An Efficient Maximal Ancestral Graph Listing Algorithm

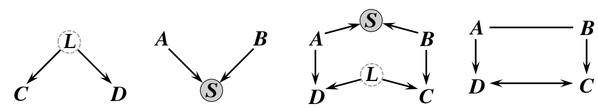
Tian-Zuo Wang [1] Wen-Bo Du [1] Zhi-Hua Zhou [1]

Abstract

Maximal ancestral graph (MAG) is a prevalent graphical model to characterize causal relations in the presence of *latent variables* including latent confounders and selection variables. Given observational data, only a Markov equivalence class (MEC) of MAGs is identifiable if without some additional assumptions. Due to this fact, *MAG listing*, listing all the MAGs in the MEC, is usually demanded in many downstream tasks. To the best of our knowledge, there are no relevant methods for MAG listing other than brute force in the literature. In this paper, we propose the first brute-force-free MAG listing method, by determining the local structures of each vertex recursively. We provide the graphical characterization for each *valid local transformation* of a vertex, and present *sound and complete* rules to incorporate the valid local transformation in the presence of latent confounders and selection variables. Based on these components, our method can efficiently output all the MAGs in the MEC with *no redundance*, that is, every intermediate graph in the recursive process is necessary for the MAG listing task. The empirical analysis demonstrates the superiority of our proposed method on efficiency and effectiveness.

1. Introduction

Causality is a vital research topic in artificial intelligence. Under Pearl's causality framework (Pearl, 2009), a key component is a graphical model to characterize the causal relations among variables. In the graphical model, each vertex denotes a variable and each edge denotes a causal relation. *Directed acyclic graph (DAG)* is one of the most widespread graphical models. Extensive studies based on DAGs are conducted, leading to the establishment of many solid results (Meek, 1995; Wienöbst et al., 2023).



(a) L (b) S (c) DAG $\mathcal{D}$ (d) MAG $\mathcal{M}$

Figure 1: Fig. 1(a): L is a latent confounder if we cannot observe L. Fig. 1(b): S is a selection variable if we cannot observe S and the collected data relies on the value of S. Fig. 1(c): a DAG $\mathcal{D}$ where L is a latent confounder and S is a selection variable. Fig. 1(d): A MAG characterizing the causal relations over the observable vertices in $\mathcal{D}$.

In real tasks, some variables relevant to the tasks are usually latent, in which case DAG is not sufficient to characterize the causal relations. Specifically, latent variables are generally classified into *latent confounders* and *selection variables*. The existence of latent variables could take additional dependence to the observable variables. See Fig. 1 for an illustration. As shown by Fig 1(a), latent confounder *influences* more than one observable variable. For example, economic policy influences both income and consumption, but it is hardly to be evaluated and thus a latent confounder. And as shown by Fig.1(b), selection variable *is influenced* by more than one observable variable. In the presence of selection variables, the collected data relies on the value of the selection variables and thus jeopardizes the representativeness of the data for the underlying population (Hünermund & Bareinboim, 2023). For example, when investigating the relationship between student talent and effort, if the student data is collected from top-tier universities, the dependence between talent and effort could be overestimated, as both talent and effort influence whether a student can enter a top-tier university.

To accommodate the presence of latent confounders and selection variables, *maximal ancestral graph (MAG)* is proposed (Richardson et al., 2002). A MAG characterizes the causal relations among the *observable variables* regardless of whether latent variables exist and where they are. Roughly speaking, a MAG can be seen as a projection graph on the observable variables of an underlying DAG contain-

[1]National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, 210023, China. Correspondence to: Zhi-Hua Zhou <zhouzh@lamda.nju.edu.cn>.

Proceedings of the 41^{st} International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

ing all relevant variables. In addition to directed edges, there could also be bi-directed and undirected edges in a MAG, which imply the presence of latent confounders and selection variables, respectively. For example, for an underlying DAG in Fig. 1(c), where S is a selection variable and L is a latent confounder, the associated MAG is as Fig. 1(d).

A DAG or MAG is generally necessary for exploiting causality in real tasks, however, they are usually not pre-known. Hence, learning the causal graph from data is a prerequisite step. Given observational data, the existing theoretical results imply that only a part of causal relations is identifiable without some additional assumptions (Spirtes et al., 2000). Specifically, we can only identify a *Markov equivalence class (MEC)* of causal graphs (DAGs or MAGs) with observational data. In the absence of latent variables, a *completed partially directed acyclic graph (CPDAG)* can be learned, representing the MEC of DAGs; and in the presence of latent variables, a *partial ancestral graph (PAG)* can be learned, representing the MEC of MAGs.

In light of the identifiability of only a MEC given observational data, listing all the graphs in the MEC is a stepping stone for many downstream applications such as intervention variable selection (He & Geng, 2008; Hauser & Bühlmann, 2014; Wang et al., 2023b), set determination of causal effects given a MEC (Maathuis et al., 2009; Malinsky & Spirtes, 2016), and graph check (Kocaoglu, 2023)[Sec. D.4]. There have been many mature algorithms for *DAG listing*, *i.e.*, listing all the DAGs given a CPDAG (Chickering, 1995; Chen et al., 2016; Wienöbst et al., 2023). These algorithms have been implemented in causality software package such as pcalg (Kalisch et al., 2012) and causaldag (Chandler Squires, 2018). However, to the best of our knowledge, there are no methods for *MAG listing* in the literature, *i.e.*, listing all the MAGs given a PAG. There are only two relevant studies with *additional requirement of no selection variables* (Malinsky & Spirtes, 2016; Wang et al., 2023b).

A primary barrier to the development of relevant methods is the absence of theoretical results supporting MAG listing. If we revisit the methods for DAG listing, there are principally two solutions. One is based on Meek rules (Meek, 1995), which provide sound and complete rules to incorporate additional causal structural knowledge into a graph. Given a CPDAG, we can recursively enumerate each unidentified edge and update the graph with the rules until all the edges are determined. The other is based on DAG transformation (Chickering, 1995), which relies on the fact that two Markov equivalent DAGs can be converted to each other through some single-edge transformations. However, the two solutions are not directly applicable for MAG listing tasks. For the first kind of method, it remains an open problem for many years that which set of rules are sound and complete to incorporate causal knowledge into an ancestral graph (Zhang, 2008); for the second kind of method, two Markov equivalent MAGs cannot necessarily be converted to each other through one-edge transformations (Zhang & Spirtes, 2005; Tian, 2005).

In this paper, we establish the first theoretical results that accommodate MAG listing in the presence of both latent confounders and selection variables, through which we propose a MAG listing method by recursively performing local transformations of each vertex. A *local transformation* of a vertex means determining the local structures of the vertex. At each iteration, our method selects a vertex, and evaluate the validity of each possible local transformation. This evaluation determines whether the local transformation can lead to a MAG in the MEC in subsequent transformations. By directly pruning invalid local transformations, we effectively avoid the unnecessary computational efforts on intermediate graphs that cannot yield a MAG in the MEC. Furthermore, we present sound and complete orientation rules to incorporate the structure knowledge of local transformation, which can be used to uncover the maximal informative graphical models under each local transformation, thereby eliminating the possibility of considering structures that are non-contributive. Our contributions are threefolds.

(1) We establish the graphical condition for determining the validity of a local transformation, through which we can prune the local transformation that cannot yield a MAG in the MEC in the subsequent transformation.

(2) We present sound and complete rules to incorporate local transformation in the presence of latent variables, through which we can uncover the maximal informative graphical models given the local transformation.

(3) We propose the algorithm to list all the MAGs in the Markov equivalence class represented by a given PAG, and validate the tremendous superiority on efficiency compared to brute-force via experiments.

2. Preliminary

In this paper, we refer to a variable/vertex with a capital letter (e.g., X) and a set of variables/vertices with boldface font (e.g., $\mathbf{X}$). Let G denote a graph and $\mathbf{V}(G)$/$\mathbf{E}(G)$ denote the set of vertices/edges in G. Given G, for any $\mathbf{V}' \subseteq \mathbf{V}(G)$, *the subgraph of G induced by* $\mathbf{V}'$ consists of the vertices in $\mathbf{V}'$ and the edges connecting the vertices in $\mathbf{V}'$ in G.

A *mixed graph* is a graph which contains undirected, directed, and bi-directed edges. The two ends of an edge are called *marks*. A *partial mixed graph (PMG)* is a graph with three kinds of marks: arrowheads, tails, and circles($\circ$). $\circ$ means that the mark here can be either arrowhead or tail but is uncertain to us. An edge $V_i \circ\!\!-\!\!\circ V_j$ is called a *circle edge*. For two vertices V_i and V_j in G, V_i is a *par-*

ent/child of V_j if there is $V_i \to V_j$/$V_i \leftarrow V_j$ in G. For a path $p = \langle V_1, V_2, \cdots, V_d\rangle$ in G, p is a *directed path* from V_1 to V_d if there is $V_i \to V_{i+1}, \forall 1 \leq i \leq d-1$; p is a *possible directed path* from V_1 to V_d if for each edge between V_i and V_{i+1}, $1 \leq i \leq d-1$, there is no arrowhead at V_i or tail at V_{i+1}; p is *uncovered* if for any consecutive triple $\langle V_{i-1}, V_i, V_{i+1}\rangle$, $2 \leq i \leq d-1$, V_{i-1} is not adjacent to V_{i+1}; p is *minimal* if any two non-consecutive vertices are not adjacent. V_i is an *ancestor/descendant* of V_j in G if $V_i = V_j$ or there is a directed path from V_i/V_j to V_j/V_i. V_i is a *possible descendant* of V_j in G if $V_i = V_j$ or there is a possible directed path from V_j to V_i. Denote the set of parents/ancestors/descendants/possible descendants of V_i in G by $\mathrm{Pa}(V_i, G)/\mathrm{Anc}(V_i, G)/\mathrm{De}(V_i, G)/\mathrm{PossDe}(V_i, G)$. If there is a directed path from V_1 to V_d and an edge $V_d \to V_1$/$V_d \leftrightarrow V_1$, we say they form a *directed cycle/almost directed cycle*. p is a *minimal possible directed path* if it is minimal and a possible directed path.

In a graph G, $*$ denotes a wildcard, which means that the mark here can be any one. If there is $V_i {*\!\!\rightarrow} V_j \leftarrow\!\!* V_k$, then they form a *collider (at V_j)*; if V_i is not adjacent to V_k as well, it is called an *unshielded collider (at V_j)* or *v-structure (at V_j)*. The *circle component* of a graph G is the subgraph of G that only remains all the vertices and all the circle edges. A graph is *chordal* if any cycle with more than four vertices has a chord that connects two vertices.

For a mixed graph G, if there is not a directed cycle or almost directed cycle, and not an edge into an undirected edge with an arrowhead, then G is *ancestral*. An ancestral graph is *maximal* if for any non-adjacent vertices, there is a set of vertices that *m-separates* them (Richardson et al., 2002). If a mixed graph is ancestral and maximal, then it is a *maximal ancestral graph (MAG)*, denoted by $\mathcal{M}$. A *partial ancestral graph (PAG)* represents a Markov equivalence class (MEC) of MAGs, denoted by $\mathcal{P}$. PAG has the same skeleton with the MAGs, and each mark in the PAG is an arrowhead/tail if the mark is an arrowhead/tail in all the Markov equivalent MAGs, but is a circle in the PAG if there are both arrowheads and tails in all the Markov equivalent MAGs. For a PMG $\mathbb{M}$ obtained from $\mathcal{P}$, a MAG $\mathcal{M}$ is *consistent with* $\mathbb{M}$ if $\mathcal{M}$ has all the non-circle marks in $\mathbb{M}$ and is a MAG in the MEC represented by $\mathcal{P}$.

3. The Proposed Method

In this section, we present the method to list all the MAGs consistent with a PAG $\mathcal{P}$. The PAG can be learned from observational data (Spirtes et al., 2000; Zhang, 2008). Denote $\mathbf{V}(\mathcal{P}) = \{V_1, \cdots, V_d\}$. In the following, we provide an overview of our method in Sec. 3.1 and propose two important components of our method in Sec. 3.2 and Sec. 3.3. The detailed algorithm is shown in Sec. 3.4.

3.1. An Overview of Our Method

For MAG listing, a direct method is to enumerate all the graphs obtained by transforming all the circles. For each generated graph, we evaluate whether it is a MAG consistent with $\mathcal{P}$ and output it if so. A comprehensive algorithm is detailed in Appendix B. However, such a brute-force method is very inefficient, since most of the enumerated graphs are not MAGs consistent with $\mathcal{P}$. In fact, when a graph is with some specific structures, we have been able to ascertain that there is not a MAG consistent with $\mathcal{P}$ containing these structures. For example, consider the PAG $\mathcal{P}$ on the first line of Fig. 2. There is an unshielded triple $D \circ\!\!-\!\!\circ A \circ\!\!-\!\!\circ B$ in $\mathcal{P}$. If we introduce a new unshielded collider $D {*\!\!\rightarrow} A \leftarrow\!\!* B$, additional conditional independences are introduced and thus any enumerated graphs with this unshielded collider cannot be a MAG consistent with $\mathcal{P}$, for this reason, we can prune the step of transforming other circles. Likewise, if there is an edge into an undirected edge with an arrowhead, it violates the ancestral property, precluding the existence of a MAG with this structure. In summary, in the process of transforming all circles, if any such invalid sub-structure emerges, further transformation of the remaining circles becomes redundant.

Building upon the insights above, we propose a MAG listing method based on *local transformation*. That is, given a PAG $\mathcal{P}$, we recursively select one vertex X with circles and introduce the local transformations of the circles at X. A local transformation of X is *valid* if there are some MAGs consistent with $\mathcal{P}$ that have the same marks at X as those indicated by the transformation. Then, we *only* consider the subsequent local transformations of other vertices for the valid local transformations of X. See $\mathcal{P}$ on the first line of Fig. 2 for an example. We first consider what the circles at A could be. There are six valid local transformations of A, which are shown on the second line. As discussed above, the local transformation $D {*\!\!\rightarrow} A \leftarrow\!\!* B$ is not valid, thus we never consider the subsequent transformation of other vertices based on this local transformation. By recursively introducing the valid local transformations of each vertex, we can find all the MAGs consistent with $\mathcal{P}$.

The idea above takes three main challenges. The first challenge is how to determine the validity of each local transformation. The validity means that there exists some MAG consistent with the PAG and the local transformation, but it is evidently impractical to determine it by firstly enumerating all the MAGs consistent with the PAG and the local transformation. In fact, it is unknown whether we can determine the validity of a local transformation based on only a PAG. Previous results have shown that it is achievable in the absence of latent variables (Maathuis et al., 2009). Fortunately, it is achievable in the presence of latent variables as well. We establish the graphical characterization for valid

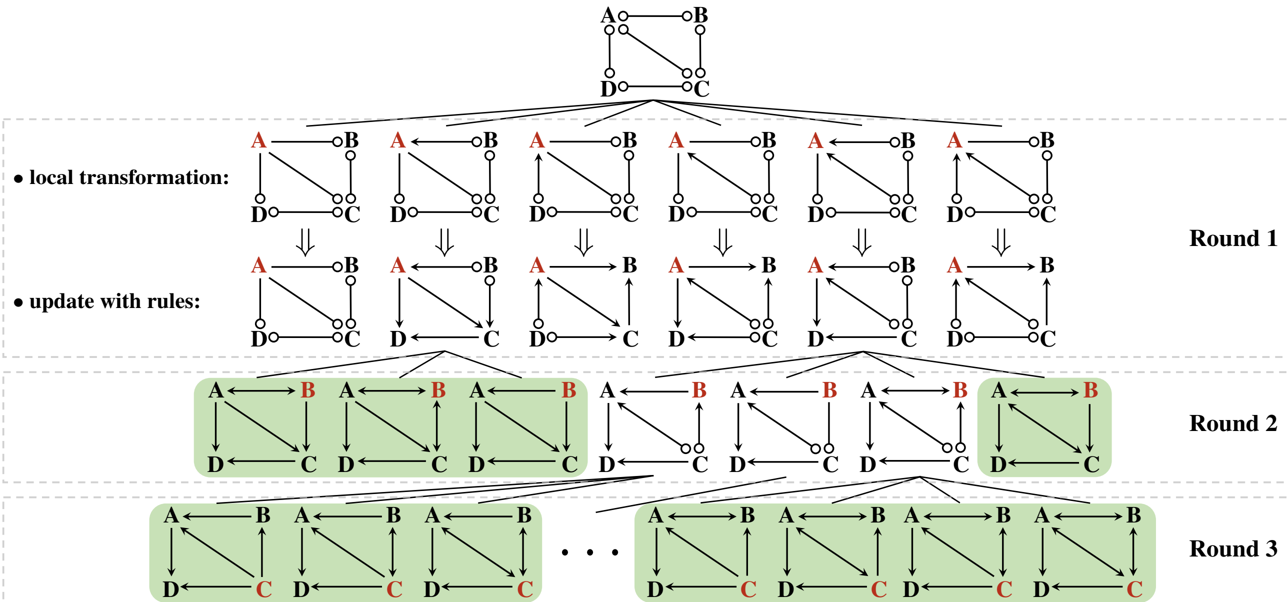


Figure 2: A realization process of MAGLIST. The graph in the root node denotes a PAG $\mathcal{P}$. The graphs in the first/second/third round are obtained from the previous round by introducing the local transformation of $A/B/C$. There are two parts in a round: introduce the local transformation and update the graph using the proposed orientation rules. Due to the space limit, we only show the two parts separately for the first round. The shaded graphs denote the MAGs that are output by MAGLIST.

local transformation, which can be evaluated in polynomial time with respect to to the number of vertices.

The second challenge lies in the fact that given a valid local transformation of a vertex X, we can identify more causal relations relative to the transformed marks at X, however, we do not know what these relations are or whether we can identify all of them. See the second graph on the second line of Fig. 2 as an example. When a valid local transformation of $B \circ\!\!\rightarrow A/A \multimap D/A \multimap C$ is introduced, $A \multimap C/A \multimap D$ can be transformed to $A \rightarrow C/A \rightarrow D$, for otherwise there is $B {*}\!\!\rightarrow A - C(D)$, which violates the ancestral property. And we can also identify $B \circ\!\!\rightarrow C \rightarrow D$ based on the ancestral property and the fact that there cannot be an additional unshielded collider. Hence, we can obtain the updated graph in the second graph on the third line of Fig. 2. However, it is still an open problem that what causal relations are identifiable in the presence of latent variables with additional causal knowledge (Zhang, 2008). Identifying the *complete* causal relations under each local transformation is necessary for our method, for otherwise it not only could take some invalid subsequent transformations which is costly, but also possibly affects the correctness of the graphical characterization for the validity of subsequent transformations.

The third challenge is that, as we recursively determine the valid local transformations and update the graph accordingly, our theoretical result needs to hold for not only a PAG, but also all the intermediate graphs in the recursive process. Hence, it is crucial to identify the common properties of all the possible graphs in this process, and utilize these properties to establish the graphical characterization for the valid local transformation and present the theoretical result for the complete causal identification.

Following the discussions above, in Sec. 3.2, we propose the graphical characterization for the valid local transformation of any vertex. In Sec. 3.3, we present sound and complete orientation rules to incorporate the causal knowledge taken by the local transformations. The two results above hold for all the graphs in the recursive process. Combining these two parts, we propose our algorithm in Sec. 3.4. All the proofs are presented in Appendix.

3.2. Valid Local Transformation

In this section, we provide the graphical characterization for valid local transformation. Given a partial mixed graph (PMG) $\mathbb{M}$, consider we want to transform the circles at vertex X. To denote each possible local transformation of X, we introduce the notation $\mathbf{C}(X)$, which is a subset of $\{V \in \mathbf{V}(\mathbb{M}) | V {*}\!\!\multimap X \text{ in } \mathbb{M}\}$. It implies that given $\mathbb{M}$, we transform $V {*}\!\!\multimap X$ to $V {*}\!\!\rightarrow X$ if $V \in \mathbf{C}(X)$ and transform $V {*}\!\!\multimap X$ to $V {*}\!\!- X$ if $V \in \{V \in \mathbf{V}(\mathbb{M}) | V {*}\!\!\multimap X \text{ in } \mathbb{M}\} \backslash \mathbf{C}(X)$. It is direct that $\mathbf{C}(X)$ uniquely determines a local transformation of X, and all subsets of $\{V \in \mathbf{V}(\mathbb{M}) | V {*}\!\!\multimap X \text{ in } \mathbb{M}\}$ can represent all the local transformations of X. As X is specified in this section, we use $\mathbf{C}$ to represent $\mathbf{C}(X)$ for brevity. At first, we provide Def. 1 to characterize the common properties of the graphs in the process where local transformation is introduced recursively. PAG is evidently a PMG compatible with local transformation (Zhang, 2008).

Definition 1. For a partial mixed graph $\mathbb{M}$, it is called a PMG *compatible with local transformation* if it satisfies the four following conditions:

(Chordal) The circle component in $\mathbb{M}$ is chordal.

(Balanced) For any three vertices A, B, C in $\mathbb{M}$, if $A \ast\!\!\rightarrow B \circ\!\!-\!\!\ast C$, then there is an edge between A and C with an arrowhead at C, namely, $A \ast\!\!\rightarrow C$. Furthermore, if the edge between A and B is $A \rightarrow B$, then the edge between A and C is either $A \rightarrow C$ or $A \circ\!\!\rightarrow C$ (i.e., it is not $A \leftrightarrow C$). And if $A -\!\!\circ B \circ\!\!-\!\!\ast C$, then A is adjacent to C. Forthermore, if $A -\!\!\circ B \circ\!\!-\!\!\circ C$, then $A -\!\!\circ C$; if $A -\!\!\circ B \circ\!\!\rightarrow C$, then $A \rightarrow C$ or $A \circ\!\!\rightarrow C$.

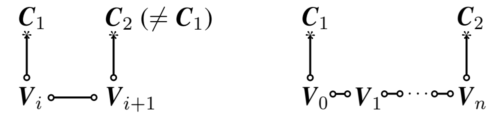
(Complete) For each circle at A on $A \circ\!\!-\!\!\ast B$ in $\mathbb{M}$, there exist MAGs $\mathcal{M}_1$ and $\mathcal{M}_2$ consistent with $\mathbb{M}$ with $A \leftarrow\!\!\ast B \in \mathbf{E}(\mathcal{M}_1)$ and $A -\!\!\ast B \in \mathbf{E}(\mathcal{M}_2)$.

(Constructive) We can always obtain a MAG consistent with $\mathbb{M}$ by transforming $-\!\!\circ$ / $\circ\!\!\rightarrow$ to $\rightarrow$ and transforming the circle component into a DAG without new unshielded colliders.

Remark 1. Recall the third challenge discussed in Sec. 3.1, it is necessary to establish theoretical results for all the graphs in the recursive process with local transformation introduced. We will show in Sec. 3.3 that the four conditions in Def. 1 are fulfilled for each graph $\mathbb{M}$ that is obtained from valid local transformations and updated with the proposed rules in the process. Hence we establish the graphical characterization of valid local transformation for the PMGs compatible with local transformation in the following.

Then, we introduce the concept *bridged* in Def. 2. It plays an important role in the following results.

Definition 2 (Bridged relative to $\mathbf{V}'$ in H, Wang et al. 2023b)**.** Let H be a partial mixed graph. Let G denote 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 the circle component of G are *bridged relative to* $\mathbf{V}'$ if in each minimal circle path from A to B in G as $V_0(= A) \circ\!\!-\!\!\circ V_1 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ V_n(= B)$, there exists one vertex $V_s, 0 \leq s \leq n$, such that $\mathcal{F}_i \subseteq \mathcal{F}_{i+1}, 0 \leq i \leq s-1$ and $\mathcal{F}_{i+1} \subseteq \mathcal{F}_i, s \leq i \leq n-1$, where $\mathcal{F}_i = \{V \in \mathbf{V}' \mid V \ast\!\!-\!\!\circ V_i \text{ in } H\}$. Evidently, both case $A = B$ and case that A and B are not connected in the circle component are the trivial cases that A and B in G are bridged relative to $\mathbf{V}'$. Further, G is *bridged relative to* $\mathbf{V}'$ *in* H if any two vertices in the circle component of G are bridged relative to $\mathbf{V}'$.



(a) (b)

Figure 3: Two unbridged cases. In (a), C_1 is not adjacent to V_{i+1} and C_2 is not adjacenct to V_i. In (b), C_2 could be C_1, $n \geq 2$, C_1 is not adjacent to $V_1, \cdots, V_n$ and C_2 is not adjacent to $V_0, \cdots, V_{n-1}$.

Remark 2. Def. 2 is somewhat complicated. Intuitively, it implies a structure that we can transform all the edges $C \ast\!\!-\!\!\circ V$ to $C \ast\!\!\rightarrow V$ for any $C \in \mathbf{V}'$ and $V \in \mathbf{V}(G)$ without generating new unshielded colliders or violating ancestral property. Fig. 3 shows two unbridged cases where $\mathbf{V}' = \{C_1, C_2\}$. If we transform $C_1 \ast\!\!\rightarrow V_i$ and $C_2 \ast\!\!\rightarrow V_{i+1}$ in case (a), no matter how we transform the circle edge $V_i \circ\!\!-\!\!\circ V_{i+1}$, there will be a new unshielded collider or an arrowhead into an undirected edge, hence the transformation is not valid. Similarly, there will be a new unshielded collider or an arrowhead into an undirected edge if we orient $C_1 \ast\!\!\rightarrow V_0$ and $C_2 \ast\!\!\rightarrow V_n$ in case (b).

Then we provide the graphical characterization for validity of a local transformation represented by $\mathbf{C}$ in Thm. 1.

Theorem 1. Suppose $\mathbb{M}$ a PMG compatible with local transformation. Given a set of vertices $\mathbf{C} \subseteq \{V | X \circ\!\!-\!\!\ast V \text{ in } \mathbb{M}\}$, let $\mathbf{Z} = \{V \in \mathbf{V}(\mathbb{M}) | V = X$, or there is $V -\!\!\circ \cdots -\!\!\circ V' -\!\!\circ X$ in $\mathbb{M}$ and $V' \notin \mathbf{C}\}$. There exists a MAG $\mathcal{M}$ consistent with $\mathbb{M}$ with $X \leftarrow\!\!\ast V$ for $\forall V \in \mathbf{C}$ and $X -\!\!\ast V$ for $\forall V \in \{V \mid X \circ\!\!-\!\!\ast V \text{ in } \mathbb{M}\} \backslash \mathbf{C}$ if and only if

(1) $\mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \cap \mathrm{Pa}(\mathbf{C}, \mathbb{M}) = \emptyset$;

(2) the subgraph $\mathbb{M}[\mathbf{C}]$ of $\mathbb{M}$ induced by $\mathbf{C}$ is a complete graph;

(3) $\mathbb{M}[\mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \backslash \mathbf{Z}]$ is bridged relative to $\mathbf{C} \cup \mathbf{Z}$ in $\mathbb{M}$;

(4) Either $\mathbf{Z} \backslash \{X\}$ or $\{V \in \mathbf{V}(\mathbb{M}) | V \ast\!\!\rightarrow X \text{ in } \mathbb{M}$ or $V \in \mathbf{C}\}$ is empty.

For a given $\mathbf{C}$, the complexity of evaluating conditions (1), (2), (4) of Thm. 1 is $\mathcal{O}(d^3)$, where d denotes the number of vertices. To evaluate condition (3), we propose Alg. 1, with soundness guarantee in Prop 1. The complexity of implementing Alg. 1 is also $\mathcal{O}(d^3)$. Hence we can evaluate the validity of a local transformation by Thm. 1 in $\mathcal{O}(d^3)$.

Proposition 1. *Suppose a PMG $\mathbb{M}$ compatible with local transformation. Given a set of vertices $\mathbf{C} \subseteq \{V | X \circ\!\!-\!\!\ast V \text{ in } \mathbb{M}\}$, let $\mathbf{Z} = \{V \in \mathbf{V}(\mathbb{M}) | V = X$, or there is $V -\!\!\circ \cdots -\!\!\circ V' -\!\!\circ X$ in $\mathbb{M}$ and $V' \notin \mathbf{C}\}$. Alg. 1 is valid to evaluate*

Algorithm 1 Evaluating the third condition of Thm. 1

Require: A PMG $\mathbb{M}$, local transformation $\mathbf{C}$

$\mathbf{Z} \leftarrow \{V \in \mathbf{V}(\mathbb{M}) | V = X$, or there is $V \multimap \cdots \multimap V' \multimap X$ in $\mathbb{M}$ and $V' \notin \mathbf{C}\}$;

For circle component in $\mathbb{M}[\text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}]$, transform as follows until no feasible updates: for any two vertices V_l and V_j such that $V_l \circ\!\!-\!\!* V_j$, orient it as $V_l -\!\!* V_j$ if **(i)** $\mathcal{F}_l \backslash \mathcal{F}_j \neq \emptyset$ or **(ii)** $\mathcal{F}_l = \mathcal{F}_j$ as well as there is a vertex $V_m \in \text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}$ not adjacent to V_j such that $V_m -\!\!* V_l \circ\!\!-\!\!* V_j$, where $\mathcal{F}_i = \{V \in \mathbf{C} \cup \mathbf{Z} \mid V *\!\!-\!\!\circ V_i \text{ in } \mathbb{M}\}, i = j, l, m$;

if in $\mathbb{M}[\text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}]$ there is an undirected edge **then**

  Return(Bridged = False)

else

  Return(Bridged = True)

end if

whether $\mathbb{M}[\text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}]$ *is bridged relative to* $\mathbf{C} \cup \mathbf{Z}$ *in* $\mathbb{M}$.

3.3. Sound and Complete Rules

In this part, we provide sound and complete rules to incorporate valid local transformations to a PMG compatible with local transformation, through which we can reveal all the causal relations that are identifiable given the information of local transformations. Suppose we incorporate the local transformations of $V_1, V_2, \cdots, V_k$ to $\mathcal{P}$.

We start by presenting sound and complete orientation rules, which include existing rules and a newly proposed rule. Ali et al. (2005); Zhang (2008) provided ten rules $\mathcal{R}_1 - \mathcal{R}_{10}$ to identify a PAG with observational data. Wang et al. (2023b) proposed a replacement rule $\mathcal{R}'_4$ of $\mathcal{R}_4$ when incorporating *local causal knowledge* in the absence of selection variable. Due to space limit, we show them in Appendix A. To accommodate the presence of selection variable, we introduce an additional rule $\mathcal{R}_{11}$. We present Prop. 2 to imply the soundness of $\mathcal{R}_{11}$. In the following, when we say the *proposed rules*, they refer to $\mathcal{R}_1 - \mathcal{R}_3, \mathcal{R}'_4, \mathcal{R}_5 - \mathcal{R}_{11}$. The soundness of these rules directly follows Zhang (2008) and Prop. 2.

$$\mathcal{R}_{11}\text{: If } A *\!\!\rightarrow B \multimap R\text{, then orient } B \multimap R \text{ as } B \rightarrow R.$$

Proposition 2. *R_{11} is sound to orient a PMG when incorporating the additional causal knowledge.*

The remaining part of this section is to prove the completeness of the proposed rules. The whole proof consists of two parts. In the first part, we present an algorithm to incorporate the causal knowledge of local transformations into a PAG and prove that the algorithm is sound and complete. In the second part, we prove that the proposed rules can orient the identical graph with the algorithm in the first part, thus the proposed rules are sound and complete.

Algorithm 2 Updating a PMG with a valid local transformation of X represented by $\mathbf{C}$

Require: A PMG $\mathbb{M}_i$, valid local transformation $\mathbf{C}$ of X

$\mathbf{Z} \leftarrow \{V \in \mathbf{V}(\mathbb{M}_i) | V = X$, or there is $V \multimap \cdots \multimap V' \multimap X$ in $\mathbb{M}_i$ and $V' \notin \mathbf{C}\}$;

For any $K \in \text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])$ and any $T \in \mathbf{C}$ such that $K \circ\!\!-\!\!* T$ in $\mathbb{M}_i$, orient $K \leftarrow\!\!* T$ (the mark at T remains); for any edges between $Z_1, Z_2 \in \mathbf{Z}$ with circles, orient the circles into tails; for all $K \in \text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])\backslash\mathbf{Z}$ and $Z \in \mathbf{Z}$ such that $Z \circ\!\!-\!\!* K$, orient $Z -\!\!* K$;

For circle component in $\mathbb{M}_i[\text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])\backslash\mathbf{Z}]$, transform as follows until no feasible updates: for any two vertices V_l and V_j such that $V_l \circ\!\!-\!\!* V_j$, orient it as $V_l -\!\!* V_j$ if **(i)** $\mathcal{F}_l \backslash \mathcal{F}_j \neq \emptyset$ or **(ii)** $\mathcal{F}_l = \mathcal{F}_j$ as well as there is a vertex $V_m \in \text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])\backslash\mathbf{Z}$ not adjacent to V_j such that $V_m -\!\!* V_l \circ\!\!-\!\!* V_j$;

Apply the proposed rules until the graph is closed under the proposed rules.

Ensure: Updated graph $\mathbb{M}_{i+1}$

We first present the sound and complete algorithm as follows. It is a recursive algorithm starting from the PMG $\mathbb{M}_0 (= \mathcal{P})$, and in the i-th round, $i = 1, 2, \cdots, k$, we incorporate the local transformation of V_i and obtain an updated graph $\mathbb{M}_i$ based on $\mathbb{M}_{i-1}$ by Alg. 2. To distinguish the vertex under local transformation and the other vertices, denote V_i in the i-th round by X. We say a PMG is *closed under the proposed rules* if no rules can be triggered to further orient the PMG. Next, we give a vital induction result in Thm. 2.

Theorem 2. Suppose $\mathbb{M}_i, 0 \leq i < k$ is a PMG compatible with local transformations. When we incorporate a valid local transformation of $X = V_{i+1}$ represented by $\mathbf{C}$ into $\mathbb{M}_i$ by Alg. 2 and obtain $\mathbb{M}_{i+1}$, (1) the arrowheads and tails in $\mathbb{M}_{i+1}$ are invariant in all the MAGs consistent with $\mathbb{M}_i$ and the local transformation of X represented by $\mathbf{C}$, and (2) $\mathbb{M}_{i+1}$ is also a PMG compatible with local transformation.

On one hand, Thm. 2 implies the soundness of Alg. 2 to incorporate any valid local transformations. More importantly, Thm. 2 provides an induction result that the property of compatibility with local transformation remains in the process of recursively incorporating valid local transformations by Alg. 2, which includes the complete property as shown in Def. 1. As $\mathbb{M}_0$ is a PMG compatible with local transformation, we conclude the desired result in Cor. 1.

Corollary 1. The recursive k-step algorithm based on Alg. 2 is sound and complete to incorporate the local transformation of $V_1, \cdots, V_k$ into a PAG. And $\mathbb{M}_i, \forall 0 \leq i \leq k$ is a PMG compatible with local transformation.

Now we are ready to present Thm. 3, which implies that the proposed rules are sound and complete to incorporate local transformations of $V_1, \cdots, V_k$. We prove it by showing that the proposed orientation rules can transform the same marks as the complete k-step recursive algorithm. Hence, under each valid local transformation, we can reveal all the causal relations that are identifiable by using the proposed rules.

Theorem 3. The orientation rules are sound and complete to transform a PAG with the local transformation of $V_1, \ldots, V_k$.

Remark 3. The proposed orientation rules are not only effective in the MAG listing task, but also applicable when we incorporate local *background knowledge* into a PAG. In causality literature, background knowledge (BK) refers to additional structural information. BK can be obtained from human expertise or interventional data. The sound and complete rules to incorporate BK into a CPDAG have been established for a long time (Meek, 1995). But it is still an open problem for the cases in the presence of latent variables. The local BK is proposed by Wang et al. (2023b), which is a special kind of BK that can imply *all* the structural knowledge at some vertex X if it implies some additional structural knowledge at X. The structural knowledge taken by the local transformation is evidently local BK since the local transformation implies all the marks at some specific vertices. Wang et al. (2023b) presented the sound and complete rules to incorporate local BK with the assumption of no selection variable. Another related study is by Andrews et al. (2020), which assumed no selection variable as well and showed that the ten rules of Zhang (2008) are complete when the BK is *tiered*, which means that the BK can divide the vertices into several disjoint parts where the causal order between the different parts is clear but the structure knowledge in each part is unknown according to the BK.

3.4. MAG Listing Algorithm

Combining the two parts above, we present MAGLIST algorithm in Alg. 3. It is a recursive algorithm. In each round, we select a vertex X with circles (on Line 7). For each possible local transformation of X (on line 8), we evaluate whether it is valid by Thm. 1, and further update the PMG under each valid local transformation with the sound and complete proposed rules (on Line 10). If we obtain a PMG without circles in the process above, it is a MAG in the MEC represented by $\mathcal{P}$ (on Line 4).

An implementation example in the format of a search tree is shown in Fig. 2. The graph in the root node denotes a PAG $\mathcal{P}$ and we aim to list all the MAGs consistent with $\mathcal{P}$. In the first round, we consider the local transformation of A. According to Thm. 1, we can determine that there are six valid local transformations of A, and thus obtain six PMGs on the second line according to the marks implied by the local transformations of A. Then we update these six graphs using the proposed rules and obtain the graphs on the third line. After the local transformation and the updates with rules, the implementation in the first round completes. In the second round, we further consider the local transformation of B, and based on the valid local transformation of B, we update the graph with the proposed rules. Due to the space limit, we no longer separately depict the stages of local transformation and update with rules, and directly show the graphs after the second round on the fourth line. There are some PMGs without any circles, which are MAGs consistent with $\mathcal{P}$. We shade these graphs with green color. For the unshaded graphs, they are updated in the third round by considering the transformation of C. We omit some branches (those unshaded but unexpanded) for brevity. The algorithm stops until there are no new unshaded leafs.

Algorithm 3 MAGLIST

Require: A PAG $\mathcal{P}$
1: $\mathcal{S} = \emptyset$ ▷ Record all the MAGs consistent with $\mathcal{P}$
2: ORIENTGRAPH($\mathcal{P}, \mathcal{S}$)
3: **function** ORIENTGRAPH($\mathbb{M}, \mathcal{S}$)
4: **if** there are no circles in $\mathbb{M}$ **then**
5: $\mathcal{S} \leftarrow \mathcal{S} \cup \{\mathbb{M}\}$
6: **else**
7: Select a variable X where there are circles in $\mathbb{M}$
8: **for** each $\mathbf{C} \subseteq \{V \in \mathbf{V}(\mathbb{M}) \mid X \circ\!\!-\!\!* V \text{ in } \mathbb{M}\}$ **do**
9: **if** the four conditions in Thm. 1 hold **then**
10: Update $\mathbb{M}$ with the local transformation represented by $\mathbf{C}$ and apply the proposed rules
11: ORIENTGRAPH($\mathbb{M}, \mathcal{S}$)
12: **end if**
13: **end for**
14: **end if**
15: **end function**
Ensure: $\mathcal{S}$

Finally, we present Cor. 2 to imply that Alg. 3 can output all and only the MAGs in the MEC represented by the PAG $\mathcal{P}$. One main limitation in the current is that, there lacks of a theoretical analysis of the time complexity of Alg. 3. It is closely related to the number of MAGs in a MEC represented by the PAG, but it remains an open problem in the literature to determine how many MAGs there are in a MEC. Many solid results have been established for the cases absence of latent variables (He et al., 2015; Wienöbst et al., 2021; Ganian et al., 2022), but the results for the cases with latent variables need to be further studied. Instead, we conduct an empirical analysis in Sec. 5.

Corollary 2. Algorithm 3 is valid to list all the MAGs consistent with $\mathcal{P}$.

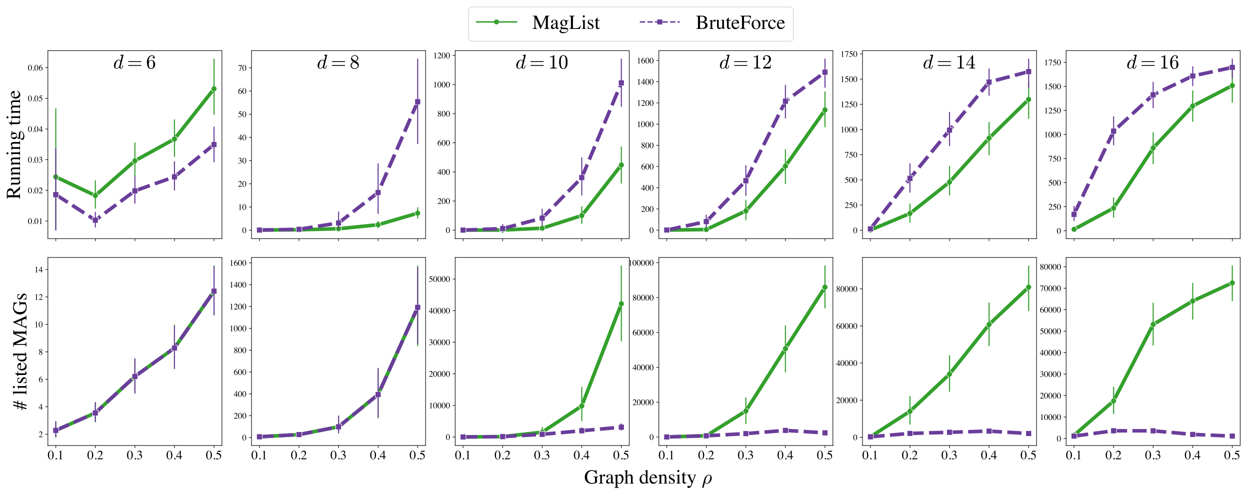


Figure 4: The running time and number of listed MAGs within 1800 seconds for MAGLIST and BRUTEFORCE in 100 simulations for each combination of the number of vertice $d \in \{6, 8, 10, 12, 14, 16\}$ including 3 latent variables and the probability of an edge between two vertices $\rho \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$. The vertical line represents the 95% confidence interval generated by bootstrap sampling. It is determined by the 2.5th and 97.5th percentiles of 1000 estimates from the random sample of equal size with replacement from the original sample.

4. Related Works

How to deal with latent variables is a crucial problem in many causality tasks. Causal discovery is one of these tasks. Numerous studies focus on causal discovery using observational data (Ali et al., 2005; Zhang, 2008). Without some further assumptions (Zhang et al., 2017; Cai et al., 2018), some causal relations are unidentifiable. To further uncover these relations, some studies additionally utilize interventional data (Yu et al., 2019; Wang et al., 2020; Jaber et al., 2020; Wang & Zhou, 2021).

The existence of latent variables takes a challenge to causal effect estimation task as well. There are many efforts on causal effect identifiability (Tian & Pearl, 2002; Shpitser & Pearl, 2006; Perkovic et al., 2017; Lee & Bareinboim, 2020), studying whether the causal effect is identifiable given a causal graph. With observational data, only a PAG is identifiable, which generally fails to identify the causal effect. To obtain the information of causal effect in this case, a series of studies are conducted to determine a set of possible causal effects via covariate adjustment (Malinsky & Spirtes, 2016; Wang et al., 2023a). However, for these methods, it is possible that the true causal effect cannot be included in the determined set, because adjustment set is not sufficient for identifying causal effects in the presence of latent variables. In contrast to these methods, Li & Pearl (2022) provided a method to determine an accurate causal effect bound for some cases. Further, Jiang et al. (2023) proposed an efficient method to approximate causal effect bounds, given the entropy of latent confounder as the available side information of latent confounder. Another kind of method for dealing with latent variables in causal effect estimation tasks is using instrumental/proximal variables (Baiocchi et al., 2014; Sverdrup & Cui, 2023). Li et al. (2024) studied how to perform nonparametric IV regression and model selection without a minimax oracle.

Recently, the integration of causality into traditional machine learning tasks attracted tremendous attention. Wang et al. (2022) provides thoughtful insights on the link between causality and robust prediction, showcasing how causal analysis can substantially enhance the performance of state-of-the-art machine learning algorithms. There are also some efforts on applying causality in decision-making (Ruan et al., 2023). Compared to traditional MDP-based methods (Sutton & Barto, 2018; Chen et al., 2021; Zhao et al., 2022; Chen et al., 2023; Jia et al., 2024) or treatment effect methods (Qin et al., 2021; Kallus, 2023), causality provides structural information. Recently, Zhou (2022; 2023) proposed the viewpoint that *influence relation*, which is a relation among *correlation* and *causation*, is more suitable for a large amount of decision problems in practice, which offers a fresh view for the study of decision methods. Following this viewpoint, Qin et al. (2023) presented a graphical model that characterizes influence relations, along with a Bayesian method for decision-making based on the graphical model.

5. Experiments

In this part, we present the empirical analysis for MAG listing tasks. We evaluate the effectiveness and efficiency of the proposed MAGLIST algorithm.

We generate simulated PAGs and record the running time for listing all the MAGs in the MEC represented by each PAG. There are two parameters here: the number of vertices $d \in \{6, 8, 10, 12, 14, 16\}$ and the probability of an edge between any two vertices $p \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$. For each parameter combination, we generate 100 Erdös-Rényi graph as the true DAGs. In each DAG, we randomly select three vertices as latent variables and then obtain the true MAG with $d - 3$ vertices based on the DAG. See Zhang (2008) for the algorithm of obtaining a MAG based on a DAG. Since the focus of this paper is not on the PAG learning, we directly use the true PAG obtained from the true MAG as the input of MAGLIST. We record the number of listed MAGs and the time spent for each PAG.

In the literature, there are no MAG listing methods. Hence we compare MAGLIST with the baseline BRUTEFORCE method, which is detailed in Appendix B. Note that BRUTEFORCE spends extremely large amount of time in most of cases. Hence, we set the maximal running time for MAG listing given each PAG by 1800 seconds. For both MAGLIST and BRUTEFORCE, when the running time exceeds this limit, the implementation will stop immediately and they output the MAGs that have been listed. The experimental results are shown in Fig. 4. When d and ρ are not large, the two methods can find the same number of MAGs. It verifies the effectiveness of MAGLIST. When d and ρ are relatively large, the number of MAGs output by MAGLIST is far more than that by BRUTEFORCE. Note it is a normal phenomenon that for BRUTEFORCE, the number of listed MAGs could become less as the graph density increases. The reason is, as d and ρ increase, the possibility of an enumerated mixed graph not being a MAG consistent with the PAG $\mathcal{P}$ increases as well. Hence, for BRUTEFORCE, most of the enumerated graph are not MAGs consistent with $\mathcal{P}$, resulting in the small number of MAGs output.

6. Conclusion

In this paper, we propose the first brute-force-free MAG listing algorithm in the presence of both latent confounders and selection variables. Our approach hinges on providing a graphical characterization for valid local transformation, and establishing sound and complete orientation rules to incorporate the structural knowledge implied by local transformation. The rules can also be applied when we introduce local BK attained from human expertise to a PAG. The experiments demonstrate the effectiveness and efficiency of the MAG listing method.

For future study, an important problem to be addressed is on the theoretical analysis of the number of MAGs in a given MEC. We believe such an analysis is vital for the further development of the relevant MAG listing methods. And another important theoretical problem is the establishment of the sound and complete orientation rules to incorporate causal background knowledge into a PAG in general. For our method, in each round, we need to consider the possible local transformations of a vertex, which takes a $\mathcal{O}(2^p)$ complexity, where p denotes the number of circles at this vertex. The establishment of the sound and complete rules for incorporating causal knowledge may further help reduce the complexity of this part, which could possibly inspire a more efficient MAG listing algorithm as the DAG listing method proposed by Wienöbst et al. (2023).

Acknowledgements

This research was supported by Leading-edge Technology Program of Jiangsu Science Foundation (BK20232003). We thank Tian Qin and Lue Tao for insightful discussions, and thank reviewers for their valuable feedback.

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

References

Ali, A. R., Richardson, T. S., Spirtes, P., and Zhang, J. Towards characterizing markov equivalence classes for directed acyclic graphs with latent variables. In *Proceedings of the 21st Conference in Uncertainty in Artificial Intelligence*, pp. 10–17, 2005.

Ali, R. A., Richardson, T. S., Spirtes, P., et al. Markov equivalence for ancestral graphs. *The Annals of Statistics*, 37(5B):2808–2837, 2009.

Andrews, B., Spirtes, P., and Cooper, G. F. 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*, pp. 4002–4011, 2020.

Baiocchi, M., Cheng, J., and Small, D. S. Instrumental variable methods for causal inference. *Statistics in medicine*, 33(13):2297–2340, 2014.

Cai, R., Qiao, J., Zhang, K., Zhang, Z., and Hao, Z. Causal discovery from discrete data using hidden compact representation. In *Advances in Neural Information Processing Systems 31*, pp. 2666–2674, 2018.

Chandler Squires. `causaldag`*: creation, manipulation, and learning of causal models*, 2018.

Chen, E. Y., Choi, A., and Darwiche, A. Enumerating equivalence classes of bayesian networks using EC graphs.

In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, pp. 591–599, 2016.

Chen, X., Yu, Y., Li, Q., Luo, F., Qin, Z. T., Shang, W., and Ye, J. Offline model-based adaptable policy learning. In *Advances in Neural Information Processing Systems 34*, pp. 8432–8443, 2021.

Chen, X., Luo, F., Yu, Y., Li, Q., Qin, Z., Shang, W., and Ye, J. Offline model-based adaptable policy learning for decision-making in out-of-support regions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45 (12):15260–15274, 2023.

Chickering, D. M. A transformational characterization of equivalent bayesian network structures. In *Proceedings of the 11st Annual Conference on Uncertainty in Artificial Intelligence*, pp. 87–98, 1995.

Ganian, R., Hamm, T., and Talvitie, T. An efficient algorithm for counting markov equivalent dags. *Artificial Intelligence*, 304:103648, 2022.

Hauser, A. and Bühlmann, P. Two optimal strategies for active learning of causal models from interventional data. *International Journal of Approximate Reasoning*, 55(4): 926–939, 2014.

He, Y. and Geng, Z. Active learning of causal networks with intervention experiments and optimal designs. *Journal of Machine Learning Research*, 9:2523–2547, 2008.

He, Y., Jia, J., and Yu, B. Counting and exploring sizes of markov equivalence classes of directed acyclic graphs. *Journal of Machine Learning Research*, 16:2589–2609, 2015.

Hu, Z. and Evans, R. Faster algorithms for markov equivalence. In *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence*, pp. 739–748, 2020.

Hünermund, P. and Bareinboim, E. Causal inference and data fusion in econometrics. *The Econometrics Journal*, pp. utad008, 2023.

Jaber, A., Kocaoglu, M., Shanmugam, K., and Bareinboim, E. Causal discovery from soft interventions with unknown targets: Characterization and learning. In *Advances in Neural Information Processing Systems 33*, pp. 9551–9561, 2020.

Jia, C., Gao, C., Yin, H., Zhang, F., Chen, X.-H., Xu, T., Yuan, L., Zhang, Z., Zhou, Z.-H., and Yu, Y. Policy rehearsing: Training generalizable policies for reinforcement learning. In *The 12th International Conference on Learning Representations, pages to appear*, 2024.

Jiang, Z., Wei, L., and Kocaoglu, M. Approximate causal effect identification under weak confounding. In *Proceedings of the 40th International Conference on Machine Learning*, pp. 15125–15143, 2023.

Kalisch, M., Mächler, M., Colombo, D., Maathuis, M. H., and Bühlmann, P. Causal inference using graphical models with the R package pcalg. *Journal of Statistical Software*, 47(11):1–26, 2012.

Kallus, N. Treatment effect risk: Bounds and inference. *Management Science*, 69(8):4579–4590, 2023.

Kocaoglu, M. Characterization and learning of causal graphs with small conditioning sets. In *Advances in Neural Information Processing Systems*, 2023.

Lee, S. and Bareinboim, E. Causal effect identifiability under partial-observability. In *Proceedings of the 37th International Conference on Machine Learning*, pp. 5692–5701, 2020.

Li, A. and Pearl, J. Bounds on causal effects and application to high dimensional data. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence*, pp. 5773–5780, 2022.

Li, Z., Lan, H., Syrgkanis, V., Wang, M., and Uehara, M. Regularized deepiv with model selection. *arXiv preprint arXiv:2403.04236*, 2024.

Maathuis, M. H., Kalisch, M., Bühlmann, P., et al. Estimating high-dimensional intervention effects from observational data. *The Annals of Statistics*, 37(6A):3133–3164, 2009.

Malinsky, D. and Spirtes, P. Estimating causal effects with ancestral graph markov models. In *Conference on Probabilistic Graphical Models*, pp. 299–309, 2016.

Meek, C. Causal inference and causal explanation with background knowledge. In *Proceedings of the 11st Annual Conference on Uncertainty in Artificial Intelligence*, pp. 403–410, 1995.

Pearl, J. *Causality*. Cambridge University Press, 2009.

Perkovic, E., Textor, J., Kalisch, M., and Maathuis, M. H. Complete graphical characterization and construction of adjustment sets in markov equivalence classes of ancestral graphs. *Journal of Machine Learning Research*, 18:220:1–220:62, 2017.

Qin, T., Wang, T.-Z., and Zhou, Z.-H. Budgeted heterogeneous treatment effect estimation. In *Proceedings of the 38th International Conference on Machine Learning*, pp. 8693–8702, 2021.

Qin, T., Wang, T.-Z., and Zhou, Z.-H. Rehearsal learning for avoiding undesired future. In *Advances in Neural Information Processing Systems 36, pages to appear*, 2023.

Richardson, T., Spirtes, P., et al. Ancestral graph markov models. *The Annals of Statistics*, 30(4):962–1030, 2002.

Ruan, K., Zhang, J., Di, X., and Bareinboim, E. Causal imitation learning via inverse reinforcement learning. In *The 11st International Conference on Learning Representations*, 2023.

Shpitser, I. and Pearl, J. Identification of joint interventional distributions in recursive semi-markovian causal models. In *Proceedings of the 21st AAAI Conference on Artificial Intelligence*, pp. 1219–1226, 2006.

Spirtes, P. and Richardson, T. A polynomial time algorithm for determining dag equivalence in the presence of latent variables and selection bias. In *Proceedings of the 6th International Workshop on Artificial Intelligence and Statistics*, pp. 489–500, 1996.

Spirtes, P., Glymour, C. N., and Scheines, R. *Causation, prediction, and search*. MIT Press, 2000.

Sutton, R. S. and Barto, A. G. *Reinforcement learning: An introduction*. MIT press, 2018.

Sverdrup, E. and Cui, Y. Proximal causal learning of conditional average treatment effects. In *Proceedings of the 40th International Conference on Machine Learning*, pp. 33285–33298. PMLR, 2023.

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

Tian, J. and Pearl, J. A general identification condition for causal effects. In *Proceedings of the 18th AAAI Conference on Artificial Intelligence*, pp. 567–573, 2002.

Wang, R., Yi, M., Chen, Z., and Zhu, S. Out-of-distribution generalization with causal invariant transformations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 375–385, 2022.

Wang, T.-Z. and Zhou, Z.-H. Actively identifying causal effects with latent variables given only response variable observable. In *Advances in Neural Information Processing Systems 34*, pp. 15007–15018, 2021.

Wang, T.-Z., Wu, X.-Z., Huang, S.-J., and Zhou, Z.-H. Cost-effectively identifying causal effects when only response variable is observable. In *Proceedings of the 37th International Conference on Machine Learning*, pp. 10060–10069, 2020.

Wang, T.-Z., Qin, T., and Zhou, Z.-H. Estimating possible causal effects with latent variables via adjustment. In *Proceedings of the 40th International Conference on Machine Learning*, pp. 36308–36335, 2023a.

Wang, T.-Z., Qin, T., and Zhou, Z.-H. Sound and complete causal identification with latent variables given local background knowledge. *Artificial Intelligence*, 322:103964, 2023b.

Wienöbst, M., Bannach, M., and Liskiewicz, M. Polynomial-time algorithms for counting and sampling markov equivalent dags. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence*, pp. 12198–12206, 2021.

Wienöbst, M., Luttermann, M., Bannach, M., and Liskiewicz, M. Efficient enumeration of markov equivalent dags. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence*, pp. 12313–12320, 2023.

Yu, K., Liu, L., and Li, J. Learning markov blankets from multiple interventional data sets. *IEEE transactions on neural networks and learning systems*, 31(6):2005–2019, 2019.

Zhang, J. On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias. *Artificial Intelligence*, 172(16-17): 1873–1896, 2008.

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

Zhang, K., Huang, B., Zhang, J., Glymour, C., and Schölkopf, B. Causal discovery from nonstationary/heterogeneous data: Skeleton estimation and orientation determination. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pp. 1347–1353, 2017.

Zhao, P., Li, L.-F., and Zhou, Z.-H. Dynamic regret of online markov decision processes. In *Proceedings of the 39th International Conference on Machine Learning*, pp. 26865–26894, 2022.

Zhou, Z.-H. Rehearsal: learning from prediction to decision. *Frontiers of Computer Science*, 16(4):164352, 2022.

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

A. Orientation Rules with Observational Data

In this section, we show the complete orientation rules proposed by Zhang (2008) for causal discovery with observational data in the presence of latent confounders and selection variables. There are eleven rules ($\mathcal{R}_0 - \mathcal{R}_{10}$). $\mathcal{R}_0$ is triggered according to the conditional independence relationship at the beginning of learning a PAG and is not triggered beyond this phase, hence we do not show it.

$\mathcal{R}_1$: If $A{*\!\!\rightarrow} B \circ\!\!-\!\!* R$, and A and R are not adjacent, then orient the triple as $A{*\!\!\rightarrow} B \rightarrow R$.

$\mathcal{R}_2$: If $A \rightarrow B{*\!\!\rightarrow} R$ or $A{*\!\!\rightarrow} B \rightarrow R$, and $A *\!\!-\!\!\circ R$, then orient $A *\!\!-\!\!\circ R$ as $A{*\!\!\rightarrow} R$.

$\mathcal{R}_3$: If $A{*\!\!\rightarrow} B \leftarrow\!\!* R$, $A *\!\!-\!\!\circ D \circ\!\!-\!\!* R$, A and R are not adjacent, and $D *\!\!-\!\!\circ B$, then orient $D *\!\!-\!\!\circ B$ as $D{*\!\!\rightarrow} B$.

$\mathcal{R}_4$: If $\langle K, \dots, A, B, R\rangle$ is a discriminating path between K and R for B, and $B \circ\!\!-\!\!* R$; then if $B \in$ Sepset(K, R), orient $B \circ\!\!-\!\!* R$ as $B \rightarrow R$; otherwise orient the triple $\langle A, B, R\rangle$ as $A \leftrightarrow B \leftrightarrow R$.

$\mathcal{R}_5$: For every (remaining) $A \circ\!\!-\!\!\circ R$, if there is an uncovered circle path $p = \langle A, B, \cdots, D, R\rangle$ between A and R *s.t.* A, D are not adjacent and B, R are not adjacent, then orient $A \circ\!\!-\!\!\circ R$ and every edge on p as undirected edges.

$\mathcal{R}_6$: If $A - B \circ\!\!-\!\!* R$ (A and R may or may not be adjacent), then orient $B \circ\!\!-\!\!* R$ as $B -\!\!* R$.

$\mathcal{R}_7$: If $A -\!\!\circ B \circ\!\!-\!\!* R$, and A, R are not adjacent, then orient $B \circ\!\!-\!\!* R$ as $B -\!\!* R$.

$\mathcal{R}_8$: If $A \rightarrow B \rightarrow R$, and $A \circ\!\!\rightarrow R$, orient $A \circ\!\!\rightarrow R$ as $A \rightarrow R$.

$\mathcal{R}_9$: If $A \circ\!\!\rightarrow R$, and $p = \langle A, B, D, \dots, R\rangle$ is an uncovered possible directed path from A to R such that R and B are not adjacent, then orient $A \circ\!\!\rightarrow R$ as $A \rightarrow R$.

$\mathcal{R}_{10}$: Suppose $A \circ\!\!\rightarrow R$, $B \rightarrow R \leftarrow D$, p_1 is an uncovered possible directed path from A to B, and p_2 is an uncovered possible directed path from A to D. Let U be the vertex adjacent to A on p_1 (U could be B), and W be the vertex adjacent to A on p_2 (W could be D). If U and W are distinct, and are not adjacent, then orient $A \circ\!\!\rightarrow R$ as $A \rightarrow R$.

Recently, Wang et al. (2023b) proposed that when we have identified a PAG and we incorporate the local background knowledge, $\mathcal{R}_4$ is replaced by $\mathcal{R}_4'$. And the ten rules $\mathcal{R}_1 - \mathcal{R}_3, \mathcal{R}_4', \mathcal{R}_5 - \mathcal{R}_{10}$ are sound and complete to incorporate the local background knowledge.

$\mathcal{R}_4'$: If $\langle K, \cdots, A, B, R\rangle$ is a discriminating path between K and R for B, and $B \circ\!\!-\!\!* R$, then orient $B \circ\!\!-\!\!* R$ as $B \rightarrow R$.

B. Brute-force MAG Listing

In this section, we present Alg. 4, a brute force MAG listing method, as a baseline. According to Thm. 2 of Zhang (2008), we can obtain a MAG $\mathcal{M}$ consistent with $\mathcal{P}$ on Line 2. On Line 3, we record the position of every circle. On Line 4, we enumerate all the mixed graphs by transforming all the circles. On Line 6, we determine whether the enumerated graph is a ancestral graph. A path p from X to Y in an ancestral graph G is an inducing path if every non-endpoint vertex on p is a collider and meanwhile an ancestor of either X or Y (Spirtes et al., 2000). It has been shown that an ancestral graph is maximal if and only if there is not an inducing path between any two non-adjacent vertices (Richardson et al., 2002). Hence, by evaluating the existence of inducing paths on Line 9, we can determine whether the enumerated graph is maximal. On Line 12, we evaluate whether the enumerated graph is a graph in the Markov equivalence class represented by the PAG $\mathcal{P}$.

C. Proofs

C.1. Proof of Theorem 1

In the proof of Thm. 1, given a PMG compatible with local transformation $\mathbb{M}$, when the four conditions are fulfilled for a local transformation represented by $\mathbf{C}$, we need to present a procedure to obtain a MAG $\mathcal{H}$ consistent with $\mathbb{M}$ and the local transformation represented by $\mathbf{C}$. Since some supporting results relevant to this procedure are needed, we present the procedure at first. We henceforth use the *procedure* to refer to the procedure of obtaining $\mathcal{H}$ based on $\mathbb{M}$ and $\mathbf{C}$.

Algorithm 4 BRUTEFORCE

Require: A PAG $\mathcal{P}$

$\mathcal{S} = \emptyset$ ▷ Record all the MAGs consistent with $\mathcal{P}$

Obtain a MAG $\mathcal{M}$ from $\mathcal{P}$ by transforming the circle component into a DAG without unshielded colliders and the edges $\circ\!\!\rightarrow$/$-\!\circ$ into $\rightarrow$

$C_{set} = \{(i, j) \mid \text{there is } V_i \ast\!\!-\!\circ V_j \text{ in } \mathcal{P}\}$ ▷ Record the indexes of all the circles in $\mathcal{P}$

for each set $\mathbf{I} \subseteq C_{set}$ **do**

  Obtain a graph G by transforming the circles in $\mathbf{I}$ to arrowheads and others (in $C_{set}\backslash\mathbf{I}$) to tails

  if there is a directed or almost directed cycle or edges into an undirected edge in G **then**

    continue ▷ It violates the ancestral property

  end if

  if there is an inducing path in G **then**

    continue ▷ It violates the maximal property

  end if

  if G is Markov equivalent to $\mathcal{M}$ by Thm. 3.7 of Ali et al. (2009) or Thm. 3.2 of Hu & Evans (2020) **then**

    $\mathcal{S} \leftarrow \mathcal{S} \cup \{G\}$ ▷ It is consistent with $\mathcal{P}$

  end if

end for

Ensure: $\mathcal{S}$

(Step 0) $\mathbf{Z} \leftarrow \{V \in \mathbf{V}(\mathbb{M}) | V = X$, or there is $V -\!\circ \cdots -\!\circ V' -\!\circ X$ in $\mathbb{M}$ and $V' \notin \mathbf{C}\}$;

(Step 1) for any $K \in$ PossDe$(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])$ and any $T \in \mathbf{C}$ such that $K \circ\!\!\ast T$ in $\mathbb{M}$, orient $K \leftarrow\!\!\ast T$ (the mark at T remains); for any edges between $Z_1, Z_2 \in \mathbf{Z}$ with circles, orient the circles into tails; for all $K \in$ PossDe$(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}$ and $Z \in \mathbf{Z}$ such that $Z \circ\!\!-\!\ast K$, orient $Z -\!\ast K$;

(Step 2) For circle component in $\mathbb{M}$[PossDe$(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}$], transform as follows until no feasible updates: for any two vertices V_l and V_j such that $V_l \circ\!\!-\!\ast V_j$, orient it as $V_l -\!\ast V_j$ if **(i)** $\mathcal{F}_l\backslash\mathcal{F}_j \neq \emptyset$ or **(ii)** $\mathcal{F}_l = \mathcal{F}_j$ as well as there is a vertex $V_m \in$ PossDe$(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}$ not adjacent to V_j such that $V_m -\!\ast V_l \circ\!\!-\!\ast V_j$, where $\mathcal{F}_i = \{V \in \mathbf{C} \cup \mathbf{Z} \mid V \ast\!\!-\!\circ V_i \text{ in } \mathbb{M}\}, i = j, k, l$;

(Step 3) for the circle component in subgraph $\mathbb{M}$[PossDe$(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}$], orient it into a DAG without new unshielded colliders;

(Step 4) for the circle component in subgraph $\mathbb{M}[-$PossDe$(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])]$, orient it into a DAG without new unshielded colliders;

(Step 5) transform edges $\circ\!\!\rightarrow$/$-\!\circ$ to $\rightarrow$.

Lemma 1. *Consider the PMG $\mathbb{M}$ compatible with local transformation. If there is $A \ast\!\!\rightarrow B$ in $\mathbb{M}$, then there is an edge as $A \ast\!\!\rightarrow V$ for any V in a connected circle component with B in $\mathbb{M}$, and A and B are not connected in a circle component.*

Proof. The proof of this part can directly refer to that of Lemma 3 of Wang et al. (2023b). □

Lemma 2. *Consider a PMG $\mathbb{M}$ compatible with local transformation. If there is an uncovered circle path $p = \langle F_1, F_2, \cdots, F_m\rangle, m \geq 3$ in $\mathbb{M}$, then it is minimal.*

Proof. Suppose it is not minimal, then there exists a sub-structure $F_j \circ\!\!-\!\!\circ F_{j+1} \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ F_k, k > j + 2$ where any non-consecutive vertices are not adjacent except for an edge between F_j and F_k. We consider the edge between F_j and F_k. It is not a circle edge, for otherwise the chordal property of $\mathbb{M}$ is not fulfilled. It is not an edge with an arrowhead, for otherwise Lemma 1 is violated. It is not an edge $F_j -\!\circ F_k$ or $F_j - F_k$, for otherwise the complete property of $\mathbb{M}$ is not fulfilled due to $\mathcal{R}_6, \mathcal{R}_7$ and the circle edge $F_k \circ\!\!-\!\!\circ F_{k-1}$. Hence any edge between F_j and F_k is invalid. The uncovered path is minimal. □

Lemma 3. *Consider a PMG $\mathbb{M}$ compatible with local transformation. 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 both $V_i \leftarrow\!\!* V_j$ and $V_i \circ\!\!- V_j$ in $\mathbb{M}$. The impossibility of $V_i \leftarrow\!\!* V_j$ has been proved by Lemma 2 of Wang et al. (2023b). We only consider $V_i \circ\!\!- V_j$. If $j-i=2$, we consider the local structure comprised of V_i, V_{i+1}, V_j. As p is possibly directed, for the edge between V_i and V_{i+1}, the mart at V_i is either circle or tail, and the mark at V_{i+1} is not tail. If it is a circle at V_i on the edge between V_i and V_{i+1}, there is $V_j -\!\!\circ V_i \circ\!\!-\!\!* V_{i+1}$. According to the balanced property of $\mathbb{M}$ and the fact that there is not a tail at V_{i+1} on the edge between V_i and V_{i+1}, there is either $V_j -\!\!\circ V_{i+1} \circ\!\!-\!\!\circ V_i$ or $V_j *\!\!\rightarrow V_{i+1} \leftarrow\!\!\circ V_i$, both of which contradicts with the fact that p is a possible directed path. If it is a tail at V_i on the edge between V_i and V_{i+1}, there must be $V_i -\!\!\circ V_{i+1} -\!\!\circ V_j -\!\!\circ V_i$, for otherwise the balanced property is not fulfilled. However, in this case the circle can only be tail, for otherwise the ancestral property cannot be fulfilled, which contradicts with the complete property of $\mathbb{M}$. Hence, for the edge between V_i and V_j, there cannot be an arrowhead at V_i or a tail at 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 4. *Consider a PMG $\mathbb{M}$ compatible with local transformation. Suppose a MAG $\mathcal{M}$ consistent with $\mathbb{M}$ and the local transformation of X represented by $\mathbf{C}$ which satisfies the fourth condition in Thm. 1. If $\mathbf{C}$ is not an empty set or there is some edge into X in $\mathbb{M}$, then $V \in \mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])$ if and only if $V \in \mathrm{De}(\mathbf{Z}, \mathcal{M})$.*

Proof. When $\mathbf{C}$ is not an empty set or there is some edge into X in $\mathbb{M}$, there is $\mathbf{Z} = \{X\}$ according to the fourth condition. Hence we will prove $V \in \mathrm{PossDe}(X, \mathbb{M}[-\mathbf{C}])$ if and only if $V \in \mathrm{De}(X, \mathcal{M})$ in the following.

We first prove the "only if" statement. If $V \in \mathrm{PossDe}(X, \mathbb{M}[-\mathbf{C}])$, there is a minimal possible directed path $p = \langle X, F_1, \ldots, F_m(=V)\rangle$ by Lemma 3. According to the condition, there is an edge $A *\!\!\rightarrow X$. Due to $F_1 \notin \mathbf{C}$, there is $X \rightarrow V_1$ or $X -\!\!\circ V_1$ in $\mathcal{M}$. For the latter case, as $A *\!\!\rightarrow X$, the edge between X and V_1 cannot be undirected in all the MAG consistent with $\mathbb{M}$ and the local transformation of X represented by $\mathbf{C}$ due to the ancestral property. Hence there is always $X \rightarrow V_1$. Hence p can only be directed in $\mathcal{M}$, otherwise there will be an edge into an undirected edge, which contradicts with the ancestral property; or there is an unshielded collider $F_{i-1} *\!\!\rightarrow F_i \leftarrow\!\!* F_{i+1}$ in $\mathcal{M}$, which has been identified in $\mathbb{M}$ and thus contradicts with the fact that p is a minimal possible directed path from X to F_m in $\mathbb{M}[-\mathbf{C}]$.

We then prove the "if" statement. According to the ancestral property, there must be a minimal directed path $X \rightarrow F_1 \cdots \rightarrow F_{m-1}, F_m(=V)$ in $\mathcal{M}$, where X is not adjacent to $F_2, \ldots, F_m$. The corresponding path in $\mathbb{M}$ of this path is a minimal possible directed path from X to V. If $V \notin \mathrm{PossDe}(X, \mathbb{M}[-\mathbf{C}])$, there can only be $F_1 \in \mathbf{C}$ due to $F_2, F_3, \ldots, F_m \notin \mathbf{C}$ as they are not adjacent to X. In this case $X \leftarrow\!\!* F_1$ should be represented by $\mathbf{C}$, which contradicts $X \rightarrow F_1$ in $\mathcal{M}$. The proof completes. □

Lemma 5. *Consider a PMG $\mathbb{M}$ compatible with local transformation and a local transformation of X represented by $\mathbf{C}$ which satisfies the fourth condition in Thm. 1. Denote $\mathcal{F}_V = \{T \in \mathbf{C} \cup \{X\} \mid T *\!\!-\!\!\circ V \text{ in } \mathbb{M}\}$ for $\forall V \in \mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}$. If for a circle edge $A \circ\!\!-\!\!\circ B$ in $\mathbb{M}[\mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}]$ such that $T \in \mathcal{F}_A\backslash\mathcal{F}_B$, then T is not adjacent to B in $\mathbb{M}$.*

Proof. Suppose T is adjacent to B. Consider the edge between B and T in $\mathbb{M}$. (1) The edge cannot be $T \leftarrow B$, for otherwise the ancestral property can never be fulfilled as T is a descendant of X but there is an edge $T *\!\!\rightarrow X$, which contradicts with the constructive property of $\mathbb{M}$. (2) Suppose an edge $T - B$. If $\mathbf{C}$ is not an empty set, then $\mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z} = \mathrm{PossDe}(X, \mathbb{M}[-\mathbf{C}])\backslash\{X\}$ and $B \in \mathrm{De}(X, \mathcal{M})$ according to Lemma 4. There is thus an edge into $T - B$, such that there cannot be a MAG consistent with $\mathbb{M}$, contradicting with the constructive property of $\mathbb{M}$. If $\mathbf{C} = \emptyset$, then $T = X$, thus there is $X - B$ in $\mathbb{M}$, in which case there is not an edge $A \circ\!\!-\!\!\circ B$ in $\mathbb{M}$. (3) Suppose an edge $T \circ\!\!- B$. If $T = X$, then there is $X \circ\!\!- B \circ\!\!-\!\!\circ A \circ\!\!-\!\!* X$, in which case the balanced property cannot hold no matter what the edge between A and X is. If $T \neq X$, then $\mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z} = \mathrm{PossDe}(X, \mathbb{M}[-\mathbf{C}])\backslash\{X\}$ and $B \in \mathrm{De}(X, \mathcal{M})$ according to Lemma 4. Hence the edge between B and T can only be $T \leftarrow B$, which leads to the same contradiction with (1) above. (4) The edge cannot be $T *\!\!\rightarrow B$, for otherwise there is $T *\!\!\rightarrow A$ in $\mathbb{M}$ due to the balanced property of $\mathbb{M}$, which contradicts with $T \in \mathcal{F}_A$. (5) The edge cannot be $T *\!\!-\!\!\circ B$, for otherwise $T \in \mathcal{F}_B$, which contradicts with $T \in \mathcal{F}_A\backslash\mathcal{F}_B$. Combining the parts above, T is not adjacent to B. □

Lemma 6. *Consider a PMG $\mathbb{M}$ compatible with local transformation. Denote $\mathbf{Z} = \{V \in \mathbf{V}(\mathbb{M}) | V = X$, or there is $V -\!\!\circ \cdots -\!\!\circ V' -\!\!\circ X$ in $\mathbb{M}$ and $V' \notin \mathbf{C}\}$. For any local transformation of vertex X in $\mathbb{M}$ represented by $\mathbf{C}$ (it is not necessarily valid), there cannot be a unshielded structure $A \circ\!\!-\!\!\circ B \circ\!\!-\!\!\circ C$ where $A, C \in \mathbf{Z}$ and $B \notin \mathbf{Z}$.*

Proof. According to the definition of $\mathbf{Z}$ and the balanced property of $\mathbb{M}$, there must be a minimal circle-tail path $A -\!\!\circ F_{n-1} -\!\!\circ \cdots -\!\!\circ F_0 (= X)$ in $\mathbb{M}$. B cannot be adjacent to F_{n-1}, for otherwise the edge can only be $B -\!\!\circ F_{n-1}$ due to the balanced property of $\mathbb{M}$, which implies that $B \in \mathbf{Z}$, contradiction. Hence, F_{n-2} cannot be adjacent to B as well, for otherwise the edge between A, B, F_{n-2}, F_{n-1} in $\mathcal{P}$ can only be undirected, which contradicts with $A \circ\!\!-\!\!\circ B$ in $\mathbb{M}$. Similarly, we can conclude that $F_{n-3}, \cdots, X$ are not adjacent to B. Since $C \in \mathbf{Z}$, we can find another minimal circle-tail path $C -\!\!\circ S_{m-1} -\!\!\circ \cdots -\!\!\circ X$ such that each vertex except for C are not adjacent to B. Hence there is a sub-structure comprised of $X, \cdots, S_{m-1}, C, B, A, F_{n-1}, \cdots, X$, where B is not adjacent to any vertices except for A, C. Hence there must be a cycle containing more than four vertices including A, B, C. In this case, the edge between A, B, C should be undirected in $\mathcal{P}$ according to $\mathcal{R}_5$, contradicting with $A \circ\!\!-\!\!\circ B$. ☐

Lemma 7. *Consider a PMG $\mathbb{M}$ compatible with local transformation and a local transformation of X represented by $\mathbf{C}$ which satisfies the fourth condition in Thm. 1. Denote $\mathcal{F}_V = \{T \in \mathbf{C} \cup \{X\} \mid T *\!\!-\!\!\circ V$ in $\mathbb{M}\}$ for $\forall V \in$ $\mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \backslash \mathbf{Z}$. If for an edge $A \circ\!\!-\!\!\circ B$, $A, B \in \mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \backslash \mathbf{Z}$ in $\mathbb{M}$, there is $\mathcal{F}_A \backslash \mathcal{F}_B \neq \emptyset$, then $\mathcal{F}_A \supset \mathcal{F}_B$.*

Proof. Suppose there is a vertex $C \in \mathcal{F}_B \backslash \mathcal{F}_A$. According to the condition, there is a vertex $D \in \mathcal{F}_A \backslash \mathcal{F}_B$. In this case, as $C, D \in \mathbf{C} \cup \{X\}$, there must be some vertex belonging to $\mathbf{C}$. Hence $\mathbf{Z} = \{X\}$ according to the fourth condition of Thm. 1 and thus $\mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \backslash \mathbf{Z} = \mathrm{PossDe}(X, \mathbb{M}[-\mathbf{C}]) \backslash \{X\}$. Hence according to Lemma 4, $A, B \in \mathrm{De}(X, \mathcal{M})$. In addition, according to Lemma 5, there is an uncovered path $\langle D, A, B, C \rangle$. Note there must be an edge $D *\!\!\rightarrow A$ in any MAG $\mathcal{M}$ consistent with $\mathbb{M}$, for otherwise the ancestral property is violated. Similarly, there is $C *\!\!\rightarrow B$. Since in $\mathcal{M}$, there cannot be new unshielded colliders relative to $\mathbb{M}$, the edge between A and B can only be $A - B$. However, the sub-structure $D *\!\!\rightarrow A - B$ violates the ancestral property. The proof completes. ☐

Lemma 8. *Consider a PMG $\mathbb{M}$ compatible with local transformation and a local transformation of X represented by $\mathbf{C}$ which satisfies the four conditions in Thm. 1. Denote $\mathcal{F}_V = \{T \in \mathbf{C} \cup \{X\} \mid T *\!\!\circ V$ in $\mathbb{M}\}$ for $\forall V \in \mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \backslash \mathbf{Z}$. For an edge $A \circ\!\!-\!\!\circ B$ in $\mathbb{M}[\mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \backslash \mathbf{Z}]$, if it is oriented as $A -\!\!\circ B$ in Step 2 of the procedure, then there is a minimal circle path in a form $F_m \circ\!\!-\!\!\circ \ldots \circ\!\!-\!\!\circ F_1 (= A) \circ\!\!-\!\!\circ F_0 (= B), m \geq 1$ in $\mathbb{M}[\mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \backslash \mathbf{Z}]$ such that $\mathcal{F}_{F_m} \supset \mathcal{F}_{F_{m-1}} = \cdots = \mathcal{F}_{F_0}$.*

Proof. In Step 2 of the procedure at the beginning of Appendix C.1, there are only two possible cases for the transformation of $A \circ\!\!-\!\!\circ B$ to $A -\!\!\circ B$. One is that $\mathcal{F}_A \backslash \mathcal{F}_B \neq \emptyset$. The other is that $\mathcal{F}_A \backslash \mathcal{F}_B = \emptyset$ as well as there is a vertex $C \in \mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \backslash \mathbf{Z}$ not adjacent to B such that $C -\!\!\circ A \circ\!\!-\!\!\circ B$. We consider these two cases. We will prove that we can find a desired path $F_m \circ\!\!-\!\!\circ \ldots \circ\!\!-\!\!\circ F_1 (= A) \circ\!\!-\!\!\circ F_0 (= B), m \geq 1$ in $\mathbb{M}[\mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \backslash \mathbf{Z}]$ such that $\mathcal{F}_{F_m} \supset \mathcal{F}_{F_{m-1}} = \cdots = \mathcal{F}_{F_0}$.

If $\mathcal{F}_A \backslash \mathcal{F}_B \neq \emptyset$, there is $\mathcal{F}_A \supset \mathcal{F}_B$ according to Lemma 7. In this case, $A \circ\!\!-\!\!\circ B$ is the desired path where $m = 1$. If $\mathcal{F}_A \backslash \mathcal{F}_B = \emptyset$, in this case there is a structure $C -\!\!\circ A \circ\!\!-\!\!\circ B$ where C is not adjacent to B. There cannot be $C -\!\!\circ A$ in $\mathbb{M}$, for otherwise there cannot be an edge $A \circ\!\!-\!\!\circ B$ in $\mathbb{M}$ due to $\mathcal{R}_7$ and the complete property of $\mathbb{M}$. There are no edges between $\mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \backslash \mathbf{Z}$ transformed in Step 0 - Step 1 of the procedure. Hence there is $F_2 \circ\!\!-\!\!\circ A$ in $\mathbb{M}$ and $F_2 -\!\!\circ A$ is obtained in Step 2 of the procedure. We consider whether $\mathcal{F}_{F_2} \backslash \mathcal{F}_A = \emptyset$. If not empty, then we can find a desired path where $m = 2$. If empty, we can find another vertice F_3 not adjacent to A with an edge $F_3 \circ\!\!-\!\!\circ F_2$ in $\mathbb{M}$. Repeat the process above, we can always find an uncovered path $F_m \circ\!\!-\!\!\circ \ldots \circ\!\!-\!\!\circ F_1 (= A) \circ\!\!-\!\!\circ F_0 (= B), m \geq 1$ in $\mathbb{M}[\mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \backslash \mathbf{Z}]$ where $\mathcal{F}_{F_m} \supset \mathcal{F}_{F_{m-1}} = \cdots = \mathcal{F}_{F_0}$. According to Lemma 2, the circle path is also minimal. We get the desired result. ☐

Lemma 9. *Consider a PMG $\mathbb{M}$ compatible with local transformation and a local transformation of X represented by $\mathbf{C}$ which satisfies the four conditions in Thm. 1. In Step 2 of the procedure, there is not an edge oriented as $J - K$.*

Proof. For simplicity, denote $\mathbb{M}[\mathrm{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \backslash \mathbf{Z}]$ by $\mathbb{M}_1$. According to the third condition of Thm. 1, there is $\mathcal{F}_J \subseteq \mathcal{F}_K$ or $\mathcal{F}_K \subseteq \mathcal{F}_J$. We only present the proof for the case $\mathcal{F}_J = \mathcal{F}_K$. The proof for the case $\mathcal{F}_J \neq \mathcal{F}_K$ is similar by deriving a contradiction through finding a minimal circle path such that the two endpoints are not bridged, thus we leave them for readers.

According to Step 2 of the procedure, if we tnansform a circle edge $J \circ\!\!-\!\!\circ K$ to an undirected edge $J - K$, then we can orient both $J -\!\!\circ K$ and $J \circ\!\!- K$ in Step 2 of the procedure. According to Lemma 8, if we orient $J -\!\!\circ K$ in Step 2, 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 $J \circ\!\!- K$ in Step 2, 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. Due to the balanced property of $\mathbb{M}$, 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 two steps. Hence the circle component in $\mathbb{M}_1$ after the first two steps is still chordal. And $V_0 \circ\!\!-\!\!\circ V_1 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ V_n$ 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 $V_0, \cdots, V_n \in \text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}$, V_0 and V_n are not bridged relative to $\mathbf{C} \cup \mathbf{Z}$, contradicting with the third condition of Thm. 1. □

Lemma 10. *Consider a PMG $\mathbb{M}$ compatible with local transformation and a local transformation of X represented by $\mathbf{C}$ which satisfies the four conditions in Thm. 1. In Step 2 of the procedure, there are only circle edges transformed to $-\!\!\circ$.*

Proof. It directly follows from Lemma 9 and Step 2 of the procedure. □

Lemma 11. *Consider a PMG $\mathbb{M}$ compatible with local transformation and a local transformation of X represented by $\mathbf{C}$ which satisfies the four conditions in Thm. 1. There is not a new unshielded collider introduced by the procedure.*

Proof. Suppose a new unshielded collider $A{*\!\!\rightarrow} B \leftrightarrow\!\!* C$ is introduced by the procedure above. There is $A *\!\!-\!\!\circ B \circ\!\!-\!\!* C$ and is not $A -\!\!\circ B$ or $B \circ\!\!- C$ in $\mathbb{M}$ due to the complete property of $\mathbb{M}$. Note there are no new arrowheads introduced in Step 2 of the procedure above. **(A.)** If $A{*\!\!\rightarrow} B$ and $B \leftarrow\!\!* C$ are introduced in Step 1 of the procedure above, there must be $A, C \in \mathbf{C}$. Due to the second condition of Thm. 1, A must be adjacent to C, contradiction. **(B.)** If one arrowhead at B is introduced in Step 1 and the other is introduced in Step 3 - Step 5, without loss of generality, suppose $A{*\!\!\rightarrow} B$ is introduced in Step 1 and $B \leftarrow\!\!* C$ is introduced in Step 3 - Step 5. There cannot be $B \circ\!\!-\!\!\circ C$ in $\mathbb{M}$, for otherwise there is $B -\!\!\circ C$ obtained in Step 2 of the procedure above as $A \in \mathcal{F}_B\backslash\mathcal{F}_C$. Since there is not $B \circ\!\!- C$ in $\mathbb{M}$ as well, there can only be $B\circ\!\!\rightarrow C$ in $\mathbb{M}$. However, according to the procedure above, the edges in the format of $B\circ\!\!\rightarrow C$ is never transformed to bi-directed edges in Step 3 - Step 5, contradiction. **C.** If both the arrowheads at B on $A{*\!\!\rightarrow} B \leftarrow\!\!* C$ are introduced in Step 5 of the procedure above, there must be $A -\!\!\circ B \circ\!\!- C$ before Step 3. **(C.1)** At first, there cannot be $A, B \in \mathbf{Z}$ due to Lemma 6. **(C.2)** If one vertex belongs to $\mathbf{Z}$ and the other does not, without loss of generality, suppose $A \in \mathbf{Z}$ and $B \notin \mathbf{Z}$. In this case, $B \circ\!\!- C$ can only be obtained in Step 2 of the procedure above. In addition, since $A \in \mathcal{F}_B\backslash\mathcal{F}_C$, $B -\!\!\circ C$ can be oriented in Step 2 of the procedure above, thus there are $B - C$ transformed in Step 2 of the procedure above, which has been proven impossible by Lemma 9. **(C.3)** If $A, B \notin \mathbf{Z}$, the only possible case is that both $A -\!\!\circ B$ and $B \circ\!\!- C$ are obtained in Step 2 of the procedure above. In this case, $A - B - C$ will be further obtained in Step 2 of the procedure, which contradicts with Lemma 9. Hence we have considered all the cases that possibly introduce a new unshielded collider and prove the impossibility of all these cases. Hence there is not a new unshielded collider introduced by the procedure. □

Lemma 12. *Suppose G be a MAG with a directed edge $A \rightarrow B$. Let G' be the graph identical to G except that the edge between A and B is $A \leftrightarrow B$. G' is a MAG Markov equivalent to G if*

(1) there is no directed path from A to B other than $A \rightarrow B$ in G;

(2) For any edge $C \rightarrow A$ in G, $C \rightarrow B$ is also in G; and for any $D \leftrightarrow A$ in G, either $D \rightarrow B$ or $D \leftrightarrow B$ is in G;

(3) there is no discriminating path for A on which B is the endpoint adjacent to A in G;

(4) there is not an undirected edge of A.

Proof. The proof completely follows the proof of Lemma 1 of Zhang & Spirtes (2005), except for the additional condition (4) which avoids that the edge $B \leftrightarrow A$ is into an undirected edge and thus violates the ancestral property. □

Proof of Theorem 1. We first prove the "if" statement.

We first introduce a new notation. Denote $\mathcal{M}(\mathbb{M})$ the set of graphs that can be obtained from $\mathbb{M}$ by transforming all edges $\circ\!\!\rightarrow$/$-\!\!\circ$ to $\rightarrow$ and orient the circle component into a DAG without new unshielded colliders. According to the constructive property of $\mathbb{M}$, all the graphs in $\mathcal{M}(\mathbb{M})$ are MAGs consistent with $\mathbb{M}$.

Denote $\mathcal{H}$ the obtained graph by the procedure at the beginning of Appendix C.1. We obtain a graph $\mathcal{H}_0$ based on $\mathcal{H}$ by transforming all the bi-directed edges $K \leftrightarrow T$ to $K \rightarrow T$ which are $K \circ\!\!\rightarrow T$ in $\mathbb{M}$, transforming all the undirected edges $J - L$ to $J \rightarrow L$ which are $J -\!\!\circ L$ in $\mathbb{M}$, and transforming the subgraph comprised of all the undirected edges $-$ which are $\circ\!\!-\!\!\circ$ in $\mathbb{M}$ into a DAG with no unshielded colliders.

The following proof is comprised of two parts: (A) we prove $\mathcal{H}_0 \in \mathcal{M}(\mathbb{M})$. (B) We prove if $\mathcal{H}_0$ is a MAG consistent with $\mathbb{M}$, then $\mathcal{H}$ is a MAG consistent with $\mathbb{M}$ and local transformation of X represented by $\mathbf{C}$. We thus get the desired result.

(A) $\mathcal{H}_0 \in \mathcal{M}(\mathbb{M})$. At first, we show that the procedure to obtain $\mathcal{H}_0$ from $\mathcal{H}$ is valid. It suffices to show that the subgraph comprised of all the undirected edges in $\mathcal{H}$ which are $\circ\!\!-\!\!\circ$ in $\mathbb{M}$ is chordal, thus we can transform this part into a DAG without new unshielded colliders by a perfect elimination order. It follows directly from the chordal property of $\mathbb{M}$ and the fact that the subgraph of any chordal graph is also chordal. When proving $\mathcal{H}_0 \in \mathcal{M}(\mathbb{M})$, we only need to prove that for the graph $\mathcal{H}_0$, the circle component in $\mathbb{M}$ is transformed to a DAG without new unshielded colliders relative to $\mathbb{M}$. We divide the circle component in $\mathbb{M}$ into $\mathbb{M}[\text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}]$, denoted by CC_1, and $\mathbb{M}[-(\text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z})]$, denoted by CC_2. Note $\mathbb{M}[-(\text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z})]$ here denotes the subgraph of $\mathbb{M}$ induced by $\mathbf{V}(\mathbb{M})\backslash(\text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z})$.

The edges in CC_1 totally follow those of $\mathcal{H}$, which is transformed by Step 2 or 3 of the procedure above. With the same proof idea as Lemma 16.1 of Wang et al. (2023b), we can prove that there are no new unshielded colliders or directed cycles comprised of these edges based on Lemma 11.

For the circle edge in $\mathbb{M}$ connecting a vertex $K \in \text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}$ and another vertex $T \in \mathbf{V}(\mathbb{M})\backslash(\text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z})$. In this case, there must be $T \in \mathbf{C} \cup \mathbf{Z}$, for otherwise $T \in \text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}$ due to the edge $K \circ\!\!-\!\!\circ T$ in $\mathbb{M}$ and $K \in \text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}$. If $\mathbf{C} \neq \emptyset$, then $\mathbf{Z} = \{X\}$ according to the fourth condition of Thm. 1. In this case, according to the procedure above, there is $K \leftarrow T$ in $\mathcal{H}$, thus there is $K \leftarrow T$ in $\mathcal{H}_0$. In this case, there will not be a directed cycle comprised of the edge between CC_1 and CC_2 since all the edges between them are directed towards CC_1. And any two edges between CC_1 and CC_2 will not introduce a new unshielded colliders since $\mathbb{M}[\mathbf{C}]$ is a complete graph according to the second condition of Thm. 1. If $\mathbf{C} = \emptyset$, there is $T \in \mathbf{Z}$ and $K \in \text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}$. According to the procedure, there is $K \leftarrow T$ in $\mathcal{H}$ and $\mathcal{H}_0$. There will not be a directed cycle comprised of more than one edge between CC_1 and CC_2 since all the edges between them are directed towards CC_1. And any two edges between CC_1 and CC_2 will not introduce a new unshielded colliders due to Lemma 6. Thus we can prove that there are no new unshielded colliders or directed cycles comprised of the edge connecting both CC_1 and CC_2.

The edges in CC_2 follow those in $\mathcal{H}$ except for that some additional undirected edges in $\mathcal{H}$ relative to $\mathbb{M}$ are directed in $\mathcal{H}_0$, which are in the form of $J - L$ such that $J -\!\!\circ L$ or $J \circ\!\!-\!\!\circ L$ in $\mathbb{M}$. According to the procedure of obtaining $\mathcal{H}_0$ based on $\mathcal{H}$, for the edges $J - L$ which are $J -\!\!\circ L$ in $\mathbb{M}$, we transform them to directed edges; for the subgraph G_1 comprised of undirected edges which are circle edges in $\mathbb{M}$, we transform them to a DAG without new unshielded colliders. As there are no edges into these undirected edges in $\mathcal{H}$, all the directed edges in $\mathcal{H}$ which are circle edges in $\mathbb{M}$ can be always obtained by a perfect elimination order where the vertices in G_1 are in the last of the order, that is, for any vertex V not in G_1 and any vertex V' in G_1, V is ahead of V' in the perfect elimination order. Hence, in this part there are no new unshielded colliders or directed cycles.

Combining the parts above, we conclude $\mathcal{H}_0 \in \mathcal{M}(\mathbb{M})$.

(B) If $\mathcal{H}_0$ is a MAG consistent with $\mathbb{M}$ and local transformation represented by $\mathbf{C}$, then $\mathcal{H}$ is a MAG consistent with $\mathbb{M}$. Suppose $\mathcal{H}_0$ is a MAG consistent with $\mathbb{M}$ and local transformation represented by $\mathbf{C}$. Since $\mathcal{H}$ has the non-circle marks represented by $\mathbf{C}$, it suffices to prove that $\mathcal{H}$ is a MAG Markov equivalent to $\mathcal{H}_0$.

At first, we construct an auxiliary graph $\mathcal{H}'_0$, such that only and all the differences between $\mathcal{H}'_0$ and $\mathcal{H}_0$ is that for the edges $K \rightarrow T, K \in \text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]), T \in \mathbf{C}$ in $\mathcal{H}_0$ such that there is $K \circ\!\!\rightarrow T$ in $\mathbb{M}$, the edges are $K \leftrightarrow T$ in $\mathcal{H}'_0$. Denote the set of different edges between $\mathcal{H}'_0$ and $\mathcal{H}_0$ in $\mathcal{H}_0 (= \mathcal{H}_0^0)$ by $Edge(\mathcal{H}_0^0) = \{K \rightarrow T \text{ in } \mathcal{H}_0^0 | K \in \text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]), T \in \mathbf{C}, K \circ\!\!\rightarrow T \text{ in } \mathcal{H}_0^0\}$. We obtain $\mathcal{H}'_0$ from $\mathcal{H}_0^0$ by transforming these edges to bi-directed edges. We transform one edge one time. At first, we select the edge $K \rightarrow T$ in $Edge(\mathcal{H}_0^0)$ 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 $Edge(\mathcal{H}_0^0)$; 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 $Edge(\mathcal{H}_0^0)$. Then we obtain

$Edge(\mathcal{H}_0^1)$ by deleting $K \to T$ from $Edge(\mathcal{H}_0^0)$. By such operation, we obtain a new graph $\mathcal{H}_0^1$ and $Edge(\mathcal{H}_0^1)$. Repeat the process above and we could obtain a series of graphs $\mathcal{H}_0^0(=\mathcal{H}_0), \mathcal{H}_0^1, \cdots, \mathcal{H}_0^m, \mathcal{H}_0^{m+1}(=\mathcal{H}_0')$. We prove the desired result by induction, that is, for any $\mathcal{H}_0^j$ and $\mathcal{H}_0^{j+1}$, where $0 \le j \le m$, if $\mathcal{H}_0^j$ is a MAG, then $\mathcal{H}_0^{j+1}$ is a MAG Markov equivalent to $\mathcal{H}_0^j$ using Lemma 12. Suppose there is $K \to T$ in $\mathcal{H}_0^j$ but $K \leftrightarrow T$ in $\mathcal{H}_0^{j+1}$. The first three conditions can be proved with the similar process as Lemma 16.1 of Wang et al. (2023b). We only present the proof of the fourth condition. Suppose there is an undirected edge $D - K$. Note $T \in \mathbf{C}$. According to the fourth condition of Thm. 1, $\mathbf{Z}\backslash\{X\}$ is an empty set. In this case, according to the procedure above, we do not generate any new undirected edges in $\mathcal{H}$ relative to $\mathbb{M}$. Hence there is $D - K$ in $\mathbb{M}$. In this case, there cannot be $K {\circ\!\!\to} T$ in $\mathbb{M}$, as the mark at K in the edge between K and T can only be a tail due to the ancestral property, the complete property of $\mathbb{M}$ is violated. Hence there cannot be an undirected edge $D - K$ in $\mathcal{H}_0^j$. We thus conclude that $\mathcal{H}_0'$ is a MAG Markov equivalent to $\mathcal{H}_0$. Next we will show $\mathcal{H}$ is a MAG Markov equivalent to $\mathcal{H}_0'$.

Note that only and all the differences between $\mathcal{H}$ and $\mathcal{H}_0'$ is that for the undirected edges $J - L$ in $\mathcal{H}$ such that there is $J {-\!\circ} L$ or $J {\circ\!\!-\!\!\circ}\, L$ in $\mathbb{M}$, the edges are $J \to L$ in $\mathcal{H}_0'$. Evidently $\mathcal{H}_0'$ and $\mathcal{H}$ have the same adjacencies. Note $\mathcal{H}$ has less arrowheads than $\mathcal{H}_0'$, and there are no new unshielded colliders or discriminating path for a vertex as a collider introduced in the process of obtaining $\mathcal{H}$, thus $\mathcal{H}_0'$ and $\mathcal{H}$ have the same unshielded colliders, and if a path π forms a discriminating path for b in $\mathcal{H}_0'$ and $\mathcal{H}$, then b is a collider on π in $\mathcal{H}_0'$ if and only if it is a collider on π in $\mathcal{H}$. It suffices to prove that $\mathcal{H}$ is a MAG. Since $\mathcal{H}$ has less arrowheads than $\mathcal{H}_0'$, $\mathcal{H}$ satisfies the maximal property and there is no directed or almost directed cycles in $\mathcal{H}$ due to the ancestral and maximal property of $\mathcal{H}_0'$. Hence it suffices to show that for each undirected edge in $\mathcal{H}$, there is not an arrowhead into this edge.

Suppose there is a structure $A{*\!\!\to} B - C$ in $\mathcal{H}$. We discuss the edge between A and B. (a) Consider there is $A{*\!\!\to} B$ in $\mathbb{M}$. There cannot be $B {-\!\circ} C$ or $B {\circ\!-} C$ in $\mathbb{M}$ due to the complete property of $\mathbb{M}$. Hence if there is $B - C$ in $\mathcal{H}$, there is $B {\circ\!\!-\!\!\circ}\, C$ in $\mathbb{M}$. According to the balanced property of $\mathbb{M}$, there is also $A{*\!\!\to} C$ in $\mathbb{M}$. If there is $B - C$ in $\mathcal{H}$ but $B {\circ\!\!-\!\!\circ}\, C$ in $\mathbb{M}$, there is either $B \in \mathbf{Z}$ or $C \in \mathbf{Z}$. Without loss of generality, suppose $B \in \mathbf{Z}$. In this case there is $A{*\!\!\to} B {-\!\circ} \cdots {-\!\circ} X$ in $\mathbb{M}$, which is impossible since the structure ${*\!\!\to}\ {-\!\circ}$ makes the complete property fail. (b) Consider there is $A {*\!-\!\circ}\, B$ in $\mathbb{M}$. At first, we note that $B \notin \mathbf{Z}$. Suppose $B \in \mathbf{Z}$, if there is $A {-\!\circ} B$, then $A \in \mathbf{Z}$ thus there is $A - B$ in $\mathcal{H}$, contradicting with $A{*\!\!\to} B$; if there is $A {\circ\!\!-\!\!\circ}\, B$ or $A {\leftarrow\!\circ} B$, then there is $A {*\!-} B$ in $\mathcal{H}$, contradicting with $A{*\!\!\to} B$ as well. In this case, if there is undirected edge $B - C$ in $\mathcal{H}$, there must be $C \in \mathbf{Z}$. Then we discuss the edge between B and C. **(b.1)** If there is $B {-\!\circ} C$ in $\mathbb{M}$, then there is $B \in \mathbf{Z}$, contradiction. **(b.2)** If there is $B {\circ\!-} C$ in $\mathbb{M}$, according to the procedure above, we never transform a circle at a vertex B on edge $B {\circ\!-} C$ to a tail if $B \notin \mathbf{Z}$, hence in this case there will not be an undirected edge $B - C$ in $\mathcal{H}$, contradiction. **(b.3)** If there is $B {\circ\!\!-\!\!\circ}\, C$ in $\mathbb{M}$, we will transform this edge to an undirected edge only if $B, C \in \mathbf{Z}$, contradiction. **(b.4)** If there is $B - C$ in $\mathbb{M}$, then there cannot be $A {*\!-\!\circ}\, B$ in $\mathbb{M}$ due to the complete property of $\mathbb{M}$, contradiction.

Combining the parts above, it is impossible that there is a structure $A{*\!\!\to} B - C$ in $\mathcal{H}$. The ancestral property holds in $\mathcal{H}$. The proof completes.

Then we prove the "only if" statement. We prove it by reduction to absurdity. Suppose a MAG $\mathcal{M}$ consistent with $\mathbb{M}$ has the non-circle marks of X represented by the local transformation $\mathbf{C}$.

If $\mathbb{M}[\mathbf{C}]$ is not complete, there are new unshielded colliders in $\mathcal{M}$ relative to $\mathbb{M}$. It is evident that $\mathcal{M}$ is not consistent with $\mathbb{M}$, contradiction.

If both $\mathbf{Z}\backslash\{X\}$ and $\{V \in \mathbf{V}(\mathbb{M}) | V{*\!\!\to} X \text{ in } \mathbb{M} \text{ or } V \in \mathbf{C}\}$ are not empty, there must be $C{*\!\!\to} X$ and $X - T$ in any MAG consistent with $\mathbb{M}$ and local transformation of X represented by $\mathbf{C}$. However, the ancestral property is not fulfilled here due to $C{*\!\!\to} X - T$, contradiction.

Hence, the fourth condition of Thm. 1 is satisfied if there is a MAG consistent with $\mathbb{M}$ and local transformation represented by $\mathbf{C}$. With this result, we can conclude the other results further.

If $\text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \cap \text{Pa}(\mathbf{C}, \mathbb{M}) \neq \emptyset$, suppose $V \to T$ where $V \in \text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])$ and $T \in \mathbf{C}$. In this case, $\mathbf{C}$ is not empty. By Lemma 4, $V \in \text{De}(X, \mathcal{M})$, thus $T \in \text{De}(X, \mathcal{M})$. According to the definition of $\mathbf{C}$, there is $X {\leftarrow\!*} T$, a directed or almost directed cycle forms, contradiction.

Suppose two vertices J, K in $\mathbb{M}[\text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}]$ are not bridged relative to $\mathbf{C} \cup \mathbf{Z}$ due to the minimal circle path $J(=V_0) {\circ\!\!-\!\!\circ}\, V_1 \cdots V_n {\circ\!\!-\!\!\circ}\, K(=V_{n+1})$ in $\mathbb{M}[\text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}])\backslash\mathbf{Z}]$. Without loss of generality, suppose for any sub-path of this path except for this path itself, the sub-path is bridged relative to $\mathbf{C}$. There are two possible cases. One is $n = 0$, the other is $n > 0$. For the first case, suppose there are two vertices $T_1, T_2 \in \mathbf{C}$ such that $T_1 \in \mathcal{F}_{V_0}\backslash\mathcal{F}_{V_1}$ and $T_2 \in \mathcal{F}_{V_1}\backslash\mathcal{F}_{V_0}$.

According to Lemma 7, T_1 is not adjacent to V_1 and T_2 is not adjacent to V_0. Since $T_1, T_2 \in \mathbf{C} \cup \{X\}$ and either $\mathbf{Z} \backslash \{X\}$ or $\{V \in \mathbf{V}(\mathbb{M}) | V {*\!\!\rightarrow} X$ in $\mathbb{M}$ or $V \in \mathbf{C}\}$ is empty, $\mathbf{C}$ is not empty and $\mathbf{Z} = \{X\}$. According to Lemma 4, V_0, V_1 are descendants of X. As there is also $T_1 {*\!\!\rightarrow} X$ and $T_2 {*\!\!\rightarrow} X$ according to the local transformation, there must be $T_1 {*\!\!\rightarrow} V_0$ and $T_2 {*\!\!\rightarrow} V_1$ in $\mathcal{M}$ due to the ancestral property. However, no matter how we transform to circle edge $V_0 \circ\!\!-\!\!\circ V_1$, there will be either a new unshilded collider or an edge into an undirected edge in $T_1 {*\!\!\rightarrow} V_0 \circ\!\!-\!\!\circ V_1 \leftarrow\!\!* T_2$, which implies that there cannot be MAGs consistent with $\mathbb{M}$ and local transformation represented by $\mathbf{C}$.

For the second case, suppose a vertex $T_1 \in \mathcal{F}_{V_0} \backslash \mathcal{F}_{V_1}$ and a vertex $T_2 \in \mathcal{F}_{V_{n+1}} \backslash \mathcal{F}_{V_n}$. If $T_1 \neq T_2$, we can get the contradiction with the same process above. If $T_1 = T_2$, there is a sub-structure comprised of $T_1, V_0, V_1, \cdots, V_{n+1}, T_1$, where each two non-consecutive vertices are not adjacenct. For this structure in $\mathcal{P}$, it holds that either all the edges are undirected by $\mathcal{R}_5$ or the edges $V_0 \rightarrow T_1 \leftarrow V_1$ by $\mathcal{R}_9$. The former case contradicts with $V_0 \circ\!\!-\!\!\circ V_1$ in $\mathbb{M}$. For the latter case, since there is an edge into T_1, T_1 cannot belong to $\mathbf{Z}$. If $T_1 = X$, then V_0 is not in $\mathbb{M}[\text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \backslash \mathbf{Z}]$, contradiction; if $T_1 \in \mathbf{C}$, according to Lemma 4, there $V_0 \in \text{De}(X, \mathcal{M})$. And since $V_0 \rightarrow T_1$, $T_1 \in \text{De}(X, \mathcal{M})$. And according to the definition of $\mathbf{C}$, there is $T_1 {*\!\!\rightarrow} X$, thus ancestral property is violated. Hence there is always a contradiction.

Combining the results above, we conclude that there does not exist a MAG consistent with $\mathbb{M}$ with the local structure of X represented by $\mathbf{C}$ when the four conditions are violated. □

C.2. Proof of Proposition 1

Proof. If $\mathbb{M}[\text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \backslash \mathbf{Z}]$ is not bridged relative to $\mathbf{C} \cup \mathbf{Z}$ in $\mathbb{M}$, according to Def. 2, it is not hard to prove that there is a path $V_0 \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ V_n$ such that $\mathcal{F}_0 \backslash \mathcal{F}_1 \neq \emptyset$ and $\mathcal{F}_n \backslash \mathcal{F}_{n-1} \neq \emptyset$. In the second step of Alg. 1, there will be an undirected edge oriented. Hence "unbridged" is returned. According to the proof of Lemma 9, if $\mathbb{M}[\text{PossDe}(\mathbf{Z}, \mathbb{M}[-\mathbf{C}]) \backslash \mathbf{Z}]$ is bridged relative to $\mathbf{C} \cup \mathbf{Z}$ in $\mathbb{M}$, there is not an undirected edge formed in Alg. 1. Hence "bridged" is returned. □

C.3. Proof of Proposition 2

Proof. If the mark at R on the edge between B and R is a tail, then there is an edge $A {*\!\!\rightarrow} B$ into an undirected edge, which contradicts with the ancestral property, impossibility. Hence there must be $B \rightarrow R$. □

C.4. Proof of Theorem 2

To prove Thm. 2, we first present some supporting results. Note the first three steps are totally same as the first three steps in the procedure at the beginning of Appendix C.1. For a valid local transformation represented by $\mathbf{C}$, it evidently fulfills the four conditions in Thm. 1. Hence some results can be directly used here.

Lemma 13. *The PMG $\mathbb{M}_{i+1}$ in Thm. 2 satisfies the closed property.*

Proof. It follows from the fourth step of Alg. 2. □

Lemma 14. *Suppose $\mathbb{M}_i$ in Thm. 2 is a PMG compatible with local transformation, there must exist a MAG consistent with $\mathbb{M}_i$.*

Proof. It follows from the complete property of $\mathbb{M}_i$. □

Lemma 15. *Suppose there is an unshielded triple $\langle A, B, C \rangle$ in a PAG $\mathcal{P}$ where B is not a collider. Then there cannot be an edge $A {*\!\!\rightarrow} B \leftarrow\!\!* C$ in any MAG consistent with $\mathcal{P}$.*

Proof. It directly follows from the ancestral property and Markov equivalence property. □

Lemma 16. *Suppose $\mathbb{M}_i$ in Thm. 2 is a PMG compatible with local transformation., for any edge $A -\!\!\circ B$ or $A - B$ in $\mathbb{M}_i$, there cannot be an edge into A or B in $\mathbb{M}_i$.*

Proof. According to Lemma 14, there exists some MAGs consistent with $\mathbb{M}_i$. Hence, there cannot be an edge into an undirected edge $A - B$ in $\mathbb{M}_i$ due to the ancestral property. It is suffices to show that there is not an edge into an edge $A -\!\!\circ B$.

Suppose there is an edge into A, that is, there is $C{*}\!\!\rightarrow A \multimap B$ in $\mathbb{M}_i$. According to the closed property and $\mathcal{R}_{11}$, there is $A \rightarrow B$ in $\mathbb{M}_i$, contradiction. Suppose there is an edge into B, that is, there is $A \multimap B \leftarrow\!\!* D$ in $\mathbb{M}_i$. According to Lemma 15, A must be adjacent to D. According to the balanced property of $\mathbb{M}_i$, there is $D{*}\!\!\rightarrow A$ in $\mathbb{M}_i$, the impossibility of this case has been proved. Hence, there is not an edge into $A \multimap B$ in $\mathbb{M}_i$. $\square$

Lemma 17. *Suppose $\mathbb{M}_i$ in Thm. 2 is a PMG compatible with local transformation. If there is a valid local transformation of X represented by $\mathbf{C}$, either $\mathbf{C}$ or $\mathbf{Z}\backslash\{X\}$ is an empty set.*

Proof. It is directly derived from Thm. 1. $\square$

Lemma 18. *The PMG $\mathbb{M}_{i+1}$ in Thm. 2 satisfies the invariant property,* i.e., *the arrowheads and tails in $\mathbb{M}_{i+1}$ are invariant in all the MAGs consistent with $\mathbb{M}_i$ and the local transformation of X represented by $\mathbf{C}$.*

Proof. When $\mathbf{C} \neq \emptyset$, $\mathbf{Z} = \{X\}$ according to Lemma 17. In this case, the non-circle marks introduced in the second step of Alg. 2 are invariant in all the MAGs consistent with $\mathbb{M}_i$ and the local transformation of X represented by $\mathbf{C}$ according to Lemma 7 of Wang et al. (2023b), thus we do not show the details here. When $\mathbf{C} = \emptyset$ and $\mathbf{Z}\backslash\{X\} \neq \emptyset$, for any vertex $Z_1 \in \mathbf{Z}$, there is a path $Z_1 \multimap \cdots Z_k \multimap X$ in $\mathbb{M}_i$. According to the local transformation represented by $\mathbf{C}$, there is $Z_k - X$ in any MAGs consistent with $\mathbb{M}_i$ and local transformation of X. Hence the path can only be $Z_1 - \cdots Z_k - X$ due to the ancestral property. In this case, for any vertex V with an edge $V \ast\!\!-\!\!\circ Z_1$, the edge can only be $V \ast\!\!- Z_1$ due to the ancestral property. Hence the tails introduced in the second step are invariant. Thus, the non-circle marks introduced in the second step of Alg. 2 are invariant.

In the fourth step, we only update the graph with the orientation rules. Since all the rules are sound, hence the non-circle marks introduced in the fourth step are also invariant. Hence it suffices to show the tail introduced in the third step is invariant.

The third step only introduces tails. Suppose the first transformed edge which introduces a tail that is not invariant is $A \multimap B$. Due to the fact that the tail is not invariant, there is a MAG $\mathcal{M}$ consistent with $\mathbb{M}_i$ and $\mathbf{C}$ with an edge $A \leftarrow\!\!* B$. According to the third step of Alg. 2, there are two possible cases for the transformation of $A \multimap B$. If $\mathcal{F}_A = \mathcal{F}_B$ and there is another vertex T not adjacent to B such that $T \multimap A \circ\!\!-\!\!\circ B$ which leads to $A \multimap B$, there is $T \rightarrow\!\!* A \leftarrow\!\!* B$ in $\mathcal{M}$ as the tail at T is invariant. According to Lemma 15, it is impossible. If $\mathcal{F}_A\backslash\mathcal{F}_B \neq \emptyset$, then there is some variable $T \in \mathcal{F}_A\backslash\mathcal{F}_B$. T is not adjacent to B due to Lemma 5. If $\mathbf{C} \neq \emptyset$, then $\text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])\backslash\mathbf{Z} = \text{PossDe}(X, \mathbb{M}_i[-\mathbf{C}])\backslash\mathbf{Z}$ and $B \in \text{De}(X, \mathcal{M})$ according to Lemma 4. According to the second step of Alg. 2, there is $T{*}\!\!\rightarrow A$ in $\mathcal{M}$. Hence, there is an additional unshielded collider $T{*}\!\!\rightarrow A \leftarrow\!\!* B$ in $\mathcal{M}$, which contradicts with the fact that $\mathcal{M}$ is consistent with $\mathcal{P}$. If $\mathbf{C} = \emptyset$, there is $T = X$, in which case there is unshielded triple $X \rightarrow\!\!* A \leftarrow\!\!* B$ in $\mathcal{M}$, which contradicts with Lemma 15. We conclude the impossibility of $A \leftarrow\!\!* B$ in $\mathcal{M}$.

Hence $\mathbb{M}_{i+1}$ satisfies the invariant property. $\square$

Lemma 19. *Consider $\mathbb{M}_{i+1}$ in Thm. 2. The subgraph $\mathbb{M}_{i+1}[\mathbf{C}]$ is a complete graph.*

Proof. It is directly derived from Thm. 1. $\square$

Lemma 20. *Suppose $\mathbb{M}_i$ in Thm. 2 is a PMG compatible with local transformation. Given a valid local transformation of X represented by $\mathbf{C}$, there is $\text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}]) \cap \text{Pa}(\mathbf{C}, \mathbb{M}_i) = \emptyset$.*

Proof. It is directly derived from Thm. 1. $\square$

Lemma 21. *Suppose $\mathbb{M}_i$ in Thm. 2 is a PMG compatible with local transformation. Given a valid local transformation of X represented by $\mathbf{C}$, there is not a new unshielded collider introduced in Alg. 2 when we incorporate the local transformation of $X(= V_{i+1})$ into $\mathbb{M}_i$.*

Proof. According to Lemma 18, if there are new unshielded colliders introduced in Alg. 2, then for any MAG $\mathcal{M}$ consistent with $\mathbb{M}_i$ and the valid local transformation of V_{i+1}, there are new unshielded colliders in $\mathcal{M}$ relative to $\mathcal{P}$, contradicting with Thm. 1. $\square$

Lemma 22. *Suppose $\mathbb{M}_i$ in Thm. 2 is a PMG compatible with local transformation. Given a valid local transformation of X represented by $\mathbf{C}$, in the fourth step of Alg. 2 to obtain $\mathbb{M}_{i+1}$ based on $\mathbb{M}_i$ and the local transformation of X represented by $\mathbf{C}$, there are only edges as $A \circ\!\!\rightarrow B$ transformed to $A \rightarrow B$ or edges as $A -\!\!\circ B$ transformed to $A \rightarrow B$.*

Proof. We can prove that when we transform a circle to an arrowhead, the edge can only be as $-\!\ominus$, with the similar process of Lemma 13 of Wang et al. (2023b) based on Lemma 9. Hence we will not show the proof of this part. Here, we will only prove that if we transform a circle to a tail, then the edge can only be as $A \circ\!\!\rightarrow B$. We denote the graph obtained after the first three steps of Alg. 2 by $\bar{\mathbb{M}}_{i+1}$.

It suffices to show that when we transform a circle to a tail, the edge cannot be as $A \circ\!\!-\!\!\circ B$ or $A -\!\!\circ B$. Suppose there are some edges as $A \circ\!\!-\!\!\circ B$ or $A -\!\!\circ B$ where a circle is transformed to a tail in the third step of Alg. 2. In the following, we will prove the impossibility of both that (i) the first aforementioned edge is $A \circ\!\!-\!\!\circ B$ and (ii) the first aforementioned edge is $A -\!\!\circ B$, which could imply that there are not edges as $A \circ\!\!-\!\!\circ B$ or $A -\!\!\circ B$ where a circle is transformed to a tail.

(i) Suppose the first aforementioned edge is as $A \circ\!\!-\!\!\circ B$. According to the proposed rules, only $\mathcal{R}_1, \mathcal{R}_4', \mathcal{R}_5, \mathcal{R}_6, \mathcal{R}_7$ can be triggered to transform $\circ\!\!-\!\!\circ$ to $-\!\ominus$. If the transformation is triggered by $\mathcal{R}_1$ or $\mathcal{R}_4'$, then there is also a circle transformed to an arrowhead, which means that there are new arrowheads introduced in the fourth step. We have shown that in the fourth step there are no new arrowheads introduced, thus conclude the impossibility of triggering $\mathcal{R}_1$ or $\mathcal{R}_4'$. $\mathcal{R}_5$ is triggered in only the process of obtaining a PAG. It suffices to consider that $\mathcal{R}_6$ or $\mathcal{R}_7$ is triggered. Note we do not orient any edges in the first step of Alg. 2.

(i.1) If $\mathcal{R}_6$ is triggered, suppose there is $C - B \circ\!\!-\!\!\circ A$ such that $\mathcal{R}_6$ is triggered. There cannot be $C - B$ in $\mathbb{M}_i$, for otherwise the complete property of $\mathbb{M}_i$ is contradicted due to $B \circ\!\!-\!\!\circ A$. Hence $C - B$ is introduced in the first three steps of Alg. 2. Note in the third step of Alg. 2, only circle edges are transformed to edges in the format of $-\!\ominus$. Hence $C - B$ is introduced in the second step. If $B \in \mathbf{Z}$, then there is either $A \in \mathbf{Z}$ or $A \in \text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}]) \backslash \mathbf{Z}$, either of which leads to $B -\!\!* A$ in the second step of Alg. 2, contradiction. If $B \notin \mathbf{Z}$, there must be $C \circ\!\!- B$ in $\mathbb{M}_i$ where $C \in \mathbf{Z}$. In this case there is $B \in \mathbf{Z}$ according to the definition of $\mathbf{Z}$, contradiction.

(i.2) If $\mathcal{R}_7$ is triggered, suppose there is $C -\!\!\circ B \circ\!\!-\!\!\circ A$ where C is not adjacent to A, such that $\mathcal{R}_7$ is triggered. There cannot be $C -\!\!\circ B$ in $\mathbb{M}_i$, for otherwise the complete property of $\mathbb{M}_i$ is contradicted. If $C -\!\!\circ B$ is introduced in the second step of Alg. 2, then there is $C \in \mathbf{Z}$ and $B \notin \mathbf{Z}$. There is also $A \notin \mathbf{Z}$, for otherwise there is $B \circ\!\!- A$ oriented in the second step of Alg. 2, contradicting with $B \circ\!\!-\!\!\circ A$. As there is $C \circ\!\!-\!\!\circ B \circ\!\!-\!\!\circ A$ in $\mathbb{M}_i$, there is $A, B \in \text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}]) \backslash \mathbf{Z}$. In this case $B -\!\!\circ A$ is oriented in the third step of Alg. 2, contradicting with $B \circ\!\!-\!\!\circ A$. If $C -\!\!\circ B$ is introduced in the third step of Alg. 2, with the similar proof process as above, $B -\!\!\circ A$ is oriented in the third step, contradiction. Hence $\mathcal{R}_7$ cannot be triggered as well.

(ii) Then we suppose the first aforementioned edge is $A -\!\!\circ B$. Only $\mathcal{R}_6$ or $\mathcal{R}_7$ could possibly trigger the transformation of $A -\!\!\circ B$ to $A - B$. Hence it suffices to consider (ii.1) $\mathcal{R}_7$ and (ii.2) $\mathcal{R}_6$.

(ii.1) If $\mathcal{R}_7$ is triggered, suppose there is $C -\!\!\circ B \circ\!\!- A$ where C is not adjacent to A such that $\mathcal{R}_7$ is triggered. It is evident that there is not $C -\!\!\circ B$ or $B \circ\!\!- A$ in $\mathbb{M}_i$ due to the complete property of $\mathbb{M}_i$. If $C -\!\!\circ B$ is introduced in the second step of Alg. 2, there is $C \in \mathbf{Z}$ and $B \notin \mathbf{Z}$. There is $A \notin \mathbf{Z}$ due to Lemma 6. Hence $C -\!\!\circ B$ is introduced in the third step of Alg. 2. Similarly, it is only possible that $B \circ\!\!- A$ is introduced in the third step. In this case, according to the third step, as C is not adjacent to A, $B - A$ can also be introduced in the third step of Alg. 2, contradicting with Lemma 9.

(ii.2) If $\mathcal{R}_6$ is triggered, suppose there is $C - B \circ\!\!- A$ such that $\mathcal{R}_6$ is triggered. We consider the case that C is adjacent to A, for otherwise $\mathcal{R}_7$ can also be triggered, the impossibility of which has been proven. It is evident that there is not $C - B$ in $\mathbb{M}_i$ due to the complete property of $\mathbb{M}_i$. In the following, we first prove that there cannot be $B \circ\!\!- A$ in $\mathbb{M}_i$. Suppose $B \circ\!\!- A$ in $\mathbb{M}_i$, there must be $B \notin \mathbf{Z}$, for otherwise there is $A \in \mathbf{Z}$ according to the definition of $\mathbf{Z}$, in which case there is $B - A$ oriented in the second step of Alg. 2. There is thus $B \notin \mathbf{Z}$. (A) If $C - B$ is oriented in the second step of Alg. 2, there must be $C \in \mathbf{Z}$ and $C \circ\!\!- B$ in $\mathbb{M}_i$, in which case there is also $B, A \in \mathbf{Z}$ due to $C \circ\!\!- B \circ\!\!- A$ in $\mathbb{M}_i$, contradiction. (B) $C - B$ is not obtained in the third step of Alg. 2 due to Lemma 9. (C) In the forth step, as $B \circ\!\!- A$ is the first edge which is not as $\circ\!\!\rightarrow$ and the circle is transformed to a tail, hence $C - B$ is not obtained in the forth step of Alg. 2. Hence, there is never $B \circ\!\!- A$ in $\mathbb{M}_i$. Hence there is $B \circ\!\!-\!\!\circ A$ in $\mathbb{M}_i$. If $B \circ\!\!- A$ is obtained in the second step of Alg. 2, there is $A \in \mathbf{Z}$ and $B \notin \mathbf{Z}$. Note that $C - B$ can only be obtained in the second step of Alg. 2 due to Lemma 9. Hence there can only be $C \circ\!\!- B$ in $\mathbb{M}_i$ and $C \in \mathbf{Z}$. In this case, there is also $B \in \mathbf{Z}$ according to the definition of $\mathbf{Z}$, contradiction. If $B \circ\!\!- A$ is obtained in the third step of Alg. 2, there is $B \notin \mathbf{Z}$. With the similar process as above, we can conclude that there is $C \circ\!\!- B$ in $\mathbb{M}_i$ and $B, C \in \mathbf{Z}$, contradiction. Hence there is always a contradiction.

Combining the parts above, we conclude that if there is an edge except for $\circ\!\!\rightarrow$ where the circle is transformed to a tail, then the first aforementioned edge can be neither $\circ\!\!-\!\!\circ$ nor $-\!\circ$. Hence all the edges where a circle is transformed to a tail can only be as $\circ\!\!\rightarrow$. $\square$

Lemma 23. *The PMG $\mathbb{M}_{i+1}$ in Thm. 2 satisfies the chordal property.*

Proof. Denote the graph obtained from $\mathbb{M}_i$ and the local transformation represented by $\mathbf{C}$ after the first three steps of Alg. 2 by $\bar{\mathbb{M}}_{i+1}$. According to Lemma 22, no circle edges are oriented in the fourth step of Alg. 2. Hence, it suffices to prove that the circle component in $\bar{\mathbb{M}}_{i+1}$ is chordal.

According to Lemma 17, when $\mathbf{C}$ is not empty, the set $\mathbf{Z}$ only contains one element X. Hence, when $\mathbf{C} \neq \emptyset$, we can prove the result by the similar proof idea with Lemma 14 of Wang et al. (2023b) based on Lemma 9. We do not show the details. It suffices to consider the case that $\mathbf{C} = \emptyset$. If there is a subgraph of $\mathbb{M}_i$ such that any vertex in this subgraph is not connected to $\mathbf{Z}$ by a path, then it is evident that no edges in this subgraph are transformed by Alg. 2, thus this subgraph must be chordal. Hence without loss of generality, we suppose that all the variables are connected to $\mathbf{Z}$ by some paths.

In the first three steps of Alg. 2, we never transform the circles at the vertices in $\mathbf{V}(\mathbb{M}_i)\backslash\text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])$. Hence it suffices to show that $\bar{\mathbb{M}}_{i+1}[\text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])]$ is chordal. It is evident that there is not a circle edge connecting $V' \in \text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])$ and $V'' \in \mathbf{V}(\mathbb{M})\backslash\text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])$, for otherwise there is $V'' \in \text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])$.

The proof of that $\bar{\mathbb{M}}_{i+1}[\text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])]$ is chordal is comprised of three parts. We will show that (a) the subgraph $\bar{\mathbb{M}}_{i+1}[\mathbf{Z}]$ is chordal; (b) there is not a circle edge connecting $\mathbf{Z}$ and $\text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])\backslash\mathbf{Z}$; (c) the subgraph $\bar{\mathbb{M}}_{i+1}[\text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])\backslash\mathbf{Z}]$ is chordal. The proof of part (c) directly refers to that of Lemma 14 of Wang et al. (2023b), we do not show the details.

(a) We will prove that $\bar{\mathbb{M}}_{i+1}[\mathbf{Z}]$ is an undirected graph. According to Alg. 2, the only possible case such that $\bar{\mathbb{M}}_{i+1}[\mathbf{Z}]$ is not an undirected graph is there is an arrowhead in $\mathbb{M}_i[\mathbf{Z}]$. We will prove the impossibility. Suppose there is $Z_1 *\!\!\rightarrow Z_2$ in $\mathbb{M}_i$ where $Z_1, Z_2 \in \mathbf{Z}$. At first $Z_2 \neq X$, for otherwise there cannot be an edge as $V -\!\circ Z_2$ in $\mathbb{M}_i$, for otherwise the complete property of $\mathbb{M}_i$ is violated since the edge can only be $V - Z_2$ as $Z_1 *\!\!\rightarrow Z_2$ in $\mathbb{M}_i$, contradiction. If $Z_2 \neq X$, it implies that there is a path $Z_2 -\!\circ Z' \cdots -\!\circ X$ in $\mathbb{M}_i$, in this case the complete property of $\mathbb{M}_i$ is violated since the edge can only be $Z_2 - Z'$ as $Z_* \!\!\rightarrow Z_2$ in $\mathbb{M}_i$, contradiction. Hence $\bar{\mathbb{M}}_{i+1}[\mathbf{Z}]$ is an undirected graph. $\bar{\mathbb{M}}_{i+1}[\mathbf{Z}]$ is chordal.

(b) If there is a circle edge connecting $\mathbf{Z}$ and $\text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])\backslash\mathbf{Z}$ in $\mathbb{M}_i$, it will be transformed to a non-circle edge in the second step of Alg. 2, hence there cannot be a circle edge connecting $\mathbf{Z}$ and $\text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])\backslash\mathbf{Z}$ in $\bar{\mathbb{M}}_{i+1}$. $\square$

Lemma 24. *The PMG $\mathbb{M}_{i+1}$ in Thm. 2 satisfies the balanced property.*

Proof. If there is $V_i *\!\!\rightarrow V_j \circ\!\!-\!\!* V_k$ or $V_i -\!\circ V_j \circ\!\!-\!\!* V_k$, V_i must be adjacent to V_k, for otherwise the circle at V_j on the edge between V_j and V_k will be transformed in the fourth step of of Alg. 2, which contradicts with Lemma 22. Denote the graph obtained from $\mathbb{M}_i$ and the local transformation represented by $\mathbf{C}$ after the first three steps of Alg. 2 by $\bar{\mathbb{M}}_{i+1}$.

Similar to the proof process of Lemma 15 of Wang et al. (2023b), we can conclude that if there is $V_i *\!\!\rightarrow V_j \circ\!\!-\!\!* V_k$ in $\mathbb{M}_{i+1}$, then there must be $V_i *\!\!\rightarrow V_k$ in $\mathbb{M}_{i+1}$.

Then, we first prove that if there is $V_i -\!\circ V_j \circ\!\!-\!\!\circ V_k$ in $\mathbb{M}_{i+1}$, then there is $V_i -\!\circ V_k$ in $\mathbb{M}_{i+1}$. **(A.)** If there is $V_i -\!\circ V_j$ in $\mathbb{M}_i$, then there is $V_i -\!\circ V_k$ in $\mathbb{M}_i$ according to the balanced property of $\mathbb{M}_i$. Next we will prove that the edge $V_i -\!\circ V_k$ cannot be transformed to $V_i \rightarrow V_k$ or $V_i - V_k$ by Alg. 2. If $V_i -\!\circ V_k$ is transformed to $V_i \rightarrow V_k$, it can be triggered in the second or the fourth step of Alg. 2. If it is triggered in the second step, if $V_k = X$, then there cannot be $V_j \circ\!\!-\!\!\circ X$; if $V_k \in \text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])\backslash\mathbf{Z}$ and $V_i \in \mathbf{C}$, then $V_i -\!\circ V_j$ is also transformed to $V_i \rightarrow V_j$, impossibility. If $V_i -\!\circ V_k$ is transformed to $V_i \rightarrow V_k$ in the fourth step, in this case there must be an edge $V_m \rightarrow V_i$, which can also lead to $V_i \rightarrow V_j$ in $\mathbb{M}_{i+1}$, contradiction. If $V_i -\!\circ V_k$ is transformed to $V_i - V_k$, there is $V_k \in \mathbf{Z}$, in which case there should be $V_j \circ\!\!- V_k$ in $\mathbb{M}_{i+1}$, contradiction. Hence there is $V_i -\!\circ V_k$ in $\mathbb{M}_{i+1}$. **(B.)** If there is not $V_i -\!\circ V_j$ in $\mathbb{M}_i$, according to Alg. 2, the edge can only be $V_i \circ\!\!-\!\!\circ V_j$ in $\mathbb{M}_i$. Since there is $V_i \circ\!\!-\!\!\circ V_j \circ\!\!-\!\!\circ V_k$ in $\mathbb{M}_i$, according to the balanced property of $\mathbb{M}_i$, there is $V_i \circ\!\!-\!\!\circ V_k$ in $\mathbb{M}_i$. According to Lemma 22, $V_i -\!\circ V_j$ is obtained in either the second or the third step of Alg. 2. If $V_i -\!\circ V_j$ is obtained by the second step of Alg. 2, there is $V_i \in \mathbf{Z}$ and $V_j \notin \mathbf{Z}$. As there is $V_j \circ\!\!-\!\!\circ V_k$ in $\mathbb{M}_{i+1}$, $V_k \notin \mathbf{Z}$. In this case, there is $V_i -\!\circ V_k$ obtained in the second step of Alg. 2. The result holds. If $V_i -\!\circ V_j$ is obtained in the third step of Alg. 2, according to Lemma 8, there must be a minimal circle path $F_m \circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ F_1(= V_i) \circ\!\!-\!\!\circ F_0(= V_j), m \geq 1$ in $\mathbb{M}_i$ where

$\mathcal{F}_{F_m} \supset \mathcal{F}_{F_{m-1}} = \cdots = \mathcal{F}_{F_0}$. Note if there is any edge connecting $F_s, 0 \leq s \leq m$ and V_k, the edge can only be circle edge in $\mathbb{M}_i$ due to the balanced property of $\mathbb{M}_i$. In this case, at first, F_m is not adjacent to V_k, for otherwise the edge is as $F_m -\!\circ V_k$ due to $\mathcal{F}_{F_m} \supset \mathcal{F}_{V_j} = \mathcal{F}_{V_k}$. And F_m cannot be adjacent to V_j, for otherwise the edge between V_k and V_j is transformed to $V_j -\!\circ V_k$ in the third step of Alg. 2, contradiction. Similarly, we can prove that $F_{m-1}, \cdots, F_2$ is not adjacent to V_k. In this case, as there is $F_2 -\!\circ F_1 (= V_i)$ transformed in the third step of Alg. 2, there is also $V_i -\!\circ V_k$ obtained. The result holds.

Finally, we will prove that if there is $V_i -\!\circ V_j \circ\!\!\rightarrow V_k$ in $\mathbb{M}_{i+1}$, then there is $V_i \rightarrow V_k$ or $V_i \circ\!\!\rightarrow V_k$ in $\mathbb{M}_{i+1}$.

(i) We first consider that there is $V_i -\!\circ V_j$ in $\mathbb{M}_i$. We discuss the situations that there is and not $V_j \circ\!\!\rightarrow V_k$ in $\mathbb{M}_i$, respectively.

(i.1) If there is also $V_j \circ\!\!\rightarrow V_k$ in $\mathbb{M}_i$, then there is $V_i \rightarrow V_k$ or $V_i \circ\!\!\rightarrow V_k$ in $\mathbb{M}_i$ according to the balanced property. According to Alg. 2, the case such that there is not an edge $V_i \rightarrow V_k$ or $V_i \circ\!\!\rightarrow V_k$ in $\mathbb{M}_{i+1}$ is either $V_i = X$, or $V_i \in \text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}]) \backslash \mathbf{Z}$ and $V_k \in \mathbf{C}$, which orients $V_i \leftrightarrow V_k$ in the second step. For the former case, there cannot be $V_i (= X) -\!\circ V_j$ in $\mathbb{M}_{i+1}$, contradiction. For the latter case, since there is $V_i -\!\circ V_j$, if $V_j \notin \mathbf{C}$, then $V_j \in \text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])$, thus $V_j \leftrightarrow V_k$ is oriented in the second step; if $V_j \in \mathbf{C}$, then according to Lemma 4, $V_i \in \text{De}(X, \mathcal{M})$ and there is $V_i -\!* V_j *\!\!\rightarrow X$, the ancestral property is violated, contradiction. Hence, the case that there is not an edge $V_i \rightarrow V_k$ or $V_i \circ\!\!\rightarrow V_k$ in $\mathbb{M}_{i+1}$ is impossible.

(i.2) If there is not $V_j \circ\!\!\rightarrow V_k$ in $\mathbb{M}_i$, there is $V_j \circ\!\!-\!\!\circ V_k$ in $\mathbb{M}_i$. According to the balanced property of $\mathbb{M}_i$, there is $V_i -\!\circ V_k$ in $\mathbb{M}_i$. $V_i -\!\circ V_k$ cannot be transformed to $V_i - V_k$ by Alg. 2, for otherwise $V_k \in \mathbf{Z}$ and thus there is also $V_j \circ\!\!- V_k$, contradicting with the edge $V_j \circ\!\!\rightarrow V_k$. Hence there must be $V_i \rightarrow V_k$ or $V_i -\!\circ V_k$ in $\mathbb{M}_{i+1}$.

(ii) Next we consider that there is not $V_i -\!\circ V_j$ in $\mathbb{M}_i$. In this case, there is $V_i \circ\!\!-\!\!\circ V_j$ in $\mathbb{M}_i$. We discuss the situations that there is and not $V_j \circ\!\!\rightarrow V_k$ in $\mathbb{M}_i$, respectively.

(ii.1) If there is $V_j \circ\!\!\rightarrow V_k$ in $\mathbb{M}_i$, according to the balanced property of $\mathbb{M}_i$, there is $V_i \rightarrow V_k$ or $V_i \circ\!\!\rightarrow V_k$ in $\mathbb{M}_i$. The case such that there is not an edge $V_i \rightarrow V_k$ or $V_i \circ\!\!\rightarrow V_k$ in $\mathbb{M}_{i+1}$ is $V_k \in \mathbf{C}$ and either (a) $V_i = X$ or (b) $V_i \in \text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}]) \backslash \mathbf{Z}$. Case (a) is impossible for otherwise there cannot be $V_i (= X) -\!\circ V_j$ in $\mathbb{M}_{i+1}$. For case (b), if $V_j \notin \mathbf{C}$, then there is $V_j \leftarrow\!\!* V_k$ oriented in the second step of Alg. 2; if $V_j \in \mathbf{C}$, there is a structure $V_j *\!\!\rightarrow X \rightarrow \cdots \rightarrow V_i -\!* V_j$, which always violates the ancestral property. Hence the case $V_i \leftrightarrow V_k$ is impossible.

(ii.2) If there is not $V_j \circ\!\!\rightarrow V_k$ in $\mathbb{M}_i$, there is $V_i \circ\!\!-\!\!\circ V_j \circ\!\!-\!\!\circ V_k \circ\!\!-\!\!\circ V_i$ in $\mathbb{M}_i$. In this case, there is either (a) $V_k = X$ or (b) $V_k \in \mathbf{C}$ and $V_j \in \text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}]) \backslash \mathbf{Z}$. Case (a) is impossible for otherwise there cannot form an edge $V_i -\!\circ V_j$ in $\mathbb{M}_{i+1}$. For case (b), there must be $V_i \in \mathbf{C}$, for otherwise there is $V_i \leftarrow\!\circ V_j$ oriented in the second step of Alg. 2, contradicting with $V_i -\!\circ V_j$. Hence there is $V_i \circ\!\!\rightarrow V_k$ oriented in the second step of Alg. 2. And since we never add an arrowhead at the vertex from $\mathbf{C}$, there can only be $V_i \rightarrow V_k$ or $V_i \circ\!\!\rightarrow V_k$ in $\mathbb{M}_{i+1}$. We get the desired result.

Combining the results above, balanced property holds in $\mathbb{M}_{i+1}$. □

Lemma 25 (Spirtes & Richardson (1996)). *Two MAGs $\mathcal{M}_1$ and $\mathcal{M}_2$ are Markov equivalent if and only if*

(1) $\mathcal{M}_1$ and $\mathcal{M}_2$ have the same adjacencies;

(2) $\mathcal{M}_1$ and $\mathcal{M}_2$ have the same unshielded colliders; and

(3) if a path π forms a discriminating path for b in $\mathcal{M}_1$ and $\mathcal{M}_2$, then b is a collider on π in $\mathcal{M}_1$ if and only if it is a collider on π in $\mathcal{M}_2$.

Lemma 26. *For the PMG $\mathbb{M}_{i+1}$ in Thm. 2, if there is $A -\!\circ B$, then there is no edge into A or B.*

Proof. If there is an edge $C \circ\!\!\rightarrow A$ in $\mathbb{M}_{i+1}$, $C \circ\!\!\rightarrow A -\!\circ B$ is not closed under $\mathcal{R}_{11}$, contradiction. If there is an edge $D \circ\!\!\rightarrow B$ in $\mathbb{M}_{i+1}$, according to the balanced property of $\mathbb{M}_{i+1}$, there is also $D *\!\!\rightarrow A$, contradiction. Hence there cannot be an edge into A or B. □

Lemma 27. *For the PMG $\mathbb{M}_{i+1}$ in Thm. 2, there is not a tail-circle cycle as $A -\!\circ B -\!\circ \cdots -\!\circ A$.*

Proof. According to Lemma A.5 of Zhang (2008), there cannot be a tail-circle cycle in $\mathcal{P}$. Suppose there is a tail-circle cycle in $\mathbb{M}_{i+1}$. At first, we show that there must be such a cycle with three edges, that is, there is $A -\!\circ B -\!\circ C -\!\circ A$. This part totally follows that of Lemma A.5 of Zhang (2008). Suppose the minimal tail-circle cycle is $V_0 -\!\circ \cdots -\!\circ V_n, n \geq 3$.

This cycle cannot be uncovered, for otherwise all the edges are transformed to undirected edges by $\mathcal{R}_6$ and $\mathcal{R}_7$. Suppose V_0 and V_2 are adjacent. According to the balanced property of $\mathbb{M}_{i+1}$, there is not a contradiction only if the edge between V_0 and V_2 is $V_0 -\!\circ V_2$ or $V_0 \circ\!- V_2$. However, either of the two cases will lead to a shorter tail-circle cycle, contradiction. Hence there must be a tail-circle cycle with three edges. Suppose $A -\!\circ B -\!\circ C -\!\circ A$ in $\mathbb{M}_{i+1}$.

(A) If there is $A -\!\circ B$ in $\mathcal{P}$, according to Lemma A.3 of Zhang (2008), there must be an uncovered path $F_0 - F_1 -\!\circ \cdots -\!\circ F_n (= A) -\!\circ B$ in $\mathcal{P}$. The path must be minimal, for otherwise the edges can be transformed to undirected edges according to $\mathcal{R}_6$ and $\mathcal{R}_7$. Note C must be adjacent to F_{n-1}, for otherwise there is $A -\!* C$ in $\mathcal{P}$ by $\mathcal{R}_7$, contradicting with $A \circ\!- C$ in $\mathbb{M}_{i+1}$. As there is $A \circ\!- C$ in $\mathbb{M}_{i+1}$, there is either $A \circ\!\!-\!\!\circ C$ or $A \circ\!- C$ in $\mathcal{P}$. If it is $A \circ\!\!-\!\!\circ C$, there is $F_{n-1} -\!\circ C$ in $\mathcal{P}$ according to the balanced property, in which case there is $C -\!* B$ in $\mathcal{P}$ by $\mathcal{R}_7$, contradicting with $B -\!\circ C$ in $\mathbb{M}_{i+1}$. Hence it is only possible that there is $A \circ\!- C$ in $\mathcal{P}$. Note that F_{n-1} is not adjacent to B, hence there cannot be $F_{n-1} -\!* C$, for otherwise there is $C -\!* B$ in $\mathcal{P}$, contradicting with $C \circ\!- B$ in $\mathbb{M}_{i+1}$. Hence there can only be $F_{n-1} \circ\!\!-\!\!\circ C$ or $F_{n-1} \circ\!- C$ in $\mathcal{P}$. Repeat the process above, we can conclude that there is $F_1 \circ\!\!-\!\!\circ C$ or $F_1 \circ\!- C$ in $\mathcal{P}$. Neither is possible due to $\mathcal{R}_7$ and the undirected edge $F_0 - F_1$.

(B) Next we consider $A -\!\circ B$ is obtained when we introduce the local transformation of V_{j+1} into $\mathbb{M}_j$, where $0 \le j \le i$. According to Alg. 2 and Lemma 22, the tail is introduced in either Step 2 or Step 3. If it is introduced in Step 2, there is $A \in \mathbf{Z}$. If there is $C -\!\circ A$ in $\mathbb{M}_j$, then there is $C \in \mathbf{Z}$ and thus $C - A$ is obtained in Step 2; if there is $C \circ\!\!-\!\!\circ A$ in $\mathbb{M}_j$, then there is $C \circ\!- A$ obtained in Step 2. Neither cannot result in an edge $A \circ\!- C$ in $\mathbb{M}_{i+1}$. If the tail is introduced in Step 3, since $\mathbb{M}_j$ fulfills the balanced property, and there is a tail-circle cycle $A -\!\circ B -\!\circ C -\!\circ A$ in $\mathbb{M}_{j+1}$, the edges among A, B, C can only be circle edges in $\mathbb{M}_j$. In this case, according to the third step of Alg. 2, the edges between A, B, C will be transformed to undirected edges, which contradicts with Lemma 9. □

Lemma 28. *The PMG $\mathbb{M}_{i+1}$ in Thm. 2 satisfies the complete property.*

In $\mathbb{M}_{i+1}$, there are three types of edges with circles: $A \circ\!\!-\!\!\circ B$, $C \circ\!\!\rightarrow D$, $E -\!\circ F$. By Lemma 28.1, we show that we can always obtain a MAG consistent with $\mathcal{P}$ and local transformation of $V_1, \cdots, V_{i+1}$ by transforming $-\!\circ$/$\circ\!\!\rightarrow$ to $\rightarrow$, transforming the circle component into a DAG without new unshielded colliders. Due to the chordal property of $\mathbb{M}_{i+1}$, each circle edge can be transformed to both $\rightarrow$ and $\leftarrow$. This part implies that the circle edges in $\mathbb{M}_{i+1}$ are unidentifiable given the PAG and the local transformation of $V_1, \cdots, V_{i+1}$.

Denote the set of all the edges in the form of $-\!\circ$ in $\mathbb{M}_{i+1}$ by TC. By Lemma 28.2, we show that for any graph obtained by the process above, we can construct a MAG consistent with $\mathcal{P}$ and local transformation of $V_1, \cdots, V_{i+1}$ with the edges in TC being undirected. The two parts above imply that the edges $-\!\circ$ in $\mathbb{M}_{i+1}$ are unidentifiable given the PAG and the local transformation of $V_1, \cdots, V_{i+1}$.

Finally, by Lemma 28.3, we show that for each edge $\circ\!\!\rightarrow$ in $\mathbb{M}_{i+1}$, it can be $\leftrightarrow$ in some MAG consistent with $\mathcal{P}$ and local transformation of $V_1, \cdots, V_{i+1}$. The proof of this part totally follows that of Thm. 3 of Zhang (2008). Combining Lemma 28.1 and Lemma 28.3, the edges $\circ\!\!\rightarrow$ in $\mathbb{M}_{i+1}$ are unidentifiable given the PAG and the local transformation of $V_1, \cdots, V_{i+1}$.

Combining the three parts above, we can conclude that $\mathbb{M}_{i+1}$ in Thm. 2 satisfies the complete property.

Lemma 28.1. *Consider $\mathbb{M}_{i+1}$ in Thm. 2. We obtain a graph $\mathcal{H}_{i+1}$ from $\mathbb{M}_{i+1}$ by transforming $\circ\!\!\rightarrow$/$-\!\circ$ to $\rightarrow$ and the circle component in $\mathbb{M}_{i+1}$ into a DAG without unshielded colliders. Then $\mathcal{H}_{i+1}$ is a MAG consistent with $\mathbb{M}_i$ and local transformation of X.*

Proof. It follows directly from the proof of Thm. 1. □

Lemma 28.2. *Denote a MAG $\mathcal{H}$ obtained by the process in Lemma 28.1. Denote $\mathbf{K} = \{V \in \mathbf{V}(\mathbb{M}_{i+1}) | V \circ\!-$ or $V -\!\circ$ in $\mathbb{M}_{i+1}\}$. Obtain a new graph $\mathcal{H}'$ based on $\mathcal{H}$ by transforming all the edges connecting two variables in $\mathbf{K}$ to undirected edges. Then $\mathcal{H}'$ is also a MAG consistent with $\mathbb{M}_i$ and local transformation of X.*

Proof. The whole proof is comprised of two parts: **A.** $\mathcal{H}'$ has the non-circle marks in $\mathbb{M}_{i+1}$;**B.** $\mathcal{H}'$ is a MAG Markov equivalent to $\mathcal{H}$.

A. $\mathcal{H}'$ has the non-circle marks in $\mathbb{M}_{i+1}$. According to the process of obtaining $\mathcal{H}'$, the only possibility that $\mathcal{H}'$ does not have the non-circle marks in $\mathbb{M}_{i+1}$ is that when we transform an edge between $K_1, K_2 \in \mathbf{K}$ to an undirected edge, an arrowhead in $\mathbb{M}_{i+1}$ is transformed to a tail. Without loss of generality, suppose there is $K_1 *\!\!\rightarrow K_2$ in $\mathbb{M}_{i+1}$. According to the definition of $\mathbf{K}$, there is an edge $A_1 -\!\circ K_2$ or $A_1 \circ\!- K_2$ in $\mathbb{M}_{i+1}$, which contradicts with Lemma 26. Hence $\mathcal{H}'$ has the non-circle marks in $\mathbb{M}_{i+1}$.

B. $\mathcal{H}'$ is a MAG Markov equivalent to $\mathcal{H}$. Note all and only differences between $\mathcal{H}'$ and $\mathcal{H}$ is that some directed edges $K_1 \rightarrow K_2$ in $\mathcal{H}$ are undirected in $\mathcal{H}'$, and there is $K_1, K_2 \in \mathbf{K}$. Since $\mathcal{H}$ is a MAG, it is direct that $\mathcal{H}'$ fulfills the maximal property and does not have directed or almost directed cycles, and $\mathcal{H}$ and $\mathcal{H}'$ have the same adjacencies, and if a path π forms a discriminating path for b in $\mathcal{H}$ and $\mathcal{H}'$, then b is a collider on π in $\mathcal{H}'$ if and only if it is a collider on π in $\mathcal{H}$. It is evident that an unshielded collider in $\mathcal{H}'$ is also an unshielded collider in $\mathcal{H}$. If there is an unshielded collider $A \ast\!\!\rightarrow B \leftarrow\!\!\ast C$ in $\mathcal{H}$ but it is not in $\mathcal{H}'$, without loss of generality, suppose there is not $A \ast\!\!\rightarrow B$ in $\mathcal{H}'$, then we conclude (1) there is $A \ast\!\!\rightarrow B \leftarrow\!\!\ast C$ in $\mathcal{P}$ and $\mathbb{M}_{i+1}$; (2) $A, B \in \mathbf{K}$. In this case, there is an edge $C \ast\!\!\rightarrow B$ into a vertex B on an edge $-\!\circ$ or $\circ\!-$ in $\mathbb{M}_{i+1}$, contradicting with Lemma 26. Hence $\mathcal{H}'$ and $\mathcal{H}$ have the same unshielded colliders. It suffices to prove that there is not a structure $A \ast\!\!\rightarrow B - C$ in $\mathcal{H}'$.

Suppose there is $A \ast\!\!\rightarrow B - C$ in $\mathcal{H}'$, where $B, C \in \mathbf{K}$. There cannot be $A \ast\!\!\rightarrow B$ in $\mathbb{M}_{i+1}$ due to Lemma 26. And there cannot be $A -\!\circ B$ in $\mathbb{M}_{i+1}$, for otherwise $A, B \in \mathbf{K}$ and thus there will be $A - B$ in $\mathcal{H}'$, contradicting with $A \ast\!\!\rightarrow B$. And in the process of obtaining $\mathcal{H}'$, we never transform an edge $\circ\!\!\rightarrow$ to $\leftrightarrow$, thus there is not $A \leftarrow\!\circ B$ in $\mathbb{M}_{i+1}$. Hence it suffices to prove the impossibility of $A \circ\!\!-\!\!\circ B$ in $\mathbb{M}_{i+1}$.

Suppose $A \circ\!\!-\!\!\circ B$ in $\mathbb{M}_{i+1}$. Since $B \in \mathbf{K}$, there is some vertex D with an edge $D -\!\circ B$ or $D \circ\!- B$ in $\mathbb{M}_{i+1}$. If there is $D -\!\circ B$, then there is $D -\!\circ A$ in $\mathbb{M}_{i+1}$ according to the balanced property of $\mathbb{M}_{i+1}$, in this case there is $A \in \mathbf{K}$ and thus there is $A - B$ in $\mathcal{H}'$, contradiction. Hence there can only be $D \circ\!- B$ in $\mathbb{M}_{i+1}$. In this case, if D is adjacent to A, then there must be $D \circ\!- A$ in $\mathbb{M}_{i+1}$ according to the balanced and closed property ($\mathcal{R}_6$) of $\mathbb{M}_{i+1}$, in which case there is $A \in \mathbf{K}$ as well. Hence D cannot be adjacent to A. Next we consider when the edge $D \circ\!- B$ is obtained from $D \circ\!\!-\!\!\circ B$.

(A) If there is $D \circ\!- B$ in $\mathcal{P}$, according to Lemma A.3 of Zhang (2008), there must be an undirected edge $F - B$ or $F -\!\circ B$ in $\mathcal{P}$ where F is not adjacent to D. Note there is also $A \circ\!\!-\!\!\circ B$ in $\mathcal{P}$. $F - B$ is impossible due to $\mathcal{R}_6$. If there is $F -\!\circ B$ in $\mathcal{P}$, there is $F -\!\circ A$ in $\mathcal{P}$ according to the balanced property of $\mathcal{P}$. Hence there is $F \rightarrow A$, $F -\!\circ A$, or $F - A$ in $\mathbb{M}_{i+1}$. Note that there is $A \circ\!\!-\!\!\circ B$ in $\mathbb{M}_{i+1}$. (A.1) If there is $F \rightarrow A$ in $\mathbb{M}_{i+1}$, there is $F \ast\!\!\rightarrow B$ according to the balanced property of $\mathbb{M}_{i+1}$, in which case $B \rightarrow D$ is obtained by $\mathcal{R}_{11}$, contradicting with $B -\!\circ D$ in $\mathbb{M}_{i+1}$. (A.2) If there is $A -\!\circ B$ in $\mathbb{M}_{i+1}$, there is $A \in \mathbf{K}$ thus there is $A - B$ in $\mathcal{H}'$, contradiction. (A.3) If there is $F - A$ in $\mathbb{M}_{i+1}$, there cannot be $A \circ\!\!-\!\!\circ B$ in $\mathbb{M}_{i+1}$ as $\mathbb{M}_{i+1}$ is closed under $\mathcal{R}_6$. Hence there cannot be $D \circ\!- B$ in $\mathcal{P}$.

(B) Next we consider $D \circ\!- B$ is obtained when we introduce the local transformation of V_{j+1} into $\mathbb{M}_j$, where $0 \leq j \leq i$. According to Alg. 2 and Lemma 22, the tail is introduced in Step 2 or Step 3. If it is introduced in Step 2, there is $B \in \mathbf{Z}$. As there is an edge $A \circ\!\!-\!\!\circ B$ in $\mathbb{M}_j$, an edge $A \circ\!- B$ or $A - B$ is also introduced by Step 2, hence there must be $A \ast\!- B$ in $\mathbb{M}_{i+1}$, in which case there cannot be an edge $A \rightarrow B$ in $\mathcal{H}'$. If it is introduced in Step 3, since there is $B -\!\circ D$ oriented in the process of obtaining $\mathbb{M}_{j+1}$ based on $\mathbb{M}_i$ by Alg. 2, there cannot be vertices with an edge into V_{j+1}. Thus, there must be an edge $V_{j+1} -\!\circ \cdots -\!\circ B$ in $\mathbb{M}_{j+1}$. In this case, since there is $A \circ\!\!-\!\!\circ B$, there must be a tail-circle path $V_{j+1} -\!\circ \cdots F_n -\!\circ A$ in $\mathbb{M}_{j+1}$. Then we discuss the edge between F_n and A in $\mathbb{M}_{i+1}$. If it is an undirected edge, then there cannot be an edge $A \circ\!\!-\!\!\circ B$ in $\mathbb{M}_{i+1}$, for otherwise $\mathcal{R}_6$ is triggered. If it is $F_n \rightarrow A$, due to the balanced property of $\mathbb{M}_{i+1}$, there is also $F_n \ast\!\!\rightarrow B$ in $\mathbb{M}_{i+1}$, in which case there cannot be an edge $B -\!\circ D$ in $\mathbb{M}_{i+1}$. If the edge between F_n and A is $F_n -\!\circ A$, then $A \in \mathbf{K}$, contradiction. Hence we can always get a contradiction. $D \circ\!- B$ in $\mathbb{M}_{i+1}$ is impossible. There cannot be a structure $A \ast\!\!\rightarrow B - C$ in $\mathcal{H}'$. We conclude that $\mathcal{H}'$ is a MAG Markov equivalent to $\mathcal{H}$. And since $\mathcal{H}'$ has the non-circle marks in $\mathbb{M}_{i+1}$, $\mathcal{H}'$ is a MAG consistent with $\mathcal{P}$ and local transformation of $V_1, \cdots, V_{i+1}$. ☐

Lemma 28.3. *Suppose there is $A \circ\!\!\rightarrow B$ in the PMG $\mathbb{M}_{i+1}$ in Thm. 2. There is a MAG $\mathcal{H}''$ consistent with $\mathbb{M}_i$ and local transformation of X with the edge $A \leftrightarrow B$.*

Proof. This part totally follows Thm. 3 of Zhang (2008) with the results we have proved before. Hence we only show the sketch. We take $\mathbb{M}_{i+1}$ as the $\mathcal{P}_{AFCI}$ of Zhang (2008). $\mathbf{P}_1$ and $\mathbf{P}_3$ are proved by Lemma 24. $\mathbf{P}_2$ and $\mathbf{P}_4$ are proved by Lemma 26 and Lemma 27. Hence Lemma B.1-Lemma B.18 of Zhang (2008) hold, which are sufficient to prove Thm. 3 of Zhang (2008). By Lemma 28.1, we have proven that when we transform the $\circ\!\!\rightarrow$/$-\!\circ$ edges to $\rightarrow$, and orient the circle component into a DAG without new unshielded colliders based on $\mathbb{M}_{i+1}$, we can always obtain a MAG $\mathcal{H}_{i+1}$ consistent with $\mathcal{P}$ and local transformation of $V_1, \cdots, V_{i+1}$. It is similar to Thm. 2 of Zhang (2008). We can construct a graph $\mathcal{H}''$ with $A \leftrightarrow B$ by the same procedure of Thm. 3 of Zhang (2008) and prove $\mathcal{H}''$ is a MAG that is Markov equivalent to a MAG $\mathcal{H}_{i+1}$. Hence $\mathcal{H}''$ is a MAG in the MEC represented by $\mathcal{P}$. And since $\mathcal{H}''$ has the non-circle edges in $\mathbb{M}_{i+1}$, $\mathcal{H}$ is a MAG consistent with $\mathcal{P}$ and local transformation of $V_1, \cdots, V_{i+1}$. ☐

Proof of Theorem 2. We can conclude that the arrowheads and tails in $\mathbb{M}_{i+1}$ are invariant in all the MAGs consistent with $\mathbb{M}_i$ and the local transformation of X represented by $\mathbf{C}$ by Lemma 18. The chordal, balanced, complete properties of $\mathbb{M}_{i+1}$ are proved by Lemma 23, 24, 28. The constructive property of $\mathbb{M}_{i+1}$ is proved by Thm. 1. $\square$

C.5. Proof of Theorem 3

Proof. Due to the soundness of the proposed rules, it suffices to prove the result by showing that the proposed rules can transform the same marks as those by the k-step algorithm based on Alg. 2. Further, it suffices to show that when we incorporate the local transformation of $X = (V_{i+1})$ into $\mathbb{M}_i$ by Alg. 2, all the marks can also be transforemd by the proposed rules.

Revisit Alg. 2, the transformation in the third step follows $\mathcal{R}_7$, and the transformation in the fourth step directly follows the rules. Hence it suffices to prove that the transformation in the second step can also follow the proposed rules or directly follow the local transformation.

There are three parts in the second step of Alg. 2. (A) For any $K \in \text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}])$ and any $T \in \mathbf{C}$ such that $K \circ\!\!-\!\!* T$ in $\mathbb{M}_i$, orient $K \leftarrow\!\!* T$ (the mark at T remains). (B) For any edges with circle between $Z_1, Z_2 \in \mathbf{Z}$, orient the circle into a tail. (C) For all $K \in \text{PossDe}(\mathbf{Z}, \mathbb{M}_i[-\mathbf{C}]) \backslash \mathbf{Z}$ and $Z \in \mathbf{Z}$ such that $Z \circ\!\!-\!\!* K$, orient $Z -\!\!* K$. We consider them one by one in the following.

For (A), when $K = X$, the transformation of $K \leftarrow T$ directly follows the local transformation of $\mathbf{C}$. When $K \neq X$, there exists a minimal possible directed path from X to K as $\langle F_0(= X), F_1, \cdots, F_{n-1}, F_n(= T)$ according to Lemma 3. In this case, $F_0 -\!\!* F_1$ can be transformed according to the local transformation of $\mathbf{C}$. And since $\mathbf{C} \neq \emptyset$, there is always $F_0 \rightarrow F_1$ which can be transformed by $\mathcal{R}_{11}$. And by $\mathcal{R}_1$, there is $F_0 \rightarrow \cdots \rightarrow F_n$. Due to $T *\!\!\rightarrow X$ and $X \rightarrow F_1$, there is always $T *\!\!\rightarrow F_1$ which could be transformed by $\mathcal{R}_2$. Similarly, there is always $T *\!\!\rightarrow F_2, F_3, \cdots, F_n$ which could be transformed by $\mathcal{R}_2$. Hence $K \leftarrow\!\!* T$ can be transformed by the proposed rules.

For (B), according to the definition of $\mathbf{Z}$, there is a path $Z_1 -\!\!\circ \cdots F_1 -\!\!\circ X$. Hence there is $F_1 - X$ according to the local transformation represented by $\mathbf{C}$. Hence there is always an undirected path $Z_1 - \cdots - F_1 - X$ which could be transformed by $\mathcal{R}_6$. Similarly, there is always an undirected path from Z_2 to X which could be transformed by $\mathcal{R}_6$. In this case, as there is an edge between Z_1 and Z_2, there is always an edge $Z_1 - Z_2$ which could be transformed by $\mathcal{R}_6$.

For (C), with the similar process of the case for (B), we can conclude that there is always an undirected path from Z to X which could be transformed by $\mathcal{R}_6$. In this case, if there is an edge $Z \circ\!\!-\!\!* K$, there is always an edge $Z -\!\!* K$ which could be transformed by $\mathcal{R}_6$.

The proof completes. $\square$