Towards a Statistical Foundation in Combining Structures of Decomposable Graphical Models[1]

Sung-Ho Kim[2]

Korea Advanced Institute of Science and Technology, Daejeon, South Korea

Summary

Graphical models offer simple and intuitive interpretations in terms of conditional independence relationships, and these are especially valuable when large numbers of variables are involved. In some settings restrictions upon experiments, number of variables, and other forms of data collection may result in our being able to estimate only parts of a large graphical model. Consider a collection $\mathcal{C}$ of submodels of a decomposable graph $\mathcal{G}$. In this article, we address the problem of combining component graphical models, and a theory is derived to the effect that one can combine the collection $\mathcal{C}$ of decomposable graphs, $\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_m$, into a larger decomposable graph, $\mathcal{H}$, of the variables that are involved in $\mathcal{G}$ so that the conditional independence relationships in $\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_m$ may be preserved in $\mathcal{H}$. It is also shown that the $\mathcal{H}$ which contains the actual graph $\mathcal{G}$ as a subgraph is determined uniquely.

Keywords: Conditional independence; Graph grafting; Graph separation; Independence graph; Maximal combined structure; Minimal connector

1 Introduction

We use the term *graphical model* to include the class of statistical models whose structures are representable via graph. The inter-relationship among the variables contained in a graphical model is interpreted in terms of conditional independence. And thus the corresponding graph is often called an independence graph (Whittaker, 1990). The independence graph is classified into two types, directed acyclic independence graph and undirected independence graph, the former being used when the relationship among the variables is in general influential or causal and the latter when the relationship is symmetrically associative. When the relationship is influential among some variables of a model and is associative among the other variables, the corresponding graph contains both directed and undirected

[1]This work was supported by Korea Research Foundation Grant(KRF-2000-DP0063)

[2]Division of Applied Mathematics, Korea Advanced Institute of Science and Technology, Daejeon, 305-701, South Korea

E-mail: sungkim@cais.kaist.ac.kr

edges. Equivalence of Markov properties among these three types of graph is discussed in Andersson et al. (1997a, 1997b) based on the results in Frydenberg (1990). Many authors (Wermuth, 1980; Wermuth and Lauritzen, 1983, Asmussen and Edwards, 1983; Kiiveri et al., 1984; Andersson et al., 1997) have shown with varying degrees of generality that the intersection of the classes of undirected independence graphs and the classes of directed acyclic independence graphs is the class of decomposable undirected independence graphs (Lauritzen, Speed, and Vijayan 1984).

Some attractive features of the decomposable graph are that it is triangulated (Darroch, Lauritzen, Speed (1980); Leimer (1989)), that decomposability is preserved in graph-collapsing (see Theorem 4.1) and that every directed acyclic graph is convertible into a decomposable graph (Lauritzen and Spiegelhalter, 1988). In this regard among others, we will confine ourselves on decomposable graphs in this article.

Graphical models are used in AI in the name of Bayesian network (Pearl, 1988) among others when the relationship among the variables involved can be interpreted as causal or influential and represented in terms of conditional independence. The relative efficiency of computational techniques for performing inference over the network makes the graphical model an extremely powerful tool for dealing with uncertainty in AI. Generating a Bayesian network from a knowledge base or a database has been an important issue in the AI research (Cooper and Herskovits, 1991; Poole, 1993; Goldman and Charniak, 1993; Bacchus, 1993; Haddawy, 1994; Chickering and Heckerman, 1999). But, researches on combining partial or conditional network models of some random variables conditional on some other variables to obtain a larger network model are rarely found in AI or Statistics. Mahoney and Laskey (1999) considered a problem of representing Bayesian networks using conditional probability models and conditioning variables but they did not address an issue of combining conditional models but an issue of combining conditioning variables for a more explicit representation of the network model. As for combining conditional model structures, Fienberg and Kim (1999) propose a theory to the effect that the graphical models can be combined in a consistent manner provided that the models are convertible into log-linear models.

In some settings restrictions upon experiments, number of variables, and other forms of data collection may result in our being able to estimate only parts of a large graphical model. In such situations, one may have to deal with building graphical models for sets of variables of moderate sizes and then combine the multiple graphical models in an effort to obtain a graphical model for the set of the variables that are involved in at least one of the "small" multiple graphical models.

Fienberg and Kim (1999) consider a problem of combining conditional graphical log-linear structures and derive a combining rule of them based on the relation between the log-linear model and its conditional version. In the graph combination, the conditioning variable is added to the set, say A, of the variables that are involved in the conditional structures and so each graph-combination ends up with a new graph that involves the conditioning variable in addition to the set A.

While Fienberg and Kim (1999) consider graphs of conditional graphical log-linear models, we will consider in this article the problem of combining graphs of marginal (as against conditional) graphical models of various types of random variables under the condition that the graph of the model of the variables that are involved in at least one of the marginal models is decomposable.

For example, suppose that we are given a pair of simple graphical models where one model is of random variables X_1, X_2, X_3 with their inter-relationship that X_1 is independent of X_3 conditional on X_2 and the other is of X_1, X_2, X_4 with their inter-relationship that X_1 is independent of X_4 conditional on X_2. From this simple pair, we can imagine a model structure for the four variables $X_1, \cdots, X_4$. The two inter-relationships are pictured in Figure 1. We will use the notation $[\cdot]\cdots[\cdot]$ as used in Fienberg (1980) to represent a model structure. X_1 and X_2 are shared in both models, and assuming that none of the variables are marginally independent of the others, the following model structures are possible corresponding to the pair of marginals:

$$[12][24][25],\ [12][24][45],\ [12][25][45],\ [12][245]. \tag{1}$$

Note that we can obtain the above pair from each of these models and that among these four models, the first three are submodels of the last one.

We will consider another pair of simple marginals, [12][23] and [24][25], where only one variable is shared. In this case, we have a longer list of possible joint model structures as follows:

$$\begin{array}{l}[12][24][23][25], [124][23][25], [124][23][35], [124][25][35], [124][235], \\ [125][23][34], [125][24][34], [125][234].\end{array} \tag{2}$$

Model structures [124][235] and [125][234] are maximal in the sense of set inclusion among these eight models.

Figure 1: Two graphs of inter-relationship

It is important to note that some variable(s) are independent of the others conditional on X_2 in each of the two pairs and in all the models in (1) and (2). That conditional independence takes place conditional on the same variable in the marginal structures and also in the joint structures underlies the main results of the article.

We address the issue of combining graphical model structures and so we can not help using independence graphs and related theories to derive desired results with more clarity and refinement. Throughout the article, graph and model terminologies are used interchangeably when confusion is not likely.

The article is organized in 8 sections. Section 2 introduces notation and graphical terminologies to use, and section 3 introduces basic notions of graph combination and presents further discussions on decomposable graphs. Section 4 then shows unique existence of a combined graph $\mathcal{H}$ which contains $\mathcal{G}$ as an edge-subgraph (defined in section 2) and which is obtained based on a collection of node-subgraphs (defined in section 2) of $\mathcal{G}$. In section 5, the graph-combining is illustrated under the condition that graphs consist of the cliques each of which is made of at most two nodes, and section 6 extends the result of section 5 to the graphs where cliques consist of more than two nodes. While we considered combining a pair of graphs in sections 5 and 6, section 7 deals with the problem of combining three or more graphs. Finally, section 8 summarizes the article with some concluding remarks.

2 Notation and graph terminologies

In the article, we will consider undirected graphs only. We denote a graph by $\mathcal{G} = (V, E)$, where V is the set of the indexes of the variables involved in $\mathcal{G}$ and E is a collection of ordered pairs, each pair representing that the nodes of the pair are connected by an edge. Since $\mathcal{G}$ is undirected, $(u, v) \in E$ is the same as (v, u) and vice versa. We say that a set of nodes of $\mathcal{G}$ forms a complete subgraph of $\mathcal{G}$ if every pair of nodes in the set are connected by an edge. Graph $\mathcal{G}$ will also be represented by a sequence of the symbol $[\cdots]$, where each pair of brackets represent a clique (i.e., a maximal complete subgraph) in $\mathcal{G}$, the clique being composed of the nodes whose indexes appearing in the brackets. For instance,

$$\mathcal{G} = [12][23][345]$$

means that graph $\mathcal{G}$ is of 5 nodes and consists of the three cliques, one consisting of nodes 1 and 2, another of nodes 2 and 3, and the rest of nodes 3, 4, and 5. Recall that the right-hand side of the above expression can also be interpreted as a model structure. In this context, the terms *graph* and *model structure* will be used in the same sense in this article.

A *path* between a pair of nodes is a sequence of edges leading from one end node of the sequence to the other. If a and b are connected by an edge, we say that a is a neighbor node of b or vice versa. A boundary of a set A in a graph $\mathcal{G}$ is the set of nodes of $\mathcal{G}$ each of which is a neighbor node of a node in A but not included in A, and we denote it by $bd(A)$. We say that a graph is *connected* if there is at least one path between every pair of nodes of the graph; otherwise we call it *disconnected*. All the graphs considered in this article are connected.

For a subset A of V, we define $\mathcal{G}_A$ to be the node-subgraph of $\mathcal{G}$ confined to A. That is, the relationship among the variables indexed in A with respect to $\mathcal{G}$ is preserved in $\mathcal{G}_A$. If $\mathcal{G} = (V, E)$, $\mathcal{G}' = (V, E')$, and $E' \subseteq E$, then we say that $\mathcal{G}'$ is an edge-subgraph of $\mathcal{G}$ and write as $\mathcal{G}' \subseteq^e \mathcal{G}$. A subgraph of $\mathcal{G}$ means either a node-subgraph or an edge-subgraph of $\mathcal{G}$. If $\mathcal{G}'$ is a subgraph of $\mathcal{G}$, we call $\mathcal{G}$ a supergraph of $\mathcal{G}'$. We will adopt a convention that $\mathcal{G}_i = (V_i, E_i)$ stands for a subgraph of $\mathcal{G}$. Suppose that there are m node-subgraphs of $\mathcal{G}$, $\mathcal{G}_1, \cdots, \mathcal{G}_m$, then, for $1 \le j \le m$, we will denote the graph $\mathcal{G}_{\cup_{i=1}^{j} V_i}$ by $\mathcal{G}_{(j)}$. We will also denote by $V(\mathcal{G})$ and $E(\mathcal{G})$ respectively the set of the nodes and the set of the edges of graph $\mathcal{G}$.

A subset $C \subseteq V$ is said to separate A from B if it intersects all paths between every $a \in A$ and every $b \in B$, and if C is a subset of every set D that separates A from B, then we say that C is a minimal separator of A from B. In a similar context, we define the notion of *connector*. We will say that $C \subseteq V$ *connects* $A_i,\ i = 1, 2, \cdots, l$, in $\mathcal{G}$ if

(i) $\cup_{i=1}^{l} A_i = V$,

(ii) $A_i \setminus C \neq \emptyset$ for $i = 1, 2, \cdots, l$,

(iii) C separates A_i from $A_j,\ j \neq i$, and C does not separate any subset of A_i from the rest of A_i for $i = 1, \cdots, l$, and

(iv) for every node in C, it has a neighbor node in C.

We will call such a set C a *connector* of $A_i,\ i = 1, 2, \cdots, l$ in $\mathcal{G}$. If $\cup_{i=1}^{l} A_i = \cup_{j=1}^{k} B_j = V$, $\{A_1, \cdots, A_l\} \neq \{B_1, \cdots, B_k\}$, and C connects $A_1, \cdots, A_l$ and D connects $B_1, \cdots, B_k$, then we say that C and D are different connectors. Two sets A and B will be said to be discrepant if neither is a subset of the other. We will call the set C a *minimal connector* of $A_i,\ i = 1, 2, \cdots, l$, in $\mathcal{G}$ if it is not a union of discrepant minimal separators and any subset of it fails to separate any of the A_i's from the others in $\mathcal{G}$; and denote by $\sigma(\mathcal{G})$ the collection of the minimal connectors in $\mathcal{G}$. Note that a (minimal) connector is a special form of a (minimal) separator. Thus, every property that is satisfied by a separator also holds for a connector. However, a (minimal) separator is not necessarily a (minimal) connector because of condition (iv).

Suppose that the graph $\mathcal{G}$ of a decomposable graphical model for a random vector $\mathbf{X}$ consists of k cliques, $C_1, \cdots, C_k$, and let $C_{(j)} = \cup_{i=1}^{j} C_i$ and $s_j = C_j \cap C_{(j-1)}$. If the cliques are labeled such that the probability model for $\mathbf{X}$ may be expressed as (see Darroch et al. (1980))

$$P(\mathbf{X} = \mathbf{x}) = P(\mathbf{x}_{C_1}) \prod_{i=2}^{k} \frac{P(\mathbf{x}_{C_i})}{P(\mathbf{x}_{s_i})},$$

we can call the sets $s_2, \cdots, s_k$, the minimal connectors of $\mathcal{G}$ because of the decomposability of the graph.

If a node is not a minimal connector nor contained in a minimal connector and if it belongs to a clique which contains only one minimal connector, then we will call the node an *exterior-node*. If a node is not an exterior-node, we will call it an *interior* node.

3 Conditional independence and combined model structures

Suppose we have a set of random variables and a list $\mathcal{L}$ of inter-relationships among them that are expressed in terms of conditional independence. Assuming that the probability model for the set of random variables is graphical, we can have an independence graph, $\mathcal{G}$ say, corresponding to the model (Whittaker, 1990).

Properties of conditional independence are described in detail in Dawid (1979, 1980) and it is well known that the separation in $\mathcal{G}$ is equivalent to the conditional independence as appearing in $\mathcal{L}$ (the separation theorem in Whittaker (1990)). Now suppose that we have two sets of random variables whose probability models each are graphical. If the probability model for the union of the two sets is also graphical, what is the possible model structure for this model? We can rewrite this statement as follows:

> Suppose that we have two lists of inter-relationships corresponding two sets of random variables and that each of the lists can be depicted by an independence graph. If the inter-relationships among the variables in the union of the two sets can also be depicted by an independence graph, what is the possible list of the inter-relationships for the union of sets?

Let the two lists be $\mathcal{L}_1$ and $\mathcal{L}_2$ and their union $\mathcal{L}_u$. We aim to derive an approach that is useful for obtaining an independence graph that reflects the conditional independence relationships displayed in $\mathcal{L}_u$. $\mathcal{L}_1$ and $\mathcal{L}_2$ are respectively about two different sets V_1 and V_2 of random variables. So obtaining an independence graph corresponding to $\mathcal{L}_u$ is the same as obtaining a new list, $\mathcal{L}_{new}$ say, of conditional independence relationships for $V_1 \cup V_2$. In this sense, the latter operation is an application of the notion of conditional independence.

As far as model structures are concerned, working with independence graphs is much easier than working with lists of conditional independence relationships. In this respect, we will look into the issue of obtaining $\mathcal{L}_{new}$ from $\mathcal{L}_1$ and $\mathcal{L}_2$ from the perspective of model structures or independence graphs.

We consider a couple of examples that may help us get an insight into the relation between a pair of graphical models and their combined version. To avoid confusion, we will call the combined version a *joint model structure*. A formal definition of the combination of model structures is given below.

Definition 3.1 *Suppose there are m marginal graphs, $\mathcal{G}_1, \cdots, \mathcal{G}_m$. Then we say that graph $\mathcal{H}$ of a set of variables V is a* combined model structure *corresponding to $\mathcal{G}_1, \cdots, \mathcal{G}_m$, if*

$$\cup_{i=1}^{m} V_i = V \tag{3}$$

and

$$\mathcal{H}_{V_i} = \mathcal{G}_i, \text{ for } i = 1, \cdots, m. \tag{4}$$

We will call $\mathcal{H}$ a maximal combined structure *(MCS) corresponding to $\mathcal{G}_1, \cdots, \mathcal{G}_m$ if any additional edge to $\mathcal{H}$ destroys equation (4) for at least one i of $1, \cdots, m$, and denote by $\mathcal{M}(\mathcal{G}_1, \cdots, \mathcal{G}_m)$ the collection of all such MCSs.*

According to the definition, a combined model structure (or combined structure, for short) is a node-supergraph of each $\mathcal{G}_i$, $i = 1, \cdots, m$. There may be many combined structures that are obtained from a collection of marginal structures. For instance, there are four combined structures of [12] and [23] such as

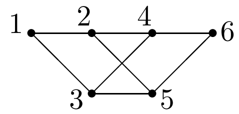
$$[123], [12][23], [12][13], \text{ and } [13][23]. \tag{5}$$

Under the decomposability assumption, one of the four graphs in (5) is the actual structure for the variables indexed in $\{1, 2, 3\}$. As more nodes get involved in marginal structures, we may obtain more combined structures from them. So, if we wish to find the true structure for V based on a collection of $\{\mathcal{G}_i\}_{i=1}^{m}$ with $V = \cup_{i=1}^{m} V_i$, we need to look into the properties that are shared in both combined structure and marginal structure.

When $\mathcal{G}$ is decomposable, its minimal connectors are complete; otherwise, its minimal connectors are not necessarily complete in $\mathcal{G}$. For instance, in the non-decomposable graph $\mathcal{G}$ as given below, each of

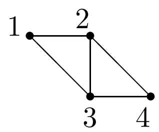
$$\{2,3,4\},\{2,4,5\},\{2,3,5\},\{3,4,5\}$$

is a minimal connector but not complete in $\mathcal{G}$.

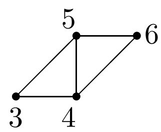


When $\mathcal{G}$ is not decomposable, $s \in \sigma(\mathcal{G})$ may not form a complete subgraph in $\mathcal{G}$ while it may be complete in a node-subgraph of $\mathcal{G}$. This sort of discrepancy of a minimal connector in graphical appearance between a graph and its node-subgraph may be a source of difficulty in searching for a graph from a collection of node-subgraphs of the graph. If we do not confine ourselves on decomposable graphs, the following node-subgraphs, $\mathcal{G}_a$ and $\mathcal{G}_b$, are possible from each of the graphs in Figure 2, for

$$a = \{1,2,3,4\} \text{ and } b = \{3,4,5,6\} : \tag{6}$$



$\mathcal{G}_a$



$\mathcal{G}_b$

In Figure 2, the graphs in panels 1, 2, 3 are decomposable, and the rest are not decomposable. Note that the minimal connectors of $\mathcal{G}_a$ and $\mathcal{G}_b$, $\{2,3\}$ and $\{4,5\}$, respectively, are complete in the graphs in panels 1, 2, and 3, while $\{2,3\}$ and $\{4,5\}$ are not necessarily complete in the non-decomposable graphs in the figure.

Theorem 3.1 *If $\mathcal{G}$ is decomposable, then each $s \in \sigma(\mathcal{G})$ is complete.*

Proof: See the proof of Proposition 2.5 of Lauritzen (1996). □

We will see below (see Theorem 4.1) that every node-subgraph of a decomposable graph is also decomposable. This implies, by Theorem 3.1, that, as for a decomposable graph, the minimal connectors are always given in the form of a complete subgraph in the graph and its node-subgraphs. This property may help us in searching for a graph based on a

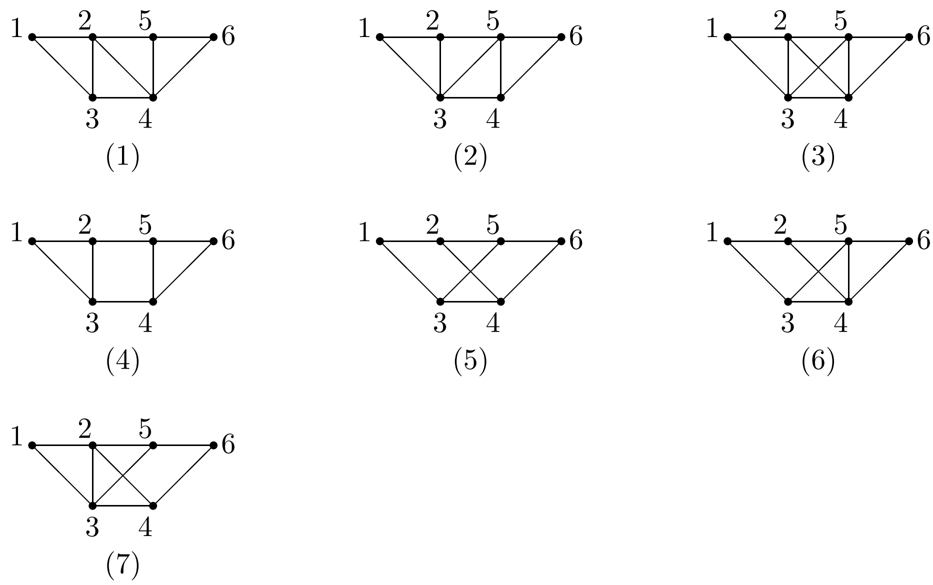
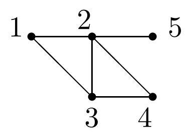


Figure 2: Graphs each of which yields the node-subgraphs, $\mathcal{G}_a$ and $\mathcal{G}_b$ for $a,\ b$ as in (6).

collection of its node-subgraphs. This is a main reason that we confine ourselves, in this article, on decomposable graphs.

There may be two minimal connectors one of which is contained in the other. For instance, in the graph below, $\{2\}$ and $\{2,3\}$ are both minimal connectors. Note that they connect different collections of subsets. $\{2\}$ connects the sets $\{1,2,3,4\}$ and $\{2,5\}$ while $\{2,3\}$ connects the three parts, $\{1,2,3\}$, $\{2,3,4\}$, and $\{2,5\}$.



We will borrow the symbol "$\perp$" from Dawid (1979) to represent conditional independence. Let $\mathbf{X}_a, \mathbf{X}_b, \mathbf{X}_c$ be three random vectors. Then $a \perp b|c$ represents that $\mathbf{X}_a$ and $\mathbf{X}_b$ are conditionally independent given $\mathbf{X}_c$. In the lemma below, $C_{\mathcal{G}}(v)$ is a clique which includes a node v in a given graph $\mathcal{G}$ and $\Xi_{\mathcal{G}}(v)$ is the collection of such $C_{\mathcal{G}}(v)$'s.

Lemma 3.1 *Let $\mathcal{G}'$ be a node-subgraph of $\mathcal{G}$ and suppose that, for three disjoint subsets a, b, c of $V(\mathcal{G}')$, $a \perp c|b$. Then*

(i) $a \perp c|b$ *in* $\mathcal{G}$.

(ii) $C_{\mathcal{G}}(a) \perp C_{\mathcal{G}}(c)|b$ *in* $\mathcal{G}$.

(iii) $\Xi_{\mathcal{G}}(a) \perp \Xi_{\mathcal{G}}(c)|b$ *in* $\mathcal{G}$.

Proof: Since

$$a \perp c|b \text{ in } \mathcal{G}', \tag{7}$$

there is no path in $\mathcal{G}'$ between a and c that bypasses b. If (i) does not hold, it is obvious that (7) does not hold either. Now suppose that result (ii) does not hold. Then there must be a path from a node in a to a node in c bypassing b. This implies negation of the condition $a \perp c|b$. Therefore, result (ii) must hold. As for (iii), suppose that (iii) does not hold. Then there must exist some cliques, say C_1 and C_2, in Ξ_a and Ξ_b, respectively, for which (ii) does not hold. Hence if (ii) holds, so must (iii). □.

This lemma states that a separator of a node-subgraph of $\mathcal{G}$ is also a separator of $\mathcal{G}$. According to the lemma, we can have

Theorem 3.2 *Suppose there are m marginal structures $\mathcal{G}_i$, $i = 1, 2, \cdots, m$. Then*

(i) for any combined structure $\mathcal{H}$ of the m marginal structures,

$$\cup_{i=1}^{m} \sigma(\mathcal{G}_i) \subseteq \sigma(\mathcal{H}).$$

(ii) for any MCS $\mathcal{H}$,

$$\cup_{i=1}^{m} \sigma(\mathcal{G}_i) = \sigma(\mathcal{H}).$$

Proof: We prove (i) first. Suppose that there exists a connector a in some of $\{\mathcal{G}_i;\ i = 1, \cdots, m\}$. Then by Lemma 3.1 a is also a connector in any combined structure of $\{\mathcal{G}_i\}_{i=1}^{m}$.

To show (ii) now, suppose there exists a set b in $\sigma(\mathcal{H}) \setminus \cup_{i=1}^{m} \sigma(\mathcal{G}_i)$. Since $b \in \sigma(\mathcal{H})$, it separates $\mathcal{H}$ into at least two graphs. Without loss of generality, we may assume that b separates $\mathcal{H}$ into two parts, say $\mathcal{H}_1$ and $\mathcal{H}_2$, and let C_1, C_2 be cliques containing b in $\mathcal{H}$. Note that $C_i \setminus \{a\} \neq \emptyset$ for $i = 1, 2$. This implies that $\mathcal{H}$ is not an MCS since we may add an edge connecting a node in $C_1 \setminus \{a\}$ to another in $C_2 \setminus \{a\}$ to obtain $\mathcal{H}'$ which still satisfies expressions (3) and (4). This contradicts that $\mathcal{H}$ is an MCS, which completes the proof of (ii). □

We know that, for two graphs $\mathcal{G}'$ and $\mathcal{G}$ with $\mathcal{G}' \subseteq^e \mathcal{G}$, every graph-separateness in $\mathcal{G}$ is preserved in $\mathcal{G}'$. As connoted from Theorem 3.2, we can thus see that the true joint structure of V is an edge-subgraph of a combined structure of a collection of marginal structures $\{\mathcal{G}_i\}_{i=1}^{m}$ with $\cup_{i=1}^{m} V_i = V$. But searching for the true joint structure from a set of marginal structures is like searching for a coin in a beach when V is "large".

Example 3.1 We consider the problem of how many joint structures we need to look at in search of the true joint structure and some related issue. Suppose we have two marginal structures, $\mathcal{G}_1 = [12][23]$ and $\mathcal{G}_2 = [24][25]$. Since $\sigma(\mathcal{G}_1) = \sigma(\mathcal{G}_2) = \{2\}$, $\{2\}$ is also a connector for $V = \{1, 2, 3, 4, 5\}$. We can then see that only the two joint structures

$$[125][234] \text{ and } [124][235] \tag{8}$$

are the MCSs corresponding to the marginal structures, $\mathcal{G}_1$ and $\mathcal{G}_2$.

There are 9 connected edge-subgraphs of each graph in (8), and since the model structure [12][23][24][25] is a common edge-subgraph of both of the graphs in (8), there are in total 19 different combined structures, including the two in (8), corresponding to $\mathcal{G}_1$ and $\mathcal{G}_2$.

Note that, in this example, there are 19 combined structures corresponding to the pair of marginal structures while there are only 2 MCSs and that each of the 19 structures is an edge-subgraph of one of two MCSs. The example calls our attention to the importance of the MCS. The actual combined structure from which a given collection of marginal structures are obtained is nothing but an edge-subgraph of an MCS. This is why it is desirable that we search for a collection of MCSs from a collection of marginal model structures.

For a given set of marginal graphs, we can easily obtain the set of minimal connectors.

Theorem 3.3 *Let $\mathcal{G}$ be a decomposable graph. Then $\sigma(\mathcal{G})$ is the collection of the intersections of the neighboring cliques of the graph.*

Proof: See Appendix B.

According to this theorem, we can now find $\sigma(\mathcal{G})$ for any decomposable graph $\mathcal{G}$ simply by taking all the intersections of the cliques of the graph.

4 Existence of maximal combined structures

In Example 3.1, the combined model structures are sought for starting from the MCSs corresponding to a given set of marginal model structures. An apparent feature of an MCS in contrast to a combined structure is stated in Theorem 3.2 (ii). The following theorem is stated as Corollary 2.8 in Lauritzen (1996).

Theorem 4.1 *Every node-subgraph of a decomposable graph is decomposable.*

It is obvious that decomposability of an edge-subgraph of a decomposable graph is not guaranteed. Suppose there are m node-subgraphs of $\mathcal{G}$. Let $S = \cup_{i=1}^{m} \cup_{s \in \sigma(\mathcal{G}_i)} s$. Then, we can see that each set s in $\sigma(\mathcal{G}_i)$ is a minimal connector of $\mathcal{G}_S$ unless it contains an exterior-node of $\mathcal{G}_S$. If s is neither a connector nor contains an exterior-node of $\mathcal{G}_S$, then neither can it be a connector of $\mathcal{G}_i$. Therefore, every set in $\cup_{i=1}^m \sigma(\mathcal{G}_i)$ is a minimal connector in $\mathcal{G}_S$ except the sets in $\cup_{i=1}^m \sigma(\mathcal{G}_i)$ each of which contains an exterior-node of $\mathcal{G}_S$.

Lemma 4.1 *Suppose that there are m node-subgraphs of $\mathcal{G}$, $\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_m$. Then*

$$\cup_{i=1}^m \sigma(\mathcal{G}_i) \subseteq \sigma(\mathcal{G}).$$

Proof: Let $\Psi = \cup_{i=1}^m \sigma(\mathcal{G}_i)$. Suppose that there exists a set s in Ψ. This implies that s is a minimal connector of a node-subgraph $\mathcal{G}_i$. Suppose, on the contrary to the result, that s is not contained in $\sigma(\mathcal{G})$. Then by definition, s can not be a connector in the node-subgraph, which is a contradiction. This completes the proof. □

Theorem 4.2 *Let $\mathcal{G}$ be decomposable and consider m node-subgraphs of $\mathcal{G}$, $\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_m$. Then every MCS $\mathcal{H}$ of $\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_m$ is decomposable.*

Proof: $\mathcal{G}_i$, $i = 1, 2, \cdots, m$, is decomposable by Theorem 4.1. So for each i, there is no chordless n-cycle, $n \geq 4$, by the definition of MCS, in the part of $\mathcal{H}$ corresponding to V_i. Suppose there is a chordless n-cycle, $n \geq 4$, in the part of $\mathcal{H}$ corresponding to $V_i \cup V_j$. For instance, let the cycle consists of 4 nodes v_1, v_2, v_3, v_4, where $\{v_1, v_2\} \subseteq V_i$ and $\{v_3, v_4\} \subseteq V_j$. Then we may create a clique of $\{v_1, v_2, v_3, v_4\}$ and replace the cycle with the clique to have an MCS of $\mathcal{G}_1, \cdots, \mathcal{G}_m$. This contradicts that $\mathcal{H}$ is an MCS. Therefore, $\mathcal{H}$ must be decomposable. □

For a set of node-subgraphs of a graph $\mathcal{G}$, there can be more than one MCS of the set of node-subgraphs. But there is only one such MCS that contains $\mathcal{G}$ as its edge-subgraph.

Theorem 4.3 *Suppose there are m node-subgraphs $\mathcal{G}_1, \cdots, \mathcal{G}_m$ of $\mathcal{G}$.*
(i) Then there exists a unique MCS $\mathcal{H}^*$ of the m node-subgraphs such that $\mathcal{G} \subseteq^e \mathcal{H}^*$.
(ii) Let $\mathcal{G}' \subseteq^e \mathcal{G}$. Then there exists a unique MCS $\mathcal{H}^*$ of the m node-subgraphs of $\mathcal{G}$ such that $\mathcal{G}' \subseteq^e \mathcal{H}^*$.

Proof: See Appendix B.

The unique existence of MCS for a given model structure throws a promising light on the searching road from a set of marginal model structures toward the MCS which contains the actual model structure as an edge-subgraph.

As illustrated in Example 3.1, the size of the set of all the possible joint structures corresponding to a set, say A, of marginal structures will be a lot larger than that of all the possible MCSs of A. Thus it will be cost efficient to search for the MCS that contains the actual joint structure corresponding to a given set of marginal model structures.

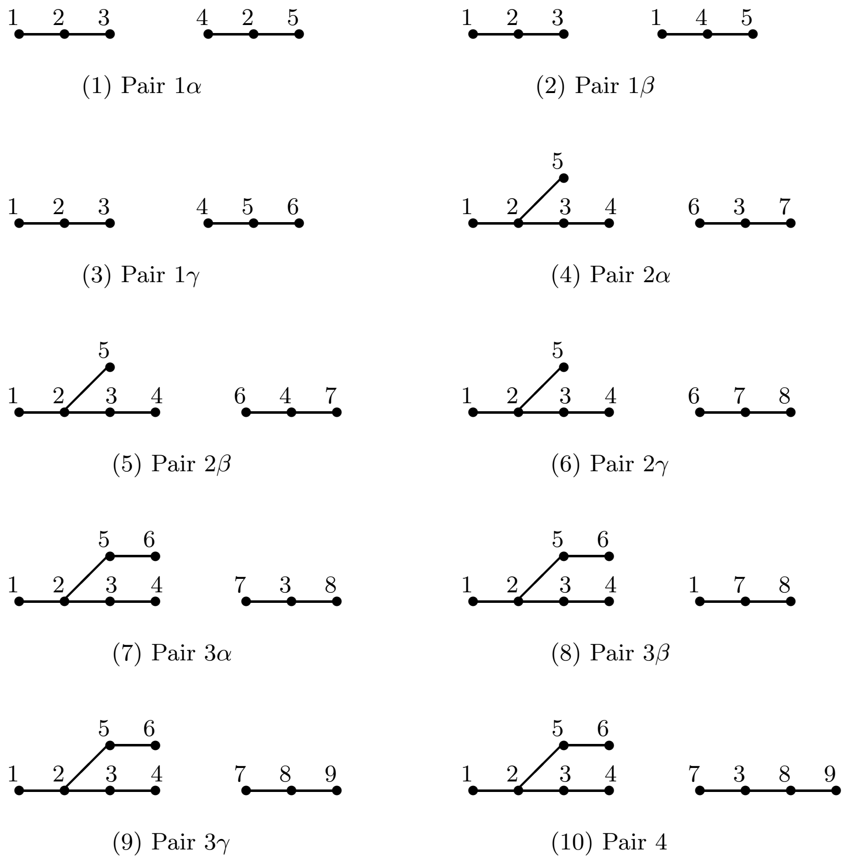


Figure 3: Pairs of marginal model structures

5 Grafting locations and illustration of combining node-subgraphs when minimal connectors are singletons of nodes

In this section, we will use several pairs of marginal structures of graphical models as in Figure 3 and propose an approach of searching for the MCS which contains an actual joint structure of which each of the given pair is a marginal structure, or equivalently a node-subgraph. When combining two node-subgraphs, we propose to do it as if we graft one subgraph onto another by grafting the minimal connectors of the former subgraph at some locations of the latter subgraph. Only the single-node connectors will be considered in this section so that we can pay more attention on diverse instances of graph-grafting. An extention to it will be discussed in next section along with a formal definition of the operation of grafting subgraphs together.

Figure 3 displays 10 pairs of model structures. Each structure of the first three pairs, Pairs 1α, 1β, and 1γ, has only one minimal connector. In Pair 1α, the marginal structures share the minimal connector, node 2; as for Pair 1β, they share node 1 that is not a connector; as for Pair 1γ, there is no node-sharing. One marginal structure of each of Pairs 2α, 2β, and 2γ contains 5 nodes with 2 minimal connectors, nodes 2 and 3, where node 2 separates the graph into three parts. The minimal connector, node 3, is shared in Pair 2α; node 4 is shared in Pair 2β, where the node is a minimal connector in only one graph of the pair; and no nodes are shared in Pair 2γ.

One more node is added in Pairs 3α, 3β, and 3γ, where one graph of each pair contains 3 separating nodes 2, 3, and 5, with node 2 separating the graph into three parts. Patterns of node-sharing for these pairs are similar to the preceding three pairs. As for Pair 4, the graphs contain 3 and 2 minimal connectors, respectively, while there is at least one graph that contains only one connector node in each of the other pairs of Figure 3. One may have noticed that the pair labels, $1\alpha, \cdots, 3\gamma, 4$, are a classifier of the pairs with respect to graph size and node-sharing. The numeric part of the label increases as the total number of the nodes contained in the graphs of the pair increases, and the letter part classifies the pattern of node-sharing, 'α' meaning that a connector is shared, 'β' meaning that a node is shared and the node is not a connector in both of the graphs, and 'γ' meaning that no nodes are shared. In this respect, Pair 4 should have been dubbed "Pair 4α". We will not consider other types of Pair 4 since they are nothing but a more complicated versions of the smaller pairs.

Theorem 4.3 suggests that we may work with MCSs in searching for the true joint structure, and Theorem 3.2 (ii) proposes that the collection of the minimal connectors of node-subgraphs of $\mathcal{G}$ plays a key role in determining MCSs of the node-subgraphs. As shown in Lemma 4.1, a minimal connector of a node-subgraph of $\mathcal{G}$ is also a minimal connector of $\mathcal{G}$. This implies that a minimal connector of $\mathcal{G}_1$ can not be separated by any $s \in \sigma(\mathcal{G}_2)$ and vice versa and so that s can not be grafted onto $\mathcal{G}_1$ at its minimal connectors but at its cliques. This will be proved in next section after looking over examples of the grafting as given below.

Consider the marginal model structures of Pair 1γ. For convenience' sake, we will call the left graph by $\mathcal{G}_1$ and the right one by $\mathcal{G}_2$; and the same applies to all the other pairs in Figure 3. There are four possible grafting locations of the minimal connector $\{5\}$ in $\mathcal{G}_1$.

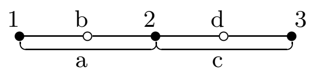


Figure 4: The grafting locations in $\mathcal{G}_1$ of Pair 1γ

Table 1: The 16 MCSs of Pair 1γ

Grafting location	The MCSs for each grafting location
a	[45][1256][23], [56][1245][23], [45][125][236], [56][125][234]
b	[145][256][23], [145][25][236], [156][245][23], [156][25][234]
c	[124][235][56], [12][2345][56], [126][235][45], [12][2356][45]
d	[124][25][356], [12][245][356], [126][25][345], [12][256][345]

They are denoted by a, b, c, d in Figure 4. b and d in this figure mean that the minimal connector $\{5\}$ is located at that points, while a and c mean that the minimal connector forms a clique with $\{1, 2\}$ and $\{2, 3\}$, respectively. The MCSs corresponding to each of the grafting locations are listed in Table 1. The graphs of the four models in row b of Table 1 are displayed in Figure 6 in the same order as they appear.

Next, we consider a pair of graphs where a connector partitions a graph into three parts as in $\mathcal{G}_1$'s of Pair 3's. We will take Pair 3β whose MCSs are not as many as for Pair 3γ for which there are 112 corresponding MCSs. The grafting locations of the minimal connector $\{7\}$ of $\mathcal{G}_2$ in $\mathcal{G}_1$ are indicated as a, $\cdots$, h in Figure 5. The MCSs corresponding to each grafting location are listed in Table 2.

It is worthwhile to note that there are two MCSs in Table 2 in each of which three minimal connectors form a clique. In general, if there are two collections, S and T, of minimal connectors where the minimal connectors in each collection are included in a clique of the corresponding node-subgraph, then there exists an MCS of the two node-subgraphs which contains a clique whose components include the minimal connectors in $S \cup T$. Also note that when more than two connectors of node-subgraphs are included in a clique of an MCS, the clique separates the MCS into more than two parts.

All the grafting locations of the other pairs in Figure 3 and the number of the MCSs for each of the grafting locations are displayed respectively in Figure 10 and Table 3 in Appendix A.

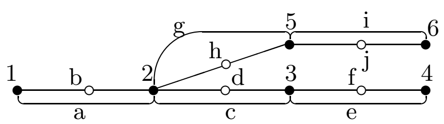


Figure 5: The grafting locations in $\mathcal{G}_1$ of Pair 3β

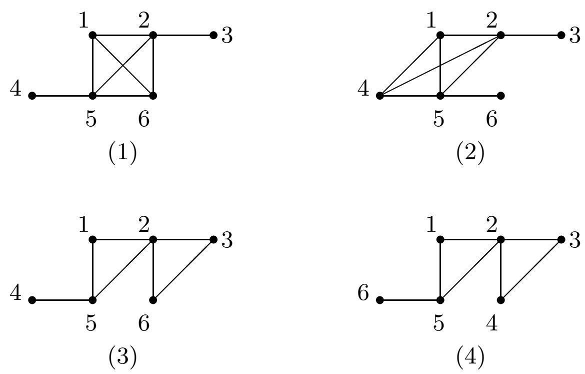


Figure 6: The graphs of the four models in row a of Table 1

Table 2: The 19 MCSs of Pair 3β

Grafting location	The MCSs for each grafting location
a	[78][127][56][25][23][34]
b	[17][278][23][34][25][56], [17][27][238][34][25][56], [17][27][23][348][25][56], [17][27][23][34][258][56], [17][27][23][34][25][568]
c	[12][237][78][34][25][56]
d	[12][27][378][34][25][56], [12][27][37][348][25][56]
e	[12][23][347][78][25][56]
f	[12][23][37][478][25][56]
g	[12][23][34][257][78][56]
h	[12][23][34][27][578][56], [12][23][34][27][57][568]
i	[12][23][34][25][567][78]
j	[12][23][34][25][57][678]

6 Combining node-subgraphs when minimal connectors are sets of nodes

In this section minimal connectors are of a general form, i.e., sets of nodes. So the grafting locations are not between nodes as in the preceeding section but between sets of nodes. The theorem below is intuitive but useful when dealing with node-subgraphs where minimal connectors are sets of nodes.

Corollary 6.1 *Let $\mathcal{G}_1$ be a node-subgraph of $\mathcal{G}$. Then for every $s \in \sigma(\mathcal{G}_1)$, there exist at least a pair of neighboring cliques in $\mathcal{G}$ that share s.*

Proof: This follows immediately from Lemma 3.1 (ii). □

At this point, it seems worthwhile to describe the notion of *graph-grafting* formally as in

Definition 6.1 *Consider two node-subgraphs of $\mathcal{G}$, $\mathcal{G}_1$ and $\mathcal{G}_2$. When a minimal connector s of $\mathcal{G}_2$ is attached to a clique, C, of $\mathcal{G}_1$ so that $s \cup C$ forms a clique, we say that s is* type-1 grafted *onto $\mathcal{G}_1$ at clique C. When s separates C into $l \geq 2$ parts $C_1, \cdots, C_l$ in $\mathcal{G}_1$ so that $s \cup C_1, \cdots, s \cup C_l$ each forms a clique, we say that s is* type-2 grafted *onto $\mathcal{G}_1$ at C.*

We have seen both types of grafting in the preceding section. The grafting locations b and d in Figure 4 and b, d, f, h, and i in Figure 5 are for type-2 grafting; and a and c and a, c, e, g, and i are the grafting locations, respectively, in Figures 4 and 5 for type-1 grafting. It is important to note that the grafting takes place in Figures 4 and 5 at cliques only either in type-1 or in type-2. Later in this section, we will see that this is always the case.

The grafting locations are counted, by Corollary 6.1, according to the same rule as in the preceding section. For example, consider the pair of subgraphs in Figure 7, where $\sigma(\mathcal{G}_1) = \{\{2,3\}, \{4,5\}\}$ and $\sigma(\mathcal{G}_2) = \{\{2,3\}, \{7,8\}\}$. Note in this pair that $\{2,3\}$ is common in both $\sigma(\mathcal{G}_1)$ and $\sigma(\mathcal{G}_2)$. By Lemma 4.1, $\{2,3\}$, $\{4,5\}$, $\{7,8\}$ are in $\sigma(\mathcal{G})$. And so the grafting locations of $\{7,8\}$ are

$$\text{between } \{2,3\} \text{ and } \{4,5\}, \text{ between } \{4,5\} \text{ and } \{6\}, \text{ and at the right-end clique of } \mathcal{G}_1. \quad (9)$$

The corresponding MCSs are given in Figure 8.

The first two MCSs in Figure 8 correspond to the first grafting location in (9) and the third and fourth MCSs correspond respectively to the second and third locations in (9).

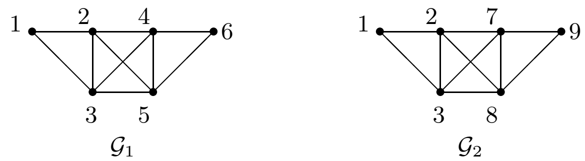


Figure 7: A graph-pair where the minimal connectors are sets of nodes

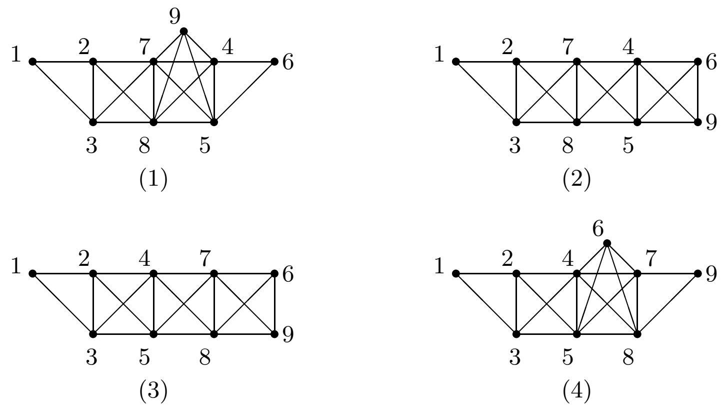


Figure 8: The MCSs corresponding to the pair in Figure 7

Note that the four edges between $\{2,3\}$ and $\{4,5\}$ are all cut by $\{7,8\}$ in graphs 1 and 2 in Figure 8. If at least one of the edges were not cut, $\{7,8\}$ could not be a connector in any combined version of the pair in Figure 7.

The theorem below considers the situation where a node is not a minimal connector in one subgraph of a graph-pair but contained in a minimal connector of the other subgraph.

Theorem 6.1 *Consider two node-subgraphs of $\mathcal{G}$, $\mathcal{G}_1$ and $\mathcal{G}_2$, and suppose that there is a node v in V_1 which is not a minimal connector of $\mathcal{G}_1$ and that there exists s in $\sigma(\mathcal{G}_2)$ such that $v \in s$. Then in every graph $\mathcal{H}$ in $\mathcal{M}(\mathcal{G}_1, \mathcal{G}_2)$,*

(i) *s forms a minimal connector in $\mathcal{H}$, and*

(ii) *in addition, if v is contained in a clique, $\mathcal{C}_v$, of $\mathcal{G}_1$, then each node in s is connected directly, in $\mathcal{H}$, to every node in C_v.*

Proof: Result (i) is obvious by Theorem 3.2 (ii). As for result (ii), suppose that there exists no edge, in $\mathcal{H}$, between the nodes v_1 and v_2, $v_2 \in s$ and $v_1 \in C_v \setminus s$. Since s is complete by Theorem 3.1, the supposition implies that there exists a new minimal connector s^* in $\mathcal{H}$

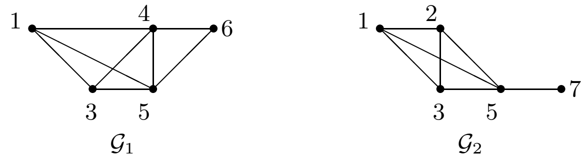


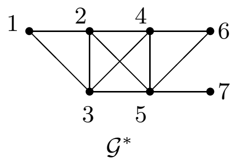
Figure 9: Two node-subgraphs of $\mathcal{G}^*$

such that $s^* \notin \sigma(\mathcal{G}_1) \cup \sigma(\mathcal{G}_2)$ since $v \in s$. This contradicts Theorem 3.2 (ii). Therefore, the desired result follows. □

This theorem states that in combining two marginal graphs $\mathcal{G}_1$ and $\mathcal{G}_2$, s is grafted onto $\mathcal{G}_1$ at the clique C_v. We will see below that actually C_v is the only clique where s is grafted.

We need not worry about the case where the node v of this theorem is itself a minimal connector of $\mathcal{G}_1$ since the minimal connectors of a graph are grafted onto another graph at the grafting locations which are determined by the minimal connectors of the latter graph. We simply treat the v as a minimal connector.

For two node-subgraphs, $\mathcal{G}_1$ and $\mathcal{G}_2$, of $\mathcal{G}$, it is possible that there are sets $s \in \sigma(\mathcal{G}_1)$ and $t \in \sigma(\mathcal{G}_2)$ such that either $s \subset t$ or vice versa. For example, consider a graph $\mathcal{G}^*$ as given in



whose node-subgraphs are given in Figure 9. In this figure, the minimal connector of $\mathcal{G}_2$ is a proper subset of that of $\mathcal{G}_1$. In this case, node 5 of $\mathcal{G}_2$ is type-1 grafted onto $\mathcal{G}_1$ at the clique $\{1, 3, 4, 5\}$ yielding a unique MCS which is the same as $\mathcal{G}^*$ except that nodes 1, 2, 3, 4, 5 now form a clique.

We close this section presenting a couple of theorems and a corollary that are very useful in combining graphical models no matter whether the minimal connectors are of mutiple nodes or not.

Theorem 6.2 *Let $\mathcal{G}_1$, $\mathcal{G}_2$ be two node-subgraphs of $\mathcal{G}$. Then a minimal connector $s \in \sigma(\mathcal{G}_2)$ is grafted onto $\mathcal{G}_1$ at a clique only.*

Proof: Let C_1 and C_2 be neighboring cliques in $\mathcal{G}_1$ and let s_1 be a minimal connector shared by the cliques. Then there can not be edges between a node in s and a node in $C_1 \setminus s_1$ and between a node in s and a node in $C_2 \setminus s_1$. If there are, it contradicts that s_1 is a minimal connector. Hence, the s must be grafted onto $\mathcal{G}_1$ at a clique only. □

Theorem 6.3 *Let $\mathcal{G}_1$, $\mathcal{G}_2$ be two node-subgraphs of $\mathcal{G}$ and assume that for some $s \in \sigma(\mathcal{G}_2)$, there is a subset t in $V(\mathcal{G}_1)$ such that $t \subseteq s$. Then*

(*i*) *t is contained in a clique C of $\mathcal{G}_1$ and is not a clique itself in $\mathcal{G}_1$, and*

(*ii*) *$s \in \sigma(\mathcal{G}_2)$ is grafted onto $\mathcal{G}_1$ at the clique C only in either type-1 or type-2, and when it is type-2 grafted, the minimal connectors contained in the clique remains as they are.*

Proof: See Appendix B.

According to the above two theorems, we graft a minimal connector of one node-subgraph onto another node-subgraph at a clique, and if a clique contains nodes of the minimal connector, then the minimal connector is grafted at the clique only.

When none of the nodes of a minimal connector of a node-subgraph are contained in another node-subgraph onto which the former is to be grafted, the minimal connector may be grafted at a clique in such a way that Lemma 3.1 and Theorem 3.2 may not be violated. Pairs 1β, 1γ, 2γ, 3β, and 3γ fall in this situation. As noted in section 5, the number of the corresponding MCSs is smaller when there are more node-sharings between a pair of node-subgraphs than when there are no or less node-sharings. For example, as for Pairs 1β and 1γ, the numbers of the possible MCSs are 5 and 16, respectively, and they are 19 and 112 as for Pairs 3β and 3γ, respectively.

When there is a minimal connector s in $\mathcal{G}_2$ with $s \cap V(\mathcal{G}_1) = \emptyset$, at which the minimal connector can be grafted is subject to the states of node-sharing between the two node-subgraphs. Although s is grafted at a fixed clique C of $\mathcal{G}_1$ in Corollary 6.2 below, the type of grafting as summarized in the corollary applies in general when several nodes are shared by both of $\mathcal{G}_1$ and $\mathcal{G}_2$ and some of the shared nodes are separated from the other shared nodes by s in $\mathcal{G}_2$.

Corollary 6.2 *Let $\mathcal{G}_1$, $\mathcal{G}_2$ be two node-subgraphs of $\mathcal{G}$ and assume that for some $s \in \sigma(\mathcal{G}_2)$, there is a subset t and a clique C in $V(\mathcal{G}_1)$ such that $t \subseteq s$ and $t \subset C$. Suppose that there are m nodes $v_1, \cdots, v_m$ in both of $\mathcal{G}_1$ and $\mathcal{G}_2$ and that v_1 is separated from the other v_i's by s in $\mathcal{G}_2$. Then, s can be grafted at C in type-2 only.*

Proof: See Appendix B.

Consider two decomposable graphical models M_1 and M_2 whose model structures are $\mathcal{G}_1$ and $\mathcal{G}_2$, respectively. According to the corollary, we can say that whenever the decomposable graphical models share variables and the variables are separated by a minimal connector s in $\mathcal{G}_2$ which is not in $\sigma(\mathcal{G}_1)$, the model structures are type-2 grafted at a clique of $\mathcal{G}_1$ by s of $\mathcal{G}_2$.

Furthermore, we can think of the case where two graphical models share some variables and none of them are separated by a minimal connector s of the model structure $\mathcal{G}_2$. We may regard this as a particular case of Corollary 6.2 taking s as a separator of an empty set from the shared variables. In this case, type-1 grafting is also possible. For instance, if v is contained in model M_2 only and conditionally independent of the shared variables given s, then s can be type-1 grafted at a clique of $\mathcal{G}_1$ and after grafting v remains separated by s from the other variables.

In this section, we have considered all the possible situations concerning grafting a minimal connector $s \in \mathcal{G}_2$ at a clique of $\mathcal{G}_1$. Denoting s as a minimal connector in $\mathcal{G}_2$, we can now summarize the results as the following grafting rules:

($R1$) If a node $v \in V_1$ is contained in s, s is grafted at a C_v of $\mathcal{G}_1$ only.

($R2$) Suppose that $s \cap V_1 = \emptyset$. If $V_1 \cap V_2 = \emptyset$, then s can be grafted at any clique of $\mathcal{G}_1$; otherwise, s is grafted according to the rule $R3$.

($R3$) Suppose that the set of nodes $\{v_1, \cdots, v_m\} = V_1 \cap V_2$. If s separates some of $v_1, \cdots, v_m$ from the others in $\mathcal{G}_2$, then s is type-2 grafted at a clique of $\mathcal{G}_1$; otherwise, s can be grafted in both of the types.

7 Combining three or more node-subgraphs

So far we have considered combining a pair of model structures. But we may often come across the cases where more than two graphical models are to be combined together. In this section we will address the issue whether the order of models in combining matters.

Theorem 4.3 proposes that for any number of node-subgraphs, there always exists a unique MCS which contains the actual supergraph. Suppose we are given $m > 2$ node-subgraphs, $\mathcal{G}_1, \cdots, \mathcal{G}_m$, for which $\cup_{i=1}^m V_i = V$, and that we combine graphs in the order of indexes of the subgraphs. Recall that $\mathcal{G}_{(j)} = \mathcal{G}_{\cup_{i=1}^j V_i}$. According to the theorem, there exists a unique MCS $\mathcal{H}_{12}$ in $\mathcal{M}(\mathcal{G}_1, \mathcal{G}_2)$ which contains $\mathcal{G}_{(2)}$ as an edge-subgraph. Let $\mathcal{F}_1 = \mathcal{G}_1$ and $\mathcal{F}_2 = \mathcal{H}_{12}$. We obtain $\mathcal{M}(\mathcal{H}_{12}, \mathcal{G}_3)$. Since $\mathcal{H}_{12} \supseteq^e \mathcal{G}_{(2)}$, there must exist, by Theorem 4.3 (ii), a unique graph $\mathcal{F}_3$ in $\mathcal{M}(\mathcal{H}_{12}, \mathcal{G}_3)$ such that $\mathcal{F}_3 \supseteq^e \mathcal{G}_{(3)}$ and so that $\mathcal{F}_{3(V_1 \cup V_2)} \supseteq^e \mathcal{G}_{(2)}$. In the same manner we can obtain a unique graph $\mathcal{F}_j$ in $\mathcal{M}(\mathcal{F}_{j-1}, \mathcal{G}_j)$ such that $\mathcal{F}_j \supseteq^e \mathcal{G}_{(j)}$ for $j = 3, 4, \cdots, m$.

Since $\mathcal{G}_{(m)} = \mathcal{G}$, $\mathcal{F}_m$ is the unique MCS containing $\mathcal{G}$ as an edge-subgraph. We will call the process of $\mathcal{F}_1$ through $\mathcal{F}_m$ *sequential structure combination* or SSC for short, and denote the collection of the MCSs such as $\mathcal{F}_m$ by $\mathcal{M}_{seq}(\mathcal{G}_1, \cdots, \mathcal{G}_m)$, where the order of subgraphs indicates the order of combination. We have just proved the following theorem.

Theorem 7.1 *Suppose we have* m *node-subgraphs of* $\mathcal{G}$, $\mathcal{G}_1, \cdots, \mathcal{G}_m$, *for which* $\cup_{i=1}^m V_i = V$. *Then there exists a unique graph,* $\mathcal{F}$, *in* $\mathcal{M}_{seq}(\mathcal{G}_1, \cdots, \mathcal{G}_m)$ *such that*

$$\mathcal{F} \supseteq^e \mathcal{G}. \tag{10}$$

The uniqueness property in Theorem 7.1 holds also when $\mathcal{G}$ is replaced in (10) by its edge-subgraph.

Theorem 7.2 *Suppose we have* m *node-subgraphs of* $\mathcal{G}$ *for which* $\cup_{i=1}^m V_i = V$, *and let* $\mathcal{G}' \subseteq^e \mathcal{G}$. *Then there exists a unique graph,* $\mathcal{F}$, *in* $\mathcal{M}_{seq}(\mathcal{G}_1, \cdots, \mathcal{G}_m)$ *such that* $\mathcal{F} \supseteq^e \mathcal{G}'$.

Proof: Since $\mathcal{G}' \subseteq^e \mathcal{G}$, it follows that $\mathcal{G}'_{(j)} \subseteq^e \mathcal{G}_{(j)}$ for $j = 1, \cdots, m$. At the jth step of the SSC, we get, by Theorem 4.3, a unique MCS $\mathcal{F}_j$ such that $\mathcal{F}_j \supseteq^e \mathcal{G}'_{(j)}$. The desired graph, $\mathcal{F}$, is obtained by putting $\mathcal{F} = \mathcal{F}_m$. □

The theorem below states that an SSC in any order of node-subgraphs yields a unique MCS, in the sense of Theorem 4.3, for a given collection of node-subgraphs of a graph.

Theorem 7.3 *Suppose there are* m *node-subgraphs,* $\mathcal{G}_1, \cdots, \mathcal{G}_m$, *for which* $\cup_{i=1}^m V_i = V$, *consider an edge-subgraph* $\mathcal{G}'$ *of* $\mathcal{G}$, *and let* $\mathcal{H}$ *be a graph in* $\mathcal{M}(\mathcal{G}_1, \cdots, \mathcal{G}_m)$ *for which* $\mathcal{H} \supseteq^e \mathcal{G}'$. *For any permutation* $r_1, r_2, \cdots, r_m$ *of* $1, 2, \cdots, m$, *the following holds:*
If $\mathcal{F} \in \mathcal{M}_{seq}(\mathcal{G}_{r_1}, \cdots, \mathcal{G}_{r_m})$ *satisfies that* $\mathcal{F} \supseteq^e \mathcal{G}'$, *then* $\mathcal{H} = \mathcal{F}$.

Proof: See Appendix B.

8 Further Discussion and Concluding Remarks

The results of the article hold without regard to the types of variables involved in models, and we confined our attention to the class of decomposable graphical models only. A major reason of it is that the class of the corresponding graphs is the intersection of the class of directed acyclic graphs (DAGs) and the class of undirected graphs. Lauritzen and Spiegelhalter (1988) showed how to convert a DAG into a decomposable undirected graph.

Every node-subgraph of a decomposable graph is also decomposable. In this article, the notion of minimal connector plays a key role in combining model structures in the context that the collection, say $\mathcal{C}$, of the minimal connectors of a given set of node-subgraphs is found in the collection of the minimal connectors of the actual graph, say $\mathcal{G}$, and that there exists a unique model structure which is an edge-supergraph of $\mathcal{G}$ and the collection of whose minimal connectors is the same as $\mathcal{C}$.

The aim of graph-combination of the node-subgraphs, $\mathcal{G}_1, \cdots, \mathcal{G}_m$, as proposed in this article, is not searching for the actual graph $\mathcal{G}$ but searching for the MCS $\mathcal{H}$ that contains $\mathcal{G}$ as an edge-subgraph. In the searching process, Theorem 3.2 (ii) plays an important role. According to this result, we have only to pay attention to the locations of the minimal connectors in combining node-subgraphs. The resulting MCSs are not unique but there exists only one MCS that contains $\mathcal{G}$ as an edge-subgraph. This unique existence makes combining node-subgraphs, i.e., combining marginal model structures theoretically sound. This is parallel to that a graphical log-linear model M is contained in the collection of the graphical log-linear models which are obtained by applying the combining method of model structures (the hypermodelling process of Fienberg and Kim (1999))to conditional log-linear models of M (This is stated as Traceability Theorem in Fienberg and Kim (1999)).

Although we have used graph terminologies, the main results are about combining the structures of decomposable graphical models. Suppose we have k collections of conditional independence relationships among variables in k different sets of variables, $V_1, \cdots, V_k$. A method of obtaining a reasonably good list of conditional independence relationships among the variables in $\cup_{i+1}^{k} V_i$ is proposed in this article based on the consistency principle that the Markov properties in a graphical model are preserved in its submodels and vice versa.

The combination of model structures is implemented through graph grafting, which takes place only at the cliques of a node-subgraph so that the conditional independence relationships in the subgraphs are preserved during the grafting process.

We have not specified any distributional assumptions for the variables in the graphical models. When all the variables are of the same kind, either continuous or discrete, Bildikar and Patil (1968) show that the probability distribution for a model and that for any of its submodels belong to the same family of distributions if the distribution is of exponential-type. But when the graphical model is of mixed variables, a model may have a probability distribution that is different from that of its submodel. For example, the distribution of a submodel of a graphical model which has a conditional Gaussian (CG) distribution (Lauritzen, 1989, 1992) is not necessarily CG. As far as the CG distribution is concerned, we may need more restriction upon the model structure and thus upon the corresponding MCSs. Any future research on this class of mixed graphical models as an extension of the results of this article deserves our attention.

It is noted in section 6 that as there are more variables (i.e., nodes) shared by submodels, we have less MCSs corresponding to the submodels. From the simple model-pair in Figure 1, we have as the corresponding MCS the last model structure, [12][245], in expression (1); as for the pair, [12][23] and [24][25], the corresponding MCSs are the maximal model structures in expression (2) as given in (8). And there are four MCSs in Figure 8 corresponding to the model pair as in Figure 7. Since we do not know the true model structure $\mathcal{G}$, neither can we see which of the models in (8) contains the true model unless we have any further information on the relationships among the variables involved. The information may be given among others in the form of conditional independence, multi-way interaction, if-then statement, cause-effect statement, the nature of variables, or observed data. For instance, if an additional piece of information is that X_2 and X_3 directly influences X_4, then the second model in (8) is the only possibility. As for the case of Figure 8, information on whether X_6 and X_9 are conditionally independent helps much in searching for the true model.

The graphs of submodels can be from experts, data or from both. In model building, it may be convenient to add variables to a given model one after another. In particular, variables may be added in the order of cause-effect sequence as for a recursive model (Wermuth and Lauritzen, 1983). But when we have to merge several recursive models, we may easily be trapped in inconsistency of inter-relationships and complexity of model structures if we merge them heuristically one node after another. The graph-grafting method as proposed in this article will save us from such inconsistency of inter-relationships since the method is grounded upon the theories of conditional independence and graph separation and also from the complexity of model structures since the grafting is operated with minimal connectors and cliques only.

Apart from the distributional assumptions on models, the results of this article are applicable to any collection of graphical models, provided that the corresponding graphs are equivalent to decomposable graphs (Andersson et al., 1997a). The results may open a road leading us from a collection of marginal model structures among random variables to a collection of global (as against "marginal") model structures among the random variables that are involved in at least one of the marginal model structures. Graphical models are widely used in various forms in the research fields such as computer science, artificial intelligence, social and biological sciences, decision science, and medical research. Model structures or knowledge structures that are representable in graphs and developed for different sets of random variables may be combined in a consistent manner towards a larger structure of random variables by applying the methods as derived in this article.

Appendix A

Pair 1α

Pair 1β

Pair 2α

Pair 2β

Pair 2γ

Pair 3α

Pair 3γ

Pair 4

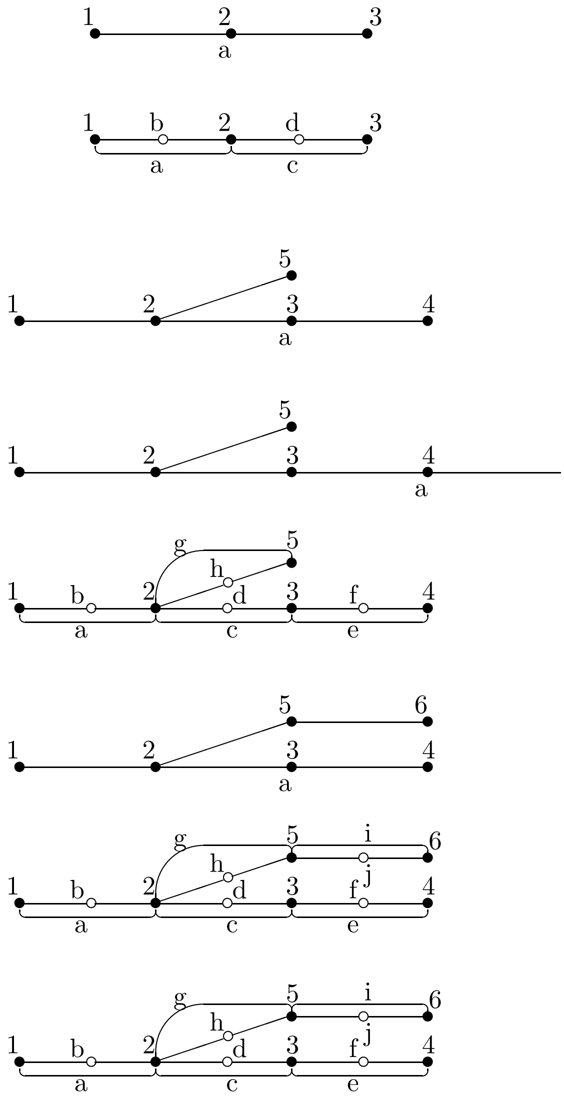


Figure 10: The grafting locations for the graph pairs except Pairs 1γ and 3β. The grafting locations for Pair 4 are of minimal connector $\{8\}$ of $\mathcal{G}_2$ of the pair.

Table 3: The numbers of the MCSs of the graph pairs in Figure 10. "GL" is the initials of "grafting location".

pair	1α	1β				2α	2β	2γ			
GL	a	a	b	c	d	a	a	a	b	c	d
#(MCSs for each GL)	2	1	2	1	1	6	8	8	8	8	12
#(MCSs for each pair)	2	5				6	8	68			

2γ				3α	3γ										4	
e	f	g	h	a	a	b	c	d	e	f	g	h	i	j	a	b
8	8	8	8	8	10	10	10	16	10	10	10	16	10	10	1	1
				8	112										18	

4							
c	d	e	f	g	h	i	j
1	4	4	4	1	2	1	1

Appendix B: Proofs

Proof of Theorem 3.3: Suppose that $\mathcal{G}$ is composed of k cliques. Then, since $\mathcal{G}$ is decomposable, we can label the cliques from 1 through k, $C_1, \cdots, C_k$, so that

$$\text{for all } i > 1 \text{ there is a } j < i \text{ such that } B_i \subseteq C_j, \tag{11}$$

where $B_i = C_i \cap \left(\cup_{l=1}^{i-1} C_l \right)$ (Theorem 7 of Pearl (1988) and Proposition 2.17 of Lauritzen (1996)). This implies that B_i is a minimal connector in $\mathcal{G}_{A_i}$ where $A_i = \cup_{l=1}^{i} C_l$. But actually, B_i is a minimal connector of $\mathcal{G}$. Suppose there are cliques, C_{i_1}, C_{i_2}, and C_{i_3}, $i_1 < i \leq i_2 < i_3$, such that there is a path from C_{i_1} to C_{i_2} passing through C_{i_3} but bypassing B_i. Then this violates the clique numbering property (11) of a decomposable graph. So such a clique C_{i_3} does not exist, implying that B_i is also a minimal connector of $\mathcal{G}$, i.e., $\{B_i\} \subseteq \sigma(\mathcal{G})$. Since there is no minimal connectors in $\mathcal{G}$ other than the B_i's, $\{B_i\} = \sigma(\mathcal{G})$. □

Proof of Theorem 4.3: We first show (i). According to Theorem 3.2 and Lemma 4.1, we have for any MCS $\mathcal{H}^{**}$ of the m node-subgraphs

$$\sigma(\mathcal{H}^{**}) \subseteq \sigma(\mathcal{G}).$$

Note that $\sigma(\cdot)$ is a collection of sets of nodes and so that this inequality does not necessarily imply that $\mathcal{G}$ is an edge-subgraph of $\mathcal{H}^{**}$. In other words, the topology of $\sigma(\mathcal{G})$ can be different from that of $\sigma(\mathcal{H}^{**})$.

If $\sigma(\mathcal{H}^{**}) = \sigma(\mathcal{G})$, then $\mathcal{G}$ itself is an MCS of the m node-subgraphs. If $\sigma(\mathcal{H}^{**}) \subset \sigma(\mathcal{G})$ and we let $\sigma' = \sigma(\mathcal{G}) \setminus \sigma(\mathcal{H}^{**})$, then we can change $\mathcal{G}$ into an MCS of the m node-subgraphs by connecting all the node-pairs of the cliques that share an element set of σ' so that the neighboring cliques may form a larger clique. By doing the same for each of the other sets in σ', we obtain an MCS whose set of minimal connectors is the same as $\sigma(\mathcal{H}^{**})$.

As for the uniqueness of MCS, suppose there are two different MCSs $\mathcal{H}'$ and $\mathcal{H}''$ which contains $\mathcal{G}$ as an edge-subgraph. By definition, neither of $\mathcal{H}'$ and $\mathcal{H}''$ is an edge-subgraph of the other. By Theorem 3.2 (ii), we have

$$\sigma(\mathcal{H}') = \cup_{i=1}^{m} \sigma(\mathcal{G}_i) = \sigma(\mathcal{H}''). \tag{12}$$

Thus we have

$$\emptyset \neq E(\mathcal{H}') \setminus E(\mathcal{G}) \neq E(\mathcal{H}'') \setminus E(\mathcal{G}) \neq \emptyset \tag{13}$$

since both $\mathcal{H}'$ and $\mathcal{H}''$ are edge-supergraphs of $\mathcal{G}$.

By Theorem 4.2, $\mathcal{H}'$ and $\mathcal{H}''$ are both decomposable. So expression (13) implies that

$$\sigma(\mathcal{H}') \neq \sigma(\mathcal{H}''), \tag{14}$$

since both $\mathcal{H}'$ and $\mathcal{H}''$ are edge-supergraphs of $\mathcal{G}$. For instance, let

$$(v_1, v_2) \in E(\mathcal{H}') \setminus E(\mathcal{G}) \text{ and } (v_1, v_2) \notin E(\mathcal{H}'') \setminus E(\mathcal{G}). \tag{15}$$

Then, there must exist a minimal connector s in $\mathcal{H}''$ such that

$$v_1 \perp v_2 | s$$

but there is no minimal connector of v_1 and v_2 in $\mathcal{H}'$. Both $\mathcal{H}'$ and $\mathcal{H}''$ are decomposable and edge-supergraphs of $\mathcal{G}$. So expression (15) implies that v_1 and v_2 are each contained in a pair of neighboring cliques in $\mathcal{H}''$ but not contained in the same clique. Thus it is impossible that $s \in \sigma(\mathcal{H}')$. This means expression (14), contradicting (12). Therefore, $\mathcal{H}' = \mathcal{H}''$.

We now prove (ii). By condition of the theorem, $\sigma(\mathcal{G}) \subseteq^e \sigma(\mathcal{G}')$. So, $\sigma(\mathcal{H}^{**}) \subseteq^e \sigma(\mathcal{G}')$. Hence, the existence is proved in the same way as the first half of the proof of result (i) by replacing $\mathcal{G}$ therein with $\mathcal{G}'$. As for the uniqueness, we may apply the same argument as the second half of the proof of result (i) by replacing $\mathcal{G}$ therein with $\mathcal{G}'$, since $\mathcal{G}' \subseteq^e \mathcal{G}$. □

Proof of Theorem 6.3: If t shares nodes with more than one cliques in $\mathcal{G}_1$, this implies that $s(\supseteq t)$ also shares nodes with multiple cliques, i.e., s can not be a minimal connector, in a resulting MCS by Lemma 3.1, which violates Theorem 3.2.

If $t \in \sigma(\mathcal{G}_1)$, it is obvious that t itself is not a clique in $\mathcal{G}_1$. Otherwise, suppose that t is a clique. Then, this means that t contains as a proper subset at least one minimal separator, say s' in $\mathcal{G}_1$ separating t from a clique, say C', and so that there is at least one node in $t \setminus C'$ which is separated from C' by s'. This implies, by Lemma 3.1, that t does not form a complete subgraph in $\mathcal{G}$, which is a contradiction to that s is a complete subgraph and $t \subseteq s$. Therefore, t can not be a clique in $\mathcal{G}_1$. This proves result (i).

We know from Theorem 6.2 that grafting takes place at cliques only. If $s \in \sigma(\mathcal{G}_1)$, then it is obvious that s is type-1 grafted at a clique which contains s in $\mathcal{G}_1$. We will next consider the case that $s \notin \sigma(\mathcal{G}_1)$. That $s \in \sigma(\mathcal{G}_2)$ means that there are at least two sets of nodes that are separated by s in $\mathcal{G}_2$. Let those sets be $A_1, \cdots, A_l$. Then the following three cases are possible:

(i) $V(\mathcal{G}_1) \cap \left(\cup_{i=1}^{l} A_i \right) = \emptyset$.

(ii) There is only one A_i which intersects with $V(\mathcal{G}_1)$. Let it be A_1.

(iii) There are more than one A_i which intersect with $V(\mathcal{G}_1)$.

In case (i), s can be either type-1 grafted or type-2 grafted at C. When s is type-1 grafted, the nodes of one of the l sets, $A_1, \cdots, A_l$, are attached to some cliques of $\mathcal{G}_1$ and the other $l-1$ A_i's are separated from each other and also from $V(\mathcal{G}_1)$ by s. By the definition of MCS (Definition 3.1), if all the l A_i's are separated from $V(\mathcal{G}_1)$, then the resulting combined structures can not be MCSs, since we can add edges still keeping expression (4) valid. When s is type-2 grafted, C can be separated into j $(2 \le j \le l)$ subsets of nodes, subject to the states of $\mathcal{G}_1$, ending up with separating $\mathcal{G}_1$ into j parts. When C is separated into two parts, the nodes of one of A_i's are attached to some cliques in one of the j parts of $\mathcal{G}_1$, those of another of A_i's are attached to the other part of the graph, and the rest of A_i's are separated from the graph and also from each other by s. Analogously, when C is separated into $k(< l)$ parts, $l-k$ A_i's are separated by s from $V(\mathcal{G}_1)$.

In case (ii), s can also be either type-1 grafted or type-2 grafted. When type-1 grafted, the nodes in $A_1 \setminus s$ are attached to some cliques of $\mathcal{G}_1$ and $A_2, \cdots, A_l$ are separated from $V(\mathcal{G}_1)$ by s. Type-2 grafting is implemented in the same way as for case (i).

In case (iii), a type-1 grafting of s is impossible, since $s \notin \sigma(\mathcal{G}_1)$ and there are nodes to be separated by s in $V(\mathcal{G}_1)$. For convenience' sake, suppose that only the first m out of $A_1, \cdots, A_l$ intersect with $V(\mathcal{G}_1)$. Then, by the definition of MCS, the clique C can be separated by s into j $(m \le j \le l)$ parts, subject to the states of $\mathcal{G}_1$ ending up with j parts of $\mathcal{G}_1$. When C is separated into $k(j \le k < l)$ parts by s, the other A_i's, $i = k+1, \cdots, l$, are separated from $V(\mathcal{G}_1)$ and from each other by s.

When a type-2 grafting takes place, s can not separate any minimal connector of $\mathcal{G}_1$ because, if such a separation takes place in $s' \in \sigma(\mathcal{G}_1)$, s' is no more a minimal connector in a resulting MCS, which violates Theorem 3.2. Therefore, all the existing minimal connectors in $\mathcal{G}_1$ are not affected during a type-2 grafting. □

Proof of Corollary 6.2: By definition, graph-grafting takes place either in type-1 or in type-2. Hence, we have only to show that type-1 grafting is impossible under the situation of the corollary. According to Theorem 6.3, s is grafted at C only. Suppose that s is type-1 grafted at C. Then there must be a new clique C' added to $\mathcal{G}_1$ which is connected to C through s, in other words, is separated from C by s. By the condition of the corollary, either $v_1 \in C'$ or $v_j \in C'$ for $j \neq 1$. But this contradicts to that C' is a new addition to $\mathcal{G}_1$ since all the v_i's are already in $\mathcal{G}_1$. Therefore, we have the desired result. □

Proof of Theorem 7.3: First of all, assume that the SSC is implemented in the order of $\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_m$ and consider the sequence $\{\mathcal{F}_j; j = 1, \cdots, k\}$ such that $\mathcal{F}_j \supseteq^e \mathcal{G}'_{(j)}$. Obviously, $\sigma(\mathcal{F}_1) = \sigma(\mathcal{G}_1)$. For j, $1 < j < m$, suppose that $\sigma(\mathcal{F}_j) = \cup_{i=1}^{j} \sigma(\mathcal{G}_i)$. Then we have, by Theorem 3.2, that

$$\begin{aligned} \sigma(\mathcal{F}_{j+1}) &= \sigma(\mathcal{F}_j) \cup \sigma(\mathcal{G}_j) \\ &= \cup_{i=1}^{j+1} \sigma(\mathcal{G}_j). \end{aligned}$$

So,

$$\sigma(\mathcal{F}) = \sigma(\mathcal{F}_m) = \cup_{i=1}^{m} \sigma(\mathcal{G}_i) = \sigma(\mathcal{H}). \qquad (16)$$

Now suppose that $\mathcal{H} \neq \mathcal{F}$. Then, by result (16), it implies that there exist at least one pair of nodes in $\mathcal{H}$ whose location is not the same between $\mathcal{H}$ and $\mathcal{F}$, which is impossible since both of them are edge-supergraphs of $\mathcal{G}'$. Therefore, it must be that $\mathcal{H} = \mathcal{F}$.

The preceeding argument is symmetric in the order of the node-subgraphs, which makes the proof complete. □

References

Andersson, S. A., Madigan, D., and Perlman, M. D. (1997a), On the Markov equivalence of chain graphs, undirected graphs, and acyclic digraphs, *Scandinavian Journal of Statistics*, **24**, 81-102.

Andersson, S. A., Madigan, D., and Perlman, M. D. (1997b). A characterization of Markov equivalence classes for acyclic digraphs, *The Annals of Statistics*, **25**, 2, 505-541.

Asmussen, S. and Edwards, D. (1983), Collapsibility and response variables in contingency tables, *Biometrika*, **70**, 3, 567-78.

Bacchus, F. (1993). Using first-order probability logic for the construction of Bayesian networks. In *Proceedings of the Ninth Conference on Uncertainty in Artificial Intelligence* (pp. 219-226). Los Angeles, CA: Morgan Kaufmann.

Bildikar, S. and Patil, G.P. (1968), Multivariate exponential-type distributions, *Annals of Mathematical Statistics*, **39**, 1316-1326.

Chickering, D.M. and Heckerman, D. (1999). Fast learning from sparse data. In *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence* (pp. 109-115). San Francisco, CA: Morgan Kaufmann.

Cooper, G. and Herskovits, E. (1991). A Bayesian method for constructing Bayesian belief networks from databases. In *Proceedings of Seventh Conference on Uncertainty in Artificial Intelligence* (pp. 86-94). Los Angeles, CA: Morgan Kaufmann.

Darroch, J. N., Lauritzen, S. L. and Speed, T. P. (1980). Markov fields and log-linear interaction models for contingency tables. *The Annals of Statistics*,**8**, 3, 522-539.

Dawid, A. P. (1979). Conditional independence in statistical theory, *Journal of Royal Statistical Society*, B, **41**, 1, 1-31.

Dawid, A. P. (1980). Conditional independence for statistical operation, *The Annals of Statistics*, **8**, 3, 598-617.

Fienberg, S.E. (1980). *The Analysis of Cross-Classified Categorical Data.* Cambridge, MA: The MIT Press.

Fienberg, S.E. and Kim, S.-H. (1999), Combining conditional log-linear structures, *Journal of The American Statistical Association*, **445**, 94, 229-239.

Frydenberg, M. (1990). The chain graph Markov property, *Scandinavian Journal of Statistics*, **17**, 333-353.

Goldman, R.P. and Charniak, E. (1993). A language for construction of belief networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **15**, 3, 196-208.

Haddawy, P. (1994). Generating Bayesian networks from probability logic knowledge bases. In *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence*, (pp. 262-269). Los Angeles, CA: Morgan Kaufmann.

Kiiveri, H., Speed, T.P., and Carlin, J.B. (1984), Recursive causal models, *Journal of The Australian Mathematical Soc. A*, **36**, 30-52.

Lauritzen, S. L. (1989), Mixed graphical association models. *The Scandinavian Journal of Statistics*, **16**, 273-306.

Lauritzen, S. L. (1992). Propagation of probabilities, means, and variances in mixed graphical association models. *Journal of The American Statistical Assoc.*, **87**, 420, 1098-1108.

Lauritzen, S.L. (1996), *Graphical Models*, Oxford: Clarendon Press.

Lauritzen, S. L., Speed, T. P., and Vijayan, K. (1984), Decomposable graphs and hypergraphs, *Journal of the Australian Mathematical Soc. A*, **36**, 12-29.

Lauritzen, S.L. and Spiegelhalter, D.J. (1988), Local Computations with Probabilities on Graphical Structures and their Application to Expert Systems, *Journal of Royal Statistical Soc.* B, **50**, 2, 157-224.

Leimer, H.-G. (1989), Triangulated graphs and marked vertices. In *Graph theory in memory of G. A. Dirac* (eds: L. D. Andersen et al.). *Annals of Discrete Mathematics*, **41**, 311-324.

Mahoney, S.M. and Laskey, K.B. (1999). Representing and combining partially specified CPTs. In K.B. Laskey and H. Prade, (Eds.), *Uncertainty in Artificial Intelligence* (391-400). San Francisco, California: Morgan Kaufmann Publishers.

Pearl, J. (1988), *Probabilistic Reasoning In Intelligent Systems: Networks of Plausible Inference,* San Mateo, CA.: Morgan Kaufmann.

Poole, D. (1993). Probabilistic horn abduction and Bayesian networks, *Artificial Intelligence*, **64**, 1, 81-129.

Wermuth, N. (1980), Linear recursive equations, covariance selection, and path analysis. *Journal of The American Statistical Assoc.*, **75**, 373, 963-972.

Wermuth, N. and Lauritzen, S. L. (1983), Graphical and recursive models for contingency tables, *Biometrika*, **70**, 3, 537-552.

Whittaker, J. (1990), *Graphical Models in Applied Multivariate Statistics*, New York: Wiley.