

MATHEMATICAL ENGINEERING TECHNICAL REPORTS

Standard Imsets for Undirected and Chain Graphical Models

Takuya KASHIMURA and Akimichi TAKEMURA

METR 2011-06

February 2011

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: <http://www.keisu.t.u-tokyo.ac.jp/research/techrep/index.html>

The METR technical reports are published as a means to ensure timely dissemination of scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the authors or by other copyright holders, notwithstanding that they have offered their works here electronically. It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author's copyright. These works may not be reposted without the explicit permission of the copyright holder.

Standard imsets for undirected and chain graphical models

Takuya Kashimura* and Akimichi Takemura*†

February 2011

Abstract

We derive standard imsets for undirected graphical models and chain graphical models. Standard imsets for undirected graphical models are described in terms of minimal triangulations for maximal prime subgraphs of the undirected graphs. For describing standard imsets for chain graphical models, we first define a triangulation of a chain graph. We then use the triangulation to generalize our results for the undirected graphs to chain graphs.

Keywords and phrases: conditional independence, decomposable graph, maximal prime subgraph, triangulation.

1 Introduction

The notion of imsets introduced by Studený [17] provides a very convenient algebraic method for encoding all conditional independence (CI) models which hold under a discrete probability distribution. However, a class of imsets does not satisfy the uniqueness property: a number of different imsets represent the same CI model.

Thus some questions related to the uniqueness property arise [17]. One of them is the task of characterizing equivalent imsets. For example, in the case of classical graphical models [10], their equivalence classes are characterized by Andersson et al. [1] and Frydenberg [4] in graphical terms. Studený [14] related a CI model induced by a imset to some face of a special polyhedral cone, and an algorithm for CI inference based on this cone is studied in [2].

*Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo.

†JST, CREST

Another question is to find a suitable representative for every equivalence class. This is motivated by a practical question about learning CI models (see Section 4.4 in [16] and Section 4 in [22]). As a subproblem of this, explicit expressions of imsets for important classes of graphical models, such as directed acyclic graphical (DAG) models and decomposable models, are given in [17]. Imsets for some chain graphical (CG) models are also known [19]. They are called standard imsets and have attractive simple forms. One of their advantages is that they give a simple method to test whether two graphs have the same CI model. Another advantage is that it provides a translation of graphical models into the framework of imsets. Thus standard imsets offer a new algebraic approach for learning graphical models [20, 8].

In this paper we derive standard imsets for undirected graphical (UG) models and general CG models. Our standard imsets generalize those for DAG models and decomposable models. For UG models we consider all minimal triangulations of an undirected graph in accordance with maximal prime subgraphs and then use the standard imsets for minimal triangulations (which are decomposable models) for defining our standard imset. For CG models we first define a triangulation of a chain graph. We then use the triangulation to generalize our results for undirected graphs to chain graphs.

The organization of the paper is as follows. In Section 2 we summarize basic definitions and known facts on imsets and graphs, including standard imsets for DAG models and decomposable models. In Section 3 we derive standard imsets for UG models. In Section 4 we introduce a notion of triangulation of a chain graph and based on the triangulation we derive standard imsets for CG models. We conclude the paper with some remarks in Section 5.

2 Preliminaries

In this section we summarize our notation, definitions and relevant preliminary results concerning conditional independence, imsets and graphical models.

2.1 Conditional independence and imsets

First we set up notation for conditional independence and imsets following Studený [17].

Let N be a finite set of variables and let $\mathcal{P}(N) = \{A : A \subseteq N\}$ denote the power set of N . For convenience, we write the union $A \cup B$ of subsets of N as AB . A singleton set $\{i\}$ is simply written as i . As usual, \mathbb{R} ,

\mathbb{Z} , \mathbb{N} denote reals, integers and natural numbers, respectively. For pairwise disjoint subsets, $A, B, C \subseteq N$, we write this triplet by $\langle A, B | C \rangle$, and the set of all disjoint triplets $\langle A, B | C \rangle$ over N by $\mathcal{T}(N)$. As usual, for a probability distribution P over N , $A \perp\!\!\!\perp B | C [P]$ denotes the conditional independence statement of variables in A and in B given the variables in C under P . The case $C = \emptyset$ corresponds to the marginal independence of A and B . In this paper, we regard a triplet $\langle A, B | C \rangle$ as an independence statement. Then the set of conditional independence statements under P is denoted as

$$\mathcal{M}_P = \{\langle A, B | C \rangle \in \mathcal{T}(N) : A \perp\!\!\!\perp B | C [P]\}.$$

We call \mathcal{M}_P the *conditional independence model* induced by P .

An *imset* over N is an integer-valued function $u : \mathcal{P}(N) \rightarrow \mathbb{Z}$, or alternatively, an element of $\mathbb{Z}^{|\mathcal{P}(N)|} = \mathbb{Z}^{2^{|N|}}$. The identifier δ_A of a set $A \subseteq N$ is defined as

$$\delta_A(B) = \begin{cases} 1, & B = A, \\ 0, & B \neq A, B \subseteq N. \end{cases}$$

For a triplet $\langle A, B | C \rangle \in \mathcal{T}(N)$, a *semi-elementary imset* $u_{\langle A, B | C \rangle}$ is defined as

$$u_{\langle A, B | C \rangle} = \delta_{ABC} + \delta_C - \delta_{AC} - \delta_{BC}.$$

If $A = a$ and $B = b$ are singletons, the imset $u_{\langle a, b | C \rangle}$ is called *elementary*. The set of all elementary imsets is denoted by $\mathcal{E}(N)$. Let $\text{cone}(\mathcal{E}(N)) \subseteq \mathbb{R}^{2^{|N|}}$ be the polyhedral cone generated by all the elementary imsets. It can be shown that every elementary imset is a generator of an extreme ray of the cone $(\mathcal{E}(N))$ [14]. A *combinatorial imset* is an imset which can be written as a non-negative integer combination of elementary imsets. The set of all combinatorial imsets is denoted by $\mathcal{C}(N)$. Let

$$\mathcal{S}(N) = \text{cone}(\mathcal{E}(N)) \cap \mathbb{Z}^{|\mathcal{P}(N)|}.$$

An element of $\mathcal{S}(N)$ is called a *structural imset*. Note that $\mathcal{C}(N) \subseteq \mathcal{S}(N)$ by definition, however, it is known that this inclusion is strict for $|N| \geq 5$ [7].

A conditional independence statement induced by a structural imset is defined as follows:

Definition 2.1. For $u \in \mathcal{S}(N)$ and a triplet $\langle A, B | C \rangle \in \mathcal{T}(N)$, we define a *conditional independence statement with respect to u* as

$$A \perp\!\!\!\perp B | C [u] \iff \exists k \in \mathbb{N}, k \cdot u - u_{\langle A, B | C \rangle} \in \mathcal{S}(N).$$

The independence model induced by u is denoted by

$$\mathcal{M}_u = \{ \langle A, B | C \rangle \in \mathcal{T}(N) : A \perp\!\!\!\perp B | C [u] \}.$$

It can be shown that the structure of conditional independence models induced by structural imsets depends only on the face lattice of $\text{cone}(\mathcal{E}(N))$, not on each imset [14]. Therefore implications of conditional independence models induced by imsets correspond to those of faces of $\text{cone}(\mathcal{E}(N))$. Next lemma, which is very useful for our proofs in later sections, follows from this fact.

Lemma 2.2. (Studený [17]) For $u, u' \in \mathcal{S}(N)$,

$$\mathcal{M}_{u'} \subseteq \mathcal{M}_u \iff \exists k \in \mathbb{N}, k \cdot u - u' \in \mathcal{S}(N).$$

The method of imsets is very powerful, because conditional independence models induced by discrete probability measures are always represented by structural imsets.

Theorem 2.3. (Studený [17]) For every discrete probability measure \mathbb{P} over N , there exists a structural imset $u \in \mathcal{S}(N)$ such that $\mathcal{M}_u = \mathcal{M}_{\mathbb{P}}$.

2.2 Graphs and graphical models

Here we summarize relevant facts on graphs and graphical models following Lauritzen [10], Studený, Roverato and Štěpánová[19], Leimer [11], and Hara and Takemura [5].

Throughout this paper, we consider a simple graph $G = (V(G), E(G))$, $V(G) = N$, $E(G) \subseteq N \times N$. An edge $(a, b) \in E(G)$ is *undirected* if $(b, a) \in E(G)$. We denote an undirected edge by $a - b$. If $(b, a) \notin E(G)$, we call (a, b) *directed* and denote $a \rightarrow b$. An *undirected graph* (UG) contains only undirected edges, while a directed graph contains only directed ones. The *underlying graph* of a graph G is the undirected graph obtained from G by replacing every directed edge with an undirected one. For a subset $S \subseteq N$, G_S denotes the subgraph of G induced by S . A graph is *complete* if all vertices are joined by an edge. A subset $K \subseteq N$ is a *clique* if G_K is complete. In particular, an empty set $K = \emptyset$ is a clique. A clique K is maximal if no proper superset $K' \supset K$ is a clique in G . \mathcal{K}_G denotes the set of maximal cliques of G .

Two vertices $a, b \in N$ are adjacent if $(a, b) \in E(G)$ or $(b, a) \in E(G)$. If $a \rightarrow b$, then a is a *parent* of b and b is a *child* of a . For a vertex $c \in N$, we denote the set of parents and the set of children of c in G by $\text{pa}_G(c)$ and

$\text{ch}_G(c)$, respectively. For a subset $C \subseteq N$, let $\text{pa}_G(C) = \bigcup_{c \in C} \text{pa}_G(c) \setminus C$ and $\text{ch}_G(C) = \bigcup_{c \in C} \text{ch}_G(c) \setminus C$. We will omit subscript G if it is obvious from the context.

A *path* of length k from a to b is a sequence $a = c_1, \dots, c_{k+1} = b$ of distinct vertices such that $(c_i, c_{i+1}) \in E(G)$ for $i = 1, \dots, k$. If a path contains only undirected edges, it is an undirected path and otherwise (i.e. it contains at least one directed edge) directed. A vertex $a \in N$ is an *ancestor* of $b \in N$ if there exists a path from a to b . Let $\text{an}_G(a)$ be the set of all ancestors of a . The ancestral set $\text{an}_G(C)$ of a subset $C \subseteq N$ is defined as $\text{an}_G(C) = \bigcup_{c \in C} \text{an}_G(c)$. Let c_1, \dots, c_k a path with $(c_k, c_1) \in E(G)$. Then we call the sequence c_1, \dots, c_k, c_1 a *cycle* of length k . Analogously to paths, a cycle is undirected if it contains only undirected edges, otherwise directed. A *directed acyclic graph* (DAG) is a directed graph containing no directed cycles.

A subset $C \subseteq N$ is said to be *connected* if there exists an undirected path from a to b for all $a, b \in C$ in the subgraph G_C . A *connectivity component* of G is a maximal connected subset in G with respect to set inclusion. The connectivity components in G form a partition of N . A *chain graph* (CG) G is a graph whose connectivity components C_1, \dots, C_m are ordered such that every edge $a \rightarrow b \in E(G)$ with $a \in C_i, b \in C_j$ for $i < j$ is directed. Equivalently, a chain graph is defined as a graph containing no directed cycles. The connectivity components of a chain graph are called *chain components*. The set of chain components of a chain graph G is denoted by \mathcal{C}_G . The chain components are most easily found by removing all directed edges from G before taking connectivity components. Both undirected graphs and directed acyclic graphs are chain graphs. In fact, a chain graph is undirected if $m = 1$, and directed acyclic if each chain component contains only one vertex. Suppose two chain graphs G, H have the same underlying graph. Then we say H is *larger* than or equal to G if $a - b \in G$ implies $a - b \in H$. In this case, we write $H \geq G$. By definition, H has more undirected edges than G if H is larger than G .

We now discuss maximal prime subgraphs of an undirected graph G . A non-empty subset $\emptyset \neq S \subset N$ is a *separator* if the number of connectivity components increases when S is removed from G . $S = \emptyset$ is a separator (and only if) G is not connected. A separator S is a *clique separator* if S is a clique. For two vertices $u, v \in N$ with $u - v \notin E(G)$, a separator S is called a (u, v) -*separator* if u and v belong to different components of $G_{N \setminus S}$. A *minimal vertex separator* is a minimal (u, v) -separator for some $u, v \in N$ with respect to set inclusion. For $\langle A, B | C \rangle \in \mathcal{T}(N)$, we say that A and B are separated by C if C is (a, b) -separator for all $a \in A$ and $b \in B$.

A graph G is *prime* if G has no clique separators. Let $G_V, V \subseteq N$, be prime. Then G_V is a *maximal prime subgraph* (mp-subgraph) and V is a *maximal prime component* (mp-component) of G , if there is no proper superset $V' \supset V$ such that $G_{V'}$ is prime. The set of mp-components of G is denoted by \mathcal{V}_G . There exists an order $V_1, \dots, V_m, m = |\mathcal{V}_G|$, of \mathcal{V}_G such that

$$\forall i \in \{2, \dots, m\}, \exists k \in \{1, \dots, i-1\}, S_i \equiv V_i \cap \bigcup_{j < i} V_j \subseteq V_k.$$

This sequence is said to be *D-ordered*, or alternatively, to have a *running intersection property* (RIP) [10]. For each i , S_i is a clique minimal vertex separator. Define $\mathcal{S}_G = \{S_2, \dots, S_m\}$. Then \mathcal{S}_G is the set of all clique minimal vertex separators in G . Moreover, the number of $S \in \mathcal{S}_G$ which appears among S_2, \dots, S_m may be more than one. This number is called the *multiplicity* of S in G , and written as $\nu_G(S)$. For any undirected graph G , $\mathcal{V}_G, \mathcal{S}_G$ and $\{\nu_G(S)\}_{S \in \mathcal{S}_G}$ are uniquely defined [11].

In graphical models, the class of models induced by *decomposable graphs* are well studied, because it has many good properties. There are several equivalent definitions of decomposable graphs. One of them is based on the decomposability of graphs. For an undirected graph G and a triplet $\langle A, B | C \rangle$ with $N = A \cup B \cup C$, we say that $\langle A, B | C \rangle$ *decomposes G into the subgraphs G_{AC} and G_{BC}* if C is a clique and separates A and B . The decomposition is proper if $A, B \neq \emptyset$. An undirected graph G is decomposable if it is complete or there exists $\langle A, B | C \rangle$ which properly decomposes G into decomposable subgraphs G_{AC} and G_{BC} . Decomposable graphs are characterized in the terms of mp-subgraphs by Leimer [11]. An undirected graph G is decomposable if and only if all mp-components of G are cliques. Furthermore, for every undirected graph G with mp-components $V_1, \dots, V_m \in \mathcal{V}_G$, there exists a decomposable graph G' such that V_1, \dots, V_m are maximal cliques of G' . The graph G' is obtained by adding edges in such a way that V_1, \dots, V_m are cliques.

Another equivalent definition is a *chordal graph*, or alternatively *triangulated graph*. An undirected graph is chordal if every cycle of length more than or equal to four has a chord. An undirected graph is chordal if and only if it is decomposable [10].

2.3 Conditional independence models induced by graphs

Here we summarize known facts on conditional independence models induced by graphs.

For directed acyclic graphs, there are two equivalent separation criteria *d-separation* [13, 21] and *moralization* [9]. However we omit their details because we do not need them in this paper. For a triplet $\langle A, B | C \rangle \in \mathcal{T}(N)$, we write $A \perp\!\!\!\perp B | C [G]$ if A and B are separated given C by these criteria. Every directed acyclic graph G induces the formal independence model

$$\mathcal{M}_G = \{\langle A, B | C \rangle \in \mathcal{T}(N) : A \perp\!\!\!\perp B | C [G]\}, \quad (1)$$

which we call a *DAG model*. A probability measure P over N is *Markovian* with respect to a directed acyclic graph G if $\mathcal{M}_G \subseteq \mathcal{M}_P$ and *perfectly Markovian* if the converse implication also holds.

For an undirected graph G and $\langle A, B | C \rangle \in \mathcal{T}(N)$, we have $A \perp\!\!\!\perp B | C [G]$ if A and B are separated by C in G [10, 13]. A *UG model* \mathcal{M}_G is again defined by (1). The definitions of a Markovian and a perfectly Markovian measure are analogous to the case of DAG models. It is known that a perfect Markovian discrete measure exists for every undirected graph [3]. A decomposable model is defined as an independence model induced by a decomposable graph. A decomposable model is simultaneously a UG model and a DAG model.

Finally we discuss chain graphs. The popular separation criterion for chain graphs is *moralization* [4]. For a chain graph G and a triplet $\langle A, B | C \rangle \in \mathcal{T}(N)$, let $H = G_{\text{an}(ABC)}$. A *moral graph* H^{mor} of H is the undirected graph obtained by adding an undirected edge $a - b$ to the underlying graph of H whenever there is a chain component $C' \in \mathcal{C}_H$ such that $a, b \in \text{pa}(C')$ and a and b are not adjacent in H . We define $A \perp\!\!\!\perp B | C [G]$ if $A \perp\!\!\!\perp B | C [H^{\text{mor}}]$ holds. The definitions of a *CG model*, a Markovian measure and a perfectly Markovian measure are analogous to the other graphs. It is known that a perfect Markovian discrete measure exists for every chain graph [15].

An important concept about chain graphs is equivalence for graphs [17]. We say that G and H are *equivalent* if $\mathcal{M}_G = \mathcal{M}_H$. Equivalent chain graphs are characterized by Frydenberg [4]. A *complex* in G is an induced subgraph of G of the form $c_0 \rightarrow c_1 - \dots - c_k \leftarrow c_{k+1}$, $k \geq 1$, and no other edges between c_0, c_1, \dots, c_{k+1} exist in G .

Theorem 2.4. (Frydenberg [4]) *Two chain graphs are equivalent if and only if their underlying graphs coincide and they have the same complexes.*

More important fact is that every equivalence class has one distinguished representative.

Theorem 2.5. (Frydenberg [4]) *Every equivalence class \mathcal{H} of chain graphs has the largest element $H_\infty \in \mathcal{H}$ such that $H \leq H_\infty$ for all $H \in \mathcal{H}$.*

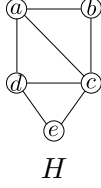


Figure 1: A decomposable graph H .

2.4 Standard insets for directed acyclic graphs and decomposable graphs

Let G be a directed acyclic graph. A *standard inset* for G is defined as follows [17]:

$$u_G = \delta_N - \delta_\emptyset + \sum_{i \in N} \{\delta_{\text{pa}(i)} - \delta_{\{i\} \cup \text{pa}(i)}\}. \quad (2)$$

This standard inset is a unique representative for equivalent graphs.

Lemma 2.6. (*Studený [17]*) *Let G a directed acyclic graph. Then $u_G \in \mathcal{C}(N)$ and $\mathcal{M}_G = \mathcal{M}_{u_G}$ hold. Moreover, for a directed acyclic graph G' , $\mathcal{M}_G = \mathcal{M}_{G'}$ if and only if $u_G = u_{G'}$.*

A standard inset for a decomposable graph H is defined by the sets of maximal cliques and clique minimal vertex separators in H [17]:

$$u_H = \delta_N - \sum_{K \in \mathcal{K}_H} \delta_K + \sum_{S \in \mathcal{S}_H} \nu_H(S) \cdot \delta_S. \quad (3)$$

Example 2.7. Put $N = \{a, b, c, d, e\}$ and consider the decomposable graph H shown in Figure 1. The sets of maximal cliques and clique minimal vertex separators in H are $\mathcal{K}_H = \{abc, acd, cde\}$ and $\mathcal{S}_H = \{ac, cd\}$ (with multiplicities $\nu_H(ac) = \nu_H(cd) = 1$). Then the standard inset for H is

$$\begin{aligned} u_H &= \delta_{abcde} - \delta_{abc} - \delta_{acd} - \delta_{cde} + \delta_{ac} + \delta_{cd} \\ &= u_{\langle b, e \mid acd \rangle} + u_{\langle a, e \mid cd \rangle} + u_{\langle b, d \mid ac \rangle}. \end{aligned}$$

For a complete graph, its standard inset is the zero inset.

Since decomposable models can be viewed as DAG models, their insets (2) and (3) lead to the same inset.

Lemma 2.8. (*Studený [17]*) *For every decomposable graph H , there exists a directed acyclic graph G such that $\mathcal{M}_G = \mathcal{M}_H$ and $u_G = u_H$.*

This implies that for a decomposable graph H , we have $u_H \in \mathcal{C}(N)$ and $\mathcal{M}_H = \mathcal{M}_{u_H}$ from Lemma 2.6.

As discussed in Section 1, these imsets for directed acyclic and decomposable graphs are not the only combinatorial ones representing their graphical models. However they are the simplest, “standard” representations [2]. A standard imset gives a simpler criterion of testing a conditional independence statement than other imsets.

Lemma 2.9. (*Bouckaert et al. [2]*) *For a directed acyclic (resp. decomposable) graph G , $\langle A, B | C \rangle \in \mathcal{M}_G$ if and only if $u_G - u_{\langle A, B | C \rangle} \in \mathcal{C}(N)$, which is also equivalent to $u_G - u_{\langle A, B | C \rangle} \in \mathcal{S}(N)$, where u_G is the standard imset in (2) or (3).*

3 Standard imsets for general undirected graphs

In this section we derive imsets for general undirected graphs. Our construction is based on a concept of a triangulation.

3.1 General undirected graphical models

For generalizing the result of decomposable graphs to general undirected graphs, consider constructing a decomposable graph from a given undirected graph by adding edges. The resulting graph is called a *triangulation* of the input graph [6]. A triangulation G' of G is minimal if there is no triangulation G'' of G such that $E(G'') \subset E(G')$. In general, there are many minimal triangulations of a graph. As for separations of an input graph and a minimal triangulation, the following lemma holds.

Lemma 3.1. *For every undirected graph H and a triplet $\langle A, B | C \rangle \in \mathcal{M}_H$, there exists a minimal triangulation H' of H such that $\langle A, B | C \rangle \in \mathcal{M}_{H'}$.*

Proof. It suffices to show the existence of a triangulation H' such that $\langle A, B | C \rangle \in \mathcal{M}_{H'}$. In fact, if H' is not minimal, we can obtain a minimal triangulation by removing edges from H' , because removing edges does not destroy the relation $A \perp\!\!\!\perp B | C$.

We construct a desired triangulation as follows (see Figure 2). Let $N = A' \cup B' \cup C' \cup D'$ be a partition of the vertex set such that

$$\begin{aligned} A' &= \{i \in N : i \text{ is connected with } A \text{ in } H_{N \setminus C}\}, \\ B' &= \{i \in N : i \text{ is connected with } B \text{ in } H_{N \setminus C}\}, \\ C' &= C \quad \text{and} \quad D' = N \setminus A'B'C'. \end{aligned}$$

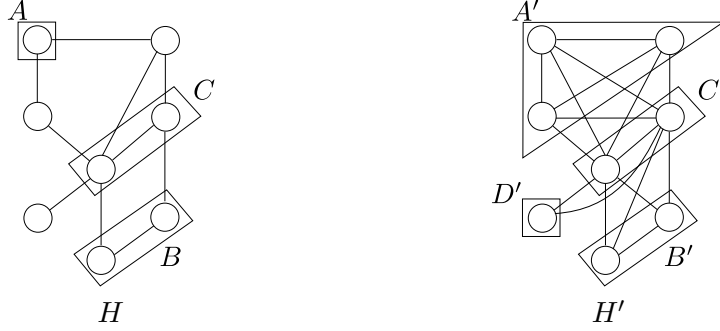


Figure 2: A construction of H' from an undirected graph H in Lemma 3.1.

Construct the graph H' by adding edges so that $H'_{A'C'}$, $H'_{B'C'}$ and $H'_{C'D'}$ are cliques. This H' is clearly decomposable, and hence, a triangulation of H . From the construction, A' and B' are not connected to each other in $H'_{N \setminus C'}$. Thus $A \subseteq A'$ and $B \subseteq B'$ are not connected to each other in $H'_{N \setminus C'}$, which means $\langle A, B \mid C \rangle \in \mathcal{M}_{H'}$. \square

For a general undirected graph, we can obtain an inset representing this UG model by using all minimal triangulations. The following theorem is the first main result of this paper.

Theorem 3.2. *Let H be an undirected graph and let $\mathfrak{T}(H)$ denote the set of all minimal triangulations of H . Put*

$$v_H = \sum_{H' \in \mathfrak{T}(H)} u_{H'}, \quad (4)$$

where $u_{H'}$ for $H' \in \mathfrak{T}(H)$ are defined by (3). Then $v_H \in \mathcal{C}(N)$ and $\mathcal{M}_H = \mathcal{M}_{v_H}$.

Proof. Since the class of combinatorial insets is closed under the addition, it is evident that the inset v_H is combinatorial.

For every undirected graph H , there exists a discrete probability measure P with $\mathcal{M}_P = \mathcal{M}_H$ [3]. Moreover, Theorem 2.3 implies that there is a structural inset $w \in \mathcal{S}(N)$ such that $\mathcal{M}_P = \mathcal{M}_w$. Then, for $H' \in \mathfrak{T}(H)$, we have

$$\mathcal{M}_{u_{H'}} = \mathcal{M}_{H'} \subseteq \mathcal{M}_H = \mathcal{M}_P = \mathcal{M}_w,$$

which implies $k_{H'} \cdot w - u_{H'} \in \mathcal{S}(N)$ for some $k_{H'} \in \mathbb{N}$ from Lemma 2.2. Therefore, putting $k = \sum_{H' \in \mathfrak{T}(H)} k_{H'}$, it follows that $k \cdot w - v_H \in \mathcal{S}(N)$. That is, $\mathcal{M}_{v_H} \subseteq \mathcal{M}_w = \mathcal{M}_H$.

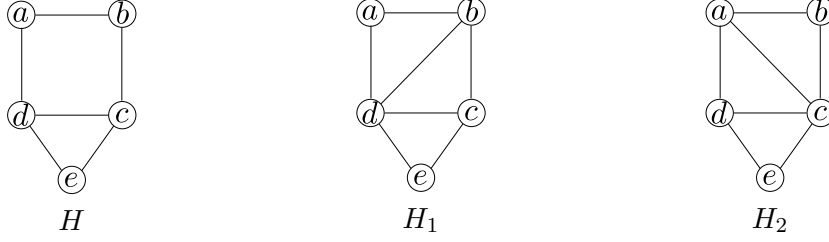


Figure 3: Non-decomposable graph H and its minimal triangulations H_1, H_2 .

Conversely, for every $\langle A, B | C \rangle \in \mathcal{M}_H$, there exists $H' \in \mathfrak{T}(H)$ such that $\langle A, B | C \rangle \in \mathcal{M}_{H'}$ from Lemma 3.1. Thus, $u_{H'} - u_{\langle A, B | C \rangle} \in \mathcal{S}(N)$ from Lemma 2.9. Hence we have

$$v_H - u_{\langle A, B | C \rangle} = \sum_{H'' \in \mathfrak{T}(H) \setminus H'} u_{H''} + (u_{H'} - u_{\langle A, B | C \rangle}) \in \mathcal{S}(N),$$

which implies $\langle A, B | C \rangle \in \mathcal{M}_{v_H}$. \square

The imset v_H in (4) is a generalization of the case of decomposable graphs, because for a decomposable graph H , the set of minimal triangulations contains H only. An example of this imset is given in the next section.

3.2 Some consideration toward a definition of standard imsets for general undirected graphs

The imset defined in the last section through all minimal triangulations has ‘extra’ additional parts as shown in the following example.

Example 3.3. Put $N = \{a, b, c, d, e\}$. Consider the graph H in Figure 3 and its minimal triangulations H_1, H_2 . Then the imset v_H in (4) is

$$\begin{aligned} v_H &= u_{H_1} + u_{H_2} \\ &= (\delta_N - \delta_{abd} - \delta_{bcd} - \delta_{cde} + \delta_{bd} + \delta_{cd}) \\ &\quad + (\delta_N - \delta_{abc} - \delta_{acd} - \delta_{cde} + \delta_{ac} + \delta_{cd}) \\ &= 2 \cdot u_{\langle ab, e | cd \rangle} + u_{\langle a, c | bd \rangle} + u_{\langle b, d | ac \rangle}. \end{aligned}$$

It can be seen that $ab \perp\!\!\!\perp e | cd$ holds in both H_1 and H_2 . This is expressed as the coefficient 2 of $u_{\langle ab, e | cd \rangle}$. Now consider an imset u_H with this coefficient 1, that is,

$$u_H = u_{\langle ab, e | cd \rangle} + u_{\langle a, c | bd \rangle} + u_{\langle b, d | ac \rangle}. \quad (5)$$

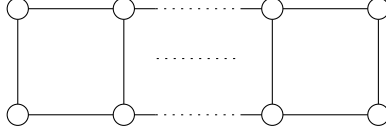


Figure 4: A graph with an exponential number of minimal triangulations.

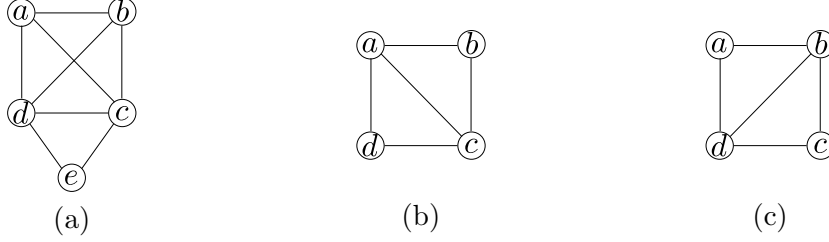


Figure 5: mp-subgraphs in the new imset.

From Lemma 3.1, $\langle A, B | C \rangle \in \mathcal{M}_H$ is equivalent to $\langle A, B | C \rangle \in \mathcal{M}_{H'}$ for some minimal triangulation H' of H . Hence, for example, letting $\langle A, B | C \rangle \in \mathcal{M}_{H_1}$, we have

$$\begin{aligned}
 u_{H_1} - u_{\langle A, B | C \rangle} &\in \mathcal{S}(N) \\
 \implies u_{H_1} + u_{\langle b, d | ac \rangle} - u_{\langle A, B | C \rangle} &\in \mathcal{S}(N) \\
 \iff u_H - u_{\langle A, B | C \rangle} &\in \mathcal{S}(N) \\
 \implies \langle A, B | C \rangle &\in \mathcal{M}_{u_H}.
 \end{aligned}$$

Since the same result holds for $\langle A, B | C \rangle \in \mathcal{M}_{H_2}$, we have $\mathcal{M}_{v_H} = \mathcal{M}_H \subseteq \mathcal{M}_{u_H}$. Also, since $v_H - u_H = u_{\langle ab, e | cd \rangle} \in \mathcal{S}(N)$, we have $\mathcal{M}_{v_H} \supseteq \mathcal{M}_{u_H}$ from Lemma 2.2. Thus $\mathcal{M}_{v_H} = \mathcal{M}_{u_H} = \mathcal{M}_H$.

Note that a graph such as the one in Figure 4 has an exponential number of minimal triangulations, which makes infeasible to calculate v_H in (4) actually.

The above examples suggest that it suffices to use only minimal triangulations of each mp-subgraph and not of the whole of the graph. In particular, the new imset (5) in Example 3.3 seems to be defined as follows: First, consider the graph obtained by adding edges to the input graph in such a way that all mp-subgraphs are complete (Figure 5-(a)), and consider its standard imset ($u_{\langle ab, e | cd \rangle}$). Next, for each mp-subgraph which is not complete, consider their minimal triangulations (Figure 5-(b), (c)) and their standard imsets ($u_{\langle a, c | bd \rangle}, u_{\langle b, d | ac \rangle}$). We show in the following sections that this idea is correct.

3.3 Minimal triangulations and mp-subgraphs

We show in this section that all minimal triangulations for an undirected graph are obtained by computing minimal triangulations for each mp-subgraph.

The following fact gives the way of adding edges to obtain a minimal triangulation:

Lemma 3.4. (*Ohtsuki et al. [12]*) *A triangulation H' of an undirected graph H is minimal if and only if for each $u - v$ added by this triangulation, no (u, v) -separators of H is a clique in H' .*

From this lemma, we have the following result about the relation between mp-subgraphs and minimal triangulations of a graph.

Lemma 3.5. *For an undirected graph H , a graph H' obtained by a minimal triangulation of each mp-subgraph is a minimal triangulation of H . Conversely, all minimal triangulations of H are obtained in this way.*

Proof. Let w be a cycle $a_1, \dots, a_n, a_{n+1} = a_1, n \geq 4$, of length more than or equal to 4 in H' . Firstly, consider the case that w contains vertices of two different mp-subgraphs of H . Then w contains two vertices of some clique minimal vertex separator of H (hence in H') which are not consecutive in w . Thus the cycle w has a chord. We next consider the case $\{a_1, \dots, a_n\} \subseteq V$ for some mp-component $V \in \mathcal{V}_H$ in H . Since H'_V is decomposable, the cycle w also has a chord in H' . Therefore H' is decomposable. Moreover, from the definition of a minimal triangulation, removing one edge from the resulting graph makes it non-decomposable. Thus the first statement is proved.

The converse is shown by contradiction. For an edge $u - v$ added by a triangulation H' of H , assume that u and v belong to different mp-subgraphs of H . Then u and v are separated by some clique minimal vertex separator of H from the definition of mp-subgraphs. Since this separator is also a clique in H' , H' is not a minimal triangulation from Lemma 3.4. \square

3.4 Definition and properties of standard imsets for undirected graphs

We define a standard imset for an undirected graph using Lemma 3.5.

Definition 3.6. For an undirected graph H , a standard imset u_H for H is defined as

$$u_H = \delta_N - \sum_{V \in \mathcal{V}_H} \delta_V + \sum_{S \in \mathcal{S}_H} \nu_H(S) \cdot \delta_S + \sum_{V \in \mathcal{V}_H} \sum_{G \in \mathfrak{T}(H_V)} u_G, \quad (6)$$

where for each $G \in \mathfrak{T}(H_V)$, $V \in \mathcal{V}_H$, u_G is the standard imset given by (3):

$$u_G = \delta_V - \sum_{K \in \mathcal{K}_G} \delta_K + \sum_{S \in \mathcal{S}_G} \nu_G(S) \cdot \delta_S.$$

Note that, if H is decomposable, the last term of u_H vanishes because all mp-components are cliques [11]. Thus this imset coincides with (3).

We show that this imset represents a UG model.

Theorem 3.7. *For an undirected graph H , define u_H as (6). Then $u_H \in \mathcal{C}(N)$ and $\mathcal{M}_H = \mathcal{M}_{u_H}$.*

Proof. The first three terms of (6) correspond to the standard imset for the decomposable graph such that all $V \in \mathcal{V}_H$ are cliques. Thus, this imset is combinatorial, and hence, $u_H \in \mathcal{C}(N)$.

Let H' be a minimal triangulation of H . Since a minimal triangulation is done in each mp-subgraph from Lemma 3.5, the following relations hold:

$$\begin{aligned} \mathcal{K}_{H'} &= \bigcup_{V \in \mathcal{V}_H} \mathcal{K}_{H'_V}, & \mathcal{S}_{H'} &= \mathcal{S}_H \cup \left(\bigcup_{V \in \mathcal{V}_H} \mathcal{S}_{H'_V} \right), \\ \mathcal{K}_{H'_{V_1}} \cap \mathcal{K}_{H'_{V_2}} &= \emptyset, & \mathcal{S}_{H'_{V_1}} \cap \mathcal{S}_{H'_{V_2}} &= \emptyset, \quad \forall V_1, V_2 \in \mathcal{V}_H, V_1 \neq V_2, \\ & & \mathcal{S}_H \cap \mathcal{S}_{H'_V} &= \emptyset, \quad \forall V \in \mathcal{V}_H, \\ & & \nu_H(S) &= \nu_{H'}(S), \quad \forall S \in \mathcal{S}_H. \end{aligned}$$

Hence a standard imset for the decomposable graph H' given by (3) is

$$\begin{aligned} u_{H'} &= \delta_N - \sum_{K \in \mathcal{K}_{H'}} \delta_K + \sum_{S \in \mathcal{S}_{H'}} \nu_{H'}(S) \cdot \delta_S \\ &= \delta_N - \sum_{V \in \mathcal{V}_H} \sum_{K \in \mathcal{K}_{H'_V}} \delta_K + \sum_{S \in \mathcal{S}_H} \nu_H(S) \cdot \delta_S + \sum_{V \in \mathcal{V}_H} \sum_{S \in \mathcal{S}_{H'_V}} \nu_{H'}(S) \cdot \delta_S \\ &= \delta_N - \sum_{V \in \mathcal{V}_H} \delta_V + \sum_{S \in \mathcal{S}_H} \nu_H(S) \cdot \delta_S \\ &\quad + \sum_{V \in \mathcal{V}_H} \left\{ \delta_V - \sum_{K \in \mathcal{K}_{H'_V}} \delta_K + \sum_{S \in \mathcal{S}_{H'_V}} \nu_{H'_V}(S) \cdot \delta_S \right\} \\ &= \delta_N - \sum_{V \in \mathcal{V}_H} \delta_V + \sum_{S \in \mathcal{S}_H} \nu_H(S) \cdot \delta_S + \sum_{V \in \mathcal{V}_H} u_{H'_V}. \end{aligned} \tag{7}$$

Let $v_H = \sum_{H' \in \mathfrak{T}(H)} u_{H'}$ given in (4). Then v_H is written as

$$v_H = \sum_{H' \in \mathfrak{T}(H)} \left\{ \delta_N - \sum_{V \in \mathcal{V}_H} \delta_V + \sum_{S \in \mathcal{S}_H} \nu_H(S) \cdot \delta_S + \sum_{V \in \mathcal{V}_H} u_{H'_V} \right\}$$

$$\begin{aligned}
&= |\mathfrak{T}(H)| \cdot \left\{ \delta_N - \sum_{V \in \mathcal{V}_H} \delta_V + \sum_{S \in \mathcal{S}_H} \nu_H(S) \cdot \delta_S \right\} + \sum_{H' \in \mathfrak{T}(H)} \sum_{V \in \mathcal{V}_H} u_{H'_V} \\
&= |\mathfrak{T}(H)| \cdot \left\{ \delta_N - \sum_{V \in \mathcal{V}_H} \delta_V + \sum_{S \in \mathcal{S}_H} \nu_H(S) \cdot \delta_S \right\} \\
&\quad + \sum_{V \in \mathcal{V}_H} \sum_{G \in \mathfrak{T}(H_V)} n_H(V, G) \cdot u_G,
\end{aligned}$$

where $n_H(V, G) = |\{H' \in \mathfrak{T}(H) : H'_V = G\}|$ for $V \in \mathcal{V}_H$ and $G \in \mathfrak{T}(H_V)$, is the number of minimal triangulations $H' \in \mathfrak{T}(H)$ such that $H'_V = G$. Note that u_H in (6) is obtained by replacing the coefficients of the right-hand side by one. Thus u_H and ν_H belong to the relative interior of the same face of $\text{cone}(\mathcal{E}(N))$. Hence we have $\mathcal{M}_{u_H} = \mathcal{M}_{\nu_H}$, which means $\mathcal{M}_{u_H} = \mathcal{M}_H$ from Theorem 3.2. \square

Example 3.8. Consider the graph H in Figure 3 again. The sets of mp-components and clique minimal vertex separators are $\mathcal{V}_H = \{abcd, cde\}$ and $\mathcal{S}_H = \{cd\}$. Since $V_2 = cde$ is a clique, the minimal triangulation of its subgraph H_{V_2} is itself. As for $V_1 = abcd$, the minimal triangulations of H_{V_1} are given in Figure 5 (b), (c). Then the standard imset for H in (6) is

$$\begin{aligned}
u_H &= \delta_{abcde} - \delta_{abcd} - \delta_{cde} + \delta_{cd} \\
&\quad + (\delta_{abcd} - \delta_{abc} - \delta_{acd} + \delta_{ac}) + (\delta_{abcd} - \delta_{abd} - \delta_{bcd} + \delta_{bd}) \\
&= u_{\langle ab, e | cd \rangle} + u_{\langle a, c | bd \rangle} + u_{\langle b, d | ac \rangle},
\end{aligned}$$

which coincides with (5).

As in the case of directed acyclic graphs and decomposable graphs, our standard imset for an undirected graph provides a simpler criterion:

Corollary 3.9. *For an undirected graph H and every triplet $\langle A, B | C \rangle \in \mathcal{T}(N)$, the followings are equivalent:*

- (i) $\langle A, B | C \rangle \in \mathcal{M}_H$,
- (ii) $u_H - u_{\langle A, B | C \rangle} \in \mathcal{C}(N)$,
- (iii) $u_H - u_{\langle A, B | C \rangle} \in \mathcal{S}(N)$.

Proof. The implication (ii) \Rightarrow (iii) \Rightarrow (i) is obvious from the definition. Thus we only need to consider the implication (i) \Rightarrow (ii). For $\langle A, B | C \rangle \in \mathcal{M}_H$, Lemma 3.1 implies that $\langle A, B | C \rangle \in \mathcal{M}_{H'}$ for some minimal triangulation H' of H . Hence $u_{H'} - u_{\langle A, B | C \rangle} \in \mathcal{C}(N)$ from Lemma 2.9. For every $V \in \mathcal{V}_H$,

some minimal triangulation G of H_V coincides with H'_V from Lemma 3.5. Thus $u_H - u_{H'} \in \mathcal{C}(N)$ from (7), which implies that

$$u_H - u_{\langle A, B | C \rangle} = (u_H - u_{H'}) + (u_{H'} - u_{\langle A, B | C \rangle}) \in \mathcal{C}(N).$$

□

Remark 3.10. In the case of directed acyclic graphs and chain graphs, some graphs may induce the same conditional independence model, and we have to consider the uniqueness of standard imsets for these graphs (cf. Lemma 2.6). However, in the case of undirected graphs, two graphs cannot have the same conditional independence model. Thus it is not necessary to consider the uniqueness.

4 Standard imsets for general chain graphs

In this section, we define a standard imset for a chain graph, which is a generalization of an undirected graph and a directed acyclic graph. Studený and Vomlel [18], and Studený, Roverato and Štěpánová [19] give standard imsets for chain graphs which are equivalent to some directed acyclic graph. Using this result, we can derive imsets for general chain graphs. Moreover we show that these imsets fully represent CG models by similar arguments as in the case of undirected graphs. In the later part of this section, we show the uniqueness of these imsets for equivalent chain graphs using the concept of feasible merging.

4.1 Generalization of a triangulation to chain graphs

Firstly, we introduce a concept which generalizes a triangulation of an undirected graph. In the case of an undirected graph, a triangulation of a graph is defined as a decomposable graph obtained by adding edges to the input graph. Since decomposable models can be interpreted as an undirected graph which is equivalent to some directed acyclic graph, we can define a triangulation of a chain graph in the same way.

Definition 4.1. A chain graph $H' = (V(H'), E(H'))$, $V(H') = V(H)$ is said to be a triangulation of a chain graph H if H' satisfies that

- (i) $a - b \in E(H')$ whenever $a - b \in E(H)$,
- (ii) $a \rightarrow b \in E(H')$ whenever $a \rightarrow b \in E(H)$, and
- (iii) H' is equivalent to some directed acyclic graph G , that is, $\mathcal{M}_{H'} = \mathcal{M}_G$.

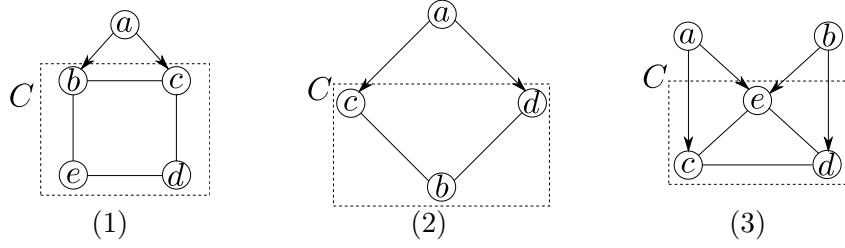


Figure 6: Examples of chain graphs violating the conditions of Lemma 4.3.

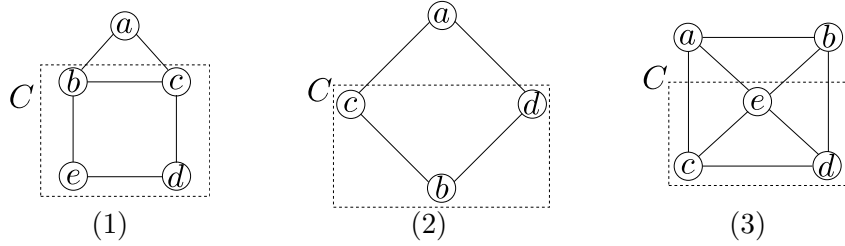


Figure 7: The closure graphs of Figure 6.

The condition (iii) has been characterized by Andersson et al. [1] in graphical terms. For a chain graph H and a chain component $C \in \mathcal{C}_H$, a closure graph for C is defined as the moral graph $\bar{H}(C) = (H_{C \cup \text{pa}(C)})^{\text{mor}}$.

Proposition 4.2. (Andersson et al. [1]) *A chain graph is equivalent to some directed acyclic graph if and only if $\bar{H}(C)$ is decomposable for every chain component $C \in \mathcal{C}_H$.*

Lemma 4.3. (cf. Remark 4.2 in [1]) *For $a \in N$ and $A \subseteq N$, let $\text{ch}_A(a) = \text{ch}(a) \cap A$ be the set of all children in H that occur in A . For any chain component $C \in \mathcal{C}_H$, the closure graph $\bar{H}(C) = (H_{C \cup \text{pa}(C)})^{\text{mor}}$ is decomposable if and only if:*

- (i) H_C is decomposable,
- (ii) for every $a \in \text{pa}(C)$, and every non-adjacent pair $c, d \in \text{ch}_C(a)$, we have $c \perp\!\!\!\perp d \mid (\text{ch}_C(a) \setminus cd) [H_C]$ (in particular $\text{ch}_C(a) \setminus cd \neq \emptyset$), and
- (iii) for every distinct pair $a, b \in \text{pa}(C)$, and every $c \in \text{ch}_C(a) \setminus \text{ch}_C(b)$, $d \in \text{ch}_C(b) \setminus \text{ch}_C(a)$, we have $c \perp\!\!\!\perp d \mid (\text{ch}_C(a) \text{ch}_C(b) \setminus cd) [H_C]$ (in particular, $\text{ch}_C(a) \text{ch}_C(b) \setminus cd \neq \emptyset$, and c, d are non-adjacent).

Example 4.4. We show in Figure 6 the examples of chain graphs which violate the conditions of Lemma 4.3. These graphs have only one chain

component C and its parent set. In Figure 6-(1), the subgraph H_C is not decomposable. In Figure 6-(2), c and d are not separated by $\text{ch}_C(a) \setminus cd = \emptyset$ because of a path $c - b - d$. In Figure 6-(3), $c \in \text{ch}_C(a) \setminus \text{ch}_C(b)$ and $d \in \text{ch}_C(b) \setminus \text{ch}_C(a)$ are adjacent. Thus c and d are not separated by $\text{ch}_C(a) \text{ch}_C(b) \setminus cd = e$. Their closure graphs are shown in Figure 7. These figures show that they are not decomposable, which implies that the graphs in Figure 6 are not equivalent to any directed acyclic graph from Proposition 4.2.

These facts show that it is sufficient to consider a minimal triangulation of $H_{C \cup \text{pa}(C)}$ for each chain component $C \in \mathcal{C}_H$ instead of the whole H . In fact, if a CG model induced by H coincides with none of DAG models, then at least one of the conditions (i), (ii) or (iii) in Lemma 4.3 is violated. When these conditions are violated, by adding edges between vertices in some C or between a vertex in C and a vertex in $\text{pa}(C)$, we can satisfy these conditions without adding any other edges. Note that the set of chain components and the parent set for each chain component in a minimal triangulation are identical with these of the input graph. Conversely, a graph obtained by minimal triangulation for each $H_{C \cup \text{pa}(C)}, C \in \mathcal{C}_H$, is clearly a minimal triangulation of H .

Remark 4.5. The above argument shows how to calculate minimal triangulations of chain graphs. Let H be a chain graph and $G = \bar{H}(C)$ be a closure graph of a chain component $C \in \mathcal{C}_H$. For a minimal triangulation G' of G , put $F = E(G') \setminus E(G)$. Then a minimal triangulation $H'_{C \cup \text{pa}(C)}$ of $H_{C \cup \text{pa}(C)}$ is obtained by adding edges (a, b) such that for $a - b \in F$

- if $a, b \in C_i$ for some i , then $a - b$, and
- if $a \in C_i, b \in C_j$ for some $i < j$, then $a \rightarrow b$,

where C_1, \dots, C_m are the ordered chain components in H . We obtain a minimal triangulation H' of H by calculating $H'_{C \cup \text{pa}(C)}$ for every $C \in \mathcal{C}_H$ in this way.

Example 4.6. Consider minimal triangulations of the graphs in Figure 6. Examples of minimal triangulations of closure graphs (Figure 7) for these graphs are shown in Figure 8. In Figure 8-(1), a minimal triangulation of the closure graph is obtained by adding the edge $c - e$. Since c and e belong to the same chain component, adding the edge $c - e$ gives a minimal triangulation (Figure 9-(1)) of the chain graph in Figure 6-(1). In Figure 8-(2), the edge $a - b$ is added. Since a and b belong to different chain components and there are directed edges from the chain component of a to

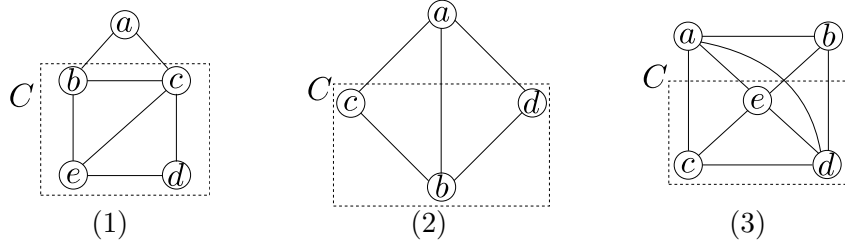


Figure 8: Examples of minimal triangulations of Figure 7.

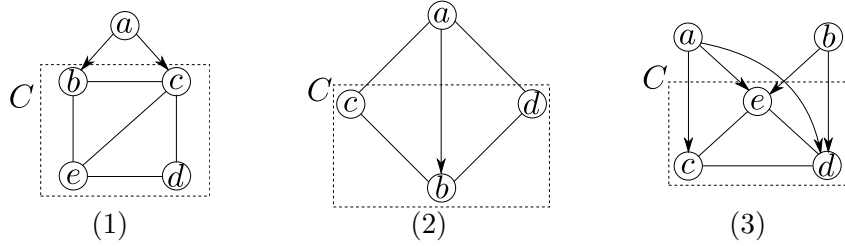


Figure 9: Examples of minimal triangulations of Figure 6.

that of b , a minimal triangulation (Figure 9-(2)) of the graph in Figure 6-(2) is obtained by adding the edge $a \rightarrow b$. As for the conditions of Lemma 4.3, $c \perp\!\!\!\perp d \mid (\text{ch}_{H,C}(a) \setminus cd) [H_C]$ holds because $\text{ch}_{H,C}(a) \setminus cd = b$. In Figure 8-(3), we add the edge $a - d$, hence, obtain the graph in Figure 9-(3) in the same way as (2). Since $\text{ch}_{H,C}(b) \setminus \text{ch}_{H,C}(a) = \emptyset$ in this graph, the condition (iii) is satisfied automatically.

The following lemma immediately holds from the above discussion.

Lemma 4.7. *For a chain graph H and a chain component $C \in \mathcal{C}_H$, assume that $\bar{H}(C)$ is decomposable. Then for every minimal triangulation H' of H , we have $H_{C \cup \text{pa}_H(C)} = H'_{C \cup \text{pa}_{H'}(C)}$.*

Corollary 4.8. *For a chain graph H and a subset $K \subseteq N$ of the vertex set, assume that $H_{\text{an}(K)}$ is equivalent to some directed acyclic graph. Then for every minimal triangulation H' of H , $H_{\text{an}_H(K)} = H'_{\text{an}_{H'}(K)}$ holds.*

Proof. Evidently, $\text{an}_H(K) = \text{an}_{H'}(K)$ holds. Also, for every chain component $C \in \mathcal{C}_{H_{\text{an}(K)}}$, the closure graph $\bar{H}(C)$ is decomposable from Proposition 4.2. Hence, from Lemma 4.7, for every minimal triangulation H' , we have $H_{C \cup \text{pa}_H(C)} = H'_{C \cup \text{pa}_{H'}(C)}$, which implies the corollary. \square

As for separations of a chain graph and its minimal triangulation, we have a similar result to Lemma 3.1 for undirected graphs.

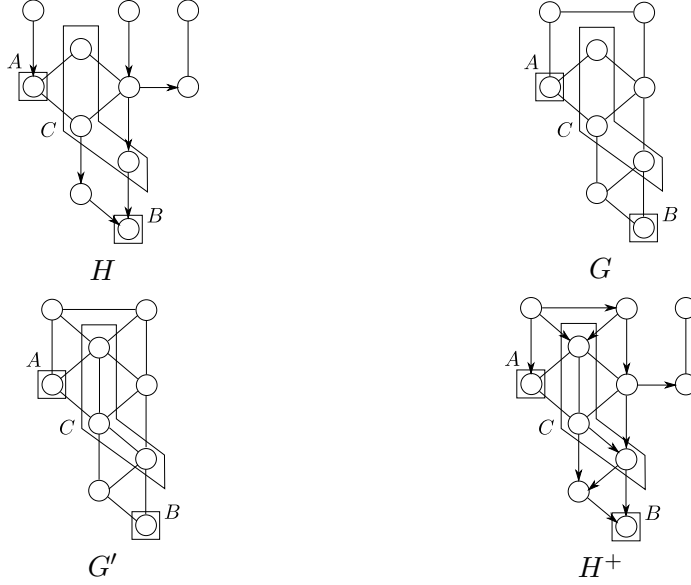


Figure 10: A construction of H^+ from a chain graph H in Lemma 4.9.

Lemma 4.9. *For every chain graph H and triplet $\langle A, B | C \rangle \in \mathcal{M}_H$, there exists a minimal triangulation H' of H such that $\langle A, B | C \rangle \in \mathcal{M}_{H'}$.*

Proof. As in the case of undirected graphs, it suffices to find a triangulation H' which satisfies $\langle A, B | C \rangle \in \mathcal{M}_{H'}$.

From the definition of the separation criterion of a chain graph, we have $\langle A, B | C \rangle \in \mathcal{M}_G$ for $G = (H_{\text{an}(ABC)})^{\text{mor}}$. Then by Lemma 3.1 there exists a minimal triangulation G' of G such that $\langle A, B | C \rangle \in \mathcal{M}_{G'}$. We construct a chain graph H^+ from G' as follows (see Figure 10):

- if $u - v \in E(G')$ and $u, v \in C_i$ for some i , then $u - v \in E(H^+)$,
- if $u - v \in E(G')$ and $u \in C_i, v \in C_j$ for some $i < j$, then $u \rightarrow v \in E(H^+)$,
- if $u - v \in E(H)$, then $u - v \in E(H^+)$,
- if $u \rightarrow v \in E(H)$, then $u \rightarrow v \in E(H^+)$,

where C_1, \dots, C_m are the ordered chain components in H . From this construction, we have $\text{an}_{H^+}(ABC) = \text{an}_H(ABC)$ and $G' = (H_{\text{an}_{H^+}(ABC)}^+)^{\text{mor}}$. Let H' be a minimal triangulation of H^+ (which may be H^+ itself). Then, since $(H_{\text{an}_{H^+}(ABC)}^+)^{\text{mor}}$ is decomposable, $H_{\text{an}_{H^+}(ABC)}^+ = H'_{\text{an}_{H'}(ABC)}$ from

Corollary 4.8. Therefore, we have $G' = (H_{\text{an}_{H^+}(ABC)}^+)^{\text{mor}} = (H'_{\text{an}_{H'}(ABC)})^{\text{mor}}$, which implies that $\langle A, B | C \rangle \in \mathcal{M}_{H'}$. \square

4.2 Definition and properties of standard imsets for chain graphs

In this section, we define a standard imset for a chain graph and show that it fully represents the CG model induced by this graph.

When a chain graph H is equivalent to some directed acyclic graph, its standard imset is defined as follows [18, 19] :

$$u_H = \delta_N - \delta_\emptyset + \sum_{C \in \mathcal{C}_H} \left\{ \delta_{\text{pa}_H(C)} - \sum_{K \in \mathcal{K}_{\bar{H}(C)}} \delta_K + \sum_{S \in \mathcal{S}_{\bar{H}(C)}} \nu_{\bar{H}(C)}(S) \cdot \delta_S \right\}. \quad (8)$$

This definition is a generalization of that of a directed acyclic graph (2) and a decomposable graph (3). Moreover, we have the following lemma about this imset:

Proposition 4.10. *(Studený et al. [19]) Assume that two chain graphs H_1, H_2 are equivalent to some directed acyclic graph. Then $\mathcal{M}_{H_1} = \mathcal{M}_{H_2}$ if and only if $u_{H_1} = u_{H_2}$.*

Therefore, for a chain graph H which is equivalent to some directed acyclic graph, we have $u_H \in \mathcal{C}(N)$ and $\mathcal{M}_H = \mathcal{M}_{u_H}$ from Lemma 2.6. Furthermore we have the following corollary from Lemma 2.9:

Corollary 4.11. *Suppose that a chain graph H is equivalent to some directed acyclic graph and let u_H be given in (8). For a triplet $\langle A, B | C \rangle \in \mathcal{T}(N)$, the followings are equivalent:*

- (i) $\langle A, B | C \rangle \in \mathcal{M}_H$,
- (ii) $u_H - u_{\langle A, B | C \rangle} \in \mathcal{C}(N)$,
- (iii) $u_H - u_{\langle A, B | C \rangle} \in \mathcal{S}(N)$.

Note that every closure graph $\bar{H}(C)$, $C \in \mathcal{C}_H$, is decomposable from Proposition 4.2. Thus (8) is also written as

$$u_H = \delta_N - \delta_\emptyset + \sum_{C \in \mathcal{C}_H} \left\{ \delta_{\text{pa}(C)} - \delta_{C \text{ pa}(C)} + u_{\bar{H}(C)} \right\},$$

where $u_{\bar{H}(C)}$ is the standard imset (2) for the decomposable graph $\bar{H}(C)$. This equation suggests that a generalization of (8) is given by replacing $u_{\bar{H}(C)}$ as (6). For $C \in \mathcal{C}_H$ and $V \subseteq C \cup \text{pa}(C)$, let $\bar{H}(C)_V$ be the induced subgraph of the closure graph $\bar{H}(C)$ by V .

Definition 4.12. A standard imset u_H for a chain graph H is defined by

$$u_H = \delta_N - \delta_\emptyset + \sum_{C \in \mathcal{C}_H} \left\{ \delta_{\text{pa}_H(C)} - \delta_{C \text{ pa}_H(C)} + u_{\bar{H}(C)} \right\}, \quad (9)$$

where $u_{\bar{H}(C)}, C \in \mathcal{C}_H$, is the standard imset for the undirected graph $\bar{H}(C)$ given by (6):

$$\begin{aligned} u_{\bar{H}(C)} &= \delta_{C \text{ pa}_H(C)} - \sum_{V \in \mathcal{V}_{\bar{H}(C)}} \delta_V + \sum_{S \in \mathcal{S}_{\bar{H}(C)}} \nu_{\bar{H}(C)}(S) \cdot \delta_S \\ &\quad + \sum_{V \in \mathcal{V}_{\bar{H}(C)}} \sum_{G \in \mathfrak{T}(\bar{H}(C)_V)} u_G. \end{aligned}$$

This imset gives a representation of CG models. The proof is similar to the case of undirected graphs.

Theorem 4.13. For a chain graph H , let a standard imset u_H for H be defined by (9). Then $u_H \in \mathcal{C}(N)$ and $\mathcal{M}_H = \mathcal{M}_{u_H}$.

Proof. The argument in the last section implies that $\mathcal{C}_H = \mathcal{C}_{H'}$ and $\text{pa}_H(C) = \text{pa}_{H'}(C), \forall C \in \mathcal{C}_H$, for a minimal triangulation H' of H . Thus a standard imset $u_{H'}$ for H' given by (8) is

$$\begin{aligned} u_{H'} &= \delta_N - \delta_\emptyset + \sum_{C \in \mathcal{C}_{H'}} \left\{ \delta_{\text{pa}_{H'}(C)} - \delta_{C \text{ pa}_{H'}(C)} + u_{\bar{H}'(C)} \right\} \\ &= \delta_N - \delta_\emptyset + \sum_{C \in \mathcal{C}_H} \left\{ \delta_{\text{pa}_H(C)} - \delta_{C \text{ pa}_H(C)} + u_{\bar{H}'(C)} \right\}. \end{aligned}$$

As in the proof (of implication $\mathcal{M}_H \subseteq \mathcal{M}_{u_H}$) of Theorem 3.7, we have $u_{\bar{H}(C)} - u_{\bar{H}'(C)} \in \mathcal{S}(N)$ for $C \in \mathcal{C}_H$, which shows that $u_H - u_{H'} \in \mathcal{S}(N)$. Also, putting $v_H = \sum_{H' \in \mathfrak{T}(H)} u_{H'}$, we have

$$\begin{aligned} v_H &= \sum_{H' \in \mathfrak{T}(H)} \left[\delta_N - \delta_\emptyset + \sum_{C \in \mathcal{C}_H} \left\{ \delta_{\text{pa}(C)} - \delta_{C \text{ pa}(C)} + u_{\bar{H}'(C)} \right\} \right] \\ &= |\mathfrak{T}(H)| \cdot \left[\delta_N - \delta_\emptyset + \sum_{C \in \mathcal{C}_H} \left\{ \delta_{\text{pa}(C)} - \delta_{C \text{ pa}(C)} \right\} \right] + \sum_{H' \in \mathfrak{T}(H)} \sum_{C \in \mathcal{C}_H} u_{\bar{H}'(C)} \\ &= |\mathfrak{T}(H)| \cdot \left[\delta_N - \delta_\emptyset + \sum_{C \in \mathcal{C}_H} \left\{ \delta_{\text{pa}(C)} - \delta_{C \text{ pa}(C)} \right\} \right] \end{aligned}$$

$$+ \sum_{C \in \mathcal{C}_H} \sum_{G \in \mathfrak{T}(\bar{H}(C))} n_H(C, G) \cdot u_G,$$

where $n_H(C, G) = |\{H' \in \mathfrak{T}(H); \bar{H}'(C) = G\}|$ for $C \in \mathcal{C}_H$ and $G \in \mathfrak{T}(\bar{H}(C))$, is the number of minimal triangulations H' of H such that $\bar{H}'(C) = G$. Therefore, as in the proof of the case of an undirected graph, u_H and v_H belong to the relative interior of the same face of $\text{cone}(\mathcal{E}(N))$. Thus we have $\mathcal{M}_{u_H} = \mathcal{M}_{v_H}$.

For every chain graph, there exists a discrete measure P over N such that $\mathcal{M}_P = \mathcal{M}_H$ [15]. Moreover Theorem 2.3 implies that $\mathcal{M}_P = \mathcal{M}_w$ for some $w \in \mathcal{S}(N)$. Hence for every $H' \in \mathfrak{T}(H)$, we have

$$\mathcal{M}_{u_{H'}} = \mathcal{M}_{H'} \subseteq \mathcal{M}_H = \mathcal{M}_P = \mathcal{M}_w,$$

which implies that $k_{H'} \cdot w - u_{H'} \in \mathcal{S}(N)$ for some $k_{H'} \in \mathbb{N}$ from Lemma 2.2. Putting $k = \sum_{H' \in \mathfrak{T}(H)} k_{H'}$, we have $k \cdot w - v_H \in \mathcal{S}(N)$. Therefore $\mathcal{M}_{u_H} = \mathcal{M}_{v_H} \subseteq \mathcal{M}_w = \mathcal{M}_H$.

Conversely, for every $\langle A, B | C \rangle \in \mathcal{M}_H$, there exists $H' \in \mathfrak{T}(H)$ such that $\langle A, B | C \rangle \in \mathcal{M}_{H'}$ from Lemma 4.9. Thus $u_{H'} - u_{\langle A, B | C \rangle} \in \mathcal{S}(N)$ from Corollary 4.11. Hence we have

$$u_H - u_{\langle A, B | C \rangle} = (u_H - u_{H'}) + (u_{H'} - u_{\langle A, B | C \rangle}) \in \mathcal{S}(N)$$

and $\langle A, B | C \rangle \in \mathcal{M}_{u_H}$. □

As in the case of undirected graphs, we have the following corollary:

Corollary 4.14. *For a chain graph H and every triplet $\langle A, B | C \rangle \in \mathcal{T}(N)$, the followings are equivalent:*

- (i) $\langle A, B | C \rangle \in \mathcal{M}_H$,
- (ii) $u_H - u_{\langle A, B | C \rangle} \in \mathcal{C}(N)$,
- (iii) $u_H - u_{\langle A, B | C \rangle} \in \mathcal{S}(N)$.

4.3 Feasible merging

From now on, we will consider the uniqueness of the standard imsets for chain graphs in Definition 4.12.

In the case of chain graphs which are equivalent to some directed acyclic graphs, the uniqueness of their standard imsets defined by (8) is given in Proposition 4.10. Its proof is based on the concept called a feasible merging [19]. In this section we review its definition and properties.

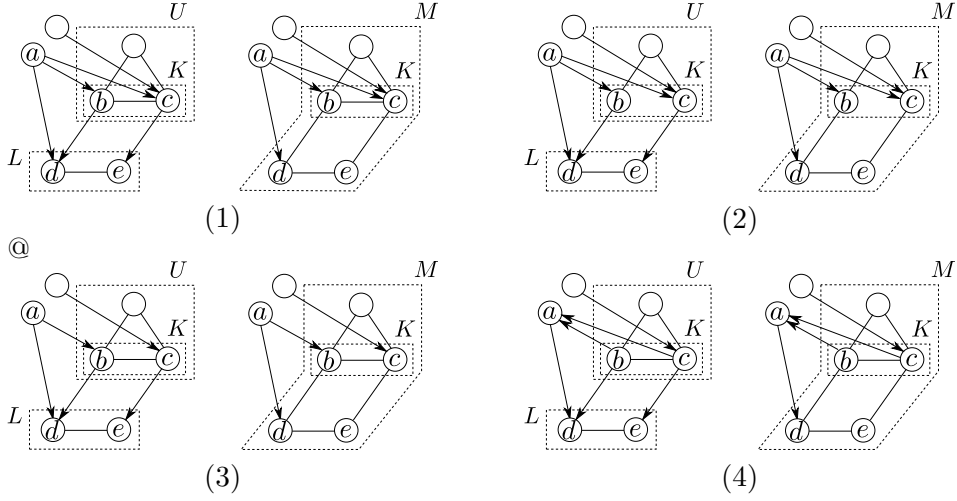


Figure 11: An examples of a feasible merging (1) and examples of infeasible merging (2), (3), (4)

Let H be a chain graph. A pair of its chain components $U, L \in \mathcal{C}_H$ is said to form a meta-arrow $U \rightrightarrows L$ if there exists a directed edge $a \rightarrow b \in E(H)$ for some $a \in U, b \in L$. *Merging of a meta-arrow* $U \rightrightarrows L$ is the operation of replacing every directed edge $a \rightarrow b \in E(H)$, $a \in U, b \in L$, with $a - b$. Merging of $U \rightrightarrows L$ is called *feasible* if the following two conditions are satisfied:

- (i) $K \equiv \text{pa}(L) \cap U$ is a clique in H , and
- (ii) $\text{pa}(L) \setminus U \subseteq \text{pa}(b)$ for any $b \in K$.

By this definition, merging is feasible if and only if $\text{pa}(L)$ is a clique in the closure graph $\bar{H}(U)$. Moreover, for the resulting graph H' and the chain component M obtained by merging of $U \rightrightarrows L$, $\text{pa}_{H'}(L)$ is a clique in $\bar{H}'(M)$.

Example 4.15. We show some examples of feasible and infeasible merging in Figure 11. The left-hand side graphs of these figures are input graphs containing $K = \{b, c\}$, $L = \{d, e\}$ and $\text{pa}(L) = \{a, b, c\}$, and the right-hand side graphs the resulting graphs obtained by merging $U \rightrightarrows L$ in the input ones. In Figure (1), K is a clique, and $\text{pa}(L) \setminus U = \{a\} = \text{pa}(b) \subset \text{pa}(c)$. Thus both conditions are satisfied, and merging is feasible. Especially, the input graph and resulting graph have the same complexes. In Figure (2), since K is not a clique, the condition (i) is not satisfied. Also in Figure

(3), the condition (ii) is violated because $\text{pa}(L) \setminus U = \{a\} \not\subseteq \text{pa}(c)$. Hence merging of $U \rightrightarrows L$ in (2) and (3) is infeasible. Note that, in Figure (2), merging of $U \rightrightarrows L$ destroys a complex $b \rightarrow d - e \leftarrow c$. Similarly, a complex $a \rightarrow d - e \leftarrow c$ vanishes in Figure (3). As in Figure (3), the condition (ii) is not satisfied in (4). In this case, the resulting graph has a directed cycle $a \rightarrow d - e - c \rightarrow a$, and hence, it is not a chain graph.

As shown in these examples, the resulting graph by feasible merging is also a chain graph and has the same complexes as the input graph. Thus, we have the following important lemma from Theorem 2.4.

Lemma 4.16. *(Studený et al. [19]) Let H be a chain graph and H' be a graph obtained by merging of $U \rightrightarrows L$ in H . Then $\mathcal{M}_H = \mathcal{M}_{H'}$ if and only if merging is feasible.*

The operation of merging can be performed without leaving the equivalence class. Especially, every larger equivalent graph is obtained by a series of feasible merging operations.

Theorem 4.17. *(Studený et al. [19]) Let G and H be chain graphs such that $\mathcal{M}_G = \mathcal{M}_H$ and $H \geq G$. Then there exists a sequence of chain graphs $G = H_1, \dots, H_r = H$, $r \geq 1$, such that H_{i+1} is obtained by the operation of feasible merging in H_i for all $i = 1, \dots, r - 1$.*

From Theorem 2.5, for proving that equivalent chain graphs have a common property, it suffices to prove that the property is shared by a pair of graphs of the class such that one is obtained by feasible merging in the other.

4.4 Uniqueness of standard imsets for chain graphs

In this section, we show that equivalent chain graphs have the same standard imset.

Theorem 4.18. *Let H_1, H_2 be chain graphs. Then $\mathcal{M}_{H_1} = \mathcal{M}_{H_2}$ if and only if $u_{H_1} = u_{H_2}$.*

To prove this theorem, the following fact is useful:

Lemma 4.19. *(cf. the proof of Theorem 20 in [19]) For a chain graph H which is equivalent to some directed acyclic graph, let H' be a graph obtained from H by feasible merging of a meta-arrow $U \rightrightarrows L$, and let M denote the merged chain component. Then $K \subseteq N$ is a maximal clique of $\bar{H}'(M)$ if and only if it is either a maximal clique of $\bar{H}(L)$ or a maximal clique of $\bar{H}(U)$ different from $\text{pa}_H(L)$.*

In a chain graph H which is equivalent to some directed acyclic graph, every mp-subgraph of $\bar{H}(C)$, $C \in \mathcal{C}_H$, is complete, because a closure graph $\bar{H}(C)$ is decomposable. As mentioned in Section 2.2, the graph obtained by adding edges to an undirected graph such that its all mp-components are cliques is decomposable. Hence, let \tilde{H} be the graph obtained by adding edges to H so that all mp-components $V \in \mathcal{V}_{\bar{H}(C)}$ for every chain component $C \in \mathcal{C}_H$ are cliques. Then \tilde{H} is a chain graph which is equivalent to some directed acyclic graph and we have the following result adapting the above lemma to \tilde{H} :

Lemma 4.20. *For a chain graph H , define H' and M as in Lemma 4.19. Then $K \subseteq N$ is an mp-component of $\bar{H}'(M)$ if and only if it is either an mp-component of $\bar{H}(L)$ or an mp-component of $\bar{H}(U)$ different from $\text{pa}_H(L)$.*

We now prove Theorem 4.18 using this result.

Proof of Theorem 4.18. Let $H = H_1$. Then the standard inset for H given by (9) is

$$u_H = \delta_N - \delta_\emptyset + \sum_{C \in \mathcal{C}_H} \left\{ \delta_{\text{pa}(C)} - \sum_{V \in \mathcal{V}_{\bar{H}(C)}} \delta_V + \sum_{S \in \mathcal{S}_{\bar{H}(C)}} \nu_{\bar{H}(C)}(S) \cdot \delta_S \right\} + \sum_{C \in \mathcal{C}_H} \sum_{V \in \mathcal{V}_{\bar{H}(C)}} \sum_{G \in \mathfrak{I}(\bar{H}(C)_V)} u_G.$$

In this equation, the first three terms are the standard inset for a graph \tilde{H} which is mentioned above. From Proposition 4.2, \tilde{H} is equivalent to some directed acyclic graph. For the graph H' , we define \tilde{H}' in the same way. Then we have $u_{\tilde{H}} = u_{\tilde{H}'}$ from Proposition 4.10. Moreover, since the closure graphs for every chain component C except for U, L, M are the same in H and H' , $u_H = u_{H'}$ is reduced to

$$\sum_{V \in \mathcal{V}_{\bar{H}(L)}} \sum_{G \in \mathfrak{I}(\bar{H}(L)_V)} u_G + \sum_{V \in \mathcal{V}_{\bar{H}(U)}} \sum_{G \in \mathfrak{I}(\bar{H}(U)_V)} u_G = \sum_{V \in \mathcal{V}_{\bar{H}'(M)}} \sum_{G \in \mathfrak{I}(\bar{H}'(M)_V)} u_G. \quad (10)$$

From Lemma 4.20, mp-components in $\bar{H}'(M)$ except for $\text{pa}_H(L)$ are identical with that of either $\bar{H}(L)$ or $\bar{H}(U)$. Also, since $\text{pa}(L)$ is a clique, letting $V = \text{pa}(L)$, we have

$$\sum_{G \in \mathfrak{I}(\bar{H}(U)_V)} u_G = \sum_{G \in \mathfrak{I}(\bar{H}(L)_V)} u_G = \sum_{G \in \mathfrak{I}(\bar{H}'(M)_V)} u_G = 0.$$

Therefore (10) holds whether $\text{pa}(L)$ is an mp-component of $\bar{H}(U)$ or $\bar{H}'(M)$ or not.

Let H_∞ be the largest chain graph in the equivalence class containing H_1, H_2 . Then we have $u_{H_1} = u_{H_\infty}$ from Theorem 4.17. Also we have $u_{H_2} = u_{H_\infty}$, which implies Theorem 4.18. \square

5 Concluding remarks

In this paper we defined standard imsets for undirected graphical models and chain graphical models. The crucial concept to derive them was a minimal triangulation. For an undirected graph, its imset was defined through all minimal triangulations of the graph. Moreover, we gave a more brief form of a standard imset using the structure of mp-subgraphs. For a chain graph, we generalized a triangulation of undirected graph. Then a standard imset for a chain graph was derived through an analogous argument as the undirected case. We also showed the uniqueness of standard imsets for equivalent chain graphs.

For directed acyclic graphs and decomposable graphs, the number of non-zero elements of their standard imsets is linear in $|N|$, while (6) and (9) may have exponential number of non-zero elements. Especially, for a prime undirected graph, imsets defined by (4) coincide with (6). Thus there is a question whether we can find an imset with smaller numbers of non-zero elements.

This is related to the degree of combinatorial imsets. The degree of a combinatorial imset is defined as the sum of positive coefficients when it is written as a non-negative integer combination of elementary imset [17]. An imset with the smallest degree is considered as a basic representative of an equivalence class in Section 7.3 in [17]. In fact, a standard imset for a directed acyclic graph has the smallest degree. Our definition of a standard imset has the smallest degree for some graphs. One of such examples is a 4-cycle graph. It is easy to see that the smallest degree in the equivalence class is 2, and (6) achieves this bound. Although, for other cycle graphs, (6) does not achieve the smallest degree, it may be possible to derive an imset with the smallest degree through our definition.

References

- [1] S. A. Andersson, D. Madigan, and M. D. Perlman: On the Markov equivalence of chain graphs, undirected graphs, and acyclic digraphs.

- Scandinavian Journal of Statistics*, **24**, no.1, pp. 81–102, 1997.
- [2] R. R. Bouckaert, R. Hemmecke, S. Linder and M. Studený: Efficient algorithms for conditional independence inference. *Journal of Machine Learning Research*, **11**, pp. 3453–3479, 2010.
 - [3] D. Geiger, and J. Pearl: Logical and algorithmic properties of conditional independence and graphical models. *Annals of Statistics*, **21**, no.4, pp. 2001–2021, 1993.
 - [4] M. Frydenberg: The chain graph Markov property. *Scandinavian Journal of Statistics*, **17**, no.4, pp. 333–353, 1990.
 - [5] H. Hara, and A. Takemura: A localization approach to improve iterative proportional scaling in Gaussian graphical models. *Communications in Statistics Theory and Methods*, **39**, no.8, pp. 1643–1654, 2010.
 - [6] P. Heggernes: Minimal triangulations of graphs: A survey. *Discrete Mathematics*, **306**, pp. 297–317, 2006.
 - [7] R. Hemmecke, J. Morton, A. Shiu, B. Sturmfels, and O. Wienand: Three counter-examples on semi-graphoids. *Combinatorics, Probability and Computing*, **17**, pp. 239–257, 2008.
 - [8] R. Hemmecke, S. Linder, and M. Studený: Learning restricted Bayesian network structures, preprint(2010). arXiv:1011.6664v1.
 - [9] S. L. Lauritzen, A. P. Dawid, B. N. Larsen, and H. G. Leimer: Independence properties of directed Markov fields. *Networks*, **20**, no.5, pp. 491–505, 1990.
 - [10] S. L. Lauritzen: *Graphical Models*. Oxford University Press, Oxford, 1996.
 - [11] H. G. Leimer: Optimal decomposition by clique separators. *Discrete Mathematics*, **113**, pp. 99–123, 1993.
 - [12] T. Ohtsuki, L. K. Cheung, and T. Fujisawa: Minimal triangulation of a graph and optimal pivoting ordering in a sparse matrix. *Journal of Mathematical Analysis and Applications*, **54**, pp. 622–633, 1976.
 - [13] J. Pearl: *Probabilistic Reasoning in Intelligent Systems - Networks of Plausible Inference*. Morgan Kaufmann, San Francisco, 1988.

- [14] M. Studený: Description of structures of stochastic conditional independence by means of faces and imsets (a series of three papers). *International Journal of General Systems*, **23**, no.2–4, pp. 123–137, 201–219, 323–341, 1994/1995.
- [15] M. Studený, and R. R. Bouckaert: On chain graph models for description of conditional independence structures. *Annals of Statistics*, **26**, no.4, pp. 1434–1495, 1998.
- [16] M. Studený: On non-graphical description of models of conditional independence structure. In *HSSS Workshop on Stochastic Systems for Individual Behaviours*, Louvain la Neuve, Belgium, 2001.
- [17] M. Studený: *Probabilistic Conditional Independence Structures*. Springer-Verlag, London, 2005.
- [18] M. Studený, and J. Vomlel: A reconstruction algorithm for the essential graph. *International Journal of Approximate Reasoning*, **50**, pp. 385–413, 2009.
- [19] M. Studený, A. Roverato, and Š. Štěpánová: Two operations of merging and splitting components in a chain graph. *Kybernetika*, **45**, no.2, pp. 208–248, 2009.
- [20] M. Studený, J. Vomlel, and R. Hemmecke: A geometric view on learning Bayesian network structures. *International Journal of Approximate Reasoning*, **51**, pp. 573–586, 2010.
- [21] T. Verma, J. Pearl: Causal networks: Semantics and expressiveness. In *Proceedings of the 4th Uncertainty in Artificial Intelligence*, North-Holland, pp. 69–76, 1990.
- [22] J. Vomlel, and M. Studený: Graphical and algebraic representatives of conditional independence models. A chapter in *Advances in Probabilistic Graphical Models*, Springer, pp. 55–80, 2007.