Journal of Machine Learning Research 7 (2006) 1045–1078 Submitted 2/06; Published 6/06

A Graphical Representation of Equivalence Classes of AMP Chain Graphs

Alberto Roverato ROVERATO@UNIMORE.IT
Department of Social, Cognitive and Quantitative Sciences
University of Modena and Reggio Emilia
Viale Allegri 9
I-42100 Reggio Emilia, Italy

Milan Studený STUDENY@UTIA.CAS.CZ
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Pod vodárenskou věží 4
18208 Prague 8 Libeň, Czech Republic

Editor: David Madigan

Abstract

This paper deals with chain graph models under alternative AMP interpretation. A new representative of an AMP Markov equivalence class, called the *largest deflagged graph*, is proposed. The representative is based on revealed internal structure of the AMP Markov equivalence class. More specifically, the AMP Markov equivalence class decomposes into finer *strong equivalence* classes and there exists a distinguished strong equivalence class among those forming the AMP Markov equivalence class. The largest deflagged graph is the largest chain graph in that distinguished strong equivalence class. A composed graphical procedure to get the largest deflagged graph on the basis of any AMP Markov equivalent chain graph is presented. In general, the largest deflagged graph differs from the AMP essential graph, which is another representative of the AMP Markov equivalence class.

Keywords: chain graph, AMP Markov equivalence, strong equivalence, largest deflagged graph, component merging procedure, deflagging procedure, essential graph

1. Introduction

This paper studies chain graph models under the alternative interpretation introduced by Andersson, Madigan and Perlman (2001). In general, a *chain graph model* is a statistical model in which a chain graph is used to represent the conditional independence structure defining the statistical model. The vertices of the graph represent random variables and the conditional independence structure is determined through the respective *Markov property*. The class of chain graphs was introduced and the original interpretation was given by Lauritzen and Wermuth (1984); see also Lauritzen and Wermuth (1989). The mathematical theory of chain graphs was developed by Frydenberg (1990), who formally defined the Markov property corresponding to the original interpretation. Following the standard literature in this field, we will refer to this property as the *LWF* Markov property.

More recently, another interpretation of chain graphs was introduced by Andersson, Madigan and Perlman (1996); see also Andersson et al. (2001). This interpretation leads to an alternative Markov property and we will refer to it as the *AMP* Markov property.

Two different chain graphs may be equivalent with respect to a considered Markov property, by which is meant that they define the same statistical model. The distinction between different interpretations of chain graphs is reflected by the different concepts of equivalence, namely the *LWF Markov equivalence* and the *AMP Markov equivalence*.

From a statistical perspective, the point of interest is a statistical model. However, if we represent a statistical model using an arbitrary graph in the respective *Markov equivalence class*, then the non-unique nature of graphical description may result in difficulties. One type of difficulty concerns problems one can meet in structural learning of graphical models; see Section 2.3 of Chickering (2002) for a review in the case of acyclic directed graphs. A solution to these problems may be provided by a suitable choice of a unique *representative* of each Markov equivalence class, that is, of a particular element in that equivalence class. The choice of a suitable representative is also important from the perspective of causal inference in chain graphs (Lauritzen, 2001, Section 11.2).

The problem of the representative choice has a natural solution in the LWF case. Frydenberg (1990) showed that every LWF equivalence class contains the *largest chain graph*, which is the graph with the largest amount of undirected edges within the LWF equivalence class. Furthermore, every arrow in the largest chain graph is an arrow with the same direction in every chain graph from the class. The largest chain graph is uniquely determined and can serve as a natural representative of the LWF equivalence class. Moreover, there exist at least two procedures that transform every chain graph into the largest chain graph of the respective LWF equivalence class Roverato (2005); Volf and Studený (1999).

However, the situation is different in the AMP case. It is not clear what is a natural representative of an AMP equivalence class and, in particular, the notion of the "largest AMP chain graph" makes no sense. Andersson et al. (2001) proposed to represent an AMP equivalence class by a so-called *AMP essential graph*. Every arrow in the AMP essential graph is either an arrow with the same direction or an undirected edge in every chain graph from the equivalence class. Their terminology was inspired by the case of acyclic directed graph models, in which case the corresponding equivalence class has a suitable representative called the *essential graph* (Andersson et al., 1997). Indeed, if an AMP equivalence class contains an equivalence class of acyclic directed graphs, then the AMP essential graph coincides with the respective essential graph; see Proposition 4.2 in Andersson and Perlman (2006). However, as of now, there is no algorithm to construct the AMP essential graph, as in the LWF case. Furthermore, in the case that the AMP equivalence class contains a completely undirected graph, it may happen that the AMP essential graph has some arrows, and this unpleasant phenomenon was already reported in Section 7 of Andersson et al. (2001).

The aim of this paper is to provide an alternative solution to the problem of unique graphical representation of AMP equivalence classes. The point is that AMP equivalence classes have a more complicated structure than LWF equivalence classes. We succeed in revealing this structure and provide a representing graph as well as an algorithm for its construction. Our solution, called the *largest deflagged graph*, is different from the AMP essential graph proposed in Andersson et al. (2001). Nevertheless, if an AMP equivalence class contains an acyclic directed graph then our representative reduces to the essential graph of the corresponding equivalence class of acyclic directed graphs. Moreover, it provides a better solution if the AMP equivalence class contains an undirected graph because then, as one would expect, that undirected graph coincides with our largest deflagged graph.

These are the main contributions of the paper:

1. We introduce a new concept, namely, the concept of *strong equivalence* of chain graphs. An important observation is that every AMP equivalence class of chain graphs decomposes into strong equivalence classes.

2. Every strong equivalence class has a unique representative given by the largest chain graph of the class. We introduce a procedure, called the *component merging procedure*, which starts from any chain graph G and, by replacing arrows with undirected edges, finds the largest graph in the class of chain graphs strongly equivalent to G.

3. There exists a unique distinguished strong equivalence class among those forming an AMP equivalence class. Its elements have the largest amount of immoralities and the least possible amount of flags. We call the graphs in this class the maximally deflagged graphs, briefly the *deflagged graphs*. We introduce a procedure, called the *deflagging procedure*, which starts from any chain graph G and, by replacing undirected edges with arrows, produces a deflagged graph $\widehat{G}$ AMP equivalent to G.

4. We propose to characterize every AMP equivalence class by means of the respective *largest deflagged graph*. This representative can be constructed by applying the component merging procedure to the chain graph $\widehat{G}$ obtained by the deflagging procedure.

The next section recalls basic graphical concepts. Then, in Section 3, the main results in the LWF case are recalled in order to let the reader see some analogy. In Section 4, we give an overview of our new results and illustrate them by an example. The results on strong equivalence of chain graphs are formulated in Section 5. They appear to be analogous to the results valid in the LWF case. In Section 6, we present a deflagging procedure to get a deflagged graph in a given AMP equivalence class. Section 7 contains some concluding remarks. Proofs of the main results are moved to the Appendix.

2. Basic Concepts

In this paper we consider graphs that admit both directed edges, called arrows, and undirected edges, called lines. Formally, given a non-empty finite set N, an *arrow* over N is an ordered pair (a,b) of distinct elements of N and a *line* over N is a subset $\{a,b\}$ of N of cardinality two, that is, an unordered pair of distinct elements of N. A *hybrid graph* is a triplet $H \equiv (N,\mathcal{A},\mathcal{L})$ where N is a finite non-empty set of *nodes*, $\mathcal{A}$ a set of arrows over N and $\mathcal{L}$ a set of lines over N such that no multiple edges are allowed, which means that if $(a,b) \in \mathcal{A}$ then $(b,a) \notin \mathcal{A}$ and $\{a,b\} \notin \mathcal{L}$. To express that N is the set of nodes of H we also say that H is a hybrid graph *over* N.

Given a hybrid graph H, we will write $a \longrightarrow b$ in H or $b \longleftarrow a$ in H to denote $(a,b) \in \mathcal{A}$. Analogously, we will write $a \,—\!\!—\, b$ in H or $b \,—\!\!—\, a$ in H if $\{a,b\} \in \mathcal{L}$. This notation is in accordance with usual pictures. An ordered pair $[a,b]$ of distinct nodes in H will be called an *edge* in H if $a \longrightarrow b$ in H, $a \,—\!\!—\, b$ in H or $a \longleftarrow b$ in H. Observe that $[a,b]$ is an edge iff $[b,a]$ is an edge: this means that edges can be viewed as unordered pairs of distinct nodes. If $[a,b]$ is an edge in H we also say that a and b are *adjacent* in H.

A set of nodes $C \subseteq N$ is *connected* in H if, for every $a, b \in C$, there exists an *undirected path* connecting them, that is, a sequence of distinct nodes $a = c_1, \dots, c_n = b$, $n \geq 1$ such that $c_i \,—\, c_{i+1}$ in H for $i = 1, \dots, n-1$. A (connectivity) *component* in H is a maximal connected set in H with respect to set inclusion. Evidently, components in a hybrid graph are pairwise disjoint.

Given a set of nodes $A \subseteq N$ in a hybrid graph H, the set of *parents* of nodes in A, denoted by $\mathrm{pa}_H(A)$, is the set

$$\mathrm{pa}_H(A) \equiv \{b \in N;\ b \longrightarrow a \ \text{ in } H \text{ for some } \ a \in A\}\,.$$

A *descending path* in H from a node a to a node b is a sequence of distinct nodes $a = c_1, \dots, c_n = b$, $n \geq 1$ such that either $c_i \longrightarrow c_{i+1}$ in H or $c_i \,—\, c_{i+1}$ in H for $i = 1, \dots, n-1$. If there exists a path of this kind in H then we say that a is an *ancestor* of b in H. The set of ancestors in H of nodes in a set $A \subseteq N$ will be denoted by $\mathrm{an}_H(A)$; observe that one has $A \subseteq \mathrm{an}_H(A)$.

An *undirected graph* is a hybrid graph without arrows, that is, $\mathcal{A} = \emptyset$. A set $K \subseteq N$ is *complete* in an undirected graph H if, $\forall\, a, b \in K$, $a \neq b$, one has $a \,—\, b$ in H. Given a hybrid graph H over N, the respective *underlying graph* is an undirected graph H^u over N such that $a \,—\, b$ in H^u iff $[a, b]$ is an edge in H.

A *directed graph* is a hybrid graph without lines, that is, $\mathcal{L} = \emptyset$. A directed graph H is *acyclic* if there is no directed cycle in H, that is, there is no sequence of nodes $d_0, \dots, d_{n-1}, d_n = d_0$, $n \geq 3$ such that $d_0, \dots, d_{n-1}$ are distinct and, $\forall\, i = 0, \dots, n-1$, $d_i \longrightarrow d_{i+1}$ in H.

The concept of a *chain graph* (CG) can be introduced in two equivalent ways. Note that we are going to use the abbreviation *CG* in the rest of the paper. The first definition is that a CG is a hybrid graph H whose components can be ordered to form a chain, that is, a sequence $C_1, \dots, C_m$, $m \geq 1$ such that

- if $a \,—\, b$ in H then $a, b \in C_i$ for some i,
- if $a \longrightarrow b$ in H then $a \in C_i, b \in C_j$ with $i < j$.

Note that this is the reason for which some authors call the components in a CG *chain components*. A consequence of this definition is that every CG has a *terminal component*, that is, a component T such that there is no arrow $a \longrightarrow b$ in H with $a \in T$. The other definition is that a CG is a hybrid graph H without *semi-directed cycles*. A semi-directed cycle of the length n is a sequence of nodes $d_0, \dots, d_{n-1}, d_n = d_0$ with $n \geq 3$ such that $d_0, \dots, d_{n-1}$ are distinct, $d_0 \longrightarrow d_1$ in H and, $\forall\, i = 1, \dots, n-1$, either $d_i \longrightarrow d_{i+1}$ in H or $d_i \,—\, d_{i+1}$ in H. See Lemma 2.1 in Studený (1997) for the proof of equivalence of both definitions of a CG. It is easy to see that every undirected graph and every acyclic directed graph is a CG.

Given a hybrid graph H over N and $\emptyset \neq A \subseteq N$, the *induced subgraph* of H for A, denoted by H_A, is the graph $H_A \equiv (A, \mathcal{A} \cap (A \times A), \mathcal{L} \cap \mathcal{P}(A))$, where $\mathcal{P}(A) \equiv \{B; B \subseteq A\}$. We will deal with a few special induced subgraphs. A *complex* in a hybrid graph H is an induced subgraph of H of the form $a \longrightarrow c_1 \,—\, \dots \,—\, c_s \longleftarrow b$, $s \geq 1$, which means that no other edge between distinct nodes $\{a, b, c_1, \dots, c_s\}$ is present in H. An example of a complex is shown in the left-hand picture of Figure 1. A special case of a complex is an *immorality*, which is a configuration $a \longrightarrow c \longleftarrow b$ in H where a, b, c are distinct nodes and $[a, b]$ is not an edge in H. An example of an immorality is shown in the middle picture of Figure 1. Two CGs G and H over N will be called *complex equivalent* iff they have the same underlying graph and complexes. Given a CG G, the class of CGs that are complex equivalent to G will be denoted by $\mathcal{G}$.

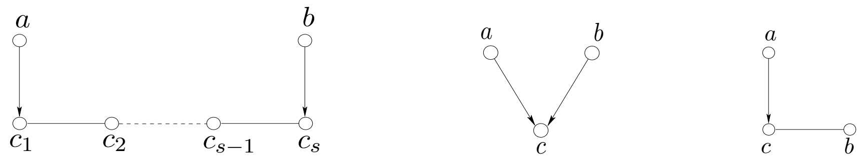


Figure 1: A complex, an immorality and a flag.

A *flag* in a hybrid graph H is another induced subgraph of H for three nodes, namely $a \longrightarrow c$ —— b where a,b,c are distinct nodes and $[a,b]$ is not an edge in H. An example of a flag is shown in the right-hand picture of Figure 1. A *triplex* in a hybrid graph H is a pair $\langle\{a,b\},c\rangle$ such that either $a \longrightarrow c \longleftarrow b$ is an immorality in H, $a \longrightarrow c$ —— b is a flag in H or a —— $c \longleftarrow b$ is a flag in H. All three different versions of a triplex are shown in Figure 2. Two CGs G and H over N will be called *triplex equivalent* iff they have the same underlying graph and triplexes. Note that coincidence of triplexes is understood as follows. If, for instance, $a \longrightarrow c \longleftarrow b$ is an immorality in G then it need not be an immorality in H but it has to be one of three versions of the triplex $\langle\{a,b\},c\rangle$. Given a CG H, the class of CGs that are triplex equivalent to H will be denoted by $\mathbb{H}$.

3. Representation of LWF Equivalence Classes

In this section we recall known results concerning the LWF case. The aim is to help the reader to realize the analogy between these former results and our new results on strong equivalence of CGs presented in Section 5. Moreover, an overview of the results in the LWF case will indicate what is the main difference from the AMP case, which is reported in Section 4.

3.1 Largest Chain Graph in a LWF Equivalence Class

In this paper, we omit the formal definition of LWF Markov property and LWF Markov equivalence; this can be found in Frydenberg (1990). Instead, we recall Frydenberg's graphical characterization of LWF equivalence of CGs (see Proposition 5.6 in Frydenberg, 1990). He showed that two CGs over the same set of nodes are LWF Markov equivalent iff they are complex equivalent.

The second crucial point is that every LWF equivalence class is endowed with a natural partial ordering. Supposing that $H = (N, \mathcal{A}_H, \mathcal{L}_H)$ and $G = (N, \mathcal{A}_G, \mathcal{L}_G)$ are two LWF equivalent CGs, we say that H is *larger* than G if $\mathcal{A}_H \subseteq \mathcal{A}_G$, that is

$$a \longrightarrow b \text{ in } H \text{ implies } a \longrightarrow b \text{ in } G, \tag{1}$$

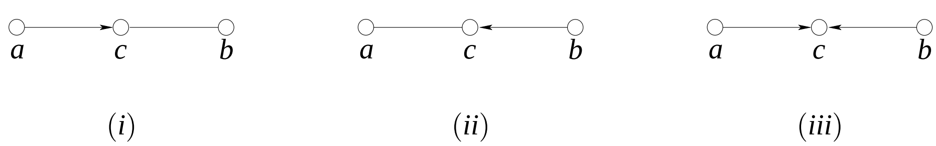


Figure 2: Three different versions of the triplex $\langle\{a,b\},c\rangle$.

for every pair a and b of distinct nodes in H. Observe that the fact that G and H have the same underlying graph necessitates that $\mathcal{L}_G \subseteq \mathcal{L}_H$, that is

$$a \,\text{---}\, b \text{ in } G \text{ implies } a \,\text{---}\, b \text{ in } H, \tag{2}$$

which means H has 'more' lines than G. One can easily show that the relation defined by (1) is a partial ordering on every LWF equivalence class; we will write $H \geq G$ if (1) is fulfilled.

Third, Frydenberg also showed (Proposition 5.7 of Frydenberg, 1990) that every LWF equivalence class $\mathcal{G}$ has the largest element with respect to this ordering, that is, $G_\infty \in \mathcal{G}$ such that for every G in $\mathcal{G}$ one has $G_\infty \geq G$. Thus, this graph G_∞, named the *largest chain graph* of $\mathcal{G}$, can serve as a natural representative of $\mathcal{G}$.

3.2 Feasible Merging of Components

The last important point is that there are procedures which allow one to get the largest CG $G_\infty \in \mathcal{G}$ on the basis of any CG $G \in \mathcal{G}$ from the LWF equivalence class. At least three procedures of this kind have been presented in the literature; however, two of them are methodologically equivalent.

One of them could be a procedure based on Theorem 3.9 of Volf and Studený (1999). The basic idea is that some arrows in a CG $G \in \mathcal{G}$ are indicated as 'protected' arrows. Then all arrows in G which are not 'protected' are replaced with lines and the largest chain graph G_∞ of $\mathcal{G}$ is obtained.

Another procedure, called the *pool-component rule*, was presented in Section 5 of Studený (1997). The basic idea is that there is an elementary operation of merging components in a CG whose result is an LWF equivalent CG. By consecutive application of this operation, the respective largest chain graph can be obtained. However, the formal description of that elementary operation given in Studený (1997) is still awkward.

The third procedure is described in Roverato (2005). Its basic idea is essentially the same; an elementary step of that procedure consists of merging components of an 'insubstantial' meta-arrow, that is, of the bunch of arrows between two certain components. It is shown in Section 4 of Roverato (2005) that, by consecutive application of that elementary step, the respective largest CG is obtained. One can show that the elementary operations presented in Studený (1997) and Roverato (2005) are equivalent (see Studený et al., 2006), but the formal description of the operation presented in Roverato (2005) is much more elegant from the mathematical point of view. We decided to take it as the basis of the following definitions.

Definition 1 *(meta-arrow)*
Let G be a CG. A pair of components (U,L) in G such that there exists an arrow $a \longrightarrow b$ in G with $a \in U$ and $b \in L$ determines a meta-arrow *in G. More specifically, the meta-arrow is the collection of all arrows $a \longrightarrow b$ with $a \in U$ and $b \in L$. The component U will be called the* upper component *and the component L the* lower component *(of the meta-arrow). We will occasionally use the notation $U \rightrightarrows L$.*

Note that the above notion is a minor modification of the concept of a meta-arrow from Roverato (2005). The essential difference is that in Definition 1 we require that at least one arrow exists from a member of U to a member of L, while in Roverato (2005) a possibly empty collection of arrows from U to L was allowed. Thus, the concept of a meta-arrow used in this paper coincides with the concept of a non-empty meta-arrow from Roverato (2005). Since empty meta-arrows play no role

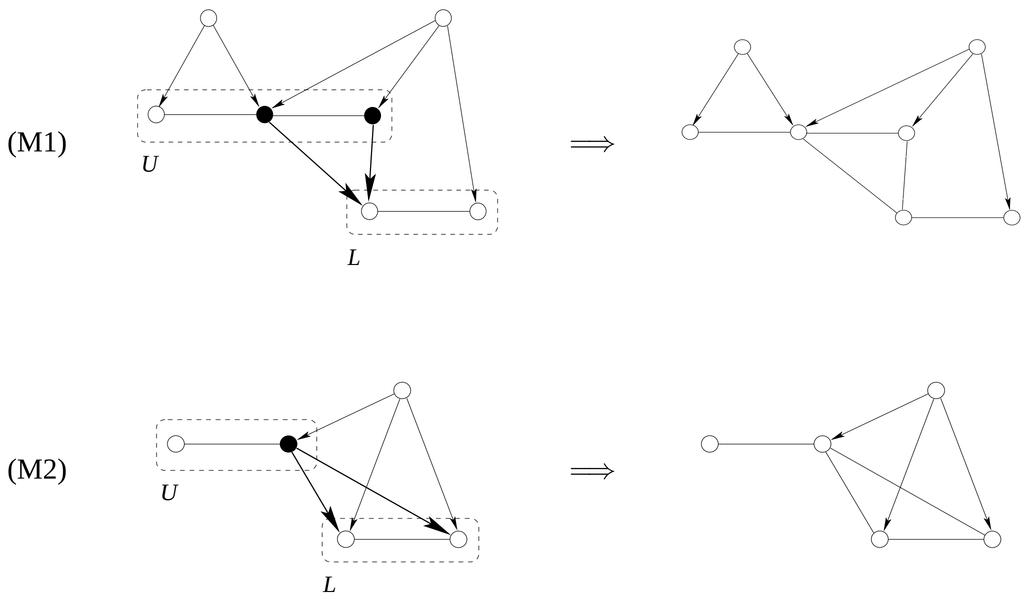


Figure 3: Two examples of feasible merging. The vertices belonging to the set K from (i) are filled in and arrows of the meta-arrow $U \rightrightarrows L$ are bold.

in a CG, we have decided to simplify our terminology. Our additional assumption also implies that the considered components U and L are different.

Definition 2 *(merging of components)*
By merging of components *in a CG G we understand the following operation applicable to G. Given a pair of components (U,L) which defines a meta-arrow, we replace all arrows of the meta-arrow $U \rightrightarrows L$ with lines and say that the resulting hybrid graph G' is obtained by merging of components U and L; more specifically, by* merging *of the upper component U and the lower component L.*

Note that the above terminology was inspired by terminology from Studený (2004). In general, the result of the operation of merging components in a CG need not be a CG. However, there are sufficient conditions for this; one of them is as follows.

Definition 3 *(feasible merging)*
Let (U,L) be a pair of components in a CG G that defines a meta-arrow in G. We say that merging of components U and L is feasible *(in G) if the following two conditions hold:*

(i) $K \equiv \mathrm{pa}_G(L) \cap U$ *is a complete set in G,*

(ii) $\forall\, b \in K \quad \mathrm{pa}_G(L) \setminus U \subseteq \mathrm{pa}_G(b)$.

Note that the assumption that (U,L) defines a meta-arrow implies that the set K in (i) is a non-empty set. Two examples of feasible merging are shown in Figure 3.

It is shown in Section 4 of Roverato (2005) that a hybrid graph G' obtained from a CG G by feasible merging of its components is a CG complex equivalent to G; actually, it is shown there that the requirements (i) and (ii) together establish a necessary and sufficient condition for this. In fact, that is the reason we decided to name this operation with CGs "feasible merging of components'' because the condition ensures that one remains in the same LWF equivalence class of CG after the merging operation. Moreover, it is also proven in Roverato (2005) that, by repeated application of this operation to a CG $G \in \mathcal{G}$, the respective largest CG $G_\infty \in \mathcal{G}$ is obtained.

4. Representation of AMP Equivalence Classes

In this section we reveal the internal structure of AMP Markov equivalence classes. First, we recall the graphical characterization of AMP equivalence. Then we introduce a special kind of equivalence of CGs, called *strong equivalence*, such that every AMP equivalence class decomposes into strong equivalence classes. Basic results on strong equivalence are postponed to Section 5. The next step is to introduce a special *flag ordering* between strong equivalence classes within a fixed AMP equivalence class. We show that the smallest element with respect to that ordering exists and, finally, we propose to represent the whole AMP equivalence class by a natural representative of that distinguished strong equivalence class, called *largest deflagged graph*.

4.1 Graphical Characterization of AMP Equivalence

The formal definitions of AMP Markov property and AMP Markov equivalence are omitted; they can be found in Andersson et al. (2001). Here we recall graphical characterization of AMP equivalence given by Andersson et al. (2001, Theorem 5). They showed that two CGs over the same set of nodes are AMP Markov equivalent iff they are triplex equivalent. An example of an AMP equivalence class is given in Figure 2. A further, less trivial, example containing ten CGs is given in Figure 4.

Given a CG H, let us consider the set $\mathbb{H}$ of all CGs triplex equivalent to H. If we consider the partial ordering of CGs in $\mathbb{H}$ defined by (1) then it may be the case that the largest CG in $\mathbb{H}$ does not exist. This is illustrated in Figure 2, where none of the three graphs is larger than the others, but also in Figure 4.

This is the main difference between the case of LWF equivalence and the case of AMP equivalence. In the LWF case, the key role is played by the ordering of CGs defined by (1). The result on the existence and uniqueness of the largest CG with respect to this ordering in each LWF equivalence class reported in Section 3 makes this object a natural representative of the LWF equivalence class. In the representation of an AMP equivalence class, the ordering defined by (1) also plays an important role, even though its use in this case is more subtle than in the LWF case. What is important is that every AMP equivalence class decomposes into some finer equivalence classes.

4.2 Definition of Strong Equivalence

Our decomposition of a given AMP equivalence class is based on the distinction between *triplex edges*, namely the arrows and lines that belong to a triplex, and non-triplex edges. More specifically, if two triplex equivalent CGs have identical triplex edges, then we say that they are strongly equivalent.

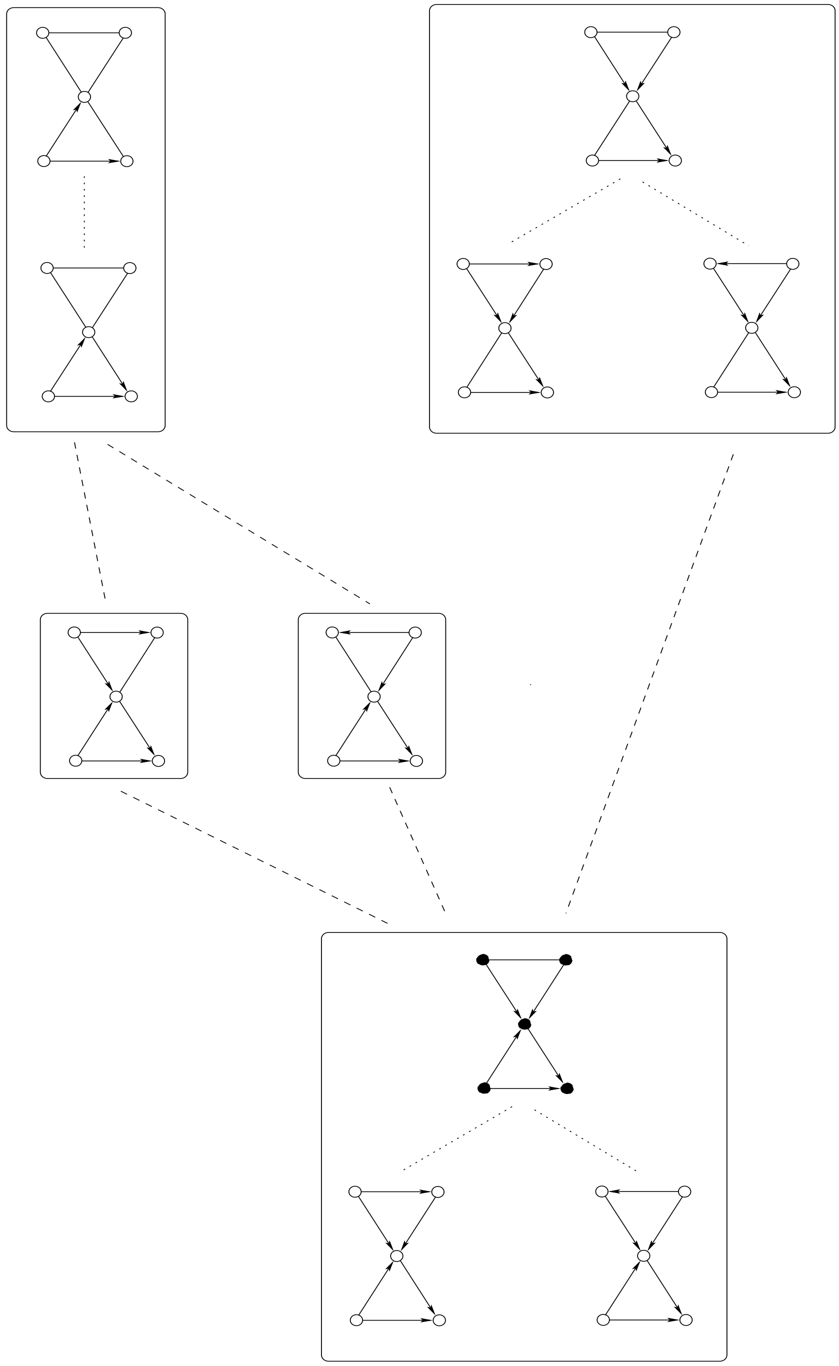


Figure 4: An example of an AMP equivalence class. The boxes represent strong equivalence classes. They are ordered by the flag ordering. There exists a unique largest graph within every strong equivalence class. The largest deflagged graph has vertices filled in.

Definition 4 *(strong equivalence of chain graphs)*
Let G,H be CGs over N. We say that they are strongly equivalent *iff*

[a] *G and H have the same underlying graph,*

[b] *an immorality $a \longrightarrow c \longleftarrow b$ occurs in G iff it occurs in H,*

[c] *a flag $a \longrightarrow c$ —— b occurs in G iff it occurs in H.*

It is easy to see that strongly equivalent CGs have the same complexes. In particular, they are both complex equivalent and triplex equivalent. On the other hand, two CGs which are both LWF and AMP Markov equivalent need not be strongly equivalent as shown, for example, by the graphs (i) and (ii) in Figure 2.

Given a CG H, the class of CGs that are strongly equivalent to H will be denoted by $\mathcal{H}$. In Figure 4, strong equivalence classes are represented by boxes. Note that, since all the graphs in $\mathcal{H}$ have the same triplex edges, it makes sense to say that $a \longrightarrow c$ is a triplex arrow in $\mathcal{H}$ if $a \longrightarrow c$ is a triplex arrow in every CG from $\mathcal{H}$, and similarly for triplex lines. We are going to show in Section 5 that, similarly to the LWF case, every strong equivalence class $\mathcal{H}$ has a unique largest element. We also present a special component merging procedure to get the largest element on basis of any graph in $\mathcal{H}$ there.

Strong equivalence is an equivalence relation that induces a partition of any AMP equivalence class $\mathbb{H}$ of CGs. We will denote the set of all strong equivalence classes included in $\mathbb{H}$ by $\overline{\overline{\mathbb{H}}} \equiv \{\mathcal{H}\,;\mathcal{H} \subseteq \mathbb{H}\}$.

4.3 Flag Ordering

Interestingly, the relation (1) restricted to triplex edges defines a partial ordering *between* strong equivalence classes from $\overline{\overline{\mathbb{H}}}$.

Definition 5 *(flag larger)*
Let $\mathbb{H}$ be an AMP equivalence class and $\mathcal{H},\mathcal{G} \in \overline{\overline{\mathbb{H}}}$. We say that $\mathcal{H}$ is flag larger *than $\mathcal{G}$ and write $\mathcal{H} \succeq \mathcal{G}$ if the following condition holds:*

$$\textit{whenever } a \longrightarrow b \textit{ is a triplex arrow in } \mathcal{H} \textit{ then } a \longrightarrow b \textit{ in } \mathcal{G}. \tag{3}$$

Observe that (3) and the fact $\mathcal{G},\mathcal{H} \in \overline{\overline{\mathbb{H}}}$ imply that

$$\text{whenever } a \;\text{——}\; b \text{ is a triplex line in } \mathcal{G} \text{ then } a \;\text{——}\; b \text{ in } \mathcal{H}. \tag{4}$$

Hence, $\mathcal{H} \succeq \mathcal{G} \succeq \mathcal{H}$ for $\mathcal{H},\mathcal{G} \in \overline{\overline{\mathbb{H}}}$ implies that $\mathcal{H}$ and $\mathcal{G}$ have the same triplex edges, that is, $\mathcal{G} = \mathcal{H}$. This allows one to see that the relation $\succeq$ is indeed an ordering on $\overline{\overline{\mathbb{H}}}$. Another point is that (4) means that $\mathcal{H}$ has 'more' triplex lines than $\mathcal{G}$. In particular, if $\mathcal{H} \succeq \mathcal{G}$ then every flag in $\mathcal{G}$ is a flag of the same type in $\mathcal{H}$. For this reason, we will refer to the ordering defined by (3) as to the *flag ordering* of strong equivalence classes. In Figure 4 we illustrated this ordering by dashed lines. Note that there exists the smallest element with respect to flag ordering. It is a natural distinguished strong equivalence class within $\overline{\overline{\mathbb{H}}}$ and, now, we prove its existence.

Proposition 6 *Given an AMP equivalence class $\mathbb{H}$, there exists a unique strong equivalence class $\mathcal{H}^{\downarrow} \in \overline{\overline{\mathbb{H}}}$ such that $\mathcal{H} \succeq \mathcal{H}^{\downarrow}$ for all $\mathcal{H} \in \overline{\overline{\mathbb{H}}}$.*

Proof As $\overline{\mathbb{H}}$ is finite and $\succeq$ is an ordering on $\overline{\mathbb{H}}$ it suffices to show that, for every $\mathcal{G}, \mathcal{H} \in \overline{\mathbb{H}}$, there exists $\mathcal{F} \in \overline{\mathbb{H}}$ with $\mathcal{G} \succeq \mathcal{F}$ and $\mathcal{H} \succeq \mathcal{F}$. Choose $G \in \mathcal{G}$ and $H \in \mathcal{H}$ and construct a hybrid graph F with the same underlying graph as G (and H) in this way: $a \longrightarrow b$ in F iff either $a \longrightarrow b$ in G or [a —— b in G and $a \longrightarrow b$ in H]. Lemma 4 in Andersson et al. (2001) says that F is a CG which is triplex equivalent to G (and H). Let $\mathcal{F}$ denote the strong equivalence class of CGs containing F. Thus, $\mathcal{F} \in \overline{\mathbb{H}}$ and the fact $G \geq F$ implies $\mathcal{G} \succeq \mathcal{F}$. The conclusion $\mathcal{H} \succeq \mathcal{F}$ can be verified directly: if $a \longrightarrow b$ is a triplex arrow in H (= in $\mathcal{H}$) then the fact that H and G are triplex equivalent implies that either $a \longrightarrow b$ in G or a —— b in G which both gives $a \longrightarrow b$ in F (= in $\mathcal{F}$). ■

4.4 Deflagged Graphs and Essential Flags

Given an AMP equivalence class $\mathbb{H}$, the symbol $\mathcal{H}^{\downarrow}$ will be used to denote the least strong equivalence class in $\overline{\mathbb{H}}$ with respect to $\succeq$. The graphs in $\mathcal{H}^{\downarrow}$ will be called *maximally deflagged graphs* or, briefly, *deflagged graphs*.

In the example in Figure 4, both triplexes in the deflagged graphs are immoralities. However, in general, not all triplex edges in $\mathcal{H}^{\downarrow}$ have to be arrows. Some flags appear to be essential for the specification of the set $\mathbb{H}$ and, therefore, their lines are shared by all graphs from $\mathbb{H}$. An example is given in Figure 5 where a single graph, which has two flags, forms the whole AMP equivalence class.

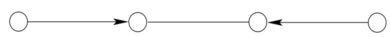


Figure 5: An example of a pair of essential flags.

Definition 7 *(essential flag)*
Let $\mathbb{H}$ be an AMP equivalence class. If $a \longrightarrow b$ —— d is a flag in H for every $H \in \mathbb{H}$ then we say that it is an essential flag *in $\mathbb{H}$.*

Actually, deflagged graphs can equivalently be introduced as follows.

Proposition 8 *Given an AMP equivalence class $\mathbb{H}$, one has $G \in \mathcal{H}^{\downarrow}$ iff $G \in \mathbb{H}$ and every flag in G is an essential flag in $\mathbb{H}$.*

Proof To verify the necessity of the condition, consider a flag $a \longrightarrow b$ —— c in G and $H \in \mathcal{H} \in \overline{\mathbb{H}}$. Then the assumption $\mathcal{H} \succeq \mathcal{H}^{\downarrow} \ni G$ implies by (4) that the triplex line b —— c in G is also in H. As $\langle\{a,c\},b\rangle$ is a triplex both in G and H it allows one to derive $a \longrightarrow b$ in H. Thus, $a \longrightarrow b$ —— c is a flag in every $H \in \mathbb{H}$.

For sufficiency, assume that $G \in \mathbb{H}$ only has essential flags. Let $\mathcal{G} \in \overline{\mathbb{H}}$ be the strong equivalence class containing G. We are to show that $\mathcal{H} \succeq \mathcal{G}$ for every $\mathcal{H} \in \overline{\mathbb{H}}$. Consider a triplex arrow $a \longrightarrow b$ in $\mathcal{H}$. It has to be a part of a triplex $\langle\{a,c\},b\rangle$. Since it has to be a triplex in G the only option which excludes $a \longrightarrow b$ in G is that a —— $b \longleftarrow c$ in G. However, then it is an essential flag in $\mathbb{H}$ and a —— $b \longleftarrow c$ in $\mathcal{H}$. This contradicts the assumption and one necessarily has $a \longrightarrow b$ in $\mathcal{G}$. ■

4.5 Largest Deflagged Graph

Let us summarize. AMP equivalence classes can effectively be handled by first considering their natural partition into strong equivalence classes (partially ordered by $\succeq$), and then by dealing with the CGs in every strong equivalence class (partially ordered by $\geq$). In this way, it is possible to identify unambiguously a graph in $\mathbb{H}$ by first considering the flag-smallest strong equivalence class and then by taking the largest graph within that class.

Definition 9 *(largest deflagged graph)*
The graph $H^{\downarrow}$ is the largest deflagged graph *of an AMP equivalence class $\mathbb{H}$ if*

(i) $H^{\downarrow} \in \mathcal{H}^{\downarrow}$,

(ii) $H^{\downarrow} \geq H$ *for all* $H \in \mathcal{H}^{\downarrow}$.

In Figure 4, the ordering of CGs within strong equivalence classes is illustrated by means of dotted lines. The largest deflagged graph is emphasized by means of vertices filled in.

Recall that the existence of the strong equivalence class $\mathcal{H}^{\downarrow}$ was proven in Proposition 6 whereas the existence and uniqueness of the largest CG in $\mathcal{H}^{\downarrow}$ is shown in Section 5. Furthermore, in Section 6, we provide a deflagging procedure which, starting from any CG G in an AMP equivalence class $\mathbb{H}$, returns a CG $\widehat{G}$ in $\mathcal{H}^{\downarrow}$. Then a component merging procedure from Section 5 can be applied to $\widehat{G}$ to get the largest deflagged graph $H^{\downarrow}$.

5. Strong Equivalence

This section is devoted to basic results on strong equivalence of CGs. These results are analogous to the results on LWF Markov equivalence recalled in Section 3. More specifically, we prove the existence of the largest CG within each strong equivalence class, introduce the respective elementary operation ascribing a larger strongly equivalent CG to a CG, and show that the largest CG in a strong equivalence class is attainable by this operation.

5.1 Largest Chain Graph in a Strong Equivalence Class

In this subsection we show the existence of the largest CG within a strong equivalence class. The first step for this is a direct construction of the supremum of two CGs with a shared underlying graph with respect to the ordering $H \geq G$ defined by (1). Note that the construction was already mentioned without further details in Frydenberg (1990). The construction utilizes the following auxiliary concept.

Definition 10 *(cyclic arrow)*
Given a hybrid graph H, we say that an arrow $a \longrightarrow b$ in H is a cyclic arrow *in H if $b \in \mathrm{an}_H(a)$. An equivalent formulation is that there exists a semi-directed cycle in H containing $a \longrightarrow b$.*

Lemma 11 *Let us consider the class E of all CGs over N with a prescribed underlying graph E, ordered by the relation $\geq$ defined by (1). Then every pair of graphs G and H from E has the supremum $G \vee H$ in $(E, \geq)$. It can be obtained directly in two steps.*

1. *Define a hybrid graph $G \cup H$ over N as follows*

$$a \longrightarrow b \text{ in } G \cup H \;\text{ iff }\; \text{both } a \longrightarrow b \text{ in } G \text{ and } a \longrightarrow b \text{ in } H\,,$$

and $a \,\text{---}\, b$ *in* $G \cup H$ *for remaining edges in* E.

2. *Replace all cyclic arrows in* $G \cup H$ *with lines and obtain* $G \vee H$.

Proof It is easy to see that $(E, \geq)$ is a partially ordered set. We need to show that $G \vee H \in E$, $G \vee H \geq G$, $G \vee H \geq H$ and, whenever there is $F \in E$ with $F \geq G, H$ then $F \geq G \vee H$.

The fact that $G \vee H$ is a CG was proven as Consequence 2.5 in Volf and Studený (1999). Hence, it is clear that $G \vee H \in E$ and that $G \vee H$ is larger than both G and H.

To show that $F \geq G \vee H$ for $F \in E$ with $F \geq G, H$, consider an arrow $a \longrightarrow b$ in F in order to verify $a \longrightarrow b$ in $G \vee H$. Since $a \longrightarrow b$ in $G \cup H$, it suffices to show $b \notin \mathrm{an}_{G \cup H}(a)$. Suppose for contradiction that there exists a descending path $\rho : b = c_1, \ldots, c_n = a$, $n \geq 2$ in $G \cup H$. There is no $1 \leq i \leq n-1$ with $c_i \longleftarrow c_{i+1}$ in F, as otherwise $c_i \longleftarrow c_{i+1}$ in $G \cup H$. Thus, ρ is a descending path in F which contradicts the assumption that F is a CG. ■

The preceding construction can be utilized to prove that every strong equivalence class of CGs is a join semi-lattice with respect to $\geq$.

Proposition 12 *Let* G *and* H *be strongly equivalent CGs over* N. *Then their supremum* $G \vee H$ *is strongly equivalent to them as well.*

Because the proof is technical, it is moved to the Appendix. Proposition 12 has the following consequence.

Corollary 13 *Given a strong equivalence class* $\mathcal{G}$ *of CGs over* N, *there exists* $G^{\dagger} \in \mathcal{G}$ *which is the largest CG in* $\mathcal{G}$.

Proof Since $\mathcal{G}$ is a finite set, one can apply Proposition 12 repeatedly to get the supremum of all graphs in $\mathcal{G}$. Of course, it is the largest CG in $\mathcal{G}$. ■

5.2 Legal Merging of Components

In this subsection we introduce an elementary operation that produces a strongly equivalent CG when applied to a CG. Here is the definition.

Definition 14 *(legal merging of components)*
Let (U, L) *be a pair of components in a CG* G *that defines a meta-arrow. We say that merging of components* U *and* L *is* legal *(in* G*) if the following three conditions hold:*

[i] $K \equiv \mathrm{pa}_G(L) \cap U$ *is a complete set in* G,

[ii] $\forall b \in K \quad \mathrm{pa}_G(L) \setminus U = \mathrm{pa}_G(b)$,

[iii] *for every* $d \in L$ *one has* $\mathrm{pa}_G(L) = \mathrm{pa}_G(d)$.

Evidently, the conditions [i]-[iii] imply the conditions (i)-(ii) from Definition 3. In brief, every legal merging (of components in a CG) is feasible. In Figure 3, (M1) is an example of feasible merging that is not legal whereas (M2) is an example of legal merging. If G is a CG without flags then the condition [iii] is always fulfilled and [ii] takes a simpler form:

$[\widetilde{\text{ii}}]$ $\mathrm{pa}_G(L) \setminus U = \mathrm{pa}_G(U)$.

Thus, the operation from Definition 14 generalizes the operation of legal merging of components (of a CG without flags) from Studený (2004). The requirement $[\text{i}]+[\widetilde{\text{ii}}]$ also coincides with the condition from Roverato (2005) demanding that the arrowhead of the meta-arrow $U \rightrightarrows L$ is strongly insubstantial.

Proposition 15 *Let G be a CG over N, and (U,L) be a pair of its components which defines a meta-arrow. Then the conditions from Definition 14 are satisfied iff the graph G' obtained by merging of components U and L is a CG strongly equivalent to G; of course, it is (strictly) larger than G.*

The proof is moved to the Appendix. Note that one has to replace the whole collection of arrows between components with lines; otherwise the obtained graph would not be a CG. This is the reason why legal merging is indeed an elementary operation yielding a larger and strongly equivalent CG.

5.3 Component Merging Procedure

An important fact is that the largest CG in a strong equivalence class $\mathcal{G}$ can be obtained from any CG in $\mathcal{G}$ by consecutive application of the operation of legal merging of components. Actually, we show the following, formally stronger, result.

Proposition 16 *Let G and H be strongly equivalent CGs over N such that $H \geq G$. Then there exists a finite sequence $G \equiv F_1, \ldots, F_m \equiv H$, $m \geq 1$ of CGs over N such that, for every $i = 1, \ldots, m-1$, the graph F_{i+1} is obtained from F_i by legal merging of components.*

The proof is technical and it is moved to the Appendix. Proposition 16 has the following consequence.

Corollary 17 *Given a strong equivalence class $\mathcal{G}$ of CGs over N and $G \in \mathcal{G}$, the largest CG $G^{\dagger}$ in $\mathcal{G}$ is attainable from G by a series of legal mergings.*

Proof We simply put $H = G^{\dagger}$ in Proposition 16. ■

6. Deflagging Procedure

In this section we describe a procedure to construct a deflagged graph $\widehat{G}$ starting from any CG G in the respective AMP equivalence class $\mathbb{H}$. We proceed as follows. First, we introduce a *labeling algorithm* that assigns some labels to endings of lines in G. Second, we introduce a *directing algorithm* which, on the basis of those labels, replaces certain lines in G with arrows. In this way, we get a CG which is both triplex equivalent to G and flag-smaller than G. Finally, we provide a deflagging procedure which consists of repeated application of these two algorithms. We show that the result is a deflagged graph.

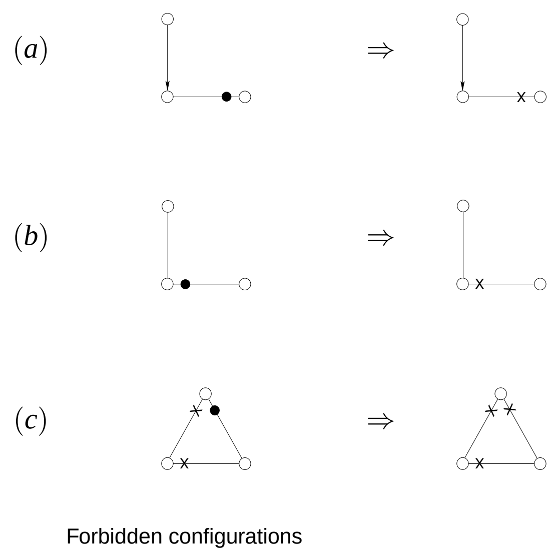


Figure 6: Three blocking rules from the labeling algorithm.

6.1 Labeling Algorithm

Let $G = (N, \mathcal{A}, \mathcal{L})$ be a CG. A *labeled graph* $G^\ell = (N, \mathcal{A}, \mathcal{L}^\ell)$ is a graph obtained by ascribing a pair of labels to every line $\{a,b\} \in \mathcal{L}$. The labels on a line a — b correspond to endings of the line: one of them is associated with a and the other with b. We use two different kinds of labels: a blocking label denoted by a cross, 'x', and a label denoted by a dot, '•', to be read as 'free'. Thus, if the blocking label is associated with a on a — b then we will say that the line is *blocked at a* and write a ×— b in G^ℓ. On the other hand, the notation a •— b in G^ℓ will mean that the line is *free at a*. The intuition behind the terminology is as follows. A blocked ending at a node a will mean that the line cannot be replaced with an arrow directed to a, for otherwise we would get a graph outside $\mathbb{H}$. A free ending at a will mean that no such conclusion has been derived so far.

Consequently, a labeled CG has three types of lines: two symmetric forms ×—× and •—•, and an asymmetric form ×—•. Let us emphasize that we only consider labeled graphs in which all lines have both endings labeled. However, in our notation, labels need not be explicitly indicated. For instance, the notation a —× b in G^ℓ will mean that either a ×—× b in G^ℓ or a •—× b in G^ℓ.

The *labeling algorithm*, whose pseudo-code is given in Algorithm 1, produces a special labeled version G^ℓ of a given CG G. Initially, all lines are replaced with labeled lines with free endings. Then, three *blocking rules*, illustrated in Figure 6, are repeatedly applied until they are not applicable. Each blocking rule modifies just one ending of one line: a free ending is blocked. In this way, we get a labeled CG in which no *forbidden configuration* (see Figure 6) is present. The labelling algorithm is the first step of the overall deflagging procedure and in the following step some lines of G are replaced by arrows; thus, the reader can possibly understand that the three forbidden configurations actually correspond to three unwanted operations: (a) corresponds to cancellation of a triplex, (b) to creation of a triplex and (c) to creation of a semi-directed cycle (of the length 3).

Algorithm 1 Pseudo-code for the `LabelingAlgorithm`(G).

input a CG $G = (N, \mathcal{A}, \mathcal{L})$
put $i = 0$
initialize $G_i^\ell = (N, \mathcal{A}, \mathcal{L}^\ell)$ by replacing every line $a \;\text{——}\; b$ in G by $a \;{\bullet\!\!-\!\!\bullet}\; b$ in G_i^ℓ
while at least one forbidden configuration is present in G_i^ℓ **do**
  $i = i + 1$
  $G_i^\ell =$ modify G_{i-1}^ℓ by applying one of the following rules (see also Figure 6):
    (a) if $a \longrightarrow b \;{-\!\!\!-\!\!\bullet}\; c$ in G_{i-1}^ℓ and a and c are not adjacent then $b \;{-\!\!\!-\!\!\times}\; c$ in G_i^ℓ
    (b) if $a \;\text{——}\; b \;{\bullet\!\!-\!\!\!-}\; c$ in G_{i-1}^ℓ and a and c are not adjacent then $b \;{\times\!\!-\!\!\!-}\; c$ in G_i^ℓ
    (c) if $a \;{-\!\!\!-\!\!\times}\; b \;{-\!\!\!-\!\!\times}\; c \;{\bullet\!\!-\!\!\!-}\; a$ in G_{i-1}^ℓ then $c \;{\times\!\!-\!\!\!-}\; a$ in G_i^ℓ
end while
return $G^\ell = G_i^\ell$

The point is that the result of the labelling algorithm is invariant with respect to the order in which the blocking rules are applied.

Proposition 18 *For any CG G, the labeled graph G^ℓ=*`LabelingAlgorithm`(G) *is unique. This means that the output of the labeling algorithm does not depend on the ordering in which the three blocking rules are applied.*

The proof can be found in the Appendix. In the rest of the paper, G^ℓ will always denote the labeled version of G resulting from the application of Algorithm 1. An example of application of the labeling algorithm is given in Figure 7.

Note that Algorithm 1 is specified so that just one single label is changed in one iteration. This is useful in the proofs of the results of this section, but may be inefficient in practice. A more efficient implementation of the procedure can be achieved by applying the rules (a) and (b) first in a multi-step, and then only applying the rule (c) iteratively. This follows from Proposition 18 and the fact that the application of the rules (a) and (b) does not depend on the result of previous iterations of Algorithm 1.

6.2 Directing Algorithm

The *directing algorithm*, described in Algorithm 2, is the second building block of the deflagging procedure. It replaces some (labeled) lines with arrows in order to possibly reduce the number of flags in the original CG. More precisely, every line of the form $a \;{\times\!\!-\!\!\bullet}\; b$ is replaced with the arrow $a \longrightarrow b$ and then the labels on other lines are removed.

Algorithm 2 Pseudo-code for the `DirectingAlgorithm`(G^ℓ).

input a labeled CG $G^\ell = (N, \mathcal{A}, \mathcal{L}^\ell)$
$G_*^\ell =$ modify G^ℓ by applying the following rule:

$$a \;{\times\!\!-\!\!\bullet}\; b \text{ in } G^\ell \quad \Rightarrow \quad a \longrightarrow b \text{ in } G_*^\ell$$

G'= unlabeled version of G_*^ℓ
return G'

We show that if the directing algorithm is applied to the result of the labeling algorithm then an AMP equivalent graph is obtained.

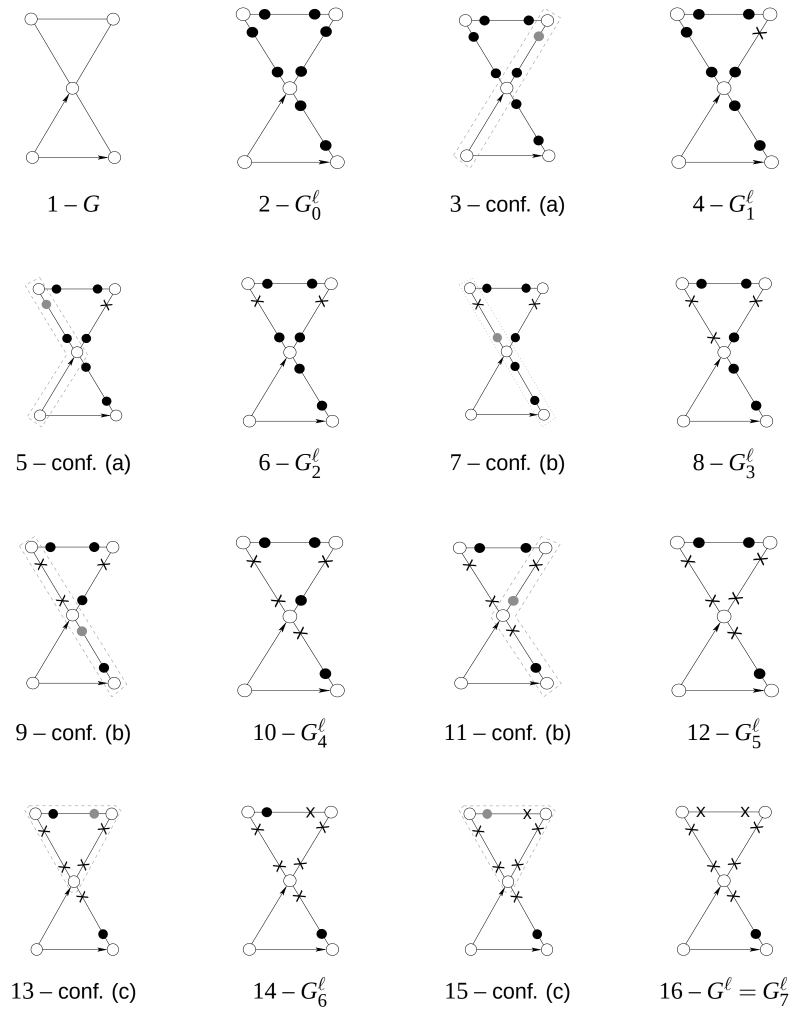


Figure 7: An example of the application of the labeling algorithm to be read following the numbering. Initially, all lines of G are replaced with labeled lines with free endings. Then, in every pair of successive pictures, a forbidden configuration is highlighted and the corresponding rule is applied.

Theorem 19 *Let G be a CG, G^ℓ denote the labeled graph obtained from G by Algorithm 1, and G' the graph resulting from G^ℓ by Algorithm 2. Then G' is a CG which is triplex equivalent to G.*

The proof is relatively long and we have placed it in the Appendix. Clearly, one has $G \geq G'$ and, hence, $\mathcal{G} \succeq \mathcal{G}'$ for the respective equivalence classes. Moreover, one has $G \neq G'$ unless no line is replaced with an arrow in the directing phase. An example of the application of the directing algorithm will be shown in the next section.

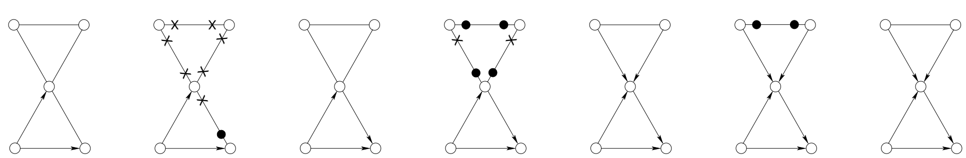
6.3 Overall Procedure

The application of the above algorithms to a CG G produces a graph G' in the same AMP equivalence class such that $G \geq G'$. However, G' still need not be a maximally deflagged graph and one can then apply the same procedure to G'. In Algorithm 3, we provide the pseudo-code of the overall deflagging procedure which consists in repeated application of both algorithms until no line is replaced with an arrow during the directing phase. Its result will be denoted by $\widehat{G}$.

Algorithm 3 Pseudo-code for the `DeflaggingProcedure` (G).

input $G = (N, \mathcal{A}, \mathcal{L})$
$j = 0$
initialize $\widehat{G}_j = G$
repeat
  $j = j + 1$
  $\widehat{G}^\ell_{j-1} = \texttt{LabelingAlgorithm}(\widehat{G}_{j-1})$
  $\widehat{G}_j = \texttt{DirectingAlgorithm}(\widehat{G}^\ell_{j-1})$
until $\widehat{G}_j$ is equal to $\widehat{G}_{j-1}$
return $\widehat{G} = \widehat{G}_j$

Since G has a finite number of lines, the procedure will return a result in finitely many steps. An example of the application of the deflagging algorithm is given in Figure 8. Note that, in this example, $\widehat{G}$ is already the largest deflagged graph from Figure 4; however, this is not true in general.



$G = \widehat{G}_0$ $\widehat{G}^\ell_0$ $\widehat{G}_1$ $\widehat{G}^\ell_1$ $\widehat{G}_2$ $\widehat{G}^\ell_2$ $\widehat{G} = \widehat{G}_3$

Figure 8: An example of the application of the deflagging procedure, where G is the top left graph in Figure 4. Note that the first application of the labeling algorithm, to obtain G^ℓ_0 from G^ℓ, is detailed in Figure 7.

We are to show that $\widehat{G}$ is a deflagged graph, that is, $\widehat{G}$ in $\mathcal{H}^\downarrow$ where $\mathcal{H}^\downarrow$ is the class of deflagged graphs in the respective AMP equivalence class. It follows from Algorithm 3 that $\widehat{G}$ is such that the directing algorithm does not direct any line if applied to the labeled version of $\widehat{G}$. This means, every line in $\widehat{G}^\ell$ is either of the type $\bullet\!\!-\!\!\bullet$ or of the type $\times\!\!-\!\!\times$.

Proposition 20 *Let G be a CG such that there is no line of asymmetric form $\times\!\!-\!\!\bullet$ in its labeled version G^ℓ. Then, every line $a \,—\, b$ in G such that $a \times\!\!-\!\!\times b$ in G^ℓ is a line in every CG F which is triplex equivalent to G.*

The proof is again postponed to the Appendix. Proposition 20 is not valid if the assumption on G is omitted. A counterexample is given in Figure 8 where $G = \widehat{G}_0$ and $F = \widehat{G}$. A consequence of Proposition 20 is that every flag in $\widehat{G}$ is an essential flag.

Corollary 21 *Given a CG G, the graph $\widehat{G} =$ `DeflaggingProcedure`(G) is a deflagged graph, formally $\widehat{G}$ in $\mathcal{H}^{\downarrow}$.*

Proof By Theorem 19, $\widehat{G}$ belongs to the same AMP equivalence class $\mathbb{H}$ as G. Owing to Proposition 8, we need to show that if $a \longrightarrow b \,—\, d$ is a flag in $\widehat{G}$ then it is an essential flag. By the blocking rule (a) $a \longrightarrow b \,—\, d$ in $\widehat{G}$ implies $a \longrightarrow b \,—\!\!\times d$ in $\widehat{G}^\ell$. Since there are no lines of the form $\bullet\!\!-\!\!\times$ in $\widehat{G}^\ell$, it necessitates $a \longrightarrow b \times\!\!-\!\!\times d$ in $\widehat{G}^\ell$. It follows from Proposition 20 that $b \,—\, d$ in H for every $H \in \mathbb{H}$. As $a \longrightarrow b \,—\, d$ has to correspond to a triplex in H, one can conclude that $a \longrightarrow b \,—\, d$ in H. ■

Note that the arguments in the proof above actually imply that a simple sufficient condition for a CG to be deflagged is that its labelled version has no line of asymmetric form.

7. Conclusions

This paper is devoted to the problem of choosing a graphical representative of the statistical model ascribed to a CG under AMP interpretation. As a matter of fact, any CG from the respective AMP Markov equivalence class provides a graphical representative of the corresponding model. However, a representative only makes sense if it complies with some properties that uniquely identify it within each class. Furthermore, in the framework of structural learning, the usefulness of a graphical representative is related to the availability of procedures which can be practically dealt with. That means, for instance, that an implementable construction procedure to obtain the representative (on the basis of any other graph in the Markov equivalence class) should be at our disposal.

Nevertheless, from the point of view of interpretation, a representative should be chosen on the basis of the information carried with respect to the corresponding statistical model. Hereafter, we address the issue of the information contained in the largest deflagged graph, which is the representative for an AMP chain graph model we have proposed.

Andersson et al. (2001, Theorem 4) showed that, for a CG H, the AMP and the LWF Markov properties coincide iff H has no flags. Thus, if there exists a CG without flags in $\mathbb{H}$ then formal distinction between the two Markov properties is not necessary. In this case, all the results derived in the LWF case can be applied. For instance, the useful factorization of conditional densities into ‘potentials’ given by Frydenberg (1990, Theorem 4.1(iii)) can be applied in the AMP case only with respect to CGs without flags. Clearly, there is a strong connection between the set of CGs without flags and the set $\mathcal{H}^{\downarrow}$ of deflagged graphs. More specifically, $\mathbb{H}$ has a CG without flags iff there are no essential flags in $\mathbb{H}$. In this case, the class of deflagged graphs $\mathcal{H}^{\downarrow}$ is just the class of CGs without flags in $\mathbb{H}$. Conversely, if there exists some essential flag in $\mathcal{H}^{\downarrow}$ then one can conclude that there is no CG in $\mathbb{H}$ for which the two Markov properties coincide. Because deflagged graphs only contain essential flags, they eliminate the ambiguity resulting from the non-unique graphical representation of triplexes, and allow an immediate comparison with the LWF case.

The above reasons justify our restriction to the class of deflagged graphs. Now we justify the choice of the largest deflagged graph in $\mathcal{H}^{\downarrow}$. If $\mathcal{H}^{\downarrow}$ contains no flags then $H^{\downarrow}$ is the largest CG without flags in $\mathbb{H}$. Thus, if $\mathbb{H}$ contains an undirected graph then the largest deflagged graph $H^{\downarrow}$ coincides with that undirected graph. Analogously, if $\mathbb{H}$ contains an acyclic directed graph D then $H^{\downarrow}$ coincides with the essential graph D^* for D (Andersson et al., 1997; Studený, 2004; Roverato, 2005). We remark that the AMP essential graph H^* proposed by Andersson et al. (2001) is a deflagged graph (see Andersson and Perlman, 2006, Lemma 3.2(a)) so that $H^* \leq H^{\downarrow}$. Nevertheless, in general, the largest deflagged graph $H^{\downarrow}$ is different from the AMP essential graph H^*: for instance, if $\mathbb{H}$ contains an undirected graph then H^* may even have some arrows (see Andersson et al., 2001, Figure 14).

Another issue related to the problem of representative choice is the topic of causal discovery in CGs (see Section 11.2 of Lauritzen, 2001). This is a controversial topic (see Section 3 of Dawid, 2002, for more discussion). The disputable question is whether one can identify some causal relationships between variables on the basis of data. Nevertheless, what we think that what is generally accepted in the field of causal discovery is the following proposition:

> If data are "generated" from a distribution which is "faithful" with respect to a CG and if an arrow $a \longrightarrow b$ is *not* invariant across the respective Markov equivalence class, then one *cannot* reveal possible causal relationship from a to b on basis of data.

In short, one cannot make causal discovery between a and b if there is an *undirected* edge between a and b in at least one of the chain graphs from the Markov equivalence class, or if there are two chain graphs such that $a \longrightarrow b$ in the one of them first and $b \longrightarrow a$ in the latter one. On the other hand, if an arrow $a \longrightarrow b$ is invariant across the respective Markov equivalence class then causal discovery *could* be possible. Consequently, from the point of view of causal discovery in chain graphs, a good representative of a Markov equivalence class should indicate that the corresponding edge is not an invariant arrow by the presence of a line. Standard representatives in the LWF case, such as the largest CGs (Studený, 1997), the essential graphs for acyclic directed graphs (Andersson et al., 1997), and the $\mathcal{B}$-essential graphs (Roverato and La Rocca, 2006), are fully informative from this point of view because they have the largest number of lines and, furthermore, they contain an arrow if and only if it is invariant. As the examples in Figures 2 and 4 show, a CG with this property may not exist in an AMP equivalence class and therefore both the AMP essential graph and the largest deflagged graph may contain some arrows that are not invariant. However, the largest deflagged graph is more informative than the AMP essential graph because it is a larger chain graph and, therefore, it has more lines.

We have not mentioned this explicitly but, in this paper, we have actually provided an algorithmic characterization of the largest deflagged graphs. More specifically, a CG G is the largest deflagged graph iff it is again obtained by the consecutive application of two procedures: the deflagging procedure is applied to G and the component merging procedure to its result $\widehat{G}$.

The results of the paper also lead to some natural open problems. For instance, we would like to know whether the converse of Proposition 20 is valid. More specifically, does the deflagging procedure identify all essential lines in $\mathbb{H}$ as double-blocked lines? Further conjecture is that the AMP essential graph is obtained if the deflagging procedure is applied to the largest deflagged graph. Another issue is as follows. We know that both LWF and AMP Markov equivalence are associated to Markov properties for CGs. Is there any Markov property for CGs which gives rise to the strong equivalence of CGs?

Acknowledgements

We are indebted to Nanny Wermuth who invited both of us to a workshop held in Wiesbaden in September 2002. We started there a discussion, which resulted in this paper. We also thank Michael Perlman for useful comments during the Barcelona meeting in 2004 and the anonymous referees for useful suggestions. Financial support to the fist author has been provided by Miur, PRIN03 and PRIN05 n. 134079. The research of the second author has been supported by GAČR n. 201/04/0393.

Appendix A. Proofs

Proof of Proposition 12

Throughout the proof we assume that G and H are strongly equivalent CGs. Let $G \cup H$ and $G \vee H$ denote the graphs introduced in Lemma 11. We start with an auxiliary observation.

Fact 1 *Let $d_0 \longrightarrow d_1$ be a cyclic arrow in $G \cup H$ and $\rho : d_0, d_1, \ldots, d_m \equiv d_0$, $m \geq 3$ a semi-directed cycle in $G \cup H$ containing it which cannot be shortened (to a semi-directed cycle in $G \cup H$ containing $d_0 \longrightarrow d_1$ of the length $l < m$). Then $d_2 \longrightarrow d_1$ in one of the graphs G and H while $d_0 \longrightarrow d_2$ in the other graph.*

Proof Since G is a CG, there exist $2 \leq j \leq m$ with $d_{j-1} \longleftarrow d_j$ in G and the same conclusion holds for H. Let us put

$$s = \min\{\, 2 \leq j \leq m \,;\, d_{j-1} \longleftarrow d_j \text{ either in } G \text{ or in } H \,\}.$$

Let us, without loss of generality, assume that $d_{s-1} \longleftarrow d_s$ in G. Then $d_0, \ldots, d_{s-1}$ is a descending path G. Moreover, observe that $d_1, \ldots, d_s$ is necessarily a descending path in the other graph, namely in H. This implies $s < m$ for otherwise ρ is a semi-directed cycle in a CG H.

The next step is to verify that $[d_{s-2}, d_s]$ is an edge in $G \cup H$. This is because otherwise $d_s \longrightarrow d_{s-1} \longleftarrow d_{s-2}$ is an immorality in G or $d_s \longrightarrow d_{s-1} \;—\; d_{s-2}$ is a flag in G, which, by strong equivalence of G and H, implies that $d_s \longrightarrow d_{s-1}$ in H and this contradicts the assumption that ρ is a semi-directed cycle in $G \cup H$.

Since H is a CG and d_{s-2}, d_{s-1}, d_s a descending path in H, one has either $d_s \longleftarrow d_{s-2}$ or $d_s \;—\; d_{s-2}$ in H, and, therefore, in $G \cup H$.

Thus, necessarily $s = 2$; otherwise ρ could be shortened in $G \cup H$ by the edge $[d_{s-2}, d_s]$ to get a shorter semi-directed cycle containing $d_0 \longrightarrow d_1$ which would contradict its definition. Thus, $d_2 \longrightarrow d_1$ in G. The facts that H is a CG, $[d_0, d_2] = [d_{s-2}, d_s]$ is an edge in H, $d_0 \longrightarrow d_1$ in H and either $d_1 \longrightarrow d_2$ or $d_1 \;—\; d_2$ in H imply that $d_0 \longrightarrow d_2$ in H. ■

Fact 2 *There is no cyclic arrow* $a \longrightarrow c$ *in* $G \cup H$ *which belongs either to an immorality* $a \longrightarrow c \longleftarrow b$ *or to a flag* $a \longrightarrow c$ —— b *in* $G \cup H$.

Proof For a contradiction, suppose that at least one such cyclic arrow exists. Choose a semi-directed cycle $\rho : d_0, d_1, \ldots, d_m \equiv d_0$, $m \geq 3$ in $G \cup H$ of shortest possible length among all semi-directed cycles containing an arrow of this kind. Assume that $d_0 = a \longrightarrow c = d_1$ is that arrow in $G \cup H$ and, using Fact 1, observe that $d_2 \longrightarrow d_1$ in one of the graph, say in G, while $d_0 \longrightarrow d_2$ in the other graph H.

Consider the induced subgraph over $\{a, c, b\}$ mentioned in the formulation of Fact 2. As $[a, b]$ is not an edge in $G \cup H$ whereas $[d_0, d_2] = [a, d_2]$ is an edge in H, one has $d_2 \neq b$. Observe that $c \longleftarrow b$ or c —— b in G. Indeed, otherwise $c \longrightarrow b$ in G implies $\neg(c \longrightarrow b \text{ in } H)$ by the assumption of Fact 2, and H has either an immorality $a \longrightarrow c \longleftarrow b$ or a flag $a \longrightarrow c$ —— b. By strong equivalence of G and H, G has the same induced subgraph for $\{a, c, b\}$, which contradicts the fact $c \longrightarrow b$ in G. By interchange of G and H derive that $c \longleftarrow b$ or c —— b in H as well.

This allows one to see that $[b, d_2]$ is an edge in $G \cup H$ as otherwise the induced subgraph of G for $\{d_2, d_1 = c, b\}$ having $d_2 \longrightarrow d_1$ coincides, by strong equivalence of G and H, with the subgraph of H and the conclusion $d_2 \longrightarrow d_1$ in H contradicts the assumption that ρ is a semi-directed cycle in $G \cup H$. Since b, c, d_2 is a descending path in H one has either $b \longrightarrow d_2$ or b —— d_2 in H.

Thus, H has either an immorality $d_0 \longrightarrow d_2 \longleftarrow b$ or a flag $d_0 \longrightarrow d_2$ —— b. Since G and H are strongly equivalent, G has the same induced subgraph for $\{d_0, d_2, b\}$. Of course, the same conclusion holds for $G \cup H$ and $d_0 \longrightarrow d_2$ is an arrow in $G \cup H$ belonging to a triplex.

Hence, it is impossible that $m > 3$ as otherwise ρ can be shortened to $d_0, d_2, \ldots, d_m = d_0$ by a cyclic arrow $d_0 \longrightarrow d_2$ of the considered type which contradicts its definition. However, if $m = 3$ then the fact $d_3 \equiv d_0 \longrightarrow d_2$ in $G \cup H$ contradicts the assumption that ρ is a semi-directed cycle in $G \cup H$. ■

Observe easily by contradiction that if G and H are strongly equivalent then

[d] if $a \longrightarrow c$ both in G and in H then an induced subgraph $a \longrightarrow c \longrightarrow b$ occurs in H iff it occurs in G.

This observation is used in the proof of the following fact and also later.

Fact 3 *There is no cyclic arrow* $c \longrightarrow b$ *in* $G \cup H$ *which belongs to an induced subgraph* $a \longrightarrow c \longrightarrow b$ *in* $G \cup H$.

Proof For a contradiction, suppose that such an arrow exists. Choose a semi-directed cycle $\rho : d_0, d_1, \ldots, d_m \equiv d_0$, $m \geq 3$ in $G \cup H$ of shortest possible length among all semi-directed cycles containing an arrow of this kind. More specifically, assume that $d_0 = c \longrightarrow b = d_1$ in $G \cup H$. By Fact 1 observe that one can assume $d_2 \longrightarrow d_1$ in G and $d_0 \longrightarrow d_2$ in H. One has either $d_2 \longleftarrow d_1$ or d_2 —— d_1 in H for otherwise $d_2 \longrightarrow d_1$ in $G \cup H$ contradicts the assumption that ρ is a semi-directed cycle. As $a \longrightarrow c$ in H whereas $d_2 \longleftarrow d_0 = c$ in H, one has $a \neq d_2$.

Observe, by contradiction, that $[a, d_2]$ is not an edge in $G \cup H$. Indeed, otherwise $a \longrightarrow c = d_0 \longrightarrow d_2$ in a CG H implies $a \longrightarrow d_2$ in H and H has either an immorality $a \longrightarrow d_2 \longleftarrow d_1$ or a flag $a \longrightarrow d_2$ —— d_1. Thus, G has the same subgraph for $\{a, d_2, d_1\}$ which contradicts the fact $d_2 \longrightarrow d_1$ in G.

Thus, H has an induced subgraph $a \longrightarrow c = d_0 \longrightarrow d_2$, which implies, by the condition [d] mentioned above Fact 3, that G has the same induced subgraph. In particular, $G \cup H$ has this induced subgraph as well. Thus, necessarily $m \leq 3$ for otherwise ρ can be shortened in $G \cup H$ by $d_0 \longrightarrow d_2$ to get a shorter cycle of the required type. If $m = 3$ then $d_3 = d_0 \longrightarrow d_2$ in $G \cup H$ implies a contradictory conclusion that ρ is not a semi-directed cycle in $G \cup H$. ■

Now, the proof of Proposition 12 follows directly from Lemma 11 and Facts 2 and 3. Since G and H are strongly equivalent an immorality or a flag in G occurs also in H and, therefore, in $G \cup H$. By Fact 2 it is preserved in $G \vee H$. Conversely, if $a \longrightarrow c \longleftarrow b$ is an immorality in $G \vee H$ then it is also in G. If $a \longrightarrow c \longrightarrow b$ is a flag in $G \vee H$ then $a \longrightarrow c$ both in G and in H. The option $c \longleftarrow b$ in one of the graphs G and H is excluded because then the graph has an immorality $a \longrightarrow c \longleftarrow b$, which is saved in $G \vee H$. If $c \longrightarrow b$ in both graphs then $G \cup H$ has an induced subgraph $a \longrightarrow c \longrightarrow b$. By Fact 3 the arrow $c \longrightarrow b$ remains in $G \vee H$ which contradicts the assumption. Thus, $c \;\text{—}\; b$ either in G or in H and this implies, by their strong equivalence, that the flag $a \longrightarrow c \;\text{—}\; b$ is in G.

Proof of Proposition 15

This proposition is analogous to the result on LWF equivalence and feasible merging given in Theorem 8 of Roverato (2005). It says this:

> Given a CG G and a meta-arrow $U \rightrightarrows L$ in G, the conditions (i) and (ii) from Definition 3 form together a necessary and sufficient condition for the graph G' obtained by merging U and L to be a CG which is complex equivalent to G.

In fact, we utilize this result in our proof of Proposition 15. Recall that the condition [i] from Definition 14 is identical to the condition (i) from Definition 3 and the condition [ii] from Definition 14 is stronger than (ii) from Definition 3.

Proof First, we are going to verify the necessity of the conditions [i]-[iii]. Since strong equivalence of CGs implies their complex equivalence the necessity of conditions (i)-(ii) follows from Theorem 8 in Roverato (2005). The conditions [i] and (i) are identical, but [ii] is stronger than (ii). Indeed, [ii] requires equality of sets $\mathrm{pa}_G(L) \setminus U$ and $\mathrm{pa}_G(b)$ for every $b \in K$ whereas (ii) only requires $\mathrm{pa}_G(L) \setminus U \subseteq \mathrm{pa}_G(b)$.

Thus, to verify [ii] it suffices to show

- $\forall b \in K \quad \mathrm{pa}_G(b) \subseteq \mathrm{pa}_G(L) \setminus U$.

Suppose for contradiction that $b \in K$ and $a \in \mathrm{pa}_G(b)$ exists with $a \notin \mathrm{pa}_G(L) \setminus U$. Then $d \in L$ exists such that $b \longrightarrow d$ in G. Of course, $a \neq d$ and, since G is a CG, the options $a \longleftarrow d$ and $a \;\text{—}\; d$ in G cannot occur. The option $a \longrightarrow d$ is excluded by the assumption $a \notin \mathrm{pa}_G(L) \setminus U$. If $[a, d]$ is not an edge in G then G has an induced subgraph $a \longrightarrow b \longrightarrow d$ while G' has a flag $a \longrightarrow b \;\text{—}\; d$ which contradicts the assumption that they are strongly equivalent.

The next step is to verify the necessity of the condition

$[\widetilde{\text{iii}}]$ for every $d \in L$ one has $\mathrm{pa}_G(L) \subseteq \mathrm{pa}_G(d)$,

which is an equivalent formulation of [iii]. Let us fix $d \in L$. Given $b \in \mathrm{pa}_G(L)$, to show that $b \longrightarrow d$ in G two cases can be distinguished.

- $b \in U$, that is, $b \in K$.
  Then there exists $g \in L$ with $b \longrightarrow g$ in G. If $g \neq d$ then one can consider a path $b \longrightarrow g \;—\; \ldots \;—\; d$ in G and its shortening ρ which cannot be shortened any more. If $[b,d]$ is not an edge in G then G has a flag $b \longrightarrow e \;—\; f$ composed of nodes of ρ. This contradicts the assumption that G and G' are strongly equivalent since one has $b \;—\; e$ in G' by definition of merging. Thus, $[b,d]$ is an edge and $b \longrightarrow d$ in G since G is a CG.

- $b \in \mathrm{pa}_G(L) \setminus U$.
  Observe that $a \in K$ exists by Definition 1 and $b \longrightarrow a$ follows from (ii). Moreover, one has $a \longrightarrow d$ in G by the previous case. If $[b,d]$ is not an edge then G has an induced subgraph $b \longrightarrow a \longrightarrow d$ while G' has a flag $b \longrightarrow a \;—\; d$ which contradicts the assumption that they are strongly equivalent. Thus, $[b,d]$ is an edge, namely $b \longrightarrow d$ in G because G is a CG.

This concludes the proof of the necessity of conditions [i]-[iii].

Second, we prove the sufficiency of those conditions. Since they imply the conditions (i)-(ii) from Definition 3, it follows from Theorem 8 in Roverato (2005) that G' is a CG which is complex equivalent to G but strictly larger. In particular, G and G' have the same immoralities and, to show that they are strongly equivalent, it suffices to verify that they have identical flags.

If $a \longrightarrow b \;—\; d$ is a flag in G then we are to show that it is a flag in G'. The only option which avoids the desired conclusion is $a \in U$ and $b \in L$. However, then $d \in L$ and by [iii] observe $a \longrightarrow d$ in G which contradicts the assumption.

If $a \longrightarrow b \;—\; d$ is a flag in G' then the fact $G' \geq G$ implies $a \longrightarrow b$ in G and the only option which avoids the desired conclusion that $a \longrightarrow b \;—\; d$ is a flag in G is that $[b,d]$ was modified. There are basically two cases.

- If $b \in L$ and $d \in U$ then $a \longrightarrow b \longleftarrow d$ is an immorality in G and, because of complex equivalence of graphs, also in G'. This contradicts the assumption.

- If $b \in U$ and $d \in L$ then observe $b \in K$ and by [ii] $a \in \mathrm{pa}_G(b) \subseteq \mathrm{pa}_G(L) \setminus U$. By [iii] get $a \in \mathrm{pa}_G(d)$ which contradicts the assumption.

Thus, the sufficiency proof is finished. ■

Proof of Proposition 16

Basic observation which is needed is as follows.

Fact 4 *Let E,F,G be CGs over N with the same underlying graph such that $E \geq F \geq G$ and the following condition holds for any $c \in N$:*

[e] *if there exists $a \in N$ with $a \;—\; c$ in E and $a \longrightarrow c$ in F then for every $b \in N$ with $c \;—\; b$ in F one has $c \;—\; b$ in G.*

If E and G are strongly equivalent then F is strongly equivalent to them as well.

Note that the conclusion of Fact 4 need not be valid if the condition [e] is omitted: consider $N = \{a,b,c\}$, E an undirected graph with $a —— c —— b$, F a CG with $a \longrightarrow c —— b$ and G a directed graph with $a \longrightarrow c \longrightarrow b$.

Proof We can show that F is strongly equivalent to E. If $a \longrightarrow c \longleftarrow b$ is an immorality in E then $E \geq F$ implies that it is an immorality in F. Conversely, if $a \longrightarrow c \longleftarrow b$ is an immorality in F then $F \geq G$ implies that it is an immorality in G and, therefore, in E.

If $a \longrightarrow c —— b$ is a flag in E, then it is a flag in G which implies $c —— b$ in F by $F \geq G$. Since $a \longrightarrow c$ in F by $E \geq F$ the graph F has a flag $a \longrightarrow c —— b$.

If $a \longrightarrow c —— b$ is a flag in F, then $F \geq G$ implies $a \longrightarrow c$ in G. We first verify $a \longrightarrow c$ in E by excluding two other variants of the edge $[a,e]$ in E. Since $E \geq F$ the case $a \longleftarrow c$ in E is excluded. The case $a —— c$ in E is also excluded, this time owing to the condition [e] from the assumption of Fact 4. Indeed, [e] says $c —— b$ in G, which implies that $a \longrightarrow c —— b$ is a flag in G and, therefore, in E, which contradicts the assumption $a —— c$ in E. Thus, $a \longrightarrow c$ in E and the aim is to show $c —— b$ in G. It can be shown by contradiction.

- If $c \longleftarrow b$ in G then $a \longrightarrow c \longleftarrow b$ is an immorality in G and, therefore, in E, which implies, by $E \geq F$, a contradictory conclusion $c \longleftarrow b$ in F.

- If $c \longrightarrow b$ in G then $a \longrightarrow c \longrightarrow b$ is an induced subgraph in G. By the condition [d] mentioned above Fact 3 applied to G and E, it is also an induced subgraph in E. The assumption $E \geq F$ then implies a contradictory conclusion $c \longrightarrow b$ in F.

Hence, $a \longrightarrow c —— b$ is a flag in G, and therefore in E. ■

The main step is the following ‘sandwich lemma’.

Fact 5 *Let G,E be strongly equivalent CGs, $E \geq G$, $E \neq G$. Then there exists a CG F which is strongly equivalent to G and E, such that $E \geq F \geq G$ and E is obtained from F by legal merging of components.*

Note that the idea of the proof of this proposition is analogous to the proof of Theorem 7 in Roverato (2005).

Proof Since $E \geq G$, every component in E is the union of components in G and the assumption $E \neq G$ implies that there exists a component C in E containing at least two components in G. As G_C is a CG one can find a terminal component T in it. By the construction $C \setminus T \neq \emptyset$ and there is an arrow from $C \setminus T$ to T in G. Let us construct a hybrid graph F from E by replacement of all lines between $C \setminus T$ and T in E by arrows from $C \setminus T$ to T. Observe the following facts.

$\{$**a**$\}$ F is a CG.

Assume for contradiction that F has a semi-directed cycle ρ. Since $F_{N \setminus C} = E_{N \setminus C}$ is a CG and F_C is a CG by construction, ρ has an edge between $N \setminus C$ and C, namely an arrow. This arrow is also an arrow in E (with the same direction); the other arrows of ρ either are kept in E or become lines, the lines of ρ retain in E. Therefore, ρ has to be a semi-directed cycle in E, which contradicts the assumption.

$\{\mathbf{b}\}$ $E \geq F \geq G$ and $F \neq E$.

The fact $E \geq F$ is evident. To see $F \geq G$ observe that if $a \longrightarrow b$ in F then either $a \longrightarrow b$ in E in which case $E \geq G$ implies $a \longrightarrow b$ in G, or $a \relbar\!\!\relbar b$ in E. In the latter case $a \in C \setminus T$ and $b \in T$ which also implies, by the definition of T, that $a \longrightarrow b$ in G.

$\{\mathbf{c}\}$ $C \setminus T$ is a connected set in F and, therefore, it is a component in F.

Indeed, suppose for contradiction that distinct $a, b \in C \setminus T$ exist which are not connected by an undirected path in $F_{C\setminus T} = E_{C\setminus T}$. Since C is a connected set in E, one can construct a path $\tilde{a} \longrightarrow c_1 \relbar\!\!\relbar \ldots \relbar\!\!\relbar c_m \longleftarrow \tilde{b}$, $m \geq 1$ in F with some $c_1, \ldots, c_m \in T$ and $\tilde{a}, \tilde{b} \in C \setminus T$ such that $[\tilde{a}, \tilde{b}]$ is not an edge in F. This path has the same form in G and can be shortened to a complex in G. This complex is not in E which contradicts the assumption that E and G are strongly equivalent since strong equivalence implies complex equivalence.

$\{\mathbf{d}\}$ F is strongly equivalent to G and E.

This follows from Fact 4 owing to $\{\mathrm{a}\}$ and $\{\mathrm{b}\}$. The condition [e] from Fact 4 holds because of the construction of F: if $a \relbar\!\!\relbar c$ in E and $a \longrightarrow c$ in F then $c \in T$ and $c \relbar\!\!\relbar b$ in F implies $b \in T$ for which reason $c \relbar\!\!\relbar b$ in G.

Now, the conclusion that E is made of F by legal merging of components is easy to see. The condition $\{\mathrm{c}\}$ implies that both $C \setminus T$ and T are components in F and E is obtained from F by merging of the upper component $U \equiv C \setminus T$ and the lower component $L \equiv T$. Since E and F are strongly equivalent is follows from Proposition 15 that the merging is legal. ■

Now, the proof of Proposition 16 is easy. The required sequence $G = F_1, \ldots, F_m = H$, $m \geq 1$ can be constructed backwards by consecutive application of Fact 5 to G and $E \equiv F_i$ to get $F_{i-1} = F$ until F_{i-1} is the graph G. Of course, one starts with $F_m = H$, where $m-1$ is the difference between the numbers of components of G and H.

Proof of Proposition 18

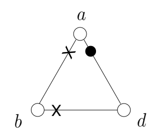
Assume for contradiction that two different orderings of applications of blocking rules leads to two different labeled graphs $G^{\ell(1)}$ and $G^{\ell(2)}$. Since they only differ in their labels, one can assume without loss of generality that $G^{\ell(1)}$ has at least one blocked label that is ‘free’ in $G^{\ell(2)}$. Let us fix a sequence of iterations $G_0^{\ell(1)}, G_1^{\ell(1)}, \ldots, G_n^{\ell(1)} = G^{\ell(1)}$, $n \geq 2$ leading to $G^{\ell(1)}$. Let $G_i^{\ell(1)}$ be the first graph in this sequence which has a blocked label that is ‘free’ in $G^{\ell(2)}$, say $a \mathrel{\times\!\!\relbar\!\!\relbar} d \in G_i^{\ell(1)}$ and $a \mathrel{\bullet\!\!\relbar\!\!\relbar} d \in G^{\ell(2)}$. In particular, $b \mathrel{\times\!\!\relbar\!\!\relbar} c$ in $G_j^{\ell(1)}$ for $j < i$ implies $b \mathrel{\times\!\!\relbar\!\!\relbar} c$ in $G^{\ell(2)}$.

We now show that $a \mathrel{\times\!\!\relbar\!\!\relbar} d \in G_i^{\ell(1)}$ and $a \mathrel{\bullet\!\!\relbar\!\!\relbar} d$ in $G^{\ell(2)}$ implies that $G^{\ell(2)}$ has a forbidden configuration, which contradicts the assumption. There are three possible cases.

1. If $a \mathrel{\times\!\!\relbar\!\!\relbar} d$ in $G_i^{\ell(1)}$ is blocked at a by the rule (a) then there exists a vertex b such that $b \longrightarrow d \relbar\!\!\relbar a$ is a flag in G (cf. Algorithm 1). In particular, $b \longrightarrow d \mathrel{\relbar\!\!\relbar\!\!\bullet} a$ in $G^{\ell(2)}$ is a forbidden configuration in $G^{\ell(2)}$.

2. If $a \mathrel{\times\!\!\relbar\!\!\relbar} d$ in $G_i^{\ell(1)}$ is blocked at a by (b) then there exists a vertex b with $b \relbar\!\!\relbar a \relbar\!\!\relbar d$ in G, while $[b, d]$ is not an edge in G. Then $b \relbar\!\!\relbar a \mathrel{\bullet\!\!\relbar\!\!\relbar} d$ is a forbidden configuration in $G^{\ell(2)}$.

3. If $a \times\!\!\!\!\relbar\!\!\relbar\, d$ in $G_i^{\ell(1)}$ is blocked at a by the rule (c) then there exists a vertex b such that the following forbidden configuration



appears in $G_{i-1}^{\ell(1)}$. As mentioned above, blocking labels in $G_j^{\ell(1)}$ for $j < i$ also occur in $G^{\ell(2)}$. Thus, that forbidden configuration is also present in $G^{\ell(2)}$.

Proof of Theorem 19

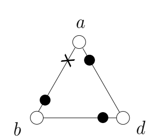
Throughout the proof we assume that G is a CG, G^ℓ the labeled version of G obtained from G by the labeling algorithm and G' the hybrid graph obtained from G^ℓ by the directing algorithm. The overall aim is to show that G' is a CG triplex equivalent to G. To improve the readability of the proof, we split it into more elementary facts. The first goal is to show that G' has no semi-directed cycle of the length 3. This is the main step to show that it has no semi-directed cycles at all, that is, it is a CG. Finally, we prove that G' is triplex equivalent to G.

We start with two auxiliary facts.

Fact 6 *If there is a semi-directed cycle in G' then it is undirected in G.*

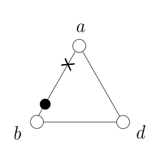
Proof Assume for contradiction that $\rho : d_0, \ldots, d_{n-1}, d_n = d_0$, $n \geq 3$ is a semi-directed cycle in G' which has an arrow $d_0 \longrightarrow d_1$ in G. Since G is a CG, there exists an arrow $d_{i-1} \longleftarrow d_i$, $2 \leq i \leq n$ in G. Basic observation is that arrows in G are kept in G' with the same direction. In particular, $d_0 \longrightarrow d_1$ and $d_{i-1} \longleftarrow d_i$ in G', which contradicts the assumption that ρ is a semi-directed cycle in G'. ■

Fact 7 *If $\rho : a, b, d, a$ is a semi-directed cycle of the length 3 in G' with $a \longrightarrow b$ in G' then it corresponds to the following configuration in G^ℓ:*

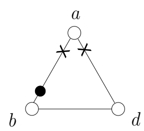


(5)

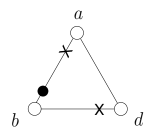
Proof By Fact 6, ρ consists of lines in G. As $a \longrightarrow b$ in G', it follows from Algorithm 2 that ρ corresponds to the following configuration



in G^ℓ. We only need to show that ρ cannot occur in either of the following two configurations in G^ℓ:

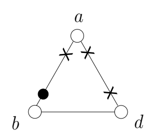


or

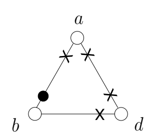


(A) (B)

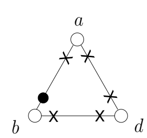
Consider the case (A) and observe that a —— d also has to have a blocked ending at d in G^ℓ. Indeed, otherwise by Algorithm 2 $a \longrightarrow d$ in G' and ρ is not a semi-directed cycle in G', which contradicts the assumption. Hence, we have



in G^ℓ. Now, it follows from Algorithm 1 that b —— d has a blocked ending at d. Indeed, otherwise a forbidden configuration $b -\!\!\!-\!\!\!\times a -\!\!\!-\!\!\!\times d \bullet\!\!\!-\!\!\!- b$ of type (c) exists in G^ℓ. Thus, the situation is as follows:



Again, b —— d has a blocked ending at b for otherwise, by Algorithm 2, $b \longleftarrow d$ in G' contradicts the assumption that ρ is a semi-directed cycle. Thus, $G^\ell_{\{a,b,d\}}$ looks like



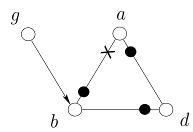
which is, however, also impossible because $a -\!\!\!-\!\!\!\times d -\!\!\!-\!\!\!\times b \bullet\!\!\!-\!\!\!- a$ is a forbidden configuration of type (c) in G^ℓ. Hence, the configuration (A) cannot occur. Using the same kind of reasoning, it is also easy to check that the configuration (B) in G^ℓ is impossible. This is left to the reader. ■

Fact 8 *G' has no semi-directed cycle of the length 3.*

Proof Suppose for contradiction that G' has a semi-directed cycle of the length 3. Thus, the set $\mathcal{A}'$ of arrows in G' belonging to (at least one of) those cycles is assumed to be non-empty. By Fact 6, every arrow $e \longrightarrow f$ in $\mathcal{A}'$ corresponds to a line e —— f in G, and, therefore, by the directing algorithm, to a labeled line $e \times\!\!\!-\!\!\!-\!\!\bullet f$ in G^ℓ. Let us fix a sequence $G^\ell_0, \ldots, G^\ell_n$, $n \geq 1$ of labeled CGs generated by the labeling algorithm. Clearly, every $e \longrightarrow f$ in $\mathcal{A}'$ is assigned the unique $1 \leq i \leq n$ such that $e \times\!\!\!-\!\!\!- f$ in G^ℓ_i and $e \bullet\!\!\!-\!\!\!- f$ in G^ℓ_j for $j < i$. Let $a \longrightarrow b$ denote that arrow in $\mathcal{A}'$ which has assigned the least such i. In particular, if $e \times\!\!\!-\!\!\!- f$ in G^ℓ_{i-1} then $e \longrightarrow f$ does not belong to $\mathcal{A}'$.

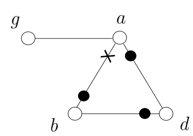
Let us fix a semi-directed cycle $\rho : a, b, d, a$ of the length 3 in G' containing $a \longrightarrow b$. By Fact 7, the subset of vertices $\{a, b, d\}$ corresponds to the configuration (5) in G^ℓ and, because of the construction of G^ℓ, also in G^ℓ_i. Now, we show that the occurrence of (5) in G^ℓ_i leads to contradiction because the line $a \mathrel{\bullet\!\!-\!\!\!-\!\!\bullet} b$ in the previous iteration G^ℓ_{i-1} cannot be blocked at a by any of the blocking rules from Algorithm 1.

1. If $b \mathrel{\bullet\!\!-\!\!\!-\!\!\times} a$ is blocked at a on basis of the rule (a) then there exists a vertex $g \in N$, not adjacent to a, such that $g \longrightarrow b$ in G. If we add this arrow to the configuration (5) in G^ℓ above then we obtain (possibly omitting an edge between g and d)



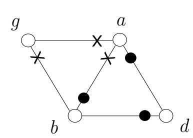
   Nodes g and d are necessarily adjacent for otherwise G^ℓ has a forbidden configuration $g \longrightarrow b \mathrel{-\!\!\!-\!\!\bullet} d$ of the type (a). Actually, one has $g \longrightarrow d$ in G as otherwise G has a semi-directed cycle g, b, d, g. However, then $g \longrightarrow d \mathrel{-\!\!\!-\!\!\bullet} a$ is a forbidden configuration of type (a) in G^ℓ, which is impossible.

2. If $a \mathrel{\times\!\!-\!\!\!-\!\!\bullet} b$ is blocked at a on basis of the rule (b) then there exists a vertex $g \in N$, not adjacent to b, such that $g \mathrel{-\!\!\!-} a$ in G. If we add this line to the configuration (5) in G^ℓ and obtain (possibly omitting an edge between g and d)

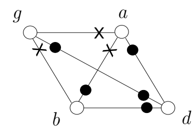


   Nodes g and d have to be adjacent for otherwise a forbidden configuration $g \mathrel{-\!\!\!-} a \mathrel{\bullet\!\!-\!\!\!-} d$ of the type (b) exists in G^ℓ. As G is a CG, one has $g \mathrel{-\!\!\!-} d$ in G. However, then $g \mathrel{-\!\!\!-} d \mathrel{\bullet\!\!-\!\!\!-} b$ is a forbidden configuration of type (b) in G^ℓ, which is impossible.

3. If $a \mathrel{\times\!\!-\!\!\!-\!\!\bullet} b$ is blocked at a on basis of the rule (c) then there exists a vertex g such that $b \mathrel{-\!\!\!-\!\!\times} g \mathrel{-\!\!\!-\!\!\times} a \mathrel{\bullet\!\!-\!\!\!-} b$ in G^ℓ_{i-1}. As $g \mathrel{-\!\!\!-\!\!\times} a$ and $d \mathrel{-\!\!\!-\!\!\bullet} a$ in G^ℓ one has $g \neq d$. Thus, the following configuration occurs in G^ℓ, where the possible edge between g and d is omitted:



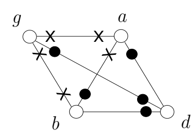
   The nodes g and d have to be adjacent for otherwise $g \mathrel{-\!\!\!-} a \mathrel{\bullet\!\!-\!\!\!-} d$ would be a forbidden configuration of type (b) in G^ℓ. Since G is a CG, one has $g \mathrel{-\!\!\!-} d$ in G. The ending of $g \mathrel{-\!\!\!-} d$ at g in G^ℓ has to be free as otherwise $d \mathrel{-\!\!\!-\!\!\times} g \mathrel{-\!\!\!-\!\!\times} a \mathrel{\bullet\!\!-\!\!\!-} d$ would be a forbidden configuration of type (c) in G^ℓ. Analogously, its ending at d in G^ℓ is also free for otherwise $b \mathrel{-\!\!\!-\!\!\times} g \mathrel{-\!\!\!-\!\!\times} d \mathrel{\bullet\!\!-\!\!\!-} b$ would be a forbidden configuration of type (c) in G^ℓ. Thus, $g \mathrel{-\!\!\!-} d$ has both endings free in G^ℓ and



To show that $a \,\text{---}\, g$ is blocked at g in G^ℓ recall that $a \mathbin{\times\!\!-\!\!-} g$ in G^ℓ_{i-1}. If $a \mathbin{\times\!\!-\!\!\bullet} g$ in G^ℓ then Algorithm 2 implies that a,g,d,a is a semi-directed cycle in G' (note that $d \mathbin{-\!\!-\!\!\bullet} a$ in G^ℓ implies that either $d \longrightarrow a$ or $d \,\text{---}\, a$ in G'). This, however, means that $a \longrightarrow g$ belongs to $\mathcal{A}'$, which contradicts the choice of $a \longrightarrow b$: as mentioned above, that choice ensures that if $e \mathbin{\times\!\!-\!\!-} f$ in G^ℓ_{i-1} then $e \longrightarrow f$ does not belong to $\mathcal{A}'$.

The conclusion that $g \,\text{---}\, b$ is blocked at b in G^ℓ can be derived analogously. If $g \mathbin{\times\!\!-\!\!\bullet} b$ in G^ℓ then Algorithm 2 implies that g,b,d,g is a semi-directed cycle in G'. Then $g \longrightarrow b$ belongs to $\mathcal{A}'$ which is not possible because of the fact $g \mathbin{\times\!\!-\!\!-} b$ in G^ℓ_{i-1}.

Hence, one has both $g \mathbin{\times\!\!-\!\!\times} b$ and $g \mathbin{\times\!\!-\!\!\times} a$ in G^ℓ, and the situation is as follows:



(6)

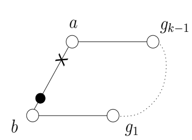
However, the configuration (6) has a subconfiguration $a \mathbin{-\!\!-\!\!\times} g \mathbin{-\!\!-\!\!\times} b \mathbin{\bullet\!\!-\!\!-} a$ which is a forbidden configuration of type (c) in G^ℓ. This contradicts the assumptions.

This completes the proof. ■

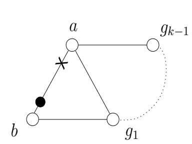
Fact 9 *The graph G' has no semi-directed cycle.*

Proof We show that if G' has a semi-directed cycle of the length $k+1$, where $k \geq 3$ then it has a semi-directed cycle of the length l, $3 \leq l \leq k$. This, together with Fact 8, implies what is desired.

Assume that $\rho : a, b, g_1, \ldots, g_{k-1}, a$, $k \geq 3$ is a semi-directed cycle in G' with $a \longrightarrow b$ in G'. By Fact 6, $a \,\text{---}\, b$ in G and Algorithm 2 implies that ρ corresponds to the following configuration



in G^ℓ, where the dotted connection stands for an undirected path and some edges are possibly omitted. It follows from Algorithm 1 that a and g_1 are adjacent in G for otherwise $g_1 \,\text{---}\, b \mathbin{\bullet\!\!-\!\!-} a$ is a forbidden configuration of the type (b) in G^ℓ. As G is a CG, $a \,\text{---}\, g_1$ in G and the situation is as follows:



If either $a \longleftarrow g_1$ or $a \;\text{——}\; g_1$ in G' then a,b,g_1,a is a semi-directed cycle of the length 3 in G'. On the other hand, if $a \longrightarrow g_1$ in G' then $a,g_1,g_2,\ldots,g_{k-1},a$ is a semi-directed cycle of the length k in G'. ■

Thus, we have verified that G' is a CG. The last step is to show that it is in the AMP equivalence class containing G.

Fact 10 *G' is triplex equivalent to G.*

Proof We already know that G and G' are CGs with the same underlying graph. Moreover, it follows from the construction of G' that $G \geq G'$.

In particular, every immorality in G remains in G'. Thus, to verify that triplexes in G are also in G' it is enough to show that every flag $a \longrightarrow b \;\text{——}\; d$ in G remains a triplex in G'. As $a \longrightarrow b$ in G', the only option of canceling the triplex $\langle\{a,d\},b\rangle$ is if $b \longrightarrow d$ in G'. Then Algorithm 2 implies $a \longrightarrow b \;\ast\!\!-\!\!\bullet\; d$ in G^ℓ, which is, however, a forbidden configuration of type (a) in G^ℓ (cf. Algorithm 1).

Now, we show that triplexes in G' are also in G. Realize that $G \geq G'$ implies that an arrow in G' cannot be an arrow with the opposite direction in G. Thus, if $a \longrightarrow b \;\text{——}\; d$ is a flag in G' then, by (2), either $a \longrightarrow b \;\text{——}\; d$ or $a \;\text{——}\; b \;\text{——}\; d$ in G. By Algorithm 2, the latter case means $a \;\ast\!\!-\!\!\bullet\; b \;\text{——}\; d$ in G^ℓ, which is a forbidden configuration of type (b). Analogously, if $a \longrightarrow b \longleftarrow d$ is an immorality in G' that does not correspond to a triplex in G then $a \;\text{——}\; b \;\text{——}\; d$ in G. Hence, $a \;\ast\!\!-\!\!\bullet\; b \;\bullet\!\!-\!\!\ast\; d$ in G^ℓ, which is also a forbidden configuration of type (b). ■

Proof of Proposition 20

Recall that G is a CG such that there is no line of the form $\ast\!\!-\!\!\bullet$ in its labeled version G^ℓ. Let F be a CG which is triplex equivalent to G. We are to show that $a \;\text{——}\; b$ in F whenever $a \;\ast\!\!-\!\!\ast\; b$ in G^ℓ. Suppose for contradiction that there exists (at least one) line of the form $e \;\ast\!\!-\!\!\ast\; f$ in G^ℓ such that $e \longrightarrow f$ in F. Thus, the set $\mathcal{A}_F$ of arrows $e \longrightarrow f$ in F of the form $e \;\ast\!\!-\!\!\ast\; f$ in G^ℓ is assumed to be non-empty.

Let us fix a chain of components $C_1,\ldots,C_m$, $m \geq 1$ in F. Let k be the highest $1 \leq k \leq m$ such that there exists an arrow $e \longrightarrow f$ from $\mathcal{A}_F$ with $f \in C_k$. Denote by $\mathcal{A}'_F$ the subset of $\mathcal{A}_F$ consisting of arrows $e \longrightarrow f$ with $f \in C_k$. Clearly, $\mathcal{A}'_F \neq \emptyset$.

The next step is to fix a sequence $G^\ell_0,\ldots G^\ell_n$, $n \geq 1$ of labeled CGs generated by Algorithm 1. Every $e \longrightarrow f$ from $\mathcal{A}'_F$ is assigned unique $1 \leq i \leq n$ such that $e \;-\!\!\!-\!\!\ast\; f$ in G^ℓ_i and $e \;-\!\!\!-\!\!\bullet\; f$ in G^ℓ_j for $j < i$. Let $a \longrightarrow b$ denote the arrow from $\mathcal{A}'_F$ which has assigned the least such i. Observe that this choice of $a \longrightarrow b$ ensures that the following two conditions are valid.

(I) If $b \longrightarrow d$ in F for some node d then $b \longrightarrow d$ does not belong to $\mathcal{A}_F$.

This is because $b \in C_k$. The fact that $C_1,\ldots,C_m$ is a chain for F implies $d \in C_l$ with $l > k$. However, k was chosen so that no arrow $e \longrightarrow f$ from $\mathcal{A}_F$ with $f \in C_l$ for $l > k$ exists.

(II) Whenever $e \;-\!\!\!-\!\!\ast\; f$ in G^ℓ_{i-1} then $\mathcal{A}'_F$ does not contain $e \longrightarrow f$.

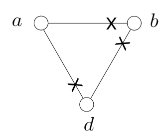
This follows from the choice of i: a necessary condition for $e \longrightarrow f$ to belong to $\mathcal{A}_F'$ is $e \mathrel{-\!\!\!-\!\!\ast} f$ in G_j^ℓ only for $j \geq i$, that is, $e \mathrel{-\!\!\!-\!\!\bullet} f$ in G_{i-1}^ℓ.

Now, we are going to derive a contradictory conclusion that $a \mathrel{\ast\!\!-\!\!\!-\!\!\ast} b$ cannot be blocked at b by any of the blocking rules from Algorithm 1.

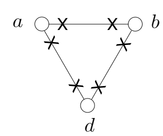
1. If $a \mathrel{\ast\!\!-\!\!\!-\!\!\ast} b$ is blocked at b on basis of the blocking rule (a) then there exists a vertex d such that $d \longrightarrow a \;\text{——}\; b$ in G and b is not adjacent to d. Hence, $d \longrightarrow a \;\text{——}\; b$ is a flag in G. As G and F are triplex equivalent, F has a triplex $\langle \{b,d\}, a \rangle$, which, however, contradicts the assumption $a \longrightarrow b$ in F.

2. If $a \mathrel{\ast\!\!-\!\!\!-\!\!\ast} b$ is blocked at b on basis of the blocking rule (b) then there exists a vertex d such that $a \;\text{——}\; b \;\text{——}\; d$ in G and a is not adjacent to d. This implies $b \longrightarrow d$ in F for otherwise the fact $a \longrightarrow b$ in F implies that $\langle \{a,d\}, b \rangle$ is a triplex in F which is not in G. Moreover, $a \;\text{——}\; b \;\text{——}\; d$ in G implies, by the blocking rule (b) from Algorithm 1, that $a \mathrel{-\!\!\!-\!\!\ast} b \mathrel{\ast\!\!-\!\!\!-} d$ in G^ℓ. Because G^ℓ has no lines of the form $\mathrel{\ast\!\!-\!\!\!-\!\!\bullet}$ this means $a \mathrel{\ast\!\!-\!\!\!-\!\!\ast} b \mathrel{\ast\!\!-\!\!\!-\!\!\ast} d$ in G^ℓ. Thus, $b \longrightarrow d$ belongs to $\mathcal{A}_F$, contradicting the condition (I) above.

3. If $a \mathrel{\ast\!\!-\!\!\!-\!\!\ast} b$ is blocked at b on basis of the blocking rule (c) then there exists a vertex d such that $a \mathrel{-\!\!\!-\!\!\ast} d \mathrel{-\!\!\!-\!\!\ast} b \mathrel{\bullet\!\!-\!\!\!-} a$ in $\widehat{G}_{i-1}^\ell$. Thus, we have

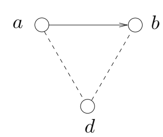


in G^ℓ. Since there is no line of the type $\mathrel{\ast\!\!-\!\!\!-\!\!\bullet}$ in G^ℓ, we have



(7)

in G^ℓ, whereas the corresponding subgraph in F is



(8)

where the dashed connection means that the nodes are adjacent. However, the configurations (7) and (8) cannot coexist because any possible type of the edge between d and b in F leads to a contradiction.

- If $d \longrightarrow b$ in F then (7) and the fact $b \in C_k$ imply $d \longrightarrow b$ is in $\mathcal{A}_F'$. As $d \mathrel{-\!\!\!-\!\!\ast} b$ in G_{i-1}^ℓ this contradicts the condition (II) above.
- If $d \;\text{——}\; b$ in F then $d \in C_k$ and $a \longrightarrow d$ in F for otherwise F has a semi-directed cycle. Hence, by (7) $a \longrightarrow d$ belongs to $\mathcal{A}_F'$. As $a \mathrel{-\!\!\!-\!\!\ast} d$ in $\widehat{G}_{i-1}^\ell$ it also contradicts the condition (II) above.
- If $b \longrightarrow d$ in F then (7) gives $b \longrightarrow d$ in $\mathcal{A}_F$ contradicting the condition (I) above.

This concludes the proof.

References

S. A. Andersson, D. Madigan and M. D. Perlman. An alternative Markov property for chain graphs. In *Uncertainty in Artificial Intelligence: Proceedings of the Twelfth Conference* (F. Jensen and E. Horvitz eds.), pages 40–48, Morgan Kaufmann, San Francisco, 1996.

S. A. Andersson, D. Madigan and M. D. Perlman. A characterization of Markov equivalence classes for acyclic digraphs. *The Annals of Statistics*, 25:505–541, 1997.

S. A. Andersson, D. Madigan and M. D. Perlman. Alternative Markov properties for chain graphs. *Scandinavian Journal of Statistics*, 28:33–85, 2001.

S. A. Andersson and M. D. Perlman. Characterizing Markov equivalence classes for AMP chain graph models. *The Annals of Statistics*, 34, 2006, forthcoming.

D. M. Chickering. Learning equivalence classes of Bayesian-network structure. *Journal of Machine Learning Research*, 2:445–498, 2002.

R. G. Cowell, A. P. Dawid, S. L. Lauritzen and D. J. Spiegelhalter. *Probabilistic Networks and Expert Systems*. Springer–Verlag, New York, 1999.

A. P. Dawid. Influence diagrams for causal modelling and inference. *International Statistical Review*, 70:161–189, 2002.

M. Frydenberg. The chain graph Markov property. *Scandinavian Journal of Statistics*, 17:333–353, 1990.

S. L. Lauritzen and N. Wermuth. Mixed interaction models. Research Report R-84-8, Inst. Elec. Sys., University of Aalborg, Denmark, 1984.

S. L. Lauritzen and N. Wermuth. Graphical models for association between variables, some of which are qualitative and some quantitative. *The Annals of Statistics*, 17:31–57, 1989.

S. L. Lauritzen. Causal inference from graphical models. In *Complex Stochastic Systems*, (O. E. Barndorff-Nielsen, D. R. Cox and C. Klüppelberg eds.), pages 63–107, Chapman and Hall/CRC, 2001.

A. Roverato. A unified approach to the characterisation of equivalence classes of DAGs, chain graphs with no flags and chain graphs. *Scandinavian Journal of Statistics*, 32:295–312, 2005.

A. Roverato and L. La Rocca. On block ordering of variables in graphical modeling. *Scandinavian Journal of Statistics*, 33:65–81, 2006.

M. Studený. A recovery algorithm for chain graphs. *International Journal of Approximate Reasoning*, 17:265–293, 1997.

M. Studený, Characterization of essential graphs by means of an operation of legal component merging. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 12:43–62, 2004.

M. Studený, A. Roverato and S. Štěpánová. Two operations of merging components in a chain graph. Submitted. Available electronically at `http://staff.utia.cas.cz/studeny/aa23.html`, 2006.

M. Volf and M. Studený. A graphical characterization of the largest chain graphs. *International Journal of Approximate Reasoning*, 20:209–236, 1999.