A \circ\!\!\rightarrow B$ is in $\mathcal{P}_{AFCI}$, the edge between D and B is either $D \rightarrow B$ or $D \circ\!\!\rightarrow B$ by Corollary 4.3.14. It can't be the former, for otherwise (ii) of Lemma 4.3.25 implies that there should not be any orientation into D (by the end of the third stage of the *Orientation Algorithm*), which contradicts $C \rightarrow D$. In the latter case, we argue that D is not a parent of K, and hence $D \circ\!\!\rightarrow B \in \mathbf{REL}(J \circ\!\!\rightarrow K)$. Suppose on the contrary that D is a parent of K. Obviously $A \circ\!\!\rightarrow K$, which is shown to be present in $\mathcal{P}_{AFCI}$ by Lemma 4.3.18, belongs to $\mathbf{REL}(J \circ\!\!\rightarrow K)$. Since $A \circ\!\!-\!\!\circ D \rightarrow K$, (ii) of Lemma 4.3.25 implies that there should not be any orientation into D (by the end of the third stage of the *Orientation Algorithm*), which contradicts $C \rightarrow D$. Therefore, D is not a parent of K, and hence $D \circ\!\!\rightarrow B \in \mathbf{REL}(J \circ\!\!\rightarrow K)$. But then $C \rightarrow D$ is an earlier violation of $\mathbf{C_3}$ than $C \rightarrow A$, a contradiction.

Case 3: $A \circ\!\!-\!\!\circ C$ is oriented as $A \leftarrow C$ by an application of $\mathbf{UR}_3$. That is, there are two vertices D, E such that $D \circ\!\!-\!\!\circ E \circ\!\!-\!\!\circ A$, $D \circ\!\!-\!\!\circ C \circ\!\!-\!\!\circ A$, $C \circ\!\!-\!\!\circ E$ are in

$\mathbf{C}(\mathcal{P}_{AFCI})$ and D, A are not adjacent. Furthermore, $D \circ\!\!-\!\!\circ E \circ\!\!-\!\!\circ A$ is already oriented as $D \rightarrow E \rightarrow A$ (before $A \leftarrow C$ is thus oriented). By the same argument as in *Case 2*, there must be an edge $E \circ\!\!\rightarrow B$ in $\mathcal{P}_{AFCI}$. Furthermore, D and B must be adjacent, otherwise $D \rightarrow E$ would contradict (ii) of Lemma 4.3.23. Corollary 4.3.14 implies that the edge between D and B is either $D \rightarrow B$ or $D \circ\!\!\rightarrow B$. But then the edge $A \circ\!\!\rightarrow B$ could be oriented as $A \rightarrow B$ by $\mathcal{R}9$ because $\langle A, C, D, B\rangle$ is an uncovered p.d. path from A to B such that C and B are not adjacent. Hence a contradiction.

Case 4: $A \circ\!\!-\!\!\circ C$ is oriented as $A \leftarrow C$ by an application of $\mathbf{UR}_1$. Then there is a chain of applications of $\mathbf{UR}_1$ (which could consist of just one application) that leads to $C \rightarrow A$ where the first edge on the chain is *not* oriented by $\mathbf{UR}_1$. So there are three subcases to consider:

Case 4.1: the first edge is oriented to satisfy one of $\mathbf{C_1}-\mathbf{C_3}$. Lemma 4.3.23 implies that it is not $\mathbf{C_1}$ or $\mathbf{C_2}$. So it must be $\mathbf{C_3}$, which, however, contradicts Lemma 4.3.24.

Case 4.2: the first edge is oriented by $\mathbf{UR}_2$. That is, there are three vertices X, Y and Z (Z could be C) such that $X \circ\!\!-\!\!\circ Y \circ\!\!-\!\!\circ Z$ is in $\mathbf{C}(\mathcal{P}_{AFCI})$, and is oriented as $X \rightarrow Y \rightarrow Z$, which in turn orients the edge $X \circ\!\!-\!\!\circ Z$ as $X \rightarrow Z$. And $X \rightarrow Z$ initiates a chain of $\mathbf{UR}_1$ applications on an uncovered circle path $u = \langle X, Z, \cdots, A\rangle$ that eventually leads to the orientation of $C \rightarrow A$. By the same induction as in Lemma 4.3.25, it is easy to show that for every vertex V on u after Z, there is an edge $Y \circ\!\!-\!\!\circ V$ in $\mathbf{C}(\mathcal{P}_{AFCI})$, oriented as $Y \rightarrow V$ before $X \circ\!\!-\!\!\circ Z$ is oriented. So in particular, Y is adjacent to A in $\mathbf{C}(\mathcal{P}_{AFCI})$ (and hence the edge between them is $Y \circ\!\!-\!\!\circ A$ in $\mathcal{P}_{AFCI}$), and the edge between them is oriented as $Y \rightarrow A$ before $X \circ\!\!-\!\!\circ Z$ is oriented. Then Y must be adjacent to B in $\mathcal{P}_{AFCI}$, for otherwise $Y \rightarrow A$ would be an earlier violation of $\mathbf{C_3}$ than $C \rightarrow A$. By Corollary 4.3.14, the edge between Y and B is either $Y \rightarrow B$ or $Y \circ\!\!\rightarrow B$ in $\mathcal{P}_{AFCI}$. If it is $Y \rightarrow B$, then according to (ii) of Lemma

4.3.25, there should not be any orientation into Y (by the end of the third stage of the *Orientation Algorithm*), which contradicts $X \to Y$. So it must be $Y \circ\!\!\to B$ in $\mathcal{P}_{AFCI}$. Furthermore, Y is not a parent of K, for otherwise there should not be any orientation into Y by (ii) of Lemma 4.3.25 (because $A \circ\!\!\to K \in \mathbf{REL}(J \circ\!\!\to K)$ by Lemma 4.3.18), which contradicts $X \to Y$. Furthermore, since $A \circ\!\!\to B \in \mathbf{REL}(J \circ\!\!\to K)$, there is a p.d. path from J to A such that no vertex on the path is a parent of K.

If this path passes Y, then the segment from J to Y is a p.d. path from J to Y with no parent of K on it. If the path does not include Y, then the path appended to $A \circ\!\!-\!\!\circ Y$ is a p.d. path from J to Y with no parent of K on it. Hence, in either case, $Y \circ\!\!\to B$ belongs to $\mathbf{REL}(J \circ\!\!\to K)$. This implies that X and B are adjacent in $\mathcal{P}_{AFCI}$, for otherwise $X \to Y$ would be an earlier violation of $\mathbf{C_3}$. The edge between X and B, furthermore, is either $X \to B$ or $X \circ\!\!\to B$ by Corollary 4.3.14. Then consider the path $\langle A, C, \cdots, X, B\rangle$, which is a p.d. path from A to B such that C is not adjacent to B and the segment between A and X is uncovered. It is easy to see that $A \circ\!\!\to B$ could be oriented as $A \to B$ by $\mathcal{R}9$, a contradiction.

Case 4.3 the first edge is oriented by $\mathbf{UR}_3$. That is, there are four vertices X, Y, Z, W (Z could be C) such that $W \circ\!\!-\!\!\circ Y \circ\!\!-\!\!\circ Z$, $W \circ\!\!-\!\!\circ X \circ\!\!-\!\!\circ Z$, $X \circ\!\!-\!\!\circ Y$ are in $\mathbf{C}(\mathcal{P}_{AFCI})$, and that W, Z are not adjacent. Furthermore, $W \circ\!\!-\!\!\circ Y \circ\!\!-\!\!\circ Z$ is oriented as $W \to Y \to Z$, which in turn orients the edge $X \circ\!\!-\!\!\circ Z$ as $X \to Z$. This then initiates a chain of $\mathbf{UR}_1$ applications on an uncovered circle path $u = \langle X, Z, \cdots, A\rangle$ that eventually leads to the orientation of $C \to A$. Notice that W, Z are not adjacent, so $\langle W, X, Z, \cdots, B\rangle$ is also an uncovered path. The rest of the argument is extremely similar to that of *Case 4.2*. $\square$

Proof of Lemma 4.3.32

Proof. Suppose **RREL** is not empty. Let

$$\mathbf{W} = \{B | \exists A \ s.t. A \to B \in \mathbf{RREL}\}$$

Let Y be a minimal vertex in $\mathbf{W}$, that is, $Y \in \mathbf{W}$ and no proper ancestor of D in $\mathcal{M}$ belongs to $\mathbf{W}$. Let X be a vertex such that $X \to Y \in \mathbf{RREL}$ and no proper descendant of X in $\mathcal{M}$ has this property. We show that $X \to Y$ satisfies the conditions (t1)-(t4) of Lemma 4.3.9.

Suppose, contrary to (t1), that X is an endpoint of an undirected edge $X — V$ in $\mathcal{M}$. Since any undirected edge in $\mathcal{M}$ is also in $\mathcal{H}_{J\circ\to K}$, $X — V$ is also in $\mathcal{P}_{AFCI}$ (see Definition 4.3.1). On the other hand, since $X \to Y \in \mathbf{RREL}$, $X \circ\!\!\to Y$ is in $\mathcal{P}_{AFCI}$. But then $X \circ\!\!\to Y$ could be oriented by $\mathcal{R}6$, a contradiction.

Suppose, contrary to (t2), that there is a directed path from X to Y in $\mathcal{M}$ that does not contain $X \to Y$. The corresponding path in $\mathcal{P}_{AFCI}$ must be potentially directed. It follows from Lemma 4.3.30 that some vertex Z on the path is adjacent to both X and Y. Since $\mathcal{M}$ is a MAG, we have $X \to Z \to Y$ in $\mathcal{M}$, and so the corresponding path $\langle X, Z, Y \rangle$ in $\mathcal{P}_{AFCI}$ is potentially directed. Notice that the edge between Z and Y can't be $Z —\!\circ\ Y$ in $\mathcal{P}_{AFCI}$ according to Lemma 4.3.1, because $X \circ\!\!\to Y$ is present. So, by the definition of p.d. path, the edge between X and Z is either $X \circ\!\!—\!\!\circ Z$ or $X \to Z$ or $X \circ\!\!\to Z$ or $X —\!\circ\ Z$, and the edge between Z and Y is either $Z \circ\!\!—\!\!\circ Y$ or $Z \to Y$ or $Z \circ\!\!\to Y$. We enumerate the possibilities below and derive a contradiction in each:

Case 1: $X \circ\!\!—\!\!\circ Z \circ\!\!—\!\!\circ Y$ appears in $\mathcal{P}_{AFCI}$. This contradicts property **CP**1 (Lemma 3.3.1) because $X \circ\!\!\to Y$ is present in $\mathcal{P}_{AFCI}$.

Case 2: $X \circ\!\!—\!\!\circ Z \to Y$ appears in $\mathcal{P}_{AFCI}$. Because $X \circ\!\!\to Y \in \mathbf{REL}(J \circ\!\!\to K)$, and $X \circ\!\!—\!\!\circ Z$ is oriented as $X \to Z$, $\mathcal{D}_{J\circ\to K}$ is not agreeable to $J \circ\!\!\to K$ ($\mathbf{C}_2$ being violated), which contradicts Corollary 4.3.29.

Case 3: $X \circ\!\!-\!\!\circ Z \circ\!\!\rightarrow Y$ appears in $\mathcal{P}_{AFCI}$. If Z is not a parent of K, then obviously $Z \circ\!\!\rightarrow Y \in \mathbf{REL}(J \circ\!\!\rightarrow K)$ (given the assumption that $X \circ\!\!\rightarrow Y \in \mathbf{REL}(J \circ\!\!\rightarrow K)$). But Z is a proper descendant of X in $\mathcal{M}$, which contradicts our choice of X. So Z must be a parent of K. Notice then that $X \circ\!\!\rightarrow K$ — which is shown to be present in $\mathcal{P}_{AFCI}$ by Lemma 4.3.18 — belongs to $\mathbf{REL}(J \circ\!\!\rightarrow K)$, but $X \circ\!\!-\!\!\circ Z$ is oriented as $X \rightarrow Z$, which means that $\mathcal{D}_{J \circ\!\rightarrow K}$ is not agreeable to $J \circ\!\!\rightarrow K$ ($\mathbf{C}_2$ being violated). This contradicts Corollary 4.3.29.

Case 4: $X \rightarrow Z \circ\!\!-\!\!\circ Y$ appears in $\mathcal{P}_{AFCI}$. Since $X \circ\!\!\rightarrow K$ is present in $\mathcal{P}_{AFCI}$ by Lemma 4.3.18, $Z \neq K$. Moreover, Z is not a parent of K, for otherwise $X \circ\!\!\rightarrow K$ could be oriented as $X \rightarrow K$ by $\mathcal{R}8$. So $Z \notin \mathbf{AR}(J \circ\!\!\rightarrow K)$. It follows that $\mathcal{D}_{J \circ\!\rightarrow K}$ is not agreeable to $J \circ\!\!\rightarrow K$, because $Z \circ\!\!-\!\!\circ Y$ is oriented as $Z \rightarrow Y$ ($\mathbf{C}_1$ being violated), which contradicts Corollary 4.3.29.

Case 5: $X \rightarrow Z \rightarrow Y$ appears in $\mathcal{P}_{AFCI}$. Then $X \circ\!\!\rightarrow Y$ could be oriented by $\mathcal{R}8$, a contradiction.

Case 6: $X \rightarrow Z \circ\!\!\rightarrow Y$ appears in $\mathcal{P}_{AFCI}$. Then Z is not a parent of K, for otherwise $X \circ\!\!\rightarrow K$ could be oriented by $\mathcal{R}8$. It follows that $Z \circ\!\!\rightarrow Y \in \mathbf{REL}(J \circ\!\!\rightarrow K)$. But Z is a proper descendant of X in $\mathcal{M}$, which contradicts our choice of X.

Case 7: $X \circ\!\!\rightarrow Z \circ\!\!-\!\!\circ Y$ appears in $\mathcal{P}_{AFCI}$. Then $Z \notin \mathbf{AR}(J \circ\!\!\rightarrow K)$, for otherwise $X \circ\!\!\rightarrow Z \in \mathbf{REL}(J \circ\!\!\rightarrow K)$, but Z is a proper ancestor of Y in $\mathcal{M}$, which contradicts our choice of Y. However, $Z \circ\!\!-\!\!\circ Y$ is oriented as $Z \rightarrow Y$, which means that $\mathcal{D}_{J \circ\!\rightarrow K}$ is not agreeable to $J \circ\!\!\rightarrow K$ ($\mathbf{C}_1$ being violated). This contradicts Corollary 4.3.29.

Case 8: $X \circ\!\!\rightarrow Z \rightarrow Y$ appears in $\mathcal{P}_{AFCI}$. We argue that Z is a parent of K in $\mathcal{P}_{AFCI}$. This is obvious if $Y = K$. Suppose $Y \neq K$, then (by the definition of "relevance") Y is a parent of K in $\mathcal{P}_{AFCI}$. It follows that Z is adjacent to K, for otherwise the edge $X \circ\!\!\rightarrow K$ (which is shown to be present by Lemma 4.3.18) could be oriented by $\mathcal{R}4$ because $\langle Z, Y, X, K\rangle$ would constitute a discriminating path for X in $\mathcal{P}_{AFCI}$. Furthermore, since $Z \to Y \to K$ is in $\mathcal{P}_{AFCI}$, the edge between Z and K is $Z \to K$ in $\mathcal{P}_{AFCI}$. Hence $X{\circ\!\!\to} Z \in \mathbf{REL}(J{\circ\!\!\to} K)$. But Z is a proper ancestor of Y in $\mathcal{M}$, which contradicts our choice of Y.

Case 9: $X{\circ\!\!\to} Z{\circ\!\!\to} Y$ appears in $\mathcal{P}_{AFCI}$. If Z is not a parent of K, then $Z{\circ\!\!\to} Y \in \mathbf{REL}(J{\circ\!\!\to} K)$. But Z is a proper descendant of X in $\mathcal{M}$, which contradicts our choice of X. So Z must be a parent of K. But then $X{\circ\!\!\to} Z \in \mathbf{REL}(J{\circ\!\!\to} K)$, and Z is a proper ancestor of Y in $\mathcal{M}$, which contradicts our choice of Y.

Case 10: $X {-\!\!\circ}\, Z\, {\circ\!\!-\!\!\circ} Y$ appears in $\mathcal{P}_{AFCI}$. This contradicts Lemma 4.3.3, because $X{\circ\!\!\to} Y$ is present.

Case 11: $X {-\!\!\circ}\, Z \to Y$ appears in $\mathcal{P}_{AFCI}$. Then $X{\circ\!\!\to} Y$ in $\mathcal{P}_{AFCI}$ could be oriented as $X \to K$ by $\mathcal{R}8$, a contradiction.

Case 12: $X {-\!\!\circ}\, Z{\circ\!\!\to} Y$ appears in $\mathcal{P}_{AFCI}$. Z is not a parent of K, for otherwise $X{\circ\!\!\to} K$ could be oriented by $\mathcal{R}8$. So $Z{\circ\!\!\to} Y \in \mathbf{REL}(J{\circ\!\!\to} K)$. But Z is a proper descendant of X in $\mathcal{M}$, which contradicts our choice of X.

Next, we show that condition (t3) holds as well. For any $W \to X$ in $\mathcal{M}$, it corresponds to either $W \to X$ or $W{\circ\!\!\to} X$ or $W\, {\circ\!\!-\!\!\circ} X$ or $W {-\!\!\circ}\, X$ in $\mathcal{P}_{AFCI}$. We argue that in any case W and Y are adjacent. In the former two cases, by Lemma 3.3.1, W and Y are adjacent (since there is a circle at X on $X{\circ\!\!\to} Y$). In the case of $W\, {\circ\!\!-\!\!\circ} X$, since it is oriented as $W \to X$ in $\mathcal{M}$, W must be adjacent to Y, for otherwise $\mathcal{D}_{J{\circ\!\!\to} K}$ is not agreeable to $J{\circ\!\!\to} K$, which contradicts Corollary 4.3.29. In the case of $W {-\!\!\circ}\, X$, by Lemma 4.3.3, W and Y are adjacent. Furthermore, the edge between W and Y must be $W \to Y$ in $\mathcal{M}$, because $W \to X \to Y$ is in $\mathcal{M}$ and $\mathcal{M}$ is a MAG.

For any $W \leftrightarrow X$ in $\mathcal{M}$, it corresponds to either $W \leftrightarrow X$ or $W{\circ\!\!\to} X$ or $W \leftarrow\!\!\circ X$

in $\mathcal{P}_{AFCI}$. In the former two cases, W and Y are adjacent by Lemma 3.3.1. In the latter case, $W \leftarrow\!\circ X \in \mathbf{REL}(J \circ\!\!\rightarrow K)$ by our assumption about bi-directed edges in $\mathcal{M}$. It then follows from Lemma 4.3.31 that W and Y are adjacent. So W and Y are adjacent in $\mathcal{M}$. Furthermore, since $\mathcal{M}$ is a MAG, the edge between W and Y is either $W \rightarrow Y$ or $W \leftrightarrow Y$ in $\mathcal{M}$.

Lastly, we show that condition (t4) in Lemma 4.3.9 is also satisfied. Suppose otherwise, that is, in $\mathcal{M}$ there is a path $p = (V_0, V_1, \cdots, V_n = X, Y)$ which is discriminating for X. Without loss of generality, suppose p is a shortest such path. Below we derive a contradiction by (eventually) showing that the corresponding path of p in $\mathcal{P}_{AFCI}$ is also a discriminating path in $\mathcal{P}_{AFCI}$, and hence the circle at X on $X \circ\!\!\rightarrow Y$ could be oriented by $\mathcal{R}4$.

Note first that the subpath $p(V_0, X)$ is into X in $\mathcal{M}$, for otherwise there is a directed path from X to Y other than the edge $X \rightarrow Y$ (which easily follows from the definition of discriminating path), which contradicts the already established fact that (t2) holds.

It follows that every edge on the subpath $p(V_1, X)$ is bi-directed in $\mathcal{M}$.

We now argue that in $\mathcal{P}_{AFCI}$ the edge between V_0 and V_1 is $V_0 {*\!\!\rightarrow} V_1$, i.e., the arrowhead at V_1 on this edge is already present in $\mathcal{P}_{AFCI}$. The following two facts will be useful: (1) $V_1 \leftrightarrow V_2$ (V_2 could be X) appears in $\mathcal{M}$; and (2) there can't be an edge between V_0 and V_2 that is into V_2 (i.e., has an arrowhead at V_2) in $\mathcal{P}_{AFCI}$. For otherwise either $\langle V_0, V_2, \cdots, V_n = X, Y\rangle$ constitutes a shorter discriminating path in $\mathcal{M}$ (if $V_2 \neq X$), or $X \circ\!\!\rightarrow Y$ in $\mathcal{P}_{AFCI}$ could be oriented as $X \rightarrow Y$ by $\mathcal{R}1$ (if $V_2 = X$), either of which is a contradiction.

Suppose for contradiction that the arrowhead at V_1 on the edge between V_0 and V_1 is not present in $\mathcal{P}_{AFCI}$. Then the mark must be a circle in $\mathcal{P}_{AFCI}$, i.e., the edge between V_0 and V_1 in $\mathcal{P}_{AFCI}$ is either $V_0 \text{---}\circ V_1$ or $V_0 \circ\text{---}\circ V_1$ or $V_0 \leftarrow\circ V_1$ (the mark at V_1 can't be a tail because in $\mathcal{M}$ it is an arrowhead).

If $V_0 \text{---}\circ V_1$ appears in $\mathcal{P}_{AFCI}$, then the arrowhead at V_1 on $V_1 \leftrightarrow V_2$ is not in $\mathcal{P}_{AFCI}$ by Lemma 4.3.1. So the edge must be $V_1 \circ\!\!\rightarrow V_2$ in $\mathcal{P}_{AFCI}$. It then follows from Lemma 4.3.3 that there is an edge $V_0 *\!\!\rightarrow V_2$ in $\mathcal{P}_{AFCI}$, which contradicts fact (2) mentioned above. So the edge between V_0 and V_1 is not $V_0 \text{---}\circ V_1$.

If the edge is $V_0 \circ\text{---}\circ V_1$, then $V_1 \leftrightarrow V_2$ is not already in $\mathcal{P}_{AFCI}$, for otherwise by Lemma 3.3.2, there would also be an edge $V_0 \leftrightarrow V_2$ in $\mathcal{P}_{AFCI}$, which again contradicts fact (2). By our assumption about bi-directed edges in $\mathcal{M}$, either $V_1 \circ\!\!\rightarrow V_2$ or $V_1 \leftarrow\circ V_2$ appears in $\mathcal{P}_{AFCI}$ and belongs to $\mathbf{REL}(J \circ\!\!\rightarrow K)$. In the former case ($V_1 \circ\!\!\rightarrow V_2$), V_0 must be adjacent to V_2, for otherwise the orientation of $V_0 \circ\text{---}\circ V_1$ (into $V_0 \rightarrow V_1$) is not agreeable to $J \circ\!\!\rightarrow K$ ($\mathbf{C}_3$ being violated). By Corollary 4.3.14, the edge between V_0 and V_2 is either $V_0 \rightarrow V_2$ or $V_0 \circ\!\!\rightarrow V_2$ in $\mathcal{P}_{AFCI}$, which contradicts fact (2). In the latter case ($V_1 \leftarrow\circ V_2$), by Lemma 3.3.2, either $V_0 \leftarrow V_2$ or $V_0 \leftarrow\circ V_2$ is in $\mathcal{P}_{AFCI}$. Now if V_0 is not a parent of K, which means $V_0 \notin \mathbf{AR}(J \circ\!\!\rightarrow K)$ ($V_0 \neq K$ because Y belongs to $\mathbf{AR}(J \circ\!\!\rightarrow K)$ but is not adjacent to V_0 by the definition of discriminating path), then the orientation of $V_0 \circ\text{---}\circ V_1$ (into $V_0 \rightarrow V_1$) is not agreeable ($\mathbf{C}_1$ being violated). So V_0 is a parent of K — which also implies that $Y \neq K$. But then the edge $V_2 \circ\!\!\rightarrow K$ — which is implied to be present in $\mathcal{P}_{AFCI}$ by Lemma 4.3.18 — could be oriented as $V_2 \rightarrow K$ by $\mathcal{R}10$ (because V_0 and Y are not adjacent, and the edge between V_2 and V_0 in $\mathcal{P}_{AFCI}$ constitutes an uncovered p.d. path from V_2 to V_0, and the edge between V_2 and Y constitutes an uncovered p.d. path in $\mathcal{P}_{AFCI}$ from V_2 to Y), a contradiction. So the edge between V_0 and V_1 in $\mathcal{P}_{AFCI}$ is not $V_0 \circ\text{---}\circ V_1$ either.

If the edge is $V_0 \leftarrow\circ V_1$ in $\mathcal{P}_{AFCI}$, the edge between them in $\mathcal{M}$ is $V_0 \leftrightarrow V_1$ by the definition of discriminating path. According to our assumption about bi-directed edges in $\mathcal{M}$, $V_0 \leftarrow\!\circ V_1 \in \mathbf{REL}(J\circ\!\!\rightarrow K)$. It follows that both V_0 and Y are parents of K (neither of them can be K because they are not adjacent by the definition of discriminating path), and $V_1\circ\!\!\rightarrow K$ is present in $\mathcal{P}_{AFCI}$ by Lemma 4.3.18. Furthermore, notice that $V_1 \rightarrow Y$ is in $\mathcal{M}$ by the definition of discriminating path, which means that the edge between V_1 and Y in $\mathcal{P}_{AFCI}$ constitutes an (uncovered) p.d. path from V_1 to Y. Since W and Y are not adjacent, it is easy to see that $V_1\circ\!\!\rightarrow K$ could be oriented as $V_1 \rightarrow K$ by $\mathcal{R}10$, a contradiction. So the edge between V_0 and V_1 in $\mathcal{P}_{AFCI}$ is not $V_0 \leftarrow\!\circ V_1$ either.

It follows that the edge between V_0 and V_1 in $\mathcal{P}_{AFCI}$ is $V_0{*}\!\!\rightarrow V_1$, i.e., the arrowhead at V_1 on this edge is already in $\mathcal{P}_{AFCI}$.

Now we argue that p is also a discriminating path for X in $\mathcal{P}_{AFCI}$. For this purpose, it suffices to show that for every $1 \leq i \leq n-1$, V_i is a collider on p in $\mathcal{P}_{AFCI}$ and is a parent of Y in $\mathcal{P}_{AFCI}$. Consider V_1 for the base case. Since we have shown that $V_0{*}\!\!\rightarrow V_1$ appears in $\mathcal{P}_{AFCI}$, and V_0 is not adjacent to Y, the edge between V_1 and Y is $V_1 \rightarrow Y$ in $\mathcal{P}_{AFCI}$ in virtue of $\mathcal{R}1$. So V_1 is a parent of Y in $\mathcal{P}_{AFCI}$. Suppose for contradiction that V_1 is not a collider on p in $\mathcal{P}_{AFCI}$. Since we have $V_0{*}\!\!\rightarrow V_1$ in $\mathcal{P}_{AFCI}$, and we have $V_1 \leftrightarrow V_2$ in $\mathcal{M}$, the edge between V_1 and V_2 must be $V_1\circ\!\!\rightarrow V_2$ in $\mathcal{P}_{AFCI}$. And by our assumption about bi-directed edges in $\mathcal{M}$, $V_1\circ\!\!\rightarrow V_2 \in \mathbf{REL}(J\circ\!\!\rightarrow K)$. Then Lemma 4.3.18 implies that there is an edge $V_1\circ\!\!\rightarrow K$ in $\mathcal{P}_{AFCI}$. But either $Y = K$ or Y is a parent of K in $\mathcal{P}_{AFCI}$, which implies that V_1 is a parent of K or can be oriented as a parent of K by $\mathcal{R}8$ in $\mathcal{P}_{AFCI}$, a contradiction. So V_1 is a collider on p in $\mathcal{P}_{AFCI}$ and is a parent of Y in $\mathcal{P}_{AFCI}$.

The inductive step can be established by a very similar argument. Suppose for all $1 \leq i < m \leq n-1$, V_i is a collider on p in $\mathcal{P}_{AFCI}$ and is a parent of Y in $\mathcal{P}_{AFCI}$. Consider V_m. By the inductive hypothesis, $\langle V_0, V_1, \cdots, V_m, Y\rangle$ is a discriminating path for V_m in $\mathcal{P}_{AFCI}$, and hence the edge between V_m and Y is $V_m \to Y$ in $\mathcal{P}_{AFCI}$ in virtue of $\mathcal{R}4$. So V_m is also a parent of Y in $\mathcal{P}_{AFCI}$. Suppose for contradiction that V_m is not a collider on p in $\mathcal{P}_{AFCI}$. It follows that either $V_{m-1} \leftarrow\!\circ V_m$ or $V_m \circ\!\!\to V_{m+1}$ appears in $\mathcal{P}_{AFCI}$. Since we have $V_{m-1} \leftrightarrow V_m \leftrightarrow V_{m+1}$ in $\mathcal{M}$, either $V_{m-1} \leftarrow\!\circ V_m$ or $V_m \circ\!\!\to V_{m+1}$ belongs to $\mathbf{REL}(J \circ\!\!\to K)$, by our assumption about bi-directed edges in $\mathcal{M}$. Then Lemma 4.3.18 implies that there is either an edge $V_{m-1} \circ\!\!\to K$ or an edge $V_m \circ\!\!\to K$ in $\mathcal{P}_{AFCI}$. But we have established that in $\mathcal{P}_{AFCI}$ both V_{m-1} and V_m are parents of Y (which in turn is either identical to K or a parent of K). So in either case we can derive a contradiction as we did in the base case. Thus we have established that p is also a discriminating path for X in $\mathcal{P}_{AFCI}$, which implies that the circle at X on $X \circ\!\!\to Y$ could be oriented by $\mathcal{R}4$, a contradiction.

Therefore, all the conditions in Lemma 4.3.9 are met. Thus changing $X \to Y$ to $X \leftrightarrow Y$ will yield a MAG Markov equivalent to $\mathcal{M}$. $\square$

4.4 A Transformational Characterization of Markov Equivalence for Directed MAGs

Markov equivalence between DAGs are characterized in various ways (e.g., Verma and Pearl 1990, Chickering 1995, Andersson et al. 1997), all of which have been found useful for various purposes. In particular, the transformational characterization provided by Chickering (1995) — that two DAGs are Markov equivalent if and only if one can be transformed to the other by a sequence of single edge reversals that preserve DAG-ness and Markov equivalence — is very useful in deriving properties shared by Markov equivalent DAGs. Moreover, when extended to provide a characterization of the submodel relation between DAGs, the transformational property implies the asymptotic correctness of a score-based search procedure over Markov equivalence classes of DAGs, known as the GES algorithm (Meek 1996, Chickering 2002).

MAGs, especially directed MAGs, are a generalization of DAGs, and it is natural to expect that a characterization of Markov equivalence between DAGs can be extended to characterize Markov equivalence between MAGs. The Verma-Pearl style characterization has indeed been generalized to MAGs (Spirtes and Richardson 1996, Ali et al. 2004), of which Proposition 3.1.2 provides an example. But Chickering's transformational characterization has not seen an extension to MAGs. We fill this gap below for *directed* MAGs (DMAGs). Specifically, we show that two DMAGs are Markov equivalent if and only if one can be transformed to the other by a sequence of single *mark changes* that preserve DMAG-ness and Markov equivalence. The reason we present this transformational characterization here is that the argument is going to depend crucially on a few results established earlier in this chapter.

One lemma that immediately stands out as relevant to this purpose is Lemma 4.3.9, which gives sufficient and necessary conditions for a kind of mark change — turning a directed edge into a bi-directed edge — to preserve Markov equivalence of MAGs. Obviously the same set of conditions also specifies when turning a bi-directed edge into a directed edge preserves Markov equivalence. Notice furthermore that if we only consider DMAGs, MAGs with no undirected edges, any single mark change is either to turn a directed edge into a bi-directed or to turn a bi-directed edge into a directed one. So Lemma 4.3.9 encompasses all possible kinds of single mark changes within the class of DMAGs. We rewrite the lemma here for easy reference:

Lemma 4.4.1. *Let $\mathcal{G}$ be an arbitrary DMAG, and $A \to B$ an arbitrary directed edge in $\mathcal{G}$. Let $\mathcal{G}'$ be the graph identical to $\mathcal{G}$ except that the edge between A and B is $A \leftrightarrow B$. (In other words, $\mathcal{G}'$ is the result of simply changing the mark at A on $A \to B$*

from an tail into an arrowhead.) $\mathcal{G}'$ *is a DMAG and Markov equivalent to* $\mathcal{G}$ *if and only if*

(t1) there is no directed path from A *to* B *other than* $A \to B$*;*

(t2) For any $C \to A$ *in* $\mathcal{G}$, $C \to B$ *is also in* $\mathcal{G}$*; and for any* $D \leftrightarrow A$ *in* $\mathcal{G}$*, either* $D \to B$ *or* $D \leftrightarrow B$ *is in* $\mathcal{G}$*;*

(t3) There is no discriminating path for A *on which* B *is the endpoint adjacent to* A.

Proof. Special case of Lemma 4.3.9. □

As we said, this lemma, by symmetry, also gives conditions for dropping an arrowhead from a bi-directed edge while preserving Markov equivalence. (The first condition in Lemma 4.3.9 is not needed as we are only concerned with DMAGs.)

We say a mark change in a DMAG is *legitimate* when the conditions in Lemma 4.4.1 are satisfied. In Chickering's result for DAGs, the basic unit of equivalence-preserving transformation is (covered) edge reversal. Here we regard an edge reversal as simply a special case of two consecutive mark changes. That is, a reversal of $A \to B$ is simply to first add an arrowhead at A (to form $A \leftrightarrow B$), and then to drop the arrowhead at B (to form $A \leftarrow B$). An edge reversal is said to be legitimate if both of the two consecutive mark changes are legitimate. Given Lemma 4.4.1, it is straightforward to check the validity of the following condition for legitimate edge reversal. (We use $\mathbf{Pa}_{\mathcal{G}}/\mathbf{Sp}_{\mathcal{G}}$ to denote the set of parents/spouses of a vertex in $\mathcal{G}$.)

Lemma 4.4.2. *Let* $\mathcal{G}$ *be an arbitrary DMAG, and* $A \to B$ *an arbitrary directed edge in* $\mathcal{G}$*. The reversal of* $A \to B$ *is legitimate if and only if* $\mathbf{Pa}_{\mathcal{G}}(B) = \mathbf{Pa}_{\mathcal{G}}(A) \cup \{A\}$ *and* $\mathbf{Sp}_{\mathcal{G}}(B) = \mathbf{Sp}_{\mathcal{G}}(A)$.

When there is no bi-directed edge in $\mathcal{G}$, that is, when $\mathcal{G}$ is a DAG, the condition in Lemma 4.4.2 is reduced to the familiar definition for *covered edge*, i.e., $\mathbf{Pa}_{\mathcal{G}}(B) = \mathbf{Pa}_{\mathcal{G}}(A) \cup \{A\}$ (Chickering 1995). The condition given by Drton and Richardson (2004) for a bi-directed edge in a bi-directed graph to be orientable as a directed edge in either direction ($\mathbf{Sp}_{\mathcal{G}}(B) = \mathbf{Sp}_{\mathcal{G}}(A)$) can be viewed as another special case of the above lemma.

Another result relevant to our current purpose is Corollary 4.3.8. Confined to DMAGs, it becomes the following proposition:

Proposition 4.4.1. *Given any DMAG $\mathcal{G}$, there exists a DMAG $\mathcal{H}$ such that*

(1) $\mathcal{H}$ is Markov equivalent to $\mathcal{G}$;

(2) every bi-directed edge in $\mathcal{H}$ is invariant;

(3) every directed edge in $\mathcal{G}$ is also in $\mathcal{H}$.

Proof. Special case of Corollary 4.3.8. □

We will call $\mathcal{H}$ in Proposition 4.4.1 a **Loyal Equivalent Graph (LEG)** of $\mathcal{G}$. In general a DMAG could have multiple LEGs. A distinctive feature of the LEGs is that they have the fewest bi-directed edges among the Markov equivalent DMAGs. Drton and Richardson (2004) explored the statistical significance of this fact for fitting bi-directed graphs, graphs that contain only bi-directed edges. They showed, roughly speaking, that if the LEGs of a bi-directed graph are DAGs, then fitting is easy; otherwise fitting is not easy (in a specific technical sense).

Another feature which is particularly relevant to our argument is that between a DMAG and any of its LEGs, only one kind of difference is possible, namely, some bi-directed edges in the DMAG are oriented as directed edges in its LEG. For a simple illustration, compare the graphs in Figure 4.3, where H1 is a LEG of G1, and H2 is a LEG of G2.

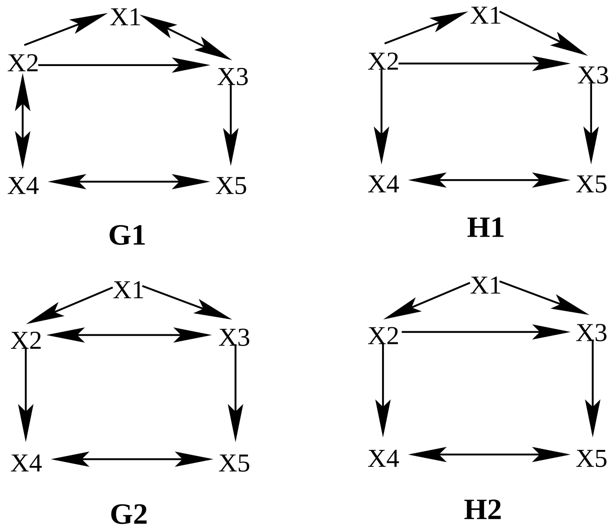

Figure 4.3: A LEG of G1 (H1) and a LEG of G2 (H2)

A directed edge in a DMAG is called **reversible** if there is another Markov equivalent DMAG in which the direction of the edge is reversed. We use the fact that **CP**1 holds of a CPAG to establish the next helpful proposition, which concerns reversible directed edges and invariant bi-directed edges in a DMAG.

Proposition 4.4.2. *Let $A \to B$ be any reversible edge in a DMAG $\mathcal{G}$. For any vertex C (distinct from A and B), there is an invariant bi-directed edge between C and A if and only if there is an invariant bi-directed edge between C and B.*

Proof. Since $A \to B$ is reversible, which means neither of the two marks of the edge is invariant, in $\mathcal{P}_\mathcal{G}$ — the CPAG of $\mathcal{G}$ — the edge between A and B would be $A \circ\!\!-\!\!\circ B$. For any C, if there is an invariant bi-directed edge between C and A in $\mathcal{G}$, $C \leftrightarrow A$ would also appear in $\mathcal{P}_\mathcal{G}$. Because **CP**1 holds of $\mathcal{P}_\mathcal{G}$, $C \leftrightarrow B$ must also appear in $\mathcal{P}$, and hence is an invariant bi-directed edge in $\mathcal{G}$. Conversely, if there is an invariant bi-directed edge between C and B in $\mathcal{G}$, the same argument shows that there would also be an invariant bi-directed edge between C and A in $\mathcal{G}$. $\square$

In particular, if $\mathcal{H}$ is a LEG of a DMAG, then $A \to B$ being reversible implies that A and B have the same set of spouses, as every bi-directed edge in $\mathcal{H}$ is invariant.

Now we have all we need to establish the transformational characterization. We first state two intermediate theorems. The first one says if all of the differences between two Markov equivalent DMAGs $\mathcal{G}$ and $\mathcal{G}'$ are of the following sort: that a directed edge is in $\mathcal{G}$ while the corresponding edge is bi-directed in $\mathcal{G}'$, then there is a sequence of legitimate mark changes that transforms one to the other. The second one says that if every bi-directed edge in $\mathcal{G}$ and every bi-directed edge in $\mathcal{G}'$ is invariant, then there is a sequence of legitimate mark changes (edge reversals) that transforms one to the other.

Theorem 4.4.1. *Let $\mathcal{G}$ and $\mathcal{G}'$ be two Markov equivalent DMAGs. If every bi-directed edge in $\mathcal{G}$ is also in $\mathcal{G}'$, and every directed edge in $\mathcal{G}'$ is also in $\mathcal{G}$, then there is a sequence of legitimate mark changes that transforms one to the other.*

Proof. We prove that there is a sequence of transformation from $\mathcal{G}$ to $\mathcal{G}'$, the reverse of which will be a transformation from $\mathcal{G}'$ to $\mathcal{G}$. Specifically we show that as long as $\mathcal{G}$ and $\mathcal{G}'$ are different, there is always a legitimate mark change that can eliminate a difference between them. The theorem then follows from a simple induction on the number of differences.

The antecedent of the theorem implies that the differences between $\mathcal{G}$ and $\mathcal{G}'$ are all of the same sort: a directed edge ($\to$) is in $\mathcal{G}$ while the corresponding edge in $\mathcal{G}'$ is bi-directed ($\leftrightarrow$). Let

$$\mathbf{Diff} = \{y|\text{ there is a } x \text{ such that } x \to y \text{ is in } \mathcal{G} \text{ and } x \leftrightarrow y \text{ is in } \mathcal{G}'\}$$

It is clear that $\mathcal{G}$ and $\mathcal{G}'$ are identical if and only if $\mathbf{Diff} = \emptyset$. We claim that if $\mathbf{Diff}$ is not empty, there is a legitimate mark change that eliminates a difference. Choose $B \in \mathbf{Diff}$ such that no proper ancestor of B in $\mathcal{G}$ is in $\mathbf{Diff}$. Let

$$\mathbf{Diff}_B = \{x | x \to B \text{ is in } \mathcal{G} \text{ and } x \leftrightarrow B \text{ is in } \mathcal{G}'\}$$

Since $B \in \mathbf{Diff}$, $\mathbf{Diff}_B$ is not empty. Choose $A \in \mathbf{Diff}_B$ such that no proper descendant of A in $\mathcal{G}$ is in $\mathbf{Diff}_B$. The claim is that changing $A \to B$ to $A \leftrightarrow B$ in $\mathcal{G}$ is a legitimate mark change.

To see this is so, let us verify the conditions stated in Lemma 4.4.1. First, suppose condition (t1) is violated, that is, suppose there is another directed path $d = \langle A, \cdots, C, B \rangle$ from A to B besides $A \to B$. d is not present in $\mathcal{G}'$, otherwise $\mathcal{G}'$ is not a MAG due to the presence of $A \leftrightarrow B$. So some edge on d in $\mathcal{G}'$ must be bi-directed. If the edge is $C \leftrightarrow B$, then C belongs to $\mathbf{Diff}_B$, but is a proper descendant of A in $\mathcal{G}$, which contradicts our choice of A. If the edge is between another pair of vertices, say $D \leftrightarrow E$ (s.t. $D \to E$ is in $\mathcal{G}$), then E is in $\mathbf{Diff}$, but is a proper ancestor of B in $\mathcal{G}$, which contradicts our choice of B. So can't be any directed path from A to B in $\mathcal{G}$ other than $A \to B$. Condition (t1) holds.

Next we check condition (t2). For the first part, let C be any parent of A in $\mathcal{G}$. C must also be a parent of A in $\mathcal{G}'$, otherwise A is in $\mathbf{Diff}$, but is a proper ancestor of B in $\mathcal{G}$, which contradicts our choice of B. It follows that C and B are adjacent, for otherwise $\langle C, A, B \rangle$ is an unshielded collider in $\mathcal{G}'$ but not in $\mathcal{G}$, contrary to the assumption that they are Markov equivalent. Then C must be a parent of B in $\mathcal{G}$, otherwise $\mathcal{G}$ is not ancestral.

For the second part, let D be any spouse of A (i.e., $D \leftrightarrow A$) in $\mathcal{G}$. D is also a spouse of A in $\mathcal{G}'$ by our assumption. It follows that D and B are adjacent, for otherwise $\langle D, A, B \rangle$ is an unshielded collider in $\mathcal{G}'$ but not in $\mathcal{G}$. But D cannot be a child of B in $\mathcal{G}$, for otherwise $\mathcal{G}$ is not ancestral. Hence D is either a parent or a spouse of B.

Finally, suppose condition (t3) is violated, that is, suppose there is a discriminating path $u = \langle U, \cdots, V, A, B \rangle$ for A. By the definition of discriminating path, V is a parent of B. It follows that the edge between A and V is not $A \rightarrow V$, for otherwise $A \rightarrow V \rightarrow B$ would be a directed path from A to B, which has been shown to be absent. Hence the edge between V and A must be bi-directed, $V \leftrightarrow A$. Further note that by our assumption about the difference between $\mathcal{G}$ and $\mathcal{G}'$, every arrowhead in $\mathcal{G}$ is also in $\mathcal{G}'$, which implies that every collider in $\mathcal{G}$ is also in $\mathcal{G}'$. In particular, every vertex between U and A on u is also a collider on u in $\mathcal{G}'$.

Now we prove by induction that every vertex between U and A on u, including A, is a parent of B in $\mathcal{G}'$, contradicting the fact that $A \leftrightarrow B$ is in $\mathcal{G}'$. Let W be the vertex next to U on u. Since U and B are not adjacent by the definition of discriminating path, $\langle U, W, B \rangle$ is an unshielded non-collider in $\mathcal{G}$ (because W is a parent of B in $\mathcal{G}$ by the definition of discriminating path). Because $\mathcal{G}$ and $\mathcal{G}'$ are Markov equivalent, $\langle U, W, B \rangle$ should also be a non-collider in $\mathcal{G}'$. But W is a collider on u in $\mathcal{G}$, and hence also a collider in $\mathcal{G}'$, which means the edge between U and W is into W. Thus $W \rightarrow B$ is in $\mathcal{G}'$, otherwise $\langle U, W, B \rangle$ would be an unshielded collider in $\mathcal{G}'$. This establishes the base case. In the inductive step, suppose the first n vertices after U on u are all parents of B in $\mathcal{G}'$, then we have a discriminating path for the $n+1$'st vertex between D and B in both graphs. Since the two graphs are Markov equivalent, the $n+1$'st vertex must be a parent of B as well, otherwise (e3) in Proposition 3.1.2 would be violated. This finishes our induction. So, in particular, A should be a parent of B in $\mathcal{G}'$, a contradiction. Thus condition (t3) also obtains.

Therefore, we can always identify a legitimate mark change to eliminate a difference as long as $\mathcal{G}$ and $\mathcal{G}'$ are still different. An induction on the number of differences between $\mathcal{G}$ and $\mathcal{G}'$ would do to complete the argument. □

Obviously a DMAG and any of its LEGs satisfy the antecedent of Theorem 4.4.1, so they can be transformed to each other by a sequence of legitimate mark changes. Steps 0-2 in Figure 4.4, for example, portrait a stepwise transformation from G1 to H1.

Theorem 4.4.2. *Let $\mathcal{G}$ and $\mathcal{G}'$ be two Markov equivalent MAGs. If every bi-directed edge in $\mathcal{G}$ and every bi-directed edge in $\mathcal{G}'$ is invariant, then there is a sequence of legitimate mark changes that transforms one to the other.*

Proof. Without loss of generality, we prove that there is a transformation from $\mathcal{G}$ to $\mathcal{G}'$. It follows from the assumption that $\mathcal{G}$ and $\mathcal{G}'$ have the same set of bi-directed edges, and hence all differences between $\mathcal{G}$ and $\mathcal{G}'$ are of the same sort: $\rightarrow$ is in $\mathcal{G}$, while $\leftarrow$ is in $\mathcal{G}'$. Let

$$\mathbf{Diff} = \{y | \text{ there is a } x \text{ such that } x \rightarrow y \text{ is in } \mathcal{G} \text{ and } x \leftarrow y \text{ is in } \mathcal{G}'\}$$

Clearly $\mathcal{G}$ and $\mathcal{G}'$ are identical if and only if $\mathbf{Diff} = \emptyset$. We claim that if $\mathbf{Diff}$ is not empty, we can always identify a legitimate edge reversal (that is, two legitimate mark changes in a row) that eliminates a difference in direction.

Suppose $\mathbf{Diff}$ is not empty. We can choose a vertex $B \in \mathbf{Diff}$ such that no proper ancestor of B in $\mathcal{G}$ is in $\mathbf{Diff}$. Let

$$\mathbf{Diff}_B = \{x | x \rightarrow B \text{ is in } \mathcal{G} \text{ and } x \leftarrow B \text{ is in } \mathcal{G}'\}$$

Since $B \in \mathbf{Diff}$, $\mathbf{Diff}_B$ is not empty. Choose $A \in \mathbf{Diff}_B$ such that no proper descendant of A in $\mathcal{G}$ is in $\mathbf{Diff}_B$. Then changing $A \rightarrow B$ to $A \leftarrow B$ in $\mathcal{G}$ is a legitimate edge reversal.

To justify this claim, we verify the conditions in Lemma 4.4.2. Note that $A \to B$, by our choice, is a reversible edge in $\mathcal{G}$ (for $A \leftarrow B$ is in $\mathcal{G}'$, which is Markov equivalent to $\mathcal{G}$). It thus follows directly from Proposition 4.4.2 (and the assumption about bi-directed edges in $\mathcal{G}$) that $\mathbf{Sp}_{\mathcal{G}}(B) = \mathbf{Sp}_{\mathcal{G}}(A)$.

The argument for $\mathbf{Pa}_{\mathcal{G}}(B) = \mathbf{Pa}_{\mathcal{G}}(A) \cup \{A\}$ is virtually the same as Chickering's proof for DAGs. For any parent C of A in $\mathcal{G}$, C is also a parent of A in $\mathcal{G}'$, otherwise A is in **Diff** and is a proper ancestor of B in $\mathcal{G}$, which contradicts our choice of B. It follows that C is adjacent to B, otherwise $\langle C, A, B\rangle$ is an unshielded collider in $\mathcal{G}'$ but not one in $\mathcal{G}$, which would contradict the Markov equivalence between $\mathcal{G}$ and $\mathcal{G}'$. Then C must be a parent of B in $\mathcal{G}$, otherwise $\mathcal{G}$ is not ancestral. Conversely, let $D \neq A$ be any other parent of B in $\mathcal{G}$. D must be adjacent to A, otherwise $\langle D, B, A\rangle$ is an unshielded collider in $\mathcal{G}$ but not one in $\mathcal{G}'$. D is not a spouse of A in $\mathcal{G}$, otherwise D is a spouse of B, according to what we just showed. So D is either a parent or a child of A in $\mathcal{G}$. Suppose it is a child of A, that is, $A \to D \to B$ is in $\mathcal{G}$. We derive a contradiction from this. Since $A \leftarrow B$ is in $\mathcal{G}'$, $A \to D \to B$ does not appear in $\mathcal{G}'$. That means either $A \leftarrow D$ or $D \leftarrow B$ (or both) is in $\mathcal{G}'$. In the former case, D is in **Diff** and is a proper ancestor of B in $\mathcal{G}$, which contradicts our choice of B. In the latter case, D is in $\mathbf{Diff}_B$ and is a proper descendant of A in $\mathcal{G}$, which contradicts our choice of A. Hence D can't be a child of A in $\mathcal{G}$, which means it is a parent of A in $\mathcal{G}$.

Note that after an edge reversal, no new bi-directed edge is introduced, so the assumption that every bi-directed edge is invariant still holds for the new graph. Hence we can always identify a legitimate edge reversal to eliminate a difference in direction as long as $\mathcal{G}$ and $\mathcal{G}'$ are still different. An easy induction on the number of differences between $\mathcal{G}$ and $\mathcal{G}'$ would do to complete the argument. $\square$

Since a LEG (of any MAG) only contains invariant bi-directed edges, two LEGs that are Markov equivalent can always be transformed to each other via a sequence of legitimate mark changes according to the above theorem. For example, steps 2-4 in Figure 4.4 constitute a transformation from H1 (a LEG of G1) to H2 (a LEG of G2).

We are ready to prove the main result of this section.

Theorem 4.4.3. *Two DMAGs $\mathcal{G}$ and $\mathcal{G}'$ are Markov equivalent if and only if there exists a sequence of single mark changes in $\mathcal{G}$ such that*

1. *after each mark change, the resulting graph is also a DMAG and is Markov equivalent to $\mathcal{G}$;*

2. *after all the mark changes, the resulting graph is $\mathcal{G}'$.*

Proof*:* The "if" part is trivial – since every mark change preserves the equivalence, the end is of course Markov equivalent to the beginning. Now suppose $\mathcal{G}$ and $\mathcal{G}'$ are equivalent. We show that there exists such a sequence of transformation. By Proposition 4.4.1, there is a LEG $\mathcal{H}$ for $\mathcal{G}$ and a LEG $\mathcal{H}'$ for $\mathcal{G}'$. By Theorem 4.4.1, there is a sequence of legitimate mark changes s_1 that transforms $\mathcal{G}$ to $\mathcal{H}$, and there is a sequence of legitimate mark changes s_3 that transforms $\mathcal{H}'$ to $\mathcal{G}'$. By Theorem 4.4.2, there is a sequence of legitimate mark changes s_2 that transforms $\mathcal{H}$ to $\mathcal{H}'$. Concatenating s_1, s_2 and s_3 yields a sequence of legitimate mark changes that transforms $\mathcal{G}$ to $\mathcal{G}'$. □

As a simple illustration, Figure 4.4 gives the steps in transforming G1 to G2 according to Theorem 4.4.3. That is, G1 is first transformed to one of its LEGs, H1; H1 is then transformed to H2, a LEG of G2. Lastly, H2 is transformed to G2.

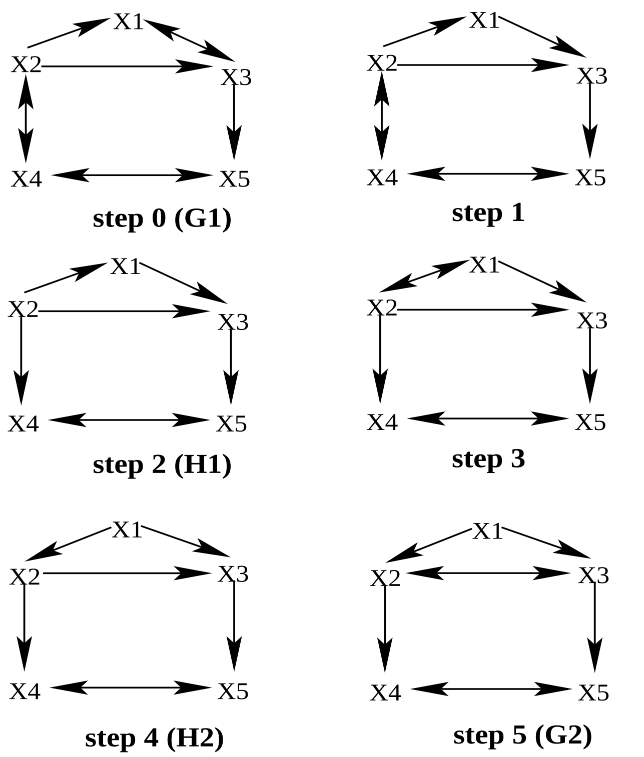

Figure 4.4: A transformation from G1 to G2

Theorems 4.4.1 and 4.4.2, as they are currently stated, are special cases of Theorem 4.4.3, but the proofs of them actually achieve a little more than what they claim. The transformations constructed in the proofs of Theorems 1 and 2 are efficient in the sense that every mark change in the transformation eliminates a difference between the current DMAG and the target. So the transformations consist of as many mark changes as the number of differences at the beginning. By contrast, the transformation constructed in Theorem 4.4.3 may take some "detours", in that some mark changes in the way actually increase rather than decrease the difference between $\mathcal{G}$ and $\mathcal{G}'$. (This is not the case in Figure 4.4, but if, for example, we chose different LEGs for G1 or G2, there would be detours.) We believe that no such detour is really necessary, that is, there is always a transformation from $\mathcal{G}$ to $\mathcal{G}'$ consisting of as many mark changes as the number of differences between them. But we are yet unable to prove this conjecture.

This transformational characterization of Markov equivalence implies that no matter how different two Markov equivalent DMAGs are, there is a sequence of Markov equivalent DMAGs in between such that the adjacent graphs differ in only one edge (mark). It could thus simplify derivations of invariance properties across a Markov equivalence class: in order to show two arbitrary Markov equivalent DMAGs share something in common, we only need to consider two Markov equivalent DMAGs with the minimal difference. Indeed, Chickering (1995) used his characterization to derive that Markov equivalent DAGs have the same number of parameters under the standard CPT parameterization (and hence would receive the same score under the typical penalized-likelihood type metrics). The discrete parameterization of DMAGs is currently under development[8]. Our results will probably be useful in showing similar facts once the discrete parameterization is available.

The characterization, however, does not hold exactly for general MAGs. A simple counterexample is given in Figure 4.5. Recall that when we include undirected edges, the requirement of ancestral graphs is that the endpoints of undirected edges are of zero in-degree. So, although the two graphs in Figure 4.5 are Markov equivalent MAGs, M1 cannot be transformed to M2 by a sequence of single legitimate mark changes, as adding any single arrowhead to M1 would make it non-ancestral. Therefore, for general MAGs, the transformation may have to include a stage of changing the undirected subgraph to a directed one in a wholesale manner.[9]

[8] Drton and Richardson (2005) provide a parameterization for bi-directed graphs with binary variables, for which the problem of parameter equivalence does not arise because no two different bi-directed graphs are Markov equivalent.

[9] As suggested by Thomas Richardson, two possible solutions are (1) to allow replacing a tail or an arrowhead with a circle temporarily, and (2) to consider Joined Graphs introduced by Ali (2002), which form a superclass of MAGs.

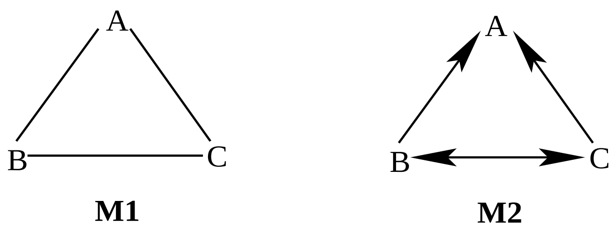

Figure 4.5: A simple counterexample with general MAGs: M1 can't be transformed into M2 by a sequence of legitimate single mark changes.

The transformational characterization for Markov equivalent DAGs was generalized, as a conjecture, to a transformational characterization for DAG I-maps by Meek (1996), which was later shown to be true by Chickering (2002). A graph is an I-map of another if the set of conditional independence relations entailed by the former is a subset of the conditional independence relations entailed by the latter. This generalized transformational property is used to prove the asymptotic correctness of the GES algorithm, a relatively efficient score-based search procedure over the Markov equivalence classes of DAGs. The reason we spent so much energy on the (augmented) FCI algorithm, an independence-constraint-based inference procedure, but did not say anything about the score-based approach in these two chapters is partly because no provably correct and feasible score-based procedure for causal inference in causally insufficient systems is yet available.[10] We expect that the GES algorithm can be extended to search over CPAGs[11], and we expect that a generalization of the transformational property just established will be useful in justifying that algorithm, but as of now these problems are open.

[10]There is obviously an utterly intractable procedure that is provably correct if a consistent score is available, i.e., to enumerate and score all possible CPAGs in the search space with a consistent scoring metric, and choose the one with the highest score.

[11]For a preliminary study, see Spirtes et al. 1997.

Chapter 5

Quantitative Reasoning with a CPAG — Prediction of Intervention Effects

An important practical reason for people to care about causation or causal explanation is the need to predict effects of actions or interventions before actually carrying them out. Sometimes we base that kind of prediction on past similar interventions or experiments, in which case Hume's Principle of Custom may be at work to prompt our projection of past experimental findings into the future without, perhaps, explicit or conscious reasoning about cause and effect. Other times, however, we do not have access to much or any controlled experimental studies for various reasons, and all we have are observations of a system before interventions or manipulations take place. In these situations, there is no simple projection of the past experience into a supposedly similar future, as the prediction is precisely about the consequences of certain changes. What we need is to go from phenomena generated from one causal process to phenomena that would be generated if the causal process were intervened to become a different one.

Let us formalize this a little bit in the language we have been using. Given a set of random variables $\mathbf{V}$ whose causal structure is properly represented by a DAG – which means at least that the set is causally sufficient and there is no feedback mechanism – we are interested in the outcome of actively controlling some variables in the system. Specifically, we may be interested in the probability distribution of some variables $\mathbf{Y}$ (possibly conditional on some other variables $\mathbf{Z}$) if a variable X were manipulated[1] to take some value in some way. Let us further assume for simplicity that we know the probability distribution of $\mathbf{V}$ when X is not manipulated (which is acceptable if, for example, every variable is observed and the number of observations is large). Is there a link between the pre-intervention probability and the post-intervention probability?

For all we know, that link needs to be assumed, or derived from certain assumptions about the intervention in question. A most common assumption is that ideally an intervention is effective and local (with no side effects). When we talk about an intervention of a variable X, we mean, among other things, that the *direct* target of the intervention is X. *Effectiveness* means that the value of X or the probability distribution of X — and if the intervention is supposed to depend on some other variables, what variables to depend upon — are completely fixed by the intervention. Since the intended effect of an intervention on its direct target is usually known, the assumption of effectiveness immediately gives us the post-intervention (conditional) probability of X. The assumption of *locality*, on the other hand, provides with regard to other variables the link between pre-intervention circumstances and post-intervention circumstances. Specifically, it requires that the intervention should not *directly* affect any variable other than the direct target, and more importantly, local mechanisms

[1] In this chapter we use *intervention* and *manipulation* interchangeably.

for other variables should remain the same as before the intervention.[2] Thus, the intervention is merely a local surgery with respect to causal mechanisms. Remote changes that occur to other variables after the intervention are due to propagation via the original causal mechanisms unaffected by the intervention. In this chapter we deal with only interventions of this kind, which we will refer to as **EL (Effective and Local) interventions**.[3]

A formal implementation of these two requirements is given by econometricians, most notably Strotz and Wold (1960) and Fisher (1970), and is nicely recounted in Pearl (2000). A causal system represented by a DAG can also be represented by a set of structural equations, in which each variable is equated with a function of its parents in the DAG and an error term. The equations are "structural" in that they represent mechanisms with causal direction that do not admit ordinary algebraic transformations. Then an EL intervention on X can be simply implemented by replacing the original equation that defines the mechanism for X with a new equation introduced by the intervention and – this is the important part – leaving all the other equations unchanged. In the simplest case where X is manipulated to a fixed value, the equation for X is simply "wiped out" (or replaced by an uninteresting equation $X = c$) but all other equations remain the same.

Graphically, the above operation amounts to erasing all arrows into X in the causal DAG, and possibly putting some of them back — arrows from those of X's parents

[2]The talk of local mechanisms assumes something called *modularity*. Modularity roughly means that each variable in the causal system is associated with a local mechanism which can be independently manipulated, i.e., manipulated without affecting other mechanisms (Woodward 2003).

[3]This restriction is surely a limitation, but just how serious this limitation is needs careful reflection. One can immediately complain that interventions in real life, for example policy interventions, almost always encounter deviants and almost always have side effects. But what this says, on my view, is simply that sometimes we do not know what the actual intervention is, because to individualize an intervention requires specifying what it does and what it does not affect directly. It is no embarrassment for any theory of intervention to concede that the consequence of an intervention is never predictable without knowing what the intervention is in the first place.

that are conditioned on in the intervention. So in general an intervention may depend on some direct causes of X in the original causal system, but nullify other ones. The resulting graph is usually called the *manipulated DAG*. We use $P_{post}(X|\mathbf{Pa}(X))$ to denote the (post-intervention) conditional distribution of X imposed by an intervention or manipulation, where $\mathbf{Pa}(X)$ (possibly empty) is assumed to be a subset of the set of direct causes of X before the intervention.[4] So an intervention, as concerns us here, can only eliminate some existing causal links, but does not add any. This implies that the manipulated DAG is a subgraph of the unmanipulated causal DAG. Lastly, an intervention on a set of variables $\mathbf{X}$ is usually assumed to consist of independent interventions[5] on the individual variables so that

$$P_{post}(\mathbf{X}|\mathbf{Pa}(\mathbf{X})) = \prod_{X \in \mathbf{X}} P_{post}(X|\mathbf{Pa}(X))$$

Based on this understanding of interventions, several authors worked out independently the fundamental link between the pre-intervention probability and the post-intervention probability (Robins 1986, Spirtes et al. 1993/2000, Pearl 2000). Here we give the formulation in Spirtes et al. (1993/2000), but instead of calling it the Manipulation Theorem we will call it the *Manipulation Principle*, as to a large extent it is a formal restatement of the restrictions on an EL intervention.

[4] This restriction can be relaxed. Intuitively, an intervention of X can depend on any variable that temporally precedes X. This suggests that formally an intervention of X can condition upon any non-descendant (rather than just parent) of X in the pre-intervention causal graph. For such more general cases it is reasonable to posit the exact same manipulation principle to be stated shortly, except that in these cases $\mathbf{Pa}(X)$ in the post-intervention causal graph may contain variables that are not direct causes of X prior to the intervention. In other words, a conditional intervention of X may not only delete some edges into X, but also add some edges into X (that were not present). The theory of invariance developed in section 5.1 is based on the assumption that conditional interventions of X will depend on only (a subset of) pre-intervention direct causes (parents) of X, which guarantees that post-intervention causal graphs are subgraphs of the pre-intervention causal graph. Without this assumption the theory has to be modified, but not dramatically.

[5] This of course is merely an assumption for convenience. In general we can have correlated interventions as well.

Manipulation Principle Given a causal DAG over $\mathbf{V}$ and a (pre-intervention) joint distribution that factorizes according to the graph (i.e., is Markov to the graph), the joint distribution of $\mathbf{V}$ after an intervention on $\mathbf{X}$ takes a similar form of factorization, as follows:

$$P_{post}(\mathbf{V}) = \prod_{X \in \mathbf{X}} P_{post}(X|\mathbf{Pa}(X)) \prod_{Y \in \mathbf{V} \backslash \mathbf{X}} P_{pre}(Y|\mathbf{Pa}(Y))$$

Where P_{post} denotes the post-intervention probability, P_{pre} denotes the pre-intervention probability, and $\mathbf{Pa}$ denotes the parent set in the manipulated graph (which, for $Y \in \mathbf{V} \backslash \mathbf{X}$, is the same as the parent set in the original graph).

Observe that every term in the above factorization is assumed to be known: the pre-intervention probabilities can be consistently estimated from observational data, and the post-intervention probabilities of $\mathbf{X}$ are assumed to be given. So the full joint distribution after intervention is predictable, given a causal DAG and a pre-intervention probability over $\mathbf{V}$. From the joint distribution we can calculate $P_{post}(\mathbf{Y}|\mathbf{Z})$ for any $\mathbf{Y}$ and $\mathbf{Z}$.

Complications come in two ways. First, our central concern in this dissertation is that the set of observed variables may be causally insufficient. Hence there may be variables that are causally relevant but are unobserved, or even worse, are unobservable. So even if the causal DAG (with latent variables) is fully known, we may not be able to predict certain intervention effects because the pre-intervention probability that we have access to is merely the marginal probability of the observed variables instead of the joint probability of all variables in the DAG. (Much of Pearl's work (1995, 1998, 2000), and more recently Tian and Pearl (2004) are paradigmatic attempts to deal with this situation.) Second, the relevant causal structure is seldom, if ever, fully known. Usually we have to infer the causal structure from observational data, and the best we can hope is to discover some features of the true causal graph. So even if a system of observed variables is known to be causally sufficient, we may not be able to predict certain intervention effects due to the insufficiency of the causal information that can be inferred from data (see Spirtes et al. 1993/2000). The situation, of course, becomes even more involved when the two complications go together, which is our major concern here.

A general formulation of the problem is this: We observe a set of variables $\mathbf{O}$, and we assume the pre-intervention probability distribution of $\mathbf{O}$ is known. $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$ are three subsets of $\mathbf{O}$ ($\mathbf{X} \cap \mathbf{Y} = \mathbf{Y} \cap \mathbf{Z} = \emptyset$). The quantity of interest is $P_{post}(\mathbf{Y}|\mathbf{Z})$ when $\mathbf{X}$ is subject to a certain EL intervention. Suppose there is no selection effect (so that the causal MAG is directed), but we do not know if $\mathbf{O}$ is causally sufficient. All we know is that $\mathbf{O}$, if not causally sufficient by itself, can be extended to a causally sufficient system $\mathbf{O} \cup \mathbf{L}$ whose causal structure is properly represented by a DAG. Of course we do not know the structure of the DAG. In fact we do not even know what those latent variables are. The hard question is whether $P_{post}(\mathbf{Y}|\mathbf{Z})$ is uniquely determined by the pre-intervention probability distribution of $\mathbf{O}$ together with the specifics of the intervention, and if so, how to calculate it in terms of the pre-intervention probability.

The problem is undoubtedly formidable, but not entirely hopeless. One way to proceed is not hard to see given what we have done in the last two chapters. Assuming the CMC and the CFC, we can infer features of the true causal DMAG (directed MAG) from the pre-intervention condition independence relations, neatly summarized in a CPAG (complete partial ancestral graph), which also represents the uncertainty there is about the true DMAG. Given the CPAG[6], the initial question is reduced

[6]Remember this means that we restrict ourselves to inferences based on conditional independence

to the following: is $P_{post}(\mathbf{Y}|\mathbf{Z})$ uniquely identifiable in terms of the pre-intervention probability $P_{pre}(\mathbf{O})$ given the CPAG?

How can this reduction help? After all the manipulation principle, the only thing we can rely on to calculate post-intervention quantities, is formulated in terms of a causal DAG. But a CPAG is far from a causal DAG. It actually represents a set of Markov equivalent DMAGs, and each DMAG, in turn, is compatible with an (infinite) number of causal DAGs with latent variables. For each of the DAGs, the quantity may or may not be calculable in terms of the probability distribution over observed variables according to the manipulation principle. And even if the quantity is calculable relative to each compatible DAG, different DAGs may give conflicting answers.

We will not try to contribute to the literature on conflict resolution, so we insist on the unanimity rule: a post-intervention quantity is identifiable given a CPAG only if it is identifiable given any causal DAG compatible with the CPAG and every such DAG gives the same answer regarding the quantity. The task is to figure out when this unanimity condition holds, without, of course, doing the impossible job of checking all compatible DAGs one by one.

This chapter by no means intends to solve this problem completely. Nor can it claim originality in providing partial solutions. The two pieces of work to be presented are either an improvement or a generalization of earlier work. Specifically, in section 5.1, we give graphical criteria for what is called *invariance* given a CPAG. The result is in most respects parallel to the theory of invariance developed by Peter Spirtes, Clark Glymour and Richard Scheines (1993/2000). However, their results are formulated with respect to a *partially oriented inducing path graph* instead of a CPAG (cf. Appendix), and their criteria are only sufficient but not necessary. By

relations.

contrast, the condition we will present is both sufficient and necessary for invariance given a CPAG. In section 5.2, we generalize Judea Pearl's celebrated *do-calculus* so that the resulting calculus is based on a CPAG rather than a single causal DAG with latent variables. Most proofs are postponed to 5.4.

5.1 Invariance Given a CPAG

A post-intervention probability is identifiable (relative to some causal information) if it can be expressed as a function of the pre-intervention probability and the known post-intervention probability of the manipulated variables. The most basic identifiable quantities are those that remain the same before and after intervention. A conditional probability $P(\mathbf{Y}|\mathbf{Z})$ is said to be **invariant** under an intervention of $\mathbf{X}$ if $P_{post}(\mathbf{Y}|\mathbf{Z}) = P_{pre}(\mathbf{Y}|\mathbf{Z})$.[7] In fact, a handful of pioneering work in observation studies is targeted on precisely this concept (e.g. Pratt and Schlaifer 1988; for a good review see Winship and Morgan 1999). Given a causal DAG, the manipulation principle enables us to calculate a post-intervention quantity in terms of pre-intervention probabilities and the known post-intervention probability of the direct targets of the intervention. This in principle can tell us whether the quantity is invariant under the intervention. The following definition thus makes sense:

[7] Though closely related, this notion of *invariance* is not to be confused with the kind of invariance we talked about in discussing the locality of interventions, or with the notion of invariance in the philosophical literature (most notably Woodward 2003.) As we said, the locality of an intervention requires that causal mechanisms for variables other than the direct targets should not be changed by the intervention and hence remain invariant. This notion of invariance applies to causal mechanisms, which is hence more fundamental than the kind of invariance we are talking about here. The latter is derived from the former. For example, by the invariance of mechanism for Y under an intervention of X, we immediately know that the probability of Y conditional on its causal parents is invariant in the sense just defined. Woodward's notion of invariance primarily refers to the stability of the mechanism that links the variable being manipulated to its effect. "Invariance" of other mechanisms is referred to as modularity instead.

Definition 5.1.1 (Invariance Given a DAG)**.** *Given a causal DAG $\mathcal{G}(\mathbf{O}, \mathbf{L})$, and three sets of variables $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \subseteq \mathbf{O}$ such that $\mathbf{X} \cap \mathbf{Y} = \mathbf{Y} \cap \mathbf{Z} = \emptyset$, $P(\mathbf{Y}|\mathbf{Z})$ is* **invariant under EL interventions of X given** *$\mathcal{G}$ if for all EL intervention of $\mathbf{X}$, the manipulation principle entails that $P_{post}(\mathbf{Y}|\mathbf{Z}) = P_{pre}(\mathbf{Y}|\mathbf{Z})$, no matter what the pre-intervention probability is.*

There is a simple graphical condition sufficient and necessary for invariance given a DAG due to Spirtes et al. (1993/2000). To state the condition, we need to introduce the now standard convention of graphically representing an EL intervention. The convention is to introduce extra *policy variables* into the causal DAG, one for each direct target variable of the intervention. A policy variable for X is simply an (extra) parent of X but otherwise not adjacent to any other variables in the DAG.[8] The resulting graph we call the $\mathbf{X}$-Policy-Augmented DAG of the original causal DAG. We have the following proposition, originally given in Spirtes et al. (1993/2000).

Proposition 5.1.1. *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 $\mathbf{X} \cap \mathbf{Y} = \mathbf{Y} \cap \mathbf{Z} = \emptyset$. $P(\mathbf{Y}|\mathbf{Z})$ is invariant under EL interventions of $\mathbf{X}$ given $\mathcal{G}$ if and only if in the $\mathbf{X}$-Policy-Augmented DAG of $\mathcal{G}$, the policy variables (of $\mathbf{X}$) are d-separated from $\mathbf{Y}$ given $\mathbf{Z}$.*[9]

This proposition is easily translated into the following theorem formulated in terms of $\mathcal{G}$ rather than the policy-augmented graph of $\mathcal{G}$.

[8]Dawid (2002) refers to policy variables as *regime indicators*.

[9]Spirtes, Glymour and Scheines argue that this condition is sufficient and *almost* necessary for invariance of $P(\mathbf{Y}|\mathbf{Z})$ under a *specific* intervention of $\mathbf{X}$. It is not exactly necessary because some intervention of $\mathbf{X}$ may leave some quantity invariant in virtue of certain *accidental* feature of the manipulation and the pre-intervention probability. For example, in our generalized version of McDermott's dog-bite case discussed in chapter 2, an intervention of the variable "button pressing" from "left hand" to "right hand" leaves the marginal probability of the variable "explosion" invariant, even though "explosion" is not d-separated from the policy variable of "button pressing" (given the empty set). (Such a case is regarded as a violation of faithfulness with respect to the policy-augmented graph.) However, the condition is truly necessary for our notion of invariance given a DAG here, as the definition quantifies over all EL interventions of $\mathbf{X}$.

Theorem 5.1.1. *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 $\mathbf{X} \cap \mathbf{Y} = \mathbf{Y} \cap \mathbf{Z} = \emptyset$. $P(\mathbf{Y}|\mathbf{Z})$ is invariant under EL interventions of $\mathbf{X}$ given $\mathcal{G}$ if and only 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.[10]

(Since $\mathbf{Z} \subseteq \mathbf{An}_{\mathcal{G}}(\mathbf{Z})$, $\mathbf{X} \cap \mathbf{Z}, \mathbf{X} \cap (\mathbf{An}_{\mathcal{G}}(\mathbf{Z}) \backslash \mathbf{Z})$ and $X \backslash \mathbf{An}_{\mathcal{G}}(\mathbf{Z})$ form a partition of $\mathbf{X}$.)

Proof. This is really just a restatement of Proposition 5.1.1. To see this, we consider the three cases separately:

Case 1: For every $X \in \mathbf{X} \cap \mathbf{Z}$, we show that there is a d-connecting path between the policy variable of X, PV_X, and a member of $\mathbf{Y}$, say, Y given $\mathbf{Z}$ in the policy-augmented graph if and only if there is a d-connecting path between X and Y given $\mathbf{Z} \backslash \{X\}$ that is into X in $\mathcal{G}$. Obviously if there is a d-connecting path between X and Y given $\mathbf{Z} \backslash \{X\}$ that is into X in $\mathcal{G}$, join that path with $PV_X \rightarrow X$ makes a d-connecting path between Y and PV_X given $\mathbf{Z}$ (because X will be a collider on that path). Conversely, if there is a path u d-connecting PV_X and Y given $\mathbf{Z}$, u has to start with $PV_X \rightarrow X$ because PV_X is otherwise not connected to other variables at all. Since $X \in \mathbf{Z}$, X is a collider on u, which means $u(X, Y)$ is into X. Now either

[10] 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 5.1.6 below is an immediate corollary of this equivalent formulation.

$u(X, Y)$ is a d-connecting path between X and Y given $\mathbf{Z}\backslash\{X\}$, or there exists a collider on the path that is not an ancestor of $\mathbf{Z}\backslash\{X\}$. In the latter case, let W be the collider closest to Y that does not have a descendant in $\mathbf{Z}\backslash\{X\}$. Then W must be an ancestor of X, otherwise u would not be d-connecting given $\mathbf{Z}$. Let d be a directed path from W to X. No vertex on d is in $\mathbf{Z}\backslash\{X\}$, for otherwise W would be an ancestor of $\mathbf{Z}\backslash\{X\}$. If no vertex on $u(Y, W)$ other than W is on d, then $u(Y, W)$ joined with d is obviously a d-connecting path between X and Y given $\mathbf{Z}\backslash\{X\}$ that is into X. Suppose there is a vertex on $u(Y, W)$ other than W that is also on d. Let Z be such a vertex closest to Y on U. Then no vertex on $u(Y, Z)$ other than Z is on $d(Z, X)$. So $u(Y, Z)$ can be joined with $d(Z, X)$ to form a d-connecting path between X and Y given $\mathbf{Z}\backslash\{X\}$ that is into X.

Case 2: For every $X \in \mathbf{X} \cap (\mathbf{An}_{\mathcal{G}}(\mathbf{Z})\backslash\mathbf{Z})$, obviously if there is a d-connecting path between X and Y given $\mathbf{Z}$ in $\mathcal{G}$, then no matter whether the path is into or out of X, joining that path with $PV_X \rightarrow X$ makes a d-connecting path between PV_X and Y given $\mathbf{Z}$. Conversely, if there is a d-connecting path between PV_X and Y given $\mathbf{Z}$, the subpath between X and Y is also d-connecting given $\mathbf{Z}$.

Case 3: For every $X \in \mathbf{X}\backslash\mathbf{An}_{\mathcal{G}}(\mathbf{Z})$, if there is a d-connecting path between X and Y given $\mathbf{Z}$ in $\mathcal{G}$ that is out of X, then joining that path with $PV_X \rightarrow X$ makes a d-connecting path between PV_X and Y given $\mathbf{Z}$, because X is a non-collider and is not in $\mathbf{Z}$. Conversely, if there is a d-connecting path between PV_X and Y given $\mathbf{Z}$, X must be a non-collider on the path, which means that the subpath between X and Y, while obviously d-connecting given $\mathbf{Z}$, is also out of X.

Therefore, (1), (2) and (3) together are equivalent to saying that in the X-policy-augmented graph of $\mathcal{G}$, the policy variables are d-separated from $\mathbf{Y}$ given $\mathbf{Z}$. □

We will take two steps to extend this result to CPAGs. First, we consider a notion of invariance given a DMAG. Recall that a causal DAG with latent variables corresponds to a unique causal DMAG over the observed variables. But many different DAGs correspond to the same MAG. So a causal DMAG actually represents an (infinite) set of causal DAGs. We define invariance given a DMAG based on the unanimity rule.

Definition 5.1.2 (Invariance Given a DMAG). *Let $\mathcal{M}$ be a causal DMAG over $\mathbf{O}$, and $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \subseteq \mathbf{O}$ be three sets of variables such that $\mathbf{X} \cap \mathbf{Y} = \mathbf{Y} \cap \mathbf{Z} = \emptyset$, $P(\mathbf{Y}|\mathbf{Z})$ is* **invariant under EL interventions of X given** $\mathcal{M}$ *if for every causal DAG $\mathcal{G}$ over $\mathbf{O}$ and (possibly) some latent variables whose MAG is $\mathcal{M}$, $P(\mathbf{Y}|\mathbf{Z})$ is invariant under EL interventions of $\mathbf{X}$ given $\mathcal{G}$.*

The question is how to judge invariance given a DMAG without checking all DAGs represented by the DMAG. In light of Theorem 5.1.1, one natural way to address this question is to set up connections between d-connecting paths in a DAG and m-connecting paths in the corresponding DMAG. To this end we need to make a distinction among directed edges in a DMAG.

Definition 5.1.3 (Visibility). *Given a DMAG $\mathcal{M}$, a directed edge $A \to B$ in $\mathcal{M}$ is* **visible** *if there is a vertex C not adjacent to B such that 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**.

Recall that under the causal interpretation of DMAGs, a directed edge means the presence of a causal pathway in the true DAG causal structure, and hence indicates a cause-effect relationship if causal transitivity is assumed (which, for one thing, follows from the assumption of CFC). What is distinctive about a visible directed edge between A and B in a causal DMAG is that it means more than that A is a cause of B. It in addition means that there is no unobserved common cause of A and B, unless mediated by some other observed variables. This implication is obvious given the following Lemma.

Lemma 5.1.1. *Let $\mathcal{G}(\mathbf{O}, \mathbf{L})$ be a DAG, and $\mathcal{M}$ be the DMAG 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 between A and B that is into A relative to $\mathbf{L}$ in $\mathcal{G}$, then there is a directed edge $A \to B$ in $\mathcal{M}$ that is invisible.*

Proof. See section 5.4. □

Lemma 5.1.1 implies that if $A \to B$ is visible in a DMAG, then in the true causal DAG, no matter which one it is, there is no inducing path between A and B relative to the set of latent variables that is into A. But if a latent variable is a common cause of A and B, then there immediately is an inducing path into A via that latent common cause. Therefore, a visible directed edge between two variables implies that they do not have a latent common cause. Conversely, if a directed edge between two variables is invisible in a DMAG, one can always construct a compatible DAG in which there is a latent common cause of the two variables.

Lemma 5.1.2. *Let $\mathcal{M}$ be any DMAG over a set of variables $\mathbf{O}$, and $A \to B$ be any directed edge in $\mathcal{M}$. If $A \to B$ is invisible in $\mathcal{M}$, then there is a DAG whose DMAG is $\mathcal{M}$ in which A and B share a latent parent, i.e., there exists a latent variable L_{AB} in the DAG such that $A \leftarrow L_{AB} \to B$ is a subgraph of the DAG.*

Proof. See section 5.4. □

The next block of lemmas, Lemmas 5.1.3-5.1.6, establish some connections between d-connecting paths in a DAG and m-connecting paths in the corresponding

DMAG. The first one, Lemma 5.1.3, records the important result that d-separation relations among observed variables in a DAG with latent variables correspond exactly to m-separation relations in the corresponding DMAG.

Lemma 5.1.3. *Let $\mathcal{G}(\mathbf{O}, \mathbf{L})$ be any DAG, and $\mathcal{M}$ be the DMAG 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 5.1.3, we know how to tell whether condition (2) of Theorem 5.1.1 holds in all DAGs compatible with a given DMAG. For the other two conditions in Theorem 5.1.1, we need to take into account orientations of d-connecting paths.

Lemma 5.1.4. *Let $\mathcal{G}(\mathbf{O}, \mathbf{L})$ be any DAG, and $\mathcal{M}$ be the DMAG 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 section 5.4. □

Lemma 5.1.5. *Let $\mathcal{M}$ be any DMAG 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 $\mathcal{G}(O, L)$ whose DMAG is $\mathcal{M}$ such that in $\mathcal{G}$ there is a path d-connecting A and B given $\mathbf{C}$ that is into A.*

Proof. See section 5.4. □

It is easy to see that these two lemmas are related to clause (1) in Theorem 5.1.1. The next lemma is related to clause (3).

Lemma 5.1.6. *Let $\mathcal{G}(\mathbf{O}, \mathbf{L})$ be any DAG, and $\mathcal{M}$ be the DMAG 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 $\mathcal{M}$, 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 section 5.4. □

We are now ready to translate Theorem 5.1.1 into a theorem about invariance given a DMAG.

Theorem 5.1.2. *Suppose $\mathcal{M}$ is the causal DMAG 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 invariant under EL interventions of $\mathbf{X}$ 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 5.1.4, if (1) holds, then for every DAG represented by $\mathcal{M}$, the first condition in Theorem 5.1.1 holds. Given Lemma 5.1.3 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 Theorem 5.1.1 holds for every DAG represented by

$\mathcal{M}$. Moreover, given 5.1.6, if (3) holds, the third condition in Theorem 5.1.1 holds for every DAG represented by $\mathcal{M}$. So (1), (2) and (3) together imply that $P(\mathbf{Y}|\mathbf{Z})$ is invariant under EL interventions of $\mathbf{X}$ given $\mathcal{M}$.

Conversely, if (1) fails, then by Lemma 5.1.5, there is a DAG represented by $\mathcal{M}$ in which the first condition in Theorem 5.1.1 fails. Likewise with conditions (2) and (3), in light of Lemmas 5.1.3 and 5.1.6, 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 invariant under EL interventions of $\mathbf{X}$ given $\mathcal{M}$. $\square$

Finally, we need to further generalize the result to invariance given a CPAG. A CPAG represents a Markov equivalence class of DMAGs, which may or may not agree upon a judgment of invariance. We will again apply the unanimity criterion.

Definition 5.1.4 (Invariance Given a CPAG)**.** *Let $\mathcal{P}$ be a causal CPAG over $\mathbf{O}$, and $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \subseteq \mathbf{O}$ be three sets of variables such that $\mathbf{X} \cap \mathbf{Y} = \mathbf{Y} \cap \mathbf{Z} = \emptyset$, $P(\mathbf{Y}|\mathbf{Z})$ is* **invariant under EL interventions of X given** *$\mathcal{P}$ if for every causal DMAG $\mathcal{M}$ in the Markov equivalence class represented by $\mathcal{P}$, $P(\mathbf{Y}|\mathbf{Z})$ is invariant under EL interventions of $\mathbf{X}$ given $\mathcal{M}$.*

Once again we aim to set up connections between m-connecting paths in a DMAG and analogous paths in its CPAG. In general a path in a CPAG may contain some vertices which cannot be unambiguously classified as colliders or non-colliders, and others that have a definite status. Let p be any path in a CPAG. A (non-endpoint) vertex is a *definite collider* on the path if both incident edges are into that vertex. A (non-endpoint) vertex C is called a *definite non-collider* on the path if one of the incident edges is out of C or it is $A *\!\!-\!\!\circ\, C \circ\!\!-\!\!* B$ on the path such that A and B are not adjacent. Likewise, a directed edge $A \rightarrow B$ in $\mathcal{P}$ is a *definitely visible* arrow if there is a vertex C not adjacent to B such that 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. Obviously these are labelled "definite" because the available informative marks in the CPAG are enough to determine their respective status, or what comes to the same thing, these colliders, non-colliders, or visible directed edges appear in all DMAGs represented by the CPAG[11]. Similarly we can define the following:

Definition 5.1.5 (Definite M-Connecting Path). *In a partial mixed graph, a path p between vertices A and B is* **definitely m-connecting** *relative to a set of vertices* $\mathbf{Z}$ *(*$A, B \notin \mathbf{Z}$*) if every non-endpoint vertex on p is either a definite non-collider or a definite collider and*

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

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

It is obvious that a definite m-connecting path in a CPAG is a m-connecting path in every DMAG represented by the CPAG. The following lemmas establish some further connections between m-connecting paths in a DMAG and definite m-connecting paths in its CPAG.

Lemma 5.1.7. *Let* $\mathcal{M}$ *be a DMAG over* $\mathbf{O}$*, and* $\mathcal{P}$ *be the CPAG 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 into A or out of A with an invisible directed edge, then there is a definite m-connecting path in* $\mathcal{P}$ *that is not out of A with a definitely visible edge.*

[11]Thus we can also call them "invariant" in line with the terminology used in Chapters 3 and 4, but this will obviously produce unnecessary terminological confusion here.

Proof. See section 5.4. $\square$

Lemma 5.1.8. *Let $\mathcal{P}$ be a CPAG 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 out of A with a definitely visible edge, then there exists a DMAG $\mathcal{M}$ represented by $\mathcal{P}$ in which there is a path m-connecting A and B given $\mathbf{C}$ that is either into A or out of A with an invisible directed edge.*

Proof. See section 5.4. $\square$

Lemma 5.1.9. *Let $\mathcal{M}$ be a DMAG over $\mathbf{O}$, and $\mathcal{P}$ be the CPAG 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.*

Proof. See section 5.4. $\square$

Lemma 5.1.10. *Let $\mathcal{P}$ be a CPAG 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 DMAG $\mathcal{M}$ 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 section 5.4. $\square$

Given a CPAG $\mathcal{P}$, call variable A a *possible ancestor* of variable B if there is a potentially directed path from A to B in $\mathcal{P}$. We use $\mathbf{PossibleAn}_{\mathcal{P}}(\mathbf{Z})$ to denote the set of possible ancestors of members of $\mathbf{Z}$. Here is the main theorem of this section.

Theorem 5.1.3. *Suppose $\mathcal{P}$ is the causal CPAG over a set of variables $\mathbf{O}$. For any $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \subseteq \mathbf{O}$ such that $\mathbf{X} \cap \mathbf{Y} = \mathbf{Y} \cap \mathbf{Z} = \emptyset$, $P(\mathbf{Y}|\mathbf{Z})$ is invariant under EL interventions of $\mathbf{X}$ 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 A 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) here are sufficient and necessary for the corresponding conditions in Theorem 5.1.2 to hold for all DMAGs represented by $\mathcal{P}$. It follows from Lemma 5.1.7 that if (1) holds, then the first condition in Theorem 5.1.2 holds for all DMAGs represented by $\mathcal{P}$. Note moreover that for every DMAG $\mathcal{M}$ represented by $\mathcal{P}$, $\mathbf{An}_{\mathcal{M}}(\mathbf{Z}) \subseteq \mathbf{PossibleAn}_{\mathcal{P}}(\mathbf{Z})$. So it again follows from Lemma 5.1.7 that if (2) holds, then the second condition in Theorem 5.1.2 holds for all DMAGs represented by $\mathcal{P}$. And it follows from Lemma 5.1.9 (and Lemma 5.1.7) that if (3) holds, the third condition in Theorem 5.1.2 holds for all DMAGs represented by $\mathcal{P}$. Hence (1), (2) and (3) are sufficient for $P(\mathbf{Y}|\mathbf{Z})$ to be invariant under EL interventions of $\mathbf{X}$ given $\mathcal{P}$.

Conversely, if (1) fails, then by Lemma 5.1.8, there exists a DMAG represented by $\mathcal{P}$ for which the first condition in Theorem 5.1.2 fails.

To show the necessity of (2), note that if X is a possible ancestor of a vertex $Z \in \mathbf{Z}$ in $\mathcal{P}$, then there exists a DMAG in which X is an ancestor of Z. (One way to see this is to recall that we can orient $\mathcal{P}$ into a DMAG by first tail-augmenting $\mathcal{P}$ and then orient the remaining circle component using Meek's algorithm. And in the latter step we can orient every $\circ\!\!-\!\!\circ$ edge incident to X to be out of X. In such a DMAG, for any W such that there is a potentially directed path in $\mathcal{P}$ from X to W, a shortest such path is oriented as a directed path from X to W.) So if (2) fails, i.e., 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 DMAG $\mathcal{M}$ 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 the the second condition in Theorem 5.1.2.

Lastly, if (3) fails, i.e., 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 5.1.10 that there exists a DMAG $\mathcal{M}$ 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 is not an ancestor of $\mathbf{Z}$ in $\mathcal{M}$, i.e., $X \in \mathbf{X} \backslash \mathbf{An}_{\mathcal{M}}(\mathbf{Z})$. So $\mathcal{M}$ fails the third condition in Theorem 5.1.2.

Therefore, (1), (2) and (3) are also necessary for $P(\mathbf{Y}|\mathbf{Z})$ to be invariant under EL interventions of $\mathbf{X}$ given $\mathcal{P}$. $\square$

This theorem is analogous in style to Theorems 7.2 and 7.3 in Spirtes et al. (1993/2000). The latter are formulated with respect to a partially oriented inducing path graph (POIPG). Without delving into the advantages CPAGs possess over POIPGs, we note a couple of improvements Theorem 5.1.3 achieves, apart from taking CPAGs as the representation of available causal information. First, Theorem 5.1.3 gives sufficient and necessary conditions, whereas the criteria given in Theorems 7.2 and 7.3 in Spirtes et al. (1993/2000) are not necessary.[12] Second, Theorems 7.2 and 7.3 are formulated in terms of possibly m-connecting paths (see the next section), which include definite m-connecting paths as special cases. In other words, there are in general more possibly m-connecting paths in a CPAG than definite m-connecting paths, and hence our Theorem 5.1.3 is probably superior from a computational per-

[12] As noted before, Spirtes, Glymour and Scheines were considering invariance given a (specific) EL intervention rather than invariance given *any* intervention. In this regard, even the criterion in Proposition 5.1.1 is not exactly necessary. The point here, however, is that even for invariance given *any* intervention, the criteria given by them are not necessary.

spective.

5.2 Generalization of the *do*-calculus

Pearl [1995] developed a *do*-calculus for identifying $P_{post}(\mathbf{Y}|\mathbf{Z})$ given a single causal DAG with latent variables, provided that interventions are simple in that variables being directly manipulated are manipulated to fixed values, i.e., $P_{post}(\mathbf{X} = \mathbf{x}) = 1$.[13] Apparently this work may be generalized in at least two directions. One is to consider more general EL interventions than simple EL interventions to fixed values. This generalization, however, is unnecessary, given the following argument presented by Pearl (2000). Suppose we are interested in $P_{post}(Y)$ given that X is manipulated to follow $P_{post}(X|\mathbf{Pa}(X))$. By the elementary probability calculus,

$$\begin{aligned} P_{post}(Y) &= \sum_{X} \sum_{\mathbf{Pa}(X)} P_{post}(Y, X, \mathbf{Pa}(X)) \\ &= \sum_{X} \sum_{\mathbf{Pa}(X)} P_{post}(Y|X, \mathbf{Pa}(X)) P_{post}(X|\mathbf{Pa}(X)) P_{post}(\mathbf{Pa}(X)) \end{aligned}$$

Note that $P_{post}(X|\mathbf{Pa}(X))$ is given, and $P_{post}(\mathbf{Pa}(X)) = P_{pre}(\mathbf{Pa}(X))$ because $\mathbf{Pa}(X)$ is a subset of the original direct causes of X (and hence the marginal probability of $\mathbf{Pa}(X)$ is invariant.) The only piece to consider then is $P_{post}(Y|X, \mathbf{Pa}(X))$. It is easy to check that the manipulation principle entails that

$$P_{post}(Y|X = x, \mathbf{Pa}(X)) = P'_{post}(Y|\mathbf{Pa}(X))$$

whenever the conditional probabilities are defined, where P'_{post} denotes the post-intervention distribution given that X is manipulated to the fixed value x. So the

[13] Pearl refers to such a simple intervention on a variable *atomic*.

problem is actually reduced to a prediction problem given simple EL interventions.[14]

The more interesting direction is to weaken the causal assumption. Suppose we do not have a unique causal DAG to start with, but rather a CPAG inferred from the pre-intervention probability distribution of the observed variables. Can Pearl's inference rules in *do*-calculus be formulated relative to a CPAG? This section aims to provide an answer.

Recall that a simple EL intervention of a variable X graphically amounts to erasing all the edges into X in the causal graph. Pearl's calculus is heavily based on such surgeries on DAGs. In particular, given a DAG $\mathcal{G}$ and a set of vertices $\mathbf{X}$, Pearl uses $\mathcal{G}_{\overline{\mathbf{X}}}$ to denote the graph resulting from deleting all edges in $\mathcal{G}$ that are into vertices in $\mathbf{X}$, and $\mathcal{G}_{\underline{\mathbf{X}}}$ to denote the graph resulting from deleting all edges in $\mathcal{G}$ that are out of vertices in $\mathbf{X}$. The following proposition summarizes Pearl's *do*-calculus, in which $P(\mathbf{Y}|do(\mathbf{X}) = \mathbf{x}, \mathbf{Z})$ is just a neat notation for $P_{post}(\mathbf{Y}|\mathbf{Z})$ under a simple EL intervention of $\mathbf{X}$ to a fixed value $\mathbf{x}$.

Proposition 5.2.1. *Let $\mathcal{G}$ be the causal DAG for $\mathbf{V}$, and $\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{W}$ be disjoint subsets of $\mathbf{V}$. The following rules are valid:*

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

$$P(\mathbf{Y}|do(\mathbf{X}), \mathbf{Z}, \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{W})$$

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

$$P(\mathbf{Y}|do(\mathbf{X}), do(\mathbf{Z}), \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{Z}, \mathbf{W})$$

[14]Note that this reduction works for $P_{post}(\mathbf{Y})$, but not for $P_{post}(\mathbf{Y}|\mathbf{Z})$. See Shpister and Pearl (2006) for a solution to the latter.

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

$$P(\mathbf{Y}|do(\mathbf{X}), do(\mathbf{Z}), \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{W})$$

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

Following the same strategy as in the last section, we generalize this result in two steps. Let us first define analogous surgeries on a DMAG.

Definition 5.2.1 (Manipulations of DMAGs)**.** *Given a DMAG* $\mathcal{M}$ *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 DMAGs may introduce new edges, i.e., replacing invisible directed edges with bi-directed edges. The reason we do this is that an invisible arrow from A to B admits a latent common parent of A and B in the underlying DAG. If so, the A-lower-manipulated DAG will correspond to a DMAG 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 DMAG of a DAG $\mathcal{G}$, we would like $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ to be the DMAG of $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, where X and Y are two (possibly empty) subsets of the observed variables. But in general this is impossible, as two DAGs represented by the same DMAG before a manipulation may correspond to different DMAGs after the manipulation. But we still have the following:

Lemma 5.2.1. *Let $\mathcal{G}(\mathbf{O}, \mathbf{L})$ be a DAG, and $\mathcal{M}$ be the DMAG 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 DMAG 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 section 5.4. $\square$

This lemma shows that Definition 5.2.1 is to a large extent appropriate. It implies that if an m-separation relation holds in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, then it holds in $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ for every $\mathcal{G}$ represented by $\mathcal{M}$. Hence the following corollary.

Corollary 5.2.2. *Let $\mathcal{M}$ be a DMAG 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 5.2.1, 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 DMAG. □

The converse of Corollary 5.2.2, however, is not true in general. The reason is roughly this. Lemma 5.2.1 is true in virtue of the fact that for every $\mathcal{G}$ represented by $\mathcal{M}$, there is a DMAG $\mathcal{M}^*$ Markov equivalent to $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ such that $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$ is a subgraph of $\mathcal{M}^*$. Often times there exists a $\mathcal{G}$ such that the DMAG of $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ is Markov equivalent to $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. But sometimes there may not be any such DAG, and that is why we do not have the converse of Lemma 5.2.1. For this limitation, however, Definition 5.2.1 is not to be blamed. Because 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 5.2.2 will not hold in general. $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, 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 all $\mathcal{G}$ represented by $\mathcal{M}$. To this end, Definition 5.2.1 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 DMAG $\mathcal{G}$ represented by $\mathcal{M}$ such that the two variables are adjacent in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$. In more plain terms, $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ does not have more adjacencies than necessary.

To give a simplest example, consider the DMAG $\mathcal{M}$ in Figure 5.1(a): $X \leftarrow Y \rightarrow Z$ (which happens to be a DAG also). The two DAGs, $\mathcal{G}1$ in 5.1(b) and $\mathcal{G}2$ 5.1(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 DMAG of $\mathcal{G}1_{\underline{Y}}$ is $X \leftrightarrow Y \quad Z$, and the DMAG 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 — say, X and Z are m-separated by the empty set — in $\mathcal{M}_{\underline{Y}}$ corresponds to a d-separation relation in both $\mathcal{G}1_{\underline{Y}}$ and $\mathcal{G}2_{\underline{Y}}$ (and, as one can show, in $\mathcal{G}_{\underline{Y}}$ for every $\mathcal{G}$ represented by $\mathcal{M}$), which is in accord with Corollary 5.2.2. By contrast, the converse of Corollary 5.2.2 fails of $\mathcal{M}$. It can be shown 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 $\mathcal{G}2_{\underline{Y}}$. (Roughly the reason is that a DAG in which there is both a latent direct cause of X and Y and a latent direct cause of Y and Z is not represented by $\mathcal{M}$, since in that DAG there is an inducing path between X and Z.) However, X and Z are not m-separated by Y in $\mathcal{M}_{\underline{Y}}$.

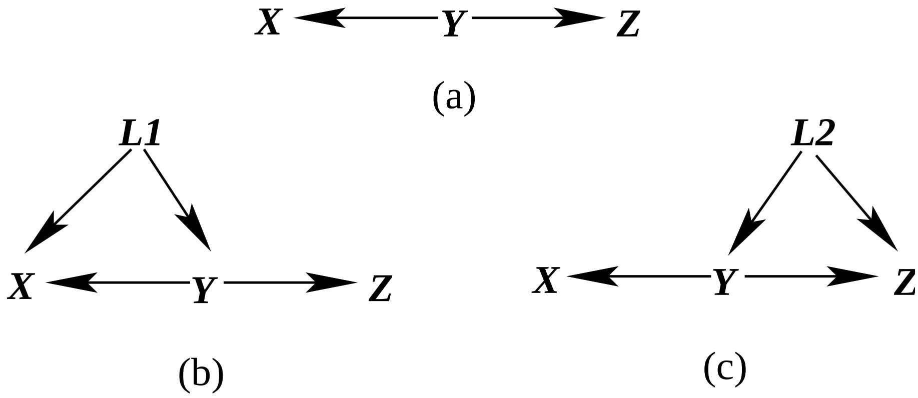

Figure 5.1: A DMAG that fails the converse of Corollary 5.2.2

But this failure is not peculiar to our definition of $\mathcal{M}_{\underline{Y}}$. In this simple example, one can easily enumerate all possible directed mixed graphs over X, Y, Z and see that for none of them are both 5.2.2 and its converse hold. Furthermore, among those graphs of which 5.2.2 holds, $\mathcal{M}_{\underline{Y}}$ is one of the graphs that have the fewest adjacencies.

We will not go into more details to illustrate this fact, which is not needed to prove the following theorem.

Theorem 5.2.1 (*do*-calculus given a DMAG)**.** *Let $\mathcal{M}$ be the causal DMAG over $\mathbf{O}$, and $\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{W}$ 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* $\mathbf{Z}$ *are m-separated by* $\mathbf{X} \cup \mathbf{W}$ *in* $\mathcal{M}_{\overline{\mathbf{X}}}$*, then*

$$P(\mathbf{Y}|do(\mathbf{X}), \mathbf{Z}, \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{W})$$

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

$$P(\mathbf{Y}|do(\mathbf{X}), do(\mathbf{Z}), \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{Z}, \mathbf{W})$$

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

$$P(\mathbf{Y}|do(\mathbf{X}), do(\mathbf{Z}), \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{W})$$

   where $\mathbf{Z}' = \mathbf{Z} \backslash \mathbf{An}_{\mathcal{M}_{\overline{\mathbf{X}}}}(\mathbf{W})$.

Proof. This readily follows from Proposition 5.2.1, Corollary 5.2.2, and the fact that for every $\mathcal{G}$ represented by $\mathcal{M}$, $\mathbf{An}_{\mathcal{G}_{\overline{\mathbf{X}}}}(\mathbf{W}) \cap \mathbf{O} = \mathbf{An}_{\mathcal{M}_{\overline{\mathbf{X}}}}(\mathbf{W})$. □

To generalize this result to CPAGs, we need to define relevant surgeries on CPAGs. They are very much like manipulations of DMAGs. Given a CPAG $\mathcal{P}$ and a set of variables $\mathbf{X}$, $\mathcal{P}_{\overline{\mathbf{X}}}$ denotes the $\mathbf{X}$-upper-manipulated graph of $\mathcal{P}$, resulting from deleting all edges in $\mathcal{P}$ that are into variables in $\mathbf{X}$, and otherwise keeping $\mathcal{P}$ as it is. $\mathcal{P}_{\underline{\mathbf{X}}}$ denotes the $\mathbf{X}$-lower-manipulated graph of $\mathcal{P}$, resulting from deleting all definitely visible edges out of variables in $\mathbf{X}$, replacing all other edges out of vertices in $\mathbf{X}$ with bi-directed edges, and otherwise keeping $\mathcal{P}$ as it is. The priority stipulation is also the same as before.

Except in very rare situations, $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ is not a CPAG any more. But from $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ we can still gain information about d-separation in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, where $\mathcal{M}$ is a DMAG in the Markov equivalence class represented by $\mathcal{P}$. For this purpose we need the following notion, already mentioned in the end of last section.

Definition 5.2.2 (Possibly M-Connecting Path). *In a partial mixed graph, a path p between vertices A and B is* **possibly m-connecting** *relative to a set of vertices $\mathbf{Z}$ ($A, B \notin \mathbf{Z}$) if*

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

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

Compare this definition with the definition of definite m-connecting paths (Definition 5.1.5), and one can immediately see that the latter is a special case of the former. Given a DMAG $\mathcal{M}$ and its CPAG $\mathcal{P}$, it is trivial to see that a m-connecting path in $\mathcal{M}$ is a possibly m-connecting path in $\mathcal{P}$. This is fortunately also true for $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ and $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$.

Lemma 5.2.3. *Let $\mathcal{M}$ be a DMAG over $\mathbf{O}$, and $\mathcal{P}$ the CPAG 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 $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$.*

Proof. See section 5.4. □

If there is no possibly m-connecting path between A and B given $\mathbf{C}$ in a partial mixed graph, we say A and B are *definitely m-separated* by $\mathbf{C}$ in the graph. Here is the main theorem of this section:

Theorem 5.2.2 (*do*-calculus given a CPAG)**.** *Let $\mathcal{P}$ be the causal CPAG for* $\mathbf{O}$*, and* $\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{W}$ *be disjoint subsets of* $\mathbf{O}$*. The following rules are valid:*

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

$$P(\mathbf{Y}|do(\mathbf{X}), \mathbf{Z}, \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{W})$$

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

$$P(\mathbf{Y}|do(\mathbf{X}), do(\mathbf{Z}), \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{Z}, \mathbf{W})$$

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

$$P(\mathbf{Y}|do(\mathbf{X}), do(\mathbf{Z}), \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{W})$$

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

Proof. It readily follows from Lemma 5.2.3, Theorem 5.2.1. The only caveat is that in general $\mathbf{An}_{\mathcal{M}_{\overline{\mathbf{X}}}}(\mathbf{W}) \neq \mathbf{PossibleAn}_{\mathcal{P}_{\overline{\mathbf{X}}}}(\mathbf{W})$ for an arbitrary $\mathcal{M}$ represented by $\mathcal{P}$. But it is always the case that $\mathbf{An}_{\mathcal{M}_{\overline{\mathbf{X}}}}(\mathbf{W}) \subseteq \mathbf{PossibleAn}_{\mathcal{P}_{\overline{\mathbf{X}}}}(\mathbf{W})$, which means that $\mathbf{Z} \backslash \mathbf{An}_{\mathcal{M}_{\overline{\mathbf{X}}}}(\mathbf{W}) \supseteq \mathbf{Z} \backslash \mathbf{PossibleAn}_{\mathcal{P}_{\overline{\mathbf{X}}}}(\mathbf{W})$ for every $\mathcal{M}$ represented by $\mathcal{P}$. So it is possible that for rule (3), $\mathcal{P}_{\overline{\mathbf{X}\mathbf{Z}'}}$ leaves more edges in than necessary, but it does not affect the validity of rule (3). $\square$

The rules in this calculus are in a sense overly restrictive. Even if a rule does not apply given a CPAG, the corresponding rule in Theorem 5.2.1 may still apply given every DMAG represented by the CPAG. This is possible because a possibly d-connecting path in a CPAG may not actualize as d-connecting in any DMAG represented by the CPAG, and because in general $\mathbf{An}_{\mathcal{M}_{\overline{\mathbf{X}}}}(\mathbf{W}) \subset \mathbf{PossibleAn}_{\mathcal{P}_{\overline{\mathbf{X}}}}(\mathbf{W})$.

Recall furthermore that the calculus based on a DMAG given in Theorem 5.2.1 is also "incomplete", as the converse of Corollary 5.2.2 does not hold. So there may well be interesting post-intervention quantities that can be identified by Pearl's *do*-calculus given any DAG compatible with a CPAG and all these DAGs give the same answer, but cannot be identified via our *do*-calculus based on the CPAG directly. We can artificially construct such examples, but we suspect that they are encountered very rarely in practice.

5.3 A Simple Example

We borrow an example from Spirtes et al. (1993/2000) to illustrate the invariance criterion and the *do*-calculus. Suppose we are able to measure the following random variables: *Income* (I), *Parents' smoking habits* (PSH), *Smoking* (S), *Genotype* (G) and *Lung cancer* (L) (The exact domain of each variable is not relevant for the illustration). The data, for all we know, are generated according to an underlying mechanism which might involve unmeasured common causes. Suppose as a matter of fact (unknown to us) the structure of the causal mechanism is the one in Figure 5.2, where *Profession* is an unmeasured common cause of *Income* and *Smoking*.

It is certainly impossible to fully recover this causal DAG from the data available, as the data alone by no means even indicate the relevance of the variable *Profession*. But we can in principle learn the CPAG shown in Figure 5.3. Although the CPAG reveals a limited amount of causal information, it is sufficient to identify some intervention effects.

For example, regarding invariance, it can be inferred that $P(L|G, S)$ is invariant under any EL intervention of I, because there is no definite m-connecting path between L and I given $\{G, S\}$. $P(L|G, S)$ is also invariant under any EL intervention of

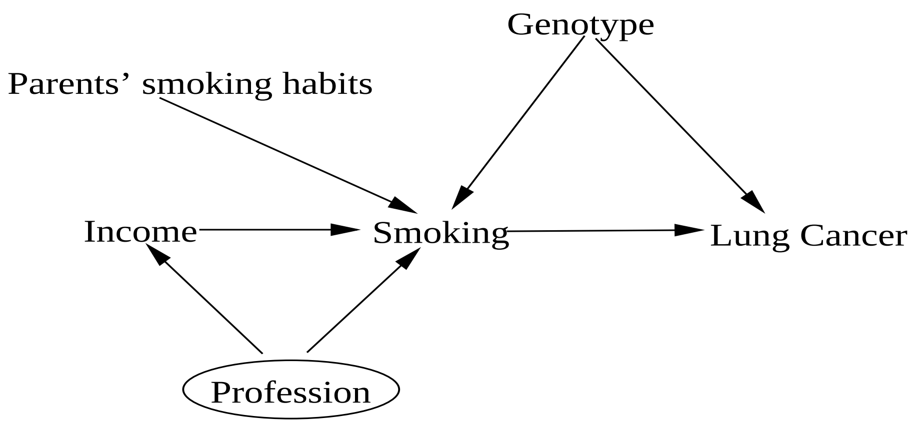

Figure 5.2: A causal DAG with a latent variable

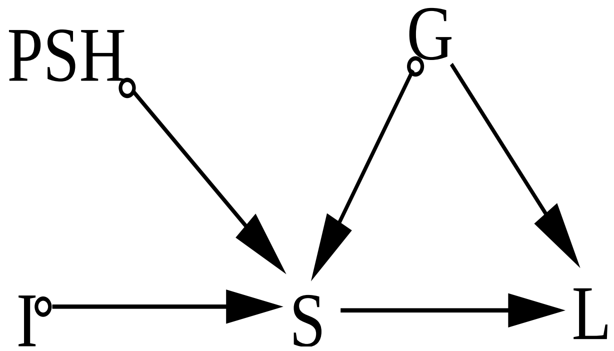

Figure 5.3: The CPAG of the causal DAG in Figure 5.2

S because the only definitely m-connecting path between L and S given $\{G\}$ is $S \rightarrow L$ which contains a definitely visible edge out of S. However, it is not invariant under EL interventions of G given the CPAG because the directed edge $G \rightarrow L$ is invisible. What this means is that there exists some DAG with latent variable compatible with this CPAG given which $P(L|G, S)$ is not invariant under some EL intervention of G.

Using the *do*-calculus presented in Theorem 5.2.2, we can infer $P(L|do(S), G) = P(L|S, G)$ by rule 2, because L and S are definitely m-separated by $\{G\}$ in $\mathcal{P}_{\underline{S}}$ (Figure 5.4(a)); and $P(G|do(S)) = P(G)$ is true by rule 3, because G and S are definitely m-separated in $\mathcal{P}_{\overline{S}}$ (Figure 5.4(b)). It follows that[15]

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

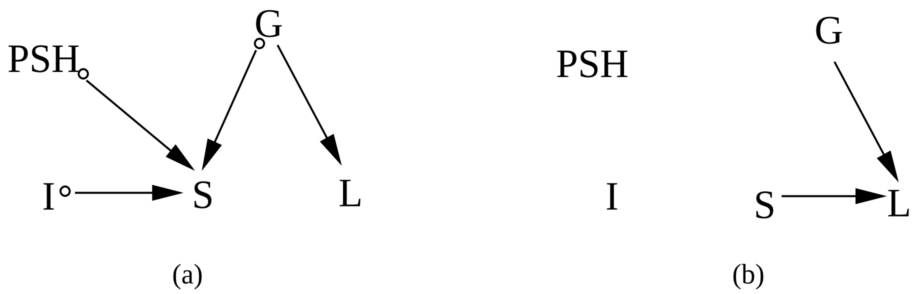

Figure 5.4: CPAG Surgery: $\mathcal{P}_{\underline{S}}$ and $\mathcal{P}_{\overline{S}}$

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 in Figure 5.5. (Notice the bi-directed edge between L and G.)

[15]The quantity $P(L)$ is not invariant given $\mathcal{P}$ under interventions of S, but as shown here using *do*-calculus, the post-intervention probability of L, i.e., $P(L|do(S))$ is identifiable in terms of quantities that are invariant under interventions of S, i.e., $P(L|S,G)$ and $P(G)$. In fact, Spirtes, Glymour and Scheines (2000) describe a prediction algorithm based on their theory of invariance that seeks to search for an expression of a post-intervention quantity in terms of invariant quantities. The prediction algorithm is recently shown by Grant Reaber to be incomplete, and in fact there are quantities that are identifiable via the *do*-calculus but not the prediction algorithm.

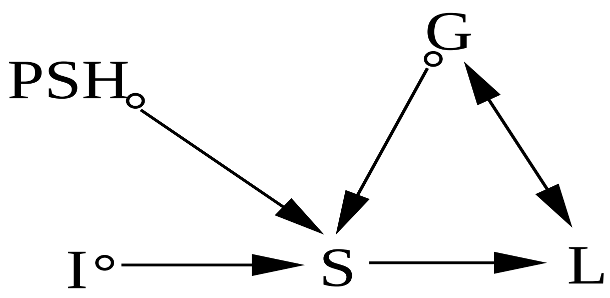

Figure 5.5: CPAG Surgery: $\mathcal{P}_{\underline{G}}$

5.4 Omitted Proofs

In our proofs, we will use the following lemma, which was proved in, for example, Spirtes, Meek and Richardson (1999, pp. 243):

Lemma 5.4.1. *Let $\mathcal{G}(\mathbf{O}, \mathbf{L})$ be a DAG, and $\langle V_0, \cdots, V_n \rangle$ be a sequence of distinct variables $\mathbf{O}$. If (1) for all $0 \leq i \leq n-1$, 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 $1 \leq i \leq n-1$, 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, Meek and Richardson (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 a same vertex multiple times. So in general it is a subsequence of the concatenation that constitutes and inducing path between V_0 and V_n.) $\square$

Lemma 5.4.1 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 DMAG. We will see several applications of this lemma in the subsequent proofs.

Proof of Lemma 5.1.1

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 5.4.1 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 $C \to V_{i+1}$. 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 5.4.1 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 $A \to B$ is invisible in $\mathcal{M}$. $\square$

Proof of Lemma 5.1.2

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 $A \leftarrow L_{AB} \to B$ 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 $Z \leftarrow L_{ZW} \to W$ to the graph.

The resulting graph we denote by $\mathcal{G}$. Obviously $\mathcal{G}$ is a DAG in which A and B share a latent parent. We need to show that $\mathcal{M} = \mathcal{M}_{\mathcal{G}}$, i.e., $\mathcal{M}$ is the DMAG of $\mathcal{G}$. For any pair of variables X and Y, there are four cases to consider:

Case 1: $X \to Y$ is in $\mathcal{M}$. Since $\mathcal{G}$ retains all directed edges in $\mathcal{M}$, $X \to Y$ 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 $X \leftarrow L_{XY} \to Y$ appears in $\mathcal{G}$. Since $X \leftarrow L_{XY} \to Y$ 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 fewest observed variables. Let $\langle X, O_1, \cdots, O_n, Y\rangle$ 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, either all of them are bi-directed, or one of them is $A \rightarrow B$ and others are bi-directed. In the former case, $\langle X, O_1, \cdots, O_n, Y\rangle$ 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 $A \to B$ 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 $A \to B$ 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 5.1.4

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.[16]

[16] Their lemma addresses the more general case where there may also be selection variables. The construction given here is an adaptation of theirs to fit our assumption that there are no selection variables.

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.[17] 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

[17] 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.

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 $A \rightarrow \mathbf{S}_K[1]$, then this edge is invisible. Given Lemma 5.1.1, it suffices to show that there is an inducing path between 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 5.4.1 (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. $\square$

Proof of Lemma 5.1.5

Proof. This lemma is fairly obvious given Lemma 5.1.2. 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. $\square$

Proof of Lemma 5.1.6

Proof. Note that because A is not an ancestor 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 $\mathbf{C}$. 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. $\square$

Proof of Lemma 5.1.7

Proof. We will use the following fact several times.

lemma 0: If a path $\langle U, \cdots, X, Y, Z \rangle$ is a discriminating path for Y in $\mathcal{M}$, and the corresponding subpath between U and Y in $\mathcal{P}$ is (also) a collider path, then the path is also a discriminating path for Y in $\mathcal{P}$.

proof: Given that $\mathcal{P}$ is a CPAG for $\mathcal{M}$, all we need to show is that for every vertex between U and Y on the path, the vertex is also a parent of Z in $\mathcal{P}$. This is easy by induction. In the base case, let V be the first vertex after U on the path. Since U and Z are not adjacent, and the edge between U and V is into V (as by assumption V is a collider on the path), so the edge between V and Z should be oriented (recall, by $\mathcal{R}1$) as $V \to Z$ in $\mathcal{P}$ (because by assumption, $V \to Z$ appears in $\mathcal{M}$. Now, for the inductive step, consider an arbitrary vertex W between U and Y. Suppose every vertex between U and W is a parent of Z in $\mathcal{P}$. It follows that there is a discriminating path between U and Z for W in $\mathcal{P}$. Hence the edge between W and Z should be oriented (recall, by $\mathcal{R}4$) as $W \to Z$ in $\mathcal{P}$. **Q.E.D.**

For the main lemma, we first show that the presence of an m-connecting path in $\mathcal{M}$ implies the presence of a definite m-connecting path in $\mathcal{P}$. We break the long argument into a couple of sub-lemmas.

lemma 1: Let p be a shortest path m-connecting A and B given $\mathbf{C}$ in $\mathcal{M}$. Let p^* denote the corresponding path constituted by the same sequence of variables in $\mathcal{P}$. Then every non-endpoint vertex on p^*, if any, is of a definite status, i.e., either a definite collider or a definite non-collider.

proof: We denote the sequence of variables on p by $\langle A = O_0, O_1, ..., O_n = B \rangle$. We first establish the following claim.

> **claim**: For every $1 \leq j \leq n-1$, if O_j is not of a definite status on p^*, then O_{j+1} is a parent of O_{j-1} in $\mathcal{M}$, and O_{j+1} is a collider on p in $\mathcal{M}$.

The claim trivially holds if every non-endpoint vertex on p^* is of a definite status. On the other hand, suppose there exists $1 \leq j \leq n-1$ such that O_j is not a definite collider or non-collider. We demonstrate the claim by induction.

In the base case, let O_K ($1 \leq K \leq n-1$) be the one closest to A on p^* that is not of a definite status, and we show that the claim holds of K. Since O_K is not of a definite status, O_{K-1} and O_{K+1} must be adjacent, for otherwise $\langle O_{K-1}, O_K, O_{K+1}\rangle$ would be an unshielded triple, and hence O_K would be of a definite status on p^*. Because p is chosen to be the shortest m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{M}$, the path $p' = p(A, O_{K-1}) \oplus \langle O_{K-1}, O_{K+1}\rangle \oplus p(O_{K+1}, B)$ (i.e., the path p with O_K passed over) is not m-connecting given $\mathbf{C}$ in $\mathcal{M}$. But the only possible vertices to block p' are O_{K-1} and O_{K+1}, as the status of other vertices are the same as they are on p. Below we rule out three of four possible cases for O_{K-1} or O_{K+1} to block p' by deriving a contradiction in each.

Case 1: O_{K-1} is a non-collider on p but a collider on p', and O_{K-1} does not have a descendant in $\mathbf{C}$ in $\mathcal{M}$. It follows that the edge between O_{K-1} and O_K in $\mathcal{M}$ is out of O_{K-1}, i.e., $O_{K-1} \rightarrow O_K$, and the edge between O_{K-1} and O_{K+1} in $\mathcal{M}$ is into O_{K-1}, i.e., $O_{K-1} \leftarrow\!\!* O_{K+1}$. So the edge between O_K and O_{K+1} in $\mathcal{M}$ is into O_K, i.e., $O_K \leftarrow\!\!* O_{K+1}$ (for otherwise $\mathcal{M}$ is not ancestral). Hence O_K is a collider on p, which implies that O_K is an ancestor of a member of $\mathbf{C}$ in $\mathcal{M}$ (because p is m-connecting). This in turn implies that O_{K-1} is an ancestor of a member of $\mathbf{C}$ in $\mathcal{M}$, which contradicts the supposition that O_{K-1} does not have a descendent in $\mathbf{C}$.

Case 2: O_{K+1} is a non-collider on p but a collider on p', and O_{K+1} does not have a descendant in $\mathbf{C}$ in $\mathcal{M}$. This case is symmetric to *Case 1*, and the same argument will produce a contradiction.

Case 3: O_{K-1} is a collider on p but a non-collider on p', and O_{K-1} is in $\mathbf{C}$. Note that this automatically implies that $O_{K-1} \neq A$. Consider O_{K-2}, i.e., the vertex next to O_{K-1} on $p(A, O_{K-1})$. Since O_{K-1} is a collider on p but a non-collider on p', we have $O_{K-2} *\!\!\rightarrow O_{K-1} \leftarrow\!\!* O_K$ and $O_{K-1} \rightarrow O_{K+1}$ in $\mathcal{M}$. Since O_k is chosen to be the vertex closest to A that is not of a definite status on p^* in $\mathcal{P}$, O_{k-1} is of a definite status on p^*, which implies that the collider, $O_{K-2} *\!\!\rightarrow O_{K-1} \leftarrow\!\!* O_K$, also occurs in $\mathcal{P}$. It follows that O_{K-2} is adjacent to O_{K+1}. For otherwise $\langle O_{K-2}, O_{K-1}, O_K, O_{K+1} \rangle$ forms a discriminating path for O_K in $\mathcal{M}$, and hence by **lemma 0**, is also a discriminating path for O_K in $\mathcal{P}$, which implies that O_K should be a definite non-collider or collider on p^* (recall $\mathcal{R}4$ in FCI), a contradiction. Moreover, the edge between O_{K-2} and O_{K+1} in $\mathcal{M}$ is into O_{K+1}, i.e., $O_{K-2} *\!\!\rightarrow O_{K+1}$, which follows from $O_{K-2} *\!\!\rightarrow O_{K-1} \rightarrow O_{K+1}$.

Note also that the edge between O_K and O_{K+1} in $\mathcal{M}$ is into O_{K+1}, which follows from $O_K *\!\!\rightarrow O_{K-1} \rightarrow O_{K+1}$. Now, the path, p'', resulting from replacing $p(O_{K-2}, Z) = O_{K-2} *\!\!\rightarrow O_{K-1} \leftarrow\!\!* O_K *\!\!\rightarrow O_{K+1}$ with $O_{K-2} *\!\!\rightarrow O_{K+1}$ on p is shorter than p and hence is not m-connecting. So O_{K-2} must block p'' (because the marks at O_{K+1} are the same). Note moreover that if the edge between O_{K-2} and O_{K-1} in $\mathcal{M}$ is $O_{K-2} \rightarrow O_{K-1}$, then the edge between O_{K-2} and O_{K+1} is $O_{K-2} \rightarrow O_{K+1}$, in which case O_{K-2} does not block p''. Hence the only possibility is that O_{K-2} is a collider on p, but is a non-collider on p'', which means that $O_{K-2} \rightarrow O_{K+1}$ is in $\mathcal{M}$.

Then clearly we can apply the exact same argument again to O_{K-3}, the vertex next to O_{K-2} on $p(A, O_{K-2})$. Indeed an obvious and rigorous argument by induction will establish that every vertex between A and O_K on p is a collider on p in both $\mathcal{M}$ and $\mathcal{P}$, and is a parent of O_{K+1} in $\mathcal{M}$. It follows that A and O_{K+1} are adjacent, for otherwise, by **lemma 0**, $p(A, O_{K+1})$ would form a discriminating path for O_K in $\mathcal{P}$, and hence O_K should be a definite non-collider or collider on p^*, a contradiction.

Now consider the edge between A and O_{K+1} in $\mathcal{M}$. It is not out of O_{K+1}, for otherwise $\mathcal{M}$ is not ancestral. But we have already shown that the edge between O_K and O_{K+1} is into O_{K+1}, so if the edge between A and O_{K+1} is into O_{K+1}, replacing $p(A, O_{K+1})$ with $A*\!\!\rightarrow O_{K+1}$ on p makes an m-connecting path between A and B given $\mathbf{C}$ in $\mathbf{M}$ shorter than p is, which is a contradiction.

Since *Cases 1, 2* and *3* are all ruled out, the only possibility is that O_{K+1} is a collider on p but a non-collider on p' (and O_{K+1} is in $\mathbf{C}$). This obviously implies that the edge between O_{K-1} and O_{K+1} in $\mathcal{M}$ is $O_{K-1} \leftarrow O_{K+1}$, i.e., O_{K+1} is a parent of O_{K-1} in $\mathcal{M}$, and O_{K+1} is a collider on p in $\mathcal{M}$. So the base case holds.

For the inductive step, suppose O_r $(1 \leq r \leq n-1)$ is not of a definite status on p^* and the **claim** holds of r, and we need to show that for the next such vertex, say, O_{r+l} on p^*, the **claim** also holds. We now argue that the following hold: (1) the edge between O_{r-1} and O_r is into O_{r-1} in $\mathcal{M}$; and (2) for every $1 \leq j \leq l$, O_{r+j} is a parent of O_{r-1} and is a collider on p in $\mathcal{M}$.

(1) follows from the fact that O_{r+1} is a parent of O_{r-1} in $\mathcal{M}$ (by the inductive hypothesis). For otherwise we have $O_{r-1} \rightarrow O_r$ in $\mathcal{M}$, which in turn implies that the edge between O_r and O_{r+1} is $O_r \leftarrow O_{r+1}$. Then we can replace $O_{r-1} \rightarrow O_r \leftarrow O_{r+1}$ by $O_{r-1} \leftarrow O_{r+1}$ on p and get a shorter m-connecting path, which is a contradiction. (The reason why the new path is also m-connecting is this: O_{r+1} obviously does not block the new path, since it does not block the old path p. For O_{r-1}, notice that it is a non-collider on p, so it is not in $\mathbf{C}$; but on the other hand, it is a parent of O_r which is a collider on p, which implies that O_{r-1} is also an ancestor of $\mathbf{C}$. Thus it does not block the new path either.)

The argument for (2) is naturally by induction. Note that an important supposition we can use here is that every other vertex between O_r and O_{r+l} is of a definite status on p^*. The base case trivially holds because of the inductive hypothesis of the outer induction, i.e., that the **claim** holds of r. Suppose for all $j < m$ it is true that O_{r+j} is a parent of O_{r-1} and is a collider on p in $\mathcal{M}$. Then all these colliders also occur in $\mathcal{P}$ because every vertex between O_r and O_{r+l} is of a definite status. So O_m must be adjacent to O_{r-1} in $\mathcal{M}$, for otherwise $p(O_{r-1}, O_m)$ would be a discriminating path for O_r in $\mathcal{M}$, and by **lemma 0**, $p^*(O_{r-1}, O_m)$ would also be a discriminating path for O_r in $\mathcal{P}$, and hence O_r would be a definite collider or non-collider, a contradiction. Now consider the edge between O_m and O_{r-1} in $\mathcal{M}$. It is into O_{r-1}, for otherwise $\mathcal{M}$ is not ancestral (because O_{m-1} is a parent of O_{r-1} and is a collider on p). Consider the path resulting from replacing $p(O_{r-1}, O_m)$ with $O_{r-1} \leftarrow\!\!* O_m$ on p. It is shorter than p, so it is not m-connecting. But by (1), $p(O_{r-1}, O_m)$ is also into O_{r-1}, so the only possibility is that O_m would block the new path. By the same argument we have used several times, it is easy to derive that O_m is a collider on p but a non-collider on the new path, which means O_m is also a parent of O_{r-1}. Hence (2) is established.

We are ready to finish the inductive step of the outer induction. Since O_{r+l} is not of a definite status on p^*, O_{r+l-1} is adjacent to O_{r+l+1}. Again, either O_{r+l-1} or O_{r+l+1} blocks the path resulting from replacing the triple $O_{r+l-1}, O_{r+l}, O_{r+l+1}$ with the edge between O_{r+l-1} and O_{r+l+1} on p. There are, similarly to the base case, four possibilities to consider, and the analogous *Case 1* and *Case 2* can be ruled out by the exact same argument we used in the base case. The analogous *Case 3* cannot be ruled out by the same argument we used, but can be ruled out as follows. suppose for contradiction that O_{r+l-1} is a collider on p and a non-collider on the new path, which implies that O_{r+l-1} is a parent of O_{r+l+1}. It then follows, by the exact same argument we used to argue for (1) above, that the edge between O_{r+l} and O_{r+l+1} is into O_{r+l+1}. And in turn it follows, by the exact same argument we used to argue for (2) above, that for every $1 \leq j \leq l$, O_{r+l-j} is a parent of O_{r+1+l} and is a collider on p in $\mathcal{M}$. This, together with (2) above, implies that $p(O_{r-1}, O_{r+l+1})$ forms an inducing path between O_{r-1} and O_{r+l+1} in $\mathcal{M}$. Because $\mathcal{M}$ is maximal, it follows that $O_{r-1} \leftrightarrow O_{r+l+1}$ occurs in $\mathcal{M}$. This means we can replace $p(O_{r-1}, O_{r+l+1})$ with $O_{r-1} \leftrightarrow O_{r+l+1}$ and get an m-connecting path between A and B given $\mathbf{C}$ shorter than p is, which is a contradiction. Therefore, the only remaining possibility is that O_{r+l+1} is a collider on p but a non-collider on the new path (resulting from replacing the triple $\langle O_{r+l-1}, O_{r+l}, O_{r+l+1} \rangle$ with the edge between O_{r+l-1} and O_{r+l+1}). It follows that O_{r+l+1} is a parent of O_{r+l-1}. Hence the **claim** holds of $r + l$ as well.

Now that we have established the **claim**, we are ready to prove **lemma 1**. The key is to consider the following claim:

> **claim***: For every $1 \leq j \leq n - 1$, if O_j is not of a definite status on p^*, then O_{j-1} is a parent of O_{j+1} in $\mathcal{M}$, and O_{j-1} is a collider on p in $\mathcal{M}$.

Obviously **claim*** is symmetric to the original **claim**. And it takes little effort to see that we can carry out an exactly symmetric argument by induction, this time starting from the vertex closest to B that is not of a definite status on p^*, to establish **claim***. But obviously **claim** and **claim*** is in conflict as long as there exists $1 \leq j \leq n - 1$ such that O_j is not of a definite status on p^*. It follows that every non-endpoint vertex on p^* is of a definite status, which concludes our proof of **lemma 1**. **Q.E.D.**

Given any path p m-connecting A and B given $\mathbf{C}$ in $\mathcal{M}$, for every collider Q on p, there is a directed path (possibly of length 0) from p to a member of $\mathbf{C}$. Define the *distance-from-*$\mathbf{C}$ of Q to be the length of a shortest directed path (possibly of length 0) from Q to $\mathbf{C}$, and define the *distance-from-*$\mathbf{C}$ of p to be the sum of the distances from $\mathbf{C}$ of the colliders on p.

lemma 2: Let p be a shortest path m-connecting A and B given $\mathbf{C}$ in $\mathcal{M}$ such that no equally short m-connecting path has a shorter distance-from-$\mathbf{C}$ than p does.[18] Let p^* denote the corresponding path constituted by the same sequence of variables in $\mathcal{P}$. Then p^* is a definite m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{P}$.

proof: Since p is a shortest m-connecting path, by **lemma 1**, every non-endpoint vertex on p^*, if any, is of a definite status, i.e., either a definite collider or a definite non-collider. Since $\mathcal{P}$ is a CPAG of $\mathcal{M}$, every definite non-collider on p^* corresponds to a non-collider on p, and hence is not in $\mathbf{C}$, for otherwise p would not be m-connecting given $\mathbf{C}$ in $\mathcal{M}$.

Similarly, for any definite collider Q on p^*, Q is also a collider on p. Hence there is a directed path (possibly of length 0) from Q to a member of $\mathbf{C}$ in $\mathcal{M}$. Let d be a shortest such path from Q to, say, $C \in \mathbf{C}$. Let d^* denote the corresponding path in $\mathcal{P}$. Because $\mathcal{P}$ is a CPAG of $\mathcal{P}$, d^* is a potentially directed path from Q to C in $\mathcal{P}$. We now show that no circle mark ($\circ$) appears on d^*, i.e., that d^* is (fully) directed. Suppose for contradiction that there is a circle on d^*. Then the mark at Q on d^* must be a circle, for otherwise, an arrowhead would meet a circle on d^*, and by the property **CP**1 (recall Lemma 3.3.1), a proper subpath of d^* would constitute a p.d. path from Q to C, which in turn implies that there is a shorter directed path from Q to C in $\mathcal{M}$ than d is, a contradiction with our choice of d.

Let $Q \circ\!\!-\!\!* S$ be the first edge on Q. Suppose S is not on p^* for the moment. Since Q is a definite collider on p^*, we have $Q_l {*\!\!\rightarrow} Q \leftarrow\!\!* Q_r$ in $\mathcal{P}$, Q_l, Q_r being the two vertices adjacent to Q on p^*. By property **CP**1, there is an edge between Q_l and S that is into S, and there is an edge between Q_r and S that is into S, i.e., $Q_l {*\!\!\rightarrow} S \leftarrow\!\!* Q_r$ in $\mathcal{P}$.

Now we show that there exists a vertex W (distinct from Q) on $p(A, Q)$ such that (i) there is an edge between W and S in $\mathcal{M}$ that is into S; and (ii) in $\mathcal{M}$, the

[18] Such an m-connecting path is named *minimal* in, for example, Richardson and Spirtes (2002).

collider/non-collider status of W on p is the same as the collider/non-colllider status of W on $p(A, W) \oplus \langle W, S\rangle$. To show this, it suffices to demonstrate that if no vertex between A and Q on p satisfies the two conditions, then A must satisfy them. Suppose no vertex between A and Q on p satisfies the two conditions. If $Q_l = A$, A satisfies (i) and (ii) trivially. Suppose $Q_l \neq A$. We argue by induction that every vertex between A and Q is a collider on p and is a parent of S in $\mathcal{M}$.

In the base case, we already established that (i) holds of Q_l. So, by our supposition, (ii) does not hold of Q_l. It follows that either Q_l is a non-collider on p but a collider on $p(A, Q_l) \oplus Q_l {*\!\!\rightarrow} S$, or Q_l is a collider on p but a non-collider on $p(A, Q_l) \oplus Q_l {*\!\!\rightarrow} S$. The former case implies that the edge between Q_l and Q is $Q_l \rightarrow Q$, and the edge between Q_l and S is into Q_l, which contradicts the fact that $Q \rightarrow S$ occurs in $\mathcal{M}$ and $\mathcal{M}$ is ancestral. So only the latter case is possible, which implies that Q_l is a parent of S in $\mathcal{M}$ and is a collider on p.

In the inductive step, suppose $Q_m \neq A$ is a vertex between A and Q_l on p, and every vertex between Q_m and Q is a collider on p and is a parent of S in $\mathcal{M}$. It follows that Q_m is adjacent to S, for otherwise $\langle Q_m, ..., Q_l, Q, S\rangle$ forms a discriminating path for Q in $\mathcal{M}$, and, by **lemma 0**, also forms a discriminating path for Q in $\mathcal{P}$ (because every non-endpoint vertex on p^* is of a definite status), and hence the circle at Q on $Q \circ\!\!-\!\!\!* S$ should have been oriented by $\mathcal{R}4$. The edge between Q_m and S in $\mathcal{M}$ is not out of S, on pain of making $\mathcal{M}$ non-ancestral. So (i) holds of Q_m. Then, by our supposition, (ii) does not hold of Q_m. By the same argument we used in the base case, Q_m is a collider on p and is a parent of S in $\mathcal{M}$.

Now we have established that every vertex between A and Q is a collider on p and is a parent of S in $\mathcal{M}$, it is easy to see that A is adjacent to S, for otherwise the circle at Q on $Q \circ\!\!-\!\!\!* S$ should have been oriented by $\mathcal{R}4$ (for the same reason stated in the last paragraph). The edge between A and S in $\mathcal{M}$ is not out of S, on pain of making $\mathcal{M}$ non-ancestral. So (i) holds of A. But (ii) holds of A trivially. Thus we have established that there exists a vertex W (distinct from Q) on $p(A, Q)$ such that (i) there is an edge between W and S in $\mathcal{M}$ that is into S; and (ii) the collider/non-collider status of W on p is the same as the collider/non-colllider status of W on $p(A, W) \oplus \langle W, S \rangle$.

By symmetry, it follows that there exists a vertex V (distinct from Q) on $p(Q, B)$ such that (i) there is an edge between V and S in $\mathcal{M}$ that is into S; and (ii) the collider/non-collider status of V on p is the same as the collider/non-colllider status of V on $\langle S, V \rangle \oplus p(V, B)$. Then the path $p' = p(A, W) \oplus \langle W, S, V \rangle \oplus p(V, B)$ (it could be that $A = W$ and/or $V = B$) is obviously m-connecting given $\mathbf{C}$ in $\mathcal{M}$. It is easy to check that either p' is shorter than p is, or p' is as long as p is (when $W = Q_l$ and $V = Q_r$) but has a shorter distance-from-$\mathbf{C}$ than p. Either case is a contradiction with our assumption about p.

Finally, if S is on p, it either lies on $p(A, Q_l)$ or lies on $p(Q_r, B)$. Without loss of generality, suppose it is on $p(Q_r, B)$. The same argument goes through to establish that there exists a vertex W (distinct from Q) on $p(A, Q)$ such that (i) there is an edge between W and S in $\mathcal{M}$ that is into S; and (ii) the collider/non-collider status of W on p is the same as the collider/non-colllider status of W on $p(A, W) \oplus \langle W, S \rangle$. Then the path $p(A, W) \oplus W {*\!\!\rightarrow} S \rangle p(S, B)$ is m-connecting given $\mathbf{C}$ but shorter than p is, a contradiction.

So our supposition that there is a circle mark on d^* is false. d^* is directed in $\mathcal{P}$ as well, and hence Q is an ancestor of $\mathbf{C}$ in $\mathcal{P}$. Therefore, p^* is a definite m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{P}$. **Q.E.D.**

It obviously follows from **lemma 2** that if there is a m-connecting path between

A and B given $\mathbf{C}$ in $\mathcal{M}$, then there is a definite m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{P}$, which is the first part of Lemma 5.1.7.

To prove the second part of Lemma 5.1.7, namely that if there is an m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{M}$ that is *into A or out of A with an invisible directed edge*, then there is a definite m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{P}$ that is not out of A with a definitely visible edge, the same argument almost applies. Basically we will establish lemmas analogous to **lemma 1** and **lemma 2**:

lemma 1': Let p be a shortest path m-connecting A and B given $\mathbf{C}$ in $\mathcal{M}$ *whose corresponding path in $\mathcal{P}$, denoted by p^*, is not out of A with a definitely visible edge.* Then every non-endpoint vertex on p^*, if any, is of a definite status, i.e., either a definite collider or a definite non-collider.

proof: The same argument for **lemma 1** applies here, except that when we derive a contradiction by showing that there is a shorter m-connecting path, we have to argue that the shorter path is also such that *its corresponding path in $\mathcal{P}$ is not out of A with a definitely visible edge.* The only place in the argument where this additional requirement is not obviously satisfied is in *Case 3* of the base case of the induction. There, an (initial) segment of p, $p(A, O_{K+1})$, is replaced by an edge $A{*\!\!\rightarrow} O_{K+1}$. We have to check that this edge does not correspond to a definitely visible $A \rightarrow O_{K+1}$ in $\mathcal{P}$. But this is not hard to show, because this replacement happens only when every vertex between A and O_{K-1} on p is a collider and is a parent of O_{K+1} in $\mathcal{M}$ (and in $\mathcal{P}$). If $A \rightarrow O_{K+1}$ is definitely visible in $\mathcal{P}$, then there is in $\mathcal{P}$ a vertex E not adjacent to O_{K+1} such that either $E{*\!\!\rightarrow} A$ or there is a collider path between E and A that is into A and every collider on the path is a parent of O_{K+1}. From this it is easy to derive that there is a discriminating path $\langle E, \cdots, O_{K-1}, O_K, O_{K+1}$ for O_K in $\mathcal{P}$, which means that O_K should be of a definite status on p^*, a contradiction. **Q.E.D.**

lemma 2': Let p be a shortest path m-connecting A and B given $\mathbf{C}$ in $\mathcal{M}$ *whose corresponding path in* $\mathcal{P}$*, denoted by* p^**, is not out of* A *with a definitely visible edge*, such that no equally short m-connecting path *whose corresponding path in* $\mathcal{P}$ *is not out of* A *with a definitely visible edge* has a shorter distance-from-$\mathbf{C}$ than p does.[19] Then p^* is a definite m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{P}$.

proof: Again, the argument for **lemma 2** applies here, except that when we derive a contradiction by showing that there is a shorter m-connecting path or a equally long but with a shorter distance-from-$\mathbf{C}$ path, we need to argue that the path is also such that *its corresponding path in* $\mathcal{P}$ *is not out of* A *with a definitely visible edge.* This again is not hard to verify. We leave the details to the reader. **Q.E.D.**

To conclude, if there is an m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{M}$ that is *into* A *or out of* A *with an invisible directed edge*, we know that its corresponding path in $\mathcal{P}$ is not out of A with a definitely visible edge, because $\mathcal{P}$ is a CPAG of $\mathcal{M}$. So there exists a path that satisfies the condition in **lemma 2'**, and hence there exists a definite m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{P}$ that is not out of A with a definitely visible edge. □

Proof of Lemma 5.1.8

Proof. This lemma is relatively easy. A path definitely m-connecting A and B given $\mathbf{C}$ in $\mathcal{P}$ is m-connecting in any DMAG represented by $\mathcal{P}$, which is an immediate consequence of the definition of CPAG. 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 $A \to D$ in $\mathcal{P}$, there exists a DMAG

[19] Such an m-connecting path is named *minimal* in, for example, Richardson and Spirtes (2002).

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 P is not $A \to D$, there exists a DMAG represented in $\mathcal{P}$ in which the edge is not $A \to D$, which follows from the completeness of P. Consider the case where the edge in $\mathcal{P}$ is $A \to D$, but it is not definitely visible. Recall that we can orient $\mathcal{P}$ into a DMAG by first tail augmenting (Definition 3.3.1) $\mathcal{P}$ and orienting the circle component into a DAG with no unshielded colliders. Moreover, it is easy to show that using, for example, Meek's algorithm we can orient the circle component into a DAG free of unshielded colliders in which every edge incident to A is oriented out of A. Let the resulting DMAG be $\mathcal{M}$. We show that $A \to D$ is invisible in $\mathcal{M}$. Suppose for 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 $\mathcal{P}$. Thus if the edge between E and E_1 is also into E_1 in $\mathcal{P}$, then the collider path appears in $\mathcal{P}$, which implies that every E_i $(1 \leq i \leq m)$ is also a parent of D in $\mathcal{P}$, and hence $A \to D$ is definitely visible in $\mathcal{P}$, a contradiction. So the edge between E and E_1 is not into E_1 in $\mathcal{P}$, but 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 property **CP1**, implies that $E \leftrightarrow E_2$ is in $\mathcal{P}$. Obviously $\langle E, E_2, \cdots, A\rangle$ makes $A \to D$ definitely visible in $\mathcal{P}$, which is a contradiction. $\square$

Proof of Lemma 5.1.9

Proof. This is an easy lemma. 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, 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 CPAG for $\mathcal{M}$ in which the path is out of A. $\square$

Proof of Lemma 5.1.10

Proof. This is again trivial. 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 DMAG represented by $\mathcal{P}$ in which the edge is out of A (which follows from the completeness of CPAG). Such a DMAG obviously satisfies the lemma. $\square$

Proof of Lemma 5.2.1

Proof. 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 $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. The edge between S and T in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ is either a bi-directed edge or an invisible directed edge.

Regarding (1), 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 DMAG of $\mathcal{G}$. Since $P \notin \mathbf{Y}$ and $Q \notin \mathbf{X}$, $P \to Q$ is also present in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$.

For (2), 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 DMAG of $\mathcal{G}$. The edge in $\mathcal{M}$ is either $S \leftrightarrow T$ or, by Lemma 5.1.1, an invisible directed edge ($S \leftarrow T$ or $S \to T$). Now since $S, T \notin \mathbf{X}$, 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 contradiction that it is visible, 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 DMAG do not create new directed edges. Also, R and T are not adjacent in $\mathcal{M}$ either, for otherwise the edge would be $R \to T$. Note that $R \notin \mathbf{Y}$ because otherwise $R \to S$ would be deleted or changed into a bi-directed edge; and $T \notin \mathbf{X}$ because otherwise $S \to T$ would be deleted. 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 $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, $R \to T$ is visible but $R \to S$ is invisible in $\mathcal{M}$. So 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 $R \to S$, from which it is not hard to derive that either $Q {*\!\!\to} S$ is in $\mathcal{M}$ or there is a vertex P on the collider path such that $P \leftrightarrow S$ is in $\mathcal{M}$.[20] In either case, $S \to T$ is visible in $\mathcal{M}$.

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 $Q \to T$ 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

[20] We have seen the precise argument for this several times in Chapter 4.

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 $S \to T$ 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 $R \to T$. 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 $\mathcal{M}$. For otherwise the edge incident to R on c has to be bi-directed as well, and hence if $R \leftrightarrow T$ is deleted, either the edge incident to R on c or the edge $S \to T$ should be deleted in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, which is a contradiction. So the edge is $R \to T$ in $\mathcal{M}$. Since $T \notin \mathbf{X}$ (for otherwise $S \to T$ would be deleted), $R \in \mathbf{Y}$, and $R \to T$ is visible in $\mathcal{M}$. But then it is easy to see that $S \to T$ is also visible in $\mathcal{M}$.

To summarize, we have shown that if $S \to T$ 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 we have established that (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 $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. The edge between S and T in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ is either a bi-directed edge or an invisible directed edge.

We are now ready to prove the lemma. Our strategy 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. We will further simplify the proof by noting that we do not need to require the transformation to preserve DMAGness, and all we need to guarantee is that each mark change results in an inducing path graph (IPG), which allows almost directed cycles. Interested readers can check the Appendix on IPGs, but the point here is that a single mark change (from $\to$ to $\leftrightarrow$) preserves Markov equivalence if the conditions (t2) and (t3) in Lemma 4.4.1 hold.[21]

Let us sketch the proof. 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 $P \to Q$ in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ (with $P \leftrightarrow Q$ in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$),

[21] This relaxation is unnecessary, but simplifies the proof greatly.

(2) implies that $P \to Q$ is invisible. It is then easy to check that conditions (t2) and (t3) in Lemma 4.4.1 hold for $P \to Q$ in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$, and thus it can be changed into $P \leftrightarrow Q$ while preserving Markov equivalence (though not necessarily DMAGness). Furthermore, it is not hard to check that 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 mixed graph (actually an IPG) that is a supergraph of $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$. Denote the supergraph by $\mathcal{I}$. We can immediately conclude 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 the supergraph of $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$, $\mathcal{I}$. 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 $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. □

Proof of Lemma 5.2.3

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. □

Chapter 6

Conclusions and Future Projects

There are, as always, many problems left untouched. Indeed, associated with every issue raised and every problem solved in this dissertation, there are further issues and open problems to be addressed. I will end with a review of what has been done and a preview of what is to be done.

The main subject of my dissertation is about the extent to which probabilistic independence and dependence relations among a set of variables can inform us about causal relations among them, and to a lesser degree, about the extent to which a limited amount of acquired causal information can be used to predict consequences of interventions. The investigation is based on two popularly adopted though also widely debated assumptions that link causality to probability, the Causal Markov Condition and the Causal Faithfulness Condition. However, for the most part, I do not assume that the two conditions as they are usually formulated apply directly to the set of observed variables whose probabilistic relations are inferrable from available data. Those variables alone may be causally insufficient. I do assume nonetheless that there is a bigger system containing the given observed variables to which the two conditions apply.

The major theme is to explore the full consequences of the Causal Markov and Faithfulness Conditions for inferring causal relations among a set of observed variables not known to be causally sufficient from facts of conditional independence and dependence among them. The primary results on this are presented in chapters 3 and 4. In a nutshell, it is shown that the FCI inference procedure, though not complete in deriving all common orientations shared by all MAGs Markov equivalent to the (unknown) true causal MAG, is complete in deriving all valid arrowheads, and hence statements about "non-cause", i.e., statements of the form: "A is not a cause of B", given a perfect oracle of conditional independence relations. Furthermore the FCI algorithm can be supplemented by a few additional inference rules so that the resulting Augmented FCI algorithm is complete with respect to all common orientations, both arrowheads and tails. The results are actually obtained in a more general setting than causal inference with possibly causal insufficiency — possible presence of selection effects is also considered. The general results lead to a characterization of Markov equivalence classes of ancestral graphical models, which may be significant in its own right for statisticians interested in graphical models.

At least two related problems, however, are left open and worth further investigation. First, the completeness result is with respect to causal sentences that are true in all causal models compatible with the Causal Markov and Faithfulness conditions alone. So implicitly it is assumed that no prior information whatsoever about the causal structure is available, and every causal model in a Markov equivalence class is possibly true. This aptly captures circumstances where only passive observation is possible, but it is not entirely adequate when some experimental control or prior causal knowledge is also available. In the semi-experimental case, extra axioms may be added stipulating certain constraints on causal structure, and validity should be judged with regard to the Causal Markov and Faithfulness conditions plus these extra axioms. If so, the Augmented FCI procedure, though obviously still sound, is probably not complete. What further inference rules are needed is an open question.

Second, the completeness result is with respect to valid causal sentences entailed by (conditional) independence and dependence facts. But dependence and independence are presumably not the only sort of probabilistic "facts" that are informative about causation. In fact, it is well known that certain causal DAGs with latent variables entail testable constraints on the marginal probability of observed variables, sometimes referred to as the Verma constraints (see Tian and Pearl 2002 for an illuminating discussion). It is both theoretically intriguing to study the nature of such "extra" evidence and practically significant to explore methodologies that may exploit such evidence.

As mentioned earlier, the approach the FCI procedure follows is referred to as the independence-constraint-based approach, as contrasted to the score-based approach (sometimes also called the Bayesian approach, though a score-based approach could also be non-Bayesian). The constraint-based approach is clearly divided into two parts: statistical inference from data and causal inference from probability, the latter of which is probably of more interest to philosophers, whereas the score-based approach does not admit a clear-cut division of this sort. From a practical point of view, however, the score-based approach is in general more stable than the constraint-based approach on small or moderate sample sizes.[1] It is thus desirable to have score-based inference procedures developed for causal inference as well.

If the set of observed variables is known to be causally sufficient (so that the

[1]In addition, the score-based approach has the advantage of always returning a "legitimate" object, unlike, for example, the PC algorithm that often returns an object with "illegitimate" bi-directed edges or directed cycles. Also, score-based algorithms can return a set of objects of high scores rather a single object. Both of these, however, may be achieved within the constraint-based approach as well, I think.

true causal structure is properly represented by a DAG), quite a few score-based algorithms have been proposed in the literature. One that is particularly relevant here is the Greedy Equivalence Search (GES) algorithm that, just as the PC algorithm, also aims to infer a Markov equivalence class of DAGs and is also provably sound in the large sample limit. There are good reasons to expect that the GES algorithm as well as its justification can be well extended to the case of causal inference without assuming causal sufficiency.

Theoretically, the transformational characterization of Markov equivalence for DMAGs given in Chapter 4 is a first, though quite small, step towards justifying a GES-style algorithm for inferring causal MAGs. A big remaining gap is to extend the transformational result to cover not only equivalent models, but independence sub-models as well. Practically, there is already a well-developed Gaussian parameterization of MAGs and an efficient algorithm to fit Gaussian MAGs using maximum likelihood estimation (Drton and Richardson 2004), so we know how to score a Gaussian MAG. We also know how to efficiently turn a CPAG into a representative MAG (with fewest bi-directed and un-directed edges) given the results in Chapters 3 and 4. We do not yet know, however, how to efficiently traverse the search space, a space of CPAGs. This is an interesting and conceivably challenging project from an algorithmic point of view.

A lot more should have been done in Chapter 5. The graphical criterion for invariance by itself can only identify a special class of inferrable post-intervention probabilities, i.e., those probabilities that remain the same before and after the intervention. How to exploit that criterion to identify more general inferrable post-intervention quantities is an interesting and important issue. In some sense, Pearl's *do*-calculus does exactly that. Upon careful inspection, the rules in *do*-calculus are based on just invariance (and probability calculus), but an iterated application of the rules could pick out all other inferrable quantities (see Shpister and Pearl 2006, Huang and Valtorta 2006 for proofs of the completeness of the *do*-calculus).

From an algorithmic point of view, however, *do*-calculus is far less efficient than a largely algebraic method recently developed to identify intervention effects given a single causal DAG with latent variables (Tian and Pearl 2004, see also Huang and Valtorta 2006). Whether that method can be adapted to the case where only a CPAG is given is probably worth investigating.

Although the primary attention of this dissertation was devoted to studying the consequences of the Causal Markov and Faithfulness conditions, I did include some discussion of the testability of the Causal Faithfulness conditions in Chapter 2. A distinction between "(asymptotically) detectable" violations of faithfulness and "undetectable" violations of faithfulness can be made assuming the Causal Markov condition holds. In the context of inferring causal DAGs, for example, the condition can be decomposed into at least two parts, and one part, the Orientation-Faithfulness condition, is in principle testable given the other, the Adjacency-Faithfulness condition. This discussion by no means exhausts the issue of characterizing detectable violations of faithfulness. Some violations of the Adjacency-Faithfulness condition, for example, are also detectable. It is desirable to have a neat characterization of all detectable violations of faithfulness, through which we can gain a good understanding of how much an empiricist has to concede to avoid being a radical skeptic about causal inference.

Conversely, it is also shown, albeit in a very preliminary fashion, how certain failure of the Causal Markov condition may be mitigated if we assume the Causal Faithfulness condition holds. I believe there is more to be said about this point, which may serve as a pragmatic basis for assuming the Causal Markov condition.

These theoretical results are not without practical implications. A modification of the familiar PC algorithm suggested by the testability result was studied empirically. It turns out that the modified algorithm, called Conservative PC, significantly improves accuracies. A conjecture for why this is the case is that the conservative version rightly suspends judgments in close-to-unfaithful situations, which thus avoids some errors the PC algorithm is very liable to.[2]

All these, however, are established in the context of causal inference with causal sufficiency. The following question is thus unavoidable given the main concern of this dissertation: can the results be extended to causal inference without causal sufficiency? There seems to be good reason for answering "yes". At least the conservative modification of the PC algorithm can be immediately carried over to the (augmented) FCI algorithm. For example, we can modify $\mathcal{R}0$ and $\mathcal{R}4$ of the FCI to the effect that the judgment of collider or non-collider relies on more conditional independence facts and the judgment may be neither given certain combination of conditional independence facts. Furthermore, the relevant Adjacency-Faithfulness and Orientation-Faithfulness should probably be formulated in terms of MAGs. Whether conservative FCI is sound under a weaker faithfulness condition and works empirically better awaits careful examination.

---

[2]This seems to be related to a big issue I wish I could have spent some time on. The issue is about what kind of reliability — in terms of the notion of *consistency* in statistics — is achievable in causal inference. Robins et al. (2003) argued, based on some canonical cases, that causal inference can at best be pointwise consistent, but not uniformly consistent under the Causal Markov and Faithfulness conditions. Zhang and Spirtes (2003) showed that under a slightly strengthened version of faithfulness, which is arguably what some social scientists implicitly assume anyway, uniform consistency is achievable in those canonical cases. However, no general uniformly consistent procedure has been proposed. Now, if the CPC algorithm really works by placing right caution on close-to-unfaithful cases, it seems to be (close to) a uniformly consistent procedure under the strengthened faithfulness assumption. I thank Peter Spirtes for pointing this out.

Appendix: Inducing Path Graphs

Much of this dissertation is rooted in the seminal work by Peter Spirtes, Clark Glymour and Richard Scheines (1993/2000, chapters 6 and 7), which is based on a graphical representation called inducing path graphs. This representation 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 directed MAGs (DMAGs). In this appendix we specify the exact relationship between them. In particular, we give an independent syntactic definition of inducing path graphs, which makes it clear that syntactically the class of DMAGs is a subclass of inducing path graphs.

An *inducing path graph (IPG)* is a directed mixed graph, defined relative to DAG $\mathcal{G}(\mathbf{O}, \mathbf{L})$ 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 a 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 is 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.[3] So $\mathcal{I}_{\mathcal{G}}$, just like the MAG for $\mathcal{G}$, represent both the conditional independence relations and (features of) the causal structure among the observed variables $\mathbf{O}$. Since the above construction obviously produces a unique graph given a DAG $\mathcal{G}$, we 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 6.0.1. *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 $\mathcal{I} = \mathcal{I}_{\mathcal{G}}$, i.e., $\mathcal{I}$ 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.

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 $c =$

[3] We will not elaborate upon the exact causal interpretation of IPGs here, which is not relevant to our purpose. See Spirtes et al. 1993/2000 for details.

$\langle O_1, \cdots, O_n, O_1 \rangle$, 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 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 $p = \langle A, O_1, \cdots, O_n, B \rangle$ 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. It is easy to check that joining all these paths together makes 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 $A \leftarrow L_{AB} \rightarrow B$.[4] 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 $\mathcal{I}$, 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 $\mathcal{I}_{\mathcal{G}}$.

[4] This is named *canonical DAG* in Richardson and Spirtes (2002).

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 $\langle A, O_1, \cdots, O_n, B\rangle$ 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 $\langle A, O_1, \cdots, O_n, B\rangle$ 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 $\mathcal{G}$. $\square$

Given this theorem, it is clear that we can define IPGs in terms of $(i1)$ and $(i2)$. So a DMAG is also an IPG, but an IPG is not necessarily a DMAG, as the former may contain an almost directed cycle. The simplest IPG which is not a DMAG is shown in Figure 6.1.

Spirtes et al. (1993/2000) 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

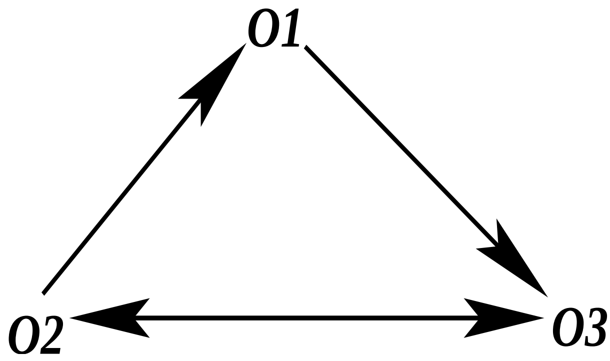

Figure 6.1: An IPG that is not a DMAG

class of IPGs. An obvious fact is that given a set of conditional independence facts that admits a faithful representation by a DMAG, the Markov equivalence class of DMAGs is included in the Markov equivalence class of IPGs. It follows that the complete POIPG cannot contain more informative marks than the CPAG. In fact we can be more precise about this. An arrowhead is in the complete POIPG if and only if it is in the CPAG, and a tail is in the complete POIPG if and only if it is a visible tail (directed edge) in the CPAG. To see the first fact, just notice that if an arrowhead appears in the true DMAG, it must also appear in the true IPG, as a tail in the latter would also imply an ancestral relation. It then follows that if an arrowhead is shared by all DMAGs in an Markov equivalence class, all IPGs Markov equivalent to these MAGs will also contain this arrowhead. To see the second fact, recall that a directed edge $A \rightarrow B$ in a DMAG implies that there is no latent common ancestor of A and B in any DAG represented by the DMAG if and only if the directed edge is visible. From this we can derive that a directed edge is in every IPG in a Markov equivalence class if and only if it is visible. We hope the idea is clear enough and will not take further pain to present the whole rigorous argument.

Thus a CPAG can reveal invisible tails whereas a POIPG cannot. Otherwise they give the exact same informative marks. Notice that all arrowheads and visible tails can be inferred from the FCI algorithm, so the FCI algorithm is complete when the output is interpreted as a POIPG. This answers the completeness question posed in Spirtes et al. (1993/2000) as well as in Neapolitan (2004).

Bibliography

Ali, R.A. (2002) Applying Graphical Models to Partially Observed Data Generating Processes. Department of Statistics, University of Washington, PhD Thesis.

Ali, R.A., T. Richardson, and P. Spirtes (2004) Markov Equivalence for Ancestral Graphs. Department of Statistics, University of Washington, Technical Report 466.

Ali, R.A., T. Richardson, P.Spirtes and J.Zhang (2005) Towards Characterizing Markov Equivalence Classes for Directed Acyclic Graphs with Latent Variables, in *Proceedings of the 21th Conference on Uncertainty in Artificial Intelligence*, Oregon: AUAI Press, pp.10-17.

Andersson, S., D. Madigan, and M. Pearlman (1997) A Characterization of Markov Equivalence Classes of Acyclic Digraphs. *Annals of Statistics* 25(2): 505-541.

Artzenius, F. (1992) The Common Cause Principle, *PSA Procceding*, Eds. D. Hull and K. Okruhlik, Vol.2, East Lansing, MI: PSA, pp. 227-37.

Becker, A., D. Geiger, and C. Meek (2000) Perfect Tree-like Markovian Distributions. *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence*, 19-23 Morgan Kaufmann.

Belnap, N. (2005) A theory of causation: Causae causantes (originating causes) as inus conditions in branching space-times, *British Journal of the Philosophy of Science*, vol. 56, pp. 221-253.

Berkovitz, J. (2002) On Causal Inference in Determinism and Indeterminism, in *Between Chance and Choice: Interdisciplinary Perspectives on Determinism*, Eds. H. Atmanspacher and R. Bishop. Charlottesville, VA: Imprint Academic, pp. 237-78.

Bickel, P.J., and K.A. Doksum (2001) *Mathematical Statistics - Basic Ideas and Selected Topics, VOL. 1* (2nd ed.). New Jersey: Prentice Hall.

Bollen, K. A. (1989) *Structural Equations with latent Variables.* New York: Wiley.

Cartwright, N. (1999) *The Dappled World.* Cambridge: Cambridge University Press.

Cartwright, N. (2004) Causation: One Word, Many Things. *Philosophy of Science*

71: 805-819.

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

Chickering, D.M. (2002) Optimal Structure Identification with Greedy Search. *Journal of Machine Learning Research*, 3:507-554.

Dawid, P. (1979) Conditional Independence in Statistical Theory. *Journal of the Royal Statistical Society, Series B* 41: 1-31.

Dawid, P. (2000) Causal inference without counterfactuals. *Journal of the American Statistical Association*, 95: 407-448.

Dawid, P. (2002) Influence Diagrams for Causal Modelling and Inference. *International Statistical Review* 70: 161-189.

Drton, M., and T. Richardson (2003). Iterative Conditional Fitting for Gaussian Ancestral Graph Models. Department of Statistics, University of Washington, Tech Report 437.

Drton, M., and T. Richardson (2004) Graphical Answers to Questions About Likelihood Inference in Gaussian Covariance Models. Department of Statistics, University of Washington, Tech Report 467.

Drton, M., and T. Richardson (2005) Binary Models for Marginal Independence. Department of Statistics, University of Washington, Tech Report 474.

Enders, W. (2003) *Applied Econometric Time Series*, 2nd ed. Danvers, MA: John Wiley.

Fisher, F.M. (1970) A Correspondence Principle for Simultaneous Equation Models. *Econometrica* 38: 73-92.

Fisher, R. A. (1935) *The Design of Experiments.* Edinburgh: Oliver and Boyd.

Frederick, E., C. Glymour, and R. Scheines (2005) On the Number of Experiments Sufficient and in the Worst Case Necessary to Identify All Causal Relations Among N Variables, in *Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence*, Oregon: AUAI Press, pp. 178-184.

Geiger, D., D. Heckerman, H. King, and C. Meek (2001) Stratified Exponential Families: Graphical Models and Model Selection. *Annals of Statistics* 29, pp. 505-529.

Glymour, C., R. Scheines, P. Spirtes, and K. Kelly (1987) *Discovering Causal Structure.* San Diego, California: Academic Press.

Goldberger, A. S. (1972) Structural equation methods in the social sciences. *Econometrica*, 40:979-1001.

Hacking, I. (1965) *Logic of Statistical Inference.* Cambridge: Cambridge University Press.

Haughton, D.A. (1988) On the choice of a model to fit data from an exponential family, in *The Annals of Statistics*, 16(1):342-355.

Hausman, D. M., and J. Woodward (1999) Independence, Invariance and the Causal Markov Condition. *British Journal for the Philosophy of Science* 50, pp. 521-83.

Hausman, D. M., and J. Woodward (2004) Manipulation and Causal Markov Condition. *Philosophy of Science* 71: 846-856.

Heckerman, D., D. Geiger, and D. Chickering (1995) Learning Bayesian networks: the combination of knowledge and statistical data. *Machine Learning* 20(3):197-243.

Heckerman, D., C. Meek, and G.F. Cooper (1999) A Bayesian Approach to Causal Discovery. *Computation, Causation, and Discovery.* Eds. C. Glymour and G.F. Cooper. Cambridge, MA: MIT Press.

Hitchcock, C. (1996) The Role of Contrast in Causal and Explanatory Claims. *Synthese*, 107, pp. 395-419.

Hitchcock, C. (2001) The intransitivity of Causation Revealed in Equations and Graphs. *Journal of Philosophy*, 98, pp. 273-299.

Holloran M.E. and C. J. Struchiner (1995) Causal Inference in Infectious Diseases. *Epidemiology*, 6(2):142-51.

Hoover, K.D. (2001) *Causality in Macroeconomics*, Cambridge: Cambridge University Press.

Hoover, K.D. (2003) Nonstationary Time Series, Cointegration, and the Principle of the Common Cause. *British Journal for the Philosophy of Science* 54, pp. 527-551.

Huang, Y., and M. Valtorta (2006) On the Completeness of an Identifiability Algorithm for Semi-Markovian Models. Tech Report, University of South Carolina, Department of Computer Science.

Huang, Y., and M. Valtorta (2006) Pearl's Calculus of Intervention Is Complete. *Proceedings of 22nd Conference on Uncertainty in Artificial Intelligence*, 217-24, Oregon: AUAI Press.

Hume, D. (1739) *A Treatise on Human Nature*. London: John Noon.

Hume, D. (1748/1984) *An Enquiry Concerning Human Understanding*, ed. C. Hendell. New York: Collier Macmillan.

Jordan, M. (1998) *Learning in Graphical Models*, Dordrecht: Kluwer.

Kelly, K. (1996) *The Logic of Reliable Inquiry.* New York: Oxford University Press.

Koster, J. (1996) Markov Properties of Nonrecursive Causal Models, *Annals of Statistics* 24: 2148-78.

Lauritzen, S.L. (1996) *Graphical Models.* Oxford, UK: Clarendon.

Lewis, D. (2000) Causation as Influence. *Journal of Philosophy* 97, pp. 182-197.

McDermott, M. (1995) Redundant Causation. *British Journal for the Philosophy of Science* 40, pp. 523-544.

Meek, C. (1995) Causal Inference and Causal Explanation with Background Knowledge. *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, 403-411. Morgan Kaufmann.

Meek, C. (1996) Graphical Models: Selecting Causal and Statistical Models. Carnegie Mellon University, Philosophy Department, PhD Thesis.

Meek, C., and C. Glymour (1994) Conditioning and Intervening. *British Journal for the Philosophy of Science* 45, pp. 1001-21.

Mill, J.S. (1843) *Systems of Logic*, vol.1. London: John Parker.

Murphy, K. P. (2001) Active Learning of Causal Bayes Net Structure, Technical Report, Department of Computer Science, UC Berkeley.

Neapolitan, R.E. (2004) *Learning Bayesian Networks.* Upper Saddle River, NJ: Prentice Hall.

Papineau, D. (1992) Can We Reduce Causal Direction to Probabilities. *PSA Procceding*, Eds. D. Hull and K. Okruhlik, Vol.2, East Lansing, MI: PSA, pp. 238-52.

Papineau, D. (1994) The Virtues of Randomization. *British Journal for the Philosophy of Science* 45, pp. 437-50.

Pearl, J. (1995) Causal Diagrams for Empirical Research. *Biometrika* 82:669-710.

Pearl, J. (1998) Graphs, Causality and Structural Equation Models. *Sociological Methods and Research* 27: 226-284.

Pearl, J. (1988) *Probabilistic Reasoning in Intelligent Systems.* San Mateo, California: Morgan Kaufmann.

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

Pearl, J., and R. Dechter (1996) Identifying independencies in causal graphs with feedback. *Proceedings of the Twelfth Annual Conference on Uncertainty in Artificial Intelligence*, 420-426. San Francisco: Morgan Kaufmann.

Pearl, J. and T. Verma. (1991) A Theory of Inferred Causation. *Proceedings of the Second International Conference of Representation and Reasoning*, San Francisco: Morgan Kaufmann.

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

Ramsey, J., P. Spirtes, and J. Zhang (2006) Adjacency-Faithfulness and Conservative Causal Inference. *Proceedings of 22nd Conference on Uncertainty in Artificial Intelligence*, 401-408, Oregon, AUAI Press.

Reichenbach, H. (1956) *The Direction of Time.* Berkeley: University of California Press.

Reid, T. (1788/1986) *Essays on the Active Powers of Man*. Lincoln-Rembrandt Publisher.

Richardson, T. (1996) Models of Feedback: Interpretation and Discovery. Carnegie Mellon University, Philosophy Department, PhD Thesis.

Richardson, T. (1998) Chain Graphs and Symmetric Associations, in *Learning in Graphical Models*, Ed. M. Jordan, Dordrecht: Kluwer.

Richardson, T. (2003) Markov Properties for Acyclic Directed Mixed Graphs. *Scandinavian Journal of Statistics* Vol. 30 Issue 1 Page 145.

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

Richardson, T., and P. Spirtes (2003) Causal Inference via Ancestral Graph Models, in *Highly Structured Stochastic Systems*. Eds. P. Green, N. Hjort, and S. Richardson. Oxford University Press.

Robins, J. (1986) 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.

Robins, J.M., Scheines, R., Spirtes, P., Wasserman, L. (2003) Uniform Consistency in Causal Inference. *Biometrika* 90(3):491-515.

Rosenbaum, P. (1995) *Observational Studies*. New York: Springer-Verlag.

Rubin, D. B. (1986) Which Ifs Have Causal Answers? *Journal of the American Statistical Association*, 81, 961-62.

Shpitser, I., and J. Pearl (2006) Identification of Conditional Interventional Distributions. *Proceedings of 22nd Conference on Uncertainty in Artificial Intelligence*, 437-44, Oregon: AUAI Press.

Silva, R., R. Scheines, C. Glymour and P. Spirtes (2003) Learning Measurement Models for Unobserved Variables. *Proceedings of the Nineteenth Conference on Uncertainty in Artificial Intelligence*, 543-555. Morgan Kaufmann.

Silva, R., R. Scheines, C. Glymour and P. Spirtes (2006) Learning the structure of linear latent variable models. *Journal of Machine Learning Research* 7:191–246.

Sober, E. (1987) The Principle of the Common Cause, *Probability and Causality*. J. Fetzer. Dordrecht: D.Reidel.

Sober, E. (2001) Venetian Sea Levels, British Bread Prices, and the Principle of the Common Cause. *British Journal for the Philosophy of Science*, 52: 1-16.

Spanos, A. (2006) Revisiting the Omitted Variables Argument: Substantive vs. Statistical Adequacy, forthcoming in *Journal of Economic Methodology*.

Spirtes, P. (1995) Directed cyclic graphical representation of feedback models. *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, 491-498. San Francisco: Morgan Kaufmann.

Spirtes, P., C. Glymour, and R. Scheines (1993/2000) *Causation, Prediction and Search*. (1993, 1st ed.) New York: Springer-Verlag. (2000, 2nd ed.) Cambridge, MA: MIT Press.

Spirtes, P., C. Meek, and T. Richardson (1999) An Algorithm for Causal Inference in the Presence of Latent Variables and Selection Bias, in *Computation, Causation, and Discovery*. Eds. C. Glymour and G.F. Cooper. Cambridge, MA: MIT Press.

Spirtes, P., and T. Richardson (1996) A Polynomial Time Algorithm For Determining DAG Equivalence in the Presence of Latent Variables and Selection Bias. *Proceedings of the 6th International Workshop on Artificial Intelligence and Statistics*.

Spirtes, P., Richardson, T., Meek, C. (1997) Heuristic Greedy Search Algorithms for Latent Variable Models. *Proceedings of the 6th International Workshop on Artificial Intelligence and Statistics*.

Spirtes, P., and T. Verma (1992) Equivalence of Causal Models with Latent Variables. Carnegie Mellon Univeriesty Philosophy Department Technical Report Phil-36.

Steck, H. (2001) Constraint-Based Structural Learning in Bayesian Networks using Finite Data Sets. Technical University of Munich, Department of Computer Science, PhD Thesis.

Steck, H. and V. Tresp (1996) Bayesian Belief Networks for Data Mining. *Proceedings of the 2nd Workshop on Data Mining and Data Warehousing*, 145-154, University of Magdeburg, Germany.

Steel, D. (2003) Making Time Stand Still: A Response to Sober's Counter-Example to the Principle of the Common Cause, *Brit. J. Phil. Sci.*, 54, pp. 309-17.

Steel, D. (2005) Indeterminism and the Causal Markov Condition. *British Journal for the Philosophy of Science* 56, pp. 3-26.

Strevens, M. (2003) Against Lewis's New Theory of Causation. *Pacific Philosophical Quarterly* 84, pp. 398-412.

Strotz, R.H., and H.A. Wold (1960) Recursive versus Nonrecursive Systems: An Attempt at Synthesis. *Econometrica* 28: 417-427

Tian, J., and J. Pearl (2004) On the Identification of Causal Effects. Tech Report, Iowa State University, Department of Computer Science.

Tian, J., and J. Pearl (2002) On the Testable Implications of Causal Models with Hidden Variables. *Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence*, 519-552. Morgan Kaufmann.

Verma, T., and J. Pearl (1990) Equivalence and Synthesis of Causal Models. *Proceedings of 6th Conference on Uncertainty in Artificial Intelligence*, 220-227.

Wiley, D.E. (1973) The identification problem for structural equation models with unmeasured variables, in *Structural Equation Models in the Social Sciences.* Eds. A.S. Goldberger and O.D.Duncan. New York: Seminar Press.

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

Whittaker, J. (1990) *Graphical Models in Applied Multivariate Statistics.* Chichester, U.K.: Wiley.

Woodward, J (2003) *Making Things Happen: A Theory of Causal Explanation.* Oxford and New York: Oxford University Press.

Wright, S. (1921) Correlation and Causation. *Journal of Agricultural Research* 20: 557-585.

Wright, S. (1934) The method of path coefficients. *Annals of Mathematical Statistics* 5:161-215.

Yule, G. U. (1903) Notes on the Theory of Association of Attributes in Statistics. *Biometrika*, 2, 121-34.

Yule, G. U. (1926) Why Do We Sometimes Get Nonsense-correlations between Time-series? A Study in Sampling and the Nature of Time-series. *Journal of Royal Statistical Society*, 89, 1-63.

Zhang, J. (2002) Consistency of Causal Inference Under a Variety of Background Assumptions. Carnegie Mellon University, Philosophy Department, Master Thesis.

Zhang, J. and Spirtes, P. (2003) Strong Faithfulness and Uniform Consistency in Causal Inference. *Proceedings of the Nineteenth Conference on Uncertainty in Artificial Intelligence*, 632-639. Morgan Kaufmann.