TECHNICAL REPORT
R-149
May, 2026

# Causal Discovery over Clusters of Variables in Non-Markovian Systems

**Tara V Anand**
Department of Biomedical Informatics
Columbia University
tara.v.anand@columbia.edu

**Adèle H Ribeiro**
Institute of Medical Informatics
University of Münster
adele.ribeiro@uni-muenster.de

**Jin Tian**
Mohamed bin Zayed University
of Artificial Intelligence
jin.tian@mbzuai.ac.ae

**George Hripcsak**
Department of Biomedical Informatics
Columbia University
gh13@cumc.columbia.edu

**Elias Bareinboim**
Causal Artificial Intelligence Laboratory
Columbia University
eb@cs.columbia.edu

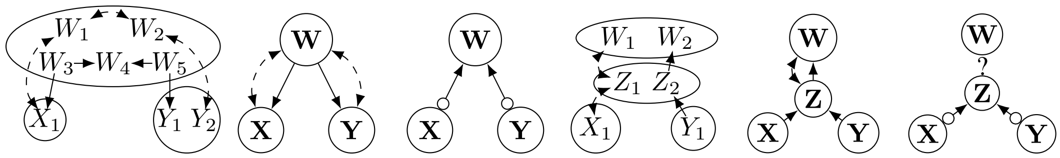
## Abstract

Causal discovery is the task of leveraging observational data to uncover causal relationships between variables. Recent work has extended these methods to operate over clusters of variables to improve scalability in high-dimensions and enable reasoning over higher-level entities. These approaches have been limited by strong assumptions including causal sufficiency. In this work, we introduce an approach for causal discovery over clusters in non-Markovian systems. First, we extend theory of graphical models in a knowledge-based context, to motivate introduction of a novel graphical equivalence class that can accommodate unobserved confounding. Then, we present a sound algorithm for causal discovery of learnable relationships between clusters of variables.

## 1 Introduction

Causal discovery is the task of using observational data to learn about causal relationships between variables [11, 17]. In causal inference, assumptions about the data-generating process are commonly represented with a causal diagram, a type of acyclic directed mixed graph (ADMG) [12, 15, 4, 13]. Because multiple causal diagrams can encode the same conditional independences, causal discovery aims to uncover a graphical Markov equivalence class of possible models, which encodes the constraints found in the data. Constraint-based methods assume faithfulness, that the observed independences reflect the true graphical structure of the model. Some algorithms, like PC [17, 7], also assume causal sufficiency, that there is no unobserved confounding in the system. When latent confounding may be present, alternative approaches are needed such as FCI [16, 22] where the associated Markov equivalence class is called a partial ancestral graph (PAG).

Existing causal discovery algorithms face scalability limitations and operate over individual variables, which may not correspond to domain-relevant entities, for example, pixels versus objects in an image. Therefore, causal discovery over clusters of variables is of interest, not only to learn relationships over semantically meaningful entities, but also as a strategy to reduce effective dimensionality. Prior work introduced cluster directed acyclic graphs (C-DAGs) to represent knowledge-based




(a) $G_1$ (b) $G_{\mathbf{C}_1}$ (c) $\mathcal{P}_{G_1}$ (d) $G_2$ (e) $G_{\mathbf{C}_2}$ (f) $\mathcal{P}_{\mathbf{C}_2}$

Figure 1: $(a)$: an ADMG with indicated clusters. $(b)$: a C-DAG of which $G_1$ is in the class. $(c)$: an attempted graphical equivalence class for $(b)$. $(d)$: an ADMG with marked clusterings. $(e)$:the C-DAG corresponding to $(d)$. $(f)$: an attempted representation over clusters where arrowheads are used to indicate independence information; the differing roles of $\mathbf{Z}$ in the triplets $\mathbf{X}, \mathbf{Z}, \mathbf{W}$ and $\mathbf{Y}, \mathbf{Z}, \mathbf{W}$ make such a representation impossible.

relationships between user-defined clusters [1], and $\alpha$C-DAGs add independence arcs to explicitly denote conditional independences and inform discovery under causal sufficiency by the algorithm CLOC [2]. Other related work focuses on leveraging clusters as intermediates for variable-level causal discovery, or on other strong assumptions of structure internal to and between clusters (e.g. all variables in adjacent clusters are connected by the same edge types) [18, 10, 5, 6, 21, 8, 3, 14, 9, 20, 19]. While $\alpha$C-DAGs and CLOC [2] address causal discovery over clusters, reliance on Markovian assumptions limits feasibility in real-world settings and the representation and algorithm break down in the presence of latent confounding. This paper generalizes those contributions to the non-Markovian case by introducing new graphical semantics, new graphical equivalence classes and a corresponding discovery algorithm.

We illustrate the need for a novel causal discovery algorithm over clusters in non-Markovian settings through observing the result from naively applying FCI and CLOC over clusters with latent confounding. Consider the C-DAG $G_{\mathbf{C_1}}$ in Figure 1, where $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}$ and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y} | \mathbf{W}$. $G_1$ illustrates an ADMG with demarcated clusters that is in the class of $G_{\mathbf{C_1}}$. $G_1$ encodes that $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$, which is not perceptible in the C-DAG. $\mathcal{P}_{G_1}$ uses the notation of PAGs to represent the independence information encoded in $G_1$, however, this representation contradicts the true edge orientations and implied ancestrality. Specifically, $\mathbf{W}$ appears as a non-ancestor of $\mathbf{X}$ and $\mathbf{Y}$, contradicting the true relationships of $G_1$, where $\mathbf{W}$ is in fact an ancestor of both $\mathbf{X}$ and $\mathbf{Y}$. Applying FCI with tests over the sets of variables in each cluster would result in $\mathcal{P}_{G_1}$ with the application of rule 0 (collider search). This example illustrates how FCI can incorrectly infer ancestrality when applied directly to clusters. If CLOC, the causal discovery algorithm for Markovian clusters, were to be applied in this case, the resulting graph would be similar, with arrowheads incorrectly oriented towards $\mathbf{W}$ (and tails at $\mathbf{X}$ and $\mathbf{Y}$). Causal diagram 1d, associated C-DAG 1e and incomplete equivalence class 1f illustrate another example of unlearnable edges between clusters, discussed further in Appendix B.1. When the goal is to graphically represent an equivalence class over clusters, edge orientations fail to accurately encode both (in)dependence structure and ancestrality.

**Summary of Contributions.** We introduce a novel graphical equivalence class that captures both independence information and ancestrality for clusters of variables in the presence of latent confounding. We define requisite preliminary graphs, and propose a sound algorithm to learn this equivalence class from observational data. Specifically we contribute the following:

1. In section 2, we extend $\alpha$C-DAGs for latent confounding (Definition 2.2), extend the d-separation criterion (Definition 2.3), and prove it is sound and complete (Theorem 2.4).
2. In section 3, we define cluster maximal ancestral graphs ($\alpha$C-MAGs), which capture ancestral relationships among clusters while preserving independence information. We further define cluster-partial ancestral graphs ($\alpha$C-PAGs) to represent a Markov equivalence class of $\alpha$C-MAGs (Definition 3.11).
3. Finally, in section 4 we introduce a sound causal discovery algorithm over clusters, LC-CLOC, to learn an $\alpha$C-PAG from data, evaluated with experiments in section 5.

### 1.1 Preliminaries

**Notation.** A single variable is denoted by a (non-boldface) uppercase letter $X$ and its realized value by a small letter $x$. A boldfaced uppercase letter $\mathbf{X}$ denotes a set (or a cluster) of variables. We use kinship relations, defined via edges in the graph. We denote by $Pa(\mathbf{X})_G$, $Ch(\mathbf{X})_G$, $An(\mathbf{X})_G$, and

$De(\mathbf{X})_G$, the sets of parents, children, ancestors, and descendants in graph $G$, respectively. A triplet $\langle V_i, V_k, V_j \rangle$ is *active* if 1) $V_k$ is a collider and $V_k$ or any of its descendants are in $\mathbf{Z}$ or 2) $V_k$ is a non-collider and is not in $\mathbf{Z}$. A path $p$ is *active* given (or conditioned on) $\mathbf{Z}$ if every triplet on $p$ is active relative to $\mathbf{Z}$. Otherwise, $p$ is said to be *inactive*. Given a graph $G$, $\mathbf{X}$ and $\mathbf{Y}$ are d-separated by $\mathbf{Z}$ if every path between $\mathbf{X}$ and $\mathbf{Y}$ is inactive given $\mathbf{Z}$. We denote this d-separation by $(\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z})_G$.

**Learned Equivalence Classes.** A partial ancestral graph (PAG) $\mathcal{P}$ can have edges of type $\circ\!\!-\!\!\circ$, $\circ\!\!\rightarrow$, $\rightarrow$, and $\leftrightarrow$. Circle marks indicate orientations that vary across the equivalence class. Arrowhead or tail marks are common for all members of the class represented by the PAG. A consecutive triplet of vertices $\langle X, Z, Y \rangle$ is *unshielded* if $X$ and $Y$ are not adjacent. If $X$ and $Y$ are adjacent, the triplet is *shielded*. In a triplet $\langle X, Z, Y \rangle$, $Z$ is a definite collider if edges from $X$ and $Y$ are into it ($X \ast\!\!\rightarrow Z \leftarrow\!\!\ast Y$). $Z$ is a definite non-collider if at least one edge is out of it or both edges have circle marks at $Z$ and the triplet is unshieleded. Otherwise, $Z$ has a non-definite status.

$\alpha$**C-DAG (Markov)** [2] Given a DAG $G(\mathbf{V}, \mathbf{E})$ and partition $\mathbf{C} = \{\mathbf{C}_1, \ldots, \mathbf{C}_n\}$ of $\mathbf{V}$, construct a graph $G_{\mathbf{C}}(\mathbf{C}, \mathbf{E}_{\mathbf{C}}, \boldsymbol{\mathcal{A}})$ over $\mathbf{C}$ as follows: An edge $\mathbf{C}_i \rightarrow \mathbf{C}_j$ is in $\mathbf{E}_{\mathbf{C}}$ if $\exists$ some $V_i \in \mathbf{C}_i$ and $V_j \in \mathbf{C}_j$ such that $V_i \in Pa(V_j)$ in $G$; The set of independence arcs $\boldsymbol{\mathcal{A}}$ is defined over all triplets $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$, by Def. A.3. For each arc trajectory, separation marks are added according to Def. A.6. For each path, connection marks are added according to Def. A.7. If for all pairs of clusters $\mathbf{C}_i, \mathbf{C}_j$ where there exists an edge $\mathbf{C}_i \rightarrow \mathbf{C}_j$, there is no directed path from $\mathbf{C}_j$ to $\mathbf{C}_i$, then $\mathbf{C}$ is an *admissible partition* of $\mathbf{V}$ and $G_{\mathbf{C}}$ is a *cluster DAG with independence arcs*, or an $\alpha$*C-DAG*, compatible with $G$. Def. A.8 details d-separation criteria over $\alpha$C-DAGs. More background in Appendix (A).

## 2 $\alpha$C-DAGs: Generalized for Latent Confounding

$\alpha$C-DAGs [2] use independence arcs, drawn per graphical triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$, and separation/connection marks to explicitly represent independence relationships, while edges denote causal relationships (see Appendix A for details). In a Markov context, there are three independence arcs: marginally-connecting, conditionally-connecting, and never-connecting. When latent confounding exists, the independence arcs/marks must be revised. In this section, we amend the graphical notation and extend reasoning tools for $\alpha$C-DAGs with latent confounding as a prerequisite for defining the Markov equivalence class under latent confounding and motivating our causal discovery approach. Figure 2 illustrates the four independence arcs needed in a non-Markov context, formalized in Definition 2.1. The new always-connecting independence arc represents, for a triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$, when $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}; \mathbf{X} \not\perp\!\!\!\perp \mathbf{Y} | \mathbf{Z}$. The asterisks denote that edges may have either arrowheads or tails.

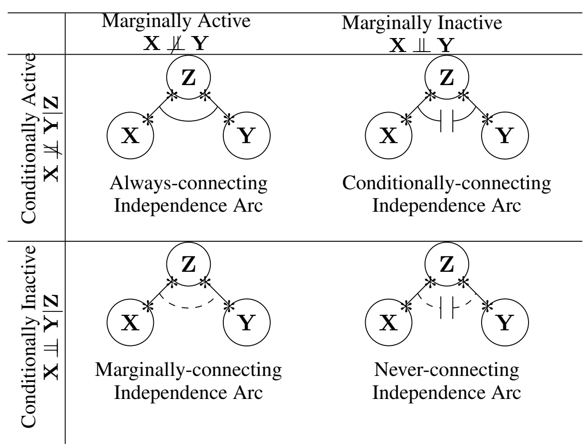



Figure 2: Independence arcs and associated marginal/conditional (in)dependences.

**Definition 2.1** (Independence Arcs)**.** *Consider a graph $\mathbf{G}_C$ over clusters $\mathbf{C} = \langle \mathbf{C}_0, ..., \mathbf{C}_n \rangle$. For any unshielded triplet $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$ (or manipulated unshielded triplet (Def. A.2) $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle^{-\mathbf{C}_i \mathbf{C}_j}$), let $\mathbf{S}$ equal a (possibly empty) set of clusters $\mathbf{S} \subset (\mathbf{C} \setminus \{\mathbf{C}_i, \mathbf{C}_j\})$ that is a minimal separating set, such that $\mathbf{C}_i \perp\!\!\!\perp \mathbf{C}_j | \mathbf{S}$, if such a set exists. For a triplet $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$, the independence arc, $\mathcal{A}_{\mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j} \in \boldsymbol{\mathcal{A}}$, that is drawn from some point on the edge between $\mathbf{C}_i$ and $\mathbf{C}_k$ to some point on the edge between $\mathbf{C}_j$ and $\mathbf{C}_k$, can be determined in the following way:*

1. *marginally-connecting iff $\mathbf{C}_k \in \mathbf{S}$. Consequently, $\mathbf{C}_i \not\perp\!\!\!\perp \mathbf{C}_j | \mathbf{S} \setminus \mathbf{C}_k$ and $\mathbf{C}_i \perp\!\!\!\perp \mathbf{C}_j | \mathbf{S}$.*
2. *conditionally-connecting iff $\mathbf{C}_k \notin \mathbf{S}$ and $\mathbf{C}_i \not\perp\!\!\!\perp \mathbf{C}_j | \mathbf{S} \cup \mathbf{C}_k$*
3. *never-connecting iff $\mathbf{C}_k \notin \mathbf{S}$ and $\mathbf{C}_i \perp\!\!\!\perp \mathbf{C}_j | \mathbf{S} \cup \mathbf{C}_k$*
4. *always-connecting iff no set $\mathbf{S}$ exists. Consequently, $\mathbf{C}_i \not\perp\!\!\!\perp \mathbf{C}_j$ and $\mathbf{C}_i \not\perp\!\!\!\perp \mathbf{C}_j | \mathbf{C}_k$.*

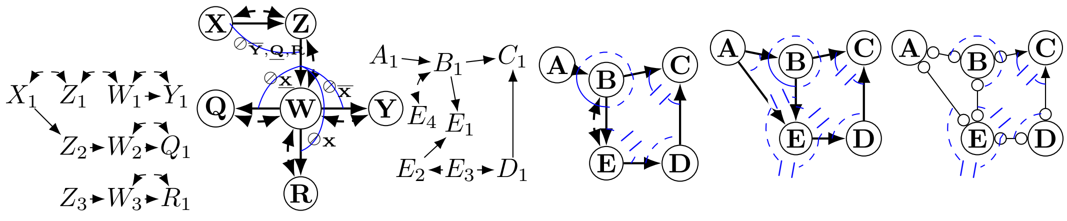



Figure 3: $G_{\mathbf{C}_1}$ is an $\alpha$C-DAG (with independence arcs between R, W, Q, and Y omitted for clarity) illustrating the three kinds of separation marks that can annotate an always-connecting arc. $G_1$ is a compatible ADMG. Independence arcs are shown in blue. $G_2$ is an ADMG, $G_{\mathbf{C}_2}$ its corresponding $\alpha$C-DAG, $\mathcal{M}_{\mathbf{C}_2}$ the $\alpha$C-MAG corresponding to $G_{\mathbf{C}_2}$, and $\mathcal{P}_{\mathbf{C}_2}$ the corresponding $\alpha$C-PAG.

*Shielded triplets are annotated according to the behavior of their respective manipulated triplets.*

For each triplet in a graph, exactly one criterion applies, so each triplet is annotated by a single independence arc.

As is in the Markov case, independence arcs alone do not fully represent all conditional (in)dependences. Separation marks and connection marks complete the representation, and are retained in the non-Markovian case. Connection marks apply unchanged, while separation marks require modification in the presence of latent confounding. For always-connecting arcs, three cases require separation marks: 1) only the marginal connection is disputed, 2) only the conditional connection is disputed, or 3) both connections are disputed. Case 1 is represented by a separation mark with an overlined subscript, $\oslash_{\overline{\mathbf{X}}}$. Case 2 uses an underlined subscript $\oslash_{\underline{\mathbf{X}}}$. Case 3 uses the standard notation, $\oslash_{\mathbf{X}}$. This albeit complex notation is introduced for completeness, since $\alpha$C-DAGs prove useful in their own right. However, many tasks do not require this complete representation. The formal definition of separation marks under latent confounding is given in Appendix B.2, Definition B.1. An example is below.

**Example 1.** *Consider Figure 3, where $G_{\mathbf{C}_1}$ is an $\alpha$C-DAG, and $G_1$ is in its class. Triplets $\langle \mathbf{X}, \mathbf{Z}, \mathbf{W} \rangle$, $\langle \mathbf{Z}, \mathbf{W}, \mathbf{Q} \rangle$, $\langle \mathbf{Z}, \mathbf{W}, \mathbf{R} \rangle$, and $\langle \mathbf{Z}, \mathbf{W}, \mathbf{Y} \rangle$ are annotated with always-connecting arcs. There is a collider path and no connecting path over variables between $\mathbf{X}$ and $\mathbf{Y}$ so separation marks with overlined subscripts are placed on $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{W}}$ and $\mathcal{A}_{\mathbf{Z},\mathbf{W},\mathbf{Y}}$. There is a connecting path and no collider path over variables between $\mathbf{X}$ and $\mathbf{Q}$, so separation marks with underlined subscripts are placed on $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{W}}$ and $\mathcal{A}_{\mathbf{Z},\mathbf{W},\mathbf{Q}}$. There are no paths over variables between $\mathbf{X}$ and $\mathbf{R}$, so separation marks with standard subscripts are placed on $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{W}}$ and $\mathcal{A}_{\mathbf{Z},\mathbf{W},\mathbf{R}}$.*

With these revisions, we now provide a complete definition for non-Markov $\alpha$C-DAGs.

**Definition 2.2** ($\alpha$C-DAGs with Latent Confounding)**.** *Given an ADMG $G(\mathbf{V}, \mathbf{E})$ and a partition $\mathbf{C} = \{\mathbf{C}_1, \ldots, \mathbf{C}_k\}$ of $\mathbf{V}$, construct a graph $G_{\mathbf{C}}(\mathbf{C}, \mathbf{E}_{\mathbf{C}}, \boldsymbol{\mathcal{A}})$ over $\mathbf{C}$. The set of edges $\mathbf{E}_{\mathbf{C}}$ is defined as follows: An edge $\mathbf{C}_i \rightarrow \mathbf{C}_j$ is in $\mathbf{E}_{\mathbf{C}}$ if there exists some $V_i \in \mathbf{C}_i$ and $V_j \in \mathbf{C}_j$ such that $V_i \in Pa(V_j)$ in $G$; A dashed bidirected edge $\mathbf{C}_i \dashleftarrow\dashrightarrow \mathbf{C}_j$ is in $\mathbf{E}_{\mathbf{C}}$ if there exists some $V_i \in \mathbf{C}_i$ and $V_j \in \mathbf{C}_j$ such that $V_i \dashleftarrow\dashrightarrow V_j$ in $G$. The set of independence arcs $\boldsymbol{\mathcal{A}}$ is defined over the set of triplets $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$, according to Definition 2.1. For each arc trajectory in $G_{\mathbf{C}}$, separation marks are added according to Definition B.1. For each path in $G_{\mathbf{C}}$, connection marks are added according to Definition A.7. If for all pairs of clusters $\mathbf{C}_i, \mathbf{C}_j$ where there exists an edge $\mathbf{C}_i \rightarrow \mathbf{C}_j$, there is no directed path $\mathbf{C}_j \rightarrow \ldots \rightarrow \mathbf{C}_i$, then we say that $\mathbf{C}$ is an* admissible partition *of $\mathbf{V}$. We then call $G_{\mathbf{C}}$ a* cluster-DAG with independence arcs*, or an* $\alpha$C-DAG*, compatible with $G$.*

The additional semantics for the non-Markov case also motivate a revised d-separation criterion, which we show is sound and complete in $\alpha$C-DAGs with latent confounding.

**Definition 2.3** (d-separation over $\alpha$C-DAGs with Latent Confounding.)**.** *A path $p_{\mathbf{C}}$ in an $\alpha$-C-DAG, $G_{\mathbf{C}}$, is said to be d-separated (or blocked) by a set of clusters $\mathbf{Z} \subset \mathbf{C}$ if and only if its corresponding arc trajectory $\mathbf{a}$ (Def. A.4) contains an independence arc $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ that is:*

1. *an always-connecting arc and (a) there exists a separation mark $\oslash_{\mathbf{C}_x}$ on $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ where $\mathbf{C}_x$ is on $p_{\mathbf{C}}$, (b) there exists a separation mark $\oslash_{\mathbf{C}_{\overline{x}}}$ on $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ where $\mathbf{C}_x$ is on $p_{\mathbf{C}}$ and $\mathbf{C}_k \notin \mathbf{Z}$, or (c) there exists a separation mark $\oslash_{\mathbf{C}_{\underline{x}}}$ on $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ where $\mathbf{C}_x$ is on $p_{\mathbf{C}}$ and $\mathbf{C}_k \in \mathbf{Z}$,*

2. *a marginally-connecting arc and (a)* $\mathbf{C}_k$ *is in* $\mathbf{Z}$ *or (b) there exists a separation mark* $\oslash_{\mathbf{C}_x}$ *on* $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ *where* $\mathbf{C}_x$ *is on* $p_{\mathbf{C}}$*,*
3. *a conditionally-connecting arc and (a)* $\mathbf{C}_k$ *is not in* $\mathbf{Z}$ *nor is any true descendant* $\mathbf{C}_d$ *of* $\mathbf{C}_k$ *(with directed and connection path* $\mathbf{C}_k \rightarrow \ldots \rightarrow \mathbf{C}_d$*) in* $\mathbf{Z}$*, (b) there exists a separation mark on* $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ $\oslash_{\mathbf{C}_x}$ *where* $\mathbf{C}_x$ *is on* $p_{\mathbf{C}}$*, or (c) for any connection mark* $\oplus_{\mathbf{C}_x}$ *on* $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$*,* $\mathbf{C}_x$ *is not in* $\mathbf{Z}$*, or*
4. *a never-connecting arc and for any connection mark* $\oplus_{\mathbf{C}_x}$ *on* $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_j,\mathbf{C}_k}$*,* $\mathbf{C}_x \notin \mathbf{Z}$*.*

*A set of clusters* $\mathbf{Z}$ *is said to d-separate two sets of clusters* $\mathbf{X}, \mathbf{Y} \subset \mathbf{C}$*, denoted by* $(\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z})_{G_{\mathbf{C}}}$*, if and only if* $\mathbf{Z}$ *blocks every path from a cluster in* $\mathbf{X}$ *to a cluster in* $\mathbf{Y}$*.*

**Theorem 2.4.** *[Soundness and completeness of d-separation in αC-DAGs with latent confounding.] In an αC-DAG* $G_{\mathbf{C}}$*, let* $\{\mathbf{X}, \mathbf{Z}, \mathbf{Y}\} \in \mathbf{C}$*.* $\mathbf{X}$ *and* $\mathbf{Y}$ *are d-separated by* $\mathbf{Z}$ *in* $G_{\mathbf{C}}$*, if and only if for any ADMG, G, compatible with* $G_{\mathbf{C}}$*,* $\mathbf{X}$ *and* $\mathbf{Y}$ *are d-separated by* $\mathbf{Z}$ *in G.*

## 3 Graphical Equivalence Classes of $\alpha$C-DAGs: $\alpha$C-MAGs and $\alpha$C-PAGs

The goal of our causal discovery algorithm is to recover a graphical equivalence class over clusters of $\alpha$C-DAGs with the same induced independence structure and ancestral relationships. This corresponds to a cluster analogue of a partial ancestral graph (PAG), which we call a cluster-PAG, or $\alpha$C-PAG. We first define cluster maximal ancestral graphs ($\alpha$C-MAGs) as a precursor to $\alpha$C-PAGs.

$\alpha$C-MAGs represent classes of $\alpha$C-DAGs that share the same partition of variables, conditional independences, and ancestral relationships over clusters. We assume no selection bias. As in maximal ancestral graphs (MAGs), $\alpha$C-MAGs have at most one edge between vertices that is either directed ($\rightarrow$) or bidirected ($\leftrightarrow$). Each unshielded triplet is labeled with one of three independence arcs: marginally-connecting, conditionally-connecting, or never-connecting. Always-connecting arcs in an $\alpha$C-DAG induce shielded triplets in the corresponding $\alpha$C-MAG. For shielded triplets, independence arcs are determined by the composite manipulated unshielded triplets. Separation and connection marks are added as needed. A *true directed cycle* occurs when $\mathbf{C}_B \rightarrow \mathbf{C}_A$ is in $G_{\mathbf{C}}$ and $\mathbf{C}_A \in An_{G_{\mathbf{C}}}(\mathbf{C}_B)$. A *true almost directed cycle* occurs when $\mathbf{C}_B \leftrightarrow \mathbf{C}_A$ is in $G_{\mathbf{C}}$ and $\mathbf{C}_A \in An_{G_{\mathbf{C}}}(\mathbf{C}_B)$. Ancestrality requires a directed path with marginally- or always-connecting arcs and no separation marks. For $\alpha$C-MAGs, *m-separation* extends from d-separation, determined by adjacencies, independence arcs, and their marks. The construction of an $\alpha$C-MAG relies on the concept of a *primitive inducing path*, defined for cluster graphs, below, with the $\alpha$C-MAG construction procedure following in Def. 3.3 and an example construction in Appendix B.3.

**Definition 3.1** (Primitive Inducing Path over Clusters.)**.** *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be two clusters with some path* $p_c$ *between them. Let* $p_v$ *and* $p'_v$ *be two possibly distinct paths over variables such that* $p_v$ *(and* $p'_v$*) are analogous (Def. A.5) to* $p_c$*. If each non-endnode variable on* $p_v$ *is a collider, and each variable on* $p'_v$ *is an ancestor of a variable in either* $\mathbf{X}$ *or* $\mathbf{Y}$*, then* $p_c$ *is a* primitive inducing path over clusters.

**Remark 3.2.** *In αC-DAGs, triplet* $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ *with an always-connecting arc forms a primitive inducing path.*

**Definition 3.3** (Procedure for construction of $\alpha$Cluster Mixed Ancestral Graphs ($\alpha$C-MAGs))**.**
***Input:*** *αC-DAG,* $G_{\mathbf{C}}$ *over partition* $\mathbf{C} = \langle \mathbf{C}_1, \ldots, \mathbf{C}_n \rangle$
***Output:*** *αC-MAG,* $\mathcal{M}_{G_{\mathbf{C}}}$

1. $\mathbf{C}_i, \mathbf{C}_j \in \mathbf{C}$ *are adjacent in* $\mathcal{M}_{G_{\mathbf{C}}}$ *iff a primitive inducing path exists between them in* $G_{\mathbf{C}}$*.*
2. *For each pair of adjacent clusters* $\mathbf{C}_i, \mathbf{C}_j \in \mathcal{M}_{G_{\mathbf{C}}}$*, orient the edge as:*
   (a) $\mathbf{C}_i \rightarrow \mathbf{C}_j$ *if* $\mathbf{C}_i \in An_G(\mathbf{C}_j)$ *and* $\mathbf{C}_j \notin An_G(\mathbf{C}_i)$
   (b) $\mathbf{C}_i \leftrightarrow \mathbf{C}_j$ *if* $\mathbf{C}_i \notin An_G(\mathbf{C}_j)$ *and* $\mathbf{C}_j \notin An_G(\mathbf{C}_i)$
3. *For each (possibly manipulated) unshielded triplet* $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$ *in* $\mathcal{M}_{G_{\mathbf{C}}}$*:*
   (a) *if present in* $G_{\mathbf{C}}$*, add the independence arc* $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ *from* $G_{\mathbf{C}}$ *with any separation or connection marks*
   (b) *Otherwise:*
      i. *if manipulated, assign a never-connecting arc*
      ii. *else, consider all paths* $p = \langle \mathbf{C}_i, \ldots, \mathbf{C}_k, \mathbf{C}_j \rangle \in G_{\mathbf{C}}$ *where* $p' = \langle \mathbf{C}_i, \mathbf{C}_k \rangle \in \mathcal{M}_{G_{\mathbf{C}}}$*.*

*A. If $\forall p$, $\mathbf{C}_i \perp\!\!\!\perp \mathbf{C}_j$ and $\mathbf{C}_i \perp\!\!\!\perp \mathbf{C}_j|\mathbf{C}_k$, $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ is never-connecting .*
*B. If $\forall p$, $\mathbf{C}_i \perp\!\!\!\perp \mathbf{C}_j$ and $\exists p' \mathbf{C}_i \not\perp\!\!\!\perp \mathbf{C}_j|\mathbf{C}_k$, $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ is conditionally-connecting.*
*C. If $\exists p'$, $\mathbf{C}_i \not\perp\!\!\!\perp \mathbf{C}_j$ and $\forall p \mathbf{C}_i \perp\!\!\!\perp \mathbf{C}_j|\mathbf{C}_k$, $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ is marginally-connecting.*

4. *Add any new marks by definitions B.1 and A.7.*

MAGs are ancestral and maximal. We extend these properties to clusters and $\alpha$C-MAGs.

**Definition 3.4** (Ancestrality in $\alpha$C-MAGs)**.** *An $\alpha$C-MAG is ancestral if it contains 1) no* true directed cycles *and 2) no* true almost directed cycles*.*

**Definition 3.5** (Maximality in $\alpha$C-MAGs)**.** *An $\alpha$C-MAG is maximal if there exists a set of clusters that m-separates every pair of non-adjacent clusters.*

Maximality is closely related to inducing paths. If an inducing path exists between $\mathbf{X}$ and $\mathbf{Y}$, no set m-separates them, so an edge between them must be present in the $\alpha$C-MAG.

**Remark 3.6.** *An $\alpha$C-MAG cannot have an always-connecting arc on an unshielded triplets.*

An $\alpha$C-PAG will ultimately be defined to represent the Markove equivalence class of $\alpha$C-MAGs. For MAGs over variables, two graphs are Markov equivalent if they encode the same m-separations. We extend a similar notion of Markov equivalence over clusters for $\alpha$C-MAGs.

**Definition 3.7** (Cluster Markov equivalence)**.** *Two $\alpha$C-MAGs $\mathcal{M}_1$, $\mathcal{M}_2$ (with the same partition $\mathbf{C}$ and variables $\mathbf{V}$) are cluster Markov equivalent if for any three disjoint sets of clusters ($\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$), $\mathbf{X}$ and $\mathbf{Y}$ are m-separated by $\mathbf{Z}$ in $\mathcal{M}_1$ if and only if $\mathbf{X}$ and $\mathbf{Y}$ are m-separated by $\mathbf{Z}$ in $\mathcal{M}_2$.*

Since m-separation in $\alpha$C-MAGs is determined by independence arcs, separation marks, and connection marks, we obtain the following proposition.

**Theorem 3.8.** *Two $\alpha$C-MAGs $\mathcal{M}_1$, $\mathcal{M}_2$ (with the same partition $\mathbf{C}$ of the same set of variables $\mathbf{V}$) are cluster Markov equivalent iff they share the same adjacencies, independence arcs on unshielded and discriminated manipulated unshielded triplets, and separation and connection marks.*

The definition of a discriminated manipulated unshielded triplet follows the definition of a discriminating path over clusters, as below.

**Definition 3.9** (Discriminating path over clusters)**.** *In an $\alpha$C-MAG, a path between $\mathbf{X}$ and $\mathbf{Y}$, $p = \langle \mathbf{X}, ..., \mathbf{W}, \mathbf{V}, \mathbf{Y} \rangle$, is a discriminating path for $\mathbf{V}$ if $p$ includes at least three edges; $\mathbf{V}$ is a non-endpoint vertex on $p$ and is adjacent to $\mathbf{Y}$ on $p$; $\mathbf{X}$ is not adjacent to $\mathbf{Y}$, and for every vertex $\mathbf{Q}_i$ between $\mathbf{X}$ and $\mathbf{V}$, $\mathcal{A}_{\mathbf{X},\mathbf{Q}_i,\mathbf{Y}}$ is a marginally-connecting arc with no separation marks subscripted with a vertex on $p$ and $\mathcal{A}_{\mathbf{X},\mathbf{Q}_i,\mathbf{Q}_{i+1}}$ is a conditionally-connecting arc with no separation marks subscripted with a vertex on $p$.*

**Definition 3.10** (Discriminated Manipulated Unshielded Triplets (Non-Markov))**.** *For a shielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ where $\mathbf{Y}$ is adjacent to some $\mathbf{W}$ and $\mathbf{X}$ and $\mathbf{Z}$ are not adjacent to $\mathbf{W}$, if $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is conditionally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, or if $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is marginally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, then $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}^{-\mathbf{XY}}}$ is a discriminated manipulated unshielded triplet. If $p = \langle \mathbf{X}, ..., \mathbf{W}, \mathbf{V}, \mathbf{Y} \rangle$ is a discriminating path for $\mathbf{V}$, $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}^{-\mathbf{WY}}}$ is a discriminated manipulated unshielded triplet.*

When all clusters are singletons containing just one variable, these definitions reduce to standard Markov equivalence for MAGs, where equivalence is characterized by adjacencies, unshielded colliders, and discriminating paths.

We now define $\alpha$C-PAGs. Similarly to PAGs over variables, $\alpha$C-PAGs may have edges of type ∘–∘, ∘→, →, and ↔. Circle marks indicate orientations that vary across the equivalence class. As with $\alpha$C-MAGs, $\alpha$C-PAGs are annotated by independence arcs, separation marks, and connection marks. Section 2 shows that independence arcs do not generally constrain edge orientations. The exception is when $\mathbf{Z}$ in a triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ is a singleton, where orientation determination reduces to the variable case. Accordingly, $\alpha$C-PAGs' cluster sizes, assessed by the partition can inform orientations.

**Definition 3.11** ($\alpha$Cluster Partial Ancestral Graph)**.** *Let $[\mathcal{G}_C]$ be the cluster Markov equivalence class of an $\alpha$C-MAG $\mathcal{G}_C$, and let $\mathbf{C} = \langle \mathbf{C}_1, ..., \mathbf{C}_n \rangle$ be the partition of variables into clusters. A cluster partial ancestral graph ($\alpha$C-PAG) for $[\mathcal{G}_C]$ is a graph $\mathcal{P}_C$ such that:*

1. *$\mathcal{P}_C$ has the same adjacencies as any member of $[\mathcal{G}_C]$*
2. *An independence arc is in $\mathcal{P}_C$ iff shared by all $\alpha$C-MAGs in $[\mathcal{G}_C]$*

3. $\mathcal{P}_C$ *has the same separation and connection marks as any member of* $[\mathcal{G}_C]$
4. *every non-circle orientation in* $\mathcal{P}$ *is an invariant mark in* $[\mathcal{G}_C]$.

*Every circle orientation in* $\mathcal{P}_C$ *corresponds to a variant orientation in* $[\mathcal{G}_C]$. $\mathcal{P}_C$ *is then a maximally informative* $\alpha$*C-PAG for* $[\mathcal{G}_C]$.

## 4 An Algorithm for Causal Discovery Over Clusters with Latent Confounding

With the definitions for a Markov equivalence class of $\alpha$C-DAGs with latent confounding established, an algorithm for constructing the graphical equivalence class can be developed. We call this algorithm *LC-CLOC*, for latent-confounding CLOC. In this proposed causal discovery algorithm over clusters in non-Markovian systems, there are three phases, as with CLOC and with similarities to FCI. Starting with a completely connected graph over clusters, with each edge of type $\circ\!\!-\!\!\circ$, in the first phase of the algorithm, edges between two nodes, $\mathbf{X}$ and $\mathbf{Y}$, are iteratively removed wherever it is possible to determine some separating set of clusters $\mathbf{S}$ such that $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}|\mathbf{S}$. This results in the skeleton of the graph. In the second phase, independence arcs are added by the arc determination rules, defined below, which follow from definitions 2.1 and 3.10, and connection and separation marks are added to independence arcs. In the final phase, logical rules inform orientation of edges. These rules include modifications of those from FCI at the variable level when relevant clusters are of size one, as well as modification of certain rules from CLOC. We note that having some clusters with only one variable is necessary for learning to occur.

**Remark 4.1.** *If all clusters contain more than one variable, it is not possible to learn any orientations.*

This is because when a cluster has more than one variable, there is no independence test that by itself determines orientations. However, independence tests paired with knowledge of a previously made orientation can inform subsequent orientations. Therefore, when some clusters are of size one, it is often possible for many orientations to be made. The rules for adding independence arcs to the skeleton are enumerated below.

**Independence arc determination rules:**

$\mathcal{AR}_1$: For an unshielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ draw a conditionally-connecting arc if $\mathbf{Z} \notin Sepset(\mathbf{X}, \mathbf{Y})$ and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Z} \cup SepSet(\mathbf{X}, \mathbf{Y})$.

$\mathcal{AR}_2$: For an unshielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ draw a marginally-connecting arc if $\mathbf{Z} \in Sepset(\mathbf{X}, \mathbf{Y})$.

$\mathcal{AR}_3$: For an unshielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ draw a never-connecting arc for remaining unmarked unshielded triplets.

Following application of these rules, separation and connection marks are added as applicable.

For a shielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ with $\mathbf{W} \circ\!\!-\!\!\circ \mathbf{Y}$ and where $\{\mathbf{X}, \mathbf{Z}\}$ and $\mathbf{W}$ are not adjacent, if $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is conditionally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, evaluate the following:

$\mathcal{AR}_4$: If $\mathbf{Z} \notin SepSet(\mathbf{X}, \mathbf{W})$, $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W}|\mathbf{Y} \cup SepSet(\mathbf{X}, \mathbf{W})$, and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W}|\mathbf{Z}, \mathbf{Y} \cup SepSet(\mathbf{X}, \mathbf{W})$, then draw an always-connecting arc for $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

$\mathcal{AR}_5$: If $\mathbf{Z} \notin SepSet(\mathbf{X}, \mathbf{W})$, and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W}|\mathbf{Z} \cup SepSet(\mathbf{X}, \mathbf{W})$, then draw a conditionally-connecting arc for $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

$\mathcal{AR}_6$: If $\mathbf{Z} \in SepSet(\mathbf{X}, \mathbf{W})$, then draw a marginally-connecting arc for $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

$\mathcal{AR}_7$: Draw a never-connecting arc for any unmarked $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

For a shielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ with $\mathbf{W} \circ\!\!-\!\!\circ \mathbf{Y}$ and where $\{\mathbf{X}, \mathbf{Z}\}$ and $\mathbf{W}$ are not adjacent, if $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is marginally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, evaluate the following:

$\mathcal{AR}_8$: If $\mathbf{Z} \notin SepSet(\mathbf{X}, \mathbf{W})$, $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W}|SepSet(\mathbf{X}, \mathbf{W}) \setminus Y$, and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W}|\mathbf{Z} \cup SepSet(\mathbf{X}, \mathbf{W}) \setminus Y$, then draw an always-connecting arc for $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

$\mathcal{AR}_9$: If $\mathbf{Z} \notin SepSet(\mathbf{X}, \mathbf{W})$, and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W}|\mathbf{Z} \cup SepSet(\mathbf{X}, \mathbf{Y})$, then draw a conditionally-connecting arc for $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

$\mathcal{AR}_{10}$: If $\mathbf{Z} \in SepSet(\mathbf{X}, \mathbf{W})$, then draw a marginally-connecting arc for $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

$\mathcal{AR}_{11}$: Draw a never-connecting arc for any unmarked $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

If $p = \langle \mathbf{X}, ..., \mathbf{W}, \mathbf{V}, \mathbf{Y} \rangle$ is a discriminating path for $\mathbf{V}$ where $\mathbf{Q}$ is the set of all nodes between $\mathbf{X}$ and $\mathbf{V}$, evaluate the following:

$\mathcal{AR}_{12}$: If $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}$ and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}, \mathbf{V}$, $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}-\mathrm{WY}}$ is always-connecting.

$\mathcal{AR}_{13}$: If $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}$ and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}, \mathbf{V}$, $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}-\mathrm{WY}}$ is marginally-connecting.

$\mathcal{AR}_{14}$: If $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}$ and $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}, \mathbf{V}$, $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}-\mathrm{WY}}$ is conditionally-connecting.

$\mathcal{AR}_{15}$: If $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}$ and $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}, \mathbf{V}$, $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}-\mathrm{WY}}$ is never-connecting.

Following the application of rules for placing the independence arcs and separation/connection marks, edges can be oriented using a combination of rules that apply to clusters of size 1 as well as clusters of any size. The complete list of orientation rules follows two additional definitions. Asterisks (e.g. $\ast\!\!-\!\!\ast$, $\ast\!\!\rightarrow$) indicate that any orientation may be possible.

**Definition 4.2** (Uncovered path)**.** *A path over clusters is uncovered if every consecutive triplet on the path is unshielded and not marked by a never-connecting arc.*

**Definition 4.3** (Possibly directed path)**.** *A path over clusters between $\mathbf{X}$ and $\mathbf{Y}$ is possibly directed (or causal) from $\mathbf{X}$ to $\mathbf{Y}$ if there is no arrowhead on the path pointing towards $\mathbf{X}$ and all triplets on the path are marked with a marginally-connecting arc with no separation marks subscripted with a node on the path.*

$\mathscr{R}_0$: For a triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$, if $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ is conditionally-connecting, and $|\mathbf{Z}| = 1$, then orient the triplet as $\mathbf{X} \ast\!\!\longrightarrow \mathbf{Z} \longleftarrow\!\!\ast \mathbf{Y}$.

$\mathscr{R}_1$: If $\mathbf{X} \ast\!\!\longrightarrow \mathbf{Z} \circ\!\!-\!\!\ast \mathbf{Y}$, $\mathbf{X}$ and $\mathbf{Y}$ are not adjacent, and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ is marginally-connecting, then orient the triplet as $\mathbf{X} \ast\!\!\longrightarrow \mathbf{Z} \rightarrow \mathbf{Y}$.

$\mathscr{R}_2$: If $\mathbf{X} \rightarrow \mathbf{Z} \ast\!\!\longrightarrow \mathbf{Y}$ or $\mathbf{X} \ast\!\!\longrightarrow \mathbf{Z} \rightarrow \mathbf{Y}$ and $\mathbf{X} \ast\!\!-\!\!\circ \mathbf{Y}$ and there are no never-connecting independence arcs on the composite manipulated unshielded triplets, then orient $\mathbf{X} \ast\!\!-\!\!\circ \mathbf{Y}$ as $\mathbf{X} \ast\!\!\longrightarrow \mathbf{Y}$.

$\mathscr{R}_3$: If $\mathbf{X} \ast\!\!\longrightarrow \mathbf{Z} \longleftarrow\!\!\ast \mathbf{Y}$, $\mathbf{X} \ast\!\!-\!\!\circ \mathbf{W} \circ\!\!-\!\!\ast \mathbf{Y}$, $\mathbf{X}$ and $\mathbf{Y}$ are not adjacent, $\mathbf{W} \ast\!\!-\!\!\circ \mathbf{Z}$, and $\mathcal{A}_{\mathbf{X},\mathbf{W},\mathbf{Y}}$ is marginally-connecting, then orient $\mathbf{W} \ast\!\!-\!\!\circ \mathbf{Z}$ as $\mathbf{W} \ast\!\!\longrightarrow \mathbf{Z}$.

$\mathscr{R}_4$: If $\mathbf{X} \circ\!\!-\!\!\circ \mathbf{Z} \circ\!\!-\!\!\circ \mathbf{Y}$ and $\mathbf{X} \longleftarrow\!\!\circ \mathbf{Y}$ and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}-\mathrm{XY}}$ is marginally-connecting or always-connecting, orient $\mathbf{X} \circ\!\!-\!\!\circ \mathbf{Z}$ as $\mathbf{X} \longleftarrow\!\!\ast \mathbf{Z}$.

$\mathscr{R}_5$: If $\mathbf{X} \ast\!\!-\!\!\ast \mathbf{Z} \ast\!\!-\!\!\ast \mathbf{Y}$, $\mathbf{Z} - \mathbf{W}$, $\mathbf{X}$ and $\mathbf{W}$ are not adjacent, $\mathbf{Y}$ and $\mathbf{W}$ are not adjacent, and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ is never-connecting or conditionally-connecting with connection mark $\oplus_{\boldsymbol{\mathcal{D}}}$ such that $\mathbf{W} \in \boldsymbol{\mathcal{D}}$, then orient $\mathbf{Z} - \mathbf{W}$ as $\mathbf{Z} \rightarrow \mathbf{W}$.

$\mathscr{R}_6$: If $\mathbf{X} \rightarrow \mathbf{Z} \rightarrow \mathbf{Y}$, $\mathbf{X} \circ\!\!\longrightarrow \mathbf{Y}$, and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}-\mathrm{XY}}$ is marginally-connecting or always-connecting, orient $\mathbf{X} \circ\!\!\longrightarrow \mathbf{Y}$ as $\mathbf{X} \rightarrow \mathbf{Y}$.

$\mathscr{R}_7$: If $\mathbf{X} \circ\!\!\longrightarrow \mathbf{Y}$, $p = \langle \mathbf{X}, \mathbf{Z}, \mathbf{W}, ...\mathbf{Y} \rangle$ is an uncovered possibly directed path from $\mathbf{X}$ to $\mathbf{Y}$ such that $\mathbf{Y}$ and $\mathbf{Z}$ are not adjacent, and $\mathcal{A}_{\mathbf{Z},\mathbf{X},\mathbf{Y}}$ is marginally-connecting, orient $\mathbf{X} \circ\!\!\longrightarrow \mathbf{Y}$ as $\mathbf{X} \rightarrow \mathbf{Y}$.

$\mathscr{R}_8$: Suppose $\mathbf{X} \circ\!\!\longrightarrow \mathbf{Y}$, $\mathbf{Z} \rightarrow \mathbf{Y} \leftarrow \mathbf{W}$, $p_1$ is an uncovered possibly directed path from $\mathbf{X}$ to $\mathbf{Z}$, and $p_2$ is an uncovered possibly directed path from $\mathbf{X}$ to $\mathbf{W}$. Let $\mathbf{V}$ be the vertex adjacent to $\mathbf{X}$ on $p_1$ ($\mathbf{V}$ could be $\mathbf{Z}$) and $\mathbf{T}$ be the vertex adjacent to $\mathbf{X}$ on $p_2$ ($\mathbf{T}$ could be $\mathbf{W}$). If $\mathbf{V}$ and $\mathbf{Z}$ ($\mathbf{T}$ and $\mathbf{W}$) are distinct, let $\mathcal{A}_{\mathbf{V},\mathbf{Z},\mathbf{Y}}$ ($\mathcal{A}_{\mathbf{T},\mathbf{W},\mathbf{Y}}$) be marginally-connecting with no separation mark subscripted by a variable on $p_1$ ($p_2$). If $\mathbf{V}$ and $\mathbf{T}$ are distinct and not adjacent, and $\mathcal{A}_{\mathbf{T},\mathbf{X},\mathbf{Y}}$ and $\mathcal{A}_{\mathbf{V},\mathbf{X},\mathbf{Y}}$ are marginally-connecting or always-connecting, then orient $\mathbf{X} \circ\!\!\longrightarrow \mathbf{Y}$ as $\mathbf{X} \rightarrow \mathbf{Y}$.

The procedure for LC-CLOC, leveraging these rules, is outlined in Algorithm 1. It will be clear from the orientation rules, that if there are no clusters of size 1, it is not possible to get the initial orientations required for downstream orientation rules to apply. LC-CLOC introduces a combination of new orientation rules and modifications of rules from FCI, which we show are sound for learning constraints for an $\alpha$C-PAG. While we conjecture that the algorithm is complete, we have yet to prove its completeness, which is an area for future work.

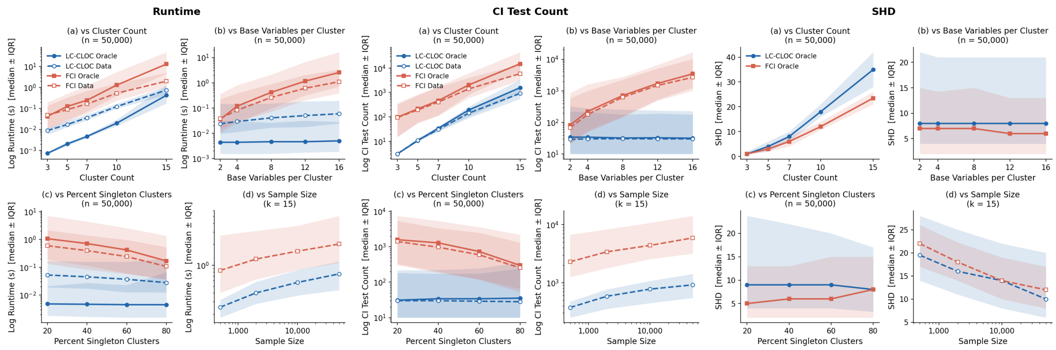



Figure 4: Runtime, conditional independence count, and SHD experimental results

**Algorithm 1** CLOC: Algorithm for Learning an $\alpha$C-PAG

**Input** : Admissible partition $\mathbf{C} = \{\mathbf{C}_1, ..., \mathbf{C}_n\}$, $P(\mathbf{C})$
**Output** : $\alpha$C-PAG, $\mathcal{P}$

1. Form complete graph $\mathcal{P}$ over $\mathbf{C}$ with $\circ\!-\!\circ$ edges.
   **for** $\mathbf{X}, \mathbf{Y} \in \mathbf{C}$ **do**
   - **for** $\mathbf{S} \subseteq \mathbf{C} \setminus \{\mathbf{X}, \mathbf{Y}\}$ **do**
     - **if** $\mathbf{X} \perp\!\!\!\perp \mathbf{Y} | \mathbf{S}$ **then**
       - SepSet $\leftarrow$ $\mathbf{S}$, SepFlag $\leftarrow$ $True$, **break**
   - **if** *SepFlag* $= True$ **then**
     - Remove the edge between $\mathbf{X}, \mathbf{Y}$ in $\mathcal{P}$
2. Determine independence arcs for unshielded triplets by applying rules $\mathcal{AR}_1$ - $\mathcal{AR}_3$ until none apply.
3. Apply separation marks and connection marks by definitions B.1 and A.7.
4. Apply rules $\mathcal{AR}_4$ - $\mathcal{AR}_{15}$ until none apply
5. Apply orientation rules $\mathcal{R}_0$ - $\mathcal{R}_8$ until none apply.

**Theorem 4.4.** *[Soundness of Independence Arc Rules, Orientation Rules and LC-CLOC] The rules for independence arc assignment and edge orientations as well as the procedure of LC-CLOC are sound.*

## 5 Experiments

We evaluate the accuracy and efficiency of LC-CLOC in comparison to an "FCI-then-Cluster" approach where FCI is applied over all variables, and then clusters are imposed over the PAG. Details of this approach and considerations for accuracy are explored further in Appendix C. We evaluate randomly generated non-Markovian cluster graphs. Each graph consists of $k$ clusters with $k \in \{3, 5, 7, 10, 15\}$ where a fraction $\rho_s \in \{0.2, 0.4, 0.6, 0.8\}$ of clusters were forced to be singletons (containing only one variable) and the remaining clusters each contained $b \in \{2, 4, 8, 12, 16\}$ variables. Oracle evaluation determines conditional independence by d-separation in the ground-truth C-DAG or DAG for LC-CLOC and FCI-then-Cluster, respectively. For finite-sample evaluation, Gaussian data are generated from a linear SEM and block CI tests are applied at sample sizes $n \in \{3{,}000,\ 10{,}000,\ 30{,}000,\ 100{,}000\}$. Our evaluation metrics include algorithm runtime, the number of conditional independence test performed, and accuracy measured by Structural Hamming Distance (SHD). For oracle evaluation, each algorithm's output $\alpha$C-PAG is compared to the ground truth C-DAG to assess how much of the graph is possible to recover, and for finite-sample evaluation, each algorithm's output $\alpha$C-PAG is compared to the corresponding oracle $\alpha$C-PAG to assess sensitivity to conditional independence estimation error.

Our results show that LC-CLOC requires orders of magnitude fewer conditional independence tests and results in shorter runtime than FCI across all settings. The gap widens with increasing cluster count and base variables per cluster. Runtime is largely insensitive to the fraction of singleton clusters and scales mildly with sample size. Oracle and data evaluations show similar trends. LC-CLOC and FCI recover comparable graph details at small $k$, with FCI holding an advantage that grows with cluster count. As the percent of clusters that are singletons increases, SHD converges for LC-CLOC and FCI. The comparisons of estimated to oracle $\alpha$C-PAGs is comparable between the two methods (not pictured) indicating that LC-CLOC is robust under multivariate conditional independence tests.

## 6 Conclusions

In this paper we introduce theoretical foundations for causal discovery over clusters where latent confounding between variables is permitted. We also propose a sound algorithm for learning the graphical equivalence class over clusters under non-Markovianity and illustrate how this algorithm offers scalability improvements over FCI. Limitations of our approach include standard assumptions in constraint-based causal discovery including faithfulness which many be violated in practice. Moreover, clustering introduces an inherent tradeoff between scalability and representation fidelity; clusters can enable tractable discovery, but may obscure variable-level relationships, so utility depends on domain. Still, the foundations introduced here provide the first framework for learning over clusters, offering flexibility within this tradeoff to improve feasibility of causal discovery in high-dimensional domains.

## References


[1] Tara V. Anand, Adèle H. Ribeiro, Jin Tian, and Elias Bareinboim. Causal effect identification in cluster dags. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 12172–12179. AAAI Press, 2023. doi: 10.1609/aaai.v37i10.26435.

[2] TV Anand, AH Ribeiro, J Tian, G Hripcsak, and E Bareinboim. Causal discovery over clusters of variables in markovian systems. In *Proceedings of the Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

[3] Bryan Andrews, Peter Spirtes, and Gregory F. Cooper. On the completeness of causal discovery in the presence of latent confounding with tiered background knowledge. In Silvia Chiappa and Roberto Calandra, editors, *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pages 4002–4011. PMLR, 26–28 Aug 2020. URL `https://proceedings.mlr.press/v108/andrews20a.html`.

[4] Elias Bareinboim and Judea Pearl. Causal inference and the data-fusion problem. *Proceedings of the National Academy of Sciences*, 113(27):7345–7352, 2016.

[5] Shaofan Chen, Yuzhong Peng, Guoyuan He, Hao Zhang, Li Cai, and Chengdong Wei. Cdsc: Causal decomposition based on spectral clustering. *Information Sciences*, 657: 119985, 2024. ISSN 0020-0255. doi: https://doi.org/10.1016/j.ins.2023.119985. URL `https://www.sciencedirect.com/science/article/pii/S0020025523015700`.

[6] Wei Chen, Yunjin Wu, Ruichu Cai, Yueguo Chen, and Zhifeng Hao. Ccsl: A causal structure learning method from multiple unknown environments. *ArXiv*, abs/2111.09666, 2021. URL `https://api.semanticscholar.org/CorpusID:244346148`.

[7] C Meek. Causal inference and causal explanation with background knowledge. In P. Besnard and S. Hanks, editors, *Uncertainty in Artificial Intelligence 11*, pages 403–410. Morgan Kaufmann, San Francisco, 1995.

[8] Xueyan Niu, Xiaoyun Li, and Ping Li. Learning cluster causal diagrams: An information-theoretic approach. In Lud De Raedt, editor, *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, pages 4871–4877. International Joint Conferences on Artificial Intelligence Organization, 7 2022. doi: 10.24963/ijcai.2022/675. URL `https://doi.org/10.24963/ijcai.2022/675`. Main Track.

[9] Pekka Parviainen and Samuel Kaski. Learning structures of bayesian networks for variable groups. *International Journal of Approximate Reasoning*, 88:110–127, 2017. ISSN 0888-613X. doi: https://doi.org/10.1016/j.ijar.2017.05.006. URL https://www.sciencedirect.com/science/article/pii/S0888613X17303134.

[10] Sepideh Pashami, Anders Holst, Juhee Bae, and Sławomir Nowaczyk. Causal discovery using clusters from observational data. In *Proceedings of the FAIM'18 Workshop on CausalML*, Stockholm, Sweden, July 2018. FAIM. URL https://urn.kb.se/resolve?urn=urn:nbn:se:hh:diva-39216. Refereed Conference Paper.

[11] Judea Pearl. Causality, 1999.

[12] Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, NY, USA, 2nd edition, 2000.

[13] Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of Causal Inference: Foundations and Learning Algorithms*. Adaptive Computation and Machine Learning. MIT Press, 2017.

[14] Eran Segal, Dana Pe'er, Aviv Regev, Daphne Koller, and Nir Friedman. Learning module networks. *Journal of Machine Learning Research*, 6(19):557–588, 2005. URL http://jmlr.org/papers/v6/segal05a.html.

[15] P Spirtes, C N Glymour, and R Scheines. *Causation, Prediction, and Search*. MIT Press, Cambridge, MA, 2nd edition, 2000.

[16] Peter Spirtes. Directed Cyclic Graphical Representation of Feedback. In P.Besnard and S Hanks, editors, *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, pages 491–498. Morgan Kaufmann, San Mateo, CA, 1995.

[17] Peter Spirtes, Clark N Glymour, and R Scheines. *Causation, Prediction, and Search*. Springer-Verlag, New York, 1993.

[18] Chandler Squires, Annie Yun, Eshaan Nichani, Raj Agrawal, and Caroline Uhler. Causal structure discovery between clusters of nodes induced by latent factors. In Bernhard Schölkopf, Caroline Uhler, and Kun Zhang, editors, *Proceedings of the First Conference on Causal Learning and Reasoning*, volume 177 of *Proceedings of Machine Learning Research*, pages 669–687. PMLR, 11–13 Apr 2022. URL https://proceedings.mlr.press/v177/squires22a.html.

[19] Jonas Wahl, Urmi Ninad, and Jakob Runge. Vector causal inference between two groups of variables. In *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence*, AAAI'23/IAAI'23/EAAI'23. AAAI Press, 2023. ISBN 978-1-57735-880-0. doi: 10.1609/aaai.v37i10.26450. URL https://doi.org/10.1609/aaai.v37i10.26450.

[20] Jonas Wahl, Urmi Ninad, and Jakob Runge. *Journal of Causal Inference*, 12(1):20230041, 2024. doi: doi:10.1515/jci-2023-0041. URL https://doi.org/10.1515/jci-2023-0041.

[21] Raanan Yehezkel and Boaz Lerner. Bayesian network structure learning by recursive autonomy identification. *Journal of Machine Learning Research*, 10(53):1527–1570, 2009. URL http://jmlr.org/papers/v10/yehezkel09a.html.

[22] Jiji Zhang. On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias. *Artificial Intelligence*, 172(16):1873–1896, 11 2008.

# A Background and Supporting Definitions

In this section, we provide formal definitions pertaining to C-DAGs and $\alpha$C-DAGs from [1, 2] for convenience.

**Definition A.1** (Cluster DAG or C-DAG)**.** *Given an ADMG $G(\mathbf{V}, \mathbf{E})$ (a DAG with bidirected edges) and a partition $\mathbf{C} = \{\mathbf{C}_1, \ldots, \mathbf{C}_k\}$ of $\mathbf{V}$, construct a graph $G_{\mathbf{C}}(\mathbf{C}, \mathbf{E}_{\mathbf{C}})$ over $\mathbf{C}$ with a set of edges $\mathbf{E}_{\mathbf{C}}$ defined as follows:*

1. *An edge $\mathbf{C}_i \rightarrow \mathbf{C}_j$ is in $\mathbf{E}_{\mathbf{C}}$ if exists some $V_i \in \mathbf{C}_i$ and $V_j \in \mathbf{C}_j$ such that $V_i \in Pa(V_j)$ in $G$;*
2. *A dashed bidirected edge $\mathbf{C}_i \dashleftarrow\dashrightarrow \mathbf{C}_j$ is in $\mathbf{E}_{\mathbf{C}}$ if exists some $V_i \in \mathbf{C}_i$ and $V_j \in \mathbf{C}_j$ such that $V_i \dashleftarrow\dashrightarrow V_j$ in $G$.*

*If $G_{\mathbf{C}}(\mathbf{C}, \mathbf{E}_{\mathbf{C}})$ contains no cycles, then we say that $\mathbf{C}$ is an* admissible partition *of $\mathbf{V}$. We then call $G_{\mathbf{C}}$ a* cluster DAG*, or* C-DAG*, compatible with $G$.*

**Definition A.2** (Manipulation of a shielded triplet)**.** *Given a shielded triplet over clusters $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$, its manipulation involves removing the edge between $\mathbf{C}_i$ and $\mathbf{C}_j$, corresponding to removal of any edges between variables in these clusters. After manipulation, the shielded triplet becomes unshielded and this manipulated unshielded triplet is referenced as $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle^{-\mathbf{C}_i\mathbf{C}_j}$.*

**Definition A.3** (Independence Arcs)**.** *Consider a graph $\mathbf{G}_C$ over clusters $\mathbf{C} = \langle \mathbf{C}_0, ..., \mathbf{C}_n \rangle$. For any unshielded triplet $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$ (or manipulated unshielded triplet $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle^{-\mathbf{C}_i\mathbf{C}_j}$), let $\mathbf{S}$ equal a (possibly empty) set of clusters $\mathbf{S} \subset (\mathbf{C} \setminus \{\mathbf{C}_i, \mathbf{C}_j\})$ such that $\mathbf{C}_i \perp\!\!\!\perp \mathbf{C}_j | \mathbf{S}$, if such a set exists. For a triplet $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$, an independence arc, $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j} \in \boldsymbol{\mathcal{A}}$, can be drawn from some point on the edge between $\mathbf{C}_i$ and $\mathbf{C}_k$ to some point on the edge between $\mathbf{C}_j$ and $\mathbf{C}_k$ in the following way:*

1. *A marginally-connecting independence arc of - - - - is drawn if and only if $\mathbf{C}_k \in \mathbf{S}$. Consequently, $\mathbf{C}_i \not\perp\!\!\!\perp \mathbf{C}_j | \mathbf{S} \setminus \mathbf{C}_k$ and $\mathbf{C}_i \perp\!\!\!\perp \mathbf{C}_j | \mathbf{S}$.*
2. *A conditionally-connecting independence arc of —||— is drawn if and only if $\mathbf{C}_k \notin \mathbf{S}$ and $\mathbf{C}_i \not\perp\!\!\!\perp \mathbf{C}_j | \mathbf{S} \cup \mathbf{C}_k$.*
3. *A never-connecting independence arc of - -||- - is drawn if and only if $\mathbf{C}_k \notin \mathbf{S}$ and $\mathbf{C}_i \perp\!\!\!\perp \mathbf{C}_j | \mathbf{S} \cup \mathbf{C}_k$.*

*Shielded triplets are annotated according to the behavior of their respective manipulated triplets.*

**Definition A.4** (Arc Trajectory)**.** *Given a graph $\mathbf{G}_C$, for some path over clusters $\langle \mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, ..., \mathbf{C}_n \rangle$, the* arc trajectory *refers to the sequence of independence arcs for each triplet along the path, $\mathbf{a} = \langle \mathcal{A}_{\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3}, ..., \mathcal{A}_{\mathbf{C}_{n-2},\mathbf{C}_{n-1},\mathbf{C}_n} \rangle$.*

**Definition A.5** (Analogous Paths)**.** *Given a C-DAG $G_{\mathbf{C}}$ and a compatible DAG $G$, we define a simple path in $G$ over variables, $p = \langle V_1, V_2, V_3, ..., V_m \rangle$ to be considered analogous to a path in $G_{\mathbf{C}}$ over clusters $p_{\mathbf{C}} = \langle \mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, ..., \mathbf{C}_n \rangle$ (and $p_{\mathbf{C}}$ analogous to $p$) if and only if the following hold: 1) for every variable $V_i$ on $p$, $V_i$ is in some cluster $\mathbf{C}_i$ on $p_{\mathbf{C}}$, 2) for every cluster $C_j$ on $p_{\mathbf{C}}$, there exists some variable $V_j \in \mathbf{C}_j$ where $V_j$ is on $p$, and 3) for any variable $V_n \in \mathbf{C}_n$, there does not exist any variable that appears after $V_n$ on $p$ that is in a cluster before $\mathbf{C}_n$ on $p_{\mathbf{C}}$.*

**Definition A.6** (**Separation Marks**). *Let $G$ be a DAG, and let $G_{\mathbf{C}}$ denote a possible C-DAG for $G$. Consider a path $p_{\mathbf{C}}$ in $G_{\mathbf{C}}$ over clusters $\langle \mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, ..., \mathbf{C}_n \rangle$ and its corresponding arc trajectory $\mathbf{a} = \langle \mathcal{A}_{\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3}, ...\mathcal{A}_{\mathbf{C}_{n-2},\mathbf{C}_{n-1},\mathbf{C}_n} \rangle$ such that:*

1. *there is no arc $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_{i+1},\mathbf{C}_{i+2}} \in \mathbf{a}$ that is a never-connecting arc,*
2. *there is no d-connecting path $p$ in $G$ over variables relative to some set of clusters $\mathbf{Z}$, analogous to $p_{\mathbf{C}}$,*
3. *there exists a d-connecting path $p'$ in $G$ over variables relative to some set of clusters $\mathbf{Z}'$ that is analogous to the path $p'_{\mathbf{C}} = \langle \mathbf{C}_1, ..., \mathbf{C}_{n-1} \rangle$ in $G_{\mathbf{C}}$, and*
4. *there exists a d-connecting path $p''$ in $G$ over variables relative to some set $\mathbf{Z}''$ of clusters that is analogous to the path $p''_{\mathbf{C}} = \langle \mathbf{C}_2, ..., \mathbf{C}_n \rangle$ in $G_{\mathbf{C}}$, .*

*Then, a separation mark, $\oslash_{\mathbf{C}_1}$ is placed on the arc $\mathcal{A}_{\mathbf{C}_{n-2},\mathbf{C}_{n-1},\mathbf{C}_n}$, and a separation mark, $\oslash_{\mathbf{C}_n}$ is placed on the arc $\mathcal{A}_{\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3}$.*

**Definition A.7** (Connection Marks). *Let $G$ be an ADMG and let $G_{\mathbf{C}}$ denote a possible C-DAG for $G$ with independence arcs. Consider a triplet over clusters in $G_{\mathbf{C}}$, $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$, and its corresponding independence arc, $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$. Consider the case where $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ is a never-connecting or conditionally-connecting independence arc and say that there exists a path $p$ in $G$ over variables through the triplet $\langle V_i, ..., V_k, ...V_j \rangle$ such that $V_i \in \mathbf{C}_i$, $V_k \in \mathbf{C}_k$, and $V_j \in \mathbf{C}_j$. Then $\forall V'_k \in \mathbf{C}_k$ and on $p$, where $V'_k$ is a collider, let $\mathbf{D}$ be the set of clusters that 1) are children of $\mathbf{C}_k$ in $G_{\mathbf{C}}$ and 2) include descendants of all colliders along the path, ($\mathbf{D} = \bigcup\{\mathbf{C}_d : V_d \in \mathbf{C}_d\}$ where $V_d \notin \{\mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j\}$ and $V_d \in Ch(V_k)$). Then the connection mark $\oplus_{\mathbf{D}}$ is added to $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$.*

**Definition A.8** (**d-separation over $\alpha$C-DAGs.**). *A path $p_{\mathbf{C}}$ in an $\alpha$C-DAG, $G_{\mathbf{C}}$, is said to be d-separated (or blocked) by a set of clusters $\mathbf{Z} \subset \mathbf{C}$ if and only if its corresponding arc trajectory $\mathbf{a}$ contains an independence arc $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ that is:*

1. *a marginally-connecting independence arc and (a) $\mathbf{C}_k$ is in $\mathbf{Z}$ or (b) there exists a separation mark $\oslash_{\mathbf{C}_x}$ on $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ where $\mathbf{C}_x$ is on $p_{\mathbf{C}}$.*
2. *a conditionally-connecting independence arc and (a) $\mathbf{C}_k$ is not in $\mathbf{Z}$ nor is any true descendant $\mathbf{C}_d$ of $\mathbf{C}_k$ (with directed and connecting path $\mathbf{C}_k \to ... \to \mathbf{C}_d$) in $\mathbf{Z}$, and (b) for any connection mark $\oplus_{\mathbf{C}_x}$ on $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$, $\mathbf{C}_x$ is not in $\mathbf{Z}$ or (c) there exists a separation mark on $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ $\oslash_{\mathbf{C}_x}$ where $\mathbf{C}_x$ is on $p_{\mathbf{C}}$.*
3. *a never-connecting independence arc and for connection mark $\oplus_{\mathbf{C}_x}$ on $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$, $\mathbf{C}_x \notin \mathbf{Z}$.*

# B Extended Discussion

## B.1 Motivation and Intuition

The motivation for the extension and introduction of $\alpha$C-DAGs and $\alpha$C-MAGs/$\alpha$C-PAGs, respectively stems from the challenge of representing independence information and ancestrality together over clusters. This challenge can also be understood by noting that the innovation of causal discovery stems from the connection between independence relationships among variables, testable from an observational dataset, and the structure of edges between the nodes corresponding to these variables in the associated causal diagram. For example, for some unshielded triplet over variables $\langle X, Z, Y \rangle$, where $X$ and $Y$ are adjacent to $Z$ and not adjacent to each other, a collider structure ($X \circ\!\!\longrightarrow Z \longleftarrow\!\!\circ Y$) always implies that $X \perp\!\!\!\perp Y$ and $X \not\perp\!\!\!\perp Y | Z$. With clusters, there is a tension between edge orientations and their implied independence relationships, where the familiar alignment of graphical structure and statistical independence is inapplicable. To witness, consider the ADMG, $G_1$ in Figure 1d with a cluster partition as indicated by the circled groups of variables. $G_{\mathbf{C}_1}$ is the corresponding C-DAG and $\mathcal{P}_{\mathbf{C}_1}$ is an attempt at representing an equivalence class of models sharing the same independences and dependences over clusters. The independence $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$ can be accurately represented with the collider structure $\mathbf{X} \circ\!\!\longrightarrow \mathbf{Z} \longleftarrow\!\!\circ \mathbf{Y}$. However, the appropriate orientations for the edge between $\mathbf{Z}$ and $\mathbf{W}$ are unclear. $\mathbf{X} \perp\!\!\!\perp \mathbf{W}$ would imply $\mathbf{W} \circ\!\!\longrightarrow \mathbf{Z}$, yet this structure would contradict the implications of the dependence $\mathbf{Y} \not\perp\!\!\!\perp \mathbf{W}$. If instead the edge were oriented as $\mathbf{W} \longleftarrow\!\!\circ \mathbf{Z}$ or $\mathbf{W} \circ\!\!-\!\!\circ \mathbf{Z}$, the dependence between $\mathbf{Y}$ and $\mathbf{W}$ would be correctly conveyed, but the dependence $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W} | \mathbf{Z}$ would

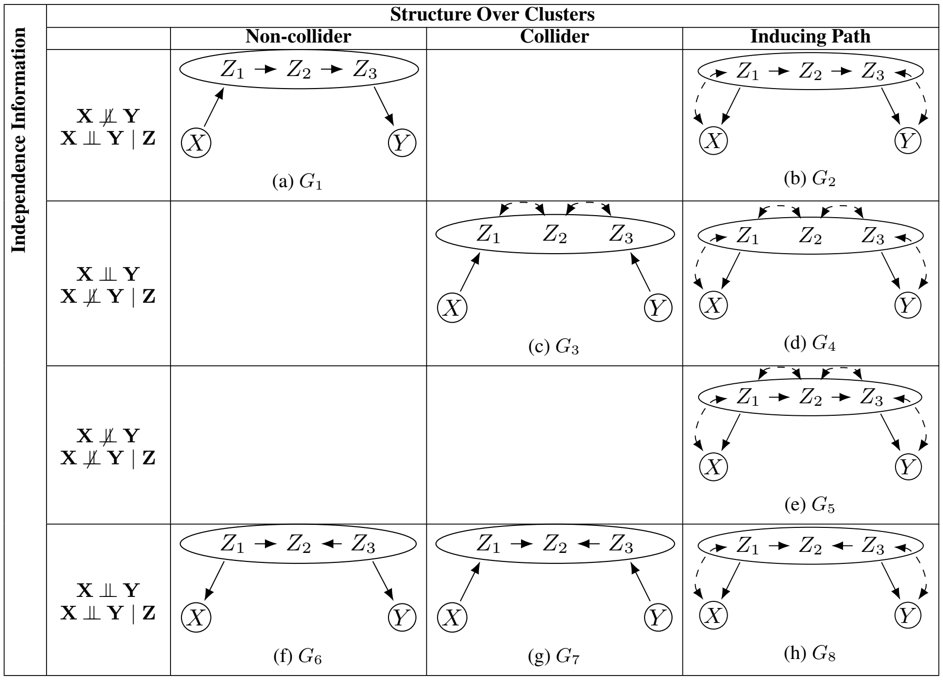



Figure 5: Eight ADMGs with indicated partitions representing triplets of non-colliders, colliders, and inducing paths over clusters, with possible independence information sets.

be misrepresented. While graphical equivalence classes typically only include one edge between each pair of nodes, an attempt to include both the edges $\mathbf{W} \circ\!\!\longrightarrow \mathbf{Z}$ and $\mathbf{W} \longleftarrow\!\!\circ \mathbf{Z}$ (or $\mathbf{W} \circ\!\!-\!\!\circ \mathbf{Z}$) would still result in ambiguity regarding which edge(s) to consider in analysis of a particular triplet.

Another way to understand the need for independence arcs and an extension of the Markov case arcs, is through observing that any direction of ancestrality could exist for multiple pairs of marginal and conditional (in)dependences. Consider the ADMGs $G_3$ and $G_7$ in Figure 5. Both graphs correspond to a C-DAG with three nodes forming a collider structure, but are each associated with their own (in)dependence information: for $G_3$, $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}; \mathbf{X} \not\perp\!\!\!\perp \mathbf{Y} | \mathbf{Z}$ as is usual with colliders, and for $G_7$, $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}; \mathbf{X} \perp\!\!\!\perp \mathbf{Y} | \mathbf{Z}$ which never occurs over variables. The collider triplet over clusters with latent confounding, then, can not be mapped to a single pair of marginal and conditional (in)dependences, as in the Markovian or variable case. In Figure 5, $G_1$ and $G_6$ both map to a C-DAG with three nodes forming a non-collider structure. As with the Markov case, it is possible to observe either $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}; \mathbf{X} \perp\!\!\!\perp \mathbf{Y} | \mathbf{Z}$ or $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}; \mathbf{X} \perp\!\!\!\perp \mathbf{Y} | \mathbf{Z}$ for a non-collider triplet over clusters, and the structure cannot be singularly associated with independence information. Finally we consider an unshielded triplet as an inducing path. For variables, a triplet as an inducing path could appear either as a fork, with bidirected edges between the center and each adjacent node, or as a chain, with at least one bidirected edge between the center node and the adjacent node with an arrowhead into it. A bidirected edge between the center node and the adjacent node with a tail from it is allowed, but not necessary. In Figure 5, ADMGs $G_2$, $G_4$, $G_5$ and $G_8$ all correspond to a C-DAG with three nodes and what appears to be an inducing path structure, however they each correspond to different marginal and conditional (in)dependences. Notably, the inducing path structure corresponding to $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}; \mathbf{X} \not\perp\!\!\!\perp \mathbf{Y} | \mathbf{Z}$ has no analog in a Markov context, motivating the need for a new independence arc.

Overall, Figure 5 summarizes possible combinations of structure and independence information. This chart can be compared with Figure 6 which illustrates examples comparing possible combinations of structure and independence information for graphs only involving variables and no clusters, where the one-to-one alignment of structure and independence information is clearly illustrated.

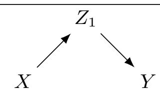
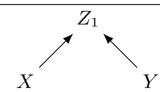
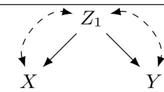
| Independence Information | Structure Over Variables | | | 
|---|---|---|---|
| | **Non-collider** | **Collider** | **Inducing Path** |
| $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}$<br>$\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}$ | <br><br>(a) $G_1$ | | |
| $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$<br>$\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}$ | | <br><br>(c) $G_2$ | |
| $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}$<br>$\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}$ | | | <br><br>(h) $G_3$ |
| $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$<br>$\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}$ | | | |

Figure 6: Three ADMGs representing triplets of non-colliders, colliders, and inducing paths over variables, with possible independence information sets.

### B.2 d-separation in non-Markov $\alpha$C-DAG

Independence arcs allow for a new representation of d-separation. Where latent confounding is allowed, for an isolated triplet with clusters $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$, the triplet is active (d-connecting) relative to the (possibly empty) set of cluster vertices $\mathbf{Z}$ if a) $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$ is marked with an always-connecting independence arc, b) $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$ is marked with a marginally-connecting independence arc and $\mathbf{C}_k \notin \mathbf{Z}$ or c) $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$ is marked with a conditionally-connecting independence arc and $\mathbf{C}_k \in \mathbf{Z}$. Otherwise, $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$ is d-separated relative to $\mathbf{Z}$. Beyond an isolated triplet in an $\alpha$C-DAG, recall that a sequence of independence arcs corresponding to a path over edges is called an *arc trajectory* (Def. A.4). With latent confounding, there may be multiple paths over the same sequence of nodes (e.g. $\mathbf{A} \leftarrow \mathbf{B} \leftarrow \mathbf{E}$ and $\mathbf{A} \leftarrow \mathbf{B} \leftrightarrow \mathbf{E}$ in $G_{\mathbf{C}_2}$ of Figure 3). However these paths share the same unique arc trajectory. Each arc trajectory must be evaluated to determine if two clusters are d-separated in an $\alpha$C-DAG.

**Example 2.** *Consider the example in Figure 3 where $G_2$ is an ADMG, and $G_{\mathbf{C}_2}$ is an $\alpha$C-DAG, where $G_2$ is in the class of $G_{\mathbf{C}_2}$. To determine whether $\mathbf{A}$ and $\mathbf{D}$ are d-separated ($\mathbf{A} \perp\!\!\!\perp \mathbf{D}$), the possible paths between $\mathbf{A}$ and $\mathbf{D}$ are identified, of which there are three: $\mathbf{A} \rightarrow \mathbf{B} \rightarrow \mathbf{C} \leftarrow \mathbf{D}$, $\mathbf{A} \rightarrow \mathbf{B} \rightarrow \mathbf{E} \rightarrow \mathbf{D}$, and $\mathbf{A} \rightarrow \mathbf{B} \leftrightarrow \mathbf{E} \rightarrow \mathbf{D}$. Consider the first path. The arc trajectory corresponding to this path is $\langle \mathcal{A}_{\mathbf{A},\mathbf{B},\mathbf{C}}, \mathcal{A}_{\mathbf{B},\mathbf{C},\mathbf{D}} \rangle$. Because there is no conditioning set in the query, only $\mathcal{A}_{\mathbf{A},\mathbf{B},\mathbf{C}}$ indicates an active triplet but not $\mathcal{A}_{\mathbf{B},\mathbf{C},\mathbf{D}}$, and therefore $\mathbf{A}$ and $\mathbf{D}$ are not connected along this path as all independence arcs in the arc trajectory must be active for two variables to be connected. Considering the second path, the arc trajectory is $\langle \mathcal{A}_{\mathbf{A},\mathbf{B},\mathbf{E}}, \mathcal{A}_{\mathbf{B},\mathbf{E},\mathbf{D}} \rangle$. While $\mathcal{A}_{\mathbf{A},\mathbf{B},\mathbf{E}}$ is an always-connecting arc, $\mathcal{A}_{\mathbf{B},\mathbf{E},\mathbf{D}}$ is a never-connecting arc, so $\mathbf{A}$ and $\mathbf{D}$ are not connected by this path either. The third path has the same arc trajectory as the second. Therefore, we can conclude that $\mathbf{A} \perp\!\!\!\perp \mathbf{D}$.*

Separation marks enrich independence arcs to offer a complete representation (with connection marks) of independence in more complex graphs. Below is the formal and complete definition of separation marks for $\alpha$C-DAGs with latent confounding.

**Definition B.1** (Separation Marks)**.** *Let $G$ be an ADMG and let $G_{\mathbf{C}}$ denote a possible C-DAG for $G$. Consider a path $p_{\mathbf{C}}$ in $G_{\mathbf{C}}$ over clusters $\langle \mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, ..., \mathbf{C}_n \rangle$ and its corresponding arc trajectory (Def. A.4) $\mathbf{a} = \langle \mathcal{A}_{\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3}, ...\mathcal{A}_{\mathbf{C}_{n-2},\mathbf{C}_{n-1},\mathbf{C}_n} \rangle$ such that:*

1. *there is no arc $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_{i+1},\mathbf{C}_{i+2}} \in \mathbf{a}$ that is a never-connecting arc,*
2. *there is no d-connecting path $p$ in $G$ over variables relative to some set of clusters $\mathbf{Z}$, that is analogous to $p_{\mathbf{C}}$,*
3. *there exists a d-connecting path $p'$ in $G$ over variables relative to some set of clusters $\mathbf{Z}'$ that is analogous to the path $p'_{\mathbf{C}} = \langle \mathbf{C}_1, ..., \mathbf{C}_{n-1} \rangle$ in $G_{\mathbf{C}}$, and*

4. *there exists a d-connecting path $p''$ in $G$ over variables relative to some set $\mathbf{Z}''$ of clusters that is analogous to the path $p''_{\mathbf{C}} = \langle \mathbf{C}_2, ..., \mathbf{C}_n \rangle$ in $G_{\mathbf{C}}$.*

*Then, a separation mark, $\oslash_{\mathbf{C}_1}$ is placed on the arc $\mathcal{A}_{\mathbf{C}_{n-2},\mathbf{C}_{n-1},\mathbf{C}_n}$, and a separation mark, $\oslash_{\mathbf{C}_n}$ is placed on the arc $\mathcal{A}_{\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3}$.*

*Now consider a path $p_{\mathbf{C}}$ in $G_{\mathbf{C}}$ over clusters $\langle \mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, ..., \mathbf{C}_n \rangle$ and its corresponding arc trajectory $\mathbf{a} = \langle \mathcal{A}_{\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3}, ...\mathcal{A}_{\mathbf{C}_{n-2},\mathbf{C}_{n-1},\mathbf{C}_n} \rangle$ such that:*

1. *there is no arc $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_{i+1},\mathbf{C}_{i+2}} \in \mathbf{a}$ that is a never-connecting arc,*
2. *there exists a path $p$ in $G$ over variables that is analogous to $p_{\mathbf{C}}$, and*
3. *$\mathcal{A}_{\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3}$ is an always-connecting arc.*

*Then,:*

1. *if there exists some $p$ that is a connecting-path, and there are no collider-paths, the separation mark $\oslash_{\underline{\mathbf{C}_n}}$ is placed on $\mathcal{A}_{\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3}$, and a separation mark $\oslash_{\underline{\mathbf{C}_1}}$ is placed on the arc $\mathcal{A}_{\mathbf{C}_{n-2},\mathbf{C}_{n-1},\mathbf{C}_n}$.*
2. *if there exists some $p$ that is a collider-path, and there are no connecting-paths, the separation mark $\oslash_{\overline{\mathbf{C}_n}}$ is placed on $\mathcal{A}_{\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3}$, and a separation mark $\oslash_{\overline{\mathbf{C}_1}}$ is placed on the arc $\mathcal{A}_{\mathbf{C}_{n-2},\mathbf{C}_{n-1},\mathbf{C}_n}$*

### B.3 $\alpha$C-MAG Semantics

#### B.3.1 Inducing Paths

The semantics of $\alpha$C-MAGs depends critically on inducing paths. Definition 3.1 for inducing paths over clusters can be understood by considering the C-DAG, $G_{\mathbf{C}}$ in Figure 7. If the definition of an inducing path over variables (see Def. B.2 below) is applied to the clusters, it may appear as if there is an inducing path between $\mathbf{X}$ and $\mathbf{Y}$ in $G_{\mathbf{C}}$. $G_1, G_2$, and $G_3$ in Figure 7 illustrate three possible ADMGs with different relationships internal to cluster $\mathbf{Z}$, and include the independence arcs that would mark the triplet in a compatible $\alpha$C-DAG. In Figure 7b, $p_v = p'_v$ between $\mathbf{X}$ and $\mathbf{Y}$ such that every node on $p_v$ is a collider and an ancestor of either $\mathbf{X}$ or $\mathbf{Y}$. In Figure 7c, $p_v \neq p'_v$, but there does exist a path $p_v$ between $\mathbf{X}$ and $\mathbf{Y}$ through $\mathbf{Z}$ where each node is a collider, and there is another path $p'_v$ where each node is an ancestor of either $\mathbf{X}$ or $\mathbf{Y}$. $G_2$ has the same independence and indescernible ancestral information over clusters as $G_1$ and will therefore have the same $\alpha$C-MAG representation. By contrast, $G_3$, while having a path $p'_v$ where all nodes on the path are ancestors of either $\mathbf{X}$ or $\mathbf{Y}$, has no path $p_v$ from $\mathbf{X}$ to $\mathbf{Y}$ through $\mathbf{Z}$ where every node on the path is a collider and so $G_3$ has different independence information from $G_1$ and $G_2$. Definition 3.1 for inducing paths over clusters will represent only and all triplets where $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}, \mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Z}$; that is, where there is no separating set over clusters. Therefore inducing paths are composed only of always-connecting independence arcs.

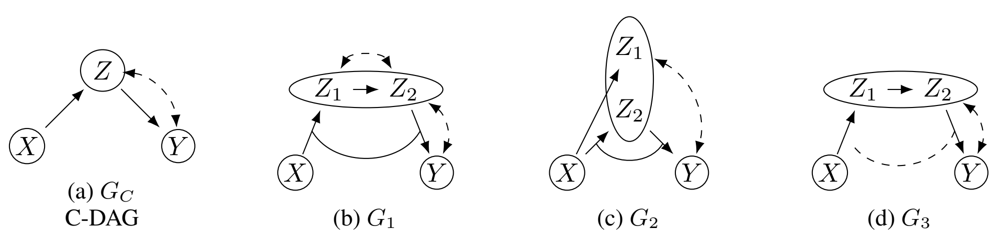



Figure 7: C-DAG $G_C$ with three compatible DAGs $G_1, G_2$ and $G_3$. $G_C$ appears to have an inducing path and $G_1$ has a true inducing path over variables. $G_2$ does not have an inducing path over variables, but the paths between $\mathbf{X}$ and $\mathbf{Y}$ through cluster $\mathbf{Z}$ cause behavior at the cluster level that is indistinguishable from that of an inducing path. While the edges over the clusters for $G_3$ create the appearance over clusters of what is familiarly interpreted as an inducing path, over variables it is clear that there are no inducing paths and the independence information does not behave as an inducing path.

**Definition B.2** (Inducing Path (over variables))**.** *Given $X, Y \in \mathbf{V}, \mathbf{L} \subseteq \mathbf{V} \setminus \{X, Y\}$, an inducing path between $X, Y$ relative to $\mathbf{L}$ is a path between $X$ and $Y$ on which every vertex not in $\mathbf{L}$ is a collider on the path and every collider is an ancestor of either $X$ or $Y$.*

### B.3.2 $\alpha$C-MAG Construction

Definition 3.3 describes how an $\alpha$C-DAG can be transformed into an $\alpha$C-MAG. An example of this process can be found below.

**Example 3.** *In Figure 3, $G_2$ is a causal diagram in the class of $\alpha$C-DAG $G_{\mathbf{C}_2}$. We generate $\alpha$C-MAG, $\mathcal{M}_{\mathbf{C}_2}$, as follows. Identifying inducing paths in $G_{\mathbf{C}}$, yields the edges of $\mathcal{M}_{\mathbf{C}}$, matching $G_{\mathbf{C}}$ with the added edge $\mathbf{A} \to \mathbf{E}$, from the always-connecting arc on and inducing path of $\langle \mathbf{A}, \mathbf{B}, \mathbf{E} \rangle$. These edges are oriented according to ancestrality by $G_{\mathbf{C}}$, and independence arcs are added to $\mathcal{M}_{\mathbf{C}}$ from $G_{\mathbf{C}}$. Adding edge $\mathbf{A} \to \mathbf{E}$ creates new triplets: $\langle \mathbf{E}, \mathbf{A}, \mathbf{B} \rangle$, $\langle \mathbf{A}, \mathbf{E}, \mathbf{B} \rangle$, and $\langle \mathbf{A}, \mathbf{E}, \mathbf{D} \rangle$. The triplets $\langle \mathbf{E}, \mathbf{A}, \mathbf{B} \rangle$ and $\langle \mathbf{A}, \mathbf{E}, \mathbf{B} \rangle$ arise from a shielded triplet and are assigned a never-connecting arc. Triplet $\langle \mathbf{A}, \mathbf{E}, \mathbf{D} \rangle$ is unshielded, and since $(\mathbf{A}, \mathbf{E})$ appears only in $\mathcal{M}_{\mathbf{C}}$ we consider paths $\langle \mathbf{A}, ..., \mathbf{E}, \mathbf{D} \rangle$ of which there is one, namely $\langle \mathbf{A}, \mathbf{B}, \mathbf{E}, \mathbf{D} \rangle$. The never-connecting arc on triplet $\langle \mathbf{B}, \mathbf{E}, \mathbf{D} \rangle$ implies the path is inactive marginally and conditionally on $\mathbf{E}$, so $\langle \mathbf{A}, \mathbf{E}, \mathbf{D} \rangle$ is assigned a never-connecting arc.*

### B.3.3 Ancestrality in $\alpha$C-MAGs

$\alpha$C-MAGs have the property of being ancestral, a property that is redefined for a cluster context in Definition 3.4. An example illustrating the definition is below.

### B.3.4 Discriminating Paths in $\alpha$C-MAGs

Markov equivalent $\alpha$C-MAGs are informed by discriminating paths, which are redefined for a cluster context in Definition 3.9. An example illustrating the definition is below.

**Example 4.** *Figure 8 illustrates two graphs, both of which are ancestral. The independence arcs inform ancestral relationships, and in the case of $\mathcal{M}_{\mathbf{C}_2}$, this allows for edges which create the appearance of an almost-directed cycle by edge orientations, but is in fact not a true almost-directed cycle, because $\mathbf{A} \notin An(\mathbf{D})$, as indicated by the independence arcs.*

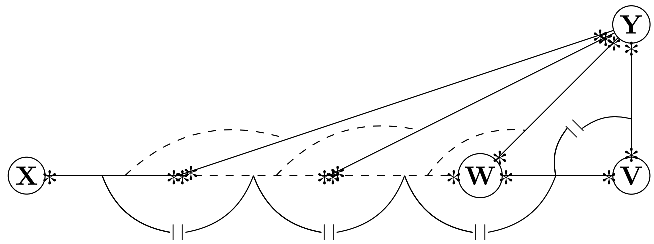
**Example 5.** *Figure 9 illustrates a canonical example of a discriminating path structure with representation over clusters. Edges with asterisks are used to indicate any possible orientation. Let $\mathbf{Q}$ be the set of all nodes between $\mathbf{X}$ and $\mathbf{V}$. If $\mathbf{X} \perp\!\!\!\perp \mathbf{Y} | \mathbf{Q}$ and $\mathbf{X} \perp\!\!\!\perp \mathbf{Y} | \{\mathbf{Q}, \mathbf{V}\}$, $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}^{-\mathbf{WY}}}$ must be a never-connecting arc. If $\mathbf{X} \perp\!\!\!\perp \mathbf{Y} | \mathbf{Q}$ and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y} | \{\mathbf{Q}, \mathbf{V}\}$, $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}^{-\mathbf{WY}}}$ must be a conditionally-connecting arc. If $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y} | \mathbf{Q}$ and $\mathbf{X} \perp\!\!\!\perp \mathbf{Y} | \{\mathbf{Q}, \mathbf{V}\}$, $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}^{-\mathbf{WY}}}$ must be a marginally-connecting arc. If $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y} | \mathbf{Q}$ and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y} | \{\mathbf{Q}, \mathbf{V}\}$, $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}^{-\mathbf{WY}}}$ must be a an always-connecting arc. If $|\mathbf{V}| = 1$, logic similar to that of the unshielded collider case will hold, in that if $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}^{-\mathbf{WY}}}$ is a conditionally-connecting arc, $\mathbf{V}$ will be a collider node.*

### B.3.5 Interpretation of $\alpha$C-MAGs

The interpretation of $\alpha$C-MAGs follows from its construction definition. An $\alpha$C-MAG can be understood to convey the following information:


(a) $\mathcal{M}_{\mathbf{C}_1}$ (b) $\mathcal{M}_{\mathbf{C}_2}$

Figure 8: $\mathcal{M}_{\mathbf{C}_1}$ and $\mathcal{M}_{\mathbf{C}_2}$ are cluster graphs that are ancestral.




(a) $G_1$

Figure 9: A discriminating path over clusters between $\mathbf{X}$ and $\mathbf{Y}$ for $\mathbf{V}$

- $\mathbf{C}_i \rightarrow \mathbf{C}_j$ means that $\exists \mathbf{V}_i \in \mathbf{C}_i$ that is a cause of some $\mathbf{V}_j \in \mathbf{C}_j$, but there is no $\mathbf{V}'_j \in \mathbf{C}_j$ that is a cause of any $\mathbf{V}'_i \in \mathbf{C}_i$
- $\mathbf{C}_i \leftrightarrow \mathbf{C}_j$ means that there is no $\mathbf{V}_i \in \mathbf{C}_i$ that is a cause of any $\mathbf{V}_j \in \mathbf{C}_j$, and there is no $\mathbf{V}'_j \in \mathbf{C}_j$ that is a cause of any $\mathbf{V}'_i \in \mathbf{C}_i$.

# C Additional Experimental Details and Results

## C.1 Experimental Setup and Design

All experiments were run on a machine with CPU: Intel i9 Chip, 32 GB of RAM, and macOS operating system. A single core was used for the experiments. Algorithms are implemented in Python.

**Graph generation.** For each configuration, we generated a random C-DAG over $k$ clusters using an Erdős–Rényi skeleton with edge density 0.2, then expanded it to a variable-level DAG. A fraction $\rho_s \in 0.2, 0.4, 0.6, 0.8$ of clusters were forced to be singletons (one variable); the remainder each contained $b$ variables. Intra-cluster variable graphs were random DAGs with edge density 0.4 and an independent 20% edge-retention dropout ($p_{\text{keep}} = 0.8$). Inter-cluster latent confounders were added to 70% of adjacent cluster pairs ($\rho_L = 0.7$), each represented by a single latent variable with directed edges into both clusters; when a latent confounder was added between two clusters, each direct variable-level edge between them was independently dropped with probability 0.2. Intra-cluster latent confounders were also added, with the fraction of confounded adjacent variable pairs drawn uniformly per cluster. We evaluated $k \in 3, 5, 7, 10, 15$ clusters with $b \in 2, 4, 8, 12, 16$ variables per non-singleton cluster, yielding graphs ranging from a handful to a few dozen observed variables. Each configuration was repeated for $R = 10$ independent random seeds.

**Algorithm Implementation.** LC-CLOC uses Wilks' Lambda independence test and FCI-then-cluster uses a Fisher-Z correlation test. FCI-then-cluster runs the standard FCI algorithm at the variable level and projects the resulting PAG to cluster-level marks. The clustering procedure is shown below.

**Definition C.1 (Clustered PAG.).** *Given a PAG, $\mathcal{P}$ over variables $\mathbf{V}$, and a partition $\mathbf{C} = \{\mathbf{C}_1, ..., \mathbf{C}_n\}$ of $\mathbf{V}$, construct a graph $\mathcal{P}_{\mathbf{C}}$ over $\mathbf{C}$ as follows.*

- *An edge between $\mathbf{C}_i$ and $\mathbf{C}j$ exists in $\mathcal{P}_{\mathbf{C}}$ if and only if there exist $V_i \in \mathbf{C}_i$ and $V_j \in \mathbf{C}_j$ such that $V_i$ and $V_j$ are adjacent in $\mathcal{P}$.*
- *For each ordered cluster pair $(\mathbf{C}_i, \mathbf{C}_j)$, let $\mathcal{M}_{ij}$ be the multiset of marks at the $\mathbf{C}_i$-end collected over all inter-cluster variable edges $V_i$–$V_j$ with $V_i \in \mathbf{C}_i$, $V_j \in \mathbf{C}_j$. The mark placed at $\mathbf{C}_i$ on the cluster-level edge $\mathbf{C}_i$—$\mathbf{C}_j$ is resolved by priority as follows:*

  *If $\mathcal{M}_{ij}$ contains both an arrowhead ($\rightarrow$) and a tail ($-$), the mark is set to $\circ$. Otherwise the mark of highest priority is used, where arrowhead$>$ tail $>$ $\circ$.*

We note that the graphical object created by the procedure above, which we refer to as a clustered PAG, determined by the FCI-then-Cluster approach, is distinct from an $\alpha$C-PAG. In particular, edges that should be variant in a cluster Markov equivalence class may be oriented in the clustered PAG, due to some feature of the distribution over variables.

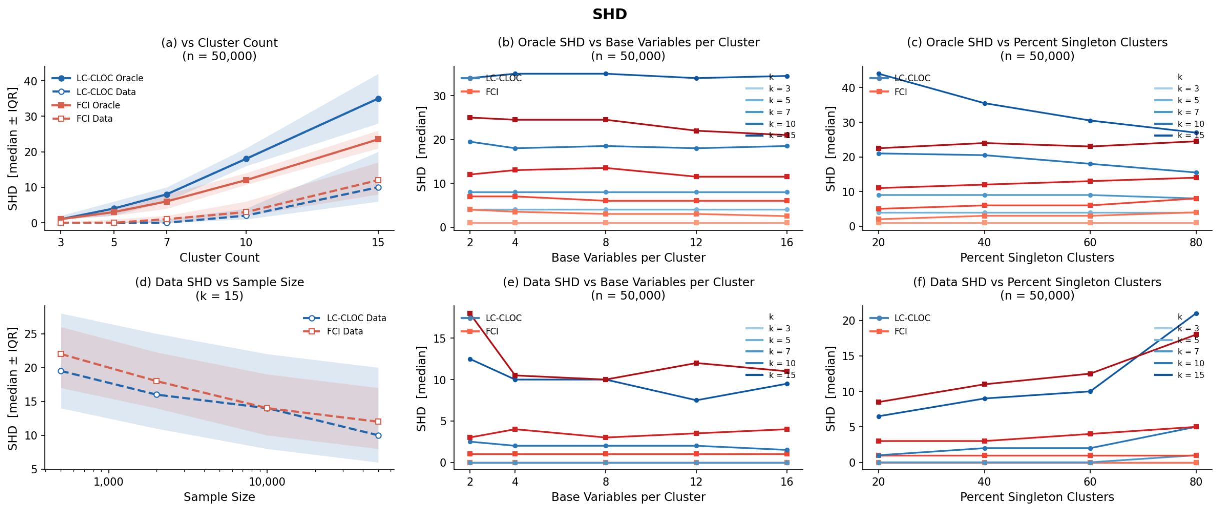



Figure 10: Stratified SHD Results

## C.2 Additional Results

Figure 10 shows SHD results stratified for the data and oracle analyses, as well as across cluster counts, sample size, base variables per cluster, and percent of singleton clusters. Panel (c) shows clearly how for a given $k$ clusters, as the percent of singleton clusters increases, the difference between the SHD of the two approaches decreases. The SHD remains constant for a given $k$ even while the base number of variables per cluster increases, showing robustness with large clusters.

# D Proofs

**Theorem 2.4.** *[Soundness and completeness of d-separation in $\alpha$C-DAGs with latent confounding.] In an $\alpha$C-DAG $G_{\mathbf{C}}$, let $\{\mathbf{X}, \mathbf{Z}, \mathbf{Y}\} \in \mathbf{C}$. $\mathbf{X}$ and $\mathbf{Y}$ are d-separated by $\mathbf{Z}$ in $G_{\mathbf{C}}$, if and only if for any ADMG, $G$, compatible with $G_{\mathbf{C}}$, $\mathbf{X}$ and $\mathbf{Y}$ are d-separated by $\mathbf{Z}$ in $G$.*

*Proof.* First we prove the soundness of d-separation by showing that if $\mathbf{X}$ and $\mathbf{Y}$ are d-separated by $\mathbf{Z}$ in $G_{\mathbf{C}}$, then, in any ADMG, $G$, compatible with $G_{\mathbf{C}}$, $\mathbf{X}$ and $\mathbf{Y}$ are d-separated by $\mathbf{Z}$ in $G$. We show by contradiction. Assume $\mathbf{X}$ and $\mathbf{Y}$ are d-separated by $\mathbf{Z}$ in $G_{\mathbf{C}}$ but in some compatible ADMG, $G$, there exists a path $p$ between a variable $X \in \mathbf{X}$ and $Y \in \mathbf{Y}$ that is active when the set of variables contained in cluster $\mathbf{Z}$ are conditioned on. By the preservation of paths and adjacencies, no connection is destroyed through clustering, so $p$ in $G$ is contained in a path $p_{\mathbf{C}}$ of $G_{\mathbf{C}}$ between clusters $\mathbf{X}$ and $\mathbf{Y}$. Since $\mathbf{X}$ and $\mathbf{Y}$ are d-separated by $\mathbf{Z}$ in $G_{\mathbf{C}}$, $p_{\mathbf{C}}$ is blocked, and $\mathbf{X}$ and $\mathbf{Y}$ are not adjacent. Therefore, by definition 2.3, there is at least one triplet of clusters in $p_{\mathbf{C}}$ that indicates a block on the path. Let this triplet be $\langle \mathbf{C}_i, \mathbf{C}_m, \mathbf{C}_j \rangle$, and let its associated independence arc be $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ where $\mathbf{C}_m$ is distinct from $\mathbf{X}$ and $\mathbf{Y}$. Consider the subpath $p_{ij}$ of $p$ contained in the triplet $\langle \mathbf{C}_i, \mathbf{C}_m, \mathbf{C}_j \rangle$ in $p_{\mathbf{C}}$. Since $p$ is active by assumption, every subpath of $p$ is active, including $p_{ij}$. The triplet $\langle \mathbf{C}_i, \mathbf{C}_m, \mathbf{C}_j \rangle$ indicates a block on the path either if 1) $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ is a never connecting arc with no connection marks $\oplus_{\mathbf{C}_d}$ such that $\mathbf{C}_d \in \mathbf{Z}$, 2) if $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ is a marginally-connecting arc where $\mathbf{C}_m \in \mathbf{Z}$, 3) if $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ is a conditionally-connecting arc such that $\mathbf{C}_m \notin \mathbf{Z}$ and with no connection mark $\oplus_{\mathbf{C}_d}$ such that $\mathbf{C}_d \notin \mathbf{Z}$ or 4) if there is a separation mark $\oslash_{\mathbf{C}_x}$, $\oslash_{\mathbf{C}_{\overline{x}}}$, or $\oslash_{\mathbf{C}_{\underline{x}}}$ on $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ such that $\mathbf{C}_x$ is on $p_{\mathbf{C}}$ and $\mathbf{C}_m \in \mathbf{Z}$ for $\oslash_{\mathbf{C}_{\overline{x}}}$ and $\mathbf{C}_m \notin \mathbf{Z}$ for $\oslash_{\mathbf{C}_{\underline{x}}}$. Cases 1-3 follow from the Markov case. In case 4, definition B.1 states that if $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ is an always-connecting path, if $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ is a marginally-connecting arc such that $\mathbf{C}_m \notin \mathbf{Z}$, or if $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ is a conditionally-connecting arc such that $\mathbf{C}_m \in \mathbf{Z}$, then $p_{ij}$ may be active, but since $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ is marked with a separation mark, there must exist some sub-path $p_{ix}$ of $p$ from some $V_i \in \mathbf{C}_i$ to some $V_x \in \mathbf{C}_x$ such that $\mathbf{C}_x$ is on $p_{\mathbf{C}}$ that is inactive. Therefore, $p$ must be inactive, there is a contradiction, and we conclude that if $\mathbf{X}$ and $\mathbf{Y}$ are d-separated by $\mathbf{Z}$ in $G_{\mathbf{C}}$, then, in any ADMG, $G$, compatible with $G_{\mathbf{C}}$, $\mathbf{X}$ and $\mathbf{Y}$ are d-separated by $\mathbf{Z}$ in $G$.

Then, we prove the completeness of d-separation by showing that if $\mathbf{X}$ and $\mathbf{Y}$ are d-separated by $\mathbf{Z}$ in an ADMG $G$, then $\mathbf{X}$ and $\mathbf{Y}$ are d-separated by $\mathbf{Z}$ in a compatible $\alpha$C-DAG $G_{\mathbf{C}}$. We prove by contradiction. Assume all paths from some $X \in \mathbf{X}$ to some $Y \in \mathbf{Y}$ are blocked by $\mathbf{Z}$ in some ADMG $G$, but $\mathbf{X}$ and $\mathbf{Y}$ are not d-separated by $\mathbf{Z}$ in $G_{\mathbf{C}}$, i.e. $(\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Z})_{G_{\mathbf{C}}}$. If all paths from any $X \in \mathbf{X}$ to any $Y \in \mathbf{Y}$ are inactive by $\mathbf{Z}$, then by preservation of paths and adjacencies, $\mathbf{X}$ and $\mathbf{Y}$ are not adjacent in $G_{\mathbf{C}}$. No connections are destroyed through clustering so any $p$ in $G$ is contained in a path $p_{\mathbf{C}}$ of $G_{\mathbf{C}}$ between clusters $\mathbf{X}$ and $\mathbf{Y}$. Because $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Z}$ in $G_{\mathbf{C}}$, by Definition 2.3, there must exist some path $p_{\mathbf{C}}$ such that 1) for any triplet $\langle \mathbf{C}_i, \mathbf{C}_m, \mathbf{C}_j \rangle$ on $p_{\mathbf{C}}$, the independence arc $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ marking it must not be marked by a separation mark $\oslash_{\mathbf{C}_k}$ where $\mathbf{C}_k$ is on $p_{\mathbf{C}}$, 2) for all marginally-connecting arcs $\mathbf{C}_m \notin \mathbf{Z}$, 3) for all conditionally connecting arcs $\mathbf{C}_m \in \mathbf{Z}$, or $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ is marked with a connection mark $\oplus_{\mathbf{C}_d}$ and $\mathbf{C}_d$ or a true descendant is in $\mathbf{Z}$, 4) for all never-connecting arcs, $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ is marked by a connection mark $\oplus_{\mathbf{C}_d}$ and $\mathbf{C}_d$ or a true descendant is in $\mathbf{Z}$.

For all paths $p$ from some $X \in \mathbf{X}$ to some $Y \in \mathbf{Y}$ in $G$ to be blocked, there must exist at least one triplet $\langle V_i, V_m, V_j \rangle$, contained either within 1 cluster (i.e. $\langle V_i, V_m, V_j \rangle \in \mathbf{C}_m$) or between 2 (i.e. $\langle V_i, V_m \rangle \in \mathbf{C}_m, V_j \in \mathbf{C}_j$ or $V_i \in \mathbf{C}_i, \langle V_m, V_j \rangle \in \mathbf{C}_m$) or 3 clusters (i.e. $V_i \in \mathbf{C}_i, V_m \in \mathbf{C}_m, V_j \in \mathbf{C}_j$) on $p_{\mathbf{C}}$, that is blocked.

1. If the triplet is a non-collider, $V_i \ast\!\!-\!\!\ast V_m \ast\!\!\longrightarrow V_j$, then $V_m$ must be in $\mathbf{Z}$, which implies that $\mathbf{C}_m \in \mathbf{Z}$. As there could be multiple paths through a cycle, the triplet over clusters, $\langle \mathbf{C}_i, \mathbf{C}_m, \mathbf{C}_j \rangle$ could still be marked by any independence arc.
   (a) If $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ is a marginally-connecting arc or never-connecting arc, because $\mathbf{C}_m \in \mathbf{Z}$ there is a contradiction with the implications of $(\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Z})_{G_{\mathbf{C}}}$.
   (b) If $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ is a conditionally-connecting arc or an always-connecting arc, then there must exist a different path, $p'$, over variables through the triplet from some some $V'_i \in \mathbf{C}_i$ to $V'_j \in \mathbf{C}_j$ through $\mathbf{C}_m$ that is a collider path. Because $\mathbf{C}_m \in \mathbf{Z}$, either there is no $X \in \mathbf{X}$ or $Y \in \mathbf{Y}$ on $p'$ or there must be another triplet, $V_q, V_r, V_w$, on $p'$ that is blocked.
2. If the triplet is a collider, $V_i \ast\!\!\longrightarrow V_m \longleftarrow\!\!\ast V_j$, then $V_m$ nor any of its descendants, $V_d$ can be in $\mathbf{Z}$, implying that $C_m \notin \mathbf{Z}$ and $C_d \notin \mathbf{Z}$ where $V_d \in \mathbf{C}_d$ and $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ is marked with the connection mark $\oplus_{\mathbf{C}_d}$.
   (a) If $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ is a marginally-connecting arc or an always connecting arc, then there must exist a different path, $p'$, over variables through the triplet from some some $V'_i \in \mathbf{C}_i$ to $V'_j \in \mathbf{C}_j$ through $\mathbf{C}_m$ that is a connecting path. Because $\mathbf{C}_m \notin \mathbf{Z}$, either there is no $X \in \mathbf{X}$ or $Y \in \mathbf{Y}$ on $p'$ or there must be another triplet, $V_q, V_r, V_w$, on $p'$ that is blocked.
   (b) If $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ is a conditionally-connecting arc or a never-connecting arc, because $\mathbf{C}_m \notin \mathbf{Z}$, and where there is a connection mark $\oplus_{\mathbf{C}_d}$, $\mathbf{C}_d \notin \mathbf{Z}$, there is a contradiction with the implications of $(\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Z})_{G_{\mathbf{C}}}$.

For any path $p'$ with a blocked triplet $\langle V_q, V_r, V_w \rangle$, either one of the conditions above leading to a contradiction (case 1a or 2b) applies, or there is a contradiction because a separation mark must exist along the path $p_{\mathbf{C}}$. By definition B.1, the separation mark would be required because by assumption, all paths between any $X \in \mathbf{X}$ and $Y \in \mathbf{Y}$ are blocked by $\mathbf{Z}$ in $G$, so it is not possible for there to be a d-connecting path relative to $\mathbf{Z}$ in $G$ analogous to $p_{\mathbf{C}}$ in $\mathbf{G}_C$. However, $p$ is a d-connecting path relative to $\mathbf{Z}$ analogous to $p'_{\mathbf{C}} = \langle \mathbf{C}_i, ..., \mathbf{C}_r \rangle$ and $p'$ is a d-connecting path relative to $\mathbf{Z}$ analogous to $p''_{\mathbf{C}} = \langle \mathbf{C}_m, ..., \mathbf{C}_w \rangle$, so by definition B.1, the criteria is met and a separation must be placed.

If $\mathbf{X}$ and $\mathbf{Y}$ are d-separated by $\mathbf{Z}$ in $G$, it is also possible that there is no path from any $X \in \mathbf{X}$ to some $Y \in \mathbf{Y}$, and $\mathbf{Z}$ would equal the empty set. In this case, by preservation of adjacencies, for any triplet $\langle \mathbf{C}_i, \mathbf{C}_m, \mathbf{C}_j \rangle$ along $p_{\mathbf{C}}$, there must be some $V_i \in \mathbf{C}_i$ adjacent to some $V_m \in \mathbf{C}_m$, and some $V'_m \in \mathbf{C}_m$ adjacent to some $V_j \in \mathbf{C}_j$. Then, there must exist some such triplet where $V_m$ is not adjacent to $V'_m$. If for all $V_m$ and $V'_m$ in $C_m$, $V_m$ and $V'_m$ are not adjacent, then $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_m,\mathbf{C}_j}$ must be marked with a never-connecting arc in $G_{\mathbf{C}}$ with no connection mark, and there would be a contradiction with the implications of $(\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Z})_{G_{\mathbf{C}}}$. Otherwise, because $\mathbf{X}$ and $\mathbf{Y}$ are d-separated by $\mathbf{Z}$ in $G$, there must exist some connecting subpaths of $p_{\mathbf{C}}$, $\mathbf{C}_i, ..., \mathbf{C}_n$ and $\mathbf{C}_{i+1}, ..., \mathbf{C}_n + 1$ such that $\mathbf{C}_i \perp\!\!\!\perp \mathbf{C}_{n+1}$, which, by definition B.1, necessitates a separation mark and then there would be a contradiction with the implications of $(\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Z})_{G_\mathbf{C}}$. □

**Remark 3.2.** *In $\alpha$C-DAGs, triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y}\rangle$ with an always-connecting arc forms a primitive inducing path.*

*Proof.* This remark follows directly from definition 2.1. □

We prove that the graph yielded by definition 3.3 is in fact ancestral and maximal by definitions 3.4 and 3.5, respectively.

*Proof.* By definition 3.4, the graphical output of definition 3.3 must not include any true directed cycles or any true almost directed cycles. Consider an $\alpha$C-MAG $\mathcal{M}_{G_\mathbf{C}}$ corresponding to $\alpha$C-DAG $G_\mathbf{C}$. If $\mathcal{M}_{G_\mathbf{C}}$ contains a true directed cycle, then by definition 3.3, there must be a cycle in the corresponding $\alpha$C-DAG. However, this is impossible by the assumption of an admissible partition for the $\alpha$C-DAG. By definition 3.3, it is also impossible to have a true almost directed cycle because if there exists some pair of variables, $\mathbf{A}, \mathbf{B}$ for which there exists an apparent directed path from $\mathbf{A}$ to $\mathbf{B}$, but $\mathbf{A} \leftrightarrow \mathbf{B}$ are connected by a bidirected edge, there must be an independence arc on the path that is not marginally-connecting and/or there is a separation mark such that there is no true directed path, and then there is no true almost directed cycle. If there is a true directed path from $\mathbf{A}$ to $\mathbf{B}$, there would be a violation of definition 3.3 as $\mathbf{A} \in An(\mathbf{B})$ so it should be $\mathbf{A} \to \mathbf{B}$. □

**Remark 3.6.** *An $\alpha$C-MAG cannot have an always-connecting arc on an unshielded triplets.*

*Proof.* This remark follows directly from remark 3.2 and definition 3.3. □

We introduce three additional remarks that will inform subsequent proofs.

**Remark D.1.** *In an $\alpha$C-DAG, marginally-connecting and always-connecting independence arcs always imply a non-collider structure in the corresponding $\alpha$C-MAG.*

*Proof.* We prove by contradiction. Consider an unshielded triplet $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j\rangle$ such that $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ is a marginally-connecting, or always-connecting independence arc. We show that orienting the triple as $\mathbf{C}_i \ast\!\!\rightarrow \mathbf{C}_k \leftarrow\!\!\ast \mathbf{C}_j$ necessarily leads to a contradiction. By the definitions of marginally-connecting and always-connecting independence arcs, we have $\mathbf{C}_i \not\perp\!\!\!\perp \mathbf{C}_j|\mathbf{S} \setminus \mathbf{C}_k$, where $\mathbf{S}$ is a minimal separating set for $\mathbf{C}_i$ and $\mathbf{C}_j$. Assume that the structure over clusters forms a collider, $\mathbf{C}_i \ast\!\!\longrightarrow \mathbf{C}_k \longleftarrow\!\!\ast \mathbf{C}_j$. There must exist at least one connecting path, $p$ between $\mathbf{C}_i$ and $\mathbf{C}_j$ through $\mathbf{C}_k$. Since the edge between $\mathbf{C}_i$ and $\mathbf{C}_k$ is assumed to have an arrowhead into $\mathbf{C}_k$, there must be a pair of nodes $V_i \in \mathbf{C}_i$ and $V_k \in \mathbf{C}_k$ on $p$ such that $V_i \ast\!\!\longrightarrow V_k$. By the admissibility of the partition and definition 3.3, an edge of the form $V_{i'} \leftarrow V_{k'}$ can not exist $\forall V_{i'} \in \mathbf{C}_i, V_{k'} \in \mathbf{C}_k$. To have a connecting path, there may be no collider on $p$, so every subsequent edge between $V_k, V_{k+1} \in \mathbf{C}_k$ along the path $p$ must be of the form $V_k \to V_{k+1}$. Because $\mathbf{C}_k \leftarrow \mathbf{C}_j$, there must also exist some $V_j \in \mathbf{C}_j$ and some $V'_k \in \mathbf{C}_k$ such that $V'_k \leftarrow V_j$ where $V'_k$ is on $p$. Because of the assumption of the admissibility of the partition and by definition 3.3, there can be no edge $V'k \to V_j$. Then there must exist a collider and there is a contradiction. Therefore, the triplet $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j\rangle$ must be a non-collider. □

**Remark D.2.** *In an $\alpha$C-DAG, a conditionally-connecting independence arc annotating triplet $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j\rangle$ could imply either a collider or a non-collider structure in its corresponding $\alpha$C-MAG whenever $|\mathbf{C}_k| > 1$ and always implies a collider structure when $|\mathbf{C}_k| = 1$.*

*Proof.* Consider an unshielded triplet $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j\rangle$ in an $\alpha$C-DAG $G_\mathbf{C}$ such that $\mathcal{A}_{\mathbf{C}_i,\mathbf{C}_k,\mathbf{C}_j}$ is a conditionally-connecting independence arc. This implies that $\mathbf{C}_i \perp\!\!\!\perp \mathbf{C}_j|\mathbf{S} \setminus \mathbf{C}_k; \mathbf{C}_i \not\perp\!\!\!\perp \mathbf{C}_j|\mathbf{C}_k \cup \mathbf{S}$ where $\mathbf{S}$ is a minimal separating set for $\mathbf{C}_i$ and $\mathbf{C}_j$ then there must exist some collider path $p = V_i, ..., V_k, ..., V_j$ in the underlying ADMG $G$ where $V_i \in \mathbf{C}_i$, $V_k \in \mathbf{C}_k$, and $V_j \in \mathbf{C}_j$, such that every non-endpoint node is a collider, and there can be no connecting paths from $\mathbf{C}_i$ to $\mathbf{C}_j$ through $\mathbf{C}_k$.

We first prove that when $|\mathbf{C}_k| = 1$, a conditionally-connecting arc implies a collider structure in the $\alpha$C-MAG. If $|\mathbf{C}_k| = 1$, then there must exist some path, $p = \langle V_i, V_k, V_j\rangle$ in $G$ where $V_i \in \mathbf{C}_i$, $V_k \in \mathbf{C}_k$ and $V_j \in \mathbf{C}_j$ in $G_\mathbf{C}$, such that every non-endpoint node is a collider for $\mathbf{C}_i \not\perp\!\!\!\perp \mathbf{C}_j|\mathbf{C}_k \cup \mathbf{S}$ to hold. Since $V_k$ is the only non-endpoint node on $p$, it must be a collider. Because $V_k$ is the only node in $\mathbf{C}_k$, if $V_k$ were an ancestor of either $V_i$ or $V_j$, then there would be a connecting path from $\mathbf{C}_i$ to $\mathbf{C}_j$ through $\mathbf{C}_k$ and there would be a contradiction. Therefore, $V_k$ is a definite non-ancestor of $V_i$ and $V_j$, and $\mathbf{C}_k$ is a definite non-ancestor of $\mathbf{C}_i$ and $\mathbf{C}_j$ in the associated $\alpha$C-MAG.

Next we show that when $|\mathbf{C}_k| > 1$, a conditionally-connecting arc could imply either a collider or non-collider structure in the corresponding $\alpha$C-MAG. If $|\mathbf{C}_k| > 1$, then there must exist some path, $p = \langle V_i, ..., V_{k_1}, ..., V_{k_n}, ..., V_j \rangle$ in $G$ where $V_i \in \mathbf{C}_i$, $\{V_{k_1}, ..., V_{k_n}\} \subseteq \mathbf{C}_k$ and $V_j \in \mathbf{C}_j$ in $G_{\mathbf{C}}$, such that every non-endpoint node is a collider so that $\mathbf{C}_i \not\perp\!\!\!\perp \mathbf{C}_j | \mathbf{C}_k \cup \mathbf{S}$ holds. Subpath $p' = \langle V_{k_1}, ..., V_{k_1} \rangle$ must then have all vertices connected with bidirected edges. The edges between $V_i$ and $V_{k_1}$ and between $V_{k_n}$ and $V_j$ would need to be oriented such that $V_{k_1}$ and $V_{k_n}$ are colliders, which could be either through directed edges, or bidirected edges. If the edges between $V_i$ and $V_{k_1}$ and between $V_{k_n}$ and $V_j$ were directed, they must be directed as $V_i \rightarrow V_{k_1}$ and $V_{k_n} \leftarrow V_j$ respectively, to preserve the collider path. In this case, edges in both the $\alpha$C-DAG and $\alpha$C-MAG would be oriented as $\mathbf{C}_i \rightarrow \mathbf{C}_k \leftarrow \mathbf{C}_j$, creating a collider structure. However, if the edges on $p$ between $V_i$ and $V_{k_1}$ and between $V_{k_n}$ and $V_j$ are bidirected, then there is a possibility of directed edges, between these or other variables in $\mathbf{C}_i$ and $\mathbf{C}_k$, and $\mathbf{C}_k$ and $\mathbf{C}_j$. If $\exists V_{i'} \in \mathbf{C}_i, \exists V_{k'} \in \mathbf{C}_k$ such that $V_{k'} \in An(V_{i'})$ or if $\exists V_{j'} \in \mathbf{C}_j, \exists V_{k'} \in \mathbf{C}_k$ such that $V_{k'} \in An(V_{j'})$, then triplet $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$ could be oriented as a non-collider in the corresponding $\alpha$C-MAG. □

**Remark D.3.** *In an $\alpha$C-DAG, a never-connecting independence arc could imply either a collider or a non-collider structure in a corresponding $\alpha$C-MAG.*

*Proof.* Consider a triplet $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$ such that $\mathcal{A}_{\mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j}$ is a never-connecting independence arc. This implies that $\mathbf{C}_i \perp\!\!\!\perp \mathbf{C}_j | \mathbf{S} \setminus \mathbf{C}_k ; \mathbf{C}_i \perp\!\!\!\perp \mathbf{C}_j | \mathbf{S} \cup \mathbf{C}_k$, where $\mathbf{S}$ is a minimal separating set for $\mathbf{C}_i$ and $\mathbf{C}_j$. Then either there is no path from any $V_i \in \mathbf{C}_i$ to some $V_j \in \mathbf{C}_j$ through $\mathbf{C}_k$, or every such path $p$ must include at least 4 nodes, $p = V_i, ..., V_{k_1}, V_{k_2}, ..., V_j$ where $V_i \in \mathbf{C}_i$, $V_{k_1}, V_{k_2}, \in \mathbf{C}_k$, and $V_j \in \mathbf{C}_j$, such that there is at least one collider triplet that is not an inducing path and at least one non-collider triplet that is not an inducing path on $p$. Consider the latter case. Let $p$ be a path of exactly 4 nodes $\langle V_i, V_{k_1}, V_{k_2}, V_j \rangle$ such that $V_i \in \mathbf{C}_i$, $V_{k_1}, V_{k_2}, \in \mathbf{C}_k$ and $V_j \in \mathbf{C}_j$. Either $V_{k_1}$ is a collider node and $V_{k_2}$ is a non-collider node or $V_{k_1}$ is a non-collider node and $V_{k_2}$ is a collider node. In the first case, $V_i \ast\!\!\longrightarrow V_{k_1} \longleftarrow\!\!\ast V_{k_2} \rightarrow V_j$ or $V_i \ast\!\!\longrightarrow V_{k_1} \leftarrow V_{k_2} \longleftarrow\!\!\ast V_j$. In the second case, $V_i \ast\!\!\longrightarrow V_{k_1} \rightarrow V_{k_2} \longleftarrow\!\!\ast V_j$ or $V_i \leftarrow V_{k_1} \ast\!\!\longrightarrow V_{k_2} \longleftarrow\!\!\ast V_j$. Then $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$ may be either a collider or a non-collider. Adding any additional node, $V_{k_i+1}$, to $p$ creates additional collider and/or non-collider nodes, but still allows for collider and non-collider structures over clusters. Now consider where there is no path from any $V_i \in \mathbf{C}_i$ to some $V_j \in \mathbf{C}_j$ through $\mathbf{C}_k$. Then the direction of any edge $V_i \ast\!\!-\!\!\ast V_k$ or $V_k' \ast\!\!-\!\!\ast V_j$ can be in any direction such that $\langle \mathbf{C}_i, \mathbf{C}_k, \mathbf{C}_j \rangle$ may be either a collider or a non-collider. □

**Remark 4.1.** *If all clusters contain more than one variable, it is not possible to learn any orientations.*

*Proof.* The proof follow from remarks D.1, D.2, D.3. Because marginally-connecting and always-connecting independence arcs are always non-colliders, and non-colliders can be oriented many ways, and because conditionally-connecting and never-connecting arcs could be oriented as either colliders or non-colliders, knowing the independence arc for a triplet does not provide sufficient information to orient any edges if the clusters have more than one variable. □

**Theorem 3.8.** *Two $\alpha$C-MAGs $\mathcal{M}_1$, $\mathcal{M}_2$ (with the same partition $\mathbf{C}$ of the same set of variables $\mathbf{V}$) are cluster Markov equivalent iff they share the same adjacencies, independence arcs on unshielded and discriminated manipulated unshielded triplets, and separation and connection marks.*

*Proof.* Adjacencies inform for which pairs of clusters there is no separating set. If an adjacency exists in $\mathcal{M}_1$ but not $\mathcal{M}_2$ then there is an m-separation in $\mathcal{M}_2$ not in $\mathcal{M}_1$. If all adjacencies are shared, then the same pairs of variables in the two graphs do not have a separating set.

It follows naturally from the definition of d-separation, which extends to $\alpha$C-MAGs, that the same independence arcs on unshielded triplets including the same separation and connection marks, are required for paths (common between the two graphs given the same adjacencies) comprised of these triplets to share the same status of being connecting or blocked. Regarding shielded triplets, if a component manipulated unshielded triplet, $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$ is discriminated, then there exists some pair of clusters, $\mathbf{A}, \mathbf{B}$ for which their marginal and conditional independence or dependence would change depending on the independence arc $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}^{-\mathbf{XY}}}$. These cases are described and the proofs are shown below for the independence arc determination rules 4-15.

If one of these cases (described in independence arc determination rules 4-15) does not apply for $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}^{-\mathbf{XY}}}$, then all d-separations in the graph will be the same regardless of the independence arc annotating the manipulated unshielded triplet. Given the shielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ in isolation, no independence arcs can be determined for the manipulated unshielded triplets as all clusters are always dependent on each other. If there is an additional variable $\mathbf{W}$ such that $\mathbf{Y}$ and $\mathbf{W}$ are adjacent and $\mathbf{X}, \mathbf{Z}$ and $\mathbf{W}$ are not, then not all triplets in the graph are shielded, potentially allowing for an independence arc to be determined. However, if 1) $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ are both marginally-connecting, 2) $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ are both conditionally connecting, or 3) $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is always connecting or never-connecting, then the paths $p_1 = \langle \mathbf{X}, \mathbf{Y}, \mathbf{W} \rangle$ and $p_2 = \langle \mathbf{X}, \mathbf{Z}, \mathbf{Y}, \mathbf{W} \rangle$ will always be blocked or open under the same conditions (with the given conditioning set) regardless of the independence arc of $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$. Then, $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$ can have any independence arc and the d-separations in the graph won't change. In all cases, whether or not cluster $\mathbf{Y}$ blocks the path determines if $\mathbf{X}$ and $\mathbf{Z}$ (and any upstream clusters on the path) are connected to $\mathbf{W}$ (and any downstream clusters on the path).

The other context where a manipulated unshielded triplet may be discriminated is in the context of a discriminating path, where a similar logic applies.

□

For the proof of soundness of LC-CLOC, first we prove the soundness and completeness of the independence arc determination rules and the soundness of the edge orientation rules. The proof for the soundness of LC-CLOC follows.

The independence arc determination rules are sound and complete for independence arc assignment. The soundness for rules $\mathcal{AR}_1, \mathcal{AR}_2$, and $\mathcal{AR}_3$ follow directly from definition 2.1. Proofs for the soundness of the remaining independence arc determination rules follow.

Consider a shielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ and manipulated unshielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$. In isolation, no independence arc can be assigned to $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$ as the information flow through the manipulated unshielded triplet cannot be isolated from edge $\mathbf{X} \circ\!\!-\!\!\circ \mathbf{Y}$. Therefore, to determine an arc for a manipulated unshielded triplet, at least one more node must be connected to the corresponding shielded triplet. Call this node $\mathbf{W}$ such that $\mathbf{W}$ is adjacent to $\mathbf{Y}$, $\mathbf{Y} \circ\!\!-\!\!\circ \mathbf{W}$ and $\mathbf{W}$ is not adjacent to $\mathbf{X}$. With this structure, there are two paths between $\mathbf{X}$ and $\mathbf{W}$. Let $p_1 = \langle \mathbf{X}, \mathbf{Y}, \mathbf{W} \rangle$ and let $p_2 = \langle \mathbf{X}, \mathbf{Z}, \mathbf{Y}, \mathbf{W} \rangle$.

$\mathcal{AR}_4$: For a shielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ with $\mathbf{W} \circ\!\!-\!\!\circ \mathbf{Y}$ and where $\{\mathbf{X}, \mathbf{Z}\}$ and $\mathbf{W}$ are not adjacent, if $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is conditionally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, $\mathbf{Z} \notin SepSet(\mathbf{X}, \mathbf{W})$, $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W} | \mathbf{Y} \cup SepSet(\mathbf{X}, \mathbf{W})$, and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W} | \mathbf{Z}, \mathbf{Y} \cup SepSet(\mathbf{X}, \mathbf{W})$, then draw an always-connecting arc for $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

*Proof.* If $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is conditionally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, then $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is either marginally-connecting or never-connecting. If $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is marginally-connecting or never-connecting, then $p_1$ is blocked conditional on $\mathbf{Y}$. On $p_2$, triplet $\langle \mathbf{Z}, \mathbf{Y}, \mathbf{W} \rangle$ is active when conditioning on $\mathbf{Y}$. Then, if $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W} | \mathbf{Y}$ and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W} | \mathbf{Y}, \mathbf{Z}$, then there must be a connecting path and a collider path from $\mathbf{X}$ to $\mathbf{Y}$ through $\mathbf{Z}$, and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ must be always-connecting. □

$\mathcal{AR}_5$: For a shielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ with $\mathbf{W} \circ\!\!-\!\!\circ \mathbf{Y}$ and where $\{\mathbf{X}, \mathbf{Z}\}$ and $\mathbf{W}$ are not adjacent, if $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is conditionally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, $\mathbf{Z} \notin SepSet(\mathbf{X}, \mathbf{W})$, and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W} | \mathbf{Z} \cup SepSet(\mathbf{X}, \mathbf{W})$, then draw a conditionally-connecting arc for $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

*Proof.* If $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is conditionally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, then $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is either marginally-connecting or never-connecting. If $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is marginally-connecting or never-connecting, then $p_1$ is blocked conditional on $\mathbf{Y}$. On $p_2$, triplet $\langle \mathbf{Z}, \mathbf{Y}, \mathbf{W} \rangle$ is active when conditioning on $\mathbf{Y}$. Then, if $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W} | \mathbf{Z}$ there must be a collider path (and no connecting path) from $\mathbf{X}$ to $\mathbf{Y}$ through $\mathbf{Z}$ and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ must be conditionally-connecting. □

$\mathcal{AR}_6$: For a shielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ with $\mathbf{W} \circ\!\!-\!\!\circ \mathbf{Y}$ and where $\{\mathbf{X}, \mathbf{Z}\}$ and $\mathbf{W}$ are not adjacent, if $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is conditionally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, and $\mathbf{Z} \in SepSet(\mathbf{X}, \mathbf{W})$, then draw a marginally-connecting arc for $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

*Proof.* If $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is conditionally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, then $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is either marginally-connecting or never-connecting. If $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is marginally-connecting or never-connecting, then $p_1$ is blocked conditional on $\mathbf{Y}$. On $p_2$, triplet $\langle \mathbf{Z}, \mathbf{Y}, \mathbf{W} \rangle$ is active when conditioning on $\mathbf{Y}$. Then, if $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W} | \mathbf{Y}$ and $\mathbf{X} \perp\!\!\!\perp \mathbf{W} | \mathbf{Y}, \mathbf{Z}$, then there must be a connecting path (and no collider path) from $\mathbf{X}$ to $\mathbf{Y}$ through $\mathbf{Z}$ and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ must be marginally-connecting. ☐

$\mathcal{AR}_7$: For a shielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ with $\mathbf{W} \circ\!\!-\!\!\circ \mathbf{Y}$ and where $\{\mathbf{X}, \mathbf{Z}\}$ and $\mathbf{W}$ are not adjacent, if $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is conditionally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, then draw a never-connecting arc for any unmarked $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

*Proof.* If $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is conditionally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, then $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is either marginally-connecting or never-connecting. If $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is marginally-connecting or never-connecting, then $p_1$ is blocked conditional on $\mathbf{Y}$. On $p_2$, triplet $\langle \mathbf{Z}, \mathbf{Y}, \mathbf{W} \rangle$ is active when conditioning on $\mathbf{Y}$. Then, if $\mathbf{X} \perp\!\!\!\perp \mathbf{W} | \mathbf{Y}$ and $\mathbf{X} \perp\!\!\!\perp \mathbf{W} | \mathbf{Y}, \mathbf{Z}$, then there must be no connecting path and no collider path from $\mathbf{X}$ to $\mathbf{Y}$ through $\mathbf{Z}$ and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ must be never-connecting. ☐

$\mathcal{AR}_8$: For a shielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ with $\mathbf{W} \circ\!\!-\!\!\circ \mathbf{Y}$ and where $\{\mathbf{X}, \mathbf{Z}\}$ and $\mathbf{W}$ are not adjacent, if $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is marginally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, $\mathbf{Z} \notin SepSet(\mathbf{X}, \mathbf{W})$, $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W} | SepSet(\mathbf{X}, \mathbf{W}) \setminus Y$, and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W} | \mathbf{Z} \cup SepSet(\mathbf{X}, \mathbf{W}) \setminus Y$, then draw an always-connecting arc for $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

*Proof.* If $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is marginally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, then $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is either conditionally-connecting or never-connecting. If $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is conditionally-connecting or never-connecting, then $p_1$ is blocked with no conditioning set. On $p_2$, triplet $\langle \mathbf{Z}, \mathbf{Y}, \mathbf{W} \rangle$ is active with no conditioning set. Then, if $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W}$ and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W} | \mathbf{Z}$, there must be a connecting and a collider path from $\mathbf{X}$ to $\mathbf{Y}$ through $\mathbf{Z}$ and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ must be always-connecting. ☐

$\mathcal{AR}_9$: For a shielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ with $\mathbf{W} \circ\!\!-\!\!\circ \mathbf{Y}$ and where $\{\mathbf{X}, \mathbf{Z}\}$ and $\mathbf{W}$ are not adjacent, if $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is marginally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, $\mathbf{Z} \notin SepSet(\mathbf{X}, \mathbf{W})$, and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W} | \mathbf{Z} \cup SepSet(\mathbf{X}, \mathbf{Y})$, then draw a conditionally-connecting arc for $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

*Proof.* If $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is marginally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, then $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is either conditionally-connecting or never-connecting. If $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is conditionally-connecting or never-connecting, then $p_1$ is blocked with no conditioning set. On $p_2$, triplet $\langle \mathbf{Z}, \mathbf{Y}, \mathbf{W} \rangle$ is active with no conditioning set. Then, if $\mathbf{X} \perp\!\!\!\perp \mathbf{W}$ and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W} | \mathbf{Z}$, there must be a collider path and no connecting path from $\mathbf{X}$ to $\mathbf{Y}$ through $\mathbf{Z}$ and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ must be conditionally-connecting. ☐

$\mathcal{AR}_{10}$: For a shielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ with $\mathbf{W} \circ\!\!-\!\!\circ \mathbf{Y}$ and where $\{\mathbf{X}, \mathbf{Z}\}$ and $\mathbf{W}$ are not adjacent, if $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is marginally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, and $\mathbf{Z} \in SepSet(\mathbf{X}, \mathbf{W})$, then draw a marginally-connecting arc for $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

*Proof.* If $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is marginally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, then $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is either conditionally-connecting or never-connecting. If $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is conditionally-connecting or never-connecting, then $p_1$ is blocked with no conditioning set. On $p_2$, triplet $\langle \mathbf{Z}, \mathbf{Y}, \mathbf{W} \rangle$ is active with no conditioning set. Then, if $\mathbf{X} \not\perp\!\!\!\perp \mathbf{W}$ and $\mathbf{X} \perp\!\!\!\perp \mathbf{W} | \mathbf{Z}$, there must be a connecting and no collider path from $\mathbf{X}$ to $\mathbf{Y}$ through $\mathbf{Z}$ and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ must be marginally-connecting. ☐

$\mathcal{AR}_{11}$: For a shielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ with $\mathbf{W} \circ\!\!-\!\!\circ \mathbf{Y}$ and where $\{\mathbf{X}, \mathbf{Z}\}$ and $\mathbf{W}$ are not adjacent, if $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is marginally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, then draw a never-connecting arc for any unmarked $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle^{-\mathbf{XY}}$.

*Proof.* If $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is marginally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is not, then $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is either conditionally-connecting or never-connecting. If $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ is conditionally-connecting or never-connecting, then $p_1$ is blocked with no conditioning set. On $p_2$, triplet $\langle \mathbf{Z}, \mathbf{Y}, \mathbf{W} \rangle$ is active with no conditioning set. Then, if $\mathbf{X} \perp\!\!\!\perp \mathbf{W}$ and $\mathbf{X} \perp\!\!\!\perp \mathbf{W}|\mathbf{Z}$, there must be no connecting and no collider path from $\mathbf{X}$ to $\mathbf{Y}$ through $\mathbf{Z}$ and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ must be never-connecting. □

$\mathcal{AR}_{12}$: If $p = \langle \mathbf{X}, ..., \mathbf{W}, \mathbf{V}, \mathbf{Y} \rangle$ is a discriminating path for $\mathbf{V}$ where $\mathbf{Q}$ is the set of all nodes between $\mathbf{X}$ and $\mathbf{V}$, if $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}$ and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}, \mathbf{V}$, $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}^{-\mathbf{WY}}}$ is always-connecting.

*Proof.* If $p = \langle \mathbf{X}, ..., \mathbf{W}, \mathbf{V}, \mathbf{Y} \rangle$ is a discriminating path for $\mathbf{V}$, then conditioning on all nodes $\mathbf{Q}$ between $\mathbf{X}$ and $\mathbf{V}$ on $p$ blocks all paths $\langle \mathbf{X}, \mathbf{Q}_i, \mathbf{Y} \rangle \forall \mathbf{Q}_i \in \mathbf{Q}$ and opens path $\langle \mathbf{X}, ..., \mathbf{W}, \mathbf{V} \rangle$. If $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}$ and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}, \mathbf{V}$, then there must be a connecting path and a collider path from $\mathbf{W}$ to $\mathbf{Y}$ through $\mathbf{V}$ and $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}^{-\mathbf{WY}}}$ must be always-connecting. □

$\mathcal{AR}_{13}$: If $p = \langle \mathbf{X}, ..., \mathbf{W}, \mathbf{V}, \mathbf{Y} \rangle$ is a discriminating path for $\mathbf{V}$ where $\mathbf{Q}$ is the set of all nodes between $\mathbf{X}$ and $\mathbf{V}$, if $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}$ and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}, \mathbf{V}$, $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}^{-\mathbf{WY}}}$ is marginally-connecting.

*Proof.* If $p = \langle \mathbf{X}, ..., \mathbf{W}, \mathbf{V}, \mathbf{Y} \rangle$ is a discriminating path for $\mathbf{V}$, then conditioning on all nodes $\mathbf{Q}$ between $\mathbf{X}$ and $\mathbf{V}$ on $p$ blocks all paths $\langle \mathbf{X}, \mathbf{Q}_i, \mathbf{Y} \rangle \forall \mathbf{Q}_i \in \mathbf{Q}$ and opens path $\langle \mathbf{X}, ..., \mathbf{W}, \mathbf{V} \rangle$. If $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}$ and $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}, \mathbf{V}$, then there must be a connecting path and no collider path from $\mathbf{W}$ to $\mathbf{Y}$ through $\mathbf{V}$ and $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}^{-\mathbf{WY}}}$ must be marginally-connecting. □

$\mathcal{AR}_{14}$: If $p = \langle \mathbf{X}, ..., \mathbf{W}, \mathbf{V}, \mathbf{Y} \rangle$ is a discriminating path for $\mathbf{V}$ where $\mathbf{Q}$ is the set of all nodes between $\mathbf{X}$ and $\mathbf{V}$, if $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}$ and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}, \mathbf{V}$, $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}^{-\mathbf{WY}}}$ is conditionally-connecting.

*Proof.* If $p = \langle \mathbf{X}, ..., \mathbf{W}, \mathbf{V}, \mathbf{Y} \rangle$ is a discriminating path for $\mathbf{V}$, then conditioning on all nodes $\mathbf{Q}$ between $\mathbf{X}$ and $\mathbf{V}$ on $p$ blocks all paths $\langle \mathbf{X}, \mathbf{Q}_i, \mathbf{Y} \rangle \forall \mathbf{Q}_i \in \mathbf{Q}$ and opens path $\langle \mathbf{X}, ..., \mathbf{W}, \mathbf{V} \rangle$. If $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}$ and $\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}, \mathbf{V}$, then there must be no connecting path and a collider path from $\mathbf{W}$ to $\mathbf{Y}$ through $\mathbf{V}$ and $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}^{-\mathbf{WY}}}$ must be conditionally-connecting. □

$\mathcal{AR}_{15}$: If $p = \langle \mathbf{X}, ..., \mathbf{W}, \mathbf{V}, \mathbf{Y} \rangle$ is a discriminating path for $\mathbf{V}$ where $\mathbf{Q}$ is the set of all nodes between $\mathbf{X}$ and $\mathbf{V}$, if $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}$ and $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}, \mathbf{V}$, $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}^{-\mathbf{WY}}}$ is never-connecting.

*Proof.* If $p = \langle \mathbf{X}, ..., \mathbf{W}, \mathbf{V}, \mathbf{Y} \rangle$ is a discriminating path for $\mathbf{V}$, then conditioning on all nodes $\mathbf{Q}$ between $\mathbf{X}$ and $\mathbf{V}$ on $p$ blocks all paths $\langle \mathbf{X}, \mathbf{Q}_i, \mathbf{Y} \rangle \forall \mathbf{Q}_i \in \mathbf{Q}$ and opens path $\langle \mathbf{X}, ..., \mathbf{W}, \mathbf{V} \rangle$. If $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}$ and $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}|\mathbf{Q}, \mathbf{V}$, then there must be no connecting path and no collider path from $\mathbf{W}$ to $\mathbf{Y}$ through $\mathbf{V}$ and $\mathcal{A}_{\mathbf{W},\mathbf{V},\mathbf{Y}^{-\mathbf{WY}}}$ must be never-connecting. □

We then prove that when none of the independence arc determination rules applies, it is not possible to determine the independence arc for the manipulated unshielded triplet.

*Proof.* There are three component manipulated unshielded triplets in a shielded triplet. In isolation, it is impossible to determine the arcs for the shielded triplet. However, with at least one additional node connected to the structure, we can leverage a path over four or more nodes to determine the arc in some cases. Consider a shielded triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$ with $\mathbf{W}$ adjacent to $\mathbf{Y}$ and not adjacent to $\mathbf{Z}$ or $\mathbf{X}$. Let $p_1 = \langle \mathbf{X}, \mathbf{Y}, \mathbf{W} \rangle$ and let $p_2 = \langle \mathbf{X}, \mathbf{Z}, \mathbf{Y}, \mathbf{W} \rangle$ If $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ is never-connecting, then, whether $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ is marginally-connecting, conditionally-connecting, or never-connecting, $p_2$ will always be inactive. Therefore $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ cannot be determined. If $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ are conditionally-connecting, then to isolate $p_2$, $\mathbf{Y}$ should not be conditioned on so that $p_1$ is inactive, but then $p_2$ will also be blocked. If $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ are marginally-connecting, then to isolate $p_2$, $\mathbf{Y}$ needs to be conditioned on to block $p_1$, but then $p_2$ will also be blocked. Whether $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ is marginally-connecting, conditionally-connecting, or never-connecting, $p_2$ will always be inactive. Therefore $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}^{-\mathbf{XY}}}$ cannot be determined.

The discriminating path follows a structure corresponding to $\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ being marginally-connecting and $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ being conditionally-connecting. With this structure, $\mathcal{A}_{\mathbf{Z},\mathbf{X},\mathbf{Y}^{-\mathbf{ZY}}}$ can be determined. If $\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{W}}$ (or any analogous triplets on the discriminating path) is not conditionally-connecting, or

$\mathcal{A}_{\mathbf{Z},\mathbf{Y},\mathbf{W}}$ (or any analogous triplets on the discriminating path) is not marginally-connecting, then $\mathcal{A}_{\mathbf{Z},\mathbf{X},\mathbf{Y}^{-\mathbf{ZY}}}$ cannot be determined.

$\mathcal{A}_{\mathbf{X},\mathbf{Y},\mathbf{Z}^{-\mathbf{ZX}}}$ can never be determined in this context because there is no simple path between $\mathbf{X}$ and $\mathbf{Z}$ beyond the shielded triplet. □

Next we show the proof for the soundness of each orientation rule.

$\mathscr{R}_0$: For a triplet $\langle \mathbf{X}, \mathbf{Z}, \mathbf{Y} \rangle$, if $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ is conditionally-connecting, and $|\mathbf{Z}| = 1$, then orient the triplet as $\mathbf{X} \ast\!\!\rightarrow \mathbf{Z} \leftarrow\!\!\ast \mathbf{Y}$.

*Proof.* The proof follows directly from remark D.2. □

$\mathscr{R}_1$: If $\mathbf{X} \ast\!\!\longrightarrow \mathbf{Z} \circ\!\!-\!\!\ast \mathbf{Y}$, $\mathbf{X}$ and $\mathbf{Y}$ are not adjacent, and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ is marginally-connecting or always-connecting, then orient the triplet as $\mathbf{X} \ast\!\!\rightarrow \mathbf{Z} \rightarrow \mathbf{Y}$.

*Proof.* The proof follows directly from remark D.1. □

$\mathscr{R}_2$: If $\mathbf{X} \rightarrow \mathbf{Z} \ast\!\!\rightarrow \mathbf{Y}$ or $\mathbf{X} \ast\!\!\rightarrow \mathbf{Z} \rightarrow \mathbf{Y}$ and $\mathbf{X} \ast\!\!-\!\!\circ \mathbf{Y}$, then orient $\mathbf{X} \ast\!\!-\!\!\circ \mathbf{Y}$ as $\mathbf{X} \ast\!\!\rightarrow \mathbf{Y}$.

*Proof.* The soundness of the rule comes from observing that if $\mathbf{X} \leftarrow\!\!\ast \mathbf{Y}$, a cycle or almost-directed cycle over clusters would be induced violating the properties of $\alpha$C-MAGs and $\alpha$C-PAGs, and violating the admissible partition criteria for the underlying $\alpha$C-DAG. □

$\mathscr{R}_3$: If $\mathbf{X} \ast\!\!\rightarrow \mathbf{Z} \leftarrow\!\!\ast \mathbf{Y}$, $\mathbf{X} \ast\!\!-\!\!\circ \mathbf{W} \circ\!\!-\!\!\ast \mathbf{Y}$, $\mathbf{X}$ and $\mathbf{Y}$ are not adjacent, $\mathbf{W} \ast\!\!-\!\!\circ \mathbf{Z}$, and $\mathcal{A}_{\mathbf{X},\mathbf{W},\mathbf{Y}}$ is marginally-connecting, then orient $\mathbf{W} \ast\!\!-\!\!\circ \mathbf{Z}$ as $\mathbf{W} \ast\!\!\rightarrow \mathbf{Z}$.

*Proof.* The soundness of the rule comes from observing that if $\mathbf{W} \leftarrow\!\!\ast \mathbf{Z}$, then by two applications of rule 2, $\mathbf{Y} \ast\!\!\rightarrow \mathbf{W}$, $\mathbf{X} \ast\!\!\rightarrow \mathbf{W}$, there would be a collider at $\mathbf{W}$. Since $\mathcal{A}_{\mathbf{X},\mathbf{W},\mathbf{Y}}$ is marginally-connecting, there is a contradiction by remark D.1. □

$\mathscr{R}_4$: If $\mathbf{X} \circ\!\!-\!\!\circ \mathbf{Z} \circ\!\!-\!\!\circ \mathbf{Y}$ and $\mathbf{X} \leftarrow\!\!\circ \mathbf{Y}$ and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}^{-\mathbf{XY}}}$ is marginally-connecting or always-connecting, and orient $\mathbf{X} \circ\!\!-\!\!\circ \mathbf{Z}$ as $\mathbf{X} \leftarrow\!\!\ast \mathbf{Z}$.

*Proof.* If $\mathbf{X} \circ\!\!-\!\!\circ \mathbf{Z}$ were oriented as $\mathbf{X} \rightarrow \mathbf{Z}$, then, due to the independence arc, $\mathbf{X}$ must be an ancestor of $\mathbf{Y}$ and there would be a contradiction. Therefore, it must be that $\mathbf{X} \leftarrow\!\!\ast \mathbf{Z}$. □

$\mathscr{R}_5$: If $\mathbf{X} \ast\!\!-\!\!\ast \mathbf{Z} \ast\!\!-\!\!\ast \mathbf{Y}$, $\mathbf{Z} - \mathbf{W}$, $\mathbf{X}$ and $\mathbf{W}$ are not adjacent, $\mathbf{Y}$ and $\mathbf{W}$ are not adjacent, and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}}$ is never-connecting or conditionally-connecting with connection mark $\oplus_{\boldsymbol{\mathcal{D}}}$ such that $\mathbf{W} \in \boldsymbol{\mathcal{D}}$, then orient $\mathbf{Z} - \mathbf{W}$ as $\mathbf{Z} \rightarrow \mathbf{W}$.

*Proof.* The soundness of the rule comes from the definition of a connection mark, $\oplus_{\mathcal{D}}$, where any cluster $\mathbf{W} \in \mathcal{D}$ must be a descendant of a collider, such that $\mathbf{Z} \rightarrow \mathbf{W}$. □

$\mathscr{R}_6$: If $\mathbf{X} \rightarrow \mathbf{Z} \rightarrow \mathbf{Y}$, $\mathbf{X} \circ\!\!\longrightarrow \mathbf{Y}$, and $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}^{-\mathbf{XY}}}$ is marginally-connecting or always-connecting, orient $\mathbf{X} \circ\!\!\longrightarrow \mathbf{Y}$ as $\mathbf{X} \rightarrow \mathbf{Y}$.

*Proof.* If $\mathcal{A}_{\mathbf{X},\mathbf{Z},\mathbf{Y}^{-\mathbf{XY}}}$ is marginally-connecting or always-connecting, then $\mathbf{X}$ is a true ancestor of $\mathbf{Y}$, so if $\mathbf{X} \circ\!\!\longrightarrow \mathbf{Y}$ were oriented as $\mathbf{X} \leftrightarrow \mathbf{Y}$ there would be a contradiction. □

$\mathscr{R}_7$: If $\mathbf{X} \circ\!\!\longrightarrow \mathbf{Y}$, $p = \langle \mathbf{X}, \mathbf{Z}, \mathbf{W}, \ldots \mathbf{Y} \rangle$ is an uncovered partially directed path from $\mathbf{X}$ to $\mathbf{Y}$ such that $\mathbf{Y}$ and $\mathbf{Z}$ are not adjacent, and $\mathcal{A}_{\mathbf{Z},\mathbf{X},\mathbf{Y}}$ is marginally-connecting, orient $\mathbf{X} \circ\!\!\longrightarrow \mathbf{Y}$ as $\mathbf{X} \rightarrow \mathbf{Y}$.

*Proof.* By definition of an uncovered possibly directed path, there can be no collider along path $p$. Then either the path is directed from $\mathbf{X}$ to $\mathbf{Y}$ such that $\mathbf{X} \in An(\mathbf{Y})$ and so $\mathbf{X} \rightarrow \mathbf{Y}$, or $\mathbf{X} \notin An(\mathbf{Y})$ and there is an arrowhead into $\mathbf{X}$ so $\mathbf{X} \rightarrow \mathbf{Y}$ to avoid a collider that would contradict the independence arc. □

$\mathscr{R}_8$: Suppose $\mathbf{X} \circ\!\!\longrightarrow \mathbf{Y}$, $\mathbf{Z} \rightarrow \mathbf{Y} \leftarrow \mathbf{W}$, $p_1$ is an uncovered possibly directed path from $\mathbf{X}$ to $\mathbf{Z}$, and $p_2$ is an uncovered possibly directed path from $\mathbf{X}$ to $\mathbf{W}$. Let $\mathbf{V}$ be the vertex adjacent to $\mathbf{X}$ on $p_1$ ($\mathbf{V}$ could be $\mathbf{Z}$) and $\mathbf{T}$ be the vertex adjacent to $\mathbf{X}$ on $p_2$ ($\mathbf{T}$ could be $\mathbf{W}$). If $\mathbf{V}$ and $\mathbf{Z}$ ($\mathbf{T}$ and $\mathbf{W}$) are distinct, let $\mathcal{A}_{\mathbf{V},\mathbf{Z},\mathbf{Y}}$ ($\mathcal{A}_{\mathbf{T},\mathbf{W},\mathbf{Y}}$) be marginally-connecting with no separation mark subscripted by a variable on $p_1$ ($p_2$). If $\mathbf{V}$ and $\mathbf{T}$ are distinct and not adjacent, and $\mathcal{A}_{\mathbf{T},\mathbf{X},\mathbf{Y}}$ and $\mathcal{A}_{\mathbf{V},\mathbf{X},\mathbf{Y}}$ are marginally-connecting or always-connecting, then orient $\mathbf{X} \circ\!\!\longrightarrow \mathbf{Y}$ as $\mathbf{X} \rightarrow \mathbf{Y}$.

*Proof.* If the edges between either $\mathbf{V}$ and $\mathbf{X}$ or $\mathbf{T}$ and $\mathbf{X}$ are oriented with arrowheads into $\mathbf{X}$, then there would be a collider structure that would contradict the marginally- or always-connecting arc of $\mathcal{A}_{\mathbf{T},\mathbf{X},\mathbf{Y}}$ or $\mathcal{A}_{\mathbf{V},\mathbf{X},\mathbf{Y}}$. If neither edge is into $\mathbf{X}$, then $\mathbf{X}$ is a definite ancestor of $\mathbf{Y}$, so $\mathbf{X} \rightarrow \mathbf{Y}$. $\square$