Journal of Machine Learning Research 18 (2018) 1-62 Submitted 6/16 Revised 4/18; Published 5/18

Complete Graphical Characterization and Construction of Adjustment Sets in Markov Equivalence Classes of Ancestral Graphs

Emilija Perković PERKOVIC@STAT.MATH.ETHZ.CH
Seminar for Statistics, ETH Zurich, Switzerland

Johannes Textor JOHANNES.TEXTOR@RADBOUDUMC.NL
Institute for Computing and Information Sciences and Department of Tumor Immunology, Radboud University Medical Center, Nijmegen, The Netherlands

Markus Kalisch KALISCH@STAT.MATH.ETHZ.CH
Marloes H. Maathuis MAATHUIS@STAT.MATH.ETHZ.CH
Seminar for Statistics, ETH Zurich, Switzerland

Editor: Christopher Meek

Abstract

We present a graphical criterion for covariate adjustment that is sound and complete for four different classes of causal graphical models: directed acyclic graphs (DAGs), maximal ancestral graphs (MAGs), completed partially directed acyclic graphs (CPDAGs), and partial ancestral graphs (PAGs). Our criterion unifies covariate adjustment for a large set of graph classes. Moreover, we define an explicit set that satisfies our criterion, if there is any set that satisfies our criterion. We also give efficient algorithms for constructing all sets that fulfill our criterion, implemented in the `R` package `dagitty`. Finally, we discuss the relationship between our criterion and other criteria for adjustment, and we provide new soundness and completeness proofs for the adjustment criterion for DAGs.

Keywords: causal effects, graphical models, covariate adjustment, latent variables, confounding

1. Introduction

Covariate adjustment is a well-known method to estimate causal effects from observational data. There are, however, still common misconceptions about what variables one should or should not adjust for. For example, it is sometimes thought that adjusting for more variables will lead to a more precise estimate, as long as the added variables are not affected by the exposure variable. While this is true for a randomized exposure, in observational data even adjustment for pre-exposure variables may lead to so-called collider bias as described in the "M-bias graph" (Shrier, 2008; Rubin, 2008). Another example is the "Table 2 fallacy" (Westreich and Greenland, 2013). In observational research papers, Table 1 often describes the data, and Table 2 shows a multiple regression analysis. By presenting all estimated coefficients in one table, it is implicitly suggested that all estimates can be interpreted similarly. This is usually not the case: some coefficients may be interpreted as a total causal effect, some may be interpreted as a direct causal effect, and some do not have any causal interpretation at all.

	DAG	MAG	CPDAG	PAG
Back-door criterion Pearl (1993)	$\Rightarrow$			
Adjustment criterion Shpitser et al. (2010), Shpitser (2012)	$\Leftrightarrow$			
Adjustment criterion van der Zander et al. (2014)	$\Leftrightarrow$	$\Leftrightarrow$		
Generalized back-door criterion Maathuis and Colombo (2015)	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$
Generalized adjustment criterion Perković et al. (2015)	$\Leftrightarrow$	$\Leftrightarrow$	$\Leftrightarrow$	$\Leftrightarrow$

Table 1: Graphical criteria for covariate adjustment: $\Rightarrow$ - sound, $\Leftrightarrow$ - sound and complete.

The practical importance of covariate adjustment has inspired a growing body of theoretical work on graphical criteria for adjustment. Pearl's back-door criterion (Pearl, 1993) is probably the most well-known, and is sound but not complete for adjustment in DAGs. Shpitser et al. (2010) and Shpitser (2012) refined the back-door criterion to a sound and complete graphical criterion for adjustment in DAGs. Others considered more general graph classes, which can represent structural uncertainty. Van der Zander et al. (2014, 2018) gave sound and complete graphical criteria for MAGs that allow for unobserved variables (latent confounding). Maathuis and Colombo (2015) presented a generalized back-door criterion for DAGs, CPDAGs, MAGs and PAGs, where CPDAGs and PAGs represent Markov equivalence classes of DAGs or MAGs, respectively, and can be inferred directly from data (see, for example, Spirtes et al., 2000; Chickering, 2002; Colombo et al., 2012; Claassen et al., 2013; Colombo and Maathuis, 2014; Nandy et al., 2018; Frot et al., 2018; Heinze-Deml et al., 2017). The generalized back-door criterion is sound but not complete for adjustment. Another line of work explores data driven covariate adjustment that does not require knowing the graph (VanderWeele and Shpitser, 2011; De Luna et al., 2011; Entner et al., 2013). Some of these data driven results are sound and complete for adjustment, but they all rely on some additional assumptions. We will not explore this direction in our paper.

In Perković et al. (2015), the preliminary conference version of the present paper, we extended the results of Shpitser et al. (2010); Shpitser (2012), van der Zander et al. (2014) and Maathuis and Colombo (2015) to derive a single sound and complete adjustment criterion for DAGs, CPDAGs, MAGs and PAGs. The different adjustment criteria are summarized in Table 1. Additionally, we note that van der Zander and Liśkiewicz (2016) showed that the generalized adjustment criterion can also be applied to the more general class of restricted chain graphs (representing a subset of a Markov equivalence class of DAGs). Furthermore, in Perković et al. (2017), we extend the generalized adjustment criterion to maximally oriented partially directed acyclic graphs (PDAGs), which represent CPDAGs with added background knowledge.

To illustrate the use of our generalized adjustment criterion, suppose we are given the CPDAG in Figure 1a and we want to estimate the total causal effect of X on Y. Our

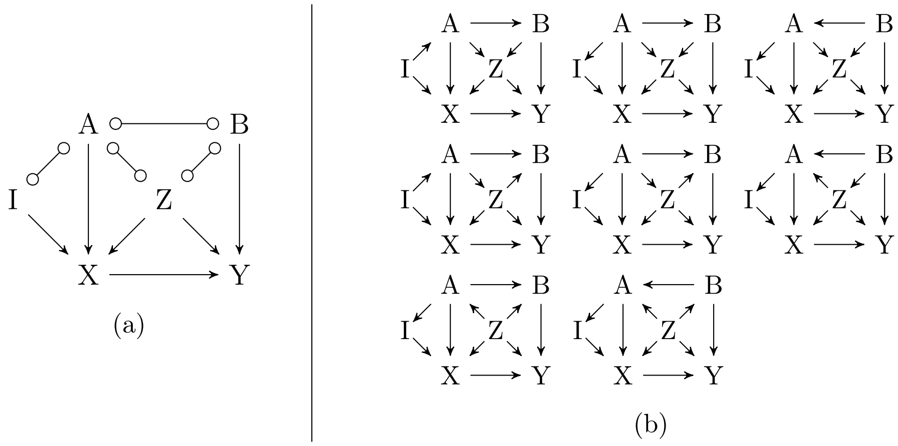

Figure 1: (a) A CPDAG in which, according to our criterion, $\{A, Z\}$ is an adjustment set for the total causal effect of X on Y. (b) The Markov equivalence class of DAGs represented by the CPDAG. An adjustment set for a CPDAG (PAG) is one that is valid for all DAGs (MAGs) in the Markov equivalence class.

criterion will inform us that the set $\{A, Z\}$ is an adjustment set for this CPDAG, meaning that it is an adjustment set in every DAG that the CPDAG represents (Figure 1b). Hence, we can estimate the causal effect without knowledge of the full causal structure. In a similar manner, by applying our criterion to a MAG or a PAG, we find adjustment sets that are valid for all DAGs represented by this MAG or PAG. Our criterion finds such adjustment sets whenever they exist; else, the causal effect is not identifiable by covariate adjustment. We hope that this ability to allow for incomplete structural knowledge, latent confounding, or both will help address concerns that graphical causal modelling "assumes that all [...] DAGs have been properly specified" (West and Koch, 2014). Moreover, our criterion for CPDAGs and PAGs can be combined with causal structure learning algorithms.

In the current paper, we give full proofs of the results in Perković et al. (2015). In addition, we provide several new results that allow us to construct sets $\mathbf{Z}$ that fulfill the generalized adjustment criterion for given sets $\mathbf{X}$ of exposures and $\mathbf{Y}$ of response variables in a DAG, CPDAG, MAG or PAG $\mathcal{G}$. In Corollary 15 we define a specific set that satisfies our criterion, if any set does. We refer to this set as a "constructive set". In Theorem 7, we show how one can express adjustment sets in terms of m-separating sets in a certain subgraph of $\mathcal{G}$. This theorem reduces the problem of finding adjustment sets to the problem of finding m-separating sets, which has been studied in detail by van der Zander et al. (2014). In Lemma 10, we prove that all adjustment sets for a CPDAG (PAG) $\mathcal{G}$ can be found in an arbitrary orientation of $\mathcal{G}$ into a valid DAG (MAG). This allows us to leverage existing implementations (van der Zander et al., 2014). We implemented the criterion itself and the construction of all adjustment sets in the software `dagitty` (Textor et al., 2016), available as a web-based GUI and an R package, and in the R package `pcalg` (Kalisch et al., 2012).

Furthermore, we explore the relationships between our generalized adjustment criterion and the previously suggested generalized back-door criterion and Pearl's back-door criterion. For both Pearl's back-door criterion and the generalized back-door criterion, a constructive set was given only in the case when the number of exposures is limited to one ($|\mathbf{X}| = 1$). We give constructive sets for each of these criteria for general $\mathbf{X}$ in Corollary 22 and Corollary 24. Moreover, in Theorem 26 we identify cases in which there exist sets satisfying two, or all three of these criteria, as well as cases in which there are only sets satisfying the generalized adjustment criterion.

Another important contribution, included in the appendix, are new soundness and completeness proofs of the adjustment criterion for DAGs as defined in Shpitser et al. (2010) and in the unpublished addendum Shpitser (2012), where the adjustment criterion in Shpitser (2012) is a revised version of the criterion in Shpitser et al. (2010) (see Definition 55 in Appendix E). Since there are no published soundness and completeness proofs for the revised criterion and since we build on this work, we felt it was important to provide these proofs. The proofs are non-trivial, but rely only on elementary concepts.

We note that, although we can find all causal effects that are identifiable by covariate adjustment, we generally do not find all identifiable causal effects, since some effects may be identifiable only by other means, using for example IDA approaches (Maathuis et al., 2009, 2010; Nandy et al., 2017; Malinsky and Spirtes, 2017), Pearl's front-door criterion (Pearl, 2009, Section 3.3.2) or the ID algorithm (Tian and Pearl, 2002; Shpitser and Pearl, 2006).

We also point out that MAGs and PAGs are in principle not only able to represent unobserved confounding, but can also account for unobserved selection variables. In this paper, however, we assume that there are no unobserved selection variables, since selection bias often rules out causal effect identification using just covariate adjustment. Bareinboim et al. (2014) discuss these problems and present creative approaches to work around them, for example by combining data from different sources. The question whether our adjustment criterion could be combined with such auxiliary methods is left for future research.

2. Preliminaries

Throughout the paper we denote sets in bold (for example $\mathbf{X}$), graphs in calligraphic font (for example $\mathcal{G}$) and nodes in a graph in uppercase letters (for example X). All omitted proofs are given in the appendix.

Nodes and edges. A graph $\mathcal{G} = (\mathbf{V}, \mathbf{E})$ consists of a set of nodes (variables) $\mathbf{V} = \{X_1, \dots, X_p\}$ and a set of edges $\mathbf{E}$. We consider simple graphs, meaning that there is at most one edge between any pair of nodes. Two nodes are called *adjacent* if they are connected by an edge. Every edge has two edge marks that can be arrowheads, tails or circles. Edges can be *directed* $\rightarrow$, *bi-directed* $\leftrightarrow$, *non-directed* $\circ\!\!-\!\!\circ$, or *partially directed* $\circ\!\!\rightarrow$. We use $\bullet$ as a stand in for any of the allowed edge marks. An edge is *into* (*out of*) a node X if the edge has an arrowhead (tail) at X. A *directed graph* contains only directed edges. A *mixed graph* may contain directed and bi-directed edges. A *partial mixed graph* may contain any of the described edges. Unless stated otherwise, definitions apply to partial mixed graphs.

Paths. A *path* p from X to Y in $\mathcal{G}$ is a sequence of distinct nodes $\langle X, \dots, Y \rangle$ in which every pair of successive nodes is adjacent in $\mathcal{G}$. If $p = \langle X_1, X_2, \dots, X_k, \rangle, k \geq 2$, then with $-p$ we denote the path $\langle X_k, \dots, X_2, X_1 \rangle$. A node V *lies on a path* p if V occurs in the sequence of nodes. If $p = \langle X_1, X_2, \dots, X_k, \rangle, k \geq 2$, then X_1 and X_k are *endpoints* of p, and any other node $X_i, 1 < i < k$, is a *non-endpoint* node on p. The *length* of a path equals the number of edges on the path. A *directed path* from X to Y is a path from X to Y in which all edges are directed towards Y, that is, $X \to \cdots \to Y$. We also refer to this as a *causal path*. A *possibly directed path* or *possibly causal path* from X to Y is a path from X to Y that does not contain an arrowhead pointing in the direction of X. A path from X to Y that is not possibly causal is called a *non-causal path* from X to Y. A directed path from X to Y together with $Y \to X$ forms a *directed cycle*. A directed path from X to Y together with $Y \leftrightarrow X$ forms an *almost directed cycle*. For two disjoint subsets $\mathbf{X}$ and $\mathbf{Y}$ of $\mathbf{V}$, a path from $\mathbf{X}$ to $\mathbf{Y}$ is a path from some $X \in \mathbf{X}$ to some $Y \in \mathbf{Y}$. A path from $\mathbf{X}$ to $\mathbf{Y}$ is *proper* (wrt $\mathbf{X}$) if only its first node is in $\mathbf{X}$. If $\mathcal{G}$ and $\mathcal{G}^*$ are two graphs with identical adjacencies and p is a path in $\mathcal{G}$, then the *corresponding path* p^* is the path in $\mathcal{G}^*$ constituted by the same sequence of nodes as p.

Subsequences, subpaths and concatenation. A *subsequence* of a path p is a sequence of nodes obtained by deleting some nodes from p without changing the order of the remaining nodes. A subsequence of a path is not necessarily a path. For a path $p = \langle X_1, X_2, \dots, X_m \rangle$, the *subpath* from X_i to X_k $(1 \leq i \leq k \leq m)$ is the path $p(X_i, X_k) = \langle X_i, X_{i+1}, \dots, X_k \rangle$. We denote the concatenation of paths by $\oplus$, so that for example $p = p(X_1, X_k) \oplus p(X_k, X_m)$. In this paper, we only concatenate paths if the result of the concatenation is again a path.

Ancestral relationships. If $X \to Y$, then X is a *parent* of Y. If there is a directed (possibly directed) path from X to Y, then X is a *ancestor* (*possible ancestor*) of Y, and Y is a *descendant* (*possible descendant*) of X. We also use the convention that every node is a descendant, possible descendant, ancestor and possible ancestor of itself. The sets of parents, descendants and ancestors of X in $\mathcal{G}$ are denoted by $\mathrm{Pa}(X, \mathcal{G})$, $\mathrm{De}(X, \mathcal{G})$ and $\mathrm{An}(X, \mathcal{G})$ respectively. The sets of possible descendants and possible ancestors of X in $\mathcal{G}$ are denoted by $\mathrm{PossDe}(X, \mathcal{G})$ and $\mathrm{PossAn}(X, \mathcal{G})$ respectively. For a set of nodes $\mathbf{X} \subseteq \mathbf{V}$, we let $\mathrm{Pa}(\mathbf{X}, \mathcal{G}) = \cup_{X \in \mathbf{X}} \mathrm{Pa}(X, \mathcal{G})$, with analogous definitions for $\mathrm{De}(\mathbf{X}, \mathcal{G})$, $\mathrm{An}(\mathbf{X}, \mathcal{G})$, $\mathrm{PossDe}(\mathbf{X}, \mathcal{G})$ and $\mathrm{PossAn}(\mathbf{X}, \mathcal{G})$.

Colliders, shields and definite status paths. If a path p contains $X_i \bullet\!\!\rightarrow X_j \leftarrow\!\!\bullet X_k$ as a subpath, then X_j is a *collider* on p. A *collider path* is a path on which every non-endpoint node is a collider. A path of length one is a trivial collider path. A path $\langle X_i, X_j, X_k \rangle$ is an *(un)shielded triple* if X_i and X_k are (not) adjacent. A path is *unshielded* if all successive triples on the path are unshielded. A node X_j is a *definite non-collider* (Zhang, 2008a) on a path p if there is at least one edge out of X_j on p, or if $X_i \bullet\!\!-\!\!\circ X_j \circ\!\!-\!\!\bullet X_k$ is a subpath of p and $\langle X_i, X_j, X_k \rangle$ is an unshielded triple. Any collider on a path is always of definite status and hence, a *definite collider*. In a DAG (MAG) we refer to definite non-colliders as *non-colliders*. A node is of *definite status* on a path if it is a collider or a definite non-collider on the path. A path p is of definite status if every non-endpoint node on p is of definite status.

m-separation and m-connection. A definite status path p between nodes X and Y is *m-connecting* given a set of nodes $\mathbf{Z}$ $(X, Y \notin \mathbf{Z})$ if every definite non-collider on p is not in $\mathbf{Z}$, and every collider on p has a descendant in $\mathbf{Z}$ (Richardson, 2003). Otherwise $\mathbf{Z}$ blocks p. If $\mathcal{G}$ is a DAG or MAG (defined later) and if $\mathbf{Z}$ blocks all paths between X and Y, we say that X and Y are m-separated given $\mathbf{Z}$ in $\mathcal{G}$. Otherwise, X and Y are m-connected given $\mathbf{Z}$ in $\mathcal{G}$. For pairwise disjoint subsets $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ of $\mathbf{V}$ in $\mathcal{G}$, we say that $\mathbf{X}$ and $\mathbf{Y}$ are m-separated given $\mathbf{Z}$ in $\mathcal{G}$ if X and Y are m-separated given $\mathbf{Z}$ in $\mathcal{G}$ for any $X \in \mathbf{X}$ and

$Y \in \mathbf{Y}$. Otherwise, $\mathbf{X}$ and $\mathbf{Y}$ are m-connected given $\mathbf{Z}$ in $\mathcal{G}$. In a DAG, m-separation and m-connection simplify to d-separation and d-connection (Pearl, 2009).

Causal Bayesian networks. A directed graph without directed cycles is a *directed acyclic graph* (DAG). A Bayesian network for a set of variables $\mathbf{V} = \{X_1, \dots, X_p\}$ is a pair $(\mathcal{G}, f)$, where $\mathcal{G}$ is a DAG, and f is a joint density for $\mathbf{V}$ that factorizes as $f(\mathbf{V}) = \prod_{i=1}^{p} f(X_i | \operatorname{Pa}(X_i, \mathcal{G}))$ (Pearl, 2009). We call a DAG *causal* if every edge $X_i \to X_j$ in $\mathcal{G}$ represents a direct causal effect of X_i on X_j. A Bayesian network $(\mathcal{G}, f)$ is a *causal Bayesian network* if $\mathcal{G}$ is a causal DAG. If a causal Bayesian network is given and all variables are observed, one can easily derive post-intervention densities. In particular, we consider interventions $do(\mathbf{X} = \mathbf{x})$, or shorthand $do(\mathbf{x})$, $(\mathbf{X} \subseteq \mathbf{V})$, which represent outside interventions that set $\mathbf{X}$ to $\mathbf{x}$, uniformly in the population (see Pearl, 2009):

$$f(\mathbf{v} \mid do(\mathbf{x})) = \begin{cases} \prod_{\{i | X_i \in \mathbf{V} \setminus \mathbf{X}\}} f(x_i \mid \operatorname{Pa}(x_i, \mathcal{G})), & \text{if } \mathbf{v} \text{ is consistent with } \mathbf{x}, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Equation (1) is known as the truncated factorization formula (Pearl, 2009), the g-formula (Robins, 1986) or the manipulated density (Spirtes et al., 2000).

Maximal ancestral graphs. A mixed graph $\mathcal{G}$ without directed cycles and almost directed cycles is called *ancestral*. A *maximal ancestral graph* (MAG) is an ancestral graph $\mathcal{G} = (\mathbf{V}, \mathbf{E})$ where every pair of non-adjacent nodes X and Y in $\mathcal{G}$ can be m-separated by a set $\mathbf{Z} \subseteq \mathbf{V} \setminus \{X, Y\}$. A DAG with unobserved variables can be uniquely represented by a MAG on the observed variables that preserves the ancestral and m-separation relationships among the observed variables (page 981 in Richardson and Spirtes, 2002). Since we consider MAGs that do not encode selection bias, the MAGs in this paper can only contain directed $(\to)$ and bi-directed $(\leftrightarrow)$ edges. The MAG of a causal DAG is a *causal* MAG.

Markov equivalence. Several DAGs can encode the same conditional independencies via d-separation. Such DAGs form a *Markov equivalence class* which can be described uniquely by a *completed partially directed acyclic graph* (CPDAG) (Meek, 1995). A CPDAG $\mathcal{C}$ has the same adjacencies as any DAG in the Markov equivalence class described by $\mathcal{C}$. A directed edge $X \to Y$ in a CPDAG $\mathcal{C}$ corresponds to a directed edge $X \to Y$ in every DAG in the Markov equivalence class described by $\mathcal{C}$. For any non-directed edge $X \circ\!\!-\!\!\circ Y$ in a CPDAG $\mathcal{C}$, the Markov equivalence class described by $\mathcal{C}$ contains a DAG with $X \to Y$ and a DAG with $X \leftarrow Y$[1]. Thus, CPDAGs only contain directed $(\to)$ and non-directed $(\circ\!\!-\!\!\circ)$ edges.

Several MAGs can also encode the same conditional independencies via m-separation. Such MAGs form a Markov equivalence class which can be described uniquely by a *partial ancestral graph* (PAG) (Richardson and Spirtes, 2002; Ali et al., 2009). A PAG $\mathcal{P}$ has the same adjacencies as any MAG in the Markov equivalence class described by $\mathcal{P}$. Any non-circle edge-mark in a PAG $\mathcal{P}$ corresponds to that same non-circle edge-mark in every MAG in the Markov equivalence class described by $\mathcal{P}$. We only consider maximally informative PAGs (Zhang, 2008b), that is, for any circle mark $X \circ\!\!-\!\!\bullet$ in a PAG $\mathcal{P}$, the Markov equivalence class described by $\mathcal{P}$ contains a MAG with $X \leftarrow\!\!\bullet Y$ and a MAG with $X -\!\!\bullet Y$. We denote all

1. The non-directed edges in a CPDAG, which we denote as $\circ\!\!-\!\!\circ$, are often denoted as $-$ in the relevant literature, see for example (Meek, 1995). We use $\circ\!\!-\!\!\circ$ instead of $-$ for the sake of consistency among different graph classes.

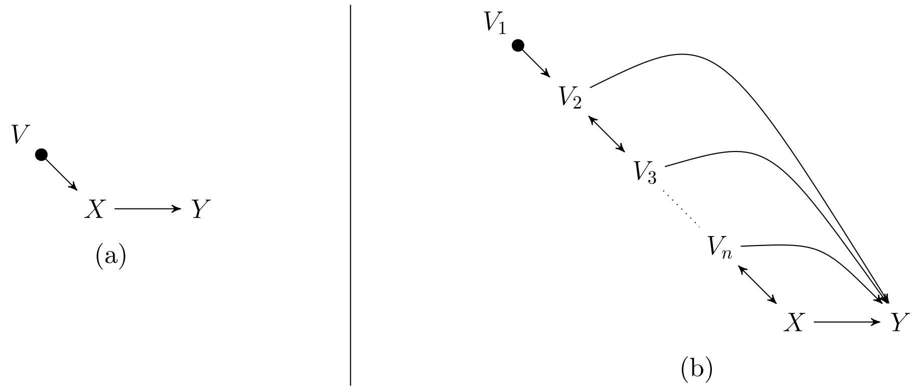

Figure 2: Two configurations where the edge $X \to Y$ is visible. Nodes V and Y must be nonadjacent in 2a, and V_1 and Y must be nonadjacent in 2b.

DAGs (MAGs) in the Markov equivalence class described by a CPDAG (PAG) $\mathcal{G}$ by $[\mathcal{G}]$. The CPDAG (PAG) of a causal DAG (MAG) is a *causal* CPDAG (PAG).

Consistent densities. A density f is *consistent* with a causal DAG $\mathcal{D}$ if the pair $(\mathcal{D}, f)$ forms a causal Bayesian network. A density f is consistent with a causal MAG $\mathcal{M}$ if there exists a causal Bayesian network $(\mathcal{D}', f')$ such that $\mathcal{M}$ represents $\mathcal{D}'$ and f is the observed marginal of f'. A density f is consistent with a causal CPDAG (PAG) $\mathcal{G}$ if it is consistent with a causal DAG (MAG) in $[\mathcal{G}]$.

Visible and invisible edges. All directed edges in DAGs and CPDAGs are said to be visible. Given a MAG $\mathcal{M}$ or a PAG $\mathcal{G}$, a directed edge $X \to Y$ is *visible* if there is a node V not adjacent to Y such that there is an edge $V \bullet\!\!\rightarrow X$, or if there is a collider path between V and X that is into X and every non-endpoint node on the path is a parent of Y, see Figure 2 (Zhang, 2006). A visible edge $X \to Y$ means that there are no latent confounders between X and Y. A directed edge $X \to Y$ that is not visible in a MAG $\mathcal{M}$ or a PAG $\mathcal{G}$ is said to be *invisible*. In the FCI algorithm, invisible edges can occur due to orientation rules **R5 - R10** of Zhang (2008b). When considering MAGs and PAGs that do not encode selection bias, invisible edges occur as a consequence of the orientation rules **R8 - R10** of Zhang (2008b).

3. The Generalized Adjustment Criterion

Throughout, let $\mathcal{G} = (\mathbf{V}, \mathbf{E})$ represent a DAG, CPDAG, MAG or PAG, and let $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ be pairwise disjoint subsets of $\mathbf{V}$, with $\mathbf{X} \neq \emptyset$ and $\mathbf{Y} \neq \emptyset$. Here, $\mathbf{X}$ represents the set of exposures and $\mathbf{Y}$ represents the set of response variables.

We will define sound and complete graphical conditions for adjustment sets relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$. Thus, if a set $\mathbf{Z}$ satisfies our conditions relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$ (see Definition 4), then it is a valid adjustment set for calculating the causal effect of $\mathbf{X}$ on $\mathbf{Y}$ (see Definition 1), and every existing valid adjustment set satisfies our conditions (see Theorem 5). First, we define what we mean by an adjustment set.

Definition 1 *(**Adjustment set**; Maathuis and Colombo, 2015) Let* $\mathbf{X}, \mathbf{Y}$ *and* $\mathbf{Z}$ *be pairwise disjoint node sets in a causal* DAG*,* CPDAG*,* MAG *or* PAG $\mathcal{G}$*. Then* $\mathbf{Z}$ *is an adjustment set relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{G}$ *if for any density*[2] f *consistent with* $\mathcal{G}$ *we have*

$$f(\mathbf{y} \mid do(\mathbf{x})) = \begin{cases} f(\mathbf{y} \mid \mathbf{x}) & \text{if } \mathbf{Z} = \emptyset, \\ \int_{\mathbf{z}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z}) f(\mathbf{z}) d\mathbf{z} & \text{otherwise.} \end{cases} \tag{2}$$

Thus, adjustment sets allow post-intervention densities involving the do-operator (left-hand side of Equation 2) to be identified as specific functions of conditional densities (right-hand side of Equation 2). The latter can be estimated from observational data. As a result, adjustment sets are important for the computation of causal effects. This is illustrated in Example 1 for the special case of multivariate Gaussian densities.

Example 1 *Suppose* f *is a multivariate Gaussian density that is consistent with a causal* DAG $\mathcal{D}$*. Let* $\mathbf{Z} \neq \emptyset$ *be an adjustment set relative to two distinct variables* X *and* Y *in* $\mathcal{D}$ *such that* $\mathbf{Z} \cap \{X \cup Y\} = \emptyset$*. Then*

$$\begin{aligned} E(Y \mid do(x)) &= \int_y y f(y \mid do(x)) dy = \int_y y \int_{\mathbf{z}} f(y \mid x, \mathbf{z}) f(\mathbf{z}) d\mathbf{z} dy \\ &= \int_{\mathbf{z}} \int_y y f(y \mid x, \mathbf{z}) dy f(\mathbf{z}) d\mathbf{z} = \int_{\mathbf{z}} E(Y \mid x, \mathbf{z}) f(\mathbf{z}) d\mathbf{z} \\ &= \int_{\mathbf{z}} (\alpha + \gamma x + \beta^T \mathbf{z}) f(\mathbf{z}) d\mathbf{z} = \alpha + \gamma x + \beta^T E(\mathbf{Z}), \end{aligned}$$

where we use the fact that all conditional expectations in a multivariate Gaussian distribution are linear, so that $E(Y \mid x, \mathbf{z}) = \alpha + \gamma x + \beta^T \mathbf{z}$*, for some* $\alpha, \gamma \in \mathbb{R}$ *and* $\beta \in \mathbb{R}^{|\mathbf{Z}|}$*. Defining the total causal effect of* X *on* Y *as* $\frac{\partial}{\partial x} E(Y \mid do(x))$*, we obtain that the total causal effect of* X *on* Y *is* γ*, that is, the regression coefficient of* X *in the regression of* Y *on* X *and* $\mathbf{Z}$*.*

Our first goal in this paper is to give a graphical criterion (see Definition 4) that is equivalent to Definition 1. To this end, we introduce some additional terminology.

Definition 2 *(**Amenability**) Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be disjoint node sets in a* DAG*,* CPDAG*,* MAG *or* PAG $\mathcal{G}$*. Then* $\mathcal{G}$ *is said to be amenable relative to* $(\mathbf{X}, \mathbf{Y})$ *if every proper possibly directed path from* $\mathbf{X}$ *to* $\mathbf{Y}$ *in* $\mathcal{G}$ *starts with a visible edge out of* $\mathbf{X}$*.*

If $\mathcal{G}$ is a MAG, then Definition 2 reduces to the notion of amenability as introduced in van der Zander et al. (2014). The intuition behind the concept of amenability is the following: In MAGs and PAGs, directed edges $X \to Y$ can represent causal effects, but also mixtures of causal effects and latent confounding. For instance, when the graph $X \to Y$ is interpreted as a DAG, the empty set is a valid adjustment set with respect to (X, Y). When the same graph is interpreted as a MAG, it can still represent the DAG $X \to Y$, but also the DAG $X \to Y$ with an additional non-causal path $X \leftarrow L \to Y$ where L is latent.

2. We use the notation for continuous random variables throughout. The discrete analogues should be obvious.

In CPDAGs and PAGs, there are edges with unknown direction. This complicates adjustment because paths containing such edges can correspond to causal paths in some represented DAGs and to non-causal paths in others. For example, the CPDAG $X \circ\!\!-\!\!\circ Y$ represents the DAGs $X \to Y$ and $X \leftarrow Y$. Amenable graphs are graphs where these problems do not occur.

Definition 3 ***(Forbidden set;*** $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$*)* *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be disjoint node sets in a* DAG*,* CPDAG*,* MAG *or* PAG $\mathcal{G}$*. Then the forbidden set relative to* $(\mathbf{X}, \mathbf{Y})$ *is defined as*

$$\begin{aligned}\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) = \{W' \in \mathbf{V} : W' \in \text{PossDe}(W, \mathcal{G}), \textit{ for some } W \notin \mathbf{X} \\ \textit{which lies on a proper possibly directed path from } \mathbf{X} \textit{ to } \mathbf{Y} \textit{ in } \mathcal{G}\}.\end{aligned}$$

Definition 4 ***(Generalized adjustment criterion)*** *Let* $\mathbf{X}, \mathbf{Y}$ *and* $\mathbf{Z}$ *be pairwise disjoint node sets in a* DAG*,* CPDAG*,* MAG *or* PAG $\mathcal{G}$*. Then* $\mathbf{Z}$ *satisfies the generalized adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{G}$ *if the following three conditions hold:*

(Amenability) $\mathcal{G}$ *is adjustment amenable relative to* $(\mathbf{X}, \mathbf{Y})$*, and*

(Forbidden set) $\mathbf{Z} \cap \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) = \emptyset$*, and*

(Blocking) *all proper definite status non-causal paths from* $\mathbf{X}$ *to* $\mathbf{Y}$ *are blocked by* $\mathbf{Z}$ *in* $\mathcal{G}$*.*

If $\mathcal{G}$ is a DAG (MAG), our criterion reduces to the adjustment criterion of Shpitser (2012), (van der Zander et al., 2014) (see Definition 55 in Appendix E). For consistency, however, we will refer to the generalized adjustment criterion for all graph types.

We note that the amenability condition does not depend on $\mathbf{Z}$. In other words, if the amenability condition is violated, then no set satisfies the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$. The forbidden set contains nodes that cannot be used for adjustment. We will try to give some intuition. For simplicity, we consider $\mathbf{X} = \{X\}$ and $\mathbf{Y} = \{Y\}$ in a DAG $\mathcal{D}$, and we are interested in estimating the total causal effect of X on Y in $\mathcal{D}$. It is clear that nodes on any causal path from X to Y in $\mathcal{D}$ should not be included in the set used for adjustment, since including such nodes would block the causal path.

To understand why we cannot include descendants of nodes on a causal path from X to Y in $\mathcal{D}$ (except for descendants of X), it is useful to consider *walks* from X to Y in $\mathcal{D}$, that is, sequences of nodes $\langle X, \ldots, Y\rangle$ in which every pair of successive nodes is adjacent in $\mathcal{D}$ (but the nodes are not necessarily distinct). A walk $r = \langle X = V_0, V_1, \ldots, V_k = Y\rangle$ is non-causal if $V_i \leftarrow\!\!\bullet V_{i+1}$ for at least one $i \in \{0, \ldots, k-1\}$. A walk r from X to Y in $\mathcal{D}$ is *connecting* given a set of nodes $\mathbf{Z}$ if $\mathbf{Z}$ contains all colliders on r and no non-collider on r is in $\mathbf{Z}$. If a walk r is not connecting given $\mathbf{Z}$, then r is *blocked* by $\mathbf{Z}$. Koster (2002) proved that there is a walk from X to Y that is connecting given $\mathbf{Z}$ in $\mathcal{D}$ if and only if there is path from X to Y that is d-connecting given $\mathbf{Z}$ in $\mathcal{D}$.

Intuitively, all non-causal walks from X to Y should be blocked in order to estimate the total causal effect of X on Y. Now, consider a path p of the form $X \to V_1 \to \cdots \to V_k \to Y$ in $\mathcal{G}$. Assume $V_i \notin \mathbf{Z}$, for all $i \in \{1, \ldots, k\}$. Including a descendant A of V_i in the set $\mathbf{Z}$

leads to walk of the form $X \to \cdots \to V_i \to \cdots \to A \leftarrow \cdots \leftarrow V_i \to \cdots \to V_k \to Y$ being connecting given $\mathbf{Z}$ in $\mathcal{G}$. Hence, including A in the adjustment set opens a non-causal walk from X to Y in $\mathcal{G}$.

We now give the main theorem of this section. Corresponding examples can be found in Section 3.1 and the proof of the theorem is given in Section 3.2.

Theorem 5 *Let $\mathbf{X}, \mathbf{Y}$ and $\mathbf{Z}$ be pairwise disjoint node sets in a causal* DAG, CPDAG, MAG *or* PAG $\mathcal{G}$. *Then $\mathbf{Z}$ is an adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$ (see Definition 1) if and only if $\mathbf{Z}$ satisfies the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$ (see Definition 4).*

Verifying the blocking condition by checking all paths requires keeping track of which paths are non-causal and hence, scales poorly to larger graphs. We therefore give an alternative definition of this condition which relies on m-separation in a so-called proper back-door graph.

Definition 6 ***(Proper back-door graph;** $\mathcal{G}^{pbd}_{\mathbf{XY}}$) Let $\mathbf{X}$ and $\mathbf{Y}$ be disjoint node sets in a* DAG, CPDAG, MAG *or* PAG $\mathcal{G}$. *The proper back-door graph $\mathcal{G}^{pbd}_{\mathbf{XY}}$ is obtained from $\mathcal{G}$ by removing all visible edges out of $\mathbf{X}$ that are on proper possibly directed paths from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$.*

If $\mathcal{G}$ is a DAG or MAG, then Definition 6 reduces to the definition of proper back-door graphs as introduced in van der Zander et al. (2014).

Theorem 7 *Replacing the Blocking condition in Definition 4 with:*

(Separation) *$\mathbf{Z}$ m-separates $\mathbf{X}$ and $\mathbf{Y}$ in $\mathcal{G}^{pbd}_{\mathbf{XY}}$,*

results in a criterion that is equivalent to the generalized adjustment criterion.

If $\mathcal{G}$ is a DAG or MAG, then Theorem 7 reduces to Theorem 4.6 in van der Zander et al. (2014).

3.1 Examples

We now provide some examples that illustrate how the generalized adjustment criterion can be applied.

Example 2 *We return to the* CPDAG $\mathcal{C}$ *in Figure 1(a). $\mathcal{C}$ is amenable relative to (X, Y) and* $\text{Forb}(X, Y, \mathcal{C}) = \{Y\}$. *One can easily verify that any superset of $\{Z, A\}$ or of $\{Z, B\}$ that does not contain X or Y satisfies the generalized adjustment criterion relative to (X, Y) in $\mathcal{C}$.*

Example 3 *To illustrate the concept of amenability, consider Figure 3 with a* PAG $\mathcal{P}$ *in (a), and two* MAG*s $\mathcal{M}_1$ and $\mathcal{M}_2$ in $[\mathcal{P}]$ in (b) and (c). The graphs $\mathcal{P}$ and $\mathcal{M}_1$ are not amenable relative to (X, Y). For $\mathcal{P}$ this is due to the path $X \circ\!\!-\!\!\circ Y$, and for $\mathcal{M}_1$ this is due*

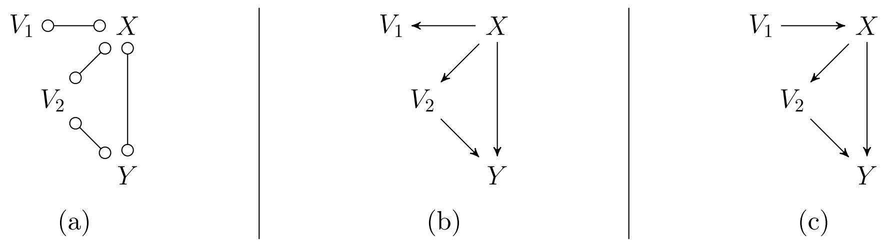

Figure 3: (a) PAG $\mathcal{P}$, (b) MAG $\mathcal{M}_1$, (c) MAG $\mathcal{M}_2$ used in Example 3.

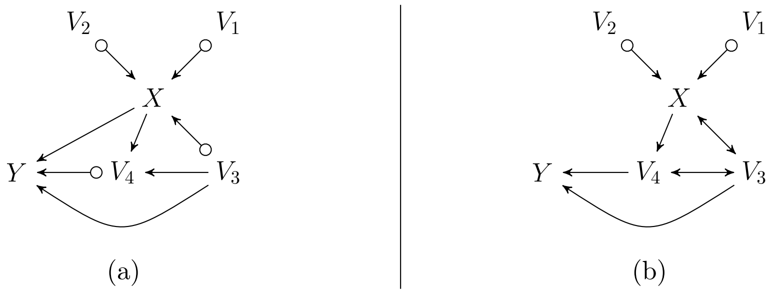

Figure 4: (a) PAG $\mathcal{P}_1$, (b) PAG $\mathcal{P}_2$ used in Example 4 and Example 6.

to the invisible edge $X \rightarrow Y$ (which implies that $\mathcal{M}_1$ also represents a DAG that contains a hidden confounder that is an ancestor of both X and Y). On the other hand, $\mathcal{M}_2$ is amenable relative to (X, Y), since the edges $X \rightarrow Y$ and $X \rightarrow V_2$ are visible due to the edge $V_1 \rightarrow X$, with V_1 not adjacent to Y or V_2. Since there are no proper definite status non-causal paths from X to Y in $\mathcal{M}_2$, it follows that the empty set satisfies the generalized adjustment criterion relative to (X, Y) in $\mathcal{M}_2$. Finally, note that $\mathcal{M}_1$ could also be interpreted as a DAG. *In that case, it would be amenable relative to (X, Y). This shows that amenability depends crucially on the interpretation of the graph.*

Example 4 *Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be the* PAG*s in Figure 4(a) and Figure 4(b), respectively. Both* PAG*s are amenable relative to (X, Y). We will show that there is an adjustment set relative to (X, Y) in $\mathcal{P}_1$ but not in $\mathcal{P}_2$. This illustrates that amenability is not a sufficient criterion for the existence of an adjustment set.*

We first consider $\mathcal{P}_1$. Note that $\mathrm{Forb}(X, Y, \mathcal{P}_1) = \{V_4, Y\}$ and that there are two proper definite status non-causal paths from X to Y: $X \leftarrow\!\!\circ V_3 \rightarrow Y$ and $X \rightarrow V_4 \leftarrow V_3 \rightarrow Y$. Path $X \leftarrow\!\!\circ V_3 \rightarrow V_4 \circ\!\!\rightarrow Y$ is not of definite status, as node V_4 is not of definite status on this path. Both proper definite status non-causal paths from X to Y are blocked by any set containing V_3. Hence, all sets satisfying the generalized adjustment criterion relative to (X, Y) in $\mathcal{P}_1$ are: $\{V_3\}$, $\{V_1, V_3\}, \{V_2, V_3\}$ and $\{V_1, V_2, V_3\}$.

In $\mathcal{P}_2$, we have $\mathrm{Forb}(X, Y, \mathcal{P}_2) = \mathrm{Forb}(X, Y, \mathcal{P}_1) = \{V_4, Y\}$, and there are three proper definite status non-causal paths from X to Y in $\mathcal{P}_2$: p_1 of the form $X \leftrightarrow V_3 \rightarrow Y$, p_2 of the form $X \leftrightarrow V_3 \leftrightarrow V_4 \rightarrow Y$ and p_3 of the form $X \rightarrow V_4 \leftrightarrow V_3 \rightarrow Y$. To block p_1, we must use V_3, and this implies that we must use V_4 to block p_2. But $V_4 \in \mathrm{Forb}(X, Y, \mathcal{P}_2)$. Hence, no set satisfies the generalized adjustment criterion $\mathbf{Z}$ relative to (X, Y) in $\mathcal{P}_2$.

3.2 Proof of Theorem 5

To prove that the generalized adjustment criterion is sound and complete for adjustment (Theorem 5), we build on the fact that the adjustment criterion for DAGs and MAGs is sound and complete for adjustment. The adjustment criterion for DAGs was first presented in Shpitser et al. (2010) and was modified in the unpublished addendum (Shpitser, 2012). In Appendix E we give the revised version of the criterion, as well as new soundness and completeness proofs, relying only on basic probability calculus, linear algebra and the do-calculus rules. The adjustment criterion for MAGs was presented and proved to be sound and complete for adjustment in van der Zander et al. (2014, see Theorem 5.8). However, van der Zander et al. (2014)'s proof assumed the soundness and completeness of the adjustment criterion for DAGs given in Shpitser (2012). This latter claim is proved here in Theorem 56 in Appendix E.

Lastly, our proof of Theorem 5 heavily relies on the three lemmas given below. Their proofs can be found in Appendix B.

Lemma 8 *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be disjoint node sets in a* CPDAG (PAG) $\mathcal{G}$. *If* $\mathcal{G}$ *is amenable (see Definition 4) relative to* $(\mathbf{X},\mathbf{Y})$*, then every* DAG (MAG) *in* $[\mathcal{G}]$ *is amenable relative to* $(\mathbf{X},\mathbf{Y})$*. On the other hand, if* $\mathcal{G}$ *violates the amenability condition relative to* $(\mathbf{X},\mathbf{Y})$*, then there is no adjustment set relative to* $(\mathbf{X},\mathbf{Y})$ *in* $\mathcal{G}$ *(see Definition 1).*

Lemma 9 *Let* $\mathbf{X},\mathbf{Y}$ *and* $\mathbf{Z}$ *be pairwise disjoint node sets in a* CPDAG (PAG) $\mathcal{G}$. *If* $\mathcal{G}$ *is amenable relative to* $(\mathbf{X},\mathbf{Y})$*, then the following statements are equivalent:*

(i) $\mathbf{Z}$ *satisfies the forbidden set condition (see Definition 4) relative to* $(\mathbf{X},\mathbf{Y})$ *in* $\mathcal{G}$.

(ii) $\mathbf{Z}$ *satisfies the forbidden set condition relative to* $(\mathbf{X},\mathbf{Y})$ *in every* DAG (MAG) *in* $[\mathcal{G}]$.

Lemma 10 *Let* $\mathbf{X},\mathbf{Y}$ *and* $\mathbf{Z}$ *be pairwise disjoint node sets in a* CPDAG (PAG) $\mathcal{G}$. *If* $\mathcal{G}$ *is amenable relative to* $(\mathbf{X},\mathbf{Y})$*, and* $\mathbf{Z}$ *satisfies the forbidden set condition relative to* $(\mathbf{X},\mathbf{Y})$*, then the following statements are equivalent:*

(i) $\mathbf{Z}$ *satisfies the blocking condition (see Definition 4) relative to* $(\mathbf{X},\mathbf{Y})$ *in* $\mathcal{G}$.

(ii) $\mathbf{Z}$ *satisfies the blocking condition relative to* $(\mathbf{X},\mathbf{Y})$ *in every* DAG (MAG) *in* $[\mathcal{G}]$.

(iii) $\mathbf{Z}$ *satisfies the blocking condition relative to* $(\mathbf{X},\mathbf{Y})$ *in a* DAG $\mathcal{D}$ (MAG $\mathcal{M}$) *in* $[\mathcal{G}]$.

Proof of Theorem 5. If $\mathcal{G}$ is a DAG (MAG), then our criterion reduces to the adjustment criterion from Shpitser (2012) (van der Zander et al., 2014) which is sound and complete for adjustment (Theorem 56 in Appendix E, Theorem 5.8 in van der Zander et al., 2014). Hence, we only consider the case that $\mathcal{G}$ is a CPDAG (PAG).

Suppose first that $\mathbf{Z}$ satisfies the generalized adjustment criterion relative to $(\mathbf{X},\mathbf{Y})$ in the CPDAG (PAG) $\mathcal{G}$. We need to show that $\mathbf{Z}$ is an adjustment set (see Definition 1) relative to $(\mathbf{X},\mathbf{Y})$ in every DAG $\mathcal{D}$ (MAG $\mathcal{M}$) in $[\mathcal{G}]$. By applying Lemmas 8, 9 and 10 in turn, it directly follows that $\mathbf{Z}$ satisfies the generalized adjustment criterion relative to $(\mathbf{X},\mathbf{Y})$ in every DAG $\mathcal{D}$ (MAG $\mathcal{M}$) in $[\mathcal{G}]$. Since the generalized adjustment criterion is sound for adjustment in DAGs (MAGs) (see Theorem 58 in Appendix E and Theorem 5.8 in van der Zander et al., 2014), $\mathbf{Z}$ is an adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ in every $\mathcal{D}$ $(\mathcal{M})$ in $[\mathcal{G}]$.

To prove the other direction, suppose that $\mathbf{Z}$ does not satisfy the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$. First, suppose that $\mathcal{G}$ violates the amenability condition relative to $(\mathbf{X}, \mathbf{Y})$. Then by Lemma 8, there is no adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$. Otherwise, $\mathcal{G}$ is amenable relative to $(\mathbf{X}, \mathbf{Y})$, but $\mathbf{Z}$ violates the forbidden set condition or the blocking condition. We need to show $\mathbf{Z}$ is not an adjustment set in at least one DAG $\mathcal{D}$ (MAG $\mathcal{M}$) in $[\mathcal{G}]$. Suppose $\mathbf{Z}$ violates the forbidden set condition. Then by Lemma 9, it follows that there exists a DAG $\mathcal{D}$ (MAG $\mathcal{M}$) in $[\mathcal{G}]$ such that $\mathbf{Z}$ does not satisfy the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ $(\mathcal{M})$. Since the generalized adjustment criterion is complete for adjustment in DAGs (MAGs) (see Theorem 57 in Appendix E and Theorem 5.8 in van der Zander et al., 2014), it follows that $\mathbf{Z}$ is not an adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ $(\mathcal{M})$. Otherwise, suppose $\mathbf{Z}$ satisfies the forbidden set condition, but violates the blocking condition. Then by Lemma 10, it follows that there is a DAG $\mathcal{D}$ (MAG $\mathcal{M}$) in $[\mathcal{G}]$ such that $\mathbf{Z}$ does not satisfy the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ $(\mathcal{M})$. Since the generalized adjustment criterion is complete for adjustment in DAGs (MAGs), it follows that $\mathbf{Z}$ is not an adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ $(\mathcal{M})$. ■

4. Constructing Adjustment Sets

We now present approaches to construct adjustment sets. First, in Theorem 11, we discuss a pre-processing of the node set $\mathbf{X}$ that in conjunction with our generalized adjustment criterion can help identify $f(\mathbf{y}|do(\mathbf{x}))$ in DAG, CPDAG, MAG or PAG $\mathcal{G}$ via adjustment.

As mentioned before, if $f(\mathbf{y}|do(\mathbf{x}))$ is not identifiable via adjustment in $\mathcal{G}$, it may be identifiable through other means. In particular, if $f(\mathbf{y}|do(\mathbf{x}))$ is not identifiable via adjustment in $\mathcal{G}$, there may be a set $\mathbf{X}' \subseteq \mathbf{X}$ such that $f(\mathbf{y}|do(\mathbf{x})) = f(\mathbf{y}|do(\mathbf{x}'))$, and $f(\mathbf{y}|do(\mathbf{x}'))$ is identifiable via adjustment in $\mathcal{G}$. One such example is given in Theorem 11.

Next, we introduce Theorem 14 that will allow us to easily construct adjustment sets that do not contain certain nodes, if any such adjustment set exists. We illustrate the results of Theorem 14 with examples in Section 4.1 and give the proof of this theorem in Section 4.2. In Section 4.3 we explain how to leverage previous results of van der Zander et al. (2014) to enumerate all (minimal) adjustment sets, and discuss how to implement this procedure efficiently.

Theorem 11 *Let $\mathbf{X}$ and $\mathbf{Y}$ be disjoint node sets in a causal DAG, CPDAG, MAG or PAG $\mathcal{G}$. Let $\mathbf{X}' \subseteq \mathbf{X}$ such that there is no possibly directed path from $\mathbf{X} \setminus \mathbf{X}'$ to $\mathbf{Y}$ that is proper with respect to $\mathbf{X}$. Then*

$$f(\mathbf{y}|do(\mathbf{x})) = \begin{cases} f(\mathbf{y}) & \text{if } \mathbf{X}' = \emptyset, \\ f(\mathbf{y}|do(\mathbf{x}')) & \text{otherwise.} \end{cases}$$

Furthermore, if $\mathbf{X}' \neq \emptyset$ and if $\mathbf{Z}$ is an adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$, then $\mathbf{Z}$ is an adjustment set relative to $(\mathbf{X}', \mathbf{Y})$ in $\mathcal{G}$.

Following Theorem 11, we recommend pre-processing the set $\mathbf{X}$ as follows: remove all nodes $X \in \mathbf{X}$ that do not have a possibly directed path to $\mathbf{Y}$ which is proper with respect to $\mathbf{X}$

$$\begin{tikzcd}
& & V_2 \arrow[dl] \arrow[dr] & & \\
V_1 \arrow[r] & X_1 \arrow[rr] & & Y \arrow[r, leftrightarrow] & X_2
\end{tikzcd}$$

Figure 5: MAG $\mathcal{M}$.

in $\mathcal{G}$. In other words, if $\mathbf{W}$ is the set of all nodes that have a possibly directed path to $\mathbf{Y}$ which is proper with respect to $\mathbf{X}$ in $\mathcal{G}$, then $\mathbf{X}' = \mathbf{X} \cap \mathbf{W}$. Then by choice of $\mathbf{W}$ and the proof of Theorem 11, all nodes in $\mathbf{X}'$ have a possibly directed path to $\mathbf{Y}$ that is proper with respect to $\mathbf{X}'$.

By Theorem 11, this pre-processing of $\mathbf{X}$ cannot hurt in identifying $f(\mathbf{y}|do(\mathbf{x}))$ via adjustment. Moreover, there are cases when this pre-processing helps to identify $f(\mathbf{y}|do(\mathbf{x}))$ via adjustment. For example, in the MAG $\mathcal{M}$ in Figure 5, there is no adjustment set relative to $(\{X_1, X_2\}, Y)$. However, there is no possibly directed path from X_2 to Y that is proper with respect to $\{X_1, X_2\}$. Furthermore, $\{V_2\}$ is an adjustment set relative to (X_1, Y). Hence, by Theorem 11 and Theorem 5, $f(y|do(x_1, x_2)) = f(y|do(x_1)) = \int_{v_2} f(y|x_1, v_2) f(v_2) dv_2$.

We now introduce two definitions that will be used in Theorem 14. First, we define the set $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ relative to disjoint node sets $\mathbf{X}$ and $\mathbf{Y}$ in a DAG, CPDAG, MAG or PAG $\mathcal{G}$.

Definition 12 *(*$\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$*) Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be disjoint node sets in a* DAG*,* CPDAG*,* MAG *or* PAG $\mathcal{G}$*. We define*

$$\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) = \text{PossAn}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G}) \setminus (\mathbf{X} \cup \mathbf{Y} \cup \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})). \tag{3}$$

If $\mathcal{G}$ is a DAG or MAG, then Definition 12 reduces to the definition of $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ in van der Zander et al. (2014), that is, $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) = \text{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G}) \setminus (\mathbf{X} \cup \mathbf{Y} \cup \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G}))$.

Definition 13 ***(Descendral set)*** *Let* $\mathbf{I}$ *be a node set in a* DAG*,* CPDAG*,* MAG *or* PAG $\mathcal{G}$*. Then* $\mathbf{I}$ *is called descendral in* $\mathcal{G}$ *if* $\mathbf{I} = \text{PossDe}(\mathbf{I}, \mathcal{G})$.

A descendral set is in a sense analogous to an ancestral set, which is a set containing all ancestors of itself. Note that $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ and $\text{PossDe}(\mathbf{X}, \mathcal{G})$ are both descendral sets. This property will be used throughout the proofs. Note that if $\mathbf{A}$ is an ancestral set and $\mathbf{B}$ is a descendral set, then $\mathbf{A} \setminus \mathbf{B}$ is an ancestral set and $\mathbf{B} \setminus \mathbf{A}$ is a descendral set.

Theorem 14 ***(Constructive set)*** *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be disjoint node sets in a* DAG*,* CPDAG*,* MAG *or* PAG $\mathcal{G}$*. Let* $\mathbf{I} \supseteq \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ *be a descendral set in* $\mathcal{G}$*. Then there exists a set* $\mathbf{Z}$ *that satisfies the generalized adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{G}$ *such that* $\mathbf{Z} \cap \mathbf{I} = \emptyset$ *if and only if* $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$ *satisfies the generalized adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{G}$*.*

The smallest set we can take for $\mathbf{I}$ in Theorem 14 is $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$. This leads to Corollary 15. Pearl's back-door criterion does not allow using descendants of $\mathbf{X}$ in a DAG $\mathcal{D}$. Moreover, the generalized back-door criterion does not allow using possible descendants of $\mathbf{X}$ in a DAG, CPDAG, MAG or PAG $\mathcal{G}$. Thus, another natural set to consider for $\mathbf{I}$ is $\text{PossDe}(\mathbf{X}, \mathcal{G})$. We will use $\mathbf{I} = \text{PossDe}(\mathbf{X}, \mathcal{G})$ and Theorem 14 in Section 5 to define sets that satisfy generalized back-door criterion and Pearl's back-door criterion.

Corollary 15 *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be disjoint node sets in a* DAG, CPDAG, MAG *or* PAG $\mathcal{G}$. *The following statements are equivalent:*

(i) *There exists an adjustment set relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{G}$.

(ii) $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ *satisfies the generalized adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{G}$.

(iii) $\mathcal{G}$ *is amenable relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{G}$ *and* $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ *satisfies the blocking condition relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{G}$.

4.1 Examples

We now provide some examples that illustrate the construction of adjustment sets.

Example 5 *Consider again the* CPDAG $\mathcal{C}$ *in Figure 1(a). As previously discussed in Example 2,* $\mathcal{C}$ *is amenable relative to* (X, Y). *The set* $\text{Adjust}(X, Y, \mathcal{C}) = \{X, Y, I, A, Z, B\} \setminus \{X, Y\} = \{I, A, Z, B\}$ *satisfies the blocking condition relative to* (X, Y) *in* $\mathcal{C}$. *Hence, by Corollary 15,* $\{I, A, Z, B\}$ *satisfies the generalized adjustment criterion relative to* (X, Y) *in* $\mathcal{C}$.

Example 6 *Consider again the* PAG*s* $\mathcal{P}_1$ *and* $\mathcal{P}_2$ *in Figure 4(a) and Figure 4(b), respectively. As previously discussed in Example 4, both* $\mathcal{P}_1$ *and* $\mathcal{P}_2$ *are amenable relative to* (X, Y).

In $\mathcal{P}_1$, $\text{Forb}(X, Y, \mathcal{P}_1) = \{V_4, Y\}$, *so* $\text{Adjust}(X, Y, \mathcal{P}_1) = \{X, Y, V_1, V_2, V_3, V_4\} \setminus \{X, Y, V_4\} = \{V_1, V_2, V_3\}$. *Since* $\{V_1, V_2, V_3\}$ *satisfies the blocking condition relative to* (X, Y) *in* $\mathcal{G}$, *it follows that* $\{V_1, V_2, V_3\}$ *satisfies the generalized adjustment criterion relative to* (X, Y) *in* $\mathcal{G}$.

In $\mathcal{P}_2$, *again* $\text{Forb}(X, Y, \mathcal{P}_2) = \{V_4, Y\}$, *so* $\text{Adjust}(X, Y, \mathcal{P}_2) = \{X, Y, V_1, V_2, V_3, V_4\} \setminus \{X, Y, V_4\} = \{V_1, V_2, V_3\}$. *Since* $\{V_1, V_2, V_3\}$ *does not block the path* $X \leftrightarrow V_3 \leftrightarrow V_4 \to Y$ *it does not satisfy the blocking condition relative to* (X, Y) *in* $\mathcal{G}$. *Hence, Corollary 15 implies that there is no adjustment set relative to* (X, Y) *in* $\mathcal{P}_2$.

4.2 Proof of Theorem 14

To prove Theorem 14 we heavily rely on Lemma 16 and Lemma 17 given below. Their proofs are given in Appendix C. Lemma 17 is related to Lemma 1 from Richardson (2003) (see Lemma 37 in Appendix A).

Lemma 16 *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be disjoint node sets in a* DAG, CPDAG, MAG *or* PAG $\mathcal{G}$. *Let* $\mathbf{I} \supseteq \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ *be a descendral set in* $\mathcal{G}$ *(see Definition 13). If there is a proper definite status non-causal path from* $\mathbf{X}$ *to* $\mathbf{Y}$ *in* $\mathcal{G}$ *that is m-connecting given* $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$, *then there is a path* p *from* $\boldsymbol{X}$ *to* $\boldsymbol{Y}$ *in* $\mathcal{G}$ *such that:*

(i) p is a proper definite status non-causal path from $\mathbf{X}$ *to* $\mathbf{Y}$ *in* $\mathcal{G}$*, and*

(ii) all colliders on p are in $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$*, and*

(iii) all definite non-colliders on p are in $\mathbf{I}$*, and*

(iv) for any collider C on p, there is an unshielded possibly directed path from C to $\mathbf{X} \cup \mathbf{Y}$*, that starts with* $\circ\!\!\rightarrow$ *or* $\rightarrow$*.*

Lemma 17 *Let* $\mathbf{X}, \mathbf{Y}$ *and* $\mathbf{Z}$ *be pairwise disjoint node sets in a* DAG*,* CPDAG*,* MAG *or* PAG $\mathcal{G}$*. Let* $\mathbf{I} \supseteq \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ *be a node set in* $\mathcal{G}$ *such that* $\mathbf{Z} \cap \mathbf{I} = \emptyset$*. Let p be a path from* $\mathbf{X}$ *to* $\mathbf{Y}$ *in* $\mathcal{G}$ *such that:*

(i) p is a proper definite status non-causal path from $\mathbf{X}$ *to* $\mathbf{Y}$ *in* $\mathcal{G}$*, and*

(ii) all colliders on p are in $\text{An}(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z}, \mathcal{G}) \setminus \mathbf{I}$*, and*

(iii) no definite non-collider on p is in $\mathbf{Z}$*.*

Then there is a proper definite status non-causal path from $\mathbf{X}$ *to* $\mathbf{Y}$ *that is m-connecting given* $\mathbf{Z}$ *in* $\mathcal{G}$*.*

Proof of Theorem 14. We only prove the non-trivial direction. Thus, assume there is a set $\mathbf{Z}$ satisfying the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$ such that $\mathbf{Z} \cap \mathbf{I} = \emptyset$. We will prove that $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$ satisfies the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$.

Since $\mathbf{Z}$ satisfies the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$, $\mathcal{G}$ is amenable relative to $(\mathbf{X}, \mathbf{Y})$. Additionally, since $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \subseteq \mathbf{I}$, $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$ satisfies the forbidden set condition relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$. It is only left to prove that $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$ satisfies the blocking condition relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$.

Suppose for a contradiction that there is a proper definite status non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ that is m-connecting given $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$. Then we can choose a path p^* in $\mathcal{G}$ that satisfies (i)$-$(iv) in Lemma 16. Then p^* also satisfies (i) in Lemma 17. By (iii) in Lemma 16, every definite non-collider on p^* is in $\mathbf{I}$. Since $\mathbf{Z} \cap \mathbf{I} = \emptyset$, no definite non-collider on p^* is in $\mathbf{Z}$. So p^* satisfies (iii) in Lemma 17. Also, since by (iv) in Lemma 16 there is a possibly directed unshielded path q^* from every collider C on p^* to $\mathbf{X} \cup \mathbf{Y}$ that starts with $C \circ\!\!\rightarrow$, Lemma 42 implies that any other edge on q^* (if there is any) is directed in $\mathcal{G}$.

Then if $\mathcal{G}$ is a DAG, CPDAG or MAG, it follows from (iv) in Lemma 16 that all colliders on p^* are in $\text{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G})$. Combining this with (ii) in Lemma 16 implies that all colliders on p^* are in $\text{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$, so that p^* satisfies (ii) in Lemma 17. Hence, if $\mathcal{G}$ is a DAG, CPDAG or MAG, all conditions of Lemma 17 are satisfied, which implies that $\mathbf{Z}$ does not satisfy the blocking condition relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$.

Thus, assume $\mathcal{G}$ is a PAG. Let $\mathcal{M} \in [\mathcal{G}]$ be a MAG obtained from $\mathcal{G}$ by first replacing all partially directed edges $\circ\!\!\rightarrow$ by directed edges $\rightarrow$, and then orienting all non-directed edges $\circ\!\!-\!\!\circ$ as a DAG without unshielded colliders (see Lemma 43 in Appendix A). Let p be the path in $\mathcal{M}$ corresponding to p^* in $\mathcal{G}$. Then p satisfies (i) and (iii) in Lemma 17. By the choice of $\mathcal{M}$ and Lemma 42, any possibly directed unshielded path q^* in $\mathcal{G}$ that starts with a partially directed edge $\circ\!\!\rightarrow$, corresponds to a directed path q in $\mathcal{M}$. Hence, by (iv) in Lemma 16, every collider on p is in $\text{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{M})$. Since p^* satisfies (ii) in Lemma 16 in $\mathcal{G}$, no collider on p^* is in $\mathbf{I}$. Hence, also no collider on p is in $\mathbf{I}$. Then all colliders on p are in $\text{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{M}) \setminus \mathbf{I}$. Additionally, since $\mathbf{I} \supseteq \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$, and $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \supseteq \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{M})$, it follows that p satisfies (ii) in Lemma 17. Thus, all conditions of Lemma 17 are satisfied, which implies that $\mathbf{Z}$ does not satisfy the blocking condition relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{M}$. This contradicts Lemma 10. ■

4.3 Implementation

We now discuss how one can implement the generalized adjustment criterion in an algorithmically efficient manner, and describe our implementation in the software `dagitty` and the R package `pcalg`.

Verification of the criterion. Given a DAG, CPDAG, MAG or PAG $\mathcal{G}$ and three disjoint node sets $\mathbf{X}, \mathbf{Y}$, and $\mathbf{Z}$, we wish to test whether $\mathbf{Z}$ fulfills the generalized adjustment criterion with respect to $(\mathbf{X}, \mathbf{Y})$. Of course, we could do this simply by verifying the three conditions of the generalized adjustment criterion (see Definition 4). However, the blocking condition is a statement about individual paths, which can pose problems for large graphs.

As a worst-case example, consider a DAG with X, Y and p remaining variables. Let every pair of variables be connected by an edge. Then the DAG contains $\sum_{i=0}^{p} p!/i! \approx p!\, e$ paths from X to Y. Thus, a direct implementation of the generalized adjustment criterion has an exponential runtime in p. Still, a verbatim implementation of the criterion can be useful for verification and didactic purposes, as well as for sparse graphs with few paths, and we therefore provide one in the function `gac` of the R package `pcalg`.

The key result for implementing the criterion in an efficient manner is Theorem 7, which replaces the path blocking condition by an m-separation condition in a subgraph of $\mathcal{G}$, the proper back-door graph $\mathcal{G}^{pbd}_{\mathbf{XY}}$. This condition can be checked efficiently by a simple depth-first or breadth-first graph traversal, known as the "Bayes-Ball algorithm" (Shachter, 1998). Specifically, for graphs represented as adjacency lists, the runtime is $O(|p| + |\mathbf{E}|)$ where $|\mathbf{E}|$ is the number of edges. Our implementation of this method can be accessed via the function `isAdjustmentSet` of the R package `dagitty`.

Constructing adjustment sets. Given a DAG, CPDAG, MAG or PAG $\mathcal{G}$ and two disjoint variable sets $\mathbf{X}, \mathbf{Y}$, we wish to find one or several sets $\mathbf{Z}$ that fulfill the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$. If a single set is sufficient, we can directly apply the main result of Section 4 and construct $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ (see Definition 12) and verify whether it satisfies the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$. Since this set is defined in terms of (possible) ancestors of $\mathbf{X}$ and $\mathbf{Y}$, it can be constructed by graph traversal in linear time. However, $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ can be a large set, since it contains all (possible) ancestors of $\mathbf{X}$ and $\mathbf{Y}$ except the forbidden nodes. This may result in a loss of statistical precision.

To avoid this, it is of interest to construct all possible adjustment sets. Again, Theorem 7 is key to achieving this: it allows us to use the algorithmic framework developed by van der Zander et al. (2014) for constructing and enumerating m-separating sets in DAGs and MAGs. For DAGs and MAGs, this can be directly applied. We propose the following procedure for CPDAGs and PAGs:

For a given CPDAG or PAG $\mathcal{G}$ and disjoint node sets $\mathbf{X}$ and $\mathbf{Y}$,

(1) Check if $\mathcal{G}$ is amenable relative to $(\mathbf{X}, \mathbf{Y})$. If not, stop because there is no adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$. Otherwise, continue.

(2) Find $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$.

(3) If $\mathcal{G}$ is a CPDAG, orient $\mathcal{G}$ into a DAG $\mathcal{D}$ in $[\mathcal{G}]$. If $\mathcal{G}$ is a PAG, orient $\mathcal{G}$ into a MAG $\mathcal{M}$ in $[\mathcal{G}]$ according to Theorem 2 from Zhang (2008b) (see Lemma 43 in Appendix A).

(4) By Lemma 10, finding all sets satisfying the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$ is equivalent to finding all sets $\mathbf{Z}$ satisfying the separation condition relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ $(\mathcal{M})$ such that $\mathbf{Z} \cap (\mathbf{X} \cup \mathbf{Y} \cup \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})) = \emptyset$. Thus, we can apply the algorithms from van der Zander et al. (2014) on $\mathcal{D}$ $(\mathcal{M})$. These algorithms are able to deal with the additional restriction that the resulting set must not contain nodes in $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$.

Thus, through this simple procedure we gain complete access to all functions in the algorithmic framework by van der Zander et al. (2014). These include listing all adjustment sets and all minimal adjustment sets (in polynomial time per set that is listed). We have implemented these features in the function `adjustmentSets` of the R package `dagitty`.

5. Relationship to (Generalized) Back-door Criteria

We now discuss the relationship between our generalized adjustment criterion and some other existing graphical criteria for covariate adjustment. In particular, we discuss Pearl's back-door criterion (see Definition 18) and the generalized back-door criterion (see Definition 20) and give constructive sets for both in Section 5.1. We use the results from Section 5.1 to precisely characterize the differences between our generalized adjustment criterion, Pearl's back-door criterion and the generalized back-door criterion in Theorem 26 of Section 5.2. We illustrate the results of Sections 5.1 and 5.2 with examples in Section 5.3.

Definition 18 *(**Back-door criterion**; Pearl, 1993) Let X and Y be distinct nodes in a* DAG $\mathcal{D}$. *A set of nodes $\mathbf{Z}$ not containing X or Y satisfies the back-door criterion relative to (X, Y) in $\mathcal{D}$ if:*

(i) no node in $\mathbf{Z}$ is a descendant of X, and

(ii) $\mathbf{Z}$ blocks every path between X and Y that contains an arrow into X.

If $\mathbf{X}$ and $\mathbf{Y}$ are two disjoint sets of nodes in $\mathcal{D}$, then $\mathbf{Z}$ is said to satisfy the back-door criterion relative to $(\mathbf{X}, \mathbf{Y})$ if it satisfies the criterion relative to any pair (X, Y) such that $X \in \mathbf{X}$, and $Y \in \mathbf{Y}$. A set $\mathbf{Z}$ that satisfies the back-door criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ is called a back-door set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$.

Definition 19 *(**Back-door path**; Maathuis and Colombo, 2015) Let X and Y be distinct nodes in a* DAG, CPDAG, MAG *or* PAG $\mathcal{G}$. *A path from X to Y in $\mathcal{G}$ is a back-door path if it does not start with a visible edge out of X.*

Definition 20 *(**Generalized back-door criterion**; Maathuis and Colombo, 2015) Let* $\mathbf{X}, \mathbf{Y}$ *and* $\mathbf{Z}$ *be pairwise disjoint node sets in a* DAG, CPDAG, MAG *or* PAG $\mathcal{G}$. *Then* $\mathbf{Z}$ *satisfies the generalized back-door criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{G}$ *if:*

(i) $\mathbf{Z}$ *does not contain possible descendants of* $\mathbf{X}$ *in* $\mathcal{G}$, *and*

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

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

5.1 Constructing (Generalized) Back-door Sets

We first focus on Pearl's back-door criterion. If $|\mathbf{X}| = 1$, the existence and construction of a back-door set in a causal DAG $\mathcal{D}$ is well understood. If $Y \in \mathrm{Pa}(X, \mathcal{D})$, then there is no back-door set relative to (X, Y) in $\mathcal{D}$, but it is obvious that $f(y \mid \mathrm{do}(x)) = f(y)$. If $Y \notin \mathrm{Pa}(X, \mathcal{D})$, then $\mathrm{Pa}(X, \mathcal{D})$ is a back-door set relative to (X, Y) in $\mathcal{D}$.

If $|\mathbf{X}| \geq 1$, the construction of a back-door set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ is less obvious. One could perhaps think that any set $\mathbf{Z}$ that satisfies our generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ such that $\mathbf{Z} \cap \mathrm{De}(\mathbf{X}, \mathcal{D}) = \emptyset$ satisfies Pearl's back-door criterion. This is not true, as shown in Lemma 21 that describes the graphical pattern that appears when there is an adjustment set but no back-door set. An example of a DAG that satisfies (i)-(iv) in Lemma 21 is given in Figure 8. Using this result and Theorem 14 we are able to define a specific set that satisfies Pearl's back-door criterion, when such a set exists. This result is given in Corollary 22.

Lemma 21 *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be disjoint node sets in a* DAG $\mathcal{D}$. *Assume there is a set satisfying the generalized adjustment criterion* $\mathbf{Z}$ *relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}$ *such that* $\mathbf{Z} \cap \mathrm{De}(\mathbf{X}, \mathcal{D}) = \emptyset$. *Then there is no back-door set relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}$ *if and only if there is a path* p *from* $\mathbf{X}$ *to* $\mathbf{Y}$ *such that:*

(i) p *is a back-door path, and*

(ii) a proper subpath of p *is a causal path, and*

(iii) there are no colliders on p, *and*

(iv) all nodes on p *are in* $\mathrm{De}(\mathbf{X}, \mathcal{D})$.

Corollary 22 *(**Constructive back-door set**) Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be disjoint node sets in a* DAG $\mathcal{D}$. *The following statements are equivalent:*

(i) There exists a set that satisfies Pearl's back-door criterion relative to $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}$.

(ii) $\mathrm{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) \setminus \mathrm{De}(\mathbf{X}, \mathcal{D})$ *satisfies Pearl's back-door criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}$.

(iii) For all $X \in \mathbf{X}$, $Y \in \mathbf{Y}$, $\mathrm{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) \setminus \mathrm{De}(\mathbf{X}, \mathcal{D})$ *satisfies condition (ii) of Pearl's back-door criterion.*

Maathuis and Colombo (2015) presented a constructive generalized back-door set for a DAG, CPDAG, MAG or PAG $\mathcal{G}$ when $|\mathbf{X}| = 1$. In Lemma 23, we show that any set $\mathbf{Z}$ that satisfies our generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$ such that $\mathbf{Z} \cap \text{PossDe}(\mathbf{X}, \mathcal{G}) = \emptyset$ is a generalized back-door set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$. Using this result and Theorem 14, we give a constructive set that satisfies the generalized back-door criterion, when such a set exists. This set is given in Corollary 24. If $|\mathbf{X}| = 1$, our constructive set for the generalized back-door criterion is a superset of the set presented in Maathuis and Colombo (2015) (Corollary 53 in Appendix D).

Lemma 23 *Let $\mathbf{X}, \mathbf{Y}$ and $\mathbf{Z}$ be pairwise disjoint node sets in a* DAG*,* CPDAG*,* MAG *or* PAG *$\mathcal{G}$. If $\mathcal{G}$ is amenable relative to $(\mathbf{X}, \mathbf{Y})$ and $\mathbf{Z}$ satisfies the blocking condition relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$, then $\mathbf{Z}$ satisfies condition (ii) of the generalized back-door criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$.*

Corollary 24 ***(Constructive generalized back-door set)*** *Let $\mathbf{X}$ and $\mathbf{Y}$ be disjoint node sets in a* DAG*,* CPDAG*,* MAG *or* PAG *$\mathcal{G}$. The following statements are equivalent:*

- *(i) There exists a set that satisfies the generalized back-door criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$.*
- *(ii)* $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \text{PossDe}(\mathbf{X}, \mathcal{G})$ *satisfies the generalized back-door criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$.*
- *(iii) $\mathcal{G}$ is amenable and* $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \text{PossDe}(\mathbf{X}, \mathcal{G})$ *satisfies condition (ii) of the generalized back-door criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$.*

5.2 Graphs for Which the Criteria Differ

We now define graphical conditions for the existence of a set satisfying one, two, or all three of the mentioned criteria. The main result of this section is given in Theorem 26, which describes the 4 graphical patterns that can appear when there is no set satisfying at least one of the criteria.

Previously, in Lemma 8 and Lemma 16, we described such patterns for our generalized adjustment criterion. In Section 5.1, we showed that any set $\mathbf{Z}$ that satisfies our generalized adjustment criterion such that $\mathbf{Z} \cap \text{PossDe}(\mathbf{X}, \mathcal{G}) = \emptyset$ satisfies the generalized back-door criterion. Thus, to describe an additional pattern that appears when there is no set that satisfies the generalized back-door criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$ we give Lemma 25. Lastly, to complete Theorem 26 we add the pattern described in Lemma 21 that additionally appears when there is no set that satisfies Pearl's back-door criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$. Thus, Theorem 26 summarizes and subsumes the results of Lemmas 8, 16, 21 and 25.

Lemma 25 *Let $\mathbf{X}$ and $\mathbf{Y}$ be disjoint node sets in a* DAG*,* CPDAG*,* MAG *or* PAG *$\mathcal{G}$ such that there exists an adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$. There is no set that satisfies the generalized back-door criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$ if and only if there is a path p from $X \in \mathbf{X}$ to $Y \in \mathbf{Y}$ in $\mathcal{G}$ and a node V on p such that:*

- *(i) p is a proper definite status non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$, and* *(ii)* $V \in \text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \cap \text{PossDe}(\mathbf{X}, \mathcal{G})$ *and is a definite non-collider on* p*, and*

(iii) *any collider on* p *is in* $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \text{PossDe}(\mathbf{X}, \mathcal{G})$ *and any definite non-collider on* $p(V, Y)$ *is in* $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$*, and*

(iv) *path* p *is of the form* $X \leftarrow \cdots \leftarrow V \leftarrow\!\!\bullet W \dots Y$*, where* $W = Y$ *is possible and if* $W \neq Y$ *then* $V \leftrightarrow W$ *is on* p*.*

Theorem 26 *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be disjoint node sets in a* DAG*,* CPDAG*,* MAG *or* PAG $\mathcal{G}$*. Consider the following criteria:*

(1) $\mathcal{G}$ *violates the amenability condition relative to* $(\mathbf{X}, \mathbf{Y})$*.*

(2) *There is a proper definite status non-causal path* p *from* $\mathbf{X}$ *to* $\mathbf{Y}$ *in* $\mathcal{G}$ *such that every collider on* p *is in* $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ *and every definite non-collider on* p *is in* $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$*.*

(3) *There is a path* p *from* $\mathbf{X}$ *to* $\mathbf{Y}$ *in* $\mathcal{G}$ *that satisfies (i)–(iv) in Lemma 25.*

(4) *There is a back-door path* p *from* $\mathbf{X}$ *to* $\mathbf{Y}$ *in* $\mathcal{G}$ *that satisfies (i)–(iv) in Lemma 21.*

The following hold:

(i) *There is no set that satisfies the generalized adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{G}$ *if and only if (1) or (2) are satisfied.*

(ii) *There is no generalized back-door set relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{G}$ *if and only if (1), (2) or (3) are satisfied.*

(iii) *If* $\mathcal{G}$ *is a* DAG*, then there is no back-door set relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{G}$ *if and only if (1), (2), (3) or (4) are satisfied.*

We now further explore condition (2) in Theorem 26 under the assumption that the DAG, CPDAG, MAG or PAG $\mathcal{G}$ is amenable relative to disjoint node sets $(\mathbf{X}, \mathbf{Y})$ (that is, (1) in Theorem 26 is violated). Condition (2) in Theorem 26 is satisfied relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$ if and only if there is no adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$ (Theorem 5). Corollary 27 provides a simple sufficient condition for condition (2) in Theorem 26 to be satisfied in DAGs, CPDAGs, MAGs and PAGs, as well as a necessary and sufficient condition for condition (2) in Theorem 26 in certain DAGs and CPDAGs.

Corollary 27 *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be disjoint node sets in a DAG, CPDAG, MAG or PAG* $\mathcal{G}$ *such that* $\mathcal{G}$ *is amenable relative to* $(\mathbf{X}, \mathbf{Y})$*. The following statements hold:*

(i) *If* $\mathbf{X} \cap \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \neq \emptyset$*, then there is no adjustment set relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{G}$*.*

(ii) *Let* $\mathcal{G}$ *be a DAG or CPDAG and* $\mathbf{Y} \subseteq \text{PossDe}(\mathbf{X}, \mathcal{G})$*. Then* $\mathbf{X} \cap \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \neq \emptyset$ *if and only if there is no adjustment set relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{G}$*.*

A necessary condition for both (3) and (4) in Theorem 26 is that $\mathcal{G}$ contains a (possibly) directed path from one node in $\mathbf{X}$ to another node in $\mathbf{X}$. Thus, if $|\mathbf{X}| = 1$, both (3) and (4) in Theorem 26 are violated. Hence, if $|\mathbf{X}| = 1$ there is a (generalized) back-door set relative to $(\mathbf{X}, \mathbf{Y})$ in a DAG (or a CPDAG, MAG or PAG) $\mathcal{G}$, if and only if there is a set satisfying the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$.

We finish this section by giving two corollaries that describe some additional simple conditions under which there exists a set satisfying two or all three of the discussed adjustment criteria. The intuition behind these results is as follows. A necessary condition for the existence of a path p from $\mathbf{X}$ to $\mathbf{Y}$ in a DAG $\mathcal{D}$ that satisfies Lemma 21 is the existence of a causal path from one node in $\mathbf{X}$ to another node in $\mathbf{X}$ in $\mathcal{D}$. This gives us Corollary 28. Similarly, a necessary condition for the existence of a path p from $\mathbf{X}$ to $\mathbf{Y}$ in a DAG, CPDAG, MAG or PAG $\mathcal{G}$ that satisfies Lemma 25 is the existence of a causal path from one node in $\mathbf{X}$ to another node in $\mathbf{X}$ that contains at least one node not in $\mathbf{X}$ in $\mathcal{G}$. Or, in the case when $\mathcal{G}$ is a DAG or CPDAG, another necessary condition for the existence of a path p from $\mathbf{X}$ to $\mathbf{Y}$ that satisfies Lemma 25 in $\mathcal{G}$ is that $\mathbf{Y} \not\subseteq \mathrm{PossDe}(\mathbf{X}, \mathcal{G})$. This gives us Corollary 29.

Corollary 28 *Let $\mathbf{X}$ and $\mathbf{Y}$ be disjoint node sets in a* DAG *$\mathcal{D}$. If there is no directed path from one node in $\mathbf{X}$ to another node in $\mathbf{X}$ in $\mathcal{D}$, then the following statements are equivalent:*

(i) There exists a set that satisfies the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$.

(ii) There exists a back-door set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$.

Corollary 29 *Let $\mathbf{X}$ and $\mathbf{Y}$ be disjoint node sets in a* DAG*,* CPDAG*,* MAG *or* PAG *$\mathcal{G}$. If $\mathcal{G}$ contains no possibly directed path $p = \langle V_1, \dots, V_k \rangle$ with $k \geq 3$ such that $\{V_1, V_k\} \subseteq \mathbf{X}$ and $\{V_2, \dots, V_{k-1}\} \cap \mathbf{X} = \emptyset$, or if $\mathcal{G}$ is a* DAG *or* CPDAG *and $\mathbf{Y} \subseteq \mathrm{PossDe}(\mathbf{X}, \mathcal{G})$, then the following statements are equivalent:*

(i) There exists a set that satisfies the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$.

(ii) There exists a generalized back-door set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$.

5.3 Examples

Figure 3 (see Example 3) in Section 3.1 shows a non-amenable graph (that is, condition (1) in Theorem 26 is satisfied). Figure 4(b) (see Example 4) in Section 3.1 shows an amenable graph for which there is no set that satisfies the generalized adjustment criterion (that is, condition (1) in Theorem 26 is violated, but condition (2) is satisfied).

We now give three additional examples. Figure 6(b) (see Example 7) shows an amenable graph for which there is no set that satisfies the generalized adjustment criterion (that is, (2) in Theorem 26 is satisfied). This example illustrates the result of Corollary 27. Figure 6(a) (see Example 7) and Figure 7 (see Example 8) show cases where there is a set that satisfies the generalized adjustment criterion, but there is no generalized back-door set (that is, conditions (1) and (2) in Theorem 26 are violated, but condition (3) is satisfied).

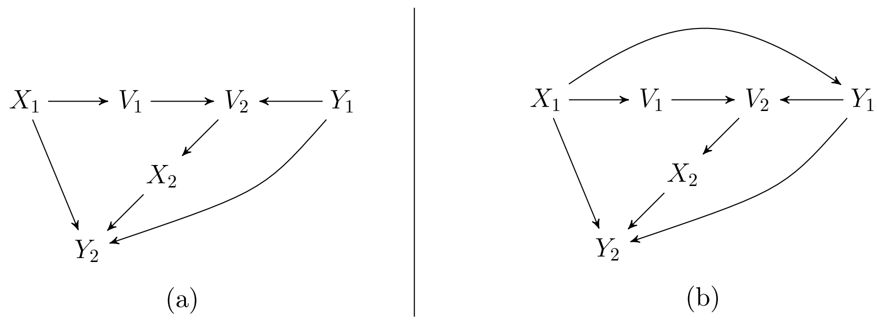

Figure 6: (a) DAG $\mathcal{D}_1$, (b) DAG $\mathcal{D}_2$ used in Example 7.

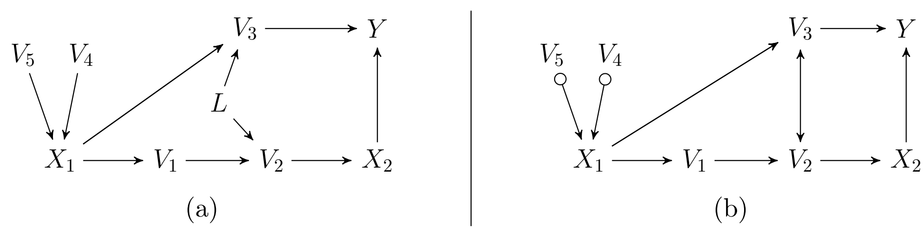

Figure 7: (a) DAG $\mathcal{D}$, (b) PAG $\mathcal{P}$ used in Example 8.

Figure 8 (see Example 9) shows an example of a DAG in which there are sets that satisfy the generalized adjustment criterion and the generalized back-door criterion, but no set satisfies Pearl's back-door criterion (that is, conditions (1), (2) and (3) in Theorem 26 are violated, but condition (4) is satisfied).

Example 7 *Let* $\mathbf{X} = \{X_1, X_2\}$ *and* $\mathbf{Y} = \{Y_1, Y_2\}$ *and consider the* DAG*s* $\mathcal{D}_1$ *and* $\mathcal{D}_2$ *in Figure 6(a) and 6(b) respectively. We first consider* DAG $\mathcal{D}_1$*. The proper non-causal path* $X_2 \leftarrow V_2 \leftarrow Y_1$ *satisfies (i)–(iv) in Lemma 25. Hence, there is no generalized back-door set relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}_1$*. However,* $\{V_1, V_2\}$ *satisfies the generalized adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}_1$*.*

We now consider DAG $\mathcal{D}_2$*. Note that the only difference between* $\mathcal{D}_1$ *and* $\mathcal{D}_2$ *is the additional edge* $X_1 \to Y_1$ *in* $\mathcal{D}_2$*. This edge implies that* $Y_1 \in \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}_2)$*. Hence, the proper non-causal path* $X_2 \leftarrow V_2 \leftarrow Y_1$ *satisfies (2) in Theorem 26 and thus, there is no set that satisfies the generalized adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}_2$*. Then by Theorem 5 there is no adjustment set relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}_2$*. Since* $\mathcal{D}_2$ *is a DAG such that* $\mathbf{Y} \subseteq \mathrm{De}(\mathbf{X}, \mathcal{D}_2)$*, (ii) in Corollary 27 implies that* $\mathbf{X} \cap \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}_2) \neq \emptyset$*. This is indeed true, since* $\mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}_2) = \{X_2, V_2, Y_1, Y_2\}$*.*

Example 8 *Let* $\mathbf{X} = \{X_1, X_2\}$ *and* $\mathbf{Y} = \{Y\}$ *and consider* DAG $\mathcal{D}$ *and* PAG $\mathcal{P}$ *in Figures 7(a) and 7(b). We first consider* DAG $\mathcal{D}$*. Any generalized back-door set relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}$ *must contain* L*. However, the same is not true for the generalized adjustment criterion. For example,* $\{V_1, V_2\}$ *satisfies the generalized adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}$*.*

We now consider $\mathcal{P}$*. Note that* $\mathcal{P}$ *is the PAG of* $\mathcal{D}$ *when* L *is unobserved. The proper non-causal path* $X_2 \leftarrow V_2 \leftrightarrow V_3 \to Y$ *satisfies (i)–(iv) in Lemma 25. Hence, there is no*

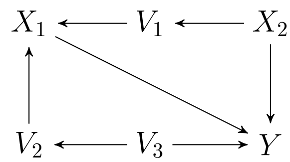

Figure 8: DAG $\mathcal{D}$ used in Example 9.

generalized back-door set relative to $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{P}$*. However, the sets* $\{V_1, V_2\}$*,* $\{V_1, V_2, V_4\}$*,* $\{V_1, V_2, V_5\}$*,* $\{V_1, V_2, V_4, V_5\}$ *all satisfy the generalized adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{P}$*.*

Example 9 *Let* $\mathbf{X} = \{X_1, X_2\}$ *and* $\mathbf{Y} = \{Y\}$ *and consider* DAG $\mathcal{D}$ *in Figure 8. The non-causal path* $X_1 \leftarrow V_1 \rightarrow X_2 \rightarrow Y$ *satisfies (i)–(iv) in Lemma 21. Hence, no set can satisfy the back-door criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}$*. However,* $\{V_2\}$*,* $\{V_3\}$*,* $\{V_2, V_3\}$*,* $\{V_1, V_2\}$*,* $\{V_1, V_3\}$*, and* $\{V_1, V_2, V_3\}$ *all satisfy the generalized adjustment criterion and* $\{V_2\}$*,* $\{V_3\}$ *and* $\{V_2, V_3\}$ *all satisfy the generalized back-door criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}$*.*

6. Discussion

We have derived a generalized adjustment criterion that is sound and complete for adjustment in DAGs, MAGs, CPDAGs and PAGs (see Definition 4, Theorem 5). This is relevant in practice, in particular in combination with algorithms that can learn CPDAGs or PAGs from observational data.

In addition to the criterion itself, we have also given all necessary ingredients for implementing efficient algorithms to test the criterion for a given set and to construct all sets that fulfill it, or to learn that no set fulfilling the criterion exists. Thus, we obtain a complete generalization of the algorithmic framework for DAGs and MAGs by van der Zander et al. (2014) to CPDAGs or PAGs. In this sense, our work presented in this paper is a theoretical contribution that closes the chapter on covariate adjustment for DAGs, CPDAGs, MAGs and PAGs without selection variables.

Correa and Bareinboim (2017) define necessary and sufficient graphical conditions for covariate adjustment in latent projection graphs in the presence of selection variables. Future work may explore how to extend their results to MAGs and PAGs with selection variables. Other future work may study the estimation accuracy of estimators based on different adjustment sets. Any valid adjustment set can be used to produce an unbiased estimator of the total causal effect. However, the efficiency of the estimators induced by distinct adjustment sets varies (Greenland et al., 1999; Kuroki and Miyakawa, 2003; Kuroki and Cai, 2004; Hahn, 2004, 1998; Guo and Dawid, 2010; De Luna et al., 2011). Moreover, Kuroki and Miyakawa (2003) and Kuroki and Cai (2004) indicate that a minimal adjustment set does not necessarily lead to the most efficient estimator. Defining a best adjustment set in terms of efficiency is still an open question.

Acknowledgments

This work was supported in part by Swiss NSF Grants 200021_149760 and 200021_172603.

Appendix A. Preliminaries

We first state various existing results and definitions.

Adjacencies, discriminating paths and d-separations. The set of all nodes adjacent to X in a graph $\mathcal{G}$ is denoted by $\mathrm{Adj}(X,\mathcal{G})$. A path $p = \langle X,\dots,Z,V,Y\rangle$ is a *discriminating path* from X to Y for V in graph $\mathcal{G}$, if it consists of at least four nodes, X is not adjacent to Y in $\mathcal{G}$ and every non-endpoint node on $p(X,V)$ is a collider on p and a parent of Y. If $\mathbf{X}$ and $\mathbf{Y}$ are d-separated given $\mathbf{Z}$ in a DAG $\mathcal{D}$, we write $\mathbf{X} \perp_d \mathbf{Y} \mid \mathbf{Z}$.

Definition 30 *(**Distance-from-$\mathbf{Z}$**) Let $\mathbf{X},\mathbf{Y}$ and $\mathbf{Z}$ be pairwise disjoint node sets in a* DAG, CPDAG, MAG *or* PAG $\mathcal{G}$*. Let p be a path between $\mathbf{X}$ and $\mathbf{Y}$ in $\mathcal{G}$ such that every collider C on p has a possibly directed path (possibly of length 0) to $\mathbf{Z}$. Define the* distance-from-$\mathbf{Z}$ *of C to be the length of a shortest possibly directed path (possibly of length 0) from C to $\mathbf{Z}$, and define the* distance-from-$\mathbf{Z}$ *of p to be the sum of the distances from $\mathbf{Z}$ of the colliders on p.*

If $\mathcal{G}$ is a MAG and p is a path from $\mathbf{X}$ to $\mathbf{Y}$ that is m-connecting given $\mathbf{Z}$ in $\mathcal{G}$, then Definition 30 reduces to the notion of *distance-from-*$\mathbf{Z}$ in Zhang (2006, p213).

Theorem 31 *(**Wright's rule** cf. Wright, 1921) Let $\mathbf{X} = \mathbf{A}\mathbf{X} + \epsilon$, where $\mathbf{A} \in \mathbb{R}^{k\times k}$, $\mathbf{X} = (X_1,\dots,X_k)^T$ and $\epsilon = (\epsilon_1,\dots,\epsilon_k)^T$ is a vector of mutually independent errors with means zero. Moreover, let $\mathrm{Var}(\mathbf{X}) = \mathbf{I}$. Let $\mathcal{D} = (\mathbf{X},\mathbf{E})$, be the corresponding* DAG *such that $X_i \to X_j$ in $\mathcal{D}$ if and only if $A_{ji} \neq 0$. A nonzero entry A_{ji} is called the edge coefficient of $X_i \to X_j$. For two distinct nodes $X_i, X_j \in \mathbf{X}$, let $p_1,\dots,p_r$ be all paths between X_i and X_j in $\mathcal{D}$ that do not contain a collider. Then $\mathrm{Cov}(X_i,X_j) = \sum_{s=1}^{r} \pi_s$, where π_s is the product of all edge coefficients along path p_s, $s \in \{1,\dots,r\}$.*

Theorem 32 *(cf. Theorem 3.2.4 Mardia et al., 1980, p63) Let $\mathbf{X} = (\mathbf{X_1}^T, \mathbf{X_2}^T)^T$ be a p-dimensional multivariate Gaussian random vector with mean vector $\mu = (\mu_{\mathbf{1}}{}^T, \mu_{\mathbf{2}}{}^T)^T$ and covariance matrix $\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma_{11}} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma_{21}} & \mathbf{\Sigma}_{22} \end{bmatrix}$, so that $\mathbf{X_1}$ is a q-dimensional multivariate Gaussian random vector with mean vector $\mu_{\mathbf{1}}$ and covariance matrix $\mathbf{\Sigma}_{11}$ and $\mathbf{X_2}$ is a $(p-q)$-dimensional multivariate Gaussian random vector with mean vector $\mu_{\mathbf{2}}$ and covariance matrix $\mathbf{\Sigma}_{22}$. Then $E[\mathbf{X_2} \mid \mathbf{X_1} = \mathbf{x_1}] = \mu_{\mathbf{2}} + \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}(\mathbf{x_1} - \mu_{\mathbf{1}})$.*

Definition 33 *(**Moralization**; Lauritzen and Spiegelhalter, 1988) Let $\mathcal{D}$ be a* DAG*. The moral graph $\mathcal{D}^m$ is formed by adding the edge $A - B$ to any structure of the form $A \to C \leftarrow B$ with $A \notin \mathrm{Adj}(B,\mathcal{D})$ (marrying unmarried parents) and subsequently making all edges in the resulting graph undirected.*

Definition 34 *(**Induced subgraph**) Let $\mathbf{X} \subseteq \mathbf{V}$ be a node set in a* DAG $\mathcal{D} = (\mathbf{V},\mathbf{E})$*. Then $\mathcal{D}_{\mathbf{X}} = (\mathbf{X},\mathbf{E}_{\mathbf{X}})$, where $\mathbf{E}_{\mathbf{X}}$ consists of all edges in $\mathbf{E}$ for which both endpoints are in $\mathbf{X}$, is the induced subgraph of $\mathcal{D}$ on $\mathbf{X}$.*

Theorem 35 *(**Reduction of d-separation to node cuts**; cf. Proposition 3 in Lauritzen et al., 1990, cf. Corollary 2 in Richardson, 2003) Let $\mathcal{D}$ be a* DAG *and let $\mathbf{X}, \mathbf{Y}$ and $\mathbf{Z}$ be pairwise disjoint node sets in $\mathcal{D}$. Then $\mathbf{Z}$ d-separates $\mathbf{X}$ and $\mathbf{Y}$ in $\mathcal{D}$ if and only if all paths between $\mathbf{X}$ and $\mathbf{Y}$ in $(\mathcal{D}_{\mathrm{An}(\mathbf{X}\cup\mathbf{Y}\cup\mathbf{Z},\mathcal{D})})^m$ contain at least one node in $\mathbf{Z}$.*

Lemma 36 *(**Basic property of** CPDAG**s and** PAG**s**; cf. Lemma 1 in Meek, 1995, cf. Lemma 3.3.1 in Zhang, 2006) Let X, Y and Z be distinct nodes in a* CPDAG *or* PAG *$\mathcal{G}$. If $X \bullet\!\!\rightarrow Y \circ\!\!-\!\!\bullet Z$, then there is an edge between X and Z with an arrowhead at Z. Furthermore, if the edge between X and Y is $X \rightarrow Y$, then the edge between X and Z is either $X \circ\!\!\rightarrow Z$ or $X \rightarrow Z$ (that is, not $X \leftrightarrow Z$).*

Lemma 37 *(cf. Lemma 1 in Richardson, 2003) Let $\mathbf{X}, \mathbf{Y}$ and $\mathbf{Z}$ be pairwise disjoint node sets in a* DAG *or* MAG *$\mathcal{G}$. If there is a path p from $X \in \mathbf{X}$ to $Y \in \mathbf{Y}$, on which no non-collider is in $\mathbf{Z}$ and every collider on p is in* $\mathrm{An}(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z}, \mathcal{G})$*, then there exists a path q from $X' \in \mathbf{X}$ to $Y' \in \mathbf{Y}$ that is m-connecting given $\mathbf{Z}$ in $\mathcal{G}$.*

Lemma 38 *(Lemma 0 in Zhang, 2006, p208) Let X and Y be distinct nodes in a* MAG *$\mathcal{M}$. If $p = \langle X, \dots, Z, V, Y\rangle$ is a discriminating path from X to Y for V in a* MAG *$\mathcal{M}$, and the corresponding path to $p(X, V)$ in the* PAG *$\mathcal{P}$ of $\mathcal{M}$ is (also) a collider path, then the corresponding path to p in $\mathcal{P}$ is also a discriminating path from X to Y for V.*

Lemma 39 *(cf. Lemma 1 in Zhang, 2006, p208) Let X and Y be distinct nodes and let $\mathbf{Z}$ be a node set that does not contain X and Y in a* MAG *$\mathcal{M}$ (*DAG *$\mathcal{D}$). Let p be a shortest path from X to Y that is m-connecting given $\mathbf{Z}$ in $\mathcal{M}$ ($\mathcal{D}$). Let $\mathcal{P}$ be the* PAG *of $\mathcal{M}$ (*CPDAG *of $\mathcal{D}$) and let p^* in $\mathcal{P}$ be the corresponding path to p in $\mathcal{M}$ ($\mathcal{D}$). Then p^* is a definite status path in $\mathcal{P}$.*

Lemma 40 *(cf. Lemma 2 in Zhang, 2006, p213) Let X and Y be distinct nodes and let $\mathbf{Z}$ be a node set that does not contain X and Y in a* MAG *$\mathcal{M}$ (*DAG *$\mathcal{D}$). Let p be a shortest path from X to Y that is m-connecting given $\mathbf{Z}$ in $\mathcal{M}$ ($\mathcal{D}$) such that no equally short m-connecting path has a shorter distance-from-$\mathbf{Z}$ (see Definition 30) than p does. Let $\mathcal{P}$ be the* PAG *of $\mathcal{M}$ (*CPDAG *of $\mathcal{D}$) and let p^* in $\mathcal{P}$ be the corresponding path to p in $\mathcal{M}$ ($\mathcal{D}$). Then p^* is a definite status path from X to Y that is m-connecting given $\mathbf{Z}$ in $\mathcal{P}$.*

Lemma 41 *(cf. Lemma B.1 in Zhang, 2008b) Let X and Y be distinct nodes in a* CPDAG *or* PAG *$\mathcal{G}$. If $p = \langle X, \dots, Y\rangle$ is a possibly directed path from X to Y in $\mathcal{G}$, then some subsequence of p forms an unshielded possibly directed path from X to Y in $\mathcal{G}$.*

Lemma 42 *(cf. Lemma B.2 in Zhang, 2008b, Lemma 7.2 in Maathuis and Colombo, 2015) Let X and Y be distinct nodes in a* CPDAG *or* PAG *$\mathcal{G}$. If $p = \langle X = V_0, \dots, V_k = Y\rangle$, $k \geq 2$ is an unshielded possibly directed path from X to Y in $\mathcal{G}$, and $V_{i-1} \bullet\!\!\rightarrow V_i$ for some $i \in \{1, \dots, k\}$, then $V_{j-1} \rightarrow V_j$ for all $j \in \{i+1, \dots, k\}$.*

Lemma 43 *(cf. Theorem 2 in Zhang, 2008b, Lemma 7.6 in Maathuis and Colombo, 2015) Let $\mathcal{G}$ be a* PAG *(*CPDAG*). Let $\mathcal{M}$ ($\mathcal{D}$) be the graph resulting from the following procedure applied to a $\mathcal{G}$:*

(1) replace all partially directed edges $\circ\!\!\rightarrow$ *in* $\mathcal{G}$ *with directed edges* $\rightarrow$*, and*

(2) orient the subgraph of $\mathcal{G}$ *consisting of all non-directed edges* $\circ\!\!-\!\!\circ$ *into a* DAG *with no unshielded colliders.*

Then $\mathcal{M}$ $(\mathcal{D})$ *is in* $[\mathcal{G}]$*. Moreover, if* X *is a node in* $\mathcal{G}$*, then one can always find an orientation of (2) that does not create any new edges into* X*.*

Lemma 44 *(Lemma 7.5 in Maathuis and Colombo, 2015) Let* X *and* Y *be two distinct nodes in a* DAG*,* CPDAG*,* MAG *or* PAG $\mathcal{G}$*. Then* $\mathcal{G}$ *cannot have both a possibly directed path from* X *to* Y *and an edge of the form* $Y \bullet\!\!\rightarrow X$*.*

Definition 45 *(*$DSEP(X,Y,\mathcal{G})$*; Spirtes et al., 2000, p136) Let* X *and* Y *be two distinct nodes in a* DAG *or* MAG $\mathcal{G}$*. We say that* $V \in DSEP(X,Y,\mathcal{G})$ *if* $V \neq X$ *and there is a collider path between* X *and* V *in* $\mathcal{G}$ *such that every node on this path is an ancestor of* X *or* Y *in* $\mathcal{G}$*.*

Definition 46 *(*$\mathcal{R}$ *and* $\mathcal{R}_{\underline{X}}$*; Maathuis and Colombo, 2015) Let* X *be a node in a* DAG*,* CPDAG*,* MAG *or* PAG $\mathcal{G}$*. Let* $\mathcal{R}$ *be a* DAG *or* MAG *represented by* $\mathcal{G}$*, in the following sense. If* $\mathcal{G}$ *is a* DAG *or* MAG*, let* $\mathcal{R} = \mathcal{G}$*. If* $\mathcal{G}$ *is a* CPDAG (PAG)*, let* $\mathcal{R}$ *be a* DAG (MAG) *in* $[\mathcal{G}]$ *as defined in Lemma 43, so that* $\mathcal{R}$ *has the same number of edges into* X *as* $\mathcal{G}$*. Let* $\mathcal{R}_{\underline{X}}$ *be the graph obtained from* $\mathcal{R}$ *by removing all directed edges out of* X *that are visible in* $\mathcal{G}$*.*

Theorem 47 *(Theorem 4.1 in Maathuis and Colombo, 2015) Let* X *and* Y *be distinct nodes in a* DAG*,* CPDAG*,* MAG *or* PAG $\mathcal{G}$*. Let* $\mathcal{R}$ *and* $\mathcal{R}_{\underline{X}}$ *be defined as in Definition 46. Then there exists a generalized back-door set relative to* (X,Y) *in* $\mathcal{G}$ *if and only if* $Y \notin \mathrm{Adj}(X,\mathcal{R}_{\underline{X}})$ *and* $DSEP(X,Y,\mathcal{R}_{\underline{X}}) \cap \mathrm{PossDe}(X,\mathcal{G}) = \emptyset$*. Moreover, if such a set exists, then* $DSEP(X,Y,\mathcal{R}_{\underline{X}})$ *is a generalized back-door set relative to* (X,Y) *in* $\mathcal{G}$*.*

A.1 Rules of the Do-calculus (Pearl, 2009, Chapter 3.4)

Let $\mathbf{X}',\mathbf{Y}',\mathbf{Z}',\mathbf{W}'$ be pairwise disjoint (possibly empty) sets of nodes in a causal DAG $\mathcal{D}$. Let $\mathcal{D}_{\overline{\mathbf{X}'}}$ denote the graph obtained by deleting all edges into $\mathbf{X}'$ from $\mathcal{D}$. Similarly, let $\mathcal{D}_{\underline{\mathbf{X}'}}$ denote the graph obtained by deleting all edges out of $\mathbf{X}'$ in $\mathcal{D}$ and let $\mathcal{D}_{\overline{\mathbf{X}'}\underline{\mathbf{Z}'}}$ denote the graph obtained by deleting all edges into $\mathbf{X}'$ and all edges out of $\mathbf{Z}'$ in $\mathcal{D}$. Then the following three rules are valid for every density function consistent with $\mathcal{D}$.

Rule 1 (Insertion/deletion of observations) If $\mathbf{Y}' \perp_d \mathbf{Z}' \mid \mathbf{X}' \cup \mathbf{W}'$ in $\mathcal{D}_{\overline{\mathbf{X}'}}$, then

$$f(\mathbf{y}' \mid \mathrm{do}(\mathbf{x}'), \mathbf{w}') = f(\mathbf{y}' \mid \mathrm{do}(\mathbf{x}'), \mathbf{z}', \mathbf{w}'). \tag{4}$$

Rule 2 (Action/observation exchange) If $\mathbf{Y}' \perp_d \mathbf{Z}' \mid \mathbf{X}' \cup \mathbf{W}'$ in $\mathcal{D}_{\overline{\mathbf{X}'}\underline{\mathbf{Z}'}}$, then

$$f(\mathbf{y}' \mid \mathrm{do}(\mathbf{x}'), \mathrm{do}(\mathbf{z}'), \mathbf{w}') = f(\mathbf{y}' \mid \mathrm{do}(\mathbf{x}'), \mathbf{z}', \mathbf{w}'). \tag{5}$$

Rule 3 (Insertion/deletion of actions)If $\mathbf{Y}' \perp_d \mathbf{Z}' \mid \mathbf{X}' \cup \mathbf{W}'$ in $\mathcal{D}_{\overline{\mathbf{X}' \cup \mathbf{Z}'(\mathbf{W}')}}$, then

$$f(\mathbf{y}' \mid \mathrm{do}(\mathbf{x}'), \mathbf{w}') = f(\mathbf{y}' \mid \mathrm{do}(\mathbf{x}'), \mathrm{do}(\mathbf{z}'), \mathbf{w}'), \tag{6}$$

where $\mathbf{Z}'(\mathbf{W}')$ denotes the set of $\mathbf{Z}'$-nodes that are not ancestors of any $\mathbf{W}'$ node in $\mathcal{D}_{\overline{\mathbf{X}'}}$. If $\mathbf{W}' = \emptyset$, then $\mathbf{Z}'(\mathbf{W}') = \mathbf{Z}'$.

$$\begin{tikzcd}[column sep=small]
 & \text{Lemma 48} \arrow[r] & \text{Lemma 49} \arrow[r] & \text{Lemma 8} \arrow[dr] & \\
 & & & \text{Lemma 9} \arrow[r] & \textbf{Theorem 5} \\
\text{Lemma 50} \arrow[r] & \text{Lemma 51} \arrow[r] & \text{Lemma 52} \arrow[r] & \text{Lemma 10} \arrow[ur] &
\end{tikzcd}$$

Figure 9: Proof structure of Theorem 5.

A.2 FCI Orientation Rules (Spirtes et al., 2000, p183)

Let A, B, C and D be distinct nodes in a PAG $\mathcal{P}$. Below, we give the first 4 orientation rules of the FCI algorithm defined in Spirtes et al. (2000).
R1 If $A \bullet\!\!\rightarrow B \circ\!\!-\!\!\bullet C$, and A and C are not adjacent, then orient the triple $\langle A, B, C\rangle$ as $A \bullet\!\!\rightarrow B \rightarrow C$.
R2 If $A \rightarrow B \bullet\!\!\rightarrow C$ or $A \bullet\!\!\rightarrow B \rightarrow C$ and $A \bullet\!\!-\!\!\circ C$, then orient $A \bullet\!\!-\!\!\circ C$ as $A \bullet\!\!\rightarrow C$.
R3 If $A \bullet\!\!\rightarrow B \leftarrow\!\!\bullet C$, $A \bullet\!\!-\!\!\circ D \circ\!\!-\!\!\bullet C$, A and C are not adjacent, and $D \bullet\!\!-\!\!\circ B$, then orient $D \bullet\!\!-\!\!\circ B$ as $D \bullet\!\!\rightarrow B$.
R4 If $\langle D, \dots, A, B, C\rangle$ is a discriminating path from D to C for B and $B \circ\!\!-\!\!\bullet C$, then orient $B \circ\!\!-\!\!\bullet C$ as $B \rightarrow C$ if B is in the separation set of D and C; otherwise orient the triple $\langle A, B, C\rangle$ as $A \leftrightarrow B \leftrightarrow C$.
These four rules were proven to be sound in Spirtes et al. (2000), meaning that edge marks oriented by these rules correspond to invariant edge marks in the maximally informative PAG for the true causal MAG. Six additional orientation rules for the FCI algorithm were defined in Zhang (2008b). The augmented FCI algorithm, including all ten orientation rules was proven to be sound and complete in Zhang (2008b).

Appendix B. Proofs for Section 3

Figure 9 shows how all lemmas in this section fit together to prove Theorem 5.

Lemma 48 *Let X be a node in a* PAG $\mathcal{P}$. *Let $\mathcal{M}$ be a* MAG $\mathcal{M}$ *in* $[\mathcal{P}]$ *that satisfies Lemma 43. Then any edge that is either $X \circ\!\!-\!\!\circ Y$, $X \circ\!\!\rightarrow Y$ or invisible $X \rightarrow Y$ in $\mathcal{P}$ is invisible $X \rightarrow Y$ in $\mathcal{M}$.*

Proof of Lemma 48. Let $\mathcal{M}$ be a MAG in $[\mathcal{P}]$ that satisfies Lemma 43. Then the edge $X \circ\!\!-\!\!\circ Y$, $X \circ\!\!\rightarrow Y$, or invisible $X \rightarrow Y$ in $\mathcal{P}$ corresponds to edge $X \rightarrow Y$ in $\mathcal{M}$. It is left to prove that $X \rightarrow Y$ is invisible in $\mathcal{M}$ in all these cases.

Suppose for a contradiction that $X \rightarrow Y$ is visible in $\mathcal{M}$. Then there is a node $D \notin \mathrm{Adj}(Y, \mathcal{M})$ such that (1) $D \bullet\!\!\rightarrow X$ is in $\mathcal{M}$, or (2) there is a collider path $\langle D, D_1, \dots, D_k, X\rangle$, $k \geq 1$, into X such that every $D_i, 1 \leq i \leq k$ is a parent of X in $\mathcal{M}$. We consider these cases separately and show that we arrive at a contradiction.

(1) Since $D \bullet\!\!\rightarrow X$ is in $\mathcal{M}$, the choice of $\mathcal{M}$ implies that $D \bullet\!\!\rightarrow X$ is in $\mathcal{P}$. Since $D \notin \mathrm{Adj}(Y, \mathcal{P})$, $X \rightarrow Y$ must be in $\mathcal{P}$, since otherwise rule **R1** of the FCI algorithm from Spirtes et al. (2000) (see Appendix A) would have been applied in $\mathcal{P}$. But then $X \rightarrow Y$ is visible in $\mathcal{P}$, which is a contradiction.

(2) Path $p = \langle D, D_1, \dots D_k, X, Y\rangle$ is a discriminating path from D to Y for X, that is into X in $\mathcal{M}$. Let p^* be the path in $\mathcal{P}$ corresponding to p in $\mathcal{M}$. Then since $p(D_1, X)$ contains only bi-directed edges in $\mathcal{M}$, the choice of $\mathcal{M}$ implies that $p^*(D_1, X)$ also contains only bi-directed edges in $\mathcal{P}$. Since $D \bullet\!\!\rightarrow D_1$ is in $\mathcal{M}$, $D \circ\!\!-\!\!\circ D_1$ or $D \bullet\!\!\rightarrow D_1$ is in $\mathcal{P}$.

Suppose first that $D \bullet\!\!\rightarrow D_1$ is in $\mathcal{P}$, then by Lemma 0 from Zhang (2006) (see Lemma 38), p^* is a discriminating path from D to Y for X, that is into X in $\mathcal{P}$. Hence, $X \rightarrow Y$ is in $\mathcal{P}$, since otherwise rule **R4** in Appendix A would have been applied. But then $X \rightarrow Y$ is a visible edge in $\mathcal{P}$, which is a contradiction.

Next, suppose that $D \circ\!\!-\!\!\circ D_1$ is in $\mathcal{P}$. Since $D_1 \leftrightarrow D_2$ is in $\mathcal{P}$, by Lemma 36 $D \leftarrow\!\!\bullet D_2$ is in $\mathcal{P}$. This edge cannot be $D \leftarrow\!\!\circ D_2$ or $D \leftarrow D_2$, otherwise $D \leftarrow\!\!\circ D_1$ would be in $\mathcal{P}$ (Lemma 36, or **R2** of the FCI orientation rules in Appendix A), contrary to our assumption. Hence, the edge $D \leftrightarrow D_2$ is in $\mathcal{P}$. Then $D \leftrightarrow D_2$ is also in $\mathcal{M}$ and $p_1 = \langle D, D_2, \dots D_k, X, Y\rangle$ is a discriminating path from D to Y for X, that is into X in $\mathcal{M}$. Additionally, since $D \leftrightarrow D_2 \leftrightarrow \dots \leftrightarrow D_k \leftrightarrow X$ is in $\mathcal{P}$, by Lemma 38 the path ${p_1}^* = \langle D, D_2, \dots D_k, X, Y\rangle$ is a discriminating path from D to Y for X, that is into X. But then $X \rightarrow Y$ is a visible edge in $\mathcal{P}$, which is a contradiction. ∎

Lemma 49 *Let X and Y be distinct nodes in a* PAG $\mathcal{P}$ *such that there is a possibly directed path p^* from X to Y in $\mathcal{P}$ that does not start with a visible edge out of X. Then there is a* MAG $\mathcal{M}$ *in* $[\mathcal{P}]$ *such that the path p in $\mathcal{M}$, consisting of the same sequence of nodes as p^* in $\mathcal{P}$, contains a subsequence p' that is a directed path from X to Y starting with an invisible edge in $\mathcal{M}$. In other words, $\mathcal{M}$ violates the amenability condition relative to* (X, Y).

Proof of Lemma 49. If $\mathcal{P}$ violates the amenability condition relative to (X, Y), then there is a possibly directed path p^* from X to Y in $\mathcal{P}$ that does not start with a visible edge out of X. Let $\mathcal{M}$ be a MAG in $[\mathcal{P}]$ that satisfies Lemma 43. We will show that $\mathcal{M}$ violates the amenability condition relative to (X, Y).

Let p'^* be a shortest subsequence of p^* such that p'^* is also a possibly directed path from X to Y in $\mathcal{P}$ that does not start with a visible edge out of X. We write $p'^* = \langle X = V_0, V_1, \dots, V_r = Y\rangle, r \geq 1$. Let p' in $\mathcal{M}$ be the path corresponding to p'^* in $\mathcal{P}$. Then by Lemma 48, $X \rightarrow V_1$ is an invisible edge in $\mathcal{M}$. It is left to show that p' is a directed path from X to Y in $\mathcal{M}$.

Suppose first that p'^* is a definite status path in $\mathcal{P}$. Then all non-endpoint nodes on p'^* are definite non-colliders. Hence, $X \rightarrow V_1$ in $\mathcal{M}$ implies that all the remaining edges on p' are oriented towards Y.

Else, p'^* is not of definite status in $\mathcal{P}$ and $r \geq 2$. Since $X \rightarrow V_1$ is in $\mathcal{M}$, it is sufficient to show that $p'(V_1, Y)$ is a directed path from V_1 to Y. Note that by the choice of p'^*, $p'^*(V_1, Y)$ is a shortest possibly directed path from V_1 to Y in $\mathcal{P}$. Hence, it is unshielded (Lemma B.1 from Zhang (2008b), see Lemma 41 in Appendix A). If $V_1 \circ\!\!\rightarrow V_2$ or $V_1 \rightarrow V_2$ is in $\mathcal{P}$, then by the choice of $\mathcal{M}$ (Lemma 43), $V_1 \rightarrow V_2$ is in $\mathcal{M}$. Additionally, since $p'(V_1, Y)$ is a possibly directed definite status path, $V_1 \rightarrow V_2$ in $\mathcal{M}$ implies that all the remaining edges on $p'(V_1, Y)$ are oriented towards Y.

Otherwise, $V_1 \circ\!\!-\!\!\circ V_2$ is in $\mathcal{P}$. Path p'^* is not of definite status, whereas $p'^*(V_1, Y)$ is of definite status, as it is unshielded. Thus, node V_1 is not of definite status on p'^* and $X \in \mathrm{Adj}(V_2, \mathcal{P})$. The edge $X \leftarrow\!\!\bullet V_2$ is not in $\mathcal{P}$ since $p'^*(X, V_2)$ is a possibly directed path from X to V_2 in $\mathcal{P}$ (Lemma 7.5 from Maathuis and Colombo (2015), see Lemma 44 in

Appendix A). Since p'^* is a shortest possibly directed path from X to Y in $\mathcal{P}$ that does not start with a visible edge out of X, and $X \leftarrow\!\!\bullet V_2$ is not in $\mathcal{P}$, it follows that $X \rightarrow V_2$ is visible in $\mathcal{P}$. Since $X \rightarrow V_2$ is visible, there is a node $D \notin \text{Adj}(V_2, \mathcal{P})$ such that (1) $D \bullet\!\!\rightarrow X$ is in $\mathcal{P}$, or (2) there is a collider path $\langle D, D_1, \dots, D_k, X\rangle, k \geq 1$, that is into X in $\mathcal{P}$ such that every $D_i, 1 \leq i \leq k$ is a parent of V_2 in $\mathcal{P}$. We consider these cases separately and show that we arrive at a contradiction, implying that $p'^*(V_1, Y)$ cannot start with $V_1 \circ\!\!-\!\!\circ V_2$.

(1) A node $D \notin \text{Adj}(V_2, \mathcal{P})$ such that $D \bullet\!\!\rightarrow X$ is in $\mathcal{P}$. Since $D \bullet\!\!\rightarrow X$ and $X \circ\!\!-\!\!\circ V_1$, $X \circ\!\!\rightarrow V_1$ or $X \rightarrow V_1$ is invisible in $\mathcal{P}$, by Lemma 36 and the definition of visibility, an edge between D and V_1 is in $\mathcal{P}$. This edge is of type $D \bullet\!\!\rightarrow V_1$, since otherwise both a possibly directed path $\langle X, V_1, D\rangle$ and $D \bullet\!\!\rightarrow X$ are in $\mathcal{P}$ (contrary to Lemma 44). Then $D \bullet\!\!\rightarrow V_1 \circ\!\!-\!\!\circ V_2$ is in $\mathcal{P}$ and Lemma 36 implies that $D \in \text{Adj}(V_2, \mathcal{P})$, a contradiction.

(2) There is a node $D \notin \text{Adj}(V_2, \mathcal{P})$ and a collider path $\langle D, D_1, \dots, D_k, X\rangle, k \geq 1$, into X such that every $D_i, 1 \leq i \leq k$ is a parent of V_2 in $\mathcal{P}$. Paths $D_i \rightarrow V_2 \circ\!\!-\!\!\circ V_1$, $i = 1, \dots, k$ are in $\mathcal{P}$, so by Lemma 36 either $D_i \circ\!\!\rightarrow V_1$ or $D_i \rightarrow V_1$ is in $\mathcal{P}$, for $i = 1, \dots, k$. If $D_1 \circ\!\!\rightarrow V_1$ is in $\mathcal{P}$, then $D \bullet\!\!\rightarrow D_1 \circ\!\!\rightarrow V_1$ implies $D \bullet\!\!\rightarrow V_1$ is also in $\mathcal{P}$ (Lemma 36). But, then $D \bullet\!\!\rightarrow V_1 \circ\!\!-\!\!\circ V_2$ implies $D \in \text{Adj}(V_2, \mathcal{P})$ (Lemma 36), a contradiction. Hence, $D_1 \rightarrow V_1$ is in $\mathcal{P}$.

This allows us to deduce that $D \notin \text{Adj}(V_1, \mathcal{P})$, otherwise $D \bullet\!\!\rightarrow D_1 \rightarrow V_1$ would imply $D \bullet\!\!\rightarrow V_1$ (Lemma 44) and we arrive at the contradiction $D \in \text{Adj}(V_2, \mathcal{P})$, as above. Hence, $\langle D, D_1, D_2, V_1\rangle$ is a discriminating path from D to V_1 for D_2, implying that D_2 is of definite status on this path (**R4** of the FCI orientation rules in Appendix A). Thus, $D_2 \circ\!\!\rightarrow V_1$ is not possible, and since $D_2 \leftrightarrow V_1$ is already ruled out by Lemma 36, $D_2 \rightarrow V_1$ is in $\mathcal{P}$. By the same reasoning, $D_i \rightarrow V_1$ is in $\mathcal{P}$, for $i = 3, \dots k$. It then follows that $\langle D, D_1, \dots, D_k, X, V_1\rangle$ is a discriminating path from D to V_1 for X in $\mathcal{P}$, so $X \rightarrow V_1$ is in $\mathcal{P}$ (**R4** in Appendix A and the fact that $X \circ\!\!-\!\!\circ V_1$, $X \circ\!\!\rightarrow V_1$ or invisible $X \rightarrow V_1$ is in $\mathcal{P}$) and $X \rightarrow V_1$ is visible. This contradicts the fact $X \circ\!\!-\!\!\circ V_1$, $X \circ\!\!\rightarrow V_1$ or invisible $X \rightarrow V_1$ is in $\mathcal{P}$. ∎

Proof of Lemma 8. First suppose that $\mathcal{G}$ is amenable relative to $(\mathbf{X}, \mathbf{Y})$, meaning that every proper possibly directed path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$ starts with a visible edge out of $\mathbf{X}$. Any visible edge in $\mathcal{G}$ is visible in all DAGs (MAGs) in $[\mathcal{G}]$, and any proper directed path in a DAG (MAG) in $[\mathcal{G}]$ corresponds to a proper possibly directed path in $\mathcal{G}$. Hence, any proper directed path from $\mathbf{X}$ to $\mathbf{Y}$ in any DAG (MAG) in $[\mathcal{G}]$ starts with a visible edge out of $\mathbf{X}$.

Next, suppose that $\mathcal{G}$ violates the amenability condition relative to $(\mathbf{X}, \mathbf{Y})$. We will show that this implies that there is no adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$. Since every directed edge in a CPDAG is visible and since the same does not hold true in a PAG, we give separate proofs for CPDAGs and PAGs.

Suppose that $\mathcal{G}$ is a PAG. Since $\mathcal{G}$ violates the amenability condition relative to $(\mathbf{X}, \mathbf{Y})$, there exists a proper possibly directed path p^* from some $X \in \mathbf{X}$ to some $Y \in \mathbf{Y}$ in $\mathcal{G}$ that does not start with visible edge out of X in $\mathcal{G}$. Then by Lemma 49 there is a MAG $\mathcal{M}$ in $[\mathcal{G}]$ such that the path p in $\mathcal{M}$, consisting of the same sequence of nodes as p^* in $\mathcal{P}$, contains a subsequence p' that is a proper directed path from X to Y starting with an invisible edge in $\mathcal{M}$. Hence, $\mathcal{M}$ violates the amenability condition relative to $(\mathbf{X}, \mathbf{Y})$. Since the generalized adjustment criterion is complete for MAGs (Theorem 5.8 in van der Zander et al., 2014) this means that there is no adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{M}$. Hence, there is no adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$.

Next, suppose $\mathcal{G}$ is a CPDAG. We now show how to find DAGs $\mathcal{D}_1$ and $\mathcal{D}_2$ in $[\mathcal{G}]$, such that a proper causal path q_1' from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{D}_1$ corresponds to a proper non-causal path q_2' from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{D}_2$ that does not contain colliders. Since $\mathcal{G}$ is not amenable relative to $(\mathbf{X},\mathbf{Y})$, there is a proper possibly directed path q^* from a node $X \in \mathbf{X}$ to a node $Y \in \mathbf{Y}$ that starts with a non-directed edge ($\circ\!\!-\!\!\circ$).

Let $q'^* = \langle X = V_0, V_1, \dots, V_k = Y\rangle, k \geq 1$, be a shortest subsequence of q^* such that q'^* is also a proper possibly directed path from X to Y starting with a non-directed edge in $\mathcal{G}$. Since we chose q'^* using the additional constraint that it must start with a non-directed edge in $\mathcal{G}$, we cannot use Lemma 41 to guarantee that q'^* is of definite status. Hence, we first show that q'^* is a definite status path, by contradiction. Thus, suppose that q'^* is not a definite status path. Then $k \geq 2$. Since the subpath $q'^*(V_1, Y)$ is a definite status path (otherwise, by Lemma 41 we can choose a shorter path), this means that V_1 is not of definite status on q'^*. This implies $X \in \mathrm{Adj}(V_2, \mathcal{G})$. Moreover, we must have $X \rightarrow V_2$, since $X \circ\!\!-\!\!\circ V_2$ contradicts the choice of q'^*, and $X \leftarrow V_2$ together with the possibly directed path $q'^*(X, V_2)$ contradicts Lemma 44. Then $X \rightarrow V_2$ implies $V_1 \rightarrow V_2$, otherwise $X \rightarrow V_2$ and a possibly directed path $-q'^*(V_2, X)$ are in $\mathcal{G}$, which contradicts Lemma 44. But then V_1 is a definite non-collider on q'^*, which contradicts that V_1 is not of definite status.

Hence, q'^* is a proper definite status possibly directed path from X to Y. By Lemma 43, there is a DAG $\mathcal{D}_1$ in $[\mathcal{G}]$ such that there are no additional arrowheads into X, as well as a DAG $\mathcal{D}_2$ in $[\mathcal{G}]$ such that there are no additional arrowheads into V_1. Let q_1' in $\mathcal{D}_1$ (q_2' in $\mathcal{D}_2$) be the path corresponding to q'^* in $\mathcal{G}$. Then q_1' is of the form $= X \rightarrow V_1 \rightarrow \dots \rightarrow Y$ and q_2' is of the form $X \leftarrow V_1 \rightarrow \dots \rightarrow Y$. Since q_1' is a proper causal path from X to Y and q_2' is a proper non-causal path from X to Y, $f(\mathbf{y} \mid do(\mathbf{x}))$ generally differs in $\mathcal{D}_1$ and $\mathcal{D}_2$. Since $\mathcal{D}_1$ and $\mathcal{D}_2$ are both represented by $\mathcal{C}$, this implies that $f(\mathbf{y} \mid do(\mathbf{x}))$ is not identifiable in $\mathcal{C}$. ∎

Proof of Lemma 9. We first prove (i)⇒(ii). Suppose that $\mathbf{Z} \cap \mathrm{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G}) = \emptyset$. Since $\mathrm{Forb}(\mathbf{X},\mathbf{Y},\mathcal{D}) \subseteq \mathrm{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$ ($\mathrm{Forb}(\mathbf{X},\mathbf{Y},\mathcal{M}) \subseteq \mathrm{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$) for any DAG $\mathcal{D}$ (MAG $\mathcal{M}$) in $[\mathcal{G}]$, it follows directly that $\mathbf{Z}$ satisfies the forbidden set condition relative to $(\mathbf{X},\mathbf{Y})$ in all DAGs (MAGs) in $[\mathcal{G}]$.

Next, we prove $\neg$ (i)⇒ $\neg$ (ii). Suppose that $\mathcal{G}$ is amenable relative to $(\mathbf{X},\mathbf{Y})$, but there is a node $V \in (\mathbf{Z} \cap \mathrm{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G}))$. Then $V \in \mathrm{PossDe}(W, \mathcal{G})$ for some $W = V_i, 1 \leq i \leq k$ on a proper possibly directed path $p = \langle X = V_0, V_1, \dots, V_k = Y\rangle, k \geq 1$. Then $q = p(X, W)$ is proper and $r = p(W, Y)$, where r is allowed to be of zero length (if $W = Y$), does not contain a node in $\mathbf{X}$. Moreover, there is a possibly directed path s from W to V, where this path is allowed to be of zero length. We will show that this implies that there is a DAG $\mathcal{D}$ (MAG $\mathcal{M}$) in $[\mathcal{G}]$ such that $\mathbf{Z} \cap \mathrm{Forb}(\mathbf{X},\mathbf{Y},\mathcal{D}) \neq \emptyset$ ($\mathbf{Z} \cap \mathrm{Forb}(\mathbf{X},\mathbf{Y},\mathcal{M}) \neq \emptyset$).

By Lemma 41, there are subsequences q', r' and s' of q, r and s that are unshielded possibly directed paths (again r' and s' are allowed to be paths of zero length). Moreover, q' is proper and must start with a directed (visible) edge, since otherwise the concatenated path $q' \oplus r'$, which is a proper possibly directed path from X to Y, would violate the amenability condition. Lemma 42 then implies that q' is a directed path from X to W in $\mathcal{G}$.

By Lemma 43, there is at least one DAG $\mathcal{D}$ (MAG $\mathcal{M}$) in $[\mathcal{G}]$ that has no additional arrowheads into W. In this graph $\mathcal{D}$ ($\mathcal{M}$), the first edge on the path corresponding to r' is oriented out of W and since r' is an unshielded possibly directed path in $\mathcal{P}$, by Lemma 42 the path in $\mathcal{D}$ ($\mathcal{M}$) corresponding to r' is a directed path from W to Y. By the same reasoning, the path corresponding to s' in $\mathcal{D}$ ($\mathcal{M}$) is a directed path from W to V. Hence, $V \in \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D})$ ($V \in \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{M})$), so that $\mathbf{Z}$ does not satisfy the forbidden set condition relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ ($\mathcal{M}$). ■

We now start the path of proving Lemma 10. The most involved part is proving the implication $\neg$ (ii)$\Rightarrow \neg$ (i), that is, if there is a proper non-causal path p from $\mathbf{X}$ to $\mathbf{Y}$ that is m-connecting given $\mathbf{Z}$ in a DAG $\mathcal{D}$ (MAG $\mathcal{M}$) in $[\mathcal{G}]$, then there must be a proper non-causal definite status path p^* from $\mathbf{X}$ to $\mathbf{Y}$ that is m-connecting given $\mathbf{Z}$ in $\mathcal{G}$ as well. We proceed in three steps. First, we show that proper non-causal paths from $\mathbf{X}$ to $\mathbf{Y}$ that are m-connecting given $\mathbf{Z}$ in $\mathcal{D}$ ($\mathcal{M}$) correspond to proper non-causal paths in $\mathcal{G}$ (Lemma 50). Second, we show that a certain shortest proper non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ that is m-connecting given $\mathbf{Z}$ in $\mathcal{D}$ ($\mathcal{M}$) corresponds to a proper definite status non-causal path p^* from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$ (Lemma 51). Lastly, we show that p^* is also m-connecting given $\mathbf{Z}$ in $\mathcal{G}$ (Lemma 52).

Lemma 50 *Let $\mathbf{X}, \mathbf{Y}$ and $\mathbf{Z}$ be pairwise disjoint node sets in a* PAG (CPDAG) *$\mathcal{P}$. Let $\mathcal{P}$ be amenable relative to $(\mathbf{X}, \mathbf{Y})$ and let $\mathbf{Z}$ satisfy the forbidden set condition relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{P}$. Let $\mathcal{M}$ be a* MAG (DAG) *in $[\mathcal{P}]$ and let $p = \langle X = V_0, V_1, \dots, V_n = Y\rangle, n \geq 2$, be a proper non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ that is m-connecting given $\mathbf{Z}$ in $\mathcal{M}$. Let p^* in $\mathcal{P}$ denote the path corresponding to p in $\mathcal{M}$. Then:*

(i) Let $i, j \in \mathbb{N}, 0 < i < j \leq n$ such that there is an edge $\langle V_i, V_j\rangle$ in $\mathcal{P}$. The path $p^*(X, V_i) \oplus \langle V_i, V_j\rangle \oplus p^*(V_j, Y)$ ($p^*(V_j, Y)$ is possibly of zero length) is a proper non-causal path in $\mathcal{P}$. For $j = i + 1$, this implies that p^* is a proper non-causal path.

(ii) If V_1 is not of definite status on p^*, then $\langle X, V_2\rangle \oplus p^*(V_2, Y)$, ($p^*(V_2, Y)$ is possibly of zero length) exists and is a proper non-causal path in $\mathcal{P}$.

(iii) If $n \geq 3$ and $V_k, 2 \leq k < n$ is not of definite status on p^*, and every non-endpoint node on $p^*(X, V_k)$ is a collider on p^* and a parent of V_{k+1} in $\mathcal{M}$, then $\langle X, V_{k+1}\rangle \oplus p^*(V_{k+1}, Y)$, ($p^*(V_{k+1}, Y)$ is possibly of zero length) exists and is a proper non-causal path in $\mathcal{P}$.

.

Proof of Lemma 50. All paths considered are proper as they are subsequences of p^*, which consists of the same sequence of nodes as the proper path p.

(i) Suppose for a contradiction that $q^* = p^*(X, V_i) \oplus \langle V_i, V_j\rangle \oplus p^*(V_j, Y)$ is possibly directed in $\mathcal{P}$. All nodes on q^* except X are in $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{P})$. Since $\mathcal{P}$ is amenable relative to $(\mathbf{X}, \mathbf{Y})$, q^* starts with a visible edge $X \to V_1$ in $\mathcal{P}$. Edge $X \to V_1$ is then also in $\mathcal{M}$ and since p is a non-causal path in $\mathcal{M}$, there is at least one collider on p. Let $V_r, r \geq 1$, be the collider closest to X on p. Then $V_r \in \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{P})$. Since p is m-connecting given $\mathbf{Z}$, a descendant of V_r is in $\mathbf{Z}$. This contradicts $\mathbf{Z} \cap \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{P}) = \emptyset$.

(ii) Since V_1 is not of definite status on p^* there is an edge between X and V_2 in $\mathcal{P}$, so path $q^* = \langle X, V_2\rangle \oplus p^*(V_2, Y)$ exists in $\mathcal{P}$. Suppose for a contradiction that q^* is a possibly directed path from X to Y in $\mathcal{P}$. Then $X \to V_2$ is in $\mathcal{P}$, since $\mathcal{P}$ is amenable relative to

$(\mathbf{X}, \mathbf{Y})$ and every node on q^* except X is in $\mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{P})$. From (i) above we know that p^* is non-causal, so since q^* is possibly directed, there is an arrowhead towards X on $p^*(X, V_2)$.

First, suppose $X \bullet\!\!-\!\!\bullet V_1 \leftarrow\!\!\bullet V_2$. Then $X \rightarrow V_2 \bullet\!\!\rightarrow V_1$ implies $X \bullet\!\!\rightarrow V_1$ is in $\mathcal{P}$, since $\mathcal{P}$ is ancestral. This contradicts that V_1 is not of definite status on p^*.

Next, suppose $X \leftarrow\!\!\bullet V_1 \bullet\!\!-\!\!\bullet V_2$ is in $\mathcal{P}$. If $\mathcal{P}$ is a CPDAG then V_1 is a definite non-collider on p^*, which contradicts that V_1 is not of definite status. If $\mathcal{P}$ is a PAG, then since $X \rightarrow V_2$ is a visible edge in $\mathcal{P}$ there is a node $D \notin \mathrm{Adj}(V_2, \mathcal{P})$ such that $D \bullet\!\!\rightarrow X$ in $\mathcal{P}$ (there is a collider path $D \bullet\!\!\rightarrow D_1 \leftrightarrow \dots \leftrightarrow D_s \leftrightarrow X, s \geq 1$, where every node $D_1, \dots, D_s$ is a parent of V_2 in $\mathcal{P}$). The path $\langle D, X, V_1, V_2 \rangle$ ($\langle D, D_1, \dots, D_s, X, V_1, V_2 \rangle$) is a discriminating path from D to V_2 for V_1 in $\mathcal{P}$ so V_1 is of definite status on p^*, contrary to the assumption.

(iii) If $\mathcal{P}$ is a CPDAG and $p(X, V_k), k \geq 2$ is a collider path, then it must be that $k = 2$ and V_k is of definite status on p. Hence, let $\mathcal{P}$ be a PAG. There is an edge between X and V_{k+1} in $\mathcal{P}$, otherwise by Lemma 38 the subpath $p^*(X, V_{k+1})$ is a discriminating path for V_k in $\mathcal{P}$, so V_k would be of definite status on p^*. Then path $q^* = \langle X, V_{k+1} \rangle \oplus p^*(V_{k+1}, Y)$ exists in $\mathcal{P}$. Suppose for a contradiction that q^* is a possibly directed path from X to Y in $\mathcal{P}$. Because $\mathcal{P}$ is amenable relative to $(\mathbf{X}, \mathbf{Y})$ the edge $X \rightarrow V_{k+1}$ is visible in $\mathcal{P}$. Also, since $V_1 \rightarrow V_{k+1}$ is in $\mathcal{M}$, edge $\langle V_1, V_{k+1} \rangle$ is possibly directed towards V_{k+1} in $\mathcal{P}$.

Consider the edge $X \bullet\!\!\rightarrow V_1$ in $\mathcal{P}$. If $X \bullet\!\!\rightarrow V_1$ is not into X in $\mathcal{P}$ then $p^*(X, V_1) \oplus \langle V_1, V_{k+1} \rangle \oplus p^*(V_{k+1}, Y)$ is a proper possibly directed path from X to Y in $\mathcal{P}$ so $V_1 \in \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{P})$. By assumption V_1 is a collider on p^* in $\mathcal{P}$ and hence also on p in $\mathcal{M}$. Since p is m-connecting given $\mathbf{Z}$, there is a node $Z \in \mathbf{Z}$ such that $Z \in \mathrm{De}(V_1, \mathcal{M})$. Since $\mathrm{De}(V_1, \mathcal{M}) \subseteq \mathrm{PossDe}(V_1, \mathcal{P}) \subseteq \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{P})$, node $Z \in \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{P})$. This contradicts the forbidden set condition.

So $X \leftrightarrow V_1$ must be in $\mathcal{P}$. Since $X \rightarrow V_{k+1}$ is a visible edge in $\mathcal{P}$ there is a node $D \notin \mathrm{Adj}(V_{k+1}, \mathcal{P})$ such that the edge $D \bullet\!\!\rightarrow X$ is in $\mathcal{P}$ (there is a collider path $D \bullet\!\!\rightarrow D_1 \leftrightarrow \dots \leftrightarrow D_s \leftrightarrow X, s \geq 1$, and every node $D_1, \dots, D_s$ is a parent of V_{k+1} in $\mathcal{P}$). By Lemma 38, path $\langle D, X, V_1, \dots, V_k, V_{k+1} \rangle$ ($\langle D, D_1, \dots, D_s, X, V_1, \dots, V_k, V_{k+1} \rangle$) is then a discriminating path from D to V_{k+1} for V_k in $\mathcal{P}$. Hence, V_k is of definite status on p^*, contrary to the original assumption. ∎

Lemma 51 *Let $\mathbf{X}, \mathbf{Y}$ and $\mathbf{Z}$ be pairwise disjoint node sets in a* PAG (CPDAG) $\mathcal{P}$. *Let $\mathcal{P}$ be amenable relative to $(\mathbf{X}, \mathbf{Y})$ and let $\mathbf{Z}$ satisfy the forbidden set condition relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{P}$. Let $\mathcal{M}$ be a* MAG (DAG) *in $[\mathcal{P}]$ and let p be a shortest proper non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ that is m-connecting given $\mathbf{Z}$ in $\mathcal{M}$. Let p^* in $\mathcal{P}$ be the corresponding path to p in $\mathcal{M}$. Then p^* is a proper definite status non-causal path in $\mathcal{P}$.*

Proof of Lemma 51. Let $p = \langle X = V_0, V_1, \dots, V_k = Y \rangle, k \geq 1$, such that $X \in \mathbf{X}, Y \in \mathbf{Y}$. It follows directly that p^* is proper and by (i) in Lemma 50, it is also non-causal in $\mathcal{P}$.

It is left to prove that p^* is of definite status in $\mathcal{P}$. For this we rely on the proof of Lemma 1 from Zhang (2006) (see Lemma 39 in Appendix A) and prove the following claims for p^*.

Claim 1 *If $V_r, 1 \leq r \leq k-1$ is not of definite status on p^*, then V_{r+1} is a parent of V_{r-1} in $\mathcal{M}$.*

Claim 2 *If $V_r, 1 \leq r \leq k-1$ is not of definite status on p^*, then V_{r-1} is a parent of V_{r+1} in $\mathcal{M}$.*

These claims contradict each other, so every node on p^* must be of definite status. Zhang (2006) proved these claims for a path q^* in $\mathcal{P}$, which is the path corresponding to a shortest path q from $\mathbf{X}$ to $\mathbf{Y}$ that is m-connecting given $\mathbf{Z}$ in $\mathcal{M}$. These claims are proven using the following argument:

If a node on q^* is not of definite status, then a subsequence q'^* that is formed by "jumping" over one, or a sequence of nodes on q^* one of which is not of definite status, constitutes a path in $\mathcal{P}$. Let q' in $\mathcal{M}$ be the path corresponding to q'^* in $\mathcal{P}$. Then q' is a path from $\mathbf{X}$ to $\mathbf{Y}$ that is shorter than q, so it is blocked by $\mathbf{Z}$ in $\mathcal{M}$. By the choice of q', the collider/definite non-collider status of all nodes on q', except two, is the same as on q. Therefore, one of these two nodes must block q' in $\mathcal{M}$. All possible cases for the status of these two nodes are considered in Zhang (2006) and a contradiction is reached in every case that does not support the claim being proven.

Almost exactly the same argument as in Zhang (2006) can be carried out to prove that Claim 1 and 2 hold for p^*. The only difference is in considering the paths that are subsequences of p. Since these paths are shorter than p and proper they are either blocked by $\mathbf{Z}$ or causal in $\mathcal{M}$. However, the subsequences of p considered in the proof of Lemma 39 are either immediately non-causal or they are constructed as in (i)-(iii) in Lemma 50 and thus non-causal by Lemma 50. The argument from Zhang (2006) then still holds for Claims 1 and 2 for p^*. ∎

Lemma 52 *Let $\mathbf{X}, \mathbf{Y}$ and $\mathbf{Z}$ be pairwise disjoint node sets in a* PAG (CPDAG) *$\mathcal{P}$. Let $\mathcal{P}$ be amenable relative to $(\mathbf{X}, \mathbf{Y})$ and let $\mathbf{Z}$ satisfy the forbidden set condition relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{P}$. Let $\mathcal{M}$ be a* MAG (DAG) *in $[\mathcal{P}]$ and let p be a path with minimal* distance-from-$\mathbf{Z}$ *among the shortest proper non-causal paths from $\mathbf{X}$ to $\mathbf{Y}$ that are m-connecting given $\mathbf{Z}$ in $\mathcal{M}$. Let p^* in $\mathcal{P}$ be the corresponding path to p in $\mathcal{M}$. Then p^* is a proper definite status non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ that is m-connecting given $\mathbf{Z}$ in $\mathcal{P}$.*

Proof of Lemma 52. By Lemma 51, p^* is a proper definite status non-causal path in $\mathcal{P}$. It is only left to prove that p^* is m-connecting given $\mathbf{Z}$ in $\mathcal{P}$.

Every definite non-collider on p^* in $\mathcal{P}$ corresponds to a non-collider on p in $\mathcal{M}$, and every collider on p^* is also a collider on p. Since p is m-connecting given $\mathbf{Z}$, no non-collider is in $\mathbf{Z}$ and every collider has a descendant in $\mathbf{Z}$. Let Q be an arbitrary collider (if there is one) on p. Then there is a directed path (possibly of zero length) from Q to a node in $\mathbf{Z}$ in $\mathcal{M}$. Let d be a shortest such path from Q to a node in $\mathbf{Z}$. Let d^* in $\mathcal{P}$ denote the corresponding path to p in $\mathcal{M}$. Then d^* is a possibly directed path from Q to Z in $\mathcal{P}$. It is only left to prove that d^* is a directed path. If d^* is of zero length, this is trivially true. Otherwise, suppose for a contradiction that there is a circle mark on d^*. Then d^* must start with a circle mark at Q (Lemma 7.2 from Maathuis and Colombo (2015) and Lemma 42, see Appendix A).

We first prove that d^* is unshielded in $\mathcal{P}$. Suppose for a contradiction that d^* is shielded. Then there exists a subpath $\langle A, B, C\rangle$ of d^* such that the edge $\langle A, C\rangle$ is in $\mathcal{P}$. The path corresponding to $d^*(Q, A) \oplus \langle A, C\rangle \oplus d^*(C, Z)$ must be a non-causal path from Q to Z in $\mathcal{M}$, otherwise we could have chosen a shorter path d. Hence, the edge $A \leftarrow\!\bullet C$ is in $\mathcal{M}$. But path d is directed from Q to Z in $\mathcal{M}$ so $A \rightarrow B \rightarrow C$ is also in $\mathcal{M}$. This contradicts that $\mathcal{M}$ is ancestral.

Let S be the first node on d after Q. If S is not a node on p, then following the proof of Lemma 2 from Zhang (2006) (Lemma 40 in Appendix 2) there exist nodes W on $p(X, Q)$ and

V on $p(Q,Y)$, distinct from Q, such that the path $W \bullet\!\!\rightarrow S \leftarrow\!\!\bullet V$ is in $\mathcal{M}$ and both W and V have the same colliders/non-collider status of both p and $p' = p(X,W) \oplus \langle W,S,V \rangle \oplus p(V,Y)$. Then p' is m-connecting given $\mathbf{Z}$. Since p' is non-causal and shorter than p, or as long as p but with a shorter distance-from-$\mathbf{Z}$ than p, p' must be non-proper, that is, $S \in \mathbf{X}$. But then $\langle S,V \rangle \oplus p(V,Y)$ is a proper non-causal m-connecting path from $\mathbf{X}$ to $\mathbf{Y}$ given $\mathbf{Z}$ that is shorter than p in $\mathcal{M}$. This contradicts our assumption about p.

If S is a node on p, then it lies either on $p(X,Q)$ or $p(Q,Y)$. Assume without loss of generality that S is on $p(Q,Y)$. Following the proof of Lemma 2 from Zhang (2006), there exists a node $W, W \neq Q$, on $p(X,Q)$ such that $W \bullet\!\!\rightarrow S$ is in $\mathcal{M}$ and W has the same collider/non-collider status on both p and $p' = p(X,W) \oplus \langle W,S \rangle \oplus p(S,Y)$. Then p' is m-connecting given $\mathbf{Z}$. Since p' is proper, and shorter than p, or as long as p but with a shorter distance-from-$\mathbf{Z}$ than p, p' must be causal in $\mathcal{M}$. Let p'^* in $\mathcal{P}$ denote the corresponding path to p in $\mathcal{M}$. Then p'^* is a possibly directed path from $\mathbf{X}$ to $\mathbf{Y}$, S is on p'^* and $Z \in \text{PossDe}(S,\mathcal{P})$, so $Z \in \text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{P}) \cap \mathbf{Z}$. This is a contradiction with $\mathbf{Z} \cap \text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{P}) = \emptyset$. ■

Proof of Lemma 10. Let the CPDAG (PAG) $\mathcal{G}$ be amenable relative to $(\mathbf{X},\mathbf{Y})$, and let $\mathbf{Z} \cap \text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G}) = \emptyset$.

We first prove $\neg$ (i)$\Rightarrow \neg$ (iii). Assume $\mathbf{Z}$ does not satisfy the blocking condition relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}$. Thus, there is a proper definite status non-causal path p from $X \in \mathbf{X}$ to $Y \in \mathbf{Y}$ that is m-connecting given $\mathbf{Z}$ in $\mathcal{G}$. Consider any DAG $\mathcal{D}$ (MAG $\mathcal{M}$) in $[\mathcal{G}]$. The path corresponding to p in $\mathcal{D}$ ($\mathcal{M}$) is a proper non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ that is m-connecting given $\mathbf{Z}$. Hence, $\mathbf{Z}$ does not satisfy the blocking condition relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{D}$ ($\mathcal{M}$) for all $\mathcal{D}$ ($\mathcal{M}$) in $[\mathcal{G}]$.

The implication $\neg$ (iii)$\Rightarrow \neg$ (ii) trivially holds, so it is only left to prove that $\neg$ (ii)$\Rightarrow \neg$ (i). Thus, assume that there is a DAG $\mathcal{D}$ (MAG $\mathcal{M}$) in $[\mathcal{G}]$ such that there is a proper non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{D}$ ($\mathcal{M}$) that is m-connecting given $\mathbf{Z}$. We choose a path p with minimal distance-from-$\mathbf{Z}$ among the shortest proper non-causal paths from $\mathbf{X}$ to $\mathbf{Y}$ that are m-connecting given $\mathbf{Z}$ in $\mathcal{D}$ ($\mathcal{M}$). By Lemma 52, the corresponding path p^* in $\mathcal{G}$ is a proper definite status non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ that is m-connecting given $\mathbf{Z}$. ■

Proof of Theorem 7. Assume that $\mathcal{G}$ is amenable relative to $(\mathbf{X},\mathbf{Y})$ and $\mathbf{Z}$ satisfies the blocking condition relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}$. Then it is left to prove that $\mathbf{Z}$ satisfies the blocking condition relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}$ if and only if it satisfies the separation condition relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}^{pbd}_{\mathbf{XY}}$.

We first prove that if $\mathbf{Z}$ satisfies the blocking condition, then $\mathbf{Z}$ m-separates $\mathbf{X}$ and $\mathbf{Y}$ in $\mathcal{G}^{pbd}_{\mathbf{XY}}$, that is, $\mathbf{Z}$ blocks all definite status paths from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}^{pbd}_{\mathbf{XY}}$. Since every definite status path from $\mathbf{X}$ to $\mathbf{Y}$ has a proper definite status path as a subpath, it is enough to show that $\mathbf{Z}$ blocks all proper definite status paths from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}^{pbd}_{\mathbf{XY}}$. Let p be a proper definite status path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}^{pbd}_{\mathbf{XY}}$. Then p must be non-causal, since otherwise $\mathcal{G}$ is not amenable. Let p^* in $\mathcal{G}$ be the path corresponding to p in $\mathcal{G}^{pbd}_{\mathbf{XY}}$, consisting of the same sequence of nodes as p. Then p^* is a proper definite status non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$. Thus, p^* is blocked by $\mathbf{Z}$. Since removing edges cannot m-connect a previously blocked path, p is blocked by $\mathbf{Z}$ in $\mathcal{G}^{pbd}_{\mathbf{XY}}$.

Next, we prove that if $\mathbf{Z}$ m-separates $\mathbf{X}$ and $\mathbf{Y}$ in $\mathcal{G}^{pbd}_{\mathbf{XY}}$, then $\mathbf{Z}$ satisfies the blocking condition. Suppose for a contradiction that there exists a proper definite status non-causal path p^* from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$ that is m-connecting given $\mathbf{Z}$. Let $\tilde{p}^*$ in $\mathcal{G}^{pbd}_{\mathbf{XY}}$ be the path corresponding to p^* in $\mathcal{G}$, constituted by the same sequence of nodes as p^*, if such a path exists in $\mathcal{G}^{pbd}_{\mathbf{XY}}$.

First, suppose $\tilde{p}^*$ does not exist in $\mathcal{G}^{pbd}_{\mathbf{XY}}$. Then since p^* is proper, it must start with a visible edge $X \to D$ in $\mathcal{G}$ such that D lies on a proper causal path from X to $\mathbf{Y}$, that is, $D \in \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$. Since p^* is non-causal and of definite status it must contain a collider $C \in \text{PossDe}(D, \mathcal{G})$. Since $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ is descendral (see Definition 13), $C \in \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ and similarly all descendants of C are in $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$. Considering that p^* is m-connecting given $\mathbf{Z}$ and C is a collider on p^*, it follows that $C \in \text{An}(\mathbf{Z}, \mathcal{G})$. This contradicts $\mathbf{Z} \cap \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) = \emptyset$.

Otherwise, $\tilde{p}^*$ is a path in $\mathcal{G}^{pbd}_{\mathbf{XY}}$. Then $\tilde{p}^*$ is a proper definite status non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}^{pbd}_{\mathbf{XY}}$ that is blocked by $\mathbf{Z}$. Since p^* is m-connecting given $\mathbf{Z}$ in $\mathcal{G}$, all colliders on p^* are in $\text{An}(\mathbf{Z}, \mathcal{G})$. Since no definite non-collider on p^* is in $\mathbf{Z}$, no definite non-collider on $\tilde{p}^*$ is in $\mathbf{Z}$. However, $\tilde{p}^*$ is blocked by $\mathbf{Z}$ in $\mathcal{G}^{pbd}_{\mathbf{XY}}$, so at least one collider C on $\tilde{p}^*$ (and therefore p^* as well) is not in $\text{An}(\mathbf{Z}, \mathcal{G}^{pbd}_{\mathbf{XY}})$. Thus, any directed path from C to $\mathbf{Z}$ must contain a visible edge $X \to D$, where $D \in \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$. This implies $D \in \text{An}(\mathbf{Z}, \mathcal{G})$, which contradicts $\mathbf{Z} \cap \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) = \emptyset$. ■

Appendix C. Proofs for Section 4

Proof of Theorem 11. The first part of the theorem follows from from Rule 3 of the do-calculus for DAGs from Pearl (2009), Rule 3 of the do-calculus for MAGs and PAGs from Zhang (2006) and the properties of the CPDAG.

Since the generalized adjustment criterion is sound and complete with respect to adjustment in DAGs, CPDAGs, MAGs and PAGs (Theorem 5), we will prove the second part of this theorem by proving that if $\mathbf{Z}$ satisfies the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$, then $\mathbf{Z}$ satisfies the generalized adjustment criterion relative to $(\mathbf{X}', \mathbf{Y})$ in $\mathcal{G}$. We prove this by showing that $\mathbf{Z}$ satisfies the three conditions of Definition 4 relative to $(\mathbf{X}', \mathbf{Y})$ in $\mathcal{G}$.

We first show that a possibly directed path from $\mathbf{X}'$ to $\mathbf{Y}$ is proper with respect to $\mathbf{X}'$ if and only if it is proper with respect to $\mathbf{X}$. Since $\mathbf{X}' \subseteq \mathbf{X}$ any path that is proper with respect to $\mathbf{X}$ will also be proper with respect to $\mathbf{X}'$. Hence, we only need to show that if p is a possibly directed path from $X' \in \mathbf{X}'$ to $Y \in \mathbf{Y}$ that is proper with respect to $\mathbf{X}'$, then p is also proper with respect to $\mathbf{X}$. Suppose for a contradiction that p is not proper with respect to $\mathbf{X}$. Let $p(X, Y)$ be the subpath of p that is proper with respect to $\mathbf{X}$. Since $X \notin \mathbf{X}'$, X must be in $\mathbf{X} \setminus \mathbf{X}'$, which contradicts the assumption that there are no possibly directed paths from $\mathbf{X} \setminus \mathbf{X}'$ to $\mathbf{Y}$ that are proper with respect to $\mathbf{X}$.

Then $\text{Forb}(\mathbf{X}', \mathbf{Y}, \mathcal{G}) \subseteq \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$, so $\mathbf{Z}$ must satisfy the forbidden set condition relative to $(\mathbf{X}', \mathbf{Y})$ in $\mathcal{G}$. Additionally, since $\mathbf{Z}$ satisfies the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$, $\mathcal{G}$ must be amenable relative to $(\mathbf{X}, \mathbf{Y})$. This means that every possibly directed path from $\mathbf{X}$ to $\mathbf{Y}$ that is proper with respect to $\mathbf{X}$ starts with a visible edge out of $\mathbf{X}$. Since $\mathbf{X}' \subseteq \mathbf{X}$ and since possibly directed paths from $\mathbf{X}'$ to $\mathbf{Y}$ that are proper with respect to $\mathbf{X}'$ are proper with respect to $\mathbf{X}$, it follows that the amenability condition is satisfied relative to $(\mathbf{X}', \mathbf{Y})$ and $\mathcal{G}$.

It is only left to show that $\mathbf{Z}$ satisfies the blocking condition relative to $(\mathbf{X}', \mathbf{Y})$ in $\mathcal{G}$. Since $\mathbf{Z}$ satisfies the blocking condition relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$ and since $\mathbf{X}' \subseteq \mathbf{X}$, $\mathbf{Z}$ blocks every definite status non-causal path from $\mathbf{X}'$ to $\mathbf{Y}$ that is proper with respect to $\mathbf{X}$. Hence, we only need to show that $\mathbf{Z}$ also blocks every definite status non-causal path from $\mathbf{X}'$ to $\mathbf{Y}$ that is proper with respect to $\mathbf{X}'$, but not proper with respect to $\mathbf{X}$. Let p be one such path from $X' \in \mathbf{X}'$ to $Y \in \mathbf{Y}$. Since p is proper with respect to $\mathbf{X}'$, but not with respect to $\mathbf{X}$, let $X \in \mathbf{X} \setminus \mathbf{X}'$ be the node on p such that $p(X, Y)$ is proper with respect to $\mathbf{X}$. Since there is no possibly directed path from $\mathbf{X} \setminus \mathbf{X}'$ to $\mathbf{Y}$ that is proper with respect to $\mathbf{X}$, $p(X, Y)$ must be a non-causal path from X to Y. Additionally, since $p(X, Y)$ is a subpath of p, $p(X, Y)$ is of definite status. Then $p(X, Y)$ is a non-causal definite status path from $\mathbf{X}$ to $\mathbf{Y}$ that is proper with respect to $\mathbf{X}$, so $p(X, Y)$ is blocked by $\mathbf{Z}$. Thus, p is also blocked by $\mathbf{Z}$. ■

Proof of Lemma 16. There is a proper definite status non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ that is m-connecting given $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$. Among all such paths consider the ones with minimal length and among those let $p = \langle X, \dots, Y \rangle, X \in \mathbf{X}, Y \in \mathbf{Y}$ be the path with a shortest distance-from-$(\mathbf{X} \cup \mathbf{Y})$ in $\mathcal{G}$. By the choice of p, (i) holds. It is left to prove that (ii)$-$(iv) also hold for p.

(ii) Since p is m-connecting given $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$ and since $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$ $= \text{PossAn}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G}) \setminus (\mathbf{X} \cup \mathbf{Y} \cup \mathbf{I})$, any collider on p is in $\text{PossAn}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G})$. Since p is proper, no collider on p is in $\mathbf{X}$. Additionally, no collider C on p is in $\mathbf{Y} \setminus \mathbf{I}$, otherwise $p(X, C)$ is a non-causal path and we could have chosen a shorter p. It is only left to show that no collider on p is in $\mathbf{I}$. Suppose for a contradiction that a collider on p is in $\mathbf{I}$. Since $\mathbf{I}$ is a descendral set, all (possible) descendants of this collider are also in $\mathbf{I}$. But then p is not m-connecting given $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$, a contradiction.

(iii) Any definite non-collider on p is a possible ancestor of an endpoint node of p or of a collider on p. Then it follows from (ii) that any definite non-collider on p is in $\text{PossAn}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G})$. Furthermore, p is m-connecting given $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$, so no definite non-collider on p is in $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$. Since $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I} = \text{PossAn}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G}) \setminus (\mathbf{X} \cup \mathbf{Y} \cup \mathbf{I})$, it follows that all definite non-colliders on p are in $\text{PossAn}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G}) \cap (\mathbf{X} \cup \mathbf{Y} \cup \mathbf{I})$. Since p is proper, no definite non-collider on p is in $\mathbf{X}$. Additionally, no definite non-collider C on p is in $\mathbf{Y} \setminus \mathbf{I}$, otherwise we could have chosen path $p(X, C)$ instead of p. Thus, all definite non-colliders on p are in $\text{PossAn}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G}) \cap \mathbf{I} \subseteq \mathbf{I}$.

(iv) Let C be a collider on p. From (ii), it follows that $C \notin \mathbf{X} \cup \mathbf{Y}$ and that there is an unshielded possibly directed path from C to a node $V \in \mathbf{X} \cup \mathbf{Y}$. Suppose that this path starts with an edge of type $C \circ\!\!-\!\!\circ Q$ (possibly $Q = V$). We derive a contradiction by constructing a proper definite status non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ that is m-connecting given $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$ and shorter than p, or of the same length as p, but with a shorter distance-from-$(\mathbf{X} \cup \mathbf{Y})$.

Let A and B be nodes on p such that $A \bullet\!\!\rightarrow C \leftarrow\!\!\bullet B$ is a subpath of p (possibly $A = X$, $B = Y$). Then paths $A \bullet\!\!\rightarrow C \circ\!\!-\!\!\circ Q$ and $B \bullet\!\!\rightarrow C \circ\!\!-\!\!\circ Q$ together with Lemma 36 imply that $A \bullet\!\!\rightarrow Q \leftarrow\!\!\bullet B$ is in $\mathcal{G}$.

Since all colliders on p are not in $\mathbf{I}$, and all definite non-colliders on p are in $\mathbf{I}$, and since $\mathbf{I}$ is a descendral set, it follows that no collider on p is a possible descendant of a definite non-collider on p. Thus, if $A \neq X$ $(B \neq Y)$, then $A \leftrightarrow C$ $(C \leftrightarrow B)$ is in $\mathcal{G}$. Moreover, if $A \leftrightarrow C$ $(C \leftrightarrow B)$ is in $\mathcal{G}$, then $A \leftrightarrow Q$ $(Q \leftrightarrow B)$ is in $\mathcal{G}$, otherwise a possibly directed path

$\langle A, Q, C\rangle$ ($\langle B, Q, C\rangle$) and $C \leftrightarrow A$ ($C \leftrightarrow B$) are in $\mathcal{G}$, which contradicts Lemma 44. Hence, if $A \neq X$ ($B \neq Y$), the collider/definite non-collider status of A (B) is the same on p and on $p(X, A) \oplus \langle A, Q\rangle$ ($\langle Q, B\rangle \oplus p(B, Y)$).

Suppose first that $Q \in \mathbf{X} \cup \mathbf{Y}$. If $Q \in \mathbf{X}$, then $\langle Q, B\rangle \oplus p(B, Y)$ is a proper definite status non-causal path that is m-connecting given $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$ in $\mathcal{G}$ and shorter than p. Otherwise, $Q \in \mathbf{Y}$. If $Q \in \mathbf{Y} \cap \mathbf{I}$, this would imply that $C \in \mathbf{I}$, which contradicts (ii). So Q must be in $\mathbf{Y} \setminus \mathbf{I}$. Then $p(X, A) \oplus \langle A, Q\rangle$ is a non-causal path. Hence, we found a proper definite status non-causal path that is m-connecting given $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$ in $\mathcal{G}$ and shorter than p, which is a contradiction.

Next, suppose that $Q \notin \mathbf{X} \cup \mathbf{Y}$. Then if Q is not on p, $p(X, A) \oplus \langle A, Q, B\rangle \oplus p(B, Y)$ is a proper definite status non-causal path that is m-connecting given $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$ in $\mathcal{G}$ and of the same length as p, but with a shorter distance-from-$(\mathbf{X} \cup \mathbf{Y})$. Otherwise, Q is on p. Then Q is a collider on p, otherwise $Q \in \mathbf{I}$ and $C \in \mathbf{I}$. Suppose first that Q is on $p(C, Y)$. Then $p(X, A) \oplus \langle A, Q\rangle \oplus p(Q, Y)$ is a non-causal path because $p(Q, Y)$ is into Q. Hence, there is a proper definite status non-causal path that is m-connecting given $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$ in $\mathcal{G}$ and shorter than p, which is a contradiction. Next, suppose that Q is on $p(X, C)$. Then $p(X, Q) \oplus \langle Q, B\rangle \oplus p(B, Y)$ is a non-causal path since it contains $Q \leftarrow B$. This path is a proper definite status non-causal path that is m-connecting given $\text{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \setminus \mathbf{I}$ in $\mathcal{G}$ and shorter than p, which is a contradiction. ■

The following proof is similar to the proof of Lemma 1 from Richardson (2003) (see Lemma 37 in Appendix A). The difference lies in the fact that Lemma 17 additionally considers CPDAGs and PAGs $\mathcal{G}$ in which we define m-separation (m-connection) only for paths of definite status, as well as the fact that we require the resulting path p to be proper and non-causal from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$.

Proof of Lemma 17. Let p be a path in $\mathcal{G}$ satisfying (i)−(ii). If there is no collider on p, or all colliders on p are in $\text{An}(\mathbf{Z}, \mathcal{G})$, then p is a proper definite status non-causal path that is m-connecting given $\mathbf{Z}$ in $\mathcal{G}$ and the lemma holds.

Hence, assume there is at least one collider C on p that is not in $\text{An}(\mathbf{Z}, \mathcal{G})$. By (ii), $C \in \text{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G}) \setminus \text{An}(\mathbf{Z}, \mathcal{G})$. We now construct a path q from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$ that is m-connecting given $\mathbf{Z}$ and prove it is proper, of definite status and non-causal.

Let D be the node closest to Y on p such that $D \in \text{An}(\mathbf{X}, \mathcal{G}) \setminus \text{An}(\mathbf{Z}, \mathcal{G})$ if such a node exists, otherwise let $D = X$. Let E be the node closest to D on $p(D, Y)$ such that $E \in \text{An}(\mathbf{Y}, \mathcal{G}) \setminus \text{An}(\mathbf{Z}, \mathcal{G})$ if such a node exists, otherwise let $E = Y$. Since at least one collider on p is in $\text{An}(\mathbf{X} \cup \mathbf{Y}) \setminus \text{An}(\mathbf{Z}, \mathcal{G})$, either $D \neq X$ or $E \neq Y$ must hold. Moreover, if $D = X$, then since there is at least one collider C on p that is in $\text{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G}) \setminus \text{An}(\mathbf{Z}, \mathcal{G})$, it follows that $E \neq D$. However, if $D \neq X$, then $D = E$ is possible.

Since $D \in \text{An}(\mathbf{X}, \mathcal{G}) \setminus \text{An}(\mathbf{Z}, \mathcal{G})$, let v_D be a shortest directed path from D to a node in $\mathbf{X}$ (possibly of length zero). Since $E \in \text{An}(\mathbf{Y}, \mathcal{G}) \setminus \text{An}(\mathbf{Z}, \mathcal{G})$, let v_E be a shortest directed path from E to a node in $\mathbf{Y}$ (possibly of length zero). Thus, all non-endpoint nodes on v_D and v_E are in $\text{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G}) \setminus \text{An}(\mathbf{Z}, \mathcal{G})$. Also, by the choice of D and E, no non-endpoint node on $p(D, E)$ is in $\text{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G}) \setminus \text{An}(\mathbf{Z}, \mathcal{G})$. Hence, no non-endpoint node on either v_D or v_E is also on $p(D, E)$.

Let $q = (-v_D) \oplus p(D, E) \oplus v_E$. We prove that q is a proper definite status non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$ that is m-connecting given $\mathbf{Z}$. Path q is of definite status by construction.

To prove that q is proper we must show that no non-endpoint node on v_D or v_E is in $\mathbf{X}$. No non-endpoint node on v_D is in $\mathbf{X}$, otherwise we could have chosen a shorter v_D. Similarly, no non-endpoint node on v_E is in $\mathbf{X}$, as this would contradict the choice of D. It is left to show that q is non-causal from $\mathbf{X}$ to $\mathbf{Y}$ and m-connecting given $\mathbf{Z}$.

We first show that q is m-connecting given $\mathbf{Z}$. By assumption, no definite non-collider on p is in $\mathbf{Z}$. Additionally, if D and E are non-endpoints on p, then all nodes on v_D and v_E are in $\mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G}) \setminus \mathrm{An}(\mathbf{Z}, \mathcal{G})$, that is, no node on either v_D or v_E is in $\mathbf{Z}$. Hence, no definite non-collider on q is in $\mathbf{Z}$. For q to be m-connecting given $\mathbf{Z}$ we still have to show that all colliders on q are in $\mathrm{An}(\mathbf{Z}, \mathcal{G})$. Any collider on q is a non-endpoint node on $p(D, E)$. Since no non-endpoint node on $p(D, E)$ is in $\mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{G}) \setminus \mathrm{An}(\mathbf{Z}, \mathcal{G})$, by choice of D and E, and all colliders on p are in $\mathrm{An}(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z}, \mathcal{G})$, by assumption, it follows that any collider on $p(D, E)$ must be in $\mathrm{An}(\mathbf{Z}, \mathcal{G})$.

It is only left to show that q is a non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$. If v_D is not of zero length, this is obviously true. If v_D is of zero length, then $q = p(X, E) \oplus v_E$. Since v_E is a directed path from E to Y, we need to show that $p(X, E)$ is non-causal from X to E. Suppose for a contradiction that $p(X, E)$ is possibly directed from X to E. Then q is a proper possibly directed path from $\mathbf{X}$ to $\mathbf{Y}$, so $E \in \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \subseteq \mathbf{I}$. By assumption, there is at least one collider on p, and in particular on $p(E, Y)$. Let C be the collider on $p(E, Y)$ closest to E. Then $C \in \mathrm{De}(E, \mathcal{G})$. Since $\mathrm{De}(E, \mathcal{G}) \subseteq \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$, it follows that $C \in \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$. Since $\mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G}) \subseteq \mathbf{I}$, this is in contradiction with (ii). ■

Appendix D. Proofs for Section 5

Proof of Lemma 21. If there is a path p from $\mathbf{X}$ to $\mathbf{Y}$ for which the listed statements hold, then no set $\mathbf{Z}$ such that $\mathbf{Z} \cap \mathrm{De}(\mathbf{X}, \mathcal{D}) = \emptyset$ can block p.

Conversely, since $\mathbf{Z} \cap \mathrm{De}(\mathbf{X}, \mathcal{D}) = \emptyset$ and $\mathbf{Z}$ satisfies the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$, then $\mathbf{Z}' = \mathrm{Adjust}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) \setminus \mathrm{De}(\mathbf{X}, \mathcal{D})$ satisfies the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ (Theorem 14). Suppose there is no back-door set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$. Then $\mathbf{Z}'$ violates condition (ii) of Pearl's back-door criterion relative to at least one $X \in \mathbf{X}$ and $Y \in \mathbf{Y}$ in $\mathcal{D}$. Hence, we can choose p to be a shortest back-door path from a node $X \in \mathbf{X}$ to a node $Y \in \mathbf{Y}$ that is d-connecting given $\mathbf{Z}'$ in $\mathcal{D}$. Then the statement in (i) holds for p. We prove that the statements (ii)$-$(iv) also hold for p.

(ii) Since $\mathbf{Z}'$ satisfies the generalized adjustment criterion and $\mathbf{Z}'$ does not block the back-door path p, it follows that p is not proper. Since p is not proper, a subpath of p forms a proper path q from $\mathbf{X}$ to $\mathbf{Y}$. If q is non-causal, then it is blocked by $\mathbf{Z}'$. But in this case $\mathbf{Z}'$ would block p as well. Hence, q is causal.

(iii) Any non-collider on p is an ancestor of an endpoint or a collider on p. Since p is d-connecting given $\mathbf{Z}'$, all colliders on p are in $\mathrm{An}(\mathbf{Z}', \mathcal{D})$. By our choice of $\mathbf{Z}'$, $\mathrm{An}(\mathbf{Z}', \mathcal{D}) \subseteq \mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D})$, so all colliders on p are in $\mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D})$. Since any non-collider on p is an ancestor of an endpoint or a collider on p, it follows that all non-colliders on p are also in $\mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D})$. Moreover, since p is d-connecting given $\mathbf{Z}'$, no non-collider on p is in $\mathbf{Z}'$. Thus, since $\mathbf{Z}' = \mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D}) \setminus (\mathrm{De}(\mathbf{X}, \mathcal{D}) \cup \mathbf{Y})$, any non-collider on p must be in $\mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D}) \cap (\mathrm{De}(\mathbf{X}, \mathcal{D}) \cup \mathbf{Y})$. Path p is a shortest back-door path from $\mathbf{X}$ to $\mathbf{Y}$, so no non-collider on p is in $\mathbf{Y}$. Hence, all non-colliders on p are in $\mathrm{De}(\mathbf{X}, \mathcal{D})$.

Now, assume that there is a collider C on p. This collider is a descendant of X or a non-collider on p, so $C \in \text{De}(\mathbf{X}, \mathcal{D})$ as well. However, since $\text{De}(\mathbf{X}, \mathcal{D}) \cap \mathbf{Z} = \emptyset$ and $\text{De}(C, \mathcal{D}) \subseteq \text{De}(\mathbf{X}, \mathcal{D})$ it follows that p is blocked by $\mathbf{Z}'$. This contradicts that p is d-connecting given $\mathbf{Z}'$.

(iv) From (ii), it follows that $Y \in \text{De}(\mathbf{X}, \mathcal{D})$. Additionally, we've shown in (iii) that there is no collider on p and that all non-collider on p are in $\text{De}(\mathbf{X}, \mathcal{D})$. Thus, all nodes on p are in $\text{De}(\mathbf{X}, \mathcal{D})$. ■

Proof of Lemma 23. Let p be a definite status back-door path from a node $X \in \mathbf{X}$ to a node $Y \in \mathbf{Y}$ in $\mathcal{G}$. Then there exists a node $X' \in \mathbf{X}$ (possibly $X' = X$) on p such that the subpath $p(X', Y)$ is a proper path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$. Since $p(X', Y)$ is a subpath of p, $p(X', Y)$ is of definite status.

If $X' \neq X$ and X' is a definite non-collider on p, then since $X' \in \mathbf{Z} \cup \mathbf{X} \setminus \{X\}$, p is blocked by $\mathbf{Z} \cup \mathbf{X} \setminus \{X\}$. Else, $X' \neq X$ and X' is a definite collider on p, or $X' = X$ and p is a proper back-door path in $\mathcal{G}$. Since $\mathcal{G}$ is amenable, all proper back-door paths from $\mathbf{X}$ to $\mathbf{Y}$ are also proper non-causal paths from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$. We prove that $\mathbf{Z} \cup \mathbf{X} \setminus \{X\}$ blocks p, by proving that it blocks $p(X', Y)$.

Suppose for a contradiction that $p(X', Y)$ is m-connecting given $\mathbf{Z} \cup \mathbf{X} \setminus \{X\}$ in $\mathcal{G}$. We show that it is then possible to construct a proper definite status non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$, that is m-connecting given $\mathbf{Z}$, which contradicts that $\mathbf{Z}$ satisfies the blocking condition relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$. Since $p(X', Y)$ is a proper back-door path and since $\mathcal{G}$ is amenable relative to $(\mathbf{X}, \mathbf{Y})$, $p(X', Y)$ is a non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$. Then since $p(X', Y)$ also of definite status, it must be blocked by $\mathbf{Z}$ in $\mathcal{G}$. Since $p(X', Y)$ is m-connecting given $\mathbf{Z} \cup \mathbf{X} \setminus \{X\}$ and blocked by $\mathbf{Z}$, it follows that no definite non-collider on $p(X', Y)$ is in $\mathbf{Z}$ and there is at least one collider on $p(X', Y)$ that is in $\text{An}(\mathbf{X} \setminus \{X\}, \mathcal{G}) \setminus \text{An}(\mathbf{Z}, \mathcal{G})$. Let C be the collider closest to Y on $p(X', Y)$ such that $C \in \text{An}(\mathbf{X}, \mathcal{G}) \setminus \text{An}(\mathbf{Z}, \mathcal{G})$. Let q be of the form $C \to \cdots \to X''$, $X'' \in \mathbf{X}$ be the shortest causal path from C to $\mathbf{X}$. Since $p(X', Y)$ is proper, it follows that $C \neq X''$.

Let D be the node closest to X'' on $-q$ such that D is also on $p(C, Y)$ (possibly $D = C$) and $r = -q(X'', D) \oplus p(D, Y)$. It is left to show that r is a proper definite status non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ that is m-connecting given $\mathbf{Z}$ in $\mathcal{G}$. Since $p(C, Y)$ does not contain a node in $\mathbf{X}$, it follows that $-q(X'', D)$ is of non-zero length. So r is a definite status non-causal path. Additionally, since q was chosen as the shortest path from C to $\mathbf{X}$, it follows that r is proper. Lastly, since both q and $p(C, Y)$ are m-connecting given $\mathbf{Z}$ and $D \notin \mathbf{Z}$ and D is a definite non-collider on r, it follows that r is m-connecting given Z. ■

Corollary 53 *Let X and Y be distinct nodes in a* DAG*,* CPDAG*,* MAG *or* PAG *$\mathcal{G}$ and let $\mathcal{R}_{\underline{X}}$ be a graph as defined in Definition 46. If there exists a generalized back-door set relative to (X, Y) in $\mathcal{G}$, then $DSEP(X, Y, \mathcal{R}_{\underline{X}}) \subseteq \text{Adjust}(X, Y, \mathcal{G}) \setminus \text{PossDe}(X, \mathcal{G})$.*

Proof of Corollary 53. Since there exists a generalized back-door set relative to (X, Y) in $\mathcal{G}$, by Theorem 47 $DSEP(X, Y, \mathcal{R}_{\underline{X}}) \subseteq \text{PossAn}(X \cup Y, \mathcal{G}) \setminus (\text{PossDe}(X, \mathcal{G}) \cup Y)$ is a generalized back-door set relative to (X, Y) in $\mathcal{G}$. Additionally, by Corollary 24 $\text{Adjust}(X, Y, \mathcal{G}) \setminus \text{PossDe}(X, \mathcal{G})$ is also generalized back-door set relative to (X, Y) in $\mathcal{G}$ and $\text{Adjust}(X, Y, \mathcal{G}) \setminus \text{PossDe}(X, \mathcal{G}) = \text{PossAn}(X \cup Y, \mathcal{G}) \setminus (\text{PossDe}(X, \mathcal{G}) \cup Y)$. ■

Proof of Lemma 25. Let there be a p from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$ for which (i)$-$(iv) hold. Then p is a proper definite status non-causal path that is m-connecting given $\text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G}) \setminus \text{PossDe}(\mathbf{X},\mathcal{G})$. Thus, $\text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G}) \setminus \text{PossDe}(\mathbf{X},\mathcal{G})$ violates the blocking condition relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}$ and Theorem 14 implies that there is no adjustment set $\mathbf{Z}$ relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}$ such that $\mathbf{Z} \cap \text{PossDe}(\mathbf{X},\mathcal{G}) = \emptyset$.

Conversely, assume there is no adjustment set $\mathbf{Z}$ relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}$ such that $\mathbf{Z} \cap \text{PossDe}(\mathbf{X},\mathcal{G}) = \emptyset$. Since there exists an adjustment set relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}$, $\mathcal{G}$ is amenable relative to $(\mathbf{X},\mathbf{Y})$. Then Theorem 14 implies that $\text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G}) \setminus \text{PossDe}(\mathbf{X},\mathcal{G})$ violates the blocking condition relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}$. Hence, there is a proper definite status non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$ that is m-connecting given $\text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G}) \setminus \text{PossDe}(\mathbf{X},\mathcal{G})$. Then we can use Lemma 16 with $\mathbf{I} = \text{PossDe}(\mathbf{X},\mathcal{G})$ to choose a shortest path p from $X \in \mathbf{X}$ to $Y \in \mathbf{Y}$ for which (i)$-$(iv) in Lemma 16 hold.

We now show that (i)$-$(iv) in Lemma 25 hold for p.

(i) Follows immediately from (i) in Lemma 16.

(ii) Any definite non-collider on p is in $\text{PossDe}(\mathbf{X},\mathcal{G})$ ((iii) in Lemma 16). Since there is an adjustment set relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}$, it follows from Corollary 15 that $\text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G})$ satisfies the generalized adjustment criterion. Thus, p is blocked by $\text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G})$ and m-connecting given $\text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G}) \setminus \text{PossDe}(\mathbf{X},\mathcal{G})$. This implies that at least one definite non-collider on p must be in $\text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G}) \cap \text{PossDe}(\mathbf{X},\mathcal{G})$.

For the remainder of the proof let V be the definite non-collider that is closest to Y on p, among all definite non-colliders on p in $\text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G}) \cap \text{PossDe}(\mathbf{X},\mathcal{G})$.

(iii) By (ii) in Lemma 16, all colliders on p are in $\text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G}) \setminus \text{PossDe}(\mathbf{X},\mathcal{G})$. It is left to show that all definite non-colliders on $p(V,Y)$ are in $\text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$.

All definite non-colliders on p are possible ancestors of an endpoint node of p or a collider on p. Hence, all definite non-colliders on p are in $\text{PossAn}(\mathbf{X} \cup \mathbf{Y},\mathcal{G})$. By (iii) in Lemma 16, all definite non-colliders are also in $\text{PossDe}(\mathbf{X},\mathcal{G})$. Hence, all definite non-colliders on p are in $\text{PossAn}(\mathbf{X} \cup \mathbf{Y},\mathcal{G}) \cap \text{PossDe}(\mathbf{X},\mathcal{G})$. Additionally, by the choice of V, no definite non-collider on $p(V,Y)$ is in $\text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G}) \cap \text{PossDe}(\mathbf{X},\mathcal{G})$. Since $\text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G}) \cap \text{PossDe}(\mathbf{X},\mathcal{G}) = (\text{PossAn}(\mathbf{X} \cup \mathbf{Y},\mathcal{G}) \cap \text{PossDe}(\mathbf{X},\mathcal{G})) \setminus (\mathbf{X} \cup \mathbf{Y} \cup \text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G}))$, it follows that all definite non-colliders on $p(V,Y)$ are in $\mathbf{X} \cup \mathbf{Y} \cup \text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$. It is only left to show that no definite non-collider on $p(V,Y)$ is in $\mathbf{X}$ or $\mathbf{Y} \setminus \text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$. Since p is proper, no definite non-collider on p is in $\mathbf{X}$. Also, no non-endpoint node C on p is in $\mathbf{Y} \setminus \text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$, otherwise $p(X,C)$ is a non-causal path and we could have chosen a shorter p. Hence, any definite non-collider on $p(V,Y)$ is in $\text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$.

(iv) Let V_2 be the node closest to Y on p such that $p(V,V_2)$ is a possibly directed path from V to V_2 (possibly of zero length). We will show that $V_2 = V$. Note that V_2 is either Y, V or a collider on p. Since $V_2 \in \text{PossDe}(V,\mathcal{G})$ and $\text{PossDe}(V,\mathcal{G}) \subseteq \text{PossDe}(\mathbf{X},\mathcal{G})$, by (iii) V_2 cannot be a collider on p. Hence, V_2 is either Y or V.

Let V_1 be the node closest to X on p such that $-p(V,V_1)$ is an possibly directed path from V to V_1 (possibly of zero length). We will show that $V_1 = X$. Note that V_1 is either X, V or a collider on p. Using the same reasoning as for V_2, V_1 cannot be a collider on p. So V_1 is either X or V. As V is a definite non-collider on p, either $V_1 \neq V$ or $V_2 \neq V$.

Suppose that $V_2 \neq V$. Then $V_2 = Y$ and $p(V,Y)$ is a possibly directed path from V to Y. Since $V \in \text{PossDe}(\mathbf{X},\mathcal{G})$, let q be a proper possibly directed path from $X' \in \mathbf{X}$ (possibly $X = X'$) to V in $\mathcal{G}$. Let W' (possibly $W' = V$) be the node closest to X' on q that is also on

$p(V, Y)$. Then $q(X', W') \oplus p(W', Y)$ is a proper possibly directed path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$, so $W' \in \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$. Since $V \in \text{PossDe}(W', \mathcal{G})$, it follows that $V \in \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$. This contradicts (ii).

Hence, $V_2 = V$. Then $V_1 = X$ and p is of the form $X \cdots \leftarrow V \leftarrow\!\!\bullet W \dots Y$, possibly $W = Y$ is in $\mathcal{G}$. Since p is of definite status, $p(X, V)$ must be of the form $X \leftarrow \cdots \leftarrow V$. If $W \neq Y$, then if W is a definite non-collider on $p(V, Y)$, (iii) implies that $W \in \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$. Hence, $V \leftrightarrow W$ is in $\mathcal{G}$, otherwise $V \in \text{PossDe}(W, \mathcal{G}) \subseteq \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$, which contradicts (ii). Else, W is a collider on $p(V, Y)$, so $V \leftrightarrow W$ must be on p. ■

Proof of Corollary 27. (i) Let $\mathbf{X}' = \mathbf{X} \cap \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$. Since $\mathbf{X}' \neq \emptyset$, there exists a proper possibly directed path $p = \langle X = V_1, \dots, V_k = Y \rangle, k > 1$, from $X \in \mathbf{X}$ to $Y \in \mathbf{Y}$ in $\mathcal{G}$ such that for some V_i, $i \in \{2, \dots, k\}$, $\text{PossDe}(V_i, \mathcal{G}) \cap \mathbf{X}' \neq \emptyset$. Let $V_j, j \in \{2, \dots, k\}$, be the node closest to Y on p such that $\text{PossDe}(V_j, \mathcal{G}) \cap \mathbf{X}' \neq \emptyset$. Let q be a shortest possibly directed path from V_j to a node X' in $\mathbf{X}'$. Since p was chosen to be proper with respect to $\mathbf{X}$, we know that $V_j \neq X'$.

By the choice of V_j on p, no other node from $p(V_j, Y)$ is on q. By Lemma 41, let $\overline{p(V_j, Y)}$ be a subsequence of $p(V_j, Y)$ that forms a possibly directed definite status path from V_j to Y. Now, we can concatenate $-q$ and $\overline{p(V_j, Y)}$ to form the path $r = (-q) \oplus \overline{p(V_j, Y)}$. We will show that r is a proper definite status non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ that does not contain a collider and consists of nodes in $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$. This means that r satisfies condition (2) in Theorem 26, so there is no set that satisfies the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$. By completeness of the generalized adjustment criterion (Theorem 5), there is no adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$.

We first show that r is proper with respect to $\mathbf{X}$. Since p is proper with respect to $\mathbf{X}$ and since $V_j \neq X$, $\overline{p(V_j, Y)}$ does not contain a node in $\mathbf{X}$. Additionally, by choice of q, X' is the only node from $\mathbf{X}$ on q. Hence, r is proper with respect to $\mathbf{X}$.

Since q and $\overline{p(V_j, Y)}$ are possibly directed paths from V_j to X' and from V_j to Y, there is no collider on r. Additionally, since $V_j \in \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ and since $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ is a descendral set in $\mathcal{G}$, all nodes on r are in $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$.

Next, we show that r is of definite status. First, $\overline{p(V_j, Y)}$ is of definite status. Since q is a shortest possibly directed path from V_j to X', it is of definite status by Lemma 41. Hence, if $V_j = Y$, then $r = (-q)$ is of definite status. If $V_j \neq Y$, it is left to show that V_j is of definite status on r.

We prove this by contradiction. Thus, suppose that V_j is not of definite status on r. Let $\langle A, V_j, B \rangle$ be a subpath of r, so that A is the node adjacent to V_j on q and B is the node adjacent to V_j on $\overline{p(V_j, Y)}$. Then $A \bullet\!\!\rightarrow V_j \circ\!\!-\!\!\bullet B$, $A \bullet\!\!-\!\!\circ V_j \leftarrow\!\!\bullet B$, or $A \bullet\!\!-\!\!\circ V_j \circ\!\!-\!\!\bullet B$ is in $\mathcal{G}$ and there is an edge $\langle A, B \rangle$ in $\mathcal{G}$. Since q and $\overline{p(V_j, Y)}$ are possibly directed paths from V_j to X' and from V_j to Y, $A \bullet\!\!\rightarrow V_j \circ\!\!-\!\!\bullet B$ and $A \bullet\!\!-\!\!\circ V_j \leftarrow\!\!\bullet B$ cannot be in $\mathcal{G}$. Hence, $A \bullet\!\!-\!\!\circ V_j \circ\!\!-\!\!\bullet B$ is a subpath of r.

Since $A \bullet\!\!-\!\!\circ V_j \circ\!\!-\!\!\bullet B$ is in $\mathcal{G}$, neither $A \bullet\!\!\rightarrow B$ nor $A \leftarrow\!\!\bullet B$ can be in $\mathcal{G}$ (Lemma 36). This implies that $A \circ\!\!-\!\!\circ B$ is in $\mathcal{G}$ and so, $\langle B, A \rangle \oplus q(A, X')$ is a possibly directed path from B to X', which contradicts the choice of V_j (since B is closer to Y on p).

It is only left to show that r is a non-causal path from X' to Y. We again use a proof by contradiction. Thus, suppose that r is a possibly directed path from X' to Y. Then, since $r(X', V_j) = (-q)$ and q are both possibly directed paths in $\mathcal{G}$, r must start with a non-directed edge. Hence, r is a proper possibly directed path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$ that starts with a non-directed edge, which contradicts that $\mathcal{G}$ is amenable relative to $(\mathbf{X},\mathbf{Y})$.

(ii) Let $\mathcal{G}$ be a DAG or CPDAG that is amenable relative to $(\mathbf{X},\mathbf{Y})$ and let $\mathbf{Y} \subseteq \text{PossDe}(\mathbf{X},\mathcal{G})$. By (i), it follows that if $\mathbf{X} \cap \text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G}) \neq \emptyset$, then there is no adjustment set relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}$. Hence, we only prove the converse statement.

If there is no adjustment set relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}$, then by the soundness of the generalized adjustment criterion (Theorem 5) there is no set that satisfies the generalized adjustment criterion relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}$. Since $\mathcal{G}$ is amenable relative to $(\mathbf{X},\mathbf{Y})$, this means that there is a path p from $X \in \mathbf{X}$ to $Y \in \mathbf{Y}$ in $\mathcal{G}$ that satisfies condition (2) in Theorem 26 i.e., p is a proper definite status non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$ such that every collider on p is in $\text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G})$ and every definite non-collider on p is in $\text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$.

We first note that, since $\mathbf{Y} \subseteq \text{PossDe}(\mathbf{X},\mathcal{G})$, $\mathbf{Y} \subseteq \text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$. We now show, by contradiction, that there is no collider on p. Thus, suppose that there is a collider C on p. Then C must be either a descendant of a non-collider on p or a descendant of both X and Y. Since $C \in \text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G})$, $C \notin \text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$. Moreover, every non-collider on p is in $\text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$ and since $\text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$ is a descendral set, C cannot be a descendant of a non-collider on p. Then p must be of the form $X \rightarrow C \leftarrow Y$. Since $Y \in \text{PossDe}(\mathbf{X},\mathcal{G})$ and $C \in \text{De}(Y,\mathcal{G})$, this contradicts that $C \notin \text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$.

Hence, p does not contain a collider. Additionally, p is a non-causal path from X to Y. This implies that there is a node A on p, $A \neq X$, such that $-p(A,X)$ is a directed path from A to X. Since A is either a non-collider on p, or $A = Y$, $A \in \text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$. Hence, $X \in \mathbf{X} \cap \text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$. ■

Proof of Corollary 28. Since there is no directed path from one node in $\mathbf{X}$ to another node in $\mathbf{X}$ in $\mathcal{D}$, it is easy to see that a path of the form $X \leftarrow V \ldots Y$, where $X \in \mathbf{X}, Y \in \mathbf{Y}$ and $V \in \text{De}(\mathbf{X},\mathcal{D})$ cannot occur in $\mathcal{D}$. Then there can be no path p from $\mathbf{X}$ to $\mathbf{Y}$ that satisfies condition (i) in Lemma 21 (condition (iv) in Lemma 25) in $\mathcal{D}$. Hence, conditions (3) and (4) in Theorem 26 are violated relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{D}$. Then by (i) and (iii) in Theorem 26 it follows that there exists a set that satisfies the generalized adjustment criterion relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{D}$ if and only if there exists a back-door set relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{D}$. ■

Proof of Corollary 29. It is enough to prove that if there exists an adjustment set relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}$ and the assumptions of the corollary hold, then there is no proper definite status non-causal path p_1 that satisfies (i)−(iv) in Lemma 25.

If $\mathcal{G}$ contains no possibly directed path $p = \langle V_1,\ldots,V_k\rangle$, with $k \geq 3, \{V_1,V_k\} \subseteq \mathbf{X}$ and $\{V_2,\ldots,V_{k-1}\} \cap \mathbf{X} = \emptyset$, it follows that there cannot be a node $V \in \text{PossDe}(\mathbf{X},\mathcal{G})$ and the path $X \leftarrow \cdots \leftarrow V$ in $\mathcal{G}$. So there cannot be a proper definite status non-causal path p_1 that satisfies (i)−(iv) in Lemma 25.

Next, suppose that $\mathcal{G}$ is a DAG or CPDAG and that $\mathbf{Y} \subseteq \text{PossDe}(\mathbf{X},\mathcal{G})$ and that there exists an adjustment set relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}$. By Corollary 15, it follows that $\text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G})$ then satisfies the blocking condition relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{G}$. Suppose for a contradiction that there is a proper definite status non-causal path p_1 that satisfies (i)−(iv) in Lemma 25. Then p_1 is of the form $X \leftarrow \cdots \leftarrow V \leftarrow Y$. By assumption $Y \in \text{PossDe}(\mathbf{X},\mathcal{G})$, so it follows that $Y \in \text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$. By the definition of the forbidden set, every other node on p is also in $\text{Forb}(\mathbf{X},\mathbf{Y},\mathcal{G})$. But then $\text{Adjust}(\mathbf{X},\mathbf{Y},\mathcal{G})$ does cannot block p. This is in contradiction with Adjust$(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ satisfying the blocking condition relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$. ■

Appendix E. Adjustment Criterion for DAGs

In this section we provide the soundness and completeness proof for the adjustment criterion from Shpitser (2012); van der Zander et al. (2014) (see Definition 55). The main result is given in Theorem 56.

This section can be read independently from the rest of the paper. Since we restrict our proof to DAGs, we first define adjustment sets (see Definition 54) and the adjustment criterion (see Definition 55) in DAGs. Thus, Definition 54 and Definition 55 are special cases of Definition 1 and Definition 4 for DAGs.

Definition 54 *(**Adjustment set**; Pearl, 2009, Chapter 3.3.1) Let* $\mathbf{X}, \mathbf{Y}$ *and* $\mathbf{Z}$ *be pairwise disjoint node sets in a causal* DAG $\mathcal{D}$*. Then* $\mathbf{Z}$ *is an adjustment set relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}$ *if for any density* f *consistent with* $\mathcal{D}$*:*

$$f(\mathbf{y} \mid do(\mathbf{x})) = \begin{cases} f(\mathbf{y} \mid \mathbf{x}) & \textit{if } \mathbf{Z} = \emptyset, \\ \int_{\mathbf{z}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z}) f(\mathbf{z}) d\mathbf{z} & \textit{otherwise.} \end{cases}$$

Definition 55 *(**Adjustment criterion**; cf. Shpitser, 2012, van der Zander et al., 2014) Let* $\mathbf{X}, \mathbf{Y}$ *and* $\mathbf{Z}$ *be pairwise disjoint node sets in a* DAG $\mathcal{D}$*. Let* Forb$(\mathbf{X}, \mathbf{Y}, \mathcal{D})$ *denote the set of all descendants in* $\mathcal{D}$ *of any* $W \notin \mathbf{X}$ *which lies on a proper causal path from* $\mathbf{X}$ *to* $\mathbf{Y}$ *in* $\mathcal{D}$*. Then* $\mathbf{Z}$ *satisfies the adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}$ *if the following two conditions hold:*

*(**Forbidden set**)* $\mathbf{Z} \cap \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) = \emptyset$*, and*

*(**Blocking**) all proper non-causal paths from* $\mathbf{X}$ *to* $\mathbf{Y}$ *in* $\mathcal{D}$ *are blocked by* $\mathbf{Z}$*.*

Definition 55 was introduced in van der Zander et al. (2014) and differs from the definition of the adjustment criterion in Shpitser (2012) in that it uses $\mathcal{D}$ in the forbidden set condition, as opposed to $\mathcal{D}_{\overline{\mathbf{X}}}$, where $\mathcal{D}_{\overline{\mathbf{X}}}$ is the graph obtained by removing all edges into $\mathbf{X}$ from $\mathcal{D}$. These two formulations of the adjustment criterion are equivalent (Remark 4.3 in van der Zander et al., 2014). We now give the main result in Theorem 56, which follows directly from Theorem 57 and Theorem 58. To prove Theorem 58 we rely on Lemma 59 and Lemma 60, which are given later in this section.

Theorem 56 *Let* $\mathbf{X}, \mathbf{Y}$ *and* $\mathbf{Z}$ *be pairwise disjoint node sets in a causal* DAG $\mathcal{D} = (\mathbf{V}, \mathbf{E})$*. Then* $\mathbf{Z}$ *satisfies the adjustment criterion (see Definition 55) if and only if* $\mathbf{Z}$ *is an adjustment set (see Definition 54).*

Theorem 57 *(**Completeness of the adjustment criterion for** DAG**s**) Let* $\mathbf{X}, \mathbf{Y}$ *and* $\mathbf{Z}$ *be pairwise disjoint node sets in a causal* DAG $\mathcal{D}$*. If* $\mathbf{Z}$ *does not satisfy the adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}$*, then there exists a density* f *consistent with* $\mathcal{D}$ *such that* $f(\mathbf{y} \mid do(\mathbf{x})) \neq \int_{\mathbf{z}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z}) f(\mathbf{z}) d\mathbf{z}$*.*

Proof of Theorem 57. Suppose that $\mathbf{Z}$ does not satisfy the adjustment criterion relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{D}=(\mathbf{V},\mathbf{E})$. It suffices to show that there is a density consistent with $\mathcal{D}$ such that $E[Y \mid do(\mathbf{X}=\mathbf{1})] \neq \int_z E[Y \mid \mathbf{X}=\mathbf{1},\mathbf{z}]f(\mathbf{z})d\mathbf{z}$ for at least one node $Y \in \mathbf{Y}$.

We consider multivariate Gaussian densities with mean vector zero, constructed using linear structural equation models (SEMs) with Gaussian noise. In particular, we let each random variable $A \in \mathbf{V}$ be a linear combination of its parents in $\mathcal{D}$ and a designated Gaussian noise variable ϵ_A with zero mean and a fixed variance. We also assume that the Gaussian noise variables $\{\epsilon_A : A \in \mathbf{V}\}$, are mutually independent. Thus, this model can be parameterized using one number per node (the residual variance) and one number per edge (the edge coefficient).

Since $\mathbf{Z}$ does not satisfy the adjustment criterion relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{D}$, $\mathbf{Z}$ violates the forbidden set condition or the blocking condition.

1 If $\mathbf{Z}$ violates the forbidden set condition, then there is a proper causal path $\langle X, V_1, \dots, V_k = Y\rangle$, $k \geq 1$, from $X \in \mathbf{X}$ to $Y \in \mathbf{Y}$ in $\mathcal{D}$ and a node $Z \in \mathbf{Z}$ such that

   (a) $Z = V_i$, for some $i \in \{1, \dots k-1\}$, or

   (b) $Z \in \mathrm{De}(\mathbf{Y}, \mathcal{D})$, or

   (c) $Z \in \mathrm{De}(V_i, \mathcal{D}) \setminus \{V_1, \dots, V_{k-1}\}$ for some $i \in \{1, \dots k-1\}$.

2 If $\mathbf{Z}$ violates the blocking condition, then there exists a proper non-causal path from $X \in \mathbf{X}$ to $Y \in \mathbf{Y}$ that is d-connecting given $\mathbf{Z}$ in $\mathcal{D}$ such that:

   (a) the path does not contain any colliders, or

   (b) the path contains at least one collider.

We now discuss these cases systematically.

(i) Suppose there is a path p from $X \in \mathbf{X}$ to $Y \in \mathbf{Y}$ that satisfies 2(a) in $\mathcal{D}$. Since p is a proper non-causal path that does not contain colliders, p starts with an edge into X, that is, p is of the form $X \leftarrow \dots Y$.

   We define our SEM so that all edge coefficients except for the ones on p are 0, and all edge coefficients on p are in $(0,1)$ and small enough so that we can choose the residual variances so that the variance of every random variable in $\mathbf{V}$ is 1. Then the density f generated by this SEM is consistent with the causal DAG $\mathcal{D}$. Moreover, f is also consistent with the causal DAG $\mathcal{D}'$ that is obtained from $\mathcal{D}$ by removing all edges except for the ones on p.

   Since $Y \perp_d \mathbf{X}$ in $\mathcal{D}'_{\overline{\mathbf{X}}}$, we use Rule 3 of the do-calculus (see Equation 6 in Appendix A), with $\mathbf{X}'=\emptyset, \mathbf{W}'=\emptyset$, $\mathbf{Z}'=\mathbf{X}$ and $\mathbf{Y}'=\{Y\}$, so that $E[Y \mid do(\mathbf{X}=\mathbf{1})] = E[Y] = 0$.

   Since p is proper, no node in $\mathbf{X}\setminus\{X\}$ is on p. Additionally, since p is d-connecting given $\mathbf{Z}$ and p does not contain colliders, no node in $\mathbf{Z}$ is on p. This implies $Y \perp_d \mathbf{Z} \cup \mathbf{X} \setminus \{X\}$ in $\mathcal{D}'$. Furthermore, $Y \perp_d \mathbf{Z} \cup \mathbf{X} \mid \mathbf{S}$ in $\mathcal{D}'$ for any subset $\mathbf{S}$ of the remaining nodes. In particular, we have $Y \perp_d \mathbf{Z} \cup \mathbf{X} \setminus \{X\} \mid X$ in $\mathcal{D}'$, so that $\int_{\mathbf{z}} E[Y \mid \mathbf{X}=\mathbf{1},\mathbf{z}]f(\mathbf{z})d\mathbf{z} = E[Y \mid X=1]$. By Theorem 32, $E[Y \mid X=1] = \mathrm{Cov}(X,Y)$. By Wright's rule (Theorem 31), $\mathrm{Cov}(X,Y)=a$, where a is the product of all edge coefficients on p. Since $\int_{\mathbf{z}} E[Y \mid \mathbf{X}=\mathbf{1},\mathbf{z}]f(\mathbf{z})d\mathbf{z} = a \neq 0$, this case is completed.

$$X \longrightarrow \overset{a}{\cdots\cdots} \longrightarrow Y \longrightarrow \overset{b}{\cdots\cdots} \longrightarrow Z$$

Figure 10: An example of path $p \oplus q$ in $\mathcal{D}$ corresponding to (iii), where $Z \in \mathbf{Z}$.

(ii) Suppose no path satisfies 2(a) in $\mathcal{D}$, but there is a path p from $X \in \mathbf{X}$ to $Y \in \mathbf{Y}$ that satisfies 1(a) in $\mathcal{D}$. Let $\tilde{\mathbf{Z}}$ be the set of all nodes in $\mathbf{Z}$ that are on p and let $Z \in \tilde{\mathbf{Z}}$.

We define our SEM so that all edge coefficients except the ones on p are 0, and all edge coefficients on p are in $(0, 1)$ and small enough so that we can choose the residual variances such that the variance of every random variable in $\mathbf{V}$ is 1. Then the density f generated by this SEM is consistent with the causal DAG $\mathcal{D}$, and also with the causal DAG $\mathcal{D}'$ that is obtained from $\mathcal{D}$ by removing all edges except for the ones on p.

Since no node from $\mathbf{X} \setminus \{X\}$ is on p, it follows that $Y \perp_d \mathbf{X} \setminus \{X\}$ in $\mathcal{D}'$. Furthermore, $Y \perp_d \mathbf{X} \setminus \{X\} \mid X$ in $\mathcal{D}'_{\overline{X}}$. We use Rule 3 of the do-calculus, with $\mathbf{X'} = \{X\}, \mathbf{W'} = \emptyset$, $\mathbf{Z'} = \mathbf{X} \setminus \{X\}$ and $\mathbf{Y'} = \{Y\}$, so that $E[Y \mid do(\mathbf{X} = \mathbf{1})] = E[Y \mid do(X = 1)]$. This means that we have reduced a joint intervention to a single intervention.

Additionally, since $Y \perp_d X$ in $\mathcal{D}'_{\underline{X}}$, we use Rule 2 of the do-calculus, with $\mathbf{X'} = \emptyset, \mathbf{W'} = \emptyset$, $\mathbf{Z'} = X$ and $\mathbf{Y'} = \{Y\}$, so that $E[Y \mid do(X = 1)] = E[Y \mid X = 1]$. By Theorem 32, $E[Y \mid X = 1] = \mathrm{Cov}(X, Y)$. Let a be the product of all edge coefficients on p. By Wright's rule (Theorem 31), we have that $\mathrm{Cov}(X, Y) = a \neq 0$.

We complete this case by showing that $\int_{\mathbf{z}} E[Y \mid \mathbf{X} = \mathbf{1}, \mathbf{z}] f(\mathbf{z}) d\mathbf{z} = 0$. Since p is proper, no node in $\mathbf{X} \setminus \{X\}$ is on p. Additionally, by our choice of $\tilde{\mathbf{Z}}$, no node in $\mathbf{Z} \setminus \tilde{\mathbf{Z}}$ is on p. Thus, $Y \perp_d \mathbf{X} \cup (\mathbf{Z} \setminus \tilde{\mathbf{Z}}) \mid \tilde{\mathbf{Z}}$ in $\mathcal{D}'$. Then $\int_{\mathbf{z}} E[Y \mid \mathbf{X} = \mathbf{1}, \mathbf{z}] f(\mathbf{z}) d\mathbf{z} = \int_{\tilde{\mathbf{z}}} E[Y \mid \tilde{\mathbf{z}}] f(\tilde{\mathbf{z}}) d\tilde{\mathbf{z}} = E[Y] = 0$.

(iii) Suppose no path satisfies 2(a) or 1(a), but there is a path p from $X \in \mathbf{X}$ to $Y \in \mathbf{Y}$ that satisfies 1(b) in $\mathcal{D}$. Choose $Z \in \mathbf{Z}$ such that the causal path q from Y to Z is a shortest causal path from Y to a node in $\mathbf{Z}$. Then Y is the only node that is on both p and q, otherwise there is a cycle in $\mathcal{D}$. Hence, $p \oplus q$ is a causal path from X to Z that contains Y (see Figure 10).

We define our SEM so that all edge coefficients except the ones on $p \oplus q$ are 0, and all edge coefficients on $p \oplus q$ are in $(0, 1)$ and small enough so that we can choose the residual variances such that the variance of every random variable in $\mathbf{V}$ is 1. Then the density f generated by this SEM is consistent with the causal DAG $\mathcal{D}$, as well as with the causal DAG $\mathcal{D}'$ that is obtained from $\mathcal{D}$ by removing all edges except the ones on $p \oplus q$.

Since there is no path that satisfies 2(a) in $\mathcal{D}$, no node from $\mathbf{X}$ is on q. Additionally, since p is proper, no node in $\mathbf{X} \setminus \{X\}$ is on $p \oplus q$. Thus, $Y \perp_d \mathbf{X} \setminus \{X\} \mid X$ in $\mathcal{D}'$. Furthermore, $Y \perp_d \mathbf{X} \setminus \{X\} \mid X$ in $\mathcal{D}'_{\overline{X}}$. Hence, we use Rule 3 of the do-calculus, with $\mathbf{X'} = \{X\}, \mathbf{W'} = \emptyset$, $\mathbf{Z'} = \mathbf{X} \setminus \{X\}$ and $\mathbf{Y'} = \{Y\}$, so that $E[Y \mid do(\mathbf{X} = \mathbf{1})] = E[Y \mid do(X = 1)]$.

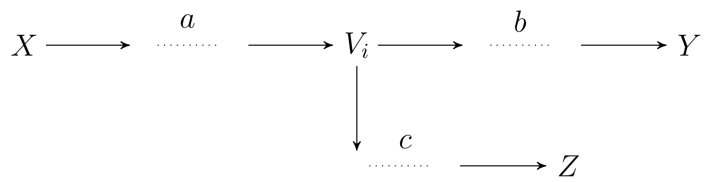

Figure 11: An example of paths p and q_i in $\mathcal{D}$ corresponding to (iv), where $Z \in \mathbf{Z}$.

Moreover, $Y \perp_d X$ in $\mathcal{D}'_{\underline{X}}$. Hence, we use Rule 2 of the do-calculus, with $\mathbf{X'} = \emptyset, \mathbf{W'} = \emptyset$, $\mathbf{Z'} = X$ and $\mathbf{Y'} = \{Y\}$, so that $E[Y \mid do(X=1)] = E[Y \mid X=1]$. Lastly, using Theorem 32 and Wright's rule (Theorem 31), we have that $E[Y \mid X=1] = \mathrm{Cov}(X,Y) = a$, where a is the product of all edge coefficients on p .

Next, we show that $\int_{\mathbf{z}} E[Y \mid \mathbf{X}=\mathbf{1}, \mathbf{z}] f(\mathbf{z}) d\mathbf{z} \neq a$. Since no path satisfies 1(a), no node from $\mathbf{Z}$ is on p. Furthermore, by the choice of q, no node from $\mathbf{Z} \setminus \{Z\}$ is on q. Hence, Z is the only node from $\mathbf{Z}$ that is on $p \oplus q$. From the above, we also know that X is the only node from $\mathbf{X}$ that is on $p \oplus q$. Hence, $Y \perp_d (\mathbf{X} \cup \mathbf{Z}) \setminus \{X, Z\} \mid \{X, Z\}$ in $\mathcal{D}'$ and we have that $\int_{\mathbf{z}} E[Y \mid \mathbf{X}=\mathbf{1}, \mathbf{z}] f(\mathbf{z}) d\mathbf{z} = \int_{z} E[Y \mid X=1, z] f(z) dz$.

Let b be the product of all edge coefficients on q. By Wright's rule (Theorem 31), we have that $\mathrm{Cov}(X,Y) = a, \mathrm{Cov}(Y,Z) = b$ and $\mathrm{Cov}(X,Z) = ab$. Now, we use Theorem 32 to calculate $E[Y \mid X=1, z]$:

$$E[Y \mid X=1, z] = \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} 1 & ab \\ ab & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ z \end{bmatrix} = \frac{a(1-b^2)}{1-(ab)^2} + \frac{b(1-a^2)}{1-(ab)^2} z.$$

$$\int_z E[Y \mid X=1, z] f(z) dz = a\frac{1-b^2}{1-(ab)^2} + \frac{b(1-a^2)}{1-(ab)^2} E[Z] = a\frac{1-b^2}{1-(ab)^2}. \tag{7}$$

Since $0 < a < 1$ and $0 < b < 1$, right-hand side of Equation (7) is strictly smaller than $a = E[Y \mid do(\mathbf{X}=\mathbf{1})]$.

(iv) Suppose no path satisfies 1(a), 1(b), or 2(a), but there is a path p from $X \in \mathbf{X}$ to $Y \in \mathbf{Y}$ that satisfies 1(c) in $\mathcal{D}$. Let V_i, $i \in \{1, \dots, k-1\}$, be a node on p that has a shortest causal path to a node in $\mathbf{Z}$. Let q_i be such a shortest causal path from V_i to $\mathbf{Z}$. Then no node except V_i is on both p and q_i, otherwise we would have chosen a different V_i (see Figure 11).

We define our SEM so that all edge coefficients which are not on p or q_i are 0, and all edge coefficients on p and q_i are in $(0,1)$ and small enough so that we can choose the residual variances so that the variance of every random variable in $\mathbf{V}$ is 1. Then the density f generated by this SEM is consistent with the causal DAG $\mathcal{D}$, as well as with the causal DAG $\mathcal{D}'$ that is obtained from $\mathcal{D}$ by removing all edges except for the ones on p and q_i.

Since there is no path that satisfies 2(a), no node from $\mathbf{X}$ is on q_i. Additionally, since p is proper, no node from $\mathbf{X} \setminus \{X\}$ is on p. Thus, $Y \perp_d \mathbf{X} \setminus \{X\} \mid X$ in $\mathcal{D}'_{\overline{X}}$ and we use Rule 3 of the do-calculus, with $\mathbf{X'} = \{X\}, \mathbf{W'} = \emptyset$, $\mathbf{Z'} = \mathbf{X} \setminus \{X\}$ and $\mathbf{Y'} = \{Y\}$, so that $E[Y \mid do(\mathbf{X} = \mathbf{1})] = E[Y \mid do(X = 1)]$.

Additionally, $Y \perp_d X$ in $\mathcal{D}'_{\underline{X}}$. Hence, we use Rule 2 of the do-calculus with $\mathbf{X'} = \emptyset, \mathbf{W'} = \emptyset$, $\mathbf{Z'} = \{X\}$ and $\mathbf{Y'} = \{Y\}$, so that $E[Y \mid do(X = 1)] = E[Y \mid X = 1]$. Let a be the product of all edge coefficients on $p(X, V_i)$ and let b be the product of all edge coefficients on $p(V_i, Y)$. Then, using Theorem 32 and Wright's rule (Theorem 31), we have that $E[Y \mid X = 1] = \mathrm{Cov}(X, Y) = ab$.

Since there is no path that satisfies 1(a), no node from $\mathbf{Z}$ is on p. By the choice of q_i, no node from $\mathbf{Z} \setminus \{Z\}$ is on q_i. Hence, no node from $(\mathbf{X} \cup \mathbf{Z}) \setminus \{X, Z\}$ is on p nor q_i. Then $Y \perp_d (\mathbf{X} \cup \mathbf{Z}) \setminus \{X, Z\} \mid \{X, Z\}$ in $\mathcal{D}'$ and it follows that $\int_{\mathbf{z}} E[Y \mid \mathbf{X} = \mathbf{1}, \mathbf{z}] f(\mathbf{z}) d\mathbf{z} = \int_z E[Y \mid X = 1, z] f(z) dz$.

Let c be the product of all edge coefficients on q_i. By Wright's rule (Theorem 31), we have that $\mathrm{Cov}(X, Y) = ab, \mathrm{Cov}(Y, Z) = bc$ and $\mathrm{Cov}(X, Z) = ac$. We can now use Theorem 32 to calculate $E[Y \mid X = 1, z]$:

$$E[Y \mid X = 1, z] = \begin{bmatrix} ab & bc \end{bmatrix} \begin{bmatrix} 1 & ac \\ ac & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ z \end{bmatrix} = \frac{ab(1 - c^2)}{1 - (ac)^2} + \frac{bc(1 - a^2)}{1 - (ac)^2} z.$$

Hence,

$$\int_z E[Y \mid X = 1, z] f(z) dz = \frac{ab(1 - c^2)}{1 - (ac)^2} + \frac{bc(1 - a^2)}{1 - (ac)^2} E[Z] = ab \frac{1 - c^2}{1 - (ac)^2}. \tag{8}$$

Since $0 < a < 1$, $0 < b < 1$ and $0 < c < 1$, right-hand side of Equation (8) is strictly smaller than $ab = E[Y \mid do(\mathbf{X} = \mathbf{1})]$.

(v) Suppose there is no path that satisfies 1(a), 1(b), 1(c), or 2(a), but there is a path that satisfies 2(b) in $\mathcal{D}$. Let p be such a path from $X \in \mathbf{X}$ to $Y \in \mathbf{Y}$ in $\mathcal{D}$ that contains the smallest number of colliders among all such paths. Since no path satisfies 2(a), there is at least one collider on p. Let $C_1, \dots, C_r$, $r \geq 1$, be all colliders on p ordered from the collider closest to X on p, which is C_1, to the collider closest to Y on p, which is C_r. Since p is d-connecting given $\mathbf{Z}$, we have $C_i \in \mathrm{An}(\mathbf{Z}, \mathcal{D})$ for all $i = 1, \dots, r$. . For each $i = 1, \dots, r$, let q_i be a shortest path (possibly of length zero) from C_i to $\mathbf{Z}$. Let $\tilde{\mathbf{Z}}$ be the collection of all nodes in $\mathbf{Z}$ that are enpoints of $q_1, \dots, q_r$.

We define our SEM so that all edge coefficients which are not on $p, q_1, \dots, q_r$ are 0, and all edge coefficients which are on $p, q_1, \dots, q_r$ are in $(0, 1)$ and are small enough so that we can choose the residual variances such that the variance of every random variable in $\mathbf{V}$ is 1. Then the density f generated by this SEM is consistent with the causal DAG $\mathcal{D}$, as well as with the causal DAG $\mathcal{D}'$ which is obtained from $\mathcal{D}$ by removing all edges except the ones on p, $q_1, \dots, q_r$. Moreover, f is a non-degenerate multivariate Gaussian density on $\mathbf{V}$.

Since there is no path that satisfies 2(a) in $\mathcal{D}$, no node from $\mathbf{X}$ is on q_r. Additionally, no node from $\mathbf{X}$ is on q_i, for any $i \in \{1, \dots, r - 1\}$, when $r > 1$, since otherwise there is a proper non-causal path p' from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{D}$ that is d-connecting given $\mathbf{Z}$ and

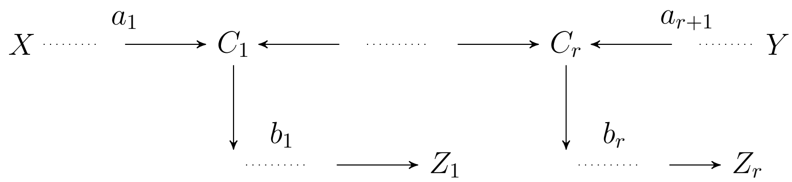

Figure 12: An example of paths p and $q_1, \dots, q_r$ in $\mathcal{D}$ corresponding to (v).

that contains a smaller number of colliders than p. Hence, X is the only node from $\mathbf{X}$ that is on $p, q_1, \dots, q_r$.

Since X is the only node from $\mathbf{X}$ that is on $p, q_1, \dots, q_r$, $Y \perp_d \mathbf{X} \setminus \{X\}$ in $\mathcal{D}'$. By assumption there is at least one collider on p. Since there is no path that satisfies 1(b) in $\mathcal{D}$, Y is not on q_1. Additionally, Y is not on q_i, for any $i \in \{2, \dots, r\}$, when $r > 1$, otherwise there is a proper non-causal path p' from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{D}$ that is d-connecting given $\mathbf{Z}$ and that contains a smaller number of colliders than p. Hence, $Y \perp_d \mathbf{X}$ in $\mathcal{D}'$. Furthermore, $Y \perp_d \mathbf{X}$ in $\mathcal{D}'_{\overline{X}}$, so we use Rule 3 of the do-calculus, with $\mathbf{X'} = \emptyset, \mathbf{W'} = \emptyset$, $\mathbf{Z'} = \mathbf{X}$ and $\mathbf{Y'} = \{Y\}$, so that $E[Y \mid do(\mathbf{X} = \mathbf{1})] = E[Y] = 0$.

By the choice of $p, q_1, \dots, q_r$, $\tilde{\mathbf{Z}}$ are the only nodes from $\mathbf{Z}$ that are on $p, q_1, \dots, q_r$. Then $\{X\} \cup \tilde{\mathbf{Z}}$ are the only nodes from $\mathbf{X} \cup \mathbf{Z}$ that are on $p, q_1, \dots, q_r$. Hence, $Y \perp_d (\mathbf{X} \cup \mathbf{Z}) \setminus (\{X\} \cup \tilde{\mathbf{Z}})$ in $\mathcal{D}'$. Furthermore, $Y \perp_d (\mathbf{X} \cup \mathbf{Z}) \setminus (\{X\} \cup \tilde{\mathbf{Z}}) \mid \{X\} \cup \tilde{\mathbf{Z}}$ in $\mathcal{D}'$. Hence, $\int_{\mathbf{z}} E[Y \mid \mathbf{X} = \mathbf{1}, \mathbf{z}] f(\mathbf{z}) d\mathbf{z} = \int_{\tilde{\mathbf{z}}} E[Y \mid X = 1, \tilde{\mathbf{z}}] f(\tilde{\mathbf{z}}) d\tilde{\mathbf{z}}$.

We now show $\int_{\tilde{\mathbf{z}}} E[Y \mid X = 1, \tilde{\mathbf{z}}] f(\tilde{\mathbf{z}}) d\tilde{\mathbf{z}} \neq 0$. For this we need the covariance matrix of $(X, \tilde{\mathbf{Z}}, Y)^T$, which we will compute by applying Wright's rule to $\mathcal{D}'$ (see Figure 12). In order to do this, we first need to show that no node on q_i is on q_j, for all $i, j \in \{1, \dots, r\}$ with $i \neq j$ and that no node on q_i except C_i is on p, for all $i \in \{1, \dots, r\}$. From this it will follow that each path q_i ends in a different node in $\tilde{\mathbf{Z}}$. We label these nodes as $\tilde{\mathbf{Z}} = (Z_1, \dots, Z_r)^T$ (see Figure 12).

We start by showing that no node on q_i, except C_i, is on p, for any $i \in \{1, \dots, r\}$. Suppose for a contradiction that for some $i \in \{1, \dots, r\}$ a node on q_i, other than C_i, is on p. Then q_i is at least of length 1, that is, $C_i \notin \mathbf{Z}$. Let D be the node closest to C_i on q_i that is also on p. Then D is either on $p(X, C_i)$ or on $p(C_i, Y)$.

Suppose first that D is on $p(X, C_i)$. Let $p' = p(X, D) \oplus (-q_i)(D, C_i) \oplus p(C_i, Y)$. From the above, we know that p' is a proper path from $\mathbf{X}$ to $\mathbf{Y}$. Since $(-q_i)(D, C_i)$ is of the form $D \leftarrow \dots \leftarrow C_i$, p' is a non-causal path from X to Y. By construction p' will have fewer colliders than p. Hence, p' is a proper non-causal path from X to Y with fewer colliders than p, since C_i is a non-collider on p'. Hence, in order to reach a contradiction, we only need to show that p' is d-connecting given $\mathbf{Z}$.

Since $p, q_1, \dots, q_r$ are d-connecting given $\mathbf{Z}$, we only need to discuss the collider/non-collider status of D and C_i on p'. Since $C_i \notin \mathbf{Z}$ and since C_i is a non-collider on p', we have that $p'(D, Y) = (-q_i)(D, C_i) \oplus p(C_i, Y)$ is d-connecting given $\mathbf{Z}$. Since D is on q_i, $D \in \mathrm{An}(\mathbf{Z}, \mathcal{D})$. So if D is a collider on p', p' is d-connecting given $\mathbf{Z}$. If D is a non-collider on p', then $(-p)(D, X)$ is out of D, so D is also a non-collider on p. Thus, in this case $D \notin \mathbf{Z}$ and hence, p' is d-connecting given $\mathbf{Z}$.

Next, suppose that D is on $p(C_i, Y)$. Let $p' = p(X, C_i) \oplus q_i(C_i, D) \oplus p(D, Y)$. From the above, we know that p' is a proper path from $\mathbf{X}$ to $\mathbf{Y}$. Since $D \in \text{An}(\mathbf{Z}, \mathcal{D})$ and since there is no path that satisfies 1(a), 1(b), or 1(c), p' cannot be a causal path. Thus, p' is a proper non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ that by construction has fewer colliders than p, since C_i is a non-collider on p'. Hence, in order to reach a contradiction, we only need to show that p' is d-connecting given $\mathbf{Z}$.

Since $p, q_1, \dots, q_r$ are d-connecting given $\mathbf{Z}$, we only need to discuss the collider/non-collider status of D and C_i on p'. Since $C_i \notin \mathbf{Z}$ and since C_i is a non-collider on p', we have that $p'(X, D) = p(X, C_i) \oplus q_i(C_i, D)$ is d-connecting given $\mathbf{Z}$. Similarly as above, $D \in \text{An}(\mathbf{Z}, \mathcal{D})$, so if D is a collider on p', p' is d-connecting given $\mathbf{Z}$. If D is a non-collider on p', then $p(D, Y)$ starts with an edge out of D, so D is also a non-collider on p. Thus, in this case $D \notin \mathbf{Z}$ and hence, p' is d-connecting given $\mathbf{Z}$.

Thus, we have shown that no node on q_i, other than C_i, is on p, for all $i \in \{1, \dots, r\}$. Next, we consider the case $r > 1$ and show, by contradiction, that no node on q_i is on q_j, for any $i, j \in \{1, \dots, r\}$ with $i \neq j$. Hence, suppose that there are q_i and q_j such that a node on q_i is also on q_j, for some $i < j$. Let D be the node closest to C_i on q_i, that is also on q_j. Note that from the above, $D \neq C_j$ and $D \neq C_i$, so that q_i and q_j are at least of length 1. Then let $p' = p(X, C_i) \oplus q_i(C_i, D) \oplus (-q_j)(D, C_j) \oplus p(C_j, Y)$. As discussed above, no node on q_i (or q_j) is in $\mathbf{X}$. Hence, p' is a proper path from $\mathbf{X}$ to $\mathbf{Y}$. Since $(-q_j)(D, C_j)$ is of the form $D \leftarrow \dots \leftarrow C_j$, p' is also a non-causal path from X to Y. Thus p' is a proper non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ that by construction has fewer colliders than p, since C_i and C_j are non-colliders on p'. Hence, in order to reach a contradiction we only need to show that it is d-connecting given $\mathbf{Z}$.

Since $p, q_1, \dots, q_r$ are d-connecting given $\mathbf{Z}$, we only need to discuss the collider/non-collider status of D, C_i and C_j on p'. Since C_i and C_j are non-colliders on p', we have that $p'(X, D)$ and $p'(D, Y)$ are both d-connecting given $\mathbf{Z}$. Since D is on q_i, $D \in \text{An}(\mathbf{Z}, \mathcal{D})$. Since D is a collider on p', p' is d-connecting given $\mathbf{Z}$.

We have now established that $\mathcal{D}'$ looks like Figure 12, where none of the paths intersect, and we can compute the covariance matrix of $(X, \tilde{\mathbf{Z}}^T, Y)^T$ using Wright's rule (Theorem 31) on $\mathcal{D}'$. For this purpose, let a_1 and a_{r+1} be the products of all edge coefficients on $p(X, C_1)$ and $p(C_r, Y)$, respectively. Let b_i, $i \in \{1, \dots, r\}$ be product of all edge coefficients on q_i. If $r > 1$, let a_j, $j \in \{2, \dots, r\}$ be the product of all edge coefficients on $p(C_{j-1}, C_j)$. Let $\boldsymbol{\Sigma}$ be the covariance matrix of $(X, \tilde{\mathbf{Z}}^T, Y)^T$. Then using Wright's rule (Theorem 31) on $\mathcal{D}'$ yields:

$$\boldsymbol{\Sigma} = \begin{array}{c} \begin{matrix} 1 & 2 & 3 & \cdots & r+1 & r+2 \end{matrix} \\ \begin{bmatrix} 1 & a_1b_1 & & & & \\ a_1b_1 & 1 & b_1a_2b_2 & & & \\ & b_1a_2b_2 & 1 & \ddots & & \mathbf{0} \\ & & \ddots & \ddots & \ddots & \\ & \mathbf{0} & & 1 & b_{r-1}a_rb_r & \\ & & & b_{r-1}a_rb_r & 1 & b_ra_{r+1} \\ & & & & b_ra_{r+1} & 1 \end{bmatrix} \end{array} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix},$$

where $\mathbf{\Sigma}_{11}$ is the covariance matrix of $(X, \tilde{\mathbf{Z}}^T)^T$, $\mathbf{\Sigma}_{22} = 1$, $\mathbf{\Sigma}_{21} = \mathbf{\Sigma}_{12}^T$ and $\mathbf{\Sigma}_{21} = \begin{bmatrix} 0 & \cdots & 0 & b_r a_{r+1} \end{bmatrix}$, $\mathbf{\Sigma}_{21} \in \mathbb{R}^{1\times(r+1)}$. Then using Theorem 32 we have that

$$E[Y \mid X = 1, \tilde{\mathbf{z}}] = \begin{bmatrix} 0 & \dots & 0 & b_r a_{r+1} \end{bmatrix} \mathbf{\Sigma}_{11}^{-1} \begin{bmatrix} 1 \\ \tilde{\mathbf{z}} \end{bmatrix}.$$

Since our multivariate Gaussian distribution is non-degenerate, $\mathrm{Det}(\mathbf{\Sigma}_{11}) > 0$. Let $t_{i,j}$ denote $(i,j)^{th}$ element of $\mathbf{\Sigma}_{11}^{-1}$, $\mathbf{\Sigma}_{11}^{-1} \in \mathbb{R}^{(r+1)\times(r+1)}$. Putting everything together, we have that

$$\begin{aligned}
\int_{\mathbf{z}} E[Y|\mathbf{X} = \mathbf{1}, \mathbf{z}] f(\mathbf{z}) d\mathbf{z} &= \int_{\tilde{\mathbf{z}}} E[Y|X = 1, \tilde{\mathbf{z}}] f(\tilde{\mathbf{z}}) d\tilde{\mathbf{z}} \\
&= \int_{\tilde{\mathbf{z}}} \begin{bmatrix} 0 & \dots & 0 & b_r a_{r+1} \end{bmatrix} \begin{bmatrix} t_{1,1} & \dots & t_{1,r+1} \\ \vdots & \ddots & \vdots \\ t_{r+1,1} & \dots & t_{r+1,r+1} \end{bmatrix} \begin{bmatrix} 1 \\ \tilde{\mathbf{z}} \end{bmatrix} f(\tilde{\mathbf{z}}) d\tilde{\mathbf{z}} \\
&= b_r a_{r+1} t_{r+1,1} + \sum_{i=1}^{r} b_r a_{r+1} t_{r+1,i+1} E[Z_i] = b_r a_{r+1} t_{r+1,1}.
\end{aligned}$$

Since $b_r \neq 0$ and $a_{r+1} \neq 0$, it is only left to show that $t_{r+1,1} \neq 0$. Using standard linear algebra, we find

$$\begin{aligned}
t_{r+1,1} &= \frac{(-1)^r}{\mathrm{Det}(\mathbf{\Sigma}_{11})} \begin{vmatrix}
a_1b_1 & 1 & b_1a_2b_2 & & & \\
 & b_1a_2b_2 & 1 & b_2a_3b_3 & & \mathbf{0} \\
 & & \ddots & \ddots & \ddots & \\
 & \mathbf{0} & & b_{r-3}a_{r-2}b_{r-2} & 1 & b_{r-2}a_{r-1}b_{r-1} \\
 & & & & b_{r-2}a_{r-1}b_{r-1} & 1 \\
 & & & & & b_{r-1}a_rb_r
\end{vmatrix} \\
&= \frac{(-1)^r a_r b_r}{\mathrm{Det}(\mathbf{\Sigma}_{11})} \prod_{i=1}^{r-1} a_i b_i^2.
\end{aligned}$$

Since $a_i \neq 0$ and $b_i \neq 0$ for all $i \in \{1, \dots, r\}$, the proof is completed.

■

Theorem 58 ***(Soundness of the adjustment criterion for*** DAG***s)*** *Let* $\mathbf{X}, \mathbf{Y}$ *and* $\mathbf{Z_0}$ *be pairwise disjoint node sets in a causal* DAG $\mathcal{D}$. *If* $\mathbf{Z_0}$ *satisfies the adjustment criterion (see Definition 55), then* $\mathbf{Z_0}$ *is an adjustment set (see Definition 54).*

Proof of Theorem 58. Assume that $\mathbf{Z_0}$ satisfies the adjustment criterion (see Definition 55) relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ and let f be a density consistent with $\mathcal{D}$. We need to show that

$$f(\mathbf{y} \mid \mathrm{do}(\mathbf{x})) = \int_{\mathbf{z_0}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}) f(\mathbf{z_0}) d\mathbf{z_0}. \tag{9}$$

Let $\mathbf{Y_D} = \mathbf{Y} \cap \mathrm{De}(\mathbf{X}, \mathcal{D})$ and $\mathbf{Y_N} = \mathbf{Y} \setminus \mathrm{De}(\mathbf{X}, \mathcal{D})$. Then $\mathbf{Y_N} \perp_d \mathbf{X}$ in $\mathcal{D}_{\overline{\mathbf{X}}}$, since $\mathcal{D}_{\overline{\mathbf{X}}}$ does not contain paths into $\mathbf{X}$ and all paths from $\mathbf{X}$ to $\mathbf{Y_N}$ that are out of $\mathbf{X}$ in $\mathcal{D}_{\overline{\mathbf{X}}}$ must contain a collider by definition of $\mathbf{Y_N}$. Hence, using Rule 3 of the do-calculus (see Equation 6 in Appendix A), with $\mathbf{X'} = \emptyset, \mathbf{W'} = \emptyset$, $\mathbf{Z'} = \mathbf{X}$ and $\mathbf{Y'} = \mathbf{Y_N}$ we have

$$f(\mathbf{y_N}) = f(\mathbf{y_N} \mid \mathrm{do}(\mathbf{x})). \tag{10}$$

Now, assume that $\mathbf{Y_D} = \emptyset$, so that $\mathbf{Y} = \mathbf{Y_N}$. Then all paths from $\mathbf{X}$ to $\mathbf{Y}$ are non-causal in $\mathcal{D}$. Since $\mathbf{Z_0}$ satisfies the adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ it blocks all proper non-causal paths from $\mathbf{X}$ to $\mathbf{Y}$. Thus, $\mathbf{Z_0}$ also blocks all non-causal paths from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{D}$, so $\mathbf{X} \perp_d \mathbf{Y} \mid \mathbf{Z_0}$ in $\mathcal{D}$. Combining the probabilistic implications of d-separation and Equation (10) with the right-hand side of Equation (9) we obtain

$$\begin{aligned} \int_{\mathbf{z_0}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}) f(\mathbf{z_0}) d\mathbf{z_0} &= \int_{\mathbf{z_0}} f(\mathbf{y} \mid \mathbf{z_0}) f(\mathbf{z_0}) d\mathbf{z_0} \\ &= f(\mathbf{y}) = f(\mathbf{y} \mid \mathrm{do}(\mathbf{x})). \end{aligned} \tag{11}$$

For the remainder of the proof we assume $\mathbf{Y_D} \neq \emptyset$. We enlarge the set $\mathbf{Z_0}$ to $\mathbf{Z} = \mathbf{Z_0} \cup \mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D}) \setminus (\mathrm{De}(\mathbf{X}, \mathcal{D}) \cup \mathbf{Y})$. Then applying (iii) in Lemma 59 to the right-hand side of Equation (9) we get

$$\int_{\mathbf{z_0}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}) f(\mathbf{z_0}) d\mathbf{z_0} = \int_{\mathbf{z}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z}) f(\mathbf{z}) d\mathbf{z}. \tag{12}$$

Let $\mathbf{Z_D} = \mathbf{Z} \cap \mathrm{De}(\mathbf{X}, \mathcal{D})$ and $\mathbf{Z_N} = \mathbf{Z} \setminus \mathrm{De}(\mathbf{X}, \mathcal{D})$. Now suppose that $\mathbf{Y} = \mathbf{Y_D}$, so that $\mathbf{Y_N} = \emptyset$. From (iii) in Lemma 60, we have $\mathbf{Y_D} \perp_d \mathbf{Z_D} \mid \mathbf{X} \cup \mathbf{Z_N}$. Using the probabilistic implications of d-separation, the right-hand side of Equation (12) equals

$$\begin{aligned} \int_{\mathbf{z_D}, \mathbf{z_N}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_D}, \mathbf{z_N}) f(\mathbf{z_D}, \mathbf{z_N}) d\mathbf{z_D} d\mathbf{z_N} &= \int_{\mathbf{z_D}, \mathbf{z_N}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_N}) f(\mathbf{z_D}, \mathbf{z_N}) d\mathbf{z_D} d\mathbf{z_N} \\ &= \int_{\mathbf{z_N}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_N}) f(\mathbf{z_N}) d\mathbf{z_N}. \end{aligned} \tag{13}$$

Since $\mathbf{Z_N}$ satisfies the generalized back-door criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ (see (iv) in Lemma 60) and since the generalized back-door criterion is sound by Theorem 3.1 in Maathuis and Colombo (2015), we have

$$\int_{\mathbf{z_N}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_N}) f(\mathbf{z_N}) d\mathbf{z_N} = f(\mathbf{y} \mid \mathrm{do}(\mathbf{x})). \tag{14}$$

Combining Equations (12), (13) and (14) completes the proof when $\mathbf{Y_N} = \emptyset$. In the remainder of the proof we assume $\mathbf{Y_D} \neq \emptyset$ and $\mathbf{Y_N} \neq \emptyset$. From (iii) in Lemma 60,

$\mathbf{Y_D} \perp_d \mathbf{Z_D} \mid \mathbf{Y_N} \cup \mathbf{X} \cup \mathbf{Z_N}$. Using the probabilistic implications of d-separation, the right-hand side of Equation (12) equals

$$\begin{aligned}
&\int_{\mathbf{z_D},\mathbf{z_N}} f(\mathbf{y_D}, \mathbf{y_N} \mid \mathbf{x}, \mathbf{z_D}, \mathbf{z_N}) f(\mathbf{z_D}, \mathbf{z_N}) d\mathbf{z_D} d\mathbf{z_N} \\
&= \int_{\mathbf{z_D},\mathbf{z_N}} f(\mathbf{y_D} \mid \mathbf{y_N}, \mathbf{x}, \mathbf{z_D}, \mathbf{z_N}) f(\mathbf{y_N} \mid \mathbf{x}, \mathbf{z_D}, \mathbf{z_N}) f(\mathbf{z_D}, \mathbf{z_N}) d\mathbf{z_D} d\mathbf{z_N} \\
&= \int_{\mathbf{z_N}} f(\mathbf{y_D} \mid \mathbf{y_N}, \mathbf{x}, \mathbf{z_N}) \int_{\mathbf{z_D}} f(\mathbf{y_N} \mid \mathbf{x}, \mathbf{z_D}, \mathbf{z_N}) f(\mathbf{z_D}, \mathbf{z_N}) d\mathbf{z_D} d\mathbf{z_N}. \qquad (15)
\end{aligned}$$

Since $\mathbf{Z}$ satisfies the adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ ((ii) in Lemma 59), $\mathbf{Z}$ blocks all proper non-causal paths from $\mathbf{X}$ to $\mathbf{Y}$. Hence, $\mathbf{Z}$ blocks all proper paths from $\mathbf{X}$ to $\mathbf{Y_N}$ in $\mathcal{D}$, so $\mathbf{X} \perp_d \mathbf{Y_N} \mid \mathbf{Z}$ in $\mathcal{D}$. Additionally, the empty set satisfies the generalized back-door criterion relative to $((\mathbf{Y_N} \cup \mathbf{X} \cup \mathbf{Z_N}), \mathbf{Y_D})$ ((v) in Lemma 60). Since the generalized back-door criterion is sound by Theorem 3.1 in Maathuis and Colombo (2015), we apply this to the right-hand side of Equation (15)

$$\begin{aligned}
&\int_{\mathbf{z_N}} f(\mathbf{y_D} \mid \mathbf{y_N}, \mathbf{x}, \mathbf{z_N}) \int_{\mathbf{z_D}} f(\mathbf{y_N} \mid \mathbf{x}, \mathbf{z_D}, \mathbf{z_N}) f(\mathbf{z_D}, \mathbf{z_N}) d\mathbf{z_D} d\mathbf{z_N} \\
&= \int_{\mathbf{z_N}} f(\mathbf{y_D} \mid \mathrm{do}(\mathbf{x}, \mathbf{y_N}, \mathbf{z_N})) \int_{\mathbf{z_D}} f(\mathbf{y_N} \mid \mathbf{z_N}, \mathbf{z_D}) f(\mathbf{z_D}, \mathbf{z_N}) d\mathbf{z_D} d\mathbf{z_N} \\
&= \int_{\mathbf{z_N}} f(\mathbf{y_D} \mid \mathrm{do}(\mathbf{x}, \mathbf{y_N}, \mathbf{z_N})) f(\mathbf{z_N} \mid \mathbf{y_N}) f(\mathbf{y_N}) d\mathbf{z_N}. \qquad (16)
\end{aligned}$$

To finish the proof we rely on the do-calculus rules. By (vi) in Lemma 60, it follows that $\mathbf{Y_N} \cup \mathbf{Z_N} \perp_d \mathbf{Y_D} \mid \mathbf{X}$ in $\mathcal{D}_{\overline{\mathbf{X}}\underline{\mathbf{Y_N} \cup \mathbf{Z_N}}}$. Using Rule 2 of the do-calculus (see Equation 5 in Appendix A) with $\mathbf{X}' = \mathbf{X}, \mathbf{W}' = \emptyset$, $\mathbf{Z}' = \mathbf{Y_N} \cup \mathbf{Z_N}$ and $\mathbf{Y}' = \mathbf{Y_D}$:

$$f(\mathbf{y_D} \mid \mathrm{do}(\mathbf{x}, \mathbf{y_N}, \mathbf{z_N})) = f(\mathbf{y_D} \mid \mathrm{do}(\mathbf{x}), \mathbf{z_N}, \mathbf{y_N}). \qquad (17)$$

Additionally, by (vii) in Lemma 60, $\mathbf{Z_N} \perp_d \mathbf{X} \mid \mathbf{Y_N}$ in $\mathcal{D}_{\overline{\mathbf{X}}}$. Using Rule 3 of the do-calculus (see Equation 6 in Appendix A) with $\mathbf{X}' = \emptyset, \mathbf{W}' = \mathbf{Y_N}$, $\mathbf{Z}' = \mathbf{X}$ and $\mathbf{Y}' = \mathbf{Z_N}$:

$$f(\mathbf{z_N} \mid \mathbf{y_N}) = f(\mathbf{z_N} \mid \mathrm{do}(\mathbf{x}), \mathbf{y_N}). \qquad (18)$$

Finally, we combine Equations (17),(18) and (10) with the right-hand side of Equation (16):

$$\begin{aligned}
&\int_{\mathbf{z_N}} f(\mathbf{y_D} \mid \mathrm{do}(\mathbf{y_N}, \mathbf{x}, \mathbf{z_N})) f(\mathbf{z_N} \mid \mathbf{y_N}) f(\mathbf{y_N}) d\mathbf{z_N} \\
&= \int_{\mathbf{z_N}} f(\mathbf{y_D} \mid \mathbf{z_N}, \mathbf{y_N}, \mathrm{do}(\mathbf{x})) f(\mathbf{z_N} \mid \mathbf{y_N}, \mathrm{do}(\mathbf{x})) f(\mathbf{y_N} \mid \mathrm{do}(\mathbf{x})) d\mathbf{z_N} \\
&= \int_{\mathbf{z_N}} f(\mathbf{y_D}, \mathbf{z_N} \mid \mathbf{y_N}, \mathrm{do}(\mathbf{x})) f(\mathbf{y_N} \mid \mathrm{do}(\mathbf{x})) d\mathbf{z_N} \\
&= f(\mathbf{y_D} \mid \mathbf{y_N}, \mathrm{do}(\mathbf{x})) f(\mathbf{y_N} \mid \mathrm{do}(\mathbf{x})) = f(\mathbf{y_D}, \mathbf{y_N} \mid \mathrm{do}(\mathbf{x})) = f(\mathbf{y} \mid \mathrm{do}(\mathbf{x})).
\end{aligned}$$

■

Lemma 59 *Let* $\mathbf{X}, \mathbf{Y}$ *and* $\mathbf{Z_0}$ *be pairwise disjoint node sets in a* DAG $\mathcal{D}$ *such that* $\mathbf{Z_0}$ *satisfies the adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}$. *Let* $\mathbf{Z_1} \subseteq \mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D}) \setminus (\mathrm{De}(\mathbf{X}, \mathcal{D}) \cup \mathbf{Y})$ *and* $\mathbf{Z} = \mathbf{Z_0} \cup \mathbf{Z_1}$. *Then:*

(i) $\mathbf{X}, \mathbf{Y}$ *and* $\mathbf{Z}$ *are pairwise disjoint, and*

(ii) $\mathbf{Z}$ *satisfies the adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}$*, and*

(iii) $\int_{\mathbf{z_0}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}) f(\mathbf{z_0}) d\mathbf{z_0} = \int_{\mathbf{z_0}, \mathbf{z_1}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}, \mathbf{z_1}) f(\mathbf{z_0}, \mathbf{z_1}) d\mathbf{z_0} d\mathbf{z_1}$*, for any density* f *consistent with* $\mathcal{D}$.

Proof of Lemma 59. (i) Since $\mathbf{X}, \mathbf{Y}$ and $\mathbf{Z_0}$ are pairwise disjoint, and $\mathbf{Z} = \mathbf{Z_0} \cup \mathbf{Z_1}$, where $\mathbf{Z_1} \cap (\mathbf{X} \cup \mathbf{Y}) = \emptyset$, it follows that $\mathbf{X}, \mathbf{Y}$ and $\mathbf{Z}$ are also pairwise disjoint.

(ii) A set that satisfies the forbidden set condition of the adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ satisfies the blocking condition if and only if it d-separates $\mathbf{X}$ and $\mathbf{Y}$ in the proper back-door graph $\mathcal{D}^{pbd}_{\mathbf{XY}}$ (van der Zander et al. (2014, Theorem 4.6); see Theorem 7 in Section 3). Thus, $\mathbf{Z_0}$ d-separates $\mathbf{X}$ and $\mathbf{Y}$ in $\mathcal{D}^{pbd}_{\mathbf{XY}}$. Then by Theorem 35, any path between $\mathbf{X}$ and $\mathbf{Y}$ in $((\mathcal{D}^{pbd}_{\mathbf{XY}})_{\mathrm{An}(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z_0}, \mathcal{D}^{pbd}_{\mathbf{XY}})})^m$ contains a node in $\mathbf{Z_0}$.

Since $\mathbf{Z_1} \cap \mathrm{De}(\mathbf{X}, \mathcal{D}) = \emptyset$, $\mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) \subseteq \mathrm{De}(\mathbf{X}, \mathcal{D})$ and $\mathbf{Z_0} \cap \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) = \emptyset$, $\mathbf{Z} \cap \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) = \emptyset$. Additionally, since $\mathbf{Z} \supseteq \mathbf{Z_0}$, all paths between $\mathbf{X}$ and $\mathbf{Y}$ in $((\mathcal{D}^{pbd}_{\mathbf{XY}})_{\mathrm{An}(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z_0}, \mathcal{D}^{pbd}_{\mathbf{XY}})})^m$ contain a node in $\mathbf{Z}$. Furthermore, $\mathbf{Z_1} \subseteq \mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D})$ implies that $\mathrm{An}(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z}, \mathcal{D}) = \mathrm{An}(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z_0}, \mathcal{D})$. Thus, all paths between $\mathbf{X}$ and $\mathbf{Y}$ in $((\mathcal{D}^{pbd}_{\mathbf{XY}})_{\mathrm{An}(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z}, \mathcal{D}^{pbd}_{\mathbf{XY}})})^m$ contain a node in $\mathbf{Z}$. Hence, $\mathbf{Z}$ satisfies the blocking condition relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ (Theorem 35, Theorem 7).

(iii) We prove this statement by induction on the number of nodes in $\mathbf{Z_1}$. Below, we prove the base case: $|\mathbf{Z_1}| = 1$. We then assume that the result holds for $|\mathbf{Z_1}| = k$, and show that it holds for $|\mathbf{Z_1}| = k + 1$. Thus, let $|\mathbf{Z_1}| = k + 1$ and take an arbitrary $Z_1 \in \mathbf{Z_1}$. Let $\mathbf{Z'_0} = \mathbf{Z_0} \cup \{Z_1\}$ and $\mathbf{Z'_1} = \mathbf{Z_1} \setminus \{Z_1\}$. The base case then implies

$$\begin{aligned} \int_{\mathbf{z_0}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}) f(\mathbf{z_0}) d\mathbf{z_0} &= \int_{\mathbf{z_0}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}) \int_{z_1} f(\mathbf{z_0}, z_1) dz_1 d\mathbf{z_0} \\ &= \int_{\mathbf{z_0}, z_1} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}, z_1) f(\mathbf{z_0}, z_1) d\mathbf{z_0} dz_1 \\ &= \int_{\mathbf{z'_0}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z'_0}) f(\mathbf{z'_0}) d\mathbf{z'_0}. \end{aligned} \tag{19}$$

By (ii) above $\mathbf{Z'_0}$ satisfies the adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$ and $\mathbf{Z'_1} \subseteq \mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D})$. Then since $|\mathbf{Z'_1}| = k$, by the induction hypothesis

$$\int_{\mathbf{z'_0}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z'_0}) f(\mathbf{z'_0}) d\mathbf{z'_0} = \int_{\mathbf{z'_0}, \mathbf{z'_1}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z'_0}, \mathbf{z'_1}) f(\mathbf{z'_0}, \mathbf{z'_1}) d\mathbf{z'_0} d\mathbf{z'_1}. \tag{20}$$

Combining 19 and 20 yields

$$\int_{\mathbf{z_0}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}) f(\mathbf{z_0}) d\mathbf{z_0} = \int_{\mathbf{z_0}, \mathbf{z_1}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}, \mathbf{z_1}) f(\mathbf{z_0}, \mathbf{z_1}) d\mathbf{z_0} d\mathbf{z_1}.$$

It is left to prove the base case of the induction. Hence, suppose $\mathbf{Z_1} = \{Z_1\}$. We show below that either (a) $\mathbf{Y} \perp_d Z_1 \mid \mathbf{X} \cup \mathbf{Z_0}$ or (b) $\mathbf{X} \perp_d Z_1 \mid \mathbf{Z_0}$ are satisfied in $\mathcal{D}$. (Note that (a) and (b) are very similar to the conditions U1 and U2, as well as U1* and U2*, from Greenland et al., 1999. These conditions are also used in Theorem 5 from Kuroki and Miyakawa, 2003, Lemma 3 from Kuroki and Cai, 2004 and are the foundation for the results in De Luna et al., 2011. Additionally, Pearl and Paz, 2014 give a discussion of these conditions, which they refer to as c-equivalence conditions, in their Theorem 1.)

If (a) $\mathbf{Y} \perp_d Z_1 \mid \mathbf{X} \cup \mathbf{Z_0}$ in $\mathcal{D}$, then by the probabilistic implications of d-separation we have that for any density f consistent with $\mathcal{D}$

$$\begin{aligned}\int_{\mathbf{z_0}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}) f(\mathbf{z_0}) d\mathbf{z_0} &= \int_{\mathbf{z_0}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}) \int_{z_1} f(\mathbf{z_0}, z_1) dz_1 d\mathbf{z_0} \\ &= \int_{\mathbf{z_0}, z_1} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}) f(\mathbf{z_0}, z_1) d\mathbf{z_0} dz_1 \\ &= \int_{\mathbf{z_0}, z_1} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}, z_1) f(\mathbf{z_0}, z_1) d\mathbf{z_0} dz_1.\end{aligned}$$

If (b) $\mathbf{X} \perp_d Z_1 \mid \mathbf{Z_0}$ in $\mathcal{D}$, then similarly

$$\begin{aligned}\int_{\mathbf{z_0}} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}) f(\mathbf{z_0}) d\mathbf{z_0} &= \int_{\mathbf{z_0}} f(\mathbf{z_0}) \int_{z_1} f(\mathbf{y}, z_1 \mid \mathbf{x}, \mathbf{z_0}) dz_1 d\mathbf{z_0} \\ &= \int_{\mathbf{z_0}, z_1} f(\mathbf{y}, z_1 \mid \mathbf{x}, \mathbf{z_0}) f(\mathbf{z_0}) d\mathbf{z_0} dz_1 \\ &= \int_{\mathbf{z_0}, z_1} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}, z_1) f(z_1 \mid \mathbf{x}, \mathbf{z_0}) f(\mathbf{z_0}) d\mathbf{z_0} dz_1 \\ &= \int_{\mathbf{z_0}, z_1} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}, z_1) f(z_1 \mid \mathbf{z_0}) f(\mathbf{z_0}) d\mathbf{z_0} dz_1 \\ &= \int_{\mathbf{z_0}, z_1} f(\mathbf{y} \mid \mathbf{x}, \mathbf{z_0}, z_1) f(z_1, \mathbf{z_0}) d\mathbf{z_0} dz_1.\end{aligned}$$

We complete the proof by showing that (a) or (b) must hold. Suppose for a contradiction that both (a) and (b) are violated. Then there is a path from $\mathbf{X}$ to Z_1 that is d-connecting given $\mathbf{Z_0}$ and a path from $\mathbf{Y}$ to Z_1 that is d-connecting given $\mathbf{X} \cup \mathbf{Z_0}$. Let p be a proper path from $X \in \mathbf{X}$ to Z_1 that is d-connecting given $\mathbf{Z_0}$ in $\mathcal{D}$ and let q be a path from Z_1 to $Y \in \mathbf{Y}$ that is d-connecting given $\mathbf{X} \cup \mathbf{Z_0}$ in $\mathcal{D}$. We will show that this contradicts that $\mathbf{Z_0}$ satisfies the adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$.

We first show, by contradiction, that q also is d-connecting given $\mathbf{Z_0}$. Thus, assume that $q = \langle Z_1, \dots, Y \rangle$ is blocked by $\mathbf{Z_0}$. Since q is d-connecting given $\mathbf{X} \cup \mathbf{Z_0}$ and blocked by $\mathbf{Z_0}$, it must contain at least one collider in $\text{An}(\mathbf{X}, \mathcal{D}) \setminus \text{An}(\mathbf{Z_0}, \mathcal{D})$. Let C be the collider closest to Y on q such that $C \in \text{An}(\mathbf{X}, \mathcal{D}) \setminus \text{An}(\mathbf{Z_0}, \mathcal{D})$. Let $r = \langle C, \dots, X' \rangle, X' \in \mathbf{X}$ be a shortest directed path from C to $\mathbf{X}$ (possibly of zero length). Then no node on $q(C, Y)$ or r, except possibly C, is in $\mathbf{X}$. We now concatenate the paths $-r(X', C)$ and $q(C, Y)$, while taking out possible loops. Hence, let V be the node closest to X' on r that is also on $q(C, Y)$. Then $-r(X', V) \oplus q(V, Y)$ is non-causal since either $-r(X', V)$ is of non-zero length, or $X' = V = C$ and $q(C, Y)$ is a path into C, because C is a collider on q. By the choice of C and r, $-r(X', V) \oplus q(V, Y)$ is a proper path that is d-connecting given $\mathbf{Z_0}$. This contradicts that $\mathbf{Z_0}$ satisfies the adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$. Thus, q is also d-connecting given $\mathbf{Z_0}$.

Let $\tilde{p}$ and $\tilde{q}$ be paths in the proper back-door graph $\mathcal{D}^{pbd}_{\mathbf{XY}}$ constituted by the same sequences of nodes as p and q in $\mathcal{D}$ respectively. We first prove that the paths $\tilde{p}$ and $\tilde{q}$ exist in $\mathcal{D}^{pbd}_{\mathbf{XY}}$. Path q is d-connecting given $\mathbf{X} \cup \mathbf{Z_0}$, so any node in $\mathbf{X}$ on q must be a collider on q. Since $\mathcal{D}^{pbd}_{\mathbf{XY}}$ is obtained from $\mathcal{D}$ by removing certain edges out of $\mathbf{X}$, no edges from q are removed and $\tilde{q}$ exists in $\mathcal{D}^{pbd}_{\mathbf{XY}}$.

Since p is proper, for $\tilde{p}$ to exist in $\mathcal{D}^{pbd}_{\mathbf{XY}}$, it is enough to show that p does not start with an edge of type $X \to W$ in $\mathcal{D}$ where W lies on a proper causal path from X to $\mathbf{Y}$ in $\mathcal{D}$. Suppose for a contradiction that p does start with $X \to W$. Then $W \in \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D})$. Then either p is a directed path from X to Z_1 so that $Z_1 \in \text{De}(W, \mathcal{D})$, or p is non-causal and there is a collider C' on p such that $C' \in \text{De}(W, \mathcal{D})$. In the former case, since $\text{De}(W, \mathcal{D}) \subseteq \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D})$, it follows that $Z_1 \in \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D})$, which contradicts the choice of Z_1 because $Z_1 \notin \text{De}(\mathbf{X}, \mathcal{D})$. In the latter case, since p is d-connecting given $\mathbf{Z_0}$, we have $\text{De}(C', \mathcal{D}) \cap \mathbf{Z_0} \neq \emptyset$. Combining this with $\text{De}(C', \mathcal{D}) \subseteq \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D})$, it follows that $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) \cap \mathbf{Z_0} \neq \emptyset$ which contradicts that $\mathbf{Z_0}$ satisfies the forbidden set condition relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$. Thus, $\tilde{p}$ exists in $\mathcal{D}^{pbd}_{\mathbf{XY}}$.

We now show that $\tilde{p}$ and $\tilde{q}$ are d-connecting given $\mathbf{Z_0}$ in $\mathcal{D}^{pbd}_{\mathbf{XY}}$. The collider/non-collider status of any node on $\tilde{p}$ and $\tilde{q}$ in $\mathcal{D}^{pbd}_{\mathbf{XY}}$ is the same as the collider/non-collider status of that same node on p and q in $\mathcal{D}$ respectively. So $\tilde{p}$ and $\tilde{q}$ are both d-connecting given $\mathbf{Z_0}$, unless every causal path from a collider on either p or q to $\mathbf{Z_0}$ contains a first edge on a proper causal path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{D}$. Any such causal path also contains a node in $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D})$, so a node in $\mathbf{Z_0}$ would be a descendant of $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D})$ in $\mathcal{D}$. Since $\text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D})$ is descendral, it follows that $\mathbf{Z_0} \cap \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) \neq \emptyset$. This contradicts that $\mathbf{Z_0}$ satisfies the adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$, specifically the forbidden set condition.

Since $\tilde{p}$ is d-connecting given $\mathbf{Z_0}$ in $\mathcal{D}^{pbd}_{\mathbf{XY}}$, by Theorem 35 it follows that there is a path a from X to Z_1 that does not contain a node in $\mathbf{Z_0}$ in the moral induced subgraph of $\mathcal{D}^{pbd}_{\mathbf{XY}}$ on nodes $\text{An}(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z_0} \cup \{Z_1\}, \mathcal{D}^{pbd}_{\mathbf{XY}})$. Similarly, $\tilde{q}$ is a d-connecting path from Z_1 to Y given $\mathbf{Z_0}$ in $\mathcal{D}^{pbd}_{\mathbf{XY}}$, so there is a path b from Z_1 to Y that does not contain a node in $\mathbf{Z_0}$ in the moral induced subgraph of $\mathcal{D}^{pbd}_{\mathbf{XY}}$ on nodes $\text{An}(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z_0} \cup \{Z_1\}, \mathcal{D}^{pbd}_{\mathbf{XY}})$. By combining paths a and b we get a path c from $\mathbf{X}$ to $\mathbf{Y}$ that does not contain a node in $\mathbf{Z_0}$ in the moral induced subgraph of $\mathcal{D}^{pbd}_{\mathbf{XY}}$ on nodes $\text{An}(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z_0} \cup \{Z_1\}, \mathcal{D}^{pbd}_{\mathbf{XY}})$. Since $Z_1 \in \text{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D})$ and $\mathcal{D}^{pbd}_{\mathbf{XY}}$ is obtained from $\mathcal{D}$ by removing certain edges out of $\mathbf{X}$, it follows that $Z_1 \in \text{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D}^{pbd}_{\mathbf{XY}})$. Then $\text{An}(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z_0} \cup \{Z_1\}, \mathcal{D}^{pbd}_{\mathbf{XY}}) = \text{An}(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z_0}, \mathcal{D}^{pbd}_{\mathbf{XY}})$. Hence, c is a path from $\mathbf{X}$ to $\mathbf{Y}$ that does not contain a node in $\mathbf{Z_0}$ in the moral induced subgraph of $\mathcal{D}^{pbd}_{\mathbf{XY}}$ on nodes $\text{An}(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z_0}, \mathcal{D}^{pbd}_{\mathbf{XY}})$. Thus, by Theorem 35, $\mathbf{X}$ and $\mathbf{Y}$ are d-connected given $\mathbf{Z_0}$ in $\mathcal{D}^{pbd}_{\mathbf{XY}}$. By Theorem 7, this contradicts that $\mathbf{Z_0}$ satisfies the adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$. ■

Lemma 60 *Let* $\mathbf{X}, \mathbf{Y}$ *and* $\mathbf{Z_0}$ *be pairwise disjoint node sets in a* DAG $\mathcal{D}$ *such that* $\mathbf{Z_0}$ *satisfies the adjustment criterion relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}$. *Let* $\mathbf{Z} = \mathbf{Z_0} \cup \text{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D}) \setminus (\text{De}(\mathbf{X}, \mathcal{D}) \cup \mathbf{Y})$. *Additionally, let* $\mathbf{Z_D} = \mathbf{Z} \cap \text{De}(\mathbf{X}, \mathcal{D})$, $\mathbf{Z_N} = \mathbf{Z} \setminus \text{De}(\mathbf{X}, \mathcal{D})$, $\mathbf{Y_D} = \mathbf{Y} \cap \text{De}(\mathbf{X}, \mathcal{D})$ *and* $\mathbf{Y_N} = \mathbf{Y} \setminus \text{De}(\mathbf{X}, \mathcal{D})$. *Then the following statements hold:*

(i) $(\mathbf{X} \cup \mathbf{Y_N} \cup \mathbf{Z}) \cap \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) = \emptyset$*, and*

(ii) *if* $p = \langle A, \dots, Y_d \rangle$ *is a non-causal path from a node* $A \in \mathbf{X} \cup \mathbf{Y_N} \cup \mathbf{Z}$ *to a node* $Y_d \in \mathbf{Y_D}$*, then* p *is blocked by* $(\mathbf{X} \cup \mathbf{Y_N} \cup \mathbf{Z_N}) \setminus \{A\}$ *in* $\mathcal{D}$*, and*

(iii) $\mathbf{Y_D} \perp_d \mathbf{Z_D} \mid \mathbf{Y_N} \cup \mathbf{X} \cup \mathbf{Z_N}$ *in* $\mathcal{D}$*, where* $\mathbf{Y_N} = \emptyset$ *is allowed, and*

(iv) *if* $\mathbf{Y_N} = \emptyset$ *then* $\mathbf{Z_N}$ *is a generalized back-door set relative to* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathcal{D}$*, and*

(v) *the empty set is a generalized back-door set relative to* $((\mathbf{X} \cup \mathbf{Y_N} \cup \mathbf{Z_N}), \mathbf{Y_D})$ *in* $\mathcal{D}$*, and*

(vi) $\mathbf{Y_N} \cup \mathbf{Z_N} \perp_d \mathbf{Y_D} \mid \mathbf{X}$ *in* $\mathcal{D}_{\overline{\mathbf{X}}\underline{\mathbf{Y_N} \cup \mathbf{Z_N}}}$*, and*

(vii) $\mathbf{X} \perp_d \mathbf{Z_N} \mid \mathbf{Y_N}$ *in* $\mathcal{D}_{\overline{\mathbf{X}}}$*.*

Proof of Lemma 60. (i) By Lemma 59, $\mathbf{Z}$ satisfies the adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{D}$, implying that $\mathbf{Z} \cap \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) = \emptyset$. By definition $\mathbf{Y_N} \cap \mathrm{De}(\mathbf{X}, \mathcal{D}) = \emptyset$ and $\mathrm{De}(\mathbf{X}, \mathcal{D}) \supseteq \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{G})$. Hence, $\mathbf{Y_N} \cap \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) = \emptyset$. It is only left to show that $\mathbf{X} \cap \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) = \emptyset$.

Suppose for a contradiction that $\mathbf{X} \cap \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) \neq \emptyset$. Let $V \notin \mathbf{X}$ be a node on a proper causal path p from $\mathbf{X}$ to $\mathbf{Y}$ (possibly $V \in \mathbf{Y}$) such that $V \in \mathrm{An}(\mathbf{X}, \mathcal{D})$. Let $q = \langle V, \dots, X \rangle$, $X \in \mathbf{X}$, be a shortest causal path from V to $\mathbf{X}$. All nodes on $p(V, Y)$ and q are in $\mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D})$. We now concatenate $-q$ and $p(V, Y)$, while taking out loops. Hence, let W be the node closest to X on q that is also on $p(V, Y)$. Then $r = -q(X, W) \oplus p(W, Y)$ is a proper non-causal path from $\mathbf{X}$ to $\mathbf{Y}$ that cannot be blocked by $\mathbf{Z}$ since $\mathbf{Z} \cap \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) = \emptyset$, which contradicts Lemma 59.

(ii) We distinguish the cases that p is (a) out of, or (b) into Y_d.

(a) Let p be out of Y_d. Since $Y_d \in \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D})$ and since by (i) node $A \notin \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D})$, there is at least one collider on p. The collider closest to Y_d on p and all of its descendants are also in $\mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D})$. It then follows from (i) that p is blocked by $(\mathbf{X} \cup \mathbf{Y_N} \cup \mathbf{Z_N}) \setminus \{A\}$.

(b) Let p be into Y_d. Since p is a non-causal path from A to Y_d, there is at least one node on p that has two edges out of it. Let B be the closest such node to Y_d on p. Then $B \in \mathrm{An}(Y_d, \mathcal{D})$. If any node on $p(B, Y_d)$ is in $\mathbf{X} \cup \mathbf{Y_N} \cup \mathbf{Z_N}$, then (ii) holds.

Hence, assume no node on $p(B, Y_d)$ is in $\mathbf{X} \cup \mathbf{Y_N} \cup \mathbf{Z_N}$. Then $B \notin \mathbf{X} \cup \mathbf{Y_N} \cup \mathbf{Z_N}$. Since $B \notin \mathbf{Z_N}$, $\mathbf{Z_N} \supseteq \mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D}) \setminus (\mathrm{De}(\mathbf{X}, \mathcal{D}) \cup \mathbf{Y})$, it follows that $B \notin \mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D}) \setminus (\mathrm{De}(\mathbf{X}, \mathcal{D}) \cup \mathbf{Y})$. Additionally, since $B \in \mathrm{An}(Y_d, \mathcal{D})$, $B \in \mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D})$. Combining $B \notin \mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D}) \setminus (\mathrm{De}(\mathbf{X}, \mathcal{D}) \cup \mathbf{Y})$ and $B \in \mathrm{An}(\mathbf{X} \cup \mathbf{Y}, \mathcal{D})$ implies $B \in \mathrm{De}(\mathbf{X}, \mathcal{D}) \cup \mathbf{Y}$. Furthermore, $B \notin \mathbf{X} \cup \mathbf{Y_N}$ and $\mathbf{Y_D} \subseteq \mathrm{De}(\mathbf{X}, \mathcal{D}) \setminus \mathbf{X}$, so $B \in \mathrm{De}(\mathbf{X}, \mathcal{D}) \setminus \mathbf{X}$. But $B \in \mathrm{An}(Y_d, \mathcal{D})$ through $p(B, Y_d)$ on which no other node is in $\mathbf{X}$, so $B \in \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D})$. Now, $-p(B, A)$ is a path out of B, where $B \in \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D})$ and $A \notin \mathrm{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D})$ by (i). Using the same reasoning as in (a), $p(A, B)$ is blocked by $(\mathbf{X} \cup \mathbf{Y_N} \cup \mathbf{Z_N}) \setminus \{A\}$. Hence, p is also blocked by $(\mathbf{X} \cup \mathbf{Y_N} \cup \mathbf{Z_N}) \setminus \{A\}$.

(iii) By (ii) every non-causal path from $\mathbf{Z_D}$ to $\mathbf{Y_D}$ is blocked by $\mathbf{X} \cup \mathbf{Y_N} \cup \mathbf{Z_N}$. Additionally, $\mathbf{Z_D} \cap \text{Forb}(\mathbf{X}, \mathbf{Y}, \mathcal{D}) = \emptyset$ and $\mathbf{Z_D} \subseteq \text{De}(\mathbf{X}, \mathcal{D})$, so no node in $\mathbf{Z_D}$ can have a proper causal path to $\mathbf{Y_D}$ in $\mathcal{D}$. Hence, any causal path p from $\mathbf{Z_D}$ to $\mathbf{Y_D}$ in $\mathcal{D}$, must be non-proper with respect to $\mathbf{X}$, that is, p has to contain a node in $\mathbf{X}$ as a non-collider. Hence, every causal path from $\mathbf{Z_D}$ to $\mathbf{Y_D}$ is also blocked by $\mathbf{X} \cup \mathbf{Y_N} \cup \mathbf{Z_N}$ in $\mathcal{D}$.

(iv) Follows directly from $\mathbf{Z_N} \cap \text{De}(\mathbf{X}, \mathcal{D}) = \emptyset$ and (ii) for $\mathbf{Y_N} = \emptyset$.

(v) Follows directly from (ii).

(vi) Since $\mathcal{D}_{\overline{\mathbf{X}}\underline{\mathbf{Y_N} \cup \mathbf{Z_N}}}$ does not contain edges into $\mathbf{X}$, all paths from $\mathbf{Y_N} \cup \mathbf{Z_N}$ to $\mathbf{Y_D}$ in $\mathcal{D}_{\overline{\mathbf{X}}\underline{\mathbf{Y_N} \cup \mathbf{Z_N}}}$ that contain a collider are blocked by $\mathbf{X}$. Hence, it is enough to prove that $\mathbf{X}$ blocks any path from $\mathbf{Y_N} \cup \mathbf{Z_N}$ to $\mathbf{Y_D}$ in $\mathcal{D}_{\overline{\mathbf{X}}\underline{\mathbf{Y_N} \cup \mathbf{Z_N}}}$ that does not contain a collider. Let $r = \langle A, \dots, Y_d \rangle$, $A \in \mathbf{Y_N} \cup \mathbf{Z_N}$, $Y_d \in \mathbf{Y_D}$ be an arbitrary path from $\mathbf{Y_N} \cup \mathbf{Z_N}$ to $\mathbf{Y_D}$ in $\mathcal{D}_{\overline{\mathbf{X}}\underline{\mathbf{Y_N} \cup \mathbf{Z_N}}}$ that does not contain a collider. Since there are no edges out of $\mathbf{Y_N} \cup \mathbf{Z_N}$ in $\mathcal{D}_{\overline{\mathbf{X}}\underline{\mathbf{Y_N} \cup \mathbf{Z_N}}}$, it follows that r is into A and that r does not contain non-colliders that are in $\mathbf{Y_N} \cup \mathbf{Z_N}$. Let r' in $\mathcal{D}$ be the path constituted by the same sequence of nodes as r in $\mathcal{D}_{\overline{\mathbf{X}}\underline{\mathbf{Y_N} \cup \mathbf{Z_N}}}$. Since removing edges cannot d-connect blocked paths, it is enough to prove that r' is blocked by $\mathbf{X}$ in $\mathcal{D}$. Since r' is into A, r' is a non-causal path from $\mathbf{Y_N} \cup \mathbf{Z_N}$ to $\mathbf{Y_D}$, so by (ii), r' is blocked by $(\mathbf{X} \cup \mathbf{Y_N} \cup \mathbf{Z_N}) \setminus \{A\}$ in $\mathcal{D}$. Since $A \in \mathbf{Y_N} \cup \mathbf{Z_N}$ and $(\mathbf{Y_N} \cup \mathbf{Z_N}) \cap \mathbf{X} = \emptyset$, it follows that $A \notin \mathbf{X}$. Then r' is blocked by $\mathbf{X} \cup ((\mathbf{Y_N} \cup \mathbf{Z_N}) \setminus \{A\})$. Furthermore, since r does not contain non-colliders in $\mathbf{Y_N} \cup \mathbf{Z_N}$, the same is true for r', so r' is blocked by $\mathbf{X}$.

(vii) Since $\mathcal{D}_{\overline{\mathbf{X}}}$ does not contain edges into $\mathbf{X}$, all paths from $\mathbf{X}$ to $\mathbf{Z_N}$ in $\mathcal{D}_{\overline{\mathbf{X}}}$ are out of $\mathbf{X}$. Since $\mathbf{Z_N} \cap \text{De}(\mathbf{X}, \mathcal{D}) = \emptyset$ and $\text{De}(\mathbf{X}, \mathcal{D}) \supseteq \text{De}(\mathbf{X}, \mathcal{D}_{\overline{\mathbf{X}}})$, all paths from $\mathbf{X}$ to $\mathbf{Z_N}$ in $\mathcal{D}_{\overline{\mathbf{X}}}$ contain at least one collider. The closest collider to $\mathbf{X}$ on any such path is then in $\text{De}(\mathbf{X}, \mathcal{D}_{\overline{\mathbf{X}}})$. Since $\mathbf{Y_N} \cap \text{De}(\mathbf{X}, \mathcal{D}) = \emptyset$ and $\text{De}(\mathbf{X}, \mathcal{D}) \supseteq \text{De}(\mathbf{X}, \mathcal{D}_{\overline{\mathbf{X}}})$, this path is blocked by $\mathbf{Y_N}$ in $\mathcal{D}_{\overline{\mathbf{X}}}$. ■

References

Ayesha R. Ali, Thomas S. Richardson, and Peter Spirtes. Markov equivalence for ancestral graphs. *Ann. Stat.*, 37:2808–2837, 2009.

Elias Bareinboim, Jin Tian, and Judea Pearl. Recovering from selection bias in causal and statistical inference. In *Proceedings of AAAI 2014*, pages 2410–2416, 2014.

David M. Chickering. Optimal structure identification with greedy search. *J. Mach. Learn. Res.*, 3:507–554, 2002.

Tom Claassen, Joris Mooij, and Tom Heskes. Learning sparse causal models is not NP-hard. In *Proceedings of UAI 2013*, pages 172–181, 2013.

Diego Colombo and Marloes H. Maathuis. Order-independent constraint-based causal structure learning. *J. Mach. Learn. Res.*, 15:3741–3782, 2014.

Diego Colombo, Marloes H. Maathuis, Markus Kalisch, and Thomas S. Richardson. Learning high-dimensional directed acyclic graphs with latent and selection variables. *Ann. Stat.*, 40:294–321, 2012.

Juan D. Correa and Elias Bareinboim. Causal effect identification by adjustment under confounding and selection biases. In *Proceedings of AAAI 2017*, pages 3740–3746, 2017.

Xavier De Luna, Ingeborg Waernbaum, and Thomas S. Richardson. Covariate selection for the nonparametric estimation of an average treatment effect. *Biometrika*, 98(4):861–875, 2011.

Doris Entner, Patrik O. Hoyer, and Peter Spirtes. Data-driven covariate selection for non-parametric estimation of causal effects. In *Proceedings of AISTATS 2013*, pages 256–264, 2013.

Benjamin Frot, Preetam Nandy, and Marloes H. Maathuis. Learning directed acyclic graphs with hidden variables via latent Gaussian graphical model selection. arXiv preprint arXiv:1708.01151, 2018.

Sander Greenland, Judea Pearl, and James M. Robins. Causal diagrams for epidemiologic research. *Epidemiology*, 10(1):37–48, 1999.

Hui Guo and Philip A. Dawid. Sufficient covariates and linear propensity analysis. In *Proceedings of AISTATS 2010*, pages 281–288, 2010.

Jinyong Hahn. On the role of the propensity score in efficient semiparametric estimation of average treatment effects. *Econometrica*, pages 315–331, 1998.

Jinyong Hahn. Functional restriction and efficiency in causal inference. *Rev. Econ. Stat.*, 86 (1):73–76, 2004.

Christina Heinze-Deml, Marloes H. Maathuis, and Nicolai Meinshausen. Causal structure learning. *Annu. Rev. Stat. Appl.*, 5:371–391, 2017.

Markus Kalisch, Martin Mächler, Diego Colombo, Marloes H. Maathuis, and Peter Bühlmann. Causal inference using graphical models with the R package pcalg. *J. Stat. Softw.*, 47(11):1–26, 2012.

Jan TA. Koster. Marginalizing and conditioning in graphical models. *Bernoulli*, 8(6):817–840, 2002.

Manabu Kuroki and Zhihong Cai. Selection of identifiability criteria for total effects by using path diagrams. In *Proceedings of UAI 2004*, pages 333–340, 2004.

Manabu Kuroki and Masami Miyakawa. Covariate selection for estimating the causal effect of control plans by using causal diagrams. *J. Roy. Stat. Soc. B*, 65(1):209–222, 2003.

Steffen L. Lauritzen and David J. Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems. *J. Roy. Stat. Soc. B*, 50(2): 157–224, 1988.

Steffen L. Lauritzen, Philip A. Dawid, Birgitte N. Larsen, and Hans-Georg Leimer. Independence properties of directed Markov fields. *Networks*, 20(5):491–505, 1990.

Marloes H. Maathuis and Diego Colombo. A generalized back-door criterion. *Ann. Stat.*, 43:1060–1088, 2015.

Marloes H. Maathuis, Markus Kalisch, and Peter Bühlmann. Estimating high-dimensional intervention effects from observational data. *Ann. Stat.*, 37:3133–3164, 2009.

Marloes H. Maathuis, Diego Colombo, Markus Kalisch, and Peter Bühlmann. Predicting causal effects in large-scale systems from observational data. *Nat. Methods*, 7:247–248, 2010.

Daniel Malinsky and Peter Spirtes. Estimating bounds on causal effects in high-dimensional and possibly confounded systems. *Int. J. of Approx. Reason.*, 88:371–384, 2017.

Kanti V. Mardia, John T. Kent, and John M. Bibby. *Multivariate Analysis (Probability and Mathematical Statistics)*. Academic Press London, 1980.

Christopher Meek. Causal inference and causal explanation with background knowledge. In *Proceedings of UAI 1995*, pages 403–410, 1995.

Preetam Nandy, Marloes H. Maathuis, and Thomas S. Richardson. Estimating the effect of joint interventions from observational data in sparse high-dimensional settings. *Ann. Stat.*, 45(2):647–674, 2017.

Preetam Nandy, Alain Hauser, and Marloes H. Maathuis. High-dimensional consistency in score-based and hybrid structure learning. *Ann. Stat.*, 2018. To appear.

Judea Pearl. Comment: Graphical models, causality and intervention. *Stat. Sci.*, 8:266–269, 1993.

Judea Pearl. *Causality: Models, Reasoning, and Inference.* Cambridge University Press, New York, NY, second edition, 2009.

Judea Pearl and Azaria Paz. Confounding equivalence in causal inference. *J. Causal Infer.*, 2(1):75–93, 2014.

Emilija Perković, Johannes Textor, Markus Kalisch, and Marloes H. Maathuis. A complete generalized adjustment criterion. In *Proceedings of UAI 2015*, pages 682–691, 2015.

Emilija Perković, Markus Kalisch, and Marloes H. Maathuis. Interpreting and using CPDAGs with background knowledge. In *Proceedings of UAI 2017*, 2017.

Thomas S. Richardson. Markov properties for acyclic directed mixed graphs. *Scand. J. Statist.*, 30:145–157, 2003.

Thomas S. Richardson and Peter Spirtes. Ancestral graph Markov models. *Ann. Stat.*, 30: 962–1030, 2002.

James M. Robins. A new approach to causal inference in mortality studies with a sustained exposure period-application to control of the healthy worker survivor effect. *Math. Mod.*, 7:1393–1512, 1986.

Donald Rubin. Author's reply. *Stat. Med.*, 27:2741–2742, 2008.

Ross D. Shachter. Bayes-ball: The rational pastime. In *Proceedings of UAI 1998*, pages 480–487, 1998.

Ilya Shpitser. Appendum to "On the validity of covariate adjustment for estimating causal effects". Personal communication, 2012.

Ilya Shpitser and Judea Pearl. Identification of joint interventional distributions in recursive semi-Markovian causal models. In *Proceedings of AAAI 2006*, pages 1219–1226, 2006.

Ilya Shpitser, Tyler VanderWeele, and James M. Robins. On the validity of covariate adjustment for estimating causal effects. In *Proceedings of UAI 2010*, pages 527–536, 2010.

Ian Shrier. Letter to the editor. *Stat. Med.*, 27:2740–2741, 2008.

Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, Prediction, and Search.* MIT Press, Cambridge, MA, second edition, 2000.

Johannes Textor, Benito van der Zander, Mark S. Gilthorpe, Maciej Liśkiewicz, and George TH Ellison. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. *Int. J. Epidemiol.*, 45(6):1887–1894, 2016.

Jin Tian and Judea Pearl. A general identification condition for causal effects. In *Proceedings of AAAI 2002*, pages 567–573, 2002.

Benito van der Zander and Maciej Liśkiewicz. Separators and adjustment sets in Markov equivalent DAGs. In *Proceedings of AAAI 2016*, pages 3315–3321, 2016.

Benito van der Zander, Maciej Liśkiewicz, and Johannes Textor. Constructing separators and adjustment sets in ancestral graphs. In *Proceedings of UAI 2014*, pages 907–916, 2014.

Benito van der Zander, Maciej Liśkiewicz, and Johannes Textor. Separators and adjustment sets in causal graphs: Complete criteria and an algorithmic framework. arXiv preprint arXiv:1803.00116, 2018.

Tyler J. VanderWeele and Ilya Shpitser. A new criterion for confounder selection. *Biometrics*, 67(4):1406–1413, 2011.

Stephen G. West and Tobias Koch. Restoring causal analysis to structural equation modeling. *Struct. Equ. Modeling*, 21:161–166, 2014.

Daniel Westreich and Sander Greenland. The table 2 fallacy: Presenting and interpreting confounder and modifier coefficients. *Am. J. Epidemiol.*, 177:292–298, 2013.

Sewall Wright. Correlation and causation. *J. Agric. Res.*, 20(7):557–585, 1921.

Jiji Zhang. *Causal Inference and Reasoning in Causally Insufficient Systems.* PhD thesis, Carnegie Mellon University, 2006.

Jiji Zhang. Causal reasoning with ancestral graphs. *J. Mach. Learn. Res.*, 9:1437–1474, 2008a.

Jiji Zhang. On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias. *Artif. Intell.*, 172:1873–1896, 2008b.