MARGINAL AMP CHAIN GRAPHS

JOSE M. PEÑA

ADIT, IDA, LINKÖPING UNIVERSITY, SE-58183 LINKÖPING, SWEDEN

JOSE.M.PENA@LIU.SE

ABSTRACT. We present a new family of graphical models that may have undirected, directed and bidirected edges. We name these new models marginal AMP (MAMP) chain graphs because each of them can be seen as the result of marginalizing out some nodes in an AMP chain graph. However, MAMP chain graphs do not only subsume AMP chain graphs but also regression chain graphs. We describe global and local Markov properties for MAMP chain graphs and prove their equivalence for compositional graphoids. We also characterize when two MAMP chain graphs are Markov equivalent.

1. INTRODUCTION

Chain graphs (CGs) are graphs with possibly directed and undirected edges, and no semidirected cycle. They have been extensively studied as a formalism to represent independence models. CGs extend Bayesian networks, i.e. directed graphs with no directed cycle, and Markov networks, i.e. undirected graphs. Therefore, they can model symmetric and asymmetric relationships between the random variables of interest, which is one of the reasons of their popularity. However, unlike Bayesian networks and Markov networks whose interpretation is unique, there are four different interpretations of CGs as independence models (Cox and Wermuth, 1993, 1996; Drton, 2009; Sonntag and Peña, 2013). In this paper, we are interested in the AMP interpretation (Andersson et al., 2001; Levitz et al., 2001) and in the multivariate regression (MVR) interpretation (Cox and Wermuth, 1993, 1996). Although MVR CGs were originally represented using dashed directed and undirected edges, we prefer to represent them using solid directed and bidirected edges.

In this paper, we combine the AMP and MVR interpretations of CGs. Specifically, we introduce a new family of graphical models that may have undirected, directed and bidirected edges. We call this new family marginal AMP (MAMP) CGs. We formally define MAMP CGs in Section 3. In that section, we also describe global and local Markov properties for MAMP CGs and prove their equivalence for compositional graphoids. In Section 4, we characterize when two MAMP CGs are Markov equivalent. Finally, we discuss in Section 5 how MAMP CGs relate to other existing graphical models such as regression CGs, maximal ancestral graphs, summary graphs and MC graphs.

2. PRELIMINARIES

In this section, we review some concepts from graphical models that are used later in this paper. All the graphs and probability distributions in this paper are defined over a finite set V. All the graphs in this paper are simple, i.e. they contain at most one edge between any pair of nodes. The elements of V are not distinguished from singletons. The operators set union and set difference are given equal precedence in the expressions.

If a graph G contains an undirected, directed or bidirected edge between two nodes V_1 and V_2, then we write that $V_1 - V_2$, $V_1 \to V_2$ or $V_1 \leftrightarrow V_2$ is in G. We represent with a circle, such

as in $\circ\!\!\rightarrow$ or $\circ\!\!-\!\!\circ$, that the end of an edge is unspecified, i.e. it may be an arrow tip or nothing. The parents of a set of nodes X of G is the set $pa_G(X) = \{V_1 | V_1 \rightarrow V_2 \text{ is in } G, V_1 \notin X \text{ and } V_2 \in X\}$. The children of X is the set $ch_G(X) = \{V_1 | V_1 \leftarrow V_2 \text{ is in } G, V_1 \notin X \text{ and } V_2 \in X\}$. The neighbors of X is the set $ne_G(X) = \{V_1 | V_1 - V_2 \text{ is in } G, V_1 \notin X \text{ and } V_2 \in X\}$. The spouses of X is the set $sp_G(X) = \{V_1 | V_1 \leftrightarrow V_2 \text{ is in } G, V_1 \notin X \text{ and } V_2 \in X\}$. The adjacents of X is the set $ad_G(X) = ne_G(X) \cup pa_G(X) \cup ch_G(X) \cup sp_G(X)$. A route between a node V_1 and a node V_n in G is a sequence of (not necessarily distinct) nodes $V_1, \ldots, V_n$ st $V_i \in ad_G(V_{i+1})$ for all $1 \le i < n$. If the nodes in the route are all distinct, then the route is called a path. The length of a route is the number of (not necessarily distinct) edges in the route, e.g. the length of the route $V_1, \ldots, V_n$ is $n-1$. A route is called descending if $V_i \rightarrow V_{i+1}$ or $V_i - V_{i+1}$ is in G for all $1 \le i < n$. A route is called strictly descending if $V_i \rightarrow V_{i+1}$ is in G for all $1 \le i < n$. The descendants of a set of nodes X of G is the set $de_G(X) = \{V_n | \text{ there is a descending route from } V_1 \text{ to } V_n \text{ in } G, V_1 \in X \text{ and } V_n \notin X\}$. The non-descendants of X is the set $nde_G(X) = V \smallsetminus de_G(X)$. The strict ascendants of X is the set $san_G(X) = \{V_1 | \text{ there is a strictly descending route from } V_1 \text{ to } V_n \text{ in } G, V_1 \notin X \text{ and } V_n \in X\}$. A route $V_1, \ldots, V_n$ in G is called a cycle if $V_n = V_1$. Moreover, it is called a semidirected cycle if $V_1 \rightarrow V_2$ is in G and $V_i \rightarrow V_{i+1}$, $V_i \leftrightarrow V_{i+1}$ or $V_i - V_{i+1}$ is in G for all $1 < i < n$. An AMP chain graph (AMP CG) is a graph whose every edge is directed or undirected st it has no semidirected cycles. A MVR chain graph (MVR CG) is a graph whose every edge is directed or bidirected st it has no semidirected cycles. A set of nodes of a graph is connected if there exists a path in the graph between every pair of nodes in the set st all the edges in the path are undirected or bidirected. A connectivity component of a graph is a connected set that is maximal (wrt set inclusion). The subgraph of G induced by a set of its nodes X, denoted as G_X, is the graph over X that has all and only the edges in G whose both ends are in X.

Let X, Y, Z and W denote four disjoint subsets of V. An independence model M is a set of statements $X \perp_M Y | Z$. Moreover, M is called graphoid if it satisfies the following properties: Symmetry $X \perp_M Y | Z \Rightarrow Y \perp_M X | Z$, decomposition $X \perp_M Y \cup W | Z \Rightarrow X \perp_M Y | Z$, weak union $X \perp_M Y \cup W | Z \Rightarrow X \perp_M Y | Z \cup W$, contraction $X \perp_M Y | Z \cup W \wedge X \perp_M W | Z \Rightarrow X \perp_M Y \cup W | Z$, and intersection $X \perp_p Y | Z \cup W \wedge X \perp_p W | Z \cup Y \Rightarrow X \perp_p Y \cup W | Z$. Moreover, M is called compositional graphoid if it is a graphoid that also satisfies the composition property $X \perp_M Y | Z \wedge X \perp_M W | Z \Rightarrow X \perp_M Y \cup W | Z$.

We now recall the semantics of AMP CGs and MVR CGs. A node B in a path ρ in an AMP CG G is called a triplex node in ρ if $A \rightarrow B \leftarrow C$, $A \rightarrow B - C$, or $A - B \leftarrow C$ is a subpath of ρ. Moreover, ρ is said to be Z-open with $Z \subseteq V$ when

- every triplex node in ρ is in $Z \cup san_G(Z)$, and
- every non-triplex node B in ρ is outside Z, unless $A - B - C$ is a subpath of ρ and $pa_G(B) \smallsetminus Z \ne \varnothing$.

Let X, Y and Z denote three disjoint subsets of V. When there is no Z-open path in G between a node in X and a node in Y, we say that X is separated from Y given Z in G and denote it as $X \perp_G Y | Z$. The independence model induced by G is the set of separation statements $X \perp_G Y | Z$.

A node B in a path ρ in a MVR CG G is called a triplex node in ρ if $A \rightarrow\!\!\!\!\!\!\circ\,\, B \,\circ\!\!\!\!\!\!\leftarrow C$ is a subpath of ρ. Moreover, ρ is said to be Z-open with $Z \subseteq V$ when

- every triplex node in ρ is in $Z \cup san_G(Z)$, and
- every non-triplex node B in ρ is outside Z.

Let X, Y and Z denote three disjoint subsets of V. When there is no Z-open path in G between a node in X and a node in Y, we say that X is separated from Y given Z in G and denote it as $X \perp_G Y | Z$. The independence model induced by G is the set of separation statements $X \perp_G Y | Z$.

3. Definition

A graph G containing possibly directed, bidirected and undirected edges is a marginal AMP (MAMP) CG if

    C1. G has no semidirected cycle,

    C2. G has no cycle $V_1, \ldots, V_n = V_1$ st $V_1 \leftrightarrow V_2$ is in G and $V_i - V_{i+1}$ is in G for all $1 < i < n$, and

    C3. if $V_1 - V_2 - V_3$ is in G and $sp_G(V_2) \neq \varnothing$, then $V_1 - V_3$ is in G too.

A set of nodes of a MAMP CG G is undirectly connected if there exists a path in G between every pair of nodes in the set st all the edges in the path are undirected. An undirected connectivity component of G is an undirectly connected set that is maximal (wrt set inclusion). We denote by $uc_G(A)$ the undirected connectivity component a node A of G belongs to.

The semantics of MAMP CGs is as follows. A node B in a path ρ in a MAMP CG G is called a triplex node in ρ if $A \rightarrowtail B \leftarrowtail C$, $A \rightarrowtail B - C$, or $A - B \leftarrowtail C$ is a subpath of ρ. Moreover, ρ is said to be Z-open with $Z \subseteq V$ when

    • every triplex node in ρ is in $Z \cup san_G(Z)$, and

    • every non-triplex node B in ρ is outside Z, unless $A - B - C$ is a subpath of ρ and $sp_G(B) \neq \varnothing$ or $pa_G(B) \smallsetminus Z \neq \varnothing$.

Let X, Y and Z denote three disjoint subsets of V. When there is no Z-open path in G between a node in X and a node in Y, we say that X is separated from Y given Z in G and denote it as $X \perp_G Y | Z$. We denote by $X \not\perp_G Y | Z$ that $X \perp_G Y | Z$ does not hold. Likewise, we denote by $X \perp_p Y | Z$ (respectively $X \not\perp_p Y | Z$) that X is independent (respectively dependent) of Y given Z in a probability distribution p. The independence model induced by G, denoted as $I(G)$, is the set of separation statements $X \perp_G Y | Z$. We say that p is Markovian wrt G when $X \perp_p Y | Z$ if $X \perp_G Y | Z$ for all X, Y and Z disjoint subsets of V. Moreover, we say that p is faithful to G when $X \perp_p Y | Z$ iff $X \perp_G Y | Z$ for all X, Y and Z disjoint subsets of V.

Note that if a MAMP CG G has a path $V_1 - V_2 - \ldots - V_n$ st $sp_G(V_i) \neq \varnothing$ for all $1 < i < n$, then $V_1 - V_n$ must be in G. Therefore, the independence model induced by a MAMP CG is the same whether we use the definition of Z-open path above or the following simpler one. A path ρ in a MAMP CG G is said to be Z-open when

    • every triplex node in ρ is in $Z \cup san_G(Z)$, and

    • every non-triplex node B in ρ is outside Z, unless $A - B - C$ is a subpath of ρ and $pa_G(B) \smallsetminus Z \neq \varnothing$.

The motivation behind the three constraints in the definition of MAMP CGs is as follows. The constraint C1 follows from the semidirected acyclicity constraint of AMP CGs and MVR CGs. For the constraints C2 and C3, note that typically every missing edge in a graphical model corresponds to a separation. However, this may not be true for graphs that do not satisfy the constraints C2 and C3. For instance, the graph G below does not contain any edge between B and D but $B \not\perp_G D | Z$ for all $Z \subseteq V \smallsetminus \{B, D\}$. Likewise, G does not contain any edge between A and E but $A \not\perp_G E | Z$ for all $Z \subseteq V \smallsetminus \{A, E\}$.

$$A \leftharpoondown B \leftharpoondown C \rightharpoondown D \rightharpoondown E$$
$$\updownarrow$$
$$F$$
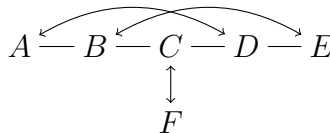
Since the situation above is counterintuitive, we enforce the constraints C2 and C3. Theorem 2 below shows that every missing edge in a MAMP CG corresponds to a separation.

Given a MAMP CG G, let $\widehat{G}$ denote the AMP CG obtained by replacing every bidirected edge $A \leftrightarrow B$ in G with $A \leftarrow L_{AB} \rightarrow B$. Note that G and $\widehat{G}$ represent the same separations

over V. Therefore, every MAMP CG can be seen as the result of marginalizing out some nodes in an AMP CG, hence the name.

Note that AMP CGs and MVR CGs are special cases of MAMP CGs. However, MAMP CGs are a proper generalization of AMP CGs and MVR CGs, as there are independence models that can be induced by the former but not by the latter. An example follows (we postpone the proof that it cannot be induced by AMP CGs and MVR CGs until after Theorem 4).

$$A \rightarrow B \longrightarrow C$$
$$\big| \qquad \updownarrow$$
$$D \leftrightarrow E$$

The theorem below shows that the independence models induced by MAMP CGs are not arbitrary in the probabilistic framework.

Theorem 1. *For any MAMP CG G, there exists a regular Gaussian probability distribution p that is faithful to G.*

Proof. It suffices to replace every bidirected edge $A \leftrightarrow B$ in G with $A \leftarrow L_{AB} \rightarrow B$ to create an AMP CG $\widehat{G}$, apply Theorem 6.1 by Levitz et al. (2001) to conclude that there exists a regular Gaussian probability distribution q that is faithful to $\widehat{G}$, and then let p be the marginal probability distribution of q over V. $\qquad \square$

Corollary 1. *Any independence model induced by a MAMP CG is a compositional graphoid.*

Proof. It follows from Theorem 1 by just noting that the set of independencies in any regular Gaussian probability distribution satisfies the compositional graphoid properties (Studený, 2005, Sections 2.2.2, 2.3.5 and 2.3.6). $\qquad \square$

Finally, we show below that the independence model induced by a MAMP CG coincides with certain closure of certain separations. We define the local separation base of a MAMP CG G as the separations

- $A \perp B | pa_G(A)$ for all non-adjacent nodes A and B of G st $B \notin de_G(A)$, and
- $A \perp B | ne_G(A) \cup pa_G(A \cup ne_G(A))$ for all non-adjacent nodes A and B of G st $A \in de_G(B)$ and $B \in de_G(A)$, i.e. $uc_G(A) = uc_G(B)$.

We define the compositional graphoid closure of the local separation base of G, denoted as $cl(G)$, as the set of separations that are in the base plus those that can be derived from it by applying the compositional graphoid properties. We denote the separations in $cl(G)$ as $X \perp_{cl(G)} Y | Z$.

Theorem 2. *For any MAMP CG G, if $X \perp_{cl(G)} Y | Z$ then $X \perp_G Y | Z$.*

Proof. Since the independence model induced by G is a compositional graphoid by Corollary 1, it suffices to prove that the local separation base of G is a subset of the independence model induced by G. We prove this next. Let A and B be two non-adjacent nodes of G. Consider the following two cases.

    Case 1: $B \notin de_G(A)$. Then, every path between A and B in G falls within one of the following cases.

        Case 1.1: $A = V_1 \leftarrow V_2 \ldots V_n = B$. Then, this path is not $pa_G(A)$-open.

        Case 1.2: $A = V_1 \leftrightarrow V_2 \ldots V_n = B$. Note that $V_2 \neq V_n$ because, by assumption, A and B are non-adjacent in G. Note also that $V_2 \notin pa_G(A)$ due to the constraint C1. Then, $V_2 \rightarrow V_3$ must be in G for the path to be $pa_G(A)$-open. By repeating this reasoning, we can conclude that $A = V_1 \leftrightarrow V_2 \rightarrow V_3 \rightarrow \ldots \rightarrow V_n = B$ is in G. However, this contradicts the assumption that $B \notin de_G(A)$.

Case 1.3: $A = V_1 - V_2 - \ldots - V_m \leftarrow V_{m+1} \ldots V_n = B$. Note that $V_m \notin pa_G(A)$ due to the constraint C1. Then, this path is not $pa_G(A)$-open.

Case 1.4: $A = V_1 - V_2 - \ldots - V_m \rightarrow V_{m+1} \ldots V_n = B$. Note that $V_{m+1} \neq V_n$ because, by assumption, $B \notin de_G(A)$. Note also that $V_{m+1} \notin pa_G(A)$ due to the constraint C1. Then, $V_{m+1} \rightarrow V_{m+2}$ must be in G for the path to be $pa_G(A)$-open. By repeating this reasoning, we can conclude that $A = V_1 - V_2 - \ldots - V_m \rightarrow V_{m+1} \rightarrow \ldots \rightarrow V_n = B$ is in G. However, this contradicts the assumption that $B \notin de_G(A)$.

Case 1.5: $A = V_1 - V_2 - \ldots - V_m \leftrightarrow V_{m+1} \ldots V_n = B$. Note that $V_m \notin pa_G(A)$ due to the constraint C1. Then, this path is not $pa_G(A)$-open.

Case 1.6: $A = V_1 - V_2 - \ldots - V_n = B$. This case contradicts the assumption that $B \notin de_G(A)$.

Case 2: $A \in de_G(B)$ and $B \in de_G(A)$, i.e. $uc_G(A) = uc_G(B)$. Then, there is an undirected path ρ between A and B in G. Then, every path between A and B in G falls within one of the following cases.

Case 2.1: $A = V_1 \leftarrow V_2 \ldots V_n = B$. Then, this path is not $(ne_G(A) \cup pa_G(A \cup ne_G(A)))$-open.

Case 2.2: $A = V_1 \leftrightarrow V_2 \ldots V_n = B$. Note that $V_2 \neq V_n$ due to ρ and the constraints C1 and C2. Note also that $V_2 \notin ne_G(A) \cup pa_G(A \cup ne_G(A))$ due to the constraint C1. Then, $V_2 \rightarrow V_3$ must be in G for the path to be $(ne_G(A) \cup pa_G(A \cup ne_G(A)))$-open. By repeating this reasoning, we can conclude that $A = V_1 \leftrightarrow V_2 \rightarrow V_3 \rightarrow \ldots \rightarrow V_n = B$ is in G. However, this together with ρ violate the constraint C1.

Case 2.3: $A = V_1 - V_2 \leftarrow V_3 \ldots V_n = B$. Then, this path is not $(ne_G(A) \cup pa_G(A \cup ne_G(A)))$-open.

Case 2.4: $A = V_1 - V_2 \leftrightarrow V_3 \ldots V_n = B$. Note that $V_3 \neq V_n$ due to ρ and the constraints C1 and C2. Note also that $V_3 \notin ne_G(A) \cup pa_G(A \cup ne_G(A))$ due to the constraints C1 and C2. Then, $V_3 \rightarrow V_4$ must be in G for the path to be $(ne_G(A) \cup pa_G(A \cup ne_G(A)))$-open. By repeating this reasoning, we can conclude that $A = V_1 - V_2 \leftrightarrow V_3 \rightarrow \ldots \rightarrow V_n = B$ is in G. However, this together with ρ violate the constraint C1.

Case 2.5: $A = V_1 - V_2 - V_3 \ldots V_n = B$ st $sp_G(V_2) = \varnothing$. Then, this path is not $(ne_G(A) \cup pa_G(A \cup ne_G(A)))$-open.

Case 2.6: $A = V_1 - V_2 - \ldots - V_m - V_{m+1} \leftarrow V_{m+2} \ldots V_n = B$ st $sp_G(V_i) \neq \varnothing$ for all $2 \leq i \leq m$. Note that $V_i \in ne_G(V_1)$ for all $3 \leq i \leq m+1$ by the constraint C3. Then, this path is not $(ne_G(A) \cup pa_G(A \cup ne_G(A)))$-open.

Case 2.7: $A = V_1 - V_2 - \ldots - V_m - V_{m+1} \leftrightarrow V_{m+2} \ldots V_n = B$ st $sp_G(V_i) \neq \varnothing$ for all $2 \leq i \leq m$. Note that $V_{m+2} \neq V_n$ due to ρ and the constraints C1 and C2. Note also that $V_{m+2} \notin ne_G(A) \cup pa_G(A \cup ne_G(A))$ due to the constraints C1 and C2. Then, $V_{m+2} \rightarrow V_{m+3}$ must be in G for the path to be $(ne_G(A) \cup pa_G(A \cup ne_G(A)))$-open. By repeating this reasoning, we can conclude that $A = V_1 - V_2 - \ldots - V_m - V_{m+1} \leftrightarrow V_{m+2} \rightarrow \ldots \rightarrow V_n = B$ is in G. However, this together with ρ violate the constraint C1.

Case 2.8: $A = V_1 - V_2 - \ldots - V_m - V_{m+1} - V_{m+2} \ldots V_n = B$ st $sp_G(V_i) \neq \varnothing$ for all $2 \leq i \leq m$ and $sp_G(V_{m+1}) = \varnothing$. Note that $V_i \in ne_G(V_1)$ for all $3 \leq i \leq m+1$ by the constraint C3. Then, this path is not $(ne_G(A) \cup pa_G(A \cup ne_G(A)))$-open.

Case 2.9: $A = V_1 - V_2 - \ldots - V_n = B$ st $sp_G(V_i) \neq \varnothing$ for all $2 \leq i \leq n-1$. Note that $V_i \in ne_G(V_1)$ for all $3 \leq i \leq n$ by the constraint C3. However, this contradicts the assumption that A and B are non-adjacent in G.

$\square$

Theorem 3. *For any MAMP CG G, if $X \perp_G Y | Z$ then $X \perp_{cl(G)} Y | Z$.*

Proof. We start by recalling some definitions from Andersson et al. (2001, Section 2). Let F be an AMP CG and F' the result of removing all the directed edges from F. Given a set $U \subseteq V$, let $W = U \cup san_G(U)$ and $W' = \bigcup_{A \in W} uc_G(A)$. Let $F[W]$ denote the graph whose nodes and edges are the union of the nodes and edges in F_W and $F'_{W'}$. $F[W]$ is called an extended subgraph of F. An undirected graph is called complete if it has an edge between any pair of nodes. In an AMP CG, a triplex $(\{A, C\}, B)$ is an induced subgraph of the form $A \to B \leftarrow C$, $A \to B - C$, or $A - B \leftarrow C$. Augmenting a triplex $(\{A, C\}, B)$ means replacing it with the complete undirected graph over $\{A, B, C\}$. In an AMP CG G, a 2-biflag $(\{A, D\}, \{B, C\})$ is a subgraph of the form $A \to B - C \leftarrow D$ st A is not adjacent to C in G and B is not adjacent to D in G. Augmenting a 2-biflag $(\{A, D\}, \{B, C\})$ means replacing it with the complete undirected graph over $\{A, B, C, D\}$. Augmenting an AMP CG F, denoted as F^a, means augmenting all its triplexes and 2-biflags and converting the remaining directed edges into undirected edges. Note that $X \perp_F Y | Z$ iff X is separated from Y given Z in $F[X \cup Y \cup Z]^a$ (Levitz et al., 2001, Theorem 4.1).

Given an undirected graph F and a set $U \subseteq V$, let F^U denote the undirected graph over U resulting from adding an edge $A - B$ to F_U if F has a path between A and B whose only nodes in U are A and B. F^U is sometimes called the marginal graph of F for U.

Now, we start the proof per se. Let $\widehat{G}$ denote the AMP CG obtained by replacing every bidirected edge $A \leftrightarrow B$ in G with $A \leftarrow L_{AB} \to B$. The node L_{AB} is called latent. Let $\overline{G} = (\widehat{G}[X \cup Y \cup Z]^a)^V$. As mentioned before, G and $\widehat{G}$ represent the same separations over V. Then, $X \perp_G Y | Z$ iff X is separated from Y given Z in $\overline{G}$. Note that the separations in $\overline{G}$ coincide with the graphoid closure of the separations $A \perp_{\overline{G}} V(\overline{G}) \smallsetminus A \smallsetminus ad_{\overline{G}}(A) | ad_{\overline{G}}(A)$ for all $A \in V(\overline{G})$, where $V(\overline{G})$ denotes the nodes in $\overline{G}$ (Bouckaert, 1995, Theorem 3.4). Therefore, to prove that $X \perp_{cl(G)} Y | Z$, it suffices to prove that $A \perp_{cl(G)} V(\overline{G}) \smallsetminus A \smallsetminus ad_{\overline{G}}(A) | ad_{\overline{G}}(A)$ for all $A \in V(\overline{G})$. Let $K_1, \dots, K_n$ denote the connectivity components of $G_{V(\overline{G})}$ st if $A \to B$ is in $G_{V(\overline{G})}$, then $A \in K_i$ and $B \in K_j$ with $i < j$. Consider the following cases.

Case 1: Assume that $A \in K_n$. Note that if $D \in ad_{\overline{G}}(A)$, then $D \in ne_G(A) \cup pa_G(A \cup ne_G(A)) \cup A' \cup ne_G(A') \cup pa_G(A' \cup ne_G(A'))$ for some $A' \in V$ st G has a path $A \dots \leftrightarrow A'$ whose every node is in K_n. To see it, note that if $D \in ad_{\overline{G}}(A) \smallsetminus ne_G(A) \smallsetminus pa_G(A \cup ne_G(A))$, then $\widehat{G}[X \cup Y \cup Z]^a$ must have a path between A and D whose every node except A and D is latent. Then, $\widehat{G}[X \cup Y \cup Z]$ must have a path of the form $A \dots L \to A' = D$, $A \dots L \to A' - D$, $A \dots L \to A' \leftarrow D$ or $A \dots L \to A' - A'' \leftarrow D$, where L is a latent node and every non-latent node between A and A' is in K_n. Note also that $A' \in de_G(A)$ iff $A \in de_G(A')$, because $A, A' \in K_n$. Let $\mathcal{B}$ contain all those A' st $A' \in de_G(A)$, i.e. $uc_G(A) = uc_G(A')$. Let $\mathcal{C}$ contain all those A' st $A' \notin de_G(A)$.

Consider any $A' \in ne_G(A)$. Note that $de_G(A) = de_G(A')$. Then, $pa_G(A \cup ne_G(A)) \subseteq nde_G(A) = nde_G(A')$. Therefore,

(1) $A \perp_{cl(G)} nde_G(A) \smallsetminus pa_G(A) | pa_G(A)$ and

(2) $A' \perp_{cl(G)} nde_G(A') \smallsetminus pa_G(A') | pa_G(A')$ follow from the local separation base of G by repeated composition, which imply

(3) $A \perp_{cl(G)} nde_G(A) \smallsetminus pa_G(A \cup ne_G(A)) | pa_G(A \cup ne_G(A))$ and

(4) $A' \perp_{cl(G)} nde_G(A') \smallsetminus pa_G(A \cup ne_G(A)) | pa_G(A \cup ne_G(A))$ by weak union. Therefore,

(5) $A \cup ne_G(A) \perp_{cl(G)} nde_G(A) \smallsetminus pa_G(A \cup ne_G(A)) | pa_G(A \cup ne_G(A))$ by repeated symmetry and composition, which implies

(6) $A \perp_{cl(G)} nde_G(A) \smallsetminus pa_G(A \cup ne_G(A)) | ne_G(A) \cup pa_G(A \cup ne_G(A))$ by symmetry and weak union. Note that

(7) $A \perp_{cl(G)} uc_G(A) \smallsetminus A \smallsetminus ne_G(A) | ne_G(A) \cup pa_G(A \cup ne_G(A))$ follows from the local separation base of G by repeated composition, which implies

(8) $A \perp_{cl(G)} [nde_G(A) \smallsetminus pa_G(A \cup ne_G(A))] \cup [uc_G(A) \smallsetminus A \smallsetminus ne_G(A)] | ne_G(A) \cup pa_G(A \cup ne_G(A))$ by symmetry and composition on (6) and (7).

Consider any $B \in \mathcal{B} \smallsetminus ne_G(A)$. By repeating the reasoning above, we can conclude

(9) $B \perp_{cl(G)} [nde_G(B) \smallsetminus pa_G(B \cup ne_G(B))] \cup [uc_G(B) \smallsetminus B \smallsetminus ne_G(B)] | ne_G(B) \cup pa_G(B \cup ne_G(B))$.

Recall that $de_G(A) = de_G(B)$ and $uc_G(A) = uc_G(B)$. Then, $pa_G(A \cup \mathcal{B} \cup ne_G(A \cup \mathcal{B})) \subseteq nde_G(A) = nde_G(B)$, and $ne_G(A \cup \mathcal{B}) \subseteq uc_G(A) = uc_G(B)$. Therefore,

(10) $A \perp_{cl(G)} [nde_G(A) \smallsetminus pa_G(A \cup \mathcal{B} \cup ne_G(A \cup \mathcal{B}))] \cup [uc_G(A) \smallsetminus A \smallsetminus \mathcal{B} \smallsetminus ne_G(A \cup \mathcal{B})] | ne_G(A \cup \mathcal{B}) \cup pa_G(A \cup \mathcal{B} \cup ne_G(A \cup \mathcal{B}))$ and

(11) $B \perp_{cl(G)} [nde_G(B) \smallsetminus pa_G(A \cup \mathcal{B} \cup ne_G(A \cup \mathcal{B}))] \cup [uc_G(B) \smallsetminus A \smallsetminus \mathcal{B} \smallsetminus ne_G(A \cup \mathcal{B})] | ne_G(A \cup \mathcal{B}) \cup pa_G(A \cup \mathcal{B} \cup ne_G(A \cup \mathcal{B}))$ by weak union and decomposition on (8) and (9). Therefore,

(12) $A \cup \mathcal{B} \perp_{cl(G)} [nde_G(A) \smallsetminus pa_G(A \cup \mathcal{B} \cup ne_G(A \cup \mathcal{B}))] \cup [uc_G(A) \smallsetminus A \smallsetminus \mathcal{B} \smallsetminus ne_G(A \cup \mathcal{B})] | ne_G(A \cup \mathcal{B}) \cup pa_G(A \cup \mathcal{B} \cup ne_G(A \cup \mathcal{B}))$ by repeated symmetry and composition, which implies

(13) $A \perp_{cl(G)} [nde_G(A) \smallsetminus pa_G(A \cup \mathcal{B} \cup ne_G(A \cup \mathcal{B}))] \cup [uc_G(A) \smallsetminus A \smallsetminus \mathcal{B} \smallsetminus ne_G(A \cup \mathcal{B})] | \mathcal{B} \cup ne_G(A \cup \mathcal{B}) \cup pa_G(A \cup \mathcal{B} \cup ne_G(A \cup \mathcal{B}))$ by symmetry and weak union.

Finally, note that $\mathcal{C} \cup ne_G(\mathcal{C}) \cup pa_G(\mathcal{C} \cup ne_G(\mathcal{C})) \subseteq nde_G(A)$. Therefore,

(14) $A \perp_{cl(G)} [nde_G(A) \smallsetminus \mathcal{C} \smallsetminus ne_G(\mathcal{C}) \smallsetminus pa_G(A \cup \mathcal{B} \cup \mathcal{C} \cup ne_G(A \cup \mathcal{B} \cup \mathcal{C}))] \cup [uc_G(A) \smallsetminus A \smallsetminus \mathcal{B} \smallsetminus ne_G(A \cup \mathcal{B})] | \mathcal{B} \cup \mathcal{C} \cup ne_G(A \cup \mathcal{B} \cup \mathcal{C}) \cup pa_G(A \cup \mathcal{B} \cup \mathcal{C} \cup ne_G(A \cup \mathcal{B} \cup \mathcal{C}))$ by weak union on (13), which implies

(15) $A \perp_{cl(G)} V(\overline{G}) \smallsetminus A \smallsetminus ad_{\overline{G}}(A) | ad_{\overline{G}}(A)$ by decomposition.

Case 2: Assume that $A \in K_{n-1}$. Let $\overline{H} = (\widehat{G}[X \cup Y \cup Z \smallsetminus K_n]^a)^V$. Note that $\overline{H}$ is a subgraph of $\overline{G}$ and, thus, $ad_{\overline{H}}(A) \subseteq ad_{\overline{G}}(A)$. Let $\mathcal{B} = ad_{\overline{G}}(A) \cap K_n$ and $\mathcal{C} = K_n \smallsetminus \mathcal{B}$. Note that $\mathcal{B}$ or $\mathcal{C}$ may be empty. Note also that $pa_G(\mathcal{B}) \subseteq ad_{\overline{G}}(A) \cup A$. To see it, note that this is evident for any $D \in ch_G(A) \cup ne_G(ch_G(A))$. On the other hand, if $D \in \mathcal{B} \smallsetminus ch_G(A) \smallsetminus ne_G(ch_G(A))$ then $\widehat{G}[X \cup Y \cup Z]^a$ must have a path between A and D whose every node except A and D is latent. Then, $\widehat{G}[X \cup Y \cup Z]$ must have a path of the form $A \ldots L \to D$ or $A \ldots L \to A' - D$, where L is a latent node and A' is a non-latent node. Then, clearly $pa_G(D) \subseteq ad_{\overline{G}}(A) \cup A$.

Consider any $B \in \mathcal{B}$. Then,

(1) $B \perp_{cl(G)} V(\overline{H}) \smallsetminus pa_G(B) | pa_G(B)$ follows from the local separation base of G by repeated composition, which implies

(2) $B \perp_{cl(G)} V(\overline{H}) \smallsetminus pa_G(\mathcal{B}) | pa_G(\mathcal{B})$ by weak union, which implies

(3) $\mathcal{B} \perp_{cl(G)} V(\overline{H}) \smallsetminus pa_G(\mathcal{B}) | pa_G(\mathcal{B})$ by repeated symmetry and composition, which implies

(4) $\mathcal{B} \perp_{cl(G)} V(\overline{H}) \smallsetminus A \smallsetminus ad_{\overline{G}}(A) | ad_{\overline{G}}(A) \smallsetminus \mathcal{B} \cup A$ by weak union and the fact, shown above, that $pa_G(\mathcal{B}) \subseteq ad_{\overline{G}}(A) \cup A$.

Note that $\overline{H}$ is a proper marginal augmented extended subgraph. Then,

(5) $A \perp_{cl(G)} V(\overline{H}) \smallsetminus A \smallsetminus ad_{\overline{H}}(A) | ad_{\overline{H}}(A)$ follows from repeating Case 1 for $\overline{H}$, which implies

(6) $A \perp_{cl(G)} V(\overline{H}) \smallsetminus A \smallsetminus ad_{\overline{G}}(A) | ad_{\overline{G}}(A) \smallsetminus \mathcal{B}$ by weak union, which implies

(7) $A \cup \mathcal{B} \perp_{cl(G)} V(\overline{H}) \smallsetminus A \smallsetminus ad_{\overline{G}}(A) | ad_{\overline{G}}(A) \smallsetminus \mathcal{B}$ by symmetry and contraction on (4) and (6), which implies

(8) $A \perp_{cl(G)} V(\overline{H}) \smallsetminus A \smallsetminus ad_{\overline{G}}(A) | ad_{\overline{G}}(A)$ by symmetry and weak union.

Finally, consider any $C \in \mathcal{C}$. Then,

(9) $C \perp_{cl(G)} V(\overline{G}) \smallsetminus C \smallsetminus ad_{\overline{G}}(C) | ad_{\overline{G}}(C)$ by Case 1, which implies

(10) $C \perp_{cl(G)} A | V(\overline{G}) \smallsetminus A \smallsetminus C$ by weak union, which implies

(11) $\mathcal{C} \perp_{cl(G)} A | V(\overline{G}) \smallsetminus A \smallsetminus \mathcal{C}$ by repeated symmetry and intersection, which implies

(12) $\mathcal{C} \perp_{cl(G)} A | V(\overline{H}) \smallsetminus A \cup \mathcal{B}$, which implies

(13) $\mathcal{C} \cup [V(\overline{H}) \smallsetminus A \smallsetminus ad_{\overline{G}}(A)] \perp_{cl(G)} A | ad_{\overline{G}}(A)$ by symmetry and contraction on (8) and (12), which implies

(14) $A \perp_{cl(G)} V(\overline{G}) \smallsetminus A \smallsetminus ad_{\overline{G}}(A) | ad_{\overline{G}}(A)$ by symmetry.

Case 3: Assume that $A \in K_i$ with $1 \le i \le n-2$. Then, repeat Case 2 replacing K_{n-1} with K_i, and letting $\mathcal{B} = ad_{\overline{G}}(A) \cap K_{i+1}$ and $\mathcal{C} = K_{i+1} \cup \ldots \cup K_n \smallsetminus \mathcal{B}$.

$\square$

4. Markov Equivalence

We say that two MAMP CGs are Markov equivalent if they induce the same independence model. In a MAMP CG, a triplex $(\{A,C\}, B)$ is an induced subgraph of the form $A \rightarrowtail B \leftarrow\!\!\circ\, C$, $A \rightarrowtail B - C$, or $A - B \leftarrow\!\!\circ\, C$. We say that two MAMP CGs are triplex equivalent if they have the same adjacencies and the same triplexes.

Theorem 4. *Two MAMP CGs are Markov equivalent iff they are triplex equivalent.*

Proof. We first prove the "only if" part. Let G_1 and G_2 be two Markov equivalent MAMP CGs. First, assume that two nodes A and C are adjacent in G_2 but not in G_1. If A and C are in the same undirected connectivity component of G_1, then $A \perp C | ne_{G_1}(A) \cup pa_{G_1}(A \cup ne_{G_1}(A))$ holds for G_1 by Theorem 2 but it does not hold for G_2, which is a contradiction. On the other hand, if A and C are in different undirected connectivity components of G_1, then $A \notin de_{G_1}(C)$ or $C \notin de_{G_1}(A)$. Assume without loss of generality that $A \notin de_{G_1}(C)$. Then, $A \perp C | pa_{G_1}(C)$ holds for G_1 by Theorem 2 but it does not hold for G_2, which is a contradiction. Consequently, G_1 and G_2 must have the same adjacencies.

Finally, assume that G_1 and G_2 have the same adjacencies but G_1 has a triplex $(\{A,C\}, B)$ that G_2 does not have. If A and C are in the same undirected connectivity component of G_1, then $A \perp C | ne_{G_1}(A) \cup pa_{G_1}(A \cup ne_{G_1}(A))$ holds for G_1 by Theorem 2. Note also that $B \notin ne_{G_1}(A) \cup pa_{G_1}(A \cup ne_{G_1}(A))$ because, otherwise, G_1 would not satisfy the constraint C1 or C2. Then, $A \perp C | ne_{G_1}(A) \cup pa_{G_1}(A \cup ne_{G_1}(A))$ does not hold for G_2, which is a contradiction. On the other hand, if A and C are in different undirected connectivity components of G_1, then $A \notin de_{G_1}(C)$ or $C \notin de_{G_1}(A)$. Assume without loss of generality that $A \notin de_{G_1}(C)$. Then, $A \perp C | pa_{G_1}(C)$ holds for G_1 by Theorem 2. Note also that $B \notin pa_{G_1}(C)$ because, otherwise, G_1 would not have the triplex $(\{A,C\}, B)$. Then, $A \perp C | pa_{G_1}(C)$ does not hold for G_2, which is a contradiction. Consequently, G_1 and G_2 must be triplex equivalent.

We now prove the "if" part. Let G_1 and G_2 be two triplex equivalent MAMP CGs. We just prove that all the non-separations in G_1 are also in G_2. The opposite result can be proven in the same manner by just exchanging the roles of G_1 and G_2 in the proof. Specifically, assume that $\alpha \perp \beta | Z$ does not hold for G_1. We prove that $\alpha \perp \beta | Z$ does not hold for G_2 either. We divide the proof in three parts.

Part 1

We say that a path has a triplex $(\{A,C\}, B)$ if it has a subpath of the form $A \rightarrowtail B \leftarrow\!\!\circ\, C$, $A \rightarrowtail B - C$, or $A - B \leftarrow\!\!\circ\, C$. Let ρ_1 be any path between α and β in G_1 that is Z-open st (i) no subpath of ρ_1 between α and β in G_1 is Z-open, (ii) every triplex node in ρ_1 is in Z, and (iii) ρ_1 has no non-triplex node in Z. Let ρ_2 be the path in G_2 that consists of the same nodes as ρ_1. Then, ρ_2 is Z-open. To see it, assume the contrary. Then, one of the following cases must occur.

Case 1: ρ_2 does not have a triplex $(\{A,C\}, B)$ and $B \in Z$. Then, ρ_1 must have a triplex $(\{A,C\}, B)$ because it is Z-open. Then, A and C must be adjacent in G_1 and G_2 because these are triplex equivalent. Let ϱ_1 be the path obtained from ρ_1 by replacing the triplex $(\{A,C\}, B)$ with the edge between A and C in G_1. Note that ϱ_1 cannot be Z-open because, otherwise, it would contradict the condition (i). Then, ϱ_1 is not Z-open because A or C do not meet the requirements. Assume without loss of generality that C does not meet the requirements. Then, one of the following cases must occur.

Case 1.1: ϱ_1 does not have a triplex $(\{A,D\},C)$ and $C \in Z$. Then, one of the following subgraphs must occur in G_1.[1]

$$A \overset{\frown}{\circ\!\!-\!\!\circ} B \overset{}{\circ\!\!-\!\!\circ} C \rightarrow D \quad A \overset{\frown}{\circ\!\!-\!\!\circ} B \leftarrow C \circ\!\!-\!\!\circ D \quad A \overset{\frown}{\rightarrow} B \leftarrow C - D \quad A \overset{\frown}{\leftrightarrow} B \leftrightarrow C - D$$

However, the first three subgraphs imply that ρ_1 is not Z-open, which is a contradiction. The fourth subgraph implies that ϱ_1 is Z-open, which is a contradiction.

Case 1.2: ϱ_1 has a triplex $(\{A,D\},C)$ and $C \notin Z \cup san_{G_1}(Z)$. Note that C cannot be a triplex node in ρ_1 because, otherwise, ρ_1 would not be Z-open. Then, one of the following subgraphs must occur in G_1.

$$A \overset{\frown}{\circ\!\!-\!\!\circ} B \overset{}{\leftharpoonup} C \leftarrow\!\!\circ D \quad A \overset{\frown}{\circ\!\!-\!\!\circ} B \overset{}{\leftharpoonup} C - D \quad A \overset{\frown}{\leftrightarrow} B \overset{}{\rightharpoonup} C - D \quad A \overset{\frown}{\rightarrow} B \overset{}{\rightharpoonup} C - D$$

However, the first and second subgraphs imply that $C \in Z \cup san_{G_1}(Z)$ because $B \in Z$, which is a contradiction. The third subgraph implies that $B - D$ is in G_1 by the constraint C3 and, thus, that the path obtained from ρ_1 by replacing $B - C - D$ with $B - D$ is Z-open, which contradicts the condition (i). For the fourth subgraph, assume that A and D are adjacent in G_1. Then, one of the following subgraphs must occur in G_1.

$$A \overset{\frown}{\rightarrow} B \overset{}{\rightharpoonup} C \overset{\frown}{\rightharpoonup} D \rightarrow E \quad A \overset{\frown}{\rightarrow} B \overset{}{\rightharpoonup} C \overset{\frown}{\rightharpoonup} D \leftarrow\!\!\circ E \quad A \overset{\frown}{\rightarrow} B \overset{}{\rightharpoonup} C \overset{\frown}{\rightharpoonup} D - E$$

However, the first subgraph implies that the path obtained from ρ_1 by replacing $A \rightarrow B - C - D$ with $A \rightarrow D$ is Z-open, because $D \notin Z$ since ρ_1 is Z-open. This contradicts the condition (i). The second subgraph implies that the path obtained from ρ_1 by replacing $A \rightarrow B - C - D$ with $A \rightarrow D$ is Z-open, because $D \in Z \cup san_{G_1}(Z)$ since ρ_1 is Z-open. This contradicts the condition (i). Therefore, only the third subgraph is possible. Thus, by repeatedly applying the previous reasoning, we can conclude without loss of generality that the following subgraph must occur in G_1, with $n \geq 4$, $V_1 = A$, $V_2 = B$, $V_3 = C$, $V_4 = D$ and where V_1 and V_n are not adjacent in G_1. Note that the subgraph below covers the case where A and D are not adjacent in the original subgraph by simply taking $n = 4$.

$$V_1 \overset{\frown}{\rightarrow} V_2 \overset{}{\rightharpoonup} V_3 \overset{}{\rightharpoonup} V_4 - \cdots - V_{n-1} - V_n$$

Since V_1 and V_n are not adjacent in G_1, G_1 has a triplex $(\{V_1, V_n\}, V_{n-1})$ and, thus, so does G_2 because G_1 and G_2 are triplex equivalent. Then, one of the following subgraphs must occur in G_2.

$$V_1 \overset{\frown}{\cdots} V_{n-1} \leftarrow\!\!\circ V_n \quad V_1 \overset{\frown}{\circ \cdots} V_{n-1} \leftarrow\!\!\circ V_n \quad V_1 \overset{\frown}{\circ \cdots} V_{n-1} - V_n$$

---

[1] If ϱ_1 does not have a triplex $(\{A,D\},C)$, then $A \leftarrow C$, $C \rightarrow D$ or $A - C - D$ must be in G_1. Moreover, recall that B is a triplex node in ρ_1. Then, $A \rightarrow B \leftarrow C$, $A \rightarrow B \leftrightarrow C$, $A \rightarrow B - C$, $A \leftrightarrow B \leftarrow C$, $A \leftrightarrow B \leftrightarrow C$, $A \leftrightarrow B - C$, $A - B \leftarrow C$ or $A - B \leftrightarrow C$ must be in G_1. However, if $A \leftarrow C$ is in G_1 then the only legal options are those that contain the edge $B \leftarrow C$. On the other hand, if $A - C - D$ is in G_1 then the only legal options are $A \rightarrow B \leftarrow C$ and $A \leftrightarrow B \leftrightarrow C$.

Note that $V_1, \ldots, V_n$ must be a path in G_2, because G_1 and G_2 are triplex equivalent. Note also that this path cannot have any triplex in G_2. To see it, recall that we assumed that ρ_2 does not have a triplex $(\{A, C\}, B)$. Recall that $V_1 = A$, $V_2 = B$, $V_3 = C$. Moreover, if the path $V_1, \ldots, V_n$ has a triplex $(\{V_i, V_{i+2}\}, V_{i+1})$ in G_2 with $2 \leq i \leq n - 2$, then V_i and V_{i+2} must be adjacent in G_1 and G_2, because such a triplex does not exist in G_1, which is triplex equivalent to G_2. Specifically, $V_i - V_{i+2}$ must be in G_1 because, as seen above, $V_i - V_{i+1} - V_{i+2}$ is in G_1. Then, the path obtained from ρ_1 by replacing $V_i - V_{i+1} - V_{i+2}$ with $V_i - V_{i+2}$ is Z-open, which contradicts the condition (i). However, if the path $V_1, \ldots, V_n$ has no triplex in G_2, then every edge in the path must be directed as $\leftarrow$ in the case of the first and second subgraphs above, whereas every edge in the path must be undirected or directed as $\leftarrow$ in the third subgraph above. Either case contradicts the constraint C1 or C2.

Case 2: Case 1 does not apply. Then, ρ_2 has a triplex $(\{A, C\}, B)$ and $B \notin Z \cup san_{G_2}(Z)$. Then, ρ_1 cannot have a triplex $(\{A, C\}, B)$. Then, A and C must be adjacent in G_1 and G_2 because these are triplex equivalent. Let ϱ_1 be the path obtained from ρ_1 by replacing the triplex $(\{A, C\}, B)$ with the edge between A and C in G_1. Note that ϱ_1 cannot be Z-open because, otherwise, it would contradict the condition (i). Then, ϱ_1 is not Z-open because A or C do not meet the requirements. Assume without loss of generality that C does not meet the requirements. Then, one of the following cases must occur.

    Case 2.1: ϱ_1 has a triplex $(\{A, D\}, C)$ and $C \notin Z \cup san_{G_1}(Z)$. Then, one of the following subgraphs must occur in G_1.[2]

$$A \overset{\frown}{\circ\!\!-\!\!\circ} B \overset{\frown}{\rightarrow} C \leftarrow\!\!\circ D \quad A \overset{\frown}{\circ\!\!-\!\!\circ} B \overset{\frown}{\rightarrow} C - D \quad A \overset{\frown}{\leftarrow} B \overset{\frown}{\rightarrow} C \leftarrow\!\!\circ D$$

$$A \overset{\frown}{\leftarrow} B \overset{\frown}{\rightarrow} C - D \quad A \overset{\frown}{\leftarrow} B \overset{\frown}{\rightarrow} C \leftarrow\!\!\circ D \quad A \overset{\frown}{-} B - C \leftarrow\!\!\circ D$$

    However, this implies that C is a triplex node in ρ_1, which is a contradiction because ρ_1 is Z-open but $C \notin Z \cup san_{G_1}(Z)$.

    Case 2.2: ϱ_1 does not have a triplex $(\{A, D\}, C)$ and $C \in Z$. Then, $A \leftarrow C$, $C \rightarrow D$ or $A - C - D$.

        Case 2.2.1: If $C \rightarrow D$ or $A - C - D$, then one of the following subgraphs must occur in G_1.

$$A \overset{\frown}{\circ\!\!-\!\!\circ} B \circ\!\!-\!\!\circ C \rightarrow D \quad A \overset{\frown}{-} B - C - D \quad A \overset{\frown}{\leftarrow} B \overset{\frown}{\rightarrow} C - D$$

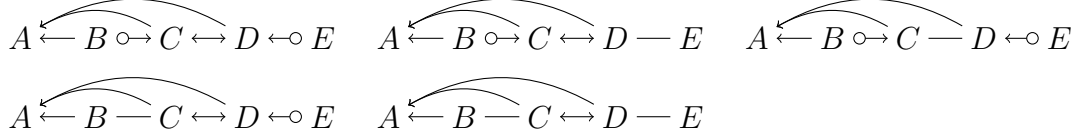   However, the first and second subgraphs imply that ρ_1 is not Z-open, which is a contradiction. The third subgraph implies that ϱ_1 is Z-open, which is a contradiction.

        Case 2.2.2: If $A \leftarrow C$ then $(\{A, D\}, C)$ is not a triplex in ϱ_1. However, note that ρ_1 must have a triplex $(\{B, D\}, C)$, because ρ_1 is Z-open and $C \in Z$. Then, one of the following subgraphs must occur in G_1.

$$A \overset{\frown}{\leftarrow} B \circ\!\!-\!\!\circ C \leftarrow\!\!\circ D \quad A \overset{\frown}{\leftarrow} B \circ\!\!-\!\!\circ C - D \quad A \overset{\frown}{\leftarrow} B - C \leftarrow\!\!\circ D$$

---

[2]If ϱ_1 has a triplex $(\{A, D\}, C)$, then $A \rightarrow C \leftarrow D$, $A \rightarrow C - D$, $A \leftrightarrow C \leftarrow D$, $A \leftrightarrow C - D$ or $A - C \leftarrow D$ must be in G_1. Moreover, recall that B is not a triplex node in ρ_1. Then, $A \leftarrow B \leftarrow C$, $A \leftarrow B \rightarrow C$, $A \leftarrow B \leftrightarrow C$, $A \leftarrow B - C$, $A \rightarrow B \rightarrow C$, $A \leftrightarrow B \rightarrow C$, $A - B \rightarrow C$ or $A - B - C$ must be in G_1. However, if $A \rightarrow C$ is in G_1 then the only legal options are those that contain the edge $B \rightarrow C$. On the other hand, if $A \leftrightarrow C$ is in G_1 then the only legal option is $A \leftarrow B \rightarrow C$. Finally, if $A - C$ is in G_1 then the only legal options are $A \leftarrow B \rightarrow C$ and $A - B - C$.

Assume that A and D are adjacent in G_1. Then, $A \leftarrow D$ must be in G_1. Moreover, $D \in Z$ because, otherwise, we can remove B and C from ρ_1 and get a Z-open path between A and B in G_1 that is shorter than ρ_1, which contradicts the condition (i). Then, D must be a triplex node in ρ_1. Then, one of the following subgraphs must occur in G_1.

$$A \overset{\frown}{\leftarrow} B \circ\!\!\rightarrow C \leftrightarrow D \leftarrow\!\!\circ E \qquad A \overset{\frown}{\leftarrow} B \circ\!\!\rightarrow C \leftrightarrow D - E \qquad A \overset{\frown}{\leftarrow} B \circ\!\!\rightarrow C - D \leftarrow\!\!\circ E$$

$$A \overset{\frown}{\leftarrow} B - C \leftrightarrow D \leftarrow\!\!\circ E \qquad A \overset{\frown}{\leftarrow} B - C \leftrightarrow D - E$$

Thus, by repeatedly applying the previous reasoning, we can conclude without loss of generality that the following subgraph must occur in G_1, with $n \geq 4$, $V_1 = A$, $V_2 = B$, $V_3 = C$, $V_4 = D$ and where V_1 and V_n are not adjacent in G_1. Note that the subgraph below covers the case where A and D are not adjacent in the original subgraph by simply taking $n = 4$.

$$V_1 \overset{\frown}{\leftarrow} V_2 \quad \cdots \quad V_{n-1} \circ\!\!-\!\!\circ V_n$$

Note that V_i is a triplex node in ρ_1 for all $3 \leq i \leq n-1$. Then, $V_i \in Z$ for all $3 \leq i \leq n-1$ by the condition (ii) because ρ_1 is Z-open. Then, V_i must be a triplex node in ρ_2 for all $3 \leq i \leq n-1$ because, otherwise, Case 1 would apply instead of Case 2. Recall that $V_2 = B$ is also a triplex node in ρ_2. Note that G_1 does not have a triplex $(\{V_1, V_n\}, V_{n-1})$ and, thus, G_2 does not have it either because these are triplex equivalent. Then, one of the following subgraphs must occur in G_2.

$$V_1 \overset{\frown}{\circ} \cdots \overset{\circ} V_{n-1} \rightarrow V_n \qquad V_1 \overset{\frown}{\leftarrow} \cdots V_{n-1} \circ\!\!-\!\!\circ V_n \qquad V_1 \overset{\frown}{} \cdots V_{n-1} - V_n$$

However, the first subgraph implies that V_{n-1} is not a triplex node in ρ_2, which is a contradiction. The second subgraph implies that G_2 has a cycle that violates the constraint C1. To see it, recall that V_i is a triplex node in ρ_2 for all $2 \leq i \leq n-1$ and, thus, $V_i \leftarrow V_{i+1}$ is not in G_2 for all $1 \leq i \leq n-2$. The third subgraph implies that $V_{n-2} \leftrightarrow V_{n-1}$ is not in G_2 because, otherwise, V_1 and V_n would be adjacent by the constraint C3. Therefore, $V_{n-2} \rightarrow V_{n-1}$ must be in G_2 because V_{n-1} is a triplex node in ρ_2. However, this implies that V_{n-2} is not a triplex node in ρ_2, which is a contradiction.

Part 2

Let ρ_1 be any of the shortest Z-open paths between α and β in G_1 st all its triplex nodes are in Z. Let ρ_2 be the path in G_2 that consists of the same nodes as ρ_1. We prove below that ρ_2 is Z-open. We prove this result by induction on the number of non-triplex nodes of ρ_1 that are in Z. If this number is zero, then Part 1 proves the result. Assume as induction hypothesis that the result holds when the number is smaller than m. We now prove it for m.

Let $\rho_1^{A:B}$ denote the subpath of ρ_1 between the nodes A and B. Let C be any of the non-triplex nodes of ρ_1 that are in Z. Note that there must exist some node $D \in pa_{G_1}(C) \smallsetminus Z$ for ρ_1 to be Z-open. If D is in ρ_1, then $\rho_1^{\alpha:D} \cup D \rightarrow C \cup \rho_1^{C:\beta}$ or $\rho_1^{\alpha:C} \cup C \leftarrow D \cup \rho_1^{D:\beta}$ is a Z-open path between α and β in G_1 that has fewer than m non-triplex nodes in Z. Then, the result holds by the induction hypothesis. On the other hand, if D is not in ρ_1, then $\rho_1^{\alpha:C} \cup C \leftarrow D$ and $D \rightarrow C \cup \rho_1^{C:\beta}$ are two paths. Moreover, they are Z-open in G_1 and they have fewer than
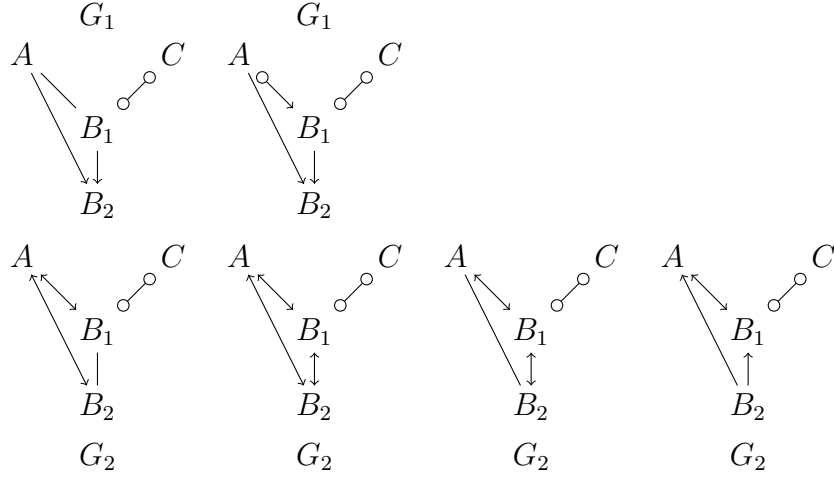
m non-triplex nodes in Z. Then, by the induction hypothesis, there are two Z-open paths $\rho_2^{\alpha:D}$ and $\rho_2^{D:\beta}$ in G_2 st the former ends with the nodes C and D and the latter starts with these two nodes. Now, consider the following cases.

Case 1: $\rho_2^{\alpha:D}$ ends with $A - C \leftarrow D$. Then, $\rho_2^{D:\beta}$ starts with $D \to C - B$ or $D \to C \leftarrowtail B$. Then, $\rho_2 = \rho_2^{\alpha:C} \cup \rho_2^{C:\beta}$ is Z-open a path in either case.

Case 2: $\rho_2^{\alpha:D}$ ends with $A - C \leftrightarrow D$. Then, $\rho_2^{D:\beta}$ starts with $D \leftrightarrow C - B$ or $D \leftrightarrow C \leftarrowtail B$. Then, $\rho_2 = \rho_2^{\alpha:C} \cup \rho_2^{C:\beta}$ is Z-open a path in either case.

Case 3: $\rho_2^{\alpha:D}$ ends with $A \rightarrowtail C - D$. Then, $\rho_2^{D:\beta}$ starts with $D - C \leftarrowtail B$, or $D - C - B$ st there is some node $E \in pa_{G_2}(C) \setminus Z$. Then, $\rho_2 = \rho_2^{\alpha:C} \cup \rho_2^{C:\beta}$ is Z-open a path in either case.

Case 4: $\rho_2^{\alpha:D}$ ends with $A \rightarrowtail C \leftarrowtail D$. Then, $\rho_2^{D:\beta}$ starts with $D \rightarrowtail C - B$ or $D \rightarrowtail C \leftarrowtail B$. Then, $\rho_2 = \rho_2^{\alpha:C} \cup \rho_2^{C:\beta}$ is Z-open a path in either case.

Case 5: $\rho_2^{\alpha:D}$ ends with $A - C - D$ st there is some node $E \in pa_{G_2}(C) \setminus Z$. Then, $\rho_2^{D:\beta}$ starts with $D - C \leftarrowtail B$, or $D - C - B$ st there is some node $F \in pa_{G_2}(C) \setminus Z$. Then, $\rho_2 = \rho_2^{\alpha:C} \cup \rho_2^{C:\beta}$ is a Z-open path in either case.

Part 3

Assume that Part 2 does not apply. Then, every Z-open path between α and β in G_1 has some triplex node B_1 that is outside Z because, otherwise, Part 2 would apply. Note that for the path to be Z-open, G_1 must have a subgraph $B_1 \to \ldots \to B_n$ st $B_1, \ldots, B_{n-1} \notin Z$ but $B_n \in Z$. Let us convert every Z-open path between α and β in G_1 into a route by replacing each of its triplex nodes B_1 that are outside Z with the corresponding route $B_1 \to \ldots \to B_n \leftarrow \ldots \leftarrow B_1$. Let ϱ_1 be any of the shortest routes so-constructed. Let ρ_1 be the path from which ϱ_1 was constructed. Note that ρ_1 cannot be Z-open because, otherwise, Part 2 would apply. Let W denote the set of all the triplex nodes in ρ_1 that are outside Z. Then, ρ_1 is one of the shortest $(Z \cup W)$-open paths between α and β in G_1 st all its triplex nodes are in $Z \cup W$. To see it, assume to the contrary that ρ_1' is a $(Z \cup W)$-open path between α and β in G_1 that is shorter than ρ_1 and st all the triplex nodes in ρ_1' are in $Z \cup W$. Let ϱ_1' be the route resulting from replacing every node B_1 of ρ_1' that is in W with the route $B_1 \to \ldots \to B_n \leftarrow \ldots \leftarrow B_1$ that was added to ρ_1 to construct ϱ_1. Clearly, ϱ_1' is shorter than ϱ_1, which is a contradiction. Let ϱ_2 and ρ_2 be the route and the path in G_2 that consist of the same nodes as ϱ_1 and ρ_1. Note that ρ_2 is $(Z \cup W)$-open by Part 2.

Consider any of the routes $B_1 \to \ldots \to B_n \leftarrow \ldots \leftarrow B_1$ that were added to ρ_1 to construct ϱ_1. This implies that ρ_1 has a triplex $(\{A, C\}, B_1)$. Assume that $B_1 \to B_2$ is in G_1 but $B_1 - B_2$ or $B_1 \leftarrowtail B_2$ is in G_2. Note that $A \rightarrowtail B_1$ or $B_1 \leftarrowtail C$ is in G_2 because, as noted above, ρ_2 is $(Z \cup W)$-open. Assume without loss of generality that $A \rightarrowtail B_1$ is in G_2. Then, $A - B_1 \to B_2$ or $A \rightarrowtail B_1 \to B_2$ is in G_1 whereas $A \rightarrowtail B_1 - B_2$ or $A \rightarrowtail B_1 \leftarrowtail B_2$ is in G_2. Therefore, A and B_2 must be adjacent in G_1 and G_2 because these are triplex equivalent. This implies that $A \to B_2$ is in G_1. Moreover, $A \in Z$ because, otherwise, we can construct a route that is shorter than ϱ_1 by simply removing B_1 from ϱ_1, which is a contradiction. This implies that $A \leftrightarrow B_1$ is in G_2 because, otherwise, ρ_2 would not be $(Z \cup W)$-open. This implies that $A \leftrightarrow B_1 - B_2$ or $A \leftrightarrow B_1 \leftarrowtail B_2$ is in G_2, which implies that $A - B_2$ or $A \leftarrowtail B_2$ is in G_2. The situation is depicted in the following subgraphs.

$$G_1 \qquad\qquad G_1$$

$$A \qquad C \qquad A \qquad C$$
$$B_1 \qquad\qquad B_1$$
$$B_2 \qquad\qquad B_2$$

$$A \qquad C \quad A \qquad C \quad A \qquad C \quad A \qquad C$$
$$B_1 \qquad B_1 \qquad B_1 \qquad B_1$$
$$B_2 \qquad B_2 \qquad B_2 \qquad B_2$$
$$G_2 \qquad\quad G_2 \qquad\quad G_2 \qquad\quad G_2$$

Now, let A' be the node that precedes A in ρ_1. Note that $A' \leftarrow A$ cannot be in ρ_1 or ρ_2 because, otherwise, these would not be $(Z \cup W)$-open since $A \in Z$. Then, $A' - A$ or $A' \rightarrowtail A$ is in G_1 and G_2. Then, $A' - A \rightarrow B_2$ or $A' \rightarrowtail A \rightarrow B_2$ is in G_1 whereas $A' - A \leftarrow B_2$, $A' - A - B_2$, $A' \rightarrowtail A \leftarrow B_2$ or $A' \rightarrowtail A - B_2$ is in G_2. These four subgraphs of G_2 imply that A' and B_2 must be adjacent in G_1 and G_2: The second subgraph due to the constraint C3 because $A \leftrightarrow B_1$ is in G_2, and the other three subgraphs because G_1 and G_2 are triplex equivalent. By repeating the reasoning in the paragraph above, we can conclude that $A' \rightarrow B_2$ is in G_1, which implies that $A' \in Z$, which implies that $A' - A$ or $A' \leftrightarrow A$ is in G_2, which implies that $A' - B_2$ or $A' \leftarrowtail B_2$ is in G_2.

By repeating the reasoning in the paragraph above,[3] we can conclude that $\alpha \rightarrow B_2$ is in G_1 and, thus, we can construct a route that is shorter than ϱ_1 by simply removing some nodes from ϱ_1, which is a contradiction. Consequently, $B_1 \rightarrow B_2$ must be in G_2.

Finally, assume that $B_1 \rightarrow B_2 \rightarrow B_3$ is in G_1 but $B_1 \rightarrow B_2 - B_3$ or $B_1 \rightarrow B_2 \leftarrowtail B_3$ is in G_2. Then, B_1 and B_3 must be adjacent in G_1 and G_2 because these are triplex equivalent. This implies that $B_1 \rightarrow B_3$ is in G_1, which implies that we can construct a route that is shorter than ϱ_1 by simply removing B_2 from ϱ_1, which is a contradiction. By repeating this reasoning, we can conclude that $B_1 \rightarrow \ldots \rightarrow B_n$ is in G_2 and, thus, that ρ_2 is Z-open. $\qquad\square$

We mentioned in the previous section that MAMP CGs are a proper generalization of AMP CGs and MVR CGs, as there are independence models that can be induced by the former but not by the latter. Moreover, we gave the following example and postponed the proof that it cannot be induced by AMP CGs and MVR CGs.

$$A \rightarrow B \;\text{---}\; C$$
$$\downarrow \qquad \updownarrow$$
$$D \leftrightarrow E$$

With the help of Theorem 4, we can now give the proof. Assume to the contrary that the independence model induced by the MAMP CG G above can be induced by an AMP CG H. Note that H is a MAMP CG too. Then, G and H must have the same triplexes by Theorem 4. Then, H must have triplexes $(\{A, D\}, B)$ and $(\{A, C\}, B)$ but no triplex $(\{C, D\}, B)$. So, $C - B - D$ must be in H. Moreover, H must have a triplex $(\{B, E\}, C)$. So, $C \leftarrow E$ must be in H. However, this implies that H does not have a triplex $(\{C, D\}, E)$, which is a contradiction because G has such a triplex. To see that no MVR CG can induce the independence model

---

[3]Let A'' be the node that precedes A' in ρ_1. For this repeated reasoning to be correct, it is important to realize that if $A' - A$ is in G_2, then $A'' \leftrightarrow A'$ must be in G_2, because $A' \in Z$ and ρ_2 is $(Z \cup W)$-open.

induced by G, simply note that no MVR CG can have triplexes $(\{A, D\}, B)$ and $(\{A, C\}, B)$ but no triplex $(\{C, D\}, B)$.

We prove below that every triplex equivalence class of MAMP CGs has a distinguished member. We say that two nodes form a directed node pair if there is a directed edge between them.

Lemma 1. *For every triplex equivalence class of MAMP CGs, there is a unique maximal (wrt to set inclusion) set of directed node pairs st some CG in the class has exactly those directed node pairs.*

Proof. Assume to the contrary that there are two such sets of directed node pairs. Let the MAMP CG G contain exactly the directed node pairs in one of the sets, and let the MAMP CG H contain exactly the directed node pairs in the other set. For every $A \to B$ in G st $A - B$ or $A \leftrightarrow B$ is in H, replace the edge between A and B in H with $A \to B$ and call the resulting graph F. We prove below that F is a MAMP CG that is triplex equivalent to G and thus to H, which is a contradiction since F has a proper superset of the directed node pairs in H.

First, note that F cannot violate the constraints C2 and C3. Assume to the contrary that F violates the constraint C1 due to a cycle ρ. Note that none of the directed edges in ρ can be in H because, otherwise, H would violate the constraint C1, since H has the same adjacencies as F but a subset of the directed edges in F. Then, all the directed edges in ρ must be in G. However, this implies the contradictory conclusion that G violates the constraint C1, since G has the same adjacencies as F but a subset of the directed edges in F.

Second, assume to the contrary that G (and, thus, H) has a triplex $(\{A, C\}, B)$ that F has not. Then, $\{A, B\}$ or $\{B, C\}$ must an directed node pair in G because, otherwise, F would have a triplex $(\{A, C\}, B)$ since F would have the same induced graph over $\{A, B, C\}$ as H. Specifically, $A \to B$ or $B \leftarrow C$ must be in G because, otherwise, G would not have a triplex $(\{A, C\}, B)$. Moreover, neither $A \leftarrow B$ nor $B \to C$ can be H because, otherwise, H would not have a triplex $(\{A, C\}, B)$. Therefore, if $A \to B$ or $B \leftarrow C$ is in G and neither $A \leftarrow B$ nor $B \to C$ is in H, then $A \to B$ or $B \leftarrow C$ must be in F. However, this implies that $B \to C$ or $A \leftarrow B$ must be in F because, otherwise, F would have a triplex $(\{A, C\}, B)$ which would be a contradiction. However, this is a contradiction since neither $B \to C$ nor $A \leftarrow B$ can be in G or H because, otherwise, neither G nor H would have a triplex $(\{A, C\}, B)$.
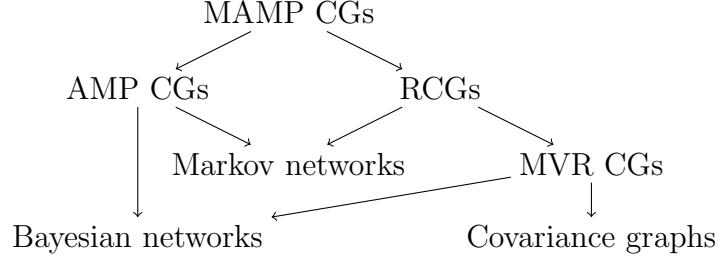
Finally, assume to the contrary that F has a triplex $(\{A, C\}, B)$ that G has not (and, thus, nor does H). Then, $A - B - C$ must be in H because, otherwise, $A \leftarrow B$ or $B \to C$ would be in H and, thus, F would not have a triplex $(\{A, C\}, B)$. However, this implies that $A \to B$ or $B \leftarrow C$ is in G because, otherwise, F would not have a triplex $(\{A, C\}, B)$. However, this implies that $B \to C$ or $A \leftarrow B$ is in G because, otherwise, G would have a triplex $(\{A, C\}, B)$. Therefore, $A \to B \to C$ or $A \leftarrow B \leftarrow C$ is in G and, thus, $A \to B \to C$ or $A \leftarrow B \leftarrow C$ must be in F since $A - B - C$ is in H. However, this contradicts the assumption that F has a triplex $(\{A, C\}, B)$. □

A MAMP CG is a directed CG (DCG) if it has exactly the maximal set of directed node pairs corresponding to its triplex equivalence class. Note that there may be several DCGs in the class. For instance, the triplex equivalence class that contains the MAMP CG $A \to B$ has two DCGs (i.e. $A \to B$ and $A \leftarrow B$).

Lemma 2. *For every triplex equivalence class of DCGs, there is a unique maximal (wrt to set inclusion) set of bidirected edges st some DCG in the class has exactly those bidirected edges.*

Proof. Assume to the contrary that there are two such sets of bidirected edges. Let the DCG G contain exactly the bidirected edges in one of the sets, and let the DCG H contain exactly the bidirected edges in the other set. For every $A \leftrightarrow B$ in G st $A - B$ is in H, replace

TABLE 1. Subfamilies of MAMP CGs.

MAMP CGs

AMP CGs   RCGs

Markov networks   MVR CGs

Bayesian networks   Covariance graphs

$A - B$ with $A \leftrightarrow B$ in H and call the resulting graph F. We prove below that F is a DCG that is triplex equivalent to G, which is a contradiction since F has a proper superset of the bidirected edges in G.

First, note that F cannot violate the constraint C1. Assume to the contrary that F violates the constraint C2 due to a cycle ρ. Note that all the undirected edges in ρ are in H. In fact, they must also be in G, because G and H have the same directed node pairs and bidirected edges. Moreover, the bidirected edge in ρ must be in G or H. However, this is a contradiction. Now, assume to the contrary that F violates the constraint C3 because $A - B - C$ and $B \leftrightarrow D$ are in F but A and C are not adjacent in F (note that if A and C were adjacent in F, then they would not violate the constraint C3 or they would violate the constraint C1 or C2, which is impossible as we have just shown). Note that $A - B - C$ must be in H. In fact, $A - B - C$ must also be in G, because G and H have the same directed node pairs and bidirected edges. Moreover, $B \leftrightarrow D$ must be in G or H. However, this implies that A and C are adjacent in G or H by the constraint C3, which implies that A and C are adjacent in G and H because they are triplex equivalent and thus also in F, which is a contradiction. Consequently, F is a MAMP CG, which implies that F is a DCG because it has the same directed edges as G and H.

Second, note that all the triplexes in G are in F too.

Finally, assume to the contrary that F has a triplex $(\{A, C\}, B)$ that G has not (and, thus, nor does H). Then, $A - B - C$ must be in H because, otherwise, $A \leftarrow B$ or $B \rightarrow C$ would be in H and thus F would not have a triplex $(\{A, C\}, B)$. However, this implies that F has the same induced graph over $\{A, B, C\}$ as G, which contradicts the assumption that F has a triplex $(\{A, C\}, B)$. $\qquad\square$

A DCG is a bidirected DCG (BDCG) if it has exactly the maximal set of bidirected edges corresponding to its triplex equivalence class. Note that there may be several BDCGs in the class. For instance, the triplex equivalence class that contains the MAMP CG $A \rightarrow B$ has two BDCGs (i.e. $A \rightarrow B$ and $A \leftarrow B$). Note however that all the BDCGs in a triplex equivalence class have the same triplex edges, i.e. the edges in a triplex.

5. DISCUSSION

In this paper we have introduced MAMP CGs, a new family of graphical models that generalize AMP CGs and MVR CGs. We have described global and local Markov properties for them and proved their equivalence for compositional graphoids. We have also characterized when two MAMP CGs are Markov equivalent. We conjecture that every Markov equivalence class of MAMP CGs has a distinguished member. We are currently working on this question. It is worth mentioning that such a result has been proven for AMP CGs (Roverato and Studený, 2006). We are also working on a constraint based algorithm for learning a MAMP CG a given probability distribution is faithful to. The idea is to combine the learning algorithms that we have recently proposed for AMP CGs (Peña, 2012) and MVR CGs (Sonntag and Peña, 2012).

We believe that the most natural way to generalize AMP CGs and MVR CGs is by allowing undirected, directed and bidirected edges. However, we are not the first to introduce a family of graphical models whose members may contain these three types of edges. In the rest of this section, we review some works that have done it before us, and explain how our work differs from them. Cox and Wermuth (1993, 1996) introduced regression CGs (RCGs) to generalize MVR CGs by allowing them to have also undirected edges. The separation criterion for RCGs is identical to that of MVR CGs. Then, there are independence models that can be induced by MAMP CGs but that cannot be induced by RCGs, because RCGs generalize MVR CGs but not AMP CGs. For instance, the independence model induced by the AMP CG G

$$A$$
$$\downarrow$$
$$B \!-\! C \!-\! D$$

cannot be induced by any RCG. To see it, assume to the contrary that it can be induced by a RCG H. Note that H is a MAMP CG too. Then, G and H must have the same triplexes by Theorem 4. Then, H must have triplexes $(\{A, B\}, C)$ and $(\{A, D\}, C)$ but no triplex $(\{B, D\}, C)$. So, $B \multimap C \to D$, $B \multimap C - D$, $B \leftarrow C \multimap D$ or $B - C \multimap D$ must be in H. However, this implies that H does not have the triplex $(\{A, B\}, C)$ or $(\{A, D\}, C)$, which is a contradiction. It is worth mentioning that, although RCGs can have undirected edges, they cannot have a subgraph of the form $A \rightarrowtail B - C$. Therefore, RCGs are a subfamily of MAMP CGs. Table 1 depicts this and other subfamilies of MAMP CGs.

Another family of graphical models whose members may contain undirected, directed and bidirected edges is maximal ancestral graphs (MAGs) (Richardson and Spirtes, 2002). The separation criterion for MAGs is identical to that of MVR CGs. Therefore, the example above also serves to illustrate that MAGs generalize MVR CGs but not AMP CGs, as MAMP CGs do. See also (Richardson and Spirtes, 2002, p. 1025). It is worth mentioning that, although MAGs can have undirected edges, they must comply with certain topological constraints. As of today, we do not know if MAGs are a subfamily of MAMP CGs. We are currently working on this question. However, even when an independence model can be induced by a MAG, a MAMP CG may be preferred, as the following example illustrates. Consider the MAMP CG $A \leftarrow B \to C - D - E$ and call it G. Consider the independence model resulting from G by marginalizing out B. This model can be induced by the MAMP CG $A \leftrightarrow C - D - E$, or by the MAGs $A \rightarrowtail C \leftarrow D \multimap E$ or $A \rightarrowtail C \leftrightarrow D \to E$. However, the MAGs suggest the existence of the causal relationship $C \leftarrow D$ or $D \to E$, although neither exists in G. On the other hand, the MAMP CG does not suggest any causal relationship and, thus, it is preferred.

Note that the result of marginalizing out B in the MAMP CG in the previous example results in another MAMP CG. Unfortunately, this property does not hold in general, i.e. MAMP CGs are not closed wrt marginalization. On the other hand, MAGs are closed wrt marginalization and conditioning. This is a desirable property, because one can compare the result of different manipulations on the system being modeled. We plan to study how to extend or modify MAMP CGs so that they become closed wrt marginalization and conditioning.

Finally, three other families of graphical models whose members may contain undirected, directed and bidirected edges are summary graphs after replacing the dashed undirected edges with bidirected edges (Cox and Wermuth, 1996), MC graphs (Koster, 2002), and loopless mixed graphs (Sadeghi and Lauritzen, 2012). As shown in (Sadeghi and Lauritzen, 2012, Sections 4.2 and 4.3), every independence model that can be induced by summary graphs and MC graphs can also be induced by loopless mixed graphs. The separation criterion for loopless mixed graphs is identical to that of MVR CGs. Therefore, the example above also serves to illustrate that loopless mixed graphs generalize MVR CGs but not AMP CGs, as

MAMP CGs do. See also (Sadeghi and Lauritzen, 2012, Section 4.1). Moreover, summary graphs and MC graphs have a rather counterintuitive and undesirable feature: Not every missing edge corresponds to a separation (Richardson and Spirtes, 2002, p. 1023). MAMP CGs, on the other hand, do not have this disadvantage (recall Theorem 2).

In summary, MAMP CGs are the only graphical models we are aware of that generalize both AMP CGs and MVR CGs. A by-product of this generalization is the following. Andersson et al. (2001, Section 5) and Kang and Tian (2009, Section 2) show that any regular Gaussian probability distribution that is Markovian wrt an AMP CG or MVR CG G can be expressed as a system of linear equations with correlated errors whose structure depends on G. We extend below this result to MAMP CGs.

Let p denote any regular Gaussian distributions that is Markovian wrt a MAMP CG G. Assume without loss of generality that p has mean 0. Let K_i denote any connectivity component of G. Let $\Omega^i_{K_i,K_i}$ and $\Omega^i_{K_i,pa_G(K_i)}$ denote submatrices of the precision matrix Ω^i of $p(K_i, pa_G(K_i))$. Then, as shown by Bishop (2006, Section 2.3.1),

$$K_i | pa_G(K_i) \sim \mathcal{N}(\beta^i pa_G(K_i), \Lambda^i)$$

where

$$\beta^i = -(\Omega^i_{K_i,K_i})^{-1} \Omega^i_{K_i,pa_G(K_i)}$$

and

$$(\Lambda^i)^{-1} = \Omega^i_{K_i,K_i}.$$

Then, p can be expressed as a system of linear equations with normally distributed errors whose structure depends on G as follows:

$$K_i = \beta^i pa_G(K_i) + \epsilon^i$$

where

$$\epsilon^i \sim \mathcal{N}(0, \Lambda^i).$$

Note that for all $A, B \in K_i$ st $uc_G(A) = uc_G(B)$ and $A - B$ is not in G, $A \perp_G B | pa_G(K_i) \cup K_i \smallsetminus A \smallsetminus B$ and thus $(\Lambda^i_{uc_G(A),uc_G(A)})^{-1}_{A,B} = 0$ (Lauritzen, 1996, Proposition 5.2). Note also that for all $A, B \in K_i$ st $uc_G(A) \neq uc_G(B)$ and $A \leftrightarrow B$ is not in G, $A \perp_G B | pa_G(K_i)$ and thus $\Lambda^i_{A,B} = 0$. Finally, note also that for all $A \in K_i$ and $B \in pa_G(K_i)$ st $A \leftarrow B$ is not in G, $A \perp_G B | pa_G(A)$ and thus $(\beta^i)_{A,B} = 0$. Let β_A contain the nonzero elements of the vector $(\beta^i)_{A,\bullet}$. Then, p can be expressed as a system of linear equations with correlated errors whose structure depends on G as follows. For any $A \in K_i$,

$$A = \beta_A pa_G(A) + \epsilon^A$$

and for any other $B \in K_i$,

$$covariance(\epsilon^A, \epsilon^B) = \Lambda^i_{A,B}.$$

It is worth mentioning that the mapping above between probability distributions and systems of linear equations is bijective. We omit the proof of this fact here, but it can be proven much in the same way as Lemma 1 in Peña (2010). Note that each equation in the system of linear equations above is a univariate recursive regression, i.e. a random variable can be a regressor in an equation only if it has been the regressand in a previous equation. This has two main advantages, as Cox and Wermuth (1993, p. 207) explain: "First, and most importantly, it describes a stepwise process by which the observations could have been generated and in this sense may prove the basis for developing potential causal explanations. Second, each parameter in the system [of linear equations] has a well-understood meaning since it is a regression coefficient: That is, it gives for unstandardized variables the amount by which the response is expected to change if the explanatory variable is increased by one unit and all other variables in the equation are kept constant." Therefore, a MAMP CG can be seen as a data generating process and, thus, it gives us insight into the system under study.

We close this section by introducing marginal-conditional AMP (MCAMP) CGs, which extend MAMP CGs with what we call convergent edges. We defer a thorough study of MCAMP CGs. Note that the edges in a MAMP CG can be split into those that have no head and no tail (undirected edges), one head and one tail (directed edges), and two heads and no tail (bidirected edges). Therefore, one may say that there is a combination missing, namely edges with no head and two tails. We call these edges convergent and represent them as $A \twoheadleftarrow B$.

In a graph, a triplex $(\{A, C\}, B)$ is an induced subgraph of the form $A \rightarrowtail B \leftarrowtail C$, $A \rightarrowtail B - C$, or $A - B \leftarrowtail C$. A graph G containing possibly directed, undirected, bidirected and convergent edges is a MCAMP CG if

C1. G has no cycle $V_1, \ldots, V_n = V_1$ st $V_1 \rightarrow V_2$ is in G and $V_i \rightarrowtail V_{i+1}$, $V_i - V_{i+1}$ or $V_i \twoheadleftarrow V_{i+1}$ is in G for all $1 < i < n$,

C2. G has no cycle $V_1, \ldots, V_n = V_1$ st $V_1 \leftrightarrow V_2$ is in G and $V_i - V_{i+1}$ or $V_i \twoheadleftarrow V_{i+1}$ is in G for all $1 < i < n$,

C3. if G has a subgraph $V_1 - V_2 - V_3$ st $V_2 \leftrightarrow V_4$ is in G, then V_1 and V_3 are adjacent in G, and

C4. if G has a triplex $(\{V_1, V_3\}, V_2)$ st $V_2 \twoheadleftarrow V_4$ or $V_2 \rightarrow \ldots \rightarrow V_4 \twoheadleftarrow V_5$ are in G, then V_1 and V_3 are adjacent in G.

The semantics of MCAMP CGs is as follows. A node B in a path ρ in a MCAMP CG G is called a triplex node in ρ if $A \rightarrowtail B \leftarrowtail C$, $A \rightarrowtail B - C$, or $A - B \leftarrowtail C$ is a subpath of ρ. Moreover, ρ is said to be Z-open with $Z \subseteq V$ when

- every triplex node B in ρ is in $Z \cup san_G(Z)$ unless $B \twoheadleftarrow D$ or $B \rightarrow \ldots \rightarrow D \twoheadleftarrow E$ is in G, and
- every non-triplex node B in ρ is outside Z unless $A - B - C$ is a subpath of ρ and $B \leftrightarrow D$ is in G or $B \leftarrow D$ is in G with $D \notin Z$.

Let X, Y and Z denote three disjoint subsets of V. When there is no Z-open path in G between a node in X and a node in Y, we say that X is separated from Y given Z in G and denote it as $X \perp_G Y | Z$. We denote by $X \not\perp_G Y | Z$ that $X \perp_G Y | Z$ does not hold. The independence model induced by G is the set of separation statements $X \perp_G Y | Z$.

The motivation behind the four constraints in the definition of MCAMP CGs is as follows. The constraint C1 generalizes the constraint C1 of MAMP CGs. For the constraints C2-C4, note that typically every missing edge in a graphical model corresponds to a separation. However, this may not be true for graphs that do not satisfy these constraints. For instance, the graph G to the left below does not contain any edge between B and D but $B \not\perp_G D | Z$ for all $Z \subseteq V \smallsetminus \{B, D\}$. Since this situation is counterintuitive, we enforce the constraint C3.
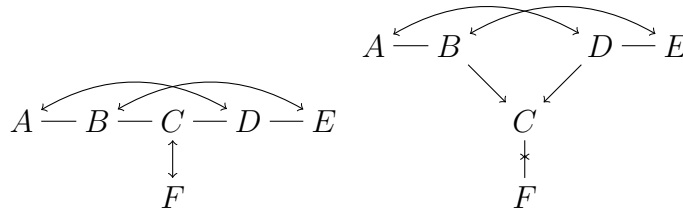
Let now G be the graph to the right above. Note that G does not contain any edge between B and D but $B \not\perp_G D | Z$ for all $Z \subseteq V \smallsetminus \{B, D\}$. Since this situation is again counterintuitive, we enforce the constraint C4 on G. We have in principle five ways to do so, i.e. there are five ways to make B and D adjacent in G. However, in this case they reduce to the following three due to the constraint C1.
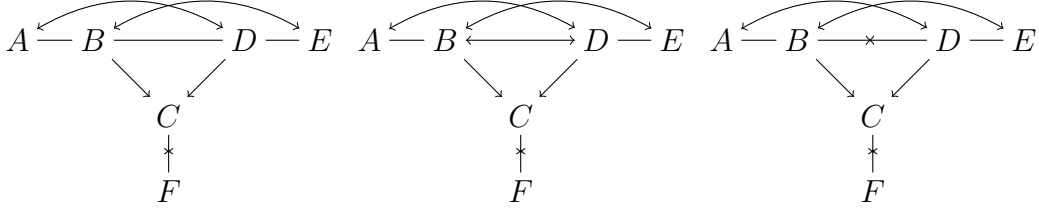
However, in each of the three graphs above, there is no edge between A and E but $A \not\perp_G E | Z$ for all $Z \subseteq V \smallsetminus \{A, E\}$. Since this is again counterintuitive, we enforce the constraint C2.

Given a MCAMP CG G, let $\widehat{G}$ denote the AMP CG obtained by replacing every bidirected edge $A \leftrightarrow B$ in G with $A \leftarrow L_{AB} \rightarrow B$ and every convergent edge $A \rightarrowtail B$ in G with $A \rightarrow S_{AB} \leftarrow B$. Let S denote all the nodes of the form S_{AB} in $\widehat{G}$. Then, note that $X \perp_G Y | Z$ iff $X \perp_{\widehat{G}} Y | Z \cup S$ for all X, Y and Z disjoint subsets of V. Therefore, every MCAMP CG can be seen as the result of marginalizing out some nodes and conditioning on some others in an AMP CG, hence the name. We conjecture that all the results proven in this paper for MAMP CGs can be extended to MCAMP CGs.

Acknowledgments

References

Andersson, S. A., Madigan, D. and Perlman, M. D. Alternative Markov Properties for Chain Graphs. *Scandinavian Journal of Statistics*, 28:33-85, 2001.

Bishop, C. M. *Pattern Recognition and Machine Learning*. Springer, 2006.

Bouckaert, R. R. *Bayesian Belief Networks: From Construction to Inference*. PhD Thesis, University of Utrecht, 1995.

Cox, D. R. and Wermuth, N. Linear Dependencies Represented by Chain Graphs. *Statistical Science*, 8:204-218, 1993.

Cox, D. R. and Wermuth, N. *Multivariate Dependencies - Models, Analysis and Interpretation*. Chapman & Hall, 1996.

Drton, M. Discrete Chain Graph Models. *Bernoulli*, 15:736-753, 2009.

Kang, C. and Tian, J. Markov Properties for Linear Causal Models with Correlated Errors. *Journal of Machine Learning Research*, 10:41-70, 2009.

Koster, J. T. A. Marginalizing and Conditioning in Graphical Models. *Bernoulli*, 8:817-840, 2002.

Lauritzen, S. L. *Graphical Models*. Oxford University Press, 1996.

Levitz, M., Perlman M. D. and Madigan, D. Separation and Completeness Properties for AMP Chain Graph Markov Models. *The Annals of Statistics*, 29:1751-1784, 2001.

Peña, J. M. Faithfulness in Chain Graphs: The Gaussian Case. In *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics*, 588-599, 2010.

Peña, J. M. Learning AMP Chain Graphs under Faithfulness. In *Proceedings of the 6th European Workshop on Probabilistic Graphical Models*, 251-258, 2012.

Richardson, T. and Spirtes, P. Ancestral Graph Markov Models. *The Annals of Statistics*, 30:962-1030, 2002.

Roverato, A. and Studený, M. A Graphical Representation of Equivalence Classes of AMP Chain Graphs. *Journal of Machine Learning Research*, 7:1045-1078, 2006.

Sadeghi, K. and Lauritzen, S. L. Markov Properties for Mixed Graphs. arXiv:1109.5909v4 [stat.OT].

Sonntag, D. and Peña, J. M. Learning Multivariate Regression Chain Graphs under Faithfulness. In *Proceedings of the 6th European Workshop on Probabilistic Graphical Models*, 299-306, 2012.

Sonntag, D. and Peña, J. M. Chain Graph Interpretations and their Relations. In *Proceedings of the 12th European Conference on Symbolic and Quantitative Approaches to Reasoning under Uncertainty*, to appear.

Studený, M. *Probabilistic Conditional Independence Structures*. Springer, 2005.