Chain Graph Interpretations and their Relations

Dag Sonntag and Jose M. Peña

ADIT, IDA, Linköping University, Sweden
dag.sonntag@liu.se, jose.m.pena@liu.se

Abstract. This paper deals with different chain graph interpretations and the relations between them in terms of representable independence models. Specifically, we study the Lauritzen-Wermuth-Frydenberg, Andersson-Madigan-Pearlman and multivariate regression interpretations and present the necessary and sufficient conditions for when a chain graph of one interpretation can be perfectly translated into a chain graph of another interpretation. Moreover, we also present a feasible split for the Andersson-Madigan-Pearlman interpretation with similar features as the feasible splits presented for the other two interpretations.
Keywords: Chain Graphs, Lauritzen-Wermuth-Frydenberg interpretation, Andersson-Madigan-Pearlman interpretation, multivariate regression interpretation.

1 Introduction

Today there exist mainly three interpretations of chain graphs (CGs). These are the Lauritzen-Wermuth-Frydenberg (LWF) interpretation presented by Lauritzen, Wermuth and Frydenberg in the late eighties [6, 7], the Andersson-Madigan-Pearlman (AMP) interpretation presented by Anderson, Madigan and Pearlman in 2001 [2] and the multivariate regression (MVR) interpretation presented by Cox and Wermuth in the nineties [3, 4]. A fourth interpretation of CGs can also be found in a study by Drton [5] but this interpretation has not been further studied and will not be discussed in this paper.

Each interpretation has a different separation criterion and do therefore represent different independence models. So far most papers have studied the different interpretations independently with a few exceptions such as the study of discrete CG models by Drton [5] and the study of CGs representing Gaussian distributions by Wermuth et al. [12]. Therefore it has not really been studied what differences and similarities that exist between the different interpretations in terms of representable independence models. Andersson et al. made a small study of this when they presented their new (AMP) interpretation and managed to show when the independence model of a CG of the AMP interpretation could be represented perfectly by a CG of the LWF interpretation. They did however not show when the opposite held and did no comparison with CGs of the MVR interpretation. Wermuth and Sadeghi did on the other hand present conditions for when a CG of the MVR interpretation could be translated into a CG of the LWF or AMP interpretation when they introduced regression graphs [11]. The conditions were however only necessary and sufficient if the two CGs contained the same connectivity components and not the more general case where the CGs could take any form.

In this paper we hope to fill this gap and hence the main contribution of this paper is a table where we show the necessary and sufficient conditions for when a CG of one interpretation can be perfectly translated into a CG of another interpretation. First we do however define a feasible split for the AMP interpretation, with similar features as the feasible splits shown for the LWF [10] and MVR [9] interpretation, that are used in these conditions. Hence this is our second contribution. Finally we also show that there for all three CG interpretations exists a minimal set of non-directed edges for each Markov equivalence class and that the CG containing these, and only these, non-directed edges can be reached through repeated feasible splits from any member of the class.

The remainder of the article is organized as follows. In the next section we present the notation we will use throughout the article. This is followed by the definitions of the feasible splits for each interpretation as well as the proof that the feasible split for CGs of the AMP interpretation is sound. In section 4 we start by presenting the conditions of when a CG of one interpretation can be perfectly represented by a CG of another interpretation. This is then followed by the proofs that these conditions are sound.

2 Notation

All graphs are defined over a finite set of variables V. If a graph G contains an edge between two nodes V_1 and V_2, we denote with $V_1 \rightarrow V_2$ a *directed edge*, with $V_1 \leftrightarrow V_2$ a *bidirected edge* and with $V_1 - V_2$ an *undirected edge*. By $V_1 \circ\!\!\rightarrow V_2$ we mean that either $V_1 \rightarrow V_2$ or $V_1 \leftrightarrow V_2$ is in G. By $V_1 -\!\!\circ V_2$ we mean that either $V_1 \rightarrow V_2$ or $V_1 - V_2$ is in G. By $V_1 \circ\!\!-\!\!\circ V_2$ we mean that there exists an edge between V_1 and V_2 in G while we with $V_1 \cdots V_2$ mean that there might or might not exist an edge between V_1 and V_2. By a *non-directed* edge we mean either a bidirected edge or an undirected edge. A set of nodes is said to be *complete* if there exist edges between all pairs of nodes in the set.

The *parents* of a set of nodes X of G is the set $pa_G(X) = \{V_1 | V_1 \rightarrow V_2$ is in G, $V_1 \notin X$ and $V_2 \in X\}$. The *children* of X is the set $ch_G(X) = \{V_1 | V_2 \rightarrow V_1$ is in G, $V_1 \notin X$ and $V_2 \in X\}$. The *spouses* of X is the set $sp_G(X) = \{V_1 | V_1 \leftrightarrow V_2$ is in G, $V_1 \notin X$ and $V_2 \in X\}$. The *neighbours* of X is the set $nb_G(X) = \{V_1 | V_1 - V_2$ is in G, $V_1 \notin X$ and $V_2 \in X\}$. The *boundary* of X is the set $bd_G(X) = pa_G(X) \cup nb_G(X) \cup sp_G(X)$. The *adjacents* of X is the set $ad_G(X) = \{V_1 | V_1 \rightarrow V_2, V_1 \leftarrow V_2, V_1 \leftrightarrow V_2$ or $V_1 - V_2$ is in G, $V_1 \notin X$ and $V_2 \in X\}$.

A *route* from a node V_1 to a node V_n in G is a sequence of nodes $V_1, \ldots, V_n$ such that $V_i \in ad_G(V_{i+1})$ for all $1 \leq i < n$. A *path* is a route containing only distinct nodes. The length of a path is the number of edges in the path. A path is called a *cycle* if $V_n = V_1$. A path is *descending* if $V_i \in pa_G(V_{i+1}) \cup sp_G(V_{i+1}) \cup nb_G(V_{i+1})$ for all $1 \leq i < n$. A path $\pi = V_1, \ldots, V_n$ is *minimal* if there exists no other path π_2 between V_1 and V_n st $\pi_2 \subset \pi$ holds. The *descendants* of a set of nodes X of G is the set $de_G(X) = \{V_n |$ there is a descending path from V_1 to V_n in G, $V_1 \in X$ and $V_n \notin X\}$. A path is *strictly descending* if $V_i \in pa_G(V_{i+1})$ for all $1 \leq i < n$. The *strict descendants* of a set of nodes X of G is the set $sde_G(X) = \{V_n |$ there is a strict descending path from V_1 to V_n in G, $V_1 \in X$ and $V_n \notin X\}$. The *ancestors* (resp. *strict ancestors*) of X is the set $an_G(X) = \{V_1 | V_n \in de_G(V_1), V_1 \notin X, V_n \in X\}$ (resp. $san_G(X) = \{V_1 | V_n \in sde_G(V_1), V_1 \notin X, V_n \in X\}$). A cycle is called a *semi-directed cycle* if it is descending and $V_i \rightarrow V_{i+1}$ is in G for some $1 \leq i < n$. A CG under the Lauritzen-Wermuth-Frydenberg (LWF) interpretation, denoted LWF CG, contains only directed and undirected edges but no semi-directed cycles. Likewise a CG under the Andersson-Madigan-Perlman (AMP) interpretation, denoted AMP CG, is a graph containing only directed and undirected edges but no semi-directed cycles. A CG under the multivariate regression (MVR) interpretation, denoted MVR CG, is a graph containing only directed and bidirected edges but no semi-directed cycles. A *connectivity component* C of a LWF CG or an AMP CG (resp. MVR CG) is a maximal (wrt set inclusion) set of nodes such that there exists a path between every pair of nodes in C containing only undirected edges (resp. bidirected edges). We denote the set of all connectivity components in a CG G by $cc(G)$ and the component to which a set of nodes X belong in G by $co_G(X)$. A *subgraph* of G is a subset of nodes and edges in G. A subgraph of G induced by a set of its nodes X is the graph over X that has all and only the edges in G whose both ends are in X. A *bidirected flag* is an induced subgraph of the form $X \leftrightarrow Y \leftrightarrow Z$ in a MVR CG. With the moral closure graph of a component C in a LWF CG G, denoted $(G_{cl(C)})^m$, we mean the subgraph of G induced by $C \cup pa_G(C)$ where every edge have been made undirected and every pair of nodes in $pa_G(C)$ have been made adjacent with undirected edges.

Let X, Y and Z denote three disjoint subsets of V. We say that X *separated* from Y given Z denoted as $X \perp_G Y | Z$ if the following criteria is met: If G is a LWF CG then X and Y are separated given Z iff there exists no route between X and Y such that every node in a non-collider section on the route is not in Z and some node in every collider section on the route is in Z. A *section* of a route is a maximal (wrt set inclusion) non-empty set of nodes $B_1 ... B_n$ such that the route contains the subpath $B_1 - B_2 - \ldots - B_n$. It is called a *collider section* if $B_1 \ldots B_n$ together with the two neighbouring nodes in the route, A and C, form the subpath $A \rightarrow B_1 - B_2 - \ldots - B_n \leftarrow C$. For any other configuration the section is a non-collider section. If G is an AMP CG then X and Y is separated given Z iff there exists no S-open path between X and Y. A path is said to be *S-open* iff every non-head-no-tail node on the path is not in Z and every head-no-tail node on the path is in Z or $san_G(Z)$. A node B is said to be a *head-no-tail* in an AMP CG G between two nodes A and C on a path if one of the following configurations exists in G: $A \rightarrow B \leftarrow C$, $A \rightarrow B - C$ or $A - B \leftarrow C$. Moreover G is also said to contain a triplex $(\{A, C\}, B)$ iff one such configuration exists in G and A and C are not adjacent in G. For any other configuration the node B is a non-collider. If G is a MVR CG then X and Y are separated given Z iff there exists no d-connecting path between X and Y. A path is said to be *d-connecting* iff every non-collider on the path is not in Z and every collider on the path is in Z or $san_G(Z)$. A node B is said to be a *collider* in a MVR CG G between two nodes A and C on a path if one of the following configurations exists in G: $A{\rightarrow}B{\leftarrow}C$, $A{\rightarrow}B{\leftrightarrow}C$,$A{\leftrightarrow}B{\leftarrow}C$ or $A{\leftrightarrow}B{\leftrightarrow}C$. For any other configuration the node B is a non-collider.

The *independence model* M induced by a graph G, denoted as $I(G)$ or $I_{PGM-class}(G)$, is the set of separation statements $X{\perp}_G Y|Z$ that hold in G according to the interpretation to which G belongs or the subscripted PGM-class. We say that two graphs G and H are *Markov equivalent* (under the same interpretation) or that they are in the same *Markov equivalence class* iff $I(G) = I(H)$.

3 Feasible Splits

For the LWF and MVR interpretation, operations for altering a CG structure without changing its Markov equivalence class have been presented [9, 10]. One such operation is called feasible split and is in this article used to prove certain theorems. Hence we repeat the definitions here. Moreover, we also present the corresponding operation, called feasible split for AMP CGs, for the AMP CG interpretation and prove that it is sound. Note that this is not the inverse operation to a legal merging presented in the deflagging procedure for AMP CGs by Roverto and Studený [8]. Their operation was applied to so called strong equivalence classes, not the more general Markov equivalence classes used here.

Definition 1. *Feasible split for LWF CGs [10]*
A connectivity component C of CG G under the LWF interpretation can be feasibly split into two disjoint sets U and L st $U \cup L = C$ by replacing every undirected edge between U and L with a directed edge orientated towards L iff:
1. $\forall A \in ne_G(L) \cap U, pa_G(L) \subseteq pa_G(A)$
2. $ne_G(L) \cap U$ is complete

Definition 2. *Feasible split for AMP CGs*
A connectivity component C of CG G under the AMP interpretation can be feasibly split into two disjoint sets U and L st $U \cup L = C$ by replacing every undirected edge between U and L with a directed edge orientated towards L iff:
1. $\forall A \in ne_G(L) \cap U, L \subseteq ne_G(A)$
2. $ne_G(L) \cap U$ is complete
3. $\forall B \in L, pa_G(ne_G(L) \cap U) \subseteq pa_G(B)$

Definition 3. *Feasible split for MVR CGs [9]*
A connectivity component C of CG G under the MVR interpretation can be feasible split into two disjoint sets U and L st $U \cup L = C$ by replacing every bidirected edge between U and L with a directed edge orientated towards L iff:
1. $\forall A \in sp_G(U) \cap L$, $U \subseteq sp_G(A)$ holds
2. $\forall A \in sp_G(U) \cap L$, $pa_G(U) \subseteq pa_G(A)$ holds
3. $\forall B \in sp_G(L) \cap U$, $sp_G(B) \cap L$ is a complete set

---

Definition 4. Maximally orientated CG
A CG G (under any interpretation) is maximally orientated iff no feasible splits can be performed on G.

Lemma 1. *A CG G of the AMP interpretation is in the same Markov equivalence class before and after a feasible split.*

Proof. Assume the contrary. Let G be a CG under the AMP interpretations and G' a graph st G' is G with a feasible split performed upon it. G and G' are in different Markov equivalence classes or G' is not a CG under the AMP interpretation iff (1) G and G' does not have the same adjacencies, (2) G and G' does not have the same triplexes or (3) G' contains semi-directed cycles.

First it is clear that G and G' contains the same adjacencies since a feasible split does not change the adjacencies of any node in G. Secondly let us assume G and G' does not have the same triplexes. First let us assume that G' contains a triplex $(\{X, Y\}, Z$ that does not exist in G. It is clear that such a triplex can only occur if $Z \in L$ since the only difference between G and G' is that G' contains some directed edges orientated towards L where G contains undirected edges. It is clear that if the triplex is a flag then the one of the node X or Y, let's say X, must be in U and the other one, let's say Y, must be in L. However, according to condition 1 Y must be adjacent to X which causes a contradiction. If the triplex is not a flag both X and Y must be in U. They also have to be in $ne_G(L)$, which, together with condition 2, contradicts that they are not adjacent. Hence we have a contradiction for that G' contains a triplex that does not exist in G.

Secondly assume G contains a triplex $(\{X, Y\}, Z)$ that does not exist in G'. It is clear that this new triplex cannot be over a node in L since these nodes only have edges orientated towards them. Instead assume $Z \in U$. This gives that one of the nodes X or Y, let's say X, must be a parent of Z and the other, let's say Y, must be in L. This does however contradict condition 3, since every parent of Z also must be a parent of Y, and hence X and Y must be adjacent. This gives us a contradiction.

Finally assume G' contain a semi-directed cycle. This means there exists two nodes X and Y st $X \in pa_{G'}(Y)$ but $X \in de_{G'}(Y) \cup co_{G'}(Y)$. It is clear that $\forall A \in V$ $de_{G'}(A) \subseteq de_G(A)$ and $co_{G'}(A) \subseteq co_G(A)$ hold. Hence we must have that $X \in de_G(Y) \cup co_G(Y)$ also hold which, together with $\forall B \in V \setminus L$ $pa_{G'}(B) = pa_G(B)$, means that Y is in L and since $\forall D \in L$ $pa_{G'}(D) = pa_G(D) \cup U$ holds X must be in U. However, at the same time $co_{G'}(Y) = co_G Y \setminus U$ and $de_{G'}(Y) \subseteq de_G Y$ must hold and hence we have a contradiction.

A maximally orientated CG can be obtained from any member of its Markov equivalence class by performing feasible splits until no more feasible splits can be performed.

Theorem 1. *A CG (under any interpretation) has the minimal set of non-directed edges for its Markov equivalence class if no feasible split is possible.*

The following theorem shows that there may exist several maximally orientated CGs in a given Markov equivalence class but all of them share the same non-directed edges.

Theorem 2. *For any Markov equivalence class of CGs (under any interpretation), there exists a unique minimal (wrt inclusion) set of non-directed edges that is shared by all members of the class.*

The proofs of the Theorem 1 and 2 for the MVR interpretation can be found in the article by Sonntag and Peña [9]. These proofs can easily be adapted for the LWF and AMP interpretations.

4 Translations between Interpretations

In this section the main result of this paper is presented, namely what the conditions are for a CG of one interpretation to be possible to translate into a CG of another interpretation. With translate we mean that the induced independence model of a CG of one interpretation can be represented perfectly by a CG of another interpretation. A summary of these results is presented in Table 1.

	LWF	AMP	MVR
LWF	-	Unidentified	$(G_{cl(K)})^m$ is chordal for all $K \in cc(G)$.
AMP	G contains no k-biflag where $k \geq 2$ [2]	-	G' does not contain any induced subgraph of the form $X - Y - Z$
MVR	G' contains no bidirected edge	G' contains no bidirected flag	-

Table 1: Given a CG G of the interpretation denoted in the row, and a maximally oriented CG G' in the Markov equivalence class of G, there exists a CG H of the interpretation denoted in the column st G and H are Markov equivalent iff the condition in the intersecting cell is fulfilled.

From the table two things can be noted. First that the conditions given in the table may include a maximally oriented CG G' in the same equivalence class as G. This is done for several reasons. First, such a graph is easy and computationally simple to find. Secondly, this allows the proofs to be based on the idea that no feasible split is possible for the interpretation in mind. Third and last, the search space of CGs is smaller and more assumptions can be made on the CG. This in turn allows for more efficient algorithms when calculating if the condition holds for some CG. The second note that can be made is that there still does not exist any necessary and sufficient condition for when a perfect translation of a LWF CG G into an AMP CG H is possible. Andersson et al. gave a necessary condition but also showed that this condition was not sufficient [2]. We have managed to prove the necessity of more elaborate conditions but still been unable to prove sufficiency for these. Hence this condition is left for future work.

The rest of this section contains the theorems stating the conditions shown in Table 1 together with their proofs.

4.1 Translation of MVR CGs to AMP CGs

Theorem 3. *Given a MVR CG G, and a maximally oriented MVR CG G' in the Markov equivalence class of G, there exists an AMP CG H st $I_{MVR}(G) = I_{AMP}(H)$ iff G' contains no bidirected flag.*

Proof. Sufficiency follows from from Lemmas 4 and 5 and necessity follows from Lemma 2.

Lemma 2. *A MVR CG G and an AMP CG H with the same structure, except that every bidirected edge in G is replaced by a undirected edge in H and where G contains no bidirected flag, represent the same independence model.*

Proof. Assume to contrary that there exists two CGs, G under the MVR interpretation and H under the AMP interpretation, st G does not contain any bidirected flag, i.e induced subgraph of the form $X \leftrightarrow Y \leftrightarrow Z$, G and H contain the same directed edges, and for every bidirected edge in G H has an undirected edge instead (and only contains those undirected edges) but $I_{MVR}(G) \neq I_{AMP}(H)$. Clearly we must have $V_G = V_H$ and that $adj_G(X) = adj_H(X), pa_G(X) = pa_H(X)$ and $co_G(X) = co_H(X)$ holds for all $X \in V_G$. Given the definition of strict descendants $san_G(X) = san_H(X)$ must also hold. Moreover note that H cannot contain any induced subgraph of the form $X - Y - Z$. Finally note that both G and H contains the same paths between X and Y.

For $I(G) \neq I(H)$ to hold there has to exist a path π in G (resp. H) that is d-connecting (resp. S-open) st there exist no path in H (resp. G) that is S-open (resp. d-connecting). Let π be a minimal d-connecting (resp. S-open) path in G (resp. H). Note that π cannot contain any contain any subpath of the form $V_1 \leftrightarrow V_2 \leftrightarrow V_3$ (resp. $V_1 - V_2 - V_3$) since the edge $V_1 \leftrightarrow V_3$ (resp. $V_1 - V_3$) must exist in G (resp. H) or G contains a bidirected flag or semi-directed cycle. This in turn would mean that π is not minimal since the path $\pi \setminus V_2$ also must be d-connecting and shorter than π. For π to be both d-connecting and S-open for any set of nodes Z it must contain the same colliders and head-no-tail nodes. A node $W \in \pi$ is a collider if it is part of the following configurations of edges in π (1) $\rightarrow W \leftarrow$, (2) $\leftrightarrow W \leftarrow$, (3) $\rightarrow W \leftrightarrow$ and (4) $\leftrightarrow W \leftrightarrow$. Clearly the fourth case cannot occur. Case 1-3 would be translated into (1) $\rightarrow W \leftarrow$, (2) $- W \leftarrow$, (3)$\rightarrow W -$ in H which are all (and the only) head-no-tail configurations. Hence π must be d-connecting in G iff π is S-open in H which contradicts the assumption.

Lemma 3. *If a maximally oriented CG G of the MVR interpretation contains a bidirected flag $X \leftrightarrow Y \leftrightarrow Z$ then G also contains an induced subgraph of the form shown in (1) Figure 1a or (2) 1b or (3) $P \multimap Q \leftrightarrow Y \leftrightarrow Z$ or (4) $P \multimap Q \leftrightarrow W \leftrightarrow Z$ st $bd_G(Q) \subseteq bd_G(Y) \cup Y$ and $Y \in sp_G(Q)$ hold.*

Proof. Assume the contrary, that no such induced subgraph exists in G even though G contains a bidirected flag and G is maximally orientated. Let C be the component of which X, Y and Z belongs. Let A be the set of nodes A_k st $A_k \in sp_G(Y)$ but $A_k \notin sp_G(Z)$. We know that X fulfills these criteria and hence $|A| \geq 1$.

First note that if there exists a node $A_k \in A$ st $bd_G(A_k) \nsubseteq bd_G(Y) \cup Y$ then there exists an induced subgraph $P \multimap A_k \leftrightarrow Y \leftrightarrow Z \cdots P$ in G for some node $P \in bd_G(A_k) \setminus bd_G(Y) \setminus Y$. Hence we have a contradiction since G either contains an induced subgraph of the form shown in Figure 1b ($P \in bd_G(Z)$) or of the form $P \multimap Q \leftrightarrow Y \leftrightarrow Z$ ($P \notin bd_G(Z)$). Therefore we must have that $bd_G(A_k) \subseteq Y \cup bd_G(Y)$ holds for all $A_k \in A$, i.e. that $bd_G(A) \subseteq Y \cup bd_G(Y)$ holds.

Secondly note that we can let B be a subset of A st B consists of the nodes in one connected subgraph in the subgraph of G induced by A (any connected subgraph will do). Let D be the set of nodes st $D = sp_G(Y) \cap sp_G(Z) \cap sp_G(B)$. With these sets we know that the spouses of Y can be either adjacent of Z or not, hence $sp_G(Y) = D \cup A$ must hold. This in turn gives that $sp_G(A) = D \cup Y$ and $bd_G(A) \subseteq D \cup Y \cup pa_G(Y)$ since $\forall A_k \in A$ $bd_G(A_k) \subseteq Y \cup bd_G(Y)$ holds. Moreover $sp_G(B) = D \cup Y$ and $bd_G(B) \subseteq D \cup Y \cup pa_G(Y)$ must also hold. Hence, if D is empty then $sp_G(B) = \{Y\}$ and $bd_G(B) \subseteq Y \cup pa_G(Y)$ must hold. This does however lead to a contradiction because a split then is possible st U consists of B and L consists of $C \setminus U$. Hence there has to exists at least one node in D.

Thirdly note that $D \cup Y$ must be complete or the induced subpath $B_k \leftrightarrow DY_i \leftrightarrow Z \leftrightarrow DY_j \leftrightarrow B_1 \leftrightarrow ... \leftrightarrow B_l \leftrightarrow B_k$, $l \geq 0$, exists in G for some nodes $B_k, B_1, ..., B_l \in B$ and $DY_i, DY_j \in D \cup Y$. This means that G contains an induced subgraph of the form shown in either Figure 1a ($l > 0$) or 1b ($l = 0$).

Fourth and finally note that there must exist a node P st $P \in bd_G(B) \cup B$ but $P \notin bd_G(D_j)$ for some $D_j \in D \cup Y$ or a split is feasible where U consists of B and L consists of $C \setminus U$. Note that $D_j \neq Y$ must hold since $bd_G(B) \cup B \subseteq bd_G(Y) \cup Y$. This means that there must exist 2 nodes B_i, D_j st $P \in bd_G(B_i), P \notin bd_G(D_j), B_i \in B$, $B_i \in sp(D_j)$ and $D_j \in D$ st the induced subgraph $P \multimap B_i \leftrightarrow D_j \leftrightarrow Z \cdots P$ exist in G. This is a contradiction either because G contains an induced subgraph of the form shown in Figure 1b ($P \in bd_G(Z)$) or $P \multimap B_i \leftrightarrow D_j \leftrightarrow Z$ ($P \notin bd_G(Z)$) where $bd_G(B_i) \subseteq bd_G(Y) \cup Y$ and $Y \in sp_G(B_i)$ holds.

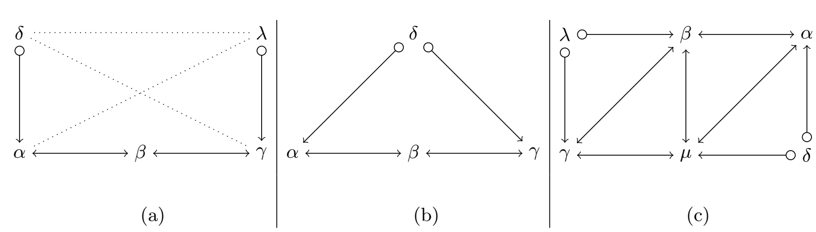

Fig. 1: MVR subgraph forms

Lemma 4. *If a maximally oriented CG G of the MVR interpretation contains a bidirected flag then G at least one of the induced subgraphs shown in Figure 1 exists in G.*

Proof. Assume the contrary, that no such induced subgraph exists in G even though G contains a bidirected flag and G is maximally orientated. Since G contains a bidirected flag we do with Lemma 3 get that G must contain an induced subgraph $X \leftrightarrow Y \leftrightarrow Z \leftarrow\!\circ W$ or a contradiction directly follows. If we now apply Lemma 3 to $X \leftrightarrow Y \leftrightarrow Z$ we get that, since for G to contain any induced subgraph of the form shown in Figure 1a or 1b is a contradiction, there exist a set of nodes (that can be renamed to) c_1, c_2, c_3 st the induced subgraph $c_1 \circ\!\!\rightarrow c_2 \leftrightarrow c_3 \leftrightarrow Z$ exists in G and $c_3 = Y$ holds or $bd_G(c_2) \subseteq bd_G(Y) \cup Y$ and $Y \in sp_G(c_2)$ hold. If $c_3 = Y$, G must contain the subgraph $c_1 \circ\!\!\rightarrow c_2 \leftrightarrow Y \leftrightarrow Z \leftarrow\!\circ W$ where $c_1 \notin adj_G(Y)$ and $W \notin adj_G(Y)$ must hold and $c_1 = W$ might hold. Clearly this subgraph takes the form of either Figure 1a ($c_1 \neq W$) or 1b ($c_1 = W$) which is a contradiction. Hence $c_3 \neq Y$, $bd_G(c_2) \subseteq bd_G(Y) \cup Y$ and $Y \in sp_G(c_2)$ must hold.

Since $W \notin adj_G(Y)$ holds and $bd_G(c_2) \subseteq bd_G(Y) \cup Y$ it is clear that $c_1, c_3 \in bd_G(Y)$ must hold. Hence $W \neq c_2$ holds since $W \notin adj_G(Y) \cup Y$. This in turn means that $W \notin bd_G(c_2)$ holds since $bd_G(c_2) \subseteq bd_G(Y) \cup Y$ and $W \notin bd_G(Y) \cup Y$. Finally we can see that $W \in bd_G(c_3)$ holds or the induced subgraph $c_1 \circ\!\!\rightarrow c_2 \leftrightarrow c_3 \leftrightarrow Z \leftarrow\!\circ W$ takes the form shown in Figure 1a ($c_1 \neq W$) or 1b ($c_1 = W$). However, if $W \in bd_G(c_3)$ holds G contains an induced subgraph of the form shown in Figure 1c (where $\delta = W$, $\lambda = c_1$, $\mu = c_3$, $\gamma = c_2$, $\beta = Y$ and $\alpha = Z$) and we have a a contradiction.

Lemma 5. *The independence model of a CG G of the MVR interpretation which contains an induced subgraph of one of the forms shown in Figure 1 cannot be perfectly represented as a CG H of the AMP interpretation.*

Proof. Assume the contrary, that there exists a CG H under the AMP interpretation that can represent these independence models.

First assume that the independence model of the graph shown in Figure 1a can be represented in a CG H of the AMP interpretation. It is clear that H must have the same skeleton, or clearly some separations or non-separations that hold in G would not hold in H. The following independence statements holds in G: $\delta \perp_G \beta | pa_G(\beta)$, $\alpha \perp_G \gamma | pa_G(\alpha)$ and $\beta \perp_G \lambda | pa_G(\beta)$. $\delta \perp_G \beta | pa_G(\beta)$ gives us that a triplex $(\{\delta, \beta\}, \alpha)$ must exist in H, since $\alpha \notin pa_G(\beta)$ i.e. that (1) $\delta \rightarrow \alpha - \beta$, (2) $\delta - \alpha \leftarrow \beta$ or (3) $\delta \rightarrow \alpha \leftarrow \beta$ exists in H. $\alpha \perp_G \gamma | pa_G(\alpha)$ does however also state that a triplex $(\{\alpha, \gamma\}, \beta)$ must exist in H, since $\beta \notin pa_G(\alpha)$. For this to happen the edge between α and β cannot be orientated towards α hence the subgraph $\delta \rightarrow \alpha - \beta \leftarrow \gamma$ must exist in H. The orientation of the edge between β and γ does however contradict the third independence statement $\beta \perp_G \lambda | pa_G(\beta)$ which implies that the triplex $(\{\beta, \lambda\}, \gamma)$ must exist in H, since $\gamma \notin pa_G(\beta)$. Hence we have a contradiction if G contains the induced subgraph shown in Figure 1a.

Secondly assume that the independence model of the graph shown in Figure 1b can be represented in a CG H of the AMP interpretation. It is clear that H must have the same skeleton, or clearly some separations or non-separations that hold in G would not hold in H. The following independence statements must then hold in G: $\delta \perp_G \beta | pa_G(\beta)$ and $\alpha \perp_G \gamma | pa_G(\alpha)$. $\delta \perp_G \beta | pa_G(\beta)$ gives us that two triplexes must exist in H, first $(\{\delta, \beta\}, \alpha)$ and secondly $(\{\delta, \beta\}, \gamma)$, since $\alpha, \gamma \notin pa_G(\beta)$. $(\{\delta, \beta\}, \alpha)$ gives that one of the following configurations must occure in H: (1) $\delta{-}\alpha{\leftarrow}\beta$, (2) $\delta{\rightarrow}\alpha{-}\beta$ or (3) $\delta{\rightarrow}\alpha{\leftarrow}\beta$. However, the independence statement $\alpha \perp_G \gamma | pa_G(\alpha)$ implies that the triplex $(\{\alpha, \gamma\}, \beta)$ must exist in H since $\beta \notin pa_G(\alpha)$. If the triplex $(\{\alpha, \gamma\}, \beta)$ should hold in H the edge between α and β cannot be orientated towards α hence the subgraph $\delta{\rightarrow}\alpha{-}\beta{\leftarrow}\gamma$ must exist in H. The orientation of the edge between β and γ does however contradict the triplex $(\{\delta, \beta\}, \gamma)$ and hence we have a contradiction for the G shown in Figure 1b.

Third and last assume that the independence model of the graph shown in Figure 1c can be represented in a CG H of the AMP interpretation. From the Figure we can read the following independence statements: $\lambda \perp_G \mu | pa_G(\mu)$, $\alpha \perp_G \gamma | pa_G(\alpha), \beta \perp_G \delta | pa_G(\beta)$. It is clear that H must have the same skeleton, or clearly some separations or non-separations that hold in G would not hold in H. $\lambda \perp_G \mu | pa_G(\mu)$ and $\alpha \perp_G \gamma | pa_G(\alpha)$ gives that the triplexes $(\{\lambda, \mu\}, \beta)$ and $(\{\alpha, \gamma\}, \mu)$ must exists in H since $\beta \notin pa_G(\mu)$ and $\mu \notin pa_G(\alpha)$. As seen above this gives that $\lambda{\rightarrow}\gamma{-}\mu{\leftarrow}\alpha$ must exist in H. Similarly $\beta \perp_G \delta | pa_G(\beta)$ and $\lambda \perp_G \mu | pa_G(\mu)$ gives that $\lambda{\rightarrow}\beta{-}\mu{\leftarrow}\delta$ must exist in H. Finally $\alpha \perp_G \gamma | pa_G(\alpha)$ and $\beta \perp_G \delta | pa_G(\beta)$ gives that the triplexes $(\{\alpha, \gamma\}, \beta)$ and $(\{\beta, \delta\}, \alpha)$ must hold in H, since $\beta \notin pa_G(\alpha)$ and $\alpha \notin pa_G(\beta)$, which in turn gives that $\gamma{\rightarrow}\beta{-}\alpha{\leftarrow}\delta$ must exist in H. This does however contradict that H is a CG since the semi-directed cycle $\gamma{\rightarrow}\beta{-}\mu{-}\gamma$ exists in H. Hence we have a contradiction.

4.2 Translation of AMP CGs to MVR CGs

Theorem 4. *Given an AMP CG G, and a maximally oriented AMP CG G' in the Markov equivalence class of G, there exists a CG H st $I_{AMP}(G) = I_{MVR}(H)$ iff G' does not contain any induced subgraph of the form $X{-}Y{-}Z$.*

Proof. Sufficiency follows from Lemma 2 while necessity follows from 6.

Lemma 6. *If a maximally orientated CG G of the AMP interpretation contains an induced subgraph of the form $X{-}Y{-}Z$ then G there exists no CG H of the MVR interpretation st $I_{AMP}(G) = I_{MVR}(H)$.*

Proof. Assume to the contrary that the lemma does not hold. Clearly G and H must have the same skeleton or some separations in H do not hold in G or vice versa. Let H have a component ordering ord for its components $c_1, ...c_k$ st $ord(c_i) < ord(c_j)$ if c_i is a parent of c_j. Let C be the component of X in G. From the assumption we know that $X \perp_G Z | nb_G(X) \cup pa_G(X \cup nb_G(X))$ holds, where $Y \in nb_G(X)$, and hence that H must contain one of the following induced subgraphs: $X{\circ\!\!\rightarrow}Y{\rightarrow}Z$, $X{\leftarrow}Y{\leftarrow\!\!\circ}Z$ or $X{\leftarrow}Y{\rightarrow}Z$. For any other configuration of edges $X \perp_H Z | nb_G(X) \cup pa_G(X \cup nb_G(X))$ does not hold. Moreover we can generalize the configurations to $X{\circ\!\!\rightarrow}Y{\rightarrow}Z$ and $X{\leftarrow}Y{\rightarrow}Z$ simply by choosing the nodes to represent X and Z accordingly. Both these structures are included in $X{\circ\!\!-\!\!\circ}Y{\rightarrow}Z$

and we will now show that this structure leads to a contradiction if a split is not feasible in G.

The proof is iterative and when a restart is noted this is where the proof restarts. For each restart it will be shown that there must exist a triplet of nodes X, Y, Z st an induced subgraph of the form $X-Y-Z$ exists in G and $X \circ\!\!-\!\!\circ Y \to Z$ in H. Apart from this we also know that $I_{AMP}(G) = I_{MVR}(H)$ holds and that no split is feasible in G. Let the set U consist of Y and every node connected by a path to Y in the subgraph of G induced by $C \setminus Z$ and the set L consist of $C \setminus U$. This separation of sets gives that $nb_G(U) = Z$. For a split not to be feasible with these sets one of the conditions in Definition 2 must fail:

Case 1, condition 1 or 2 fails. This means that there exist two nodes $W \in C$ and $P \in C$ st the induced subgraph $P-Z-W$ exists in G. Note that one of the nodes might be Y. This means that $P \perp_G W | nb_G(W) \cup pa_G(W \cup nb_G(W))$ holds, where $Z \in nb_G(W)$ and hence that $P \circ\!\!\to Z \to W$, $P \leftarrow Z \leftarrow\!\!\circ W$ or $P \leftarrow Z \to W$ must exist in H as described above. Without losing generality we can say that either $P \circ\!\!\to Z \to W$ or $P \leftarrow Z \to W$ exists in H and that $W \neq Y$ by choosing P and W appropriately. This means that we can restart the proof with the structure $P \circ\!\!-\!\!\circ Z \to W$ in H (and $P-Z-W$ in G). The number of restarts is bounded since (1) the number of nodes in V is bounded and that $ord(co_H(Z)) > ord(co_H(Y))$.

Case 2, condition 1 and 2 hold but condition 3 fails. This means that there exists two nodes $W \in U$ and $P \notin C$ st the induced subgraph $Z-W \leftarrow P$ exists in G. First let us cover the case where $W = Y$. This means that $Z \perp_G P | pa_G(Z)$ holds. Since H have the same skeleton as G this means that H must contain an induced subgraph of the form $P \circ\!\!\to Y \leftarrow\!\!\circ Z$ since $Y \notin pa_G(Z)$. At the same time we know that H contains the edge $Y \to Z$ which causes a contradiction and hence $Y \neq W$ must hold. Therefore, $P \notin pa_G(Y)$ holds which generalized means that that $pa_G(Y) \subseteq pa_G(Z)$ must hold. For $Z \perp_G P | pa_G(Z)$ to hold in H there must exist an unshielded collider between Z and P over W and hence that the induced subgraph $Z \circ\!\!\to W \leftarrow\!\!\circ P$ exists in H. Similarly we have that $Y \perp_G P | pa_G(Y)$ gives that H contains an induced subgraph of the form $Y \circ\!\!\to W \leftarrow\!\!\circ P$. Note that $Y \in adj_G(W)$ must hold since condition 2 holds. Moreover for H not to contain a semi directed cycle over $Y \to Z \circ\!\!\to W \leftarrow\!\!\circ Y$ we can see that $Y \to W \leftarrow\!\!\circ P$ must exist in H. Finally note that $X \neq W$ must hold since $X \notin adj_G(Z)$ holds.

Now assume $X \in nb_G(W)$. For $X \perp_H Z | nb_G(X) \cup pa_G(X \cup nb_G(X))$ to hold, together with $W \in nb_G(X)$ and $Z \circ\!\!\to W$, it is easy to see that the induced subgraph $Z \circ\!\!\to W \to X$ must be in H. We can now see that $P \in pa_G(X)$ must hold or the induced subgraph $X \leftarrow W \leftarrow\!\!\circ P$ in H contradicts that $X \perp_G P | pa_G(X)$ holds in H. Moreover, for $X \circ\!\!-\!\!\circ Y \to W \to X$ not to form a semi-directed cycle in H the edge between X and Y must be orientated to $X \leftarrow Y$. We can therefore restart the proof by replacing X with Z, i.e. with the induced subgraph $X \leftarrow Y \circ\!\!-\!\!\circ Z$ in H (and $X-Y-Z$) in G. Since we know that $Z \circ\!\!\to W \to X$ exists in H we know that $ord(co_H(X)) > ord(co_H(Z))$. Hence we cannot get back to this subcase again (or we would have that $ord(co_H(X)) < ord(co_H(Z))$ which is a contradiction). This, together with that Y is kept the same and that $|V|$ is finite gives that the number of restarts is bounded. Hence $X \notin nb_G(W)$ must hold.

Now assume that $pa_G(Z) \subseteq pa_G(W)$. We can now restart the proof with $X {\circ\!\!-\!\!\circ} Y \rightarrow W$. The number of iterations is then bounded since $|V|$ is finite and case 2 cannot occur with Z as W again, or $pa_G(Z) \nsubseteq pa_G(W)$ would have to hold which is a contradiction. Hence $pa_G(Z) \nsubseteq pa_G(W)$ must hold. Let Q be the parent of Z not shared by W. Since $W \perp_G Q | pa_G(W)$ holds, and we know that H contains the induced subgraph $Q {\circ\!\!-\!\!\circ} Z {\circ\!\!\rightarrow} W {\leftarrow\!\!\circ} P$, we can draw the conclusion that H must contain the induced subgraph $Q {\circ\!\!\rightarrow} Z \leftrightarrow W {\leftarrow\!\!\circ} P$ since $Z \notin pa_G(W)$. Note that if there exist two different nodes W_1 and W_2 such that both have the properties described for W in case 2 W_1 and W_2 must be adjacent. If this were not the case we would have that both $W_1 \perp_G W_2 | nb_G(W_1) \cup pa_G(W_1 \cup nb_G(W_1))$ and $W_1 \not\perp_H W_2 | nb_G(W_1) \cup pa_G(W_1 \cup nb_G(W_1))$ would hold, since $Z \in nb_G(W_1)$. Also note that since $W_1 \leftrightarrow Z$ and $W_2 \leftrightarrow Z$ exists in H and the edge between W_1 and W_2 must be bidirected or H contains a semi-directed cycle. Let D be a set of nodes containg Z as well as all nodes that have the properties described for W. From the description above we can see that D must be complete and that the subgraph induced by D in H must only contain bidirected edges. We will now show that a split must be feasible in G with D as L and $C \setminus D$ as U. For a split not to be feasible one of the constraints in Definition 2 must fail.

Assume condition (1) or (2) fails. Then there exists three nodes $R \in C$, $T \in C$ and $D_j \in D$ st the induced subgraph $T - D_j - R$ exists in G. Since $T \perp_G R | nb_G(R) \cup pa_G(R \cup nb_G(R))$ holds we must, without losing generalization, have that H contains the induced subgraph $T {\circ\!\!-\!\!\circ} D_j \rightarrow R$, since $D_j \in nb_G(R)$. If this is the case we can however restart the proof with this induced subgraph and know that the number of iterations is bounded since $|V|$ is finite and $ord(co_H(D_j)) > ord(co_H(Y))$.

Assume condition (1) and (2) holds but (3) fails. Then there exists two nodes $R \in U$ and $T \notin C$ st the induced subgraph $D_i - R \leftarrow T$ exists in G for some $D_i \in D$. First note that R must be adjacent of all nodes in D or condition 1 would have failed in this split. Secondly note that $R - Y$ must exist in G or condition 2 would fail if we restart the proof with $X {\circ\!\!-\!\!\circ} Y \rightarrow D_i$ and a contradiction follows from there. Thirdly note that $R \notin adj_G(X)$ must hold or the proof could be restarted with $X {\circ\!\!-\!\!\circ} Y \rightarrow D_i$, for which condition 3 would fail with R as W and a contradiction would follow as shown above. Finally note that $pa_G(R) \subseteq pa_G(Z)$ must hold or R would be in D. This means that $pa_G(W) \nsubseteq pa_G(R)$, and hence that $P \notin pa_G(R)$, holds. Moreover we know that the edge $D_i \leftrightarrow W$ exists in G. For $D_i \perp_G T | pa_G(D_i)$ to hold in H it is clear that H must contain the induced subgraph $D_i {\circ\!\!\rightarrow} R {\leftarrow\!\!\circ} T$ since $R \notin pa_G(D_i)$. Similarly we have that for $R \perp_H P | pa_G(R)$ to hold H must contain an induced subgraph of the form $R {\circ\!\!\rightarrow} W {\leftarrow\!\!\circ} P$ since $W \notin pa_G(R)$. This means that for $R {\circ\!\!\rightarrow} W \leftrightarrow D_i {\circ\!\!\rightarrow} R$ not to form a semi-directed cycle in H the edge $R \leftrightarrow D_i$ must exist in H. Moreover, since $\forall D_m \in D \setminus D_i$ $R \in adj_G(D_m)$, $R \leftrightarrow D_i$ and $D_i \leftrightarrow D_m$ hold, clearly $R \leftrightarrow D_m$ must also hold or G contains a semi-directed cycle. Hence the subgraph of H induced by $D \cup R$ is complete and contains only bidirected edges. This in turn means that for $Y \rightarrow D_i \leftrightarrow R {\circ\!\!-\!\!\circ} Y$ to not form a semi-directed cycle $Y \rightarrow R$ must exist in H.

Hence we can move R into D and redo do the last split again. The number of restarts are bounded since $|V|$ is finite.

Hence each condition in Definition 2 must hold and we have a contradiction.

4.3 Translation of MVR CGs to LWF CGs

Theorem 5. *Given a MVR CG G, and a maximally oriented MVR CG G' that is in the same Markov equivalence class as G, there exist a LWF CG H st $I_{MVR}(G) = I_{LWF}(H)$ iff G' contains no bidirected edge, i.e. can be represented as a BN.*

Proof. From Lemma 7 it follows that a maximally oriented CG G' of the MVR interpretation with a bidirected edge must have a subgraph of the form shown in Figure 2. If it does not contain any bidirected edge in the maximally oriented model it trivially follows that it is a BN (and hence it can be represented as a CG of the LWF interpretation). From Lemma 8 it then follows that no CG G of the MVR interpretation which contains a subgraph of the form shown in Figure 2 can be represented as a CG of the LWF interpretation.

Lemma 7. *If a bidirected edge exists in a maximally oriented CG G of the MVR interpretation then G must contain an induced subgraph of the form shown in Figure 2.*

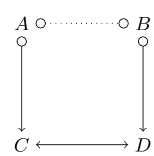

Fig. 2: Included subgraph in Lemma 7 and 8.

Proof. Assume to the contrary that a CG G of the MVR interpretation exists where (1) no induced subgraph of the form shown in Figure 2 exists, (2) no split is feasible and (3) at least one bidirected edge exists. From this assumption we can see that there has to exist at least two nodes X and Y st $X \leftrightarrow Y$ exists in G. Let C be the connectivity component to which X and Y belongs. Separate the nodes of C into two sets U and L st X and every node connected by a path to X in the subgraph of G induced by $C \setminus Y$ belongs to L and $C \setminus L$ belongs to U. This separation of nodes allows us to know that $sp_G(L)$ only contains Y. For a split not to be feasible at least one condition in Definition 3 has to fail.
Case 1. Assume constraint 1 fails. This means a node $Z \in L$ exist st $Z \leftrightarrow Y \leftrightarrow X$ occurs in G where $Z \notin adj_G(X)$ must hold, or Z would be in U. Now remove Y from U and add it to L as well as all nodes not connected by a path with Z in the subgraph of G induced by $U \setminus Y$ and attempt another split. This

separation of nodes allows us, since we previously had $nb_G(L) = Y$ and Y now have changed sets, to say that $sp_G(U) = Y$ must hold and hence that constraint 3 cannot fail. However, if constraint 1 or 2 fails we know there exists a node W st $W{\multimap}Z{\leftrightarrow}Y{\leftrightarrow}X$ is a subgraph of G but where $W \notin adj_G(Y)$, and $Z \notin adj_G(X)$ and $W \notin adj_G(X)$ by definition of the initial split, which implies a contradictory induced subgraph. Hence constraint 1 cannot fail in the initial split.
Case 2. If constraint 2 or 3 fails in the initial split we know there exists two nodes V_1 and P_1 st $P_1{\multimap}Y{\leftrightarrow}V_1$ exists in G but where $V_1 \notin adj_G(P_1)$ (note that V_1 might be X). Now let L consist of every node connected by a path to Y in the subgraph of G induced by $C \setminus V_1$ and the nodes $C \setminus L$ belong to U. This separation of nodes allows us to know that $sp_G(L)$ only contains V_1. If constraint 1 fails when performing a split with these sets it is clear from case 1 that a contradiction occurs. If constraint 2 or 3 fails we know there exists two new nodes V_2 and P_2 st $P_2{\multimap}V_1{\leftrightarrow}V_2$ exists in G but where $P_2 \notin adj_G(V_2)$. Note that V_2 or P_2 cannot be P_1 since $P_1 \notin adj_G(P_1)$. We now get that V_2 cannot be Y or an induced subgraph like that in Figure 2 occurs. $V_2 \in adj(Y)$ and $P_2 \in adj(Y)$ must also hold or the induced subgraph V_2 (resp. P_2)${\multimap}V_1{\leftrightarrow}Y{\leftarrow\!\circ}P_1$ occurs. By replacing V_1 with V_2, setting the proper U and L as described above it we can now repeat this procedure iteratively. Moreover, for every repetition i we must have that V_i and P_i must be adjacent of every $V_j (j < i)$ as well as Y or a contradiction occurs. This means that any nodes V_i and P_i already used in a previous repetition cannot be used in a later one, or both $P_i \in adj(V_i)$ and $P_i \notin adj(V_i)$ would have to hold. This in turn means that the number of repetitions is bounded since $|C|$ is finite and hence we have a contradiction that condition 2 or 3 can fail.

This means that all three conditions in Definition 3 must hold and hence a split must be feasible if the induced subgraph shown in Figure 2 does not occur.

Lemma 8. *If a CG G of the MVR interpretation contains an induced subgraph of the form shown in Figure 2 then G cannot be translated into a CG H of the LWF interpretation.*

Proof. Assume to the contrary that there exists a CG H, of the LWF interpretation, with the same independence model as G while G contains an induced subgraph of the form shown in Figure 2. Clearly H and G must contain the same nodes and adjacencies or some separations or non-separations must exist in G but not in H.

From Figure 2 we can read that $A \perp_G D | pa_G(D)$ and $C \perp_G B | pa_G(C)$ hold. For $A \perp_G D | pa_G(D)$ to hold in H C must be a collider between A and D and hence H must contain the induced subgraph $A{\rightarrow}C{\leftarrow}D$. Similarly $C \perp_G B | pa_G(C)$ gives that H must contain the induced subgraph $C{\rightarrow}D{\leftarrow}B$ and hence we have a contradiction.

4.4 Translation of LWF CGs to MVR CGs

Theorem 6. *Given a LWF CG G there exists a CG H st* $I_{LWF}(G) = I_{MVR}(H)$ *iff* $(G_{cl(K)})^m$ *is chordal for all* $K \in cc(G)$.

Proof. To prove the "if" part, note that if $(G_{cl(K)})^m$ is chordal for all $K \in cc(G)$, then there is a DAG D st $I_{LWF}(G) = I_{BN}(D)$ [1, Proposition 4.2] and, thus, it suffices to take $H = D$.

To prove the "only if" part, assume to the contrary that $V_1 - \ldots - V_n$ is a chordless undirected cycle in $(G_{cl(K)})^m$ for some $K \in cc(G)$. Note that H has the same adjacencies as G. Therefore, $V_{i-1} \leftarrow V_i$ and/or $V_i \rightarrow V_{i+1}$ must be in H because, otherwise, $V_{i-1} \perp_G V_{i+1} | Z \in I_{LWF}(G)$ for some Z st $V_i \in Z$ whereas $V_{i-1} \perp_H V_{i+1} | Z \notin I_{MVR}(H)$, which contradicts that $I_{LWF}(G) = I_{MVR}(H)$. Assume without loss of generality that $V_i \rightarrow V_{i+1}$ is in H. Then, $V_{i+1} \rightarrow V_{i+2}$ must be in H too, by an argument similar to the previous one. Repeated application of this reasoning implies that H has a semi-directed cycle, which contradicts the definition of CG.

Acknowledgments

This work is funded by the Center for Industrial Information Technology (CENIIT) and a so-called career contract at Linköping University, by the Swedish Research Council (ref. 2010-4808), and by FEDER funds and the Spanish Government (MICINN) through the project TIN2010-20900-C04-03.

References

1. Andersson, S.A., Madigan, D., Pearlman, M.D.: A Characterization of Markov Equivalence Classes for Acyclic Digraphs. Annals of Statistics 25:2, 505–541 (1997)
2. Andersson, S.A., Madigan, D., Pearlman, M.D.: An Alternative Markov Property for Chain Graphs. Scandianavian Journal of Statistics 28, 33–85 (2001)
3. Cox, D.R., Wermuth, N.: Linear Dependencies Represented by Chain Graphs. Statistical Science 8, 204–283 (1993)
4. Cox, D.R., Wermuth, N.: Multivariate Dependencies: Models, Analysis and Interpretation. Chapman and Hall (1996)
5. Drton, M.: Discrete Chain Graph Models. Bernoulli 15(3), 736–753 (2009)
6. Frydenberg, M.: The Chain Graph Markov Property. Scandinavian Journal of Statistics 17, 333–353 (1990)
7. Lauritzen, S.L., Wermuth, N.: Graphical Models for Association Between Variables, Some of Which are Qualitative and Some Quantitative. Annals of Statistics 17, 31–57 (1989)
8. Roverto, A., Studený, M.: A Graphical Representation of Equivalence Classes of AMP Chain Graphs. Journal of Machine Learning Research 7(6), 1045–1078 (2006)
9. Sonntag, D., Peña, J.M.: Learning Multivariate Regression Chain Graphs under Faithfulness. In: Proceedings of the 6th European Workshop on Probabilistic Graphical Models. pp. 299–306 (2012)
10. Studený, M., Roverto, A., Štěpánová, Š.: Two Operations of Merging and Splitting Components in a Chain Graph. Kybernetika 45, 208–248 (1997)
11. Wermuth, N., Sadeghi, K.: Sequences of Regression and their Indepdendences. TEST 21(2), 215–252 (2012)
12. Wermuth, N., Wiedenbeck, M., Cox, D.: Partial Inversion for Linear Systems and Partial Closure of Independence Graphs. BIT Numerical Mathematics 46, 883–901 (2006)