INTERVENTIONAL MARKOV EQUIVALENCE FOR MIXED GRAPH MODELS

LIAM SOLUS

ABSTRACT. We study the problem of characterizing Markov equivalence of graphical models under general interventions. For DAGs, this problem is solved using data from an interventional setting to refine MECs of DAGs into smaller, interventional MECs. A recent graphical characterization of interventional MECs of DAGs relates to their global Markov property. Motivated by this, we generalize interventional MECs to all loopless mixed graphs via their global Markov property and generalize the graphical characterization given for DAGs to ancestral graphs. We also extend the notion of interventional Markov equivalence probabilistically: via invariance properties of distributions Markov to acyclic directed mixed graphs (ADMGs). We show that this generalization aligns with the standard causal interpretation of ADMGs. Finally, we show the two generalizations coincide at their intersection, thereby completely generalizing the characterization for DAGs to directed ancestral graphs.

1. Introduction

A major goal in statistics is to predict the effects of interventions amongst jointly distributed random variables. This is relevant in fields such as medicine, computational biology, and economics where practitioners are interested in altering predicted outcomes [11, 15, 19]. Oftentimes, the goal is to learn a causal DAG, which encodes a set of conditional independence (CI) relations and whose edges encode causal relations amongst the variables. However, using observational data alone, it is only possible to learn a causal DAG up to the CI relations it encodes, which are the same for all DAGs in its Markov equivalence class (MEC) [2].

By further sampling from interventional distributions, i.e., distributions produced by modifying variables within the system, identifiability of a causal DAG can be improved. For example, perfect (or hard) interventions [4], which force the targeted variables to the values of independent variables, partition an MEC of DAGs into perfect $\mathcal{I}$-MECs [6]. Similarly, general interventions, which do not necessarily force the targeted variables to become independent, partition an MEC of DAGs into $\mathcal{I}$-MECs [23]. Recently, [6] and [23] gave graphical characterizations of perfect $\mathcal{I}$-MECs and (general) $\mathcal{I}$-MECs for DAGs, respectively. These characterizations play a key role in algorithms that use observational and interventional data to learn causal DAGs [6, 22, 23], with the latter two being the only such provably consistent algorithms; i.e., they return a member of the $\mathcal{I}$-MEC of the true DAG if the data-generating distribution is faithful to $\mathcal{G}$.

While DAG models are fundamental, they make restrictive assumptions that can be alleviated by working with more general mixed graph models. For example, biologists are often interested in learning protein synthesis networks that use both directed and bidirected arrows [16]. Mixed graph models associate CI relations to a graph with edges that may be directed, bidirected, or undirected, and thereby incorporate latent confounders and selection variables.
into the model. Very recently, [8] gave some first results towards extending the theory of $\mathcal{I}$-Markov equivalence to such models by characterizing $\mathcal{I}$-Markov equivalence for directed ancestral graphs (directed AGs) under the assumption that the interventions are controlled. The methods presented in this paper generalize and extend the results of [8] in several ways: The main results presented here show that any graphical characterization of Markov equivalence for directed AGs yields a graphical characterization of $\mathcal{I}$-Markov equivalence, one of which corresponds to the characterization of [8]. Moreover, the results presented here apply to all ancestral graphs (i.e., not only directed AGs) and do not require that the interventions be controlled. Additionally, we show that the interventional distributions associated to directed AGs admit a factorization that is subject to invariance properties analogous to those satisfied by an interventional distribution associated to a DAG.

In this paper, we first generalize the notion of $\mathcal{I}$-MECs to all formal independence models for loopless mixed graphs by way of their global Markov property. We prove the graphical characterization of $\mathcal{I}$-MECs of [23] can be extended to all ancestral graphs (AGs) under reasonable assumptions on the intervention targets. We then show that the notion of general interventional distributions can be extended to all acyclic directed mixed graphs (ADMGs) via their factorization criterion, and we use this to prove a probabilistic version of $\mathcal{I}$-Markov equivalence for ADMGs. We then prove that any distribution arising as the marginal of a distribution Markov to a causal DAG over the observed variables following an intervention will be an interventional distribution with respect to the ADMG that is the latent projection of the DAG onto the observed variables. This generalizes the results of [8] for directed AGs to all ADMGs. Finally, we prove that our generalization of $\mathcal{I}$-MECs for AGs via formal independence models and the probabilistic version for ADMGs coincide at their intersection; i.e., for all directed AGs. This completely generalizes the result of [23] to a family of mixed graph models that allow for the incorporation of latent confounders and avoids the assumption of [8] that the interventions need to be controlled.

2. Independence Models and Mixed Graphs

We first recall the necessary basics of mixed graph models, formal independence models, and interventions. All necessary graph theory terms are defined in the appendix. A mixed graph is a graph $\mathcal{G} = ([p], E)$ in which the set of edges $E$ contains a mixture of undirected $i - j$, bidirected $i \leftrightarrow j$, and directed $i \rightarrow j$ edges. If it is ambiguous as to whether or not there is an arrowhead at the endpoint $i$ of an edge $e$, we will place a question mark after $i$; for example $i \circ \rightarrow j$. A loopless mixed graph (LMG) is a mixed graph that does not contain any loops; i.e. edges which have both endpoints being the same node. We will call an LMG simple if between any two nodes there is at most one edge. We can define a notion of separation within an LMG that generalizes the notion of $d$-separation in DAGs [13]: Let $\mathcal{G} = ([p], E)$ be an LMG and $C \subset [p]$. A path $\pi$ is $m$-connecting given $C$ in $\mathcal{G}$ if all of its collider subpaths $\langle v_1, v_2, v_3 \rangle$ have $v_2 \in C \cup \text{an}_G(C)$ and all of its nodes that are not $v_2$ for some collider subpath are not in $C$. For two disjoint subsets $A, B \subset [p]$, we say that $C$ $m$-separates $A$ and $B$ in $\mathcal{G}$ if there is no $m$-connecting path between any node in $A$ and any node in $B$ given $C$ in $\mathcal{G}$.

To an LMG $\mathcal{G} = ([p], E)$, we then associate a collection of triples $\mathcal{J}(\mathcal{G})$, where $\langle A, B \mid C \rangle \in \mathcal{J}(\mathcal{G})$ if and only if $A$ and $B$ are $m$-separated given $C$ in $\mathcal{G}$. We call the set of triples $\mathcal{J}(\mathcal{G})$ the (formal) independence model for $\mathcal{G}$. In general, $\mathcal{G}$ is not uniquely determined by $\mathcal{J}(\mathcal{G})$, which leads to the notion of Markov equivalence:
Definition 1. Two LMGs $G$ and $H$ are Markov equivalent, denoted $G \approx H$, and belong to the same Markov equivalence class (MEC) if and only if $\mathcal{J}(G) = \mathcal{J}(H)$.

A distribution $\mathbb{P}$ over random variables $X_1, \ldots, X_p$ is Markov with respect to $G$ if $\mathbb{P}$ entails $X_A \perp \perp X_B \mid X_C$ whenever $\langle A, B \mid C \rangle \in \mathcal{J}(G)$. We let $\mathcal{M}(G)$ denote the collection of all distributions $\mathbb{P}$ that are Markov to $G$. If there exists $\mathbb{P}$ such that $\mathcal{J}(G)$ encodes precisely the set of CI relations entailed by $\mathbb{P}$ then we call $\mathcal{J}(G)$ a probabilistic independence model. A number of families of LMGs yield well-studied probabilistic independence models.

The structure of an LMG $G = ([p], E)$ is often intimately tied to the factorizations of the distributions within $\mathcal{M}(G)$. The LMG $G$ is called a directed acyclic graph (DAG) if it contains only directed edges and no directed cycles. In this case, a distribution $f$ factorizes in a simple way:

\[
f(x) = \prod_{i \in [p]} f(x_i \mid x_{\text{pa}_G(i)}).
\]

In the case of DAGs, typically all members of an MEC only have a subset of their directed arrows in common. Thus, we can only learn a subset of the causal relations between nodes by sampling from the observational distribution. One way to improve identifiability of the causal DAG, i.e., learn more of the directed arrows, is to perform interventions on variables in the system and sample data from the resulting interventional distributions. Let $X = (X_1, \ldots, X_p)$ be a random vector with joint distribution $\mathbb{P}$ Markov to a DAG $G = ([p], E)$, and let $f^{(0)}$ denote its density function. Given a subset $I \subset [p]$, called an intervention target, a distribution with density $f^{(I)}$ admitting a factorization

\[
f^{(I)}(x) = \prod_{i \in I} f^{(I)}(x_i \mid x_{\text{pa}_G(i)}) \prod_{i \notin I} f^{(0)}(x_i \mid x_{\text{pa}_G(i)})
\]

is called an interventional distribution. The intervention is called perfect whenever $f^{(I)}(x_i \mid x_{\text{pa}_G(i)}) = f^{(I)}(x_i)$ for all $i \in I$.

Let $I$ be a multiset of intervention targets, and $(f^{(I)})_{I \in I}$ be a sequence of distributions over $X = (X_1, \ldots, X_p)$ indexed by the elements of $I$. For a DAG $G = ([p], E)$, define the set of intervention settings for $G$ with respect to $I$ as

\[
\mathcal{M}_I(G) := \{ (f^{(I)})_{I \in I} \mid \forall I, J \in I: f^{(I)} \in \mathcal{M}(G), \text{ and } f^{(I)}(x_i \mid x_{\text{pa}_G(i)}) = f^{(J)}(x_i \mid x_{\text{pa}_G(i)}) \text{ for all } i \notin I \cup J \}.
\]

In \cite{Glymour2016}, an intervention setting $(f^{(I)})_{I \in I} \in \mathcal{M}_I(G)$ is called controlled if $f^{(I)}(x_i \mid x_{\text{pa}_G(i)}) = f^{(J)}(x_i \mid x_{\text{pa}_G(i)})$ whenever $i \in I \cap J$. Since $\mathcal{M}_I(G)$ represents the collection of all sequences of distributions that can be generated from the interventions $I$, \cite{Glymour2016} used this to formally define Markov equivalence of DAGs under general interventions:

Definition 2. \cite{Glymour2016} Definition 3.4] For a collection of intervention targets $I$, two DAGs $G$ and $H$ are $I$-Markov equivalent and belong to the same $I$-Markov equivalence class ($I$-MEC) if $\mathcal{M}_I(G) = \mathcal{M}_I(H)$.

Notice that this definition, which relies on invariance properties of the factorization \cite{Pearl1988}, is fundamentally different than the definition of Markov equivalence of DAGs given by Definition \cite{Pearl1988} which is phrased in terms of global Markov properties. However, the result of \cite{Glymour2016} Theorem 3.9] demonstrates that $I$-Markov equivalence of DAGs can also be defined using the global Markov property of DAGs. This suggests a generalization to mixed graph
models, for which it is easier to work with global Markov properties since the corresponding factorization criteria generalizing (1) are not as simple (or simply unknown).

2.1. Interventional Graphs and Global $\mathcal{I}$-Markov Properties. We now recall the main result of [23] so as to set the stage for our corresponding generalization.

**Definition 3.** Let $\mathcal{G} = ([p], E)$ be an LMG and $\mathcal{I}$ a collection of intervention targets. The *interventional graph* for $\mathcal{I}$ ($\mathcal{I}$-LMG) is the graph $\mathcal{G}^\mathcal{I}$ with nodes $[p] \cup W_\mathcal{I}$, where $W_\mathcal{I} := \{\omega_\mathcal{I} | \mathcal{I} \in \mathcal{I} \setminus \{\emptyset\}\}$, and edges $E \cup E_\mathcal{I}$, where $E_\mathcal{I} := \{\omega_\mathcal{I} \rightarrow i | i \in \mathcal{I} \setminus \{\emptyset\}\}$.

Figure 1 shows an LMG and its $\mathcal{I}$-LMG.$^1$ [21] proved two DAGs are Markov equivalent if and only if they have the same adjacencies and the same $v$-structures. The characterization of $\mathcal{I}$-MECs given by [23] extends this result via the combinatorics of $\mathcal{I}$-DAGs. When considered in the context of [21], their characterization frames $\mathcal{I}$-Markov equivalence for DAGs in terms of the global Markov property for DAGs. In the following, if $(i, j, k)$ is a $v$-structure in an $\mathcal{I}$-LMG $\mathcal{G}^\mathcal{I}$ such that $i$ or $k$ is in $W_\mathcal{I}$, we call it an $\mathcal{I}$-$v$-structure.

**Theorem 1.** [23, Theorem 3.9] Let $\mathcal{I}$ be a collection of interventional targets for which $\emptyset \in \mathcal{I}$. Two DAGs $\mathcal{G}$ and $\mathcal{H}$ are $\mathcal{I}$-Markov equivalent if and only if $\mathcal{G}^\mathcal{I}$ and $\mathcal{H}^\mathcal{I}$ have the same $\mathcal{I}$-$v$-structures.

It is worth noting that [10] also consider $\mathcal{I}$-LMGs in their Joint Causal Inference (JCI) framework, in which they use the $\mathcal{I}$-LMG to learn causal structure by applying classic causal discovery algorithms to the $\mathcal{I}$-LMG. On the other hand, they do not provide explicit graphical characterizations of $\mathcal{I}$-MECs as in Theorem 1 or as is the main goal of this paper.

By [21], Theorem 1 equivalently states that two DAGs $\mathcal{G}$ and $\mathcal{H}$ are $\mathcal{I}$-Markov equivalent if and only if $\mathcal{G}^\mathcal{I}$ and $\mathcal{H}^\mathcal{I}$ are Markov equivalent. The former interpretation (stated in Theorem 1) captures the intuition that intervening on nodes of a graph should “lock” more arrows in place; that is, determine more causal relations. However, the latter interpretation suggests a natural generalization of $\mathcal{I}$-Markov equivalence to formal independence models for LMGs, free of a factorization criterion, and based solely on global Markov properties:

$^1$As in [23], we treat interventions as parameters instead of random variables so as to avoid problems with faithfulness violations as discussed in [10]. To make this distinction in our depictions of $\mathcal{I}$-LMGs, we draw nodes corresponding to random variables as open circles containing labels, and we draw interventional nodes in $W_\mathcal{I}$ as filled, black circles.
Definition 4. Let $\mathcal{I}$ be a collection of intervention targets. Two LMGs $\mathcal{G}$ and $\mathcal{H}$ are $\mathcal{I}$-Markov equivalent and belong to the same $\mathcal{I}$-Markov equivalence class ($\mathcal{I}$-MEC) if

$$\mathcal{J}(\mathcal{G}^I) = \mathcal{J}(\mathcal{H}^I).$$

Remark 1. It may seem counterintuitive to define a notion of interventional equivalence based solely on the graph structure and without regard to the invariance properties of the underlying distribution. However, Theorem 1 in fact gives a completely graph-theoretical way of viewing these invariance properties. Recall that the $\mathcal{I}$-DAG $\mathcal{G}^I$ adds a new source node $\omega_I$ to $\mathcal{G}$ such that $\omega_I \rightarrow i$ for all $i \in I$. Theorem 1 then states that the intervention represented by the addition of this node will not change the causal mechanisms of nodes from which the point of intervention can be $d$-separated. For example, the causal mechanism $f(x_j \mid x_{\text{pa}_G(j)})$ for any ancestor $j \notin I$ of a node $i \in I$ will remain unchanged by the intervention $I$, and this is captured graphically by the fact that $\omega_I$ is $d$-separated from $j$ in $\mathcal{G}^I$ given only $W_I \setminus w_I$. Definition 4 is the natural generalization of equivalence under this completely graph-theoretical description of the defining invariance properties of interventions in DAGs.

In Section 4, our choice of abstraction to this level will again be justified since we will show that any two directed AGs are $\mathcal{I}$-Markov equivalent in terms of Definition 4 if and only if they encode the same interventional settings from a more traditional causal perspective.

In the case of DAGs, notice that Theorem 1 would be trivial if we take Definition 4 as the definition of $\mathcal{I}$-Markov equivalence instead of Definition 2. This is because an $\mathcal{I}$-DAG is again a DAG. However, as seen in Figure 1, an $\mathcal{I}$-LMG $\mathcal{G}^I$ need not always be in the same subclass of LMGs as $\mathcal{G}$. This makes generalizing Theorem 1 with respect to Definition 4 nontrivial for classes of LMGs other than DAGs. In particular, we would like to generalize Theorem 1 to one such family of LMGs, called ancestral graphs (AGs), that would allow for modeling interventions in the same fashion as DAGs, but now in the presence of latent confounders and selection variables. A complete generalization of Theorem 1 to all AGs with arbitrary intervention targets is impossible (as we will see in Example 1). However, it will extend to all AGs under reasonable assumptions on the intervention targets. Notice that the generalization given in Definition 4 is for formal independence models, and hence any corresponding generalization of Theorem 1 such as the one we will prove in Theorem 10 will be a purely mathematical result at first. To give such a result probabilistic meaning, it is also necessary to generalize the collection of intervention settings associated to a DAG to more general LMGs and relate such generalizations to Definition 4. This is done in the last section of the paper (Theorem 19), so as to completely generalize Theorem 1 to a more general family of LMGs than DAGs; namely, the directed AGs.

2.2. Ancestral, Ribbonless, Anterior, and Maximal Graphs. We now introduce families of LMGs, and some of their properties, that will be necessary to generalize Theorem 1. A ribbon in $\mathcal{G}$ is a collider path $\langle i, j, k \rangle$ such that in $\mathcal{G}$:

1. There is no endpoint-identical edge between $i$ and $k$; i.e., there is no $i \leftrightarrow k$ edge in the case $i \leftrightarrow j \leftrightarrow k$, no $i \rightarrow k$ edge in the case $i \rightarrow j \leftrightarrow k$, and no $i \leftarrow k$ edge in the case $i \rightarrow j \leftarrow k$, and

2. $j$ or a descendant of $j$ is an endpoint of an undirected edge or is on a directed cycle.

A ribbonless graph (RG) is an LMG containing no ribbons. RGs are probabilistic; i.e., $\mathcal{J}(\mathcal{G})$ is a probabilistic independence model. In fact, they induce the same independence models as MC-graphs [9], and each such model can be induced by marginalizing and conditioning a
A second family of graphs admitting this property are the ancestral graphs. An LMG is an ancestral graph (AG) if:

1. \(G\) contains no directed cycles,
2. whenever \(G\) contains a bidirected edge \(i \leftrightarrow j\), there is no directed path from \(i\) to \(j\) or \(j\) to \(i\) in \(G\), and
3. whenever \(G\) contains an undirected edge \(i \sim j\), neither \(i\) nor \(j\) has an arrowhead pointing towards it.

AGs were introduced by as a generalization of DAG models that are closed under marginalization and conditioning. They offer a means by which to model causation in the presence of latent confounders and selection variables. The bidirected edge \(i \leftrightarrow j\) can be interpreted as representing a triple \(i \leftarrow k \rightarrow j\) where the node \(k\) corresponds to an unobserved random variable. Similarly, the undirected edge \(i \sim j\) can be thought of as a triple \(i \rightarrow k \leftarrow j\) in which the variable corresponding to \(k\) has been conditioned upon (a selection variable).

Thus, we call the endpoints of an undirected edge \(x \sim y\) selection adjacent nodes, and we let \(\text{sa}(G)\) denote the collection of selection adjacent nodes in \(G\).

To generalize Theorem 1 to AGs, we will need to examine the interplay between RGs and AGs. In doing so, we will use the notions of anterior and maximal graphs. A graph \(G\) is called anterior if \(G\) contains no triples of the form \(i \circ \rightarrow j \sim k\). Given an LMG \(G\), the anterior graph of \(G\), denoted \(G^*\), is the graph produced by consecutively removing all arrowheads pointing into nodes adjacent to an undirected edge; i.e. \(i \leftrightarrow j \sim k\) becomes \(i \leftarrow j \sim k\) and \(i \rightarrow j \sim k\) becomes \(i \sim j \sim k\). We make use of the following lemma from [18]:

Lemma 2. [18, Proposition 1] If \(G\) is an RG then \(\mathcal{J}(G) = \mathcal{J}(G^*)\).

An LMG \(G\) is called maximal if adding any edge to \(G\) changes the collection \(\mathcal{J}(G)\). An inducing path \(\pi\) between vertices \(i\) and \(j\) of an AG \(G\) is a path on which any non-endpoint vertex is both a collider on \(\pi\) and an ancestor of at least one of \(i\) or \(j\) in \(G\). By [13, Theorem 5.1], we know that every AG \(G\) is contained in a unique maximal ancestral graph (MAG) \(\overline{G}\), and that \(\overline{G}\) is produced from \(G\) by adding bidirected arrows between the endpoints of the inducing paths in \(G\). A similar result holds for RGs [17]. We will require a few lemmas about maximal graphs, the first of which is immediate:

Lemma 3. An LMG \(G = ([p], E)\) is maximal if and only if for every pair of nonadjacent nodes \(i, j \in [p]\) there exists a subset \(C \subset [p] \setminus \{i, j\}\) such that \(\langle i, j \mid C \rangle \in \mathcal{J}(G)\).

The next lemma takes a bit more work, but is a natural extension of an observation made in [1]. For the sake of completeness, a proof is given in the appendix.

Lemma 4. Let \(G\) and \(H\) be Markov equivalent maximal simple LMGs on node set \([p]\). Then \(G\) and \(H\) have the same adjacencies and v-structures.

Note that Definition 4 makes sense for general LMGs, but Lemma 4 only holds for simple LMGs. This, however, is not a problem since our main result in this paper generalizes Theorem 1 to directed AGs, which are by definition simple.

3. I-Markov Equivalence for AGs

With the definitions and observations of the previous section, we can now describe exactly how Theorem 1 generalizes to AGs. To understand the context in which the desired generalization holds, we consider the following example:
Figure 2. Two I-MAGs $G^I$ and $H^I$ such that $G$ and $H$ are Markov equivalent and $G^I$ and $H^I$ have the same $I$-v-structures, but $G$ and $H$ are not I-Markov equivalent.

Example 1. Let $G$ and $H$ be the Markov equivalent AGs depicted in Figure 2. As seen here, by taking $I = \{\{1, 4\}, \{1, 5\}\}$, the LMGs $G^I$ and $H^I$ have the same $I$-v-structures. However, $G$ and $H$ are not I-Markov equivalent. For example, $\langle \omega_{\{1,5\}}, \omega_{\{1,4\}} | 2 \rangle \in J(G^I)$, but $\langle \omega_{\{1,5\}}, \omega_{\{1,4\}} | 2 \rangle \not\in J(H^I)$. The induced subgraph of $G^I$ on the node set $\{\omega_{\{1,4\}}, \omega_{\{1,5\}}, 1, 2\}$ is a special type of ribbon in $G^I$ called a straight ribbon [18]. The underlying problem appears to be the “double intervention” on selection adjacent nodes, which results in a straight ribbon in the interventional graph. As we will see, assuming that we do not doubly-intervene on selection adjacent nodes is equivalent to assuming that $G^I$ is ribbonless. This will be a sufficient condition for generalizing Theorem 1 to AGs.

Even with the assumption mentioned in Example 1, we still capture many useful cases in causal modeling. For example, it does not rule out cases in which we can target single variables, such as when we generate interventional distributions for studying gene regulatory networks via targeted gene deletions using the CRISPR/CAS-9 system [3]. It is also a trivial assumption if there are no selection variables; i.e. in the case of directed AGs.

It is also worth noting that intervention settings that doubly-intervene on nodes arise naturally in practical contexts. For instance, the intervention setting considered in [16] considers a multiset of intervention targets when learning a specific protein signaling network. Here, researchers are interested in modeling an interventional setting that includes different interventional experiments for up-regulation and down-regulation of targeted molecules. The result is a model containing doubly-intervened nodes, as depicted in Figure 3 in the appendix. In a similar fashion, doubly-intervened nodes such as those depicted in Figure 2 can arise when the reagents used for regulating gene expression target more than one molecule at a time.

Definition 5. Let $G = ([p], E)$ be an LMG and $\mathcal{I}$ a collection of intervention targets. We say that $\mathcal{I}$ doubly-intervenes on a node $i \in [p]$ if there exist two distinct elements $I$ and $J$ in the multiset $\mathcal{I}$ such that $i \in I \cap J$.

Lemma 5. Let $G = ([p], E)$ be an AG, and $\mathcal{I}$ a collection of intervention targets. Then $\mathcal{I}$ does not doubly-intervene on any selection adjacent nodes of $G$ if and only if $G^\mathcal{I}$ is ribbonless.

Remark 2. Lemma 5 does not hold in the case that $G$ is only assumed to be ribbonless. For example, if we consider the RG $G$ from Figure 1 and we intervene with targets $\mathcal{I} = \{\{4\}, \{1, 4\}\}$, then $G^\mathcal{I}$ is not ribbonless even though $\mathcal{I}$ does not doubly-intervene on any selection-adjacent nodes in $G$. This is due to the cyclic ribbon $\langle \omega_{\{1,4\}}, 4, \omega_{\{1\}} \rangle$. 

\[ \omega_{\{1,4\}} \quad 1 \quad \omega_{\{1,5\}} \]
\[ 4 \quad 2 \quad 5 \]
$G^I$

\[ \omega_{\{1,4\}} \quad 1 \quad \omega_{\{1,5\}} \]
\[ 4 \quad 2 \quad 5 \]
$H^I$
In order to generalize Theorem 1 to AGs, we will work with maximal AGs, or MAGs. Our first goal is to show that this is sufficient in the sense that two AGs $G$ and $H$ are $I$-Markov equivalent if and only if their unique maximal extensions $\overline{G}$ and $\overline{H}$ are $I$-Markov equivalent. To do so, we require the following two lemmas.

**Lemma 6.** Let $G = ([p], E)$ be an AG, and $I$ a collection of intervention targets that does not doubly-intervene on any selection adjacent nodes of $G$. Then $\overline{G} = \overline{G^{I}}([p])$.

**Lemma 7.** Let $G = ([p], E)$ be an AG, and $I$ a collection of intervention targets that does not doubly-intervene on any selection adjacent nodes of $G$. If $i \in W_I$ and $j \in [p]$ such that $i$ and $j$ are not adjacent in $G^{I}$, then $i$ and $j$ are not adjacent in $\overline{G^{I}}$.

Using Lemmas 6 and 7, we get the following:

**Proposition 8.** Let $G = ([p], E)$ be an AG, and $I$ a collection of intervention targets that does not doubly-intervene on any selection adjacent nodes in $G$. Then $\overline{G^{I}} = \overline{G^{I}}^{I}$.

Proposition 8 tells us that if $I$ is a collection of intervention targets that does not doubly-intervene on any nodes of $G$ or $H$, then $G$ and $H$ are $I$-Markov equivalent if and only if $G^{I}$ and $H^{I}$ have the same $I$-v-structures. Hence, it suffices to characterize $I$-Markov equivalence when $G$ and $H$ are assumed to be MAGs. Another immediate consequence of Proposition 8 is the following:

**Lemma 9.** Suppose $G = ([p], E)$ is an AG and $I$ a collection of intervention targets that does not doubly-intervene on any nodes of $G$. Then $G$ is maximal if and only if $(G^{I})^{*}$ is a MAG that is Markov equivalent to $G^{I}$.

We can now prove our main theorem in this section.

**Theorem 10.** Let $G$ and $H$ be AGs with node sets $[p]$, and let $I$ be a collection of intervention targets such that no selection adjacent node is doubly-intervened in either $G$ or $H$. The following are equivalent:

1. $G$ and $H$ are $I$-Markov equivalent.
2. $\overline{G}$ and $\overline{H}$ are $I$-Markov equivalent.
3. $\overline{G}$ and $\overline{H}$ are Markov equivalent and $\overline{G^{I}}$ and $\overline{H^{I}}$ have the same $I$-v-structures.

**Proof.** The equivalence of (1) and (2) is an immediate consequence of Proposition 8. So it only remains to show (2) if and only if (3). For notational simplicity, in the remainder of this proof we assume $G$ and $H$ are MAGs. Suppose first that $G$ and $H$ are $I$-Markov equivalent. Then $J(G^{I}) = J(H^{I})$. Hence, $J(G) = J(H)$, so $G$ and $H$ are Markov equivalent MAGs. By Proposition 8 we know that $G^{I}$ and $H^{I}$ are maximal graphs. So by Lemma 4 and the fact that $G^{I}$ and $H^{I}$ are Markov equivalent, we know that $G^{I}$ and $H^{I}$ have the same v-structures. Hence, they have the same $I$-v-structures.

Conversely, suppose that $G$ and $H$ are Markov equivalent and have the same $I$-v-structures. Since $G$ and $H$ are MAGs and $G^{I}$ and $H^{I}$ contain no doubly-intervened selection adjacent nodes, then by Lemma 5 and Proposition 8 we know that $G^{I}$ and $H^{I}$ are maximal RGs.
Thus, by \[18\] Proposition 1, we know that \(G^I\) and \(H^I\) are Markov equivalent to their anterior graphs. Hence,

\begin{equation}
(3) \quad \mathcal{J}(G^I^*) = \mathcal{J}(G^I) = \mathcal{J}(H^I) = \mathcal{J}(H^I^*) .
\end{equation}

Moreover, by Lemma\[9\] we know that \((G^I)^*\) and \((H^I)^*\) are MAGs. So by equality \((3)\) and \[1\] Theorem 3.7, in order to show that \(G\) and \(H\) are \(I\)-Markov equivalent, it suffices to observe that \((G^I)^*\) and \((H^I)^*\) have the same adjacencies and the same colliders with order. Since \(G\) and \(H\) are Markov equivalent MAGs, by the same theorem we know that \(G\) and \(H\) are, respectively, subgraphs of \((G^I)^*\) and \((H^I)^*\) given by deleting all nodes in \(W_I\). Since the nodes \(\omega_I \in W_I\) are, by definition, adjacent to the same nodes in both \((G^I)^*\) and \((H^I)^*\), then it suffices to check that \((G^I)^*\) and \((H^I)^*\) have the same colliders with order that contain nodes in \(W_I\).

By \[1\] Definition 3.11], any collider with order 0 using an interventional node \(\omega_I \in W_I\) is an \(I\)-v-structure. Hence, by assumption, these are the same in both \((G^I)^*\) and \((H^I)^*\). On the other hand, suppose that \((a, b, \omega_I)\) is a triple with order \(i > 0\) in \((G^I)^*\) or \((H^I)^*\). Then by \[1\] Definition 3.11], there is a discriminating path \(\pi = (x, q_1, \ldots, q_m, b, y)\) for \(b\) with either \((a, b, \omega_I) = (q_m, b, y)\) or \((a, b, \omega_I) = (y, b, q_m)\), and the \(m\) colliders

\[\langle x, q_1, q_2 \rangle, \langle q_1, q_2, q_3 \rangle, \ldots, \langle q_{m-1}, q_m, q_b \rangle\]

are all of order at most \(i - 1\). However, by definition of a discriminating path, this would imply that the interventional node \(\omega_I\) has either parents or spouses in \((G^I)^*\) or \((H^I)^*\), which is impossible. Hence, no triple with order \((a, b, \omega_I)\) for \(\omega_I \in W_I\) can have order greater than zero. Thus, we conclude that \((G^I)^*\) and \((H^I)^*\) have the same adjacencies and colliders with order. Therefore, \(G\) and \(H\) are \(I\)-Markov equivalent. 

Condition 3 in Theorem\[10\] is actually a graphical characterization of \(I\)-Markov equivalence. This is because there already exist graphical characterizations of Markov equivalence of MAGs, such as \[1\] Theorem 3.2] or \[24\] Theorem 2.1]. Thus, even though \(G^I\) and \(H^I\) are not necessarily AGs, by \(3\) we can still apply these graphical characterizations and the condition on the \(I\)-v-structures in \(G^I\) and \(H^I\) to get graphical characterizations of \(I\)-Markov equivalence for the AGs \(G\) and \(H\). On the other hand, in the special case of Theorem\[10\] when no selection adjacent nodes are intervened upon in either \(G^I\) or \(H^I\), condition \(3\) is equivalent to saying that \(G\) and \(H\) are \(I\)-Markov equivalent if and only if \(G^I\) and \(H^I\) are Markov equivalent MAGs, which can be deduced by the graphical characterizations of \[1\] Theorem 3.2] or \[24\] Theorem 2.1]. One special case of this that we will consider is the case when \(G\) and \(H\) are directed AGs.

For directed AGs, we also note that Theorem\[10\] places no restrictions on the intervention targets \(I\). Since this case will be of special interest, we state this as corollary below. In the coming section we will see that directed AGs, and more generally ADMGs, also allow for a generalization of the interventional settings \(M_2(G)\) via a known factorization criterion for such models. This will allow us to give probabilistic meaning to Theorem\[10\] and Corollary\[11\]. Using this, we will recover a complete generalization of Theorem\[1\] to directed AGs.

**Corollary 11.** Let \(G\) and \(H\) be directed AGs on node set \([p]\), and let \(I\) be any collection of intervention targets. The following are equivalent:

1. \(G\) and \(H\) are \(I\)-Markov equivalent.
2. \(G^I\) and \(H^I\) are \(I\)-Markov equivalent.
3. $\mathcal{G}$ and $\mathcal{H}$ are Markov equivalent and $\mathcal{G}^I$ and $\mathcal{H}^I$ have the same $\mathcal{I}$-$v$-structures.

Theorem 10 and Corollary 11 can be used to describe $\mathcal{I}$-MECs for families of AGs in a number of contexts arising in practice: Theorem 10 can be used when practitioners are able to design interventional experiments where each interventional target is a singleton. On the other hand, Corollary 11 is applicable when practitioners are not able to place such restrictions on their intervention targets but are willing to model without the incorporation of selection variables.

4. Factorization Criteria and $\mathcal{I}$-ADMGs

We have now characterized a generalization of $\mathcal{I}$-Markov equivalence for DAGs to AGs from the perspective of global Markov properties as specified by Definition 4. However, it is typical to think of interventions as manipulations of variables, and the global Markov property perspective can obscure this. For example, in [2], we see that an interventional density $f(I)$ Markov to a DAG $\mathcal{G}$ can be viewed as a modification of the conditional factors $f(\emptyset)(x_i \mid x_{pa_G(i)})$ of the observational density as specified by the intervention target $I$. In Definition 2 $\mathcal{I}$-Markov equivalence for DAGs is defined in terms of these modified densities, which is fundamentally different than the perspective taken in Definition 4. Since recursive factorizations of the distributions Markov to a general LMG are more complex and less well-understood, it can be beneficial to instead analyze such models from the perspective of global Markov properties, as we did in the previous section. However, the family of *acyclic directed mixed graphs* (ADMGs) admits a relatively nice recursive factorization which we can use to generalize the set of interventional settings $\mathcal{M}_I(\mathcal{G})$ to these more general models. ADMGs are the class of graphs that contain only directed or bidirected edges and no directed cycles. In particular, they contain all directed AGs. In this section, we will extend the definition of the interventional settings $\mathcal{M}_I(\mathcal{G})$ to ADMGs via their factorization criteria, and characterize when two ADMGs $\mathcal{G}$ and $\mathcal{H}$ satisfy $\mathcal{M}_I(\mathcal{G}) = \mathcal{M}_I(\mathcal{H})$. As a corollary, we will prove that for directed AGs, such an equality holds if and only if the two graphs satisfy the conditions of Corollary 11 thereby completely generalizing Theorem 1.
Each $H \in [A]_G$ is a path-connected subgraph of $G_*$, and satisfies $H = \text{barren}_G(\text{an}_G(H))$. For $H \in [A]_G$, we define

$$\text{tail}(H) := \left(\text{dis}_G(H) \setminus H\right) \cup \text{pa}_G\left(\text{dis}_G(H) \setminus \{H\}\right).$$

In [12], the following factorization criterion for ADMGs is proven.

**Theorem 12.** [12, Theorem 4] A probability distribution $\mathbb{P}$ with density function $f$ is Markov to an ADMG $G$ if and only if for every $A \in \mathcal{A}(G)$, the marginal density $f(x_A)$ factors as

$$f(x_A) = \prod_{H \in [A]_G} f(x_H | x_{\text{tail}(H)}).$$

Using Theorem 12, we can extend the notion of an interventional distribution captured in equation (2) to ADMGs:

**Definition 6.** A distribution with density function $f^{(I)}$ is called an interventional distribution with respect to the intervention target $I$ and ADMG $G$ if for all $A \in \mathcal{A}(G)$

$$f(x_A) = \prod_{H \in [A]_G : I \cap H \neq \emptyset} f^{(I)}(x_H | x_{\text{tail}(H)}) \prod_{H \in [A]_G : I \cap H = \emptyset} f^{(0)}(x_H | x_{\text{tail}(H)}),$$

where $f^{(0)}$ is the density of a distribution Markov to $G$.

The collection of sequences of interventional distributions $(f^{(I)})_{I \in \mathcal{I}}$ that can be generated by intervening on a collection of targets $\mathcal{I}$ with respect to an observational density Markov to an ADMG $G$ is then

$$\mathcal{M}_\mathcal{I}(G) := \{ (f^{(I)})_{I \in \mathcal{I}} : \forall I, J \in \mathcal{I} \text{ and } A \in \mathcal{A}(G) : f^{(I)} \in \mathcal{M}(G) \text{ and } f^{(I)}(x_H | x_{\text{tail}(H)}) = f^{(J)}(x_H | x_{\text{tail}(H)}) \forall H \in [A]_G : H \cap (I \cup J) = \emptyset\}.$$  

This fact is seen in the following lemma, whose proof is analogous to that of [23, Lemma A.1].

**Lemma 13.** Suppose $\emptyset \in \mathcal{I}$ and $G$ is an ADMG. Then $(f^{(I)})_{I \in \mathcal{I}} \in \mathcal{M}_\mathcal{I}(G)$ if and only if there exists $f^{(0)} \in \mathcal{M}(G)$ such that for all $I \in \mathcal{I}$, $f^{(I)}$ factorizes according to equation (5).

**Remark 3.** The causally-minded reader may wonder whether our chosen definition for general interventions with respect to an ADMG (Definition 6) and our corresponding generalization of the collection $\mathcal{M}_\mathcal{I}(G)$ to ADMGs capture the intuitive notion of intervention given by interpreting a distribution Markov to an ADMG as the marginal over the observed variables in a distribution Markov to a causal DAG. As we will see with Theorem 17, the answer to this question is “yes.” However, it turns out that this fact is most easily proven by first characterizing when $\mathcal{M}_\mathcal{I}(G) = \mathcal{M}_\mathcal{I}(H)$ for two ADMGs $G$ and $H$. We will now give the necessary characterization.

Definition 6 and (5) are the natural extensions of equations (2) and $\mathcal{M}_\mathcal{I}(G)$ to the family of ADMGs with respect to Theorem 12. Equipped with these definitions, we can now generalize the $\mathcal{I}$-Markov property for DAGs [23, Definition 3.6] by passing from $d$-separation to $m$-separation.

**Definition 7** ($\mathcal{I}$-Markov Property). Let $\mathcal{I}$ be a collection of intervention targets such that $\emptyset \in \mathcal{I}$ and $(f^{(I)})_{I \in \mathcal{I}}$ a set of strictly positive distributions over $X_1, \ldots, X_p$. For an ADMG $G = ([p], E)$, we say that $(f^{(I)})_{I \in \mathcal{I}}$ satisfies the $\mathcal{I}$-Markov property with respect to $G^\mathcal{I}$ if and only if

$$\forall H \in [A]_G \text{ with } I \cap H \neq \emptyset, \forall J \neq H \text{ satisfying } H \cap (I \cup J) = \emptyset, \forall K \in [A]_G \text{ satisfying } H \cap K \neq \emptyset, \forall L \neq K \text{ satisfying } H \cap (K \cup L) = \emptyset,
\text{ and } H \cap (I \cup J) = \emptyset, \forall M \in [A]_G \text{ satisfying } H \cap M \neq \emptyset, \forall N \neq M \text{ satisfying } H \cap (M \cup N) = \emptyset.$$

This $m$-separation property is a natural extension of the notion of $d$-separation for DAGs to ADMGs, and it allows us to characterize the interventional Markov equivalence for mixed graph models in a similar way to the causal DAG case.
1. $X_A \perp A_B \mid X_C$ in $f^{(I)}$ for all $I \in \mathcal{I}$ and any $A, B, C \subset [p]$ for which $(A, B \mid C) \in \mathcal{F}(\mathcal{G})$.
2. $f^{(\emptyset)}(X_B \mid X_C) = f^{(I)}(X_B \mid X_C)$ for any $I \in \mathcal{I}$ and $B, C \subset [p]$ such that $(B, \{\omega_I\} \mid C \cup W_I \setminus \{\omega_I\}) \in \mathcal{F}(\mathcal{G}^I)$.

In [23, Proposition 3.8], the authors showed that for a DAG $\mathcal{G}$, if $\emptyset \in \mathcal{I}$, then $(f^{(I)})_{I \in \mathcal{I}} \in \mathcal{M}_\mathcal{I}(\mathcal{G})$ if and only if $(f^{(I)})_{I \in \mathcal{I}}$ satisfies the $\mathcal{I}$-Markov property with respect to $\mathcal{G}^I$. A key ingredient in the proof of this statement is to argue that if $(f^{(I)})_{I \in \mathcal{I}} \in \mathcal{M}_\mathcal{I}(\mathcal{G})$, then the conditional factors in the product

$$f^{(I)} = \prod_{i \in [p]} f^{(I)}(x_i \mid x_{pa_G(i)})$$

can be partitioned in such a way that, upon marginalization and conditioning, we can recover property 2 of Definition 7. This argument generalizes to ADMGs with the factorization given in Theorem 12. However, it requires that we take a few extra steps to account for the fact that the heads $H \in [A]_G$ are no longer necessarily singletons if $\mathcal{G}$ is not a DAG. In particular, we require the following lemma.

**Lemma 14.** Suppose that $(f^{(I)})_{I \in \mathcal{I}} \in \mathcal{M}_\mathcal{I}(\mathcal{G})$, we have two sets $B, C \subset [p]$, and $I \in \mathcal{I}$ such that $B$ is $m$-separated from $\omega_I$ given $C \cup W_I \setminus \{\omega_I\}$ in $\mathcal{G}^I$. Let $A := \text{ang}(B \cup C) \in \mathcal{A}(\mathcal{G})$. Let $Z \subset A$ denote all vertices in $A$ that are $m$-connected to $\omega_I$ given $C \cup W_I \setminus \{\omega_I\}$ in $\mathcal{G}^I$, and let $B' := A \setminus (Z \cup C)$. Finally, let

$$C' := \{i \in C \mid (pa_G(i) \cup sp_G(i)) \cap B' = \emptyset\}, \quad \text{and}$$

$$C'' := \{i \in C \mid (pa_G(i) \cup sp_G(i)) \cap B' = \emptyset\}.$$

Then

1. No district in $\mathcal{G}$ contains both a node of $B'$ and a node of $Z$.
2. No district in $\mathcal{G}$ contains both a node of $C''$ and a node of $Z$.

**Remark 4.** By a similar argument as in the proof of Lemma 14 (see the appendix), we see that the parents of heads $H$ containing elements of $B'$ and/or $C''$ must be contained in $B' \cup C$. Similarly, the parents of heads $H$ containing elements of $Z$ and/or $C''$ are contained in $Z \cup C''$.

Given Lemma 14 and Remark 4, we can prove the following theorem generalizing [23, Proposition 3.8] to ADMGs.

**Theorem 15.** Suppose $\emptyset \in \mathcal{I}$ and $\mathcal{G}$ is an ADMG. Then $(f^{(I)})_{I \in \mathcal{I}} \in \mathcal{M}_\mathcal{I}(\mathcal{G})$ if and only if $(f^{(I)})_{I \in \mathcal{I}}$ satisfies the $\mathcal{I}$-Markov property with respect to $\mathcal{G}^I$.

4.1. **Interventions in ADMGs Causally.** Since ADMGs can be used to represent causal DAGs with latent confounders, we would like to see that our definition of interventional setting given by Definition 6 and 7 captures the intuitive notion interventional setting arising from this perspective. To formalize and prove this statement we first recall the connection between causal DAGs including latent confounders and ADMGs.

**Definition 8.** Let $\mathcal{G}$ be a DAG with vertex set $[p] \sqcup L$, where the vertices in $[p]$ are observed and those in $L$ are latent, and $\sqcup$ denotes a disjoint union. The *latent projection* $\mathcal{G}([p])$ is a mixed graph with vertex set $[p]$ such that for every pair of distinct vertices $i, j \in [p]$.
1. \( \mathcal{G}([p]) \) contains an edge \( i \to j \) if there exists a directed path in \( \mathcal{G} \) \( i \to \cdots \to j \) on which every non-endpoint vertex is \( L \), and

2. \( \mathcal{G}([p]) \) contains an edge \( i \leftrightarrow j \) if there exists a path in \( \mathcal{G} \) between \( i \) and \( j \) such that the endpoints are all non-colliders in \( L \) and such that the edge adjacent to \( i \) and the edge adjacent to \( j \) both have arrowheads pointing towards \( i \) and \( j \), respectively.

Since \( \mathcal{G} = ([p] \sqcup L, E) \) is a DAG, we see that \( \mathcal{G}([p]) \) is an ADMG. In the following, given \( f \in \mathcal{M}(\mathcal{G}) \), we let \( \tilde{f} := \int_L f \, dL \) denote the marginal density of \( f \) over \([p]\). It is well-known that \( \tilde{f} \) is Markov to \( \mathcal{G}([p]) \) (see for instance [13, Theorem 7.1]).

Given an ADMG \( \mathcal{H} = ([p], E) \) we can always produce a DAG \( \mathcal{G} = ([p] \sqcup L, E') \) such that \( \mathcal{H} = \mathcal{G}([p]) \) by replacing each bidirected arrow \( i \leftrightarrow j \) in \( \mathcal{H} \) with \( i \leftrightarrow \ell_{i,j} \to j \) and letting \( L := \{ \ell_{i,j} : i \leftrightarrow j \text{ an edge in } \mathcal{H} \} \). Thus, given a DAG \( G = ([p] \sqcup L, E) \) and an interventional setting \( (f(I))_{I \in \mathcal{I}} \in \mathcal{M}_\mathcal{I}(\mathcal{G}) \) for \( \mathcal{I} = \{ \emptyset, I_1, \ldots, I_K \} \) with \( I_k \subseteq [p] \) for all \( k \in [K] \) we would like to see that \( (\tilde{f}(I))_{I \in \mathcal{I}} \in \mathcal{M}_\mathcal{I}(\mathcal{G}([p])) \). This would imply that the invariance properties used to define interventional densities and interventional settings for ADMGs in Definition 6 and 3 align with the standard notion given by Definition 2 and the causal interpretation of ADMGs as latent projections of causal DAGs. To prove this, we will make use of the following proposition.

**Lemma 16.** [14, Proposition 4] Let \( \mathcal{G} = ([p] \sqcup L, E) \) be a DAG where \([p] \sqcup L \) is a disjoint union. For disjoint subsets \( A, B, C \subseteq [p] \), with \( C \) possibly empty, \( A \) and \( B \) are \( d \)-separated given \( C \) in \( \mathcal{G} \) if and only if \( A \) and \( B \) are \( m \)-separated given \( C \) in the latent projection \( \mathcal{G}([p]) \).

Combining Lemma 16 with our Theorem 15, we can prove that our definition of interventional setting for ADMGs aligns with the causally intuitive notion.

**Theorem 17.** Let \( \mathcal{G} = ([p] \sqcup L, E) \) be a DAG, where \([p] \sqcup L \) denotes a disjoint union. Let \( \mathcal{I} \) be a collection of intervention targets satisfying \( \emptyset \in \mathcal{I} \) and \( I \subset [p] \) for all \( I \in \mathcal{I} \). If \( (f(I))_{I \in \mathcal{I}} \in \mathcal{M}_\mathcal{I}(\mathcal{G}) \) then \( (\tilde{f}(I))_{I \in \mathcal{I}} \in \mathcal{M}_\mathcal{I}(\mathcal{G}([p])) \).

**Proof.** By Theorem 15 it suffices to show that \( (\tilde{f}(I))_{I \in \mathcal{I}} \) satisfies the \( \mathcal{I} \)-Markov property with respect to \( \mathcal{G}^\mathcal{I} \). Since \( f(I) \in \mathcal{M}(\mathcal{G}) \) for all \( I \in \mathcal{I} \) then \( \tilde{f}(I) \) is Markov to \( \mathcal{G}([p]) \) for all \( I \in \mathcal{I} \) [13, Theorem 7.1]. Hence, \( (\tilde{f}(I))_{I \in \mathcal{I}} \) satisfies the first condition of the \( \mathcal{I} \)-Markov property stated in Definition 7. To see that \( (\tilde{f}(I))_{I \in \mathcal{I}} \) satisfies the second condition, we must show that whenever \( C \cup W_T \setminus \{ w_I \} \) \( m \)-separates \( B \) and \( w_I \) in \( \mathcal{G}([p])^\mathcal{I} \) then \( \tilde{f}(0)(X_B \mid X_C) = \tilde{f}(I)(X_B \mid X_C) \).

To this end, notice first that since \( B, C \subset [p] \) we have for all \( I \in \mathcal{I} \) that

\[
\tilde{f}(I)(X_B, X_C) = \int_{L \setminus [p] \setminus \{ B, C \}} f(I) = \int_{[p] \setminus \{ B, C \}} f(I) = \tilde{f}(I)(X_B, X_C)
\]

and similarly \( \tilde{f}(I)(X_C) = \tilde{f}(I)(X_C) \). It then follows that

\[
f(I)(X_B \mid X_C) = \tilde{f}(I)(X_B \mid X_C)
\]

for all \( I \in \mathcal{I} \). Moreover, since \( \mathcal{G}^\mathcal{I} \) is a DAG and \( \mathcal{G}^\mathcal{I}([p]) = \mathcal{G}([p])^\mathcal{I} \), then by Lemma 16 we know that \( B \) is \( d \)-separated from \( w_I \) given \( C \cup W_T \setminus \{ w_I \} \) in \( \mathcal{G}^\mathcal{I} \) if and only if \( B \) is \( m \)-separated from \( w_I \) given \( C \cup W_T \setminus \{ w_I \} \) in \( \mathcal{G}([p])^\mathcal{I} \). Since \( (f(I))_{I \in \mathcal{I}} \in \mathcal{M}_\mathcal{I}(\mathcal{G}) \), it follows from [23, Proposition 3.8] that \( f(0)(X_B \mid X_C) = f(I)(X_B \mid X_C) \) whenever \( B \) is \( d \)-separated from \( w_I \) given \( C \cup W_T \setminus \{ w_I \} \) in \( \mathcal{G}^\mathcal{I} \). So by Lemma 16 and (7), we conclude that \( \tilde{f}(0)(X_B \mid X_C) = \tilde{f}(I)(X_B \mid X_C) \) whenever \( B \) is \( d \)-separated from \( w_I \) given \( C \cup W_T \setminus \{ w_I \} \) in \( \mathcal{G}([p])^\mathcal{I} \). This completes the proof. \( \square \)
Remark 5. Very recently, another definition of $\mathcal{I}$-Markov equivalence for directed AGs appeared in [8, Definition 2]. In [8], the authors defined an interventional distribution associated to a directed AG $\mathcal{G}$ to be a distribution $\bar{f}^{(1)}$ given by marginalizing over the observed variables in an interventional distribution $f^{(1)}$ with respect to a DAG $\mathcal{H} = ([p] \sqcup L, E)$ for which $\mathcal{G} = \mathcal{H}([p])$. Theorem [17] shows that the definition of interventional distribution given in Definition [6] captures the causally intuitive definition used by [8]. Moreover, it shows that the invariance properties defining the intervention in $\mathcal{H}$ before marginalization over the observed variables are translated directly into analogous invariance properties in the factorization criterion for all ADMGs given by Theorem [12].

4.2. Directed Ancestral Graphs. Since we now know that our definition of interventional settings for ADMGs aligns with the causal interpretation of ADMGs as latent projections of DAGs and their associated interventional settings, our last goal is to connect this notion with the combinatorial characterization of $\mathcal{I}$-Markov equivalence for formal independence models given in Theorem [10]. Theorem [15] says two ADMGs $\mathcal{G}$ and $\mathcal{H}$ satisfy $\mathcal{M}_I(\mathcal{G}) = \mathcal{M}_I(\mathcal{H})$ if and only if they have the same $\mathcal{I}$-Markov properties. Whereas, Theorem [10] characterizes $\mathcal{I}$-MECs of AGs via global Markov properties. Our last theorem says these two generalizations coincide for LMGs that are both AGs and ADMGs; that is, for directed AGs. To prove this, we require a lemma whose proof is derived from that of [23, Lemma 3.10]:

Lemma 18. Suppose that $\mathcal{G} = ([p], E)$ is an ADMG, $A, C \subset [p]$ are disjoint, and $I \in \mathcal{I}$ is such that $\omega_I$ and $A$ are m-connected given $C \cup W_I \setminus \{\omega_I\}$ in $\mathcal{G}^I$ by a path $\pi$ that contains no bidirected edges. Then there exists a sequence $(f^{(I)})_{i \in \mathcal{I}}$ satisfying the $\mathcal{I}$-Markov property with respect to $\mathcal{G}^I$ such that $f^{(0)}(x_A \mid x_C) \neq f^{(1)}(x_A \mid x_C)$.

The following theorem then unifies all our previous results, and it completely generalizes Theorem [1] to a context in which we can incorporate the presence of latent confounders in our models:

Theorem 19. Let $\mathcal{G}$ and $\mathcal{H}$ be two directed AGs and let $\mathcal{I}$ be a collection of intervention targets. Suppose $\emptyset \in \mathcal{I}$. Then the following are equivalent:

1. $\mathcal{G}$ and $\mathcal{H}$ are $\mathcal{I}$-Markov equivalent,
2. $\mathcal{G}$ and $\mathcal{H}$ are Markov equivalent and $\mathcal{G}^I$ and $\mathcal{H}^I$ have the same $\mathcal{I}$-v-structures,
3. $\mathcal{M}_I(\mathcal{G}) = \mathcal{M}_I(\mathcal{H})$.

Proof. The equivalence of 1 and 2 was shown in Corollary [11]. So it only remains to show the equivalence of 1 and 3. Suppose first that $\mathcal{G}$ and $\mathcal{H}$ are $\mathcal{I}$-Markov equivalent. Then $\mathcal{G}^I$ and $\mathcal{H}^I$ have the same set of m-separation statements. Hence, by Theorem [15], $\mathcal{M}_I(\mathcal{G}) = \mathcal{M}_I(\mathcal{H})$.

Conversely, suppose that $\mathcal{G}$ and $\mathcal{H}$ are not $\mathcal{I}$-Markov equivalent. By Corollary [11], it then follows that either $\mathcal{G}$ and $\mathcal{H}$ are not Markov equivalent or $\mathcal{G}$ and $\mathcal{H}$ are Markov equivalent and (without loss of generality) $\mathcal{G}^I$ contains an $\mathcal{I}$-v-structure that is not in $\mathcal{H}^I$. Since $\mathcal{G}$ and $\mathcal{H}$ are maximal Markov equivalent graphs then, by Lemma [8] we know that they have the same adjacencies. Thus, $\mathcal{G}^I$ and $\mathcal{H}^I$ also have the same adjacencies. Consequently, if $\omega_I \rightarrow j \leftarrow k$ is an $\mathcal{I}$-v-structure in $\mathcal{G}^I$ that is not in $\mathcal{H}^I$, we know that we have the induced path $\omega_I \rightarrow j \rightarrow k$ in $\mathcal{H}^I$. Moreover, since $\mathcal{G}$ and $\mathcal{H}$ are directed MAGs, they contain no selection adjacent nodes. So by Proposition [8], we have that $\mathcal{G}^I$ and $\mathcal{H}^I$ are also MAGs. Thus, by [13, Corollary 5.3], we know that $\omega_I$ and $k$ are m-separated given $\text{an}_{\mathcal{G}}(k) \cup W_I \setminus \{\omega_I\}$ in $\mathcal{G}^I$. On the other hand, we also know that $\omega_I$ and $k$ are m-separated given $\text{an}_{\mathcal{H}}(k) \cup W_I \setminus \{\omega_I\}$, which is a contradiction.
but $m$-connected given $\arg_{\omega}(k) \cup W_k \setminus \{\omega_I\}$ in $\overline{\mathcal{G}}$. Hence, by Lemma 18 there exists $\{f(I)\}_{I \in \mathcal{I}}$ that is $\mathcal{I}$-Markov to $\overline{\mathcal{H}}$ with $f(\emptyset)(x_k | x_{\arg_{\omega}(k)}) \neq f(I)(x_k | x_{\arg_{\omega}(k)})$. However, since $\omega_I$ and $k$ are $m$-separated given $\arg_{\omega}(k)$ in $\overline{\mathcal{G}}$, it follows that $\{f(I)\}_{I \in \mathcal{I}}$ cannot satisfy the $\mathcal{I}$-Markov property with respect to $\overline{\mathcal{G}}$. This completes the proof. □

Remark 6. Finally, we note that Theorem 19 provides a generalization of the very recent result [8, Theorem 2]. In [8], the authors work with controlled interventions, whereas Theorem 19 applies to the more general context of interventional settings defined in (6). Additionally, Theorem 19 does not specifically characterize $\mathcal{I}$-MECs for directed AGs in terms of their discriminating paths, but it instead allows one to use any combinatorial characterization of Markov equivalence of AGs to produce a combinatorial characterization of $\mathcal{I}$-AGs. By applying Theorem 19 to the characterization of Markov equivalence of AGs given in [1], we recover [8, Theorem 2]. On the other hand, if we instead apply Theorem 19 to the characterization given in [24], we recover an alternative characterization of $\mathcal{I}$-Markov equivalence for directed AGs.

5. Discussion

Here, we studied the problem of characterizing Markov equivalence of mixed graph models under general interventions. First, we extended the notion of interventional Markov equivalence of DAG models to formal independence models for LMGs via their global Markov properties. From this formal, mathematical perspective, we proved that the graphical characterization of $\mathcal{I}$-MECs identified for DAG models by [23] generalizes to AGs under reasonable assumptions on the interventional targets $\mathcal{I}$.

We then generalized the notion of interventional settings $\mathcal{M}_{\mathcal{I}}(\mathcal{G})$ for DAGs to ADMGs via their factorization criterion. We proved that a sequence of distributions $\{f(I)\}_{I \in \mathcal{I}}$ is an intervention setting for a collection of targets $\mathcal{I}$ and an ADMG $\mathcal{G}$ if and only if it satisfies the $\mathcal{I}$-Markov property for $\overline{\mathcal{G}}$, thereby generalizing [23, Proposition 3.8] to all ADMGs. We then used this fact to show that our choice of definition of general intervention with respect to an ADMG is justified from a causal perspective. Namely, we showed that any distribution arising as the marginal of a distribution Markov to a causal DAG $\mathcal{G}$ over the observed variables, following an intervention at some of the observed variables, will be an interventional distribution with respect to the latent projection of $\mathcal{G}$ onto the observed variables in terms of Definition 6. Finally, in Theorem 19 we showed that our graphical characterization of $\mathcal{I}$-Markov equivalence for two directed AGs (Corollary 11) is equivalent to the directed AGs having the same set of interventional settings. In this case, our results place no restrictions on the intervention targets $\mathcal{I}$, and hence yield a complete generalization of Theorem 1 to directed AGs. Since Theorem 19 places no restrictions on the interventional settings, and since Corollary 11 provides a mechanism for producing multiple characterizations of $\mathcal{I}$-Markov equivalence, this additionally generalizes the recent result [8, Theorem 2] in two ways. This result allows for the improved modeling of causal systems under general interventions in the presence of latent confounders.

The results presented here suggest a number of directions for future work: First, if we could identify a graphical characterization of MECs of all simple ADMGs, or conversely a factorization criterion for all AGs, we could then hope to generalize Theorem 19 to these broader families of LMGs. Second, the characterizations given here for $\mathcal{I}$-MECs set the
stage for extending causal discovery algorithms that rely on graphical characterizations of Markov equivalence to similar algorithms that allow for latent confounders. Such algorithmic extensions could improve our ability to provide accurate causal models based on new interventional data sets arising in computational biology and genomics.

References

[1] A. R. Ali, T. S. Richardson, and P. Spirtes. Markov equivalence for ancestral graphs. *The Annals of Statistics*, 37(5B):2808–2837, 2009.
[2] S. A. Andersson, D. Madigan, and M. D. Perlman. A characterization of markov equivalence classes for acyclic digraphs. *The Annals of Statistics*, 25(2):505–541, 1997.
[3] A. Dixit, O. Parnas, B. Li, J. Chen, C. P. Fuoco, L. Jerby-Arnon, N. D. Marjanovic, D. Dionne, T. Burks, R. Raychowdhury, and B. Adamson. Perturb-seq: dissecting molecular circuits with scalable single-cell rna profiling of pooled genetic screens. *Cell*, 167(7):1853–1866, 2016.
[4] F. Eberhardt, C. Glymour, and R. Scheines. On the number of experiments sufficient and in the worst case necessary to identify all causal relations among n variables. *Uncertainty in Artificial Intelligence*, pages 178–184, 2005.
[5] S. B. Gillispie and M. D. Perlman. Enumerating markov equivalence classes of acyclic digraphs. *Uncertainty in Artificial Intelligence*, pages 171–177, 2001.
[6] A. Hauser and P. Buhlmann. Characterization and greedy learning of interventional markov equivalence classes of directed acyclic graphs. *Journal of Machine Learning Research*, 77(1):291–318, 2012.
[7] A. Hauser and P. Buhlmann. Jointly interventional and observational data: estimation of interventional markov equivalence classes of directed acyclic graphs. *J. R. Statist. Soc. B*, 13(Aug):2409–2464, 2015.
[8] M. Kocaoglu, A. Jaber, K. Shanmugam, and E. Barenboim. Characterization and learning of causal graphs with latent variables from soft interventions. *The proceedings of the 32-nd Conference on Advances in Neural Information Processing Systems (NeurIPS)*.
[9] J. T. A. Koster. Marginalizing and conditioning in graphical models. *Bernoulli*, 8(6):817–840, 2002.
[10] J. M. Mooij, S. Magliacane, and T. Claassen. Joint causal inference from multiple contexts. *arXiv:1611.10351v4*, 2019.
[11] J. Pearl. Causality: Models, reasoning, and inference. *Cambridge University Press, Cambridge*, 2000.
[12] T. Richardson. A factorization criterion for acyclic directed mixed graphs. *The proceedings of the 25-th Conference on Uncertainty in Artificial Intelligence*, 2009.
[13] T. Richardson and P. Spirtes. Ancestral graph markov models. *The Annals of Statistics*, 30(4):962–1030, 2002.
[14] T. S. Richardson, R. J. Evans, J. M. Robins, and I. Shpitser. Nested markov properties for acyclic directed mixed graphs. 2017. arXiv preprint arXiv:1701.06686.
[15] J. M. Robins, M. A. Hernan, and B. Brumback. Marginal structural models and causal inference in epidemiology. *Epidemiology*, 2000.
[16] K. Sachs, O. Perez, D. Pe’er, D. A. Lauffenburger, and G. P. Nolan. Causal protein-signaling networks derived from multiparameter single-cell data. *Science*, 2005.
[17] K. Sadeghi. Graphical representation of independence structures. *Ph.D. thesis, University of Oxford*, 2012.
[18] K. Sadeghi and S. Lauritzen. Markov properties for mixed graphs. *Bernoulli*, 20(2):676–696, 2014.
[19] P. Spirtes, C. N. Glymour, and R. Scheines. Causation, prediction, and search. *MIT Press, Cambridge*, 2001.
[20] J. Tian and J. Pearl. A general identification condition for causal effects. *AAAI/IAAI*, pages 567–573, 2002.
[21] T. Verma and J. Pearl. Equivalence and synthesis of causal models. *Uncertainty in Artificial Intelligence*, 1991.
[22] Y. Wang, L. Solus, K. D. Yang, and C. Uhler. Permutation-based causal inference algorithms with interventions. *Advances in Neural Information Processing Systems*, 2017.
[23] K. D. Yang, A. Katcoff, and C. Uhler. Characterizing and learning equivalence classes of causal dags under interventions. *Proceedings of the 35-th International Conference on Machine Learning*, 2018.
A.1. Graph Theory Definitions. The following is a complete dictionary of the graph theoretical terms used in the manuscript. A mixed graph is a graph \(G = ([p], E)\) in which the set of edges \(E\) contains a mixture of undirected \(i - j\), bidirected \(i \leftrightarrow j\), and directed \(i \rightarrow j\) edges. A mixed graph \(G\) is called simple if there exists at most one edge of any type between each pair of nodes in \(G\). For each edge-type, we call \(i\) and \(j\) the endpoints of the edge, and we say \(i\) and \(j\) are adjacent \(G\) if they are the two endpoints of any single edge in \(G\). We call \(i \in [p]\) a source if all edges with endpoint \(i\) have no arrowhead at \(i\), and we call \(i\) a sink if every edge with endpoint \(i\) has an arrowhead at \(i\). If it is ambiguous as to whether or not there is an arrowhead at the endpoint \(i\) of an edge \(e\), we will place a question mark after \(i\); for example \(i \rightarrow? j\). If \(A \subset [p]\), the induced-subgraph \(G\) on \(A\), denoted \(G(A)\), is the graph with node set \(A\) and all edges in \(E\) which have both endpoints in \(A\). A loopless mixed graph (LMG) is a mixed graph that does not contain any loops; i.e., edges which have both endpoints being the same node. For an LMG \(G = ([p], E)\), and \(i \in [p]\), we say that \(j \in [p]\) is a parent of \(i\) in \(G\) if \(j \rightarrow i \in E\), we say \(j\) is a spouse of \(i\) if \(i \leftrightarrow j \in E\), we say \(j\) is a child of \(i\) if \(i \rightarrow j \in E\), and we let \(pa_G(i), sp_G(i), and de_G(i)\) denote the parents, spouses, and children of \(i\) in \(G\), respectively. A path in \(G\) is a sequence \(\pi = \langle v_1, e_1, v_2, e_2, \ldots, v_m, e_m, v_m \rangle\) where \(v_k \in [p]\) and \(e_k \in E\) for all \(k\) and \(e_k\) has endpoints \(v_k\) and \(v_{k-1}\). We call \(v_1\) and \(v_m\) the endpoints of \(\pi\). If \(A \subset [p]\) such that for each \(a, b \in A\) there exists some path \(\pi\) in \(G\) between \(a\) and \(b\) using only vertices in \(A\), then \(A\) (or the induced subgraph \(G(A)\)) is called path-connected in \(G\). For a path \(\pi\), if \(e \in E\) is an edge with endpoints \(v_1\) and \(v_m\) such that \(e\) has an arrowhead at the endpoint \(v_1\) if and only if \(v_1\) has an arrowhead at \(v_1\), and similarly for \(v_m\), then we call \(e\) an endpoint-identical edge for \(\pi\). Any subsequence of \(\pi\) that is also a path is called a subpath of \(\pi\). A path \(\pi\) is called directed if \(e_k = v_k \rightarrow v_{k+1}\) for all \(k\), and it is called anterior if there exists \(t \in [m]\) such that \(e_k = v_k \rightarrow v_{k+1}\) for all \(k \leq t\) and \(e_k = v_k \rightarrow v_{k+1}\) for all \(k > t\). A cycle is a sequence \(\pi = \langle v_1, e_1, v_2, e_2, \ldots, v_m, e_m, v_m \rangle\) where \(v_k \in [p]\) and \(e_k \in E\) for all \(k\), \(e_k\) has endpoints \(v_k\) and \(v_{k-1}\), and \(v_m = v_1\). It is called directed if \(e_k = v_k \rightarrow v_{k+1}\) for all \(k\). When the edges in a path or cycle \(\pi\) are understood, we will simply write \(\pi\) as the sequence of vertices \(\pi = \langle v_1, v_2, \ldots, v_m \rangle\). A node \(j \in [p]\) is an ancestor of \(i \in [p]\) in \(G\) if there exists a directed path \(\pi\) in \(G\) with \(v_1 = j\) and \(v_m = i\). In this case \(i\) is called a descendant of \(j\) in \(G\). Similarly, \(j\) is anterior to \(i\) if the path \(\pi\) is an anterior path. We let \(ang_G(i), de_G(i), and ant_G(i)\) denote the collection of nodes in \(G\) that are ancestors, descendants, or anterior to \(i\), respectively.

A vertex \(v\) on a path \(\pi = \langle v_1, v_2, \ldots, v_m \rangle\) that is not an endpoint of \(\pi\) is called a collider if two arrowheads point to \(v\) on \(\pi\). A collider path is a path on which every vertex is a collider except for its endpoints. A collider path on three nodes \(\langle v_1, v_2, v_3 \rangle\) is called a \(v\)-structure (or an unshielded collider or an immorality) \([1][5]\) if \(v_1\) and \(v_3\) are not adjacent in \(G\).

When discussing ancestral graphs (AGs), we will also utilize the notion of discriminating paths, whose existence in a pair of MAGs implies that a certain collider path \(i, j, k\) will be the same in both graphs, even though \(i\) and \(k\) are adjacent. A path \(\pi = \langle v_0, v_1, \ldots, v_m, j, k \rangle\) is a discriminating path for \(\langle v_m, j, k \rangle\) in a MAG \(G\) if \(v_0\) is not adjacent to \(k\), and for all \(i \in [m]\), the vertex \(v_i\) is a collider on \(\pi\) and a parent of \(k\). Using discriminating paths, we can also define triples with order. A triple \(\langle i, j, k \rangle\) has order \(0\) if \(i\) and \(k\) are not adjacent. For \(t \geq 0\),

[24] H. Zhao, Z. Zheng, and B. Liu. On the markov equivalence of maximal ancestral graphs. Science in China Series A: Mathematics, 48(4):548–562, 2005.
if we let $\mathcal{D}_t$ denote the set of all triples of order $t$ in a given MAG $\mathcal{G}$, then $\langle i, j, k \rangle \in \mathcal{D}_{t+1}$ if it is not a triple with order $s < t + 1$, and there is a discriminating path $\langle v_0, v_1, \ldots, v_m, b, c \rangle$ with either $\langle i, j, k \rangle = \langle v_m, b, c \rangle$ or $\langle k, j, i \rangle = \langle v_m, b, c \rangle$ such that

$$\langle v_0, v_1, v_2 \rangle, \langle v_1, v_2, v_3 \rangle, \ldots, \langle v_{m-1}, v_m, b \rangle \in \bigcup_{s \leq t} \mathcal{D}_s.$$

**Appendix B. Proofs for Section 2**

B.1. **Proof of Lemma 4.** Suppose first that $i$ and $j$ are adjacent in $\mathcal{G}$ but not in $\mathcal{H}$. Since they are not adjacent in $\mathcal{H}$ then, by [18, Lemma 4], there exists a subset $C$ of $[p] \setminus \{i, j\}$ such that $i$ is $m$-separated from $j$ given $C$ in $\mathcal{H}$. Therefore, since $\mathcal{J}(\mathcal{G}) = \mathcal{J}(\mathcal{H})$ and $\langle i, j \mid C \rangle \in \mathcal{J}(\mathcal{H})$, it follows that $i$ is $m$-separated from $j$ given $C$ in $\mathcal{G}$. However, this contradicts the fact that $i$ and $j$ are adjacent, since an adjacency is always an $m$-connecting path. By symmetry of this argument, we deduce that $\mathcal{G}$ and $\mathcal{H}$ have the same adjacencies.

Suppose now that $i \rightarrow j \leftarrow c$ forms a $v$-structure in $\mathcal{G}$ but not in $\mathcal{H}$. Since $i$ and $k$ are not adjacent, then by maximality of $\mathcal{G}$ we know that there exists a triple $\langle i, k \mid C \rangle \in \mathcal{J}(\mathcal{G})$ for some $C \subset [p] \setminus \{i, j\}$. By Markov equivalence of $\mathcal{G}$ and $\mathcal{H}$, we know that $\langle i, k \mid C \rangle \in \mathcal{J}(\mathcal{H})$ as well. Moreover, by the previous argument, we know that $i$ and $k$ are not adjacent in $\mathcal{H}$. Note now that we must have $j \notin C$, since otherwise $i \rightarrow j \leftarrow c$ would form an $m$-connecting path given $C$ in $\mathcal{G}$. Thus, since $j \notin C$, $i$ and $k$ are not adjacent in $\mathcal{H}$, and the tripath $\langle i, j, k \rangle$ in $\mathcal{H}$ is not a collider path (since it is not a $v$-structure), then the path $\langle i, j, k \rangle$ must be $m$-connecting given $C$ in $\mathcal{H}$. However, this contradicts the fact that $\langle i, j, k \rangle \in \mathcal{J}(\mathcal{H})$. Thus, we conclude that $\mathcal{G}$ and $\mathcal{H}$ must have the same $v$-structures.

**Appendix C. Proofs for Section 3**

C.1. **Proof of Lemma 5.** Suppose that $\mathcal{G}^T$ contains no doubly intervened selection adjacent nodes. If $\mathcal{G}^T$ contains a ribbon then there exists a collider path $\langle i, j, k \rangle$ such that (1) $i$ and $k$ have no endpoint identical edge between them, and (2) $j$ or a descendant of $j$ is an endpoint of a line, or $j$ is in a directed cycle. Since $\mathcal{G}$ is an AG, it contains no directed cycle. So any ribbon $\langle i, j, k \rangle$ in $\mathcal{G}^T$ must have $j$ or a descendant of $j$ as the endpoint of an undirected edge. Since $\mathcal{G}$ is an AG, this could only happen if $i, k \in W_T$, meaning that $j$ must be a doubly-intervened selection adjacent node. Hence, $\mathcal{G}^T$ must be ribbonless.

Conversely, suppose that $\mathcal{G}^T$ is ribbonless. Let $j$ be a selection adjacent node in $\mathcal{G}$. If $\omega_I, \omega_J \in W_T$ are such that $j \in I \cap J$, then the collider path $\langle \omega_I, j, \omega_J \rangle$ would be a ribbon in $\mathcal{G}^T$. Hence, $\mathcal{G}^T$ contains no doubly-intervened selection adjacent nodes.

C.2. **Proof of Lemma 6.** Note first, by Lemma 5 $\overline{\mathcal{G}^T}$ is well-defined since $\mathcal{G}^T$ is ribbonless. We now show the inclusion $\overline{\mathcal{G}^T} \supseteq \overline{\mathcal{G}^T}([p])$. First, recall from [13, Theorem 5.1] that, in an AG $\mathcal{G}$, if two nodes $i$ and $j$ are adjacent in $\overline{\mathcal{G}^T}$ but not in $\mathcal{G}$ then they are connected by a bidirected edge in $\overline{\mathcal{G}^T}$ and an inducing path containing at least two edges in $\mathcal{G}$. Hence, to prove the desired inclusion, we must show that if $i, j \in [p]$ are connected by a bidirected edge in $\overline{\mathcal{G}^T}$ but are not adjacent in $\mathcal{G}^T$, then $i$ and $j$ are connected by a bidirected edge in $\overline{\mathcal{G}^T}$. To this end, note that if $i, j \in [p]$ are adjacent in $\overline{\mathcal{G}^T}$ but not in $\mathcal{G}^T$ then they are connected by an inducing path in $\mathcal{G}^T$ containing at least two edges, say $\pi = \langle v_0 = i, v_1, \ldots, v_m = j \rangle$ where $m \geq 2$. Thus, to prove that $i$ and $j$ are also connected by a bidirected edge in $\overline{\mathcal{G}^T}$, it suffices to show that $\pi$ is an inducing path in $\mathcal{G}$. However, by [13, Lemma 4.5 (iii)], we know that every edge on $\pi$ is bidirected. Hence, since each node in $W_T$ is a source node, we know...
Figure 3. An accepted ground-truth protein signaling network studied in primary human immune system cells [16]. Red and green nodes represent the points of intervention, with red nodes denoting inhibitory interventions and green nodes denoting activating interventions. When all nodes and arrows, black, red, and green, are considered, this is an example of an $\mathcal I$-LMG that is not an $\mathcal I$-DAG. As this is the accepted ground-truth in the biology, researchers need to distinguish it from other Markov equivalent LMGs based on their interventional experiments (see for instance [16, 7]). Since it is not a DAG we cannot apply the graphical characterization of [23]. Hence, a more general theory of $\mathcal I$-Markov equivalence is needed. Developing such a theory is the focus of this paper.
C.3. Proof of Lemma 7. By Lemma 3 it suffices to prove that
\[ i \perp j \mid (\text{ant}(j) \cup W_T) \setminus \{i, j\}. \]
Suppose on the contrary, that there exists an \( m \)-connecting path \( \pi = \langle i, q_1, \ldots, q_m, j \rangle \) given \( C = (\text{ant}(j) \cup W_T) \setminus \{i, j\} \) in \( G^T \). Notice first that since \( i \in W_T \), we know that \( i \rightarrow q_1 \).
Moreover, since all nodes in \( W_T \) are source nodes in \( G^T \) and \( W_T \setminus \{i\} \subseteq C \), then \( q_1, \ldots, q_m \in [p] \); that is, the path \( \langle q_1, \ldots, q_m \rangle \) is contained in \( G \). Since \( \text{ant}(j) \setminus \{j\} \subseteq C \), we know that \( \pi \) is not of the form
\[ i \rightarrow q_1 \rightarrow q_2 \rightarrow \cdots \rightarrow q_{k-1} \rightarrow q_k \rightarrow q_{k+1} \rightarrow \cdots \rightarrow q_m \rightarrow j. \]
Hence, there must be some collider on \( \pi \). Take the collider path \( \langle q_t-1, q_t, q_{t+1} \rangle \) on \( \pi \) with \( t \) maximum. Since \( \pi \) is \( m \)-connecting then \( q_t \in C \cup \text{an}(C) \). Since every node in \( W_T \) is a source node, it must be that \( q_t \in \text{ant}(j) \setminus \{j\} \). Since \( t \) is maximum, then \( q_{t+1}, \ldots, q_m \) are noncolliders on \( \pi \), and hence not in \( \text{ant}(j) \setminus \{j\} \). Hence, the subpath \( \langle q_{t+1}, \ldots, q_m, j \rangle \) of \( \pi \) must be of the form
\[ q_{t+1} \leftarrow q_{t+2} \leftarrow \cdots \leftarrow j. \]
It follows that \( q_t \in \text{an}(j) \setminus \{j\} \). This is because the node \( q_t \) could only be in \( \text{ant}(j) \setminus (\text{an}(j) \cup \{j\}) \) if \( q_{t+1} \in W_T \), but this is impossible since we have already seen that \( q_1, \ldots, q_m, j \in [p] \).
As well, it also follows that the edge between \( q_t \) and \( q_{t+1} \) cannot be bidirected, as this would contradict the maximality of \( t \). Therefore, the edge between \( q_t \) and \( q_{t+1} \) must be of the form \( q_t \leftarrow q_{t+1} \). However, since \( q_t \in \text{an}(j) \setminus \{j\} \), it then follows that there is a directed cycle in \( G \), which contradicts the assumption that \( G \) is ancestral. Thus, we conclude that no such \( m \)-connecting path exists, completing the proof.

C.4. Proof of Proposition 8. By Lemmas 6 and 7 we know that the only possible adjacencies in \( \overline{G^T} \) that may not be in \( \overline{G^T} \) would be between nodes in \( W_T \). By Lemma 5 we know \( G^T \) is ribbonless. Hence, by [17] Algorithm 2.2, we know that \( \overline{G^T} \) is produced from \( G^T \) by adding endpoint identical edges between the endpoints of inducing paths in \( G^T \). Since every node in \( W_T \) is a source node in \( G^T \), we also know there are no arrows pointing to these nodes. Thus, by [17] Proposition 2.6, we know that there is no inducing path between nodes in \( W_T \). Thus, we conclude that \( \overline{G^T} = G^T \).

C.5. Proof of Lemma 9. First assume that \( G \) is maximal; i.e., that \( G \) is a MAG. By Lemma 4 we know that \( G^T \) is ribbonless. By [18] Proposition 1, we know that \( G^T \) is Markov equivalent to \( (G^T)^* \). To see that \( (G^T)^* \) is a MAG, notice by Proposition 8 \( G^T \) is maximal since \( G = \overline{G} \) implies \( G^T = \overline{G^T} = \overline{G}^T \). It follows that \( (G^T)^* \) is also maximal, since removing arrowheads pointing into selection adjacent nodes cannot generate new inducing paths. Moreover, since \( G \) is a MAG, then \( G^T \) contains no directed cycles nor bidirected edges with a directed path between its endpoints. Hence, neither does \( (G^T)^* \). Thus, \( (G^T)^* \) is a MAG.
Conversely, suppose that \( (G^T)^* \) is a MAG that is Markov equivalent to \( G^T \). Since \( G \) is ancestral and \( I \) does not doubly-intervene on any selection adjacent nodes in \( G \), we know that \( G^T \) is produced from \( (G^T)^* \) by replacing undirected edges of the form \( \omega_I \rightarrow j \) for some \( \omega_I \in W_T \) and \( j \in \text{sa}(G) \) a selection adjacent node, with directed arrows \( \omega_I \rightarrow j \). Since \( G \) is ancestral and \( I \) does not doubly-intervene on any selection adjacent nodes, doing so cannot produce any new colliders, and hence \( G^T \) does not containing any inducing paths that are not in \( (G^T)^* \). Since, by Lemma 4 \( G^T \) is ribbonless, it follows from [18] Theorem 2 that \( G^T \)
is maximal; i.e., $G^T = G^T$. So by Proposition 8, $G^T = G^T$. Thus, $G = G$; i.e., $G$ is maximal.

**Appendix D. Proofs for Section 4**

D.1. **Proof of Lemma 13.** Suppose first that there exists an $f^0 \in \mathcal{M}(G)$ such that for all $I \in \mathcal{I}$, $f^I$ factorizes according to equation (5). It is then immediate from Theorem 12 that $f^I \in \mathcal{M}(G)$ for all $I \in \mathcal{I}$. It then remains to check that for all $A \in \mathcal{A}(G)$ that $f^I(x_H | x_{\text{tail}(H)}) = f^J(x_H | x_{\text{tail}(H)})$ for all $H \in [A]_G$ such that $H \cap (I \cup J) = \emptyset$. To see this, note that since $H \cap (I \cup J) = \emptyset$, then $H \cap I = \emptyset$ and $H \cap J = \emptyset$. So by equation (5),

$$f^I(x_H | x_{\text{tail}(H)}) = f^J(x_H | x_{\text{tail}(H)}).$$

Hence, $\{f^I\}_{I \in \mathcal{I}} \in \mathcal{M}_{\mathcal{I}}(G)$.

Conversely, suppose that $\{f^I\}_{I \in \mathcal{I}} \in \mathcal{M}_{\mathcal{I}}(G)$. We need to find $f^0$ such that $f^0 \in \mathcal{M}(G)$ and for all $I \in \mathcal{I}$, $f^I$ factorizes according to equation (5). To this end, for $A \in \mathcal{A}(G)$ and $H \in [A]_G$, pick (if it exists) $I_H \in \mathcal{I}$ such that $I_H \cap H = \emptyset$, and set

$$f^0(x_H | x_{\text{tail}(H)}) := f^{(I_H)}(x_H | x_{\text{tail}(H)}).$$

Note that this choice is well-defined by the second defining condition of $\mathcal{M}_{\mathcal{I}}(G)$ and the fact that this choice is independent of $A$. On the other hand, if no such $I_H$ exists, then set $f^0(x_H | x_{\text{tail}(H)})$ equal to any strictly positive density. Since this operation is well-defined, we get a distribution with density function $f^0$ such that for all $A \in \mathcal{A}(G)$,

$$f^0(x_A) = \prod_{H \in [A]_G} f^0(x_H | x_{\text{tail}(H)}).$$

Hence, by Theorem 12, $f^0 \in \mathcal{M}(G)$. Moreover, for all $I \in \mathcal{I}$ and $A \in \mathcal{A}(G)$,

$$f^I(x_A) = \prod_{H \in [A]_G} f^I(x_H | x_{\text{tail}(H)}),$$

$$= \prod_{H \in [A]_G : I \cap H \neq \emptyset} f^I(x_H | x_{\text{tail}(H)}) \prod_{H \in [A]_G : I \cap H = \emptyset} f^I(x_H | x_{\text{tail}(H)}),$$

$$= \prod_{H \in [A]_G : I \cap H \neq \emptyset} f^I(x_H | x_{\text{tail}(H)}) \prod_{H \in [A]_G : I \cap H = \emptyset} f^0(x_H | x_{\text{tail}(H)}).$$

Hence, $f^I$ factorizes according to equation (5).

D.2. **Proof of Lemma 14.** For statement (1), suppose on the contrary that there exists a district $D \in \text{dis}(G)$ that contains $z \in Z$ and $b' \in B'$. By definition of a district, since $z, b' \in D$, we know that they are path-connected in $G$, by a path consisting of only bidirected arrows. Let $\pi = \langle v_0 := z, v_1, \ldots, v_m := b' \rangle$ denote such a path. Then since $z \in A \setminus (B \cup C)$, we know that $z$ is ancestral to either some $b \in B$ or $c \in C$. If $z$ is ancestral to $b \in B$, then the directed path from $z$ to $b$ must be blocked by some $c \in C$. This is because, by assumption, every $b \in B$ is $m$-separated from $\omega_I$ given $C \cup W_T \setminus \{\omega_I\}$ in $G^T$. Since $z \notin C$ and $z$ is $m$-connected to $\omega_I$ given $C \cup W_T \setminus \{\omega_I\}$ in $G^T$, concatenating such an $m$-connecting path from $\omega_I$ to $z$ with the directed path from $z$ to $b$ would result in an $m$-connecting path from $\omega_I$ to $b$, which would be a contradiction. Hence, $z$ is ancestral to some $c \in C$, and thus there exists some directed path from $z$ to $c$. Therefore, $v_1$ is $m$-connected to $\omega_I$ given $C \cup W_T \setminus \{\omega_I\}$ in $G^T$. In other words, $v_0, v_1 \in Z$. Repeating this argument in an inductive fashion demonstrates
that in fact \( v_0, \ldots, v_m \in Z \). However, this contradicts the assumption that \( b' \in B' \). Thus, no node in \( Z \) can be in the same district as a node in \( B' \).

To see statement (2) holds, we can use a similar inductive argument. Suppose on the contrary that there is a district \( D \in \text{dis}(\mathcal{G}) \) containing a node \( z \in Z \) and a node \( c' \in C' \). Since \( c' \) and \( z \) are in \( D \), a district in \( \mathcal{G} \), then they are path-connected by a path \( \pi = \langle v_0 := z, v_1, \ldots, v_m := c' \rangle \) consisting of only bidirected arrows. By the argument for statement (1), we know that \( \pi \) is an \( m \)-connecting path given \( C \cup W_I \setminus \{ \omega_I, c' \} \). Since \( c' \in C' \), then there exists \( b' \in B' \) and a path \( \pi' = \langle v_0 := z, v_1, \ldots, v_m, b' \rangle \) such that \( v_m \) is a collider on \( \pi' \). Hence, \( \pi' \) is an \( m \)-connecting path given \( C \cup W_I \setminus \{ \omega_I \} \) in \( \mathcal{G}^I \), which contradicts the fact that \( b' \in B' \). Hence, a node of \( Z \) and a node of \( C' \) cannot be in the same district in \( \mathcal{G} \).

### D.3. Proof of Theorem [15]

Suppose first that \( \{ f^{(I)} \}_{I \in \mathcal{I}} \) satisfies the \( \mathcal{I} \)-Markov properties with respect to \( \mathcal{G}^I \). Since \( \mathcal{G} \) is an ADMG, then by Theorem [12] and condition (1) of the \( \mathcal{I} \)-Markov properties, we know that for all \( A \in [A]_{\mathcal{G}} \)

\[
(8) \quad f^{(I)}(x_A) = \prod_{H \in [A]_{\mathcal{G}}} f^{(I)}(x_H \mid x_{\text{tail}(H)})
\]

Consider now \( H \in [A]_{\mathcal{G}} \) for which \( H \cap I = \emptyset \). Note that since \( \mathcal{G}^I \) is also an ADMG, and since \( H \cap I = \emptyset \), then \( \text{tail}_{\mathcal{G}^I}(H) = \text{tail}_{\mathcal{G}}(H) \). Since \( \text{tail}_{\mathcal{G}^I}(H) \) is the Markov blanket of \( H \) in \( \mathcal{G}^I \) [12 Section 3.1], then \( \omega_I \) is \( m \)-separated from \( H \) given \( \text{tail}(H) \cup W_I \setminus \{ \omega_I \} \) in \( \mathcal{G}^I \). So by condition (2) of the \( \mathcal{I} \)-Markov properties, we can replace \( f^{(I)}(x_H \mid x_{\text{tail}(H)}) \) with \( f^{(\emptyset)}(x_H \mid x_{\text{tail}(H)}) \) in the product in equation (8). Since this is independent of the choice of \( I \in \mathcal{I} \), Lemma [13] implies that \( \{ f^{(I)} \}_{I \in \mathcal{I}} \in \mathcal{M}_{\mathcal{I}}(\mathcal{G}) \).

Conversely, suppose that \( \{ f^{(I)} \}_{I \in \mathcal{I}} \in \mathcal{M}_{\mathcal{I}}(\mathcal{G}) \). Notice first that since for all \( I \in \mathcal{I} \), we have \( f^{(I)} \in \mathcal{M}(\mathcal{G}) \), then \( \{ f^{(I)} \}_{I \in \mathcal{I}} \) satisfies condition (1) of the \( \mathcal{I} \)-Markov properties. So it remains to check that \( \{ f^{(I)} \}_{I \in \mathcal{I}} \) satisfies condition (2) of the \( \mathcal{I} \)-Markov properties with respect to \( \mathcal{G}^I \). To this end, assume that \( A, B, B', C, C', C'', Z \subset [p] \) are subsets as specified in Lemma [14] and Remark [4] and set \( \hat{\mathcal{I}} := \{ \emptyset, I \} \). Note, by Lemma [14] and Remark [4] we can partition \( [A]_{\mathcal{G}} \) as \( [A]_{\mathcal{G}} = Q \sqcup Y \sqcup M \sqcup N \), where

- \( Q := \{ H \in [A]_{\mathcal{G}} \mid B' \cap H \neq \emptyset \} \),
- \( Y := \{ H \in [A]_{\mathcal{G}} \mid Z \cap H \neq \emptyset \} \),
- \( M := \{ H \in [A]_{\mathcal{G}} \mid B' \cap H = \emptyset, C' \cap H \neq \emptyset \} \),
- \( N := \{ H \in [A]_{\mathcal{G}} \mid Z \cap H = \emptyset, C'' \cap H \neq \emptyset \} \).
By Lemma 13, we know that for $\hat{I} \in \hat{I}$, the marginal density $f^{(\hat{I})}(x_A)$ factors according to equation (5), and so

$$f^{(\hat{I})}(x_A) = f^{(\hat{I})}(x_{B'}, x_{C'}, x_{C''}, x_Z),$$

$$= \prod_{H \in Q} f^{(\hat{I})}(x_H | x_{\text{tail}(H)}) \prod_{H \in Y} f^{(\hat{I})}(x_H | x_{\text{tail}(H)}) \prod_{H \in M} f^{(\hat{I})}(x_H | x_{\text{tail}(H)}),$$

$$= \prod_{H \in Q} f^{(0)}(x_H | x_{\text{tail}(H)}) \prod_{H \in Y} f^{(0)}(x_H | x_{\text{tail}(H)}) \prod_{H \in M} f^{(\hat{I})}(x_H | x_{\text{tail}(H)}),$$

$$= g(x_{B'}, x_{C'}) h(x_Z, x_{C''}; \hat{I}),$$

where $g(x_{B'}, x_{C'})$ is the product of the conditional factors over $Q$ and $Y$, and $h(x_Z, x_{C''}; \hat{I})$ is the product of the conditional factors over $M$ and $N$. From this point we can now use an analogous argument to that used in the proof of Proposition 3.8 in [23]. Namely, marginalizing out the variables for nodes in $Z$ and $B' \setminus B$ gives

$$f^{(\hat{I})}(x_B, x_C) = \hat{g}(x_B, x_{C'}) \hat{h}(x_{C''}; \hat{I}),$$

where $\hat{g}(x_B, x_{C'}) := \int_{B' \setminus B} g(x_{B'}, x_{C'})$ and $\hat{h}(x_{C''}; \hat{I}) := \int_{Z} h(x_Z, x_{C''}; \hat{I})$. By conditioning on $C$, it is then straightforward to check that $f^{(\hat{I})}(x_B | x_C) = f^{(0)}(x_B | x_C)$ for $\hat{I} \in \hat{I}$. Hence, $\{f^{(\hat{I})}\}_{\hat{I} \in \hat{I}}$ satisfies condition (2) of the $Z$-Markov properties, completing the proof.

Matematik, KTH, SE-100 44 Stockholm, Sweden
E-mail address: solus@kth.se