The Annals of Statistics
2015, Vol. 43, No. 3, 1060–1088
DOI: 10.1214/14-AOS1295

A GENERALIZED BACK-DOOR CRITERION[1]

BY MARLOES H. MAATHUIS AND DIEGO COLOMBO

ETH Zurich

We generalize Pearl's back-door criterion for directed acyclic graphs (DAGs) to more general types of graphs that describe Markov equivalence classes of DAGs and/or allow for arbitrarily many hidden variables. We also give easily checkable necessary and sufficient graphical criteria for the existence of a set of variables that satisfies our generalized back-door criterion, when considering a single intervention and a single outcome variable. Moreover, if such a set exists, we provide an explicit set that fulfills the criterion. We illustrate the results in several examples. R-code is available in the R-package `pcalg`.

1. Introduction. Causal Bayesian networks are widely used for causal reasoning [e.g., Glymour et al. (1987), Koller and Friedman (2009), Pearl (1995, 2000, 2009), Spirtes, Glymour and Scheines (1993, 2000)]. In particular, if the causal structure is known and represented by a directed acyclic graph (DAG), this framework allows one to deduce post-intervention distributions and causal effects from the pre-intervention (or observational) distribution. Hence, if the causal DAG is known, one can estimate causal effects from observational data. Covariate adjustment is often used for this purpose. The *back-door criterion* [Pearl (1993)] is a graphical criterion that is sufficient for adjustment, in the sense that a set of variables can be used for covariate adjustment if it satisfies the back-door criterion for the given graph.

In practice, there are two important complications. First, the underlying DAG may be unknown. In this case one can try to estimate the DAG, but in general one cannot identify the underlying DAG uniquely. Instead, one can identify its Markov equivalence class, which consists of all DAGs that encode the same conditional independence relationships as the underlying DAG. Such a Markov equivalence class can be represented uniquely by a different type of graph, called a completed partially directed acyclic graph (CPDAG) [Andersson, Madigan and Perlman (1997), Meek (1995), Spirtes, Glymour and Scheines (1993)]. Second, it is often the case that some important variables were not measured, meaning that we do not have causal sufficiency. In this case, one can work with maximal ancestral graphs (MAGs) instead of DAGs [Richardson and Spirtes (2002, 2003)]. Finally,

---

Received July 2013; revised November 2014.

[1]Supported in part by Swiss NSF Grant 200021-129972.

MSC2010 subject classifications. 62H99.

Key words and phrases. Causal inference, covariate adjustment, hidden confounders, DAG, CPDAG, MAG, PAG.

the underlying MAG may be unknown, so that it must be estimated from data. Again, there is an identifiability problem here, as we can generally only identify the Markov equivalence class of the underlying MAG, which can be represented uniquely by a partial ancestral graph (PAG) [Ali, Richardson and Spirtes (2009), Richardson and Spirtes (2002)].

In this paper, we therefore consider generalizations of the back-door criterion to the following three scenarios:

(1) we assume causal sufficiency, and we only know the CPDAG, that is, the Markov equivalence class of the underlying DAG;

(2) we do *not* assume causal sufficiency, and we know the MAG on the observed variables;

(3) we do *not* assume causal sufficiency, and we only know the PAG, that is, the Markov equivalence class of the underlying MAG on the observed variables.

In scenarios 2 and 3, we allow for arbitrarily many hidden (or unmeasured) variables. We do not, however, allow for selection variables, that is, for unmeasured variables that determine whether a unit is included in the sample.

Since the back-door criterion is a simple criterion that is widely used for DAGs, it seems useful to have similar criteria for CPDAGs, MAGs and PAGs. We also hope that our generalized back-door criterion will make working with MAGs and PAGs less daunting, and more accessible to people in practice.

Our generalized back-door criterion for DAGs, CPDAGs, MAGs and PAGs is given in Section 3; see especially Definition 3.7 and Theorem 3.1. Corresponding R-code is available in the function `backdoor` in the R-package `pcalg` [Kalisch et al. (2012)]. Our results are derived by first formulating invariance conditions that are sufficient for adjustment, and then using the graphical criteria for invariance derived by Zhang (2008a). We also show that the generalized back-door criterion is equivalent to Pearl's back-door criterion for single interventions in DAGs, and is slightly more general for multiple interventions in DAGs (Lemma 3.1 and Example 1). In Section 4, we give necessary and sufficient criteria for the existence of a set that satisfies the generalized back-door criterion relative to a pair of variables (X, Y) and a DAG, MAG, CPDAG or PAG. Moreover, if a generalized back-door set exists, we provide an explicit such set. These results are summarized in Theorem 4.1, using a general framework that covers DAGs, CPDAGs, MAGs and PAGs. Corollaries 4.1–4.3 specialize the results for DAGs, CPDAGs and MAGs, respectively. We illustrate our results with several examples in Section 5. All proofs are given in Section 7.

We close this introduction by discussing related work. For a given causal DAG, identifiability of causal effects in general or via covariate adjustment has been studied by various authors. In particular, there are complete graphical criteria for the identification of causal effects if a causal DAG with unmeasured variables is given [e.g., Huang and Valtorta (2006), Shpitser and Pearl (2006a, 2006b, 2008), Tian and Pearl (2002)]. Shpitser, Van der Weele and Robins (2010a, 2010b) studied effects that are identifiable via covariate adjustment, and provided necessary and sufficient graphical criteria for this purpose, again if the causal DAG is given. Their results can be viewed as an improvement on the back-door criterion, which is only sufficient for adjustment. Textor and Liśkiewicz (2011) studied covariate adjustment for a given DAG from an algorithmic perspective. Among other things, they showed that the back-door criterion and the adjustment criterion of Shpitser, Van der Weele and Robins (2010a) are equivalent if one is interested in minimal adjustment sets for a certain subclass of graphs. Van der Zander, Liśkiewicz and Textor (2014) extended these necessary and sufficient graphical criteria for covariate adjustment to MAGs.

There are also existing approaches that do not make the assumption that the causal DAG or MAG is given. The prediction algorithm [Spirtes, Glymour and Scheines (2000), Chapter 7] roughly starts from a PAG and uses invariance results. In this sense it is probably closest to our work. The main difference between this method and our results is that the prediction algorithm is more complex. In particular, it searches over all possible orderings of the variables, which quickly becomes infeasible for large graphs. The prediction algorithm may, however, be more informative, in the sense that certain distributions may be identifiable by the prediction algorithm but not by the generalized back-door criterion. Studying the exact relationship between these two approaches would be an interesting topic for future work.

Other work on data driven methods for selection of adjustment variables for the estimation of causal effects does not assume that the causal structure is known, but does make some assumptions about causal relationships between the variables of interest and/or about the existence of a set of variables that can be used for covariate adjustment [de Luna, Waernbaum and Richardson (2011), Entner, Hoyer and Spirtes (2013), VanderWeele and Shpitser (2011)]. In the current paper, we do not make any such assumptions. On the other hand, we start from a given DAG, CPDAG, MAG or PAG. We do not see this as a genuine restriction of our approach, however, since there are algorithms to estimate CPDAGs and PAGs from data (e.g., the PC algorithm [Spirtes, Glymour and Scheines (2000)], greedy equivalence search [Chickering (2002)] and versions of the FCI algorithm [Claassen, Mooij and Heskes (2013), Colombo et al. (2012), Spirtes, Glymour and Scheines (2000)]). These algorithms have been shown to be consistent, even in certain sparse high-dimensional settings [Colombo et al. (2012), Kalisch and Bühlmann (2007)]. In practice, one could therefore first employ such an algorithm, and then apply the results in the current paper.

2. Preliminaries.

Throughout this paper, we denote sets in a bold font (e.g., $\mathbf{X}$) and graphs in a calligraphic font (e.g., $\mathcal{D}$ or $\mathcal{M}$).

2.1. *Basic graphical definitions.* A graph $\mathcal{G} = (\mathbf{V}, \mathbf{E})$ consists of a set of vertices $\mathbf{V} = \{X_1, \dots, X_p\}$ and a set of edges $\mathbf{E}$. The vertices represent random variables, and the edges describe conditional independence and causal (ancestral) relationships. There is at most one edge between every pair of vertices, and the edge set $\mathbf{E}$ can contain (a subset of) the following four edge types: $\rightarrow$ (*directed*), $\leftrightarrow$ (*bi-directed*), $\circ\!\!-\!\!\circ$ (*nondirected*) and $\circ\!\!\rightarrow$ (*partially directed*). A *directed graph* contains only directed edges, a *mixed graph* can contain directed and bi-directed edges and a *partial mixed graph* can contain all four edge types. The endpoints of an edge are called *marks*, and they can be *tails*, *arrowheads* or *circles*. We use the symbol "$*$" to denote an arbitrary edge mark. If we are only interested in the presence or absence of edges, and not in the edge marks, then we refer to the *skeleton* of a graph.

Two vertices are *adjacent* if there is an edge between them. The adjacency set of a vertex X in $\mathcal{G}$, denoted by $\text{adj}(X, \mathcal{G})$, consists of all vertices adjacent to X in $\mathcal{G}$. A *path* is a sequence of distinct adjacent vertices. The *length of a path* $p = \langle X_i, X_{i+1}, \dots, X_{i+\ell}\rangle$ equals the corresponding number of edges, in this case ℓ. The path p is said to be *out of* (*into*) X_i if the edge between X_i and X_{i+1} has a tail (arrowhead) at X_i. A sub-path of p from X_j to $X_{j'}$ is denoted by $p(X_j, X_{j'})$. We denote the concatenation of paths by $\oplus$, so that, for example, $p = p(X_i, X_{i+k}) \oplus p(X_{i+k}, X_{i+\ell})$ for $k \in \{1, \dots, \ell - 1\}$. We use the convention that we remove any loops that may occur due to the concatenation, so that the result does not contain duplicate vertices and is again a path. The path p is a *directed path* from X_i to $X_{i+\ell}$ if for all $k \in \{1, \dots, \ell\}$, the edge $X_{i+k-1} \rightarrow X_{i+k}$ occurs, and it is a *possibly directed path* if for all $k \in \{1, \dots, \ell\}$, the edge between X_{i+k-1} and X_{i+k} is not into X_{i+k-1}. A *cycle* occurs if there is a path between X_i and X_j of length greater than one, and X_i and X_j are adjacent. A directed path from X_i to X_j forms a *directed cycle* together with the edge $X_j \rightarrow X_i$, and an *almost directed cycle* together with the edge $X_j \leftrightarrow X_i$. A *directed acyclic graph* (DAG) is a directed graph without directed cycles. An *ancestral* graph is a mixed graph without directed and almost directed cycles.

If $X_j \rightarrow X_i$, we say that X_i is a *child* of X_j, and X_j is a *parent* of X_i. The corresponding sets of parents and children are denoted by $\text{pa}(X_i, \mathcal{G})$ and $\text{ch}(X_i, \mathcal{G})$. If there is a (possibly) directed path from X_i to X_j or if $X_i = X_j$, then X_i is a (*possible*) *ancestor* of X_j and X_j a (*possible*) *descendant* of X_i. The sets of ancestors, descendants, possible ancestors, and possible descendants of a vertex X_i in $\mathcal{G}$ are denoted by $\text{an}(X_i, \mathcal{G})$, $\text{de}(X_i, \mathcal{G})$, $\text{possibleAn}(X_i, \mathcal{G})$, and $\text{possibleDe}(X_i, \mathcal{G})$, respectively. These definitions are applied disjunctively to a set $\mathbf{Y} \subseteq \mathbf{V}$, for example, $\text{an}(\mathbf{Y}, \mathcal{G}) = \{X_i | X_i \in \text{an}(X_j, \mathcal{G}) \text{ for some } X_j \in \mathbf{Y}\}$.

A path $\langle X_i, X_j, X_k\rangle$ is an *unshielded triple* if X_i and X_k are not adjacent. A nonendpoint vertex X_j on a path is a *collider* on the path if the path contains $*\!\!\rightarrow X_j \leftarrow\!\!*$. A nonendpoint vertex on a path which is not a collider is a *noncollider* on the path. A *collider path* is a path on which every nonendpoint vertex is a collider. A path of length one is a trivial collider path.

2.2. *Causal Bayesian networks.* A Bayesian network for a set of variables $\mathbf{V} = \{X_1, \dots, X_p\}$ is a pair $(\mathcal{D}, f)$, where $\mathcal{D} = (\mathbf{V}, \mathbf{E})$ is a DAG, and f is a joint probability density for $\mathbf{V}$ (with respect to some dominating measure) that factorizes according to $\mathcal{D}$: $f(\mathbf{V}) = \prod_{i=1}^{p} f(X_i | \mathrm{pa}(X_i, \mathcal{D}))$. If the DAG is interpreted causally, in the sense that $X_i \to X_j$ means that X_i has a (potential) direct causal effect on X_j, then we talk about a *causal DAG* and a *causal Bayesian network*.

One can easily derive post-intervention densities if the causal Bayesian network is given and all variables are observed. In particular, we consider interventions $\mathrm{do}(\mathbf{X} = \mathbf{x})$ for $\mathbf{X} \subseteq \mathbf{V}$ [Pearl (2000)], which represent outside interventions that set the variables in $\mathbf{X}$ to their respective values in $\mathbf{x}$. We assume that such interventions are effective, meaning that $\mathbf{X} = \mathbf{x}$ after the intervention. Moreover, we assume that the interventions are local, meaning that the generating mechanisms of the other variables, and hence their conditional distributions given their parents, do not change. We then have

$$f\bigl(\mathbf{V} | \mathrm{do}(\mathbf{X} = \mathbf{x})\bigr) = \begin{cases} \prod_{X_i \in \mathbf{V} \setminus \mathbf{X}} f\bigl(X_i | \mathrm{pa}(X_i, \mathcal{D})\bigr), & \text{for values of } \mathbf{V} \text{ consistent with } \mathbf{x}, \\ 0, & \text{otherwise.} \end{cases}$$

This is known as the g-formula or the truncated factorization formula [Pearl (2000), Robins (1986), Spirtes, Glymour and Scheines (1993)].

In a Bayesian network $(\mathcal{D}, f)$, the DAG $\mathcal{D}$ encodes conditional independence relationships in the density f via d-separation [Pearl (2000); see also Definition 3.5]. Several DAGs can encode the same conditional independence relationships. Such DAGs form a Markov equivalence class which can be uniquely represented by a CPDAG. A CPDAG is a graph with the same skeleton as each DAG in its equivalence class, and its edges are either directed ($\to$) or nondirected ($\circ\!\!-\!\!\circ$). An edge $X_i \to X_j$ in such a CPDAG means that $X_i \to X_j$ is present in every DAG in the Markov equivalence class, while an edge $X_i \circ\!\!-\!\!\circ X_j$ represents uncertainty about the edge marks, in the sense that the Markov equivalence class contains at least one DAG with $X_i \to X_j$ and at least one DAG with $X_i \leftarrow X_j$. (Note that many authors use $X_i - X_j$ instead of $X_i \circ\!\!-\!\!\circ X_j$; we use $\circ\!\!-\!\!\circ$ to ensure that the CPDAG satisfies the syntactic properties of a PAG; see below.)

If some of the variables in a DAG are unobserved, one can transform the DAG into a unique *maximal ancestral graph* (MAG) on the observed variables; see Richardson and Spirtes [(2002), page 981] for an algorithm. In particular, two vertices X_i and X_j are adjacent in a MAG if and only if no subset of the remaining observed variables makes them conditionally independent. Moreover, a tail mark $X_i \!-\!\!* X_j$ in a MAG $\mathcal{M}$ means that X_i is an ancestor of X_j in all DAGs represented by $\mathcal{M}$, while an arrowhead $X_i \leftarrow\!\!* X_j$ means that X_i is not an ancestor of X_j in all DAGs represented by $\mathcal{M}$. Thus an edge $X_i \to X_j$ in $\mathcal{M}$ means that there is a directed path from X_i to X_j in all DAGs represented by $\mathcal{M}$, but we emphasize that it does not represent a direct effect with respect to the observed variables, in the sense that there may be other observed variables on the directed path. Several different DAGs can lead to the same MAG, and a MAG represents a class of (infinitely many) DAGs that have the same d-separation and ancestral relationships among the observed variables. The MAG of a causal DAG is called a *causal MAG*.

A MAG encodes conditional independence relationships via the concept of m-separation (Definition 3.5). Again, several MAGs can encode the same conditional independence relationships. Such MAGs are called Markov equivalent, and can be uniquely represented by a *partial ancestral graph* (PAG). This is a partial mixed graph with the same skeleton as each MAG in its Markov equivalence class. A tail mark (arrowhead) at an edge $X_i \multimap X_j$ ($X_i \leftarrow\!\!* X_j$) in such a PAG means that $X_i \multimap X_j$ ($X_i \leftarrow\!\!* X_j$) in every MAG in the Markov equivalence class, while a circle mark at an edge $X_i \circ\!\!-\!\!* X_j$ represents uncertainty about the edge mark, in the sense that the Markov equivalence class contains at least one MAG with $X_i \multimap X_j$, and at least one MAG with $X_i \leftarrow\!\!* X_j$.

We say that a density f is *compatible with a DAG* $\mathcal{D}$ if the pair $(\mathcal{D}, f)$ forms a causal Bayesian network. A density f is *compatible with a CPDAG* $\mathcal{C}$ if it is compatible with a DAG in the Markov equivalence class described by $\mathcal{C}$. A density f is *compatible with a MAG* $\mathcal{M}$ if there exists a causal Bayesian network $(\mathcal{D}^*, f^*)$ (including hidden variables), such that $\mathcal{M}$ is the MAG of $\mathcal{D}^*$ and f is the corresponding marginal of f^*. Finally, f is *compatible with a PAG* $\mathcal{P}$ if it is compatible with a MAG in the Markov equivalence class described by $\mathcal{P}$.

3. Generalized back-door criterion.

We now present our generalized back-door criterion in Definition 3.7 and Theorem 3.1, where the name "generalized back-door criterion" is motivated by Lemma 3.1. We first introduce some more specialized definitions.

Zhang (2008a) introduced the concept of (*definitely*) *visible* edges in MAGs and PAGs. The reason for this is as follows. A directed edge $X \to Y$ in a DAG, CPDAG, MAG or PAG always means that X is a cause (or ancestor) of Y, because of the tail mark at X. However, if we allow for hidden variables (i.e., in MAGs and PAGs), there may be a hidden confounding variable between X and Y. Visible edges refer to situations where there cannot be such a hidden confounder between X and Y. Invisible edges, on the other hand, are possibly confounded in the sense that there is a DAG represented by the MAG or PAG with $X \leftarrow L \to Y$, where L is not measured (in addition to $X \to \cdots \to Y$).

DEFINITION 3.1 [Visible and invisible edges; cf. Zhang (2008a)]. All directed edges in DAGs and CPDAGs are said to be *visible*. Given a MAG $\mathcal{M}$/PAG $\mathcal{P}$, a directed edge $A \to B$ in $\mathcal{M}/\mathcal{P}$ 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 nonendpoint vertex on the path is a parent of B. Otherwise $A \to B$ is said to be invisible.

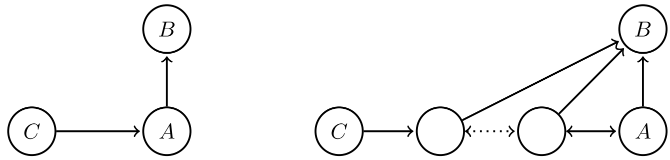

FIG. 1. *Edge configurations in MAGs and PAGs for a visible edge $A \to B$; cf. Zhang (2008a), Figure 6. Instead of the tail mark at C, one can also have an arrowhead or circle mark.*

Figure 1 illustrates the different graphical configurations that can lead to a visible edge. We note that Zhang (2008a) used slightly different terminology, referring to *definitely visible* edges in a PAG, while we simply say *visible* for both MAGs and PAGs. Borboudakis, Triantafillou and Tsamardinos (2012) used the term *pure-causal* edges instead of *visible* edges in MAGs.

We can now generalize the concept of a *back-door path* in Definition 3.2.

DEFINITION 3.2 (Back-door path). Let (X, Y) be an ordered pair of vertices in $\mathcal{G}$, where $\mathcal{G}$ is a DAG, CPDAG, MAG or PAG. We say that a path between X and Y is a *back-door path from X to Y* if it does not have a visible edge out of X.

In a DAG, this definition reduces to a path between X and Y that starts with $X \leftarrow$, which is the usual back-door path as defined by Pearl (1993). In a CPDAG, a back-door path from X to Y is a path between X and Y that starts with $X \leftarrow$ or $X \circ\!\!-\!\!\circ$. In a MAG, it is a path between X and Y that starts with $X \leftrightarrow$, $X \leftarrow$ or an invisible edge $X \rightarrow$. Finally, in a PAG, it is a path between X and Y that starts with $X \leftarrow\!\!*$, $X \circ\!\!-\!\!*$ or an invisible edge $X \rightarrow$.

We also need generalizations of the concept of d-separation in DAGs [Definition 1.2.3 of Pearl (2000)]. In MAGs, one can use m-separation [Section 3.4 of Richardson and Spirtes (2002)]. In CPDAGs and PAGs, there is the additional complication that it may be unclear whether a vertex is a collider or a noncollider on the path. We therefore need the following definitions:

DEFINITION 3.3 [Definite noncollider; Zhang (2008a)]. A nonendpoint vertex X_j on a path $\langle \dots, X_i, X_j, X_k, \dots \rangle$ in a partial mixed graph $\mathcal{G}$ is a *definite noncollider* on the path if (i) there is a tail mark at X_j, that is, $X_i *\!\!- X_j$ or $X_j -\!\!* X_k$, or (ii) $\langle X_i, X_j, X_k \rangle$ is unshielded and has circle marks at X_j, that is, $X_i *\!\!-\!\!\circ X_j \circ\!\!-\!\!* X_k$ and X_i and X_k are not adjacent in $\mathcal{G}$.

The motivation for conditions (i) and (ii) is straightforward. A tail mark out of X_j on the path ensures that X_j is a noncollider on the path in any graph obtained by orienting any possible circle marks. Condition (ii) comes from the fact that the collider status of unshielded triples is known in CPDAGs and PAGs. Hence, if the graph contains an unshielded triple that was not oriented as a collider, then it must be a noncollider in all underlying DAGs or MAGs. If $\mathcal{G}$ is a DAG or a MAG, then only condition (i) applies and reduces to the usual definition of a noncollider.

DEFINITION 3.4 (Definite status path). A nonendpoint vertex X_j on a path p in a partial mixed graph is said to be of a *definite status* if it is either a collider or a definite noncollider on p. The path p is said to be of a *definite status* if all nonendpoint vertices on the path are of a definite status.

A path of length one is a trivial definite status path. Moreover, in DAGs and MAGs, all paths are of a definite status.

We now define m-connection for definite status paths.

DEFINITION 3.5 (m-connection). A definite status path p between vertices X and Y in a partial mixed graph is *m-connecting* given a (possibly empty) set of variables $\mathbf{Z}$ ($X, Y \notin \mathbf{Z}$) if the following two conditions hold:

(a) every definite noncollider on the path is not in $\mathbf{Z}$;
(b) every collider on the path is an ancestor of some member of $\mathbf{Z}$.

If a definite status path p is not m-connecting given $\mathbf{Z}$, then we say that $\mathbf{Z}$ blocks p.

If $\mathbf{Z} = \varnothing$, we usually omit the phrase "given the empty set." Definition 3.5 reduces to m-connection for MAGs and d-connection for DAGs. We note that Zhang (2008a) used the notions of *possible m-connection* and *definite m-connection* in PAGs, where his notion of definite m-connection is the same as our notion of m-connection for definite status paths.

We now define an adjustment criterion for DAGs, CPDAGs, MAGs and PAGs. Throughout, we think of $\mathbf{X}$ and $\mathbf{Y}$ as nonempty sets.

DEFINITION 3.6 (Adjustment criterion). Let $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{W}$ be pairwise disjoint sets of vertices in $\mathcal{G}$, where $\mathcal{G}$ represents a DAG, CPDAG, MAG or PAG. Then we say that $\mathbf{W}$ satisfies the adjustment criterion relative to ($\mathbf{X}$, $\mathbf{Y}$) and $\mathcal{G}$ if for any density f compatible with $\mathcal{G}$, we have

$$f\big(\mathbf{y}|\,\mathrm{do}(\mathbf{x})\big) = \begin{cases} f(\mathbf{y}|\mathbf{x}), & \text{if } \mathbf{W} = \varnothing, \\ \displaystyle\int_{\mathbf{w}} f(\mathbf{y}|\mathbf{w},\mathbf{x}) f(\mathbf{w})\, d\mathbf{w} = E_{\mathbf{W}}\big\{f(\mathbf{y}|\mathbf{w},\mathbf{x})\big\}, & \text{otherwise.} \end{cases}$$

If $\mathbf{X} = \{X\}$ and $\mathbf{Y} = \{Y\}$, we simply say that a set satisfies the criterion relative to (X, Y) [rather than $(\{X\}, \{Y\})$] and the given graph.

We now propose our generalized back-door criterion for DAGs, CPDAGs, MAGs and PAGs. We will show in Theorem 3.1 that this criterion is sufficient for adjustment.

DEFINITION 3.7 (Generalized back-door criterion and generalized back-door set). Let $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{W}$ be pairwise disjoint sets of vertices in $\mathcal{G}$, where $\mathcal{G}$ represents a DAG, CPDAG, MAG or PAG. Then $\mathbf{W}$ satisfies the *generalized back-door criterion* relative to $(\mathbf{X}, \mathbf{Y})$ and $\mathcal{G}$ if the following two conditions hold:

(B-i) $\mathbf{W}$ does not contain possible descendants of $\mathbf{X}$ in $\mathcal{G}$;

(B-ii) for every $X \in \mathbf{X}$, the set $\mathbf{W} \cup \mathbf{X} \setminus \{X\}$ blocks every definite status back-door path from X to any member of $\mathbf{Y}$, if any, in $\mathcal{G}$.

A set $\mathbf{W}$ that satisfies the generalized back-door criterion relative to $(\mathbf{X}, \mathbf{Y})$ and $\mathcal{G}$ is called a *generalized back-door set* relative to $(\mathbf{X}, \mathbf{Y})$ and $\mathcal{G}$.

REMARK 3.1. Condition (B-i) in Definition 3.7 is equivalent to the following:

(B-i)′ $\mathbf{W}$ does not contain possible descendants of $\mathbf{X}$ *along a definite status path* in $\mathcal{G}$.

Condition (B-i)′ may be easier to check computationally than (B-i). The equivalence of (B-i) and (B-i)′ is shown in the proof of Theorem 3.1, using Lemma 7.2.

THEOREM 3.1. *Let* $\mathbf{X}$, $\mathbf{Y}$ *and* $\mathbf{W}$ *be pairwise disjoint sets of vertices in* $\mathcal{G}$, *where* $\mathcal{G}$ *represents a DAG, MAG, CPDAG or PAG. If* $\mathbf{W}$ *satisfies the generalized back-door criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *and* $\mathcal{G}$ (*Definition* 3.7), *then it satisfies the adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *and* $\mathcal{G}$ (*Definition* 3.6).

The proof of Theorem 3.1 consists of two steps. First, we formulate invariance criteria that are sufficient for adjustment (Theorem 7.1). Next, we translate the invariance criteria into the graphical criteria given in Definition 3.7, using results of Zhang (2008a) (Theorem 7.3).

We refer to Definition 3.7 as generalized back-door criterion because its conditions are closely related to Pearl's original back-door criterion [Pearl (1993, 2000)].

DEFINITION 3.8 [Pearl's back-door criterion; Definition 3.3.1 of Pearl (2000)]. A set of variables $\mathbf{W}$ satisfies the back-door criterion relative to an ordered pair of variables (X, Y) in a DAG $\mathcal{D}$ if the following two conditions hold:

(P-i) no vertex in $\mathbf{W}$ is a descendant of X in $\mathcal{D}$;

(P-ii) $\mathbf{W}$ blocks every path between X and Y in $\mathcal{D}$ that is into X.

Similarly, if $\mathbf{X}$ and $\mathbf{Y}$ are two disjoint subsets of vertices in $\mathcal{D}$, then $\mathbf{W}$ is said to satisfy the back-door criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ if it satisfies the criterion relative to any pair (X, Y) such that $X \in \mathbf{X}$ and $Y \in \mathbf{Y}$.

In particular, the conditions in Definition 3.7 are equivalent to Pearl's back-door criterion for a DAG with a single intervention ($|\mathbf{X}| = 1$). For a DAG with multiple interventions, any set that satisfies Pearl's back-door criterion also satisfies the generalized back-door criterion, but not necessarily the other way around. In this sense, our criterion is slightly better; see Lemma 3.1 and Example 1.

LEMMA 3.1. *Let* $\mathbf{X}$, $\mathbf{Y}$ *and* $\mathbf{W}$ *be pairwise disjoint sets of vertices in a DAG* $\mathcal{D}$. *If* $\mathbf{W}$ *satisfies Pearl's back-door criterion* (*Definition* 3.8) *relative to* $(\mathbf{X}, \mathbf{Y})$ *and* $\mathcal{D}$, *then* $\mathbf{W}$ *satisfies the generalized back-door criterion* (*Definition* 3.7) *relative to* $(\mathbf{X}, \mathbf{Y})$ *and* $\mathcal{D}$.

4. Finding a set that satisfies the generalized back-door criterion. An important reason for the popularity of Pearl's back-door criterion is the following. Consider two distinct vertices X and Y in a DAG $\mathcal{D}$. Then $\mathrm{pa}(X, \mathcal{D})$ satisfies the back-door criterion relative to (X, Y) and $\mathcal{D}$, unless $Y \in \mathrm{pa}(X, \mathcal{D})$. In the latter case, there is no set that satisfies the back-door criterion relative to (X, Y) and $\mathcal{D}$, but it is easy to see that $f(y|\,\mathrm{do}(x)) = f(y)$ for any density f compatible with $\mathcal{D}$, since there cannot be a directed path from X to Y in $\mathcal{D}$.

In this section, we formulate similar results for the generalized back-door criterion. In particular, we consider the following problem. Given two distinct vertices X and Y in a DAG, CPDAG, MAG or PAG, can we easily determine if there exists a generalized back-door set relative to (X, Y) and the given graph? Moreover, if this question is answered positively, can we give an explicit set that satisfies the criterion? Theorem 4.1 addresses these questions in general, while Corollaries 4.1–4.3 give specific results for DAGs, CPDAGs and MAGs.

We emphasize that throughout this section, we focus on the setting with a single intervention variable X and a single variable of interest Y. The setting with multiple interventions (i.e., a set $\mathbf{X}$) is considerably more difficult, even for DAGs [Shpitser, Van der Weele and Robins (2010a)]. It therefore seems challenging to generalize the results in this section to sets $\mathbf{X}$. Handling sets $\mathbf{Y}$ seems less difficult, and we plan to study this in future work.

In a DAG, the following result is well known. If X and Y are not adjacent in a DAG $\mathcal{D}$ and $X \notin \mathrm{an}(Y, \mathcal{D})$, then $\mathrm{pa}(X, \mathcal{D})$ blocks all paths between X and Y. In MAGs, we have a similar result, but we need to use $\text{D-SEP}(X, Y, \mathcal{M})$ instead of the parent set; see Definition 4.1 and Lemma 4.1.

DEFINITION 4.1 [$\text{D-SEP}(X, Y, \mathcal{G})$; cf. page 136 of Spirtes, Glymour and Scheines (2000)]. Let X and Y be two distinct vertices in a mixed graph $\mathcal{G}$. We say that $V \in \text{D-SEP}(X, Y, \mathcal{G})$ if $V \neq X$, and there is a collider path between X and V in $\mathcal{G}$, such that every vertex on this path (including V) is an ancestor of X or Y in $\mathcal{G}$.

LEMMA 4.1. *Let X and Y be two distinct vertices in an ancestral graph $\mathcal{G}$. Then the following statements are equivalent*: (i) X *and* Y *are m-separated in* $\mathcal{G}$ *by some subset of the remaining variables*, (ii) $Y \notin$ D-SEP$(X, Y, \mathcal{G})$, *and* (iii) X *and* Y *are m-separated in* $\mathcal{G}$ *by* D-SEP$(X, Y, \mathcal{G})$. *Moreover, if* $\mathcal{G}$ *is a MAG, a fourth equivalent statement is* (iv) X *and* Y *are not adjacent in* $\mathcal{G}$.

We now introduce important definitions that are needed to formulate our generalized back-door criterion in Theorem 4.1.

DEFINITION 4.2 ($\mathcal{R}^*$ and $\mathcal{R}_{\underline{X}}$). Let X be a vertex in $\mathcal{G}$, where $\mathcal{G}$ is a DAG, CPDAG, MAG or PAG.

Let $\mathcal{R}^* = \mathcal{R}^*(\mathcal{G}, X)$ be a class of DAGs or MAGs, defined as follows. If $\mathcal{G}$ is a DAG or a MAG, we simply let $\mathcal{R}^* = \{\mathcal{G}\}$. If $\mathcal{G}$ is a CPDAG/PAG, we let $\mathcal{R}^*$ be the subclass of DAGs/MAGs in the Markov equivalence class described by $\mathcal{G}$ that have the same number of edges into X as $\mathcal{G}$.

For any $\mathcal{R} \in \mathcal{R}^*$, let $\mathcal{R}_{\underline{X}} = \mathcal{R}_{\underline{X}}(\mathcal{R}, \mathcal{G}, X)$ be the graph obtained from $\mathcal{R}$ by removing all directed edges out of X that are visible in $\mathcal{G}$; see Definition 3.1.

For any given $\mathcal{G}$ and X, we say that a graph $\mathcal{R}_{\underline{X}}$ satisfies Definition 4.2 if there exists an $\mathcal{R} \in \mathcal{R}^*(\mathcal{G}, X)$ such that $\mathcal{R}_{\underline{X}} = \mathcal{R}_{\underline{X}}(\mathcal{R}, \mathcal{G}, X)$.

Lemma 7.6 shows that the class $\mathcal{R}^*$ is always nonempty. The definition of $\mathcal{R}_{\underline{X}}$ is related to the X-lower manipulated MAGs that were used by Zhang (2008a). It is important to note, however, that $\mathcal{R}_{\underline{X}}$ is obtained from $\mathcal{R}$ by removing the edges out of X that are visible in $\mathcal{G}$ (rather than $\mathcal{R}$). Moreover, Zhang replaced invisible edges by bi-directed edges, but that is not needed for our purposes (although it would not hurt to do so). Finally, we note that $\mathcal{R}_{\underline{X}}$ is ancestral, since any $\mathcal{R} \in \mathcal{R}^*$ is ancestral.

We can now present the main result of this section.

THEOREM 4.1 (Generalized back-door set). *Let X and Y be two distinct vertices in $\mathcal{G}$, where $\mathcal{G}$ is a DAG, CPDAG, MAG or PAG. Let $\mathcal{R}_{\underline{X}}$ be any graph satisfying Definition* 4.2. *Then there exists a generalized back-door set relative to* (X, Y) *and* $\mathcal{G}$ *if and only if* $Y \notin \text{adj}(X, \mathcal{R}_{\underline{X}})$ *and* D-SEP$(X, Y, \mathcal{R}_{\underline{X}}) \cap \text{possibleDe}(X, \mathcal{G}) = \emptyset$. *Moreover, if such a generalized back-door set exists, then* D-SEP$(X, Y, \mathcal{R}_{\underline{X}})$ *is such a set.*

The definitions of $\mathcal{R}^*$ and $\mathcal{R}_{\underline{X}}$ in Definition 4.2 are needed in Theorem 4.1 to ensure that D-SEP$(X, Y, \mathcal{R}_{\underline{X}}) \cap \text{possibleDe}(X, \mathcal{G}) \neq \emptyset$ implies that there does not exist a generalized back-door set relative to (X, Y) and $\mathcal{G}$; see also Example 8.

For DAGs, CPDAGs and MAGs we can simplify Theorem 4.1 somewhat; see Corollaries 4.1–4.3. Corollary 4.1 is the well-known result for DAGs that we discussed earlier. Corollary 4.3 is given without proof, since it follows straightforwardly from Theorem 4.1.

COROLLARY 4.1 (Generalized back-door set for a DAG). *Let X and Y be two distinct vertices in a DAG $\mathcal{D}$. Then there exists a generalized back-door set relative to (X, Y) and $\mathcal{D}$ if and only if $Y \notin \mathrm{pa}(X, \mathcal{D})$. Moreover, if such a generalized back-door set exists, then $\mathrm{pa}(X, \mathcal{D})$ is such a set.*

COROLLARY 4.2 (Generalized back-door set for a CPDAG). *Let X and Y be two distinct vertices in a CPDAG $\mathcal{C}$. Let $\mathcal{C}_{\underline{X}}$ be the graph obtained from $\mathcal{C}$ by removing all directed edges out of X. Then there exists a generalized back-door set relative to (X, Y) and $\mathcal{C}$ if and only if $Y \notin \mathrm{pa}(X, \mathcal{C})$ and $Y \notin \mathrm{possibleDe}(X, \mathcal{C}_{\underline{X}})$. Moreover, if such a generalized back-door set exists, then $\mathrm{pa}(X, \mathcal{C})$ is such a set.*

COROLLARY 4.3 (Generalized back-door set for a MAG). *Let X and Y be two distinct vertices in a MAG $\mathcal{M}$. Then there exists generalized backdoor set relative to (X, Y) and $\mathcal{M}$ if and only if $Y \notin \mathrm{adj}(X, \mathcal{M}_{\underline{X}})$ and $\text{D-SEP}(X, Y, M_{\underline{X}}) \cap \mathrm{de}(X, \mathcal{M}) = \varnothing$. Moreover, if such a generalized back-door set exists, then $\text{D-SEP}(X, Y, \mathcal{M}_{\underline{X}})$ is such a set.*

5. Examples. We now give several examples to illustrate the theory for DAGs, CPDAGs, MAGs and PAGs.

5.1. *DAG examples.* We start with an example that shows that the generalized back-door criterion for DAGs is weaker than Pearl's back-door criterion for DAGs, in the sense that it can happen that there is no set that satisfies Pearl's back-door criterion, while there is a set that satisfies the generalized back-door criterion.

EXAMPLE 1. Consider the DAG $\mathcal{D}$ in Figure 2(a) with $\mathbf{X} = \{X_1, X_3, X_4\}$ and $\mathbf{Y} = \{Y\}$. We first show that $\mathbf{W} = \varnothing$ is a generalized back-door set relative to $(\mathbf{X}, \mathbf{Y})$ and $\mathcal{D}$. Note that we cannot use Theorem 4.1 since $\mathbf{X}$ is a set. We therefore work with Definition 3.7 directly. We only need to check that the back-door path from X_4 to Y is blocked by $\mathbf{W} \cup \mathbf{X} \setminus \{X_4\} = \{X_1, X_3\}$, which is the case since X_3 is a noncollider on the path. Indeed, we have that $f(y|\,\mathrm{do}(x_1, x_3, x_4)) = f(y|x_1, x_3, x_4)$ in Figure 2(a), which can be further simplified to $f(y|x_3)$.

On the other hand, there is no set that satisfies Pearl's back-door criterion (Definition 3.8) with respect to $(\mathbf{X}, \mathbf{Y})$. To see this, note that $\{X_2, X_3, X_4\} \subseteq \mathrm{de}(X_1, \mathcal{D})$.

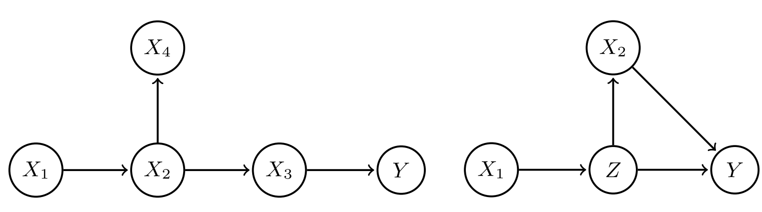

FIG. 2. *DAG examples.* (a) *The DAG $\mathcal{D}$ for Example 1.* (b) *The DAG $\mathcal{D}$ for Example 3.*

Hence, the only possible candidate set is $\mathbf{W} = \varnothing$. But this set does not block the back-door path from X_4 to Y, since there is no collider on this path.

Next, we note that the generalized back-door criterion is not necessary for identifying post-intervention distributions. Two simple examples are given below.

EXAMPLE 2. Let X and Y be two distinct vertices in $\mathcal{G}$, where $\mathcal{G}$ represents a DAG, CPDAG, MAG or PAG. If $X \leftrightarrow\!\!* Y$ in $\mathcal{G}$, then $Y \in \text{adj}(X, \mathcal{R}_{\underline{X}})$ for any $\mathcal{R}_{\underline{X}}$ satisfying Definition 4.2. Hence Theorem 4.1 implies that there does not exist a generalized back-door set relative to (X, Y) and $\mathcal{G}$.

On the other hand, it is clear that $f(y|\,\text{do}(x)) = f(y)$ for any density f compatible with $\mathcal{G}$, since the edge $X \leftrightarrow\!\!* Y$ implies that there cannot be a possibly directed path from X to Y in $\mathcal{G}$; see Lemma 7.5 below.

EXAMPLE 3. Let $\mathcal{D}$ be the DAG in Figure 2(b), and let $\mathbf{X} = \{X_1, X_2\}$ and $\mathbf{Y} = \{Y\}$. Then there does not exist a generalized back-door set relative to $(\mathbf{X}, \mathbf{Y})$ and $\mathcal{D}$. To see this, note that the only candidate variable Z cannot be used, since it is a descendant of X_1. Moreover, $\mathbf{W} = \varnothing$ violates condition (B-ii) in Definition 3.7 for X_2, since $\mathbf{W} \cup \mathbf{X} \setminus \{X_2\} = \{X_1\}$ does not block the back-door path $X_2 \leftarrow Z \rightarrow Y$.

On the other hand, $f(y|\,\text{do}(x_1, x_2)) = \int f(z|x_1) f(y|x_2, z)\, dz$ for any density f compatible with $\mathcal{D}$, by the g-formula.

5.2. *CPDAG examples*. We now illustrate the theory for CPDAGs. In Example 5, there is a set that satisfies the generalized back-door criterion, while in Example 4 there is none.

EXAMPLE 4. In the CPDAG $\mathcal{C}$ in Figure 3(a), $f(y|\,\text{do}(x))$ is not identifiable. To see this, note that the Markov equivalence class represented by this CPDAG contains three DAGs. Without loss of generality, we denote these by $\mathcal{D}_1, \mathcal{D}_2$

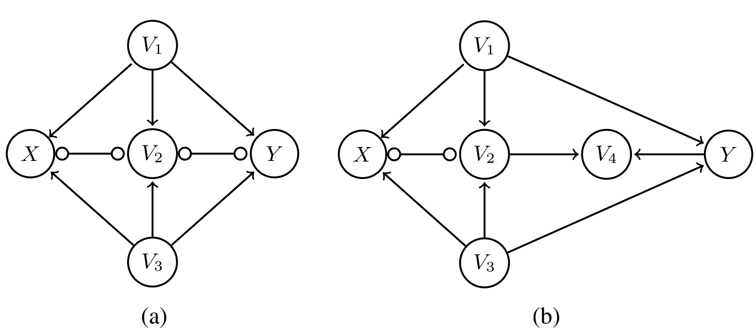

FIG. 3. *CPDAG examples*. (a) *The CPDAG* $\mathcal{C}$ *for Example* 4. (b) *The CPDAG* $\mathcal{C}'$ *for Example* 5.

and $\mathcal{D}_3$, where we assume that $\mathcal{D}_1$ contains the sub-graph $X \leftarrow V_2 \rightarrow Y$, $\mathcal{D}_2$ contains the sub-graph $X \leftarrow V_2 \leftarrow Y$, and $\mathcal{D}_3$ contains the sub-graph $X \rightarrow V_2 \rightarrow Y$. In $\mathcal{D}_1$ and $\mathcal{D}_2$ there is no directed path from X to Y, so that $f(y|\,\mathrm{do}(x)) = f(y)$ for any density f compatible with $\mathcal{D}_1$ or $\mathcal{D}_2$. In $\mathcal{D}_3$, however, there is a directed path from X to Y. Hence, one can easily construct a density f that is compatible with $\mathcal{D}_3$ such that $f(y|\,\mathrm{do}(x)) \neq f(y)$. This implies that $f(y|\,\mathrm{do}(x))$ is not identifiable. This implies that there cannot be a generalized back-door set relative to (X, Y) and $\mathcal{C}$.

We now apply Theorem 4.1 to the CPDAG $\mathcal{C}$ to check if this leads to the same conclusion. Note that $\mathcal{G} = \mathcal{C}$ and $\mathcal{R}^* = \{\mathcal{D}_3\}$. Hence, we take $\mathcal{R} = \mathcal{D}_3$ and the corresponding $\mathcal{R}_{\underline{X}} = \mathcal{D}_3$. We then have D-SEP$(X, Y, \mathcal{R}_{\underline{X}}) = \{V_1, V_2, V_3\}$ and possibleDe$(X, \mathcal{G}) = \{V_2, Y\}$. Hence, D-SEP$(X, Y, \mathcal{R}_{\underline{X}}) \cap$ possibleDe$(X, \mathcal{G}) = \{V_2\}$, and Theorem 4.1 correctly says that it is impossible to satisfy the generalized back-door criterion relative to (X, Y) and $\mathcal{C}$.

Finally, we check if Corollary 4.2 also yields the same result. Note that $\mathcal{C}_{\underline{X}} = \mathcal{C}$ and $Y \in$ possibleDe$(X, \mathcal{C}_{\underline{X}}) = \{V_2, Y\}$. Hence, we again find that it is impossible to satisfy the generalized back-door criterion relative to (X, Y) and $\mathcal{C}$.

EXAMPLE 5. In the CPDAG $\mathcal{C}'$ in Figure 3(b), $f(y|\,\mathrm{do}(x))$ is identifiable and equals $f(y)$, since there is no possibly directed path from X to Y in $\mathcal{C}'$.

We now check if we also arrive at this conclusion by applying Theorem 4.1. Note that there are two DAGs in the Markov equivalence class described by $\mathcal{C}'$, namely $\mathcal{D}'_1$ with the edge $X \rightarrow V_2$ and $\mathcal{D}'_2$ with the edge $X \leftarrow V_2$. Thus in Theorem 4.1, we have $\mathcal{G} = \mathcal{C}'$ and $\mathcal{R}^* = \{\mathcal{D}'_1\}$. Hence we take $\mathcal{R} = \mathcal{D}'_1$ and the corresponding $\mathcal{R}_{\underline{X}} = \mathcal{D}'_1$. Note that $Y \notin \mathrm{adj}(X, \mathcal{R}_{\underline{X}}) = \{V_1, V_2, V_3\}$ and D-SEP$(X, Y, \mathcal{R}_{\underline{X}}) = \{V_1, V_3\}$ and possibleDe$(X, \mathcal{G}) = \{V_2, V_4\}$. Hence, D-SEP$(X, Y, \mathcal{R}_{\underline{X}}) \cap$ possibleDe$(X, \mathcal{G}) = \varnothing$, and D-SEP$(X, Y, \mathcal{R}_{\underline{X}}) = \{V_1, V_3\}$ satisfies the generalized back-door criterion relative to (X, Y) and $\mathcal{C}'$. We can indeed check that the set $\{V_1, V_3\}$ satisfies the conditions in Definition 3.7.

Finally, we also apply Corollary 4.2. Note that $\mathcal{C}'_{\underline{X}} = \mathcal{C}'$. Moreover, $Y \notin$ pa$(X, \mathcal{C}')$ and $Y \notin$ possibleDe$(X, \mathcal{C}'_{\underline{X}})$. Hence, pa$(X, \mathcal{C}') = \{V_1, V_3\}$ satisfies the generalized back-door criterion relative to (X, Y) and $\mathcal{C}'$.

5.3. *MAG examples.* Next, we illustrate the theory for MAGs. In Examples 6 and 7, there does not exist a generalized back-door set relative to (X, Y) and the given MAGs. In Example 6, this is due to $Y \in \mathrm{adj}(X, \mathcal{M}_{\underline{X}})$, while in Example 7, it is due to D-SEP$(X, Y, \mathcal{M}_{\underline{X}}) \cap \mathrm{de}(X, \mathcal{M}) \neq \varnothing$.

EXAMPLE 6. Consider the MAG $\mathcal{M}$ consisting of the invisible edge $X \rightarrow Y$, and suppose we are interested in $f(y|\,\mathrm{do}(x))$. Then underlying DAG could be as in Figure 4(a), where L is unobserved. This is a well-known example where $f(y|\,\mathrm{do}(x))$ is not identifiable.

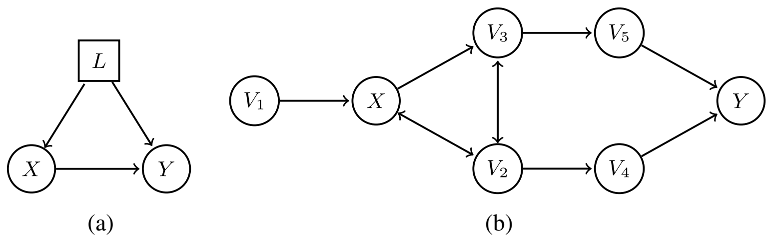

FIG. 4. *MAG examples*. (a) *A possible DAG described by the MAG in Example* 6, *where L is latent*. (b) *The MAG* $\mathcal{M}$ *for Example* 7.

We now apply Corollary 4.3 to check if we indeed find that it is impossible to satisfy the generalized back-door criterion relative to (X, Y) and $\mathcal{M}$. We have that $\mathcal{M} = \mathcal{M}_{\underline{X}}$ is the graph $X \to Y$. Hence, $Y \in \mathrm{adj}(X, \mathcal{M}_{\underline{X}})$, which leads to the correct conclusion.

EXAMPLE 7. Consider the MAG $\mathcal{M}$ in Figure 4(b) and apply Corollary 4.3 with $\mathbf{X} = \{X\}$ and $\mathbf{Y} = \{Y\}$. Since the edge $X \to V_3$ is visible, $\mathcal{M}_{\underline{X}}$ is constructed from $\mathcal{M}$ by removing this edge. We then have $\text{D-SEP}(X, Y, \mathcal{M}_{\underline{X}}) = \{V_1, V_2, V_3\}$ and $\mathrm{de}(X, \mathcal{M}) = \{V_3, V_5, Y\}$. Hence the intersection of $\mathrm{de}(X, \mathcal{M})$ and $\text{D-SEP}(X, Y, \mathcal{M}_{\underline{X}})$ is nonempty, and it follows that there is no generalized back-door set relative to (X, Y) and $\mathcal{M}$.

Indeed, we see that it is impossible to satisfy conditions (B-i) and (B-ii) in Definition 3.7. In order to block the back-door path $\langle X, V_2, V_4, Y\rangle$, we must include V_2 or V_4 in our set $\mathbf{W}$, but doing so opens the collider V_2 on the back-door path $\langle X, V_2, V_3, V_5, Y\rangle$. Hence, the latter path must be blocked by V_3 or V_5. But both these vertices are descendants of X in $\mathcal{M}$, and are therefore not allowed by condition (B-i).

5.4. *PAG example*. Finally, Example 8 is an example where there exists a generalized back-door set relative to some (X, Y) and a PAG. This example also illustrates that there may be subsets of $\text{D-SEP}(X, Y, \mathcal{R}_{\underline{X}})$ in Theorem 4.1 that satisfy the generalized back-door criterion. In other words, Theorem 4.1 may yield a non-minimal set. Hence, if one is interested in a minimal generalized back-door set, one could consider all subsets of $\text{D-SEP}(X, Y, \mathcal{R}_{\underline{X}})$. Example 8 also illustrates why $\mathcal{R}_{\underline{X}}$ is required to satisfy Definition 4.2.

EXAMPLE 8. Consider the PAG $\mathcal{P}$ in Figure 5(a), and suppose we are interested in $f(y|\,\mathrm{do}(x))$. Note that the MAG $\mathcal{R} = \mathcal{M}$ as given in Figure 5(b) is in $\mathcal{R}^*$; see Definition 4.2. We will apply Theorem 4.1 using the corresponding graph $\mathcal{R}_{\underline{X}}$, which is as $\mathcal{M}$ but without the edge $X \to Y$. We then have $Y \notin \mathrm{adj}(X, \mathcal{R}_{\underline{X}})$ and $\text{D-SEP}(X, Y, \mathcal{R}_{\underline{X}}) \cap \mathrm{possibleDe}(X, \mathcal{G}) = \{V_1, V_2\} \cap \{V_3, V_4, Y\} = \varnothing$. Hence Theorem 4.1 implies that $\{V_1, V_2\}$ is a generalized back-door set relative to (X, Y)

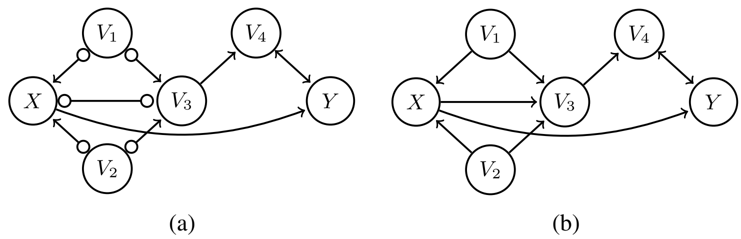

FIG. 5. *PAG example.* (a) *The PAG $\mathcal{P}$ for Example* 8. (b) *A possible MAG $\mathcal{M}$ for Example* 8.

and $\mathcal{P}$. One can easily verify that all subsets of $\{V_1, V_2\}$ are also generalized back-door sets relative to (X, Y) and $\mathcal{P}$, since all back-door paths from X to Y are blocked by the collider V_4 on these paths. This shows that D-SEP$(X, Y, \mathcal{R}_{\underline{X}})$ is not minimal.

This example also shows the importance of Definition 4.2. To see this, let $\mathcal{R}'$ be as $\mathcal{R}$, but with the edge $X \leftarrow V_3$ instead of $X \rightarrow V_3$, so that there is an additional edge into X. Then D-SEP$(X, Y, \mathcal{R}'_{\underline{X}}) = \{V_1, V_2, V_3\}$, and we get D-SEP$(X, Y, \mathcal{R}'_{\underline{X}}) \cap$ possibleDe$(X, \mathcal{G}) = \{V_3\} \neq \varnothing$. This shows that applying Theorem 4.1 with $\mathcal{R}'_{\underline{X}}$ instead of $\mathcal{R}_{\underline{X}}$ leads to incorrect results.

6. Discussion. In this paper, we generalize Pearl's back-door criterion [Pearl (1993)] to a generalized back-door criterion for DAGs, CPDAGs, MAGs and PAGs. We also provide easily checkable necessary and sufficient criteria for the existence of a generalized back-door set, when considering a single intervention variable and a single outcome variable. Moreover, if such a set exists, we provide an explicit set that satisfies the generalized back-door criterion. This set is not necessarily minimal, so if one is interested in a minimal set, one could consider all subsets.

Although effects that can be computed via the generalized back-door criterion are only a subset of all identifiable causal effects, we hope that the generalized back-door criterion will be useful in practice, and will make it easier to work with CPDAGs, MAGs and PAGs. Moreover, combining our results for CPDAGs and PAGs with fast causal structure learning algorithms such as the PC algorithm [Spirtes, Glymour and Scheines (2000)] or the FCI algorithm [Claassen, Mooij and Heskes (2013), Colombo et al. (2012), Spirtes, Glymour and Scheines (2000)] yields a computationally efficient way to obtain information on causal effects when assuming that the observational distribution is faithful to the true unknown causal DAG with or without hidden variables. To our knowledge, the prediction algorithm of Spirtes, Glymour and Scheines (2000) is the only alternative approach under the same assumptions, but the prediction algorithm is computationally much more complex.

The IDA algorithm [Maathuis et al. (2010), Maathuis, Kalisch and Bühlmann (2009)] has been designed to obtain bounds on causal effects when assuming that the observational distribution is faithful to the true underlying causal DAG without hidden variables. IDA roughly combines the PC algorithm with Pearl's back-door criterion. We could now apply a similar approach in the setting with hidden variables, by combining the FCI algorithm with the generalized back-door criterion for MAGs.

Possible directions for future work include studying the exact relationship between the prediction algorithm and our generalized back-door criterion, generalizing the results in Section 4 to allow for sets $\mathbf{X}$ and $\mathbf{Y}$ and extending the recent results of Van der Zander, Liśkiewicz and Textor (2014) to CPDAGs and PAGs.

7. Proofs.

7.1. *Proofs for Section 3.* In order to prove Theorem 3.1, we formulate so-called invariance conditions that will turn out to be sufficient for adjustment; see Definition 7.1 and Theorem 7.1 below. First, we briefly define what is meant by invariance. We refer to Zhang (2008a) for full details.

Let $\mathbf{Y}$, $\mathbf{Z}$ and $\mathbf{X}$ be three subsets of vertices in a causal DAG $\mathcal{D}$, where $\mathbf{X} \cap \mathbf{Y} = \mathbf{Y} \cap \mathbf{Z} = \varnothing$. Then a density $f(\mathbf{y}|\mathbf{z})$ is said to be entailed to be invariant under interventions on $\mathbf{X}$ given $\mathcal{D}$ if $f_{\mathbf{X}:=\mathbf{x}}(\mathbf{y}|\mathbf{z}) = f(\mathbf{y}|\mathbf{z})$ for all causal Bayesian networks $(\mathcal{D}, f)$, where the subscript $\mathbf{X} := \mathbf{x}$ denotes $\text{do}(\mathbf{X} = \mathbf{x})$. (This notation is used since $\mathbf{X}$ and $\mathbf{Z}$ are allowed to overlap.) The density $f(\mathbf{y}|\mathbf{z})$ is said to be entailed to be invariant under interventions on $\mathbf{X}$ given a CPDAG $\mathcal{C}$, a MAG $\mathcal{M}$ or a PAG $\mathcal{P}$ if it is entailed to be invariant under interventions on $\mathbf{X}$ given all DAGs represented by $\mathcal{C}$, $\mathcal{M}$ or $\mathcal{P}$, respectively.

DEFINITION 7.1 (Invariance criterion). Let $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{W}$ be pairwise disjoint sets of vertices in $\mathcal{G}$, where $\mathcal{G}$ is a DAG, CPDAG, MAG or PAG. Then $\mathbf{W}$ satisfies the *invariance criterion* relative to $(\mathbf{X}, \mathbf{Y})$ and $\mathcal{G}$ if the following two conditions hold for any density f compatible with $\mathcal{G}$:

(I-i) $f(\mathbf{w}|\,\text{do}(\mathbf{x})) = f(\mathbf{w})$;
(I-ii) $f(\mathbf{y}|\,\text{do}(\mathbf{x}), \mathbf{w}) = f(\mathbf{y}|\mathbf{x}, \mathbf{w})$.

In other words, conditions (I-i) and (I-ii) state that $f(\mathbf{w})$ and $f(\mathbf{y}|\mathbf{x}, \mathbf{w})$ are entailed to be invariant under interventions on $\mathbf{X}$ given $\mathcal{G}$. The conditions are also closely related to the conditions in equation (9) of Pearl (1993). We note that condition (I-i) is trivially satisfied if $\mathbf{W} = \varnothing$.

THEOREM 7.1. *Let* $\mathbf{X}$, $\mathbf{Y}$ *and* $\mathbf{W}$ *be pairwise disjoint sets of vertices in* $\mathcal{G}$, *where* $\mathcal{G}$ *is a DAG, CPDAG, MAG or PAG. If* $\mathbf{W}$ *satisfies the invariance criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *and* $\mathcal{G}$, *then it satisfies the adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *and* $\mathcal{G}$.

PROOF. If $\mathbf{W} = \varnothing$, condition (I-ii) immediately gives $f(\mathbf{y}|\,\mathrm{do}(\mathbf{x})) = f(\mathbf{y}|\mathbf{x})$. Otherwise, we have

$$(1)\quad f\big(\mathbf{y}|\,\mathrm{do}(\mathbf{x})\big) = \int_{\mathbf{w}} f\big(\mathbf{y}, \mathbf{w}|\,\mathrm{do}(\mathbf{x})\big)\,d\mathbf{w} = \int_{\mathbf{w}} f\big(\mathbf{y}|\mathbf{w}, \mathrm{do}(\mathbf{x})\big) f\big(\mathbf{w}|\,\mathrm{do}(\mathbf{x})\big)\,d\mathbf{w}.$$

Under conditions (I-i) and (I-ii), the right-hand side of (1) simplifies to $\int_{\mathbf{w}} f(\mathbf{y}|\mathbf{w}, \mathbf{x}) f(\mathbf{w})\,d\mathbf{w}$. □

Spirtes, Glymour and Scheines (1993, 2000), Zhang (2008a) formulated invariance results for DAGs, MAGs and PAGs. We derive a similar result for CPDAGs and then summarize the results for all these types of graphs in Theorem 7.2.

THEOREM 7.2 (Graphical criteria for invariance). *Let $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ be three subsets of observed vertices in $\mathcal{G}$, where $\mathcal{G}$ represents a DAG, CPDAG, MAG or PAG. Moreover, let $\mathbf{X} \cap \mathbf{Y} = \mathbf{Y} \cap \mathbf{Z} = \varnothing$. Then $f(\mathbf{y}|\mathbf{z})$ is entailed to be invariant under interventions on $\mathbf{X}$ given $\mathcal{G}$ if and only if*:

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

(2) *for every $X \in \mathbf{X} \cap (\mathrm{possibleAn}(\mathbf{Z}, \mathcal{G}) \setminus \mathbf{Z})$, there is no m-connecting definite status path between X and any member of $\mathbf{Y}$ given $\mathbf{Z}$*;

(3) *for every $X \in \mathbf{X} \setminus \mathrm{possibleAn}(\mathbf{Z}, \mathcal{G})$, every m-connecting definite status path, if any, between X and any member of $\mathbf{Y}$ given $\mathbf{Z}$ is into X.*

PROOF. One can easily check that the conditions reduce to the appropriate conditions for DAGs, MAGs and PAGs [Zhang (2008a), Proposition 18, Theorem 24 and Theorem 30]. The result for CPDAGs can be proved analogously. □

Note that $\mathbf{X} \cap \mathbf{Z}$, $\mathbf{X} \cap (\mathrm{possibleAn}(\mathbf{Z}, \mathcal{G}) \setminus \mathbf{Z})$ and $\mathbf{X} \setminus \mathrm{possibleAn}(\mathbf{Z}, \mathcal{G})$ form a partition of $\mathbf{X}$. Hence, only one of the conditions in Theorem 7.2 is relevant for a given $X \in \mathbf{X}$.

We also need the following basic property of PAGs and CPDAGs:

LEMMA 7.1 [Basic property of CPDAGs and PAGs; Lemma 1 of Meek (1995) for CPDAGs, and Lemma 3.3.1 of Zhang (2006) for PAGs]. *For any three vertices A, B and C in a CPDAG $\mathcal{C}$ or PAG $\mathcal{P}$, the following holds: if $A \ast\!\!\rightarrow B \circ\!\!-\!\!\ast C$, then there is an edge between A and C with an arrowhead at C, namely $A \ast\!\!\rightarrow C$. Furthermore, if the edge between A and B is $A \rightarrow B$, then the edge between A and C is either $A \circ\!\!\rightarrow C$ or $A \rightarrow C$ (i.e., not $A \leftrightarrow C$).*

We now show that the invariance conditions in Definition 7.1 are equivalent to the graphical conditions of Definition 3.7.

THEOREM 7.3. *The generalized back-door criterion* (*Definition* 3.7) *is equivalent to the invariance criterion* (*Definition* 7.1).

PROOF. We first show that condition (B-ii) of Definition 3.7 is equivalent to condition (I-ii) of Definition 7.1. We use Theorem 7.2 with $(\mathbf{X}', \mathbf{Y}', \mathbf{Z}')$, where $\mathbf{X}' = \mathbf{X}$, $\mathbf{Y}' = \mathbf{Y}$ and $\mathbf{Z}' = \mathbf{X} \cup \mathbf{W}$. Then $\mathbf{X}' \subseteq \mathbf{Z}'$, and clause (1) of the theorem yields that (I-ii) is equivalent to the following: for every $X \in \mathbf{X}$, every m-connecting definite status path, if any, between X and any member of $\mathbf{Y}$ given $(\mathbf{X} \cup \mathbf{W}) \setminus \{X\}$ is out of X with a visible edge. This is equivalent to condition (B-ii) by our definition of a back-door path; see Definition 3.2.

By Lemma 7.2 (below), condition (B-i) of Definition 3.7 is equivalent to condition (B-i)$'$ of Remark 3.1.

We now show that condition (B-i)$'$ is equivalent to condition (I-i) in Definition 7.1. We use Theorem 7.2 with $(\mathbf{X}', \mathbf{Y}', \mathbf{Z}')$, where $\mathbf{X}' = \mathbf{X}$, $\mathbf{Y}' = \mathbf{W}$ and $\mathbf{Z}' = \varnothing$. Then $\mathbf{Z}' = \text{possibleAn}(\mathbf{Z}', \mathcal{G}) = \varnothing$ and clause (3) of the theorem yields that (I-i) is equivalent to the following condition (I-i)$'$: for every $X \in \mathbf{X}$, every m-connecting definite status path, if any, between X and any member of $\mathbf{W}$ is into X. We now show that (I-i)$'$ is equivalent to (B-i)$'$.

First suppose that $\mathbf{W}$ violates (B-i)$'$. Then there are $W \in \mathbf{W}$ and $X \in \mathbf{X}$ such that there is a possibly directed definite status path p from X to W. Since p is possibly directed, it is not into X and it cannot contain colliders. Hence, it is an m-connecting definite status path between X and W that is not into X. This violates (I-i)$'$.

Now suppose that $\mathbf{W}$ violates (I-i)$'$. Then there are $W \in \mathbf{W}$ and $X \in \mathbf{X}$ such that there is an m-connecting definite status path between X and W that is not into X. Let $p = \langle X = U_1, \dots, U_k = W\rangle$ be such a path. Then every nonendpoint vertex on p must be a definite noncollider. Suppose that p is not a possibly directed path from X to W, meaning that there exists an $i \in \{2, \dots, k\}$ such that the edge between U_{i-1} and U_i is into U_{i-1}. If $i = 2$, this means that the path is into X, which is a contradiction. If $i > 2$, then the edge between U_{i-2} and U_{i-1} must be out of U_{i-1}, since U_{i-1} is a definite noncollider. But this means that the edge must be into U_{i-2}, since edges of the form $\circ\!\!-$ or $-$ are not allowed. Continuing this argument, we find that for all $j \in \{2, \dots, i\}$, the edge between U_{j-1} and U_j is into U_{j-1}. But this means that the path is into $U_1 = X$, which is a contradiction. Hence, p is a possibly directed path from X to W. Together with the fact that p is of a definite status, this violates (B-i)$'$. □

LEMMA 7.2. *Let X and Y be two distinct vertices in $\mathcal{G}$, where $\mathcal{G}$ is a DAG, CPDAG, MAG or PAG. If $Y \in$ possibleDe$(X, \mathcal{G})$, then there is a possibly directed definite status path $p = \langle X = U_1, \dots, U_k = Y\rangle$ from X to Y. Moreover, if $U_{i-1} \ast\!\!\rightarrow U_i$ for some $i \in \{2, \dots, k\}$, then $U_{j-1} \rightarrow U_j$ for all $j \in \{i + 1, \dots, k\}$.*

PROOF. If $\mathcal{G}$ is a DAG or a MAG, the lemma is trivially true. So let $\mathcal{G}$ be a CPDAG or a PAG, and assume that $Y \in \text{possibleDe}(X, \mathcal{G})$. This implies that there is a possibly directed path from X to Y in $\mathcal{G}$. Let $p = \langle X = U_1, \dots, U_k = Y \rangle$ be a shortest such path. If p is of length one, then the Lemma is trivially true. So assume that the length of p is at least two, that is, $k \geq 3$.

We first show that p is a definite status path. Note that p can contain the following edges $U_{i-1} \circ\!\!-\!\!\circ U_i$, $U_{i-1} \circ\!\!\rightarrow U_i$ and $U_{i-1} \rightarrow U_i$ $(i = 2, \dots, k)$. We now consider a sub-path $p(U_{i-1}, U_{i+1}) = \langle U_{i-1}, U_i, U_{i+1} \rangle$ of p, for some $i \in \{2, \dots, k-1\}$.

This sub-path cannot be of the form $U_{i-1} \circ\!\!\rightarrow U_i \circ\!\!-\!\!\circ U_{i+1}$ or $U_{i-1} \circ\!\!\rightarrow U_i \circ\!\!\rightarrow U_{i+1}$. To see this, suppose that the sub-path takes such a form. Then Lemma 7.1 implies the edge $U_{i-1} *\!\!\rightarrow U_{i+1}$. Suppose that this edge is into U_{i-1}; that is, it is $U_{i-1} \leftrightarrow U_{i+1}$. Then Lemma 7.1 applied to $U_{i+1} \leftrightarrow U_{i-1} \circ\!\!\rightarrow U_i$ implies the edge $U_{i+1} *\!\!\rightarrow U_i$, which is a contradiction. If the edge $U_{i-1} *\!\!\rightarrow U_{i+1}$ is not into U_{i-1}, then p is not a shortest possibly directed path.

Similarly, the sub-path cannot be of the form $U_{i-1} \rightarrow U_i \circ\!\!-\!\!\circ U_{i+1}$ or $U_{i-1} \rightarrow U_i \circ\!\!\rightarrow U_{i+1}$. To see this, suppose that the sub-path takes such a form. Then Lemma 7.1 implies the edge $U_{i-1} \circ\!\!\rightarrow U_{i+1}$ or $U_{i-1} \rightarrow U_{i+1}$. In either case, p is not a shortest possibly directed path.

Moreover, if the sub-path is of the form $U_{i-1} \circ\!\!-\!\!\circ U_i \circ\!\!-\!\!\circ U_{i+1}$, $U_{i-1} \circ\!\!-\!\!\circ U_i \circ\!\!\rightarrow U_{i+1}$ or $U_{i-1} \circ\!\!-\!\!\circ U_i \rightarrow U_{i+1}$, then it must be unshielded. To see this, suppose that the sub-path takes such a form and is not unshielded. If the edge between U_{i-1} and U_{i+1} is into U_{i-1}, then Lemma 7.1 applied to $U_{i+1} *\!\!\rightarrow U_{i-1} \circ\!\!-\!\!\circ U_i$ implies the edge $U_{i+1} *\!\!\rightarrow U_i$, which is a contradiction. If the edge between U_{i-1} and U_{i+1} is not into U_{i-1}, then p is not a shortest possibly directed path.

Hence, p can only contain triples of the form $U_{i-1} \circ\!\!\rightarrow U_i \rightarrow U_{i+1}$ or $U_{i-1} \rightarrow U_i \rightarrow U_{i+1}$, or of the form $U_{i-1} \circ\!\!-\!\!\circ U_i \circ\!\!-\!\!\circ U_{i+1}$, $U_{i-1} \circ\!\!-\!\!\circ U_i \circ\!\!\rightarrow U_{i+1}$ or $U_{i-1} \circ\!\!-\!\!\circ U_i \rightarrow U_{i+1}$ where U_{i-1} and U_{i+1} are not adjacent. In all these cases, the middle vertex U_i is a definite noncollider, so that p is a definite status path. Finally, if $U_{i-1} *\!\!\rightarrow U_i$ for some $i \in \{2, \dots, k\}$, it follows that $U_{j-1} \rightarrow U_j$ for all $j \in \{i+1, \dots, k\}$. $\square$

PROOF OF THEOREM 3.1. This follows directly from Theorems 7.1 and 7.3. $\square$

PROOF OF LEMMA 3.1. Conditions (P-i) and (B-i) are trivially equivalent for DAGs. We therefore only show that (P-ii) implies (B-ii), by contradiction. Thus, suppose that $\mathbf{W}$ blocks all back-door paths between $X \in \mathbf{X}$ and $Y \in \mathbf{Y}$ in $\mathcal{D}$, but there exist $X \in \mathbf{X}$ and $Y \in \mathbf{Y}$ such that there is a back-door path p from X to Y that is not blocked by $\mathbf{W} \cup \mathbf{X} \setminus \{X\}$. This means that: (i) no noncollider on p is in $\mathbf{W} \cup \mathbf{X} \setminus \{X\}$, (ii) all colliders on p have a descendant in $\mathbf{W} \cup \mathbf{X} \setminus \{X\}$, (iii) there is at least one collider on p that has a descendant in $\mathbf{X} \setminus \{X\}$ but not in $\mathbf{W}$. Among all colliders satisfying (iii), let Q be the one that is closest to Y on p, and let X' denote a descendant of Q in $\mathbf{X} \setminus \{X\}$. Then the directed path $q(Q, X')$ from Q to

X' is m-connecting given $\mathbf{W}$, since it is a path consisting of noncolliders and none of its vertices are in $\mathbf{W}$. Moreover, the sub-path $p(Q, Y)$ of p is m-connecting given $\mathbf{W}$ by construction. But this means that $q(X', Q) \oplus p(Q, Y)$ is a back-door path from X' to Y that is m-connecting given $\mathbf{W}$. This contradicts (P-ii). □

7.2. *Proofs for Section 4*. We first give several lemmas, starting with a result about m-connection in MAGs. This result basically says that replacing condition (b) in Definition 3.5 by "every collider on the path is an ancestor of some member of $\mathbf{Z} \cup \{X, Y\}$" does not change the m-separation relations in a MAG.

LEMMA 7.3 [Richardson (2003), Corollary 1]. *Let X and Y be two distinct vertices and $\mathbf{Z}$ be a subset of vertices in a mixed graph $\mathcal{M}$, with $\mathbf{Z} \cap \{X, Y\} = \varnothing$. If there is a path between X and Y in $\mathcal{M}$ on which no noncollider is in $\mathbf{Z}$ and every collider is in* $\mathrm{an}(\mathbf{Z} \cup \{X, Y\}, \mathcal{M})$, *then there is a path* (*not necessarily the same path*) *m-connecting X and Y given $\mathbf{Z}$ in $\mathcal{M}$.*

PROOF OF LEMMA 4.1. Let $\mathcal{G}$ be an ancestral graph. First, we note that (iii) trivially implies (i). Next, we show that (i) implies (ii), or equivalently, that not (ii) implies not (i). Thus, suppose that $Y \in$ D-SEP$(X, Y, \mathcal{G})$. Then there is a collider path between X and Y such that every vertex on the path is an ancestor of $\{X, Y\}$ in $\mathcal{G}$. This path is m-connecting given any subset of the remaining vertices, by Lemma 7.3.

Next, we show that (ii) implies (iii). Suppose that $Y \notin$ D-SEP$(X, Y, \mathcal{G})$. If there is no path between X and Y in $\mathcal{G}$, then X and Y are trivially m-separated by any subset of the remaining vertices. Thus, assume that there is at least one path between X and Y. Consider an arbitrary such path, and call it p. Since $Y \notin$ D-SEP$(X, Y, \mathcal{G})$, we have $Y \notin \mathrm{adj}(X, \mathcal{G})$. Hence the length of p must be at least two. We will show that p is blocked by D-SEP$(X, Y, \mathcal{G})$.

Suppose p starts with $X \leftarrow V$. Then $V \in$ D-SEP$(X, Y, \mathcal{G})$, since $V \in \mathrm{an}(X, \mathcal{G})$. Since V is a noncollider on p, this implies that p is blocked by D-SEP$(X, Y, \mathcal{G})$.

Suppose p is of the form $X {*\!\!\rightarrow} V \rightarrow \cdots \rightarrow Y$. Then $V \in \mathrm{an}(Y, \mathcal{G})$, so that $V \in$ D-SEP$(X, Y, \mathcal{G})$. Since V is a noncollider on p, this implies that p is blocked by D-SEP$(X, Y, \mathcal{G})$.

Suppose p starts with $X {*\!\!\rightarrow} V \rightarrow \cdots$ and the sub-path $p(V, Y)$ of p contains at least one collider. Let C be the collider closest to V on p. Then $V \in \mathrm{an}(C, \mathcal{G})$. If $C \notin \mathrm{an}($D-SEP$(X, Y, \mathcal{G}), \mathcal{G})$, then p is blocked by D-SEP$(X, Y, \mathcal{G})$. Hence, suppose $C \in \mathrm{an}($D-SEP$(X, Y, \mathcal{G}), \mathcal{G})$. Since any vertex in D-SEP$(X, Y, \mathcal{G})$ is an ancestor of $\{X, Y\}$ in $\mathcal{G}$, this implies $C \in \mathrm{an}(\{X, Y\}, \mathcal{G})$ and hence $V \in \mathrm{an}(\{X, Y\}, \mathcal{G})$ and $V \in$ D-SEP$(X, Y, \mathcal{G})$. Since V is a noncollider on p, p is blocked by D-SEP$(X, Y, \mathcal{G})$.

Suppose p is a collider path of the form $X {*\!\!\rightarrow} \leftrightarrow \cdots \leftarrow\!\!* Y$. Then at least one of the colliders is not in $\mathrm{an}(\{X, Y\}, \mathcal{G})$, since otherwise $Y \in$ D-SEP$(X, Y, \mathcal{G})$.

Let C be the collider closest to X on p that is not in $\text{an}(\{X, Y\}, \mathcal{G})$. Then $C \notin \text{an}(\text{D-SEP}(X, Y, \mathcal{G}), \mathcal{G})$. Hence, p is blocked by $\text{D-SEP}(X, Y, \mathcal{G})$.

Suppose p is of the form $X {*\!\!\rightarrow} \leftrightarrow \cdots \leftrightarrow V \leftarrow W \cdots Y$, with $W \neq Y$ ($W = Y$ was treated in the previous case) and the sub-path $p(X, V)$ is allowed to be of length one (i.e., $X {*\!\!\rightarrow} V$). If $W \in \text{D-SEP}(X, Y, \mathcal{G})$, then p is blocked by $\text{D-SEP}(X, Y, \mathcal{G})$. So suppose that $W \notin \text{D-SEP}(X, Y, \mathcal{G})$. Then there does not exist a collider path between X and W such that each vertex on the path is in $\text{an}(\{X, Y\}, \mathcal{G})$. This implies that there is a collider on the sub-path $p(X, W)$ of p that is not in $\text{an}(\{X, Y\}, \mathcal{G})$. Among such vertices, let Z be the one that is closest to X on $p(X, W)$. Then $Z \notin \text{an}(\text{D-SEP}(X, Y, \mathcal{G}), \mathcal{G})$. Hence, p is blocked by $\text{D-SEP}(X, Y, \mathcal{G})$.

Finally, if $\mathcal{G}$ is a MAG, two vertices are adjacent if and only if no subset of the remaining variables can m-separate them. Hence, (i) and (iv) are equivalent for MAGs. □

The following lemma says that we can check the existence of m-connecting definite status back-door paths in $\mathcal{G}$ by checking the existence of m-connecting paths in $\mathcal{R}_{\underline{X}}$, where $\mathcal{R}_{\underline{X}}$ is any graph satisfying Definition 4.2. This lemma is closely related to Lemma 5.1.7 of Zhang (2006) and Lemmas 26 and 27 of Zhang (2008a).

LEMMA 7.4. *Let X and Y be two distinct vertices and $\mathbf{Z}$ be a subset of vertices in $\mathcal{G}$, where $\mathcal{G}$ is a DAG, CPDAG, MAG or PAG. Let $\mathcal{R}_{\underline{X}}$ be any graph satisfying Definition 4.2. Then there is a definite status m-connecting back-door path from X to Y given $\mathbf{Z}$ in $\mathcal{G}$ if and only if there is an m-connecting path between X and Y given $\mathbf{Z}$ in $\mathcal{R}_{\underline{X}}$.*

PROOF. Let $\mathcal{R} \in \mathcal{R}^*$ and $\mathcal{R}_{\underline{X}}$ satisfy Definition 4.2. We first prove the "only if" statement. Suppose there is a definite status m-connecting back-door path p from X to Y given $\mathbf{Z}$ in $\mathcal{G}$. Let p' and p'' be the corresponding paths in $\mathcal{R}$ and $\mathcal{R}_{\underline{X}}$, consisting of the same sequence of vertices. (Note that p'' exists by the definition of $\mathcal{R}_{\underline{X}}$ and the fact that p is a back-door path in $\mathcal{G}$.) Then the path p' is m-connecting given $\mathbf{Z}$ in $\mathcal{R}$. The path p'', however, is not necessarily m-connecting in $\mathcal{R}_{\underline{X}}$, since it may happen that there is a collider Q on the path such that $Q \in \text{an}(\mathbf{Z}, \mathcal{R})$ but $Q \notin \text{an}(\mathbf{Z}, \mathcal{R}_{\underline{X}})$. But this can only occur if $Q \in \text{an}(X, \mathcal{R}_{\underline{X}})$. Hence, p'' satisfies the following properties: no noncollider on p'' is in $\mathbf{Z}$ and every collider on p'' is in $\text{an}(\mathbf{Z} \cup \{X\}, \mathcal{R}_{\underline{X}})$. It then follows from Lemma 7.3 that there is an m-connecting path between X and Y given $\mathbf{Z}$ in $\mathcal{R}_{\underline{X}}$.

We now prove the "if" statement. Suppose that there is an m-connecting path p'' between X and Y given $\mathbf{Z}$ in $\mathcal{R}_{\underline{X}}$. Let p' be the corresponding path in $\mathcal{R}$, consisting of the same sequence of vertices. Then p' is also m-connecting given $\mathbf{Z}$ in $\mathcal{R}$. Moreover, p does not start with a visible edge out of X in $\mathcal{G}$, because p'' exists in $\mathcal{R}_{\underline{X}}$. By Lemma 2′ in the proof of Lemma 5.1.7 of Zhang (2006), it then follows that there exists an m-connecting definite status back-door path between X and Y given $\mathbf{Z}$ in $\mathcal{G}$. □

The next lemma is used several times to derive a contradiction.

LEMMA 7.5. *Let U and V be two distinct vertices in $\mathcal{G}$, where $\mathcal{G}$ denotes a DAG, CPDAG, MAG or PAG. Then $\mathcal{G}$ cannot have both a possibly directed path from U to V and an edge of the form $V \ast\!\!\rightarrow U$.*

PROOF. This lemma is trivial for DAGs and MAGs, since they cannot contain (almost) directed cycles. So we only show the result for CPDAGs and PAGs. Let $\mathcal{G}$ denote the CPDAG or PAG, and suppose that $\mathcal{G}$ contains an edge of the form $V \ast\!\!\rightarrow U$ as well as a possibly directed path from U to V in $\mathcal{G}$. Then there is also a possibly directed definite status path $p = \langle U = U_1, \dots, U_k = V\rangle$ from U to V in $\mathcal{G}$, by Lemma 7.2. The path p has the following properties: if $U_{i-1} \ast\!\!\rightarrow U_i$ for some $i \in \{2, \dots, k\}$, then $U_{j-1} \rightarrow U_j$ for all $j \in \{i+1, \dots, k\}$, and the length of p must be at least two, because of the edge $V \ast\!\!\rightarrow U$.

If p is fully directed, there is an (almost) directed cycle in any DAG or MAG in the Markov equivalence class described by $\mathcal{G}$, which violates the ancestral property.

Otherwise, if p contains a directed sub-path, let $p(U_d, V)$ be the longest directed sub-path. Then the sub-path $p(U, U_d)$ must be of the form $U \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ U_d$ or $U \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ \circ\!\!\rightarrow U_d$. In either case, the edge $V \ast\!\!\rightarrow U$ and repeated applications of Lemma 7.1 imply the edge $V \ast\!\!\rightarrow U_d$. This gives an (almost) directed cycle together with the directed path $p(U_d, V)$ in any DAG or MAG in the Markov equivalence class described by $\mathcal{G}$. This again contradicts the ancestral property.

Otherwise, p does not contain a directed sub-path. Let T be the vertex preceding V on the path. Then the path has one of the following two forms: $U \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ T \circ\!\!-\!\!\circ V$ or $U \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ T \circ\!\!\rightarrow V$. The edge $V \ast\!\!\rightarrow U$ and repeated applications of Lemma 7.1 yield the edge $V \ast\!\!\rightarrow T$, which contradicts $T \circ\!\!-\!\!\circ V$ or $T \circ\!\!\rightarrow V$. □

Theorem 4.1 requires a DAG or MAG in $\mathcal{R}^*$; see Definition 4.2. The following lemma establishes such a DAG or MAG exists, since $\mathcal{R}^*$ is always nonempty. This result is closely related to constructions in Ali et al. (2005), Theorem 2 of Zhang (2008b) and Lemma 27 of Zhang (2008a).

LEMMA 7.6. *Let $\mathcal{G}$ be a PAG (CPDAG) with k edges into X, $k \in \{0, 1, \dots\}$. Then there exists at least one MAG (DAG) $\mathcal{R}$ in the Markov equivalence class represented by $\mathcal{G}$ that has k edges into X.*

PROOF. Building on the work of Meek (1995), Theorem 2 of Zhang (2008b) gives a procedure to create a MAG (DAG) in the Markov equivalence class represented by a PAG (CPDAG) $\mathcal{G}$. One first replaces all partially directed ($\circ\!\!\rightarrow$) edges in $\mathcal{G}$ by directed ($\rightarrow$) edges. Next, one considers the circle component $\mathcal{G}^C$ of $\mathcal{G}$, that is, the sub-graph of $\mathcal{G}$ consisting of nondirected ($\circ\!\!-\!\!\circ$) edges and orients this into a directed graph without directed cycles and unshielded colliders. The first step of this procedure only creates tail marks, and hence cannot yield an additional edge into X. For the second step, we will argue that we can construct such a graph that does not have any edges into X.

First, we note that $\mathcal{G}^C$ is chordal; that is, any cycle of length four or more has a chord, which is an edge joining two vertices that are not adjacent in the cycle; see the proof of Lemma 4.1 of Zhang (2008b). Any chordal graph with more than one vertex has two simplicial vertices, that is, vertices V such that all vertices adjacent to V are also adjacent to each other [e.g., Golumbic (1980)]. Hence, $\mathcal{G}^C$ must have at least one simplicial vertex that is different from X. We choose such a vertex V_1 and orient any edges incident to V_1 into V_1. Since V_1 is simplicial, this does not create unshielded colliders. We then remove V_1 and these edges from the graph. The resulting graph is again chordal [e.g., Golumbic (1980)] and therefore again has at least one simplicial vertex that is different from X. Choose such a vertex V_2, and orient any edges incident to V_2 into V_2. We continue this procedure until all edges are oriented. The resulting ordering is called a perfect elimination scheme for $\mathcal{G}^C$. By construction, this procedure yields an acyclic directed graph without unshielded colliders. Moreover, since X is chosen as the last vertex in the perfect elimination scheme, we do not orient any edges into X. $\square$

LEMMA 7.7. *Let X and Y be two distinct vertices in $\mathcal{G}$, where $\mathcal{G}$ is a DAG, CPDAG, MAG or PAG. Let $\mathcal{R}_{\underline{X}}$ be any graph satisfying Definition 4.2. If $V \in$ D-SEP$(X, Y, \mathcal{R}_{\underline{X}}) \cap \text{possibleDe}(X, \mathcal{G})$, then $V \in \text{an}(Y, \mathcal{R}_{\underline{X}})$.*

PROOF. Let $\mathcal{R}_{\underline{X}}$ satisfy Definition 4.2, and let $V \in$ D-SEP$(X, Y, \mathcal{R}_{\underline{X}}) \cap$ possibleDe$(X, \mathcal{G})$. This means that there is a collider path p_1 between X and V in $\mathcal{R}_{\underline{X}}$ such that every vertex on the path is an ancestor of X or Y in $\mathcal{R}_{\underline{X}}$. In particular, $V \in \text{an}(\{X, Y\}, \mathcal{R}_{\underline{X}})$.

We first show that $V \in \text{pa}(X, \mathcal{R}_{\underline{X}})$ leads to a contradiction. Thus, suppose there is an edge $X \leftarrow V$ in $\mathcal{R}_{\underline{X}}$. By construction of $\mathcal{R}_{\underline{X}}$, $\mathcal{G}$ then contains an edge of the form $X \leftarrow\!\!\circ V$ or $X \leftarrow V$, but this forms a contradiction together with $V \in$ possibleDe$(X, \mathcal{G})$, by Lemma 7.5.

We now show that $V \in \text{an}(X, \mathcal{R}_{\underline{X}}) \setminus \text{pa}(X, \mathcal{R}_{\underline{X}})$ leads to a contradiction. Thus suppose there is a directed path from V to X in $\mathcal{R}_{\underline{X}}$ of the form $\langle V, \dots, W, X\rangle$, where $V \neq W$ and $W \neq X$. By construction of $\mathcal{R}_{\underline{X}}$, the edge $W \rightarrow X$ must also be into X in $\mathcal{G}$, so that $\mathcal{G}$ contains $W \circ\!\!\rightarrow X$ or $W \rightarrow X$. Since $V \in$ possibleDe$(X, \mathcal{G})$, there is a possibly directed path p_{xv} from X to V in $\mathcal{G}$. Since $\mathcal{R}_{\underline{X}}$ contains a directed path from V to W, $\mathcal{G}$ must also contain a possibly directed path p_{vw} from V to W. This implies that $p_{xv} \oplus p_{vw}$ is a possibly directed path from X to W in $\mathcal{G}$, so that $W \in$ possibleDe$(X, \mathcal{G})$. But this forms a contradiction with $W \circ\!\!\rightarrow X$ or $W \rightarrow X$ in $\mathcal{G}$, by Lemma 7.5.

Hence, we must have $V \in \mathrm{an}(Y, \mathcal{R}_{\underline{X}})$. $\square$

We can now prove the main result in Section 4.

PROOF OF THEOREM 4.1. Let $\mathcal{R}_{\underline{X}}$ satisfy Definition 4.2. We first show that $Y \in \mathrm{adj}(X, \mathcal{R}_{\underline{X}})$ or $\text{D-SEP}(X, Y, \mathcal{R}_{\underline{X}}) \cap \mathrm{possibleDe}(X, \mathcal{G}) \neq \varnothing$ implies that there does not exist a generalized back-door set relative to (X, Y) and $\mathcal{G}$, since no set $\mathbf{W}$ can satisfy conditions (B-i) and (B-ii) in Definition 3.7. Thus suppose that $Y \in \mathrm{adj}(X, \mathcal{R}_{\underline{X}})$. Then there is a definite status back-door path of length one in $\mathcal{G}$ that cannot be blocked. Hence condition (B-ii) cannot be satisfied by any set $\mathbf{W}$. Next, suppose that there exists some vertex $V \in \text{D-SEP}(X, Y, \mathcal{R}_{\underline{X}}) \cap \mathrm{possibleDe}(X, \mathcal{G}) \neq \varnothing$. Then there is a collider path p_1 between X and V in $\mathcal{R}_{\underline{X}}$ such that every vertex on the path is in $\mathrm{an}(\{X, Y\}, \mathcal{R}_{\underline{X}})$. Moreover, by Lemma 7.7, there is a directed path p_2 from V to Y in $\mathcal{R}_{\underline{X}}$. Now consider $p = p_1 \oplus p_2$. All nonendpoint vertices on p that are not on p_2 are colliders on p and in $\mathrm{an}(\{X, Y\}, \mathcal{R}_{\underline{X}})$. The remaining nonendpoint vertices on p are noncolliders and in $\mathrm{possibleDe}(X, \mathcal{G})$ [since $V \in \mathrm{possibleDe}(X, \mathcal{G})$], so that including them in $\mathbf{W}$ violates condition (B-i). It then follows by Lemma 7.3 that for any subset $\mathbf{W}$ satisfying condition (B-i), there exists an m-connecting path between X and Y given $\mathbf{W}$ in $\mathcal{R}_{\underline{X}}$. By Lemma 7.4, this means that we cannot block all definite status back-door paths from X to Y in $\mathcal{G}$ without violating condition (B-i).

We now prove the other direction. Thus suppose that $Y \notin \mathrm{adj}(X, \mathcal{R}_{\underline{X}})$ and $\text{D-SEP}(X, Y, \mathcal{R}_{\underline{X}}) \cap \mathrm{possibleDe}(X, \mathcal{G}) = \varnothing$. Then we need to show that $\text{D-SEP}(X, Y, \mathcal{R}_{\underline{X}})$ satisfies conditions (B-i) and (B-ii) of Definition 3.7. Condition (B-i) is satisfied trivially, since $\text{D-SEP}(X, Y, \mathcal{R}_{\underline{X}}) \cap \mathrm{possibleDe}(X, \mathcal{G}) = \varnothing$. To prove that condition (B-ii) is satisfied as well, we first show $Y \notin \text{D-SEP}(X, Y, \mathcal{R}_{\underline{X}})$, by contradiction. Thus, suppose $Y \in \text{D-SEP}(X, Y, \mathcal{R}_{\underline{X}}) \subseteq \text{D-SEP}(X, Y, \mathcal{R})$. By Lemma 4.1, this implies $Y \in \mathrm{adj}(X, \mathcal{R})$. Since $Y \notin \mathrm{adj}(X, \mathcal{R}_{\underline{X}})$, this implies that $X \to Y$ in $\mathcal{G}$ with a visible edge. But this means that $Y \in \mathrm{possibleDe}(X, \mathcal{G})$, so that $\text{D-SEP}(X, Y, \mathcal{R}_{\underline{X}}) \cap \mathrm{possibleDe}(X, \mathcal{G}) \neq \varnothing$. This is a contradiction, which implies $Y \notin \text{D-SEP}(X, Y, \mathcal{R}_{\underline{X}})$. Hence $\text{D-SEP}(X, Y, \mathcal{R}_{\underline{X}})$ m-separates X and Y in $\mathcal{R}_{\underline{X}}$ by Lemma 4.1 (we use here that $\mathcal{R}_{\underline{X}}$ is ancestral). By Lemma 7.4, this implies that $\text{D-SEP}(X, Y, \mathcal{R}_{\underline{X}})$ blocks all definite status back-door paths from X to Y in $\mathcal{G}$, so that condition (B-ii) is satisfied. $\square$

PROOF OF COROLLARY 4.1. Although this result for DAGs is well known, we show how one can derive this from Theorem 4.1. Note that $\mathcal{D}_{\underline{X}}$ is the graph obtained by removing all directed edges out of X from $\mathcal{D}$. Moreover, $\text{D-SEP}(X, Y, \mathcal{D}_{\underline{X}}) = \mathrm{pa}(X, \mathcal{D})$ and $\mathrm{possibleDe}(X, \mathcal{D}) = \mathrm{de}(X, \mathcal{D})$. Now the condition $Y \notin \mathrm{adj}(X, \mathcal{D}_{\underline{X}})$ is equivalent to $Y \notin \mathrm{pa}(X, \mathcal{D})$. The other condition $\text{D-SEP}(X, Y, \mathcal{D}_{\underline{X}}) \cap \mathrm{possibleDe}(X, \mathcal{D}) = \varnothing$ reduces to $\mathrm{pa}(X, \mathcal{D}) \cap \mathrm{de}(X, \mathcal{D}) = \varnothing$, and this is fulfilled automatically by the acyclicity of $\mathcal{D}$. Hence Theorem 4.1 reduces to the given statement. $\square$

PROOF OF COROLLARY 4.2. Let $\mathcal{D}$ be a DAG in the Markov equivalence class represented by $\mathcal{C}$, constructed without orienting additional edges into X. Let $\mathcal{D}_{\underline{X}}$ be obtained from $\mathcal{D}$ by removing all directed edges out of X that were directed out of X in $\mathcal{C}$. Let $\mathcal{C}_{\underline{X}}$ be obtained from $\mathcal{C}$ by removing all directed edges out of X.

We first show that $Y \in \text{pa}(X, \mathcal{C})$ or $Y \in \text{possibleDe}(X, \mathcal{C}_{\underline{X}})$ imply $Y \in \text{adj}(X, \mathcal{D}_{\underline{X}})$ or $\text{D-SEP}(X, Y, \mathcal{D}_{\underline{X}}) \cap \text{possibleDe}(X, \mathcal{C}) \neq \varnothing$. Thus suppose $Y \in \text{pa}(X, \mathcal{C})$. Then $Y \in \text{adj}(X, \mathcal{D}_{\underline{X}})$. Next, suppose $Y \in \text{possibleDe}(X, \mathcal{C}_{\underline{X}})$. It can be easily shown that $\mathcal{C}_{\underline{X}}$ satisfies the basic property of Lemma 7.1, that $A \to B \circ\!\!-\!\!\circ C$ implies $A \to C$ (since all edges that are removed are directed edges out of X). Hence, Lemma 7.2 applies to $\mathcal{C}_{\underline{X}}$, and it follows that there is a possibly directed definite status path from X to Y in $\mathcal{C}_{\underline{X}}$. All nonendpoint vertices on this path must be definite noncolliders. By construction of $\mathcal{C}_{\underline{X}}$, the first edge on this path must be nondirected in $\mathcal{C}_{\underline{X}}$, and by construction of $\mathcal{D}_{\underline{X}}$, this edge must be oriented out of X in $\mathcal{D}_{\underline{X}}$. This implies that the entire path must be directed from X to Y in $\mathcal{D}_{\underline{X}}$, since all nonendpoint vertices are noncolliders. Let V be the vertex adjacent to X on the path. Then $V \in \text{D-SEP}(X, Y, \mathcal{D}_{\underline{X}})$. Moreover, $V \in \text{possibleDe}(X, \mathcal{C})$. Hence $\text{D-SEP}(X, Y, \mathcal{D}_{\underline{X}}) \cap \text{possibleDe}(X, \mathcal{C}) \neq \varnothing$.

We now show that $\text{D-SEP}(X, Y, \mathcal{D}_{\underline{X}}) \cap \text{possibleDe}(X, \mathcal{C}) \neq \varnothing$ or $Y \in \text{adj}(X, \mathcal{D}_{\underline{X}})$ imply $Y \in \text{pa}(X, \mathcal{C})$ or $Y \in \text{possibleDe}(X, \mathcal{C}_{\underline{X}})$. Thus suppose that $Y \in \text{pa}(X, \mathcal{C})$ imply $Y \in \text{pa}(X, \mathcal{C})$ or $Y \in \text{possibleDe}(X, \mathcal{C}_{\underline{X}})$. Thus suppose $Y \in \text{adj}(X, \mathcal{D}_{\underline{X}})$. Then either $X \leftarrow Y$ or $X \circ\!\!-\!\!\circ Y$ in $\mathcal{C}$. This implies that $Y \in \text{pa}(X, \mathcal{C})$ or $Y \in \text{possibleDe}(X, \mathcal{C}_{\underline{X}})$. Next, suppose that there exists a vertex $V \in \text{D-SEP}(X, Y, \mathcal{D}_{\underline{X}}) \cap \text{possibleDe}(X, \mathcal{C})$. Note that $V \in \text{D-SEP}(X, Y, \mathcal{D}_{\underline{X}})$ implies: (i) $V \in \text{pa}(X, \mathcal{D}_{\underline{X}})$ or (ii) $V \in \text{ch}(X, \mathcal{D}_{\underline{X}}) \cap \text{an}(Y, \mathcal{D}_{\underline{X}})$ or (iii) $V \in \text{pa}(\text{ch}(X, \mathcal{D}_{\underline{X}}) \cap \text{an}(Y, \mathcal{D}_{\underline{X}}))$. By construction of $\mathcal{D}_{\underline{X}}$, case (i) implies $V \in \text{pa}(X, \mathcal{C})$. But this is in contradiction with $V \in \text{possibleDe}(X, \mathcal{C})$, by Lemma 7.5. In case (ii), we have $X \to V$ and a directed path from V to Y in $\mathcal{D}_{\underline{X}}$, so that $Y \in \text{de}(X, \mathcal{D}_{\underline{X}})$. Similarly, we can obtain $Y \in \text{de}(X, \mathcal{D}_{\underline{X}})$ in case (iii). This implies $Y \in \text{possibleDe}(X, \mathcal{C}_{\underline{X}})$ in cases (ii) and (iii).

The above shows the following: if $Y \in \text{pa}(X, \mathcal{C})$ or $Y \in \text{possibleDe}(X, \mathcal{C}_{\underline{X}})$, then it is impossible to satisfy the generalized back-door criterion relative to (X, Y) and $\mathcal{C}$. On the other hand, if $Y \notin \text{pa}(X, \mathcal{C})$ and $Y \notin \text{possibleDe}(X, \mathcal{C}_{\underline{X}})$, then $\text{D-SEP}(X, Y, \mathcal{D}_{\underline{X}})$ satisfies the generalized back-door criterion relative to (X, Y) and $\mathcal{C}$. It is left to show that in the latter case, we can replace $\text{D-SEP}(X, Y, \mathcal{D}_{\underline{X}})$ by $\text{pa}(X, \mathcal{C})$. Since $\text{pa}(X, \mathcal{C}) \subseteq \text{D-SEP}(X, Y, \mathcal{D}_{\underline{X}})$, it is clear that $\text{pa}(X, \mathcal{C})$ satisfies condition (B-i) of Definition 3.7. We will now show that it also satisfies condition (B-ii).

Thus, suppose that $Y \notin \text{pa}(X, \mathcal{C})$ and $Y \notin \text{possibleDe}(X, \mathcal{C}_{\underline{X}})$. Consider a definite status back-door path $p = \langle X = U_1, \dots, U_k = Y \rangle$ from X to Y in $\mathcal{C}$. Since p is a back-door path, it must start with $X \leftarrow U_2$ or $X \circ\!\!-\!\!\circ U_2$. Moreover, the length of p is at least two. If $X \leftarrow U_2$, then it is clear that $\text{pa}(X, \mathcal{C})$ blocks p. If $X \circ\!\!-\!\!\circ U_2$, then p cannot have a sub-path of the form $U_{i-1} \circ\!\!-\!\!\circ U_i \leftarrow U_{i+1}$, $i \in \{2, \dots, k-1\}$, because U_i is of a definite status. Moreover, p cannot be possibly directed, because then $Y \in \text{possibleDe}(X, \mathcal{C}_{\underline{X}})$. Hence, there must be at least one collider on p. Let Q be the collider on p that is closest to X. Then the sub-path $p(X, Q)$ is a possibly directed path from X to Q in $\mathcal{C}$. Suppose that Q is an ancestor of some vertex $W \in \text{pa}(X, \mathcal{C})$ in $\mathcal{C}$. Then there is a possibly directed path from X to W in $\mathcal{C}$, as well as an edge $W \to X$. But this is impossible by Lemma 7.5. Hence, Q cannot be an ancestor of any member of $\text{pa}(X, \mathcal{C})$ in $\mathcal{C}$. This implies that p is blocked by $\text{pa}(X, \mathcal{C})$. □

Acknowledgements. We are very grateful to Markus Kalisch, Thomas Richardson and two anonymous referees for their comments and suggestions that have significantly improved the paper.

REFERENCES

ALI, A. R., RICHARDSON, T. S., SPIRTES, P. L. and ZHANG, J. (2005). Towards characterizing Markov equivalence classes for directed acyclic graphs with latent variables. In *Proceedings of the 21st Annual Conference on Uncertainty in Artificial Intelligence* (*UAI* 2005) 10–17. AUAI Press, Arlington, VA.

ALI, R. A., RICHARDSON, T. S. and SPIRTES, P. (2009). Markov equivalence for ancestral graphs. *Ann. Statist.* **37** 2808–2837. MR2541448

ANDERSSON, S. A., MADIGAN, D. and PERLMAN, M. D. (1997). A characterization of Markov equivalence classes for acyclic digraphs. *Ann. Statist.* **25** 505–541. MR1439312

BORBOUDAKIS, G., TRIANTAFILLOU, S. and TSAMARDINOS, I. (2012). Tools and algorithms for causally interpreting directed edges in maximal ancestral graphs. In *Proceedings of the 6th European Workshop on Probabilistic Graphical Models* (*PGM* 2012) 35–42. DECSAI, Univ. Granada.

CHICKERING, D. M. (2002). Learning equivalence classes of Bayesian-network structures. *J. Mach. Learn. Res.* **2** 445–498. MR1929415

CLAASSEN, T., MOOIJ, J. and HESKES, T. (2013). Learning sparse causal models is not NP-hard. In *Proceedings of the 29th Annual Conference on Uncertainty in Artificial Intelligence* (*UAI* 2013) 172–181. AUAI Press, Corvallis, OR.

COLOMBO, D., MAATHUIS, M. H., KALISCH, M. and RICHARDSON, T. S. (2012). Learning high-dimensional directed acyclic graphs with latent and selection variables. *Ann. Statist.* **40** 294–321. MR3014308

DE LUNA, X., WAERNBAUM, I. and RICHARDSON, T. S. (2011). Covariate selection for the non-parametric estimation of an average treatment effect. *Biometrika* **98** 861–875. MR2860329

ENTNER, D., HOYER, P. O. and SPIRTES, P. (2013). Data-driven covariate selection for nonparametric estimation of causal effects. *J. Mach. Learn. Res. Workshop Conf. Proc.* **31** 256–264.

GLYMOUR, C., SCHEINES, R., SPIRTES, P. and KELLY, K. (1987). *Discovering Causal Structure: Artificial Intelligence, Philosophy of Science, and Statistical Modeling*. Academic Press, San Diego, CA.

GOLUMBIC, M. C. (1980). *Algorithmic Graph Theory and Perfect Graphs*. Academic Press, New York.

HUANG, Y. and VALTORTA, M. (2006). Identifiability in causal Bayesian networks: A sound and complete algorithm. In *Proceedings of the 21st National Conference on Artificial Intelligence* (*AAAI* 2006) 1149–1154. AAAI Press, Menlo Park, CA.

KALISCH, M. and BÜHLMANN, P. (2007). Estimating high-dimensional directed acyclic graphs with the PC-algorithm. *J. Mach. Learn. Res.* **8** 613–636.

KALISCH, M., MÄCHLER, M., COLOMBO, D., MAATHUIS, M. H. and BÜHLMANN, P. (2012). Causal inference using graphical models with the R package pcalg. *J. Stat. Softw.* **47** 1–26.

KOLLER, D. and FRIEDMAN, N. (2009). *Probabilistic Graphical Models*: *Principles and Techniques*. MIT Press, Cambridge, MA. MR2778120

MAATHUIS, M. H., COLOMBO, D., KALISCH, M. and BÜHLMANN, P. (2010). Predicting causal effects in large-scale systems from observational data. *Nat. Methods* **7** 247–248.

MAATHUIS, M. H., KALISCH, M. and BÜHLMANN, P. (2009). Estimating high-dimensional intervention effects from observational data. *Ann. Statist.* **37** 3133–3164. MR2549555

MEEK, C. (1995). Causal inference and causal explanation with background knowledge. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence* (*UAI* 1995) 403–410. Morgan Kaufmann, San Francisco, CA.

PEARL, J. (1993). Comment: Graphical models, causality and intervention. *Statist. Sci.* **8** 266–269.

PEARL, J. (1995). Causal diagrams for empirical research. *Biometrika* **82** 669–710. MR1380809

PEARL, J. (2000). *Causality*: *Models*, *Reasoning*, *and Inference*. Cambridge Univ. Press, Cambridge. MR1744773

PEARL, J. (2009). *Causality*: *Models*, *Reasoning*, *and Inference*, 2nd ed. Cambridge Univ. Press, Cambridge. MR2548166

RICHARDSON, T. (2003). Markov properties for acyclic directed mixed graphs. *Scand. J. Stat.* **30** 145–157. MR1963898

RICHARDSON, T. and SPIRTES, P. (2002). Ancestral graph Markov models. *Ann. Statist.* **30** 962–1030. MR1926166

RICHARDSON, T. S. and SPIRTES, P. (2003). Causal inference via ancestral graph models. In *Highly Structured Stochastic Systems*. *Oxford Statist. Sci. Ser.* **27** 83–113. Oxford Univ. Press, Oxford. MR2082407

ROBINS, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math. Modelling* **7** 1393–1512. MR0877758

SHPITSER, I. and PEARL, J. (2006a). Identification of conditional interventional distributions. In *Proceedings of the* 22*nd Annual Conference on Uncertainty in Artificial Intelligence* (*UAI* 2006) 437–444. AUAI Press, Corvallis, OR.

SHPITSER, I. and PEARL, J. (2006b). Identification of joint interventional distributions in recursive semi-Markovian causal models. In *Proceedings of the* 21*st National Conference on Artificial Intelligence* (*AAAI* 2006) 1219–1226. AAAI Press, Menlo Park, CA.

SHPITSER, I. and PEARL, J. (2008). Complete identification methods for the causal hierarchy. *J. Mach. Learn. Res.* **9** 1941–1979. MR2447308

SHPITSER, I., VAN DER WEELE, T. and ROBINS, J. (2010a). On the validity of covariate adjustment for estimating causal effects. In *Proceedings of the* 26*th Conference on Uncertainty and Artificial Intelligence* (*UAI* 2010) 527–536. AUAI Press, Corvallis, OR.

SHPITSER, I., VAN DER WEELE, T. and ROBINS, J. (2010b). Appendum to *On the validity of covariate adjustment for estimating causal effects*. Personal communication.

SPIRTES, P., GLYMOUR, C. and SCHEINES, R. (1993). *Causation*, *Prediction*, *and Search*. *Lecture Notes in Statistics* **81**. Springer, New York. MR1227558

SPIRTES, P., GLYMOUR, C. and SCHEINES, R. (2000). *Causation*, *Prediction*, *and Search*, 2nd ed. MIT Press, Cambridge, MA. MR1815675

TEXTOR, J. and LIŚKIEWICZ, M. (2011). Adjustment criteria in causal diagrams: An algorithmic perspective. In *Proceedings of the* 27*th Conference on Uncertainty in Artificial Intelligence* (*UAI* 2011) 681–688. AUAI Press, Corvallis, OR.

TIAN, J. and PEARL, J. (2002). A general identification condition for causal effects. In *Proceedings of the* 18*th National Conference on Artificial Intelligence* (*AAAI* 2002) 567–573. AAAI Press, Menlo Park, CA.

VANDERWEELE, T. J. and SHPITSER, I. (2011). A new criterion for confounder selection. *Biometrics* **67** 1406–1413. MR2872391

VAN DER ZANDER, B., LIŚKIEWICZ, M. and TEXTOR, J. (2014). Constructing separators and adjustment sets in ancestral graphs. In *Proceedings of the* 30*th Annual Conference on Uncertainty in Artificial Intelligence* (*UAI* 2014) 907–916. AUAI Press, Corvallis, OR.

ZHANG, J. (2006). Causal inference and reasoning in causally insufficient systems. Ph.D. thesis, Carnegie Mellon Univ., Pittsburgh, PA.

ZHANG, J. (2008a). Causal reasoning with ancestral graphs. *J. Mach. Learn. Res.* **9** 1437–1474. MR2426048

ZHANG, J. (2008b). On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias. *Artificial Intelligence* **172** 1873–1896. MR2459793

SEMINAR FOR STATISTICS
ETH ZURICH
RÄMISTRASSE 101
8092 ZURICH
SWITZERLAND
E-MAIL: maathuis@stat.math.ethz.ch
colombo@stat.math.ethz.ch