Artificial Intelligence 347 (2025) 104387

Contents lists available at ScienceDirect

Artificial Intelligence

journal homepage: www.elsevier.com/locate/artint

Estimating possible causal effects with latent variables via adjustment and novel rule orientation

Tian-Zuo Wang [a,b], Lue Tao [a,b], Tian Qin [a,b], Zhi-Hua Zhou [a,b,*]

[a] National Key Laboratory for Novel Software Technology, Nanjing University, China
[b] School of Artificial Intelligence, Nanjing University, China

ARTICLE INFO

Keywords:
Causal effects
Latent confounders
Partial ancestral graph
Orientation rules
Background knowledge

ABSTRACT

Causal effect estimation from observational data is a fundamental task in artificial intelligence and has been widely studied given known causal relations. However, in the presence of latent confounders, only a part of causal relations can be identified from observational data, characterized by a partial ancestral graph (PAG), where some causal relations are indeterminate. In such cases, the causal effect is often unidentifiable, as there could be super-exponential number of potential causal graphs consistent with the identified PAG but associated with different causal effects. In this paper, we target on *set determination* within a PAG, *i.e.*, determining the set of possible causal effects of a specified variable X on another variable Y via covariate adjustment. We develop the first set determination method that does not require enumerating any causal graphs. Furthermore, we present two novel orientation rules for incorporating structural background knowledge (BK) into a PAG, which facilitate the identification of additional causal relations given BK. Notably, we show that these rules can further enhance the efficiency of our set determination method, as certain transformed edges during the procedure can be interpreted as BK and enable the rules to reveal further causal information. Theoretically and empirically, we demonstrate that our set determination methods can yield the same results as the enumeration-based method with *super-exponentially less* computational complexity.

1. Introduction

Causal effect estimation is a fundamental task in artificial intelligence. Given a causal graph that characterizes the causal relations among relevant variables, a lot of methods have been proposed for estimating causal effect from observational data [1–5]. In practice, however, there is generally not a priori causal graph. Without specific assumptions, only a part of causal relations can be identified from observational data, which generally renders the causal effect unidentifiable.

The issue of unidentifiability is further exacerbated when there are latent variables, in which case one can only learn a graphical model called *partial ancestral graph (PAG)* from observational data. Essentially, a PAG represents a Markov equivalence class (MEC) of *maximal ancestral graphs (MAG)* which characterize the causal relations among *observed variables*. For example, as shown in Fig. 1, the MAG in Fig. 1(b) belongs to the MEC represented by the PAG in Fig. 1(a). Roughly speaking, a MAG is a “projection graph on the observed variables” of an underlying *directed acyclic graph (DAG)* characterizing the causal relations among both observed and latent

* Corresponding author at: National Key Laboratory for Novel Software Technology, Nanjing University, China.
E-mail address: zhouzh@lamda.nju.edu.cn (Z.-H. Zhou).

https://doi.org/10.1016/j.artint.2025.104387
Received 17 August 2024; Received in revised form 11 May 2025; Accepted 9 June 2025

Available online 13 June 2025

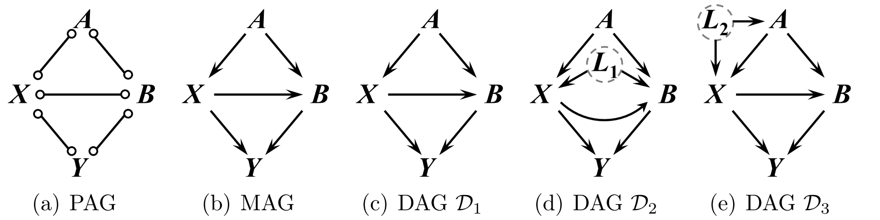

Fig. 1. Fig. 1(a): PAG $\mathcal{P}$. Fig. 1(b): MAG $\mathcal{M}$ in the MEC represented by $\mathcal{P}$. Fig. 1(c)/ 1(d)/ 1(e): DAG $\mathcal{D}_1/\mathcal{D}_2/\mathcal{D}_3$ that is projected to $\mathcal{M}$ over the observed variables $\{A, B, X, Y\}$. L_1, L_2: latent variables. As there can be *any* number of latent variables, there can be infinite DAGs projected to $\mathcal{M}$. $f(Y|do(X))$ is identifiable via covariate adjustment in $\mathcal{D}_1$ and $\mathcal{D}_3$ but not in $\mathcal{D}_2$, thus it is not identifiable in $\mathcal{M}$ and $\mathcal{P}$.

variables.[1] Notably, many DAGs, even infinite ones, can be projected to one MAG. For instance, Fig. 1(c)/ 1(d)/ 1(e) illustrate three underlying DAGs corresponding to the MAG in Fig. 1(b). Hereafter, we say that a DAG is *consistent with a PAG* if it can be projected to a MAG in the MEC represented by the PAG.

Given the general unidentifiability of causal effects, Maathuis et al. [7] focused on the problem of *set determination*, *i.e.*, determining the set of all possible causal effects, where a causal effect is considered possible if there exists a DAG, consistent with observational data, that yields this effect. Although such a set is less informative than identifying a unique causal effect, it still provides valuable information and, more importantly, does not require prior structural knowledge or experimental data, making it broadly applicable in practice. For set determination in the absence of latent variables, Maathuis et al. [7] proposed the first method that avoids enumerating any causal graphs, and subsequent works have further improved its efficiency and reduced estimation error [7–10]. These methods obtain each possible causal effect by *(covariate) adjustment*, *i.e.*, by adjusting for certain observed variables known as the *adjustment set* [11–16]. Despite the extensive literature on set determination without latent confounders, it remains an open problem that how to determine the set if we allow for the existence of latent variables [7]. Towards this ambitious target, Malinsky and Spirtes [17] proposed the method LV-IDA to obtain each possible causal effect via adjustment. However, in the worst case, this approach requires enumerating all MAGs, which is computationally prohibitive.

In the presence of latent confounders, the main challenge for set determination lies in the large number of DAGs consistent with a given PAG. Specifically, a PAG can represent a super-exponential number of MAGs and each MAG can be a projection graph of infinite DAGs. The enumerations of either MAGs or DAGs are computationally prohibitive. In this paper, we propose the first method to determine the **PAG**-consistent **caus**al **e**ffect **s**et *without any enumerations of DAGs or MAGs*, called by PAGcauses for short. Our approach is grounded in two key theoretical results. One is a graphical condition for determining adjustment sets comprised of observed variables in *all* the DAGs projected to a given MAG, through which we can circumvent the enumerations of DAGs and find adjustment sets based on mere MAGs. The other is a graphical characterization for adjustment set in MAGs consistent with a given PAG, through which we can convert the problem of *searching adjustment sets in all the MAGs* to a decision problem that for each set of variables, *whether it can be an adjustment set in some MAG*. The complexity of PAGcauses is super-exponentially less than that of enumeration-based methods [17]. Further, we prove that PAGcauses can output the set of all possible causal effects that are identifiable via adjustment. Note that there are possibly some DAGs consistent with the PAG where the causal effect cannot be identified with observational data via adjustment. We do not try to return such value since it is beyond the ability of current observations when latent variables exist.

In addition, we present two novel orientation rules for incorporating causal background knowledge (BK) into a PAG, one of which is fundamentally different in form from existing rules [18,6,19,20]. These rules not only inspire the establishment of sound and complete rules to incorporate BK in the future - which could address the open problem of *causal identification* [6], *i.e.*, determining which causal relations are *identifiable* from observational data and BK - but also, interestingly, accelerate set determination, even in the absence of BK. The reason is, certain transformed edges during the set determination procedure can be interpreted as hypothesis BK, which enable the rules to reveal further causal information. Furthermore, while there have been a few studies on incorporating specific types of BK, such as tiered BK [21] or local BK [19], a comprehensive solution remains elusive. Unlike existing rules, which typically identify causal relations based on a small number of *edges* or *paths*, our proposed rules can identify causal relations that depend on a broader *subgraph*. Our findings suggest that, in the presence of latent variables, causal identification with BK requires more complicated rules, highlighting the intrinsic complexity of this problem. Experiments verify the effectiveness and significant efficiency improvement of PAGcauses as well as its enhanced version that utilizes the proposed orientation rules.

A preliminary version of this work appeared in a conference paper [22]. Compared with the original version, we make three additional contributions.

(1) We propose Algorithm 2 to improve the set determination efficiency, through which we could largely reduce the size of candidate sets that could possibly be an adjustment set for estimating a possible causal effect.
(2) We propose two novel orientation rules to incorporate BK into a PAG, which are beneficial for solving the open problem of what causal relations are identifiable given BK in the presence of latent confounders in the future.

[1] The "projection" is elaborated on Page 1877 of Zhang [6].

(3) Based on the novel orientation rules, we present an improved method for set determination. The application of the rules further reduces the computational complexity of set determination.

2. Preliminary

2.1. Graphs notations

A graph $G = (\mathbf{V}, \mathbf{E})$ consists of a set of vertices $\mathbf{V} = \{V_1, \cdots, V_p\}$ and a set of edges $\mathbf{E}$. For any subset $\mathbf{V}' \subseteq \mathbf{V}$, the *subgraph induced by* $\mathbf{V}'$ is $G[\mathbf{V}'] = (\mathbf{V}', \mathbf{E}_{\mathbf{V}'})$, where $\mathbf{E}_{\mathbf{V}'}$ is the set of edges in $\mathbf{E}$ whose both endpoints are in $\mathbf{V}'$. $\mathbf{V}(G)$ denotes the set of vertices in G. G is a complete graph if there is an edge between any two vertices. The subgraph induced by an empty set is trivially a complete graph. $G[-\mathbf{V}']$ denotes the subgraph induced by $\mathbf{V}(G)\backslash\mathbf{V}'$. Usually, bold letter (*e.g.*, $\mathbf{V}$) denotes a set of vertices and normal letter (*e.g.*, V) denotes a vertex. A graph is *chordal* if any cycle of length four or more has a chord, which is an edge joining two vertices that are not adjacent in the cycle. If $G = (\mathbf{V}, \mathbf{E})$ is chordal, then all subgraphs of G induced by $\mathbf{V}' \subseteq \mathbf{V}$ are chordal. A vertex A of G is called *simplicial* if its adjacency set induces a complete subgraph of G. As shown by Dirac [23], Golumbic [24], there are at least two non-adjacent simplicial vertices in any a non-complete chordal graph with more than one vertex. A *perfect elimination order* of a graph G is an ordering $\sigma = (V_1, \cdots, V_n)$ of its vertices, so that each vertex V_i is a simplicial vertex in the induced subgraph $G_{V_i,\cdots,V_n}$.

A graph G is *mixed* if the edges in the graph are either directed $\rightarrow$ or bi-directed $\leftrightarrow$. The two ends of an edge are called *marks* and have two types *arrowhead* or *tail*. A graph is a *partial mixed graph* (*PMG*) if it contains directed edges, bi-directed edges, and edges with *circles* ($\circ$). The circle implies that the mark here could be either arrowhead or tail but is indeterminate. An edge $V_i \circ\!\!-\!\!\circ V_j$ is a *circle edge*. The circle component in G is the subgraph consisting of all the $\circ\!\!-\!\!\circ$ edges in G. V_i is *adjacent* to V_j in G if there is an edge between V_i and V_j. A vertex V_i is a *parent/child* of a vertex V_j if there is $V_i \rightarrow V_j / V_i \leftarrow V_j$. A *path* in a graph G is a sequence of distinct vertices $\langle V_0, \cdots, V_n \rangle$ such that for $0 \le i \le n-1$, V_i and V_{i+1} are adjacent in G. $\bigoplus$ denotes concatenation of paths. For a path p, let $p[V_i, V_j]$ denote the sub-path of p from V_i to V_j. A *directed path* from V_i to V_j is a path comprised of directed edges pointing to the direction of V_j. A *possible directed path* from V_i to V_j is a path without arrowhead at the mark near V_i on every edge. V_i is an *ancestor/possible ancestor* of V_j if there is a *directed path/possible directed path* from V_i to V_j or $V_i = V_j$. Define *descendant/possible descendant* similarly. Denote the set of *parents/ancestors/possible ancestors/descendants/possible descendants* of V_i in G by $\mathrm{Pa}(V_i, G)/\mathrm{Anc}(V_i, G)/\mathrm{PossAn}(V_i, G)/\mathrm{De}(V_i, G)/\mathrm{PossDe}(V_i, G)$. If $V_i \in \mathrm{Anc}(V_j, G)$ and there is $V_i \leftarrow V_j / V_i \leftrightarrow V_j$ in G, it forms a *directed cycle/almost directed cycle*. $*$ is a wildcard that denotes any of arrowhead, tail, or circle. We make a convention that if an edge is $\circ\!\!-\!\!*$, the $*$ here cannot be a tail for otherwise the circle can be replaced by an arrowhead due to no edges with two tails under the assumption of no selection bias.

A non-endpoint vertex V_i is a *collider* on a path if the path contains $*\!\!\rightarrow V_i \leftarrow\!\!*$. A path p from V_i to V_j is a *collider path* if V_i and V_j are adjacent or all the passing vertices are colliders on p. p is a *minimal path* if there are no edges between any two non-consecutive vertices. A path p from V_i to V_j is a *minimal collider path* if p is a collider path and there is not a proper subset $\mathbf{V}'$ of the vertices in p such that there is a collider path from V_i to V_j comprised of $\mathbf{V}'$. A triple $\langle V_i, V_j, V_k \rangle$ on a path is *unshielded* if V_i and V_k are not adjacent. p is an *uncovered path* if every consecutive triple on p is unshielded. A path p is a *minimal possible directed path* if p is minimal and a possible directed path. A graph is a *directed acyclic graph (DAG)* if it contains only directed edges and has no directed cycles, denoted by $\mathcal{D}$. $\mathbf{X}$ and $\mathbf{Y}$ are *m-separated* by $\mathbf{Z}$ if there is no active path between any vertex in $\mathbf{X}$ and any vertex in $\mathbf{Y}$ relative to $\mathbf{Z}$.

2.2. Maximal ancestral graph and partial ancestral graph

A mixed graph is an *ancestral graph* if there is no directed or almost directed cycle.[2] An ancestral graph is a *maximal ancestral graph (MAG, denoted by $\mathcal{M}$)* if it is *maximal*, i.e., for any two non-adjacent vertices, there is a set of vertices that m-separates them. A path p between X and Y in an ancestral graph is an *inducing path* if every non-endpoint vertex on p is a collider and meanwhile an ancestor of either X or Y. An ancestral graph is maximal if and only if there is no inducing path between any two non-adjacent vertices [25]. In a MAG, a path $p = \langle X, \cdots, W, V, Y \rangle$ is a *discriminating path for* V if (1) X and Y are not adjacent, and (2) every vertex between X and V on the path is a collider on p and a parent of Y.

Two MAGs are *Markov equivalent* if they share the same m-separations. A class comprised of all the Markov equivalent MAGs is a *Markov equivalence class (MEC)*. A *partial ancestral graph (PAG, denoted by $\mathcal{P}$)* represents an MEC, where a tail/arrowhead occurs if the corresponding mark is tail/arrowhead in all the Markov equivalent MAGs, and a circle occurs otherwise. We say a MAG $\mathcal{M}$ is *consistent with a PAG* $\mathcal{P}$ if $\mathcal{P}$ represents the Markov equivalence class of $\mathcal{M}$. Further, we say a DAG $\mathcal{D}$ *is represented by* or *can be projected to* $\mathcal{M}$ if $\mathcal{M}$ can be obtained from $\mathcal{D}$ by the construction process on Page 1877 of Zhang [6].

In a mixed graph, a directed edge $A \rightarrow B$ is *visible* if there a vertex C not adjacent to B, such that there is an edge $C \rightarrow A$, or there is a collider path between C and A that is into A and every non-endpoint is a parent of B. Otherwise $A \rightarrow B$ is said to be *invisible*. For a MAG $\mathcal{M}$, let $\mathcal{M}_{\underline{X}}$ denote the graph obtained from $\mathcal{M}$ by removing all visible directed edges out of X in $\mathcal{M}$ [26]. For a PAG $\mathcal{P}$, let $\mathcal{M}$ be any MAG consistent with $\mathcal{P}$ that has the same number of edges into X as $\mathcal{P}$, and let $\mathcal{P}_{\underline{X}}$ denote the graph obtained from $\mathcal{M}$ by removing all directed edges out of X that are visible in $\mathcal{M}$ (it is not required to be unique). $\mathcal{M}_{\underset{\sim}{X}}$ denotes the graph obtained from $\mathcal{M}$ by deleting the directed edges out of X.

[2] Since we assume no selection bias, we do not consider undirected edges in this paper.

Ali et al. [27], Zhang [6] proposed ten rules $\mathcal{R}_1 - \mathcal{R}_{10}$ to learn a PAG with observational data. Further, given a PAG, when incorporating *local* BK, *i.e.*, the full structural knowledge of some specific variables, Wang et al. [19] proposed a rule $\mathcal{R}'_4$ to replace $\mathcal{R}_4$ and an additional rule $\mathcal{R}_{11}$.

2.3. Causal effect estimation

Definition 1 *(Adjustment set; Pearl [3], van der Zander et al. [11]).* Given a DAG, MAG, or PAG G, $\mathbf{Z}$ is called an adjustment set relative to (X,Y) if for any density f compatible with G, the causal effect of X on Y

$$P(Y|do(X)) = \begin{cases} P(Y|X), & \text{if } \mathbf{Z} = \varnothing, \\ \int_{\mathbf{Z}} P(Y|\mathbf{Z},X)P(\mathbf{Z})\,\mathrm{d}\mathbf{Z}, & \text{otherwise.} \end{cases} \tag{1}$$

A common method to estimate causal effect is *covariate adjustment*, *i.e.*, estimating by (1) with *adjustment sets* in Definition 1. *Adjustment criterion* in Definition 2 characterizes adjustment sets. Given a DAG, a set $\mathbf{Z}$ satisfies adjustment criterion if and only if $\mathbf{Z}$ is an adjustment set [4,28,11].

Definition 2 *(Adjustment criterion; Shpitser et al. [4], VanderWeele and Shpitser [28], van der Zander et al. [11]).* Let $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$ be pairwise disjoint sets of vertices in a DAG $\mathcal{D}$. Let $\mathrm{Forb}(\mathbf{X},\mathbf{Y},\mathcal{D})$ 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 $\mathbf{Y}$, *i.e.*, only the first node is in $\mathbf{X}$, in $\mathcal{D}$. Then $\mathbf{Z}$ satisfies adjustment criterion relative to $(\mathbf{X},\mathbf{Y})$ in $\mathcal{D}$ if the following two conditions hold:

(Forbidden set) $\mathbf{Z} \cap \mathrm{Forb}(\mathbf{X},\mathbf{Y},\mathcal{D}) = \emptyset$, and
(Blocking) all proper non-causal paths from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{D}$ are blocked by $\mathbf{Z}$.

Definition 3 *($\text{D-SEP}(X,Y,G)$; Spirtes et al. [29], Colombo et al. [30], Maathuis et al. [31]).* Let X and Y be two distinct vertices in a mixed graph G. We say that $V \in \text{D-SEP}(X,Y,G)$ if $V \neq X$, and there is a collider path between X and V in G, such that every vertex on this path (including V) is an ancestor of X or Y in G.

Maathuis et al. [31] proposed *generalized back-door criterion* to identify causal effect of $\mathbf{X}$ on $\mathbf{Y}$ in a MAG/PAG by adjusting for a *generalized back-door set*. When X is singleton, there is a generalized back-door set relative to (X,Y) if and only if there is an adjustment set relative to (X,Y) [12]. We combine them in Proposition 1.

Proposition 1 *(Maathuis et al. [31], Perkovic et al. [12]). Let X and Y be two distinct vertices in G, where G is a MAG or PAG. There exists an adjustment set relative to (X,Y) in G if and only if $Y \notin \mathrm{Adj}(X,G_{\underline{X}})$ and $\text{D-SEP}(X,Y,G_{\underline{X}}) \cap \mathrm{PossDe}(X,G) = \emptyset$. Moreover, if an adjustment set exists, then $\text{D-SEP}(X,Y,G_{\underline{X}})$ is such a set. Denote $\text{D-SEP}(X,Y,G_{\underline{X}})$ by $\mathbf{D}$, then*

$$f(Y|do(X=x)) = \int_{\mathbf{D}} f(\mathbf{D})f(Y|\mathbf{D},X=x)\,\mathrm{d}\mathbf{D}. \tag{2}$$

3. Proposed method

In this section, we aim to determine the set of possible causal effects of a variable X on outcome Y via adjustment given a PAG $\mathcal{P}$. $\mathcal{P}$ can be learned by FCI algorithm from observational data [29]. Let d denote the number of vertices in $\mathcal{P}$. We assume the absence of selection bias. If the causal effect is identifiable in $\mathcal{P}$ by Proposition 1, an unbiased estimate can be directly returned, thus there is no need to determine a set. Hence we focus on the case when it is unidentifiable. And our attention is only on finding all *valid adjustment sets* given $\mathcal{P}$, without touching upon practical calculation of causal effects by (1). Here, a set of vertices is a *valid adjustment set* if there exists a DAG $\mathcal{D}$ consistent with $\mathcal{P}$ such that the causal effect of X on Y in $\mathcal{D}$ is identifiable by adjusting for this set.

A direct method is first enumerating all the MAGs in the MEC represented by $\mathcal{P}$, then enumerating all the DAGs for each above MAG. However, the enumerations of either DAGs or MAGs are computationally prohibitive. In Sec. 3.1, we present the theoretical result for circumventing the enumerations of DAGs. We provide a graphical characterization for the adjustment set comprised of observable variables in DAGs represented by a given MAG, through which we can find all valid adjustment sets based on mere MAGs instead of DAGs. In Sec. 3.2, we find all the valid adjustment sets given a PAG without enumerating MAGs. The whole process includes two steps. First, we determine all possible local structures at X. The local structure is not sufficient for determining an adjustment set, but it can lead to a graphical characterization of valid adjustment sets relative to (X,Y), which can be evaluated in $\mathcal{O}(d^3)$ for each set. A primary algorithm and an improved one for set determination are presented in Sec. 3.3, associated with a worst-case complexity analysis in Sec. 3.4. All the proofs are in Sec. 3.5.

3.1. Adjustment sets in DAGs represented by a MAG

In this part, given any MAG $\mathcal{M}$, we provide a graphical characterization for adjustment sets relative to (X,Y) in *all* DAGs represented by $\mathcal{M}$. The causal effect of X on Y in $\mathcal{M}$ is *identifiable via covariate adjustment* if and only if all the DAGs represented by

$\mathcal{M}$ are associated with the same causal effect and can be estimated with observational data via covariate adjustment [11, Def. 5.3]. Proposition 1 presents the sufficient and necessary condition for the identifiability. Hence, if the graphical conditions of Proposition 1 are satisfied for $\mathcal{M}$, it is direct that all the DAGs represented by $\mathcal{M}$ are associated with the same causal effect, and we can obtain the adjustment set according to Proposition 1. However, when it is unidentifiable in $\mathcal{M}$, different DAGs represented by $\mathcal{M}$ are possibly associated with different causal effects. Perhaps surprisingly, we find that even if the causal effect is unidentifiable by adjustment in $\mathcal{M}$, for each DAG represented by $\mathcal{M}$, it is either associated with a *common* causal effect value that can be estimated, or with a causal effect value that cannot be estimated with observational data by covariate adjustment. Concretely, we provide Theorem 1, establishing a graphical characterization for the adjustment sets comprised of $\mathbf{V}(\mathcal{M})$ in the DAGs.

Theorem 1. *Suppose a MAG $\mathcal{M}$ where $X \in \mathrm{Anc}(Y, \mathcal{M})$. There exists a DAG $\mathcal{D}$ represented by $\mathcal{M}$ such that the causal effect of X on Y in $\mathcal{D}$ can be identified by adjusting for a set comprised of $\mathbf{V}(\mathcal{M})$ if and only if $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) \cap \mathrm{De}(X, \mathcal{M}) = \emptyset$. Furthermore, if such a set exists, $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ is an adjustment set.*

Remark 1. We restrict X to be an ancestor of Y in Theorem 1 because $X \notin \mathrm{Anc}(Y, \mathcal{M})$ is a trivial case that X has no causal effect on Y in any DAG $\mathcal{D}$ represented by the MAG $\mathcal{M}$.

Theorem 1 is similar to Proposition 1 in form, but they are quite different in both the implications and proofs. Proposition 1 implies an adjustment set in a MAG $\mathcal{M}$ when all the DAGs represented by $\mathcal{M}$ are associated with the same causal effect. Here, since we aim to determine the set of possible causal effects, *regardless of whether* all the DAGs represented by $\mathcal{M}$ are associated with the same causal effect, if there are some DAGs represented by the MAG associated with the causal effect that is identifiable by adjusting for observed variables, then we want to include the causal effect value in the set. See the MAG $\mathcal{M}$ in Fig. 1(b) for an example, where $\text{D-SEP}(X, Y, \mathcal{M}_{\underline{X}}) \cap \mathrm{PossDe}(X, \mathcal{M}) = \{B\}$ and $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) \cap \mathrm{PossDe}(X, \mathcal{M}) = \emptyset$. According to Proposition 1, the causal effect of X on Y in $\mathcal{M}$ is not identifiable by covariate adjustment. It is true since the DAGs in Fig. 1(c), 1(d), and 1(e) represented by $\mathcal{M}$ are associated with different causal effects. Nevertheless, according to Theorem 1, there exists some DAG represented by $\mathcal{M}$ where the causal effect can be identified and $\{A\}$ is an adjustment set. It is the case in DAGs as Fig. 1(c), 1(e).

Theorem 1 has two implications. One is that it provides a graphical condition based on a mere MAG for the existence of adjustment sets comprised of observed variables in the DAGs represented by the MAG, thus we can find adjustment sets on the level of MAGs. The other is that it implies that the causal effects are *identical* in the DAGs where the causal effect is identifiable by adjusting for a set of observed variables. Hence, if $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) \cap \mathrm{De}(X, \mathcal{M}) = \emptyset$, $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ is an adjustment set in *all the DAGs* where the causal effect is identifiable by adjusting for a set comprised of $\mathbf{V}(\mathcal{M})$, thus finding $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ as the adjustment set is sufficient. These two aspects ensure that we can obtain the adjustment set without enumerating DAGs. Henceforth, we make a convention that when we say *an adjustment set in* $\mathcal{M}$, it implies an adjustment set comprised of $\mathbf{V}(\mathcal{M})$ in some DAG represented by $\mathcal{M}$ as Theorem 1, without restricting that all the DAGs represented by $\mathcal{M}$ are associated with the same causal effect as Proposition 1.

3.2. Adjustment sets in MAGs consistent with a PAG

In this part, given a PAG $\mathcal{P}$, we aim to find all the sets that can be adjustment sets relative to (X, Y) in some MAG consistent with $\mathcal{P}$. This task is hard due to the large MAG spaces encoded by the PAG. Hence, we first introduce Proposition 2 of Wang et al. [19] to determine all the *valid local structures at* X, *i.e.*, what orientations of the circles at X in $\mathcal{P}$ can be in the MAGs consistent with $\mathcal{P}$. Before that, we introduce *bridged* in Definition 4.

Definition 4 *(Bridged relative to $\mathbf{V}'$ in H, Wang et al. [19]).* Let H be a partial mixed graph. Denote G a subgraph of H induced by a set of vertices $\mathbf{V}$. Given a set of vertices $\mathbf{V}'$ in H that is disjoint of $\mathbf{V}$, two vertices A and B in a connected circle component of G are *bridged relative to* $\mathbf{V}'$ if either $A = B$ or in each minimal circle path $A(= V_0) \circ\!\!-\!\!\circ V_1 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ V_n \circ\!\!-\!\!\circ B(= V_{n+1})$ from A to B in G, there exists one vertex $V_s, 0 \le s \le n+1$, such that $\mathcal{F}_{V_i} \subseteq \mathcal{F}_{V_{i+1}}, 0 \le i \le s-1$ and $\mathcal{F}_{V_{i+1}} \subseteq \mathcal{F}_{V_i}, s \le i \le n$, where $\mathcal{F}_i = \{V \in \mathbf{V}' \mid V *\!\!-\!\!\circ V_i \text{ in } H\}$. Further, G is *bridged relative to* $\mathbf{V}'$ *in* H if any two vertices in a connected circle component of G are bridged relative to $\mathbf{V}'$.

Proposition 2 *(Wang et al. [19]). Given a PAG $\mathcal{P}$, for any set of vertices $\mathbf{C} \subseteq \{V \mid X \circ\!\!-\!\!* V \text{ in } \mathcal{P}\}$, there exists a MAG $\mathcal{M}$ consistent with $\mathcal{P}$ with $X \leftarrow\!\!* V$, $\forall V \in \mathbf{C}$ and $X \to V$, $\forall V \in \{V \mid X \circ\!\!-\!\!* V \text{ in } \mathcal{P}\} \backslash \mathbf{C}$ if and only if*

(1) $\mathrm{PossDe}(X, \mathcal{P}[-\mathbf{C}]) \cap \mathrm{Pa}(\mathbf{C}, \mathcal{P}) = \emptyset$*;*
(2) *the subgraph $\mathcal{P}[\mathbf{C}]$ of $\mathcal{P}$ induced by $\mathbf{C}$ is a complete graph;*
(3) $\mathcal{P}[\mathrm{PossDe}(X, \mathcal{P}[-\mathbf{C}]) \backslash \{X\}]$ *is bridged relative to $\mathbf{C} \cup \{X\}$ in $\mathcal{P}$.*

Remark 2. Given any set $\mathbf{C} \subseteq \{V \mid X \circ\!\!-\!\!* V \text{ in } \mathcal{P}\}$, we transform $X \circ\!\!-\!\!* V$ to $X \leftarrow\!\!* V$ for $\forall V \in \mathbf{C}$ and transform $X \circ\!\!-\!\!* V$ to $X \!-\!\!* V$ for others, the marks at X are definite. Hence, each set $\mathbf{C}$ dictates a local structure at X.

Hence, for each local structure dictated by $\mathbf{C}$, we can determine the validity by Proposition 2. After enumerating each set $\mathbf{C}$ and using Proposition 2, we can obtain all valid local structures at X. Given each valid local structure, we use the sound and complete

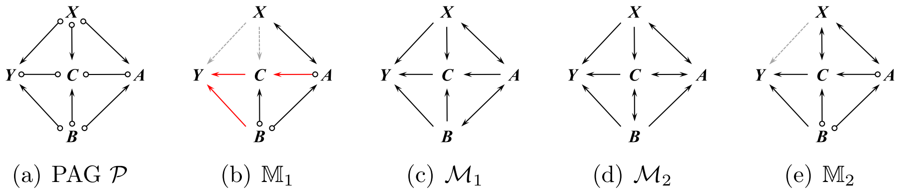

Fig. 2. Fig. 2(a) depicts a PAG $\mathcal{P}$. We first consider all the valid local structures at X. For the local structure dictated by $\mathbf{C} = \{A\}$, we could obtain a maximal local MAG $\mathbb{M}_1$ as Fig. 2(b), where the edges $C \to Y \leftarrow B, A \circ\!\!\to C$ colored by red denotes those oriented by the rules of Wang et al. [19]. Fig. 2(c) and 2(d) depict two MAGs valid to $\mathbb{M}$, where the adjustment sets are $\{A, B\}$ and $\emptyset$, respectively. Fig. 2(e) depicts another maximal local MAG $\mathbb{M}_2$ obtained from the local structure dictated by $\mathbf{C} = \{A, C\}$. The edges with solid line in Fig. 2(b)/ 2(e) are those remained after deleting the edges out of X. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

orientation rules [19] to further orient the PAG with the local structures. We give Fig. 2(a) as an example.[3] The local structure dictated by $\mathbf{C} = \{A\}$ is valid according to Proposition 2. When we introduce this local structure, it can be seen as background knowledge, and thus we further orient the partial graph as Fig. 2(b) with the complete orientation rules of Wang et al. [19] until no rules are triggered. In the following, we call the graph obtained from $\mathcal{P}$ with a valid local structure dictated by $\mathbf{C}$ and the complete orientation rules by *maximal local MAG based on $\mathcal{P}$ and* $\mathbf{C}$, and denote it by $\mathbb{M}$. When it is clear from the context, we call it *maximal local MAG* for short. A MAG $\mathcal{M}$ is *valid to* $\mathbb{M}$ if $\mathcal{M}$ is consistent with $\mathcal{P}$ and has the non-circle marks in $\mathbb{M}$. For each valid local structure at X, we can obtain a maximal local MAG. The graph in Fig. 2(b) is an example of $\mathbb{M}$. By obtaining all valid local structures of X and using the complete orientation rules, we can obtain all maximal local MAGs.

Given a maximal local MAG $\mathbb{M}$, we propose the method to find the adjustment sets in the MAGs valid to $\mathbb{M}$ in the following. We only consider the non-trivial case $Y \in \mathrm{PossDe}(X, \mathbb{M})$, otherwise X has no causal effect on Y according to Lemma 8 in Sec. 3.5.2. A question here is, could we determine a common adjustment set for all the MAGs valid to $\mathbb{M}$? Unfortunately, it is not the case in the presence of latent variables. See $\mathbb{M}$ in Fig. 2(b) for an example, there are MAGs valid to $\mathbb{M}$ in Fig. 2(c) and 2(d) with distinct adjustment sets $\{A, B\}$ and $\emptyset$, respectively. The local structure at X is not sufficient for determining a common adjustment set. We need to consider further what adjustment sets can be in the MAGs valid to $\mathbb{M}$.

A trivial method is to enumerate each MAG $\mathcal{M}$ valid to $\mathbb{M}$, and obtain $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ as the adjustment set if the conditions of Theorem 1 hold. However, the space of MAGs is extremely large. In the worst case, the size is still $\mathcal{O}(3^{d^2/2})$, due to $(d-1)(d-2)/2$ circle edges with 3 types. To circumvent the enumeration, we convert the problem of finding $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ in each enumerated MAG to the problem of determining for any given subset $\mathbf{W}$ of the observed variables, whether there is a MAG $\mathcal{M}$ valid to $\mathbb{M}$ such that $\mathbf{W}$ is an adjustment set in $\mathcal{M}$ as Theorem 1, *i.e.*, $\mathbf{W} = \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ and $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) \cap \mathrm{De}(X, \mathcal{M}) = \emptyset$. The benefit of the conversion is apparent. The size of the space of subset $\mathbf{W}$ is $\mathcal{O}(2^d)$, which is super-exponentially less than that of the space of MAGs $\mathcal{O}(3^{d^2/2})$. Note when $\mathbb{M}$ is a complete graph and all the marks at X are arrowheads, there are exactly 2^{d-2} adjustment sets that result in different causal effects. Hence $\mathcal{O}(2^d)$ is a lower bound of the complexity of finding adjustment sets in MAGs valid to $\mathbb{M}$.

The key here is, for any maximal local MAG $\mathbb{M}$ and set $\mathbf{W}$, we need to determine the existence of MAGs $\mathcal{M}$ valid to $\mathbb{M}$ fulfilling the two conditions (1) $\mathbf{W} = \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ and (2) $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) \cap \mathrm{De}(X, \mathcal{M}) = \emptyset$. To achieve it, we propose a graphical characterization. The idea is that for any given $\mathbb{M}$ and $\mathbf{W}$, we determine *the feasibility of constructing* a MAG $\mathcal{M}$ fulfilling the two conditions. If we could construct such a MAG, then there exist the MAGs aforementioned, otherwise the MAGs do not exist.

For the construction process, we introduce $\bar{\mathbf{W}}$ and *block set* $\mathbf{S}$ in Definition 5. Intuitively, $\bar{\mathbf{W}}$ denotes the set of vertices disjoint of $\mathbf{W}$ that are not allowed to be ancestors of Y in the constructed $\mathcal{M}$, otherwise $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) \backslash \mathbf{W} \neq \emptyset$; and $\mathbf{S}$ denotes a set of vertices located at the paths from $\bar{\mathbf{W}}$ to Y. To prevent $\bar{\mathbf{W}}$ to be ancestors of Y in the constructed $\mathcal{M}$, we need to orient the edges of $\mathbf{S}$ in the paths mentioned above. See Fig. 2(b) for an example. Given $\mathbf{W} = \emptyset$, to construct a MAG $\mathcal{M}$ such that $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) = \mathbf{W}$, it is necessary to restrict that $\{A\}$ is not an ancestor of Y in $\mathcal{M}$, otherwise $A \in \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) \backslash \mathbf{W}$. To achieve it, $A \circ\!\!\to C$ must be oriented as bi-directed. Hence we define $\bar{\mathbf{W}} = \{A\}$ and $\mathbf{S} \supseteq \{C\}$ in Definition 5.

Definition 5. Given a set of vertices $\mathbf{W}$ in a maximal local MAG $\mathbb{M}$, we define a set of vertices $\bar{\mathbf{W}}$ as $V \in \bar{\mathbf{W}}$ if and only if $V \in \mathrm{PossAn}(Y, \mathbb{M}) \backslash \mathbf{W}$ and there exists a collider path beginning with an arrowhead from X to V where each non-endpoint vertex belongs to $\mathbf{W}$. We say $\mathbf{S}$ is a *block set* if $\mathrm{Anc}(Y \cup \mathbf{W}, \mathbb{M}) \cap [\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}) \backslash \bar{\mathbf{W}}] \subseteq \mathbf{S} \subseteq \mathrm{PossAn}(Y \cup \mathbf{W}, \mathbb{M}) \cap [\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}) \backslash \bar{\mathbf{W}}]$.

We then present *potential adjustment set* in Definition 6. According to Definition 3 and Theorem 1, there is $\mathbf{W} = \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ and $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) \cap \mathrm{De}(X, \mathcal{M}) = \emptyset$ in some MAG $\mathcal{M}$ valid to $\mathbb{M}$ *only if* $\mathbf{W}$ is a potential adjustment set. The proof is included in that of a later result Theorem 3.

[3] The examples in Fig. 2(b) and 2(e) of the preliminary version [22] are flawed, as $B \circ\!\!\to Y$ should be $B \to Y$, transformed by $\mathcal{R}'_4$. We revise the examples in this paper.

Definition 6. In a maximal local MAG $\mathbb{M}$, $\mathbf{W}$ is a *potential adjustment set* if

(1) $\forall V \in \mathbf{W}$, there is a collider path $X \leftrightarrow \cdots \leftarrow\!\!* V$ such that each non-endpoint belongs to $\mathbf{W}$, and there is a possible directed path from V to Y that does not go through the vertices in $\bar{\mathbf{W}}$;
(2) $\mathbf{W} \cap \mathrm{PossDe}(X, \mathbb{M}) = \emptyset$;
(3) $\bar{\mathbf{W}} \cap \mathrm{Anc}(Y \cup \mathbf{W}, \mathbb{M}) = \emptyset$.

We present the condition for the existence of MAGs with $\mathbf{W}$ being an adjustment set in Theorem 2. To prove the existence of MAGs, we show the construction method by Algorithm 3 in Sec. 3.5.2. The complexity of verifying the conditions in Theorem 2 for a block set $\mathbf{S}$ is $\mathcal{O}(d^3)$.

Theorem 2. *Given a maximal local MAG $\mathbb{M}$, for any potential adjustment set $\mathbf{W}$, there exists a MAG $\mathcal{M}$ valid to $\mathbb{M}$ such that $\mathbf{W}$ is an adjustment set in $\mathcal{M}$ if there exists a block set $\mathbf{S}$ such that*

(1) $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}]) \cap \mathrm{Pa}(\mathbf{S}, \mathbb{M}) = \emptyset$;
(2) $\mathbb{M}[\mathbf{S}_V]$ *is a complete graph for any* $V \in \bar{\mathbf{W}}$, *where* $\mathbf{S}_V = \{V' \in \mathbf{S} | V \circ\!\!\rightarrow\!* V' \text{ in } \mathbb{M}\}$;
(3) $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ *is bridged relative to* $\mathbf{S}$ *in* $\mathbb{M}$.

Remark 3. The initial target of Theorem 2 is to present a sufficient condition for the existence of MAG $\mathcal{M}$ such that $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) = \mathbf{W}$, where $\mathbf{W}$ is a given potential adjustment set. However, following our MAG construction method as Algorithm 3 in Sec. 3.5.2, we can only construct a MAG $\mathcal{M}$ where $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) \subseteq \mathbf{W}$ and $\mathbf{W} \cap \mathrm{De}(X, \mathcal{M}) = \emptyset$, as Algorithm 3 cannot ensure that each vertex V in $\mathbf{W}$ is an ancestor of Y in the constructed MAG. Nevertheless, there is always a good property that $\mathbf{W} \backslash \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) \perp Y \mid \{X, \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})\}$ if $\mathbf{W} \backslash \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ is non-empty in the constructed $\mathcal{M}$. With this property, it is direct that using $\mathbf{W}$ and $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ as the adjustment sets can lead to the *same* causal effect, *i.e.*, $\mathbf{W}$ is equivalent to $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ in the sense of estimating the causal effect value. See the proof of Theorem 2 in Sec. 3.5.2 for details.

Hence, for each subset $\mathbf{W}$ of $\mathbf{V}(\mathcal{P}) \backslash \{X, Y\}$, we can determine whether it can be an adjustment set in some MAG valid to $\mathbb{M}$ by determining whether it is a potential adjustment set and whether the three conditions in Theorem 2 are satisfied for a block set $\mathbf{S}$ if so. By considering all the subsets, we can find all the valid adjustment sets in MAGs valid to $\mathbb{M}$.

Finally, there is one issue remaining to address: for the adjustment set $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ in any MAG $\mathcal{M}$ valid to $\mathbb{M}$, can we always find it by the process above? We present Theorem 3 to show that the adjustment set $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ in any MAG $\mathcal{M}$ valid to $\mathbb{M}$ is a potential adjustment set and satisfies the three conditions in Theorem 2 for a block set $\mathbf{S}$.

Theorem 3. *Given a maximal local MAG $\mathbb{M}$, suppose a MAG $\mathcal{M}$ valid to $\mathbb{M}$ such that there exists an adjustment set relative to (X, Y). Let $\mathbf{W}$ be $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$. Then $\mathbf{W}$ is a potential adjustment set in $\mathbb{M}$ and there exists a block set $\mathbf{S}$ such that*

(1) $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}]) \cap \mathrm{Pa}(\mathbf{S}, \mathbb{M}) = \emptyset$;
(2) $\mathbb{M}[\mathbf{S}_V]$ *is a complete graph for any* $V \in \bar{\mathbf{W}}$, *where* $\mathbf{S}_V = \{V' \in \mathbf{S} | V \circ\!\!\rightarrow\!* V' \text{ in } \mathbb{M}\}$;
(3) $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ *is bridged relative to* $\mathbf{S}$ *in* $\mathbb{M}$.

3.3. The algorithm for set determination

Based on the theoretical results above, we present a direct method in Algorithm 1, to find all the valid adjustment sets given a PAG $\mathcal{P}$ and then determine the set of possible causal effects by (1). If X is not a possible ancestor of Y in $\mathcal{P}$, X has no effect on Y in every MAG consistent with $\mathcal{P}$, thereby returning no causal effect. If the causal effect is identifiable in $\mathcal{P}$ by Proposition 1, we obtain the adjustment set directly according to Proposition 1. When it is not identifiable, we first find all valid local structures at X based on Proposition 2, and obtain the corresponding maximal local MAGs. In each maximal local MAG $\mathbb{M}$, for any set $\mathbf{W} \subseteq \mathbf{V}(\mathcal{P}) \backslash \{X, Y\}$, if $\mathbf{W}$ is a potential adjustment set as Definition 6, we determine whether it is an adjustment set in some $\mathcal{M}$ valid to $\mathbb{M}$ by Theorem 2. Note the sufficient condition in Theorem 2 is that *there exists a block set* such that the three conditions are fulfilled, hence we search for each set $\mathbf{S}$ on Line 10. Theorem 4 indicates that the set of causal effects by Algorithm 1 is equal to that by enumerating all the causal graphs consistent with $\mathcal{P}$.

Theorem 4. *Given a PAG $\mathcal{P}$, denote the set of (possible) causal effects in the DAGs consistent with $\mathcal{P}$ which can be estimated with observational data by covariate adjustment and the set of (possible) causal effects obtained from Algorithm 1 by $\mathrm{CE}(\mathcal{P})$ and $\widehat{\mathrm{CE}}(\mathcal{P})$. There is $\mathrm{CE}(\mathcal{P}) \overset{set}{=} \widehat{\mathrm{CE}}(\mathcal{P})$.*

Proof. The result can be directly concluded according to Theorem 1, 2, and 3. As we consider all the valid local structures at X in Algorithm 1, it suffices to show that the set of causal effects $\widehat{\mathrm{CE}}(\mathbb{M})$ identified with $\widehat{\mathrm{AS}}(\mathbb{M})$ by (1) is equal to the set $\mathrm{CE}(\mathbb{M})$, where

Algorithm 1: PAGcauses.

Input: PAG $\mathcal{P}$, X, Y
1 $\widehat{\mathrm{AS}}(\mathcal{P}) = \emptyset$ // Record all the valid adjustment sets;
2 **if** $X \notin \mathrm{PossAn}(Y, \mathcal{P})$ **then return** No causal effects;
3 **if** *the conditions in Proposition 1 are satisfied for* $\mathcal{P}$ **then**
4   **return** $\widehat{\mathrm{AS}}(\mathcal{P}) \leftarrow \{\text{D-SEP}(X, Y, \mathcal{P}_{\underset{\sim}{X}})\}$ // Prop. 1
5 **for** *each set* $\mathbf{C} \subseteq \{V \mid V \ast\!\!\multimap X \text{ in } \mathcal{P}\}$ **do**
6   **if** *the three conditions in Proposition 2 are satisfied* **then**
7   Obtain a maximal local MAG $\mathbb{M}$ based on $\mathcal{P}$ and $\mathbf{C}$;
8   Find all potential adjustment sets $\mathbf{W}_1, \mathbf{W}_2, \cdots$ given $\mathbb{M}$ according to Definition 6;
9   **for** *each potential adjustment set* $\mathbf{W}_i$ **do**
10   **for** *each block set* $\mathbf{S}$ **do**
11   **if** *the three conditions in Theorem 2 are satisfied given* $\mathbf{S}$ **then**
12                     $\widehat{\mathrm{AS}}(\mathcal{P}) \leftarrow \widehat{\mathrm{AS}}(\mathcal{P}) \cup \{\mathbf{W}_i\}$;
13   **break** // Break the loop of $\mathbf{S}$;

Output: Set of causal effects via adjustment in the given PAG $\mathcal{P}$ identified with $\widehat{\mathrm{AS}}(\mathcal{P})$ by (1)

$\widehat{\mathrm{AS}}(\mathbb{M})$ denotes the obtained adjustment sets on Line 8-13 of Algorithm 1, and $\mathrm{CE}(\mathbb{M})$ denotes the set of causal effects that can be identified with observational data in the DAGs represented by any MAGs valid to $\mathbb{M}$.

As shown by Theorem 1, the set of causal effects $\mathrm{CE}(\mathbb{M})$ in the DAGs represented by the MAG $\mathcal{M}$ valid to $\mathbb{M}$ is equal to that identified with the set of $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ for each $\mathcal{M}$ valid to $\mathbb{M}$ such that $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) \cap \mathrm{De}(X, \mathcal{M}) = \emptyset$. According to Theorem 3, there is evidently $\mathrm{CE}(\mathbb{M}) \subseteq \widehat{\mathrm{CE}}(\mathbb{M})$. And Theorem 2 implies that for each set $\mathbf{V}'$ in $\widehat{\mathrm{AS}}(\mathbb{M})$, there is some MAG $\mathcal{M}$ valid to $\mathbb{M}$ such that the set $\mathbf{V}'$ is an adjustment set in $\mathcal{M}$, *i.e.*, the adjustment set $\mathbf{V}'$ implies the same causal effect by (1) as that of $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ in some MAG $\mathcal{M}$ valid to $\mathbb{M}$. Hence $\widehat{\mathrm{CE}}(\mathbb{M}) \subseteq \mathrm{CE}(\mathbb{M})$. We conclude $\widehat{\mathrm{CE}}(\mathbb{M}) \overset{set}{=} \mathrm{CE}(\mathbb{M})$, thus $\widehat{\mathrm{CE}}(\mathcal{P}) \overset{set}{=} \mathrm{CE}(\mathcal{P})$. $\square$

We note that the enumerations of all the subsets $\mathbf{W} \subseteq \mathbf{V}(\mathcal{P}) \backslash \{X, Y\}$ on Line 8 of Algorithm 1 takes a $\mathcal{O}(2^d)$ computation complexity. In fact, it is usually not necessary to consider all the subsets. In the following, we further improve the efficiency of this part from two aspects, through which the number of sets considered on Line 8 grows far more slowly than exponentially in general.

Aspect 1. Given a maximal local MAG $\mathbb{M}$, we can determine some vertices that *must* belong to $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ in *any* MAGs $\mathcal{M}$ valid to $\mathbb{M}$. See $\mathbb{M}$ in Fig. 2(e) for an example, there is a collider path $X \leftrightarrow C$ and C is an ancestor of Y, thus $C \in \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ according to Definition 3 for any $\mathcal{M}$ valid to $\mathbb{M}$. Hence, for any set $\mathbf{W}$, if there exists some MAG $\mathcal{M}$ valid to $\mathbb{M}$ in Fig. 2(e) such that $\mathbf{W} = \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$, $\mathcal{M}$ contains the vertices aforementioned. Following this idea, we present $\text{DD-SEP}(X, Y, \mathbb{M}_{\underset{\sim}{X}})$, short for *definite* $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$, in Definition 7. $\mathbb{M}_{\underset{\sim}{X}}$ denotes the graph obtained from $\mathbb{M}$ by deleting all directed edges out of X. Proposition 3 implies that for any MAG $\mathcal{M}$ valid to $\mathbb{M}$ such that an adjustment set exists, there is $\text{DD-SEP}(X, Y, \mathbb{M}_{\underset{\sim}{X}}) \subseteq \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$. Hence, we only need to consider all the sets $\mathbf{W}$ that contains $\text{DD-SEP}(X, Y, \mathbb{M}_{\underset{\sim}{X}})$ on Line 8 of Algorithm 1.

Definition 7 *(*$\text{DD-SEP}(X, Y, \mathbb{M}_{\underset{\sim}{X}})$*).* Let $\mathbb{M}$ be a maximal local MAG. $V \in \text{DD-SEP}(X, Y, \mathbb{M}_{\underset{\sim}{X}})$ if and only if there is $X \leftrightarrow V_1 \leftrightarrow \cdots \leftrightarrow V_{k-1} \leftarrow\!\!\ast\, V$ in $\mathbb{M}_{\underset{\sim}{X}}$, where (1) $Y \in \mathrm{PossDe}(X, \mathbb{M})$; (2) $V_1, \cdots, V_{k-1} \in \text{DD-SEP}(X, Y, \mathbb{M}_{\underset{\sim}{X}})$; (3) $V \in \mathrm{Anc}(Y, \mathbb{M})$ or the subgraph $\mathbb{M}[\mathcal{Q}_V]$ is not a complete graph, where $\mathcal{Q}_V = \{V' \in \mathrm{Anc}(Y, \mathbb{M}) \mid V \circ\!\!\!-\!\!\ast\, V' \text{ in } \mathbb{M}\}$.

Proposition 3. *Given a maximal local MAG* $\mathbb{M}$ *based on* $\mathcal{P}$ *and* $\mathbf{C}$*, if* $V \in \text{DD-SEP}(X, Y, \mathbb{M}_{\underset{\sim}{X}})$*, then* $V \in \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ *in any MAG* $\mathcal{M}$ *valid to* $\mathbb{M}$ *such that there exists an adjustment set relative to* (X, Y)*.*

Aspect 2. On Line 8 of Algorithm 1, we aim to find mere all the potential adjustment sets $\mathbf{W} \subseteq \mathbf{V}(\mathcal{M}) \backslash \{X, Y\}$. According to the first property of Definition 6, for each potential adjustment set $\mathbf{W}$, all the vertices in $\mathbf{W}$ and X form a connected component, where there exists a collider path between any two vertices in $\mathbf{W} \cup \{X\}$. See Fig. 2(b) for an example. $\mathbf{W} = \{B\}$ cannot be a potential adjustment set, as $\{X, B\}$ does not form a connected component where there exists a collider path between X and B; while $\mathbf{W} = \{A\}$ can be a potential adjustment set. Hence, instead of enumerating all the subsets, we can first recursively find all the subsets which form a connected component with X such that there exists a collider path between any two vertices. It suffices to determine whether these sets are potential adjustment sets.

Combining the two aspects above, we propose an $\mathbb{M}$-dependent method in Algorithm 2 to execute Line 8-13 of Algorithm 1. We first obtain $\text{DD-SEP}(X, Y, \mathbb{M}_{\underset{\sim}{X}})$, then find all the sets $\mathbf{W}$ fulfilling $\text{DD-SEP}(X, Y, \mathbb{M}_{\underset{\sim}{X}}) \subseteq \mathbf{W} \subseteq \mathbf{V}(\mathbb{M}) \backslash \{X, Y\}$ which form a connected component with X such that there is a collider path between any two vertices. See the maximal local MAG in Fig. 2(b) as an example, where $\text{DD-SEP}(X, Y, \mathbb{M}_{\underset{\sim}{X}}) = \emptyset$. The edges with solid lines are those kept in $\mathbb{M}_{\underset{\sim}{X}}$. Only $\emptyset, \{A\}, \{A, B\}$ can form a connected component with X, we thus consider whether the three sets are potential adjustment sets and whether the conditions of Theorem 2 are fulfilled for some block set if they are.

Algorithm 2: $\mathbb{M}$-dependently finding adjustment sets in MAGs valid to $\mathbb{M}$ (Line 8-13 of Algorithm 1).

Input: a maximal local MAG $\mathbb{M}$, X, Y
$\widehat{\mathrm{AS}}(\mathbb{M}) = \emptyset$ // To store the adjustment set for (X, Y) in each MAG $\mathcal{M}$ valid to $\mathbb{M}$;
$\mathbf{W} \leftarrow$ DD-SEP$(X, Y, \mathbb{M})$, and obtain $\bar{\mathbf{W}}$ as Definition 5;
`Adjust`$(\mathbb{M}, X, Y, \mathbf{W}, \bar{\mathbf{W}})$;
Algorithm `Adjust`$(\mathbb{M}, X, Y, \mathbf{W}, \mathbf{I})$
  if $\mathbf{I} = \emptyset$ *and* $\mathbf{W}$ *is a potential adjustment set* **then**
    for *each block set* $\mathbf{S}$ **do**
      if *the three conditions in Theorem 2 are satisfied given* $\mathbf{S}$ **then**
        $\widehat{\mathrm{AS}}(\mathbb{M}) \leftarrow \widehat{\mathrm{AS}}(\mathbb{M}) \cup \{\mathbf{W}\}$;
        break // Break the loop of $\mathbf{S}$;
  else if $\mathbf{I} \neq \emptyset$ **then**
    for $\emptyset \subset \mathbf{V}' \subseteq \mathbf{I}$, *denote* $\mathbf{V}'' = \mathbf{I} \backslash \mathbf{V}'$ **do**
      $\mathbf{W} \leftarrow \mathbf{W} \cup \mathbf{V}'$, and obtain $\bar{\mathbf{W}}$ in Definition 5;
      `Adjust`$(\mathbb{M}, X, Y, \mathbf{W}, \bar{\mathbf{W}} \backslash \mathbf{V}'')$;
Output: $\widehat{\mathrm{AS}}(\mathbb{M})$ // Adjustment sets in MAGs valid to $\mathbb{M}$

3.4. Complexity analysis

In this part, we present a worst-case complexity analysis of PAGcauses and the baseline method LV-IDA [17] to determine the set of possible causal effects of X on Y in a complete graph including d vertices. In this case, the PAG is a complete circle component without any arrowheads or tails. Note the number of causal effect values in the worst case grows exponentially with respect to d. Hence the task of set determination cannot be finished within polynomial time. To evaluate the complexity in general cases, we conduct empirical analysis in Sec. 5.

For LV-IDA, there are $3^{(d^2-d)/2}$ MAGs enumerated. For each MAG, we need to first judge whether the enumerated MAG is consistent with $\mathcal{P}$, it spends $\mathcal{O}(d^3)$ at least [32,33]. Then if the MAG is consistent with $\mathcal{P}$, we find D-SEP$(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ in $\mathcal{O}(d)$ at least. Hence, the total complexity is $\mathcal{O}(d^4 3^{(d^2-d)/2})$. Another method of LV-IDA is obtaining the MAG consistent with $\mathcal{P}$ by transformation, where the obtained MAG is always consistent with $\mathcal{P}$ [34,35]. If we adopt this method, it is hard to analyze the complexity accurately as it is unavoidable that many repeatable MAGs are obtained in the process and it is hard to know when the transformation should stop. In the most ideal case, even if we assume that we *never* obtain a repeatable MAG in the transformation process and there is an oracle tell us the time point when we find all the MAGs consistent with $\mathcal{P}$, we could obtain a lower bound of computational complexity as $\Omega(d 3^{(d^2-d)/2}) \times d^2 = \Omega(d^3 3^{(d^2-d)/2})$, where the extra d^2 is the complexity of verifying the transformation characterization by Lemma 1 of Zhang and Spirtes [34].

Next we consider the complexity of PAGcauses. Note the first and second conditions of Theorem 2 can be determined in $\mathcal{O}(d^3)$ for a given block set $\mathbf{S}$. For the determination of the third condition, we exploit Lemma 5 to achieve it. The orientation in Algorithm 3 and the testing of whether there are edges oriented with different directions or new unshielded colliders can be achieved in $\mathcal{O}(d^3)$. Hence the complexity of judging the three conditions of Theorem 2 given a maximal local MAG and a block set $\mathbf{S}$ is $\mathcal{O}(d^3)$.

As our method first determines all valid local structures at X, then finds adjustment sets in the MAGs valid to each maximal local MAG, we analyze the complexity of finding adjustment sets in the MAGs valid to a given maximal local MAG at first. Suppose in a maximal local MAG $\mathbb{M}$ which is obtained from $\mathcal{P}$ and $\mathbf{C}$ which dictates a local structure of X, there are i vertices with edges $*\!\!\rightarrow X$ and $d-i-1$ vertices with edges $*\!\!- X$. In this case, the complexity $T(i)$ of finding all adjustment sets in $\mathbb{M}$ is

$$T(i) \leq \sum_{k=0}^{i} C_i^k 2^{k+d-i-1} \mathcal{O}(d^3) = 2^{d-i-1} 3^i \mathcal{O}(d^3),$$

where k denotes the number of vertices in $\mathbf{W}$ and C_i^k is because there are C_i^k sets $\mathbf{W}$ such that there are k vertices in $\mathbf{W}$. When there is k number of vertices in $\mathbf{W}$, since every vertex is adjacent to X, there must be $i-k$ vertex in $\bar{\mathbf{W}}$. Hence there are at most $d-1-i+k$ vertices in each block set $\mathbf{S}$ and thus there are at most $2^{k+d-i-1}$ sets of $\mathbf{S}$. $\mathcal{O}(d^3)$ is the complexity of judging the conditions in Theorem 2.

Next we consider the complexity $C(d)$ of finding adjustment sets in the MAGs consistent with $\mathcal{P}$ by PAGcauses. According to the result above, it directly concludes that the complexity is

$$C(d) \leq \sum_{i=0}^{d-1} C_{d-1}^i \mathcal{O}(d^3)(2^{d-i-1} 3^i \mathcal{O}(d^3)) = 2^{d-1} \sum_{i=0}^{d-1} (3/2)^i \mathcal{O}(d^6) = \mathcal{O}(5^d d^6),$$

where C_{d-1}^i is because there are C_{d-1}^i sets $\mathbf{C}$ which dictates a local structure at X such that there are i vertices with edges $*\!\!\rightarrow X$, the first $\mathcal{O}(d^3)$ is the complexity of judging Lemma 2.

Hence, the complexity of PAGcauses is super-exponentially less than that by (locally) enumerating MAGs.

3.5. Proof

3.5.1. Proof of Theorem 1

Proof. We first prove the "if" statement. We could construct a DAG $\mathcal{D}$ by remaining the directed edges in $\mathcal{M}$ and adding a sub-structure $V_i \leftarrow L_{ij} \rightarrow V_j$ with a latent variable L_{ij} if there is $V_i \leftrightarrow V_j$ in $\mathcal{M}$. The paths from X to Y in $\mathcal{M}_{\underset{\sim}{X}}$ are *totally same* as the back-door paths from X to Y in $\mathcal{D}$, *i.e.*, the paths from X to Y in $\mathcal{D}_{\underline{X}}$. Since X is an ancestor of Y, it is evident that X and Y are not adjacent in $\mathcal{M}_{\underset{\sim}{X}}$. And $\mathcal{M}_{\underset{\sim}{X}}$ is also a MAG according to Prop. 3.5 and Corollary 4.6 of Richardson et al. [25]. By Lemma 4.1 of Maathuis et al. [31], X and Y are m-separated by D-SEP$(X,Y,\mathcal{M}_{\underset{\sim}{X}})$ in $\mathcal{M}_{\underset{\sim}{X}}$. It is direct that X and Y are d-separated by D-SEP$(X,Y,\mathcal{M}_{\underset{\sim}{X}})$ in $\mathcal{D}_{\underline{X}}$. The reason is that the only difference between $\mathcal{M}_{\underset{\sim}{X}}$ and $\mathcal{D}_{\underline{X}}$ is that the bi-directed edges $V_i \leftrightarrow V_j$ in $\mathcal{M}_{\underset{\sim}{X}}$ are $V_i \leftarrow L_{ij} \rightarrow V_j$ in $\mathcal{D}_{\underline{X}}$, thus all the paths from X to Y in $\mathcal{M}_{\underset{\sim}{X}}$ and $\mathcal{D}_{\underline{X}}$ have the same colliders and non-colliders, and each latent variable $L_{i,j}$ in $\mathcal{D}_{\underline{X}}$ is a non-collider so that it does not influence whether a path is active or d-separated by D-SEP$(X,Y,\mathcal{M}_{\underset{\sim}{X}})$ in $\mathcal{D}_{\underline{X}}$. Hence the Blocking condition of adjustment criterion in Definition 2 is fulfilled in $\mathcal{D}$ by D-SEP$(X,Y,\mathcal{M}_{\underset{\sim}{X}})$. Due to D-SEP$(X,Y,\mathcal{M}_{\underset{\sim}{X}}) \cap \mathrm{De}(X,\mathcal{M}) = \emptyset$, the Forbidden set condition is evidently fulfilled. Hence D-SEP$(X,Y,\mathcal{M}_{\underset{\sim}{X}})$ is an adjustment set in $\mathcal{D}$.

For the "only if" statement, we will first prove that D-SEP$(X,Y,\mathcal{M}_{\underset{\sim}{X}})$ can block all non-causal paths from X to Y in DAG $\mathcal{D}$. And then we show D-SEP$(X,Y,\mathcal{M}_{\underset{\sim}{X}}) \cap \mathrm{De}(X,\mathcal{M}) = \emptyset$.

Denote D-SEP$(X,Y,\mathcal{M}_{\underset{\sim}{X}})$ by $\mathbf{D}$. For contradiction, suppose in $\mathcal{D}$ there is a non-causal active path p as $X \leftarrow \cdots S_1, \cdots, S_2, \cdots, S_i, \cdots, S_n, \cdots, Y$ relative to $\mathbf{D}$, where only $S_0(=X), S_1, S_2, S_3, \cdots, S_i, \cdots, S_n, S_{n+1}(=Y)$ denotes the vertices in $\mathbf{V}(\mathcal{M})$ and there could be vertices except for $S_0(=X), S_1, S_2, S_3, \cdots, S_i, \cdots, S_n, S_{n+1}(=Y)$ in the path that do not belong to $\mathbf{V}(\mathcal{M})$. $\bigoplus$ denotes the concatenation of paths. For a path p, let $p[V_i, V_j]$ denote the sub-path of p from V_i to V_j.

We first present some facts given the conditions above.

Fact 1. *All the colliders in p belongs to $\mathbf{D}$, and non-colliders belong to $\mathbf{V}(\mathcal{M}) \backslash \mathbf{D}$.*

Proof of Fact 1. It is directly concluded according to the definition of active path.

Fact 2. *All the vertices in $\mathbf{D}$ are ancestors of either X or Y in $\mathcal{M}_{\underset{\sim}{X}}$.*

Proof of Fact 2. It is due to the definition of D-SEP$(X,Y,\mathcal{M}_{\underset{\sim}{X}})$.

Fact 3. *If in $\mathcal{D}$ there exists a directed path from V to V' where $V \in \mathbf{V}(\mathcal{M})$, $V' \in \mathbf{D}$, then V is ancestor of either X or Y in $\mathcal{M}_{\underset{\sim}{X}}$.*

Proof of Fact 3. It is concluded directly by Fact 2 and the algorithm to obtain a MAG based on $\mathcal{D}$ of Zhang [6].

Fact 4. *Suppose in p there is a sub-path $X \leftarrow \cdots \rightarrow S_1 \leftarrow \cdots \rightarrow S_2 \leftarrow \cdots \cdots \rightarrow S_{k-1} \leftarrow \cdots S_k$ where $S_1, S_2, \cdots, S_{k-1} \in \mathbf{D}$, if S_k is an ancestor of X or Y in $\mathcal{M}_{\underset{\sim}{X}}$, then X cannot be an ancestor of S_k in $\mathcal{D}$.*

Proof of Fact 4. Suppose X is an ancestor of S_k in $\mathcal{D}$. According to the condition, S_k is an ancestor of Y in $\mathcal{M}_{\underset{\sim}{X}}$. Hence there is a directed path from X to Y across S_k in $\mathcal{D}$. Denote the directed path from S_k to Y in $\mathcal{D}$ by p_1. According the adjustment criterion, if there exists an adjustment set $\mathbf{W}$ comprised of some vertices in $\mathbf{V}(\mathcal{M})$, $\mathbf{W}$ cannot contain the vertices in p_1. We consider the path $p[X, S_k] \bigoplus p_1$, where $S_1, S_2, \cdots, S_{k-1}$ in $p[X, S_k]$ are colliders and S_k is non-colliders due to the directed path from S_k to Y. $\mathbf{W}$ is required to block this path, hence there exists some vertex V in $S_1, S_2, \cdots, S_{k-1}$ such that all the descendants of V does not belong to $\mathbf{W}$. Suppose the nearest vertex to X in $p[X, S_k]$ whose descendants do not belong to $\mathbf{W}$ is S_j. If there is a directed path p_2 from S_j to Y in $\mathcal{D}$ that does not go through X, then $p[X, S_j] \bigoplus p_2$ is active relative to $\mathbf{W}$, contradiction with the blocking condition of adjustment criterion; if there is not a directed path from S_j to Y in $\mathcal{D}$ that does not go through X, there must be a directed path p_3 from S_j to X in $\mathcal{D}$ since S_j is an ancestor of either X or Y in $\mathcal{D}$. In this case we consider a new path $p_3 \bigoplus p[S_j, S_k] \bigoplus p_2$. To block this path, there is another vertex V in $S_{j+1}, S_{j+2}, \cdots, S_{k-1}$ such that all the descendants of V does not belong to $\mathbf{W}$. Suppose the nearest vertex to S_j whose descendants do not belong to $\mathbf{W}$ in $p[S_j, S_k]$ is S_t. If there is a directed path p_4 from S_t to Y that does not go through X, there is an active path $p_3 \bigoplus p[S_j, S_t] \bigoplus p_4$ relative to $\mathbf{W}$, contradicting with the blocking condition of adjustment criterion, hence there is directed path p_5 from S_t to X where each vertex does not belong to $\mathbf{W}$. As the process above, we consider $p_5 \bigoplus p[S_t, S_k] \bigoplus p_2$ instead. Repeat the process above, if there is not a contradiction, there must be a vertex S_m such that there is a directed path p_6 from S_m to X where each vertex does not belong to $\mathbf{W}$, and for each non-endpoint vertex V in $p[S_m, S_k]$ there is at least one descendant of V belonging to $\mathbf{W}$. In this case we have an active non-causal path $p_6 \bigoplus p[S_m, S_k] \bigoplus p_2$, contradicting with the blocking condition of adjustment criterion. Hence, X cannot be an ancestor of S_k.

Fact 5. *If $S_1, S_2, \cdots, S_k$ are colliders in p, then there is a collider path from X to S_{k+1} beginning with an arrowhead at X in $\mathcal{M}$, i.e., in the form of $X \leftrightarrow \cdots \leftrightarrow \leftarrow\!\!* \; S_{k+1}$.*

Proof of Fact 5. According to fact 1, $S_1, \cdots, S_k \in \mathbf{D}$. We consider the sub-path $p[X, S_{k+1}]$. Note the sub-path $p[S_i, S_{i+1}], \forall 0 \le i \le k$ is an inducing path relative to $\langle \mathbf{V}(\mathcal{D}) \backslash \mathbf{V}(\mathcal{M}), \emptyset \rangle$. Hence we can always find a longest inducing path relative to $\langle \mathbf{V}(\mathcal{D}) \backslash \mathbf{V}(\mathcal{M}), \emptyset \rangle$ which is a sub-path of p starting by X. Denote the longest inducing path by $p[X, S_i]$. If $i = k+1$, there is evidently $X \leftarrow\!\!* S_{k+1}$ because $p[X, S_{k+1}]$ is an inducing path and fact 4 (there cannot be an edge $X \to S_{k+1}$ in $\mathcal{M}$ according to fact 4), the result holds. Hence we only consider $i \le k$ below. We will prove there is (a) $X \leftrightarrow S_i$; (b) $S_{i+1} *\!\!\rightarrow S_i$ in $\mathcal{M}$.

For the proof of (a), $X \leftarrow\!\!* S_i$ in $\mathcal{M}$ can be directly concluded because $p[X, S_i]$ is an inducing path and fact 4. If $S_i \to X$ in $\mathcal{M}$, S_i is an ancestor of X in $\mathcal{D}$, thus $p[X, S_i] \bigoplus p[S_i, S_{i+1}]$ is also an inducing path because S_i is a collider in p, contradicting with the premise that $p[X, S_i]$ is the longest inducing path staring by X. Hence there can only be $X \leftrightarrow S_i$ in $\mathcal{M}$. For the proof of (b), suppose $S_{i+1} \leftarrow S_i$ in $\mathcal{M}$, hence S_{i+1} is a descendant of S_i in $\mathcal{D}$. Considering the inducing path $p[X, S_i]$, each collider is an ancestor of either X or S_i, thus each collider is an ancestor of either X or S_{i+1}. Hence $p[X, S_i] \bigoplus p[S_i, S_{i+1}]$ is also an inducing path because S_i is a collider in p, contradicting with the premise that $p[X, S_i]$ is the longest inducing path staring by X. Hence S_{i+1} is not a descendant of S_i in $\mathcal{D}$, thus in $\mathcal{M}$ there is $S_i \leftarrow\!\!* S_{i+1}$.

According to the result above, if $i = k$, there is $X \leftrightarrow S_k \leftarrow\!\!* S_{k+1}$ in $\mathcal{M}$, we get the desired result. Hence we consider $i < k$ below.

Then, we could find the longest inducing path relative to $\langle \mathbf{V}(\mathcal{D}) \backslash \mathbf{V}(\mathcal{M}), \emptyset \rangle$ which is a sub-path of $p[S_i, S_{k+1}]$ starting by S_i. Suppose the path is $p[S_i, S_j]$. S_i and S_j is adjacent in $\mathcal{M}$ due to the inducing path. We first prove there is not $S_i \to S_j$ in $\mathcal{M}$. Otherwise, all the vertices in $p[S_i, S_j]$ are ancestors of S_j according to the definition of inducing path. Since all the vertices in $p[X, S_i]$ are ancestors of X or S_i, we have a new inducing path $p[X, S_j]$ since each vertex is an ancestor of X or S_j and meanwhile a collider because $S_1, S_2, \cdots, S_k$ are ancestors in p. It contradicts the fact that $p[X, S_i]$ is the longest inducing path starting by X. Hence there is $S_i \leftarrow\!\!* S_j$ in $\mathcal{M}$. If $j = k+1$, we get the desired result. If $j < k+1$, we will prove there is $S_i \leftrightarrow S_j$ in $\mathcal{M}$. Otherwise, if it is $S_j \to S_i$ in $\mathcal{M}$, S_j is an ancestor of S_i in $\mathcal{D}$, hence $p[S_i, S_j] \bigoplus p[S_j, S_{j+1}]$ is an inducing path relative to $\langle \mathbf{V}(\mathcal{D}) \backslash \mathbf{V}(\mathcal{M}), \emptyset \rangle$, contradicting with the fact that $p[S_i, S_j]$ is the longest inducing path that is a sub-path of $p[S_i, S_{k+1}]$ starting by S_i. Hence, there is $X \leftrightarrow S_i \leftrightarrow S_j$ in $\mathcal{M}$. Similarly, we can find a longest inducing path $p[S_j, S_t]$ starting by S_j such that $S_j \leftarrow\!\!* S_t$ (otherwise $p[S_i, S_j] \bigoplus p[S_j, S_t]$ is an inducing path), and in $\mathcal{M}$ there is $X \leftrightarrow S_i \leftrightarrow S_j \leftarrow\!\!* S_{k+1}$ if $t = k+1$ and there is $X \leftrightarrow S_i \leftrightarrow S_j \leftrightarrow S_t$ if $t < k+1$. Repeat the process above, we could always find a collider path $X \leftrightarrow \cdots \leftarrow\!\!* S_{k+1}$ in $\mathcal{M}$.

With the five facts above, we first prove that all of $S_0, S_1, S_2, \cdots, S_n$ are colliders in p by mathematical induction. Since p is active relative to $\mathbf{W}$, where only $S_0, S_1, S_2, \cdots, S_n, S_{n+1}$ are observed variables in $\mathcal{M}$. It is evident that there is an edge between S_i and S_{i+1} in $\mathcal{M}$, for $\forall 0 \le i \le n$.

We first prove that S_1 is a collider on the path. If not, the path is either $X \leftarrow \cdots \leftarrow S_1 \leftarrow \cdots Y$ or $X \leftarrow \cdots \to S_1 \to \cdots Y$. For the first case, it is evident that S_1 is an ancestor of X in $\mathcal{M}_{\underset{\sim}{X}}$; for the second case, if the path from S_1 to Y is directed, S_1 is ancestor of Y, otherwise there is a directed path from S_1 to a collider in p, according to fact 1 and fact 2 the collider is an ancestor of either X or Y in $\mathcal{M}_{\underset{\sim}{X}}$, hence S_1 is an ancestor of either X or Y in $\mathcal{M}_{\underset{\sim}{X}}$. That is, S_1 is always an ancestor of X or Y in $\mathcal{M}_{\underset{\sim}{X}}$. Note X is adjacent to S_1 in $\mathcal{M}$. According to fact 4, S_1 cannot be a descendant of X in $\mathcal{D}$. Hence there is $X \leftarrow\!\!* S_1$ in $\mathcal{M}$. According to the definition of $\mathbf{D} = \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$, we conclude $S_1 \in \mathbf{D}$. In this case if S_1 is not a collider in p, it does not belong to $\mathbf{D}$ by fact 1, contradicting with $S_1 \in \mathbf{D}$. Hence S_1 is a collider in p. Since the path p is active relative to $\mathbf{D}$, there is $S_1 \in \mathbf{D}$.

Then we suppose any vertex S_i in $S_1, S_2, \cdots, S_k, k \ge 2$ is a collider in p and $S_i \in \mathbf{D}$. We could prove S_{k+1} is a collider in p and $S_{k+1} \in \mathbf{D}$. If not, the sub-path from S_k to Y in p is either $S_k \leftarrow \cdots \leftarrow S_{k+1} \leftarrow \cdots Y$ or $S_k \leftarrow \cdots \to S_{k+1} \to \cdots Y$. For the first case, S_{k+1} is an ancestor of X or Y in $\mathcal{M}_{\underset{\sim}{X}}$ according to fact 3; for the second case, if the path from S_{k+1} to Y is directed, S_{k+1} is ancestor of Y, otherwise there is a directed path from S_{k+1} to a collider in p, according to fact 1 and fact 2 the collider is an ancestor of either X or Y in $\mathcal{M}_{\underset{\sim}{X}}$, hence S_{k+1} is an ancestor of either X or Y in $\mathcal{M}_{\underset{\sim}{X}}$. Hence in both cases, S_{k+1} is an ancestor of X or Y in $\mathcal{M}_{\underset{\sim}{X}}$. By fact 5, there is a collider path $X \leftrightarrow \cdots \leftarrow\!\!* S_{k+1}$ in $\mathcal{M}$. Hence $S_{k+1} \in \mathbf{D}$ according to the definition of $\mathbf{D} = \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$. In this case if S_{k+1} is not a collider in p, it does not belong to $\mathbf{D}$ by fact 1, contradicting with $S_{k+1} \in \mathbf{D}$. Hence S_{k+1} is a collider in p. Since the path p is active relative to $\mathbf{D}$, there is $S_{k+1} \in \mathbf{D}$. The induction step completes.

By induction, we conclude that $S_1, S_2, \cdots, S_n$ are colliders in p. With fact 4, X cannot be an ancestor of $S_{n+1} (= Y)$ in $\mathcal{D}$, thus X is not an ancestor of Y in $\mathcal{M}$, contradicting with the condition $X \in \text{Anc}(Y, \mathcal{M})$. Hence there is always a contradiction if there is an active non-causal path p relative to $\mathbf{D}$ from X to Y. Hence, we conclude that there is not an active non-causal path relative to $\mathbf{D}$ from X to Y in DAG $\mathcal{D}$.

Next we prove $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ does not contain a vertex in $\text{De}(X, \mathcal{D})$. Suppose $V \in \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) \cap \text{De}(X, \mathcal{D})$, suppose there is $X (= V_0) \leftrightarrow V_1 \leftrightarrow \cdots \leftrightarrow V_{t-1} \leftarrow\!\!* V (= V_t)$ in $\mathcal{M}_{\underset{\sim}{X}}$ where each non-endpoint is an ancestor of X or Y in $\mathcal{M}_{\underset{\sim}{X}}$. Evidently each non-endpoint is an ancestor of either X or Y in $\mathcal{D}$, and there is a path p_1 in the form of $X \leftarrow \cdots \to V_1 \leftarrow \cdots \to V_{t-1} \leftarrow \cdots V$ in $\mathcal{D}$ where each sub-path $p_1[V_i, V_{i+1}], 0 \le i \le t-1$ is an inducing path relative to $\langle \mathbf{V}(\mathcal{D}) \backslash \mathbf{V}(\mathcal{M}), \emptyset \rangle$ and p_2 is a directed path from V to Y. If there exists an adjustment set by adjustment criterion, the set $\mathbf{W}$ does not contain the vertex in p_2. To block the path $p_1 \bigoplus p_2$, there exists some vertex in p_1 that does not belong to $\mathbf{W}$. Similar to the part of fact 4, we could prove there is always a path unblocked by $\mathbf{W}$, contradicting the Blocking condition of adjustment criterion. Hence we prove $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) \cap \text{De}(X, \mathcal{D}) = \emptyset$, the Forbidden set condition in adjustment criterion is evidently fulfilled. Hence $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ fulfills adjustment criterion in $\mathcal{D}$. We get the desired result. $\square$

3.5.2. Proof of Theorem 2

We first show some facts before presenting the proof for Theorem 2.

Algorithm 3: Orient a maximal local MAG of X as a MAG.

Input: Maximal local MAG $\mathbb{M}$, potential adjustment set $\mathbf{W}$ and corresponding $\bar{\mathbf{W}}$ according to Definition 5, block set $\mathbf{S}$
1: for $\forall K \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $\forall T \in \mathbf{S}$ such that $K \circ\!\!-\!\!* T$ in $\mathbb{M}$, orient it as $K \leftarrow\!\!* T$ (the mark at T remains);
2: update the subgraph $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ as follows until no feasible updates: for any two vertices V_i and V_j such that $V_i \circ\!\!-\!\!\circ V_j$, orient it as $V_i \rightarrow V_j$ if (1) $\mathcal{F}_{V_i} \backslash \mathcal{F}_{V_j} \neq \emptyset$ or (2) $\mathcal{F}_{V_i} = \mathcal{F}_{V_j}$ as well as there is a vertex $V_k \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ not adjacent to V_j such that $V_k \rightarrow V_i \circ\!\!-\!\!\circ V_j$, where $\mathcal{F}_{V_i} = \{V \in \mathbf{S} \mid V *\!\!-\!\!\circ V_i$ in $\mathbb{M}\}$;
3: orient the circles on the remaining $\circ\!\!\rightarrow$ edges as tails;
4: in subgraph $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$, orient the circle component into a DAG without new unshielded colliders;
5: in subgraph $\mathbb{M}[-\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$, orient the circle component into a DAG without new unshielded colliders.
Output: A MAG $\mathcal{M}$

Lemma 1. *Given a maximal local MAG $\mathbb{M}$ of X, the following properties are satisfied:*

(Closed) *the circle component in $\mathbb{M}$ is chordal;*
(Invariant) *The arrowheads and tails in $\mathbb{M}$ are invariant in all the MAGs consistent with $\mathcal{P}$ with the local marks at X.*
(Chordal) *the circle component in $\mathbb{M}$ is chordal;*
(Balanced) *for any three vertices A, B, C, if $A *\!\!\rightarrow B \circ\!\!-\!\!* C$, then there is an edge between A and C with an arrowhead at C, namely, $A *\!\!\rightarrow C$. Furthermore, if the edge between A and B is $A \rightarrow B$, then the edge between A and C is either $A \rightarrow C$ or $A \circ\!\!\rightarrow C$ (i.e., it is not $A \leftrightarrow C$);*
(Complete) *For each circle at vertex A on any edge $A \circ\!\!-\!\!* B$ in $\mathbb{M}_s$, there exist MAGs $\mathcal{M}_1$ and $\mathcal{M}_2$ consistent with $\mathcal{P}$ and BK regarding $V_1, \ldots, V_s$ with $A \leftarrow\!\!* B \in \mathbf{E}(\mathcal{M}_1)$ and $A \rightarrow B \in \mathbf{E}(\mathcal{M}_2)$;*
(P6) *We can always obtain a MAG consistent with $\mathcal{P}$ by transforming the circle component into a DAG without unshielded colliders and transforming $A \circ\!\!\rightarrow B$ as $A \rightarrow B$.*

Proof. The first five properties directly follow Thm.1 of Wang et al. [19]. The property $P6$ follows Lemma 16.1 of Wang et al. [19]. □

Since the proof of existence of MAGs is involved in Theorem 2, we first present an algorithm to obtain a MAG valid to $\mathbb{M}$ in Algorithm 3, with the proof for the validity of MAG construction in Lemma 7.

At first, we present Lemma 2, 3 and 4 following Wang et al. [19].

Lemma 2. *Consider a maximal local MAG $\mathbb{M}$. If there is a possible directed path from A to B in $\mathbb{M}$, then there is a minimal possible directed path from A to B in $\mathbb{M}$.*

Proof. Suppose the possible directed path $p = \langle V_0(=A), V_1, \ldots, V_m(=B)\rangle$. If p is minimal, the result trivially holds. If not, we can always find a subpath $\langle V_i, V_{i+1}, \ldots, V_j\rangle, j-i \geq 2$ such that any non-consecutive vertices are not adjacent except for an edge between V_i and V_j. We will show the impossibility of $V_i \leftarrow\!\!* V_j$ in $\mathbb{M}$. Suppose $V_i \leftarrow\!\!* V_j$ in $\mathbb{M}$. Note there is a circle/tail at V_i on the edge between V_i and V_{i+1} due to the possible directed path p. If $j-i=2$, there is always an edge $V_{i+1} \leftarrow\!\!* V_{i+2}(=V_j)$ due to the balance/closed property of $\mathbb{M}$, contradicting the possible directed path p. If $j-i>2$, due to the non-adjacency of the V_j and V_{i+1}, there is either $V_i \rightarrow V_{i+1} \rightarrow \ldots V_j$ or $V_i \leftarrow\!\!* V_{i+1}$ identified in $\mathcal{P}$. The latter case is impossible due to the possible directed path p. For the former case, there is an almost directed or directed cycles, contradiction. Hence, the edge between V_i and V_j is either $V_i \rightarrow V_j$ or $V_i \circ\!\!-\!\!* V_j$, we thus find a shorter possible directed path $\langle V_0, V_1, \ldots, V_i, V_j, V_{j+1}, \ldots, V_m\rangle$ in $\mathbb{M}$. Repeat this process until obtaining a possible directed path such that there is not a proper sub-structure where any non-consecutive vertices are not adjacent except for an edge between endpoints. This path is a minimal possible directed path. □

Lemma 3. *Consider a maximal local MAG $\mathbb{M}$. If there is $A *\!\!\rightarrow B$ in $\mathbb{M}$, then there is an edge as $A *\!\!\rightarrow V$ for any V in a connected circle component with B in $\mathbb{M}$, and A and B are not in a connected circle component.*

Proof. It is a direct conclusion of the balanced property of $\mathbb{M}$. We first consider any vertex V_1 that has a circle edge with B, there is $A *\!\!\rightarrow B \circ\!\!-\!\!\circ V_1$ in $\mathbb{M}$. According to the balanced property, there is $A *\!\!\rightarrow V_1$. Similarly, we can conclude that the result holds for all the vertices in a circle component with B. Hence A and B cannot be in a connected circle component. □

Lemma 4. *Consider a maximal local MAG $\mathbb{M}$ of X and a block set $\mathbf{S}$. Denote $\mathcal{F}_{V_i} = \{V \in \mathbf{S} \mid V *\!\!-\!\!\circ V_i$ in $\mathbb{M}\}$ for $\forall V_i \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$. For an edge $J \circ\!\!-\!\!\circ K$ in $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$, if it is oriented as $J \rightarrow K$ in the second step of Algorithm 3, then there is a vertex $V_m \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ such that there is a minimal path $V_m \circ\!\!-\!\!\circ \ldots \circ\!\!-\!\!\circ V_1(=J) \circ\!\!-\!\!\circ V_0(=K), m \geq 1$ in $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ where $\mathcal{F}_{V_m} \supset \mathcal{F}_{V_{m-1}} = \cdots = \mathcal{F}_{V_0}$.*

Proof. A directed edge $J \rightarrow K$ is oriented in the second step only if in two situations: (1) $\mathcal{F}_K \subset \mathcal{F}_J$; (2) $\mathcal{F}_K = \mathcal{F}_J$ and there is another vertex $L \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ that is not adjacent to K and there is $L \rightarrow J$. Note $L \rightarrow J$ can only be oriented in the second step since

(1) the edges connecting $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $\mathbf{S}$ is not oriented in the first step; (2) $L \to J$ cannot appear in $\mathbb{M}$ for otherwise either $J \to K$ or $J \leftrightarrow\!* K$ is identified in $\mathbb{M}$ due to the closed property of $\mathbb{M}$.

If $\mathcal{F}_{V_0} \subset \mathcal{F}_{V_1}$, there is a desired path where $m = 1$. If $\mathcal{F}_{V_0} = \mathcal{F}_{V_1}$, we could find $V_2 \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ that is not adjacent to V_0 and there is $V_2 \to V_1$ oriented in the second step. Similarly, we conclude either $\mathcal{F}_{V_1} \subset \mathcal{F}_{V_2}$, in which case there is a desired path where $m = 2$; or $\mathcal{F}_{V_1} = \mathcal{F}_{V_2}$, in which case there is $V_3 \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ that is not adjacent to V_1 and there is $V_3 \to V_2$ oriented in the second step. Repeat the process and we can always find an uncovered path $V_m \circ\!-\!\circ \dots \circ\!-\!\circ V_1(= J) \circ\!-\!\circ V_0(= K), m \geq 1$ in $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ where $\mathcal{F}_{V_0} = \cdots = \mathcal{F}_{V_{m-1}} \subset \mathcal{F}_{V_m}$. Finally, it suffices to prove that the path is minimal. If not, there exists a sub-structure $V_i \circ\!-\!\circ V_{i+1} \circ\!-\!\circ \cdots \circ\!-\!\circ V_j, j > i + 2$ where any two non-consecutive vertices are not adjacent except for an edge between V_i and V_j. Since only the edges containing vertices in $\mathbf{S}$ are transformed in the first step, if there is a non-circle edge between V_i and V_j before the second step, the edge is non-circle in $\mathbb{M}$, in which case V_i and V_j cannot be in a circle component according to Lemma 3, contradicting with the circle path comprised of $V_i, V_{i+1}, \dots, V_j$. Hence there is $V_i \circ\!-\!\circ V_j$ in $\mathbb{M}$, in which case the chordal property of $\mathbb{M}$ is not fulfilled. Thus the path can only be minimal. □

Lemma 5. *Given the first two conditions of Theorem 2 fulfilled, there are neither edges oriented with different directions nor new unshielded colliders generated in the second step of Algorithm 3 if and only if* $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ *is bridged relative to* $\mathbf{S}$ *in* $\mathbb{M}$.

Proof. We first prove the "if" statement. We prove that there are neither edges oriented with different directions nor new unshielded colliders generated in the second step of Algorithm 3, respectively. For simplicity, denote $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ by $\mathbb{M}_1$.

Suppose we orient both $J \to K$ and $J \leftarrow K$ in the second step. According to Lemma 4, if we orient $J \to K$ in the second step, there is a minimal circle path $V_0 \circ\!-\!\circ V_1 \circ\!-\!\circ \cdots V_{m-1}(= J) \circ\!-\!\circ V_m(= K)$ where $\mathcal{F}_{V_0} \supset \mathcal{F}_{V_1} = \cdots = \mathcal{F}_{V_m}$. If we also orient $J \leftarrow K$ in the second step, there is another minimal circle path $V_{m-1}(= J) \circ\!-\!\circ V_m(= K) \circ\!-\!\circ \cdots \circ\!-\!\circ V_n, n > m$ in $\mathbb{M}_1$ where $\mathcal{F}_{V_{m-1}} = \mathcal{F}_{V_m} = \cdots = \mathcal{F}_{V_{n-1}} \subset \mathcal{F}_n$. Note V_{m+1} is adjacent to V_m but not adjacent to V_{m-1}, while V_{m-2} is adjacent to V_{m-1} but not adjacent to V_m, hence V_{m-2}, V_{m-1}, V_m, V_{m+1} are distinct vertices. According to Lemma 3, there cannot be non-circle edge between the variables in the circle path. Also note no circle edges in $\mathbb{M}_1$ are oriented in the first step. Hence the circle component in $\mathbb{M}_1$ after the first step is still chordal. Hence $V_0 \circ\!-\!\circ V_1 \circ\!-\!\circ \cdots \circ\!-\!\circ V_n$ is also a minimal circle path, otherwise there is a circle cycle whose length is larger than 3 without a chord because this cycle must contain $V_{m-2}, V_{m-1}, V_m, V_{m+1}$ where V_{m-2} is not adjacent to V_m and V_{m-1} is not adjacent to V_{m+1}. Since $V_0, \cdots, V_n \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$, V_0 and V_n are not bridged relative to $\mathbf{S}$, contradicting with that $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ is bridged relative to $\mathbf{S}$ in $\mathbb{M}$.

Suppose there is a new unshielded collider $A \to B \leftarrow C$ generated in the second step. According to Lemma 4 there is a minimal path $F_1 \to \cdots \to F_m(= A) \to B, m \geq 2$ and a minimal path $V_1 \to \cdots V_n(= C) \to B, n \geq 2$ such that $\mathcal{F}_{F_1} \supset \mathcal{F}_{F_2} = \cdots = \mathcal{F}_B$ and $\mathcal{F}_{V_1} \supseteq \mathcal{F}_{V_2} = \cdots = \mathcal{F}_B$. A and C are evidently different vertices that are not adjacent. In this case there is a circle path $p : F_1 \circ\!-\!\circ \cdots \circ\!-\!\circ F_m(= A) \circ\!-\!\circ B \circ\!-\!\circ V_n(= C) \circ\!-\!\circ \cdots \circ\!-\!\circ V_1$ in $\mathbb{M}$ such that $\mathcal{F}_{F_1} \supset \mathcal{F}_{F_2} = \cdots = \mathcal{F}_B = \cdots = \mathcal{F}_{V_2} \subset \mathcal{F}_{V_1}$. According to Lemma 3, there are no non-circle edges between the variables in p. In this case, there is always a minimal circle path from F_1 to V_1 such that F_1 and V_1 are not bridged relative to $\mathbf{S}$ in $\mathbb{M}$, contradiction.

We then prove the "only if" statement. Suppose $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ is not bridged relative to $\mathbf{S}$ in $\mathbb{M}$.

If $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ is not bridged relative to $\mathbf{S}$ in $\mathbb{M}$, we will prove the result by showing that there are either edges oriented with difference direction or new unshielded colliders generated in the second step of Algorithm 3.

Suppose two vertices J, K in $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ are not bridged relative to $\mathbf{S}$ due to the minimal circle path $J(= V_0) \circ\!-\!\circ V_1 \cdots V_n \circ\!-\!\circ K(= V_{n+1})$ in $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$. There are two possible cases (they possibly happen simultaneously). One is that there exists $0 \leq s \leq n$ such that $\mathcal{F}_{V_s} \not\subseteq \mathcal{F}_{V_{s+1}}$ and $\mathcal{F}_{V_{s+1}} \not\subseteq \mathcal{F}_{V_s}$. The other is that there exists $1 \leq s \leq n$ such that $\mathcal{F}_{V_s} \subset \mathcal{F}_{V_{s-1}}$ and $\mathcal{F}_{V_s} \subset \mathcal{F}_{V_{s+1}}$.

For the first case, suppose there are two vertices $T_1, T_2 \in \mathbf{S}$ such that $T_1 \in \mathcal{F}_{V_s} \backslash \mathcal{F}_{V_{s+1}}$ and $T_2 \in \mathcal{F}_{V_{s+1}} \backslash \mathcal{F}_{V_s}$. In the second step of Algorithm 3, we will orient both $V_i \to V_j$ and $V_i \leftarrow V_j$. For the second case, suppose a vertex $T_1 \in \mathcal{F}_{V_{s-1}} \backslash \mathcal{F}_{V_s}$ and a vertex $T_2 \in \mathcal{F}_{V_{s+1}} \backslash \mathcal{F}_{V_s}$. In the second step of Algorithm 3, there is $V_{s-1} \to V_s \leftarrow V_{s+1}$ oriented. As the path is minimal, a new unshielded collider is generated. □

Lemma 6. *Given the three conditions in Theorem 2 fulfilled with* $\mathbf{S}$*, for any* $K \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ *and any* $T_1, T_2 \in \mathbf{S}$ *such that there is* $T_1 *\!\!-\!\circ K \circ\!\!-\!* T_2$ *in* $\mathbb{M}$*,* T_1 *is adjacent to* T_2.

Proof. If $K \in \bar{\mathbf{W}}$, according to Definition 5, $T_1, T_2 \in \mathbf{S}_K$. According to the second condition of Theorem 2, T_1 is adjacent to T_2. It suffices to prove the result for $K \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}]) \backslash \bar{\mathbf{W}}$.

According to the definition of $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}]) \backslash \bar{\mathbf{W}}$, we can always find a minimal possible directed path p_1 from a vertex in $\bar{\mathbf{W}}$ to K in $\mathbb{M}[-\mathbf{S}]$ where each non-endpoint vertex does not belong to $\bar{\mathbf{W}}$. We suppose p_1 is from $F_t \in \bar{\mathbf{W}}$ to K comprised of $V_0(= F_t), V_1, \cdots, V_s(= K)$. And according to the definition of $\bar{\mathbf{W}}$, there is a collider path as $X(= F_0) \leftrightarrow \cdots F_{t-1} \leftarrow\!\circ($ or $\leftrightarrow) F_t$, where each non-endpoint belongs to $\mathbf{W}$. Next we will prove that there is always an edge $V_0(= F_t) \circ\!\!-\!* T_1$. At first, we name a fact. For any vertex $V_i, 0 \leq i \leq s$, there is not an edge as $V_i \to T_1$ in $\mathbb{M}$ due to the first condition of Theorem 2. We discuss the possible edge between F_t and V_1: **(1).** $F_t \to V_1$; **(2).** $F_t \circ\!-\!\circ V_1$; **(3).** $F_t \circ\!\!\to V_1$. We will prove that for any case, there is $T_1 *\!\!-\!\circ V_0(= F_t)$.

(1). If there is $F_t \to V_1$ in $\mathbb{M}$, there is $F_t \to V_1 \to \cdots \to V_s(= K)$ in $\mathbb{M}$ since p_1 is a minimal possible directed path in $\mathbb{M}$ and the closed property of $\mathbb{M}$. Note there is $V_s(= K) \circ\!\!-\!* T_1$ in $\mathbb{M}$, T_1 is adjacent to V_{s-1} because there is not a structure as $A \to B \circ\!\!-\!* C$ in $\mathbb{M}$ where A and C are not adjacent as a result of closed property of $\mathbb{M}$. Since there is $T_1 *\!\!-\!\circ V_s(= K)$ in $\mathbb{M}$ and due to the complete property of $\mathbb{M}$, there must be $T_1 \leftarrow V_{s-1}$ or $T_1 \ast\!\!-\!\!\circ V_{s-1}$. Since for any vertex $V_i, 0 \leq i \leq s$ there is not an edge as $T_1 \leftarrow V_i$ in $\mathbb{M}$, the edge can only be $T_1 \ast\!\!-\!\!\circ V_{s-1}$. Repeat this process for $V_{s-1}, V_{s-2}, \cdots, F_t(=V_0)$, we can prove that there is $T_1 \ast\!\!-\!\!\circ V_0(=F_t)$.

(2). If there is $F_t \circ\!\!-\!\!\circ V_1$ in $\mathbb{M}$, p_1 must be in the form of $V_0(F_t) \circ\!\!-\!\!\circ V_1 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ V_j \circ\!\!\rightarrow V_{j+1} \rightarrow \cdots \rightarrow V_s, 1 \leq j \leq s$. Note there is a possible directed path $p_1[V_1, V_s] \bigoplus V_s \circ\!\!-\!\!\ast T_1$ from V_1 to T_1, V_1 is a possible ancestor of T_1 in $\mathbb{M}$. And according to the definition of $\mathbf{S}$, V_1 is a possible ancestor of Y in $\mathbb{M}$. At first, V_1 cannot be adjacent to F_{t-1}. Otherwise, there is $F_{t-1} \leftrightarrow\!\!\ast V_1$ due to the balanced property of $\mathbb{M}$ and $F_{t-1} \leftrightarrow\!\!\ast F_t \circ\!\!-\!\!\circ V_1$. And since V_1 is a possible ancestor of Y in $\mathbb{M}$, $V_1 \in \mathbf{W} \cup \bar{\mathbf{W}}$. However, V_1 cannot belong to $\mathbf{W}$ for otherwise $V_1 \in \mathbf{S}$ according to Definition 5 and thus V_1 cannot be in $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}]) \backslash \bar{\mathbf{W}}$. V_1 cannot belong to $\bar{\mathbf{W}}$ since p_1 does not go through any vertex in $\bar{\mathbf{W}}$ except for the endpoints. Hence V_1 cannot belong to $\mathbf{W} \cup \bar{\mathbf{W}}$, contradiction. Hence V_1 is not adjacent to F_{t-1}. In this case, according to the third condition of Theorem 2, there is $F_t \circ\!\!-\!\!\ast T_1$ in $\mathbb{M}$, otherwise F_t and K are not bridged relative to $\mathbf{S}$ in light of the facts that (1) $p_1[F_t, V_s]$ is a minimal circle path; (2) V_1 is not adjacent to $F_{t-1} \in \mathbf{S}$; (3) there is $V_s \circ\!\!-\!\!\ast T_1 \in \mathbf{S}$ in $\mathbb{M}$.

(3). If there is $F_t \circ\!\!\rightarrow V_1$, p_1 is as $F_t \circ\!\!\rightarrow V_1 \rightarrow \cdots V_s$ due to the closed property of $\mathbb{M}$. We can conclude that there is $V_0(=F_t) \circ\!\!-\!\!\ast T_1$ with the same proof for the case **(1)**.

Hence, we conclude that there is always $V_0(=F_t) \circ\!\!-\!\!\ast T_1$. Similarly, there is always $V_0(=F_t) \ast\!\!-\!\!\circ T_2$. In this case, according to Definition 5, there is $T_1, T_2 \in \mathbf{S}_{F_t}$. Hence that T_1 is adjacent to T_2 directly follows the second condition of Theorem 2. □

Lemma 7. *Given the three conditions in Theorem 2 fulfilled with* $\mathbf{S}$*, we can obtain a MAG valid to* $\mathbb{M}$ *by Algorithm 3.*

Proof. The whole proof in this part is comprised of two parts. **A.** we can obtain a unique graph $\mathcal{H}$ without circles by Algorithm 3. **B.** $\mathcal{H}$ is a MAG valid to $\mathbb{M}$. The unique graph means that we will not orient an edge with two directions.

A. We can obtain a unique graph $\mathcal{H}$ without circles by Algorithm 3.

We have shown in Lemma 5 that there are neither edges oriented with different directions nor new unshielded colliders generated in the second step of Algorithm 3. In the following, we will first prove that **A.1.** there are not circle edges connecting $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ and $\mathbb{M}[-\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$. Then, we prove that **A.2.** the circle component in $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ and **A.3.** the circle component in $\mathbb{M}[-\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ are chordal, respectively. Given these results, the fourth and fifth steps can be executed since every chordal graph has a perfect elimination order, through which we can orient the chordal graph as a DAG without unshielded colliders. We conclude that we can obtain a unique graph $\mathcal{H}$ without circles by Algorithm 3.

A.1. Suppose there is a circle edge $A \circ\!\!-\!\!\circ B$, where $A \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $B \notin \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$, after the first third steps of Algorithm 3. There must be $B \in \mathbf{S}$ for otherwise there is $B \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ due to $A \circ\!\!-\!\!\circ B$. However, for such a case, the circle edge should have been transformed to $A \leftarrow\!\!\circ B$ in the first step of Algorithm 3, contradiction.

A.2. Denote $\bar{\mathbb{M}}$ the obtained graph after the first three steps of Algorithm 3. We will prove that the circle component in $\bar{\mathbb{M}}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ is chordal.

Suppose the circle component in $\bar{\mathbb{M}}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ is not chordal, there is $V_0 \circ\!\!-\!\!\circ V_1 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ V_n \circ\!\!-\!\!\circ V_0, n \geq 3$, where there is not a circle edge between every two unconsecutive vertices. There must exist non-circle edges between the unconsecutive vertices in this cycle, otherwise it is a cycle of length four or more without a chord in $\mathbb{M}$, contradicting with the chordal property of $\mathbb{M}$. Hence, we can always find a minimal sub-structure $V_k \circ\!\!-\!\!\circ V_{k+1} \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ V_m \leftarrow V_k, 0 \leq k < m \leq n$ without other directed edges between any two vertices among $V_k, \cdots, V_m$ except for a directed edge between V_m and V_k (we suppose it $V_m \leftarrow V_k$ without loss of generality), and there is not a proper sub-structure satisfying the conditions above. According to Lemma 3, $V_k \rightarrow V_m$ can only be a circle edge in $\mathbb{M}$. Hence in $\mathbb{M}$ there is $V_k \circ\!\!-\!\!\circ V_{k+1} \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ V_m \circ\!\!-\!\!\circ V_k$. Since the circle component in $\mathbb{M}$ in chordal, the length of the sub-structure can only be three. Hence it holds $m = k+2$ and there is $V_k \circ\!\!-\!\!\circ V_{k+1} \circ\!\!-\!\!\circ V_{k+2} \leftarrow V_k$ in $\bar{\mathbb{M}}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$. Next, we will prove its impossibility.

We first prove that the edge $V_{k+2} \leftarrow V_k$ cannot be oriented by the first step of Algorithm 3. If it is, then there is $V_{k+2} \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $V_k \in \mathbf{S}$. Note there is $V_{k+1} \circ\!\!-\!\!\circ V_{k+2}$ in $\mathbb{M}$, there must be $V_{k+1} \in \mathbf{S}$ for otherwise there is $V_{k+1} \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and thus $V_{k+1} \leftarrow V_k$ is oriented in the first step of Algorithm 3. However, when $V_{k+1} \in \mathbf{S}$, there is $V_{k+2} \leftarrow V_{k+1}$ oriented in the first step of Algorithm 3. Hence that $V_{k+2} \leftarrow V_k$ is oriented in the first step of Algorithm 3 is impossible to obtain a sub-structure $V_k \circ\!\!-\!\!\circ V_{k+1} \circ\!\!-\!\!\circ V_{k+2} \leftarrow V_k$ in $\bar{\mathbb{M}}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$.

Thus $V_{k+2} \leftarrow V_k$ is only possible oriented in the second step. Due to $V_k \circ\!\!-\!\!\circ V_{k+1} \circ\!\!-\!\!\circ V_{k+2}$ in $\bar{\mathbb{M}}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$, there is $\mathcal{F}_{V_k} = \mathcal{F}_{V_{k+1}} = \mathcal{F}_{V_{k+2}}$. As $V_k \rightarrow V_{k+2}$ is oriented in the second step, according to Lemma 4, there exists a minimal path $F_t \circ\!\!-\!\!\circ \cdots F_1(=V_k) \circ\!\!-\!\!\circ F_0(=V_{k+2})$ in $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ such that $\mathcal{F}_{F_t} \supset \mathcal{F}_{F_{t-1}} = \cdots = \mathcal{F}_{F_1} = \mathcal{F}_{F_0}$. Evidently F_2 is adjacent to V_{k+1}, otherwise $V_k \rightarrow V_{k+1}$ is also oriented. Since (1) $\mathcal{F}_{F_1}(\mathcal{F}_{V_k}) = \mathcal{F}_{F_2} = \mathcal{F}_{V_{k+1}}$, (2) there is not an edge oriented as different directions in the second step of Algorithm 3 according to Lemma 5, and (3) F_2 is not adjacent to V_{k+2}, there can only be $F_2 \circ\!\!-\!\!\circ V_{k+1}$ in $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$. We find a sub-structure $F_2 \circ\!\!-\!\!\circ V_{k+1} \circ\!\!-\!\!\circ V_k \leftarrow F_2$ such that $\mathcal{F}_{F_2} = \mathcal{F}_{V_{k+1}} = \mathcal{F}_{V_k}$. Similar to the previous proof, there is $F_3 \circ\!\!-\!\!\circ V_{k+1}$ in $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ where $\mathcal{F}_{F_3} = \mathcal{F}_{V_{k+1}}$. Repeat this process until F_t and we conclude that there must be $F_t \circ\!\!-\!\!\circ V_{k+1}$. Since $\mathcal{F}_{F_t} \supset \mathcal{F}_{F_{t-1}} = \cdots = \mathcal{F}_{F_1} = \mathcal{F}_{F_0}$, it is oriented as $F_t \rightarrow V_{k+1}$ in the second step of Algorithm 3, in which case $V_{k+1} \rightarrow V_{k+2}$ is also oriented in the second step of Algorithm 3 due to the non-adjacency of F_t and V_{k+2}, contradiction.

A.3. In the first four steps of Algorithm 3, we never transform the circle edges in $\mathbb{M}[-\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$. Due to the chordal property of $\mathbb{M}$ and the fact that the subgraph of a chordal graph is also chordal, the circle component in $\mathbb{M}[-\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ is chordal.

According the three parts above, it is guaranteed that a graph could be output without different orientations on one edge by Algorithm 3.

B. $\mathcal{H}$ is a MAG valid to $\mathbb{M}$.

It evidently follows that $\mathcal{H}$ has the non-circle marks in $\mathbb{M}$. Hence, it suffices to prove that $\mathcal{H}$ is a MAG consistent with $\mathcal{P}$. To prove it, we construct an auxiliary graph $\mathcal{H}_0$ by transforming all the bi-directed edges $K \leftrightarrow T$ in $\mathcal{H}$ which are $K \circ\!\!\rightarrow T$ in $\mathbb{M}$ to $K \rightarrow T$. According to Algorithm 3 to obtain $\mathcal{H}$ from $\mathbb{M}$, there is $K \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $T \in \mathbf{S}$. We first show that **(B.1.)** $\mathcal{H}_0$ is a MAG consistent with $\mathcal{P}$. Then we show that **(B.2.)** $\mathcal{H}$ can be obtained from $\mathcal{H}_0$ by transformation that preserves the property being a MAG consistent with $\mathcal{P}$.

B.1. We will prove $\mathcal{H}_0$ is a MAG consistent with $\mathcal{P}$ by showing that $\mathcal{H}_0$ can be seen as a graph obtained from $\mathbb{M}$ by transforming $\circ\!\!\rightarrow$ to $\rightarrow$ and transforming the circle component into a DAG without new unshielded colliders, through which we can get the desired result by Property $(P6)$ of $\mathbb{M}$ according to Lemma 1.

It is direct that $\mathcal{H}_0$ has the non-circle marks in $\mathbb{M}$ and there are no new bi-directed edges in $\mathcal{H}_0$ relative to $\mathbb{M}$ since all additional bi-directed edges in $\mathcal{H}$ relative to $\mathbb{M}$ are possibly introduced in only the first step of orientation process of $\mathcal{H}$, which have been transformed to directed edges in $\mathcal{H}_0$. Besides, all the circles on $\circ\!\!\rightarrow$ edges in $\mathbb{M}$ are oriented as tails in $\mathcal{H}_0$. In the following it suffices to show that $\mathcal{H}_0$ is also a graph oriented from $\mathbb{M}$ by orienting the circle component in $\mathbb{M}$ into a DAG without unshielded colliders.

Hence, we only consider the circle component in $\mathbb{M}$. We divide it into two parts, one is the circle component in $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$, denoted by CC_1; and the other is the circle component in $\mathbb{M}[-\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$, denoted by CC_2. Evidently, the set of vertices in CC_2 is $\mathbf{V}(\mathbb{M})\backslash\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$. In the following we will prove that in $\mathcal{H}_0$ both CC_1 and CC_2 are oriented to DAGs without new unshielded colliders and there are no new unshielded colliders or directed or almost directed cycles comprised of the vertices in both CC_1 and CC_2.

We first consider the orientation in $\mathcal{H}_0$ of the edges in CC_1 in the process of obtaining $\mathcal{H}$ from $\mathbb{M}$ and transforming $\mathcal{H}$ to $\mathcal{H}_0$. In Step 3, if there is an edge $\circ\!\!\rightarrow$, then the two vertices cannot be connected in the circle component, hence the edges in CC_1 cannot be oriented in Step 3. Hence, according to Algorithm 3, the edges in CC_1 can only be oriented by either Step 2 or Step 4. There are no new unshielded colliders or directed or almost directed cycles oriented in the edges of CC_1 by the three following facts. **(1).** There are no new unshielded colliders or directed or almost directed cycles in the edges of CC_1 oriented by Step 2 according to Lemma 5. **(2).** There are no unshielded colliders or directed or almost directed cycles in the edges of CC_1 oriented in Step 4. **(3).** There are no new unshielded colliders or directed or almost directed cycles in edges of CC_1 oriented by both Step 2 and Step 4 due to the balanced property of $\mathbb{M}$ and the impossibility of the transformation of circle edges to bi-directed edges.

Then we consider the orientation in $\mathcal{H}_0$ of the edges in CC_2. The edges in CC_2 totally follows Step 5 of Algorithm 3, which evidently does not introduce new unshielded colliders or directed or almost directed cycles.

Finally, we consider circle edges in $\mathbb{M}$ connecting $K \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $T \in \mathbf{V}(\mathbb{M})\backslash\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$. Suppose there is a circle edge $K \circ\!\!-\!\!\circ T$, where $K \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $T \in \mathbf{V}(\mathbb{M})\backslash\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$. In this case there can only be $T \in \mathbf{S}$, otherwise there is also $T \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ due to $K \circ\!\!-\!\!\circ T$ where $K \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$, contradicting with $T \in \mathbf{V}(\mathbb{M})\backslash\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$. Thus for all the circle edges in $\mathbb{M}$ connecting $K \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $T \in \mathbf{V}(\mathbb{M})\backslash\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$, in Step 1 of Algorithm 3 we will orient it as $K \leftarrow T$ due to $T \in \mathbf{S}$, and we remain this directed edge when obtaining $\mathcal{H}_0$ from $\mathcal{H}$. Hence there cannot be a directed or almost directed cycle containing the vertices in both $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $\mathbf{V}(\mathbb{M})\backslash\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ in $\mathcal{H}_0$. Next we prove that there is not a new unshielded collider containing the vertices in both $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $\mathbf{V}(\mathbb{M})\backslash\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$. According to the construction process, for any circle edge $K \circ\!\!-\!\!\circ T$ where $K \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $T \in \mathbf{S}$, there is $K \leftarrow T$ oriented. Hence, it there is a new unshielded collider, it is as $T \rightarrow K_1 \leftarrow K_2$ or $T_1 \rightarrow K \leftarrow T_2$. We will prove the impossibility of both the cases. $T \rightarrow K_1 \leftarrow K_2$ is evidently impossible for otherwise there is $K_1 \rightarrow K_2$ oriented in Step 2, which contradicts with Lemma 5 that one edge cannot be oriented as different direction. The impossibility of $T_1 \rightarrow K \leftarrow T_2$ is due to Lemma 6. Hence there are not new unshielded colliders or directed or almost directed cycles comprised of the verttices in both CC_1 and CC_2.

Hence, we prove that the graph $\mathcal{H}_0$ constructed based on $\mathcal{H}$ can also be seen as a graph obtained from $\mathbb{M}$ by transforming all edges $\circ\!\!\rightarrow$ to $\rightarrow$ and orienting the circle component into a DAG without unshielded colliders. By Property $P6$ of $\mathbb{M}$ according to Lemma 1, $\mathcal{H}_0$ is a MAG consistent with $\mathcal{P}$.

B.2. We will prove that $\mathcal{H}$ can be obtained from $\mathcal{H}_0$ by transformation that preserves the property being a MAG consistent with $\mathcal{P}$.

Note that the only difference between $\mathcal{H}$ and $\mathcal{H}_0$ is that, for $\forall K \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $\forall T \in \mathbf{S}$ such that $K \circ\!\!\rightarrow T$ in $\mathbb{M}$, there is $K \rightarrow T$ in $\mathcal{H}_0$ but $K \leftrightarrow T$ in $\mathcal{H}$. Denote the set of different edges in $\mathcal{H}_0$ by $Edge(\mathcal{H}_0) = \{K \rightarrow T \text{ in } \mathcal{H} \mid K \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}]), T \in \mathbf{S}, K \circ\!\!\rightarrow T \text{ in } \mathbb{M}\}$. We could obtain $\mathcal{H}$ from $\mathcal{H}_0$ by transforming these edges to bi-directed edges. We transform one edge one time. At first, we select the edge $K \rightarrow T$ in $Edge(\mathcal{H}_0)$ according to the selection criterion that (1) we select K that is not an ancestor of any other V_1 such that there is an edge $V_1 \rightarrow V_2$ in $Edge(\mathcal{H}_0)$; and (2) given K selected in the first step, we select T that is not a descendant of any other V_2 such that there is an edge $K \rightarrow V_2$ in $Edge(\mathcal{H}_0)$. Then we obtain $Edge(\mathcal{H}_1)$ by deleting $K \rightarrow T$ from $Edge(\mathcal{H}_0)$. By such operation, we obtain a new graph $\mathcal{H}_1$ and $Edge(\mathcal{H}_1)$. Repeat the process above and we could obtain a series of graphs $\mathcal{H}_0, \mathcal{H}_1, \cdots, \mathcal{H}_m, \mathcal{H}_{m+1}(= \mathcal{H})$. We prove the desired result by induction. Given $\mathcal{H}_0$ is a MAG consistent with $\mathcal{P}$, we will show that for any $\mathcal{H}_i$ and $\mathcal{H}_{i+1}$, where $0 \leq i \leq m$, if $\mathcal{H}_i$ is a MAG, then $\mathcal{H}_{i+1}$ is a MAG Markov equivalent to $\mathcal{H}_i$. Suppose the edge that will be transformed in $\mathcal{H}_i$ is $K \rightarrow T$. According to Lemma 1 of Zhang and Spirtes [34], given $\mathcal{H}_i$ is a MAG, it suffices to show that (1) there is no directed path from K to T in $\mathcal{H}_i$ other than $K \rightarrow T$; (2) for any $A \rightarrow K$ in $\mathcal{H}_i$, $A \rightarrow T$ is also in $\mathcal{H}_i$; and for any $B \leftrightarrow K$ in $\mathcal{H}_i$, either $B \rightarrow T$ or $B \leftrightarrow T$ is in $\mathcal{H}_i$; (3) there is no discriminating path for K on which T is the endpoint adjacent to K in $\mathcal{H}_i$.

(1) For the sake of contradiction, suppose there is a directed path from K to T in $\mathcal{H}_i$ other that $K \rightarrow T$, we suppose the minimal directed path of this path is $K(= F_0) \rightarrow F_1 \rightarrow \cdots \rightarrow F_m \rightarrow T(= F_{m+1})$. Since we only transform directed edges to bi-directed edges in the whole process, the directed path is also in $\mathcal{H}_0$. We first prove that there must be a vertex $F_n, 1 \leq n \leq m$ such that $F_n \in \mathbf{S}$. Otherwise, all of $F_1, \cdots, F_m$ belong to $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ since $F_0 \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and there is a possible directed path comprised of $F_0, F_1, \cdots, F_m$ in $\mathbb{M}$. **(i.)** If there is $F_m \rightarrow T$ in $\mathbb{M}$, it contradicts with the first condition of Theorem 2. **(ii.)** If there is $F_m \circ\!\!-\!\!\circ T$ in $\mathbb{M}$, there is $F_m \leftarrow T$ oriented in the first step of Algorithm 3. Since we never reverse an edge in the process from $\mathcal{H}_0$ to $\mathcal{H}$, there cannot

be an edge $F_m \to T$ in $\mathcal{H}_i$. **(iii.)** If there is $F_m \circ\!\!\rightarrow T$ in $\mathbb{M}$, there is $F_m \to T$ in $\mathcal{H}_0$ and $\mathcal{H}_i$. According to the edge selection criterion, when there is both $F_m \to T$ and $K \to T$ in $\mathcal{H}_i$, we transform $F_m \to T$ ahead of $K \to T$ due to $K \to F_1 \to \cdots \to F_m$, contradiction. For the other situations for the edge between F_m and T in $\mathbb{M}$, there cannot form an edge $F_m \to T$ in $\mathcal{H}_i$. Hence we conclude there is a vertex $F_n, 1 \le n \le m$ such that $F_n \in \mathbf{S}$.

Without loss of generality, we suppose $F_n \in \mathbf{S}$ and $F_l \notin \mathbf{S}, \forall 1 \le l \le n-1$. We first prove there is not a vertex $F_l, 1 \le l \le n-1$ adjacent to T. If there is, since $F_l \to \cdots \to F_m \to T$ in $\mathcal{H}_0$, there is $F_l \to T$ in $\mathcal{H}_0$ due to the ancestral property. In this case there is a directed path $F_1 \to \cdots F_l \to T$ without vertices in $\mathbf{S}$ in $\mathcal{H}_0$, which implies that there is a possible directed path where the sub-path from F_1 to F_l is minimal and any variables on the path do not belong to $\mathbf{S}$, contradicting the result we prove above. Hence F_l cannot be adjacent to T for $\forall 1 \le l \le n-1$. **(i.)** If $n \ge 2$, **(i.a.)** if there $F_n \circ\!\!-\!\!* T$ or $F_n \to T$ in $\mathbb{M}$, there is an uncovered possible directed path comprised of $K, F_1, \cdots, F_n, T$ in $\mathbb{M}$ where F_1 is not adjacent to T. In this case $K \circ\!\!\rightarrow T$ has been oriented as $K \to T$ in $\mathbb{M}$ by $\mathcal{R}_9$ since $\mathbb{M}$ is closed under the orientation rules, contradiction. **(i.b.)** If there is $F_n \leftarrow\!\!* T$ in $\mathbb{M}$, note the non-adjacency of T and F_{n-1}. Due to the edge $T *\!\!\rightarrow F_n$ and the complete property of $\mathbb{M}$, the mark at F_n on the edge between F_{n-1} and F_n is identifiable in $\mathbb{M}$. And due to the possible directed path, there is $F_{n-1} \to F_n$ in $\mathcal{H}_0$, there can only be $F_{n-1} \to F_n$ or $F_{n-1} \circ\!\!\rightarrow F_n$ in $\mathbb{M}$. The former case contradicts with the first condition of Theorem 2 due to $F_{n-1} \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $F_n \in \mathbf{S}$. For the latter case, the edge $F_{n-1} \to F_n$ should be transformed to bi-directed edge ahead of $K \to T$, hence there cannot be an edge $F_{n-1} \to F_n$ in $\mathcal{H}_i$, contradiction. **(ii.)** If $n = 1$, there is $K \to T' \to T$ in $\mathcal{H}$, where $T' \in \mathbf{S}$. In this case if there is not $K \circ\!\!\rightarrow T'$ in $\mathbb{M}$, there cannot be an edge $K \to T'$ in $\mathcal{H}_i$; if there is $K \circ\!\!\rightarrow T'$ in $\mathbb{M}$, there is thus both $K \to T'$ and $K \to T$ in $\mathcal{H}_0$, $K \circ\!\!\rightarrow T'$ is transformed to bi-directed ahead of $K \to T$ due to $T' \to T$, thereby there is not an edge $K \to T'$ in $\mathcal{H}_i$. Hence there cannot be a sub-structure $K \to T' \to T$ in $\mathcal{H}_i$, contradiction. Hence, there is always a contradiction if there is a directed path from K to T in $\mathcal{H}_i$.

(2) In this part, we prove that if there is an edge $A \to K$ in $\mathcal{H}_i$, there is $A \to T$ in $\mathcal{H}_i$; if there is $B \leftrightarrow K$ in $\mathcal{H}_i$, either $B \to T$ or $B \leftrightarrow T$ is in $\mathcal{H}_i$. Note there is $K \circ\!\!\rightarrow T$ in $\mathbb{M}$, where $K \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $T \in \mathbf{S}$.

It suffices to show that for A such that $A \to K$ or $A \leftrightarrow K$ in $\mathcal{H}_i$, A is adjacent to T. According to the ancestral property of $\mathcal{H}_i$, we get the desired result due to $K \to T$ in $\mathcal{H}_i$.

We discuss the possible cases of the edge between A and K in $\mathbb{M}$. If there is $A *\!\!\rightarrow K \circ\!\!\rightarrow T$ in $\mathbb{M}$, due to the closed property of $\mathbb{M}$, A is adjacent to T. Hence the result evidently holds.

If there is $A \circ\!\!-\!\!\circ K$ in $\mathbb{M}$, we discuss whether $A \in \mathbf{S}$. If not, then $A \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ due to $K \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$. Suppose T is not adjacent to A for contradiction. In this case, we orient $K \to A$ in the second step due to $T \in \mathcal{F}_K \backslash \mathcal{F}_A$, there is thus $K \to A$ in $\mathcal{H}_0$. Considering we never reverse a directed edge in the whole procedure, there is not $A \to K$ in $\mathcal{H}_i$. And since only the directed edge connecting a vertex in $\mathbf{S}$ and a vertex in $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ is possibly converted to a bi-directed edge in the process from $\mathcal{H}_0$ to $\mathcal{H}_i$, $A \leftarrow K$ cannot be transformed to $A \leftrightarrow K$ due to $A, K \in \mathrm{PossDe}(X, \mathbb{M}_i[-\mathbf{S}])$, so that $A \leftrightarrow K$ is not in $\mathcal{H}_i$. Hence when $A \circ\!\!-\!\!\circ K$ in $\mathbb{M}_i$ and $A \notin \mathbf{S}$, there is not an edge $A \to K$ or $A \leftrightarrow K$ in $\mathcal{H}_i$. If $A \in \mathbf{S}$, according to Algorithm 3 to obtain $\mathcal{H}$ from $\mathbb{M}$, there is $A \leftarrow K$ in $\mathcal{H}$. And since in the process of transforming $\mathcal{H}_0$ to $\mathcal{H}$ we only transform directed edges to bi-directed edges, there is $A \leftarrow K$ in $\mathcal{H}_i$, in which case there is not $A \to K$ or $A \leftrightarrow K$ in $\mathcal{H}_i$.

If there is $A \leftarrow\!\!\circ K$ in $\mathbb{M}$, there is $A \leftarrow K$ in $\mathcal{H}_0$. Since we never reverse a directed edge in the whole process, and only the directed edge connecting a vertex in $\mathbf{S}$ and a vertex in $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ is possibly converted to a bi-directed edge in the process from $\mathcal{H}_0$ to $\mathcal{H}$, we only need to consider there is $A \leftrightarrow K$ in $\mathcal{H}_i$, where $A \in \mathbf{S}$. In this case, there is $A, T \in \mathbf{S}_K$. According to Lemma 6, A is adjacent to T. The result holds.

For the other cases for the edge between A and K in $\mathbb{M}$ except for $A *\!\!\rightarrow K$, $A \circ\!\!-\!\!\circ K$, and $A \leftarrow\!\!\circ K$, there cannot be an edge $A \to K$ or $A \leftrightarrow K$ in $\mathcal{H}_i$. We thus have considered all the possible cases and conclude that if there is $A \to K$ in $\mathcal{H}_i$, there is $A \to T$ in $\mathcal{H}_i$; if there is $A \leftrightarrow K$ in $\mathcal{H}_i$, either $A \to T$ or $A \leftrightarrow T$ is in $\mathcal{H}_i$ according to the balanced property.

(3) In this part, we prove that there is no discriminating path for K on which T is the endpoint adjacent to K in $\mathcal{H}_i$. The proof of this part refers to the proof of (T3) of Theorem 3 by Zhang [6] with some modifications.

Suppose a path $p = (V_0, V_1, \cdots, V_n = K, T)$ in $\mathcal{H}_i$ which is a discriminating path for K. Without loss of generality, suppose p is the shortest path. According to the construction of $\mathrm{Edge}(\mathcal{H}_i)$, there is $K \circ\!\!\rightarrow T$ in $\mathbb{M}$. We derive a contradiction by showing that p is already a discriminating path in $\mathbb{M}$. Hence there cannot be an edge $K \circ\!\!\rightarrow T$ in $\mathbb{M}$, otherwise if $i \ge 1$ (there is local BK) it will be oriented as $K \to T$ by $\mathcal{R}_4'$ due to the closed property of $\mathbb{M}$. There is $V_{n-1} \leftrightarrow K$ in $\mathcal{H}_i$, for otherwise there would be a directed path $K \to V_{n-1} \to T$ from K to T other than the edge $K \to T$ in $\mathcal{H}_i$, contradiction. It follows that every edge on the subpath from V_1 to K is bi-directed in $\mathcal{H}_i$.

Next we will prove that there is an edge $V_0 *\!\!\rightarrow V_1$ in $\mathbb{M}$. Suppose for contradiction, the edge is either $V_0 \circ\!\!-\!\!\circ V_1$ or $V_0 \leftarrow\!\!\circ V_1$.

(i). Suppose $V_0 \circ\!\!-\!\!\circ V_1$ in $\mathbb{M}$. There cannot be an edge $V_1 \leftrightarrow V_2$ in $\mathbb{M}$, for otherwise there is $V_0 \leftrightarrow V_2$ in $\mathbb{M}$ due to the balanced property of $\mathbb{M}$, which contradicts with the shortest discriminating path p. Since we do not transform a circle edge in $\mathbb{M}$ to a bi-directed edge, the edge between V_1 and V_2 are either $V_1 \circ\!\!\rightarrow V_2$ or $V_1 \leftarrow\!\!\circ V_2$. For the former case, V_0 is adjacent to V_2, for otherwise $V_0 *\!\!\rightarrow V_1 \leftarrow\!\!* V_2$ is identifiable in $\mathcal{P}$ and $\mathbb{M}$ since $V_0 *\!\!\rightarrow V_1 \leftrightarrow V_2$ in $\mathcal{H}_i$ and $\mathcal{H}_i$ is a MAG Markov equivalent to $\mathcal{H}_0$ which is consistent with $\mathcal{P}$, contradicting with $V_0 \circ\!\!-\!\!\circ V_1$ in $\mathbb{M}$. According to the balanced property of $\mathbb{M}$, there is $V_0 *\!\!\rightarrow V_2$ in $\mathbb{M}$ thus there is $V_0 *\!\!\rightarrow V_2$ in $\mathcal{H}_i$, in which case there is a shorter discriminating path without V_1, contradiction. For the latter case, there is $V_0 \circ\!\!-\!\!\circ V_1 \leftarrow\!\!\circ V_2$ in $\mathbb{M}$. As shown by the orientation procedure, we only add an arrowhead at the vertex in $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$, and we never orient an edge connecting two vertices from $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ as bi-directed, hence $V_0 *\!\!\rightarrow V_1$ and $V_1 \leftrightarrow V_2$ cannot be oriented at the same time in the process of obtaining $\mathcal{H}$ from $\mathcal{H}_0$.

(ii). Suppose $V_0 \leftarrow\!\!\circ V_1$. Due to the fact that a bi-directed edge is oriented in $\mathcal{H}_i$ compared to $\mathbb{M}$ only if the edge connects a vertex in $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and a vertex in $\mathbf{S}$, and the fact that an arrowhead is added only at the vertex in $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$, there is $V_0 \in \mathbf{S}$ and $V_1 \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$. According to the first condition of Theorem 2, there is not $V_1 \to T$ in $\mathbb{M}(= \mathcal{H})$. In light of the fact that we never transform an edge to a directed edge in the process from $\mathcal{H}_0$ to $\mathcal{H}$, there can only be $V_1 \circ\!\!\rightarrow T$ in $\mathbb{M}$. According to Definition 5, $V_0, T \in \mathbf{S}_{V_1}$. According to the definition of discriminating path, V_0 is not adjacent to T, which contradicts with Lemma 6. Hence $V_0 \leftarrow\!\!\circ V_1$ in $\mathbb{M}$ is impossible.

We conclude that there is $V_0 \ast\!\!\rightarrow V_1$ in $\mathbb{M}$. The remaining part is to prove by induction that for every $1 \le i \le n-1$, V_i is a collider and a parent of T in $\mathbb{M}$. $V_1 \rightarrow T$ is evident due to the non-adjacency of V_0 and T. Note $T \in \mathbf{C}$ and $V_1 \rightarrow T$ in $\mathbb{M}$, thus $V_1 \notin \text{PossDe}(X, \mathbb{M}[-\mathbf{C}])$ due to $\text{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}]) \cap \text{Pa}(\mathbf{S}, \mathbb{M}) = \emptyset$ according to the first condition of Theorem 2. Hence, there cannot be an edge $V_1 \rightarrow V_2$ in $\mathbb{M}$ since the edge cannot be oriented as $V_1 \leftrightarrow V_2$ in $\mathcal{H}_i$. If there is not a collider at V_1 in $\mathbb{M}$, there is $V_1 \circ\!\!\rightarrow V_2$. It is impossible because we never transform it to bi-directed in the process from $\mathbb{M}$ to $\mathcal{H}_i$ as $V_1 \notin \text{PossDe}(X, \mathbb{M}[-\mathbf{C}])$. Hence the collider is identifiable in $\mathbb{M}$. Similarly, we could prove $V_2 \rightarrow T$ in $\mathbb{M}$. Then we prove there is $V_2 \leftarrow\!\!\ast V_3$ in $\mathbb{M}$. If the edge is a circle edge, then there must be $V_1 \leftrightarrow V_3$ according to the balance property, in which case there is a shorter discriminating path, contradiction. Then we consider the edge is $V_2 \ast\!\!\rightarrow V_3$. Due to $T \in \mathbf{C}$ and $\text{PossDe}(X, \mathbb{M}[-\mathbf{C}]) \cap \text{Pa}(\mathbf{C}, \mathbb{M}) = \emptyset$, $V_2 \notin \text{PossDe}(X, \mathbb{M}[-\mathbf{C}])$. Hence $V_2 \ast\!\!\rightarrow V_3$ in $\mathbb{M}$ can never be transformed to bi-directed since arrowhead is added at only the vertex in $\text{PossDe}(X, \mathbb{M}[-\mathbf{C}])$. Thus $V_1 \leftrightarrow V_2 \leftarrow\!\!\ast V_3$ is identifiable in $\mathbb{M}$. By such way, we prove that the path is a discriminating path for K in $\mathbb{M}$. Thus there cannot be an edge $K \circ\!\!\rightarrow T$ in $\mathbb{M}$, otherwise it will be oriented as $K \rightarrow T$ by $\mathcal{R}'_4$ if $i \ge 1$ and oriented as $K \rightarrow T$ or $K \leftrightarrow T$ if $i = 0$ since $\mathbb{M}$ is closed under the orientation rules, contradicting with $K \circ\!\!\rightarrow T$ in $\mathbb{M}_i$.

Hence, we conclude that $\mathcal{H}$ is a MAG Markov equivalent to $\mathcal{H}_0$. Since we have proven in **B.1.** that $\mathcal{H}_0$ is a MAG consistent with $\mathcal{P}$, $\mathcal{H}$ is a MAG consistent with $\mathcal{P}$. And according to Algorithm 3 to obtain $\mathcal{H}$, $\mathcal{H}$ has the non-circle marks in $\mathbb{M}$. Hence $\mathcal{H}$ is a MAG valid to $\mathbb{M}$. $\square$

We then show two results of Wang and Zhou [36], which are used in the main proof.

Definition 8 *(Critical vertex for (X, Y); Wang and Zhou [36]).* In a maximal local MAG $\mathbb{M}$, F_t is called a *critical vertex* for (X, Y) if there is a path $X \leftrightarrow F_1 \leftrightarrow \cdots \leftrightarrow F_{t-1} \leftrightarrow F_t$ or $X \leftrightarrow F_1 \leftrightarrow \cdots \leftrightarrow F_{t-1} \leftarrow\!\!\circ F_t$, $t \ge 1$, where each non-endpoint variable is an ancestor of X or Y in $\mathbb{M}$, and there is a non-empty variable set $\mathbf{S}$ relative to F_t defined as follows: $S \in \mathbf{S}$ if and only if in M (1) S is a child of $X, F_1, \cdots, F_{t-1}$, (2) there is one minimal possible directed path from F_t to Y in the form of $F_t \circ\!\!-\!\!\ast S \cdots Y$.

Proposition 4 *(Wang and Zhou [36]). Let $\mathbb{M}$ be a maximal local MAG of X based on a PAG $\mathcal{P}$, suppose $\mathcal{M}$ a MAG valid to $\mathbb{M}$. Denote* $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ *by* $\mathbf{D}$. *If F_t is not a critical vertex in $\mathbb{M}$ and $F_t \notin \text{Anc}(Y, \mathcal{M})$, then $F_t \perp Y \mid \mathbf{D}, X$ in $\mathcal{M}$.*

Note Proposition 4 is the part 1 of Theorem 3 of Wang and Zhou [36]. There are two modifications. One is in their paper they consider local MAG. Here the maximal local MAG is a special case of local MAG, hence the result holds also. The other is that a new notion $\text{PD-SEP}(X, Y, M)$ is used in their paper for their settings. Such a notion is redundant here. Hence we do not mention that.

Lemma 8. *Suppose a maximal local MAG $\mathbb{M}$ and a MAG $\mathcal{M}$ valid to $\mathbb{M}$. There is* $\text{PossDe}(X, \mathbb{M}) = \text{De}(X, \mathcal{M})$.

Proof. It is evident that $\text{PossDe}(X, \mathbb{M}) \supseteq \text{De}(X, \mathcal{M})$. It suffices to show $\text{PossDe}(X, \mathbb{M}) \subseteq \text{De}(X, \mathcal{M})$. Suppose a vertex $V \in \text{PossDe}(X, \mathbb{M})$. Then there is a possible directed path from X to V in $\mathbb{M}$. According to Lemma 2, there is a minimal possible directed path p_1 from X to V in $\mathbb{M}$. Since the mark at X is definite in $\mathbb{M}$, p_1 starts with a directed edge out of X. According to the completeness property of $\mathbb{M}$, p_1 can only be a directed path in $\mathbb{M}$. Hence $V \in \text{De}(X, \mathcal{M})$. We thus conclude $\text{PossDe}(X, \mathbb{M}) \subseteq \text{De}(X, \mathcal{M})$. $\square$

Lemma 9. *Given a maximal local MAG $\mathbb{M}$ where $X \in \text{Anc}(Y, \mathbb{M})$ and a potential adjustment set $\mathbf{W}$. Suppose the three conditions in Theorem 2 are satisfied and we obtain a MAG $\mathcal{M}$ according to Algorithm 3. If there is some vertex $V \in \mathbf{W} \cup \bar{\mathbf{W}}$, there always exists a collider path from X to V beginning with an arrowhead at X where each non-endpoint belongs to* $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$.

Proof. Denote the constructed MAG by $\mathcal{M}$. According to the definition of potential adjustment set $\mathbf{W}$ and $\bar{\mathbf{W}}$, there is a collider path as $X (= F_0) \leftrightarrow F_1 \leftrightarrow \cdots \leftrightarrow F_{t-1} \leftarrow\!\!\ast V$, where $F_1, F_2, \cdots, F_{t-1} \in \mathbf{W}$. If each vertex in this path is an ancestor of Y in $\mathcal{M}$, the result evidently holds. Hence we consider if there exists a vertex that is not an ancestor of Y in the collider path $X \leftrightarrow F_1 \leftrightarrow \cdots \leftrightarrow F_{t-1} \leftarrow\!\!\ast V$. We prove that in this case, we could *always* find another collider path from X to Y beginning with an arrowhead at X where each non-endpoint is an ancestor of Y.

Without loss of generality, suppose $F_i, 1 \le i \le t-1$ be the first vertex from X to V that is not an ancestor of Y in $\mathcal{M}$. According to Definition 6, (1) there cannot be an edge as $X \rightarrow F_s, 1 \le s \le i-1$, (2) there exists one minimal possible directed path p from F_i to Y that do not go through the vertex in $\bar{\mathbf{W}}$. Suppose $p = \langle F_i, V_1, \cdots, Y \rangle$. We first show that there is $F_i \circ\!\!-\!\!\ast V_1$ in $\mathbb{M}$. It is evidently not in the form of $F_i \leftarrow\!\!\ast V_1$ since p is a minimal possible directed path. If it is $F_i \rightarrow V_1$ in $\mathbb{M}$, since p is minimal possible directed and $\mathbb{M}$ is closed under orientation rules, there must be $F_i \rightarrow V_1 \rightarrow \cdots \rightarrow Y$ in $\mathbb{M}$, thus F_i is an ancestor of Y in $\mathcal{M}$, contradiction. Hence the edge can only be $F_i \circ\!\!-\!\!\ast V_1$ in $\mathbb{M}$.

Considering $F_{i-1} \leftrightarrow F_i \circ\!\!-\!\!\ast V_1$ and $F_{i+1} \ast\!\!\rightarrow F_i \circ\!\!-\!\!\ast V_1$ in $\mathbb{M}$, there is $F_{i-1} \ast\!\!\rightarrow V_1$ and $F_{i+1} \ast\!\!\rightarrow V_1$ in $\mathbb{M}$ according to the balance property of $\mathbb{M}$. At first, we prove that F_i is not a critical vertex here. Otherwise, suppose a minimal possible directed path from F_i to Y comprised of $F_i, J_1, \cdots, J_{k-1}, Y$ where $J_1 \in \text{Chd}(X, \mathcal{M})$. By Algorithm 3, we only transform $K \circ\!\!\rightarrow T$ as $K \leftrightarrow T$ when $K \in \text{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $T \in \mathbf{S}$. Since there is $F_i \circ\!\!\rightarrow J_1$ in $\mathbb{M}$ and $F_i \in \mathbf{W}$, $F_i \circ\!\!\rightarrow J_1$ is transformed to $F_i \rightarrow J_1$ in $\mathcal{M}$. In this case, there must be $F_i \rightarrow J_1 \rightarrow \cdots \rightarrow Y$

in $\mathcal{M}$ since the path is a minimal path, otherwise there will be new unshielded colliders in $\mathcal{M}$ relative to $\mathbb{M}$, impossibility. There is thus $F_i \in \mathrm{Anc}(Y, \mathcal{M})$, contradiction. F_i cannot be a critical vertex here. It is easy to prove that there exists a vertex $F_j, 0 \le j \le i-1$ such that there is an edge $F_j \leftrightarrow V_1$, and a vertex $F_k, i < k \le t-1$ such that there is an edge $V_1 \leftrightarrow F_k$ or an edge $F_t \ast\!\!\rightarrow V_1$ (otherwise there is a discriminating path for F_i, hence the circle at $F_i \circ\!\!-\!\!\ast V_1$ should be identified in $\mathcal{P}$, we do not show the details). In this case, if $V_1 \notin \mathbf{W}$, then $V_1 \in \bar{\mathbf{W}}$, contradiction. Hence V_1 could only belong to $\mathbf{W}$. Then we consider whether $V_1 \in \mathrm{Anc}(Y, \mathbb{M})$, in which case $V_1 \in \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$. If it is, we find a new collider path $X \leftrightarrow \cdots \leftrightarrow F_j \leftrightarrow V_1 \leftrightarrow F_k \leftrightarrow \cdots F_{t-1} \leftarrow\!\!\ast F_t$ or $X \leftrightarrow \cdots \leftrightarrow F_j \leftrightarrow V_1 \leftarrow\!\!\ast F_t$. If $V_1 \notin \mathrm{Anc}(Y, \mathbb{M})$, we conclude that V_1 is not a critical vertex as the process above. And similarly, we conclude $V_2 \in \mathbf{W}$ and further consider V_2. By such way, we could always find a vertex $V_t, 1 \le t \le k-1$ such that there is a collider path $X \leftrightarrow \cdots \leftrightarrow V_t \leftrightarrow \cdots \leftarrow\!\!\ast F_t$ and the first possible vertex that is not ancestor of Y in $\mathcal{M}$ is only possible in the sub-path from F_{i+1} to F_t. If such a vertex exists, $F_u, i+1 \le u \le t$ for example, repeat the process above and we could find a new collider path. Repeat the process above and we could always find a collider path beginning with arrowhead at X satisfying that each non-endpoint is an ancestor of Y in $\mathcal{M}$. It is evident that each non-endpoint in this path belongs to $\mathbf{D}$. Hence there is a collider path from X to V beginning with an arrowhead at X where each non-endpoint belongs to $\mathbf{D}$. □

Lemma 10. *If there exists a minimal collider path in a MAG $\mathcal{M}$ consistent with a PAG $\mathcal{P}$, then it is also a collider path in $\mathcal{P}$.*

Proof. Suppose a minimal collider path p in $\mathcal{M}$, we consider its corresponding path in $\mathcal{P}$. If there exists a circle or tail at the non-endpoint vertex on this path, according to the completeness of FCI [6], there exists a MAG Markov equivalent to $\mathcal{M}$ that has a tail there, which contradicts Theorem 2.1 of Zhao et al. [37] that Markov equivalent MAGs have the same minimal collider paths. Hence the corresponding path of p in $\mathcal{P}$ is also a collider path. □

Theorem 2. *Given a maximal local MAG $\mathbb{M}$, for any potential adjustment set $\mathbf{W}$, there exists a MAG $\mathcal{M}$ valid to $\mathbb{M}$ such that $\mathbf{W}$ is an adjustment set in $\mathcal{M}$ if there exists a block set $\mathbf{S}$ such that*

(1) $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}]) \cap \mathrm{Pa}(\mathbf{S}, \mathbb{M}) = \emptyset$*;*
(2) $\mathbb{M}[\mathbf{S}_V]$ *is a complete graph for any* $V \in \bar{\mathbf{W}}$*, where* $\mathbf{S}_V = \{V' \in \mathbf{S} | V \circ\!\!-\!\!\ast V' \text{ in } \mathbb{M}\}$*;*
(3) $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ *is bridged relative to* $\mathbf{S}$ *in* $\mathbb{M}$.

Proof. We will prove that $\mathbf{W}$ is an adjustment set relative to (X, Y) in the MAG $\mathcal{M}$ obtained by Algorithm 3.

The whole proof process is comprised of three parts: (1) for $\forall V \in \mathbf{W}$, there is $V \in \mathbf{D}$ or $V \perp Y \mid \mathbf{D}, X$; (2) for $\forall V \in \bar{\mathbf{W}}$, $V \notin \mathrm{Anc}(Y, \mathcal{M})$; (3) If $V \in \mathbf{D}$, then $V \in \mathbf{W}$. Then by the first and third parts, we could get the desired results.

(1) For $\forall V \in \mathbf{W}$, there is $V \in \mathbf{D}$ or $V \perp Y \mid \mathbf{D}, X$.

In this part, we first prove that if $V \in \mathbf{W}$, there is either $V \in \mathrm{Anc}(Y, \mathcal{M})$ or $V \perp Y \mid \mathbf{D}, X$. According to Definition 6 and Lemma 8, $V \notin \mathrm{De}(X, \mathcal{M})$.

If $V \in \mathrm{Anc}(Y, \mathbb{M})$, it is evident that $V \in \mathrm{Anc}(Y, \mathcal{M})$. Hence we only consider $V \notin \mathrm{Anc}(Y, \mathbb{M})$ in the following. According to the condition of $\mathbf{W}$, by Lemma 9, there is a collider path as $X(=F_0) \leftrightarrow F_1 \leftrightarrow \cdots \leftrightarrow F_{t-1} \leftarrow\!\!\ast V$ where each non-endpoint belongs to $\mathbf{D}$.

If the edge between F_{t-1} and V is as $F_{t-1} \leftarrow\!\!\circ V$ in $\mathbb{M}$, by Algorithm 3, there is $F_{t-1} \leftarrow V$ in $\mathcal{M}$, thereby $V \in \mathrm{Anc}(Y, \mathcal{M})$. If the edge is $F_{t-1} \leftrightarrow V$ in $\mathbb{M}$, according to Definition 6, there exist some minimal possible directed paths from V to Y in $\mathbb{M}$. **(i).** If V is a critical vertex, suppose a minimal possible directed path from V to Y comprised of $V, V_1, \cdots, V_{k-1}, Y$ where $V_1 \in \mathrm{Chd}(X, \mathcal{M})$. In this case there must be $V \circ\!\!\rightarrow V_1$ in $\mathbb{M}$ due to the balanced as well as closed property of $\mathbb{M}$ and the definition of critical vertex. By Algorithm 3, we orient $V \circ\!\!\rightarrow V_1$ as $V \rightarrow V_1$, in this case there is $V \in \mathrm{Anc}(Y, \mathbb{M})$ (a new bi-directed edge is additionally introduced only if $V \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$, $V_1 \in \mathbf{S}$, and there is $V \circ\!\!\rightarrow V_1$ in $\mathbb{M}$. However, when $V \in \mathbf{W}$, $V \notin \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$, hence the edge can only be oriented as $V \rightarrow V_1$.). **(ii).** If V is not a critical vertex, if $V \in \mathrm{Anc}(Y, \mathcal{M})$, there is $V \in \mathbf{D}$. If $V \notin \mathrm{Anc}(Y, \mathcal{M})$, by Proposition 4, it holds $V \perp Y \mid \mathbf{D}, X$.

(2) for $\forall V \in \bar{\mathbf{W}}$, $V \notin \mathrm{Anc}(Y, \mathcal{M})$.

Suppose there is a vertex $V \in \bar{\mathbf{W}} \cap \mathrm{Anc}(Y, \mathcal{M})$ such that $V \in \mathrm{Anc}(Y, \mathcal{M})$ for contradiction. According to the definition of $\bar{\mathbf{W}}$, for $\forall V \in \bar{\mathbf{W}}$, there exists a collider path $X(=F_0) \leftrightarrow F_1 \leftrightarrow \cdots \leftrightarrow F_{t-1} \leftarrow\!\!\ast V$ where each non-endpoint belongs to $\mathbf{W}$. If there exists one directed path from V to Y in $\mathcal{M}$, thus there is a minimal directed path from V to Y in $\mathcal{M}$. According to the condition $\bar{\mathbf{W}} \cap \mathrm{Anc}(Y, \mathbb{M}) = \emptyset$, there does not exist one directed path from V to Y in $\mathbb{M}$, hence the minimal directed path in $\mathcal{M}$ is not a directed path in $\mathbb{M}$.

If there exists some vertex $V' \in \bar{\mathbf{W}}$ that does not go through the vertex in $\bar{\mathbf{W}}$ except for V' in the path, there is $V' \in \mathrm{Anc}(Y, \mathcal{M})$. We consider V' instead of V. By such operations we could always find a vertex in $\bar{\mathbf{W}}$ that has a minimal possible directed path to Y in $\mathcal{M}$ where each non-endpoint does not belong to $\bar{\mathbf{W}}$. Hence, we suppose there does not exist a vertex in $\bar{\mathbf{W}}$ in the minimal possible directed path from V to Y without loss of generality. That is, each non-endpoint in the path does not belong to $\bar{\mathbf{W}}$.

Then there is a minimal directed path from V to Y in $\mathcal{M}$. We consider its corresponding path p in $\mathbb{M}$. Suppose $p = \langle V, V_1, \cdots, V_k(= Y)\rangle, k \ge 1$. p must be a minimal possible directed path from V to Y. Note p is not a directed path in $\mathbb{M}$, for otherwise it contradicts with the third condition of Definition 6. In light of the closed property of $\mathbb{M}$, there is a vertex $V_s, 1 \le s \le k$ in p which is an ancestor of Y in $\mathbb{M}$ and any vertex in $V, V_1, \cdots, V_{s-1}$ is not an ancestor of Y in $\mathbb{M}$. If for any V_i in $V_1, \cdots, V_{s-1}$ there is $V_i \notin \mathbf{S}$, then there is $V_s \in \mathbf{S}$ according to Definition 5. According to the first step of Algorithm 3, there is $V_{s-1} \leftarrow\!\!\ast V_s$ oriented in $\mathcal{M}$, which contradicts with the fact that the corresponding path of p in $\mathcal{M}$ is a directed path. If there is some V_i in $V_1, \cdots, V_{s-1}$ such that $V_i \in \mathbf{S}$. Without loss of generality, suppose $V_1, V_2, \cdots, V_{j-1} \notin \mathbf{S}$. In this case there is $V_{j-1} \leftarrow V_j$ oriented in the first step of Algorithm 3 thus there is

$V_{j-1} \leftarrow V_j$ in $\mathcal{M}$, which contradicts with the fact that the corresponding path of p in $\mathcal{M}$ is a directed path. Hence both of the cases are impossible. Hence for any $V \in \bar{\mathbf{W}}$, there is $V \notin \mathrm{Anc}(Y, \mathcal{M})$.

(3) $\mathbf{D} \subseteq \mathbf{W}$. Suppose $V \in \mathbf{D}\backslash\mathbf{W}$, then there exists a minimal collider path $p = X \leftrightarrow F_1 \cdots \leftrightarrow F_{t-1} \leftarrow\!\!* V(= F_t)$ where each vertex is an ancestor of Y in $\mathcal{M}$. Without loss of generality, we suppose $F_1, \cdots, F_{t-1} \in \mathbf{W}$ and $V \in \mathbf{D}\backslash\mathbf{W}$ (if $F_i \notin \mathbf{W}$, we consider F_i instead of V), we will prove the result by showing $V \in \mathbf{W}$, which contradicts with $V \in \mathbf{D}\backslash\mathbf{W}$. If there is not an edge as $X \to V$, due to the minimal collider path, the collider path p is identifiable in $\mathcal{P}$ by Lemma 10. According to $V \in \mathbf{D}$, it follows $V \in \mathrm{PossAn}(Y, \mathbb{M})$. Since $F_1, \cdots, F_{t-1} \in \mathbf{W}$, there is $V \in \mathbf{W} \cup \bar{\mathbf{W}}$. Hence if $V \notin \mathbf{W}$, there is $V \in \bar{\mathbf{W}}$. However, as we have shown in the previous part, for $V \in \bar{\mathbf{W}}$, $V \notin \mathrm{Anc}(Y, \mathcal{M})$, contradiction.

If there is $X \to V$, due to $F_1, \cdots, F_{t-1} \in \mathbf{W}$ and $\mathbf{W} \cap \mathrm{PossDe}(X, \mathbb{M}) = \emptyset$, there is not an edge as $X \to F_i, 1 \le i \le t-1$. Note in Algorithm 3 we never orient an edge $A \circ\!\!\to B$ where $A, B \in \mathbf{W} \cup \{X\}$ as bi-directed, thus if $X \leftrightarrow F_1 \leftrightarrow \cdots \leftrightarrow F_{t-1}$ in $\mathcal{M}$ is obtained by Algorithm 3, there is $X \leftrightarrow F_1 \leftrightarrow \cdots \leftrightarrow F_{t-1}$ in $\mathbb{M}$. Also note we never add an arrowhead at a vertex in $\mathbf{W}$ in a non-circle edge connecting a vertex in $\mathbf{W}$ and a vertex not in $\mathbf{W}$. Hence if $F_{t-1} \leftarrow\!\!* V$ in $\mathcal{M}$, there is either $F_{t-1} \leftarrow\!\!* V$ in $\mathbb{M}$, or $F_{t-1} \circ\!\!-\!\!\circ V$ in $\mathbb{M}$. For the former case, since $X \leftrightarrow F_1 \cdots \leftrightarrow F_{t-2} \leftrightarrow F_{t-1} \leftarrow\!\!* V$ is in $\mathbb{M}$, there is $V \in \mathbf{W} \cup \bar{\mathbf{W}}$. And since $V \in \mathbf{D}\backslash\mathbf{W}$, there is $V \in \bar{\mathbf{W}} \cap \mathbf{D}$. However, as we have shown in the previous part, for $V \in \bar{\mathbf{W}}$, $V \notin \mathrm{Anc}(Y, \mathcal{M})$, contradiction. For the latter case, since circle edge will only be transformed to directed edge by Algorithm 3, and there is $V_{t-1} \leftarrow\!\!* V_t$ in $\mathcal{M}$, there is $F_{t-1} \leftarrow V$ in $\mathcal{M}$. Due to $X \to V$ in $\mathbb{M}$, there is $F_{t-1} \in \mathrm{PossDe}(X, \mathbb{M})$. Hence $F_{t-1} \in \mathbf{W} \cap \mathrm{PossDe}(X, \mathbb{M})$, contradiction with the second condition of Definition 6. Hence both of the cases are impossible. We thus conclude $\mathbf{D} \subseteq \mathbf{W}$.

Combining the first the third results, $\mathbf{D} \subseteq \mathbf{W}$; and for $V \in \mathbf{W}\backslash\mathbf{D}$, there is $V \perp Y \mid \mathbf{D}, X$. Since $\mathbf{W} \cap \mathrm{De}(X, \mathbb{M}) = \emptyset$, there is $\mathbf{D} \cap \mathrm{De}(X, \mathcal{M}) = \emptyset$, hence in this case $\mathbf{D}$ is an adjustment set relative to (X, Y) in $\mathcal{M}$. And since $V \perp Y \mid \mathbf{D}, X$ for $V \in \mathbf{W}\backslash\mathbf{D}$, $\mathbf{W}$ is also an adjustment set relative to (X, Y) in $\mathcal{M}$ by the following equations

$$
\begin{aligned}
\int_{\mathbf{W}} f(\mathbf{W}) f(Y \mid \mathbf{W}, X)\,\mathrm{d}\mathbf{W} &= \int_{\mathbf{W}} f(\mathbf{W}) f(Y \mid \mathbf{D}, X)\,\mathrm{d}\mathbf{W} \\
&= \int_{\mathbf{D}} f(\mathbf{D}) f(Y \mid \mathbf{D}, X)\,\mathrm{d}\mathbf{D} \quad (\because \mathbf{W}\backslash\mathbf{D} \perp Y \mid \mathbf{D}, X) \\
&= f(Y \mid do(X)). \quad (\because \text{Thm. 1 and (1)}) \quad \square
\end{aligned}
$$

3.5.3. Proof of Theorem 3

Theorem 3. *Given a maximal local MAG $\mathbb{M}$, suppose a MAG $\mathcal{M}$ valid to $\mathbb{M}$ such that there exists an adjustment set relative to (X, Y). Let $\mathbf{W}$ be $\mathrm{D\text{-}SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$. Then $\mathbf{W}$ is a potential adjustment set in $\mathbb{M}$ and there exists a block set $\mathbf{S}$ such that*

(1) $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}]) \cap \mathrm{Pa}(\mathbf{S}, \mathbb{M}) = \emptyset$*;*
(2) $\mathbb{M}[\mathbf{S}_V]$ *is a complete graph for any* $V \in \bar{\mathbf{W}}$*, where* $\mathbf{S}_V = \{V' \in \mathbf{S} | V \circ\!\!\to\!\!* V' \text{ in } \mathbb{M}\}$*;*
(3) $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ *is bridged relative to* $\mathbf{S}$ *in* $\mathbb{M}$.

Proof. Since there exists an adjustment set relative to (X, Y) in $\mathcal{M}$, there is $\mathrm{D\text{-}SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) \cap \mathrm{De}(X, \mathcal{M}) = \emptyset$ according to Theorem 1.

At first, we show that $\mathbf{W} = \mathrm{D\text{-}SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ is a potential adjustment set as Definition 6 in $\mathbb{M}$, by proving $\mathbf{W} = \mathrm{D\text{-}SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ satisfied the three conditions in Definition 6. Note $\mathbf{W} \cap \bar{\mathbf{W}}$ cannot be a non-empty set according to the definition of $\mathbf{W}$ and $\bar{\mathbf{W}}$.

For the first condition, consider $V \in \mathbf{W} = \mathrm{D\text{-}SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$. According to the definition of $\mathrm{D\text{-}SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$, there is a collider path from X to V where each non-endpoint belongs to $\mathrm{D\text{-}SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$. Since $V \in \mathbf{W}$, there is a directed path from V to Y. Thus there is a possible directed path from V to Y in $\mathbb{M}$. If there exists a minimal possible directed path from V to Y in $\mathbb{M}$ which does not go through the vertex in $\bar{\mathbf{W}}$, then there is $V \in \mathbf{W}$. If not, it follows that for each possible directed path from V to Y in $\mathbb{M}$, the path goes through the vertex in $\bar{\mathbf{W}}$. Since there is at least one minimal possible directed path in $\mathbb{M}$ is directed in $\mathcal{M}$, there exists some vertex $V' \in \bar{\mathbf{W}}$ which is an ancestor of Y. According to Definition 5, there is $V' \in \mathbf{W}$, thus it holds that $V' \in \mathbf{W} \cap \bar{\mathbf{W}}$, contradiction.

For the second condition, suppose $V \in \mathbf{W} \cap \mathrm{PossDe}(X, \mathbb{M})$. According to Lemma 8, there is $V = \mathbf{W} \cap \mathrm{De}(X, \mathcal{M}) \neq \emptyset$, contradicting with Theorem 1.

For the third condition, suppose $V \in \bar{\mathbf{W}} \cap \mathrm{Anc}(Y, \mathbb{M})$. There is $V \in \mathrm{Anc}(Y, \mathcal{M})$. According to the definition of $\bar{\mathbf{W}}$, there is a collider path in the form of $X \leftrightarrow \cdots \leftarrow\!\!* V$ in $\mathbb{M}$, hence there is $V \in \mathbf{D}(= \mathbf{W})$ due to $V \in \mathrm{Anc}(Y, \mathbb{M})$, thus it holds $\mathbf{W} \cap \bar{\mathbf{W}} \neq \emptyset$, contradiction.

Hence $\mathbf{W}$ is a potential adjustment set. Then we set $\mathbf{S} = \mathrm{Anc}(Y \cup \mathbf{W}, \mathcal{M}) \cap [\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M})\backslash\bar{\mathbf{W}}]$ which evidently fulfills that $\mathrm{Anc}(Y \cup \mathbf{W}, \mathbb{M}) \cap [\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M})\backslash\bar{\mathbf{W}}] \subseteq \mathbf{S} \subseteq \mathrm{PossAn}(Y \cup \mathbf{W}, \mathbb{M}) \cap [\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M})\backslash\bar{\mathbf{W}}]$. It suffices to show that the three conditions in Theorem 3 are satisfied for such $\mathbf{S}$.

For the first condition, suppose $V \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}]) \cap \mathrm{Pa}(\mathbf{S}, \mathbb{M})$. Suppose $V \to S$ where $S \in \mathbf{S}$. According to the selected $\mathbf{S}$ and $\mathbf{D} = \mathbf{W}$, there is $V \in \mathrm{Anc}(Y, \mathcal{M})$. In this case, there can only be $V \in \bar{\mathbf{W}}$, for otherwise there is $V \in \mathbf{S}$, which contradicts with $V \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$. Note V is an ancestor of Y in $\mathcal{M}$, hence $V \in \mathbf{W}$. Thus $V \in \mathbf{W} \cap \bar{\mathbf{W}}$, contradiction.

For the second condition, if for $V \in \bar{\mathbf{W}}$, $\mathbb{M}[\mathbf{S}_V]$ is not a complete graph, to generate no new unshielded colliders in $\mathcal{M}$ relative to $\mathbb{M}$, there is a directed edge $V \to V'$ where $V' \in \mathbf{S}_V$, in this case there must be a directed path from V to Y in $\mathcal{M}$ and thus

$V \in \mathrm{Anc}(Y, \mathcal{M})$. According to the definition of $\bar{\mathbf{W}}$, there is $V \in \mathbf{W}$, thus it holds that $V \in \mathbf{W} \cap \bar{\mathbf{W}}$, contradiction. Hence the second condition is satisfied.

For the third condition, if it is not bridged, then there must be vertice $V \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$ and $S \in \mathbf{S}$ such that there is $V \to S$ in $\mathcal{M}$. Note $\mathbf{S} = \mathrm{Anc}(Y \cup \mathbf{W}, \mathcal{M}) \cap [\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}) \backslash \bar{\mathbf{W}}]$. Hence there can only be $V \in \bar{\mathbf{W}}$, for otherwise there is $V \in \mathbf{S}$, which contradicts with $V \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])$. Similar to the proof of the first condition, in this case there is $V \in \mathbf{W} \cap \bar{\mathbf{W}}$, contradiction. $\square$

3.5.4. Proof of Proposition 3

Proof. We prove it by mathematical induction. For $V \in \text{DD-SEP}(X, Y, \mathbb{M}_{\underset{\sim}{X}})$, we consider the minimal collider path satisfying the three conditions of Definition 7 in $\mathbb{M}$. If the length is 1, there is $X \leftarrow\!\!* \underset{\sim}{V}$ in $\mathbb{M}_{\underset{\sim}{X}}$ since there cannot be $X \multimap\!\!\!\!\!* V$ in $\mathbb{M}_{\underset{\sim}{X}}$ according to the definition of $\mathbb{M}_{\underset{\sim}{X}}$. If there is $V \in \mathrm{Anc}(Y, \mathbb{M})$, it trivially concludes that $V \in \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ according to the definition. If $\mathbb{M}[\mathcal{Q}_V]$ is not a complete graph, evidently there are at least two vertices in $\mathcal{Q}_V$. Suppose $S_1, S_2 \in \mathcal{Q}_V$ are not adjacent. It is evident that there is either $V \to S_1$ or $V \to S_2$ in $\mathcal{M}$, otherwise there is a new unshielded collider in $\mathcal{M}$ relative to $\mathbb{M}$ and $\mathcal{P}$, which contradicts with that $\mathcal{M}$ is valid to $\mathbb{M}$. Thus there is also $V \in \mathrm{Anc}(Y, \mathcal{M})$ such that $V \in \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$. Hence if the length is 1, there is $V \in \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$. Suppose the result holds when the length of the minimal collider path mentioned above is k. For the vertex V with a minimal collider path satisfying the three conditions of Definition 7 whose length is $k+1$, suppose the path is comprised of $X, V_1, V_2, \cdots, V_{k+1}$. We have $V_1, V_2, \cdots, V_k \in \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$. For V_{k+1}, similar to the proof above, no matter whether $V_{k+1} \in \mathrm{Anc}(Y, \mathbb{M})$ or $\mathbb{M}[\mathcal{Q}_{V_{k+1}}]$ is not a complete graph, there is always $V_{k+1} \in \mathrm{Anc}(Y, \mathbb{M})$, thus $V_{k+1} \in \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ due to the collider path where each non-endpoint belongs to $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$. $\square$

4. Refined method with novel orientation rules

In this section, we present two novel rules to incorporate background knowledge (BK) into a partial mixed graph (PMG) H. As there have established sound and complete rules to obtain a PAG with observational data in the literature [6], we do not consider the stage of identifying a PAG. Hence, we restrict that H is a PAG or a graph obtained from a PAG by orienting some circles. Also, when establishing the rules for incorporating BK, we assume that the introduced BK is correct, *i.e.*, there exist MAGs consistent with the PMG and BK. Based on the novel orientation rules, we further enhance PAGcauses by using the rules to prevent enumerating block sets on Line 10 of Algorithm 1, which can reduce the complexity exponentially.

4.1. Proposed orientation rules

In this part, we present two novel rules to incorporate BK into a PMG H. At first, we introduce an important concept, *unbridged path relative to* $\mathbf{V}'$ in a PMG H, in Definition 9, where $\mathbf{V}'$ is a subset of vertices in H. Intuitively, an unbridged path p relative to $\mathbf{V}'$ is a path with an *intriguing property*: if every vertex in p is not an ancestor of $\mathbf{V}'$ in H, then every vertex in p *must be* an ancestor of $\mathbf{V}'$ in any MAG consistent with H.

Definition 9 *(Unbridged path relative to* $\mathbf{V}'$*).* Denote H a PMG. If there is an uncovered circle path $p : V_0 \circ\!\!-\!\!\circ V_1 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ V_n, n \geq 1$ in $H[-\mathbf{V}']$ such that $\mathcal{F}_{V_0} \backslash \mathcal{F}_{V_1} \neq \emptyset$ and $\mathcal{F}_{V_n} \backslash \mathcal{F}_{V_{n-1}} \neq \emptyset$, where $\mathcal{F}_{V_i} = \{V \in \mathbf{V}' \mid V *\!\!-\!\!\circ V_i \text{ or } V *\!\!\rightarrow V_i \text{ in } H\}$, then p is an unbridged path relative to $\mathbf{V}'$.

Remark 4. One may wonder why the abovementioned property holds for unbridged path p if every vertex in p is not an ancestor of $\mathbf{V}'$ in H. The reason is, in any MAG $\mathcal{M}$ consistent with H, there cannot be additional unshielded colliders relative to H, which introduce additional conditional independence such that the graphs do not belong to the MEC. Suppose $C_1 \in \mathcal{F}_{V_0} \backslash \mathcal{F}_{V_1}$ and $C_2 \in \mathcal{F}_{V_n} \backslash \mathcal{F}_{V_{n-1}}$ according to Definition 9. Since (1) $C_1 \notin \mathcal{F}_{V_1}$ and (2) V_1 is not an ancestor of $C_1 \in \mathbf{V}'$ in H, we can conclude that C_1 is not adjacent to V_1. Similarly, C_2 is not adjacent to V_{n-1}. Hence, to avoid generating unshielded colliders, the corresponding path in $\mathcal{M}$ of p as well as C_1 and C_2 can only be $C_1 *\!\!-\!\!* V_0 \to \cdots \to V_n \to C_2$, $C_1 \leftarrow V_0 \leftarrow \cdots \leftarrow V_n *\!\!-\!\!* C_2$, or $C_1 \leftarrow V_0 \leftarrow \cdots \leftarrow\!\!* V_i \to \cdots V_n \to C_2$. In any case, any vertex in p is an ancestor of either C_1 or C_2. See Fig. 3(a) for an example. $D \circ\!\!-\!\!\circ E$ is an unbridged path relative to $\mathbf{V}' = \{C_1, C_2\}$ due to $C_1 \in \mathcal{F}_D \backslash \mathcal{F}_E$ and $C_2 \in \mathcal{F}_E \backslash \mathcal{F}_D$. If we transform all the circles in $C_1 \leftarrow\!\!\circ D \circ\!\!-\!\!\circ E \circ\!\!\rightarrow C_2$ without generating unshielded colliders, D and E must be ancestors of either C_1 or C_2.

Next, we present rule $\mathcal{R}_{12}$ inspired by the property above, and then $\mathcal{R}_{13}$ as a supplement of the case of $\mathcal{R}_{12}$ when some vertex in the unbridged path *has been* an ancestor of $\mathbf{V}'$ in H.

- $\mathcal{R}_{12}$ Suppose an edge $A \circ\!\!-\!\!* B$ in a PMG H. Let $\mathbf{S}_A = \{V \in \mathbf{V}(H) | V *\!\!\rightarrow A \text{ in } H\} \cup \{A\}$. If there is an unbridged path $\langle K_1, \cdots, K_m \rangle$ relative to $\mathbf{S}_A$ in $H[-\mathbf{S}_A]$ and for every vertex $K_i \in \{K_1, \cdots, K_m\}$, there exists an uncovered possible directed path $\langle A, B, \cdots, K_i \rangle$ $(B \neq K_i)$, then orient $A \circ\!\!-\!\!* B$ as $A \leftarrow\!\!* B$.
- $\mathcal{R}_{13}$ Suppose an edge $A \circ\!\!-\!\!* B$ in a PMG H. Let $\mathbf{S}_A = \{V \in \mathbf{V}(H) | V *\!\!\rightarrow A \text{ in } H\} \cup \{A\}$. If in H there is an uncovered possible directed path $\langle A, B, \cdots, K \rangle$, where $K \in \mathrm{Anc}(\mathbf{S}_A, H)$, then orient $A \circ\!\!-\!\!* B$ as $A \leftarrow\!\!* B$.

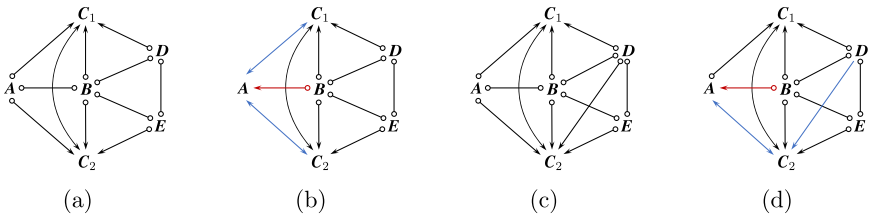

Fig. 3. Two examples for $\mathcal{R}_{12}$ and $\mathcal{R}_{13}$. Two PAGs are shown in Fig. 3(a) and 3(c). Blue lines denote the edges transformed according to BK, red lines denote the edges transformed by $\mathcal{R}_{12}$ and $\mathcal{R}_{13}$. Fig. 3(b) shows a PMG transformed from Fig. 3(a) with additional BK and $\mathcal{R}_{12}$. Fig. 3(d) shows a PMG transformed from Fig. 3(c) with additional BK and $\mathcal{R}_{13}$.

Algorithm 4: Implementation of $\mathcal{R}_{12}$ and $\mathcal{R}_{13}$.

```
   Input: PMG H
1  for each vertex A with circles in H do
2  |  for each edge A o-* B in H do
3  |  |  Obtain S_A = {V ∈ V(H) | V *-> A in H} ∪ {A};
4  |  |  Obtain a set of vertices D defined as V ∈ D if and only if V ∈ V(H)\S_A and there is an uncovered path p from A to V where B is the vertex
   |  |    adjacent to A in p;
5  |  |  if there exists V ∈ D such that V ∈ Anc(S_A, H) then
6  |  |  |  Transform A o-* B to A <-* B;
7  |  |  |  Restart the loop on Line 2;
8  |  |  else
9  |  |  |  Obtain graph H' based on H by transforming V o-* V' to V <-* V', ∀V ∈ D, ∀V' ∈ S_A;
10 |  |  |  Update the circle component in H'[D] as follows until no updates: for V_i, V_j ∈ D, transform V_i o-o V_j into V_i -> V_j if either of the two conditions
   |  |  |    holds (1) F_{V_i}\F_{V_j} ≠ ∅; or (2) there is a vertex V_k ∈ D such that there is V_k -> V_i and V_k is not adjacent to V_j, where
   |  |  |    F_V = {V' ∈ S_A | V' *-o V or V' *-> V in H};
11 |  |  |  if there are new unshielded colliders in H' then
12 |  |  |  |  Transform A o-* B to A <-* B in H;
13 |  |  |  |  Restart the loop on Line 2
   Output: Updated H
```

We present two examples for $\mathcal{R}_{12}$ and $\mathcal{R}_{13}$ in Fig. 3. Consider PAG in Fig. 3(a) and BK $C_1 *\!\!\rightarrow A \leftarrow\!\!* C_2$ in Fig. 3(b). See $\mathcal{R}_{12}$, $\mathbf{S}_A = \{C_1, C_2, A\}$, and there exist uncovered possible directed paths $p_1 = \langle A, B, D\rangle$ and $p_2 = \langle A, B, E\rangle$ from A to D and E, respectively. $D \circ\!\!-\!\!\circ E$ is unbridged relative to $\mathbf{S}_A$. Hence, the edge between A and B is transformed to $A \leftarrow\!\!\circ B$ by $\mathcal{R}_{12}$. This transformation is intuitive after knowing the property of unbridged path. In Fig. 3(b), no vertex in the unbridged path is an ancestor of $\mathbf{S}_A$, thus D and E are ancestors of $\mathbf{S}_A$ in any MAG consistent with H. Without loss of generality, suppose D is ancestor of C_1. Due to the uncovered possible directed path p_1, if there is $A \rightarrow B$, p_1 can only be a directed path from A to D, and thus there is an almost directed cycle $A \rightarrow B \rightarrow D \rightarrow C_1 \leftrightarrow A$, which violates the ancestral property. For $\mathcal{R}_{13}$, see a PAG in Fig. 3(c). If BK is $C_2 \leftrightarrow A$ and $D \rightarrow C_2$ as Fig. 3(d), there is $\mathbf{S}_A = \{A, C_2\}$ and an uncovered possible directed path $\langle A, B, D\rangle$ where $D \in \mathrm{Anc}(\mathbf{S}_A, H)$. Hence we transform $A \circ\!\!-\!\!\circ B$ to $A \leftarrow\!\!\circ B$ for the same reason as above.

We present Theorem 5 to imply the soundness of $\mathcal{R}_{12}$ and $\mathcal{R}_{13}$ to incorporate BK. Note previous rules [6,19] cannot trigger these two transformations. Recently, Venkateswaran and Perkovic [20] independently discover $\mathcal{R}_{13}$, along with some fundamental results, while $\mathcal{R}_{12}$ is not involved.

Theorem 5. *$\mathcal{R}_{12}$ and $\mathcal{R}_{13}$ are sound to incorporate BK.*

One remaining issue is the implementation of $\mathcal{R}_{12}$ and $\mathcal{R}_{13}$. We provide Algorithm 4 to implement $\mathcal{R}_{12}$ and $\mathcal{R}_{13}$,[4] with theoretical guarantee in Proposition 5. Essentially, the edges transformed on Line 6 of Algorithm 4 are triggered by $\mathcal{R}_{13}$. Line 11 involves detecting the presence of unbridged paths, during which new unshielded colliders are generated if such paths are found. The edges transformed on Line 12 are triggered by $\mathcal{R}_{12}$. Suppose there are m edges and d vertices in H, the complexity of implementing Algorithm 4 is $\mathcal{O}(m^3 d)$: there are m edges possibly transformed. Since the loop on Line 2 could be repeated by d times at most due to Line 7 and Line 13, the number of loops on Line 1 and Line 2 is less than md. In each round, suppose we detect whether $A \circ\!\!-\!\!* B$ can be transformed by $\mathcal{R}_{12}$ or $\mathcal{R}_{13}$. The complexity of implementing Line 3 and Line 4 is $\mathcal{O}(d)$. And determining $\mathrm{Anc}(\mathbf{S}_A, H)$ takes a $\mathcal{O}(md)$ complexity. Implementing Line 9 and 11 takes a $\mathcal{O}(m)$ complexity. Implementing Line 10 takes a $\mathcal{O}(m^2)$ complexity. Hence the complexity of Algorithm 4 is $\mathcal{O}(m^3 d)$.

Proposition 5. *Given a PMG H, Algorithm 4 can transform all and only the edges that can be transformed by $\mathcal{R}_{12}$ or $\mathcal{R}_{13}$.*

[4] The main focus of our paper is not on implementation. There is possibly a more efficient method to implement $\mathcal{R}_{12}$ and $\mathcal{R}_{13}$, which we leave for future work.

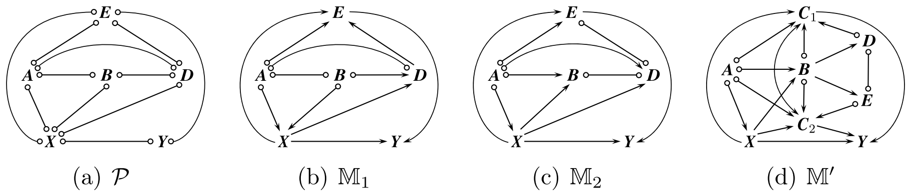

Fig. 4. Fig. 4(a) shows a PAG $\mathcal{P}$. $\mathbb{M}_1$ and $\mathbb{M}_2$ in Fig. 4(b) and 4(c) are two maximal local MAGs obtained after different local transformations of X. Fig. 4(c) and 4(d) show two examples for implementing Algorithm 5 given $\mathbf{W} = \emptyset$, where $\mathbf{S} = \{E\}$ and $\mathbf{S} = \{C_1, C_2, B\}$ are returned, respectively.

Note $\mathcal{R}_{12}$ is quite different from existing rules of Zhang [6], Wang et al. [19], Venkateswaran and Perkovic [20]. The existing rules transform an edge based on just few edges or paths. $\mathcal{R}_{12}$ is more complicated, as it considers not only the unbridged path, but also all the paths from A (in $\mathcal{R}_{12}$) to every vertex in the unbridged path, which form a sub-graph. The establishment of $\mathcal{R}_{12}$ implies the intrinsic hardness of causal identification from observational data and BK in the presence of latent variables.

Interestingly, we find that $\mathcal{R}_{12}$ and $\mathcal{R}_{13}$ can be seen as two *generalizations* of $\mathcal{R}_3$ and $\mathcal{R}_2$. Suppose a PMG H. Consider there is $E \to C \ast\!\!\to A$: $\mathcal{R}_2$ says if there is an edge $A \circ\!\!-\!\!\ast E$, then we orient it as $A \leftarrow\!\!\ast E$; while $\mathcal{R}_{13}$ says if there is an uncovered possible directed path from A to E beginning with $A \circ\!\!-\!\!\ast B$, then we orient $A \circ\!\!-\!\!\ast B$ as $A \leftarrow\!\!\ast B$. $\mathcal{R}_{13}$ generalizes an edge $A \circ\!\!-\!\!\ast E$ in $\mathcal{R}_2$ to an uncovered possible directed path from A to E beginning with $A \circ\!\!-\!\!\ast B$. $\mathcal{R}_{12}$ is also a generalization of $\mathcal{R}_3$. Consider there is an unshielded triple $C \ast\!\!\to A \leftarrow\!\!\ast D$ in a PMG H: $\mathcal{R}_3$ says if there is $C \ast\!\!-\!\!\circ B \circ\!\!-\!\!\ast D$ and an edge $A \circ\!\!-\!\!\ast B$, then we orient $A \circ\!\!-\!\!\ast B$ to $A \leftarrow\!\!\ast B$. Here the reason for the transformation is, although B is not an ancestor of $\{C, D\}$ in H, B must be an ancestor of either C or D in any MAG consistent with H. That is, the vertex B here has the same property as the unbridged path we discuss above. Consider $C \ast\!\!\to A$ and $D \ast\!\!\to A$ in $\mathcal{R}_3$, if there is $A \to B$, either A, B, C or A, B, D will form a directed or almost directed cycle, which violates the ancestral property. While $\mathcal{R}_{12}$ says if there is an unbridged path p relative to $\{A, C, D\}$ and B is the vertex adjacent to A in the uncovered path from A to every vertex in p, then we orient $A \circ\!\!-\!\!\ast B$ to $A \leftarrow\!\!\ast B$. $\mathcal{R}_{13}$ generalizes the one vertex B in $\mathcal{R}_3$ to a subgraph induced by all the vertices V such that there is an uncovered possible directed path from A to V beginning with $A \circ\!\!-\!\!\ast B$.

4.2. Utilizing proposed rules in set determination

In this part, we introduce the rules proposed in Sec. 4.1 to improve PAGcauses for set determination.

We first recall Theorem 2, which provides a sufficient condition for the existence of MAGs valid to $\mathbb{M}$ such that a given set $\mathbf{W}$ is an adjustment set. The intuition of the sufficient condition is to construct a MAG $\mathcal{M}$ such that the adjustment set is $\mathbf{W}$. To ensure it, some vertices (characterized by $\bar{\mathbf{W}}$) are required to not be ancestors of $\mathbf{W} \cup \{Y\}$ in $\mathcal{M}$. Hence, it is necessary to introduce some additional arrowheads to *block* all the possible directed paths from $\bar{\mathbf{W}}$ to $\mathbf{W} \cup \{Y\}$. For this purpose, intuitively, the block set $\mathbf{S}$ in Definition 5 is introduced to characterize the position to introduce arrowheads to *block* the possibly directed paths (See Algorithm 3 for MAG construction algorithm). There are an exponential number (with respect to d) of ways to introduce arrowheads to achieve it, hence there are an exponential number of block sets in Definition 5. Thus, given $\mathbf{W}$, evaluating the existence of MAGs in Theorem 2 needs to enumerate every block set $\mathbf{S}$. Essentially, the process of enumerating $\mathbf{S}$ and evaluating the three conditions is to determine the existence of a kind of way to introduce arrowheads such that a MAG can be constructed with $\mathbf{W}$ being the adjustment set. Hence, given a potential adjustment set $\mathbf{W}$, an exponential complexity of enumerating all the block sets is involved in using Theorem 2 in Algorithm 1, because there are exponential number of possible ways to block all the possible directed paths from $\bar{\mathbf{W}}$ to $\mathbf{W} \cup \{Y\}$ such that no vertex in $\bar{\mathbf{W}}$ is an ancestor of $\mathbf{W} \cup \{Y\}$ in the constructed MAG. In the following, we show that enumerating block sets can be circumvented by utilizing $\mathcal{R}_{12}$, thus can avoid the possibly exponential computational burden here. We first define a set of vertices $\mathbf{S}_0$ in Definition 10.

Definition 10. Given a maximal local MAG $\mathbb{M}$ and a potential adjustment set $\mathbf{W}$, we define $\mathbf{S}_0$ as $V' \in \mathbf{S}_0$ if and only if there exists a vertex $V \in \bar{\mathbf{W}}$ and there exists a minimal possible directed path p from V to $\mathrm{Anc}(\mathbf{W} \cup \{Y\}, \mathbb{M})$ such that V' is the vertex adjacent to V in p and each non-endpoint in p does not belong to $\bar{\mathbf{W}}$.

It is evident that if we want to obtain a MAG $\mathcal{M}$ valid to $\mathbb{M}$ such that $\bar{\mathbf{W}}$ are not ancestors of $\mathbf{W} \cup \{Y\}$ in $\mathcal{M}$, all the edges $V \circ\!\!-\!\!\ast S, V \in \bar{\mathbf{W}}, S \in \mathbf{S}_0$ in $\mathbb{M}$ must be transformed to $V \leftarrow\!\!\ast S$. Initialize $\mathbf{S} = \mathbf{S}_0$, characterizing the positions to introduce arrowheads, similar to the block set. The transformed edges above can be seen as BK, and can possibly trigger $\mathcal{R}_{12}$ to further update the graph. The updates triggered by $\mathcal{R}_{12}$ help us enlarge the set $\mathbf{S}$.

Following the idea above, we present Algorithm 5. Note there is a premise in $\mathcal{R}_{12}$ that BK is correct, but it does not necessarily hold here. Hence, in the process of triggering $\mathcal{R}_{12}$, we need to evaluate whether the BK is possible, which is achieved on Line 4 and 9. When $\mathcal{R}_{12}$ can no longer be triggered (Line 8), roughly speaking, if there is not an unbridged path relative to $\mathbf{S}$, $\mathbf{S}$ is a block set that satisfies the three conditions of Theorem 2. Fig. 4(c) and Fig. 4(d) shows two maximal local MAGs. If $\mathbf{W} = \emptyset$, then $\bar{\mathbf{W}} = \{A\}$ for the graphs in Fig. 4(c) and Fig. 4(d). In this case, for Fig. 4(c), we obtain $\mathbf{S}_0 = \{E\}$ according to Definition 10, and obtain $\mathbf{S} = \mathbf{S}_0 = \{E\}$ by Algorithm 5. For Fig. 4(d), we obtain $\mathbf{S}_0 = \{C_1, C_2\}$. By Algorithm 5, $A \circ\!\!\to B$ will be transformed to $A \leftrightarrow B$ on Line 6, thus $\mathbf{S} = \{C_1, C_2, B\}$ is output.

Then, we present relevant theoretical guarantees for Algorithm 5. Theorem 6 implies that given a maximal local MAG $\mathbb{M}$ and a potential adjustment set $\mathbf{W}$, if we can obtain a set by Algorithm 5, then there is a MAG $\mathcal{M}$ valid to $\mathbb{M}$ such that $\mathbf{W}$ is an adjustment

Algorithm 5: Updating $\mathbf{S}$.

Input: Maximal local MAG $\mathbb{M}$, X, Y, $\mathbf{W}$
1 $\mathbf{S}$ is initialized as $\mathbf{S}_0$ in Definition 10, and $\bar{\mathbf{W}}$ is determined as Definition 5;
2 $\mathbf{T} = \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}]) \backslash \bar{\mathbf{W}}$, $H = \mathbb{M}$;
3 **while** *1* **do**
4 **if** *$H[\mathbf{S}_V]$ is not a complete graph for some $V \in \bar{\mathbf{W}}$, where $\mathbf{S}_V = \{V' \in \mathbf{S} | V \circ\!\!-\!\!* V'$ in $\mathbb{M}\}$; or* $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}]) \cap \mathrm{Pa}(\mathbf{S}, \mathbb{M}) \neq \emptyset$ **then return** "No";
5 Update H by transforming $V \circ\!\!-\!\!* S$ to $V \leftarrow\!\!* S$ for any $V \in \bar{\mathbf{W}}$ and $S \in \mathbf{S}$;
6 **if** *an edge $A \leftarrow\!\!* B$ can be transformed by $\mathcal{R}_{12}$ in H* **then**
7 $\mathbf{S} = \mathbf{S} \cup (\mathrm{Anc}(B, \mathbb{M}) \cap \mathbf{T})$
8 **else**
9 **if** *there is not an unbridged path relative to* $\mathbf{S}$ *in* $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ **then**
10 **return** $\mathbf{S}$
11 **else**
12 **return** "No"
Output: $\mathbf{S}$

set in $\mathcal{M}$. According to Theorem 6, whether Algorithm 5 outputs a set of vertices or "No" is an indicator of whether the input $\mathbf{W}$ is an adjustment set in some MAG or not. Hence, by executing Algorithm 5 for each potential adjustment set $\mathbf{W}$, we can find a set of adjustment sets for (X, Y) in all the MAGs valid to $\mathbb{M}$. We further provide Theorem 7 to indicate that if there exists a MAG $\mathcal{M}$ valid to $\mathbb{M}$ such that the causal effect is identifiable by covariate adjustment, then there is an adjustment set in $\mathcal{M}$ being a potential adjustment set such that Algorithm 5 can return a set of vertices. It ensures that via using Algorithm 5 for each potential adjustment set $\mathbf{W}$, we can estimate all possible causal effects that are identifiable by adjustment.

Theorem 6. *Given a maximal local MAG $\mathbb{M}$, for any potential adjustment set $\mathbf{W}$, if Algorithm 5 could return a set of vertices $\mathbf{S}$, then there exists a MAG $\mathcal{M}$ valid to $\mathbb{M}$ such that $\mathbf{W}$ is an adjustment set in $\mathcal{M}$.*

Theorem 7. *Given a maximal local MAG $\mathbb{M}$, suppose a MAG $\mathcal{M}$ valid to $\mathbb{M}$ such that there exists an adjustment set relative to (X, Y). Let $\mathbf{W}$ be* D-SEP$(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ *as Definition 3. Then $\mathbf{W}$ is a potential adjustment set in $\mathbb{M}$ and Algorithm 5 can return a set of vertices $\mathbf{S}$ given $\mathbf{W}$ as the input.*

Remark 5. There are at most $d(d-1)/2$ circles that could be transformed on Line 6 of Algorithm 5, where d denotes the number of vertices. And the transformation on Line 6 is a necessary condition for entering the next round of loop. Hence, the number of loop in Algorithm 5 is at most $\mathcal{O}(d^2)$. For the other parts in Algorithm 5, the complexity is at most $\mathcal{O}(d^3)$. Hence Algorithm 5 can be implemented in polynomial time.

Remark 6. One may wonder given a maximal local MAG $\mathbb{M}$ and a potential adjustment set $\mathbf{W}$, whether we could determine whether $\mathbf{W}$ can be an adjustment set in some MAG valid to $\mathbb{M}$ by judging the three conditions in Theorem 2 with a prescribed $\mathbf{S}$, such as $\mathbf{S} = \mathbf{S}_{min}$ or $\mathbf{S} = \mathbf{S}_{max}$ in Definition 5. It is infeasible. Consider $\mathbb{M}_2$ in Fig. 4(c) and $\mathbf{W} = \emptyset$, the three conditions (in Theorem 2) hold when $\mathbf{S} = \mathbf{S}_{min} = \{E\}$, but do not hold when $\mathbf{S} = \mathbf{S}_{max} = \{B, D, E\}$. While consider $\mathbb{M}'$ in Fig. 4(d) and $\mathbf{W} = \emptyset$, the three conditions do not hold when $\mathbf{S} = \mathbf{S}_{min} = \{C_1, C_2\}$, but hold when $\mathbf{S} = \{C_1, C_2, B\}$. Hence, it is not direct that which block set can satisfy the three conditions. Intuitively, the benefit taken by utilizing $\mathcal{R}_{12}$ is that it implies which vertex should be added into $\mathbf{S}$, instead of enumerating all block sets as Theorem 2. For example, when using Algorithm 5 for $\mathbb{M}_2$ in Fig. 4(c) given $\mathbf{W} = \emptyset$, there is $\mathbf{S}_0 = \{E\}$ and no unbridged paths, thus $\mathbf{S} = \{E\}$ is returned; while for $\mathbb{M}'$ in Fig. 4(d) given $\mathbf{W} = \emptyset$, there is $\mathbf{S}_0 = \{C_1, C_2\}$ and an unbridged path $D \circ\!\!-\!\!\circ E$, thus $\mathbf{S} = \{C_1, C_2, B\}$ is returned.

With the intuition above, we could introduce the rules to circumvent the enumerations of block sets in set determination. Based on the framework of Algorithm 1 and Algorithm 2, we provide *PAGrules*, short for **PAG**-consistent causal effect set with **rules**, in Algorithm 6. The only difference is, from Line 9-12, we evaluate whether each potential adjustment set is an adjustment set by Algorithm 5, instead of enumerating each block sets. Corollary 1 implies that PAGrules can return the set of causal effects that are identifiable by covariate adjustment in all the DAGs consistent with $\mathcal{P}$.

Corollary 1. For a PAG $\mathcal{P}$, denote CE$(\mathcal{P})$ and $\widehat{\mathrm{CE}}(\mathcal{P})$ the set of causal effects in the DAGs represented by the MAGs consistent with $\mathcal{P}$ that can be estimated by covariate adjustment with observable variables and the set of causal effects output by Algorithm 6. It holds that $\mathrm{CE}(\mathcal{P}) \overset{set}{=} \widehat{\mathrm{CE}}(\mathcal{P})$.

4.3. Proof

4.3.1. Proof of Theorem 5

Proof. We first prove the soundness of $\mathcal{R}_{13}$, then prove the soundness of $\mathcal{R}_{12}$.

Algorithm 6: PAGrules.

Input: PAG $\mathcal{P}$, X, Y
$\widehat{\mathrm{AS}}(\mathcal{P}) = \emptyset$ // Record all the valid adjustment sets;
if $X \notin \mathrm{PossAn}(Y, \mathcal{P})$ **then return** No causal effects;
if *the conditions in Proposition 1 are satisfied for* $\mathcal{P}$ **then**
  return $\widehat{\mathrm{AS}}(\mathcal{P}) \leftarrow \{\text{D-SEP}(X, Y, \mathcal{P}_{\underline{X}})\}$ // Prop. 1
for *each set* $\mathbf{C} \subseteq \{V \mid V \ast\!\!-\!\!\circ X \text{ in } \mathcal{P}\}$ **do**
  if *the three conditions in Proposition 2 are satisfied* **then**
    Obtain a maximal local MAG $\mathbb{M}$ based on $\mathcal{P}$ and $\mathbf{C}$;
    Find all potential adjustment sets $\mathbf{W}_1, \mathbf{W}_2, \cdots$, that contains DD-SEP$(X, Y, \mathbb{M})$ given $\mathbb{M}$;
    for *each potential adjustment set* $\mathbf{W}_i$ **do**
      Obtain $\mathbf{S}_0$ and $\bar{\mathbf{W}}$ as Definition 10 and Definition 5;
      if *Alg. 5 can return a set of vertices given* $\mathbf{W}_i$ *and* $\mathbb{M}$ **then**
        $\widehat{\mathrm{AS}}(\mathcal{P}) \leftarrow \widehat{\mathrm{AS}}(\mathcal{P}) \cup \{\mathbf{W}_i\}$;
Output: $\widehat{\mathrm{AS}}(\mathcal{P})$ // Adjustment sets in MAGs consistent with $\mathcal{P}$

For $\mathcal{R}_{13}$, suppose there is an MAG $\mathcal{M}$ with $A \rightarrow B$. It is evident that the uncovered path is $A \rightarrow B \rightarrow \cdots \rightarrow K$. According to the definition of $\mathbf{S}_A$ and the conditions in $\mathcal{R}_{13}$, there must be a vertex $C \in \mathbf{S}_A$ such that there is $C \ast\!\!\rightarrow A \rightarrow \cdots K \rightarrow \cdots C$, which contradicts with the ancestral property, contradiction.

For $\mathcal{R}_{12}$, suppose there is an MAG $\mathcal{M}$ with $A \rightarrow B$. According to the condition of $\mathcal{R}_{12}$, it is evident that for each vertex $K_i, 1 \le i \le m$, there is a minimal directed path $A \rightarrow B \rightarrow \cdots \rightarrow K_i$ in $\mathcal{M}$. And for any $T \in \mathbf{S}_A \backslash \{A\}$ and $K_i, 1 \le i \le m$ such that there is an edge between T and K_i in the PMG H, there must be $T \ast\!\!\rightarrow K_i$ in $\mathcal{M}$, for otherwise there is $A \rightarrow B \rightarrow \cdots \rightarrow K_i \rightarrow T \ast\!\!\rightarrow A$, contradicting with ancestral property. According to Definition 9, there is an uncovered path $p : K_1 \circ\!\!-\!\!\circ K_2 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ K_m$ such that $\mathcal{F}_{K_1} \backslash \mathcal{F}_{K_2} \neq \emptyset$ and $\mathcal{F}_{K_m} \backslash \mathcal{F}_{K_{m-1}} \neq \emptyset$.

Suppose $C_1 \in \mathcal{F}_{K_1} \backslash \mathcal{F}_{K_2}$ and $C_2 \in \mathcal{F}_{K_m} \backslash \mathcal{F}_{K_{m-1}}$. If C_1 is adjacent to K_2, as $C_1 \notin \mathcal{F}_{K_2}$, there is $K_2 \rightarrow C_1$ in H. In this case there must be $A \leftarrow\!\!\ast B$ according to $\mathcal{R}_{13}$, the soundness of which has been proven, thus $A \rightarrow B$ is impossible. In the following, we consider the case that C_1 is not adjacent to K_2 and C_2 is not adjacent to K_{m-1}. We have shown before that if there is an edge $A \rightarrow B$ in $\mathcal{M}$, there must be $C_1 \ast\!\!\rightarrow K_1$ and $C_2 \ast\!\!\rightarrow K_m$ in $\mathcal{M}$. In this case, no matter how we transform the circles in the path $C_1 \ast\!\!\rightarrow K_1 \circ\!\!-\!\!\circ K_2 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ K_m \leftarrow\!\!\ast C_2$, there will be a new unshielded collider in $\mathcal{M}$, which contradicts with the fact that $\mathcal{M}$ is consistent with H. □

4.3.2. Proof of Proposition 5

Proof. For $\mathcal{R}_{13}$, the implementation on Line 6 of Algorithm 4 as well as the conditions on Line 1 - Line 3 totally follow the conditions in $\mathcal{R}_{13}$. Hence it evidently that for the edges that can be transformed by $\mathcal{R}_{13}$, Algorithm 4 can transform all these edges on Line 5. It suffices to show that except for the edges by $\mathcal{R}_{13}$, Algorithm 4 can transform and only transform the edges triggered by $\mathcal{R}_{12}$ in the following.

We first prove that if Algorithm 4 transforms an edge $A \circ\!\!-\!\!\ast B$ to $A \leftarrow\!\!\ast B$ on Line 12, then $\mathcal{R}_{12}$ can be triggered to transform $A \circ\!\!-\!\!\ast B$ to $A \leftarrow\!\!\ast B$. Suppose there is an edge $A \circ\!\!-\!\!\ast B$ to $A \leftarrow\!\!\ast B$ on Line 12 which cannot be triggered by $\mathcal{R}_{12}$. Without loss of generality, suppose in the k-th transformation of Algorithm 4, there is the first edge $A \circ\!\!-\!\!\ast B$ transformed to $A \leftarrow\!\!\ast B$ on Line 12 which cannot be triggered by $\mathcal{R}_{12}$. Since the rigorous proof is somewhat tedious, we just show a proof sketch here. If Algorithm 4 transforms an edge $A \circ\!\!-\!\!\ast B$ to $A \leftarrow\!\!\ast B$ on Line 12, then there is an unshielded collider formed in H' on Line 11. In this case, according to the proof idea of Lemma 5, we can prove that $H'[\mathbf{D}]$ is not bridged relative to $\mathbf{S}_A$. Then, we will prove that there is an unbridged path in $H'[\mathbf{D}]$ relative to $\mathbf{S}_A$.

Suppose there is a minimal path $p = V_0 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ V_{n+1}$ in $H'[\mathbf{D}]$ where there is not a vertex V_s such that $\mathcal{F}_{V_i} \subseteq \mathcal{F}_{V_{i+1}}, 0 \le i \le s-1$ and $\mathcal{F}_{V_{i+1}} \subseteq \mathcal{F}_{V_i}, s \le i \le n$. As Line 6 is not triggered, no vertex in $\mathbf{D}$ belongs to $\mathrm{Anc}(\mathbf{S}_A, H')$. Hence if there is some vertex $V \notin \mathcal{F}_{V_i}$, V is not adjacent to V_i. Next, we consider the path p. We discuss whether $\mathcal{F}_{V_0} \backslash \mathcal{F}_{V_1} = \emptyset$. If empty, we consider whether $\mathcal{F}_{V_1} \backslash \mathcal{F}_{V_2} = \emptyset$ instead. We repeat the process above until we find the first index j such that $\mathcal{F}_{V_j} \backslash \mathcal{F}_{V_{j+1}} \neq \emptyset$. Note such j must exist, for otherwise, there is $\mathcal{F}_{V_0} \subseteq \mathcal{F}_{V_1} \subseteq \mathcal{F}_{V_2} \subseteq \cdots \subseteq \mathcal{F}_{V_{n+1}}$, in which case there is $s = n+1$ such that $\mathcal{F}_{V_i} \subseteq \mathcal{F}_{V_{i+1}}, 0 \le i \le n$, contradiction. Then, we consider the sub-path $V_j \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ V_{n+1}$. Note according to the process above, there is $\mathcal{F}_{V_0} \subseteq \mathcal{F}_{V_1} \subseteq \mathcal{F}_{V_2} \subseteq \cdots \subseteq \mathcal{F}_{V_j}$ and $\mathcal{F}_{V_j} \backslash \mathcal{F}_{V_{j+1}} \neq \emptyset$. Then, we consider whether $\mathcal{F}_{V_{n+1}} \backslash \mathcal{F}_{V_n} = \emptyset$. If empty, we consider whether $\mathcal{F}_{V_n} \backslash \mathcal{F}_{V_{n-1}} = \emptyset$ instead. We repeat the process above until we find the first index k such that $\mathcal{F}_{V_k} \backslash \mathcal{F}_{V_{k-1}} \neq \emptyset$. Similar to the proof above, such index k must exist. And there is $\mathcal{F}_{V_k} \supseteq \mathcal{F}_{V_{k+1}} \supseteq \cdots \supseteq \mathcal{F}_{V_{n+1}}$ and $\mathcal{F}_{V_{k-1}} \neq \emptyset$. Next we discuss the relationship between j and k. We will prove the impossibility of $k \le j$. Suppose $k \le j$. According to the result above, there is $\mathcal{F}_{V_0} \subseteq \mathcal{F}_{V_1} \subseteq \mathcal{F}_{V_2} \subseteq \cdots \subseteq \mathcal{F}_{V_j}$ and $\mathcal{F}_{V_k} \supseteq \mathcal{F}_{V_{k+1}} \supseteq \cdots \supseteq \mathcal{F}_{V_{n+1}}$. Hence, there is $\mathcal{F}_{V_0} \subseteq \mathcal{F}_{V_1} \subseteq \mathcal{F}_{V_2} \subseteq \cdots \subseteq \mathcal{F}_{V_k} = \mathcal{F}_{V_{k+1}} = \cdots = \mathcal{F}_{V_j} \supseteq \cdots \supseteq \mathcal{F}_{V_{n+1}}$. In this case, there is $s = j$ such that $\mathcal{F}_{V_i} \subseteq \mathcal{F}_{V_{i+1}}, 0 \le i \le s-1$ and $\mathcal{F}_{V_{i+1}} \subseteq \mathcal{F}_{V_i}, s \le i \le n$, contradiction. Hence there is $k \ge j+1$. In this case, there is a minimal path $V_j \circ\!\!-\!\!\circ V_{j+1} \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ V_k$ such that $\mathcal{F}_{V_j} \backslash \mathcal{F}_{V_{j+1}} \neq \emptyset$ and $\mathcal{F}_{V_k} \backslash \mathcal{F}_{V_{k-1}} \neq \emptyset$ in $H'[\mathbf{D}]$. It is an unbridged path relative to $\mathbf{S}_A$ in $H'[\mathbf{D}]$. And since there are uncovered possible directed paths from A to any vertice in $\mathbf{D}$ and at the same time B is the vertex adjacent to A in the paths, $\mathcal{R}_{12}$ can be triggered to transform $A \circ\!\!-\!\!\ast B$ to $A \leftarrow\!\!\ast B$ in H'.

Then we prove that if there is only one edge $A \circ\!\!-\!\!\ast B$ that could be transformed to $A \leftarrow\!\!\ast B$ by $\mathcal{R}_{12}$ in a graph H', then Line 12 of Algorithm 4 will transform this edge. Suppose there is an unbridged path $V_0 \circ\!\!-\!\!\circ V_1 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ V_n$ relative to $\mathbf{S}_A$ in H'. Then there is $\mathcal{F}_{V_0} \backslash \mathcal{F}_{V_1} \neq \emptyset$ and $\mathcal{F}_{V_n} \backslash \mathcal{F}_{V_{n-1}} \neq \emptyset$. Suppose $S_1 \in \mathcal{F}_{V_0} \backslash \mathcal{F}_{V_1}$ and $S_2 \in \mathcal{F}_{V_n} \backslash \mathcal{F}_{V_{n-1}}$. According to Line 9 of Algorithm 4, there is $S_1 \ast\!\!\rightarrow V_0$ and $V_n \leftarrow\!\!\ast S_2$ in H'. Note S_1 cannot be adjacent to V_1, for otherwise there can only be $V_1 \rightarrow S_1$ in H, for otherwise there is $S_1 \in \mathcal{F}_{V_1}$. However, in this case Line 6 of Algorithm 4 has been executed such that $A \leftrightarrow\!\!* B$ in H', contradiction. Similarly, S_2 is not adjacent to V_{n-1}. Hence there is an uncovered path $S_1 *\!\!\rightarrow V_0 \circ\!\!-\!\!\circ V_1 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ V_n \leftarrow\!\!* S_2$ in H'. According to the update of Line 10 of Algorithm 4, there must be an unshielded collider in this path after the update. Thus Line 12 will be implemented and $A \circ\!\!-\!\!* B$ can be transformed by $A \leftarrow\!\!* B$. The proof completes. □

4.3.3. Proof of Theorem 6

Proof of Theorem 6. Since the algorithm returns a set of vertices $\mathbf{S}$ after T rounds, according to Line 4, 6, and 9 of Algorithm 5, we conclude that (1) $\mathbb{M}[\mathbf{S}_V]$ is a complete graph for any $V \in \bar{\mathbf{W}}$, where $\mathbf{S}_V = \{V' \in \mathbf{S} | V \circ\!\!-\!\!* V' \text{ in } \mathbb{M}\}$, (2) $\text{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}]) \cap \text{Pa}(\mathbf{S}, \mathbb{M}) = \emptyset$, (3) there is no unbridged path relative to $\mathbf{S}$ in $\mathbb{M}[\text{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$. According to (3) above and Definition 4, $\mathbb{M}[\text{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}])]$ is bridged relative to $\mathbf{S}$ in $\mathbb{M}$. Thus the three conditions in Theorem 2 are satisfied given the set $\mathbf{S}$. We can conclude the desired result by Theorem 2. □

4.3.4. Proof of Theorem 7

We first present some supporting results.

Lemma 11. *Given a maximal local MAG $\mathbb{M}$, suppose a MAG $\mathcal{M}$ valid to $\mathbb{M}$ such that there exists an adjustment set relative to (X, Y). Let $\mathbf{W}$ be $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$. Suppose there is a minimal possible directed path $p = \langle J_0(= V), J_1, \cdots, J_s(= T)\rangle$ from $V \in \bar{\mathbf{W}}$ to a vertex T in $\mathbb{M}$, where each non-endpoint in p does not belong to $\mathbf{W} \cup \bar{\mathbf{W}}$. If $T \in \text{Anc}(Y, \mathcal{M})$, then p can only be as $J_0 \circ\!\!\rightarrow J_1 \rightarrow \cdots \rightarrow J_s$ in $\mathbb{M}$. And there exists a collider path $X(= F_0) \leftrightarrow F_1 \leftrightarrow \cdots \leftrightarrow F_{t-1} \leftarrow\!\!* V$ in $\mathbb{M}$ with edges $F_i \rightarrow J_1, 0 \le i \le n-1$.*

Proof. According to Definition 5, there exists a collider path $X(= F_0) \leftrightarrow F_1 \leftrightarrow \cdots \leftrightarrow F_{t-1} \leftarrow\!\!* V$ in $\mathbb{M}$, where $F_1, \cdots, F_{t-1} \in \mathbf{W}$. There cannot be an edge $F_i \leftarrow\!\!* J_1$ in $\mathbb{M}$ for any $0 \le i \le t-1$, for otherwise $J_1 \in \mathbf{W} \cup \bar{\mathbf{W}}$.

Since V is not an ancestor of Y in $\mathcal{M}$ and p is a minimal possible directed path, there must be $F_{t-1} \leftrightarrow V \leftarrow\!\!* J_1$ in $\mathcal{M}$. Hence F_{t-1} is adjacent to J_1, for otherwise there is a new unshielded collider in $\mathcal{M}$ relative to $\mathbb{M}$. Since (1) for each $F_i, 0 \le i \le t-1$, there cannot be $F_i \leftarrow\!\!* J_1$ in $\mathbb{M}$, and (2) the balanced property is fulfilled in $\mathbb{M}$, we can conclude that there is $F_i \rightarrow J_1$ or $F_i \circ\!\!\rightarrow J_1, \forall 1 \le i \le n-1$ and $X \rightarrow J_1$, otherwise there is always a discriminating path for V which leads to a non-circle mark at V on the edge between V and J_1 in $\mathcal{P}$. There is either $F_{n-1} \rightarrow J_1$ or $F_{n-1} \circ\!\!\rightarrow J_1$. Due to the balanced property of $\mathbb{M}$ and $\mathcal{R}_2$ of Zhang [6], there is always $V \circ\!\!\rightarrow J_1$. Since the path p is a minimal possible directed path, the path can only be as $V \circ\!\!\rightarrow J_1 \rightarrow \cdots \rightarrow J_s$. □

Lemma 12. *Given a maximal local MAG $\mathbb{M}$, suppose a MAG $\mathcal{M}$ valid to $\mathbb{M}$ such that there exists an adjustment set relative to (X, Y). Let $\mathbf{W}$ be $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$. For any S incorporated into the set of vertices $\mathbf{S}$ in the process of Algorithm 5 (on Line 7), there is $S \in \text{Anc}(Y, \mathcal{M})$.*

Proof. For $S \in \mathbf{W}$, since $\mathbf{W} = \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$, there is $S \in \text{Anc}(\mathbf{W} \cup \{Y\}, \mathbb{M})$ and $S \in \text{Anc}(Y, \mathcal{M})$. For $S \in \mathbf{S}_0$ defined in Definition 10, according to Lemma 11, since S is the vertex adjacent to a vertex $V \in \bar{\mathbf{W}}$ in a minimal possible directed path from V to a vertex in $\text{Anc}(Y, \mathbb{M})$, there is $S \in \text{Anc}(Y, \mathbb{M})$. Suppose there are T rounds in Algorithm 5, where in the i-th round, $1 \le i \le T-1$, an edge $A_i \leftarrow\!\!* S_i$ is transformed by $\mathcal{R}_{12}$ on Line 6 of Algorithm 5, and thus $\text{Anc}(S_i, \mathbb{M}) \cap \mathbf{T}$ is incorporated to $\mathbf{S}$ on Line 7 of Algorithm 5. For brevity, denote $\mathbf{T} = \text{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}]) \setminus \bar{\mathbf{W}})$ like Line 2 of Algorithm 5.

We first prove $S_1 \in \text{Anc}(Y, \mathcal{M})$. Since $\mathcal{R}_{12}$ is triggered, there is an unbridged path $p = \langle K_1, \cdots, K_m\rangle$ relative to $\mathbf{S}_0$ and there exists an uncovered possible directed path $\langle V, S_1, \cdots, K_j\rangle$ for $\forall 1 \le j \le m$ in $\mathbb{M}$, where $V \in \bar{\mathbf{W}}$. Without loss of generality, suppose p is an unbridged path. In the unbridged path, there is $C *\!\!-\!\!\circ K_1 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ K_m \circ\!\!-\!\!* D$ in $\mathbb{M}$, where $C \in \mathcal{F}_{K_1} \setminus \mathcal{F}_{K_2}$, $D \in \mathcal{F}_{K_m} \setminus \mathcal{F}_{K_{m-1}}$, $\mathcal{F}_{K_i} = \{V \in \mathbf{V}' \mid V *\!\!-\!\!\circ K_i \text{ in } \mathbb{M}\}$. Next we prove there is not $K_1 \rightarrow C$ in $\mathbb{M}$. Suppose $K_1 \rightarrow C$ in $\mathbb{M}$ for contradiction. According to Lemma 11, $C \in \mathbf{S}_0 \subseteq \text{Anc}(\mathbf{W} \cup \{Y\}, \mathbb{M}) \subseteq \text{Anc}(Y, \mathcal{M})$. And since there is a directed path $S_1 \rightarrow \cdots \rightarrow K_1 \rightarrow C$ in $\mathbb{M}$ and $C \in \mathbf{S}_0 \subseteq \text{Anc}(\mathbf{W} \cup \{Y\}, \mathbb{M})$, $S_1 \in \text{Anc}(\mathbf{W} \cup \{Y\}, \mathbb{M})$. Since $V \in \bar{\mathbf{W}}$ is adjacent to S_1, S_1 should belong to $\mathbf{S}_0$, contradiction. Hence there cannot be an edge $K_1 \rightarrow C$ in $\mathbb{M}$.

Hence, in $\mathcal{M}$ valid to $\mathbb{M}$, for each vertex $F_j, 1 \le j \le m$, it is easy to prove that F_j is an ancestor of either C or D, for otherwise there will be an unshielded collider in the path $\langle C, K_1, \cdots, K_m, D\rangle$ in $\mathcal{M}$. And since $C, D \in \mathbf{S}_0$, it holds $C, D \in \text{Anc}(Y, \mathbb{M})$. And since there is $V \circ\!\!\rightarrow S_1 \rightarrow \cdots \rightarrow K_j$ in $\mathbb{M}$, S_1 is an ancestor of Y in $\mathcal{M}$. All the vertices in $\text{Anc}(S_1, \mathbb{M}) \cap \mathbf{T}$ are evidently ancestors of Y in $\mathcal{M}$.

Next we prove the induction result. Suppose in the first i round, each vertex in $\mathbf{S}_0, \text{Anc}(S_1, \mathbb{M}), \cdots, \text{Anc}(S_i, \mathbb{M})$ is an ancestor of Y in $\mathcal{M}$. We will prove $S_{i+1} \in \text{Anc}(Y, \mathcal{M})$.

Since $\mathcal{R}_{12}$ is triggered, there is an unbridged path $p = \langle T_1, \cdots, T_f\rangle$ relative to $\mathbf{S}_0 \cup \bigcup_{1 \le q \le i}(\text{Anc}(S_q, \mathbb{M}) \cap \mathbf{T})$ and there exists an uncovered possible directed path $\langle V, S_{i+1}, \cdots, T_s\rangle$ for $1 \le s \le f$ in $\mathbb{M}$ where $V \in \bar{\mathbf{W}}$. Without loss of generality, suppose p is an unbridged path. In the unbridged path, there is $J *\!\!-\!\!\circ T_1 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ T_f \circ\!\!-\!\!* K$ in $\mathbb{M}$, where $J \in \mathcal{F}_{T_1} \setminus \mathcal{F}_{T_2}$, $K \in \mathcal{F}_{T_f} \setminus \mathcal{F}_{T_{f-1}}$. Next we prove there is not $T_1 \rightarrow J$ in $\mathbb{M}$. Suppose $T_1 \rightarrow J$ in $\mathbb{M}$ for contradiction. Note in the whole process of Algorithm 5, we never add a tail, hence there is $T_1 \rightarrow J$ in $\mathbb{M}$. In this case, there is a directed path $S_{i+1} \rightarrow \cdots \rightarrow T_1 \rightarrow J$ in $\mathbb{M}$ and $J \in \mathbf{S}_0 \cup \bigcup_{1 \le q \le i}(\text{Anc}(S_q, \mathbb{M}) \cap \mathbf{T})$. If $J \in \mathbf{S}_0$, then $S_{i+1} \in \text{Anc}(\mathbf{W} \cup \{Y\}, \mathbb{M})$, S_{i+1} should belong to $\mathbf{S}_0$, contradiction. If $S_{i+1} \in \text{Anc}(S_q, \mathbb{M}) \cap \mathbf{T}, 1 \le q \le i$, then S_{i+1} should have been incorporated into $\mathbf{S}$ in the q-round of Algorithm 5, contradiction. Hence there cannot be an edge $T_1 \rightarrow J$ in $\mathbb{M}$.

Hence, in $\mathcal{M}$ valid to $\mathbb{M}$, for each vertex $T_s, 1 \le s \le f$, it is an ancestor of either J or K, for otherwise there will be an unshielded collider in the path $\langle J, T_1, \cdots, T_f, K\rangle$ in $\mathcal{M}$. And since $J, K \in \mathbf{S}_0 \cup \bigcup_{1 \le q \le i}(\text{Anc}(S_q, \mathbb{M}) \cap \mathbf{T})$, and J, K are ancestors of Y in $\mathcal{M}$, it holds

$J, K \in \mathrm{Anc}(Y, \mathcal{M})$. And since there is $V \circ\!\!\rightarrow S_{i+1} \rightarrow \cdots \rightarrow T_s$ in $\mathbb{M}$, S_{i+1} is an ancestor of Y. Thus $S_{i+1} \in \mathrm{Anc}(Y, \mathcal{M})$. Hence all the vertices in $\mathrm{Anc}(S_{i+1}, \mathbb{M}) \cap \mathbf{T}$ are ancestors of Y in $\mathcal{M}$.

By induction, we can prove that all the incorporated vertices in $\mathbf{S}$ are ancestors of Y in $\mathcal{M}$. □

Theorem. 7. Theorem 3 has implied that $\mathbf{W} = \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ is a potential adjustment set. We will prove that a set of vertices $\mathbf{S}$ will be returned by Algorithm 5.

For $V \in \bar{\mathbf{W}}$ and $S \in \mathbf{S}_0$ in Definition 10, there must be $V \leftarrow\!\!* S$ in $\mathcal{M}$ if there is $V \circ\!\!-\!\!* S$ in $\mathbb{M}$. Hence the arrowheads introduced in the first round of Algorithm 5 must exist in $\mathcal{M}$. And due to the soundness of $\mathcal{R}_{12}$ by Theorem 5, all the arrowheads introduced in Algorithm 5 exist in $\mathcal{M}$.

Denote $\mathbf{T} = \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}]) \backslash \bar{\mathbf{W}}$ like Line 2 of Algorithm 5. Suppose there are J rounds in Algorithm 5, where in the i-th round, $1 \leq i \leq J-1$, an edge $A_i \leftarrow\!\!* S_i$ is transformed by $\mathcal{R}_{12}$ on Line 6 of Algorithm 5, and thus $\mathrm{Anc}(S_i, \mathbb{M}) \cap \mathbf{T}$ is incorporated to $\mathbf{S}$ on Line 7 of Algorithm 5. Hence there is evidently $\mathbf{S}_{i+1} = \mathbf{S}_i \cup (\mathrm{Anc}(S_{i+1}, \mathbb{M}) \cap \mathbf{T})$, for $0 \leq i \leq J-1$.

It suffices to show that in the i-th round $1 \leq i \leq J$, there is (1) $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}_i]) \cap \mathrm{Pa}(\mathbf{S}_i, \mathbb{M}) = \emptyset$, (2) $\mathbb{M}[\mathbf{S}_V]$ is a complete graph for any $V \in \bar{\mathbf{W}}$, where $\mathbf{S}_V = \{V' \in \mathbf{S}_i | V \circ\!\!-\!\!* V' \text{ in } \mathbb{M}\}$, and (3) there is not an unbridged path relative to $\mathbf{S}_i$ in $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}_i])]$. As when these three conditions are satisfied in each round, Algorithm 5 could output a set of vertices. Suppose in the i round, the algorithm output "No". According to Algorithm 5, at least one of the three conditions is violated. We will prove the impossibility of the violations of the three conditions in the following.

If $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}_i]) \cap \mathrm{Pa}(\mathbf{S}_i, \mathbb{M}) \neq T$, suppose there is a minimal possible directed path p from $V \in \bar{\mathbf{W}}$ to T in $\mathbb{M}[-\mathbf{S}_i]$ such that each non-endpoint does not belong to $\bar{\mathbf{W}}$, and there is an edge $T \rightarrow S$ in $\mathbb{M}$ for $S \in \mathbf{S}_i$. Hence $T \in \mathrm{Anc}(S, \mathbb{M})$. According to Lemma 11, p is as $V \circ\!\!\rightarrow\rightarrow \cdots \rightarrow T$ in $\mathbb{M}$. In this case, if $S \in \mathbf{S}_0$, according to Definition 10 and Lemma 11, there is $S \in \mathrm{Anc}(\mathbf{W} \cup \{Y\}, \mathbb{M})$. Thus $T \in \mathrm{Anc}(\mathbf{W} \cup \{Y\}, \mathbb{M})$, which implies that there is a minimal possible directed path p' from $V \in \bar{\mathbf{W}}$ to $T \in \mathrm{Anc}(\mathbf{W} \cup \{Y\}, \mathbb{M})$ such that p' is a sub-path of p. However, according to Definition 10, in the case above, the vertex adjacent to V in p' should belong to $\mathbf{S}_0$, contradicting with the fact that the path p is in $\mathbb{M}[-\mathbf{S}_i]$. If S is incorporated into $\mathbf{S}$ in Algorithm 5 in the $j, j < i$ round, since there is $T \in \mathrm{Anc}(S, \mathbb{M})$, T should belong to $\mathbf{S}_{j+1}, \mathbf{S}_{j+2}, \cdots, \mathbf{S}_i$, contradicting with the fact that the path p is in $\mathbb{M}[-\mathbf{S}_i]$. Hence there is always a contradiction if there is $\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}_i]) \cap \mathrm{Pa}(\mathbf{S}_i, \mathbb{M}) \neq \emptyset$.

Since $\mathbf{W} = \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$, for any $V \in \bar{\mathbf{W}}$, V is not an ancestor of Y in $\mathcal{M}$. If $\mathbb{M}[\mathbf{S}_V]$ is not a complete graph for some $V \in \bar{\mathbf{W}}$, there must be an edge $V \rightarrow S$ in $\mathcal{M}$, for otherwise there will be new unshielded collider at V. Due to Lemma 12, S is an ancestor of Y. Thus V is an ancestor of Y, thus $V \in \mathbf{W} \cap \bar{\mathbf{W}}$, contradicting with $\mathbf{W} \cap \bar{\mathbf{W}} = \emptyset$.

Finally, we prove that there is not an unbridged path relative to $\mathbf{S}_i$ in $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}_i])]$. Suppose $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}_i])]$ is not bridged relative to $\mathbf{S}_i$ in $\mathbb{M}$. Since $\mathbb{M}[\mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}_i])]$ is not bridged relative to $\mathbf{S}_i$ in $\mathbb{M}$, without loss of generality, suppose an unbridged path $K_1 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ K_m$ relative to $\mathbf{S}_i$, and there is $A \in \bar{\mathbf{W}}$. According to Lemma 12, all the vertices in $\mathbf{S}_i$ are ancestors of Y in $\mathcal{M}$. At first, we prove $K_1, \cdots, K_m$ are ancestors of Y in $\mathcal{M}$. As there are vertices $S_1, S_2 \in \mathbf{S}_i$ such that $S_1 \in \mathcal{F}_{K_1} \backslash \mathcal{F}_{K_2}$ and $S_2 \in \mathcal{F}_{K_m} \backslash \mathcal{F}_{K_{m-1}}$, where $\mathcal{F}_V = \{V' \in \mathbf{S}_i | V \circ\!\!-\!\!* V' \text{ in } \mathbb{M}\}$ as Definition 9.

Here, we prove that for any $K_j, 1 \leq j \leq m$ and $S \in \mathbf{S}_i$ in the process of Algorithm 5, there cannot be an edge $S \leftarrow K_j$. Suppose there is such an edge. If $S \in \mathbf{S}_0$, according to Definition 10, there is a minimal possible directed path $p = \langle V', S, \cdots, T \rangle$ from $V' \in \bar{\mathbf{W}}$ to $T \in \mathrm{Anc}(\mathbf{W} \cup \{Y\}, \mathbb{M})$ where each non-endpoint does not belong to $\bar{\mathbf{W}}$ in $\mathbb{M}$. According to Definition 5, there exists a collider path $X \leftrightarrow F_1 \leftrightarrow \cdots \leftrightarrow F_{t-1} \leftarrow\!\!* V'$ where $F_1, \cdots, F_{t-1} \in \mathbf{W}$. And according to the result (2) above, $\mathbb{M}[\mathbf{S}_{V'}]$ is a complete graph. Hence F_{t-1} is adjacent to S. In this case if there is an edge $F_{t-1} \leftarrow\!\!* S$ in $\mathbb{M}$, there is $S \in \mathbf{W} \cup \bar{\mathbf{W}}$. And due to $S \notin \bar{\mathbf{W}}$, there is $S \in \mathbf{W}$. And if there is an edge $F_{t-1} \circ\!\!-\!\!* S$ or $F_{t-1} \rightarrow S$ in $\mathbb{M}$, there is $F_{t-1} \circ\!\!-\!\!* S$ or $F_{t-1} \rightarrow S$ in $\mathbb{M}$, there must be $V' \circ\!\!\rightarrow S$ in $\mathbb{M}$. Since p is a minimal possible directed path in $\mathbb{M}$, p is as $V' \circ\!\!\rightarrow S \rightarrow \cdots \rightarrow T$ in $\mathbb{M}$, thus $S \in \mathrm{Anc}(\mathbf{W} \cup \{Y\}, \mathbb{M})$. Hence, no matter what the edge is between S and F_{t-1}, there is $S \in \mathrm{Anc}(\mathbf{W} \cup \{Y\}, \mathbb{M})$. In this case, if there is $K_j \rightarrow S$ in $\mathbb{M}$, there is $K_j \in \mathbf{S}_0$ due to $S \in \mathbf{S}_0$, contradicting with $K_j \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}_i])$. If S is incorporated into $\mathbf{S}$ in the $j, j < i$ round on Line 7 of Algorithm 5, K is also incorporated into $\mathbf{S}$ in this round due to $K \in \mathrm{Anc}(S, \mathbb{M})$, contradicting with $K_j \in \mathrm{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}_i])$. Hence for any $K_j, 1 \leq j \leq m$ and $S \in \mathbf{S}_i$ in the process of Algorithm 5, there cannot be an edge $S \leftarrow K_j$.

Hence, consider the uncovered path $p_1 = \langle S_1, K_1, K_2, \cdots, K_m, S_2 \rangle$ in $\mathbb{M}$ where the sub-path from K_1 to K_m is a circle path. Note the non-circle marks in $\mathbb{M}$ also exist in $\mathcal{M}$ due to the soundness of $\mathcal{R}_{12}$ according to Theorem 5. Since there cannot be new unshielded colliders in $\mathcal{M}$ relative to $\mathbb{M}_i$, and each vertex in $K_1, K_2, \cdots, K_m$ is an ancestor of either S_1 or S_2 in $\mathcal{M}$. Since S_1 and S_2 are ancestors of Y according to Lemma 12, any vertex in $K_1, \cdots, K_m$ are ancestors of Y in $\mathcal{M}$. And since $S_1 \in \mathcal{F}_{K_1} \backslash \mathcal{F}_{K_2}$, $S_2 \in \mathcal{F}_{K_m} \backslash \mathcal{F}_{K_{m-1}}$, the complete property of $\mathbb{M}$ according to Lemma 1, and for any $K_j, 1 \leq j \leq m$ and $S \in \mathbf{S}_i$ in the process of Algorithm 5, there cannot be an edge $S \leftarrow K_j$, the uncovered path p_1 in $\mathbb{M}$ is in the form of $S_1 *\!\!-\!\!\circ K_1 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ K_m \circ\!\!-\!\!* S_2$, that is, there cannot be $S_1 *\!\!\rightarrow K_1$ or $K_m \leftarrow\!\!* S_2$ in $\mathbb{M}$.

Next, for any $K_j, 1 \leq j \leq m$, consider the minimal possible directed path $p = \langle A, B_j, \cdots, K_j \rangle$ from A to K_j. Note we use notation B_j to denote the vertex adjacent to A in the minimal possible directed path from A to $K_j, 1 \leq j \leq m$. Without loss of generality, we suppose each non-endpoint in p does not belong to $\bar{\mathbf{W}}$, since if there is another vertex $A' \in \bar{\mathbf{W}}$ in p, we can consider A' instead of A, it is evidently that $K_1, \cdots, K_m$ are possible descendants of A' as well since there is a minimal possible directed path from A' to K_j and there are circle paths from K_j to each vertex in $K_1, \cdots, K_m$.

Note it is possible that there are many minimal possible directed paths from A to K_j. Next, we prove that for any $B_k, B_j, 1 \leq k < j \leq m$, there is either (1) B_k and B_j denote the same vertex, or (2) B_k is adjacent to B_j and B_k is adjacent to each vertex in $\mathbf{S}$ in $\mathbb{M}$. Suppose B_k is not adjacent to B_j or $S \in \mathbf{S}$ in $\mathbb{M}$. Since S is an ancestor of Y in any MAG $\mathcal{M}$ valid to $\mathbb{M}$ such that $\mathbf{W} = \text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$, and $A \in \bar{\mathbf{W}}$, there is $A \leftarrow\!\!* S$ in $\mathcal{M}$. Since there is not $A \leftarrow\!\!* B_j$ or $A \leftarrow\!\!* B_k$ in $\mathbb{M}$, to prevent generating an unshielded collider at A in $\mathcal{M}$, there is either $A \to B_j$ or $A \to B_k$ in $\mathcal{M}$. And since B_j and B_k are located at minimal possible directed paths from A to K_j and K_k, respectively, there must be $A \in \text{Anc}(K_k, \mathcal{M})$ or $A \in \text{Anc}(K_j, \mathcal{M})$. Since we have proven that K_k, K_j are ancestors of Y in $\mathcal{M}$, A is an ancestors of Y in $\mathcal{M}$, in which case there is $A \in \mathbf{W} \cap \bar{\mathbf{W}}$, contradicting with $\mathbf{W} \cap \bar{\mathbf{W}} = \emptyset$ in Definition 5. Similarly, we could conclude that B_k is adjacent to each vertex in $\mathbf{S}$. Hence, for any $B_k, B_j, 1 \le k < j \le m$, there is either (1) B_k and B_j denote the same vertex, or (2) B_k is adjacent to B_j and B_k is adjacent to each vertex in $\mathbf{S}$ in $\mathbb{M}$.

Next, we prove that $\forall 1 \le j \le m$, there is not $B_j = K_j$. That is, the minimal possible directed path from A to K_j cannot be $A \circ\!\!\to\!\!* K_j$ in $\mathbb{M}$. Suppose there is $A \circ\!\!\to\!\!* K_j$ in $\mathbb{M}$. According to the result above, for any $S \in \mathbf{S}$, K_j is adjacent to S. In this case, there must be $m \ge 3$, for otherwise if the unbridged path is just $K_1 \circ\!\!-\!\!\circ K_2$, suppose $j = 1$, then there must be $\mathcal{F}_{K_1} \supseteq \mathcal{F}_{K_2}$, contradicting with the definition of unbridged path in Definition 9. We consider the circle path $K_j \circ\!\!-\!\!\circ K_{j+1} \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ K_m$ in $\mathbb{M}$ (If $j \ge m-1$, then we consider the circle path $K_1 \circ\!\!-\!\!\circ \cdots K_j$ instead. And it is impossible that $m = 3$ and $j = 2$, for otherwise the path cannot be unbridged). There is $S_2 \in \mathcal{F}_{K_m} \backslash \mathcal{F}_{K_{m-1}}$. Since there cannot be an edge $K_{m-1} \to S_2$ in $\mathbb{M}$, which we have proven before, and $S_2 \in \mathcal{F}_{K_m} \backslash \mathcal{F}_{K_{m-1}}$, S_2 cannot be adjacent to K_{m-1}. Next, we can conclude that K_{m-2} is not adjacent to S_2, for otherwise in the substructure comprised of $K_{m-2}, K_{m-1}, K_m, S_2$, there must be $K_m \to S_2 \leftarrow K_{m-2}$ oriented by $\mathcal{R}_9$ in $\mathcal{P}$, which leads to $K_{m-2} \to S_2$ in $\mathbb{M}$, contradiction. Similarly, we can conclude that no vertices in $K_j, K_{j+1}, \cdots, K_m$ are adjacent to S_2. However, we have proven that K_j is adjacent to any S in $\mathbf{S}$, contradiction. Hence, for any $B_j, 1 \le j \le m$, $B_j \ne K_j$.

Next, we will prove that for any $B_j, 1 \le j \le m$, B_j is also in the minimal possible directed path from A to K_k, where $1 \le k \le m$ and $i \ne j$. Without loss of generality, suppose $k > j$. Consider the minimal possible directed path $\langle A, B_{j+1}, \cdots, K_{j+1} \rangle$ from A to K_{j+1} in $\mathbb{M}$. We will prove that B_j is also the vertex adjacent to A in a minimal possible directed path from A to K_{j+1}. If B_{j+1} and B_j denote the same vertex, the result evidently holds. We just consider the case $B_{j+1} \ne B_j$. We have proven that B_j is adjacent to B_{j+1} before.

Note each vertex in $K_1, \cdots, K_m$ is an ancestor of Y in $\mathcal{M}$. According to Lemma 11, there must be $p_1 = A \circ\!\!\to B_j \to \cdots \to K_j$ and $p_2 = A \circ\!\!\to B_{j+1} \to \cdots \to K_{j+1}$. Since we have proven $K_{j+1} \ne B_{j+1}$ above, A cannot be adjacent to K_{j+1}, for otherwise p_2 is not a minimal possible directed path. If B_j is adjacent to K_{j+1} in $\mathbb{M}$, it is evident that there is a minimal possible directed path $\langle A, B_j, K_{j+1} \rangle$, thus B_j is also the vertex adjacent to A in a minimal possible directed path from A to K_{j+1}. If B_j is not adjacent to K_{j+1}, due to the possible directed path $B_j \to \cdots \to K_j \circ\!\!-\!\!\circ K_{j+1}$ in $\mathbb{M}$, there must be a minimal possible directed path p' from B_j to K_{j+1} in $\mathbb{M}$ according to Lemma 2. And A is not adjacent to any non-endpoint in p' since $A \circ\!\!\to B_j \to \cdots \to K_j$ is a minimal possible directed path and A is not adjacent to K_{j+1}. Hence, we have a new minimal possible directed path $A \circ\!\!\to B_j \bigoplus p'$ from A to K_{j+1}. Similarly, we can prove that B_j is also the vertex adjacent to A in a minimal possible directed path from A to K_{j+2}. Repeat the process, we can prove that B_j is also the vertex adjacent to A in a minimal possible directed path from A to K_k, for any $1 \le k \le m$ and $k \ne j$.

Till now, we have proven that there exists a minimal possible directed path from A to each vertex in $K_1, \cdots, K_m$ such that B_j is the common vertex adjacent to A in all paths. And it is evidently that the minimal possible directed path is an uncovered path. Hence, if $B_j \notin \mathbf{S}_i$, the edge $A \circ\!\!\to\!\!* B_j$ should be transformed by $\mathcal{R}_{12}$ in $\mathbb{M}$ on Line 6 of Algorithm 5, thus the algorithm will enter the next loop, contradiction. Hence there is not an unbridged path relative to $\mathbf{S}_i$ in $\mathbb{M}[\text{PossDe}(\bar{\mathbf{W}}, \mathbb{M}[-\mathbf{S}_i])]$. □

4.3.5. Proof of Corollary 1

Proof. The proof follows Theorem 1 based on Theorem 6 and Theorem 7. Theorem 6 and Theorem 7 can ensure that by using Algorithm 5 for each potential adjustment set, we can find the set of causal effects in all the DAGs represented by the MAGs valid to $\mathbb{M}$. And since in PAGrules, all possible local transformation are considered on Line 5, PAGrules can return the set of causal effects in all the DAGs represented by the MAGs consistent with $\mathcal{P}$. □

5. Experiments

In this part, we evaluate the effectiveness and efficiency of our proposed methods for set determination, PAGcauses and PAGrules, as presented in Algorithm 1 and Algorithm 6, respectively. Compared to PAGcauses, PAGrules introduces two main improvements. First, it significantly reduces the size of candidate sets that could potentially serve as adjustment sets. Second, it incorporates novel orientation rules to further decrease the computational complexity of set determination. To assess the individual contributions of these improvements, we conduct an ablation study. Specifically, we propose a variant, PAGcauses-pruning, which reduces the size of candidate sets by incorporating Algorithm 2. In PAGcauses-pruning, we replace Lines 8–13 of Algorithm 1 in PAGcauses with Algorithm 2. We also take LV-IDA [17] as a baseline, which determines the set by enumerating the MAGs in a subspace of MEC. Note LV-IDA returns the causal effect in the MAG $\mathcal{M}$ where the causal effect is *identifiable* via adjustment, *i.e.*, $\text{D-SEP}(X, Y, \mathcal{M}_{\underline{X}}) \cap \text{De}(X, \mathcal{M}) = \emptyset$, while PAGcauses/PAGcauses-pruning/PAGrules also consider the MAG $\mathcal{M}$ where the causal effect is not identifiable but there is some DAG $\mathcal{D}$ represented by $\mathcal{M}$ such that the causal effect is identifiable by some observed variables, *i.e.*, $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}}) \cap \text{De}(X, \mathcal{M}) = \emptyset$. Hence, to compare the methods fairly, we modify LV-IDA by finding $\text{D-SEP}(X, Y, \mathcal{M}_{\underset{\sim}{X}})$ instead of $\text{D-SEP}(X, Y, \mathcal{M}_{\underline{X}})$. In this way, all the methods tackle the same task. Note this modification does not require any additional computation for LV-IDA.

We generate random DAGs with vertex number $d = 8, 10, 12, 14, 16$ and each edge between two vertices occurs with probability $\rho = 0.2, 0.3, 0.4, 0.5$, which is called *graph density*. The DAG is parameterized as a linear Gaussian structural equation model. The weight of each edge is drawn from Uniform([1, 2]). For each graph, we randomly select four vertices as latent variables and the others as observed variables. To prevent degeneration to a trivial case where X has no causal effect on Y, we select the last vertex in the causal order as Y. We randomly select a variable X. For each set of parameters, we generate 100 causal graphs and obtain the output set in the time limit of 3000 seconds for each graph. For LV-IDA, when the running time reaches the limit, it stops and returns the

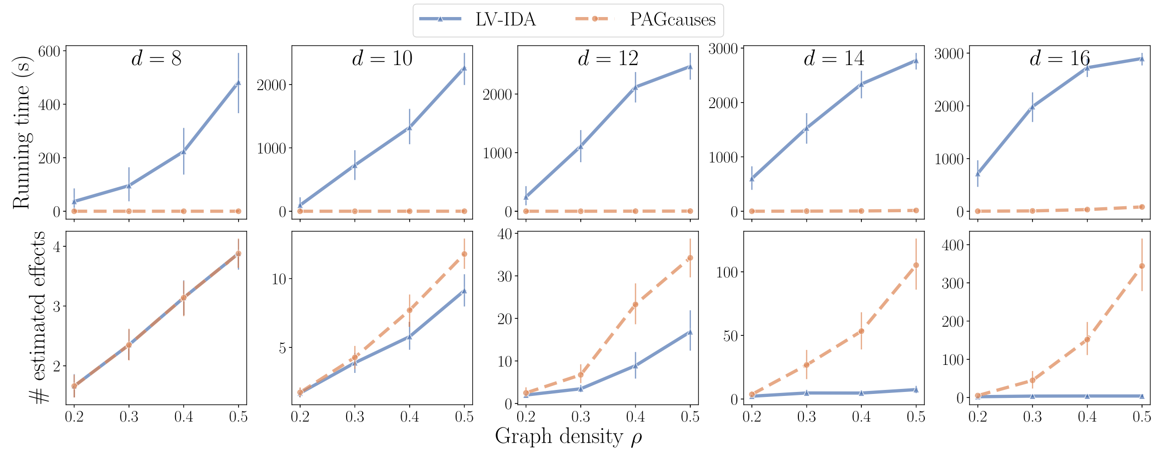

Fig. 5. Results of the number of returned causal effects values and running time over 100 simulations for each vertice number(including 4 latent ones)/graph density by PAGcauses and LV-IDA. The vertical line represents the 95% confidence interval generated by bootstrap sampling. The maximum running time for each simulation is 3000 seconds.

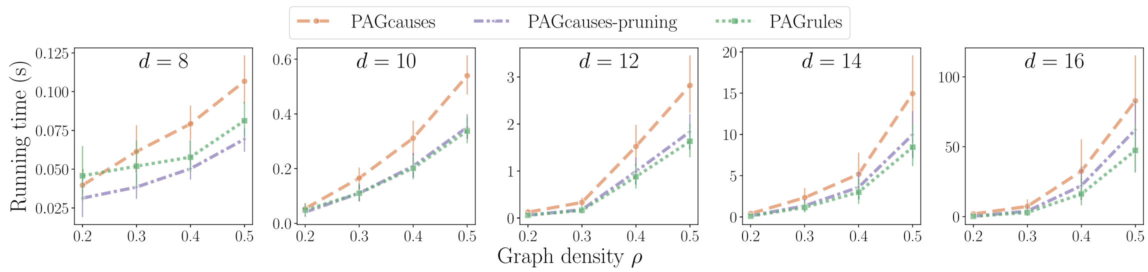

Fig. 6. Results of the running time over 100 simulations for each vertex number(including 4 latent ones)/graph density by PAGcauses, PAGcauses-pruning and PAGrules. The vertical line represents the 95% confidence interval generated by bootstrap sampling.

causal effects in the enumerated MAGs. For PAGcauses, it returns no effects if the time is used up. Since the main focus is on the set determination based on a PAG, we obtain PAG directly with the true covariance matrix of the observable variables.

We show in Fig. 5 the average number of the output set of causal effects and running time in 100 graphs under each parameter by PAGcauses and LV-IDA. The tremendous superiority of the proposed methods demonstrate the advantage on complexity taken by the proposed graphical characterization of possible adjustment set which helps circumvent the enumerations of MAGs.

Note when d and ρ is small, PAGcauses and LV-IDA obtain the same set. When $d \geq 10$, the sets are different because usually LV-IDA cannot enumerate all the MAGs within the limited time. When d grows, the number of returned causal effects by LV-IDA tends to 0. We give a rough analysis for this phenomenon. Suppose within the limited time we can enumerate N MAGs at most. 2^d is the rough number of causal effects and $3^{d^2/2}$ is the rough number of MAGs. As d grows, the expected number of returned causal effects $N \times 2^d/3^{d^2/2}$ tends to 0. Hence, it implies that if we determine the set of possible causal effects by MAG enumerations, due to the sparsity of the causal effect relative to the MAG space, it is even possibly hard to get just a few causal effect values in limited time when the space of MAGs is large. In contrast, the methods that does not need to enumerate MAGs will not suffer the sparsity of the causal effect, thus could always return some causal effect estimates.

Next, we will only compare the three methods PAGcauses, PAGcauses-pruning, and PAGrules. The reason that we do not include the results of LV-IDA anymore is, the complexity of the three methods is super-exponentially less than LV-IDA, such that the difference between the three methods cannot be highlighted. In addition, according to our experiments, the three methods can find the same set of causal effects. Hence we will only show the average running time in 100 graphs under each parameter, but do not show the number of the output set of causal effects. The experimental results are given in Fig. 6. In general, PAGrules is more efficient than PAGcauses-pruning, and PAGcauses-pruning is more efficient than PAGcauses, demonstrating the effectiveness of utilizing the novel rules and the proposed method to reduce the candidate adjustment set in set determination. Note when $d = 8$, there are only four observed variables in a PAG, in which case the proposed rule $\mathcal{R}_{12}$ cannot be triggered due to the insufficient number of variables. However, detecting whether $\mathcal{R}_{12}$ is triggered still takes a $\mathcal{O}(m^3 d)$ complexity, leading to unnecessary computational complexity. In this case, PAGcauses-pruning finishes with less time than PAGrules.

Finally, we conduct experiments to compare the three methods in a high-dimensional setting. Specifically, we set $d = 100$ and vary the graph density as $\rho = \{0.15, 0.2, 0.25, 0.3\}$. All other experimental settings are kept consistent with the previous experiments, except that this implementation is executed in serial. The results, presented in Table 1, demonstrate the effectiveness of Algorithm 2 and the proposed rules in accelerating set determination. Notably, PAGcauses fails to return any valid results within the allotted time

Table 1
Results of the running time over 100 simulations by PAGcauses, PAGcauses-pruning and PAGrules with vertex number $d = 100$ and varying graph density ρ in the format of mean ± std. Since enumeration of all sets $\mathbf{W}$ satisfying Definition 6 is involved in PAGcauses, no result can be returned within 3600 seconds.

Method-Time(s)-ρ	0.15	0.2	0.25	0.3
PAGcauses	-	-	-	-
PAGcauses-pruning	0.252 ± 0.055	0.304 ± 0.058	0.384 ± 0.073	0.465 ± 0.096
PAGrules	0.214 ± 0.050	0.267 ± 0.052	0.349 ± 0.075	0.418 ± 0.090

in these cases. This is because PAGcauses first enumerates all sets $\mathbf{W}$ that satisfy Definition 6. When $d = 100$, the number of such sets is approximately $\mathcal{O}(2^{98})$, which is computationally prohibitive.

6. Conclusion

In this paper, we present the first method for set determination that does not require enumerating any DAGs or MAGs. By introducing new graphical conditions, our approach determines the set of possible causal effects with super-exponentially less complexity compared to previous methods. In addition, we propose two novel orientation rules for incorporating BK to a PAG, one of which is fundamentally different in form from existing rules. These rules offer a promising direction toward addressing the open problem of causal identification given BK. By utilizing the proposed rules, we present a refined method for set determination that achieves improved efficiency. Experiments validate the effectiveness and efficiency of our proposed methods.

There are two promising directions for future research. The first is causal effect estimation that integrates both data and knowledge. In recent years, machine learning methods that leverage data and knowledge have attracted extensive attention [38,39]. In the context of causal inference, the availability of domain knowledge enables the identification of additional causal relations [40,18,41,9], thereby facilitating more precise causal effect estimation. Hence, developing set determination methods that effectively leverage data and knowledge is an intriguing problem, and our proposed rules can be applicable. The second direction is to develop decision methods when sufficient interactions with environments are unavailable. Some existing approaches attempt to address this by estimating causal effects [42–44]. However, causal effect estimation relies on identifying underlying causal relations, which is a fundamentally difficult problem - potentially even more challenging than the decision-making task itself. Recently, Zhou [45] highlighted the importance of identifying an intermediate relation between association and causation for decision-making, termed *influence* [46]. Essentially, influence relations represent a form of knowledge that, when combined with data, can support effective decision-making. Based on this idea, Qin et al. [47], Du et al. [48] presented novel decision methods. Exploring such influence-based decision-making methods is a valuable direction for future research.

CRediT authorship contribution statement

Tian-Zuo Wang: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **Lue Tao:** Writing – review & editing, Visualization. **Tian Qin:** Writing – review & editing, Visualization. **Zhi-Hua Zhou:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by NSFC (62495092), JiangsuSF (BK20232003, BK20241201), and the Collaborative Innovation Center of Novel Software Technology and Industrialization, the AI & AI for Science Project of Nanjing University. We are grateful for the anonymous reviewers from ICML and AIJ for their valuable comments. Tian-Zuo Wang was supported by National Postdoctoral Program for Innovative Talents and the Xiaomi Foundation.

Data availability

No data was used for the research described in the article.

References

[1] J. Tian, J. Pearl, A general identification condition for causal effects, in: Proceedings of the 18th National Conference on Artificial Intelligence and 14th Conference on Innovative Applications of Artificial Intelligence, 2002, pp. 567–573.

[2] I. Shpitser, J. Pearl, Identification of joint interventional distributions in recursive semi-markovian causal models, in: Proceedings of the 21st National Conference on Artificial Intelligence and the 18th Innovative Applications of Artificial Intelligence Conference, 2006, pp. 1219–1226.

[3] J. Pearl, Causality, Cambridge University Press, 2009.

[4] I. Shpitser, T.J. VanderWeele, J.M. Robins, On the validity of covariate adjustment for estimating causal effects, in: Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence, 2010, pp. 527–536.

[5] Y. Jung, J. Tian, E. Bareinboim, Estimating identifiable causal effects on markov equivalence class through double machine learning, in: Proceedings of the 38th International Conference on Machine Learning, 2021, pp. 5168–5179.

[6] J. Zhang, On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias, Artif. Intell. 172 (2008) 1873–1896.

[7] M.H. Maathuis, M. Kalisch, P. Bühlmann, et al., Estimating high-dimensional intervention effects from observational data, Ann. Stat. 37 (2009) 3133–3164.

[8] E. Perkovic, M. Kalisch, M.H. Maathuis, Interpreting and using cpdags with background knowledge, in: Proceedings of the 33rd Conference on Uncertainty in Artificial Intelligence, 2017.

[9] Z. Fang, Y. He, IDA with background knowledge, in: Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence, 2020, pp. 270–279.

[10] J. Witte, L. Henckel, M.H. Maathuis, V. Didelez, On efficient adjustment in causal graphs, J. Mach. Learn. Res. 21 (2020) 9956–10000.

[11] B. van der Zander, M. Liskiewicz, J. Textor, Constructing separators and adjustment sets in ancestral graphs, in: Proceedings of the 30th Conference on Uncertainty in Artificial Intelligence, 2014, pp. 907–916.

[12] E. Perkovic, J. Textor, M. Kalisch, M.H. Maathuis, Complete graphical characterization and construction of adjustment sets in markov equivalence classes of ancestral graphs, J. Mach. Learn. Res. 18 (2017) 220:1–220:62.

[13] B. van der Zander, M. Liskiewicz, J. Textor, Separators and adjustment sets in causal graphs: complete criteria and an algorithmic framework, Artif. Intell. 270 (2019) 1–40.

[14] R. Guo, E. Perkovic, Minimal enumeration of all possible total effects in a markov equivalence class, in: The 24th International Conference on Artificial Intelligence and Statistics, 2021, pp. 2395–2403.

[15] L. Henckel, E. Perković, M.H. Maathuis, Graphical criteria for efficient total effect estimation via adjustment in causal linear models, J. R. Stat. Soc. Ser. B 84 (2022) 579–599.

[16] Z. Fang, R. Zhao, Y. Liu, Y. He, On the representation of causal background knowledge and its applications in causal inference, CoRR, arXiv:2207.05067, 2022.

[17] D. Malinsky, P. Spirtes, Estimating causal effects with ancestral graph markov models, in: Conference on Probabilistic Graphical Models, 2016, pp. 299–309.

[18] C. Meek, Causal inference and causal explanation with background knowledge, in: Proceedings of the 11th Annual Conference on Uncertainty in Artificial Intelligence, 1995, pp. 403–410.

[19] T.-Z. Wang, T. Qin, Z.-H. Zhou, Sound and complete causal identification with latent variables given local background knowledge, Artif. Intell. 322 (2023) 103964.

[20] A. Venkateswaran, E. Perkovic, Towards complete causal explanation with expert knowledge, preprint, arXiv:2407.07338, 2024.

[21] B. Andrews, P. Spirtes, G.F. Cooper, On the completeness of causal discovery in the presence of latent confounding with tiered background knowledge, in: Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics, 2020, pp. 4002–4011.

[22] T.-Z. Wang, T. Qin, Z.-H. Zhou, Estimating possible causal effects with latent variables via adjustment, in: Proceedings of the 40th International Conference on Machine Learning, 2023, pp. 36308–36335.

[23] G.A. Dirac, On rigid circuit graphs, in: Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg, vol. 25, Springer, 1961, pp. 71–76.

[24] M.C. Golumbic, Algorithmic Graph Theory and Perfect Graphs, Elsevier, 2004.

[25] T. Richardson, P. Spirtes, et al., Ancestral graph markov models, Ann. Stat. 30 (2002) 962–1030.

[26] J. Zhang, Causal reasoning with ancestral graphs, J. Mach. Learn. Res. 9 (2008) 1437–1474.

[27] R.A. Ali, T. Richardson, P. Spirtes, J. Zhang, Orientation rules for constructing markov equivalence classes of maximal ancestral graphs, Technical Report, 2005.

[28] T.J. VanderWeele, I. Shpitser, A new criterion for confounder selection, Biometrics 67 (2011) 1406–1413.

[29] P. Spirtes, C.N. Glymour, R. Scheines, Causation, Prediction, and Search, MIT Press, 2000.

[30] D. Colombo, M.H. Maathuis, M. Kalisch, T.S. Richardson, Learning high-dimensional directed acyclic graphs with latent and selection variables, Ann. Stat. (2012) 294–321.

[31] M.H. Maathuis, D. Colombo, et al., A generalized back-door criterion, Ann. Stat. 43 (2015) 1060–1088.

[32] Z. Hu, R. Evans, Faster algorithms for markov equivalence, in: Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence, 2020, pp. 739–748.

[33] M. Wienöbst, M. Bannach, M. Liskiewicz, A new constructive criterion for markov equivalence of mags, in: Proceedings of the 38th Conference on Uncertainty in Artificial Intelligence, 2022, pp. 2107–2116.

[34] J. Zhang, P. Spirtes, A transformational characterization of markov equivalence for directed acyclic graphs with latent variables, in: Proceedings of the 21st Conference in Uncertainty in Artificial Intelligence, 2005, pp. 667–674.

[35] J. Tian, Generating markov equivalent maximal ancestral graphs by single edge replacement, in: Proceedings of the 21st Conference in Uncertainty in Artificial Intelligence, 2005, pp. 591–598.

[36] T.-Z. Wang, Z.-H. Zhou, Actively identifying causal effects with latent variables given only response variable observable, in: Advances in Neural Information Processing Systems, 2021, pp. 15007–15018.

[37] H. Zhao, Z. Zheng, B. Liu, On the markov equivalence of maximal ancestral graphs, Sci. China Ser. A, Math. 48 (2005) 548–562.

[38] Z.-H. Zhou, Abductive learning: towards bridging machine learning and logical reasoning, Sci. China Inf. Sci. 62 (2019) 76101:1–76101:3.

[39] Z.-H. Zhou, Y.-X. Huang, Abductive learning, in: P. Hitzler, M.K. Sarker (Eds.), Neuro-Symbolic Artificial Intelligence: The State of the Art, IOS Press, Amsterdam, 2022, pp. 353–369.

[40] T.-Z. Wang, W.-B. Du, Z.-H. Zhou, An efficient maximal ancestral graph listing algorithm, in: Proceedings of the 41st International Conference on Machine Learning, 2024, pp. 50353–50378.

[41] Y. He, Z. Geng, Active learning of causal networks with intervention experiments and optimal designs, J. Mach. Learn. Res. 9 (2008) 2523–2547.

[42] T. Qin, T.-Z. Wang, Z.-H. Zhou, Budgeted heterogeneous treatment effect estimation, in: Proceedings of the 38th International Conference on Machine Learning, 2021, pp. 8693–8702.

[43] U. Shalit, F.D. Johansson, D.A. Sontag, Estimating individual treatment effect: generalization bounds and algorithms, in: Proceedings of the 34th International Conference on Machine Learning, 2017, pp. 3076–3085.

[44] S. Lee, E. Bareinboim, Structural causal bandits: where to intervene?, in: Advances in Neural Information Processing Systems, 2018, pp. 2573–2583.

[45] Z.-H. Zhou, Rehearsal: learning from prediction to decision, Front. Comput. Sci. 16 (4) (2022) 164352.

[46] Z.-H. Zhou, Rehearsal: learning from prediction to decision, in: Keynote at the CCF Conference on AI, Urumqi, China, 2023.

[47] T. Qin, T.-Z. Wang, Z.-H. Zhou, Rehearsal learning for avoiding undesired future, Adv. Neural Inf. Process. Syst. 36 (2023) 80517–80542.

[48] W.-B. Du, T. Qin, T.-Z. Wang, Z.-H. Zhou, Avoiding undesired future with minimal cost in non-stationary environments, Adv. Neural Inf. Process. Syst. 37 (2024) 135741–135769.