Extended Formulations for Higher Order Polytopes in Combinatorial Optimization

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

von Mirjam Friesen, M.Sc.

geb. am 6. Oktober 1988 in Hamburg

genehmigt durch die Fakultät für Mathematik
der Otto-von-Guericke-Universität Magdeburg

Gutachter: Prof. Dr. Volker Kaibel
Prof. Dr. Christoph Buchheim

eingereicht am: 17. Juni 2019
Verteidigung am: 24. September 2019

Zusammenfassung

Wir sind interessiert an der konvexe Hülle von Vektoren $(x, y) \in \{0,1\}^n$, bei denen y ein Monom in x ist. Diese Vektoren können als *charakteristische Vektoren höherer Ordnung* kombinatorischer Strukturen betrachtet werden. Wir nennen diese Polytope entsprechend *Polytope höherer Ordnung*.

Mit Hilfe von linearer Optimierung über diese Art von Polytopen können polynomielle Optimierungsprobleme wie zum Beispiel das *quadratische minimale Spannbaumproblem (QMST-Problem)* gelöst werden. Diese Probleme sind häufig NP-schwer und in der Regel sind keine vollständigen Beschreibungen der zugehörigen Polytope höherer Ordnung in Form von Gleichungen und Ungleichungen bekannt.

Es gibt Beschreibungen für Matroidpolytope höherer Ordnung, allerdings nur für sehr spezielle Mengen von Monomen [15] [16]. Diese Beschreibungen brauchen exponentiell (in der Größe der Grundmenge) viele Ungleichungen.

In dieser Arbeit erforschen wir *erweiterte Formulierungen*. Um Monome zu modellieren, nutzen wir kleine erweiterte Formulierungen des Spannbaumpolytops. Mit klein meinen wir Formulierungen, welche nur polynomiell (in der Anzahl der Graphknoten) viele Ungleichungen haben. Die erweiterten Formulierungen beinhalten zusätzliche strukturelle Informationen, mit deren Hilfe wir kleine erweiterte Formulierungen der Waldpolytope höherer Ordnung mit verschachtelten Monomen, welche Bäumen entsprechen, und mit verschachtelten Monomen vom maximalen Grad 3 modellieren. Das beinhaltet den Fall von einem Monom vom Grad 2 oder 3 und impliziert Formulierungen für die zugehörigen Spannbaumpolytope höherer Ordnung.

Der Fall mit einem Monom vom Grad 2 ist durch seine Verbindung zum QMST-Problem besonders interessant. Indem wir die Beschreibung der Spannbaumpolytope mit einem Monom vom Grad 2 für alle möglichen grad-2 Monome kombinieren, erhalten wir eine Relaxierung des *quadratischen Spannbaumpolytopes*. Nutzen wir als Beschreibungen unsere erweiterten Formulierungen, modellieren wir auf implizierte Weise eine zusätzliche Beziehung zwischen den Monomen und verbessern die Relaxierung im Vergleich zu jener, welche wir mit den Beschreibungen im Originalraum erhalten. Als Nebenresultat finden wir neue Facetten des adjazenten quadratischen Waldpolytopes und des adjazenten quadratischen Spannbaumpolytopes. Mit Hilfe von Computerexperimenten veranschaulichen wir den Grad der Verbesserung in den Relaxierungen.

Bezüglich gerichteter Graphen wissen wir von keiner vollständigen

Beschreibung für Arboreszenzpolytope höherer Ordnung, solange die Monommenge nichtleer ist. Wir vergleichen zwei erweiterte Formulierungen des Arboreszenzpolytopes bezüglich der Möglichkeiten einzelne grad-2 Monome zu modellieren. Die erweiterten Formulierungen projizieren auf neue Relaxierungen der zugehörigen Arboreszenzpolytope höherer Ordnung.

Summary

We are interested in the convex hull of vectors $(x, y) \in \{0,1\}^n$, where y is a monomial in x. Those vectors can be considered as higher order characteristic vectors of combinatorial structures. Accordingly, we call those polytopes higher order polytopes.

Linear optimization over those polytopes solves polynomial combinatorial optimization problems like for example the quadratic minimum spanning tree problem (QMST-problem). Those problems are often NP-hard and complete descriptions of the corresponding higher order polytopes in terms of equations and inequalities are usually unknown.

There are descriptions of higher order matroid polytopes, but only for special sets of monomials [15] [16]. Those descriptions have exponentially (in the size of the ground set) many inequalities.

In this work, we investigate extended formulations. To model monomials, we use small extended formulations for the spanning tree polytope. By small we mean formulations that do only have polynomially (in the number of graph nodes) many inequalities. The extended formulations provide additional structural information, which we use to model small extended formulations for higher order forest polytopes with nested monomials that are trees and with nested monomials up to degree-three. This includes the cases of one degree-two or degree-three monomial and implies formulations for the corresponding higher order spanning tree polytopes.

The degree-two case is of special interest due to its relation to the QMST-problem. Combining the descriptions of higher order spanning tree polytopes with one degree-two monomial for all possible degree-two monomials, we obtain a relaxation of the quadratic spanning tree polytope. Doing this with our extended formulations for one degree-two monomial we model in an implicit way a further relation between the monomials and improve the relaxation compared to those we obtain using the descriptions in the original space. As a side effect, we find new facets of the adjacent quadratic forest polytope and the adjacent quadratic spanning tree polytope. Via computational experiments we visualize the amount of improvement of the relaxations.

Considering directed graphs we do not know a complete description of higher order arborescence polytopes for any nonempty set of monomials. We compare two extended formulations for the arborescence polytope regarding their ability to model degree-two monomials. The extended formulations project onto new relaxations of the corresponding higher order arborescence polytopes.

Acknowledgments

— Due to General Data Protection Regulation this chapter is slashed in the electronic version. —

I thank the Deutsche Forschungsgesellschaft (DFG, German Research Foundation), which funded most of this work. [1] [2]

[1]"Extended formulations in combinatorial optimization", KA1616/4-1
[2]DFG-GRK 2297 MathCoRe

Contents

1 Introduction

In combinatorial optimization we optimize over a finite set of elements. Such elements are often subsets of some basic set, like for example all cycle free edge sets of a graph (known as forests). Let E be a finite basic set and $\mathcal{T} \subseteq 2^E$ be a set of combinatorial structures. For a set $\mathcal{M} \subseteq 2^E$ we define the *higher order polytope*

$$P_{\mathcal{T}}(\mathcal{M}) := \operatorname{conv}\left\{(x,y) \in \{0,1\}^E \times \{0,1\}^{\mathcal{M}} \;\middle|\; \begin{aligned} x &= \chi(T), && T \in \mathcal{T} \\ y_M &= \prod_{e \in M} x_e, && M \in \mathcal{M} \end{aligned}\right\},$$

where $\chi(T) \in \{0,1\}^E$ is the *characteristic vector* of T with $\chi(T)_e = 1$ if and only if $e \in T$. The higher order polytope of the empty set $P_{\mathcal{T}}(\varnothing)$ is the polytope $P_{\mathcal{T}} = \operatorname{conv}\{\chi(T) \mid T \in \mathcal{T}\}$.

We observe that the y-variables are linearization variables for monomials in x as well as characteristics for sets $M \in \mathcal{M}$ with $y_M = 1$ if and only if $M \subseteq T$. Due to this identification we call the sets $M \in \mathcal{M}$ *monomials* and the vectors (x,y) *higher order characteristic vectors*.

Our investigations are motivated by the fact that for all $c \in \mathbb{Q}^E$ and $q \in \mathbb{Q}^{\mathcal{M}}$ we can solve the polynomial optimization problem

$$\min\left\{\sum_{e \in E} c_e x_e + \sum_{M \in \mathcal{M}} q_M \prod_{e \in M} x_e \;\middle|\; x = \chi(T),\ T \in \mathcal{T}\right\}$$

by linear optimization over $P_{\mathcal{T}}(\mathcal{M})$.

Depending on $\mathcal{M}$ it might be hard to describe $P_{\mathcal{T}}(\mathcal{M})$ directly and much easier to describe $P_{\mathcal{T}}(\mathcal{M}_i)$ for subsets $\mathcal{M}_i \subset \mathcal{M}$ for $i \in [k]$, where $[k] := \{1,\dots,k\}$. Using this we can create a relaxation of $P_{\mathcal{T}}(\mathcal{M})$ defined as

$$\mathcal{R}(\mathcal{M}_1,\dots,\mathcal{M}_k) := \left\{(x,y) \in \mathbb{R}^E \times \mathbb{R}^{\mathcal{M}} \;\middle|\; \begin{aligned} &(x, y|_{\mathcal{M}_i}) \in P_{\mathcal{T}}(\mathcal{M}_i) \\ &\text{for all } i \in [k] \end{aligned}\right\}. \tag{1.1}$$

A *relaxation* of an integer polytope P (i.e., $P = \operatorname{conv} X$ for some finite set $X \subset \mathbb{Z}^n$) is a polytope $R \supseteq P$ such that $R \cap \mathbb{Z}^n = P \cap \mathbb{Z}^n$. Those relaxations are used to model optimization problems as *(mixed)* *integer programs* (MIP or IP respectively). They can then be solved with the popular *branch and cut algorithm* implemented in several optimization solvers. (See Chapter 4 for a short introduction or [6] for a survey about integer programming.)

The idea of building relaxations like in (1.1) is due to Buchheim and Klein, who did this in the context of spanning trees and forests for single degree-two monomials $\mathcal{M}_i = \{M_i\}$ with $|M_i| = 2$ [4]. They found complete descriptions of the higher order forest polytope $P_F(\mathcal{M}_i)$ and its face the higher order spanning tree polytope $P_{ST}(\mathcal{M}_i)$. Furthermore, they observed an improvement of the root gap compared to the relaxation given by a description of P_F and McCormick's linearization [30] in computational experiments.

Their descriptions were independently shown to be complete by Fischer and Fischer, who continued the work with McCormick and developed descriptions of higher order matroid polytopes with nested monomials ($M_1 \subset M_2 \subset \cdots \subset M_k$) and with monotone monomials ($\mathcal{M} = 2^{\overline{E}}$ for some set $\overline{E} \subseteq E$) [14] [15] [16]. We use their description for nested monomials in Chapter 2 to prove our extended formulations for higher order forest polytopes with nested monomials.

In her dissertation Klein also studied higher order branching and arborescence polytopes as well as higher order matching polytopes all restricted to single degree-two monomials [24]. She had a conjecture for a complete description of the higher order matching polytope with one degree-two monomial in bipartite graphs, which leads to a relaxation of the quadratic assignment polytope. Later her conjecture was proved by Walter [35]. Hupp, Klein and Liers used facets of the higher order matching polytope with one degree-two monomial in their implementations to solve the quadratic matching problem [22].

In this work we investigate *extended formulations* for such polytopes. An *extension* of a polytope P is a polytope Q in a higher dimension that can be projected onto P. Instead of optimizing over P one can then optimize over Q. A description of Q in terms of equations and inequalities is called an *extended formulation* for P. We measure the *size* of an extended formulation as the number of inequalities and denote by $\mathrm{xc}(P)$ the *extension complexity*, i.e., the size of a smallest possible extension of P.

Extended formulations were successfully used to decrease the size of formulations in many cases. (See [5] and [23] for surveys.) We analyse an additional effect. Using extended formulations for $P_{\mathcal{T}}(\mathcal{M}_i)$ of sizes σ_i we can clearly combine them to an extended formulation for $\mathcal{R}(\mathcal{M}_1, \ldots, \mathcal{M}_k)$ of size $\sigma_1 + \cdots + \sigma_k$. In this combined formulation we can use linear relations between the additional variables, like identifying some of them, to obtain an extended formulation $\mathcal{R}'(\mathcal{M}_1, \ldots, \mathcal{M}_k)$ with

$$P_{\mathcal{T}}(\mathcal{M}) \subseteq \text{proj}\left(\mathcal{R}'(\mathcal{M}_1, \ldots, \mathcal{M}_k)\right) \subsetneq \mathcal{R}(\mathcal{M}_1, \ldots, \mathcal{M}_k),$$

where proj describes the projection onto $\mathbb{R}^E \times \mathbb{R}^{\mathcal{M}}$. Hence, the relaxation we obtain using extended formulations can improve the relaxation build with the descriptions in the original space.

We show this effect in the (hopefully) prototypical example of spanning trees and forests with degree-two monomials in Chapter 3. Moreover, we analyse the practical impact of this result via computational experiments in Chapter 4. Therefore, we generate random instances of the quadratic minimum spanning tree problem (QMST-problem) and analyse the root gap and other measurements for several IP formulations. The formulations in the original space are based on Buchheim and Klein's description and the other formulations are build out of our new extended formulations from Chapter 2.

The projection of our combined extended formulations lead to new valid inequalities of the quadratic forest and the quadratic spanning tree polytope. Those including only adjacent monomials are actually facets of the adjacent quadratic forest polytope and the adjacent quadratic spanning tree polytope as we show in Chapter 3.

Considering rooted arborescences we are not aware of a description of the higher order arborescence polytopes for any nonempty set of monomials. In Chapter 5 we compare two different extended formulations regarding their capability to model degree-two monomials. Here, the structural information given by the formulations helps to understand a few facets of $P_{\mathcal{T}}(\mathcal{M})$ that we can imply by our formulations. Contrary to the polytopes that Klein studied in her dissertation [24], we consider arborescences with a fixed root node.

A general upper bound for the extension complexity of higher order polytopes arises from Balas' disjunctive programming bound [3]. Therefore, we consider monomials $\mathcal{M}$ with a constant width in the view of partially ordered sets via inclusion and assume that we know an extended formulation for $P_{\mathcal{T}}$.

Theorem 1. *There exists an extended formulation for $P_{\mathcal{T}}(\mathcal{M})$ of size $(\zeta + 1)^{\omega}\sigma$ where $\zeta = \max_{M \in \mathcal{M}} |M|$, ω is the width of $\mathcal{M}$ viewed as a partially ordered set (by inclusion) and σ is the size of some extended formulation for $P_{\mathcal{T}}$.*

Proof. By Dilworth's Theorem [11] there exist ω chains

$$M^i_1 \subsetneq M^i_2 \subsetneq \cdots \subsetneq M^i_{\ell_i} \qquad \text{for all } i \in [\omega]\,,$$

with $\mathcal{M} = \left\{ M^i_j \middle| \, i \in [\omega]\,, j \in [\ell_i] \right\}$.

Since $P_{\mathcal{T}}(\mathcal{M})$ is the (coordinate) projection of $P_{\mathcal{T}}(\mathcal{M}')$ for $\mathcal{M} \subseteq \mathcal{M}'$, we can assume (with $M^i_0 := \varnothing$)

$$M^i_j \setminus M^i_{j-1} = \{e^i_j\}$$

for some unique $e^i_j \in E$, for all $i \in [\omega]$ and for all $j \in [\ell]$. Let $T \in \mathcal{T}$, we define the signature $s(T) \in \{0, \ldots, \zeta\}^\omega$ by

$$s_i := \max \left\{ j \in [\ell_i] \,\middle|\, e^i_j \in T \right\} \qquad \text{for all } i \in [\omega].$$

For each possible signature $s \in \{0, \ldots, \zeta\}^\omega$ we consider the face $P_{\mathcal{T}}(\mathcal{M})^s$ of $P_{\mathcal{T}}(\mathcal{M})$ defined by

$$x_{e^i_j} = 1 \qquad \text{for all } i \in [\omega] \text{ and } j \in [s_i] \tag{1.2}$$

$$x_{e^i_{s_i+1}} = 0 \qquad \text{if } s_i < \ell_i \tag{1.3}$$

$$y_{M^i_j} = 1 \qquad \text{for all } i \in [\omega] \text{ and } j \in [s_i]$$

$$y_{M^i_j} = 0 \qquad \text{for all } i \in [\omega] \text{ and } j \in [\ell_i] \setminus [s_i]\,.$$

Due to the fact that y_M is fixed for all $M \in \mathcal{M}$ the polytope $P_{\mathcal{T}}(\mathcal{M})^s$ is isomorphic to the face of $P_{\mathcal{T}}$ defined by equations (1.2) and (1.3). Consequently, $P_{\mathcal{T}}(\mathcal{M})$ is isomorphic to the convex hull of m faces of $P_{\mathcal{T}}$, where m is the number of possible signatures s, which is at most $(\zeta + 1)^\omega$.

Using Balas's [3] extended formulation with the well known disjunctive programming bound

$$\mathrm{xc}\left(P_1 \cup P_2 \cup \cdots \cup P_m\right) \leq \sum_{i=1}^{m} \max \left\{ \mathrm{xc}\left(P_i\right), 1 \right\}$$

we can build an extended formulation for size $(\zeta + 1)^\omega \sigma$. $\square$

Our aim was to find smaller formulations than those in Theorem 1. Given an undirected graph $G = (V, E)$ we build extended formulations for the higher order forests polytope $P_F(\mathcal{M})$ and its face the higher order spanning tree polytope $P_{ST}(\mathcal{M})$ for some specific sets of monomials $\mathcal{M} \subseteq 2^E$ in Chapter 2. Therefore, we use an extended formulation for

P_{ST} by Martin that has size $\Theta\left(|V||E|\right)$ [29, Section 3.1]. Our formulations increase this formulation only by a summand of size $O\left(|\mathcal{M}|\zeta^2\right)$ instead of a factor as in Theorem 1. The formulations can be used black box with other descriptions of P_F or P_{ST} respectively.

For planar graphs there exists a much smaller extended formulation for P_{ST} by Williams of size $\Theta\left(|E|\right)$ [37]. Using this formulation with our black box approach based on Martin's formulation for one degree-two monomial we got a formulation of size $\Theta\left(|E|\right)$, which is asymptotically the same size as the formulation from Theorem 1, but smaller by a factor of two. We discuss this and a further formulation for single adjacent degree-two monomials directly based on Williams' in Section 2.1.

Our formulations for higher order arborescence polytopes are not complete descriptions and project onto relaxations of $P_{\mathcal{T}}(\mathcal{M})$, but similarly to the formulations for forests they are very small, since they only add a small number of inequalities to the extended formulations for the arborescence polytope that we build on.

Preliminaries In this work we assume basic knowledge about convex geometry, polyhedra and mathematical optimization. Additionally we introduce important notations and expressions on the first appearance and provide a list of notations and a glossary at the end.

2 Extended Formulations for Higher Order Forest Polytopes

Let $G = (V, E)$ be a graph. A *forest* in G is a cycle-free set $F \subseteq E$. If F connects all nodes in V we call it a *spanning tree*. The forest polytope P_F of a graph G is the polytope $P_{\mathcal{T}}$ as defined in Chapter 1 where $\mathcal{T}$ is the set of all forests in G, i.e.,

$$P_F := \operatorname{conv}\left\{\chi(F) \,\middle|\, F \text{ is a forest in } G\right\}.$$

Analogously for connected graphs the spanning tree polytope P_{ST} is defined as

$$P_{ST} := \operatorname{conv}\left\{\chi(T) \,\middle|\, T \text{ is a spanning tree of } G\right\}.$$

We omit G from the notation, since the graph G should be clear from the context.

For $S \subseteq V$ we denote by $E(S)$ all edges in E that have both end nodes in S. Furthermore we define $x(D) := \sum_{e \in D} x_e$ for all $D \subseteq E$.

Proposition 1 (Edmonds [13])**.** *The forest polytope P_F is described by*

$$x(E(S)) \leq |S| - 1 \qquad \textit{for all } S \subseteq V \textit{ with } S \neq \emptyset \tag{2.1}$$

$$x \geq 0 \tag{2.2}$$

and for connected graphs the spanning tree polytope P_{ST} is its face defined by

$$x(E) = |V| - 1. \tag{2.3}$$

Edmonds' constraints (2.1) are also called *rank inequalities* in the context of matroid theory or *subtour elimination constraints* in the context of the traveling salesman problem.

As you can see Edmonds' description has exponential size $\Theta\left(2^{|E|}\right)$, although the minimum spanning tree problem can be solved in polynomial time in $|V|$ with algorithms like those of Prim and Kruskal [33] [26].

One very nice and small extended formulation for P_{ST} is due to Martin and has size $\Theta(|V||E|)$ [29]. It provides a lot of additional structural information that we will use in order to design extended formulations for $P_F(\mathcal{M})$ and $P_{ST}(\mathcal{M})$ for some specific sets of monomials $\mathcal{M}$ later.

For $S \subseteq V$ we denote by $\delta(S)$ all edges adjacent to S, i.e.,

$$\delta(S) := \left\{\{v, w\} \in E \,|\, v \in S \text{ and } w \notin S\right\}.$$

We simply write $\delta(v)$ instead of $\delta(\{v\})$ for a single node $v \in V$. When dealing with directed graphs we denote the edges directed towards S by $\delta^{in}(S)$ and the edge directed out of S by $\delta^{out}(S)$. An *arborescence* of a directed graph is a cycle-free set of edges, such that $\delta^{in}(v) = 1$ for all nodes v except of one root node r where $\delta^{in}(r) = 0$.

Proposition 2 (Martin [29]). *Let $G = (V, E)$ be a connected graph. The following system together with the projection onto x is an extended formulation for P_{ST}.*

$$z^u_{v,w} + z^u_{w,v} = x_{\{v,w\}} \qquad \text{for all } u \in V \text{ and } \{v,w\} \in E \tag{2.4}$$

$$z^u\left(\delta^{in}(v)\right) = 1 \qquad \text{for all } u, v \in V \text{ with } u \neq v \tag{2.5}$$

$$z^u\left(\delta^{in}(u)\right) = 0 \qquad \text{for all } u \in V \tag{2.6}$$

$$z \geq 0, \tag{2.7}$$

where $z^u\left(\delta^{in}(v)\right) := \sum_{\{v,w\}\in\delta(v)} z^u_{w,v}$.

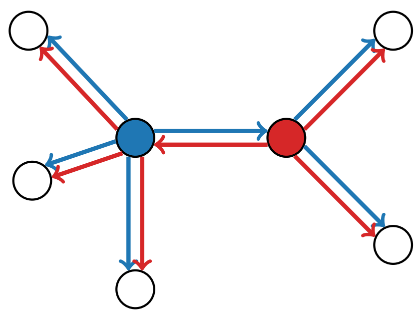


Figure 2.1: Two arborescences with the same underlying tree

The z-variables are related to arborescences in the following way: For a given spanning tree $T \subseteq E$ with the characteristic vector x we can define valid z^u as the characteristic vectors of the corresponding u-arborescence, i.e., the directed version of T with $\delta^{in}(v) = 1$ for all $v \in V \setminus \{u\}$ and $\delta^{in}(u) = 0$. In Figure 2.1 you can see two of these arborescences.

To build extended formulations for forests instead of trees we modify Martin's formulation by replacing equation (2.4) by

$$x_{\{v,w\}} \leq z^a_{v,w} + z^a_{w,v} \qquad \text{for all } \{v,w\} \in E. \tag{2.4'}$$

Proposition 3 (Martin [29]). *Let $G = (V, E)$ be a connected graph. The system* (2.4′),(2.5)-(2.7) *together with the projection onto x is an extended formulation for P_F.*

To construct the z-variables for a given forest $F \subseteq E$ we extend F to a spanning tree T with $F \subseteq T$ and construct the z-variables as the characteristic vectors of u-arborescence as before.

We will present extended formulations for $P_F(\mathcal{M})$ based on Martin's extended formulation for specific monomials $\mathcal{M}$. They all imply extended formulations for the corresponding higher order spanning tree polytope $P_{ST}(\mathcal{M})$ using a formulation for P_{ST} instead of P_F and equation (2.4) instead of inequality (2.4').

Our extended formulations as well as the descriptions of $P_F(\mathcal{M})$ that we work with contain the following linearization constraints.

Proposition 4 (McCormick [30]). *Let $x \in \{0,1\}^E$. For a set of nested monomials $\varnothing = M_0 \subset M_1 \subset \cdots \subset M_k \subseteq E$ and $y \in \{0,1\}^k$ defined by $y_i := \prod_{e \in M_i} x_e$ for all $i \in [k]$ the following linearization constraints are valid.*

$$y_i \le x_e \qquad \text{for all } i \in [k] \text{ and } e \in M_i \setminus M_{i-1} \tag{2.8}$$

$$y_i \le y_{i-1} \qquad i \in [k] \setminus \{1\} \tag{2.9}$$

$$y_1 \ge \sum_{e \in M_1} x_e - |M_1| + 1 \tag{2.10}$$

$$y_i \ge \sum_{e \in M_i \setminus M_{i-1}} x_e + y_{i-1} - |M_i \setminus M_{i-1}| \qquad i \in [k] \setminus \{1\}. \tag{2.11}$$

2.1 Degree-Two Monomials

We consider the case of one degree-two monomial $\mathcal{M} = \{M\} \subset 2^E$ with $|M| = 2$.

In the *adjacent case*, where $M = \{\{a,b\},\{b,c\}\}$ for pairwise distinct $\{a,b,c\} \subseteq V$, we will find smaller descriptions than in the *general case*. Additionally, they are, in our opinion, easier to understand. Therefore, we consider this case separately.

In this section we write y instead of y_M, since we only consider one monomial at a time.

Proposition 5 (Buchheim and Klein [4]). *Let $G = (V,E)$ be a graph and $\mathcal{M} = \{M\} \subset 2^E$ and $|M| = 2$. A description of $P_F(\mathcal{M})$ is given by Edmonds' subtour elimination constraints* (2.1)*, McCormick's linearization constraints* (2.8) *and* (2.10) *combined with*

$$x \ge 0 \tag{2.2}$$

$$y \ge 0 \tag{2.12}$$

and the quadratic subtour elimination constraints

$$x(E(S)) + y \le |S| - 1 \qquad \text{for all } S \subset V \text{ with } a, c \in S \text{ and } b \notin S \tag{2.13}$$

for the adjacent case $M = \{\{a,b\},\{b,c\}\}$ *and*

$$x(E(S_1)) + x(E(S_2)) + y \le |S_1| + |S_2| - 2 \tag{2.14}$$

for all $S_1, S_2 \subset V$ *with* $\{a,b\}$ *and* $\{c,d\}$ *both have one end node in* S_1 *and one in* S_2 *for the general case* $M = \{\{a,b\},\{c,d\}\}$.

Regarding Martin's z-variables we observe: If two adjacent edges $\{a,b\}$ and $\{b,c\}$ are contained in a tree, the edge $\{b,c\}$ is directed from b to c in the corresponding a-arborescence as illustrated in Figure 2.2. Thus, we can add

$$y \le z^a_{b,c} \tag{2.15}$$

to our formulation.

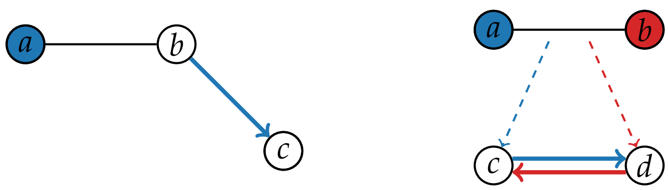


Figure 2.2: Direction of arcs in specific arborescences

In the general case let $T \subseteq E$ be a spanning tree that contains both edges $\{a,b\}$ and $\{c,d\}$. Obviously the direction of the edge $\{c,d\}$ is the same in the corresponding a- and b-arborescences. Thus, if x and z are the characteristic vectors of T and its arborescences, we know with (2.4′)

$$z^a_{c,d} + z^b_{d,c} = z^a_{c,d} + z^a_{d,c} \ge x_{\{c,d\}}.$$

Using the linearization constraint (2.8) we can add

$$y \le z^a_{c,d} + z^b_{d,c} \tag{2.16a}$$
$$y \le z^a_{d,c} + z^b_{c,d} \tag{2.16b}$$

to our formulation.

When we consider spanning trees, where we have the projection constraint (2.4) instead of (2.4′), we observe that (2.15) and (2.4) imply

$$y \le z^a_{b,c} \le x_{b,c}$$

and (2.16) with (2.4) imply

$$\begin{aligned} y &\leq z^a_{c,d} + z^b_{d,c} = x_{c,d} + z^a_{c,d} - z^b_{c,d} \\ y &\leq z^a_{d,c} + z^b_{c,d} = x_{c,d} - z^a_{c,d} + z^b_{c,d}, \end{aligned}$$

that we can combine to

$$y \leq x_{c,d} - |z^a_{c,d} - z^b_{c,d}| \leq x_{c,d}.$$

Altogether, (2.15) or (2.16) respectively imply with (2.4) McCormick's linearization constraint (2.8). Accordingly, we call (2.15) and (2.16) *extended linearization constraints*.

Remark 1. The constraints of the general case do imply the constraints for the adjacent case: Setting $b = d$ in (2.16b) leads to

$$y \leq z^a_{b,c} + z^b_{c,b},$$

which is equal to (2.15) since $z^b_{c,b} = 0$ by (2.6).

Theorem 2 (adjacent case). *Let $G = (V, E)$ be a connected graph and*

$$\mathcal{M} = \{M\} \subset 2^E \qquad \text{with} \qquad M = \big\{\{a,b\},\{b,c\}\big\}.$$

An extended formulation for $P_F(\mathcal{M})$ is given by Martin's inequalities (2.4′), (2.5) - (2.6) *for $u = a$ and McCormick's linearization constraints* (2.8) *and* (2.10) *together with*

$$\begin{aligned} & x \in P_F \\ & y \leq z^a_{b,c} && (2.15) \\ & y \geq 0 && (2.12) \end{aligned}$$

and the coordinate projection onto (x, y).

We will prove generalizations of Theorem 2 in the next theorem and in Section 2.3.

Theorem 3 (general case). *Let $G = (V, E)$ be a connected graph and*

$$\mathcal{M} = \{M\} \subset 2^E \qquad \text{with} \qquad M = \big\{\{a,b\},\{c,d\}\big\}.$$

An extended formulation for $P_F(\mathcal{M})$ *is given by Martin's inequalities* (2.4′), (2.5) - (2.6) *for all* $u \in \{a, b\}$ *and McCormick's linearization constraints* (2.8) *and* (2.10) *together with*

$$
\begin{aligned}
& x \in P_F \\
& y \leq z^a_{c,d} + z^b_{d,c} && \text{(2.16a)} \\
& y \leq z^a_{d,c} + z^b_{c,d} && \text{(2.16b)} \\
& y \geq 0 && \text{(2.12)}
\end{aligned}
$$

and the coordinate projection onto (x, y).

Proof. Let $x = \chi(F)$ for a forest $F \subseteq E$ and $y = x_{a,b}\, x_{c,d}$. We extend F to a spanning tree $T \subseteq E$ with $F \subseteq T$ and choose for all $u \in \{a, b\}$ the variable z^u as the characteristic vector of the u-arborescence induced by T. This choice is obviously valid for the formulation as discussed before.

To prove that $P_F(\mathcal{M})$ is contained in the projection it suffices to show that the constraints in our formulation imply the quadratic subtour elimination constraint (2.14). Using (2.4) and (2.7) from Martin's formulation we obtain for all $i \in \{1, 2\}$ and $u \in V$

$$x\left(E\left(S_i\right)\right) \leq \sum_{v \in S_i} z^u\left(\delta^{in}\left(v\right)\right) - z^u\left(\delta^{in}\left(S_i\right)\right).$$

Choosing u such that $u \in S_i$ we receive with (2.5) and (2.6)

$$x\left(E\left(S_i\right)\right) \leq |S_i| - 1 - z^u\left(\delta^{in}\left(S_i\right)\right) \qquad \text{for all } i \in \{1, 2\}.$$

Assuming without loss of generality $a \in S_1$ and $b \in S_2$ we can combine it to

$$x\left(E\left(S_1\right)\right) + x\left(E\left(S_2\right)\right) \leq |S_1| + |S_2| - 2 - z^a\left(\delta^{in}\left(S_1\right)\right) - z^b\left(\delta^{in}\left(S_2\right)\right).$$

If $c \in S_1$ and $d \in S_2$, we obtain

$$x\left(E\left(S_1\right)\right) + x\left(E\left(S_2\right)\right) \leq |S_1| + |S_2| - 2 - z^a_{d,c} - z^b_{c,d}$$

and if $d \in S_1$ and $c \in S_2$, we obtain

$$x\left(E\left(S_1\right)\right) + x\left(E\left(S_2\right)\right) \leq |S_1| + |S_2| - 2 - z^a_{c,d} - z^b_{d,c},$$

which does with (2.16) imply (2.14). $\square$

Relation to subgraph polytopes

Martin's extended formulation is originally constructed as the dual of a separation problem for Edmonds' subtour elimination constraints.[29] This approach was reviewed by Conforti et al. using non-empty subgraph polytopes.[7]

In this section we construct our new formulations analogously.

The *subgraph polytope* of a graph $G = (V, E)$ is defined as

$$Q_{sub} := \operatorname{conv}\left\{ \left(\chi(D), \chi(S)\right) \in \{0,1\}^E \times \{0,1\}^V \,\middle|\, D \subseteq E(S),\, S \subseteq V \right\}.$$

Proposition 6 (Conforti et al. [7])**.** *The subgraph polytope Q_{sub} is described by*

$$\begin{aligned} \alpha_e - \beta_v &\leq 0 && \textit{for all } v \in V \textit{ and } e \in \delta(v) && (2.17)\\ \beta &\leq 1 && && (2.18)\\ \alpha &\geq 0 && && (2.19) \end{aligned}$$

For disjunct node sets $A, B \subseteq V$ let $Q_{A,B}$ be the face of Q_{sub} that is defined by

$$\begin{aligned} \beta_v &= 1 && \text{for all } v \in A\\ \beta_v &= 0 && \text{for all } v \in B. \end{aligned}$$

Hence, in $Q_{A,B}$ we restrict the node sets $S \subseteq V$ in the definition of Q_{sub} to those sets, where we have $A \subseteq S$ and $B \cap S = \varnothing$.

Adjacent Case

In the adjacent case $M = \big\{\{a,b\},\{b,c\}\big\}$ we consider the superset of the quadratic subtour elimination constraints (2.13)

$$x(D) - |S| \leq -1 - y \quad \text{for all } D \subseteq E(S),\, S \subset V,\, a, c \in S \text{ and } b \notin S. \tag{2.20}$$

Those inequalities are all valid for $P_F(\mathcal{M})$ and can be separated by solving

$$\max\left\{ \sum_{e \in E} \alpha_e x_e - \sum_{v \in V} \beta_v \,\middle|\, (\alpha, \beta) \in Q_{\{a,c\},\{b\}} \right\}. \tag{2.21}$$

If the solution of this optimization problem is less or equal to $-1 - y$, all constraints in (2.20) are fulfilled by x.

Let A be the set of directed arcs corresponding to E with both directions for each edge in E. Applying strong duality to the system (2.20) is equivalent to the existence of $\sigma \in \mathbb{R}^A$ and $\tau \in \mathbb{R}^V$ such that

$$\tau\big(V \setminus \{b\}\big) \leq -1 - y \tag{2.22}$$
$$\sigma_{v,w} + \sigma_{w,v} \geq x_{\{v,w\}} \qquad \text{for all } \{v,w\} \in E \tag{2.23}$$
$$-\sigma\left(\delta^{in}(v)\right) + \tau_v = -1 \qquad \text{for all } v \in V \tag{2.24}$$
$$\sigma \geq 0 \tag{2.25}$$
$$\tau_v \geq 0 \qquad \text{for all } v \in V \setminus \{a,b,c\}. \tag{2.26}$$

To eliminate τ we insert (2.24) in (2.22) and (2.26) and obtain

$$\sum_{v \in V \setminus \{b\}} \sigma\left(\delta^{in}(v)\right) \leq |V| - 2 - y \tag{2.22$'$}$$
$$\sigma\left(\delta^{in}(v)\right) \geq 1 \qquad \text{for all } v \in V \setminus \{a,b,c\}. \tag{2.26$'$}$$

Now we replace σ by z^a (that will turn out to be the same variables that we know from Martin's extended formulation in Proposition 2) via the following relations:

$$\begin{aligned} \sigma_{b,c} &= 0 \\ \sigma_{c,b} &= z^a_{b,c} + z^a_{c,b} \\ \sigma_{v,w} &= z^a_{v,w} \qquad \text{for all } \{v,w\} \in E \setminus \big\{\{b,c\}\big\} \end{aligned}$$

This leads to

$$\sum_{v \in V \setminus \{b\}} z^a\left(\delta^{in}(v)\right) - z^a_{b,c} \leq |V| - 2 - y \tag{2.22$''$}$$
$$z^a_{v,w} + z^a_{w,v} \geq x_{\{v,w\}} \qquad \text{for all } \{v,w\} \in E \tag{2.23$'$}$$
$$z^a\left(\delta^{in}(v)\right) \geq 1 \qquad \text{for all } v \in V \setminus \{a,b,c\} \tag{2.26$''$}$$
$$z^a \geq 0. \tag{2.25$'$}$$

Using z-variables, which are characteristic vectors of arborescences as described after Proposition 2, we can replace (2.26″) by equations and insert $z^a\left(\delta^{in}(v)\right) = 1$ for $v \in V \setminus \{a\}$ and $z^a\left(\delta^{in}(a)\right) = 0$ in (2.22″) to obtain

$$y \leq z^a_{b,c}, \tag{2.22$'''$}$$

which is our extended linearization constraint (2.15).

General Case

In the general case $M = \{\{a,b\},\{c,d\}\}$ we regard the following superset of the quadratic subtour elimination constraint (2.14)

$$x(D_1) + x(D_2) - |S_1| - |S_2| \leq -2 - y \tag{2.27}$$

for all $D_i \subseteq E(S_i)$, $S_1, S_2 \subset V$, $i = 1,2$ such that $\{a,b\}$ and $\{c,d\}$ have one endpoint in S_1 and one in S_2.

This is valid for $P_F(\mathcal{M})$ and can be separated by the two optimization problems

$$\max\left\{ \sum_{e\in E} \alpha_e^1 x_e - \sum_{v\in V} \beta_v^1 + \sum_{e\in E} \alpha_e^2 x_e - \sum_{v\in V} \beta_v^2 \,\middle|\, \left(\alpha^1,\beta^1\right) \in Q_{\{a,c\},\{b,d\}},\ \left(\alpha^2,\beta^2\right) \in Q_{\{b,d\},\{a,c\}} \right\}$$

and

$$\max\left\{ \sum_{e\in E} \alpha_e^3 x_e - \sum_{v\in V} \beta_v^3 + \sum_{e\in E} \alpha_e^4 x_e - \sum_{v\in V} \beta_v^4 \,\middle|\, \left(\alpha^3,\beta^3\right) \in Q_{\{a,d\},\{b,c\}},\ \left(\alpha^4,\beta^4\right) \in Q_{\{b,c\},\{a,d\}} \right\}$$

If the maxima of both problems are less or equal to $-2-y$ all constraints in (2.27) are fulfilled by x.

Thus, using strong duality system (2.27) is equivalent to the existence of $\sigma^k \in \mathbb{R}^A$ and $\tau^k \in \mathbb{R}^V$, $k \in [4]$ (A is the directed version of E as before) with

$$\tau^1(V\setminus\{b,d\}) + \tau^2(V\setminus\{a,c\}) \leq -2 - y \tag{2.28a}$$

$$\tau^3(V\setminus\{b,c\}) + \tau^4(V\setminus\{a,d\}) \leq -2 - y \tag{2.28b}$$

$$\sigma_{v,w}^k + \sigma_{w,v}^k \geq x_{\{v,w\}} \qquad \text{for all } \{v,w\} \in E,\ k \in [4] \tag{2.29}$$

$$-\sigma^k\left(\delta^{in}(v)\right) + \tau_v^k = -1 \qquad \text{for all } v \in V,\ k \in [4] \tag{2.30}$$

$$\sigma^k \geq 0 \qquad \text{for all } k \in [4] \tag{2.31}$$

$$\tau_v^k \geq 0 \qquad \text{for all } v \in V\setminus\{a,b,c,d\},\ k \in [4]. \tag{2.32}$$

We eliminate τ from the system by using equation (2.30) and obtain

$$\sum_{v\in V\setminus\{b,d\}} \sigma^1\left(\delta^{in}(v)\right) + \sum_{v\in V\setminus\{a,c\}} \sigma^2\left(\delta^{in}(v)\right) \leq 2|V| - 6 - y \tag{2.28a'}$$

$$\sum_{v\in V\setminus\{b,c\}} \sigma^3\left(\delta^{in}(v)\right) + \sum_{v\in V\setminus\{a,d\}} \sigma^4\left(\delta^{in}(v)\right) \leq 2|V| - 6 - y \tag{2.28b'}$$

$$\sigma^k\left(\delta^{in}(v)\right) \geq 1 \qquad \text{for all } v \in V \setminus \{a,b,c,d\},\ k \in [4]\,. \tag{2.32'}$$

Now we can replace

$$\begin{aligned}
\sigma^1_{v,w} &= \sigma^3_{v,w} = z^a_{v,w} && \text{for all } \{v,w\} \in E \setminus \big\{\{a,b\},\{c,d\}\big\} \\
\sigma^1_{c,d} &= \sigma^3_{d,c} \ = z^a_{c,d} + z^a_{d,c} \\
\sigma^1_{d,c} &= \sigma^3_{c,d} \ = 0 \\
\sigma^2_{v,w} &= \sigma^4_{v,w} = z^b_{v,w} && \text{for all } \{v,w\} \in E \setminus \big\{\{a,b\},\{c,d\}\big\} \\
\sigma^2_{c,d} &= \sigma^4_{d,c} \ = 0 \\
\sigma^2_{d,c} &= \sigma^4_{c,d} \ = z^b_{c,d} + z^b_{d,c}
\end{aligned}$$

and obtain

$$\sum_{v \in V \setminus \{b,d\}} z^a\left(\delta^{in}(v)\right) - z^a_{d,c} + \sum_{v \in V \setminus \{a,c\}} z^b\left(\delta^{in}(v)\right) - z^b_{c,d} \leq 2|V| - 6 - y \tag{2.28a''}$$

$$\sum_{v \in V \setminus \{b,c\}} z^a\left(\delta^{in}(v)\right) - z^a_{c,d} + \sum_{v \in V \setminus \{a,d\}} z^b\left(\delta^{in}(v)\right) - z^b_{d,c} \leq 2|V| - 6 - y. \tag{2.28b''}$$

Using $z^u\left(\delta^{in}(v)\right) = 1$ and $z^u\left(\delta^{in}(u)\right) = 0$ for all $u, v \in V$ with $u \neq v$ we receive

$$y \leq z^a_{d,c} + z^b_{c,d} \tag{2.28a'''}$$

$$y \leq z^a_{c,d} + z^b_{d,c}, \tag{2.28b'''}$$

which is equal to our extended linearization constraints (2.16).

Formulations for planar graphs

Let $G = (V, E)$ be a *planar graph*, i.e., there exists an embedding in the plane such that no two edges cross each other. For such a plane representation of G we can define the *dual graph* $G_d := (V_d,\ E_d)$, where V_d are the regions defined by E as boundaries and E_d are the dual edges. For each primal edge in E there exists one crossing dual edge that connects two regions. (See Figure 2.3 for an example.)

For the spanning tree polytope P_{ST} corresponding to a planar graph there exists an extended formulation by Williams of size $\Theta(|E|)$ [37]. It is significantly smaller than Martin's formulation of size $\Theta(|V||E|)$.

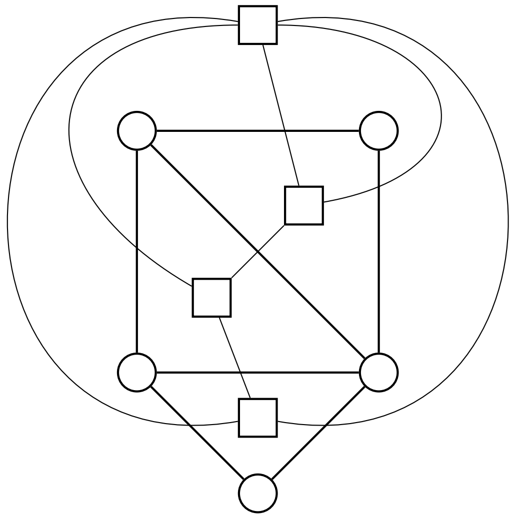


Figure 2.3: Plane embedding of a planar graph and its dual graph

Williams' formulation requires to know a planar embedding of the graph G and thus the dual graph G_d.

We choose root nodes $r \in V$ and $R \in V_d$ such that r is on the boundary of R. Furthermore, we consider directed arcs, where we have both directions for each edge in $E \setminus \{\delta(r)\}$ and $E_d \setminus \{\delta(R)\}$ and the arcs corresponding to the edges in $\delta(r)$ and $\delta(R)$ are directed out of r and R respectively.

For all $e \in E$ and $v \in e$ we denote $z_{e,v}$ as the variable for the directed arc corresponding to e and directed towards v. Furthermore, let $\{I, J\} \in E_d$ be the edge crossing e. We denote $\omega_{e,I}$ as the variable corresponding to the directed dual arc (J, I).

Proposition 7 (Williams [37]). *Let $G = (V, E)$ be a planar graph and $G_d = (V_d, E_d)$ be a dual graph of G corresponding to some embedding. Furthermore let $r \in V$ and $R \in V_d$, such that r is at the boundary of R. An extended formulation for P_{ST} is given by the projection defined by*

$$\operatorname{proj}(z,\omega)_e := \begin{cases} z_{e,v} + z_{e,w} & \text{if } e = \{v,w\} \in E,\ w \neq r \\ z_{e,v} & \text{if } e = \{r,v\} \in E \end{cases}$$

for all $e \in E$,

$$z_{e,v} + z_{e,w} + \omega_{e,I} + \omega_{e,J} = 1 \tag{2.33}$$

for all $e = \{v, w\} \in E$ *with crossing dual edge* $\{I, J\} \in E_d$ *and*

$$z\left(\delta^{in}(v)\right) = 1 \qquad \text{for all } v \in V \setminus \{r\} \tag{2.34}$$

$$\omega\left(\delta^{in}(I)\right) = 1 \qquad \text{for all } I \in V_d \setminus \{R\} \tag{2.35}$$

$$z \geq 0 \tag{2.36}$$

$$\omega \geq 0 \tag{2.37}$$

where $z\left(\delta^{in}(v)\right) = \sum_{e\in\delta(v)} z_{e,v}$, $\omega\left(\delta^{in}(I)\right) = \sum_{e\in\overline{\delta}(I)} z_{e,I}$ *and* $\overline{\delta}(I)$ *are the primal edges in* E *that surround* I.

Using our formulation based on Martin for the higher order spanning tree polytope with one adjacent degree-two monomial from Theorem 2 with Williams' formulation as the description of P_{ST}, we need to add one set of directed arcs leading to $2|E|$ additional inequalities. For the general degree-two case we need $4|E|$ additional inequalities. Balas' formulation in Theorem 1 with Williams' formulation has size $12|E|$. Hence, our formulation based on the extended formulation by Martin leads to a smaller extended formulation than disjunctive programming even in the planar case, where we have this very small extended formulation by Williams. (See also Table 2.1 for a summary of the sizes.)

Formulation based on	adjacent	general
Martin	$6\|E\|$	$8\|E\|$
Balas	$12\|E\|$	$12\|E\|$
Williams	$4\|E\| + 4\|\delta(V_d(b))\|$	—

Table 2.1: Size of extended formulations for $P_{ST}(\mathcal{M})$ with $\mathcal{M} = \{M\}$ and $|M| = 2$ that use Williams' formulation as the description of P_{ST}.

Now we will construct a third formulation that is purely based on Williams and that is smaller than those based on Martin in most cases.

This formulation only works for the adjacent case, where we have $M = \{\{a, b\}, \{b, c\}\}$ for pairwise distinct $\{a, b, c\} \subseteq V$ and we assume $r \notin \{a, b, c\}$.

Consider the graph $\overline{G} := \left(V \setminus \{b\}, \overline{E}\right)$ that we build by removing the node b from G and adding a new edge $e_M = \{a, c\}$ instead as illustrated in Figure 2.4. Hence, $\overline{E} := \left(E \setminus \delta(b)\right) \cup e_M$. It is possible that the edge $\{a, c\}$ is contained in $\overline{G}$ twice. We join the regions in G adjacent to b to $\overline{A}$ on the one side of the path $\{\{a, b\}, \{b, c\}\}$ and $\overline{B}$ on the other

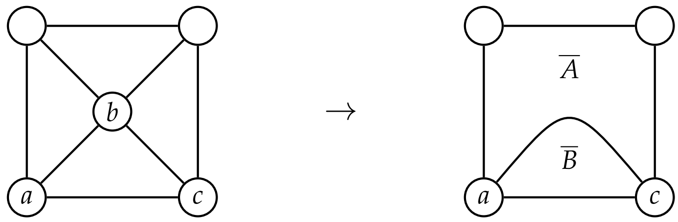


Figure 2.4: Transformation from G to $\overline{G}$

side. We refer to $\overline{A}$ or $\overline{B}$ simultaneously as regions in $\overline{G}$ and sets of regions in G. For $v \in V$ we denote by $V_d(v)$ all regions touching v, e.g., $\overline{A} \cup \overline{B} = V_d(b)$.

Each tree $T \subseteq E$ in G naturally implies a forest $F \subset \overline{E}$ with $\overline{E} \cap T \subset F$ and $e_M \in F$ if and only if $M \subseteq T$. To apply Williams' formulation for this setup, we extend F to a spanning tree $\overline{T} \subset \overline{E}$ using the following lemma.

Lemma 1. *For each spanning tree $T \subseteq E$ with characteristic vector x and $y = x_{a,b}\, x_{b,c}$ we can construct a spanning tree $\overline{T} \subseteq \overline{E}$ with characteristic vector $\overline{x}$ such that*

$$\overline{x}_e = x_e \qquad \text{for all } e \in \overline{E} \setminus \left(\overline{\delta}\left(V_d(b)\right) \cup \{e_M\}\right) \tag{2.38}$$

$$\overline{x}_{e_M} = y \tag{2.39}$$

$$\overline{x}_e \geq x_e \qquad \text{for all } e \in \overline{\delta}\left(V_d(b)\right). \tag{2.40}$$

For the characteristic vectors ω and $\overline{\omega}$ of the corresponding dual R-arborescence it holds furthermore

$$\overline{\omega}_{e,I} = \omega_{e,I} \qquad \text{for all } I \notin V_d(b) \tag{2.41}$$

$$\overline{\omega}_{e,I} \leq \omega_{e,I} \qquad \text{for } e \in \overline{\delta}\left(V_d(b)\right) \text{ and } I \in V_d(b). \tag{2.42}$$

As before $\overline{\delta}\left(V_d(b)\right)$ are the primal edges surrounding $V_d(b)$.

Proof. Let $T \subseteq E$ be a spanning tree in G.

Case 1: If $|\delta(b) \cap T| \leq 1$, we choose $\overline{T} = T \cap \overline{E}$.

Case 2: If $\delta(b) \cap T = M$, we choose $\overline{T} = \left(T \cap \overline{E}\right) \cup \{e_M\}$.

Case 3: Otherwise, we first transform T and construct $\overline{T}$ afterwards. Let $T^0 = T$. As long as we have $|\delta(b) \cap T^i| \geq 2$ and $\delta(b) \cap T^i \neq M$ construct T^{i+1} as illustrated in Figure 2.5:

Choose one edge $e_i \in \delta(b) \cap T$, $e_i \notin M$ with the crossing dual edge $\{I_i\ ,\ J_i\}$. We know that I_i and J_i are connected in the dual tree

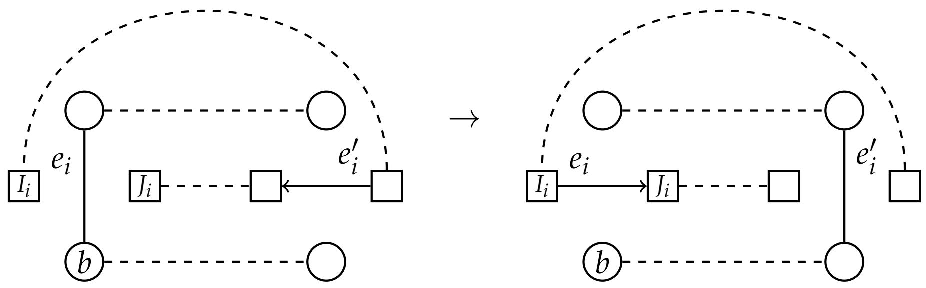


Figure 2.5: Transformation from T^i to T^{i+1}

of T^i. The path between J_i and I_i enters $V_d(b)$with some edge, due to $|\delta(b) \cap T^i| \geq 2$. We call the crossing primal edge e'_i and obtain T^{i+1} by replacing e_i with e'_i, i.e., $T^{i+1} = \left(T^i \setminus \{e_i\}\right) \cup \{e'_i\}$.

After at most $|\delta(b) - 1|$ steps we obtain T^k with $\left|\delta(b) \cap T^k\right| \leq 1$ or $\delta(b) \cap T^k = M$ and can transform it to $\overline{T}$ as in Case 1 or 2.

Our construction affects only edges in $\delta(b)$ and $\delta(V_d(b))$. In $\delta(V_d(b))$ it increase primal edges and decrease dual edges. Furthermore we add e_M to $\overline{T}$ if and only if $M \subseteq T$. Thus, the constraints (2.38)-(2.42) are fulfilled. □

To obtain an extended formulation for $P_{ST}(\mathcal{M})$ we combine Williams' extended formulation from Proposition 7 for G and $\overline{G}$.

Theorem 4. *Let $G = (V, E)$ be a planar graph, $\mathcal{M} = \{M\} \subseteq 2^E$ with $M = \big\{\{a,b\},\{b,c\}\big\}$ and $\overline{G} = (V \setminus \{b\}, \overline{E})$ as described before. An extended formulation for $P_{ST}(\mathcal{M})$ is given by Williams' formulation in Proposition 7 for G and $\overline{G}$ with the corresponding variables z, $\overline{z}$, ω and $\overline{\omega}$ together with*

$$x_e = \begin{cases} z_{e,v} + z_{e,w} & \text{if } e = \{v,w\} \in E,\ w \neq r \\ z_{e,v} & \text{if } e = \{r,v\} \in E \end{cases} \qquad \text{for all } e \in E,$$

the constraints (2.38)-(2.42) *from Lemma 1, McCormick's linearization constraints* (2.8) *and* (2.10) *and the coordinate projection onto (x, y).*

Proof. For each spanning tree in T we find a spanning tree $\overline{T}$ with the properties from Lemma 1. We can choose z, $\overline{z}$, w and $\overline{w}$ as the characteristic vectors of the corresponding r- and R-arborescences.

To verify that the formulation is complete we only have to check, whether the quadratic subtour elimination constraints (2.13) are fulfilled.

If $a, c \in S$ we have $e_M \in \overline{E}(S)$. Using Lemma 1 we obtain for all $S \subset V$ with $a, c \in S$ and $b \notin S$

$$x\left(E(S)\right) + y \leq \overline{x}\big(\overline{E}(S)\big)$$

and due to the fact that $\overline{x} \in P_{ST}$ for the graph $\overline{G}$ we know that Edmonds' rank constraints (2.1) are fulfilled. Hence, we have

$$x\left(E(S)\right) + y \leq \overline{x}\big(\overline{E}(S)\big) \leq |S| - 1. \qquad \square$$

Remark 2. One can rewrite the formulation in Theorem 4 such that it increases Williams' formulation only by $4\big|\delta\left(V_d(b)\right)\big|$ inequalities using equations (2.33) and (2.36).

2.2 Degree-Three Monomials

Let again $G = (V, E)$ be an undirected connected graph and let now $\mathcal{M} = \{M\} \subset 2^E$ with $M = \big\{\{a_i, b_i\} \,|\, i \in [3]\big\}$. As before we write y instead of y_M.

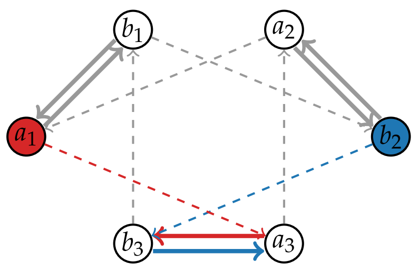


Figure 2.6: Illustration for inequality (2.43)

Consider a tree $T \subseteq E$ in G. If $M \subseteq T$ the edges $\{a_i, b_i\}$ for $i \in [3]$ must be connected. Regarding Figure 2.6 this means exactly two of the possible connections illustrated as dashed lines are part of T.

We consider the sum of the z-variables corresponding to edges in M with both directions for each edge and the root node defined as the source of the dashed arrows in Figure 2.6, e.g., the root node corresponding to (a_3, b_3) is a_1.

Choosing z^u as the characteristic vectors of the u-arborescences induced by T, the z-variable corresponding to a dashed arrow is 1 if there is a path in T connecting the end nodes of this arrow and the corresponding edge e_i is contained in T.

If $M \subseteq T$, two connections are part of the tree T and the sum of z-variables is at least two. Hence, we can add the following inequality to our formulation.

$$2y \leq z^{a_2}_{a_1,b_1} + z^{b_3}_{b_1,a_1} + z^{a_3}_{a_2,b_2} + z^{b_1}_{b_2,a_2} + z^{a_1}_{a_3,b_3} + z^{b_2}_{b_3,a_3}. \tag{2.43}$$

To build a complete description of $P_F(\mathcal{M})$ we need several constraints of this type. Let

$$\tau : [3] \to \{a_i | \, i \in [3]\} \cup \{b_i | \, i \in [3]\}$$

such that $\tau(i) \in \{a_i, b_i\}$ for all $i \in [3]$ and

$$\overline{\tau}(i) := \begin{cases} a_i & \text{if } \tau(i) = b_i \\ b_i & \text{if } \tau(i) = a_i \end{cases}$$

For all such τ we rewrite (2.43) as

$$\begin{aligned} 2y \leq \; & z^{\tau(2)}_{\tau(1),\overline{\tau}(1)} + z^{\overline{\tau}(3)}_{\overline{\tau}(1),\tau(1)} + z^{\tau(3)}_{\tau(2),\overline{\tau}(2)} + z^{\overline{\tau}(1)}_{\overline{\tau}(2),\tau(2)} \\ & + z^{\tau(1)}_{\tau(3),\overline{\tau}(3)} + z^{\overline{\tau}(2)}_{\overline{\tau}(3),\tau(3)}, \end{aligned} \tag{2.44}$$

which leads to 2^3 inequalities.

Proposition 8. *Let $M = \{\{a_i, b_i\} | \, i \in [3]\}$ and let z^u be the characteristic vector of u-arborescences with the same underlying undirected tree $T \subseteq E$ for all $u \in V$, then inequality* (2.44) *is valid for*

$$y = \begin{cases} 1 & \textit{if } M \subseteq T \\ 0 & \textit{otherwise} \end{cases}.$$

Proof. The case $y = 0$ is evident, since $z \geq 0$.

In case $y = 1$, we know that the edges $e_i \in M \subseteq T$ for $i \in [3]$ are connected in T.

Regarding the minimal subtree of T contains M we see that at least two of the end nodes of the edges in M are leafs. For each such leaf ℓ exists an $i \in [3]$ such that we have either $\ell = \tau(i)$ or $\ell = \overline{\tau}(i)$ for all τ that fulfill the requirements of (2.44). Hence, either

$$z^{\overline{\tau}(j)}_{\tau(i),\overline{\tau}(i)} = 1 \quad \text{or} \quad z^{\tau(j)}_{\overline{\tau}(i),\tau(i)} = 1 \qquad \text{for all } j \in [3] \setminus \{i\}.$$

This sums up to the right-hand side of (2.44) being at least two. □

To obtain a complete description of $P_F(\mathcal{M})$ we have to add the constraints we know from one degree-two monomial for all degree-two submonomials of M.

Theorem 5. *Let $G = (V,E)$ be a connected graph and $\mathcal{M} = \{M\}$ with $M = \{\{a_i, b_i\} \,|\, i \in [3]\}$. An extended formulation for $P_F(\mathcal{M})$ is given by Martin's inequalities* (2.4'), (2.5) - (2.6) *for $u \in M$, McCormick's linearization constraints* (2.8) *and* (2.10) *together with*

$$x \in P_F$$

$$y \leq z^{a_i}_{a_j,b_j} + z^{b_i}_{b_j,a_j} \qquad \text{for all } i \in [2],\ j \in [3] \setminus [i] \tag{2.16a}$$

$$y \leq z^{a_i}_{b_j,a_j} + z^{b_i}_{a_j,b_j} \qquad \text{for all } i \in [2],\ j \in [3] \setminus [i] \tag{2.16b}$$

$$\begin{aligned} 2y \leq\ & z^{\tau(2)}_{\tau(1),\overline{\tau}(1)} + z^{\overline{\tau}(3)}_{\overline{\tau}(1),\tau(1)} + z^{\tau(3)}_{\tau(2),\overline{\tau}(2)} + z^{\overline{\tau}(1)}_{\overline{\tau}(2),\tau(2)} \\ & + z^{\tau(1)}_{\tau(3),\overline{\tau}(3)} + z^{\overline{\tau}(2)}_{\overline{\tau}(3),\tau(3)} \qquad \text{for all } \tau, \overline{\tau} \text{ as described above} \end{aligned} \tag{2.44}$$

$$y \geq 0 \tag{2.12}$$

and the coordinate projection onto (x,y).

The proof of Theorem 5 is analog to the proof of Theorem 7 in the next section. We just have to replace y_1 and y_2 by y and use the observation that the coefficient $\alpha_{1,3}$ related to y in the extended rank inequalities (2.45) is $\alpha_{1,2} + \alpha_{3,3}$.

2.3 Nested Monomials

We consider the case that $\mathcal{M}$ consists of several monomials $M_1, \ldots, M_k$. The corresponding polynomial forest problem is hard in general, but if the monomials are nested, i.e., $M_1 \subset M_2 \subset \ldots \subset M_k$, it is solvable in polynomial time in $|V|$. We can see this for example by using Theorem 1 with the extended formulation by Martin to bound the extension complexity of $P_F(\mathcal{M})$ by

$$\mathrm{xc}\left(P_F(\mathcal{M})\right) \leq \left(|M_k| - 1\right) \mathrm{xc}\left(P_F\right) \in O\left(|V|^2|E|\right).$$

We do only consider monomials that are cycle-free and thus it holds $|M_k| \leq |V| - 1$.

Fischer et al. provided a complete description not only of $P_F(\mathcal{M})$ but more general of the higher order matroid polytope with nested monomials $\mathcal{M}$.[15] For this formulation we need an order of the elements in M_k such that $M_i = \{e_1, \ldots, e_{|M_i|}\}$ for all $i \in [k]$.

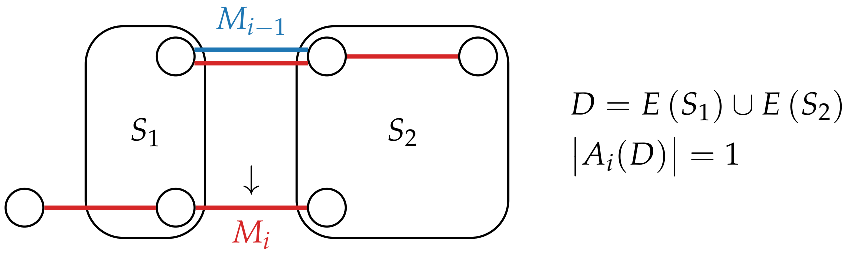


Figure 2.7: An example illustrating $A_i(D)$. The edges that belong to elements in $A_i(D)$ are marked with $\downarrow$.

Throughout the whole section we write y_i instead of y_{M_i}.

Proposition 9 (Fischer, Fischer, McCormick [16]). *The matroid polytope with nested monomials $M_1 \subset M_2 \subset \cdots \subset M_k$ is described by McCormick's linearization* (2.8)-(2.11) *and*

$$x(D) + \sum_{i=1}^{k} \alpha_{k_{i-1}+1,k_i}(D) y_i \le r(D) \quad \textit{for all } D \subseteq E \textit{ with } \operatorname{cl}(D) = D \tag{2.45}$$

$$x \ge 0 \tag{2.2}$$

$$y \ge 0, \tag{2.12}$$

where

$$k_0 := 0, \quad k_i := |M_i| \qquad \textit{for all } i \in [k] \tag{2.46}$$

$$\overline{E}_{k_i} := M_i = \{e_1, \ldots, e_{k_i}\} \qquad \textit{for all } i \in [k] \tag{2.47}$$

$$\overline{E}_m := \{e_1, \ldots, e_m\} \qquad \textit{for all } m \in [|M_k|] \tag{2.48}$$

$$\alpha_{i,j}(D) := \sum_{m=i}^{j} \alpha_m(D) \qquad \textit{for all } j \in [|M_k|],\, i \in [j-1],\, D \subseteq E \tag{2.49}$$

$$\alpha_m(D) := \big|\{e_m\} \setminus D\big| + r\big(D \cup \overline{E}_{m-1}\big) - r\big(D \cup \overline{E}_m\big) \qquad \textit{for all } m \in [|M_k|] \tag{2.50}$$

$$\operatorname{cl}(D) := D \cup \Big\{ e \in E \,\Big|\, r(D) = r\big(D \cup \{e\}\big) \Big\} \tag{2.51}$$

and $r(D)$ describes the rank of D.

The constraints (2.45) are called *extended rank constraints* due to their relation to Edmonds' rank constraints (2.1). If we consider forests and single degree-two monomials, the inequalities (2.45) are equal to the quadratic subtour elmination constraints (2.13) and (2.14).

To understand (2.45) in general imagine that we add the elements $\{e_1, \ldots, e_{|M_k|}\}$ successively to D. Now the coeffiencts $\alpha_{i,j}$ are equal to the number of elements e_m with $e_m \notin D$ and $i < m \leq j$ that do not increase the rank, i.e.,

$$r(D \cup \overline{E}_{m-1}) = r(D \cup \overline{E}_m).$$

In the context of forests we have

$$r(D) = \sum_{\substack{S \subseteq V \\ S \text{ component of } D}} (|S| - 1)$$

and find

$$\alpha_{k_{i-1}+1,k_i}(D) = |A_i(D)| \tag{2.52}$$

with

$$\begin{aligned} A_i(D) := \Big\{ m \in [k_i] \setminus [k_{i-1}] \;\Big|\; & e_m \notin D \text{ and both end nodes of } e_m \\ & \text{are in the same component of } (V, D \cup \overline{E}_{m-1}) \Big\}. \end{aligned} \tag{2.53}$$

An example on how to count the elements in $A_i(D)$ is given in Figure 2.7.

Nested trees

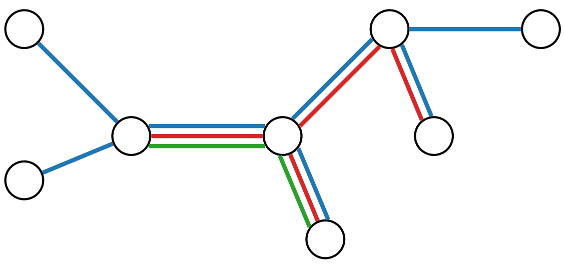


Figure 2.8: Three nested trees

We consider the case that the monomials $M_1, \ldots, M_k$ are trees like in Figure 2.8. This is a generalization of the adjacent case in Section 2.1.

For $D \subseteq E$ we denote by $V(D)$ the nodes in D, i.e.,

$$V(D) := \{v \in V \,|\, v \in e \text{ for some } e \in D\}.$$

For all $i \in [k]$ and for each node $u \in V(M_i)$ and each edge $\{v, w\} \in M_i$ the direction of $\{v, w\}$ in the u-arborescence with underlying undirected tree M_i is defined by the unique u-w-path (or u-v-path respectively) in M_i. If all edges of M_i are contained in a forest we know for the corresponding z-variables that $z^u_{v,w} = 1$ if and only if v lays on the path from u to w in M_i. This leads to a generalized form of inequality (2.15) and furthermore to an extended formulation for $P_F(\mathcal{M})$.

Theorem 6. *Let $G = (V, E)$ be a connected graph and $\mathcal{M} = \{M_1, \dots, M_k\}$ with $M_1 \subset M_2 \subset \dots \subset M_k \subseteq E$ and M_i are trees for all $i \in [k]$. An extended formulation for $P_F(\mathcal{M})$ is given by Martin's inequalities* (2.4′), (2.5)-(2.6) *for $u \in V(M_k)$ and McCormick's inequalities* (2.8)-(2.11) *together with*

$$x \in P_F$$
$$y \geq 0 \tag{2.12}$$
$$y_i \leq z^u_{v,w} \tag{2.15′}$$

for all $i \in [k]$, $u \in V(M_i)$ and $\{v, w\} \in M_i$ where v is on the path from u to w in M_i.

To prove Theorem 6 we will use the complete description by Fischer et al. in Proposition 9.

Without loss of generality we consider the order of the edges in M_k such that each subset $\overline{E}_m$ for $m \in [|M_k|]$ is connected. This order of edges implies an order of the nodes in $V(M_k) =: \{0, \dots, |M_k|\}$ defined by $e_1 = \{0, 1\}$ and $e_m = \{s(m), m\}$, where $s(m) \in V\left(\overline{E}_{m-1}\right)$ is the source and m is the target of e_m for all $m \in [|M_k|]$.

We observe that in our context of forests $\mathrm{cl}(D) = D$ for $D \subseteq E$ is equivalent to the existence of pairwise disjunct $S_1, \dots, S_\ell$ with $S_j \subseteq V$ for $j \in [\ell]$ such that $D = \dot\bigcup_{j=1}^{\ell} E\left(S_j\right)$. Lets assume that the S_j are ordered such that $S_1, \dots, S_{\ell'}$ intersect with $V(M_k)$ and $S_{\ell'+1}, \dots, S_\ell$ and $V(M_k)$ are disjunct.

For all $j \in [\ell']$ we define the first node in $V(M_k)$ that intersects with S_j as

$$f_j := \min\left\{v \in V(M_k) \,\middle|\, v \in S_j\right\}.$$

To obtain a more visual impression imagine e_m to be the first edge that enters S_j. Then, we have $f_j = m$. The only exception to this is the case $0 \in S_j$ where we have $f_j = 0$.

In the next lemma we count the edges in $M_i \setminus M_{i-1}$ entering any S_j after the first one to receive an alternative description of $A_i(D)$. An illustrative example can be found in Figure 2.9.

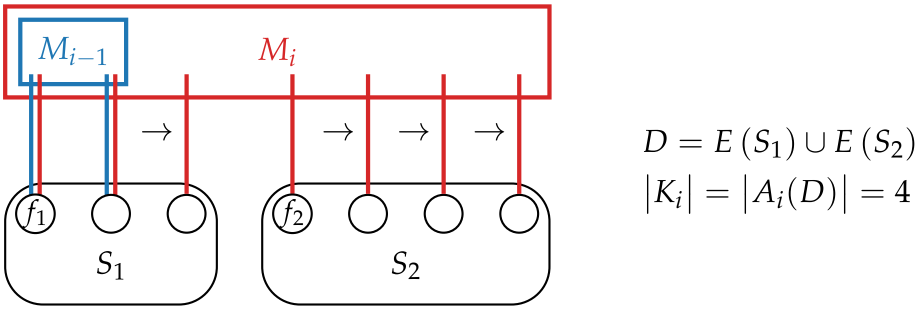


Figure 2.9: An example illustrating K_i. The edges that belong to elements in K_i are marked with $\rightarrow$.

Recall $k_i = |M_i|$.

Lemma 2. *Let $D = \dot\cup_{j=1}^{\ell} E(S_j)$ for each $i \in [k]$ we have $A_i(D) = K_i$ with $A_i(D)$ as defined in* (2.53) *and*

$$K_i := \Big\{ m \in [k_i] \setminus [k_{i-1}] \,\Big|\, \text{there exist a } j \in [\ell'] \text{ with } m \in S_j \setminus \{f_j\} \text{ and } s(m) \notin S_j \Big\}.$$

Proof. For each $m \in [|M_k|]$ we define the component of $(V, D \cup \overline{E}_m)$ that includes $V(\overline{E}_m)$ as

$$U_m = V(\overline{E}_m) \cup \bigcup_{j \in [\ell'], f_j \leq m} S_j.$$

In order to show the inclusion $A_i(D) \subseteq K_i$ let m be in $A_i(D)$. Hence, the end nodes of $e_m = \{s(m), m\}$ are in the same component of the subgraph $(V, D \cup \overline{E}_{m-1})$, which is U_{m-1}. Since $e_m \notin D$ and $m \notin V(\overline{E}_{m-1})$, there exists $j \in [\ell']$ such that $m \in S_j$. Furthermore, $f_j \leq m-1$ and thus $m \neq f_j$. With $e_m \notin D$ we see that $s(m) \notin S_j$ and thus $m \in K_i$.

To establish the reverse inclusion $K_i \subseteq A_i(D)$ choose now $m \in K_i$ and $j \in [\ell']$ such that $m \in S_j \setminus \{f_j\}$. Clearly we have $s(m) \in U_{m-1}$. Due to $m \neq f_j$ (with $m \in S_j$ this implies $m > f_j$) we also have $S_j \subseteq U_{m-1}$ and thus $m \in U_{m-1}$. This means both end nodes of e_m are in the same component of $(V, D \cup \overline{E}_{m-1})$ and thus $m \in A_i(D)$. $\square$

Lemma 3. *The constraints in Theorem 6 imply*

$$x(D) + \sum_{i=1}^{k} |K_i| y_i \leq r(D) \qquad \text{for all } D \subseteq E \text{ with } \mathrm{cl}(D) = D$$

with K_i as defined in Lemma 2.

Proof. With Martin's constraints (2.4) and (2.5) we obtain

$$x\left(E\left(S_j\right)\right) = \underbrace{\sum_{v \in S_j} z^{f_j}\left(\delta^{in}\left(v\right)\right)}_{=\left|S_j\right|-1} - z^{f_j}\left(\delta^{in}\left(S_j\right)\right) \qquad \text{for all } j \in \left[\ell'\right]$$

and due to $x \in P_F$ we have

$$x\left(E\left(S_j\right)\right) \leq \left|S_j\right| - 1 \qquad \text{for all } j \in [\ell] \setminus \left[\ell'\right].$$

Combining this we obtain

$$\begin{aligned} x\left(D\right) &= \sum_{j=1}^{\ell'} x\left(E\left(S_j\right)\right) + \sum_{j=\ell'+1}^{\ell} x\left(E\left(S_j\right)\right) \\ &= r(D) - \sum_{j=1}^{\ell'} z^{f_j}\left(\delta^{in}\left(S_j\right)\right) \end{aligned}$$

We observe $\bigcup_{i=1}^{k} K_i \subseteq \bigcup_{j=1}^{\ell'} S_j$ and for all $m \in K_i$ exists $j(m) \in [\ell']$ with $m \in S_{j(m)}$. Furthermore, we have $\{s(m), m\} \in \delta^{in}\left(S_{j(m)}\right)$ and $s(m)$ lays on the path from $f_{j(m)}$ to m in M_k. Hence, we can apply inequality (2.15′) to obtain

$$\sum_{i=1}^{k} |K_i| y_i = \sum_{i=1}^{k} \sum_{m \in K_i} y_i \leq \sum_{i=1}^{k} \sum_{m \in K_i} z_{s(m),m}^{f_{j(m)}} \leq \sum_{j=1}^{\ell'} z^{f_j}\left(\delta^{in}\left(S_j\right)\right).$$

□

To complete the proof of Theorem 6 let P be the projection of the polytope described by Theorem 6 onto (x, y).

For each forest F we can construct the vectors z as characteristic vectors of arborescences induced by a spanning tree including F. Those vectors together with $x = \chi(F)$ and $y_i = \prod_{e \in M_i} x_i$ for $i \in [k]$ fulfill the constraints in Theorem 6 as described before and thus we have $P_F(\mathcal{M}) \subseteq P$.

The inverse inclusion $P \subseteq P_F(\mathcal{M})$ follows directly from Lemma 2 and Lemma 3 with the description by Fischer et al. in Proposition 9.

Nested monomials up to degree-three

In the case of general nested monomials, we restrict to monomials of degree less or equal to 3, i.e., $\mathcal{M} = \{M_1, M_2\}$ with $M_1 \subset M_2$ and $|M_2| \leq 3$. The cases with $M_1 = \varnothing$ are covered in the former sections, where we consider single degree-two monomials (Section 2.1) and single degree-three monomials (Section 2.2). Hence, the remaining case is $|M_1| = 2$ and $|M_2| = 3$.

Theorem 7. *Let $G = (V, E)$ be a connected graph, $\mathcal{M} = \{M_1, M_2\} \subset 2^E$ with $M_1 = \{e_1, e_2\}$, $M_2 = \{e_1, e_2, e_3\}$ and $e_i = \{a_i, b_i\}$ for $i \in [3]$. An extended formulation for $P_F(\mathcal{M})$ is given by Martin's constraints* (2.4'),(2.5)-(2.7)*, McCormick's linearization* (2.8)-(2.11) *and*

$$x \in P_F$$

$$y_k \leq z^{a_i}_{a_j,b_j} + z^{b_i}_{b_j,a_j} \tag{2.16a}$$

$$y_k \leq z^{a_i}_{b_j,a_j} + z^{b_i}_{a_j,b_j} \tag{2.16b}$$

for all $(k, i, j) \in \big\{(1, 1, 2), (2, 1, 3), (2, 2, 3)\big\}$

$$\begin{aligned} 2y_2 \leq \quad & z^{\tau(2)}_{\tau(1),\overline{\tau}(1)} + z^{\overline{\tau}(3)}_{\overline{\tau}(1),\tau(1)} + z^{\tau(3)}_{\tau(2),\overline{\tau}(2)} + z^{\overline{\tau}(1)}_{\overline{\tau}(2),\tau(2)} \\ & + z^{\tau(1)}_{\tau(3),\overline{\tau}(3)} + z^{\overline{\tau}(2)}_{\overline{\tau}(3),\tau(3)} \end{aligned} \tag{2.44}$$

for all τ and $\overline{\tau}$ as introduced before Theorem 5 and

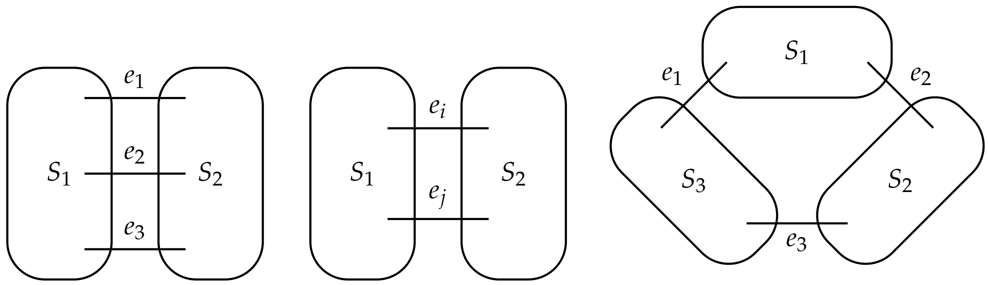
$$y \geq 0. \tag{2.12}$$

To apply the formulation by Fischer et.al in Proposition 9 we need to understand the cases where the coefficients $\alpha_{k_{i-1}+1,k_i}$ of y_i in (2.45) are nonzero. In the current setting we have $k_1 = 2$ and $k_2 = 3$ and are interested in the coefficient $\alpha_{1,2}$ and $\alpha_{3,3}$.

As before $\mathrm{cl}(D) = D$ is equivalent to the existence of pairwise disjunct $S_1, \ldots, S_\ell$ such that $D = \dot\bigcup_{j=1}^{\ell} E\left(S_j\right)$.

Using the correlation $\alpha_{k_{i-1}+1,k_i} = \left|A_i(D)\right|$ in (2.52) we have $\alpha_{1,2} \neq 0$ if and only if $e_2 \notin D$ and the end nodes of e_2 are in the same component in $\big(V, D \cup \{e_1\}\big)$ as in Figure 2.10 (i) and (ii) for $i = 1$ and $j = 2$. Furthermore, $\alpha_{3,3} \neq 0$ if and only if $e_3 \notin D$ and the end nodes of e_3 are in the same component of $\big(V, D \cup \{e_1, e_2\}\big)$ as in Figure 2.10 (i), (ii) (for $i \in \{1, 2\}$ and $j = 3$) and (iii).

All in all, we can combine it to the three cases illustrated in Figure 2.10:



(i) $\alpha_{1,2} = 1$ and $\alpha_{3,3} = 1$ (ii) $\alpha_{1,2} + \alpha_{3,3} = 1$ (iii) $\alpha_{1,2} = 0$ and $\alpha_{3,3} = 1$

Figure 2.10: The three cases where $\alpha_{1,2}$ or $\alpha_{3,3}$ are nonzero

(i) In this case we have $\{e_1, e_2, e_3\} \cap D = \varnothing$ and the end nodes of all three edges are in two of the sets S_j. Here we have $\alpha_{1,2} = 1$ and $\alpha_{3,3} = 1$.

(ii) Consider only two edges $\{e_i, e_j\} \cap D = \varnothing$ for $i \in \{1,2\}$, $j \in \{2,3\}$ and $i < j$. The end nodes of those edges are in two of the set S_j. The remaining edge should be somewhere else, such that we do not have Case (i). Here we have either $\alpha_{1,2} = 1$ (if $j = 2$) or $\alpha_{3,3} = 1$ (if $j = 3$).

(iii) As in Case (i) we have $\{e_1, e_2, e_3\} \cap D = \varnothing$. Now e_1, e_2, e_3 are included in one cycle in $D \cup \{e_1, e_2, e_3\}$ and we have $\alpha_{1,2} = 0$ and $\alpha_{3,3} = 1$.

Lemma 4. *The formulation in Theorem 7 implies the extended rank inequalities* (2.45).

Proof. If $\alpha_{1,2}(D) = \alpha_{3,3}(D) = 0$ inequality (2.45) is a combination of Edmond's rank constraints (2.1) and thus fulfilled by $x \in P_F$.

If on the other hand $\alpha_{1,2}(D)$ or $\alpha_{3,3}(D)$ are nonzero we have one of the three cases discussed before.

Case (i): Let

$$\overline{D} = E(S_1) \dot{\cup} E(S_2)$$

and without loss of generality

$$a_i = e_i \cap S_1 \qquad b_i = e_i \cap S_2.$$

Using Martin's constraints (2.4),(2.5),(2.7) together with our extended linearization constraints (2.16b) we obtain

$$x(\overline{D}) \leq \overbrace{\sum_{v \in S_1} z^{a_1}\left(\delta^{in}(v)\right)}^{=|S_1|-1} + \overbrace{\sum_{v \in S_2} z^{b_1}\left(\delta^{in}(v)\right)}^{=|S_2|-1}$$
$$\underbrace{-z^{a_1}_{b_2,a_2} - z^{b_1}_{a_2,b_2}}_{\leq -y_1} \underbrace{-z^{a_1}_{b_3,a_3} - z^{b_1}_{a_3,b_3}}_{\leq -y_2}.$$

Due to $r(D) = \sum_{j=1}^{\ell}\left(|S_j| - 1\right)$ and $x(D) = x(D \setminus \overline{D}) + x(\overline{D})$ this combines to

$$x(D) + y_1 + y_2 \leq r(D).$$

Case (ii): The proof is analog to the proof for single degree-two monomials in Theorem 3.

Case (iii): Let

$$\overline{D} = \dot\bigcup_{j=1}^{3} E\left(S_j\right).$$

We choose τ such that

$$\overline{\tau}(i) := e_i \cap S_i \qquad \text{for } i \in \{1, 2, 3\}.$$

Using Martin's constraints (2.4), (2.5) and (2.7) we obtain

$$x(\overline{D}) \leq +\frac{1}{2}\sum_{j=1}^{3} \overbrace{\sum_{v \in S_j}\left(z^{\overline{\tau}(j)}\left(\delta^{in}(v)\right) + z^{\tau(j+1 \mod 3)}\left(\delta^{in}(v)\right)\right)}^{=2\left(|S_j|-1\right)}$$
$$-\frac{1}{2}\left(z^{\tau(2)}_{\tau(1),\overline{\tau}(1)} + z^{\overline{\tau}(3)}_{\overline{\tau}(1),\tau(1)} + z^{\tau(3)}_{\tau(2),\overline{\tau}(2)} + z^{\overline{\tau}(1)}_{\overline{\tau}(2),\tau(2)} \right.$$
$$\left. + z^{\tau(1)}_{\tau(3),\overline{\tau}(3)} + z^{\overline{\tau}(2)}_{\overline{\tau}(3),\tau(3)}\right).$$

Now with $r(D) = \sum_{j=1}^{\ell}\left(|S_j| - 1\right)$, $x(D) = x(D \setminus \overline{D}) + x(\overline{D})$ and inequality (2.44) we receive

$$x(D) + y_2 \leq r(D) \qquad \square$$

In order to complete the proof of Theorem 7 let P be the projection of the polytope described by Theorem 7 onto (x, y).

For each forest F we can construct the vectors z as characteristic vectors of arborescences induced by a spanning tree including F. Those vectors together with $x = \chi(F)$ and $y_i = \prod_{e \in M_i} x_i$ for $i \in [k]$ fulfill the constraints in Theorem 7 as described before and thus we have $P_F(\mathcal{M}) \subseteq P$.

The inverse inclusion $P \subseteq P_F(\mathcal{M})$ follows directly from Lemma 4 with the description by Fischer et al. in Proposition 9.

3 Relaxations and Facets of Quadratic Forest Polytopes

We investigate our formulations for higher order forest polytopes with one degree-two monomial (Theorem 2 and 3) to point out further properties and correlations.

3.1 Improving the One Quadratic Term Technique

Part of our motivation to investigate higher order polytopes was the QMST-problem and the related polytopes.

Let

$$\mathcal{M} = \left\{ M \in 2^E \middle| \, |M| = 2 \right\}. \tag{3.1}$$

The polytope $P_{QF} := P_F(\mathcal{M})$ is called the quadratic forest polytope.

Buchheim and Klein used their description of $P_F(\{M\})$ for $M \in \mathcal{M}$ to build a relaxation of P_{QF} defined by

$$\mathcal{R}(\mathcal{M}) := \left\{ (x,y) \in \mathbb{R}^E \times \mathbb{R}^{\mathcal{M}} \middle| \, (x, y_M) \in P_F(\{M\}), \, M \in \mathcal{M} \right\}. \tag{3.2}$$

[4]

For all $M \in \mathcal{M}$ let $Q(M)$ be the extension of $P_F(\{M\})$ defined as in Theorem 2 and 3. For simplicity we use Martin's extended formulation (Proposition 3) as description of P_F. We define

$$\mathcal{R}'(\mathcal{M}) := \left\{ (x,y,z) \in \mathbb{R}^n \middle| \, (x, y_M, z) \in Q(M), \, M \in \mathcal{M} \right\} \tag{3.3}$$

where $n = |E| + |\mathcal{M}| + 2|V||E|$. By identifying the z-variables used in our descriptions of $Q(M)$ we model a new relation between the monomials in $\mathcal{M}$ that improves the relaxation in the following way.

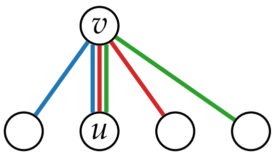
Theorem 8. *Let $\mathcal{M}$, $\mathcal{R}(\mathcal{M})$ and $\mathcal{R}'(\mathcal{M})$ be defined as before in* (3.1), (3.2) *and* (3.3). *Furthermore, let* proj *be the coordinate projection onto (x,y). Then*

$$\operatorname{proj}\left(\mathcal{R}'(\mathcal{M})\right) \subsetneq \mathcal{R}(\mathcal{M})$$

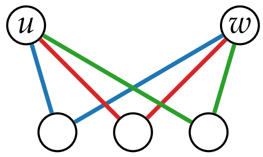
holds for some graphs $G = (V,E)$.

To prove Theorem 8 we will present new inequalities that we obtain by projecting $\mathcal{R}'(\mathcal{M})$ onto (x, y). Using simple examples we will show that those are not valid for $\mathcal{R}(\mathcal{M})$.

The first inequalities that we introduce only use adjacent monomials and describe facets of the *adjacent quadratic forest polytope*, i.e. the higher order forest polytope with the monomial set consisting of all adjacent degree-two monomials. We will prove it in Section 3.2.



(i) Edge case $S = V \setminus \{v\}$



(ii) Edge case $S = \{u, w\}$

Figure 3.1: Edge pairs that appear in Inequality (3.4)

For an adjacent monomial $M = \{\{a, b\}, \{b, c\}\}$ we define the feet of M as $f(M) := \{a, c\}$.

Let $S \subsetneq V, u \in S$ and

$$\mathcal{M}_S^u := \left\{ M \in \mathcal{M} \middle| M \text{ is adjacent, } f(M) = S \cap V(M) \text{ and } u \in f(M) \right\}.$$

We observe

$$\mathcal{M}_S^u = \left\{ \{\{u, j\}, \{j, i\}\} \in \mathcal{M} \middle| i \in S \text{ and } j \in V \setminus S \right\}.$$

Using Martin's constraint (2.4′), (2.5), (2.6) for all $i \in S$ and our extended linearization constraints (2.15) we obtain

$$\begin{aligned} & x(E(S)) + \sum_{M \in \mathcal{M}_S^u} y_M \\ & = \sum_{i \in S} \Bigg(\underbrace{z^u \left(\delta^{in}(i) \right)}_{=1 \text{ for } i \neq u \text{ and } 0 \text{ for } i = u} + \sum_{\substack{j \in V \setminus S \\ \{\{u,j\},\{j,i\}\} \in \mathcal{M}_S^u}} \Big(\underbrace{-z_{j,i}^u + y_{\{\{u,j\},\{j,i\}\}}}_{\leq 0} \Big) \Bigg) \\ & \leq |S| - 1 \end{aligned} \tag{3.4}$$

The edge pairs for the edge cases $|S| = |V| - 1$ and $|S| = 2$ are illustrated in Figure 3.1.

The following two examples show that (3.4) ist not among the description of $\mathcal{R}(\mathcal{M})$.

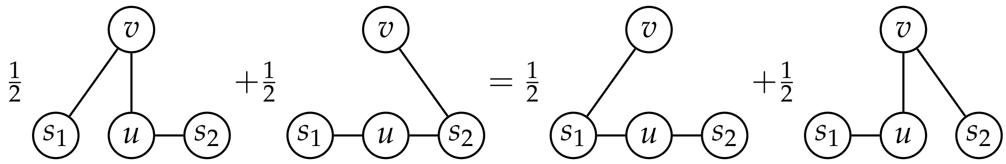


Figure 3.2: Illustration of the convex combinations in Example 1

Example 1. Consider x as the vector corresponding to the convex combinations of spanning trees in Figure 3.2. Let $M_i = \{\{u,v\},\{v,s_i\}\}$ for $i \in [2]$. Regarding the convex combinations in Figure 3.2 it is obvious that $(x, \frac{1}{2}) \in P(\{M_i\})$ for $i \in [2]$. Hence, $(x,y) \in \mathcal{R}(\mathcal{M})$ for $y_1 = y_2 = \frac{1}{2}$, but inserting the same values in (3.4) for $S = \{u, s_1, s_2\}$ we obtain

$$x(E(S)) + y_1 + y_2 = 2\frac{1}{2} > 2 = |S| - 1.$$

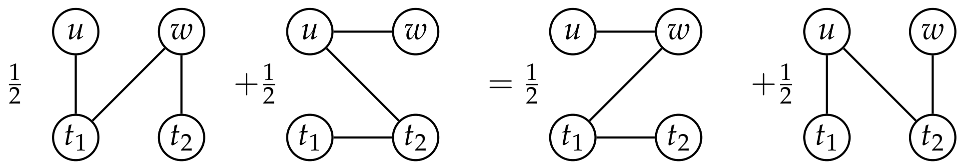


Figure 3.3: Illustration of the convex combinations in Example 2

Example 2. Consider x as the vector corresponding to the convex combinations of spanning trees in Figure 3.3. Let $M_i = \{\{u,t_i\},\{w,t_i\}\}$ for $i \in [2]$. Regarding the convex combinations in Figure 3.3 it is obvious that $(x, \frac{1}{2}) \in P(\{M_i\})$ for $i \in [2]$. Hence, $(x,y) \in \mathcal{R}(\mathcal{M})$ for $y_1 = y_2 = \frac{1}{2}$, but inserting the same values in (3.4) for $S = \{u, w\}$ we obtain

$$x(E(S)) + y_1 + y_2 = 1\frac{1}{2} > 1 = |S| - 1.$$

The derivation of (3.4) works the same if we consider spanning trees instead of forests. The next constraint only arises form our formulation for spanning trees, although it is also valid for forests. In this case we use (2.4) from Martin's formulation for spanning trees in Proposition 2.

For pairwise different $u, v, w \in V$ we add (2.15) for the monomials $\{\{u,v\},\{v,w\}\}$ and $\{\{u,v\},\{u,w\}\}$ (illustrated in Figure 3.4) and use (2.6) and (2.4) to obtain

$$y_{\{\{u,v\},\{v,w\}\}} + y_{\{\{u,v\},\{u,w\}\}} \leq z^w_{v,u} + z^w_{u,v} = x_{\{u,v\}} \tag{3.5}$$

The following example shows that (3.5) is not among the description of $\mathcal{R}(\mathcal{M})$.

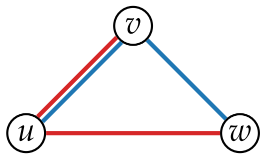


Figure 3.4: Edge pairs appearing in Inequality (3.5)

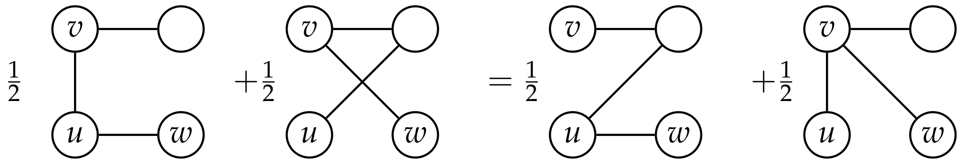


Figure 3.5: Illustration of the convex combinations in Example 3

Example 3. Consider x as the vector corresponding to the convex combinations of spanning trees in Figure 3.5. Let $M_1 = \{\{u,v\},\{v,w\}\}$ and $M_2 = \{\{u,v\},\{u,w\}\}$. Regarding the convex combinations in Figure 3.5 it is obvious that $(x, \frac{1}{2}) \in P(\{M_i\})$ for $i \in [2]$. Hence, $(x,y) \in \mathcal{R}(\mathcal{M})$ for $y_1 = y_2 = \frac{1}{2}$, but inserting the same values in (3.5) we obtain

$$y_1 + y_2 = 1 > \frac{1}{2} = x_{\{u,v\}}.$$

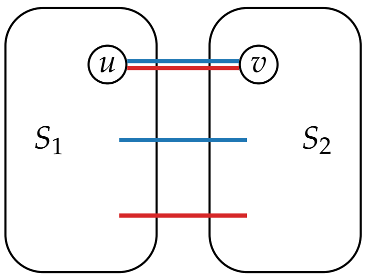


Figure 3.6: Edge pairs appearing in Inequality (3.6)

In order to derive an inequality using also nonadjacent monomials, let $S_1, S_2 \subset V$ with $S_1 \cap S_2 = \emptyset$ and $u \in S_1$, $v \in S_2$ with $\{u,v\} \in E$. Combining (2.4'),(2.5)-(2.7) and (2.16) for edge pairs consisting of $\{u,v\}$

and any other edge between S_1 and S_2 like in Figure 3.6 we obtain

$$\begin{aligned}
&x(E(S_1)) + x(E(S_2)) + \sum_{\substack{i \in S_1 \setminus \{u\}, j \in S_2 \setminus \{v\} \\ \text{with } \{i,j\} \in E}} y_{\{\{u,v\},\{i,j\}\}} \\
&\leq \sum_{i \in S_1} \underbrace{z^u\left(\delta^{in}(i)\right)}_{=1 \text{ for } i \neq u \text{ and } 0 \text{ for } i=u} + \sum_{j \in S_2} \underbrace{z^v\left(\delta^{in}(j)\right)}_{=1 \text{ for } i \neq v \text{ and } 0 \text{ for } i=v} \\
&\quad + \sum_{\substack{i \in S_1 \setminus \{u\}, j \in S_2 \setminus \{v\} \\ \text{with } \{i,j\} \in E}} \underbrace{\left(y_{\{\{u,v\},\{i,j\}\}} - z^u_{j,i} - z^v_{i,j}\right)}_{\leq 0} \\
&\leq |S_1| + |S_2| - 2.
\end{aligned} \tag{3.6}$$

The following example shows that (3.6) is not among the description of $\mathcal{R}(\mathcal{M})$.

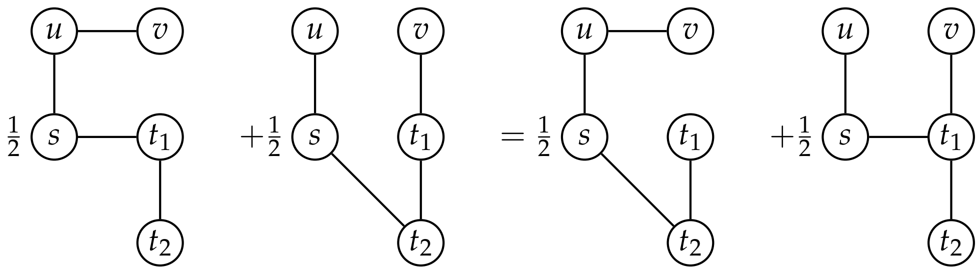


Figure 3.7: Illustration of the convex combinations in Example 4

Example 4. Consider x as the vector corresponding to the convex combinations of spanning trees in Figure 3.7. Let $M_i = \big\{\{u,v\},\{s,t_i\}\big\}$ for $i \in [2]$. Regarding the convex combinations in Figure 3.7 it is obvious that $(x, \frac{1}{2}) \in P\big(\{M_i\}\big)$ for $i \in [2]$. Hence, $(x,y) \in \mathcal{R}(\mathcal{M})$ for $y_1 = y_2 = \frac{1}{2}$. Inserting the values in (3.6) with $S_1 = \{u,s\}$ and $S_2 = \{v,t_1,t_2\}$ we obtain

$$x\big(E(S_1)\big) + x\big(E(S_2)\big) + y_1 + y_2 = 3\frac{1}{2} > 3 = |S_1| + |S_2| - 2.$$

Altogether, we see that the combination of our extended formulations with the one quadratic term technique leads to a better relaxation.

3.2 The Adjacent Quadratic Forest Polytope

Let $G = (V,E)$ be a complete graph and

$$\mathcal{M} = \left\{ M \in 2^E \,\middle|\, |M| = 2 \text{ and the edges in } M \text{ are adjacent} \right\}$$

be the set of all adjacent degree-two monomials. We investigate the adjacent quadratic forest polytope $P_{AQF} := P_F(\mathcal{M})$ and its face the adjacent quadratic spanning tree polytope $P_{AQST} := P_{ST}(\mathcal{M})$.

In contrast to the quadratic spanning tree polytope (including also nonadjacent monomials) the dimension of P_{AQST} is $n-1$ (where n is the full dimension) and the affine hull is described by $x(E) = |V| - 1$. (See [28, Proposition 11] and [32, Corollary 1].)

Recently, Pereira and da Cunha showed that the inequalities (3.4) for $|S| = 2$ as well as the inequalities (3.5) induce facets of P_{AQST} (for $|V| \geq 6$) [32]. Inequality (3.5) also describes a facet of the quadratic forest polytope P_{QF} as proved by Lee and Leung before [28].

Our inequalities (3.4) are generalizations of one facet class described by Pereira and da Cunha as well as of the quadratic subtour elimination constraints (2.13). Hence, the question arises how far we can generalize this type of constraint.

Question 1. *Which inequalities of the form*

$$x(E(S)) + \sum_{M \in \mathcal{M}(S)} y_M \leq |S| - 1 \qquad \textit{for } S \subseteq V \textit{ and } \mathcal{M}(S) \subseteq \mathcal{M}$$

are valid for P_{AQF}?

Let $f(M) := \{a, c\}$ be the foots of $M = \{\{a,b\},\{b,c\}\}$ and $h(M) := b$ the head. Since $x(E(S)) = |S| - 1$ as soon as the corresponding forest is connected in S we do only consider $\mathcal{M}(S)$ with:

- $f(M) \subseteq S$ for all $M \in \mathcal{M}(S)$ and
- $h(M) \notin S$ for all $M \in \mathcal{M}(S)$.

This way, the monomials play the role of one edge in $E(S)$, such that if a monomial is part of a forest $F \subseteq E$, the foots must lay in different components of $\left(S, F \cap E(S)\right)$ and thus $x(E(S)) \leq |S| - 2$.

Furthermore, we have to fulfill $\sum_{M \in \mathcal{M}(S)} y_M \leq |S| - 1$ for all forests $F \subseteq E$. Therefore, we define the graph $G_{\mathcal{M}(S)} := \left(S, E_{\mathcal{M}(S)}\right)$ with

$$E_{\mathcal{M}(S)} := \left\{ f(M) \,\middle|\, M \in \mathcal{M}(S) \right\}.$$

For all forests $F \subseteq E$ we define

$$E_{\mathcal{M}(S)}(F) := \left\{ f(M) \in E_{\mathcal{M}(S)} \,\middle|\, M \subseteq F \right\}.$$

Now, we see that $y_M = 1$ if and only if $f(M) \in E_{\mathcal{M}(S)}(F)$ for the corresponding forest $F \subseteq E$. Hence, $\sum_{M\in\mathcal{M}(S)} y_M \le |S| - 1$ if and only if $E_{\mathcal{M}(S)}(F)$ is cycle free.

All in all, we can answer our question above with the following result.

Proposition 10. *Let $G = (V, E)$ be a graph, $S \subset V$ and $\mathcal{M}(S)$ be a set of adjacent degree-two monomials with $f(M) = S \cap V(M)$ for all $M \in \mathcal{M}(S)$, such that $E_{\mathcal{M}(S)}(F)$ as defined before is cycle-free for all forests $F \subseteq E$, then*

$$x(E(S)) + \sum_{M\in\mathcal{M}(S)} y_M \le |S| - 1$$

is valid for P_{AQF}.

To find new facets of P_{AQF} we are interested in sets $\mathcal{M}(S)$, which are maximal in the sense that adding any further monomial would hurt the discussed properties. One class of those sets was given by the definition of $\mathcal{M}^u_S$ in Section 3.1. (See also Theorem 9.)

There we have $f(M) = S \cap V(M)$ for all monomials $M \in \mathcal{M}^u_S$ by definition. Assume now that we have a cycle in

$$E_{\mathcal{M}^u_S} \subseteq \big\{\{u, v\}\big|\, v \in S\big\}.$$

Since all monomials in $\mathcal{M}^u_S$ have u as one of their foots, the monomials corresponding to the cycle have the same two foots and build a cycle in E. Thus, they can not be part of the same forest and $E_{\mathcal{M}^u_S}(F)$ is cycle free for all forests $F \subseteq E$.

We leave the classification of further $\mathcal{M}(S)$ with the discussed properties open for further research.

New facets

We will show that (3.4) for any $S \subset V$ with $2 \le |S| \le |V| - 1$ describes a facet of P_{AQST} and P_{AQF}.

Theorem 9. *Let $G = (V, E)$ be a complete graph with $|V| \ge 5$ and let $S \subsetneq V$ with $|S| \ge 2$. The face $\mathcal{F}$ of P_{AQST} given by*

$$x(E(S)) + \sum_{M\in M^u_S} y_M = |S| - 1, \tag{3.7}$$

where $M^u_S := \big\{M \in \mathcal{M}\big|\, f(M) = S \cap V(M) \text{ and } u \in f(M)\big\}$ is a facet for all $u \in S$.

To prove Theorem 9 let

$$ax + by = c \tag{3.8}$$

be valid for all (x, y) that lay in $\mathcal{F}$. Our strategy is to add and subtract equation (3.8) for higher order characteristic vectors $(x, y) \in \mathcal{F}$ of trees, which are equal except for a few edges. This way, we obtain step by step more information about a and b.

Lemma 5.

$$b_M = 0 \qquad \text{for all } M \in \mathcal{M} \setminus \mathcal{M}_S^u \tag{3.9}$$

Proof. We distinguish three cases.

Case 1: In the case $f(M) \not\subseteq S$, we consider two monomials $\{e_1, e_2\}$ and $\{e_3, e_4\}$ with the same foots, where at least one foot is not contained in S. We want to add and subtract (3.8) for trees that differ only in the edges e_1, e_2, e_3 and e_5 like in Figure 3.8.

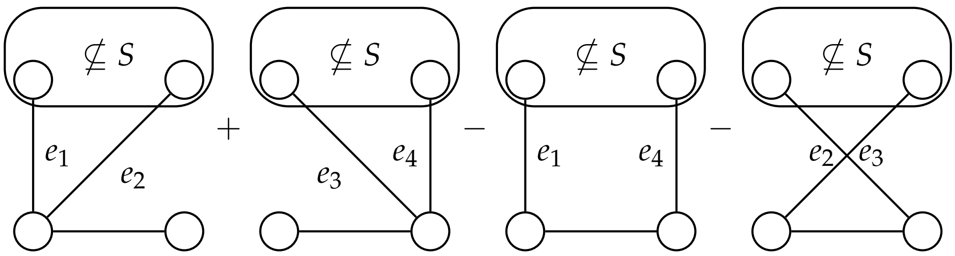


Figure 3.8

If $u \in e_1$ and $e_1 \in E(S)$, we choose the trees such that they are connected in S and we have $x(E(S)) = |S| - 1$ for the first and third tree and such that they have two components belonging to the end nodes of e_1 in S leading to $x(E(S)) = |S| - 2$ for the second and fourth tree. In the other cases we can choose the trees connected in S with $x(E(S)) = |S| - 1$ for all four trees. This way, we can choose the trees such that $(x, y) \in \mathcal{F}$ if:

- Both foots are not in S.
- Only the foot belonging to e_1 is in S and:
  - Both heads are in S.
  - None of the heads is in S.
  - The node u is in e_1.

In all this cases we obtain

$$b_{\{e_1,e_2\}} + b_{\{e_3,e_4\}} = 0.$$

Since we consider a complete graph with $|V| \geq 5$ we can find for any pair of foots that are not both in S three monomials M_1,M_2 and M_3 of that kind and obtain

$$b_{M_1} = -b_{M_2} = b_{M_3} = -b_{M_1}$$

which implies

$$b_M = 0 \qquad \text{for all } M \in \mathcal{M} \text{ with } f(M) \not\subseteq S.$$

Case 2: In the case $f(M) \subset S$ and $u \notin f(M)$ the addition and subtraction of (3.8) for trees like in Figure 3.9 leads to

$$b_{\{e_1,e_2\}} + b_{\{e_3,e_4\}} = 0.$$

Using trees like in Figure 3.10 we obtain

$$b_{\{e_1,e_2\}} + b_{\{e_5,e_6\}} = 0 \quad \text{and analog} \quad b_{\{e_3,e_4\}} + b_{\{e_5,e_6\}} = 0,$$

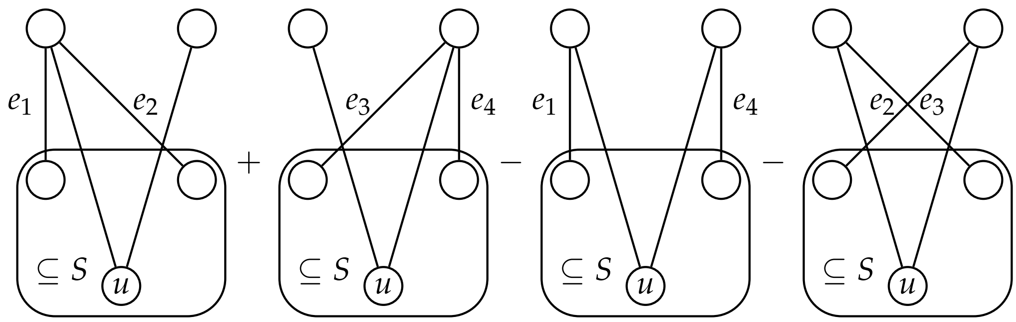


Figure 3.9

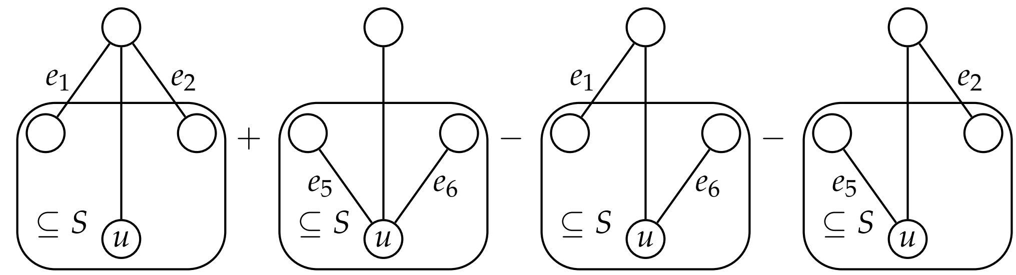


Figure 3.10

which we combine to

$$b_{\{e_1,e_2\}} = -b_{\{e_3,e_4\}} = b_{\{e_5,e_6\}} = -b_{\{e_1,e_2\}}.$$

This implies

$$b_M = 0 \qquad \text{for all } M \in \mathcal{M} \text{ with } f(M) \subset S \text{ and } u \notin f(M).$$

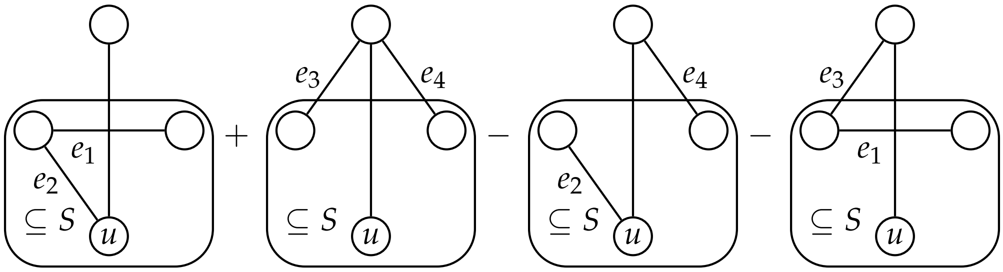


Figure 3.11

Case 3: In the case $M \subset E(S)$ with $u \in f(M)$ we consider trees like in Figure 3.11 and see

$$b_{\{e_1,e_2\}} + b_{\{e_3,e_4\}} = 0,$$

which does with Case 2 lead to

$$b_M = 0 \qquad \text{for all } M \in \mathcal{M} \text{ with } M \subset E(S).$$

□

Lemma 6. *It exist constants $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}$ such that*

$$a_e = \alpha \qquad \text{for all } e \in E \setminus E(S) \tag{3.10}$$

$$a_e = \beta \qquad \text{for all } e \in E(S) \tag{3.11}$$

Proof. We subtract (3.8) for two trees that differ only in one edge as in Figure 3.12

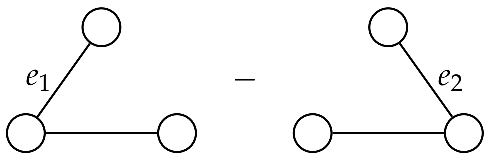


Figure 3.12

Case 1: In the case $\{e_1, e_2\} \subset E(V \setminus S)$ the exchange of e_1 and e_2 does not affect any monomial in $\mathcal{M}_S^u$ and we obtain

$$a_{e_1} - a_{e_2} = 0.$$

Hence, we can find $\alpha \in \mathbb{R}$ such that

$$a_e = \alpha \qquad \text{for all } e \in E(V \setminus S).$$

Case 2: In the case $\{e_1, e_2\} \subset \delta(S)$ we choose the rest of the spanning trees such that $x(E(S)) = |S| - 1$, which means there is no monomial $M \in \mathcal{M}_S^u$ that is also contained in one of the trees. Hence, we obtain for some $\gamma \in \mathbb{R}$

$$a_e = \gamma \qquad \text{for all } e \in \delta(S).$$

Case 3: In the case $e_1 \in \delta(S)$ and $e_2 \in E(V \setminus S)$ we can choose the spanning trees, such that no monomial $M \in \mathcal{M}_S^u$ is contained in one of the trees. Hence, we have

$$\alpha = \gamma$$

and can combine the former two cases to (3.10)

Case 4: Consider the case $\{e_1, e_2\} \subset E(S)$. As in Case 1 the exchange of e_1 and e_2 does not affect any monomial in $\mathcal{M}_S^u$ and we can find $\beta \in \mathbb{R}$ such that

$$a_e = \beta \qquad \text{for all } e \in E(S).$$

□

Lemma 7. *It exists a constant $\delta \in \mathbb{R}$ such that*

$$b_M = \delta \qquad \text{for all } M \in \mathcal{M}_S^u. \tag{3.12}$$

Proof. In order to prove this lemma we distinguish two cases:

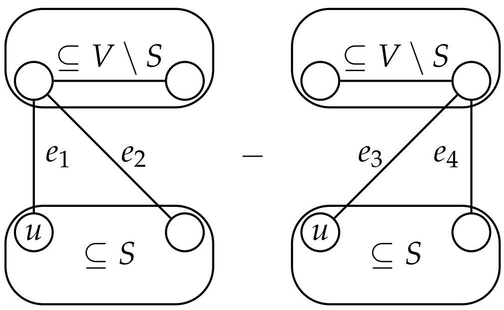


Figure 3.13

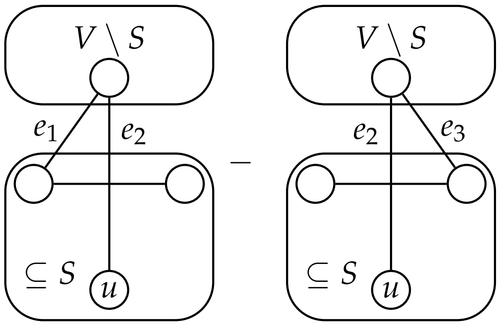


Figure 3.14

Case 1: If $|V \setminus S| \geq 2$, we consider two trees as in Figure 3.13 and obtain

$$a_{e_1} + a_{e_2} + b_{\{e_1,e_2\}} - a_{e_3} - a_{e_4} - b_{\{e_3,e_4\}} = 0.$$

Hence, with Lemma 6 we can find a $\delta \in \mathbb{R}$ such that (3.12) holds.

Case 2: If $|V \setminus S| = 1$ and $|S| \geq 3$, we consider two trees like in Figure 3.14 and obtain

$$a_{e_1} + b_{\{e_1,e_2\}} - a_{e_3} - b_{\{e_2,e_3\}} = 0.$$

Thus, with Lemma 6 we can find $\delta \in \mathbb{R}$ such that (3.12) holds. □

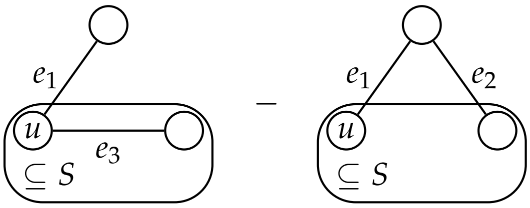


Figure 3.15

To build a relation between α,β and δ we consider trees like in Figure 3.15 and obtain

$$a_{e_3} - a_{e_2} - b_{\{e_1,e_2\}} = 0.,$$

which due to Lemma 6 and Lemma 7 implies

$$\beta = \alpha + \delta.$$

Altogether each valid equation for $\mathcal{F}$ can be written as

$$\alpha\Big(x(E)\Big) + \delta\Big(x\left(E\left(S\right)\right) + \sum_{M \in \mathcal{M}_S^u} y_M\Big) = c = \alpha\big(|V| - 1\big) + \delta\big(|S| - 1\big)$$

for some $\alpha \in \mathbb{R}$ and $\delta \in \mathbb{R}$. This completes the proof of Theorem 9.

Proposition 11. *Let $G = (V, E)$ be a complete graph with $|V| \geq 5$ and let $S \subsetneq V$ with $|S| \geq 2$. The face $\mathcal{F}$ of P_{AQF} defined by*

$$x\left(E\left(S\right)\right) + \sum_{M \in M_S^u} y_M \leq |S| - 1 \qquad \text{for all } u \in S, \tag{3.7}$$

where $M_S^u := \left\{ M \in \mathcal{M} \middle| f(M) = S \cap V(M) \text{ and } u \in f(M) \right\}$ is a facet.

Proof. Consider a forest $F_1 \subseteq E\left(S\right)$ that is connected in S. As a second forest we choose $F_2 = F_1 \cup \{e\}$ for some $e \in \delta\left(S\right)$. It is easy to see that the corresponding higher order characteristic vectors of F_1 and F_2 both lay in $\mathcal{F}$.

From the proof of Theorem 9 we know that each valid inequality for $\mathcal{F}$ can be written as

$$\alpha\Big(x(E)\Big) + \delta\Big(x\left(E\left(S\right)\right) + \sum_{M \in \mathcal{M}_S^u} y_M\Big) = \alpha\big(|V| - 1\big) + \delta\big(|S| - 1\big)$$

for some $\alpha \in \mathbb{R}$ and $\delta \in \mathbb{R}$.

The difference of this equation for the higher order characteristic vectors of F_1 and F_2 leads to

$$\alpha = 0,$$

which means that all equations valid for $\mathcal{F}$ are multiples of (3.7). □

4 Solving the QMST-Problem with Integer Programming

For a graph $G = (V, E)$ we consider the set of all degree-two monomials

$$\mathcal{M} = \{M \in 2^E \mid |M| = 2\}.$$

Optimization over $P_{QST} := P_{ST}(\mathcal{M})$ solves the QMST-problem, which is known to be NP-hard even for accordion graphs (special class of planar graphs) [9]. For an earlier and general hardness result see [1].

In Section 3.1 we used our extended formulations for $P_{ST}(\{M\})$ for $M \in \mathcal{M}$ to improve the one quadratic term technique by strengthening the relaxation $\mathcal{R}(\mathcal{M})$ of P_{QST}. To get a better idea of the amount of improvement we will compare those relaxations with computational experiments. Our main interest lays in the relative gap between the integer solution and the solution of the LP relaxation called *root gap*.

Another benefit of the extended formulations was the reduced size. The original descriptions of P_{QST} had $\Theta(2^{|E|})$ many inequalities whereas the size of the extended formulations is polynomial in $|E|$.

Separation routines decide for a given point x^* whether x^* is in a polytope and if not provide a violated constraint. Instead of adding all constraints at the start of the optimization, we can use separation routines to find and add only those constraints that are needed during the optimization process.

Due to the size, the inequalities of the extended formulations can be separated simply by enumeration in polynomial time in $|E|$. For the formulations in the original space enumeration would be very slow, but there exist other separation routines running in polynomial time in $|E|$. Hence, it is not clear which approach performs better in practice, the small extended formulations or the large formulations in the original space with fast separation routines. In our experiments we will measure different values to discuss this question. Therefore, the author implemented the separation routines and routines to build the extended formulations in Gurobi's Python API [20][17].

4.1 Introduction to Branch and Cut Solver

Before we discuss our implementation and experiments regarding the QMST-problem we will introduce the *branch and cut algorithm* and the

Algorithm 1 Branch and bound (minimization)

$\gamma_u \leftarrow \inf$
$\mathcal{L} \leftarrow \{\emptyset\}$ ▷ set of branch and bound nodes
while $\mathcal{L} \neq \emptyset$ **do**
    Select $\mathcal{N} \in \mathcal{L}$ ▷ select branch and bound node
    $\mathcal{L} \leftarrow \mathcal{L} \setminus \{\mathcal{N}\}$
    Solve LP relaxation of $\mathcal{N}$
    if LP relaxation of $\mathcal{N}$ is feasible **then**
        $x^* \leftarrow$ solution of LP relaxation
        if $x^* \in \mathbb{Z}^n$ and $\gamma_u > \langle c, x^* \rangle$ **then**
            $\gamma_u \leftarrow \langle c, x^* \rangle$ ▷ new upper bound
            $x^I \leftarrow x^*$ ▷ new MIP incumbent
        else if $\gamma_u > \langle c, x^* \rangle$ **then** ▷ continue branching
            choose i with $x_i^* \notin \mathbb{Z}$
            $\mathcal{L} \leftarrow \mathcal{L} \cup \left\{ \mathcal{N} \cup \{x_i \leq \lfloor x_i^* \rfloor\} \right\}$
            $\mathcal{L} \leftarrow \mathcal{L} \cup \left\{ \mathcal{N} \cup \{x_i \geq \lceil x_i^* \rceil\} \right\}$
        end if
    end if
end while
if $\gamma_u < \inf$ **then**
    return x^I
else
    return no integer solution found
end if

general concepts of modern solvers that use it.

An *integer program (IP)* has the canonical form

$$\min \left\{ \langle c, x \rangle \;\middle|\; Ax \leq b,\, x \geq 0,\, x \in \mathbb{Z}^n \right\}$$

for $c \in \mathbb{Q}^n$ and $\langle c, x \rangle := \sum_{i=1}^n c_i x_i$. If only some of the variables are constrained to be integral we call it a *mixed integer program (MIP)*. The *LP relaxation* of it is the corresponding linear program (LP) were we drop all integrality constraints.

In 1958 Ralph Gomory proposed a *cutting plane method* to solve MIPs [19]. It first solves the LP relaxation and then adds constraints to cut of non-integer points. Shortly afterwards in 1960 Land and Doig proposed a *branch and bound algorithm* [27]. A simple version can be found in Algorithm 1.

For a minimization problem it finds upper bounds by solving LP relaxations. There are different branching strategies to build the tree of *branch and bound nodes*. One popular strategy is to choose an index i were the entry x_i^* of the solution in the current node is not integral. Now we know that each integer solution is either less or equal to x_i^* rounded down ($\lfloor x_o^* \rfloor$) or greater or equal to x_i^* rounded up ($\lceil x_i^* \rceil$). Hence, we add the node with the constraint $x_i \leq \lfloor x_i^* \rfloor$ and the node with the constraint $x_i \geq \lceil x_i^* \rceil$ to the branch and bound tree as we can see in Algorithm 1 line 13-15.

If a branch and bound node has an integer solution this provides an upper bound to the optimization problem and there is no need for further branching in this node. Nodes with a larger solution than the best known upper bound or where the LP relaxation is infeasible can also be pruned.

To find better upper bounds and thus reduce the number of branch and bound nodes one can use cutting planes to solve the LP relaxations. This combination of the cutting plane method and branch and bound is called *branch and cut*. Modern branch and cut solvers additionally use heuristics to find further integer solutions that improve the upper bound.

At each time point the best known mixed integer solution x^I is called *MIP incumbent* and additionally to the upper bound $\gamma_u = \langle c, x^I \rangle$ we have a lower bound defined as the minimum of the objectives $\langle c, x \rangle$ over all current leaf nodes.

For a more detailed introduction into integer programming and the branch and cut algorithm we refer to [6].

By now branch and cut is the most popular algorithm in mixed integer programming and implemented in several solvers like *CPLEX*, *Gurobi* and *SCIP*.

The author decided to use Gurobi, due to its clear documentation and the easy to use Python API. According to Gurobi's benchmarks, it is the fastest available MIP-solver [21].

Most modern branch and cut solvers allow to intervene in the behaviour of the algorithm via callback functions. A *callback* is a routine that is called by the solver at specific points during the branch and cut algorithm. In the following we explain the most important callbacks as they are defined in Gurobi [20]. The names and usages in other solvers are very similar.

Lazy constraints are used if the number of constraints is very large. Instead of adding all constraints a-priori, one adds only the violated ones during the optimization process. The separation can be done in a callback. To verify that the solution is correct one should check the constraints every time a new MIP incumbent is found. Additionally it is possible to add them also for continuous solutions.

User cuts are used to strengthen the LP relaxation by cutting of non-integer points. Contrary to lazy constraints they are not allowed to cut of integer solutions. User cuts are add when a continuous solution is found.

Heuristics can be used to find a starting MIP incumbent or to improve the current best incumbent.

4.2 Implementation

Formulations

The author implemented a python module that provides routines that can be used to build different MIP formulations for the QMST-problem using Gurobi. The formulations that we will compare are:

`martin` Martin's extended formulation (Proposition 2 with linearization constraints (2.8)-(2.10)).

`aq-m` Martin's extended formulation and our extended linearization constraints for all *adjacent* degree-two monomials (Proposition 2 with (2.15) and (2.10) for adjacent monomials and (2.8),(2.10) for nonadjacent monomials).

`q-m` Martin's extended formulation and our extended linearization constraints for *all* degree-two monomials (Proposition 2 with (2.15) for all adjacent monomials, (2.16) for all nonadjacent monomials and (2.10) for all monomials).

`sub` Subtour elimination constraints (Proposition 1 with linearization constraints (2.8) and (2.10)).

`aq-sub` Subtour elimination constraints with the quadratic ones for adjacent monomials (Proposition 1 with (2.13), (2.8) and (2.10)).

`q-sub` Subtour elimination constraints with the quadratic ones for all monomials (Proposition 1 with (2.13)-(2.14), (2.8) and (2.10).

Furthermore, we add

$$\sum_{M \in \mathcal{M}} y_M = \binom{|V|-1}{2}$$

for those test instances that contain all possible monomials as Buchheim and Klein did in [4]. They observed a "positive impact on bounds".

Laziness

To build the extended formulations (`martin`, `aq-m` and `q-m`) the author wrote several routines to add different sets of inequalities that can be combined to build the formulations. We distinguish four sets of constraints:

`lin` McCormick's linearization constraints (2.8) and (2.10).

`ef` Martin's extended formulation (Proposition 2).

`adj` Our extended linearization constraints for adjacent degree-two monomials (2.15).

`nonadj` Our extended linearization constraints for general degree-two monomials (2.16) that we use for nonadjacent monomials here.

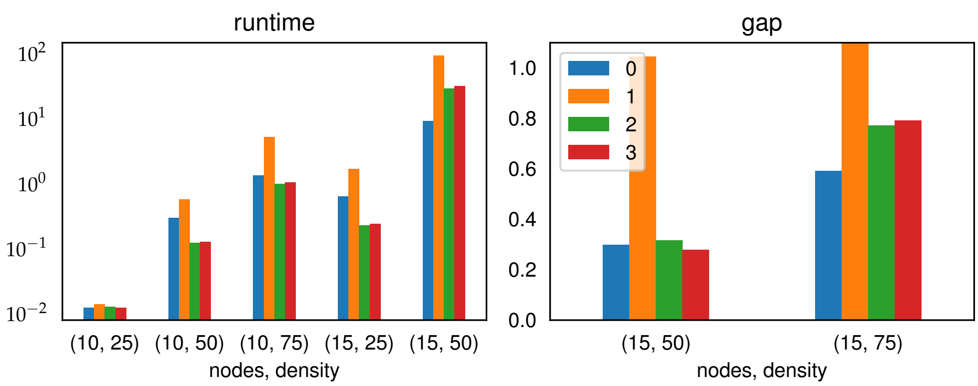


Figure 4.1: Test results for different lazy parameters for `lin` using the model `martin` with the lazy parameter 0 for `ef`

Since we only have polynomially many inequalities in all four models, we can separate them by enumeration. Anyway it might be advantageous to treat them as lazy constraints instead of adding them all a priori

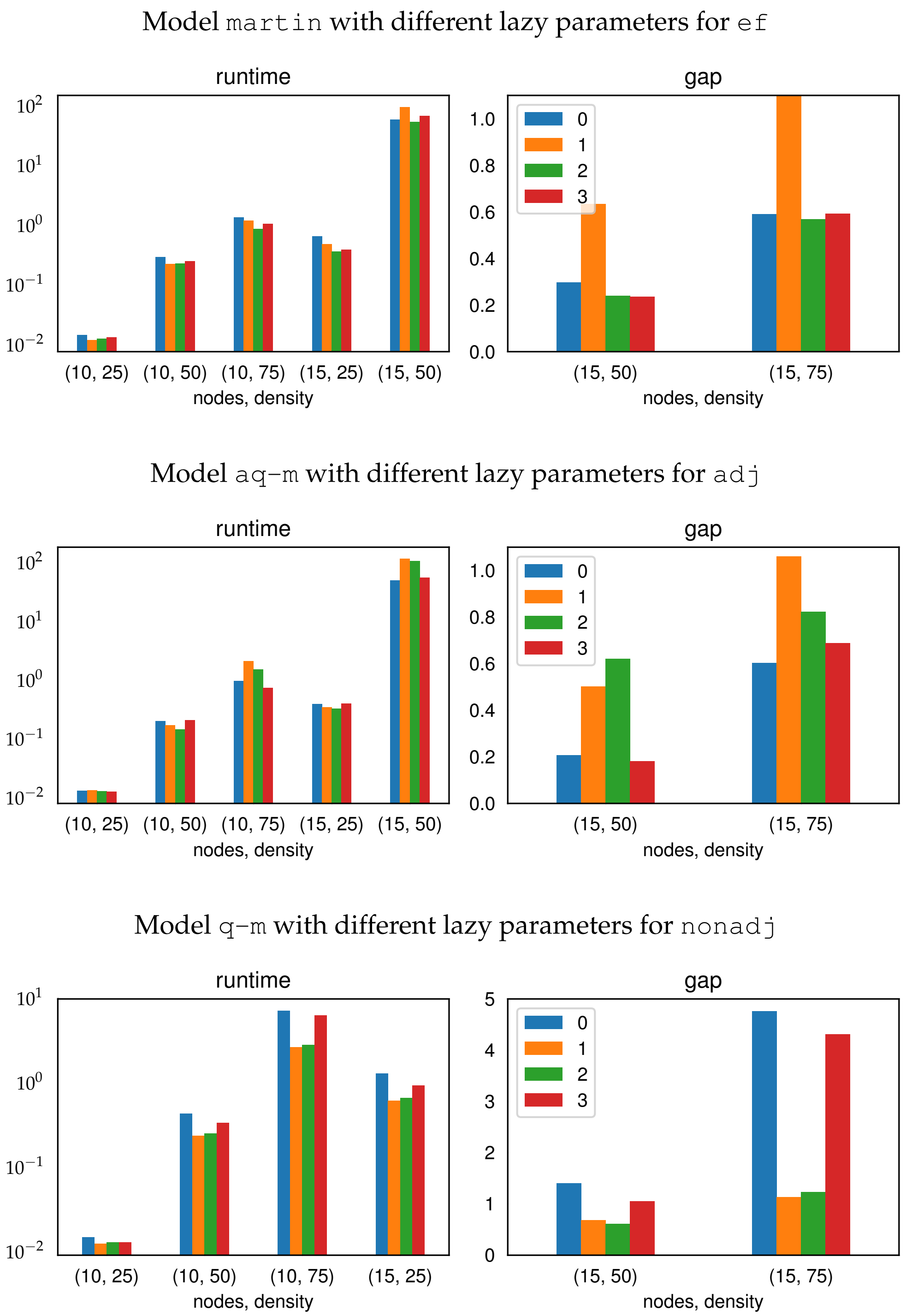


Figure 4.2: Test results for different lazy parameters (the remaining constraints use the default parameters see Table 4.1)

to the model. Gurobi offers the possibility to set a lazy parameter to a constraint. One can choose between 0 (not lazy) and 1-3 (lazy). A higher value refers to a more aggressive pulling in of the constraint. This way, we do not have to implement the enumeration ourself, but leave it to Gurobi.

To choose the default value we consider some little experiments using our random test instances with all possible monomials, which we will describe later. Each instance was run twice with a time limit of 180s and the plots refer to the run with the minimal runtime. This was done because the author observed peeks in the runtime on the same instances, that occur rarely and could affect the results considering those small values. The results presented in the next figures show the mean over ten instances for the gap and over those of the ten instances that were solved with all parameters in the time limit for the runtime.

The first test used the formulation `martin` with lazy parameter 0 for `ef` and variable lazy parameter for `lin` to find the best value regarding the linearization constraints. (Figure 4.1) Since the parameter 0 provides the best results for the larger instances it was chosen as the default for `lin`.

The next tests used this value and varied the lazy parameter for `ef` in the model `martin`. (Figure 4.2) Here the best results were obtained with the parameter 2, which was then chosen as the default one for the remaining tests.

Continuing this workflow `aq-m` was run to find the best lazy parameter for `adj` and `q-m` to obtain the lazy parameter for `nonadj`. (Figure 4.2)

You can find a summary of the default lazy parameters that were finally chosen in Table 4.1.

constraints	lazy parameter
`lin` (2.8) and (2.10)	0
`ef` (2.4)-(2.7)	2
`adj` (2.15)	0
`nonadj` (2.16)	1

Table 4.1: Default values of the lazy parameter

Separation strategies

The subtour elimination constraints (2.1) can be separated by solving specific max flow problems. This method is based on the work of Dantzig et al. [10] (See also [34, Section 58.2].) Buchheim and Klein proposed similar separation routines for the quadratic subtour elimination constraints [4].

To implement those routines the author used the powerful python modul *graph-tool* [31], which includes implementations of max flow and min cut algorithms.

Since all routines work on the same network graph, this graph was added as an attribute to the Gurobi model. This way, we only have to create the separation graph once before we start the optimization. In the callbacks it remains to set the capacities as described in [4].

To separate the (linear) subtour elimination constraints (2.1) the algorithm enumerates over all nodes in V to assure that the resulting sets $S \subset V$ are not empty. For each node the max flow routine finds the set $S \subset V$ with $v \in S$ that provides the most violated inequality among (2.1). We have to decide between several possibilities on how to handle the number of added constraints in our callback routine. Therefore, different strategies were tested, namely:

`node` Add all violated constraints that were found. (There are at most $|V|$.)

`one` Add only the first violated constraint that was found and stop looking for further constraints.

`ord` Add only the first violated constraint and move the node to the end of the list of nodes to start the search using the other nodes in the next call.

`most` Enumerate over all nodes and add only the most violated constraint.

To make sure that we fulfill all constraints we have to add the inequalities every time we found a new MIP incumbent. Additionally, it is possibly to add them also for continues solutions as in the following strategy.

`ip-cut` Add constraints also for continuous solutions. (We only add one per call as in strategy `one`.)

Figure 4.3 presents the mean of the runtime and gap over the ten instances from our test set with all possible degree-two monomials; for the runtime we calculated the mean only over the solved instances. You can see in the first plot (Model `sub`) that adding the constraints too often as

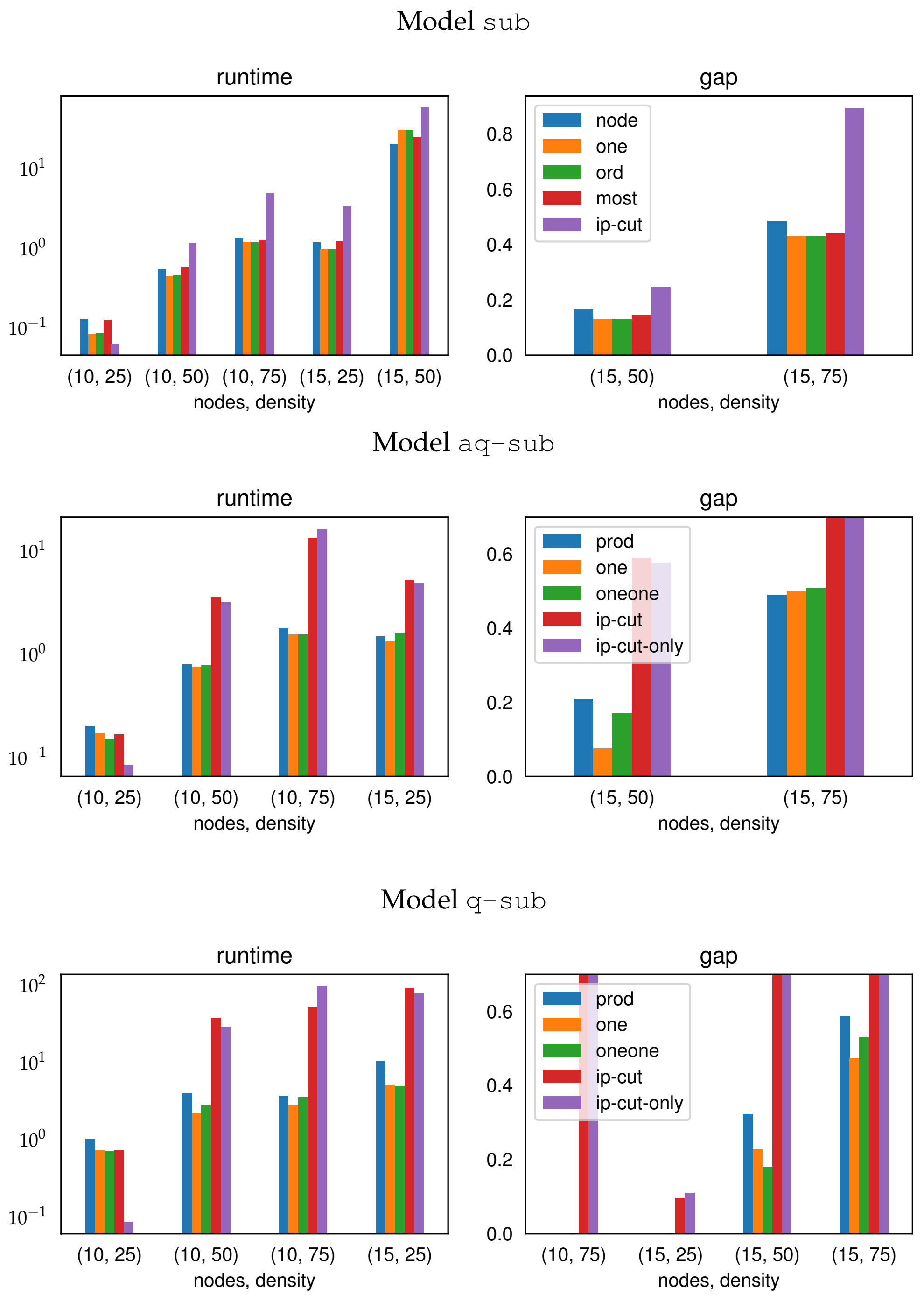


Figure 4.3: Comparison of different separation strategies

in `ip-cut` increases the runtime and gap significantly. Furthermore, it seems to be better to stop the enumeration, when we found a constraint as we did in `one` and `ord`. Those two have nearly the same runtime and gap. The author decided to use the more simple method `one`, where we do not modify the iteration list.

Regarding the quadratic subtour elimination constraints (2.13) and (2.14) we have the option to only add the constraints for adjacent monomials. First we decide on our strategy for only adjacent monomials and then for adjacent and nonadjacent monomials together. The callback searches for violated quadratic subtour elimination constraints first and calls the callback for the linear constraints afterwards. To find violated constraints the algorithm enumerates over all monomials. For the case of only adjacent monomials as well as the case including all monomials we compare the different strategies:

`prod` Add violated constraints for each product.

`one` Add only the first violated constraint that was found and stop the enumeration.

`oneone` Do not look for (linear) subtour elimination constraints, if a quadratic one was found.

`ip-cut-only` Use the quadratic subtour elimination constraints only as cutting planes where we add `one` per call (this is possible since `sub` alone is already a complete MIP formulation).

`ip-cut` Use the constraints as lazy constraints and also as cutting planes. (We only add one per call as in strategy `one`.)

As you can see in Figure 4.3 it turned out to only add the first violated quadratic subtour elimination constraint and look for violated subtour elimination constraints afterwards every time a new MIP incumbent is found (`one`) is prevailing the fastest method. So we use this method in the following experiments.

Test instances

Former computational experiments regarding the QMST-problem usually use random generated instances. (See, e.g., [1],[4], [32].) We used the instances by Cordone and Passeri [8]. [1] They are generated randomly and are split into two sets. The smaller one has between 10 and 30 nodes and the bigger one has between 35 and 50 nodes. There are instances for edge density 33%, 67% and 100% with different ranges for the linear and quadratic cost. The cost functions are all positive. The following tests run on the instances with less or equal to 20 nodes. Even the small instances needed quite a long time in our tests and there are only 4 instances per node density pair.

To get more comparable results using more instances, the author generated her own set. Therefore, 10 random connected graphs per node-density pair using 10,15 and 20 as number of nodes and edge densities of 25%,50% and 75% were created. For each graph two instances with random objective on the edges and edge pairs were generated. One instance includes all edge pairs and the other only adjacent edge pairs. The objective was chosen as random integers between -100 and 100.

4.3 Experiments and Results

Test setup

All test were done on a Intel® Core™i7-2600 CPU running at 3.4GHz on 4 cores and 8 threads.

To compare the different formulations different values were measured and calculated:

root gap The root gap (for solved instances) calculated as

$$\frac{|MIP^* - LP^*|}{|MIP^*|},$$

where MIP^* is the optimal solution of the MIP and LP^* is the optimal solution of the LP relaxation. To calculate this we solved the LP relaxation separately.

node count The number of branch and cut nodes as reported by Gurobi.

runtime The *runtime* in seconds reported by Gurobi (wall-clock time).

[1] Cordone and Passeri's instances be downloaded from `https://homes.di.unimi.it/cordone/research/qmst.html`

gap The current relative MIP optimality *gap* reported by Gurobi. It is computed as

$$\frac{|\gamma_u - \gamma_\ell|}{|\gamma_\ell|},$$

where γ_u is the upper bound given by the objective of the current best MIP incumbent and γ_ℓ is the lower bound given by the minimum of the objective of all actual leaf nodes in the branch and cut tree.

In all following tests the time limit was 3600s.

Results

nodes	density	sub	aq-sub	q-sub	martin	aq-m	q-m
10	25	10	10	10	10	10	10
	50	10	10	10	10	10	10
	75	10	10	10	10	10	10
15	25	10	10	10	10	10	10
	50	10	10	10	10	10	9
	75	10	9	9	6	6	0
20	25	10	10	10	10	10	10
	50	0	0	0	0	0	0
	75	0	0	0	0	0	0

Table 4.2: Number of solved instances form our test set with *all* possible monomials using the different IP-formulations

Considering our instances with all possible degree-two monomials we see in Table 4.2 that one could solve more instance with the formulations in the original space than with the extended formulations. Including the quadratic subtour elimination constraints however has a negative impact regarding the number of solved instances (`aq-sub`, `q-sub` compared to `sub`). We can not observe any impact of the extended linearization constraints for adjacent monomials considering the number of solved instances (`aq-m` compared to `martin`), but using also the extended linearization constraints for nonadjacent monomials (`q-m`) decreased this number significantly.

nodes	density	sub	aq-sub	martin	aq-m
10	25	10	10	10	10
	50	10	10	10	10
	75	10	10	10	10
15	25	10	10	10	10
	50	10	10	10	10
	75	10	10	10	10
20	25	10	10	10	10
	50	10	10	10	10
	75	1	1	5	6

Table 4.3: Number of solved instances form our test set with all *adjacent* monomials using the different IP-formulations

Considering instances with only adjacent monomials (Table 4.3) one could solve more instances using extended formulations than using formulations in the original space with separation routines(`sub`,`aq-sub`). This time we can observe a positive impact of the extended linearization constraint for adjacent monomials. So our first guess is that the extended formulations lead to better performances in the case of only adjacent monomials and we will confirm this considering the more detailed plots in Figure 4.5.

In Figure 4.4 and 4.5 you can see the average measurements. We were able to calculate the LP relaxation for all instances. Accordingly, the plot of the root gap is the mean over those instances, where we could solve the problem with at least one of our formulations. The plots of the node count and the runtime present the mean over those instances that were solved with all formulations. The gap refers to the mean over all ten instances and we plotted only those, where this value is greater than 0.

Considering the results for the instances with all possible monomials in Figure 4.4, we observe that the quadratic subtour elimination constraints in `aq-sub` and `q-sub` provide only a small improvement of the root gap compared to the improvement in the extended formulations `aq-m` and `q-m`. Especially the difference between `aq-sub` and `q-sub` is very small, whereas `q-m` provides a significant improvement compared to `aq-m`.

Although `q-m` provides such a good root gap it performs much worse than the other formulations in practice regarding the node count, the runtime and the gap. Hence, we obtain the impression that for some reason

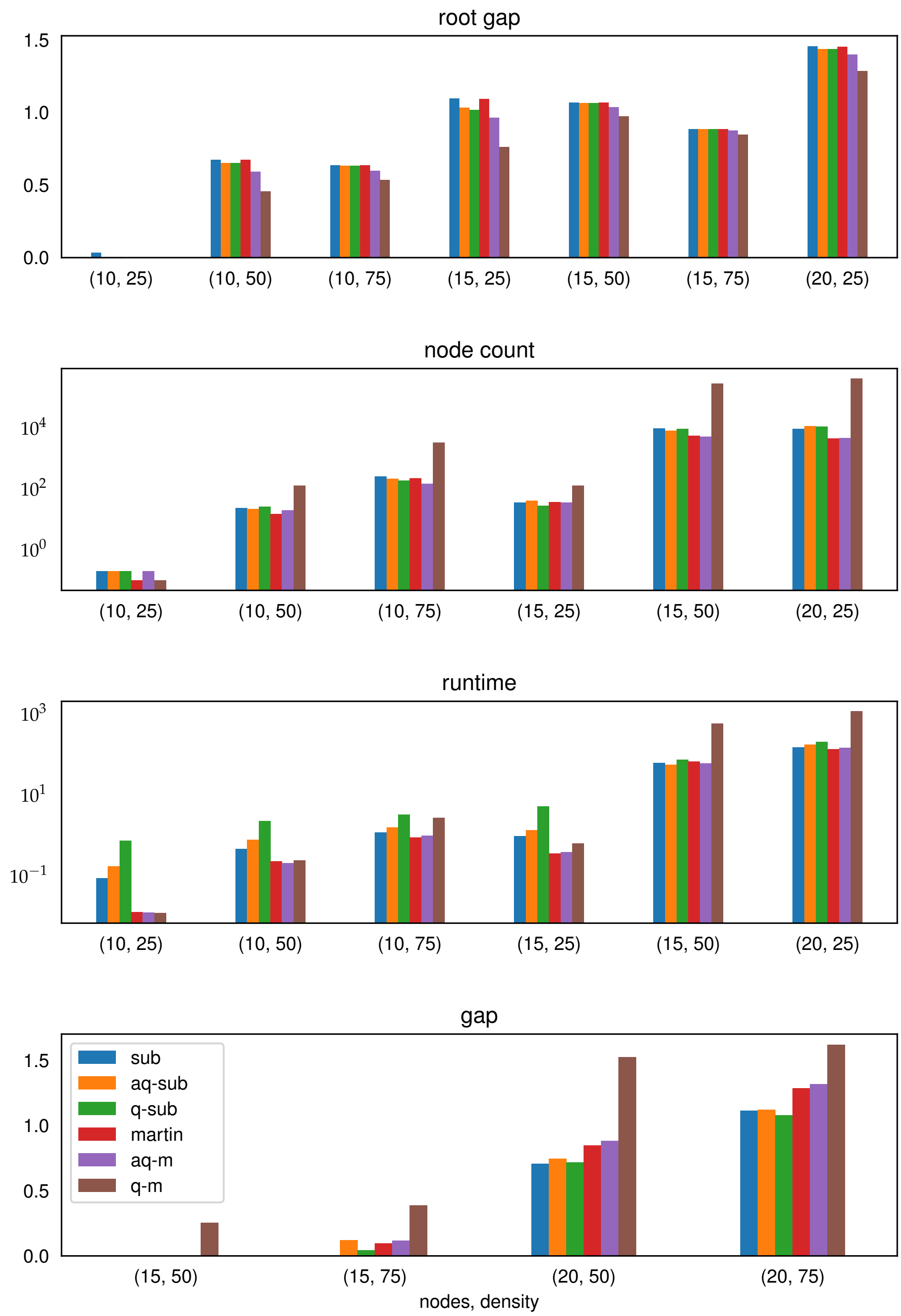


Figure 4.4: The mean of different measured values over our test instances with *all* possible monomials

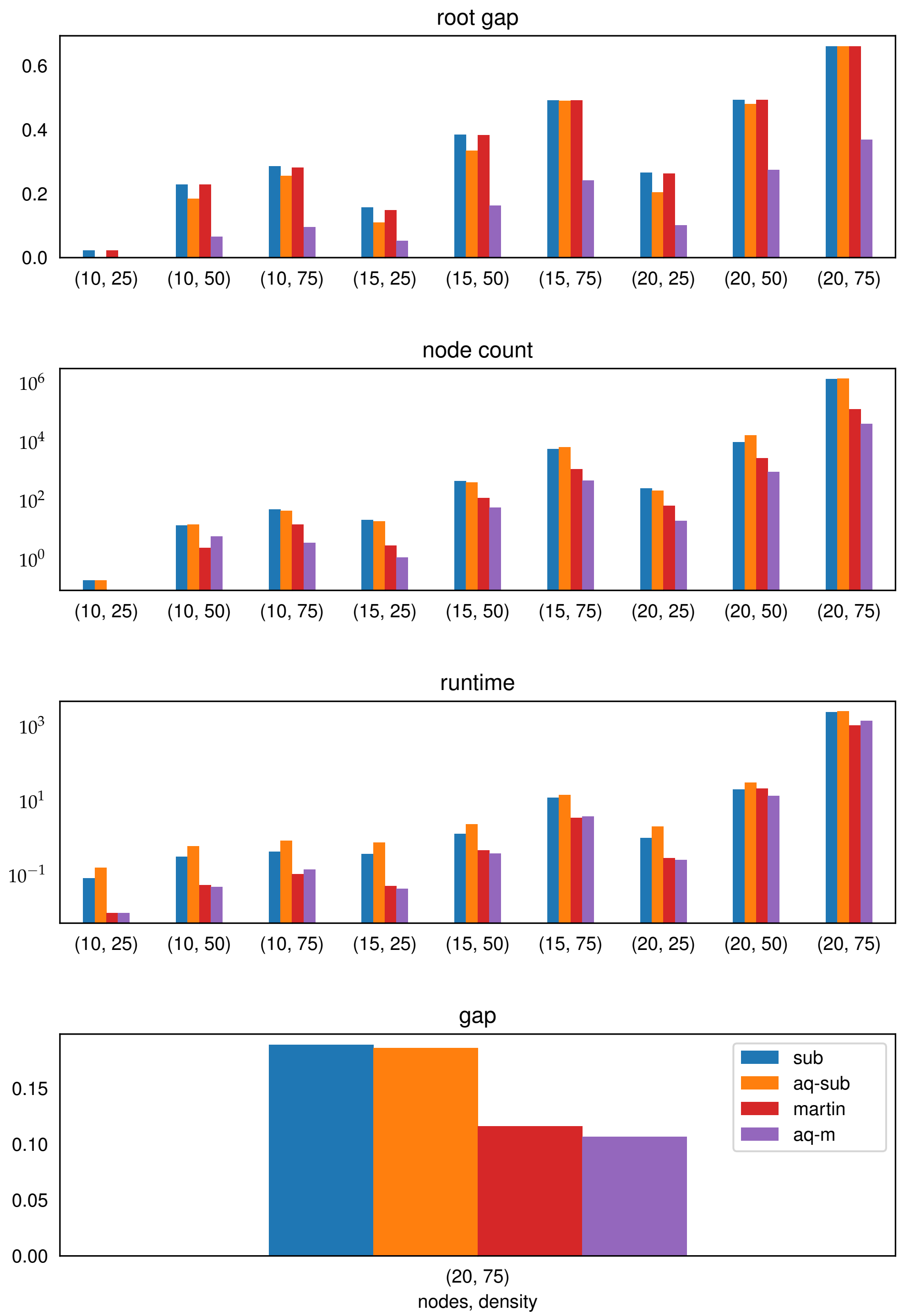


Figure 4.5: The mean of different measured values over our test instances with all *adjacent* monomials

(which would need further investigations) Gurobi can not handle this formulation as well as the others.

The other two extended formulations `martin` and `aq-m` have slightly better node counts and runtimes than the separation based formulations, but regarding the gap we guess that `sub`,`aq-sub` and `q-sub` might perform better for large and dense instances. This fits perfectly with our observation regarding the number of solved instances.

This tendency is not that obvious regarding the instances with only adjacent monomials (Figure 4.5). Here the extended formulations lead to better results than the formulations using separation routines in all measured values, like guessed before regarding the number of solved instances. Especially the improvement of the root gap in `aq-m` compared to `aq-sub` is remarkable.

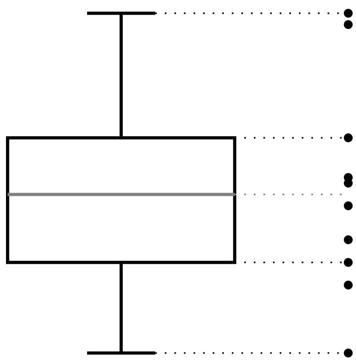


Figure 4.6: Boxplot for 10 example points

To get an impression of the dispersion of the data we consider some box plots for a subset of the values. The whiskers mark the minimum and maximum value, the grey line is the middle quantile (i.e., the median) and the box spans the region of the middle 50% of data. In our case we only have 10 data point and thus we have 6 points inside the box and 4 points outside the box, 2 above and 2 below as illustrated in Figure 4.6.

We consider the instance set with 15 nodes and the edge density of 50% from the instances including all possible monomials and the instance set with 20 nodes and the edge density of 50% from the instances with only adjacent monomials in Figure 4.7 and Figure 4.8.

In both cases we have similar dispersions for all formulations regarding the root gap. Considering the fact that one formulation implies another as we proved in Chapter 2, we expect the root gap to decrease corresponding to this implications. The box plots suggest that this improvement is fairly evenly for all instances.

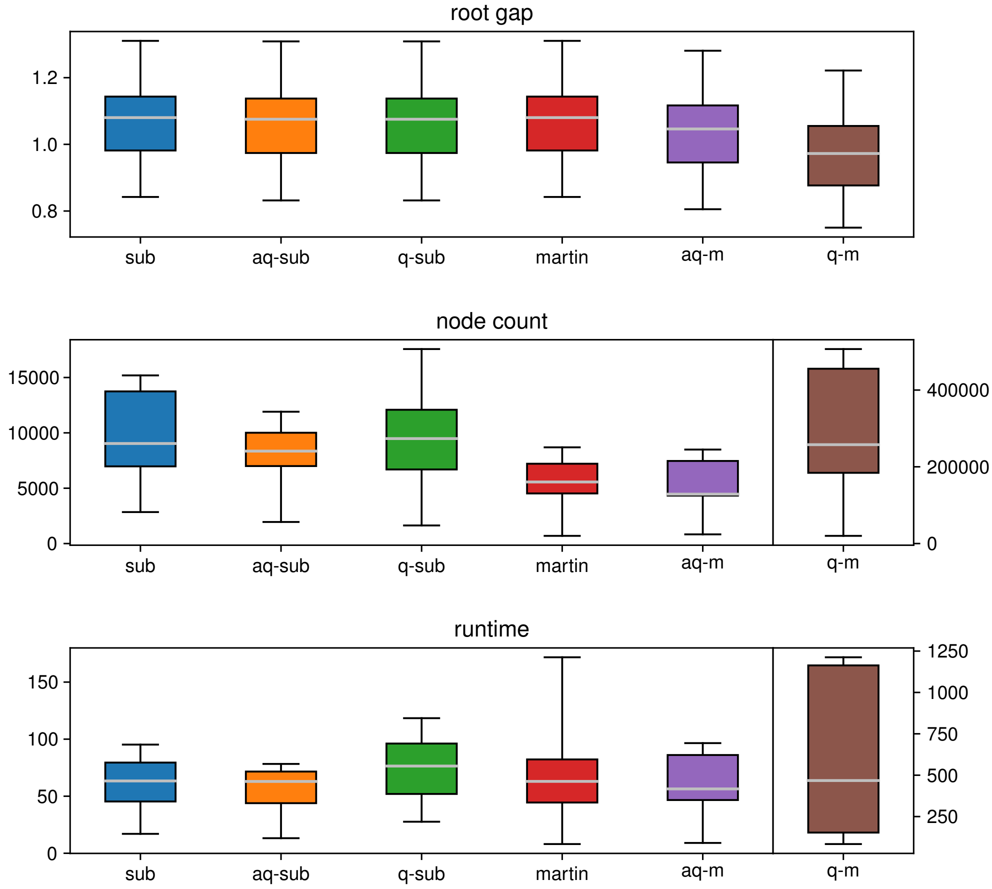


Figure 4.7: Box plots for the instances with *all* possible monomials and $n = 15, d = 50$

The other measurements vary more. Keeping in mind that the box plots refer to only 10 instances and Gurobi is a huge solver with a lot of heuristics, cutting planes and other routines that handles each instance differently good, it is not very surprising that the dispersion varies. Since the formulation `q-m` resulted in much worse performance than the other formulations we plotted the node count and the runtime with a different scale. Considering the instances with only adjacent monomials, there is one noticeable outlier considering the formulation `aq-sub`.

All in all, the box plots go well to what one should expect.

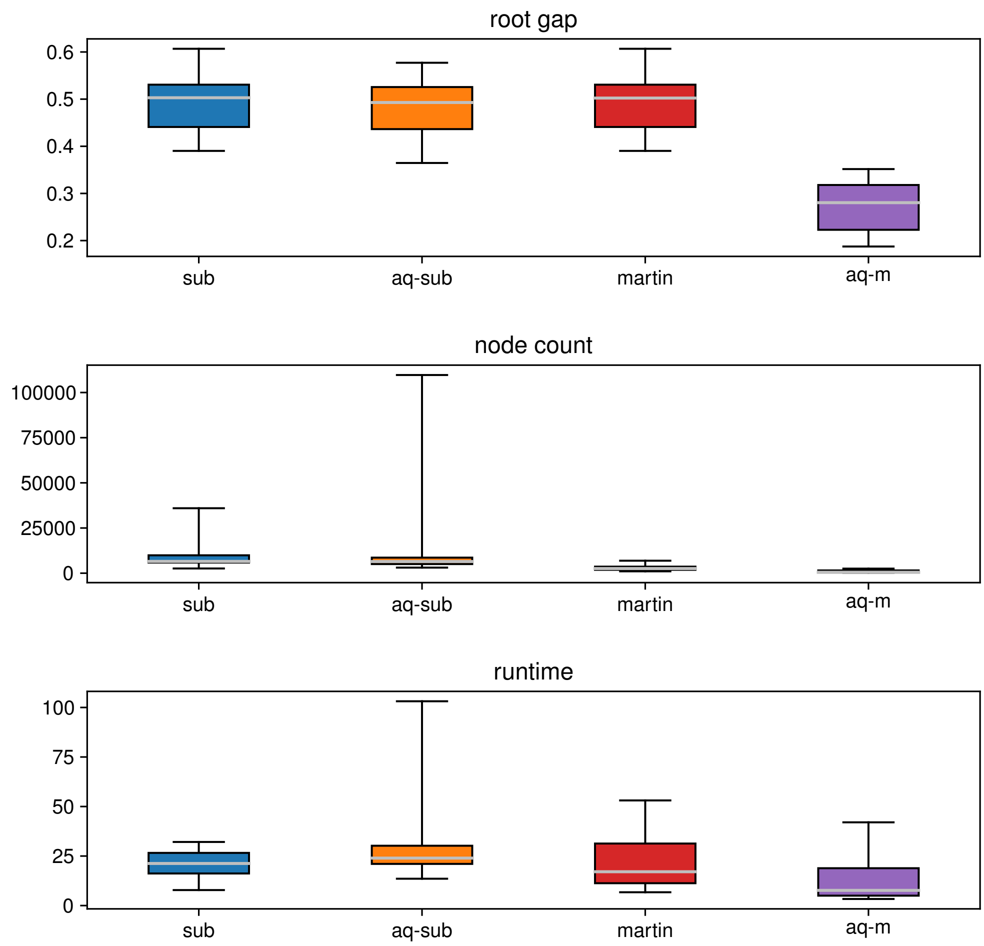


Figure 4.8: Box plots for the instances with all *adjacent* monomials and $n = 20, d = 50$

Considering the instances provided by Cordone and Passeri we obtain similar results as with our instances including all possible monomials. (See Table 4.4 and Figure 4.9.) Using the formulations with the subtour elimination constraints one could solve more instances than with our extended formulations in the time limit of one hour. This time, the quadratic subtour elimination constraint for nonadjacent monomials improved the number of solved instances, whereas the extended linearization constraints for nonadjacent monomials still has a negative impact on this number.

In Figure 4.9 we see that for large and dense instances `q-sub` leads to the best results regarding runtime, node count and gap. This is different

n	d	sub	aq-sub	q-sub	martin	aq-m	q-m
10	33	4	4	4	4	4	4
	67	4	4	4	4	4	4
	100	4	4	4	4	4	4
15	33	4	4	4	4	4	4
	67	3	3	4	1	1	1
	100	1	1	1	1	1	0
20	33	1	1	1	1	1	0
	67	0	0	0	0	0	0
	100	0	0	0	0	0	0

Table 4.4: Number of solved instances from Cordone and Passeri's test set

to what we observed before. The reason for this could be that we allow the objective also to be negative or that our instances are smaller and less dense.

Our conjecture that the formulations `sub`,`aq-sub` and `q-sub` can be solved faster than the extended formulations for large and dense instances got confirmed here.

The improvement of the root gap however is as before significantly better using extended formulations.

All in all, it is not easy to say which approach performs better. Considering only adjacent monomials extended formulations can lead to better performance, whereas the large and dense instances with all monomials can be solved faster using the separation based approach. Considering the root gap we could observe an evenly improvement by the extended formulations.

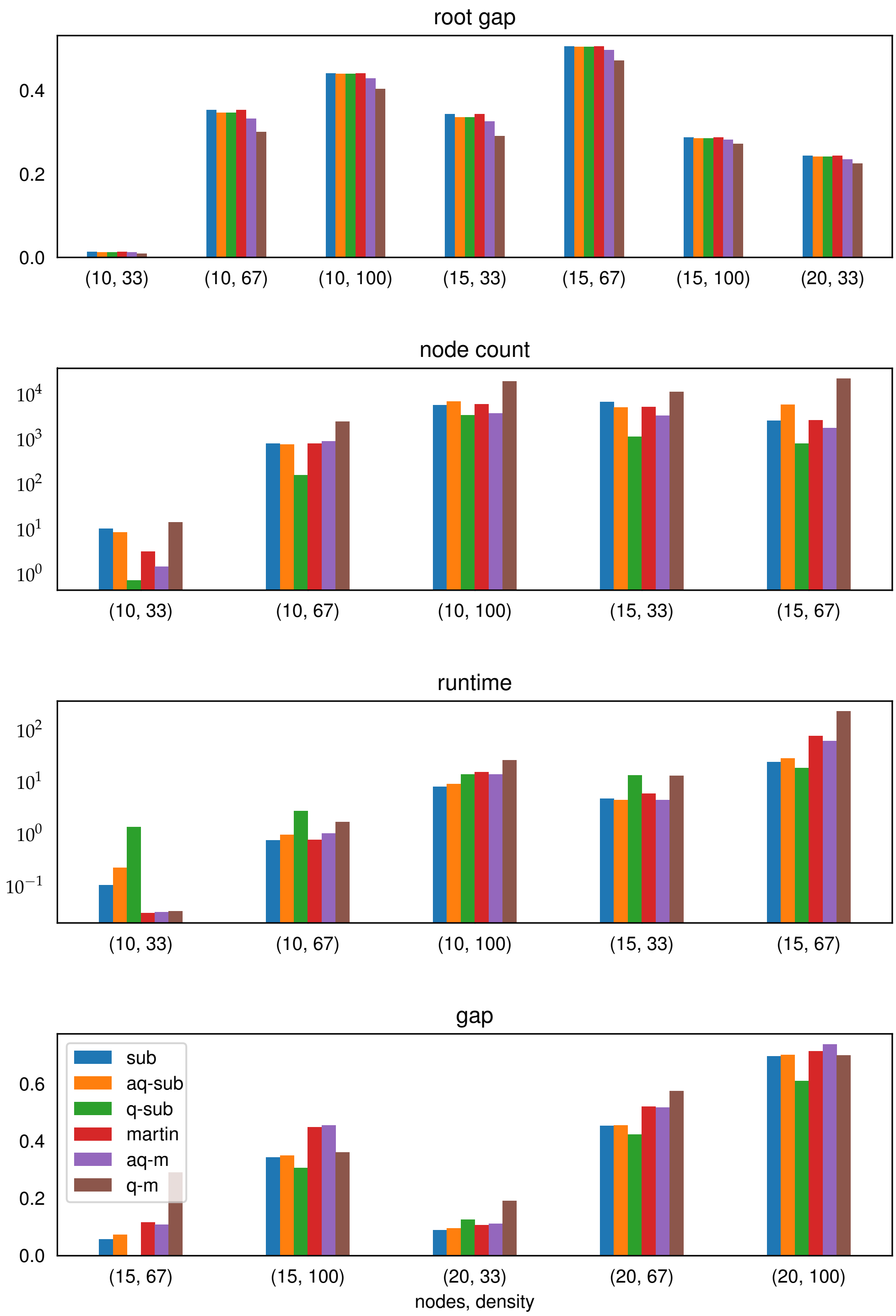


Figure 4.9: Testinstances of Cordone and Passeri

5 Higher Order Arborescence Polytopes

In Chapter 2 we modeled extended formulations for higher order forest polytopes in undirected graphs. In this chapter we investigate directed graphs and arborescences, which are the counterparts to spanning trees in undirected graphs.

Martin's extended formulation for the spanning tree polytope P_{ST} is based on arborescences in the sense that for each spanning tree we obtain valid z^u- variables by choosing the characteristic vector of the induced u-arborescence.

Let $G = (V, E)$ be a directed graph. We recall that for a root node $r \in V$ an *r-arborescence* $A \subseteq E$ is a cycle free set of edges, such that r has no ingoing edge in A and all other nodes have exactly one ingoing edge in A. The *r-arborescence polytope* of G is defined by

$$P_{arb} := \operatorname{conv}\left\{\chi(A) \middle| A \text{ is an } r\text{-arborescence in } G\right\}$$

and a complete description was given by to Edmonds [12]. You can find it for example in [34, Section 51.4] and [25, Corollar 6.15].

Proposition 12 (Edmonds [12])**.** *The r-arborescence polytope P_{arb} is described by*

$$x\left(\delta^{in}(S)\right) \geq 1 \qquad \textit{for all } S \subseteq V \setminus \{r\} \textit{ with } S \neq \emptyset \tag{5.1}$$

$$x\left(\delta^{in}(v)\right) = 1 \qquad \textit{for all } v \in V \setminus \{r\} \tag{5.2}$$

$$x\left(\delta^{in}(r)\right) = 0 \tag{5.3}$$

$$x \geq 0 \tag{2.2}$$

It is possible to exchange (5.1) by Edmonds' rank constraints, which we know from the spanning tree polytope, to obtain the following alternative description. (See, e.g., [25, Chapter 6, Exercise 23].)

Remark 3. The description in Proposition 12 is equivalent to

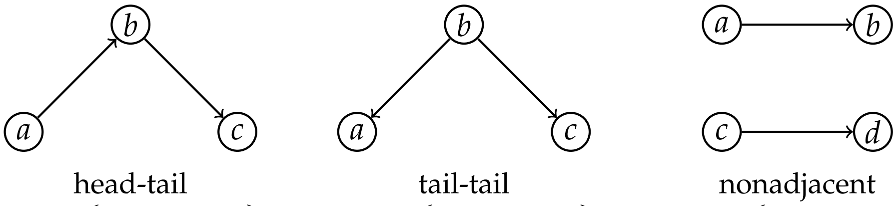
$$x(E(S)) \leq |S| - 1 \qquad \text{for all } S \subset V \text{ with } S \neq \emptyset \tag{2.1}$$

$$x\left(\delta^{in}(v)\right) = 1 \qquad \text{for all } v \in V \setminus \{r\} \tag{5.2}$$

$$x\left(\delta^{in}(r)\right) = 0 \tag{5.3}$$

$$x \geq 0. \tag{2.2}$$

For $\mathcal{M} \subseteq 2^E$ the polytopes $P_{arb}(\mathcal{M})$ are called *higher order r-arborescence polytopes*. In this work we only consider single degree-two monomials $\mathcal{M} = \{M\}$ with $|M| = 2$ and we distinguish the three cases



$M = \{(a,b),(b,c)\}$ $\quad$ $M = \{(b,a),(b,c)\}$ $\quad$ $M = \{(a,b),(c,d)\}$

Figure 5.1: Possible degree-two monomials in directed graphs

illustrated in Figure 5.1 where a, b, c and d are pairwise distinct. We assume $r \notin \{a,b,c,d\}$. In the next two sections we analyze two extended formulations for P_{arb} and compare their capability to model extended formulations for $P_{arb}(\mathcal{M})$.

5.1 Extended Formulation Based on R. Kipp Martin

Martin's extended formulation for P_{ST} in Proposition 2 can be easily adapted for directed graphs:

$$z^u_{v,w} + z^u_{w,v} = \overline{x}_{v,w} \qquad \text{for all } u \in V,\ (v,w) \in E \tag{2.4''}$$

$$z^u\left(\delta^{in}(v)\right) = 1 \qquad \text{for all } u, v \in V,\ u \neq v \tag{2.5}$$

$$z^u\left(\delta^{in}(u)\right) = 0 \qquad \text{for all } u \in V \tag{2.6}$$

$$z \geq 0 \tag{2.7}$$

where for all $(v,w) \in E$

$$\overline{x}_{v,w} := \begin{cases} x_{v,w} + x_{w,v} & \text{if } (w,v) \in E \\ x_{v,w} & \text{if } (w,v) \notin E. \end{cases}$$

Proposition 13. *Let $G = (V,E)$ be a connected directed graph. The projection of* (2.4″),(2.5)-(2.7) *onto z^r is the r-arborescence polytope P_{arb}.*

Proof. For a given r-arborescence A let $z^r = x$ be the characteristic vector of A and z^u be the characteristic vectors of the induced u-arborescences that have the same underlying undirected tree. This choice is valid for the system (2.4″),(2.5)-(2.7).

On the other hand, let (x, z) be valid for the system. We will show that (2.4″),(2.5)-(2.7) imply the description of P_{arb} from Remark 3. It is easy to see that (2.5)-(2.7) directly imply (5.2),(5.3) and (2.2) for $x = z^r$. To obtain Edmond's rank inequalities (2.1) we use (2.4″) to observe

$$z^r(E(S)) = z^u(E(S)) \qquad \text{for all } S \subseteq V \text{ and } u \in V$$

and with (2.5) and (2.6) we receive

$$z^u(E(S)) \leq \sum_{v \in S} z^u\left(\delta^{in}(v)\right) = |S| - 1 \quad \text{for all } S \subseteq V \text{ and } u \in S. \tag{2.1}$$

□

For the ease of notation we can simply add

$$x = z^r. \tag{5.4}$$

To extend the formulation for a degree-two monomial, we can use the extended linearization constraints from Chapter 2. Those are

$$y \leq z^c_{b,a} \tag{2.15a}$$
$$y \leq z^a_{b,c} \tag{2.15b}$$

for the adjacent cases (head-tail and tail-tail) and

$$y \leq z^a_{c,d} + z^b_{d,c} \tag{2.16a}$$
$$y \leq z^a_{d,c} + z^b_{c,d} \tag{2.16b}$$
$$y \leq z^c_{a,b} + z^d_{b,a} \tag{2.16c}$$
$$y \leq z^c_{b,a} + z^d_{a,b}. \tag{2.16d}$$

for the nonadjacent case. Again, we need McCormick's linearization constraints

$$y \leq x_{e_i} \qquad \text{for } i \in \{1, 2\} \tag{2.8}$$
$$y \geq x_{e_1} + x_{e_2} - 1 \tag{2.10}$$

where $\{e_1, e_2\} := M$ [30]. The following propositions illustrate a strong relationship to the quadratic subtour elmination constraints (2.13) and (2.14).

Proposition 14 (nonadjacent case). *Let $G = (V, E)$ be a connected directed graph and let Q be the polytope described by Martin's constraints* (2.4″),(2.5)-(2.7) *for $u \in \{a, b, c, d\}$, McCormick's linearization* (2.8)-(2.10)*, our extended linearization constraints* (2.16a)-(2.16d) *and*

$$\begin{aligned} & x \in P_{arb} \\ & y \geq 0. \end{aligned} \tag{2.12}$$

The projection of Q onto (x, y) is the polytope P described by McCormick's linearization (2.8)-(2.10) *and*

$$x \in P_{arb}$$

$$x(E(S_1)) + x(E(S_2)) + y \leq |S_1| + |S_2| - 2 \qquad \textit{for } S_1, S_2 \subset V \quad \textit{with } (a, b), (c, d) \textit{ each have one endnode in } S_1 \textit{ and one in } S_2 \tag{2.14}$$

$$y \geq 0. \tag{2.12}$$

Proof. To show that the projection of Q is contained in P we combine constraints (2.4″),(2.5)-(2.7) from Martin's formulation with our extended linearization constraints (2.16a)-(2.16d) the same way as in the proof of Theorem 3 to imply the quadratic subtour elimination constraint (2.14).

For the reverse inclusion we have to show that for each $(x, y) \in P$ we can find a vector z such that (x, y, z) is valid for Martin's formulation (2.4″),(2.5)-(2.7) and our additional constraints (2.16a)-(2.16d). Let for all $(v, w) \in E$

$$\overline{x}_{\{v,w\}} := \begin{cases} x_{v,w} + x_{w,v} & \text{if } (w, v) \in E \\ x_{v,w} & \text{if } (w, v) \notin E \end{cases}.$$

Since $\overline{x} \in P_{ST}$ of the corresponding undirected graph we can choose $\overline{y}$ maximal such that $(\overline{x}, \overline{y}) \in P_{ST}(\mathcal{M})$ for the monomial $\mathcal{M} = \{M\}$ with $M = \{\{a, b\}, \{c, d\}\}$.

Now we have $y \leq \overline{y}$, because otherwise we would have

$$y > \overline{y} \quad \geq \overline{x}_{a,b} + \overline{x}_{c,d} - 1$$

and due to (2.8) and the definition of $\overline{x}$

$$\begin{aligned} & y \leq x_{a,b} \leq \overline{x}_{a,b} \\ & y \leq x_{c,d} \leq \overline{x}_{c,d}. \end{aligned}$$

Using Buchheim and Klein's formulation in Proposition 5 together with $x(E(S)) = \overline{x}(E(S))$ this implies $(\overline{x}, y) \in P_{ST}(\mathcal{M})$, which is a contradiction to the maximality of $\overline{y}$.

Due to $(\overline{x}, \overline{y}) \in P_{ST}(\mathcal{M})$ there exists a convex combination of trees

$$\overline{x} = \sum_{i \in I} \lambda_i \chi(T_i)$$

for some index set I, such that

$$\overline{y} = \sum_{i \in I \text{ with } M \subseteq T_i} \lambda_i.$$

Each tree T_i and each $v \in V \setminus \{r\}$ induce a unique v-arborescence A_i^v as

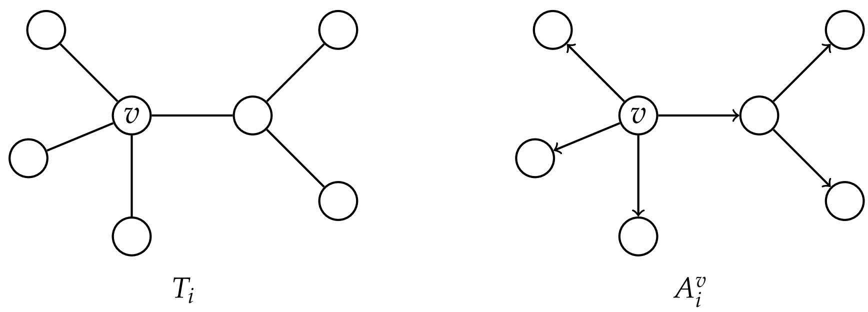


Figure 5.2: A tree and its induced v-arborescence

illustrated in Figure 5.2. This way, we can define

$$z^v := \sum_{i \in I} \lambda_i \chi(A_i^v).$$

Additionally, we set $z^r = x$. This choice of z is valid for Martin's formulation. For each tree T_i with $M \subseteq T_i$ we have either $(c, d) \in A_i^a$ or $(d, c) \in A_i^b$ and thus

$$z_{c,d}^a + z_{d,c}^b \geq \sum_{i \in I \text{ with } M \subseteq T_i} \lambda_i = \overline{y} \geq y. \tag{2.16a}$$

The proof works analogously for inequalities (2.16b)-(2.16d). □

Proposition 15 (adjacent cases). *Let $G = (V, E)$ be a connected directed graph and let Q be the polytope described by Martin's constraints* (2.4''),(2.5)-(2.7) *for $u \in \{a, c\}$, McCormick's linearization* (2.8),(2.10), *our additional constraints* (2.15a),(2.15b) *and*

$$\begin{aligned} & x \in P_{arb} \\ & y \geq 0. \end{aligned} \tag{2.12}$$

The projection of Q onto (x, y) is the polytope P described by (2.8), (2.10) *and*

$$\begin{aligned} & x \in P_{arb} \\ x(E(S)) + y \le |S| - 1 \quad & \text{for all } S \subset V \text{ with } \{a, c\} \subseteq S,\, b \notin S \qquad (2.13) \\ & y \ge 0. \qquad (2.12) \end{aligned}$$

Proof. To show that the projection of Q is contained in P we use (2.4″) to obtain

$$x(E(S)) = \sum_{v \in S} z^a\left(\delta^{in}(v)\right) - z^a\left(\delta^{in}(S)\right)$$

for all $S \subseteq V$ with $\{a, c\} \subseteq S$, $b \notin S$. We combine it with (2.5)-(2.7) to

$$x(E(S)) \le |S| - 1 - z^a_{b,c}$$

and use (2.15a) to imply (2.13).

For the inverse inclusion we use analog argumentations and notation as in the proof of Proposition 14. Now with $M = \{(a, b), (b, c)\}$ or $M = \{(b, a), (b, c)\}$ this leads to

$$\overline{x} = \sum_{i \in I} \lambda_i \chi(T_i)$$

and

$$\overline{y} = \sum_{i \in I \text{ with } M \subseteq T_i} \lambda_i.$$

For all T_i with $M \subset T_i$ we have $(b, a) \in A^c_i$ and $(b, c) \in A^a_i$ and thus

$$z^c_{b,a} \ge \sum_{i \in I \text{ with } M \subseteq T_i} \lambda_i = \overline{y} \ge y \tag{2.15a}$$

$$z^a_{b,c} \ge \sum_{i \in I \text{ with } M \subseteq T_i} \lambda_i = \overline{y} \ge y. \tag{2.15b}$$

□

Proposition 16. *Let $\mathcal{M} = \{M\} \subseteq 2^E$ with $|M| = 2$. There exists a connected directed graph $G = (V, E)$ such that the formulations in Propositions 14 and 15 do not describe $P_{arb}(\mathcal{M})$ completely.*

Proof. We will provide three counter examples, one for each of the three cases presented before in Figure 5.1. They all use the observation that the quadratic subtour elimination constraints (2.14) and (2.13) in Proposition 5 are fulfilled by x and y as soon as the are fulfilled in the underlying undirected graph. Let x be a convex combination of characteristic vectors of r-arborescences

$$x = \sum_{i \in I} \lambda_i \chi(A_i),$$

then the formulations only enforce

$$y \leq \min\left\{x_{e_1}, x_{e_2}, \sum_{i \in I \text{ with } \{\bar{e}_1, \bar{e}_2\} \subset T_i} \lambda_i\right\},$$

where $\{e_1, e_2\} := M$, T_i is the spanning tree corresponding to A_i in the underlying undirected graph and $\bar{e}_j$ is the undirected version of e_j for $j \in \{1,2\}$.

Head-tail case: Let x be as in Figure 5.3 and $y = \frac{1}{2}$. This is valid for the formulation in Proposition 15 using the former argumentation with the convex combination shown in the figure, but it is impossible to express x as a convex combination of r-arborescences where one arborescence includes (a,b) and (b,c), since the only edges in $\delta^{in}(a)$ are (b,a) and (c,a) and each of them would combined with the edges (a,b) and (b,c) induce a cycle. This means, y must be 0.

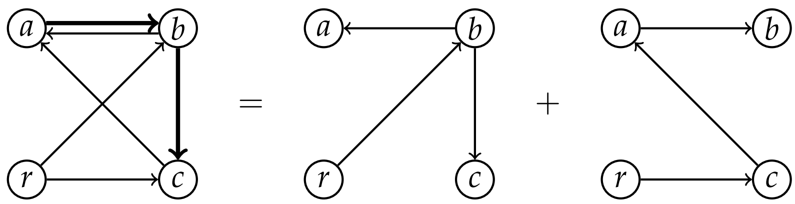


Figure 5.3: Convex combination of r-arborescences. All edge weights in this figure are $\frac{1}{2}$.

Tail-tail case: Let x be as in Figure 5.4 and $y = \frac{1}{2}$. This is valid for the formulation in Proposition 15 using the former argumentation with the convex combination shown in the figure, but it is impossible to express x as a convex combination of r-arborescences where one arborescence includes (b,a) and (b,c), since the only edges is $\delta^{in}(b)$ are (a,b) and (c,b). Hence, y must be 0.

Nonadjacent case: Let x be as in Figure 5.5 and $y = \frac{1}{2}$. This is valid for the formulation in Proposition 14 using the former argumentation with the

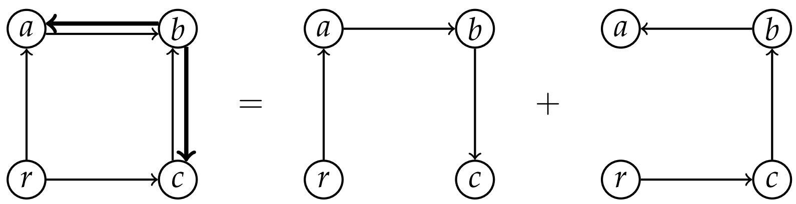


Figure 5.4: Convex combination of r-arborescences. All edge weights in this figure are $\frac{1}{2}$.

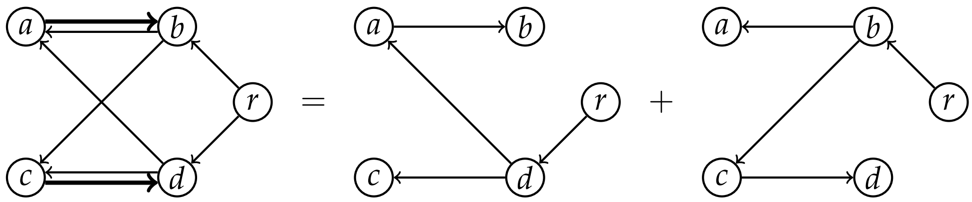


Figure 5.5: Convex combination of r-arborescences. All edge weights in this figure are $\frac{1}{2}$.

convex combination shown in the figure, but it is not possible to express x as a convex combination of r-arborescences where one arborescence includes (a, b) and (c, d), since the only edges in $\delta^{out}(r)$ are (r, b) and (r, d) and this would lead to $x\left(\delta^{in}(b)\right) > 1$ or $x\left(\delta^{in}(d)\right) > 1$ respectively. Therefore, y must be 0. $\square$

5.2 Extended Formulation Based on Richard T. Wong

In the former section we were not able to build a complete description of $P_{arb}(\mathcal{M})$ for $\mathcal{M} = \{M\}$ and $|M| = 2$ using Martin's formulation with our additional constraints. The problem is that the information about the direction of the edges in the arborescence is hidden in the interaction of the z-variables and can not be used directly in the additional constraints for the monomial. This is different in another extended formulation introduced by Richard T. Wong originally for the traveling salesman problem [39], but often also mentioned for spanning trees [38].

Each r-arborescence contains unique directed paths from r to each node $v \in V \setminus \{r\}$. Such r-v-paths can be regarded as r-v-flows of capacity 1. Those flows are expressed by the variables w^v for $v \in V \setminus \{r\}$ in Wong's formulation.

Proposition 17 (Wong [39]). *The following constraints provide an extended formulation for P_{arb}.*

$$x(E) = |V| - 1 \tag{2.3}$$
$$w_e^u \leq x_e \qquad \text{for all } e \in E \text{ and } u \in V \setminus \{r\} \tag{5.5}$$
$$w^u\left(\delta^{in}(v)\right) = w^u\left(\delta^{out}(v)\right) \qquad \text{for all } u, v \in V \setminus \{r\} \text{ with } v \neq u \tag{5.6}$$
$$w^u\left(\delta^{in}(u)\right) = 1 \qquad \text{for all } u \in V \setminus \{r\} \tag{5.7}$$
$$w^u\left(\delta^{out}(r)\right) = 1 \qquad \text{for all } u \in V \setminus \{r\} \tag{5.8}$$
$$w \geq 0. \tag{5.9}$$

Proof. For a given r-arborescence A we define w^u as the characteristic vector of the unique r-u path in A. This choice is obviously valid for the formulation (2.3),(5.5)-(5.9) and thus P_{arb} is contained in the projection of the polytope defined by Wong's formulation.

Let now (x, w) be valid for (2.3),(5.5)-(5.9). We will show that (x, w) fulfills all constraints of the description of P_{arb} in Proposition 12. Inequality (5.5) with (5.9) lead to

$$x_e \geq 0 \tag{2.2}$$

and (5.5) with (5.7) to

$$x\left(\delta^{in}(v)\right) \geq w^v\left(\delta^{in}(v)\right) = 1 \qquad \text{for all } v \in V \setminus \{r\}.$$

Together with

$$x(E) = \sum_{v \in V} x\left(\delta^{in}(v)\right) = |V| - 1, \tag{2.3}$$

this implies

$$x\left(\delta^{in}(v)\right) = 1 \qquad \text{for all } v \in V \setminus \{r\}. \tag{5.2}$$

Since w^v describes a flow we see with (5.5)

$$x\left(\delta^{in}(r)\right) \geq w^v\left(\delta^{in}(r)\right) = 0. \tag{5.3}$$

To complete the proof we observe that w^u is an r-u flow of capacity 1 and $\delta^{in}(S)$ is an r-u cut for each $u \in S$ if $r \notin S$. Thus, using weak duality between flows and cuts we have

$$x\left(\delta^{in}(S)\right) \geq w^u\left(\delta^{in}(S)\right) \geq 1 \qquad \text{for all } S \subseteq V \setminus \{r\} \text{ and } u \in S. \tag{5.1}$$

□

Head-tail case

Let $\mathcal{M} = \{M\}$ with $M = \{(a,b),(b,c)\}$. An evident idea to extend the formulation in this case is the inequality

$$y \leq w^c_{a,b}. \tag{5.10}$$

Let x be a characteristic vector of some r-arborescence A, then $M \subseteq A$ implies that the unique r-c-path in A includes the edges (a,b) and (b,c) that leads to $w^c_{a,b} = 1$, where w^c is the characteristic vector of the r-c-path in A. Hence, (5.10) is valid for $P_{Arb}(\mathcal{M})$.

The following proposition shows that Wong's formulation with (5.10) is at least as good as the formulation in the former section based on Martin in the sense that the projection of the formulation based on Wong is contained in the projection of the formulation based on Martin.

Proposition 18. *Wong's formulation* (2.3),(5.5)-(5.9) *with* (5.10) *implies*

$$x\left(E\left(S\right)\right) + y \leq |S| - 1 \qquad \textit{for all } S \subset V \textit{ with } a,c \in S, b \notin S. \tag{2.13}$$

Proof. We use the observation that for all $S \subseteq V$ we have

$$\begin{aligned} w^u\left(E\left(S\right)\right) &= \sum_{v \in S\setminus\{r\}} w^u\left(\delta^{in}\left(v\right)\right) - w^u\left(\delta^{in}\left(S\right)\right) \\ &= \sum_{v \in S} w^u\left(\delta^{out}\left(v\right)\right) - w^u\left(\delta^{out}\left(S\right)\right). \end{aligned}$$

This implies for all $S \subseteq V \setminus \{r\}$ and $u \in S$

$$\begin{aligned} w^u\left(\delta^{in}\left(S\right)\right) &= w^u\left(\delta^{out}\left(S\right)\right) - \sum_{v \in S\setminus\{u\}} \left(w^u\left(\delta^{out}\left(v\right)\right) - w^u\left(\delta^{in}\left(v\right)\right)\right) \\ &\quad + w^u\left(\delta^{in}\left(u\right)\right) \\ &= w^u\left(\delta^{out}\left(S\right)\right) + 1 \end{aligned}$$

and for all $S \subseteq V$ with $r,u \in S$ and $u \neq r$ we obtain

$$\begin{aligned} w^u\left(\delta^{in}\left(S\right)\right) &= w^u\left(\delta^{out}\left(S\right)\right) - \sum_{v \in S\setminus\{r,u\}} \left(w^u\left(\delta^{out}\left(v\right)\right) - w^u\left(\delta^{in}\left(v\right)\right)\right) \\ &\quad - w^u\left(\delta^{out}\left(r\right)\right) + w^u\left(\delta^{in}\left(u\right)\right) \\ &= w^u\left(\delta^{out}\left(S\right)\right). \end{aligned}$$

Altogether we combine it with (5.2) and (5.5) to

$$\begin{aligned}
x\left(E\left(S\right)\right) &= \sum_{v\in S\setminus\{r\}} x\left(\delta^{in}\left(v\right)\right) - x\left(\delta^{in}\left(S\right)\right) \\
&\leq \begin{cases} |S| - w^{c}\left(\delta^{in}\left(S\right)\right) & \text{if } r \notin S \\ |S| - 1 - w^{c}\left(\delta^{in}\left(S\right)\right) & \text{if } r \in S \end{cases} \\
&= |S| - 1 - w^{c}\left(\delta^{out}\left(S\right)\right) \\
&\leq |S| - 1 - w^{c}_{a,b},
\end{aligned}$$

which does with (5.10) imply

$$x\left(E\left(S\right)\right) \leq |S| - 1 - y. \tag{2.13}$$

□

Unfortunately, the formulation in Proposition 18 does also not lead to a complete formulation for $P_{arb}(\mathcal{M})$ as the author verified using polymake [18][2] for the complete graph with four and five nodes, i.e., the graph $G = (V, E)$ with $E = \left\{(u,v) \middle| u \in V, v \in V \setminus \{r\}\right\}$. The calculated vertices of the projection showed that we need further constraints bounding y from the upper and lower side.

Next the author calculated the facets of $P_{arb}(\mathcal{M})$ for the complete graphs with four and five nodes. The new upper bounds on y are

$$x\left(\delta^{in}\left(S_a\right) \setminus \left(\delta^{out}\left(S_b\right) \cup \delta^{out}\left(S_c\right)\right)\right) + x\left(\delta^{in}\left(S_b\right)\right) + x\left(\delta^{in}\left(S_c\right)\right) \geq 2 + y \tag{5.11}$$

for all $S_i \subseteq V \setminus \{r\}$ pairwise disjunct with $i \in S_i$ for $i \in \{a,b,c\}$ and the new lower bounds are

$$x\left(\delta^{in}\left(S \cup \{b\}\right)\right) + x\left(\delta^{in}(S) \cap \delta^{out}(b)\right) + y \geq x_{a,b} + x_{b,c} + 1 \tag{5.12}$$

for all $S \subseteq V \setminus \{r,a,b\}$ with $c \in S$ and

$$x_{r,b} + x_{r,d} + x_{a,d} + y \geq x_{b,c} \tag{5.13}$$

$$x_{r,a} + x_{r,b} + x_{r,d} + x_{c,a} + x_{d,a} + y \geq x_{a,b} + x_{b,c} \tag{5.14}$$

for $V = \{r,a,b,c,d\}$.

To model the new upper bounds (5.11) we add a new variable σ. It refers to a sub flow of w^a with capacity y that does not flow through b or c. For integral values with $y = 1$ this σ is equal to w^a and we have $\sigma\left(\delta^{out}\left(r\right)\right) = \sigma\left(\delta^{in}\left(a\right)\right) = 1$.

Proposition 19. *Wong's formulation* (5.5)-(5.9) *together with*

$$\sigma_{v,w} \leq w^a_{v,w} \qquad \textit{for all } (v,w) \in E \tag{5.15}$$
$$\sigma\left(\delta^{in}(v)\right) = \sigma\left(\delta^{out}(v)\right) \qquad \textit{for all } v \in V \setminus \{r,a\} \tag{5.16}$$
$$\sigma\left(\delta^{out}(r)\right) = y \tag{5.17}$$
$$\sigma\left(\delta^{in}(a)\right) = y \tag{5.18}$$
$$\sigma\left(\delta^{in}(b)\right) = 0 \tag{5.19}$$
$$\sigma\left(\delta^{in}(c)\right) = 0 \tag{5.20}$$
$$\sigma \geq 0 \tag{5.21}$$

implies (5.11).

Proof. Using constraints (5.5),(5.16),(5.21) and $\sigma\left(\delta^{in}(S_i)\right) = \sigma(\delta^{out}(S_i))$ for $i \in \{b,c\}$ we obtain

$$\begin{aligned}
&x\left(\delta^{in}(S_a) \setminus \left(\delta^{out}(S_b) \cup \delta^{out}(S_c)\right)\right) + x\left(\delta^{in}(S_b)\right) + x\left(\delta^{in}(S_c)\right) \\
&\quad \geq \sigma\left(\delta^{in}(S_a)\right) - \sigma\left(\delta^{in}(S_b)\right) - \sigma\left(\delta^{in}(S_c)\right) + w^b\left(\delta^{in}(S_b)\right) \\
&\qquad + w^c\left(\delta^{in}(S_c)\right).
\end{aligned}$$

Considering the flow decomposition of w^i for $i \in \{b,c\}$ we observe that those path that are also contained in $\sigma\left(\delta^{in}(S_i)\right)$ have to leave S_i again before they enter i, since the flow corresponding to σ does not include i. Thus, we have

$$w^i\left(\delta^{in}(S_i)\right) - \sigma\left(\delta^{in}(S_i)\right) \geq w^i\left(\delta^{in}(i)\right) \qquad \text{for } i \in \{b,c\}$$

and with (5.18) and (5.7) we obtain (5.11)

$$\begin{aligned}
&x\left(\delta^{in}(S_a) \setminus \left(\delta^{out}(S_b) \cup \delta^{out}(S_c)\right)\right) + x\left(\delta^{in}(S_b)\right) + x\left(\delta^{in}(S_c)\right) \\
&\quad \geq \sigma\left(\delta^{in}(a)\right) + w^b\left(\delta^{in}(b)\right) + w^c\left(\delta^{in}(c)\right) \\
&\quad \geq 2 + y. \qquad \square
\end{aligned}$$

To model the lower bound (5.12) we can directly use Wong's variables and only need one additional constraint that is similar to the last of McCormick's constraints (2.10) but not directly related.

Proposition 20. *Let $G = (V, E)$ be a connected directed graph, $r \in V$ and A be an r-arborescence in G. Furthermore, let $x = \chi(A)$ and for $u \in V$ let w^u be the characteristic vectors of the r-u-paths in A. For $y = x_{a,b}x_{b,c}$ the inequality*

$$y \geq w^c_{a,b} + w^c_{b,c} - w^c\left(\delta^{in}(b)\right). \tag{5.22}$$

is fulfilled by w and y.

Proof. In the case $y = 1$ we have

$$w^c_{a,b} = w^c_{b,c} = w^c\left(\delta^{in}(b)\right) = 1$$

and (5.22) is fulfilled.

For the case $y = 0$ observe

$$w^c\left(\delta^{in}(b)\right) \geq w^c_{a,b}$$

and

$$w^c\left(\delta^{in}(b)\right) = w^c\left(\delta^{out}(b)\right) \geq w^c_{b,c}$$

and thus we have

$$w^c_{a,b} + w^c_{b,c} - w^c\left(\delta^{in}(b)\right) \leq \min\left\{w^c_{a,b}, w^c_{b,c}\right\} = 0$$

due to $y = 0$. □

Proposition 21. *Wong's formulation* (5.5)-(5.9) *for $u = c$ together with* (5.22) *implies* (5.12).

Proof. With $(a, b) \in \delta^{in}(S \cup \{b\})$ and $(b, c) \in \delta^{in}(S) \cap \delta^{out}(b)$ and (5.5) we obtain

$$\begin{aligned}
&x\left(\delta^{in}(S \cup \{b\})\right) + x\left(\delta^{in}(S) \cap \delta^{out}(b)\right) - x_{a,b} - x_{b,c} + y \\
&\geq w^c\left(\delta^{in}(S \cup \{b\}) \setminus \{(a, b)\}\right) + w^c\left(\left(\delta^{in}(S) \cap \delta^{out}(b)\right) \setminus \{(b, c)\}\right) + y
\end{aligned}$$

and with (5.22)

$$\begin{aligned}
&x\left(\delta^{in}(S \cup \{b\})\right) + x\left(\delta^{in}(S) \cap \delta^{out}(b)\right) - x_{a,b} - x_{b,c} + y \\
&\geq w^c\left(\delta^{in}(S \cup \{b\})\right) + w^c\left(\delta^{in}(S) \cap \delta^{out}(b)\right) - w^c\left(\delta^{in}(b)\right) \\
&= \sum_{v \in S \cup \{b\}} w^c\left(\delta^{in}(v)\right) - w^c\left(E(S \cup \{b\})\right) + w^c\left(\delta^{in}(S) \cap \delta^{out}(b)\right) \\
&\quad - w^c\left(\delta^{in}(b)\right).
\end{aligned}$$

We now use (5.6),(5.7) and $w^c\left(\delta^{out}(c)\right) = 0$ to transform it to

$$\begin{aligned}
&x\left(\delta^{in}(S \cup \{b\})\right) + x\left(\delta^{in}(S) \cap \delta^{out}(b)\right) - x_{a,b} - x_{b,c} + y \\
&\geq 1 + \sum_{v \in S} w^c\left(\delta^{out}(v)\right) - w^c\left(E(S \cup \{b\})\right) + w^c\left(\delta^{in}(S) \cap \delta^{out}(b)\right) \\
&= 1 + w^c\left(\delta^{out}(S)\right) - w^c\left(\delta^{in}(b) \cap \delta^{out}(S)\right).
\end{aligned}$$

With the observation that $\delta^{in}(b) \cap \delta^{out}(S) \subseteq \delta^{out}(S)$ and (5.9) this shows

$$x\left(\delta^{in}(S \cup \{b\})\right) + x\left(\delta^{in}(S) \cap \delta^{out}(b)\right) - x_{a,b} - x_{b,c} + y \geq 1. \qquad \square$$

Unfortunately, constraint (5.22) does not imply the remaining lower bounds (5.13) and (5.14) as the following two examples show.

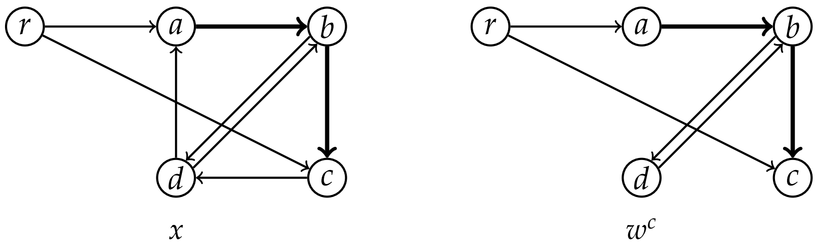


Figure 5.6: Illustration of one $x \in P_{arb}$ and one possible corresponding variable w^c. All edge weights in this figure are $\frac{1}{2}$.

Example 5. Inserting the values from Figure 5.6 in (5.13) we obtain

$$y \geq \frac{1}{2},$$

whereas the values inserted in McCormick's constraint (2.10) and the new constraint (5.22) only enforce

$$y \geq \frac{1}{2} + \frac{1}{2} - 1 = 0.$$

Example 6. Inserting the values from Figure 5.7 in (5.14) we obtain

$$\frac{1}{2} + y \geq 1,$$

whereas the values inserted in McCormick's constraint (2.10) and the new constraint (5.22) only enforce

$$y \geq \frac{1}{2} + \frac{1}{2} - 1 = 0 \qquad \text{and} \qquad y \geq \frac{1}{2} - \frac{1}{2} = 0.$$

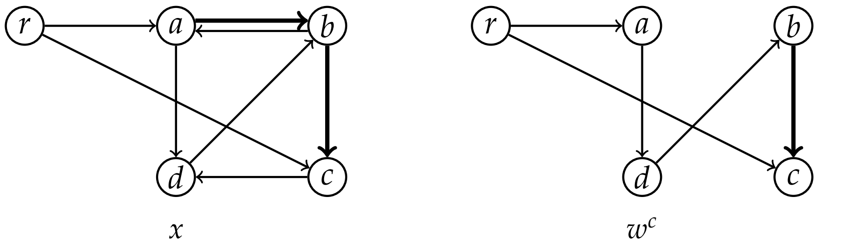


Figure 5.7: Illustration of one $x \in P_{arb}$ and one possible corresponding variable w^c. All edge weights in this figure are $\frac{1}{2}$.

Tail-tail case

Our studies focus on the head-tail case and we will regard only one idea for the tail-tail case and one for the nonadjacent case. They both result from studying Wong's variables related to the end nodes of the edges in the monomial to obtain upper bounds for y.

Given an r-arborescence A the corresponding w^a and w^c variables are the characteristic vectors for the r-a-and r-c-paths. If $w^a\left(\delta^{in}(c)\right) = 1$ the arborescence A contains an r-c-a-path that would build a cycle with $\left\{(b,a),(b,c)\right\}$ or any r-a-c-path thus $y = 0$ and $w^c\left(\delta^{in}(a)\right) = 0$. Analog $w^c\left(\delta^{in}(a)\right) = 1$ implies $y = 0$ and $w^a\left(\delta^{in}(a)\right) = 0$. Hence, we can add

$$y \leq 1 - w^a\left(\delta^{in}(c)\right) - w^c\left(\delta^{in}(a)\right). \tag{5.23}$$

to Wong's formulation.

Regarding the example from the proof of Proposition 16 in the former section (Figure 5.4) we can see in Figure 5.8 that (5.23) enforces $y \leq 0$. The following proposition shows how we can build an extended formu-

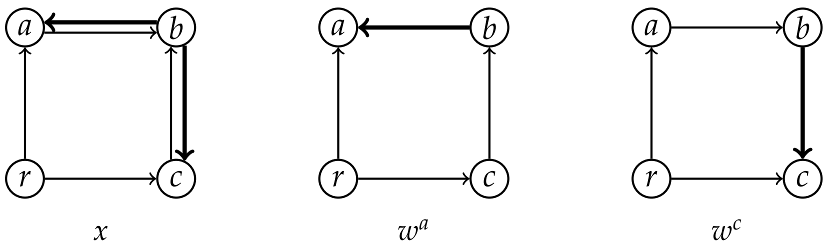


Figure 5.8: Illustration of one $x \in P_{arb}$ and the corresponding variables w^a and w^c. All edge weights in this figure are $\frac{1}{2}$.

lation base on Wong that is stronger than the formulation based on Martin in the sense that the projection of the first is contained in the projection of the second.

Proposition 22. *Wong's extended formulation* (5.5)-(5.9) *together with* (5.23) *and McCormick's linearization constraint* (2.8) *implies the quadratic subtour elimination constraints*

$$x(E(S)) + y \leq |S| - 1 \qquad \textit{for all } S \subset V \textit{ with } a, c \in S, b \notin S. \tag{2.13}$$

Proof. If $r \in S$ we simply combine (5.2) and (2.8) to

$$\begin{aligned} x(E(S)) &= \sum_{v \in S} x\left(\delta^{in}(v)\right) - x\left(\delta^{in}(S)\right) \\ &\leq |S| - 1 - x_{b,a} \\ &\leq |S| - 1 - y. \end{aligned}$$

Otherwise if $r \notin S$, we use the fact that w^a and w^c are flows of capacity 1 and can be decomposed into paths. Since S is an r-a cut and an r-c cut we observe with (5.5) that $x\left(\delta^{in}(S)\right)$ is at least the capacity of w^a without the paths including c plus the capacity of w^c without the paths including a. This leads to

$$x\left(\delta^{in}(S)\right) \geq 1 - w^a\left(\delta^{in}(c)\right) + 1 - w^c\left(\delta^{in}(a)\right)$$

and with (5.2) and (5.23)

$$\begin{aligned} x(E(S)) &= \sum_{v \in S} x\left(\delta^{in}(v)\right) - x\left(\delta^{in}(S)\right) \\ &\leq |S| - 1 - (1 - w^a\left(\delta^{in}(c)\right) - w^c\left(\delta^{in}(a)\right) \\ &\leq |S| - 1 - y. \end{aligned}$$

□

Nonadjacent case

We consider the case $M = \{(a,b),(c,d)\}$ for pairwise distinct a,b,c and d. Let A be an r-arborescence with the corresponding characteristic vectors of the r-b- and r-d-path w^b and w^d. If $w^b\left(\delta^{in}(d) \setminus \{(c,d)\}\right) = 1$, we know $(c,d) \notin A$ and $w^d\left(\delta^{in}(b) \setminus \{(a,b)\}\right) = 0$. Analogously, if $w^d\left(\delta^{in}(b) \setminus \{(a,b)\}\right) = 0$, it holds $w^b\left(\delta^{in}(d) \setminus \{(c,d)\}\right) = 0$ and $(a,b) \notin A$.

Hence, we can add

$$y \leq 1 - w^b\left(\delta^{in}(d) \setminus \{(c,d)\}\right) - w^d\left(\delta^{in}(b) \setminus \{(a,b)\}\right) \tag{5.24}$$

to Wong's formulation.

Regarding the example from the proof of Proposition 16 in the former section we see in Figure 5.9

$$w^b\left(\delta^{in}(d)\setminus\{(c,d)\}\right) = w^d\left(\delta^{in}(b)\setminus\{(a,b)\}\right) = \frac{1}{2}.$$

Inserting this into (5.24) we got $y \le 0$. Hence, this example lays in the

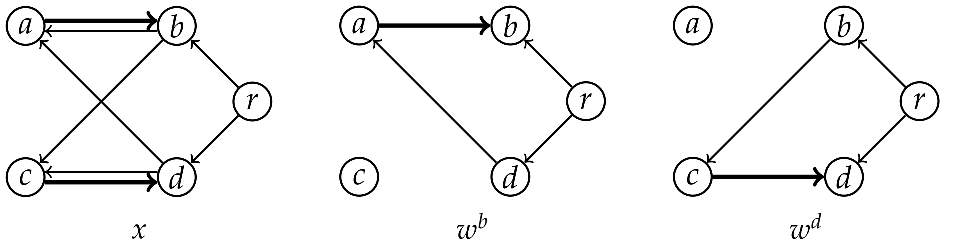


Figure 5.9: Illustration of one $x \in P_{arb}$ and the corresponding variables w^a and w^b. All edge weights in this figure are $\frac{1}{2}$.

projection of the formulation based on Martin but not in the projection of the formulation here. Unfortunately, we found an example the other way around, such that we can not say which formulation is better regarding the projection.

Proposition 23. *Wong's formulation* (5.5)-(5.9) *together with* (5.24) *does not imply the quadratic subtour elimination constraints*

$$x(E(S_1)) + x(E(S_2)) + y \le |S_1| + |S_2| - 2 \tag{2.14}$$

for all $S_1, S_2 \subset V$ where (a,b) and (b,c) have one end node in S_1 and the other in S_2.

Proof. We regard the example in Figure 5.10 and observe

$$w^b\left(\delta^{in}(d)\setminus\{(c,d)\}\right) = w^d\left(\delta^{in}(b)\setminus\{(a,b)\}\right) = 0.$$

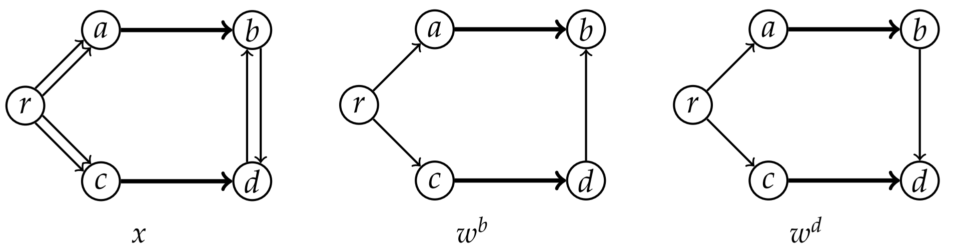


Figure 5.10: Illustration of one $x \in P_{arb}$ and the corresponding variables w^b and w^d. All edge weights in this figure are $\frac{1}{2}$.

Hence, Wong's formulation with (5.24) only implies $y \leq 1$, whereas (2.14) for $S_1 = \{r, a, c\}$ and $S_2 = \{b, d\}$ enforces

$$y \leq 0. \qquad \square$$

6 Conclusion and Outlook

We investigated higher order polytopes related to combinatorial optimization problems with polynomials as objective functions. Our main focus laid in forest and spanning tree problems.

In Chapter 2 we presented extended formulations for higher order forest polytopes that are based on known extended formulations for the spanning tree polytope. Those formulations imply extended formulations for higher order spanning tree polytopes. To model the monomials we used the structural information provided by the known formulations and constructed new constraints. For the polytopes with one degree-two monomial we only needed two new inequalities. Hence, we got very small and easy formulations.

In Section 2.2 we generalized the constraints for degree-three monomials. To build a complete description we also needed the constraints for all degree-two submonomials leading to 14 additional inequalities, six for the degree-two submonomials and eight degree-three specific ones.

One might wonder whether we can generalize this for monomials of higher degree. The author did indeed generalize inequality (2.44) for degree-four monomials, but did not prove, whether this also leads to a complete description of the corresponding higher order forest polytopes. We guess that in this case one would again need the constraints for all degree-two and degree-three submonomials leading to $12 + 32$ additional inequalities only for the submonomials. Hence, the formulation becomes rather large.

Considering the unexpected bad performance of Gurobi on our extended formulation with degree-two monomials including the general form of the extended linearization constraints in Chapter 4, we question, whether investigations in generalizations of this formulation for monomials of higher degree are worth the effort. Before discussing more complicated formulations, we would prefer to understand why the performance of Gurobi using the formulation `q-m` including nonadjacent monomials was that much worse than the other formulations, although the root gap was better.

In Section 3.1 we showed that the combination of our small extended formulations for higher order forest polytopes with only one degree-two monomial for all degree-two monomials leads to a better relaxation of the quadratic forest polytope than the combination of the descriptions in the original space. This effect is due to the natural identification of the additional variables in our extended formulations. Thus, we modeled a relation between the monomials in an implicit and automatic way.

To obtain a feeling of the amount of improvement in the relaxations we measured the root gap of random QMST instances in Chapter 4. We observed that the improvement arising from that relation we build in the extension variables is clearly stronger than the improvement obtained in the original space by adding the quadratic subtour elimination constraints to the relaxation that only uses McCormick's linearization constraints and a description of the spanning tree polytope.

Considering the measurements related to the performance in Gurobi (node count, runtime and gap) we observed that a smaller root gap does not necessarily imply a better performance. The performance using our small extended formulations was in many cases worse than the performance on the larger formulations in the original space using fast separation routines. On the instances including only adjacent monomials the performance on our extended formulations was slightly better than on the other formulations, but this might be due to the fact that the instances are smaller (since they include less monomials) than the instance including all monomials. Overall, we have to keep in mind that we ran Gurobi with standard parameters. Although one might improve the performance adjusting them, there always remains a lot of randomness in the performance and we should be very careful in judging a formulation based on such restricted tests. It might be that on different solvers or with different parameters we would observe very different performances.

The projection of our combined extended formulations led to new valid inequalities for the quadratic forest polytope as described in Section 3.1. Those inequalities can be considered as strengthening of Buchheim and Klein's quadratic subtour elimination constraints [4]. In the adjacent case both have the form

$$x\left(E\left(S\right)\right)+\sum_{M\in\mathcal{M}(S)} y_M \leq |S|-1 \tag{6.1}$$

for some node sets S and for specific sets of adjacent monomials $\mathcal{M}(S)$. For general monomials they have the form

$$x\left(E\left(S_1\right)\right)+x\left(E\left(S_2\right)\right)+\sum_{M\in\mathcal{M}(S_1,S_2)} y_M \leq |S_1|+|S_2|-2 \tag{6.2}$$

for some node sets S_1, S_2 and for specific sets of monomials $\mathcal{M}(S_1, S_2)$.

In Section 3.2 we asked for which sets $\mathcal{M}(S)$ the inequalities (6.1) are valid for the adjacent quadratic forest polytope. We elaborated properties

$\mathcal{M}(S)$ should fulfill. One class of such sets arises from the projection of our combined extended formulations in Section 3.1. An open question is how we can describe further sets of monomials that are maximal in the sense that adding a monomial would hurt the desired properties. Such sets could lead to new facets of the adjacent quadratic forest polytope and the adjacent quadratic spanning tree polytope as it does for those we described in this work.

One might wonder whether the inequalities (6.2) in the general case imply facets of the quadratic forest polytope or the quadratic spanning tree polytope, too. The author used IPO [36] to calculate random facets of those polytopes for the complete graphs with four, five and six nodes and could not find any facet of the form described above. Hence, our guess is that in the general case, where we also have nonadjacent monomials, there are no inequalities of the form in (6.2) that describe facets of the quadratic forest polytope or the quadratic spanning tree polytope.

In Chapter 5 we compared two extended formulations for the arborescence polytope with respect to their potential to model degree-two monomials. We build extended formulations that project onto relaxations of the higher order arborescence polytopes with one degree-two monomial ($P_{arb}(\{M\})$). In the case of adjacent monomials our formulations based on Wong's extended formulation is stronger than our formulation based on Martin's formulation. In the nonadjacent case neither the projection of the formulation based on Martin is contained in the projection of the formulation based on Wong nor the other way around. To improve the relaxation in the nonadjacent case one might combine both formulations or improve the formulation based on Wong in the future.

All in all, Wong's variables contain more information regarding the direction of the edges in the monomials and we were able to model more relations. In the head-tail case we illustrated this by implying further facets of $P_{arb}(\{M\})$. Although we were not able to build complete descriptions we could observe that Wong's formulation contains a lot of structural information, which can be used to model extended formulations of relaxations of $P_{arb}(\{M\})$. This can help to find or understand facets of $P_{arb}(\mathcal{M})$ where we do not know a complete description for any nonempty set of monomials yet.

Generally, we hope that this work motivates to use extended formulations and especially the structural information they provide in modelling further structures or relations. We did this for monomials and relations between them. Those relations became clear in the extended space using the provided structural information of known extended formulations. This can be seen for example in the head-tail case in Section 5.2, where we calculated facets in the original space, which are hard to understand with respect to their combinatorial meaning, but we could imply those facets by easy to understand relations in the extended space.

Bibliography

[1] Arjang Assad and Weixuan Xu. The quadratic minimum spanning tree problem. *Naval Research Logistics (NRL)*, 39(3):399–417, 1992.

[2] Benjamin Assarf, Ewgenij Gawrilow, Katrin Herr, Michael Joswig, Benjamin Lorenz, Andreas Paffenholz, and Thomas Rehn. Computing convex hulls and counting integer points with `polymake`. *Mathematical Programming Computation*, 9(1):1–38, 2017.

[3] E. Balas. Disjunctive programming and a hierarchy of relaxations for discrete optimization problems. *SIAM Journal on Algebraic Discrete Methods*, 6(3):466–486, 1985.

[4] Christoph Buchheim and Laura Klein. Combinatorial optimization with one quadratic term: Spanning trees and forests. *Discrete Applied Mathematics*, 177(0):34 – 52, 2014.

[5] Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. Extended formulations in combinatorial optimization. *4OR*, 8(1):1–48, 2010.

[6] Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. *Integer Programming*. 2012.

[7] Michele Conforti, Volker Kaibel, Matthias Walter, and Stefan Weltge. Subgraph polytopes and independence polytopes of count matroids. *Operations Research Letters*, 43(5):457 – 460, 2015.

[8] Roberto Cordone and Gianluca Passeri. Solving the quadratic minimum spanning tree problem. *Applied Mathematics and Computation*, 218(23):11597–11612, 2012.

[9] Ante Ćustić, Ruonan Zhang, and Abraham P. Punnen. The quadratic minimum spanning tree problem and its variations. *arXiv preprint arXiv:1603.04451*, 2016.

[10] George Dantzig, Ray Fulkerson, and Selmer Johnson. Solution of a large-scale traveling-salesman problem. *Journal of the operations research society of America*, 2(4):393–410, 1954.

[11] Robert P Dilworth. A decomposition theorem for partially ordered sets. *Annals of Mathematics*, pages 161–166, 1950.

[12] Jack Edmonds. Optimum branchings. *Journal of Research of the National Bureau of Standards B*, 71:233–240, 1967.

[13] Jack Edmonds. Matroids and the greedy algorithm. *Mathematical programming*, 1(1):127–136, 1971.

[14] Anja Fischer and Frank Fischer. Complete description for the spanning tree problem with one linearised quadratic term. *Operations Research Letters*, 41(6):701 – 705, 2013.

[15] Anja Fischer, Frank Fischer, and S. Thomas McCormick. Matroid optimization problems with monotone monomials in the objective. *Preprint-Serie des Instituts für Numerische und Angewandte Mathematik Universität Göttingen*, 2017-6, 2017.

[16] Anja Fischer, Frank Fischer, and S. Thomas McCormick. Matroid optimisation problems with nested non-linear monomials in the objective function. *Mathematical Programming*, 169(2):417–446, 2018.

[17] Python Software Foundation. Python language reference. `http://www.python.org`, 2018.

[18] Ewgenij Gawrilow and Michael Joswig. `polymake`: a framework for analyzing convex polytopes. In *Polytopes—combinatorics and computation*, pages 43–73. Birkhäuser, Basel, 2000.

[19] Ralph E Gomory et al. Outline of an algorithm for integer solutions to linear programs. *Bulletin of the American Mathematical society*, 64(5):275–278, 1958.

[20] LLC Gurobi Optimization. Gurobi optimizer reference manual. `http://gurobi.com`, 2018.

[21] LLC Gurobi Optimiztation. Benchmarks. `http://www.gurobi.com/resources/benchmarks`, 2019.

[22] Lena Hupp, Laura Klein, and Frauke Liers. An exact solution method for quadratic matching: The one-quadratic-term technique and generalisations. *Discrete Optimization*, 18:193 – 216, 2015.

[23] Volker Kaibel. Extended formulations in combinatorial optimization. *Optima*, 85:2–7, 2011.

[24] Laura Klein. *Combinatorial optimization with one quadratic term*. PhD thesis, TU Dortmund, 2014.

[25] Bernhard Korte and Jens Vygen. *Combinatorial Optimization*. Springer, 2012.

[26] Joseph B Kruskal. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical society*, 7(1):48–50, 1956.

[27] Ailsa H Land and Alison G Doig. An automatic method for solving discrete programming problems. In *50 Years of Integer Programming 1958-2008*, pages 105–132. Springer, 2010.

[28] Jon Lee and Janny Leung. On the boolean quadric forest polytope. *INFOR: Information Systems and Operational Research*, 42(2):125–141, 2004.

[29] R Kipp Martin. Using separation algorithms to generate mixed integer model reformulations. *Operations Research Letters*, 10(3):119–128, 1991.

[30] Garth P McCormick. Computability of global solutions to factorable nonconvex programs: Part I- Convex underestimating problems. *Mathematical programming*, 10(1):147–175, 1976.

[31] Tiago P. Peixoto. The graph-tool python library. *figshare*, 2014.

[32] Dilson Lucas Pereira and Alexandre Salles da Cunha. Polyhedral results, branch-and-cut and lagrangian relaxation algorithms for the adjacent only quadratic minimum spanning tree problem. *Networks*, 71(1):31–50, 2018.

[33] Robert Clay Prim. Shortest connection networks and some generalizations. *Bell Labs Technical Journal*, 36(6):1389–1401, 1957.

[34] Alexander Schrijver. *Combinatorial Optimization – Polyhedra and Efficiency*. Springer, 2003.

[35] Matthias Walter. Complete description of matching polytopes with one linearized quadratic term for bipartite graphs. *arXiv preprint arXiv:1607.01880*, 2016.

[36] Matthias Walter. Ipo – investigating polyhedra by oracles. `http://polyhedra-oracles.bitbucket.org`, 2016.

[37] Justin C. Williams. A linear-size zero-one programming model for the minimum spanning tree problem in planar graphs. *Networks*, 39(1):53–60, 2002.

[38] Laurence A. Wolsey. Using extended formulations in practice. *Optima*, 85:7–9, 2011.

[39] Richard T Wong. Integer programming formulations of the traveling salesman problem. In *Proceedings of the IEEE international conference of circuits and computers*, pages 149–152. IEEE Press Piscataway, NJ, 1980.

List of Notations

Sets

2^E The power set of E. 1, 2, 9, 11, 20, 21, 29, 33, 37, 47, 68, 72, 96

$[k]$ The set $\{1, \ldots, k\}$. 1, 4, 9, 15, 16, 21–28, 32

Vectors

$\chi(T)$ The characeristic vector of T with $\chi(T) \in \{0,1\}^n$ and $\chi(T)_e = 1$ if and only if $e \in T$. 1, 7, 12, 13, 28, 32, 71–73, 79

$\langle c, x \rangle$ The scalar product $\sum c_i x_i$. 48, 49

$x(D)$ The sum $\sum_{e \in D} x_e$. 7, 8, 10, 12–16, 18, 21, 28, 31, 34, 35, 37–40, 43–45, 67–70, 72, 74–83, 86

Polytopes

conv X The convex hull of X. 1, 7, 13, 67

$P_{\mathcal{T}}$ The polytope $P_{\mathcal{T}}(\varnothing)$. 1–5, 7–9, 11, 12, 16–18, 21, 23, 28, 33, 67–72, 75, 80, 81, 83, *see* $P_{\mathcal{T}}(\mathcal{M})$

$P_{\mathcal{T}}(\mathcal{M})$ The higher order polytope with monomials $\mathcal{M} \subseteq 2^E$, where E is some finite basic set and $\mathcal{T} \subseteq 2^E$ is a set of combinatorial structures like for example

$P_F(\mathcal{M})$ forests,

$P_{ST}(\mathcal{M})$ spanning trees or

$P_{arb}(\mathcal{M})$ arborescences

of a graph $G = (V, E)$. 1–5, 7, 9, 11–13, 15, 18, 20, 22, 23, 26, 28, 29, 32, 33, 38, 47, 68, 70–72, 74, 76, 77, 87, *see* spanning tree, forest & arborescence

$\mathrm{xc}(P)$ The extension complexity of P, i.e., the smallest possible size of an extension of P. 2, 4, 23, *see* extension complexity

Graph theory

$\delta(v)$ The set of all edges adjacent to a node $v \in V$. For a set of nodes $S \subset V$ we define

$$\delta(S) := \bigcup_{v \in S} \delta(v) \setminus E(S).$$

7, 8, 17–21, 43, 45, *see* $E(S)$

$\delta^{in}(v)$ The set of all edges adjacent to v and directed into v. For a set of nodes $S \subset V$ we define

$$\delta^{in}(S) := \bigcup_{v \in S} \delta^{in}(v) \setminus E(S).$$

8, 12, 14–16, 18, 28, 31, 34, 37, 67–69, 72–83, *see* $E(S)$

$\delta^{in}(v)$ The set of all edges adjacent to v and directed into v. For a set of nodes $S \subset V$ we define

$$\delta^{out}(S) := \bigcup_{v \in S} \delta^{out}(v) \setminus E(S).$$

8, 74–80, *see* $E(S)$

$E(S)$ The set of all edges in E that have both end nodes in S. 7, 10, 12, 13, 15, 21, 24, 26–31, 35, 38–40, 42–45, 67, 69, 70, 72, 76, 77, 82, 83, 86

$V(D)$ The set of all nodes in D, i.e.,

$$\{v \in V \,|\, v \in e \text{ for some } e \in D\}.$$

25–27, 34, 39

Complexity theory

$O(g)$ The big o-notation defined as: $f \in O(g)$ if and only if f is bounded above by g asymptotically. 5, 23

$\Theta(g)$ The big theta notation defined as: $f \in \Theta(g)$ if and only if f is bounded above and below by g asymptotically. 5, 7, 16, 47

Rounding operators

$\lceil x \rceil$ The smallest integer $\overline{x}$ with $\overline{x} \geq x$. 48, 49

$\lfloor x \rfloor$ The biggest integer $\underline{x}$ with $\underline{x} \leq x$. 48, 49

Glossary

arborescence A cycle free edge set such that each node $v \neq r$ has exactly one ingoing edge and the root node r has no ingoing edge. To explicitly name the root node we call it also r-arborescence. 3, 8–10, 12, 14, 19–22, 26, 28, 32, 67, 68, 71, 73–76, 79, 81, 82

characteristic vector The 0-1 vector $\chi(T)$. 1, 8–10, 12, 14, 19–22, 28, 32, 67, 68, 73, 75, 76, 79, 81, 82, *see* $\chi(T)$

extended formulation A description of an extension in terms of equations and inequalities. iii, 2–5, 7–9, 11–14, 16–18, 20, 23, 26, 29, 33, 37, 47, 50, 51, 58, 59, 62, 64, 65, 67, 68, 74, 75, 81, 82, 85–87, *see* extension

extension An extension of a polytope P is a polytope in a higher dimension that can be projected onto P. 2, 33

extension complexity The extension complexity of a polytope is the size of the smallest possible extension. 2, 3, 23, *see* $\mathrm{xc}(P)$, extension & size

forest A cycle free edge set. 1–3, 5, 7–9, 12, 19, 24–26, 28, 32, 35, 38, 39, 45

higher order characteristic vector A vector (x, y), where x is a characteristic vector and the entries in y are monomials in x. iii, 1, 40, 45, *see* characteristic vector & $\chi(T)$

higher order polytope The convex hull of higher order characteristic vectors. iii, 1–5, 9, 18, 23, 33, 34, 67, 68, 85, 87, *see* higher order characteristic vector & $P_{\mathcal{T}}(\mathcal{M})$

integer program (IP) A linear program together with the integrality constraint for all variables. 2, 3, 48, 58, 59, 96, *see* LP

linear program (LP) An optimization problem formulated with a linear objective and linear constraints in terms of equations and inequalities. 48

LP relaxation The linear program arising from a (mixed) integer program by dropping the integrality constraint. 47–50, 57, 59, *see* LP, IP & MIP

mixed integer program (MIP) Like IP but only some variables are constrainted to by integral. 1, 2, 48–50, 54, 56–58, 96, *see* IP

monomial In this work we call a set $M \in 2^E$ monomial due to its relation to $\prod_{e \in M} x_e$. iii, 1–5, 7, 9, 18, 23–25, 29, 31, 33–36, 38–41, 43, 47, 50, 51, 53, 54, 56, 58–65, 68–70, 74, 81, 85–87

planar graph A graph that has an embedding in the plane such that no two edges crosses each other. 5, 16, 17, 20

QMST-problem The quadratic minimum spanning tree problem. iii, 3, 33, 47, 50, 56, 86

relaxation A relaxation of an (integer) polytope P is a polytope $R \supseteq P$ such that $R \cap \mathbb{Z}^n = P \cap \mathbb{Z}^n$. iii, 1–3, 5, 33, 37, 47, 85–87

root gap The relative gap between the (mixed) integer solution and the solution of the LP relaxation of a MIP or IP. 2, 3, 47, 57, 59, 62, 65, 86, *see* LP relaxation, MIP & IP

separation Given a polytope $P \subset \mathbb{R}^n$ and some point $x^* \in \mathbb{R}^n$ a separation solves the problem to decide whether $x^* \in P$ and if $x^* \notin P$ it provides a violated constraint. 13, 47, 49, 54, 55, 59, 62, 65, 86

size In the context of extended formulations the size is the number of inequalities. 2–5, 7, 16, 18, 47, *see* extended formulation

spanning tree A tree that spans all nodes of a graph. 2, 3, 7–10, 12, 19, 20, 28, 32, 35–37, 43, 67, 73, 74, *see* tree

tree A forest that is connected. iii, 8, 10, 19, 21, 22, 25, 26, 40–44, 69, 71, *see* forest