Chain graphs which are maximal ancestral graphs are recursive causal graphs

Thomas Richardson

January 17, 2001

Abstract

In this note we prove that a chain graph is Markov equivalent to some DAG under marginalizing and conditioning if and only if it is Markov equivalent to a recursive causal graph.

1 Introduction

This note considers chain graph models under the standard Markov property introduced by Lauritzen, Wermuth and Frydenberg. For the purposes of interpretation it is of interest to know when a chain graph is Markov equivalent to some DAG under marginalizing and conditioning, since such a DAG may represent a generating process. In this note we prove that a chain graph is Markov equivalent to a DAG model under marginalizing and conditioning if and only if it is Markov equivalent to a 'recursive causal graph', introduced by Kiiveri et al. (1984).

The proof proceeds in two stages: first, we give sufficient conditions for a chain graph to be Markov equivalent to a recursive causal graph; second, we show that any chain graph that does not satisfy these conditions is not Markov equivalent to any DAG under marginalizing and conditioning. Since recursive causal graphs are ancestral, it follows as an immediate corollary of an existing result, that any recursive causal graph is Markov equivalent to a DAG under marginalizing and conditioning.

2 Basic concepts and definitions

For definitions of basic graphical concepts, chain graphs and the associated Markov property see Lauritzen (1996).

A *recursive causal graph* (Kiiveri et al., 1984) is a chain graph in which the following configuration

$$x \rightarrow y - z$$

never occurs (regardless of whether x and z are adjacent. It should be noted that although these graphs are termed 'causal', no explicit causal interpretation or manipulation theory is given to the directed edges present in the graph; in this context the term implies that the graph includes directed edges. A recursive causal graph can always be decomposed into an undirected graph and a DAG, with any edges connecting the components pointing towards the DAG.

Let $\mathfrak{I}(\mathcal{CG})$ denote the independence model given by applying the global Markov property to a chain graph, $\mathcal{CG}$.

3 Markov equivalence of a chain graph and a recursive causal graph

We say that two chain graphs $\mathcal{CG}_1$ and $\mathcal{CG}_2$ are *Markov equivalent* if $\mathfrak{I}(\mathcal{CG}_1) = \mathfrak{I}(\mathcal{CG}_2)$. If $\mathcal{CG}$ is a chain graph, then we let $(\mathcal{CG})^{\sim}$ denote the undirected graph formed by replacing all directed edges with undirected edges.

We define the *moral graph* $(\mathcal{CG})^m$ to be the undirected graph in which two vertices x and y are adjacent if either (i) x and y are adjacent in $\mathcal{CG}$, or (ii) there exists a path of the form

$$x \rightarrow v_1 - \cdots - v_r \leftarrow y$$

in the chain graph $\mathcal{CG}$.

A triple $\langle \alpha, B, \beta \rangle$ is said to form a *minimal complex* in $\mathcal{CG}$ if the induced subgraph on $B \cup \{\alpha, \beta\}$ is:

$$\alpha \rightarrow b_1 - \cdots - b_r \leftarrow \beta$$

with $B = \{b_1, \ldots, b_r\}$. We call a minimal complex in which B is a single vertex an *unshielded collider*.

Frydenberg (1990) Proposition 5.5 presents the following characterization of Markov equivalence for chain graphs:

Theorem 3.1 *Two chain graphs $\mathcal{CG}_1$, $\mathcal{CG}_2$ are Markov equivalent if and only if*

(a) $(\mathcal{CG}_1)^{\sim} = (\mathcal{CG}_2)^{\sim}$, *and*

(b) $\mathcal{CG}_1$ *and* $\mathcal{CG}_2$ *have the same minimal complexes.*

A *chordless cycle* in a chain graph is a sequence of vertices $\langle v_1, \ldots, v_n \rangle$ in which v_i and v_j are adjacent if and only if they occur consecutively (i.e. $|i - j| \equiv 1 \pmod{n}$). We define a chordless cycle to be *collider free* if the induced subgraph on $\{v_1, \ldots, v_n\}$ does not contain the configuration $\rightarrow v_i \leftarrow$. Let $\text{CF}_{\mathcal{CG}}$ denote the set of vertices in the chain graph $\mathcal{CG}$ which are in collider free chordless cycles containing 4 or more edges.

Theorem 3.2 *A chain graph $\mathcal{CG}$ is Markov equivalent to a recursive causal graph $\mathcal{R}$ if the following conditions hold:*

(a) *all minimal complexes in $\mathcal{CG}$ are colliders, and*

(b) $(\mathcal{CG}_A)^m = (\mathcal{CG}_A)^{\sim}$, *where* $A = \text{An}(\text{CF}_{\mathcal{CG}})$.

If condition (a) holds, then condition (b) requires that no vertex in an unshielded collider should be in a collider free chordless cycle or an ancestor of a vertex in a collider free chordless cycle. (Note that we are using *ancestor* as used in Lauritzen (1996); in the language of Richardson and Spirtes (2000) and Frydenberg (1990) the term 'anterior' would be used instead.)

Proof: Suppose that a chain graph, $\mathcal{CG}$, satisfies conditions (a) and (b). We will show that $\mathcal{CG}$ may be transformed into a recursive causal graph, and that at each step we are not changing the Markov properties of the graph. There are three steps:

Step 1 Undirect all edges in $\mathcal{CG}$ between vertices in A. Let the resulting graph be $\mathcal{CG}(1)$.

Step 2 Exhaustively apply the following rule to $\mathcal{CG}(i)$:

For a given vertex α, let $\mathcal{E}_\alpha$ be the set of edges $c - x$, such that $c \in \text{ch}_{\mathcal{CG}(i)}(\alpha)$ and $x \notin \text{ch}_{\mathcal{CG}(i)}(\alpha)$. Transform the graph by replacing all edges $c - x$ in $\mathcal{E}_\alpha$ with edges directed as $c \rightarrow x$.

Let the resulting graph be $\mathcal{CG}(r)$.

Step 3 For each chain component τ, for which $\text{pa}_{\mathcal{CG}(r)}(\tau) \neq \emptyset$, direct all edges in τ, so that no directed cycles and no unshielded colliders are introduced in τ.

We first show that each step is well-defined, and that after each transformation the resulting graph will still be a chain graph:
Step 1: Suppose for a contradiction that there is a partially directed cycle in the graph resulting from the transformation. Let $x \rightarrow y$ be a directed edge in this cycle. Since this partially directed cycle was not present in the original graph, it follows that there is an edge $a_1 - a_2$ in this cycle, where $a_1, a_2 \in A$. However, in this case x, y are also in A which is a contradiction.

Step 2: Suppose for a contradiction, that a partially directed cycle is introduced via an application of the transformation rule to a chain graph. Consider the first instance of this transformation which introduces a partially directed cycle and let $\langle x_1, \ldots, x_n, x_1 \rangle$ be $x_1 \rightarrow x_n$ be an instance. Since there were no partially directed cycles in the graph prior to applying the rule, it follows that all directed edges in the cycle were previously undirected edges in $\mathcal{E}_\alpha$. Hence $x_1 \in \text{ch}(\alpha)$, $x_n \notin \text{ch}(\alpha)$, where α is the vertex mentioned in the transformation rule. Let x_j be the first vertex after x_1 for which $x_j \notin \text{ch}(\alpha)$; such a vertex is guaranteed to exist since $x_n \notin \text{ch}(\alpha)$. Since $x_{j-1} \in \text{ch}(\alpha)$ it follows that $x_{j-1} \rightarrow x_j$, hence $\langle x_1, \ldots, x_n, x_1 \rangle$ does not form a partially directed cycle, which is a contradiction.

Step 3: First note that if a chain component τ in $\mathcal{CG}$ contains a chordless (undirected) four-cycle then the vertices in τ are contained in A, hence any edges $x \rightarrow v$, where v is an ancestor of a vertex in τ, or in τ, will be undirected after Step 1. As a consequence the transformation rule in Step 2 will not orient any edges between vertices in A. In addition, since the transformation in Step 2 replaces undirected edges with directed edges, no new non-decomposable chain components are introduced. Consequently, if after step 2 there is a chain component τ, with $\text{pa}_{\mathcal{CG}(r)}(\tau) \neq \emptyset$ then τ is decomposable, hence, by standard results, it is possible to direct the edges in τ without introducing any directed cycles or unshielded colliders. Thus this step is well-defined. The graph clearly remains a chain graph after each step in this transformation since each chain component is replaced with a directed acyclic graph.

We next show that after each transformation the resulting chain graph is Markov equivalent to the chain graph prior to the transformation. Since the transformations do not change $(\mathcal{CG})^{\sim}$, by Theorem 3.1, it is sufficient to show that the set of minimal complexes does not change. Note that by condition (a) any minimal complex in $\mathcal{CG}$ is an unshielded collider.

Step 1: It follows directly from condition (b) that there are no minimal complexes in the induced subgraph $\mathcal{CG}_A$ hence no minimal complexes are removed in this operation. Further, if $a \in A$ and $x \to a$ in $\mathcal{CG}$ then $x \in A$ and $x - a$ after the transformation. Consequently no new minimal complexes are introduced via this transformation.

Step 2: By condition (a) all minimal complexes are unshielded colliders, and hence do not involve any undirected edges. Hence no minimal complexes are removed during this step. Suppose for a contradiction that a new minimal complex is introduced by some application of the transformation rule, which directs an edge $\beta \to \gamma$ which was previously undirected. In this case there exists a vertex δ such that $\langle \beta, B, \delta \rangle$ forms a minimal complex with $\gamma \in B$, i.e.

$$\text{either} \quad \beta \to \gamma \leftarrow \delta \quad \text{or} \quad \beta \to \gamma - \cdots - \leftarrow \delta.$$

Note that in both cases δ and β are not adjacent. If α and δ are not adjacent then

$$\alpha \to \beta - \gamma \cdots \leftarrow \delta$$

formed a minimal complex in the chain graph prior to the transformation, contradicting condition (a). If α and δ are adjacent then $\{\alpha, \beta, \gamma, \delta\} \subseteq \mathrm{AF}_{\mathcal{CG}}$ since these vertices lie on a chordless collider free cycle containing more than 3 edges. Hence the edge between α and β was undirected in step 1. Further, any edge between vertices which were ancestors of α and β was also undirected in step 1. It follows from this that the edge between α and β could not have become directed through an application of the transformation in step 2. This is a contradiction.

Step 3: Since no directed edges are undirected by the transformations made in this step, and all minimal complexes in the graph are unshielded colliders, no minimal complexes are removed. Further, by construction, no new unshielded colliders are introduced involving two edges in τ. In addition, since all the edges in a chain component are replaced with directed edges, no minimal complexes involving undirected edges are introduced. We now observe that after step 2 if there are edges $\alpha \to \beta - \gamma$ then there is an edge

$\alpha \to \gamma$, since otherwise the edge $\beta - \gamma$ would have been oriented as $\beta \to \gamma$. Consequently, no new unshielded colliders involving one edge in τ and one edge that was already directed are introduced by the transformation.

The conclusion of the Theorem now follows directly from the observation that after step 3 the resulting graph is a recursive causal graph. □

It will follow directly from results proved in the next section that the conditions given in Theorem 3.2 are necessary as well as sufficient.

4 Markov equivalence of a chain graph and a DAG under marginalizing and conditioning

In this section we first prove that a chain graph which does not satisfy conditions (a) and (b) given in Theorem 3.2 is not Markov equivalent to any DAG under marginalizing and conditioning. We then show that any recursive causal graph is Markov equivalent to a DAG under marginalizing and conditioning.

4.1 Marginalizing and conditioning independence models

An independence model $\mathfrak{I}$ with vertex set V after marginalizing out a subset L, is simply the subset of triples which do not involve any vertices in L. More formally we define:

$$\mathfrak{I}[_L \equiv \left\{ \langle X, Y \mid Z \rangle \;\middle|\; \langle X, Y \mid Z \rangle \in \mathfrak{I};\; (X \cup Y \cup Z) \cap L = \emptyset \right\}.$$

If $\mathfrak{I}$ contains the independence relations present in a distribution P, then $\mathfrak{I}[_L$ contains the subset of independence relations remaining after marginalizing out the '*L*atent' variables in L.

An independence model $\mathfrak{I}$ with vertex set V after conditioning on a subset S is the set of triples defined as follows:

$$\mathfrak{I}[^S \equiv \left\{ \langle X, Y \mid Z \rangle \;\middle|\; \langle X, Y \mid Z \cup S \rangle \in \mathfrak{I};\; (X \cup Y \cup Z) \cap S = \emptyset \right\}.$$

Thus if $\mathfrak{I}$ contains the independence relations present in a distribution P then $\mathfrak{I}[^S$ constitutes the subset of independencies holding among the remaining variables after conditioning on S. (Note that the set S is suppressed in the conditioning set in the independence relations in the resulting independence model.) The letter S is used because Selection effects represent one context in which conditioning may occur.

Combining these definitions we obtain:

$$\mathfrak{I}[^S_L \equiv \Big\{\langle X, Y \mid Z\rangle \Big| \langle X, Y \mid Z \cup S\rangle \in \mathfrak{I};\ (X \cup Y \cup Z) \cap (S \cup L) = \emptyset\Big\}.$$

Proposition 4.1 *For an independence model $\mathfrak{I}$ over V containing disjoint subsets S_1, S_2, L_1, L_2,*

(i) $\mathfrak{I}[^{\emptyset}_{\emptyset} = \mathfrak{I}$,

(ii) $\left(\mathfrak{I}[^{S_1}_{L_1}\right)[^{S_2}_{L_2} = \mathfrak{I}[^{S_1 \cup S_2}_{L_1 \cup L_2}$.

4.2 Chain graphs containing minimal complexes that are not unshielded colliders

In this section we prove the following Lemma:

Lemma 4.2 *If a chain graph $\mathcal{CG}$ contains a minimal complex that is not an unshielded collider, then there is no DAG $\mathcal{D}$ such that*

$$\mathfrak{I}(\mathcal{D})[^S_L = \mathfrak{I}(\mathcal{CG})$$

We first require a definition and a result from Richardson (1998).

Distinct vertices x and y are said to be *inseparable* in independence model $\mathfrak{I}$ if there is no set W such that $x \perp\!\!\!\perp y \mid W$ in $\mathfrak{I}$. If x and y are not inseparable in $\mathfrak{I}$, they are *separable*. Let $[\mathfrak{I}]^{\text{Ins}}$ be the undirected graph in which there is an edge $x - y$ if and only if x and y are inseparable in $\mathfrak{I}$.

A vertex b will be said to be *between* x and y in $\mathfrak{I}$ if and only if there exists a sequence of *distinct* vertices $\langle x \equiv x_0, x_1, \ldots, x_n \equiv b, x_{n+1}, \ldots, x_{n+m} \equiv y\rangle$ in $[\mathfrak{I}]^{\text{Ins}}$ such that each consecutive pair of vertices x_i, x_{i+1} in the sequence are inseparable in $\mathfrak{I}$. Clearly b will be between x and y in $\mathcal{G}$ if and only if b lies on a path between x and y in $[\mathfrak{I}]^{\text{Ins}}$. The set of vertices between x and y in $\mathfrak{I}$, is denoted $\mathit{Between}(\mathfrak{I}; x, y)$.

An independence model $\mathfrak{I}$ is *between-separated*, if for all pairs of vertices x, y and sets W $(x, y \notin W)$:

$$\text{If} \quad x \perp\!\!\!\perp y \mid W \text{ in } \mathfrak{I} \quad \text{then} \quad x \perp\!\!\!\perp y \mid W \cap \text{Between}(\mathfrak{I}; x, y) \text{ in } \mathfrak{I}.$$

Intuitively an independence model is between-separated if whenever a set W makes a pair of vertices, x, y independent, x and y remain independent when W is restricted to those vertices in W that are between x and y.

Theorem 4.3 *For an arbitrary DAG, $\mathcal{D}$, with vertex set V containing disjoint subsets S, L the independence model $\mathfrak{I}(\mathcal{D})[^S_L$ is between-separated.*

Lemma 4.4 *If $\mathcal{CG}$ is a chain graph consisting of a minimal complex, $\langle \alpha, \{b_1, b_2\} \beta \rangle$ then $\mathfrak{I}(\mathcal{CG})$ is not between-separated.*

Proof: It follows directly from the pairwise Markov property for chain graphs that any pair of non-adjacent vertices are separable in $\mathfrak{I}(\mathcal{CG})$. Conversely, any pair of adjacent vertices are inseparable. Hence, $[\mathfrak{I}(\mathcal{CG})]^{\text{Ins}} = (\mathcal{CG})^{\sim}$, in particular, $\text{Between}(\mathfrak{I}(\mathcal{CG}); b_1, \beta) = \{b_2\}$. However, although $b_1 \perp\!\!\!\perp \beta \mid \{\alpha, b_2\}$ in $\mathfrak{I}(\mathcal{CG})$, since $\{\alpha, b_2\} \cap \text{Between}(\mathfrak{I}(\mathcal{CG}); b_1, \beta) = \{b_2\}$, $b_1 \perp\!\!\!\perp \beta \mid \{\alpha, b_2\} \cap \text{Between}(\mathfrak{I}(\mathcal{CG}); b_1, \beta)$ is not in $\mathfrak{I}(\mathcal{CG})$. Thus $\mathfrak{I}(\mathcal{CG})$ is not between-separated. □

We are now in a position to prove Lemma 4.2

Proof: Suppose for a contradiction that there is a chain graph $\mathcal{CG}$ with vertex set V containing a minimal complex $\langle \alpha, \{b_1, \ldots, b_r\}, \beta \rangle$, with $r > 1$, and $\mathfrak{I}(\mathcal{D})[^S_L = \mathfrak{I}(\mathcal{CG})$, for some DAG $\mathcal{D}$ with vertex set $V \cup S \cup L$. Now let $S^* = an_{\mathcal{CG}}(\{b_1, \ldots, b_r\}) \setminus \{\alpha, \beta\}$ and $L^* = V \setminus (S^* \cup \{\alpha, \beta\} \cup \{b_1, b_r\})$. It now follows directly from the global Markov property for chain graphs that

$$\mathfrak{I}(\mathcal{CG})[^{S^*}_{L^*} = \mathfrak{I}(\mathcal{CG}^*)$$

where $\mathcal{CG}^*$ is the graph:

$$\alpha \to b_1 - b_r \leftarrow \beta.$$

By Proposition 4.1(ii),

$$\mathfrak{I}(\mathcal{D})[^{S \cup S^*}_{L \cup L^*} = \mathfrak{I}(\mathcal{CG})[^{S^*}_{L^*} = \mathfrak{I}(\mathcal{CG}^*).$$

But this is a contradiction, since, by Theorem 4.3, $\mathfrak{I}(\mathcal{D})[^{S \cup S^*}_{L \cup L^*}$ is between-separated, but by Lemma 4.4, $\mathfrak{I}(\mathcal{CG}^*)$ is not between-separated. □

4.3 Chain graphs violating condition (b)

In this subsection we prove the following Lemma:

Lemma 4.5 *If $\mathcal{CG}$ is a chain graph which satisfies condition (a), but not condition (b) in Theorem 3.2 then there is no DAG $\mathcal{D}$ such that*

$$\mathfrak{I}(\mathcal{D})[^{S}_{L} = \mathfrak{I}(\mathcal{CG})$$

In order to prove this Lemma we need the following:

Lemma 4.6 *If $x \perp\!\!\!\perp y \mid W$ is in $\mathfrak{I}(\mathcal{D})[^{S}_{L}$ for some set W, but x and z, and y and z are not separable in $\mathfrak{I}(\mathcal{D})[^{S}_{L}$, then $z \in W$ if and only if $z \in an_{\mathcal{D}}(\{x, y\} \cup S)$.*

Proof Sketch: Theorem 4.18 of Richardson and Spirtes (2000) shows that there exists a maximal ancestral graph $\mathcal{G}$ such that

$$\mathfrak{I}(\mathcal{D})[^{S}_{L} = \mathfrak{I}(\mathcal{G})$$

under the ancestral graph Markov property. (See Richardson and Spirtes (2000) for the relevant definitions.) Since x and z, and y and z are inseparable in $\mathfrak{I}(\mathcal{D})[^{S}_{L}$ it follows by Theorem 4.2 in Richardson and Spirtes (2000) that there are edges between x and z, and y and z in $\mathcal{G}$. Further, since $x \perp\!\!\!\perp y \mid W$ in $\mathfrak{I}(\mathcal{G})$, $\langle x, z, y \rangle$ forms a collider in $\mathcal{G}$ if and only if $z \notin W$. If $\langle x, z, y \rangle$ forms a collider in $\mathcal{G}$ then since there are arrowheads present at z in $\mathcal{G}$, it follows from Lemmas 4.7 and 4.9 of Richardson and Spirtes (2000), that $z \notin \mathrm{an}_{\mathcal{D}}(\{x, y\} \cup S)$. Conversely, if $\langle x, z, y \rangle$ forms a non-collider in $\mathcal{G}$ then again by Lemma 4.7 of Richardson and Spirtes (2000), $z \in \mathrm{an}_{\mathcal{D}}(\{x, y\} \cup S)$.□

This Lemma is the basis of the orientation rules used in the FCI algorithm described in Spirtes et al. (1993, 1995).

We are now ready to prove Lemma 4.5:

Proof: Let $\mathcal{CG}$ be a chain graph in which every minimal complex is an unshielded collider, but condition (b) is not satisfied, hence there is some unshielded collider $\langle x, z, y \rangle$, in which $z \in \mathrm{An}_{\mathcal{CG}}(c)$ for some vertex c in the collider free cycle. Suppose for a contradiction that there is some DAG $\mathcal{D}$ such that

$$\mathfrak{I}(\mathcal{D})[^S_L = \mathfrak{I}(\mathcal{CG}).$$

The proof is in two parts. We will first show that every vertex in a collider free chordless cycle in $\mathcal{CG}$ is an ancestor of S in $\mathcal{D}$. We will then show that there is an unshielded collider $\langle x, z^*, y\rangle$ in $\mathcal{CG}$ and that $z^* \in \mathrm{An}_{\mathcal{D}}(c)$. This leads directly to a contradiction, since it follows from Lemma 4.6 that $z^* \notin \mathrm{An}_{\mathcal{D}}(S)$.

Let the vertices on the collider free chordless cycle be $\langle c_1, \ldots, c_n\rangle$. By definition, c_i and c_j are adjacent in $\mathcal{CG}$ if and only if they occur consecutively in the cycle. Consequently, c_i and c_j are inseparable in $\mathfrak{I}(\mathcal{CG})$ if and only if they occur consecutively. Since the cycle is chordless, contains more than 3 edges, and is collider free, it further follows that if $j \equiv i + 2 \pmod{n}$ then $c_i \perp\!\!\!\perp c_j \mid W_i$ in $\mathfrak{I}(\mathcal{CG})$ for some set W_i containing the vertex that occurs between c_i and c_j on the cycle (more precisely, c_k for $k \equiv i + 1 \pmod{n}$). Consequently, by Lemma 4.6, in $\mathcal{D}$ each vertex in the cycle is either an ancestor of one of the vertices adjacent to it in the cycle, or an ancestor of a vertex in S.

Suppose for a contradiction that some vertex in the cycle is not an ancestor of S in $\mathcal{D}$. Without much loss of generality we may suppose $c_n \notin \mathrm{an}_{\mathcal{D}}(S)$. By Lemma 4.6, $c_n \in \mathrm{an}_{\mathcal{D}}(\{c_{n-1}, c_1\} \cup S)$. Suppose that $c_n \in \mathrm{an}_{\mathcal{D}}(\{c_1\})$. It now follows that $c_1 \notin \mathrm{an}_{\mathcal{D}}(S)$, since $c_n \notin \mathrm{an}_{\mathcal{D}}(S)$, and $c_1 \notin \mathrm{an}_{\mathcal{D}}(c_n)$, since otherwise there would be a directed cycle in $\mathcal{D}$. Since, by Lemma 4.6, $c_1 \in \mathrm{an}_{\mathcal{D}}(\{c_n, c_2\} \cup S)$, it follows that $c_1 \in \mathrm{an}_{\mathcal{D}}(\{c_2\})$. Arguing in the same way, it follows that $c_i \in \mathrm{an}_{\mathcal{D}}(\{c_{i+1}\})$ for $i = 1, \ldots, n-1$. However, this is a contradiction since c_1 is an ancestor of c_n and c_n is an ancestor of c_1. The argument in the case in which c_n is an ancestor of c_{n-1} is symmetric. This establishes the first claim.

We now show that there is an unshielded collider $\langle x, z^*, y\rangle$ in $\mathcal{CG}$, and $z^* \in \mathrm{An}_{\mathcal{D}}(c)$. Consider the unshielded collider $\langle x, z, y\rangle$. Since $z \in \mathrm{An}_{\mathcal{CG}}(c)$ for some vertex c in the collider free cycle in $\mathcal{CG}$, it follows that in $\mathcal{CG}$ either there is a path from z to c on which any directed edges point towards c, or $z = c$. If $z = c$ then let $z^* = z$ and we are done, hence suppose that $z \neq c$, and let $\boldsymbol{\pi} = \langle z = z_1, \ldots, z_m = c\rangle$ be a shortest path from z to c in $\mathcal{CG}$. Let z_j be the vertex furthest from z on the path $\boldsymbol{\pi}$ which is adjacent to both x and y (such a vertex is guaranteed to exist since z is adjacent to both x and y). It follows that $x \to z_j \leftarrow y$ in $\mathcal{CG}$ since otherwise there would be a partially directed cycle. Again, if $j = m$ then $z^* = z_j = c$ and again we are done. If $j < m$ then by construction z_{j+1} is not adjacent to both x and y. Suppose without loss that x and z_{j+1} are not adjacent (the other case is symmetrical). Since we have $x \rightarrow z_j \leftarrow y$ in $\mathcal{CG}$, it follows from Lemma 4.6 that $z_j \notin \text{an}_{\mathcal{D}}(\{x, y\} \cup S)$, since $x \perp\!\!\!\perp y \mid W$ in $\mathfrak{I}(\mathcal{CG})$ for some W with $z_j \notin W$. At the same time, $x \rightarrow z_j \rightarrow z_{j+1}$, and x and z_{j+1} are not adjacent, so by Lemma 4.6, $z_j \in \text{an}_{\mathcal{D}}(\{x, z_{j+1}\} \cup S)$. Consequently, $z_j \in \text{an}_{\mathcal{D}}(\{z_{j+1}\})$ and $z_{j+1} \notin \text{an}_{\mathcal{D}}(\{z_j\} \cup S)$. Since $\boldsymbol{\pi}$ is a shortest path, it follows that the only vertices on the path which are adjacent in $\mathcal{CG}$ are those that occur consecutively on the path; were any other vertices adjacent a shorter path could be formed. Consequently, by repeated applications of Lemma 4.6 it follows that $z_k \in \text{an}_{\mathcal{D}}(\{z_{k+1}\})$ and $z_k \notin \text{an}_{\mathcal{D}}(\{z_{k-1}\} \cup S)$, for $j < k < m$. Hence $z_j \in \text{an}_{\mathcal{D}}(\{z_m\})$ and so letting $z^* = z_j$ we are done.

Finally, since $\langle x, z^*, y \rangle$ forms an unshielded collider in $\mathcal{CG}$, by Lemma 4.6 $z^* \notin \text{An}_{\mathcal{D}}(S)$. This is a contradiction since we have shown that $z^* \in \text{An}_{\mathcal{D}}(c)$, and $c \in \text{an}_{\mathcal{D}}(S)$. □

4.4 Recursive causal graphs and DAGs under marginalizing and conditioning

In Theorem 6.3 of Richardson and Spirtes (2000) it is proved that for any ancestral graph $\mathcal{G}$ with vertex set V, there exists an associated DAG, $\mathcal{D}(\mathcal{G})$, with vertex set $V \cup S \cup L$ such that

$$\mathfrak{I}(\mathcal{G}) = \mathfrak{I}(\mathcal{D}(\mathcal{G}))[^{S_\mathcal{G}}_{L_\mathcal{G}}.$$

Since recursive causal graphs are a subclass of ancestral graphs, the following result is an immediate consequence:

Theorem 4.7 *If $\mathcal{CG}$ is Markov equivalent to a recursive causal graph then there exists a DAG $\mathcal{D}$ such that*

$$\mathfrak{I}(\mathcal{G}) = \mathfrak{I}(\mathcal{D})[^{S}_{L}.$$

In the case of a recursive causal graph, the construction of $\mathcal{D}(\mathcal{G})$ takes a simple form: each undirected edge $x - y$ is replaced with a common child, $x \rightarrow \sigma_{xy} \leftarrow y$, S contains all of these vertices σ_{xy}, and $L = \emptyset$.

4.5 Characterization of chain graphs which are Markov equivalent to a DAG under marginalization and conditioning

We are now able to establish the main result of this note:

Theorem 4.8 *If $\mathcal{CG}$ is a chain graph, then the following three conditions are equivalent:*

(i) *$\mathcal{CG}$ is Markov equivalent to a recursive causal graph;*

(ii) *there exists a DAG $\mathcal{D}$ such that*

$$\mathfrak{I}(\mathcal{CG}) = \mathfrak{I}(\mathcal{D})[^{S}_{L};$$

(iii) *the chain graph $\mathcal{CG}$ satisfies:*

(a) *all minimal complexes in $\mathcal{CG}$ are colliders, and*

(b) $(\mathcal{CG}_A)^m = (\mathcal{CG}_A)^\sim$, *where* $A = \mathrm{An}(\mathrm{CF}_{\mathcal{CG}})$.

Proof: Theorem 3.2 establishes that (iii) $\Rightarrow$ (i). Theorem 4.7 establishes that (i) $\Rightarrow$ (ii). Finally, Lemmas 4.2 and 4.5 establish that (ii) $\Rightarrow$ (iii), by contraposition. □

5 Discussion

Note that recursive causal graphs lead to the same independence model under the alternative Markov property, given by Andersson et al. (1996), as under the original Markov property considered here. Hence it remains true under the alternative property that a chain graph is Markov equivalent to a DAG under marginalization and conditioning if it is Markov equivalent to a recursive causal DAG. However, this condition is no longer necessary under the alternative Markov property. For example, any chain graph in which all chain components are cliques is Markov equivalent, under the AMP property, to some DAG under marginalizing and conditioning.

References

Andersson, S. A., D. Madigan, and M. D. Perlman (1996). An alternative Markov property for chain graphs. In F. V. Jensen and E. Horvitz (Eds.), *Uncertainty in Artificial Intelligence: Proceedings of the* 12^{th} *Conference*, San Francisco, pp. 40–48. Morgan Kaufmann.

Frydenberg, M. (1990). The chain graph Markov property. *Scandin. J. Statist. 17*, 333–353.

Kiiveri, H., T. Speed, and J. Carlin (1984). Recursive causal models. *Journal of the Australian Mathematical Society, Ser. A 36*, 30–52.

Lauritzen, S. (1996). *Graphical Models.* Number 17 in Oxford Statistical Science Series. Oxford, UK: Clarendon Press.

Richardson, T. (1998). Chain graphs and symmetric associations. In M. Jordan (Ed.), *Learning in Graphical Models*, Number 89 in NATO Science Series D, pp. 231–260. Kluwer.

Richardson, T. and P. Spirtes (2000, June). Ancestral graph Markov models. Technical Report 375, Department of Statistics, University of Washington.

Spirtes, P., C. Glymour, and R. Scheines (1993). *Causation, Prediction and Search.* Number 81 in Lecture Notes in Statistics. Springer-Verlag.

Spirtes, P., C. Meek, and T. Richardson (1995). Causal inference in the presence of latent variables and selection bias. In P. Besnard and S. Hanks (Eds.), *Uncertainty in Artificial Intelligence: Proceedings of the* 11^{th} *Conference*, San Francisco, pp. 403–410. Morgan Kaufmann.