# Proving the NP-completeness of optimal moral graph triangulation


## Abstract

Moral graphs were introduced in the 1980s as an intermediate step when transforming a Bayesian network to a junction tree, on which exact belief propagation can be efficiently done. The moral graph of a Bayesian network can be trivially obtained by connecting non-adjacent parents for each node in the Bayesian network and dropping the direction of each edge. Perhaps because the moralization process looks simple, there has been little attention on the properties of moral graphs and their impact in belief propagation on Bayesian networks. This paper addresses the mistaken claim that it has been previously proved that optimal moral graph triangulation with the constraints of minimum fill-in, treewidth or total states is NP-complete. The problems are in fact NP-complete, but they have not previously been proved. We now prove these.


## 1 Introduction

One way to conduct an exact inference on a given Bayesian network (BN) is to transform it to a tree-like structure, called *junction tree* (JT), and conduct inference on a junction tree instead. The process of obtaining a JT from a BN consists of moralization, triangulation and tree decomposition. *Moralization* was introduced by Lauritzen and Spiegelhalter (1988) as connecting non-adjacent parents for each node in the BN and dropping all the directions. *Tree decomposition* maps a graph $G = (V, E)$ to a tree $T$, in which each tree node $t$ is a subset $V_t$ of vertices in $V$ and satisfying the following three conditions: (1) $\cup_{t \in T} V_t = V$; (2) for each edge $e \in E$ there exists a tree node $t \in T$ s.t. $V(e) \subseteq t$; (3) if $V_{t_i} \cap V_{t_k} = I$ then $I \subseteq V_{t_j}$ for each $t_j$ that appears on the path between $t_i$ and $t_k$.

Any graph has a tree decomposition, but not all decomposed trees are JTs. A JT is a tree decomposition s.t. each tree node is a complete subgraph. To ensure every DAG can be transformed to a JT, which represents a family of distributions that contains the distribution of the given BN, it is necessary to triangulate the DAG's moral graph. Here, *triangulation* finds a set of fill-in edges, whose addition makes a graph triangulated. When the marginal distribution of individual variables is of interest, the JT algorithm sums over the other variables in a tree node. Hence, the complexity of the JT algorithm is exponential in the size of a tree node.

Generally, a DAG may have more than one way of being triangulated. Then an optimal triangulation can be defined in terms of the following three constraints:

- minimum fill-in: deciding whether a graph can be triangulated by at most $\lambda$ fill in edges [1];
- treewidth: deciding whether a graph has treewidth at most $\omega$;
- total states: deciding whether a graph can be triangulated s.t. the total number of states is at most $\delta$ when summing over all maximal cliques in the triangulated graph.

The minimum fill-in problem is appealing because the treewidth usually increases exponentially in the number of fill-in edges. The total states problem incorporates both clique size and the number of states per variable, so is also essential to the complexity of the JT algorithm. The minimum fill-in and the treewidth problems for graphs were proved to be NP-complete by Yannakakis (1981) and Arnborg et al. (1987), respectively. Each proof stated a polynomial time reduction from a

[1] Originally, the minimum fill-in problem for graphs was presented and proved by Yannakakis (1981) as an optimization problem. But it can be revised to a decision problem, for which the original proof also works.

known NP-complete problem to optimally triangulating specially constructed graphs that are not moral. These reductions are sufficient to show the NP-completeness of the minimum fill-in and treewidth problems for graphs, but the difficulty of these problems do not automatically carry over to moral graphs. These works, however, were cited in Lauritzen and Spiegelhalter (1988) (Section 6 and discussion with Augustin Kong) during the discussion of triangulating moral graphs. So it gives the impression that the minimum fill-in and treewidth problems for moral graphs were proved to be NP-complete. Based on a similar reduction, Wen (1990) presented a proof of the NP-completeness of the total states problem for moral graphs. The proof is insufficient to support his claim, for the same reason above. Since then, all three works have being inaccurately cited as proving the NP-completeness of optimally triangulating moral graphs, e.g., Kjærulff (1990); Larrañaga et al. (1997); Amir (2001); Flores and Gámez (2007); Ottosen and Vomlel (2012); Li and Ueno (2017), etc.

This paper proves that the minimum fill-in, treewidth or total states problems for moral graphs are indeed NP-complete. It applies an additional step to each polynomial transformation to ensure the built graphs are moral after revision. Section 2 introduces equivalent properties to graph morality and the necessary concepts for the proofs. Section 3 demonstrates why the original constructions cannot produce moral graphs and presents a fix to each problem.

## 2 Preliminary

Throughout this report, unless mentioned otherwise, all graphs are assumed to be simple, connected and undirected. $G = (V, E)$ is used to denote a graph, whose vertex set is $V$ and edge set is $E$. For $uv \in E$, the subtraction $G - u$ denotes the induced subgraph $G[V \setminus \{u\}]$ and $G - uv$ denotes the subgraph $(V, E \setminus \{uv\})$.

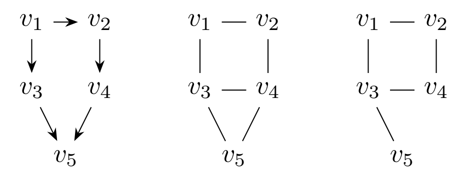
**Definition 2.1.** *The **deficiency** of a vertex $x$ in a graph $G = (V, E)$ is $D_G(x) = \{uv \notin E \mid u, v \in N_G(x)\}$.*

**Definition 2.2.** *A vertex $x$ in a graph $G$ is **simplicial** if $D_G(x) = \emptyset$.*

**Definition 2.3.** *An **ordering** of a graph $G = (V, E)$ is a bijection $\alpha : \{1, \dots, |V|\} \leftrightarrow V$.*

For convenience, let $\alpha(0) = \emptyset$. Then define the subgraph $G^i = G - \{\alpha(0), \dots, \alpha(i)\}$ for $i \in [0, |V|]$. It is called the *elimination graph* (w.r.t. $\alpha$) if for each $j \in [1, i]$ the node $\alpha(j)$ is simplicial in $G^{j-1}$.

**Definition 2.4.** *The **triangulation** of a graph $G$ w.r.t. an ordering $\alpha$ is the set of edges $H_G(\alpha) = \{D_{G^{i-1}}(\alpha(i)) \mid i \in [1, |V|]\}$.*



(a) (b) (c)

Figure 1: (a) a DAG, (b) its moral graph, (c) a non-weakly recursively simplicial graph.

The above definition implies that in the triangulated graph $F = G + H_G(\alpha)$, the node $\alpha(i)$ is simplicial in $F^{i-1}$ for each $i \in [1, |V|]$.

To distinguish them from undirected edges, the directed edge from $u$ to $v$ is denoted by $\overrightarrow{uv}$. The *skeleton* of a (partially) directed graph is the undirected graph obtained by dropping the direction of each directed edges. In a directed graph, $u$ is a *parent* of $v$, denoted by $u \in P_G(v)$, if there is a directed edge $\overrightarrow{uv}$.

**Definition 2.5.** *The **moral graph** of a directed acyclic graph $G = (V, E)$ is the skeleton of the graph $(V, E \cup F)$, where $F = \{uv \mid u, v \in P_G(x) \wedge \{\overrightarrow{uv}, \overrightarrow{vu}\} \cap E = \emptyset\}$.*

**Definition 2.6.** *A graph $G = (V, E)$ is **weakly recursively simplicial** (WRS) if there exists a simplicial vertex $x \in V$ and $E' \subseteq E(G[N(x)])$ s.t. the subgraph $G' = G - x - E'$ is weakly recursively simplicial.*

**Definition 2.7.** *A set of **excesses** of a graph $G = (V, E)$ w.r.t. an ordering $\alpha$ is a bijection $\epsilon_\alpha : \{\alpha(1), \dots, \alpha(|V|)\} \leftrightarrow \{\epsilon_\alpha(\alpha(1)), \dots, \epsilon_\alpha(\alpha(|V|))\}$, where each excess $\epsilon_\alpha(\alpha(i)) \subseteq E(G[N(\alpha(i))])$ consists of some edges between the neighbours of $v_i$.*

The composition $\kappa = (\alpha, \epsilon_\alpha)$ of an ordering of a graph $G$ and a set of excesses (w.r.t. $\alpha$) is called an *elimination kit* of $G$. Let $\kappa(0) = \emptyset$ and $\kappa(i) = \{\alpha(i), \epsilon_\alpha(\alpha(i))\}$ be the $i^{th}$ elimination kit. The concept of elimination graph can be generalized to $G^i = G - \{\kappa(0), \dots, \kappa(i)\}$ for $i \in [0, |V|]$.

**Definition 2.8.** *Let $G = (V, E)$ be a graph and $\kappa = (\alpha, \epsilon_\alpha)$ be an elimination kit of $G$. Then $\kappa$ is a **perfect elimination kit** (PEK) of $G$ if each node $x \in V$ satisfies $D_{G^{\alpha^{-1}(x)-1}}(x) = \emptyset$.*

**Theorem 2.1.** *Let $G$ be a graph. The following are equivalent:*

1. *$G$ is moral.*
2. *$G$ is weakly recursively simplicial.*
3. *$G$ has a perfect elimination kit.*

*Proof.* The proof is contained in a paper that is currently under review. $\square$

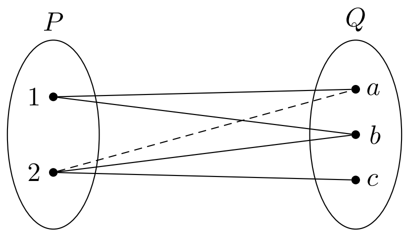



Figure 2: A bipartite non-chain graph with only the solid edges and a bipartite chain graph with all edges. The node 2 is a saturated node in the chain graph.

**Example 2.1.** *Figure 1(a) is a DAG $G$. Figure 1(b) is the moral graph of $G$. The edge $v_3v_4$ is a filled-edge by the moralization process. The moral graph has a PEK $\kappa = (\alpha, \epsilon_\alpha)$, where $\alpha = \{v_5, v_3, v_4, v_1, v_2\}$ and $\epsilon_\alpha = \{\{v_3v_4\}, \emptyset, \emptyset, \emptyset, \emptyset\}$.*

*Figure 1(c) is a non-WRS graph. $v_5$ is the only simplicial node and $E(G[N(x)]) = \emptyset$. The subgraph after removing $v_5$ and the empty set of edges is a 4-cycle that has no simplicial node.*

**Corollary 2.1.** *If a graph has no simplicial node, then it is not moral.*

**Corollary 2.2.** *If a graph is chordal, then it is moral.*

Denote a bipartite graph by $G = (P \sqcup Q, E)$, where $P \sqcup Q$ represents the disjoint union of two sets in vertices of $G$.

**Definition 2.9.** *A bipartite graph $G = (P \sqcup Q, E)$ is **chain** if there is an ordering $\alpha : \{1, \ldots, |P|\} \leftrightarrow P$ s.t. the neighbours of the vertices in $P$ form a chain $N_G(\alpha(|P|)) \subseteq \cdots \subseteq N_G(\alpha(1))$.*

The definition is also well defined for the vertices in $Q$.

**Definition 2.10.** *Let $G = (P \sqcup Q, E)$ be a bipartite graph. The **partition completion** of $G$ is a function $C(\cdot)$ that makes each $P$ and $Q$ a clique.*

In particular, the partition completion $C_P(\cdot)$ restricted to $P$ only makes $P$ a clique.

**Definition 2.11.** *A vertex in a bipartite graph is **saturated** if it is connected to every vertex in the other partition.*

**Example 2.2.** *The bipartite graph in Figure 2 with only the solid edges is not a chain graph, because the neighbour sets of $P$'s nodes do not form a chain. But the bipartite graph with all the edges is a chain graph, because $N(1) \subseteq N(2)$ w.r.t. the ordering $\alpha = \{2, 1\}$. In the chain graph, the node 2 is saturated, because it is adjacent to all nodes in $Q$.*

## 3 Optimal moral graph triangulation

Yannakakis (1981), Arnborg et al. (1987) and Wen (1990) proved a polynomial reduction from a NP-complete problem to the minimum fill-in, treewidth or total states problems for graphs, respectively. The NP-complete problems used in these proofs are the optimal linear arrangement (OLA), minimum cut linear arrangement (MCLA) and elimination degree sequence (EDS), respectively. These problems are listed below and can be found in (Garey and Johnson, 1979, page. 200-201).

**OPTIMAL LINEAR ARRANGEMENT**
INSTANCE: Graph $G = (V, E)$, positive integer $k \leq |V|$.
QUESTION: Is there an ordering $\alpha : \{1, \ldots, |V|\} \leftrightarrow V$ s.t. $c(\alpha) = \sum_{uv \in E} |\alpha^{-1}(u) - \alpha^{-1}(v)| \leq k$?

**MINIMUM CUT LINEAR ARRANGEMENT**
INSTANCE: Graph $G = (V, E)$, positive integer $k \leq |V|$.
QUESTION: Is there an ordering $\alpha : \{1, \ldots, |V|\} \leftrightarrow V$ s.t. $\forall i \in [1, |V|], |\{uv \in E \mid \alpha^{-1}(u) \leq i < \alpha^{-1}(v)\}| \leq k$?

**ELIMINATION DEGREE SEQUENCE**
INSTANCE: Graph $G = (V, E)$, sequence $< d_1, \ldots, d_{|V|} >$ of non-negative integers not exceeding $|V| - 1$.
QUESTION: Is there an ordering $\alpha : \{1, \ldots, |V|\} \leftrightarrow V$ s.t. $\forall i \in [1, |V|]$, if $\alpha^{-1}(v) = i$ then there are exactly $d_i$ vertices $u$ s.t. $\alpha^{-1}(u) > i$ and $uv \in E$?

Each of the above problems asks if there exists an ordering $\alpha$ of a given graph s.t. a certain constraint is satisfied. By fixing a node $u \in V$ at an arbitrary place in an ordering $\alpha$, each of the OLA, MCLA and EDS problems seeks for a restricted ordering from the subset $A = \{\alpha \mid \alpha(i) = u\}$ of orderings of $G$ to satisfy its constraint. It is easy to verify that these restricted problems remain NP-complete. Because if there is an $O(|V|^k)$ time algorithm to answer the question within the restricted domain $A$, it takes $|V| \times O(|V|^k)$ time to answer the original question in the entire set of orderings.

The motivation for creating a bipartite graph is the relation between chain graphs and chordal graphs stated next.

**Lemma 3.1.** *(Yannakakis, 1981). $C(G')$ is chordal if and only if $G'$ is a bipartite chain graph.*

It follows from the lemma that triangulation of the graph $C(G')$ is equivalent to making $G'$ a chain graph. To address the matter that the original transformed graphs by Yannakakis (1981), Arnborg et al. (1987) or Wen (1990) are not moral, Lemma 3.3 proves an equivalent relation between bipartite graphs and moral graphs. In addition, it gives insights of how to revise the transformations to get moral graphs, on which the optimal triangulation can be solved.

**Lemma 3.2.** *If $G' = (P \sqcup Q, E')$ is a bipartite graph, then $C_p(G')$ is triangulated.*

*Proof.* Trivial. □

**Lemma 3.3.** *Let $G' = (P \sqcup Q, E')$ be a bipartite graph. The graph $C(G')$ is moral if and only if $\exists u \in P \sqcup Q$ s.t. $\forall v \in N_{G'}(u)$, the node $v$ is saturated.*

*Proof.* Without loss of generality, consider a node $u \in Q$. The neighbour set $N_{C(G')}(u) = N_{G'}(u) \cup \{Q \setminus u\}$. The partition completion implies that both $N_{G'}(u)$ and $\{Q \setminus u\}$ are cliques in $C(G')$.

For any node in $P \sqcup Q$, if it has at least one unsaturated neighbour in the bipartite graph $G'$, then $\exists v \in N_{G'}(u)$ s.t. $vw \notin E'$ for some node $w \in \{Q \setminus u\}$. It then implies that $N_{C(G')}(u)$ is not a clique in $C(G')$, so $u$ is not a simplicial node in $C(G')$. This is true for all nodes in $C(G')$, so the graph has no simplicial node and by Corollary 2.1 it is not moral.

If all neighbours of $u$ are saturated in the bipartite graph $G'$, then the neighbour set $N_{C(G')}(u)$ is a clique. It follows that $u$ is a simplicial node in $C(G')$. The subgraph $H = G - u - \{vw \mid v, w \in Q \wedge v, w \neq u\}$ is the same as the subgraph $C_p(G - u)$, in which $G - u$ is bipartite. By Lemma 3.2, the subgraph $H$ is triangulated. Therefore the graph $C(G')$ is moral by Corollary 2.2 and Theorem 2.1. □

### 3.1 Minimum fill-in

Yannakakis (1981) presented a polynomial transformation from an instance of the OLA problem into an instance of the minimum fill-in problem for graphs. The process first takes a graph $G = (V, E)$ (Figure 3(a)) and transforms it into a bipartite graph $G' = (P \sqcup Q, E')$ (Figure 3(b)) by the following steps:

Y1. $P = \{u \mid u \in V\}$,

Y2. $Q = \{e^j \mid e \in E \wedge j \in \{1, 2\}\} \cup \{R(u) \mid u \in V\}$, where $R(u) = \{r_u^j \mid j \in \{1, \ldots, |V| - d_G(u)\}\}$,

Y3. $E' = \{ue^j \mid u \in V(e) \wedge e \in E \wedge j \in \{1, 2\}\} \cup \{uv \mid u \in P \wedge v \in R(u)\}$.

It then applies the partition completion on the bipartite graph $G'$ to obtain the graph $C(G')$, on which the minimum fill-in problem is solved. As can be seen, each edge node $e_i^j \in Q$ is incident to exactly two nodes in $P$, so $G'$

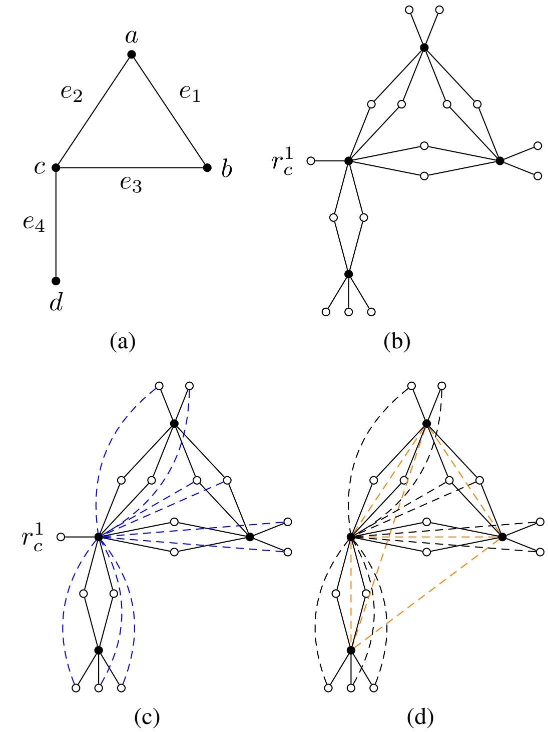



Figure 3: (a) a graph $G$; (b) the bipartite graph $G'$ transformed from $G$ by Y1-Y3; (c) the bipartite graph $\hat{G}$ transformed from $G'$ by saturating the node $c$ using L4; (d) the subgraph obtained from $C(\hat{G})$ by removing the simplicial node $r_c^1$ and its excess $\{vw \mid v, w \in Q \wedge v, w \neq r_c^1\}$.

has no saturated node, unless $G = uv$. It follows from Lemma 3.3 that $C(G')$ is not moral. The key to make $C(G')$ moral is to pick a node $u \in V$ and apply the additional step

L4. for a given node $u \in v$, let $\hat{E} = E' \cup S(u)$, where $S(u) = \{uv \notin E' \mid v \in Q\}$

that makes the corresponding node $u \in P$ saturated in $\hat{G}$ (Figure 3(c)).

**Lemma 3.4.** *Let $\hat{G} = (P \sqcup Q, \hat{E})$ be the bipartite graph constructed from a graph $G = (V, E)$ by Y1-Y3 & L4 for a given node $u \in V$. Then $C(\hat{G})$ is moral.*

*Proof.* The proof follows from Lemma 3.3. □

It is easy to see that the modified transformation can still be done in polynomial time, because the number of edges added by L4 is linear to the number of nodes in $Q$. It remains to show that a *Yes* instance of the restricted OLA problem is also a *Yes* instance of the minimum fill-in problem for moral graphs and vice versa.

To do so, the next lemma calculates the difference between the cost of a graph $G$ w.r.t. an ordering $\alpha$ and the number of fill-in edges that triangulates the corresponding moral graph $C(\hat{G})$ w.r.t. $\alpha$. And it proves that the difference is constant for any restricted ordering $\alpha$. Define the cost of an edge $e = uv \in E$ w.r.t. an ordering $\alpha$ to be $\delta(e, \alpha) = |\alpha^{-1}(u) - \alpha^{-1}(v)|$.

**Lemma 3.5.** *Given a graph $G = (V, E)$ and a positive integer $k \leq |V|$, for any node $w \in V$ the minimum cost is $k$ w.r.t. an ordering $\alpha$ of $G$ s.t. $\alpha^{-1}(w) = 1$ if and only if the corresponding moral graph $C(\hat{G})$ can be triangulated by $\lambda = k + \frac{|V|(|V|-1)(|V|-2)}{2} - 2|E| + d_G(w)$ fill in edges w.r.t. $\alpha$.*

*Proof.* Let $\hat{G}$ be the polynomial transformed graph from $G = (V, E)$ using Y1-Y3 & L4. For a given node $w \in V$, define $A = \{\alpha \mid \alpha(1) = w\}$ to be a subset of orderings of $G$. Then any ordering $\alpha \in A$ uniquely specifies a set $F_{\hat{G}}(\alpha)$ of fill-in edges to make $\hat{G}$ a chain by the following two steps:

(a) for each node $u \in Q$ calculate $\sigma(u) = \max\{i \mid u\alpha(i) \in E\}$,

(b) for any ordering $\alpha$, define $F_{\hat{G}}(\alpha) = \{u\alpha(j) \notin \hat{E} \mid u \in Q \wedge j < \sigma(u)\}$.

It is easy to see that $F_{\hat{G}}(\alpha)$ is minimal because any edge deletion from it stops the neighbour sets of $P$'s nodes in $\hat{G} + F_{\hat{G}}(\alpha)$ from forming a chain. Lemma 3.1 implies $C(\hat{G}) + F_{\hat{G}}(\alpha)$ is triangulated, so $F_{\hat{G}}(\alpha)$ is a minimal triangulation of $C(\hat{G})$. It remains to show that every ordering $\alpha \in A$ yields a set of fill in edges with cardinality

$$f_{\hat{G}}(\alpha) = c(\alpha) + \frac{|V|(|V|-1)(|V|-2)}{2} - 2|E| + d_G(w), \tag{1}$$

where $c(\alpha)$ is the total cost of $G$ w.r.t. $\alpha$. Yannakakis (1981) proved that for every ordering $\pi$ (not necessarily in $A$), the number of fill-in edges

$$f_{G'}(\pi) = c(\pi) + \frac{|V|^2(|V|-1)}{2} - 2|E|. \tag{2}$$

The following is a brief explanation of Yannakakis (1981)'s proof of equation (2). For every $v \in V$, each $x \in R(v)$ connects to $\pi^{-1}(v) - 1$ nodes in $P$, whose orderings are smaller than $\sigma(x)$. For any edge $e = uv \in E$, assume without loss of generality that $\pi^{-1}(u) < \pi^{-1}(v)$. Since each $e^j$ in $\hat{G}$ is adjacent to the two end nodes of the edge $e$, the fill in edges in $F_{G'}(\alpha)$ connect $e^j$ to $\pi^{-1}(v) - 2 = \pi^{-1}(u) + [\pi^{-1}(v) - \pi^{-1}(u)] - 2 = \pi^{-1}(u) + \delta(e, \pi) - 2$ nodes in $P$. Hence, $F_{G'}(\pi)$ contains $\pi^{-1}(v) + \pi^{-1}(u) + \delta(e, \pi) - 4$ edges incident to both $e^1$ and $e^2$. Therefore, the number of fill in edges w.r.t to $\pi$ is calculated by

$$\begin{aligned}
f_{G'}(\pi) = &\sum_{v \in V} \sum_{x \in R(v)} [\pi^{-1}(v) - 1] + \\
&\sum_{e=uv \in E} [\pi^{-1}(v) + \pi^{-1}(u) + \delta(e, \pi) - 4] \\
= &\sum_{v \in V} [|V| - d_G(v)][\pi^{-1}(v) - 1] + \\
&\sum_{v \in V} d_G(v)\pi^{-1}(v) + \sum_{e \in E} \delta(e, \pi) - 4|E| \\
= &\sum_{v \in V} |V|[\pi^{-1}(v) - 1] + \\
&\sum_{v \in V} d_G(v) + c(\pi) - 4|E| \\
= &c(\pi) + \frac{|V|^2(|V|-1)}{2} - 2|E|,
\end{aligned}$$

where by definition $\sum_{e \in E} \delta(e, \pi) = c(\pi)$ and the last equality obtains because $\sum_{v \in V} d_G(v) = 2|E|$ and $\sum_{v \in V}(\pi^{-1}(v) - 1) = |V|(|V|-1)/2$. Note that the reason to make two edge nodes and the residual nodes is to cancel the term $d_G(v)$ during the derivation of $f_{G'}(\pi)$, so that the difference between $f_{G'}(\pi)$ and $c(\pi)$ is constant, regardless of the corresponding ordering.

The size difference between the two sets of edges $\hat{E}$ and $E'$ is

$$\begin{aligned}
|S(w)| = &2(|E| - d_G(w)) + \sum_{u \in V}^{u \neq w} |R(u)| \\
= &2(|E| - d_G(w)) + \sum_{u \in V}^{u \neq w} (|V| - d_G(u)) \\
= &2(|E| - d_G(w)) + |V|(|V|-1) - \\
&\sum_{u \in V} d_G(u) + d_G(w) \\
= &|V|(|V|-1) - d_G(w).
\end{aligned} \tag{3}$$

Since equation (2) is true for every ordering, it certainly holds for orderings in $A$. It follows from $\alpha^{-1}(w) = 1$ that $S(w) \subseteq F_{G'}(\alpha)$. Hence, equation (1) is obtained by subtracting equation (3) from equation (2). If there exists an ordering $\alpha$ of $G$, w.r.t. which the minimum cost of $G$ is $k$, then $\alpha$ produces a set of fill-in edges that triangulates the moral graph $C(\hat{G})$ with $\lambda$ edges. Conversely, if the moral graph $C(\hat{G})$ can be triangulated w.r.t. an ordering $\alpha$ with $\lambda$ fill-in edges, $\alpha$ indicates the minimum cost of the graph $G$ is $k$. □

**Theorem 3.1.** *The minimum fill-in problem for moral graphs is NP-complete.*

*Proof.* Since any set of fill-in edges that triangulates a moral graph can be verified for whether or not it contains at most $\lambda$ edges in polynomial time, the minimum fill-in problem for moral graphs is in NP. Given a graph $G$ can be polynomially transformed to the corresponding moral graph $C(\hat{G})$, Lemma 3.5 proves the NP-hardness of the problem. $\square$

### 3.2 Treewidth

Arnborg et al. (1987) reduced the MCLA problem to the decision problem of whether or not a graph has a bounded treewidth. Below are the steps (A1 through A3) of Arnborg et al. (1987)'s polynomial transformation from a graph $G = (V, E)$ (Figure 3(a)) that is an instance of the MCLA problem to a bipartite graph $G' = (P \sqcup Q, E')$ (Figure 4(a)), w.r.t. which $C(G')$ is an instance of the treewidth problem for graphs.

A1. $P = \{u^i \mid u \in V \wedge i \in \{1, \ldots, \Delta(G) + 1\}\}$,

A2. $Q = \{e^j \mid e \in E \wedge j \in \{1, 2\}\} \cup \{R(u) \mid u \in V\}$, where $R(u) = \{r_u^j \mid j \in \{1, \ldots, \Delta(G) + 1 - d_G(u)\}\}$,

A3. $E' = \{u^i e^j \mid u^i \in P \wedge u \in V(e) \wedge e \in E \wedge j \in \{1, 2\}\} \cup \{u^i v \mid u \in P \wedge v \in R(u)\}$.

The transformation is similar to that of Yannakakis (1981) but produces multiple copies for the nodes in $G$. The bipartite graph $G'$ built via the above three steps has no saturated node (unless $G = uv$) for the same reason discussed in the preceding subsection, so $C(G')$ is not moral by Lemma 3.3. To make it moral, the following step

L*4. for a given node $u \in V$, let $\hat{E} = E' \cup \{S(u^j) \mid j \in [1, \Delta(G) + 1]\}$, where $u^j \in P$ is the corresponding node to $u$ and $S(u^j) = \{u^j v \notin E' \mid v \in Q\}$

is applied to each copy $u^j$ of a given node $u \in V$ (Figure 4(b)) to make valid of any residual node of $u$ being simplicial in $C(\hat{G})$.

**Lemma 3.6.** *Let $\hat{G} = (P \sqcup Q, \hat{E})$ be the bipartite graph constructed from a graph $G = (V, E)$ by A1-A3 & L*4 for a given node $u \in V$. Then $C(\hat{G})$ is moral.*

*Proof.* The proof follows from Lemma 3.3. $\square$

Although this additional step is applied to all copies of $u$, the polynomial time is guaranteed by the bounded number $\Delta(G)+1$ of copies of $u$. Before proceeding, it is necessary to draw the connection between an ordering w.r.t.

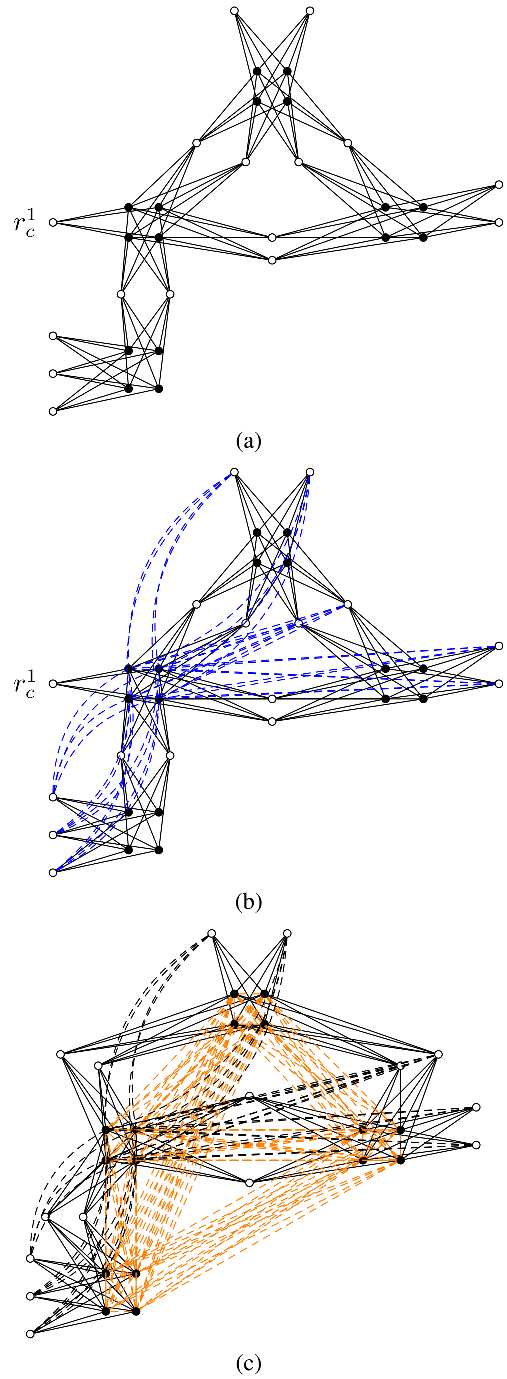



Figure 4: (a) the bipartite graph $G'$ transformed from $G$ (Figure 3(a)) by A1-A3; (b) the bipartite graph $\hat{G}$ transformed from $G'$ by saturating the nodes $\{c^1, c^2, c^3, c^4\}$ using L*4; (c) the subgraph obtained from $C(\hat{G})$ by removing the simplicial node $r_c^1$ and its excess $\{vw \mid v, w \in Q \wedge v, w \neq r_c^1\}$.

which a chain graph $G'$ is defined and a perfect elimination ordering (PEO) of the corresponding triangulated graph $C(G')$.

**Lemma 3.7.** *Let $G' = (P \sqcup Q, E')$ be a chain graph w.r.t. an ordering $\alpha$ of $P$ and $\pi_P$ be the reverse of $\alpha$. Then for any ordering $\pi_Q$ of $Q$, the elimination ordering $\{\pi_P, \pi_Q\}$ is perfect for the graph $C(G')$.*

*Proof.* The neighbour set $N_{C(G')}(\alpha(|P|)) = \{P \setminus \alpha(|P|)\} \cup N_{G'}(\alpha(|P|))$, where each of the two subsets is a clique because of the partition completion. $G'$ is a chain graph implies that for each $i \in [1, |P| - 1]$, $N_{G'}(\alpha(|P|)) \subseteq N_{G'}(\alpha(i))$. It follows that each $\alpha(i)$ is adjacent to all nodes in $N_{G'}(\alpha(|P|))$, so $\alpha(|P|)$ is simplicial in $C(G')$. By the same argument, the node $\alpha(|P|-1)$ is simplicial in the subgraph $C(G') - \alpha(|P|)$. Hence, the subgraph of $C(G')$ induced by $P$ can be eliminated recursively according to $\pi_P$. The remaining part is a complete subgraph over $Q$. Hence, any ordering of $Q$ appends to $\pi_P$ forms a PEO of $C(G')$. □

It has been shown that A1-A3, L*4 and partition completion polynomially transform an instance of the restricted MCLA problem to an instance of the treewidth problem for moral graphs. Based on this transformation, the next lemma proves that a *Yes* answer to the restricted MCLA problem is also a *Yes* answer to the treewidth problem for moral graphs and vice versa. Define the *linear cut value* of $G$ w.r.t. an ordering $\alpha$ as $\max_{1 \leq i < |V|} |\{uv \in E \mid \alpha(u) \leq i < \alpha(v)\}|$.

**Lemma 3.8.** *Given a graph $G = (V, E)$ and a positive integer $k \leq |V|$, for any node $v \in V$ the minimum linear cut value is $k$ w.r.t. an ordering $\alpha$ of $G$ s.t. $\alpha^{-1}(v) = |V|$ if and only if the treewidth of the corresponding moral graph $C(\hat{G})$ is $\omega = (\Delta(G) + 1) \times (|V| + 1) + k$.*

*Proof.* $\hat{G} = (P \sqcup Q, \hat{E})$ is the bipartite graph constructed from $G$ using A1-A3 & L*4 for a given node $v \in V$. Let $\pi_P$ be an ordering of $P$ s.t. for any node $u_i = \alpha(i) \in V$, the corresponding node $u_i^j \in P$ has order

$$\pi_P^{-1}(u_i^j) = (\Delta(G) + 1) \times i - j + 1, \qquad (4)$$

where $j \in [1, \Delta(G) + 1]$. Furthermore, let $\beta$ be the reverse order of $\pi_P$. Then steps (a) and (b) in Lemma 3.5 specify a set $F_{\hat{G}}(\beta)$ of fill-in edges w.r.t. $\beta$ to triangulate $C(\hat{G})$, because the neighbour sets of $P$'s nodes in $\hat{G} + F_{\hat{G}}(\beta)$ form the chain $N(\beta(|P|)) \subseteq \cdots \subseteq N(\beta(1))$. By Lemma 3.7, for any ordering $\pi_Q$ of $Q$, the ordering $\{\pi_P, \pi_Q\}$ is a PEO of the triangulated graph $T = C(\hat{G}) + F_{\hat{G}}(\beta)$. For each $i \in [1, (\Delta(G) + 1) \times |V|]$, the node $\pi_P(i)$ and its neighbours in the elimination graph $T^{i-1}$ form a clique $K^i$. By going through $\{\pi_P, \pi_Q\}$, we get a list of cliques that include all maximal cliques in $T$ and consequently the maximum clique. Since each $v^j$ is a saturated node in $P$, all maximal cliques correspond to nodes in $P$ only.

To calculate the size of each corresponding maximal clique, consider the node $u_i^1$ in the elimination graph w.r.t. $\pi_P$ by removing from $T$ the initial $(\Delta(G)+1) \times (i-1)$ nodes in $P$. The partition completion $C(\hat{G})$ connects $u_i^1$ to $\Delta(G)$ nodes corresponding to $u_i$ and $\Delta(G) + 1$ nodes corresponding to each of the remaining $|V| - i$ nodes in $V$. In addition, the edge set $\hat{E}$ and the fill-in edges $F_{\hat{G}}(\beta)$ connects $u_i^1$ to $\Delta(G) + 1 - d_G(u_l)$ residual nodes in each $R(u_l)$ for $\sigma(u_l) \geq \alpha^{-1}(u_i)$ and the two edge nodes for each edge $e \in E$ for $\sigma(e^j) \geq \alpha^{-1}(u_i)$. Let $e = xy \in E$ and assume without loss of generality that $\alpha^{-1}(x) < \alpha^{-1}(y)$. Then define $E_1^i = \{xy \in E \mid \alpha^{-1}(x) \leq i < \alpha^{-1}(y)\}$ and $E_2^i = \{xy \in E \mid \alpha^{-1}(y) \leq i\}$. The degree of $u_i^j$ in the corresponding elimination graph can be calculated by

$$\begin{aligned} d(u_i^j) =& \Delta(G) + [(\Delta(G) + 1) \times (|V| - i)] + \\ & \left[ (\Delta(G) + 1) \times i - \sum_{k=1}^{i} d_G(u_k) \right] + \\ & 2|E_1^i| + 2|E_2^i| \\ =& (\Delta(G) + 1) \times (|V| + 1) - 1 + |E_1^i|, \end{aligned} \qquad (5)$$

because $\sum_{k=1}^{i} d_G(u_k) = |E_1^i| + 2|E_2^i|$. Note that the reason for having $\Delta(G) + 1$ copies of each node in $P$ and two edge nodes for each edge in $G$ is to cancel the terms containing $i$ and $E_2^i$ in the final answer.

It is obvious that $\max\{E_1^i \mid i \in [1, |V|]\}$ is the linear cut value of $G$. If an ordering $\alpha$ gives the *Yes* answer to the restricted MCLA problem of a graph $G$, the treewidth of the corresponding moral graph $C(\hat{G})$ is equal to $\omega$ when triangulating it w.r.t. the ordering $\{\pi_P, \pi_Q\}$, where $\pi_P$ is generated according to $\alpha$ by equation (4). Conversely, if the treewidth of the moral graph $C(\hat{G})$ is $\omega$ w.r.t. the ordering $\{\pi_P, \pi_Q\}$, the minimum linear cut value of $G$ is $k$ w.r.t. the ordering of $V$ induced from $\pi_P$. □

**Theorem 3.2.** *The treewidth problem for moral graphs is NP-complete.*

*Proof.* Let $F$ be a set of fill-in edges to triangulate a moral graph $G$. It takes polynomial time to find the maximum clique in $G + F$ and test if it is at most $k$, so the treewidth problem is in NP. Hence, the theorem follows from Lemma 3.8 and the polynomial transformation from $G$ to $C(\hat{G})$. □

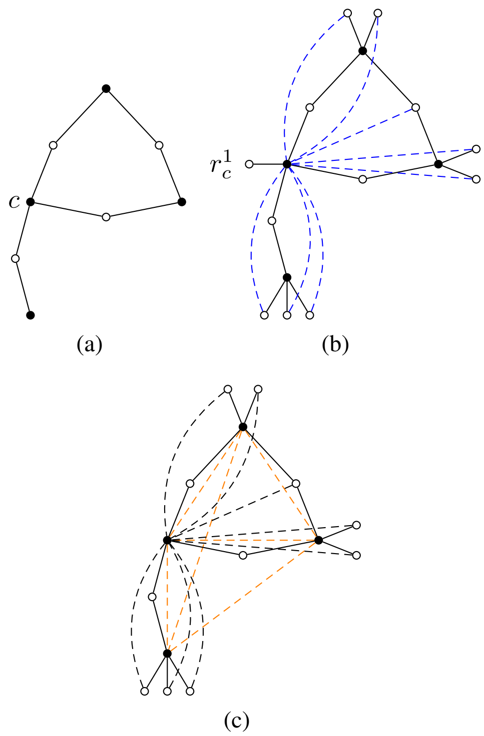



Figure 5: (a) the bipartite graph $G'$ transformed from $G$ (Figure 3(a)) by W1-W3; (b) the bipartite graph $\hat{G}$ transformed from $G$ by W1 & L2-L4 for a given node $c$; (c) the subgraph obtained from $C(\hat{G})$ by removing the simplicial node $r_c^1$ and its excess $\{vw \mid v, w \in Q \wedge v, w \neq r_c^1\}$.

### 3.3 Total states

By considering a constraint on the number of states per variable, Wen (1990) introduced the total states problem for moral graphs and presented a proof for its difficulty by polynomially reducing the EDS problem to it. His transformation is simpler than the previous two cases, using one copy of the nodes in $G$ and one edge node for each edge in $G$, without creating residual nodes (Figure 5(a)). The detailed transformation from a graph $G$ (Figure 3(a)) that is an instance of the EDS problem to a bipartite graph $G'$ is given by the following steps:

W1. $P = \{u \mid u \in V\}$,

W2. $Q = \{e^1 \mid e \in E\}$,

W3. $E' = \{ue^1 \mid u \in V(e) \wedge e \in E\}$.

Applying the partition completion on $G'$ to transform it to $C(G')$, yields an instance of the total states problem for graphs but not for moral graphs by Lemma 3.3. Therefore, Wen (1990)'s reduction does not prove that the total states problem for moral graphs is NP-complete.

To prove the EDS problem is reducible to the total states problem for moral graphs in polynomial time, W2 and W3 need to be replaced by the following two steps

L2. $Q = \{e^1 \mid e \in E\} \cup \{R(u) \mid u \in V\}$, where $R(u) = \{r_u^j \mid j \in \{1, \ldots, \Delta(G) + 1 - d_G(u)\}\}$,

L3. $E' = \{ue^1 \mid u \in V(e) \wedge e \in E\} \cup \{uv \mid u \in P \wedge v \in R(u)\}$,

to create residual nodes before applying the same step L4 (as in Section 3.1) for a given node to get the bipartite graph $\hat{G}$ (Figure 5(b)).

**Lemma 3.9.** *Let $\hat{G} = (P \sqcup Q, \hat{E})$ be the bipartite graph constructed from a graph $G = (V, E)$ by W1 & L2-L4 for a given node $u \in V$. Then $C(\hat{G})$ is moral.*

*Proof.* The proof follows from Lemma 3.3. □

For simplicity, define $N(i) = |\{u \mid u\alpha(i) \in E \wedge \alpha^{-1}(u) > i\}|$. Assume all variables considered are binary, the following lemma proves the hardness of the total states problem for moral graphs by polynomially reducing it from the restricted EDS problem.

**Lemma 3.10.** *Given a graph $G = (V, E)$ and a sequence of non-negative integers $< d_1, \ldots, d_{|V|} >$ not exceeding $|V| - 1$, for any node $w \in V$ each value in the sequence satisfies $d_i = N(i)$ w.r.t. an ordering $\alpha$ of $G$ s.t. $\alpha^{-1}(w) = |V|$ if and only if the corresponding moral graph $C(\hat{G})$ has the total number of states equal to $\delta = \sum_{i=1}^{|V|} 2^{\delta_i}$, where $\delta_i = |V| + \Delta(G) \times i + 1 + \sum_{j=1}^{i} [d_j - d_G(\alpha(j))]$.*

*Proof.* Let $\hat{G} = (P \sqcup Q, \hat{E})$ be the bipartite graph constructed from $G$ using W1 & L2-L4. Let $\beta$ be the reverse order of $\alpha$. According to steps (a) and (b) in Lemma 3.5, $\beta$ specifies a set $F_{\hat{G}}(\beta)$ of fill-in edges to make $\hat{G}$ a chain graph. Hence, $T = C(\hat{G}) + F_{\hat{G}}(\beta)$ is a triangulated graph. By Lemma 3.7, for any ordering $\alpha_Q$ of $Q$ the ordering $\{\alpha, \alpha_Q\}$ is a PEO of $T$. As stated in the proof of Lemma 3.8, the maximal cliques of $T$ only correspond to nodes in $P$, so the rest of this proof does not consider the degrees of the nodes in $Q$.

For $i \in [1, |V|]$, the neighbour set $N_{T^{i-1}}(\alpha(i)) = N_P \cup N_Q$, where $N_P = \{u \in P \mid \alpha^{-1}(u) > i\}$ and $N_Q = \cup_{j=1}^{i} (N_{T^{j-1}}(\alpha(j)) \cap Q)$ because $\hat{G} + F_{\hat{G}}(\beta)$ is a chain graph. The cardinality of $N_P$ can be easily calculated by $|V| - i + 1$. The set $N_Q$ consists of the union of the neighbours of $\alpha(j)$ restricted to $Q$ in the elimination graph $T^{j-1}$ for all $j \in [1, i]$. The restricted neighbour set $N_{T^{j-1}}(\alpha(j)) \cap Q$ contains $N(j)$ edge nodes incident to $\alpha(j)$ because there is exactly one edge node for each edge in $G$, and $\Delta(G)+1-d_G(\alpha(j))$ residual nodes of $\alpha(j)$. So the total size $|N_Q| = \sum_{j=1}^{i} N(j)+\Delta(G)+1-d_G(\alpha(j))$. To show $N_P \cup N_Q$ forms a clique, it is easy to see that each of these subsets is a clique because of the partition completion. For any $u \in N_P$, the condition $\alpha^{-1}(u) > i$ implies $\beta^{-1}(u) < i$. The definition of $\sigma(\cdot)$ (Lemma 3.5 step (a)) implies that $\sigma(v) \geq i$ for any $v \in N_Q$. It follows that $\sigma(v) > \beta^{-1}(u)$ for any $u \in N_P$ and $v \in N_Q$, so each node in $N_P$ is connected to each node in $N_Q$ by the edges in $F_{\hat{G}}(\beta)$. Therefore, the closed neighbourhood $N_{T^{i-1}}[\alpha(i)] = N_{T^{i-1}}(\alpha(i)) \cup \{\alpha(i)\}$ is a clique in $T^{i-1}$. It is in fact a maximal clique, because the set $R(\alpha(i))$ is not incident to $\alpha(i-1)$ for $i \in [2, |V|]$. The size $k_i$ of the maximal clique corresponding to $\alpha(i)$ is thus

$$\begin{aligned} k_i &= |V| - i + 1 + \sum_{j=1}^{i}[N(j) + \Delta(G) + 1 - d_G(\alpha(j))] \\ &= |V| + \Delta(G) \times i + 1 + \sum_{j=1}^{i}[N(j) - d_G(\alpha(j))]. \end{aligned} \tag{6}$$

Since all variables are binary, the total number of states summing over all $|V|$ maximal cliques is $\sum_{i=1}^{|V|} 2^{k_i}$.

If there exists an ordering $\alpha$ of $G$, w.r.t. which the EDS answer is *Yes*, substituting $N(i)$ by $d_i$ in equation (6) entails that the total states of $C(\hat{G})$ is $\delta$ w.r.t. $\alpha$. That is, $\alpha$ is a *Yes* answer to the total states problem for the corresponding moral graph $C(\hat{G})$. Conversely, if the answer to the total states problem for a moral graph $C(\hat{G})$ is *Yes* w.r.t. an ordering $\{\alpha, \alpha_Q\}$, it follows from equation (6) that $d_i = N(i)$, so $\alpha$ gives a *Yes* answer to the EDS problem for the graph $G$. □

**Theorem 3.3.** *The total states problem for moral graphs is NP-complete.*

*Proof.* Since the maximal cliques of a triangulated graph can be found in polynomial time, the total states of any triangulated graph can be verified in polynomial time to be greater than $\delta$ or not. Hence, the problem is in NP. Given the moral graph $C(\hat{G})$ can be transformed from a graph $G$ by W1 & L2-L4 in polynomial time, Lemma 3.10 proves the NP-hardness of the total states problem for moral graphs. □

## 4 Conclusion

Optimal moral graph triangulation plays an important role in determining the computational complexity of the junction tree algorithm for belief propagation on Bayesian networks. The minimum number of fill-in edges is closely related to the maximum clique size of the triangulated moral graph. The treewidth of a moral graph directly determines the efficiency of the junction tree algorithm when computing probabilities of unobserved variables by marginalizing out observed variables in the largest clique. The total number of states when summing over all maximal cliques in a triangulated moral graph takes into account the number of states per variable. The optimal moral graph triangulation with the objective of minimizing the number of fill-in edges, the maximum clique size or the total number of states has proved to be NP-complete in each case in this paper. Thus, we show that previous claims that optimal moral graph triangulation is NP-complete were in fact correct, by supplying the missing proofs.

## References


E. Amir. Efficient approximation for triangulation of minimum treewidth. In *Proceedings of the Seventeenth Annual Conference on Uncertainty in Artificial Intelligence*, pages 7–15. Morgan Kaufmann Publishers Inc., 2001.

S. Arnborg, D. G. Corneil, and A. Proskurowski. Complexity of finding embeddings in a k-tree. *SIAM Journal on Algebraic Discrete Methods*, 8(2):277–284, 1987.

M. J. Flores and J. A. Gámez. A review on distinct methods and approaches to perform triangulation for Bayesian networks. In *Advances in Probabilistic Graphical Models*, pages 127–152. Springer, 2007.

M. R. Garey and D. S. Johnson. *Computers and intractability*. W.H. Freeman, San Francisco, CA, 1979.

U. Kjærulff. Triangulation of graphs-algorithms giving small total state space. Technical report, Aalborg University, Denmark, 1990.

P. Larrañaga, C. M. H. Kuijpers, M. Poza, and R. H. Murga. Decomposing Bayesian networks: triangulation of the moral graph with genetic algorithms. *Statistics and Computing*, 7(1):19–34, 1997.

S. L. Lauritzen and D. J. Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 157–224, 1988.

C. Li and M. Ueno. An extended depth-first search algorithm for optimal triangulation of Bayesian networks. *International Journal of Approximate Reasoning*, 80: 294–312, 2017.

T. J. Ottosen and J. Vomlel. All roads lead to Rome-New search methods for the optimal triangulation problem. *International Journal of Approximate Reasoning*, 53 (9):1350–1366, 2012.

W. X. Wen. Optimal decomposition of belief networks. In *Proceedings of the Sixth Annual Conference on Uncertainty in Artificial Intelligence*, pages 209–224. Elsevier Science Inc., 1990.

M. Yannakakis. Computing the minimum fill-in is NP-complete. *SIAM Journal on Algebraic Discrete Methods*, 2(1):77–79, 1981.