Dual Formulation of the Chordal Graph Conjecture

Milan Studený STUDENY@UTIA.CAS.CZ
Department of Decision-Making Theory, the Institute of Information Theory and Automation of the Czech Academy of Sciences, Prague, 18208 Pod Vodárenskou věží 4, Czech Republic

James Cussens JAMES.CUSSENS@BRISTOL.AC.UK
Department of Computer Science, University of Bristol, UK

Václav Kratochvíl VELOREX@UTIA.CAS.CZ
Department of Decision-Making Theory, the Institute of Information Theory and Automation of the Czech Academy of Sciences, Prague, 18208 Pod Vodárenskou věží 4, Czech Republic

Abstract

The idea of an integer linear programming approach to structural learning of decomposable graphical models led to the study of the so-called *chordal graph polytope*. An open mathematical question is what is the minimal set of linear inequalities defining this polytope. Some time ago we came up with a specific conjecture that the polytope is defined by so-called *clutter* inequalities. In this theoretical paper we give a dual formulation of the conjecture. Specifically, we introduce a certain *dual polyhedron* defined by trivial equality constraints, simple monotonicity inequalities and certain inequalities assigned to incomplete chordal graphs. The main result is that the list of (all) vertices of this bounded polyhedron gives rise to the list of (all) facet-defining inequalities of the chordal graph polytope. The original conjecture is then equivalent to a statement that all vertices of the dual polyhedron are zero-one vectors. This dual formulation of the conjecture offers a more intuitive view on the problem and allows us to disprove the conjecture.

Keywords: Learning decomposable models; chordal graph polytope; clutter inequalities; dual polyhedron; chordal graph inequalities.

1. Introduction

The motivation for this theoretical paper is learning *decomposable models*, which are assigned to chordal undirected graphs. These are fundamental graphical models (Lauritzen, 1996) and one of the reasons for this is that their elegant mathematical properties are at the core of a well-known method of local computation (Cowell et al., 1999). Decomposable models can also be interpreted as special cases of Bayesian network models (Pearl, 1988) assigned to directed acyclic graphs.

Thus, most of the methods for structural learning of decomposable models follow the standard methods for learning Bayesian networks (Neapolitan, 2004). We are specifically interested in the *integer linear programming* (ILP) approach to structural learning of decomposable models. It is a special case of the score-based approach, where the goal is to maximize an additively decomposable score, like the BIC score (Schwarz, 1978) or the BDeu score (Heckerman et al., 1995) and methods of ILP are used for this purpose. The idea behind the ILP approach is to encode graphical models by vectors whose components are integers in such a way that the standard scores turn into linear functions of the vector representatives, up to a constant term. Note that the number of components of such vectors is inevitably exponential in the number of nodes of graphs. Different ways to encode Bayesian network models were compared by Cussens et al. (2017).

Our approach to learning decomposable models is based on encoding the models by vectors named *characteristic imsets*, which were originally introduced and tested in the context of learning Bayesian networks (Hemmecke et al., 2012; Studený and Haws, 2014). This mode of representation leads to a particularly elegant way of encoding chordal graphs. Nonetheless, an additional invertible linear mapping, named superset Möbius inversion, allows one to transform characteristic imsets into yet other vector representatives which have only a few non-zero components; these components correspond to cliques and separators of the graphs and are, therefore, close to junction tree representations of the graphs.

We have defined in (Studený and Cussens, 2016) the so-called *chordal graph polytope*, which is the convex hull of all characteristic imsets for chordal graphs over a fixed set N of nodes, $|N| \geq 2$. Note, however, that the polytope itself has already been introduced under slightly longer name by Lindner (2012) in her thesis, where she derived some basic observations on it.

We have also introduced in (Studený and Cussens, 2016) special linear *clutter inequalities* valid for the vectors in the polytope, which correspond to singleton-containing clutters (= classes of inclusion incomparable sets, alternatively named Sperner families or anti-chains) of subsets of N. Then in (Studený and Cussens, 2017), we have shown that each clutter inequality is a facet-defining inequality for the chordal graph polytope, which means it belongs to the minimal set of linear inequalities defining the polytope. Note in this context that we do not distinguish between an inequality and its multiple by a positive factor and that the uniqueness of the minimal defining set of linear inequalities is relative to the affine (= shifted linear) space generated by the polytope.

Finally, we have raised a *chordal graph conjecture* in (Studený and Cussens, 2017) saying that the minimal set of inequalities defining the chordal graph polytope consists of the clutter inequalities and only one additional *lower bound inequality* which requires the non-negativity of the component for N. The conjecture has been confirmed in case $|N| \leq 5$; since then we have been trying intensively to confirm or disprove the conjecture.

In this paper we give an alternative dual formulation of the conjecture. Specifically, we will introduce a certain bounded polyhedron in $[0,1]^{\mathcal{P}(N)}$, where $\mathcal{P}(N)$ denotes the power set of N, specified by some elementary monotonicity inequalities and special *inequalities assigned to incomplete chordal graphs* over N. The main result of this paper is that the vertices of this *dual polyhedron* give rise to the minimal set of inequalities characterizing the chordal graph polytope. The original chordal graph conjecture is then equivalent to the statement that all vertices of the dual polyhedron are zero-one vectors, and therefore, are indicators of singleton-containing set filters of subsets of N. Our dual formulation allowed us to disprove the conjecture at the end of the paper.

Our approach is not the only ILP approach to learning decomposable models. Other authors have used vector representatives of these models different to ours. Sesh Kumar and Bach (2013) used special codes for junction trees that correspond to chordal graphs, while Pérez et al. (2014, 2018) used zero-one vector encodings of certain special coarsenings of maximal hyper-trees. Note that they all have been interested in learning decomposable models with a given clique size limit.

There are also approaches to structural learning of decomposable models which are not based on ILP but some of them used encodings of junction trees as well. Corander et al. (2013) used a constraint satisfaction approach and expressed their search space of models in terms of logical constraints. Kangas et al. (2014, 2015) applied the idea of decomposing junction trees into subtrees and used the method of dynamic programming. Rantanen et al. (2017) have used a branch and bound method and integrated dynamic programming.

Let us note in this context that above mentioned superset Möbius inversions of characteristic imsets are naturally related to the junction tree representatives of decomposable models used either by Sesh Kumar and Bach (2013) or by Kangas et al. (2014, 2015). They are also linearly related to the so-called *standard imsets* for chordal graphs from (Studený, 2005, § 7.2.2). We emphasize those links here because the coefficients of the inequalities assigned to chordal graphs in our dual formulation of the conjecture are just components of these vector representatives.

1.1 Content of this paper

In this paper we omit the concepts related to statistical learning of graphical models because these are not needed to present our mathematical result; the reader can find them in (Studený and Cussens, 2017). We assume that the reader is familiar with elementary concepts from polyhedral geometry which can be found in standard textbooks like (Barvinok, 2002), (Chvátal, 1983) or (Wolsey, 1998).

In Section 2 we recall basic definitions and facts. In Section 3 we introduce the dual polyhedron and formulate the main result, proved in Appendix A. This allows us to disprove the conjecture, which is done in the last concluding Section 4.

2. Basic concepts

Let N be a finite set of *variables* which correspond to the nodes of our graphs and in statistical context are interpreted as random variables. To avoid the trivial case we assume $n := |N| \geq 2$. Let $\mathcal{P}(N) := \{S \,:\, S \subseteq N\}$ denote the power set of N. We call a subset $S \subseteq N$ a *singleton* if $|S| = 1$. Given a set system $\mathcal{L} \subseteq \mathcal{P}(N)$, the symbol $\bigcup \mathcal{L}$ will denote the union of sets from $\mathcal{L}$. Given a predicate $\mathbf{P}$, the symbol $\delta(\mathbf{P})$ will denote the zero-one indicator of $\mathbf{P}$, that is, $\delta(\mathbf{P}) = 1$ if $\mathbf{P}$ holds and $\delta(\mathbf{P}) = 0$ if $\mathbf{P}$ does not hold.

2.1 Chordal graphs

An undirected graph G *over* N (= a graph G having N as the set of nodes) is called *chordal* if every cycle in G of length at least 4 has a chord, that is, an edge connecting non-consecutive nodes in the cycle. A set $S \subseteq N$ is complete in G if every two distinct nodes from S are connected by an edge in G. Maximal complete sets with respect to inclusion are called the *cliques* of G. A chordal graph G over N is complete if N is a clique, otherwise it is called *incomplete*.

A well-known equivalent definition of a chordal graph is that the collection of its cliques can be ordered into a sequence $C_1, \ldots, C_m$, $m \geq 1$, satisfying the *running intersection property* (RIP):

$$\forall\, i \geq 2 \;\; \exists\, j < i \quad \text{such that} \;\; S_i := C_i \cap \left(\bigcup_{\ell < i} C_\ell \right) \subseteq C_j \,.$$

The sets $S_i = C_i \cap (\bigcup_{\ell<i} C_\ell)$, $i = 2, \ldots, m$ are the respective separators. The *multiplicity* $\nu_G(S)$ of a separator S is the number of indices $2 \leq i \leq m$ such that $S = S_i$; the separators and their multiplicities are known not to depend on the choice of the ordering satisfying RIP, see (Studený, 2005, Lemma 7.2). In the sequel, the collection of *cliques of* G will be denoted by $\mathcal{C}(G)$, the collection of its *separators* by $\mathcal{S}(G)$.

A *junction tree* for G is a hyper-tree $\mathcal{J}$ having $\mathcal{C}(G)$ as the set of hyper-nodes in $\mathcal{J}$ and satisfying the condition that, for every pair $C, K \in \mathcal{C}(G)$, the intersection $C \cap K$ is contained in every clique on the (unique) path between C and K in $\mathcal{J}$. Another equivalent definition of a chordal graph is that it is an undirected graph which has a junction tree; see (Cowell et al., 1999, Theorem 4.6). A hyper-edge between C and K in a junction tree $\mathcal{J}$ can be labeled by the intersection $C \cap K$. Another well-known fact is that labels of these hyper-edges are just the separators of G and every separator S occurs as many times as its multiplicity $\nu_G(S)$.

2.2 Characteristic imsets for chordal graphs and the chordal graph polytope

Given a chordal graph G over N, the *characteristic imset* of G is a zero-one vector c_G whose components are indexed by subsets S of N:

$$\mathsf{c}_G(S) = \begin{cases} 1 & \text{if } S \text{ is a complete set in } G, \\ 0 & \text{for remaining } S \subseteq N. \end{cases}$$

Thus, c_G is formally a vector in $\mathbb{R}^{\mathcal{P}(N)}$ whose components for the empty set and singletons have always the value 1. Nonetheless, the roles of the empty set and singletons in our later linear inequalities for characteristic imsets differ: while the component for the empty set plays no role in our inequalities, it is useful to distinguish between components for different singletons because this step allows us to identify our inequalities with certain set systems. Therefore, we will understand every characteristic imset as a vector c in the linear space $\mathbb{R}^{\mathcal{P}(N)\setminus\{\emptyset\}}$ belonging to its affine subspace $\mathsf{A} := \{\, \mathsf{c} \in \mathbb{R}^{\mathcal{P}(N)\setminus\{\emptyset\}} \,:\, \mathsf{c}(\{i\}) = 1 \text{ for any } i \in N \}$ specified by those equality constraints.

Let us introduce the *chordal graph polytope* D_N over N as follows:

$$D_N \;:=\; \mathrm{conv}\,(\,\{\mathsf{c}_G \,:\, G \text{ chordal graph over } N\}\,) \qquad \text{where conv}\,(\cdot) \text{ denotes the convex hull.}$$

Since A is the affine hull of D_N the dimension of D_N is $2^n - n - 1$, where $n = |N|$.

2.3 Clutter inequalities and the chordal graph conjecture

By a *clutter* we call any set system $\mathcal{L} \subseteq \mathcal{P}(N)$ such that the sets in $\mathcal{L}$ are inclusion incomparable: if $L, R \in \mathcal{L}$ then $L \subseteq R$ implies $L = R$. A (set) *filter* is a set system $\mathcal{F} \subseteq \mathcal{P}(N)$ closed under supersets: if $S \in \mathcal{F}$ and $S \subseteq T \subseteq N$ then $T \in \mathcal{F}$. Clutters and filters are in a 1-to-1 correspondence: any clutter $\mathcal{L}$ generates a filter $\mathcal{L}^{\uparrow} := \{\, T \subseteq N \,:\, \exists\, L \in \mathcal{L} \quad L \subseteq T \}$ and conversely, given a filter $\mathcal{F}$, the class $\mathcal{F}_{\min}$ of inclusion minimal sets in $\mathcal{F}$ is a clutter generating $\mathcal{F}$. Note that a clutter of non-empty sets contains a singleton iff the corresponding filter contains a singleton.

Definition 1 (clutter inequality) Given a clutter $\mathcal{L} \subseteq \mathcal{P}(N)$ which contains at least one singleton the corresponding *clutter inequality* for $\mathsf{c} \in \mathsf{A} \subseteq \mathbb{R}^{\mathcal{P}(N)\setminus\{\emptyset\}}$ has the form

$$1 \le \sum_{\emptyset \neq \mathcal{B} \subseteq \mathcal{L}} (-1)^{|\mathcal{B}|+1} \cdot \mathsf{c}\left(\bigcup \mathcal{B}\right) . \tag{1}$$

Recall that because we assume $\mathsf{c} \in \mathsf{A}$ one has to substitute $\mathsf{c}(L) = 1$ in (1) whenever $L \subseteq N$ is a singleton. Note also that the inequality for a clutter consisting of only one set is superfluous: then $\mathcal{L} = \{L\}$ with $L \subseteq N$, $|L| = 1$, and (1) holds with equality since $\mathsf{c} \in \mathsf{A}$.

One can re-write (1) in a standardized form:

$$1 \;\le\; \sum_{\emptyset \neq S \subseteq N} \kappa_{\mathcal{L}}(S) \cdot \mathsf{c}(S) \qquad \text{where } \kappa_{\mathcal{L}}(S) = \sum_{\emptyset \neq \mathcal{B} \subseteq \mathcal{L} \,:\, \bigcup \mathcal{B} = S} (-1)^{|\mathcal{B}|+1} \text{ for any } S \subseteq N. \tag{2}$$

It implies that the coefficients $\kappa_{\mathcal{L}}(-)$ vanish outside the class $\mathcal{U}(\mathcal{L}) := \{\bigcup\mathcal{B} \,:\, \emptyset \neq \mathcal{B} \subseteq \mathcal{L}\}$ of unions of sets from $\mathcal{L}$. As shown in (Studený and Cussens, 2017, Lemma 1), they can be computed recursively within this class as follows:

$$\kappa_{\mathcal{L}}(S) = 1 - \sum_{T\in\mathcal{U}(\mathcal{L}):\, T\subset S} \kappa_{\mathcal{L}}(T) \qquad \text{for any } S \in \mathcal{U}(\mathcal{L}).$$

In particular, $\kappa_{\mathcal{L}}(L) = 1$ for $L \in \mathcal{L}$. The example below illustrates the procedure.

Notational convention In our examples, clutters of subsets of N will be denoted by the lists of sets in the clutters separated by straight lines. The sets are encoded by lists of their elements without commas. Thus, a clutter $\mathcal{L} = \{\ \{a,b\},\ \{a,c\},\ \{d\}\ \}$ will be denoted $|d|ab|ac|$. With a little abuse of notation, we will also use this notation for a clutter to identify the corresponding filter. Thus, $\lambda_{|d|ab|ac|} \in \mathbb{R}^{\mathcal{P}(N)}$ will denote the indicator of the filter generated by the clutter $|d|ab|ac|$. Chordal graphs over N will be analogously denoted by the lists of (all) their cliques separated by colons; thus, the empty graph over $N = \{a,b,c,d\}$ will be denoted $a : b : c : d$.

Example 1 Take $N = \{a,b,c,d\}$ and a clutter $\mathcal{L}$ specified by $|d|ab|ac|$. The fact that $\kappa_{\mathcal{L}}(L) = 1$ for $L \in \mathcal{L}$ gives $\kappa_{\mathcal{L}}(\{a,b\}) = \kappa_{\mathcal{L}}(\{a,c\}) = \kappa_{\mathcal{L}}(\{d\}) = 1$. The remaining elements in $\mathcal{U}(\mathcal{L})$ are $\{a,b,c\}$, $\{a,b,d\}$, $\{a,c,d\}$, and N. The recursive formula above applied to $\{a,b,c\}$ yields

$$\kappa_{\mathcal{L}}(\{a,b,c\}) \;=\; 1 - \kappa_{\mathcal{L}}(\{a,b\}) - \kappa_{\mathcal{L}}(\{a,c\}) = 1 - 1 - 1 = -1\,.$$

Analogously, $\kappa_{\mathcal{L}}(\{a,b,d\}) = \kappa_{\mathcal{L}}(\{a,c,d\}) = -1$. Finally, N has all other sets in $\mathcal{U}(\mathcal{L})$ as proper subsets which gives

$$\begin{aligned}\kappa_{\mathcal{L}}(N) \;=\; & 1 - \kappa_{\mathcal{L}}(\{a,b,c\}) - \kappa_{\mathcal{L}}(\{a,b,d\}) - \kappa_{\mathcal{L}}(\{a,c,d\}) \\ & - \kappa_{\mathcal{L}}(\{a,b\}) - \kappa_{\mathcal{L}}(\{a,c\}) - \kappa_{\mathcal{L}}(\{d\}) = 1 + 3 - 3 = 1\,.\end{aligned}$$

Taking into consideration that $\mathsf{c}(\{d\}) = 1$ one gets from (2):

$$1 \leq \mathsf{c}(\{a,b\}) + \mathsf{c}(\{a,c\}) + 1 - \mathsf{c}(\{a,b,c\}) - \mathsf{c}(\{a,b,d\}) - \mathsf{c}(\{a,c,d\}) + \mathsf{c}(N)\,,$$

which can be re-written in the form

$$\mathsf{c}(\{a,b,c\}) + \mathsf{c}(\{a,b,d\}) + \mathsf{c}(\{a,c,d\}) \leq \mathsf{c}(\{a,b\}) + \mathsf{c}(\{a,c\}) + \mathsf{c}(N)\,.$$

Our conjecture was that all facet-defining inequalities for D_N were as follows.

Conjecture 2 (chordal graph conjecture) *For any $n = |N| \geq 2$, the least set of inequalities for $\mathsf{c} \in \mathsf{A}$ defining D_N consists of the lower bound $0 \leq \mathsf{c}(N)$ and the inequalities (1) for those clutters $\mathcal{L}$ of subsets of N that contain at least one singleton and at least one another set.*

Note that the conjecture was known to be true when $|N| \leq 5$.

2.4 Two different Möbius inversions

Given a vector $\mathsf{c} \in \mathbb{R}^{\mathcal{P}(N)}$, its *superset Möbius inversion* is the vector $\mathsf{m} \in \mathbb{R}^{\mathcal{P}(N)}$ given by

$$\mathsf{m}(T) \;:=\; \sum_{S:\, T\subseteq S} (-1)^{|S\setminus T|} \cdot \mathsf{c}(S) \quad \text{for any } \; T \subseteq N. \tag{3}$$

The inverse formula to (3) is $\mathsf{c}(S) \;=\; \sum_{T:\, S\subseteq T} \mathsf{m}(T)$ for any $S \subseteq N$, which is easy to verify by a direct substitution and re-arranging sums. Note that in both formulas we sum over *supersets* of the set for which we compute the value. By (Studený and Cussens, 2017, Lemma 3), the superset Möbius inversion of the characteristic imset c_G of a chordal graph G over N has the form

$$\mathsf{m}_G(T) \;=\; \sum_{C\in\mathcal{C}(G)} \delta(T=C) - \sum_{S\in\mathcal{S}(G)} \nu_G(S)\,\cdot\,\delta(T=S) \qquad \text{for any } \; T \subseteq N\,. \tag{4}$$

Every linear inequality for c_G can be re-written in terms of its superset Möbius inversion m_G and conversely; however, the relation of the respective coefficient vectors is given by a dual Möbius inversion. Specifically, given a vector $\lambda \in \mathbb{R}^{\mathcal{P}(N)}$ its *subset Möbius inversion* is given by

$$\kappa_\lambda(S) \;:=\; \sum_{T:\, T\subseteq S} (-1)^{|S\setminus T|} \cdot \lambda(T) \quad \text{for any } \; S \subseteq N\,. \tag{5}$$

Note that the inverse formula to (5) is $\lambda(T) \;=\; \sum_{S:\, S\subseteq T} \kappa_\lambda(S)$ for any $T \subseteq N$ and that we sum over *subsets* in both these formulas. Provided that $\lambda \in \mathbb{R}^{\mathcal{P}(N)}$ is a coefficient vector for a linear inequality, a vector $\mathsf{m} \in \mathbb{R}^{\mathcal{P}(N)}$ is a superset Möbius inversion of a vector $\mathsf{c} \in \mathbb{R}^{\mathcal{P}(N)}$ one has:

$$\sum_{T\subseteq N} \lambda(T)\cdot \mathsf{m}(T) = \sum_{S\subseteq N} \kappa_\lambda(S)\cdot \mathsf{c}(S)\,, \quad \text{which particularly implies} \tag{6}$$

$$1 \;\le\; \sum_{T\subseteq N} \lambda(T)\cdot \mathsf{m}(T) \quad \Longleftrightarrow \quad 1 \;\le\; \sum_{S\subseteq N} \kappa_\lambda(S)\cdot \mathsf{c}(S)\,.$$

The verification of this fact can be done by direct substitution into the formulas and re-arranging the sums; it is left to the reader. A special case of the equivalence in (6) is a standard re-writing of the clutter inequalities in terms of superset Möbius inversion m_G from (Studený and Cussens, 2017, Lemma 2). Specifically, given a clutter $\mathcal{L} \subseteq \mathcal{P}(N)$ (containing a singleton), consider the indicator $\lambda \in \mathbb{R}^{\mathcal{P}(N)}$ of the filter $\mathcal{L}^{\uparrow}$ generated by $\mathcal{L}$. Then the subset Möbius inversion κ_λ of λ is just the coefficient vector $\kappa_{\mathcal{L}}$ of the respective clutter inequality:

$$1 \;\le\; \sum_{T\subseteq N} \delta(T\in\mathcal{L}^{\uparrow})\cdot \mathsf{m}_G(T) \qquad \Longleftrightarrow \qquad 1 \;\le\; \sum_{S\subseteq N} \kappa_{\mathcal{L}}(S)\cdot \mathsf{c}_G(S)\,.$$

In our later dual formulation of the problem we are going to assign certain linear inequalities to incomplete chordal graphs G over N. The coefficients of those inequalities are given by slightly modified superset Möbius inversions m_G; specifically, given an *incomplete chordal graph G* over N, we introduce a vector $\overline{\mathsf{m}}_G \in \mathbb{R}^{\mathcal{P}(N)}$ as follows:

$$\overline{\mathsf{m}}_G(N) = -1 \qquad \text{and} \quad \overline{\mathsf{m}}_G(S) := \mathsf{m}_G(S) \quad \text{for remaining } S \subset N. \tag{7}$$

Note that the vector $\overline{\mathsf{m}}_G$ is nothing else than (-1) multiple of the *standard imset* for the chordal graph G from (Studený, 2005, § 7.2.2).

3. Dual formulation of the problem

We assign an inequality for $\lambda \in \mathbb{R}^{\mathcal{P}(N)}$ to every *incomplete chordal graph* G over N:

$$\sum_{C\in\mathcal{C}(G)} \lambda(C) \,-\, \sum_{S\in\mathcal{S}(G)} \nu_G(S)\,\cdot\,\lambda(S) \,-\, \lambda(N) \geq 0\,, \tag{8}$$

$$\text{that is,}\quad \langle \overline{\mathsf{m}}_G, \lambda\rangle := \sum_{S\subseteq N} \overline{\mathsf{m}}_G(S)\cdot\lambda(S) \geq 0 \quad \text{using the formulas (4) and (7)}\,.$$

Definition 3 (dual polyhedron for the chordal graph polytope) We define the *dual polyhedron* $\mathsf{P} \subseteq \mathbb{R}^{\mathcal{P}(N)}$ as the set of vectors $\lambda \in \mathbb{R}^{\mathcal{P}(N)}$ satisfying $\lambda(\emptyset) = 0$, $\lambda(N) = 1$, simple monotonicity inequalities $\lambda(N) - \lambda(N \setminus \{i\}) \geq 0$ for $i \in N$ and the above-mentioned inequalities (8) assigned to incomplete chordal graphs over N.

Lemma 4 (basic facts on the dual polyhedron)
Every $\lambda \in \mathsf{P}$ is non-decreasing: $\lambda(S) \leq \lambda(T)$ whenever $S \subseteq T$. In particular, $\mathsf{P} \subseteq [0,1]^{\mathcal{P}(N)}$ is a non-empty bounded polyhedron. One has $\lambda \in \{0,1\}^{\mathcal{P}(N)} \cap \mathsf{P}$ iff λ is an indicator of singleton-containing filter $\mathcal{F} \subseteq \mathcal{P}(N)$ with $\emptyset \notin \mathcal{F}$. Moreover, $\{0,1\}^{\mathcal{P}(N)} \cap \mathsf{P} \subseteq \text{ext}\,(\mathsf{P})$, where $\text{ext}\,(\mathsf{P})$ denotes the set of vertices (= extreme points) of P.

Proof Given $T \subset N$ and $i \in T$, consider a chordal graph G with cliques T and $N \setminus \{i\}$. The inequality (8) gives $\lambda(T) + \lambda(N \setminus \{i\}) - \lambda(T \setminus \{i\}) - \lambda(N) \geq 0$. Add the simple inequality $\lambda(N) - \lambda(N \setminus \{i\}) \geq 0$ to get $\lambda(T \setminus \{i\}) \leq \lambda(T)$. Using an inductive argument we observe that any $\lambda \in \mathsf{P}$ is non-decreasing, which easily implies $\mathsf{P} \subseteq [0,1]^{\mathcal{P}(N)}$. As P is a bounded polyhedron, by (Barvinok, 2002, Corollary 8.7) it is a polytope and has finitely many vertices (= extreme points).

Given $\lambda \in \{0,1\}^{\mathcal{P}(N)} \cap \mathsf{P}$ we put $\mathcal{F} := \{T \subseteq N \,:\, \lambda(T) = 1\}$. Clearly, $\mathcal{F}$ is a filter. To see that $\mathcal{F}$ contains a singleton use (8) for the empty graph over N to obtain $\sum_{i\in N} \lambda(\{i\}) \geq 1$.

Conversely, let λ be the indicator of a singleton-containing filter $\mathcal{F} \subseteq \mathcal{P}(N)$. The validity of simple monotonicity inequalities for λ is evident. To verify (8) for λ and some incomplete chordal graph G over N consider a junction tree $\mathcal{J}$ for G and introduce a sub-forest of $\mathcal{J}$ determined by $\mathcal{F}$: it has those hyper-nodes of $\mathcal{J}$ which belong to $\mathcal{F}$ and those hyper-edges in $\mathcal{J}$ which are labeled by sets from $\mathcal{F}$. The point is that the left hand side of (8) is equal to the number of connectivity components of the $\mathcal{F}$-sub-forest reduced by 1. Indeed, recall that hyper-edges of $\mathcal{J}$ are labeled by separators and the expression $\sum_{C\in\mathcal{C}(G)} \lambda(C) \,-\, \sum_{S\in\mathcal{S}(G)} \nu_G(S)\,\cdot\,\lambda(S)$ equals the difference between the number of hyper-nodes of the $\mathcal{F}$-sub-forest and the number of its hyper-edges, which is just the number of its components. The assumption that $\mathcal{F}$ contains a singleton implies that the $\mathcal{F}$-sub-forest has at least one hyper-node, and, therefore, the left-hand side of (8) is non-negative.

To show that $\lambda \in \{0,1\}^{\mathcal{P}(N)} \cap \mathsf{P}$ is an extreme point of P assume that $\lambda = \alpha\cdot\lambda_1 + (1-\alpha)\cdot\lambda_2$ with $\alpha \in (0,1)$ and $\lambda_i \in \mathsf{P}$. For $S \subseteq N$ with $\lambda(S) = 0$ one has $0 = \alpha\cdot\lambda_1(S) + (1-\alpha)\cdot\lambda_2(S)$ with $\lambda_i(S) \geq 0$, which enforces $\lambda_i(S) = 0$. If $S \subseteq N$, $\lambda(S) = 1$, then $1 = \alpha\cdot\lambda_1(S) + (1-\alpha)\cdot\lambda_2(S)$ with $\lambda_i(S) \leq 1$, which enforces $\lambda_i(S) = 1$. Hence, $\lambda_i = \lambda$ means that $\lambda \in \text{ext}\,(\mathsf{P})$. ■

Example 2 Consider $N = \{a, b, c\}$. Then the dual polyhedron P has the dimension $6 = 2^3 - 2$ and is specified by 10 inequalities breaking into 4 permutational types. One has 3 simple monotonicity inequalities of the form $\lambda(N) \geq \lambda(N \setminus \{i\})$, and 7 inequalities assigned to incomplete chordal graphs falling into 3 permutational types: $i:j:k$, $i:jk$, and $ij:ik$. The number of vertices of P is also 10 and they fall into 4 permutational types. All of them are the indicators of singleton-containing filters; their types are $|i|$, $|i|jk|$, $|i|j|$, and $|i|j|k|$.

Note that we have confirmed computationally that in case $|N|\leq 5$ all vertices of P are indicators of singleton-containing filters. Here is the main result of the paper, proved in Appendix A.

Theorem 5 (dual characterization of the chordal graph polytope)
Assume $|N|\geq 2$. Let $\mathrm{ext}\,(\mathsf{P})$ denote the set of vertices of the dual polyhedron. Given $\mathsf{c}\in\mathsf{A}$, one has $\mathsf{c}\in D_N$ if and only if $\mathsf{c}(N)\geq 0$ and $\sum_{S\subseteq N}\kappa_\lambda(S)\cdot\mathsf{c}(S)\geq 1$ holds for every $\lambda\in\mathrm{ext}\,(\mathsf{P})$.

Corollary 6 (dual formulation of the chordal graph conjecture)
The chordal graph conjecture holds for $|N|\geq 2$ iff $\{0,1\}^{\mathcal{P}(N)}\cap\mathsf{P}=\mathrm{ext}\,(\mathsf{P})$.

Proof If $\{0,1\}^{\mathcal{P}(N)}\cap\mathsf{P}=\mathrm{ext}\,(\mathsf{P})$ then, by Theorem 5, Lemma 4 and (6), one has $\mathsf{c}\in D_N$ iff $\mathsf{c}(N)\geq 0$ and $\sum_{S\subseteq N}\kappa_{\mathcal{L}}(S)\cdot\mathsf{c}(S)\geq 1$ for every singleton-containing clutter $\mathcal{L}\subseteq\mathcal{P}(N)$. As mentioned earlier, the inequalities for clutters consisting of one (singleton) set only are superfluous.

Because of the page limit we omit the proof of the necessity of $\{0,1\}^{\mathcal{P}(N)}\cap\mathsf{P}=\mathrm{ext}\,(\mathsf{P})$ for the validity of the conjecture. Its steps correspond to the steps of the proof from Appendix A; however, the order of the steps is inverse and $\mathrm{ext}\,(\mathsf{P})$ is replaced by $\{0,1\}^{\mathcal{P}(N)}\cap\mathsf{P}$ in the claims. ■

4. Conclusion: rebuttal of the conjecture

We have just found a counter-example to the validity of the chordal graph conjecture when $|N|=6$. In the following example we only present the main idea which leads to the conjecture rebuttal; because of the page limit we do not give complete reasoning for our conclusion.

Example 3 Take $N=\{a,b,c,d,e,f\}$. We present an example of $\lambda\in\mathsf{P}$ which does not belong to the convex hull of $\{0,1\}^{\mathcal{P}(N)}\cap\mathsf{P}$. In fact, our vector λ belongs to a special 15-dimensional face F of P. Here is the definition of λ (below we abbreviate notation and write ab instead of $\{a,b\}$):

$$\begin{aligned}
0 &= \lambda(\emptyset)=\lambda(a)=\lambda(b)=\lambda(c)=\lambda(d)=\lambda(ad)=\lambda(bc),\\
\frac{1}{2} &= \lambda(e)=\lambda(f)=\lambda(ae)=\lambda(af)=\lambda(be)=\lambda(bf)=\lambda(ce)=\lambda(cf)=\lambda(de)=\lambda(df)\\
&= \lambda(ab)=\lambda(ac)=\lambda(bd)=\lambda(cd)=\lambda(abe)=\lambda(acf)=\lambda(abd)=\lambda(acd)=\lambda(bcd),\\
1 &= \lambda(L) \qquad \text{for remaining } L\subseteq N\text{; let us denote by } \mathcal{Q} \text{ the class of these remaining sets.}
\end{aligned}$$

Note that λ is non-decreasing: $\lambda(S)\leq\lambda(T)$ whenever $S\subseteq T\subseteq N$. It also belongs to the affine subspace A' of $\mathbb{R}^{\mathcal{P}(N)}$ specified by 45 equalities from the first and last line above and 4 equalities

$$\lambda^*(ab)=\lambda^*(abe),\ \ \lambda^*(ac)=\lambda^*(acf),\ \ \lambda^*(bd)=\lambda^*(abd),\ \ \lambda^*(cd)=\lambda^*(acd) \tag{9}$$

required for $\lambda^*\in\mathbb{R}^{\mathcal{P}(N)}$. Let us put (we are using our earlier notational convention)

$$\lambda' := \frac{1}{2}\cdot\lambda_{|e|ab|bd|adf|bcf|cdf|}+\frac{1}{2}\cdot\lambda_{|f|ac|cd|ade|bce|bde|} \qquad \text{and observe that } \lambda=\lambda'-\frac{1}{2}\cdot\delta_{bcd},$$

where δ_{bcd} is the indicator of the set bcd. The verification of (8) for λ breaks into 3 cases:

- If G is an incomplete chordal graph over N which has a clique $K \in \mathcal{Q}$ then the term $\lambda(K)$ in (8) cancels against the term $-\lambda(N)$. It follows from the existence of a junction tree for G that the remaining terms can be divided into pairs $[\lambda(C), -\lambda(S)]$, where $C \in \mathcal{C}(G)$, $S \in \mathcal{S}(G)$ and $S \subseteq C$. Thus, the validity of (8) follows from that fact that λ is non-decreasing.
- There exist 72 chordal graphs G over N which have the clique bcd but no clique in $\mathcal{Q}$. The facts that λ is non-decreasing and satisfies (9) allows one to observe (by a detailed analysis) that the assigned inequalities (8) follow from the following 6 inequalities

  $$\lambda(af)-\lambda(a), \lambda(ae)-\lambda(a), \lambda(cf)-\lambda(c), \lambda(be)-\lambda(b), \lambda(e)-\lambda(\emptyset), \lambda(f)-\lambda(\emptyset) \geq 1-\lambda(bcd)\,,$$

  which are clearly valid for λ with equality.
- If G is a chordal graph over N which has no clique in $\mathcal{Q} \cup \{bcd\}$ then bcd does not occur as a separator G. Hence, one can observe that $\langle \overline{\mathsf{m}}_G, \lambda\rangle = \langle \overline{\mathsf{m}}_G, \lambda'\rangle$. Moreover, by Lemma 4, $\lambda' \in \mathsf{P}$. Hence, one has $\langle \overline{\mathsf{m}}_G, \lambda'\rangle \geq 0$, which implies the same inequality (8) for λ.

Thus, $\lambda \in \mathsf{P}$. The fact that $\lambda \notin \text{conv}\,(\{0,1\}^{\mathcal{P}(N)} \cap \mathsf{P})$ can be derived by a contradiction. Otherwise, λ is a convex combination of such $\lambda^* \in \{0,1\}^{\mathcal{P}(N)} \cap \mathsf{P}$. Note that λ belongs to $\mathsf{F} := \mathsf{P} \cap \mathsf{A}'$, which is a face of P (use Lemma 4 saying that every $\lambda' \in \mathsf{P}$ is non-decreasing). Hence, all these $\lambda^* \in \{0,1\}^{\mathcal{P}(N)} \cap \mathsf{P}$ must belong to F. The equalities defining A' and Lemma 4 allow one to observe that one $\lambda^*(bcd) = 1$ for any such $\lambda^* \in \{0,1\}^{\mathcal{P}(N)} \cap \mathsf{F}$. The same holds for their convex combinations, but the fact $\lambda(bcd) = \frac{1}{2}$ contradicts this observation.

Acknowledgments

The research of Milan Studený and Václav Kratochvíl has been supported by the grant GAČR number 19-04579S.

Appendix A. Proof of Theorem 5

Proof Let cone$\,(\mathsf{P})$ denote the cone generated by the dual polyhedron P. Since P is a subset of the affine space specified by $\lambda(N) = 1$ one has cone$\,(\mathsf{P})$ = cone$\,(\text{ext}\,(\mathsf{P}))$. Observe that $\lambda \in \mathbb{R}^{\mathcal{P}(N)}$ belongs to cone$\,(\mathsf{P}) \equiv$ cone$\,(\text{ext}\,(\mathsf{P}))$ iff $\lambda(\emptyset) = 0$ and λ satisfies both the simple monotonicity inequalities and the graphical inequalities (8) for incomplete chordal graphs G over N.

Indeed, by repeating the arguments in the 1st paragraph of the proof of Lemma 4 we observe that the above inequalities imply that λ is non-decreasing; hence, $\lambda(\emptyset) = 0$ implies $\lambda(N) \geq 0$.

Let lin$\,(\mathsf{S})$ denote the linear hull of a set $\mathsf{S} \subseteq \mathbb{R}^{\mathcal{P}(N)}$. Introduce for every $j \in N$ a vector $\lambda_{|j|} \in \mathbb{R}^{\mathcal{P}(N)}$ as the indicator of the filter generated by a trivial clutter $\mathcal{L} := \{\,\{j\}\,\}$, that is, $\lambda_{|j|}(T) = \delta(j \in T)$ for any $T \subseteq N$. Then we observe, for every $\lambda \in \mathbb{R}^{\mathcal{P}(N)}$ satisfying $\lambda(\emptyset) = 0$, that one has $\lambda \in$ cone$\,(\text{ext}\,(\mathsf{P}) \cup \text{lin}\,(\{\lambda_{|j|} : j \in N\}))$ iff λ satisfies the inequalities (8) for all incomplete chordal graphs G over N.

Indeed, every $\lambda \in \text{ext}\,(\mathsf{P}) \subseteq \mathsf{P}$ satisfies inequalities (8) for all incomplete chordal graphs G over N. Since the conical combination preserves the validity of (8), the necessity of the inequalities follows from the fact that every $\lambda_{|j|}$ satisfies (8) with equality. To this end one can repeat, for every $j \in N$, the arguments in the

3rd paragraph of the proof of Lemma 4 with $\mathcal{F}_j := \{T \subseteq N \,:\, j \in T\}$ *and realize that the* $\mathcal{F}_j$*-sub-forest of* $\mathcal{J}$ *has only one connectivity component (use the definition of a junction tree for this purpose). For the sufficiency of the inequalities* (8) *we assume* $\lambda \in \mathbb{R}^{\mathcal{P}(N)}$ *satisfying them and put* $\beta_j := \lambda(N) - \lambda(N \setminus \{j\})$ *for* $j \in N$*. Then* $\lambda' := \lambda - \sum_{j \in N} \beta_j \cdot \lambda_{|j|}$ *satisfied both* $\lambda'(\emptyset) = 0$ *and the simple monotonicity inequalities with equality. Since* λ' *also satisfies all inequalities* (8)*, by the previous observation,* $\lambda' \in \text{cone}\,(\text{ext}\,(\mathsf{P}))$*, which allows us to observe that* $\lambda \in \text{cone}\,(\text{ext}\,(\mathsf{P}) \cup \text{lin}\,(\{\lambda_{|j|} : j \in N\}))$.

One can interpret any $\lambda \in \mathbb{R}^{\mathcal{P}(N)}$ satisfying $\lambda(\emptyset) = 0$ as a vector in $\mathbb{R}^{\mathcal{P}(N)\setminus\{\emptyset\}}$ and understand the inequalities (8) in this context (= ignore the component for the empty set because it plays no role). The same convention will concern the superset Möbius inversions m_G for incomplete graphs G over N and their modified versions $\overline{\mathsf{m}}_G$. In this case, the component for the empty set is determined by the remaining components and does not occur in any linear inequality of our interest. The next step is to observe using a duality consideration that, for every $\overline{\mathsf{m}} \in \mathbb{R}^{\mathcal{P}(N)\setminus\{\emptyset\}}$, $\overline{\mathsf{m}} \in \text{cone}\,(\{\overline{\mathsf{m}}_G : G \text{ is an incomplete chordal graph over } N\})$ iff $\sum_{T: j \in T} \overline{\mathsf{m}}(T) = 0$ for any $j \in N$ and $\sum_{\emptyset \neq L \subseteq N} \overline{\mathsf{m}}(L) \cdot \lambda(L) \geq 0$ for any $\lambda \in \text{ext}\,(\mathsf{P})$.

Consider two cones in $\mathbb{R}^{\mathcal{P}(N)\setminus\{\emptyset\}}$*: put* $\mathsf{K} := \text{cone}\,(\{\overline{\mathsf{m}}_G : G \textit{ is an incomplete chordal graph over } N\})$ *and* $\mathsf{L} := \text{cone}\,(\{+\lambda_{|j|} : j \in N\} \cup \{-\lambda_{|j|} : j \in N\} \cup \text{ext}\,(\mathsf{P}))$*. By definition, they are both polyhedral cones, and, therefore, closed convex cones. Consider the scalar product* $\langle \overline{\mathsf{m}}, \lambda \rangle := \sum_{\emptyset \neq S \subseteq N} \overline{\mathsf{m}}(S) \cdot \lambda(S)$ *in the space* $\mathbb{R}^{\mathcal{P}(N)\setminus\{\emptyset\}}$*, which allows one to define the dual cone* $\mathsf{K}^* := \{\lambda \in \mathbb{R}^{\mathcal{P}(N)\setminus\{\emptyset\}} \,:\, \langle \overline{\mathsf{m}}, \lambda \rangle \geq 0 \textit{ for all } \overline{\mathsf{m}} \in \mathsf{K}\}$ *to every subset* $\mathsf{K} \subseteq \mathbb{R}^{\mathcal{P}(N)\setminus\{\emptyset\}}$*. Since* $\mathsf{L} = \text{cone}\,(\text{ext}\,(\mathsf{P}) \cup \text{lin}\,(\{\lambda_{|j|} : j \in N\}))$*, by the previous observation and the equality* $\mathsf{S}^* = (\text{cone}\,(\mathsf{S}))^*$ *for any* $\mathsf{S} \subseteq \mathbb{R}^{\mathcal{P}(N)\setminus\{\emptyset\}}$ *one gets* $\mathsf{L} = \mathsf{K}^*$*. A well-known fact is that* $\mathsf{K} = \mathsf{K}^{**}$ *for any closed cone; see for example (Studený, 1993, Consequence 1). In particular,* $\mathsf{L} = \mathsf{K}^*$ *gives* $\mathsf{L}^* = \mathsf{K}^{**} = \mathsf{K}$*. Again using* $(\text{cone}\,(\mathsf{S}))^* = \mathsf{S}^*$ *observe that* $\mathsf{K} = \mathsf{L}^*$ *consists of those* $\overline{\mathsf{m}} \in \mathbb{R}^{\mathcal{P}(N)\setminus\{\emptyset\}}$ *for which* $\langle \overline{\mathsf{m}}, \lambda_{|j|} \rangle = 0$ *for any* $j \in N$ *and* $\langle \overline{\mathsf{m}}, \lambda \rangle \geq 0$ *for any* $\lambda \in \text{ext}\,(\mathsf{P})$.

This allows us to characterize the convex hull of our graphical vectors: for every $\overline{\mathsf{m}} \in \mathbb{R}^{\mathcal{P}(N)\setminus\{\emptyset\}}$, $\overline{\mathsf{m}} \in \text{conv}\,(\{\overline{\mathsf{m}}_G : G \text{ is an incomplete chordal graph over } N\})$ iff $\overline{\mathsf{m}}(N) = -1$, $\sum_{T: j \in T} \overline{\mathsf{m}}(T) = 0$ for any $j \in N$ and $\sum_{\emptyset \neq L \subseteq N} \overline{\mathsf{m}}(L) \cdot \lambda(L) \geq 0$ for any $\lambda \in \text{ext}\,(\mathsf{P})$.

By (7)*, one has* $\overline{\mathsf{m}}_G(N) = -1$ *for each incomplete chordal graph* G *over* N*, which implies, together with the above characterization of the conic hull, the necessity of our linear constraints. For their sufficiency also use that observation saying that* $\overline{\mathsf{m}}$ *belongs to the conic hull:* $\overline{\mathsf{m}} = \sum_k \alpha_k \cdot \overline{\mathsf{m}}_{G_k}$*, where* G_k *are (all) incomplete chordal graphs over* N *and* $\alpha_k \geq 0$*. The substitution* $-1 = \overline{\mathsf{m}}(N) = \sum_k \alpha_k \cdot \overline{\mathsf{m}}_{G_k}(N) = \sum_k \alpha_k \cdot (-1)$ *gives* $\sum_k \alpha_k = 1$*, that is,* $\overline{\mathsf{m}}$ *belongs to the convex hull.*

By definition, $\mathsf{m}_G(N) = 0$ for incomplete chordal graphs G; using (7), for every $\mathsf{m} \in \mathbb{R}^{\mathcal{P}(N)\setminus\{\emptyset\}}$, $\mathsf{m} \in \text{conv}\,(\{\mathsf{m}_G : G \text{ is an incomplete chordal graph over } N\})$ iff $\mathsf{m}(N) = 0$, $\sum_{T: j \in T} \mathsf{m}(T) = 1$ for any $j \in N$ and $\sum_{\emptyset \neq L \subseteq N} \lambda(L) \cdot \mathsf{m}(L) \geq \lambda(N) = 1$ for any $\lambda \in \text{ext}\,(\mathsf{P})$.

One can re-write that in terms of the vector c whose superset Möbius inversion is m, just use the formula (6): for every $\mathsf{c} \in \mathbb{R}^{\mathcal{P}(N)\setminus\{\emptyset\}}$, $\mathsf{c} \in \text{conv}\,(\{\mathsf{c}_G : G \text{ is an incomplete chordal graph over } N\})$ iff $\mathsf{c}(N) = 0$, $\mathsf{c}(\{j\}) = 1$ for any $j \in N$ and $\sum_{\emptyset \neq S \subseteq N} \kappa_\lambda(S) \cdot \mathsf{c}(S) \geq 1$ for any $\lambda \in \text{ext}\,(\mathsf{P})$.

Finally observe that $\mathsf{c} \in \text{conv}\,(\{\mathsf{c}_G : G \text{ is a chordal graph over } N\})$ iff $\mathsf{c}(N) \geq 0$, $\mathsf{c}(\{j\}) = 1$ for any $j \in N$ and $\sum_{\emptyset \neq S \subseteq N} \kappa_\lambda(S) \cdot \mathsf{c}(S) \geq 1$ for any $\lambda \in \text{ext}\,(\mathsf{P})$.

The characteristic imset c_H *for the complete graph* H *over* N *clearly satisfies those linear constraints even with equalities except for* $\mathsf{c}(N) \geq 0$*: by* (4) *its superset Möbius inversions* m_H *is the indicator of the set* N*, use* (6) *and the fact that* $\lambda(N) = 1$ *for every* $\lambda \in \text{ext}\,(\mathsf{P}) \subseteq \mathsf{P}$*. This, together with the former observation about incomplete graphs implies the necessity of the linear constraints.*

For their sufficiency we first realize that they imply that c *is non-increasing:* $\mathsf{c}(S) \geq \mathsf{c}(T)$ *whenever* $\emptyset \neq S \subseteq T \subseteq N$. *To this end, for every fixed* $\emptyset \neq S \subset N$ *and* $i \in N \setminus S$ *consider the clutter* $\mathcal{L} := \{\, \{i\}, S \,\}$ *and the indicator* $\lambda_{|i|S|}$ *of the corresponding filter. By Lemma 4,* $\lambda_{|i|S|} \in \mathrm{ext}\,(\mathsf{P})$. *The inequality for* $\lambda_{|i|S|}$, *that is, the clutter inequality for* $\mathcal{L}$, *then gives* $\mathsf{c}(\{i\}) + \mathsf{c}(S) - \mathsf{c}(\{i\} \cup S) \geq 1$ *(see Section 2.3). Substitute* $\mathsf{c}(\{i\}) = 1$ *to get* $\mathsf{c}(S) \geq \mathsf{c}(\{i\} \cup S)$; *an inductive argument can be used to conclude that* c *is non-increasing. Because* $\mathsf{c}(\{j\}) = 1$ *for arbitrary* $j \in N$, *it implies* $0 \leq \mathsf{c}(N) \leq 1$. *Let us put* $\alpha := \mathsf{c}(N)$. *In case* $\alpha = 0$ *we use the former observation to conclude that* c *is in the convex hull. In case of* $\alpha = 1$ *use that fact that* c *is non-increasing to realize that* $\mathsf{c} = \mathsf{c}_H$, *which again implies the desired conclusion. In case* $0 < \alpha < 1$ *we put* $\mathsf{c}' := \frac{1}{1-\alpha} \cdot (\mathsf{c} - \alpha \cdot \mathsf{c}_H)$, *which means that* $\mathsf{c} = \alpha \cdot \mathsf{c}_H + (1-\alpha) \cdot \mathsf{c}'$ *is a convex combination. It implies, for every* $\lambda \in \mathrm{ext}\,(\mathsf{P})$, *that*

$$\sum_{\emptyset \neq S \subseteq N} \kappa_\lambda(S) \cdot \mathsf{c}'(S) = \frac{1}{1-\alpha} \cdot \Big(\underbrace{\sum_{\emptyset \neq S \subseteq N} \kappa_\lambda(S) \cdot \mathsf{c}(S)}_{\geq 1} - \alpha \cdot \underbrace{\sum_{\emptyset \neq S \subseteq N} \kappa_\lambda(S) \cdot \mathsf{c}_H(S)}_{=1} \Big) \geq \frac{1}{1-\alpha} \cdot (1-\alpha) = 1\,.$$

Since one has both $\mathsf{c}'(N) = 0$ *and* $\mathsf{c}'(\{j\}) = 1$ *for any* $j \in N$, *by the former claim,* c' *is in the convex hull of* c_G*'s for incomplete chordal graphs* G *over* N. *Hence,* c *is in the convex hull for all chordal graphs over* N.

One has $D_N = \mathrm{conv}\,(\{\mathsf{c}_G : G \text{ is a chordal graph over } N\})$ and $\mathsf{c} \in \mathsf{A}$ iff $\mathsf{c}(\{j\}) = 1$ for any $j \in N$. Thus, the last claim above implies what is said in Theorem 5. ■

References

A. Barvinok. *A Course in Convexity*. American Mathematical Society, Providence, 2002.

V. Chvátal. *Linear Programming*. W. H. Freeman, 1983.

J. Corander, T. Janhunen, J. Rintanen, H. Nyman, and J. Pensar. Learning chordal Markov networks by constraint satisfaction. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weiberger, editors, *Advances in Neural Information Processing Systems 26*, pages 1349–1357. Curran Associates, 2013.

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

J. Cussens, M. Studený, and D. Haws. Polyhedral aspects of score equivalence in Bayesian network structure learning. *Mathematical Programming A*, 164(1/2):285–324, 2017.

D. Heckerman, D. Geiger, and D. M. Chickering. Learning Bayesian networks: the combination of knowledge and statistical data. *Machine Learning*, 20:194–243, 1995.

R. Hemmecke, S. Lindner, and M. Studený. Characteristic imsets for learning Bayesian network structure. *International Journal of Approximate Reasoning*, 53:1336–1349, 2012.

K. Kangas, T. Niinimäki, and M. Koivisto. Learning chordal Markov networks by dynamic programming. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 27*, pages 2357–2365. Curran Associates, 2014.

K. Kangas, T. Niinimäki, and M. Koivisto. Averaging decomposable graphs by dynamic programming and sampling. In M. Meila and T. Heskes, editors, *Uncertainty in Artificial Intelligence 31*, pages 415–424. AUAI Press, 2015.

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

S. Lindner. *Discrete optimization in machine learning: learning Bayesian network structures and conditional independence implication*. PhD thesis, TU Munich, 2012.

R. E. Neapolitan. *Learning Bayesian Networks*. Pearson Prentice Hall, Upper Saddle River, 2004.

J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufman, San Mateo, 1988.

A. Pérez, C. Blum, and J. A. Lozano. Learning maximum weighted $(k+1)$-order decomposable graphs by integer linear programming. In L. C. van der Gaag and A. J. Feelders, editors, *Lecture Notes in AI 8754: PGM 2014*, pages 396–408, 2014.

A. Pérez, C. Blum, and J. A. Lozano. Approximating the maximum weighted decomposable graph problem with applications to probabilistic graphical models. In V. Kratochvíl and M. Studený, editors, *JMLR Workshop and Conference Proceedings 72: PGM 2018*, pages 320–331, 2018.

K. Rantanen, A. Hyttinen, and M. Järvisalo. Learning chordal Markov networks via branch and bound. In *Information Processing Systems 31: NISP 2017*, pages 1848–1858, 2017.

G. E. Schwarz. Estimation of the dimension of a model. *Annals of Statistics*, 6:461–464, 1978.

K. S. Sesh Kumar and F. Bach. Convex relaxations for learning bounded-treewidth decomposable graphs. In S. Dasgupta and D. McAlester, editors, *JMLR Workshop and Conference Proceedings 28: ICML 2013*, volume 1, pages 525–533, 2013.

M. Studený. Convex cones in finite-dimensional real vector spaces. *Kybernetika*, 2:180–200, 1993.

M. Studený. *Probabilistic Conditional Independence Structures*. Springer, London, 2005.

M. Studený and J. Cussens. The chordal graph polytope for learning decomposable models. In A. Antonucci, G. Corani, and C. P. de Campos, editors, *JMLR Workshop and Conference Proceedings 52: PGM 2016*, pages 499–510, 2016.

M. Studený and J. Cussens. Towards using the chordal graph polytope in learning decomposable models. *International Journal of Approximate Reasoning*, 88:259–281, 2017.

M. Studený and D. Haws. Learning Bayesian network structure: towards the essential graph by integer linear programming tools. *International Journal of Approximate Reasoning*, 55:1043–1071, 2014.

L. A. Wolsey. *Integer Programming*. John Wiley, New York, 1998.