Describing Orbitopes by Linear Inequalities and Projection Based Tools

Dissertation
zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

von Dipl. Jour. Andreas Loos

geboren am 27. April 1973 in Erlangen
genehmigt durch die Fakultät für Mathematik
der Otto-von-Guericke-Universität Magdeburg
Gutachter: Prof. Dr. Volker Kaibel
Prof. Dr. Marc Pfetsch
eingereicht am 1. Dezember 2010
Verteidigung am 28. April 2011

Zusammenfassung

Es sei $\mathcal{M}_{p,q}$ die Menge aller 0/1-Matrizen mit p Zeilen und q Spalten. Lässt man eine Permutationsgruppe auf den Spalten der Elemente von $\mathcal{M}_{p,q}$ operieren, dann zerfällt $\mathcal{M}_{p,q}$ in eine Menge von Orbits. Die lexikografisch maximalen Repräsentanten all dieser Orbits bilden die Ecken der Polytope, die in dieser Doktorarbeit untersucht werden, der so genannten *Orbitope*. Spezifiziert man die Anzahl der 1-Einträge in jeder Zeile, dann lassen sich Packungs- (Zeilensumme $\leqslant 1$), Partitionierungs- (Zeilensumme $= 1$), Überdeckungs- (Zeilensumme $\geqslant 1$) und volle Orbitope (beliebige Zeilensumme) unterscheiden. Ein Hauptaugenmerk dieser Doktorarbeit liegt auf der Untersuchung von Orbitopen über der vollen symmetrischen Gruppe, insbesondere im Hinblick auf deren *lineare Beschreibungen*.

Wir zeigen, dass es möglich ist, über vollen Orbitopen über der vollen symmetrischen Gruppe in polynomieller Zeit zu optimieren, während die Optimierung über den anderen Orbitopen im Allgemeinen $\mathcal{NP}$-schwer ist. Um Polytope linear zu beschreiben, entwickeln wir Methoden in Zusammenhang mit der Projektion erweiterter Beschreibungen, etwa das *faithful sectioning* oder *branched polyhedral systems*. Wir erhalten für einen Spezialfall der vollen Orbitope, den Orbisack, vollständige lineare Beschreibungen. Wie gezeigt werden wird, sind die entwickelten Werkzeuge nicht nur auf die Verwendung mit Orbitopen beschränkt, sondern von allgemeinerem Interesse, etwa im Zusammenhang mit Stabile-Mengen-Polytopen.

Abstract

Let $\mathcal{M}_{p,q}$ the set of all 0/1-matrices with p rows and q columns. If you let a permutation group operate on the columns of the elements of $\mathcal{M}_{p,q}$, then $\mathcal{M}_{p,q}$ can be partitioned by a set of orbits. The lexicographic maximal representatives of all these orbits are the vertices of the polytopes covered in this thesis, the so called *orbitopes*. Specifying the number of 1s in each row, one distinguishes packing (rowsum $\leqslant 1$), partitioning (rowsum $= 1$), covering (rowsum $\geqslant 1$) and full orbitopes (arbitrary rowsum). One main interest of this thesis is to study orbitopes over the full symmetric group, particularly with regard to their *linear descriptions*.

We show that it is possible to optimize over full orbitopes over the full symmetric group in polynomial time, while optimization over the other orbitopes is in general $\mathcal{NP}$-hard. To linearly describe polytopes, we develop methods in connection with the projection of extended formulations, for instance the *faithful sectioning* or *branched polyhedral systems*. For a special case of a full orbitope, the orbisack, we obtain complete linear descriptions. As will be shown, the tools are not restricted to be used with orbitopes only. Instead, they are of more general interest, for instance in connection with stable set polytopes.

Acknowledgements

Without any doubt, there are not many advisors like Volker Kaibel. In the six years since I met him, I can barely remember one moment when he didn't have time for me and my ideas. His brain is almost always open (as Erdös would have said), and his optimism is irrepressible. Apart from combinatorics, linear optimization and tools to tackle problems, he taught me a lot about how mathematics is presented. And he reminded me more than once that doing mathematics includes knowing how to pace oneself. I would like to thank him for all of this.

I would also like to thank Günter M. Ziegler: he granted me a fellowship from his Leibniz Fund and supplied me with a desktop, coffee and hints. I really enjoyed the atmosphere at his institute which is international, energetic, witty and (especially in summer) hot.

Matthias Peinhardt did some mathematics with me at the blackboard and in the train to and from Magdeburg. I would like to thank him, as well as Rüdiger Stephan, who talked to me about mathematics at rather unusual places, for instance at the Zoo and at several playgrounds in Berlin. He also read some part of this thesis and gave me some very helpful hints.

I would like to thank Katie Bedford for reading the English between the formulas and pointing out typos and mistakes.

Last, I would like to thank Tina Heidborn who sometimes had a hard job to give me unfailing support.

Berlin, June 2011 Andreas Loos

Contents

Chapter 1

Introduction

1.1 Outline

Solving a combinatorial optimization problem basically means choosing an optimal subset from a given ground set of elements, on condition that the subset satisfies certain properties. Often, one describes the choice of elements by binary variables indicating whether or not an element has been chosen, that is: any feasible solution is encoded by a 0/1-vector in a space of appropriate dimension. A set of linear constraints can then describe the feasible domain. In many cases, one finds that after assigning weights to the elements a solution is optimal if and only if the total sum of weights of chosen elements is maximal or minimal. In these cases, one is able to express the combinatorial problem as an integer program (IP). In lucky cases, one can even formulate it as a linear program (LP).

Probably the most prominent example of a combinatorial problem with LP-formulation is the problem of weighted perfect matching. For a given graph $G = (\mathcal{V}, \mathcal{E})$ with weights on the edges, a matching is a subset of pairwise disjunct edges in $\mathcal{E}$. If the matching has cardinality $\frac{|\mathcal{V}|}{2}$, it is called perfect. Its weight is the total sum of the weights of the edges in the matching. Jack Edmonds ([31]) showed almost half a century ago that a vector $\boldsymbol{x} \in \mathbb{R}^{\mathcal{E}}$ is the incidence vector of a perfect matching if and only if it is an extreme point of the domain described by the following set of inequalities

$$\begin{aligned} x_e &\geqslant 0 && \forall\, e \in \mathcal{E} \\ \textstyle\sum_{e\in\delta(v)} x_e &= 1 && \forall\, v \in \mathcal{V} \\ \textstyle\sum_{e\in\delta(\mathcal{U})} x_e &\geqslant 1 && \forall\, \mathcal{U} \subseteq \mathcal{V} \text{ odd}, \end{aligned} \tag{1.1}$$

where $\delta(v)$ denotes the set of edges incident with node v. So, a formerly combinatorial problem of finding a perfect matching of maximum edge weight has been turned into a linear problem and can be solved with standard tools like the simplex algorithm, the ellipsoid method or interior point methods. Using an appropriate separation algorithm, one can therefore solve the problem in polynomial time (despite the fact that the number of inequalities is growing exponentially with the number of edges). On the other hand, the inequalities from above can be obtained from Edmonds' combinatorial algorithm for the weighted matching problem (see [107], 26.3b; for more information about polyhedral techniques in combinatorial optimization, see for instance Aardal and Hoesel's overview [1] and [2] for an annotated bibliography).

Generally speaking, geometry, solving the problem algorithmically and gaining structural (combinatorial) insights to the problem go often hand in hand.

It is not unusual for these (IP or LP-)formulation of combinatorial problems to bear symmetries. That means: in every solution, one can perform certain permutations of the components of every solution without changing the objective value and feasibility. In 2005, Volker Kaibel and Marc Pfetsch focused on the shape of constraints that break certain classes of symmetry, and how knowledge about it could be exploited when solving symmetric problems. At first, they considered packing and partitioning problems, where objects are assigned to at most or exactly one from a set of properties, respectively, on the assumption that certain permutations of properties also yield a feasible solution. The standard example is the problem of coloring a graph $G = (\mathcal{V}, \mathcal{E})$ with as few colors as possible taken from a set $\mathcal{C}$ of colors. The default IP-model uses binary variables $x_{v,c}$ telling whether node v has been colored with color c, as well as binary variables y_c telling whether color c has been used at all:

$$\begin{aligned} &\text{minimize } \sum_{c \in \mathcal{C}} y_c \text{ such that} \\ &\qquad \sum_{c \in \mathcal{C}} x_{v,c} = 1 && \forall v \in \mathcal{V} \\ &\qquad x_{v,c} + x_{w,c} \leqslant y_c && \forall \{v, w\} \in \mathcal{E} \text{ and } c \in \mathcal{C} \\ &\qquad \boldsymbol{x} \in \{0,1\}^{\mathcal{V} \times \mathcal{C}} \\ &\qquad \boldsymbol{y} \in \{0,1\}^{\mathcal{C}} \end{aligned}$$

Clearly, relabeling the colors does not change the structure of the solution. Hence, the space of feasible solutions is larger than necessary, which can in particular become interesting when using branching algorithms. To refine the search by breaking the symmetry, one could consider only solutions $(\boldsymbol{x}, \boldsymbol{y})$ where $\boldsymbol{x}$ has lexicographically ordered columns. (Note that vector $\boldsymbol{x}$ has the form of a matrix.)

This immediately raises some questions: which inequalities have to be added to the inequality system above to break this class of symmetry? Can we gain structural insights in symmetry breaking from these inequalities, and can this knowledge be exploited, for example for improving branch and bound algorithms?

Kaibel and Pfetsch called the polytopes defined by the symmetry-breaking inequalities "orbitopes", since the respective inequalities select one representative from each orbit arising from the operation of the permutations on the variables the problem is described in.

They obtained non-redundant linear descriptions of packing and partitioning orbitopes (that is: orbitopes connected with assignment problems of the packing and partitioning type) with the full symmetric group and the cyclic group operating on the columns of the solutions ([65]). Their study of packing and partitioning orbitopes also led to ways to fix variables in branch and bound algorithms in order to break these symmetries in the solutions ([64]). The question remained whether other orbitopes can be linearly described, in particular the convex hull of all 0/1-matrices with lexicographically ordered columns (the so-called "full orbitope over the full symmetric group"). This case is particularly interesting since the group is large and has a more complex structure than for instance the cyclic group.

This thesis investigates the possibilities for linear descriptions of certain classes of orbitopes. For a special case of the full orbitope, the orbisack, the complete non-redundant linear description is derived.

The tools used for these purposes are presented in a general overview in chapter 2. They are not customized to orbisacks, although they fit well with them. Therefore, we develop in section 2.4 the linear descriptions of a number of other polytopes with the method of *faithful sectioning*, in particular clique polytopes for bounded clique size and path set polytopes on acyclic digraphs. This work arose in collaboration with Volker Kaibel and Matthias Peinhardt.

In chapter 3, we define orbitopes in detail and show that for many classes of orbitopes, (the decision problem associated with) optimization over orbitopes is in fact $\mathcal{NP}$-complete. This is in general true for k-packing, k-partitioning, and k-covering orbitopes over the full symmetric group (which are the convex hulls of 0/1-matrices with columns in lexicographic order and less than, exactly, or more than k 1-entries in each row, respectively, where $k \geqslant 1$).

In contrast to this, we show that it is possible to optimize in polynomial time over full orbitopes over the full symmetric group. We develop a dynamic programming algorithm for linear optimization over $\mathbf{O}_{p,q}$ running in time $\mathcal{O}(pq^3)$. Figure 3.2 shows an overview over complexity of optimization and linear descriptions for orbitopes over the full symmetric group. Note that the parts of the work concerning $\mathcal{NP}$-completeness of optimization over orbitopes also emerged from joint work with Volker Kaibel and Matthias Peinhardt.

The algorithm for optimization over full orbitopes is interesting not because of any potential practical applications, but in connection with the search for "nice" linear descriptions. This does not concern the number of facets: Grötschel, Lovász and Schrijver showed ([52]) that a linear problem can be solved in polynomial time, provided there is an appropriate separation oracle. This implies that even for polynomial problems, the number of facets can definitely grow exponentially with the problem size — the perfect matching problem from above is an example for this.

Instead, the question is if one can say in polynomial time that a given inequality is part of the complete linear description of the feasible domain. Karp and Papadimitriou ([67]) were putting this question and the complexity of optimization in the following relation (see also [107] for an outline of their results): Let Π a combinatorial problem with linear objective, and let each instance $\sigma \in \Pi$ be formulated as an IP over a polyhedron $\mathbf{P}_\sigma \in \mathbb{R}^{m_\sigma}$. Then, the decision problem

$$\text{Given } \sigma \in \Pi,\ \boldsymbol{c} \in \mathbb{Q}^{m_\sigma},\ \text{and } k \in \mathbb{Q},\ \text{is there an } \boldsymbol{x} \in \mathbf{P}_\sigma \text{ with } \langle \boldsymbol{c}, \boldsymbol{x} \rangle > k?$$

is in co-$\mathcal{NP}$ if and only if for each σ, a description $\mathcal{I}_\sigma$ of $\mathbf{P}_\sigma$ exists such that problem

$$\text{Given } \sigma \in \Pi,\ \boldsymbol{c} \in \mathbb{Q}^{m_\sigma},\ \text{and } \ell \in \mathbb{Q},\ \text{does inequality } \langle \boldsymbol{c}, \boldsymbol{x} \rangle \leqslant \ell \text{ belong to } \mathcal{I}_\sigma?$$

belongs to $\mathcal{NP}$. This implies that if the (decision problem associated with the) combinatorial problem is $\mathcal{NP}$-complete and $\mathcal{NP} \neq$ co-$\mathcal{NP}$, then it is an $\mathcal{NP}$-complete problem to decide whether an inequality belongs to the linear description of the feasible domain of the problem. Roughly speaking, a "nice" linear description can be expected only for polytopes over which one can optimize in polynomial time.

Hence, there is hope for a "nice" linear description of full orbitopes over the full symmetric group, although computer experiments indicate that for these orbitopes, the linear descriptions seem to be much more complicated than for packing and partitioning orbitopes. A complete linear description for the full orbitope $\mathbf{O}_{p,q}$ exists so far only in case that the vertices have two columns, i.e. $q = 2$.

This special case is the so-called *orbisack*. In chapter 4, we present three different ways to obtain linear descriptions of these orbitopes. The motivation was to find a proof that could be inductively or in other ways extended to orbitopes with more than two columns. However, this hope has not been fulfilled. Any of the three proofs has its idiosyncrasies that prohibit to extend the proof to full orbitopes in general.

- The combinatorial proof relies on the fact that for $q = 2$, the coefficients of facet inducing inequalities follow a certain sign pattern. This is not true for $q > 2$ anymore, since our experiments show that for $q > 2$, there exist different facets with the same sign pattern.
- The proof by faithful sectioning relies on the fact that each vertex of an orbisack has either only rows $(1, 1)$ or $(0, 0)$, or it has a unique row $(1, 0)$ that

"ensures" that the columns of the vertex are in lexicographic order. This fact – leading to the notion of "critical row" – enables us to find nice extended formulations for the orbisack, which are the prerequisite for the method of faithful sectioning. However, for $q > 2$, our search for extended formulations brought only two formulations to light: one is derived from the dynamic programming algorithm, and the other one is related to branched polyhedral systems (BPS). Unfortunately, it is already in the case of orbisack unclear how to construct the linear description *directly* from these extended formulations.
- The last proof is an application of work of Weismantel and Pochet concerning the linear description of sequential knapsack polytopes ([117]). It relies on the fact that $\mathbf{O}_{p,2}$ is a sequential knapsack polytope — however, $\mathbf{O}_{p,q}$ is for $q > 2$ and $p > 1$ not even a knapsack anymore.

Note that only the first two proofs lead to a non-redundant description. We conclude chapter 4 by studying the graph of the orbisack and characterizing the adjacency structure. The hope was to show that the edge expansion of the orbisack is bounded from below by 1. However, despite the fact that the graph has an appealing structure, we did not succeed in proving this lower bound (which we conjecture to be tight for orbisacks).

In chapter 5, we present the tool of *branched polyhedral systems* (BPS), which can be used to obtain an extended formulation for orbitopes. Different from the other extended formulations presented in the chapters before, it is an open question how to use the extended formulation from BPS for obtaining a linear description of orbitopes. Therefore, BPS is treated in a chapter of its own. It would be a bit out of the scope of this thesis to fathom out all possibilities of this method. For more information on BPS, we refer to [63], where the material from this chapter can also be found.

1.2 Notation, Wording, Basic Definitions

The following section will fix notation for a range of objects from linear algebra, graph theory and polyhedral theory that will be useful for our following work. Note that the aim of the section is to clarify notation to prevent later ambiguities; it is by no means a basic introduction to the respective fields of mathematics. For these purposes, we refer the reader to common textbooks like [40], [75, 76] or [114, 115] for linear algebra, [21] or [13] for graph theory, [71] for combinatorial optimization, [106, 107], [94] or [113] for (combinatorial) linear programming and [120, 121, 53] for polyhedral aspects, in particular for 0/1-polytopes.

1.2.1 General Notation

Important Sets

By $\mathbb{N}$, we denote the set of natural numbers including 0. $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ denote the set of integers, rational numbers and real numbers, respectively. The set of positive integers $\mathbb{N} \setminus \{0\}$ is denoted by $\mathbb{N}_>$. For some $x \in \mathbb{R}$, we define $\lfloor x \rfloor := \max(\{j \in \mathbb{Z} \mid j \leqslant x\})$ and $\lceil x \rceil := \min(\{j \in \mathbb{Z} \mid j \geqslant x\})$. For some set $\mathcal{S} \subseteq \mathbb{R}$, we denote by $\mathcal{S}_+$ the set $\{x \in \mathcal{S} \mid \boldsymbol{x} \geqslant 0\}$ and by $\mathcal{S}_-$ the set $\{x \in \mathcal{S} \mid x \leqslant 0\}$. The empty set is denoted by $\varnothing$.

We denote by $2^{\mathcal{S}}$ the power set of set $\mathcal{S}$, that is $\{\mathcal{T} \mid \mathcal{T} \subseteq \mathcal{S}\}$. Any set $\mathcal{F} \subseteq 2^{\mathcal{S}}$ is called a *family*.

Ranges

For $n \in \mathbb{N}_>$, the *(integral) range* $[n]$ is the subset of $\mathbb{N}_>$ containing all natural numbers from 1 through n. For $n \in \mathbb{R} \setminus \mathbb{N}_>$, we define $[n] := \varnothing$. Furthermore, we write $[n]_0 := [n] \cup \{0\}$ and for $1 \leqslant i < n$, $i, n \in \mathbb{N}_>$, we write $[i..n] := [n] \setminus [i-1]$. The cardinality of a range is called its *length*.

Note that the closed interval $\{x \in \mathbb{R} \mid a \leqslant x \leqslant b\}$ bounded by some real numbers $a, b \in \mathbb{R}$ is written $[a, b]$.

object	style	examples
general sets and families (sets of sets)	calligraphic, capital letter	$\mathcal{S}$
scalars	normal, lower case	x, λ
vectors, tupels	bold, lower case	$\boldsymbol{x}$, $\boldsymbol{\ell}$
matrices	bold, capital letter	$\boldsymbol{A}$, $\boldsymbol{M}$
graphs, digraphs and hypergraphs	normal, capital letter	D, G, H
nodes, edges, arcs and hyperarcs	normal, lower case	v, a
polyhedra, polytopes, faces, facets	bold, capital letter	$\mathbf{P}$, $\mathtt{F}$

Figure 1.1: Overview over the typesetting styles for the most frequently used mathematical objects throughout the document.

Symmetric Difference

The *symmetric difference* of two sets $\mathcal{A}$ and $\mathcal{B}$ is defined as $(\mathcal{A} \setminus \mathcal{B}) \cup (\mathcal{B} \setminus \mathcal{A})$ and denoted by $\mathcal{A}\Delta\mathcal{B}$.

For two finite sets $\mathcal{A}$ and $\mathcal{B}$, the *Cartesian set product* is defined as

$$\mathcal{A} \times \mathcal{B} = \{(a, b) \mid a \in \mathcal{A} \text{ and } b \in \mathcal{B}\}.$$

Cartesian Set Product

Notation "$\times$" is also used for the direct product of groups and for the Cartesian graph product (see below). Let $\mathcal{I}$ be a nonempty index set and let for each $i \in \mathcal{I}$ set $\mathcal{S}_i := \mathcal{S}$ be a copy of some nonempty *ground set* $\mathcal{S}$. Iterating the Cartesian set product, we write then

$$\mathcal{S}^{\mathcal{I}} := \bigtimes_{i \in \mathcal{I}} \mathcal{S}_i.$$

If $\mathcal{I} = [n]$, we will also write $\mathcal{S}^n$ instead of $\mathcal{S}^{[n]}$.

Vectors

The elements of $\mathcal{S}^{\mathcal{I}}$ are called *vectors* or tupels, that is, vector $\boldsymbol{v} = (v_i)_{i \in \mathcal{I}}$ with *entries* or components v_i. If $|\mathcal{I}| = n$, then the vector (tupel) is an n-vector (n-tupel). The ground set $\mathcal{S}$ will usually be $\mathbb{R}$, $\mathbb{Q}$ or $\{0, 1\}$ for our purposes. If all components of a vector are in $\mathbb{Z}$, the vector is called *integral*. Note that for our purposes, the index set of vectors will often be two dimensional, for instance $\mathcal{I} = [m] \times [n]$.

Incidence Vector

A special vector that will frequently be used is the *incidence vector* or *characteristic vector* of a subset $\mathcal{T} \subseteq \mathcal{S}$ of a set $\mathcal{S}$; it is defined as the vector $\boldsymbol{x}[\mathcal{T}] \in \{0, 1\}^{\mathcal{S}}$ of $\mathcal{T}$ with $x_i = 1$ if and only if $i \in \mathcal{T}$.

Scalar Product

The *scalar product* of two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{\mathcal{I}}$, denoted by $\langle ., . \rangle$, is defined as

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \sum_{i \in \mathcal{I}} x_i y_i.$$

Modifying Vectors

We sometimes have to change single components of some vector $\boldsymbol{v} \in \mathbb{R}^{\mathcal{I}}$ and leave the remaining components as they are. This will be referred to as a *modification*. More precisely, we call $\widetilde{\boldsymbol{v}} \in \mathbb{R}^{\mathcal{I}}$ a modification of $\boldsymbol{v}$ in component $\widetilde{v}_\ell := s \in \mathbb{R}$, if

$$\widetilde{v}_i = \begin{cases} s, & \text{if } i = \ell \\ v_i, & \text{otherwise} \end{cases} \qquad \forall i \in \mathcal{I}.$$

Note that a modification in more than one component is also possible. Moreover, it can happen that $\boldsymbol{v} = \widetilde{\boldsymbol{v}}$.

For two vectors $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^{\mathcal{I}}$, we write $\boldsymbol{v} > \boldsymbol{w}$ if and only if $v_i > w_i$ for all $i \in \mathcal{I}$.

Comparing Vectors

We also compare vectors lexicographically. For two vectors $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^{\mathcal{I}}$ with an index set $\mathcal{I}$ ordered by "$>$", we say that $\boldsymbol{v}$ is lexicographically larger than $\boldsymbol{w}$, in short $\boldsymbol{v} \succ \boldsymbol{w}$, if and only if there is an index $k \in \mathcal{I}$ such that $v_k > w_k$ and for all $i \in \mathcal{I}$ with $i < k$, it holds that $v_i = w_i$.

Clearly, for two vectors $\boldsymbol{v}, \boldsymbol{w} \in \{0,1\}^n$, it holds that $\boldsymbol{v} \succ \boldsymbol{w}$, if and only if

$$\sum_{i=1}^{n} 2^{n-i} v_i > \sum_{i=1}^{n} 2^{n-i} w_i.$$

Matrices

While vectors are denoted by small bold letters, matrices are denoted by capitals. In a matrix $\boldsymbol{A} \in \mathbb{R}^{[m]\times[n]}$, the entry or component in row i and column j is denoted by $a_{i,j}$.

Let $\boldsymbol{A} \in \mathbb{R}^{\mathcal{I}\times\mathcal{J}}$ be a matrix, and let $\mathcal{I}' \subseteq \mathcal{I}$ and $\mathcal{J}' \subseteq \mathcal{J}$ be subsets of index sets $\mathcal{I}$ and $\mathcal{J}$, respectively. Then $\boldsymbol{A}_{\mathcal{I}'\times\mathcal{J}'}$ denotes the submatrix that arises from deleting all entries in $\boldsymbol{A}$ which do not have a row index in $\mathcal{I}'$ or a column index in $\mathcal{J}'$. Moreover, we use "wild cards to describe columns and rows of matrices. That is: the vector $\boldsymbol{a}_{i,*}$ is the ith row of matrix $\boldsymbol{A}$ and analogously, $\boldsymbol{a}_{*,j}$ denotes its jth column.

Note that we use the same notation for vectors with a two-dimensional index set. Example: Let $\boldsymbol{v} \in \mathbb{R}^{[5]\times[5]}$, then vector $\boldsymbol{v}_{*,2}$ is the second column of vector $\boldsymbol{v}$ and vector $\boldsymbol{w} := \boldsymbol{v}_{[1..2]\times[2..5]}$ is the vector in $\mathbb{R}^{[1..2]\times[2..5]}$ that arises from deleting rows 3 through 5 and the first column in $\boldsymbol{v}$. Note that $\boldsymbol{w}_{*,1}$ is not defined.

Matrix Product

The matrix product of two matrices $\boldsymbol{A} \in \mathbb{R}^{[m]\times[n]}$ and $\boldsymbol{B} \in \mathbb{R}^{[n]\times[p]}$ is matrix $\boldsymbol{AB} \in \mathbb{R}^{[m]\times[p]}$ defined by entries

$$(\boldsymbol{AB})_{i,j} := \langle \boldsymbol{a}_{i,*}, \boldsymbol{b}_{*,j} \rangle \quad \forall (i,j) \in [m]\times[p]$$

Depending on the context, a vector $\boldsymbol{v} \in \mathcal{S}^{\mathcal{I}}$ can be seen as a column vector (that is, a $\mathcal{I}\times 1$-matrix) or as a row vector (a $1\times\mathcal{I}$-matrix). If not explicitly denoted, we implicitly assume compatibility of sizes when using notation like

$$\boldsymbol{Ax} \leqslant \boldsymbol{b} \quad \text{or} \quad \boldsymbol{yA} = \boldsymbol{c}.$$

Sometimes, we explicitly denote by $\boldsymbol{M}^\intercal$ the transpose of a matrix $\boldsymbol{M} \in \mathbb{R}^{\mathcal{A}\times\mathcal{B}}$, that is:

$$\boldsymbol{A} = \boldsymbol{M}^\intercal \text{ if and only if } a_{j,i} = m_{i,j} \text{ for all } (i,j) \in \mathcal{A}\times\mathcal{B}.$$

(Similarly, for vectors.)

Special Matrices and Vectors

We denote by $\mathbb{0}_{[m]\times[n]}$ or $\mathbb{1}_{[m]\times[n]}$ a $[m]\times[n]$-matrix with only entries 0 or 1, respectively. The matrix $\boldsymbol{A} = \mathbb{I}^{[m]} \in \{0,1\}^{[m]\times[m]}$ is the *unit matrix* with m rows and m columns, i.e. the matrix with entries $a_{i,j} = 1$ if and only if $i = j$. Similarly, $\mathbb{0}_{[m]}$ and $\mathbb{1}_{[m]}$ denote m-vectors with all entries 0 and 1, respectively, and $\mathbb{e}^k_{[m]}$ is the kth *unit vector* in $\{0,1\}^{[m]}$, that is a vector with an entry 1 in the kth component and 0s, otherwise. If the context is clear, we drop the dimension in the index, writing for instance $\mathbb{0}$ instead of $\mathbb{0}_{[m]\times[n]}$.

Support

The *support* of a vector $\boldsymbol{x} \in \mathbb{R}^{\mathcal{I}}$ (and analogously for matrices) is defined as

$$\operatorname{supp}(\boldsymbol{x}) := \{i \in \mathcal{I} \mid x_i \neq 0\}.$$

Moreover, we define

$$\operatorname{supp}^+(\boldsymbol{x}) := \{i \in \mathcal{I} \mid x_i > 0\}.$$

and

$$\operatorname{supp}^-(\boldsymbol{x}) := \{i \in \mathcal{I} \mid x_i < 0\}.$$

Size of Numbers and Vectors

Mainly for estimating computational complexity, we need the *size* of numbers. The size of a rational number $\frac{p}{q} \in \mathbb{Q}$ with $p, q \in \mathbb{Z}$ co-prime is defined by

$$\operatorname{size}\left(\frac{p}{q}\right) := 1 + \lceil \log_2(|p|+1) \rceil + \lceil \log_2(|q|+1) \rceil,$$

while the size of a n-tupel of rational numbers $\boldsymbol{x} \in \mathbb{Q}^{[n]}$ is defined by

$$\mathrm{size}(\boldsymbol{x}) := n + \sum_{i=1}^{n} \mathrm{size}(x_i),$$

and, analogously, the size of a rational $m \times n$-matrix $\boldsymbol{A} \in \mathbb{Q}^{[m]\times[n]}$ is defined by

$$\mathrm{size}(\boldsymbol{A}) := mn + \sum_{(i,j)\in[m]\times[n]} \mathrm{size}(a_{i,j}).$$

Landau Notation and Polynomial Time Algorithms

We denote by $\mathcal{O}$ the Landau ("big O") notation to estimate the growth of a function. For real valued functions $f, g : \mathbb{R} \to \mathbb{R}$, it holds that $f = \mathcal{O}(g)$ if there is a scalar $\lambda \in \mathbb{R}_+ \setminus \{0\}$ and an index $x_0 \in \mathbb{R}$ such that

$$|f(x)| \leqslant \lambda |g(x)| \quad \forall x \in \mathbb{R} \text{ with } x \geqslant x_0.$$

Using this notation, we say that an algorithm for some problem $\mathcal{P}$ *runs in time of* $\mathcal{O}(f(x))$, if the worst case running time for an instance $\mathcal{P}$ of size x is of $\mathcal{O}(f(x))$. If not otherwise stated, we use the unit time model, i.e. we assume that multiplications, divisions, additions and subtractions cost a constant unit of time. If $f(x)$ is a polynomial, we say that the algorithm has polynomial running time or is *polynomial*. For more information concerning complexity theory, we refer the reader to [46].

1.2.2 Notation related to Graphs and Hypergraphs

1.2.2.1 Common Graphs and Digraphs

Undirected Graph

An *undirected graph* is a pair $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is some set and $\mathcal{E}$ is a family of unordered pairs of elements of $\mathcal{V}$. The elements of $\mathcal{V}$ are called *vertices* or *nodes* of the graph, the set $\mathcal{E}$ contains the *edges* of G. If not explicitly stated, we will assume that there are no loops, i.e. $v \neq w$ for each edge $\{v, w\} \in \mathcal{E}$, and no parallel edges, i.e. $\mathcal{E}$ does not contain two identical subsets. If we want to emphasize that $\mathcal{V}$ is the set of vertices of graph G, we write $\mathcal{V}_G$. (Similarly for edges.)

We say that if edge $e = \{v, w\} \in \mathcal{E}$, then vertices v and w are *adjacent*. If for two distinct edges e, f, it holds that $e \cap f = \{v\} \neq \varnothing$, then we say that e and f are *incident* with each other as well as with vertex v; otherwise, we say the edges are *disjoint*.

(Induced) Subgraph

A graph $G' = (\mathcal{V}', \mathcal{E}')$ is a *subgraph* of G if $\mathcal{V}' \subseteq \mathcal{V}$ and $\mathcal{E}' \subseteq \mathcal{E}$. If $\mathcal{E}'$ is the set of all edges from $\mathcal{E}$ with both end nodes in $\mathcal{V}'$, then G' is *induced* by $\mathcal{V}'$ and we write $G' = G[\mathcal{V}']$. At the same time, G' is induced by all edges in $\mathcal{E}'$ and $G' = G[\mathcal{E}']$. By $\mathcal{V}[\mathcal{E}']$, we denote the set of nodes $\mathcal{V}' \subseteq \mathcal{V}$ that are incident with the edges in $\mathcal{E}'$, and by $\mathcal{E}[\mathcal{V}']$, we denote the edges with both ends in $\mathcal{V}'$ (which is $\mathcal{E}'$).

Complete Graph

A graph is called *complete*, if $\mathcal{E}$ contains an edge $\{v, w\}$ for each $v, w \in \mathcal{V}$ with $v \neq w$. We denote by K_n the complete graph with n vertices.

Directed Graph

A *directed graph*, *digraph* or *network* is a pair $D = (\mathcal{V}, \mathcal{A})$ of a vertex set $\mathcal{V}$ and a family $\mathcal{A}$ of ordered pairs of elements of $\mathcal{V}$, called the *arcs* of the digraph. An arc $a = (v, w)$ leads from its start node or *tail* v to its end node or *head* w. The *reverse arc* to (u, v) is (v, u). In arc (v, w), node w is called the *successor* of v, and, conversely, node v is the *predecessor* of w.

Common (Di)Graph

Sometimes, we will write *common graph* or *common digraph*, respectively, to emphasize that we are not speaking of a directed hypergraph (see below).

Successors and Predecessors

Let $\mathcal{S} \subseteq \mathcal{V}$. We denote by

$$\mathrm{succ}_D(\mathcal{S}) := \{w \in \mathcal{V} \mid \exists (v, w) \in \mathcal{A} \text{ for some } v \in \mathcal{S}\}$$

the set of *successors* in D and by

$$\mathrm{pred}_D(\mathcal{S}) := \{w \in \mathcal{V} \mid \exists (w, v) \in \mathcal{A} \text{ for some } v \in \mathcal{S}\}$$

the set of *predecessors* in D. If $\mathcal{S}$ contains one node v only, we also write $\mathrm{succ}_D(v)$ or $\mathrm{pred}_D(v)$, respectively, and we drop index D if the context is clear.

Subdigraph

Subdigraphs and *induced subdigraphs* are defined analogously as for undirected graphs, by replacing edges by arcs. As for graphs, we assume that digraphs do not contain parallel arcs or loops unless explicitly stated.

Canonical Digraph

One can always obtain a digraph D from an undirected graph G by replacing every edge $\{u, v\}$ in $\mathcal{E}$ by two arcs, namely arc (u, v) and reverse arc (v, u). We will call this digraph the *canonical digraph* of G.

Graph Isomorphism

Moreover, we call two graphs $G = (\mathcal{V}, \mathcal{E})$, $G' = (\mathcal{V}', \mathcal{E}')$ *isomorphic*, denoted by $G \simeq G'$, if there is a bijection $\pi : \mathcal{V} \to \mathcal{V}'$ such that $\{\pi(v), \pi(w)\} \in \mathcal{E}'$ if and only if $\{v, w\} \in \mathcal{E}$. Similarly, digraphs $D = (\mathcal{V}, \mathcal{A})$ and $D' = (\mathcal{V}', \mathcal{A}')$ are isomorphic if there is a bijection $\pi : \mathcal{V} \to \mathcal{V}'$ such that $(\pi(v), \pi(w)) \in \mathcal{A}'$ if and only if $(v, w) \in \mathcal{A}$.

k-Clique

A (possibly empty) subset of nodes of a graph that induces a subgraph which is isomorphic to a complete graph K_k is called a *k-clique*.

Complement

The *complement* of a graph $G = (\mathcal{V}, \mathcal{E})$ is denoted by $\overline{G} = (\mathcal{V}, \overline{\mathcal{E}})$; for edge set $\overline{\mathcal{E}}$, it holds that

$$\{v, w\} \in \overline{\mathcal{E}} \text{ if and only if } \{v, w\} \notin \mathcal{E}.$$

The complement of a digraph is defined analogously.

Stable Set, k-Partite Graph

A subset $\mathcal{S}$ of pairwise not adjacent nodes of some graph or digraph is *stable* or *independent*. A graph or digraph is *k-partite*, if its vertex set can be partitioned into k stable sets.

Paths, Cycles

Let $G = (\mathcal{V}, \mathcal{E})$ be a graph. A sequence of edges $\mathcal{P} = (e_1, \ldots, e_n)$ is called a *(simple) s-t-path* if

(i) $e_i = \{v_i, v_{i+1}\}$ for all $i \in [n]$,
(ii) $v_1 = s$ and $v_{n+1} = t$, and
(iii) $v_i \neq v_j$ for $i \neq j$.

If $s = t$ (condition (iii) is not satisfied), then $\mathcal{P}$ is called a *cycle*. If additionally all nodes in $\mathcal{P}$ are distinct, then $\mathcal{P}$ is called a *simple cycle*. We say that node v is *used* by a path if $v \in \bigcup_{i=1}^{n} e_i$. Directed s-t-paths and cycles are defined analogously as for undirected graphs as a sequence of arcs $(a_1, \ldots, a_n)$ with $a_i = (v_i, v_{i+1})$ for all $i \in [n]$. n is the *length* of the path or cycle. A cycle of length n is a *n-cycle*. Note that $\mathcal{P}$ can also be the empty set, which means that for any node v, there is a v-v-path.

A digraph that has no cycles is called *acyclic*. Sometimes, we write in short *DAG* instead of "directed acyclic digraph".

Connected Graph

Two nodes u, v in an undirected graph are *connected*, if there is a u-v-path in G. Connectedness defines an equivalence relation on the vertices: two nodes in the graph are equivalent if they are connected. The equivalence classes are called the *connected components* of the graph. A graph with one single connected component is called *connected*.

Subsets of Nodes and Edges/Arcs

For any undirected graph $G = (\mathcal{V}, \mathcal{E})$ and a subset $\mathcal{S} \subseteq \mathcal{V}$, we define:

- the *neighborhood*

  $\mathrm{N}_G(\mathcal{S}) \ := \{w \in \mathcal{V} \setminus \mathcal{S} \mid \exists \{w, v\} \in \mathcal{E} \text{ for some } v \in \mathcal{S}\}$,
- the *star*

  $\delta_G(\mathcal{S}) \ := \{e \in \mathcal{E} \mid e = \{v, w\} \text{ with } v \in \mathcal{S} \text{ and } w \notin \mathcal{S}\}$, and

- the *reach*
  $r_G(\mathcal{S}) \;\; := \{w \in \mathcal{V} \mid \exists v\text{-}w\text{-path in } G \text{ for some } v \in \mathcal{S}\}$.

Analogously, we define for digraphs $D = (\mathcal{V}, \mathcal{A})$:

- the *in-neighborhood*
  $\mathrm{N}_D^{\mathrm{in}}(\mathcal{S}) := \{w \in \mathcal{V} \setminus \mathcal{S} \mid \exists (w, v) \in \mathcal{A} \text{ for some } v \in \mathcal{S}\}$,
- the *out-neighborhood*
  $\mathrm{N}_D^{\mathrm{out}}(\mathcal{S}) := \{w \in \mathcal{V} \setminus \mathcal{S} \mid \exists (v, w) \in \mathcal{A} \text{ for some } v \in \mathcal{S}\}$,
- the *in-star*
  $\delta_D^{\mathrm{in}}(\mathcal{S}) \;\; := \{a \in \mathcal{A} \mid a = (w, v) \text{ with } w \notin \mathcal{S} \text{ and } v \in \mathcal{S}\}$,
- the *out-star*
  $\delta_D^{\mathrm{out}}(\mathcal{S}) := \{a \in \mathcal{A} \mid a = (w, v) \text{ with } w \in \mathcal{S} \text{ and } v \notin \mathcal{S}\}$,
- the the *in-reach*
  $r_D^{\mathrm{in}}(\mathcal{S}) \;\; := \{w \in \mathcal{V} \mid \exists w\text{-}v\text{-path in } D \text{ for some } v \in \mathcal{S}\}$, and
- the *out-reach*
  $r_D^{\mathrm{out}}(\mathcal{S}) := \{w \in \mathcal{V} \mid \exists v\text{-}w\text{-path in } D \text{ for some } v \in \mathcal{S}\}$.

In-star, out-star and star are sometimes also referred to as *in-cuts*, *out-cuts* and *cuts*, respectively. If the set $\mathcal{S}$ contains only one single element v, we write $.(v)$ instead of $.(\{v\})$, for instance $\delta_D^{\mathrm{in}}(v)$ instead of $\delta_D^{\mathrm{in}}(\{v\})$. Note that for single nodes, in-neighborhood and the set of predecessors are identical, and similarly out-neighborhood and set of successors.

For some digraph $D = (\mathcal{V}, \mathcal{E})$ and subsets $\mathcal{S}, \mathcal{T}$ of the node set $\mathcal{V}$, we denote by

$$\mathcal{S} : \mathcal{T} := \{(v, w) \in \mathcal{A} \mid v \in \mathcal{S} \text{ and } w \in \mathcal{T}\}$$

the set of arcs with tail in $\mathcal{S}$ and head in $\mathcal{T}$. For undirected graphs, $\mathcal{S} : \mathcal{T}$ is similarly defined as the set of edges with one end in $\mathcal{S}$ and the other end in $\mathcal{T}$. Note that in directed graphs, $\mathcal{S} : (\mathcal{V} \setminus \mathcal{S}) = \delta_D^{\mathrm{out}}(\mathcal{S})$ and $(\mathcal{V} \setminus \mathcal{S}) : \mathcal{S} = \delta_D^{\mathrm{in}}(\mathcal{S})$, and analogously, in an undirected graph G, $\mathcal{S} : (\mathcal{V} \setminus \mathcal{S}) = \delta_G(\mathcal{S})$.

Degree of a Node

In an undirected graph G, the *degree* of a node v is defined as $\deg(v) := |\delta_G(v)|$. Similarly, in a digraph D, the in-degree and out-degree of v are $\deg^{\mathrm{in}}(v)_D := |\delta_D^{\mathrm{in}}(v)|$ and $\deg^{\mathrm{out}}(v)_D := |\delta_D^{\mathrm{out}}(v)|$, respectively.

For all subsets of nodes, edges and arcs defined so far, we will always drop indices D and G, if it is clear which (di)graph is underlying.

Trees, Arborescences

A *tree* $T = (\mathcal{V}, \mathcal{E})$ is an undirected connected graph containing no cycles. An *arborescence* is a directed graph $D = (\mathcal{V}, \mathcal{A})$ arising from tree T by choosing an arbitrary node $r \in \mathcal{V}$ as *root* and replacing each edge $\{v, w\} \in \mathcal{E}$ either by an arc (v, w) or by an arc (w, v) such that there is a directed path from v to r for each $v \in \mathcal{V}$ in D.

Cartesian Graph Product

Let G and H be be two graphs. Then the *Cartesian graph product* $G \times H$ is defined as the graph with the following properties:

- the vertex set of $G \times H$ is the Cartesian set product $V_G \times V_H$ and
- any two vertices (u, u') and (v, v') are adjacent in $G \times H$ if and only if either
  - $u = v$ and u' is adjacent with v' in H, or
  - $u' = v'$ and u is adjacent with v in G.

Flows and Circulations

We will also use *flows and circulations*. Let $D = (\mathcal{V}, \mathcal{A})$ be some digraph. We assign to each arc $a \in \mathcal{A}$ a lower and an upper bound $\ell_a, u_a \in \mathbb{R} \cup \{-\infty, +\infty\}$ defining the *capacity* of a. A vector $\boldsymbol{y} \in \mathbb{R}^{\mathcal{A}}$ with

$$u_a \leqslant y_a \leqslant \ell_a \qquad \forall a \in \mathcal{A}$$

is called a (feasible) *flow* on D. If it additionally holds that

$$\sum_{a \in \delta_D^{\mathrm{in}}(v)} y_a = \sum_{a \in \delta_D^{\mathrm{out}}(v)} y_a \qquad \forall v \in \mathcal{V},$$

then the flow is called a *circulation*.

1.2.2.2 Hypergraphs

Generalizing the definitions concerning graphs and digraphs, we will now proceed with definitions of hypergraphs, hyperpaths, and related objects. Here, we follow in many aspects Ausiello et al. [5], in particular in their definition of hyperpaths.

Hypergraphs

Let $\mathcal{U} \neq \varnothing$ be some finite set of vertices. A hyperarc is an ordered tupel

$$(\mathcal{S}, \mathcal{T}) \in (2^{\mathcal{U}} \setminus \{\varnothing\}) \times (2^{\mathcal{U}} \setminus \{\varnothing\})$$

with $\mathcal{S} \cap \mathcal{T} = \varnothing$. Arcs and vertices define the directed hypergraph $H = (\mathcal{U}, \mathcal{A})$. Each arc $a = (\mathcal{S}, \mathcal{T})$ in $\mathcal{A}$ has a (nonempty) set $\operatorname{tail}(a) := \mathcal{S}$ of *tail nodes* and a nonempty set $\operatorname{head}(a) := \mathcal{T}$ of *head nodes*. The arc is directed from its tail nodes to its head nodes.

F-Arcs, B-Arcs, BF-Arcs

A hyperarc $a \in \mathcal{A}$ with $|\operatorname{tail}(a)| = 1$ is called *forward arc* (F-arc), an arc with $|\operatorname{head}(a)| = 1$ is a *backward arc* (B-arc), and arcs with $|\operatorname{tail}(a)| = |\operatorname{head}(a)| = 1$ are called *backward-forward arcs* (BF-arcs). If all arcs in $\mathcal{A}$ are F-arcs (B-arcs, BF-arcs), then the graph is called F-hypergraph (B-hypergraph, BF-hypergraph), respectively. (Note that a BF-hypergraph is a common digraph.) For B-arcs, we write $(\mathcal{S}, t)$ instead of $(\mathcal{S}, \{t\})$, and for F-arcs, we write $(s, \mathcal{T})$ instead of $(\{s\}, \mathcal{T})$.

Relaxation of Hypergraph

The *relaxation* of a directed hypergraph is the digraph $D = (\mathcal{U}, \bar{\mathcal{A}})$ with vertex set $\mathcal{U}$ and arc set $\bar{\mathcal{A}}$, where $(u, v) \in \bar{\mathcal{A}}$ if and only if there is a hyperarc $(\mathcal{S}, \mathcal{T}) \in \mathcal{A}$ with $u \in \mathcal{S}$ and $v \in \mathcal{T}$.

Induced Subgraphs

For some subset $\mathcal{L} \subseteq \mathcal{A}$ of hyperarcs, we denote by $\mathcal{U}[\mathcal{L}]$ the *node set* of $\mathcal{L}$, i.e. those nodes in $\mathcal{U}$ that are incident with the hyperarcs in $\mathcal{L}$:

$$\mathcal{U}[\mathcal{L}] := \bigcup_{a \in \mathcal{L}} \big(\operatorname{tail}(a) \cup \operatorname{head}(a)\big)$$

For some subset $\mathcal{L} \subseteq \mathcal{A}$ of arcs, we call $(\mathcal{U}[\mathcal{L}], \mathcal{L})$ a *subgraph* of H.

Hyperpath

A $\mathcal{W}$-t-hyperpath in a hypergraph $H = (\mathcal{U}, \mathcal{A})$ is an ordered subset of arcs $\mathcal{L} = (a_1, \ldots, a_n) \subseteq \mathcal{A}$ such that

[p1] $t \in \operatorname{head}(a_n)$

[p2] For all $i \in [n]$, $\operatorname{tail}(a_i) \subseteq \mathcal{W} \cup \big(\bigcup_{k<i} \operatorname{head}(a_k)\big)$.

[p3] Any proper subgraph of $(\mathcal{U}[\mathcal{L}], \mathcal{L})$ violates conditions [p1] or [p2].

For each $t \in \mathcal{U}$, there is an empty hyperpath $\mathcal{L} = \varnothing$ leading from t to t. We say that hyperpath $\mathcal{L}$ *is using* its node set $\mathcal{U}[\mathcal{L}]$.

Subsets of Nodes and Hyperarcs

Similarly as for common digraphs, we define for any hypergraph $H = (\mathcal{U}, \mathcal{A})$ and any $v \in \mathcal{U}$

- the *in-neighborhood*
  $\mathrm{N}_H^{\mathrm{in}}(v) := \{u \in \mathcal{U} \mid \exists (\mathcal{S}, \mathcal{T}) \in \mathcal{A} \text{ with } u \in \mathcal{S} \text{ and } v \in \mathcal{T}\}$,
- the *out-neighborhood*
  $\mathrm{N}_H^{\mathrm{out}}(v) := \{u \in \mathcal{U} \mid \exists (\mathcal{S}, \mathcal{T}) \in \mathcal{A} \text{ with } v \in \mathcal{S} \text{ and } u \in \mathcal{T}\}$,
- the *in-star*
  $\delta_H^{\mathrm{in}}(v) := \{a \in \mathcal{A} \mid a = (\mathcal{S}, \mathcal{T}) \text{ with } v \in \mathcal{T}\}$,
- the *out-star*
  $\delta_H^{\mathrm{out}}(v) := \{a \in \mathcal{A} \mid a = (\mathcal{S}, \mathcal{T}) \text{ with } v \in \mathcal{S}\}$,
- the *in-reach*
  $r_H^{\mathrm{in}}(v) := \{u \in \mathcal{U} \mid \exists \mathcal{W}\text{-}v\text{-hyperpath with } u \in \mathcal{W}\}$, and
- the *out-reach*
  $r_H^{\mathrm{out}}(v) := \{u \in \mathcal{U} \mid \exists \mathcal{W}\text{-}u\text{-hyperpath with } v \in \mathcal{W}\}$.

If the context is clear, we will drop index H.

Last, a directed hypergraph is called *acyclic* if and only if its relaxation is an acyclic digraph. This implies the existence of a (topological) order on the nodes U.

Remark 1.1

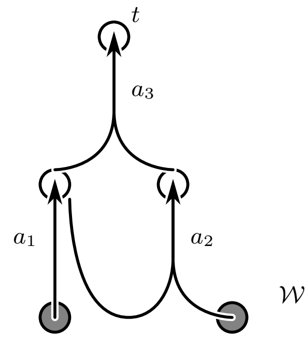

Figure 1.2: Undirected cycle in the relaxation of a $\mathcal{W}$-t-hyperpath.

- Each $\mathcal{W}$-t-hyperpath $\mathcal{L}$ starts from a nonempty set $\mathcal{U}[\mathcal{L}] \cap \mathcal{W} \subseteq \mathcal{W}$.
- Definition [p1] through [p3] generalizes the definition for paths in common digraphs.
- Even in B-hypergraphs, the relaxation of the subgraph induced by a $\mathcal{W}$-t-hyperpath can contain undirected cycles, as figure 1.2 shows.
- [p3] ensures in particular that the relaxation of a $\mathcal{W}$-t-hyperpath is a connected graph.
- In literature, a couple of alternative definitions for hyperpaths can be found. (For alternatives, see e.g. [42, 101] and for a discussion [95, 58].)

Lemma 1.2 *If H is a B-hypergraph, then the relaxation of the subgraph induced by a nonempty $\mathcal{W}$-t-hyperpath $\mathcal{L} = (a_1, \dots, a_n)$ cannot contain directed (simple) cycles.*

Proof. Let $\mathcal{L}$ be a $\mathcal{W}$-t-hyperpath. We call an arc $a_k \in \mathcal{L}$ *dispensable*,
(a) if $\mathrm{head}(a_k) \subseteq \mathcal{W}$ or
(b) if there is an arc $a_i \in \mathcal{L}$ with $\mathrm{head}(a_k) = \mathrm{head}(a_i)$ and $i < k$ or
(c) if all arcs $a_i \in \mathcal{L}$ with $\mathrm{head}(a_k) \in \mathrm{tail}(a_i)$ are dispensable.
We denote the set of dispensable arcs in $\mathcal{L}$ by $\mathcal{D}$.

First, we claim that set $\mathcal{D}$ is empty if $\mathcal{L}$ is a hyperpath. Otherwise, $\mathcal{L}' := \mathcal{L} \setminus \mathcal{D}$ will be satisfying properties [p1] and [p2], as will be shown in the following. By [p3], this is in contradiction to the assumption that $\mathcal{L}$ is a $\mathcal{W}$-t-hyperpath.

If $t \in \mathcal{W}$, then $\mathcal{D} = \mathcal{L}$; so, if $\mathcal{D} \neq \varnothing$, then $\mathcal{L}$ is indeed no hyperpath since $\mathcal{L}' = \varnothing$ is a hyperpath.

If $t \notin \mathcal{W}$, then the arc from $\{a_j \in \mathcal{L} \mid \mathrm{head}(a_j) = \{t\}\}$ with minimal index m will not be dispensable and is therefore in $\mathcal{L}'$. Because of (b), $\mathcal{L}'$ cannot contain any arcs with index larger than m. So [p1] is satisfied for $\mathcal{L}'$. Let now a_i be an arbitrary arc in $\mathcal{L}'$. We consider the tail nodes of a_i that are not in $\mathcal{W}$. Assume there is a $v \in \mathrm{tail}(a_i) \setminus \mathcal{W}$ with

$$v \notin \bigcup_{\substack{a_k \in \mathcal{L}' \\ k < i}} \mathrm{head}(a_k).$$

In this case, all arcs in $\mathcal{L}$ with head v must have been regarded as dispensable. However, this could not happen because of (a), because $v \notin \mathcal{W}$. Nor could it happen because of (c), as a_i is not dispensable. The application of (b) always leaves at least one arc not dispensable. So any node in the tail of a_i must either be in $\mathcal{W}$ or it has in-neighbors with lower index in $\mathcal{L}'$. Hence, [p2] is satisfied for $\mathcal{L}'$.

Last, any proper subgraph of $(\mathcal{U}[\mathcal{L}'], \mathcal{L}')$ is a proper subgraph of $(\mathcal{U}[\mathcal{L}], \mathcal{L})$. Hence, no proper subset of $\mathcal{L}'$ can be a hyperpath by assumption that $\mathcal{L}$ was a hyperpath and [p3] is satisfied. Therefore $\mathcal{L}'$ is a hyperpath which is different from $\mathcal{L}$ if and only if $\mathcal{D} \neq \varnothing$.

For the second part of the proof, we assume that the relaxation of $\mathcal{L}$ contains a simple directed k-cycle $(\tilde{a}_1, \ldots, \tilde{a}_k)$. We select an ordered set of arcs $\mathcal{A} := (a(1), \ldots, a(k))$ from $\mathcal{L}$ such that the relaxation of each $a(i)$ contains arc $\tilde{a}_i$. So, for any $j \in [k]$, it holds for $a(j)$ and $a((j+1) \bmod k)$ that

$$\text{head}\left(a(j)\right) \cap \text{tail}\left(a((j+1) \bmod k)\right) \neq \varnothing.$$

There must be at least one pair of cyclically subsequent arcs a_k, a_i in $\mathcal{A}$ with $k > i$. As we are in a B-hypergraph, head(a_k) and tail(a_i) intersect in one single node v. As $\mathcal{L}$ is by assumption a hyperpath, we get from [p2] that either $v \in \mathcal{W}$ or $\{v\} = \text{head}(a_j)$ for some arc $a_j \in \mathcal{L} \setminus \mathcal{C}$ with $j < i$ or both. In any case, $a_k \in \mathcal{A}$ is an dispensable arc. Therefore, $\mathcal{L}$ contains a dispensable arc contradicting the fact that $\mathcal{L}$ is hyperpath.

□

1.2.3 Notation related to Polyhedral Theory

Hulls

Let $\mathcal{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ be a nonempty set of k points in $\mathbb{R}^n$, and let $\boldsymbol{y}$ be some point in $\mathbb{R}^n$ that can be expressed in the following way:

$$\boldsymbol{y} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i.$$

$$\text{If} \begin{cases} \lambda_i \in \mathbb{R} \ \forall\, i \in [k], \\ \lambda_i \in \mathbb{R} \ \forall\, i \in [k] \text{ and } \sum_{i=1}^{k} \lambda_i = 1, \\ \lambda_i \in \mathbb{R}_+ \ \forall\, i \in [k], \\ \lambda_i \in \mathbb{R}_+ \ \forall\, i \in [k] \text{ and } \sum_{i=1}^{k} \lambda_i = 1, \end{cases} \quad \text{then } \boldsymbol{y} \text{ is a} \begin{cases} \text{linear} \\ \text{affine} \\ \text{conic} \\ \text{convex} \end{cases} \quad \text{combination of } \mathcal{X}.$$

The set of all linear, affine, conic or convex combinations of finite subsets of a nonempty (not necessarily finite) set $\mathcal{X} \subset \mathbb{R}^n$ is called the *linear*, *affine*, *conic* or *convex hull* of $\mathcal{X}$ and is denoted by $\text{lin}(\mathcal{X})$, $\text{aff}(\mathcal{X})$, $\text{cone}(\mathcal{X})$ or $\text{conv}(\mathcal{X})$, respectively. In each case, the set $\mathcal{X}$ is said to *generate* the hull.

Let $\boldsymbol{a} \in \mathbb{R}^n$ and $b \in \mathbb{R}$. The set $\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{a}\boldsymbol{x} = b\}$ is called a *hyperplane* in $\mathbb{R}^n$; the set $\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{a}\boldsymbol{x} \leqslant b\}$ is a *halfspace* in $\mathbb{R}^n$. For arbitrary b, both are called *affine*; if $b = 0$, they are *linear*.

Polyhedra

A set $\mathbf{P}$ of points in $\mathbb{R}^n$ is called a *(convex) polyhedron* if there is some matrix $\boldsymbol{A} \in \mathbb{Q}^{m \times n}$ and some vector $\boldsymbol{b} \in \mathbb{Q}^m$ such that

$$\mathbf{P} = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\}.$$

In this case, $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ is called a (complete) linear description for $\mathbf{P}$. The *size of the description* is the encoding length of $(\boldsymbol{A}, \boldsymbol{b})$, that is

$$\text{size}(\boldsymbol{A}) + \text{size}(\boldsymbol{b}).$$

Obviously, a polyhedron $\mathbf{P}$ is the intersection of finitely many affine halfspaces in $\mathbb{R}^n$. It follows immediately that the intersection of two polyhedra is again a polyhedron.

A central fact of polyhedral optimization is that each nonempty polyhedron can be decomposed into the Minkowski sum of the conic hull of a finite set of points

$\mathcal{C} \subset \mathbb{R}^n$ (the so-called recession cone) and the convex hull of a finite set of points $\mathcal{V} \subset \mathbb{R}^n$:

$$\mathbf{P} = \operatorname{conv}(\mathcal{V}) + \operatorname{cone}(\mathcal{C}),$$

and conversely, any Minkowski sum of a finitely generated cone and the convex hull of some finite set of points is a polyhedron. Let $v \in \mathbf{P}$ be a point in the polyhedron. If for all $x, y \in \mathbf{P} \setminus \{v\}$, it holds that $v \notin \operatorname{conv}\{x, y\}$, then v is an *extreme point* or *vertex* of $\mathbf{P}$. The *lineality space* of polyhedron $\mathbf{P}$ is the largest linear subspace

$$\operatorname{lineal}(\mathbf{P}) = \operatorname{cone}(\mathcal{C}) \cap - \operatorname{cone}(\mathcal{C})$$

contained in $\operatorname{cone}(\mathcal{C})$. If $\operatorname{lineal}(\mathbf{P}) = \{\mathbb{0}\}$, then polyhedron $\mathbf{P}$ is called *pointed*.

For pointed polyhedra, it holds that if $\mathcal{V}$ is minimal (i.e. removing any point from $\mathcal{V}$ leads to a polyhedron different from $\mathbf{P}$), then it is the set of *vertices* of the polyhedron, denoted by $\mathcal{V}_{\mathbf{P}}$, and, similarly, if set $\mathcal{C}$ is minimal, then each $\boldsymbol{c} \in \mathcal{C}$ *generates an extreme ray* of $\mathbf{P}$.

Polytopes

If $\mathcal{C} = \{\varnothing\}$, then the polyhedron is bounded and it is called a *polytope*.

The above said implies that there are two equivalent ways to describe any pointed polyhedron: one can either specify its vertices and the generators of extreme rays of its recession cone or one can list a set of affine halfspaces that determines the polyhedron. In the first case, we speak of a $\mathcal{V}$-representation and in the second case of an $\mathcal{H}$-representation of the polyhedron.

Dimension of Polytope

The *dimension* of a polytope, in short $\dim(\mathbf{P})$, is the dimension of its affine hull. If $\mathbf{P} \subset \mathbb{R}^n$ and $\dim(\mathbf{P}) = n$, then we call polytope $\mathbf{P}$ *full dimensional*.

(Standard) Simplex

A prominent example for equivalent $\mathcal{V}$- and $\mathcal{H}$-descriptions is the $d-1$-dimensional *standard simplex* $\triangle_{d-1}$ in $\mathbb{R}^d$. This polytope can either be described as the convex hull of the d unit vectors $\mathbb{e}^i$ in $\mathbb{R}^d$, that is

$$\triangle_{d-1} = \operatorname{conv}\left(\left\{\mathbb{e}^i \in \{0,1\}^d \,\middle|\, i \in [d]\right\}\right)$$

or, equivalently, as

$$\triangle_{d-1} = \left\{\boldsymbol{x} \in \mathbb{R}^d \,\middle|\, \sum_{i=1}^{d} x_i = 1 \text{ and } x_i \geqslant 0 \ \forall i \in [d]\right\}.$$

(Note that the non-standard d-dimensional *simplex* $\widetilde{\triangle}_d$ in $\mathbb{R}^d$ is described by

$$\begin{aligned}\widetilde{\triangle}_d = \operatorname{conv}\left(\left\{\mathbb{e}^i \in \{0,1\}^d \,\middle|\, i \in [d]\right\} \cup \left\{\mathbf{0}\right\}\right) = \\ = \left\{\boldsymbol{x} \in \mathbb{R}^d \,\middle|\, \sum_{i=1}^{d} x_i \leqslant 1 \text{ and } x_i \geqslant 0 \ \forall i \in [d]\right\}.\end{aligned}$$

This is — in contrast to the standard simplex — a full dimensional but non-regular polytope, since it has edges of different lengths.)

Hypercube

Another example for the equivalence of $\mathcal{H}$- and $\mathcal{V}$-description is the *d-hypercube* which can either be described as

$$\operatorname{conv}\left(\{0,1\}^d\right)$$

or as

$$\left\{\boldsymbol{x} \in \mathbb{R}^d \,\middle|\, -x_i \leqslant 0 \text{ and } x_i \leqslant 1 \ \forall i \in [d]\right\}.$$

Knapsack Polyhedra

A third class of polyhedra we will encounter in this work is the class of *knapsack polyhedra*. Let vector $\boldsymbol{a} \in \mathbb{R}^m$ and scalar $b \in \mathbb{R}$ be given. Moreover, let for each $i \in [m]$ a value $s_i \in \mathbb{N}_{>} \cup \{+\infty\}$ be given, which enables us to define a set

$$\mathcal{S} := \{\boldsymbol{x} \in \mathbb{N}^m \mid 0 \leqslant x_i \leqslant s_i \quad \forall i \in [m]\}.$$

The associated knapsack polyhedron is then defined as

$$\operatorname{conv}(\{\boldsymbol{x} \in \mathcal{S} \mid \boldsymbol{a}\boldsymbol{x} \leqslant b\}).$$

It is therefore the convex hull of those points in $\mathcal{S}$ that lie in the halfspace specified by $\boldsymbol{a}$ and b.
The idea is to fill a knapsack of maximal capacity b (which may not exceeded) with items that have each a certain weight a_1 through a_m; one can choose up to s_i copies of item i.

If $s_i = 1$ for all $i \in [m]$, then the associated polytope is called a 0/1-knapsack polytope. Note that the d-dimensional simplex $\widetilde{\Delta}_d$ from above is a very basic 0/1-knapsack polytope.

Valid Inequalities, Faces

We say that an inequality $\boldsymbol{a}\boldsymbol{x} \leqslant b$ with $\boldsymbol{a} \in \mathbb{Q}^m$ and $b \in \mathbb{Q}$ is *valid* for a set of points $\mathcal{V} \in \mathbb{R}^m$, if $\boldsymbol{a}\boldsymbol{v} \leqslant b$ for all $\boldsymbol{v} \in \mathcal{V}$. In other words, if $\boldsymbol{a}\boldsymbol{x} \leqslant b$ is valid for $\mathcal{V}$, then the halfspace defined by the inequality completely contains set $\mathcal{V}$. A vector $\boldsymbol{v} \in \mathbb{R}^m$ is called *active* for inequality $\boldsymbol{a}\boldsymbol{x} \leqslant b$, if

$$\boldsymbol{a}\boldsymbol{v} = b,$$

that is, $\boldsymbol{v}$ lies in the hyperplane defined by $\boldsymbol{a}$ and b.

Let now $\mathbf{P}$ be a d-dimensional polytope in $\mathbb{R}^n$ and $\boldsymbol{a}\boldsymbol{x} \leqslant b$ some inequality valid for $\mathbf{P}$. We call the set

$$\mathtt{F}_{\mathbf{P}}(\boldsymbol{a}\boldsymbol{x} \leqslant b) := \mathbf{P} \cap \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{a}\boldsymbol{x} = b\}$$

a *face* of polytope $\mathbf{P}$ defined by inequality $\boldsymbol{a}\boldsymbol{x} \leqslant b$. (If the context is clear, we drop index $\mathbf{P}$.) By definition, the polytope $\mathbf{P}$ itself is also added to the set of faces. Clearly, any face is again a polytope. All faces apart from the empty set and $\mathbf{P}$ are called *proper faces*. Proper faces of dimension 0, 1, and $d-1$ are called vertices, edges, and facets, respectively. If $\dim(\mathtt{F}_{\mathbf{P}}(\boldsymbol{a}\boldsymbol{x} \leqslant b)) = d-1$, then $\boldsymbol{a}\boldsymbol{x} \leqslant b$ is called a *facet defining inequality*. It is a well-known fact that any proper face of $\mathbf{P}$ and $\mathbf{P}$ itself can be described as the intersection of a nonempty set of facets of $\mathbf{P}$. Among the proper faces, facets are inclusion maximal. For full dimensional polytopes, a *complete linear description* $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ with $\boldsymbol{A} \in \mathbb{Q}^{[m]\times[n]}$ and $b \in \mathbb{Q}$ is *irredundant* if every inequality in the description is facet inducing. This implies that no inequality in the description can be obtained from a conic combination of other inequalities in the description, that is: for all $i \in [m]$, there exists no vector $\boldsymbol{y} \in \mathbb{R}_+^{[m]}$, $\boldsymbol{y} \neq \mathbb{e}^i$ such that

Vertices, Edges, Facets

$$\boldsymbol{y}(\boldsymbol{A}, \boldsymbol{b}) = (\boldsymbol{a}_{i,*}, b_i).$$

When describing a polyhedron $\mathbf{P} \subseteq \mathbb{R}^{[m]}$ linearly, we distinguish two kinds of inequalities, *trivial inequalities* and *non-trivial inequalities*. An inequality $\boldsymbol{a}\boldsymbol{v} \leqslant b$ is called trivial if $|\operatorname{supp}(\boldsymbol{a})| = 1$; otherwise, the inequality is called *non-trivial*. If $\mathbf{P}$ is a 0/1-polytope, then trivial inequalities ($x_i \leqslant 1$, $x_i \geqslant 0$) are also referred to as *cube inequalities*. Facets defined by (non-)trivial inequalities are called (non-)trivial facets, respectively.

Let $\boldsymbol{a}\boldsymbol{x} \leqslant b$ define a face $\mathtt{F}$ of polytope P. Then we denote by $\mathcal{V}_{\mathbf{P}}[\boldsymbol{a}\boldsymbol{x} \leqslant b]$ the set of those vertices in $\mathcal{V}_{\mathbf{P}}$ that are active for $\boldsymbol{a}\boldsymbol{x} \leqslant b$. Note that $\mathcal{V}_{\mathbf{P}}[\boldsymbol{a}\boldsymbol{x} \leqslant b]$ is exactly the set of vertices $\mathcal{V}_{\mathtt{F}}$ of the face (and polytope) $\mathtt{F}$. In particular, if $\mathtt{F}$ is an edge of $\mathbf{P}$, then $\mathtt{F}$ is the convex hull of exactly two vertices of $\mathbf{P}$.

For inequalities $\boldsymbol{a}\boldsymbol{x} \leqslant b$ with $\boldsymbol{a} \in \mathbb{Q}^{[p]\times[2]}$ and $b \in \mathbb{Q}$, we will from time to time use the following representation in graphic form:

$$\begin{array}{|c|c||c|c|l}
\hline
a_{1,1} & a_{1,2} & x_{1,1} & x_{1,2} & (\leqslant)b \\
\hline
a_{2,1} & a_{2,2} & x_{2,1} & x_{2,2} & \\
\hline
a_{3,1} & a_{3,2} & x_{3,1} & x_{3,2} & \\
\hline
\vdots & \vdots & \vdots & \vdots & \\
 & & & & \\
\hline
\end{array}$$

Consequently, we say for general vectors $\boldsymbol{v} \in \mathbb{R}^{[p]\times[q]}$ that a row $\boldsymbol{v}_{i,*}$ is *lower* than row $\boldsymbol{v}_{j,*}$ if and only if $i > j$.

Graph of a Polytope

Let $\mathbf{P}$ be some polytope with vertex set $\mathcal{V}_{\mathbf{P}}$, and let $G_{\mathbf{P}} = (\mathcal{V}_{\mathbf{P}}, \mathcal{E})$ be some undirected graph with the property that there is an edge $\{u, v\} \in \mathcal{E}$ if and only if there is an edge in polytope $\mathbf{P}$ containing $\boldsymbol{u}$ and $\boldsymbol{v}$. Then $G_{\mathbf{P}}$ is called the *graph of the polytope*. As the nodes of the graph of the polytope and the vertices of the polytope are identified with each other, we will write both $v \in \mathcal{E}$ or $\boldsymbol{v} \in \mathcal{E}$, depending on whether we are stressing that v is node of $G_{\mathbf{P}}$ or that the respective vertex of the polytope has certain properties.

Simple Polytopes

A polytope $\mathbf{P}$ with dimension d is called *simple*, if each vertex is adjacent to exactly d edges, i.e. $\deg(v) = d$ of each node v in the graph of the polytope.

Chapter 2

The Setting, our Toolbox and its Origins

2.1 Symmetry Breaking and Orbitopes

Many linear programs exhibit symmetry under certain permutations of variables. What is meant by that? Let

$$\mathbf{P} = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\}$$

be some polyhedron in $\mathbb{R}^n$, where $\boldsymbol{A} \in \mathbb{Q}^{m\times n}$, $\boldsymbol{b} \in \mathbb{Q}^m$ and let

$$\min(\{\langle \boldsymbol{c}, \boldsymbol{x}\rangle \mid \boldsymbol{x} \in \mathbf{P}\}) \tag{*}$$

be a linear problem over $\mathbf{P}$ with cost vector $\boldsymbol{c} \in \mathbb{Q}^n$. We denote by Π_n the group of permutations of n elements and let Π_n operate on the variables of (*):

$$\Pi_n \times \mathbb{R}^n \to \mathbb{R}^n, \quad (\sigma, \boldsymbol{x}) \mapsto \sigma(\boldsymbol{x}),$$

such that $\sigma \in \Pi_n$ permutes the entries in $\boldsymbol{x}$.

Definition 2.1 The *symmetry group* G of the linear problem (*) is defined as

$$G := \{\sigma \in \Pi_n \mid \sigma(\boldsymbol{x}) \in \mathbf{P} \text{ and } \langle \boldsymbol{c}, \sigma(\boldsymbol{x})\rangle = \langle \boldsymbol{c}, \boldsymbol{x}\rangle \text{ for all } \boldsymbol{x} \in \mathbf{P}\},$$

G is obviously a subgroup of Π_n.

If $\boldsymbol{x}^\star$ is a feasible solution to (*), then we can pick an arbitrary permutation $\sigma \in G$ and obtain with $\sigma(\boldsymbol{x}^\star)$ another feasible solution to (*) with the same objective value. In other words: each element in the orbit $G(\boldsymbol{x}^\star)$ does the same job on the cost functional as $\boldsymbol{x}^\star$.

Example 2.2 A well-known example of an IP with symmetries is the Bin Packing Problem. Let $\mathcal{S} \subset \mathbb{Q}$ be a set of rational numbers ("items of size") s_1 through s_m with $0 < s_i \leqslant 1$ for all $1 \leqslant i \leqslant m$. The problem consists in partitioning $[m]$ into n subsets ("bins") $\mathcal{I}_1$ through $\mathcal{I}_n$ such that $\sum_{i\in\mathcal{I}_j} s_i \leqslant 1$ for all $1 \leqslant j \leqslant n$. A common IP model for this problem uses binary variables $x_{i,j}$ that are 1 if and only if item i is put into bin $\mathcal{I}_j$:

$$\begin{aligned}
\max \sum_{i\in[m]}\sum_{j\in[n]} x_{i,j} \qquad & \text{s.t.} & \\
\sum_{j\in[n]} x_{i,j} \leqslant 1 \qquad & & \forall i \in [m]\\
\sum_{i\in[m]} s_i x_{i,j} \leqslant 1 \qquad & & \forall j \in [n]\\
x_{i,j} \in \{0,1\} \qquad & & \forall (i,j) \in [m]\times[n]
\end{aligned}$$

It is possible to pack all items into the bins, if and only if there exists a solution $\boldsymbol{x}^\star$ with objective value m. However, the formulation of the bin-packing problem from above bears a lot of intrinsic symmetry, since all bins have the same size, and all $x_{i,j}$ are weighted the same in the objective functional. So, the symmetry group of the IP contains any permutation of bins, as any permutation of columns in $\boldsymbol{x}^\star$ leads to another solution with same objective value.
Note that a similar IP can be used to model the problem of Graph Partitioning, see for instance [64] or [32].

Symmetry becomes in particular a problem when solving IPs with branch-and-bound methods, because it makes the branching tree unnecessarily large. Therefore, several techniques have been developed to reduce or even eliminate symmetry in integer and linear programs. (For an overview, see for instance [85].)

(1) In some cases, it is possible to find a problem *reformulation* with reduced symmetry, for example by using a different set of variables. Examples are bin packing, see [39], or the social golfers problem, see [110, 54]. However, breaking the symmetry often leads to larger models (more variables, for instance).
(2) Another approach to break symmetry is *perturbation*. One can for instance add a small random vector to the objective vector. Before, all elements in an orbit $G(\boldsymbol{x})$ are equivalent as they yielded the same objective value. After, they become distinguishable in this sense. However, the solution space is in this case the same and the computational effort to answer the question whether a solution is feasible or not is not reduced.
(3) The third method consists in narrowing the set of feasible solutions by *disregarding symmetric solutions*, either in advance or during computation.

There are several possibilities to go for option (3). Prerequisite is the choice of a *representative* (or a set of specific points) for each orbit. To take the (unique) lexicographic maximal or minimal solution over other solutions is an idea which has been popular for quite a long time in combinatorics (*isomorphism-free backtracking*, see [103, 87]) or in constraint programming (see e.g. [36, 48]).

When solving IPs, there are two major ways to deal with unwanted solutions: one can cut them off by adding inequalities; or one can use knowledge about them such that respective subtrees in the branching tree will not be entered. In practice, a mix of both methods can be used. Note that simply adding inequalities may impair the model, since the number of inequalities grows.

In 2000, Rothberg ([104]) presented ideas for a systematic way to generate cuts a priori by identifying dominated subsets of variables. (His method even applied to mixed integer programs.) However, Rothbergs "domination cuts" were a rather ad hoc approach coming from practice.

Friedman ([41]) proposed a different approach to generate cuts dynamically. Here, H is given as a finite group of affine transformations of some 0/1-polytope $\mathbf{P} \subseteq \mathbb{R}^n$. If for some set $\mathcal{F} \subset \mathbf{P}$, the set

$$\{\boldsymbol{x} \in \mathbf{P} \mid \exists\, \boldsymbol{y} \in \mathcal{F},\ \exists\, \delta \in H \text{ such that } \boldsymbol{x} = \delta(\boldsymbol{y})\}.$$

equals $\mathbf{P}$, then $\mathcal{F}$ is called a *fundamental domain*. More concrete, fundamental domains are constructed on the basis of an ordering vector $\boldsymbol{c} \in \mathbb{R}^n$

$$\mathcal{F}_{\boldsymbol{c}} := \{\boldsymbol{x} \in \mathbf{P} \mid \langle \boldsymbol{c}, \boldsymbol{x} \rangle \geqslant \langle \boldsymbol{c}, \delta(\boldsymbol{x}) \rangle\ \forall \delta \in H\}.$$

Friedman's idea is to find separating hyperplanes between a fundamental domain and the remaining points in $\mathbf{P}$. However, since the choice of cost vector $\boldsymbol{c}$ is delicate and the fundamental domains are usually not as refined as orbitopes (see below), his framework is so far only of theoretical interest.

Another approach was undertaken by Margot in 2002. He transferred isomorphism-free backtracking to branch-and-cut by pruning subtrees containing symmetric solutions during computation (*isomorphism pruning*, see [82, 83]) and applied this method later to coloring problems ([84]). Margot generated cutting inequalities to fix subsets of variables to certain values while running the branching algorithm. A similar approach, called *orbital branching*, has been formulated by Linderoth et al. ([97], further developed in [98]).

Herr and Bödi ([56, 55]) create projections of the constraints matrix and objective vectors depending on the symmetry to obtain smaller IPs or LPs that with optimal solutions of the same value as for the original problems. They apply their method to IPs with full symmetric or alternating groups.

A somewhat different perspective has been taken by Kaibel and Pfetsch. Coming from packing and partitioning problems, they were interested in the structure of the convex hulls over certain two-dimensional 0/1-vectors that are lexicographic maximal subject to some group operating on the columns of the vectors. Kaibel and Pfetsch called these objects *orbitopes* ([65]. Of course, adding the inequalities describing an orbitope to the IP formulation would break symmetries in the IP formulation and therefore narrow the solution space. But the idea was that orbitopes could do more: The linear description of orbitopes is not only separating the lexicographic maximal representative from the remaining points in each orbit but also ensuring integrality of the extremal points. Since Kaibel and Pfetsch had observed that symmetry also seems to weaken the bounds from the LP-bounds relaxation of the IP, their idea was that linearly describing orbitopes could lead to a deeper structural understanding of how to strengthen these bounds by an appropriate subset of the inequalities describing the orbitope.

We will give a formal description of orbitopes in Definition 3.3.

Kaibel, Pfetsch and Peinhardt drew also combinatorial information from their separation algorithm over packing and partitioning orbitopes to fix variables when optimizing over subsets of the vertices of these orbitopes, for instance in graph partitioning problems ([64]). Another approach was to add some of the facet defining inequalities to the graph coloring polytope and to compare the resulting polytope with a similar polytope defined by Méndéz-Díaz and Zabala ([88, 90]. The latter had enhanced the classic IP-model for graph coloring by inequalities inducing an order on the color labels: before choosing color j, all colors 1 through $j-1$ have to be in use. (Recently, they developed a branch-and-cut-algorithm using their polyhedral results, see [89].)

In our work, we will exclusively focus on the geometric aspects and polyhedral properties of orbitopes. It is a different question (and would go too far) to investigate if and how one can use those results to actually improve computations in practice.

2.2 Extended Formulations

Often, the facial structure of a polyhedron $\mathbf{P} \subset \mathbb{R}^n$ associated with a linear problem is too complex to be handled directly. This may become a problem when trying to study the geometry of the polyhedron or when optimizing over it. If the linear problem describes a combinatorial problem, it can also happen that the linear description of the feasible domain is growing exponentially in the size of the combinatorial problem. One way to overcome the problem is to search for a new set of variables polynomial in size of the combinatorial problem, to formulate a new polyhedron $\mathbf{Q}$ with easier facetial structure and/or smaller size of linear description that can be linearly projected onto $\mathbf{P}$. This is the general idea behind *extended formulations*. (Note however that the face lattice of polyhedron $\mathbf{P}$ is isomorphic to a sublattice of polyhedron $\mathbf{Q}$. Therefore, $\mathbf{Q}$ will have at least as many faces as $\mathbf{P}$; see [62].)

Definition 2.3 (Projection) Let $\mathbf{Q}$ be a polyhedron in $\mathbb{R}^n \times \mathbb{R}^d$ defined by

$$\mathbf{Q} := \{(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^n \times \mathbb{R}^d \mid \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \leqslant \boldsymbol{b}\},$$

with matrices $\boldsymbol{A} \in \mathbb{Q}^{m\times n}$ and $\boldsymbol{B} \in \mathbb{Q}^{m\times d}$, and vector $\boldsymbol{b} \in \mathbb{Q}^m$. The polyhedron

$$\mathrm{proj}_x(\mathbf{Q}) := \{\boldsymbol{x} \in \mathbb{R}^n \mid \exists \boldsymbol{y} \in \mathbb{R}^d \text{ such that } (\boldsymbol{x}, \boldsymbol{y}) \in \mathbf{Q}\}$$

is called the *projection of* $\mathbf{Q}$ onto the x-space or in $\mathbb{R}^n$.

Let now $\widetilde{\mathbf{Q}} = \{\boldsymbol{y} \in \mathbb{R}^d \mid \widetilde{\boldsymbol{B}}\boldsymbol{y} \leqslant \widetilde{\boldsymbol{b}}\}$ be some polyhedron in $\mathbb{R}^d$ described by $\widetilde{\boldsymbol{B}} \in \mathbb{Q}^{m\times d}$ and $\widetilde{b} \in \mathbb{Q}^m$, and let matrix $\boldsymbol{S} \in \mathbb{Q}^{n\times d}$ induce a linear map $\sigma : \mathbb{R}^d \to \mathbb{R}^n$ projecting $\widetilde{\mathbf{Q}}$ to $\sigma(\widetilde{\mathbf{Q}}) = \mathbf{P}$ in $\mathbb{R}^n$. Then for the polyhedron

$$\mathbf{Q} := \{(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^n \times \mathbb{R}^d \mid \widetilde{\boldsymbol{B}}\boldsymbol{y} \leqslant \widetilde{\boldsymbol{b}},\ -\mathbb{1}\boldsymbol{x} + \boldsymbol{S}\boldsymbol{y} = \mathbb{0}\}$$

it holds that $\mathrm{proj}_x(\mathbf{Q}) = \mathbf{P}$.

Definition 2.4 (Extended formulation) Let polyhedron $\mathbf{Q}$ be as in Definition 2.3. If polyhedron $\mathbf{P} = \mathrm{proj}_x(\mathbf{P})$, then we call a linear description of $\mathbf{Q}$ an *extended formulation* for $\mathbf{P}$. The extended formulation is called *compact* if the size of the matrix $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{b})$ is polynomial in n.

Sometimes, one can use an extended formulation for a polyhedron $\mathbf{P}$ to derive a linear description for $\mathbf{P}$ by characterizing the extreme rays of the associated projection cone.

Definition 2.5 The *projection cone* of $\mathbf{Q}$ is the polyhedral cone

$$\mathbf{C} := \{\boldsymbol{v} \in \mathbb{R}^m \mid \boldsymbol{v}\boldsymbol{B} = \mathbb{0},\ \boldsymbol{v} \geqslant \mathbb{0}\},$$

with matrix $\boldsymbol{B}$ as defined in Definition 2.3.

Suppose, the set $\mathrm{extr}(\mathbf{C})$ of extreme rays of the projection cone is known, then a linear description of polyhedron $\mathbf{P}$ can be easily obtained by the following classical theorem:

Theorem 2.6 ([10]) *Let polyhedron* $\mathbf{Q} = \{(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^n \times \mathbb{R}^d \mid \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \leqslant \boldsymbol{b}\}$ *and projection cone* $\mathbf{C} = \{\boldsymbol{v} \in \mathbb{R}^m \mid \boldsymbol{v}\boldsymbol{B} = \mathbb{0},\ \boldsymbol{v} \geqslant \mathbb{0}\}$ *be given as above. Then*

$$\mathrm{proj}_x(\mathbf{Q}) = \{\boldsymbol{x} \in \mathbb{R}^n \mid (\boldsymbol{v}\boldsymbol{A})\boldsymbol{x} \leqslant \langle \boldsymbol{v}, \boldsymbol{b}\rangle, \boldsymbol{v} \in \mathrm{extr}(\mathbf{C})\}$$

Proof. The projection cone $\mathbf{C}$ is pointed, because $\mathbf{C} \subseteq \mathbb{R}^m_+$. Therefore, $\mathbf{C}$ is generated by its extreme rays.

$\subseteq$ Let $\boldsymbol{x} \in \mathrm{proj}_x(\mathbf{Q})$. Then by definition of projection, there is a $\boldsymbol{y}$ such that $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbf{Q}$. Hence, $\boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \leqslant \boldsymbol{b}$. Now let $\boldsymbol{v} \in \mathrm{extr}(\mathbf{C})$; therefore, in particular $\boldsymbol{v} \geqslant \mathbb{0}$ and $\boldsymbol{v}\boldsymbol{B} = \mathbb{0}$ hold. Hence, multiplication from left gives

$$\boldsymbol{v}(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y}) = (\boldsymbol{v}\boldsymbol{A})\boldsymbol{x} \leqslant \boldsymbol{v}\boldsymbol{b}.$$

$\supseteq$ Let $\boldsymbol{x}$ satisfy $(\boldsymbol{v}\boldsymbol{A})\boldsymbol{x} \leqslant \langle \boldsymbol{v}, \boldsymbol{b}\rangle$ for all $\boldsymbol{v} \in \mathrm{extr}(\mathbf{C})$. Then there is no conic combination of extreme rays $\boldsymbol{v}' \geqslant 0$ such that $\boldsymbol{v}'\boldsymbol{B} = \mathbb{0}$ and $\boldsymbol{v}'(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) > 0$, because otherwise $(\boldsymbol{v}'\boldsymbol{A})\boldsymbol{x} > \langle \boldsymbol{v}', \boldsymbol{b}\rangle$. Using Farkas' Lemma, there must therefore exist a $\boldsymbol{y}$ such that $\boldsymbol{B}\boldsymbol{y} \leqslant \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$, which means that $\boldsymbol{x} \in \mathrm{proj}_x(\mathbf{Q})$. $\square$

One reason why extended formulations are interesting is that optimization over $\mathbf{P}$ can be replaced by optimization over $\widetilde{\mathbf{Q}}$: as soon as the linear map $\sigma : \boldsymbol{y} \mapsto \boldsymbol{x} = \boldsymbol{S}\boldsymbol{y}$ is known, we can use it to define the map

$$\sigma^\star : \mathbb{R}^n \to \mathbb{R}^d, \boldsymbol{x} \mapsto \boldsymbol{y} = \boldsymbol{S}^\intercal \boldsymbol{x}$$

adjoined to σ. Let now $\boldsymbol{c} \in \mathbb{R}^n$ be some cost vector. Then

$$\max(\{\langle \boldsymbol{c}, \boldsymbol{x}\rangle \mid \boldsymbol{x} \in \mathbf{P}\}) = \max\{\langle \sigma^\star(\boldsymbol{c}), \boldsymbol{y}\rangle \mid \boldsymbol{y} \in \widetilde{\mathbf{Q}}\},$$

using the identity $\langle \boldsymbol{S}^\intercal \boldsymbol{c}, \boldsymbol{y}\rangle = \langle \boldsymbol{c}, \boldsymbol{S}\boldsymbol{y}\rangle = \langle \boldsymbol{c}, \boldsymbol{x}\rangle$.

Extended formulations can be seen in a context with the classic method of lift-and-project (for an overview, see for example [9, 10]). In literature, one can find a large number of approaches to generating extended formulations more or less automatically. We will only list a few of them:

- Lovász and Schrijver provide ways to construct extended formulations for general 0/1-problems [78] and apply their method to develop a polynomial time algorithm for the weighted stable set problem on certain graphs. A similar, but more elaborate approach has been developed by Sherali and Adams [109]. In both approaches, the authors introduce new variables corresponding to *products of the original variables*.
- Another common idea is the derivation of extended formulations from *dynamic programming* algorithms due to Martin et al. ([86]).
- Also, *disjunctive programming* (for instance [8]) can be seen as a method to generate canonical extended formulations for a certain class of polytopes. Let a set of k nonempty polyhedra $\mathbf{P}_i = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}_i \boldsymbol{x} \leqslant \boldsymbol{b}_i\}$ be given, as well as sets $\mathcal{V}_i$ of vertices and sets $\mathcal{R}_i$ of generators of the cones for each $\mathbf{P}_i$, $i \in [k]$. Then the polyhedron

  $$\{\boldsymbol{y} \in \mathbb{R}^n \mid \boldsymbol{y} = \sum_{i=1}^k \boldsymbol{\mu}_i,\ \boldsymbol{A}_i \boldsymbol{\mu}_i \leqslant \boldsymbol{b}_i \mu_i^0,\ \sum_{i=1}^k \mu_i^0 = 1,\ (\boldsymbol{\mu}_i, \mu_i^0) \in \mathbb{R}^{n+1},\ \mu_i^0 \geqslant 0\ \forall i \in [k]\}$$

  is an extended formulation for the polyhedron

  $$\operatorname{conv}(\bigcup_{i=1}^k \mathcal{V}_i) + \operatorname{cone}(\bigcup_{i=1}^k \mathcal{R}_i).$$

- Haus, Köppe and Weismantel construct extended formulations when relaxing *knapsack relaxations* in IPs (see [69]).

The use of extended formulations is a classic approach for proofs of the integrality of a linear description. (See for example the proof of Balas and Pulleyblank concerning the linear description of the perfectly matchable subgraphs of bipartite ([11]) and arbitrary graphs ([12].)

Extended formulations have been used in connection with the linear description of hop-constrained path polytopes (Stephan, [111]), Gomory corner polyhedra (Köppe et al. [70]), the stable set polytope for distance claw free graphs (Pulleyblank and Shepherd [102]), the maximal clique problem with edge weights (Lee et al. [100]), mixed integer programs in general (Wolsey [118]) or packing and partitioning orbitopes (Faenza and Kaibel, [35]).

An overview on examples of the use of extended formulations can be found in the excellent survey of Conforti, Cornuéjols and Zambelli ([26]).

We will mainly use extended formulations in connection with faithful sectionings (see section 2.4).

Remark 2.7 Note that the fact that any polytope is the projection of a higher dimensional standard simplex gives a trivial extended formulation for any polytope $\mathbf{P}$: one simply maps the i-th unit vector $\mathbb{e}^i \in \mathbb{R}^{|\mathcal{V}_\mathbf{P}|}$ to the i-th vertex in the vertex set $\mathcal{V}_\mathbf{P}$ of the polytope $\mathbf{P}$. However, this extended formulation does not yield much of an insight into the structure of the original problem. Moreover, as the number of vertices of $\mathbf{P}$ is often exponential in the dimension of the polytope, the same holds for the dimension of the simplex. On the other hand, the simplex can be easily described as

$$\{\boldsymbol{y} \in \mathbb{R}^{|\mathcal{V}_\mathbf{P}|} \mid \sum_{i=1}^{|\mathcal{V}_\mathbf{P}|} y_{\boldsymbol{v}_i} = 1,\ y_{\boldsymbol{v}_i} \geqslant 0\ \forall \boldsymbol{v}_i \in \mathcal{V}_\mathbf{P}\}.$$

To sum up, one can say that extended formulations are a way to spread difficulty on two shoulders: instead of directly focusing on the linear description of a polyhedron, one deals with the linear description of the extension polyhedron on the one hand and with its linear projection on the other.

2.3 Dynamic Programming

In general, a combinatorial optimization problem is characterized by a finite ground set $\mathcal{G}$, a set of feasible solutions $\mathcal{S} \subseteq 2^{\mathcal{G}}$ of subsets of $\mathcal{G}$ and an objective function $f : \mathcal{S} \to \mathbb{Q}$ which maps each solution to a specific rational value. The problem can then be formulated as

$$\min_{x \in \mathcal{S}} f(x).$$

A problem solvable by dynamic programming (DP) is a special case of a combinatorial optimization problem. Here, the problem is solved by breaking it into finitely many subproblems that can be solved in a stacked process. The subproblems are usually referred to as *stages* and the solutions to the subproblems are called *states*. A *decision* chooses states from earlier stages and transforms them into a new state. Correspondingly, each state has a value assigned that is composed from the values of a combination of earlier states and the cost of the transition. This value is computed by the Bellman equations, following the optimality principle which can be stated as follows: any sequence of decisions (also referred to as *policy*) $d_1, \ldots, d_n$ is optimal only if for all $k \in [n]$, subsequence $d_k, \ldots, d_n$ is optimal under the assumption that decisions $d_1, \ldots, d_{k-1}$ are done. (Detailed descriptions of dynamic programming can be found for instance in [16], [17, 18] or [74].)

In the easiest case, every state arises from transforming a *single* predecessor state. In this case, the policies correspond to simple s-t-paths in a common acyclic network, where each node corresponds to a state and each decision is represented by an arc. Node s represents an initial state and t a final (global) state. The optimal policy corresponds to a path that is shortest with respect to the total sum of arc weights in the path, where the arc weights correspond to the costs of transitions. As a consequence, many problems solvable by DP are in fact solved by computing a shortest path in a directed acyclic graph. Classical examples include the binary knapsack problem (Bellman and Dantzig, [16]), the Dreyfus-Wagner-algorithm for the computation of Minimum Steiner Trees ([30]), the dynamic programming approach to solve TSP, algorithms for sequence alignment, and subsequence counting problems which have a certain practical importance in connection with DNA sequencing ([93, 72, 34]) or simple lot-sizing problems ([119]).

However, these models turn out to be too simple in cases where *more than one* state must be combined for the computation of an output state (which can already happen for more advanced lot-sizing problems). To model problems of this kind in terms of graph theory, one makes use of directed acyclic *hyper*graphs instead of common digraphs. This insight came up in the 1980s (see [51]; overviews in [6, 43, 73]). Examples are multi echelon lot sizing problems ([7]) and some dynamic programming algorithms derived from tree decompositions, for example Maximum Stable Set on Trees (see e.g. [29, 3]). The following small example shows the idea:

Example 2.8 Problem CHAIN MULTIPLICATION ([68]): given a sequence of matrices $\boldsymbol{A}_1 \in \mathbb{R}^{d_0 \times d_1}$ through $\boldsymbol{A}_n \in \mathbb{R}^{d_{n-1} \times d_n}$, one has to find a parenthesization such that the product $\prod_{i=1}^{n} \boldsymbol{A}_i$ is computed with as few computations as possible.

For this problem, the underlying hypergraph is $H = (\mathcal{U}, \mathcal{A})$, where node set $\mathcal{U}$ contains for each $1 \leqslant k < \ell \leqslant n$ one node $u_{k,\ell}$. In an optimal solution, state $u_{k,\ell}$ has as value the optimal number of computations for the matrix product of subsequence $\boldsymbol{A}_k$ through $\boldsymbol{A}_\ell$. The arc set $\mathcal{A}$ contains for each $1 \leqslant j \leqslant k < \ell \leqslant n$ one hyperarc $(\{u_{j,k}, u_{k+1,\ell}\}, u_{j,\ell})$. The weight of the hyperarc is the number of computations when multiplying product $\prod_{i=j}^{k} \boldsymbol{A}_i$ with product $\prod_{i=k+1}^{\ell} \boldsymbol{A}_i$, i.e. value $d_{j-1} d_k d_\ell$.

A solution corresponds to a $\mathcal{W}$-t-hyperpath in this hypergraph, from $\mathcal{W} := \{u_{i,i} \in \mathcal{U} \mid i \in [n]\}$ to final state $t := u_{1,n}$. The length of the hyperpath is determined by the total sum of its arc weights.

Looking at the matter from a geometric point of view, the convex hull over incidence vectors characterizing the $\mathcal{W}$-t-hyperpaths in a hypergraph $H = (\mathcal{U}, \mathcal{A})$ associated to the underlying problem becomes interesting. Several choices of variables can be thought of to describe the hyperpaths in H. It immediately suggests itself to encode the hyperpaths by incidence vectors $\boldsymbol{x} \in \{0,1\}^{\mathcal{A}}$, setting $x_a = 1$ if and only if arc a is used in the hyperpath. However, one could also describe the hyperpath by the nodes it is passing through or think of a mix of both descriptions.

DEFINITION 2.9 Let $H = (\mathcal{V}, \mathcal{A})$ be some directed hypergraph with $\mathcal{W} \subseteq \mathcal{V}$ and $t \in \mathcal{V}$. We will call the convex hull over all incidence vectors in $\{0,1\}^{\mathcal{V}}$ of subsets $\mathcal{V}[\mathcal{L}] \subseteq \mathcal{V}$ induced by $\mathcal{W}$-t-hyperpaths $\mathcal{L}$ in H the *hyperpath set polytope in node space*, and the convex hull of all incidence vectors in $\{0,1\}^{\mathcal{A}}$ of $\mathcal{W}$-t-hyperpaths in H the *hyperpath polytope in arc space*.

Remark 2.10 We remark that if H is a B-F-hypergraph and $\mathcal{W} = \{s\}$, then $\mathcal{W}$-t-hyperpaths become in fact simple s-t-paths in a common digraph. In this case, the hyperpath set polytope is what is in literature referred to as path set polytope (see e.g. [112]).

Moreover, we note that a description of the $\mathcal{W}$-t-hyperpath polytope in a combination of arc and node variables is provided by means of the more general concept of branched polyhedral systems which will be studied in chapter 5.

If there is a linear projection σ projecting any characteristic vector of a hyperpath (policy) with respect to nodes to the corresponding solution, then the hyperpath set polytope together with σ provides an extended formulation for the polyhedron of solutions.

However, it is $\mathcal{NP}$-hard to determine shortest hyperpaths in general directed (hyper)graphs (for common digraphs, it is obvious that HAMILTON PATH is a special case); hence, a "nice" linear description of the hyperpath set polytope can not be expected in general. (See page 3.) The situation looks different when we consider classes of hypergraphs, in particular the following class that conforms to the properties of many dynamic programming algorithms.

DEFINITION 2.11 (DP hypergraphs) A hypergraph $H = (\mathcal{U}, \mathcal{A})$ will be called *DP hypergraph* if it has the following properties:

[dph1] H has a unique sink t, i.e. one single node t such that there is no arc $a \in A$ with $t \in \operatorname{tail}(a)$.

[dph2] H is an acyclic B-hypergraph.
As $\mathcal{U}$ is finite and H is acyclic, there must exist a set of nodes

$$\varnothing \neq \mathcal{W} = \{w \in \mathcal{U} \mid \nexists a \in \mathcal{A} \text{ such that } w \in \operatorname{head}(a)\}.$$

We call $\mathcal{W}$ the *source nodes* of H. Moreover, from the acylicity and the uniqueness of sink t follows that H is connected.

[dph3] For any $\mathcal{W}$-t hyperpath $\mathcal{L}$ in H, it holds that the relaxation of $(\mathcal{U}[\mathcal{L}], \mathcal{L})$ is an arborescence rooted at t.

It is easy to see that DP hypergraphs reproduce the properties of many dynamic programming algorithms.

- A DP hypergraph $H = (\mathcal{U}, \mathcal{A})$ is a B-hypergraph. This reflects the fact that in each computational step, subsets of states are transformed into single output states.
- H has a unique sink t, which serves as a "final state", along with a set of sources $\mathcal{W} \subset \mathcal{U}$ such that none of the nodes in $\mathcal{W}$ is head node of any arc in $\mathcal{A}$.
- H is acyclic; we are not interested in cycling algorithms.

Remark 2.12 In a hypergraph $H = (\mathcal{U}, \mathcal{A})$ complying to [dph1] and [dph2], there exists for every node $u \in \mathcal{U}$ a $\mathcal{W}$-t-hyperpath passing through node u.

Why is this so? Certainly, we can bring $\mathcal{U}$ in a topological order such that for each arc $a \in \mathcal{A}$, $v \succ w$ for $v \in \text{head}(a)$ and for all $w \in \text{tail}(a)$. We will now construct a hyperpath $\mathcal{L} \subseteq \mathcal{W}$ containing some node u. If $u = t$, we initialize $\mathcal{L}$ by setting $\mathcal{L} = \varnothing$. Otherwise, we let $\mathcal{L}$ initially consist of a set of arcs $a_1, \ldots, a_n$ with the following properties:

- $u \in \text{tail}(a_1)$
- $\text{head}(a_i) \cap \text{tail}(a_{i+1}) \neq \varnothing$, $i = 1, \ldots, n-1$
- $\text{head}(a_n) = \{t\}$

This series of arcs must exist as H is an acyclic hypergraph with unique sink t. Note that for each node $v \in \mathcal{U}[\mathcal{L}]$, there cannot be more than one arc in $\mathcal{L}$ having v in its head because of the acyclicity of H.

Next, we define the set of "open ends"

$$\mathcal{U}^\circ_{\mathcal{L}} := \{v \in \mathcal{U}[\mathcal{L}] \mid v \notin \mathcal{W} \text{ and } \nexists a \in \mathcal{L} \text{ such that } v \in \text{head}(a)\}$$

We extend $\mathcal{L}$ to become a hyperpath by repeatedly applying the following steps:

(1) If $\mathcal{U}^\circ_{\mathcal{L}} \neq \varnothing$, choose the topologically largest node $v \in \mathcal{U}^\circ_{\mathcal{L}}$,
(2) add an arc $a \in \mathcal{A} \setminus \mathcal{L}$ to $\mathcal{L}$ with $\text{head}(a) = \{v\}$, and
(3) update $\mathcal{U}^\circ_{\mathcal{L}}$ accordingly.

Note that node v chosen in (1) cannot be in $\mathcal{W}$ by definition of $\mathcal{U}^\circ_{\mathcal{L}}$. But by Definition 2.11 of $\mathcal{W}$, it holds that for every node $v \notin \mathcal{W}$, there must be an arc a with $\{v\} = \text{head}(a)$, and $a \notin \mathcal{L}$ by definition of $\mathcal{U}^\circ_{\mathcal{L}}$. Hence, the arc a added in (2) must exist. After the update, $\mathcal{U}^\circ_{\mathcal{L}}$ contains only nodes topological smaller than v. As the vertex set $\mathcal{U}$ is finite, the algorithm must therefore terminate with $\mathcal{U}^\circ_{\mathcal{L}} = \varnothing$.

By construction, there is for each node $v \in \mathcal{U}[\mathcal{L}] \setminus \mathcal{W}$ exactly one arc in $\mathcal{L}$ with $\text{head}(a) = \{v\}$. Moreover, for any arc $a \in \mathcal{L}$, it holds that $\text{head}(a) \cap \mathcal{W} = \varnothing$. So, we can order the arcs in $\mathcal{L}$ according to the topological order of their head nodes. This ordered arc set obviously satisfies [p1] and [p2]. Moreover, it is not possible to remove a subset of arcs from $\mathcal{L}$ without violating [p1] and [p2], because then, there exist nodes that are not head of any arc in $\mathcal{L}$. Hence, $\mathcal{L}$ is a hyperpath.

From Lemma 1.2, we get already that the relaxation $D_{\mathcal{L}}$ of $(\mathcal{U}[\mathcal{L}], \mathcal{L})$ is an acyclic digraph for any hyperpath $\mathcal{L}$. [dph3] ensures that $D_{\mathcal{L}}$ does even not contain any undirected cycles.

Remark 2.13 Condition [dph3] can be alternatively formulated by introducing a *reference set system* $\mathcal{R}$, as Martin et al. do. (See [86]):

[dph3'] Each node $u \in \mathcal{U}$ of the hypergraph is labeled with a nonempty set $\mathcal{R}(u) \in \mathcal{R}$ such that for each arc $a \in \mathcal{A}$, it holds:

- $u, v \in \text{tail}(a)$ with $u \neq v \Rightarrow \mathcal{R}(u) \cap \mathcal{R}(v) = \varnothing$
- $u \in \text{tail}(a)$, $v \in \text{head}(a) \Rightarrow \mathcal{R}(u) \subseteq \mathcal{R}(v)$

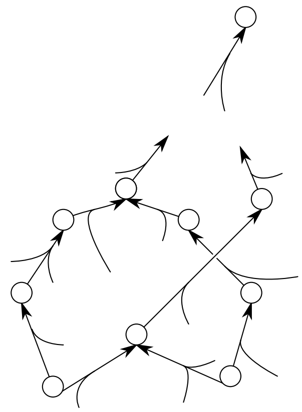

Figure 2.1: Condition [dph3'] forbids in particular cycles with more than one sink in $\mathcal{W}$-t-hyperpaths.

Proof. ⇒ We develop $\mathcal{R}$ from [dph3]. Since H is acyclic, we can bring the node set $\mathcal{U}$ in topological order. Starting in that order from the last node in $\mathcal{W}$, we set $\mathcal{R}(v) := v$ if $v \in \mathcal{W}$ and $\mathcal{R}(v) := \bigcup_{u \in \text{pred}_H(v)} \mathcal{R}(u)$, otherwise. Obviously, we get for any $a \in \mathcal{A}$ with $\text{head}(a) = v$ that for all $u \in \text{tail}(a)$, the inclusion $\mathcal{R}(u) \subseteq \mathcal{R}(v)$ is satisfied.

Now assume that there exists an arc a with $\text{head}(a) = v$, $u, w \in \text{tail}(a)$, $u \neq w$, and $\mathcal{R}(u) \cap \mathcal{R}(w) \neq \varnothing$. Then we can choose two sequences of arcs $\mathcal{L}_\alpha := (\alpha_1, \ldots, \alpha_k)$ and $\mathcal{L}_\beta := (\beta_1, \ldots, \beta_\ell)$ such that

(i) $\alpha_1 = \beta_1 = a$,
(ii) $\text{head}(\alpha_i) \in \text{tail}(\alpha_{i-1})$ for all $1 < i \leqslant k$ and, similarly, $\text{head}(\beta_i) \in \text{tail}(\beta_{i-1})$ for all $1 < i \leqslant \ell$,
(iii) $\text{tail}(\alpha_i) \cap \text{tail}(\beta_j) = \varnothing$ for all $i < k$, $j < \ell$, and
(iv) $\text{tail}(\alpha_k) \cap \text{tail}(\beta_\ell) \neq \varnothing$.

Furthermore, we can as in the proof of remark 2.12 choose a set $\mathcal{L}_t \subseteq \mathcal{A}$ connecting node v with node t. We define initially a set $\mathcal{L}$ as $\mathcal{L}_t \cup \mathcal{L}_\alpha \cup \mathcal{L}_\beta$. Proceeding algorithmically as in the proof of remark 2.12, we end up with a set $\mathcal{L}$ of arcs that has all open tail nodes in $\mathcal{W}$ and a single open head node in node t. We claim that $\mathcal{L}$ is a hyperpath. [p1] and [p2] are obviously satisfied, and since by construction deleting a subset of arcs from $\mathcal{L}$ leaves open tail nodes not in $\mathcal{W}$, we get that [p3] is satisfied, too. So we found an hyperpath which has a relaxation that is not an arborescence, contradicting [dph3].

⇐ Assume that [dph3'] holds for DP hypergraph H. It suffices to show that for any $\mathcal{W}$-$\boldsymbol{t}$-hyperpath $\mathcal{L}$, $D_\mathcal{L} = (\mathcal{U}[\mathcal{L}], \mathcal{L})$ does not contain undirected cycles. Assume the contrary. Each undirected cycle $\mathcal{C}$ in the undirected version $G_\mathcal{L}$ of digraph $D_\mathcal{L}$ must contain a unique sink. Otherwise, there must be an arc a in $\mathcal{L}$ violating condition $\mathcal{R}(u) \cap \mathcal{R}(w) = \varnothing$ for all $u, w \in \text{tail}(a)$, $u \neq w$ (see figure 2.1). This implies that there is also a unique source in $\mathcal{C}$. Moreover, each cycle $\mathcal{C}$ must contain one unique node $v_\mathcal{C}^{\max}$ ($v_\mathcal{C}^{\min}$) that is maximal (minimal) among all nodes in $\mathcal{C}$ with respect to the topological order of $D_\mathcal{L}$. Since there is only one source in $\mathcal{C}$, for all nodes v in $\mathcal{C}$ holds that $\mathcal{R}(v_\mathcal{C}^{\min}) \subseteq \mathcal{R}(v)$.

We proceed by induction along the topological order of $D_\mathcal{L}$ to show that no node in $\mathcal{U}[\mathcal{L}]$ can serve as a maximal node in a cycle $\mathcal{C}$. Clearly, there can not be a cycle $\mathcal{C}$ with $v_\mathcal{C}^{\max}$ among the sources $\mathcal{W}$ by definition of $\mathcal{W}$.

Now assume that there is an undirected cycle $\mathcal{C}$ such that $v_\mathcal{C}^{\max} \notin \mathcal{W}$ is maximal node in $\mathcal{C}$. Let u, w be the direct predecessors of $v_\mathcal{C}^{\max}$ such that arcs $(u, v_\mathcal{C}^{\max})$ and $(w, v_\mathcal{C}^{\max})$ are both in $\mathcal{C}$. As $\varnothing \neq \mathcal{R}(v_\mathcal{C}^{\min}) \subseteq \mathcal{R}(u) \cap \mathcal{R}(w)$, there can be no arc a in $\mathcal{L}$ with $\{u, w\} \subseteq \text{tail}(a)$ because of [dph3']. So u and w must be in the tails of two different arcs a_u and a_w with $\text{head}(a_u) = \text{head}(a_w) = \{v_\mathcal{C}^{\max}\}$. The proof of Lemma 1.2 shows that one of these arcs must be dispensable. Hence, $\mathcal{L}$ cannot be a hyperpath. Contradiction. ◻

Martin et al. gave a complete linear description of the path polytope in arc variables for DP-hypergraphs ([86]):

Theorem 2.14 ([86]) *Let $H = (\mathcal{U}, \mathcal{A})$ be a DP hypergraph. Then the following inequalities provide a complete linear description of the path polytope of $\mathcal{W}$-t-hyperpaths in arc space:*

$$\sum_{a \in \delta_H^{\text{in}}(t)} x_a = 1 \tag{2.1}$$

$$\sum_{a \in \delta_H^{\text{in}}(u)} x_a - \sum_{a \in \delta_H^{\text{out}}(u)} x_a = 0 \qquad \forall u \in \mathcal{U} \setminus (\mathcal{W} \cup \{t\}) \tag{2.2}$$

$$x_a \geqslant 0 \qquad \forall a \in \mathcal{A} \tag{2.3}$$

If we additionally postulate that H is a BF-hypergraph (i.e. a common digraph), then this description becomes the well-known linear description of the path polytope (in arc variables) for acyclic digraphs (see [57]).

As will be shown later (see page 55), the DP algorithm for optimization over orbitopes (see figure 3.9) allows the construction of a corresponding DP-hypergraph. For the special case of orbisacks, it is even possible to achieve a complete description of the hyperpath set polytope (in node variables on this hypergraph).

2.4 Faithful Sectioning

Faithful sectioning can be a useful tool to derive the $\mathcal{H}$-representation of a polyhedron from its $\mathcal{V}$-representation. The idea is not completely new. However, as far as we know, it has never been introduced systematically. In the literature, faithful sectionings are not explicitly tagged, but informally described as a reduction of one polyhedron to another (see for instance [107]).

Definition 2.15 Let $\sigma : \mathbb{R}^d \to \mathbb{R}^n$ be a linear projection from $\mathbb{R}^d$ to $\mathbb{R}^n$. We call a map $s : \mathbb{R}^n \to \mathbb{R}^d$ a *σ-section*, if $\sigma(s(\boldsymbol{x})) = \boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^n$.

Let now $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ be a set of linear inequalities with $\boldsymbol{A} \in \mathbb{Q}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{Q}^m$ that is valid for polyhedron $\mathbf{P}$, and let $\mathbf{Q}$ be a polyhedron in $\mathbb{R}^d$ which is projected to $\mathbf{P}$ by means of the projection σ. The main idea is to show that for each $\boldsymbol{x} \in \mathbb{R}^n$ satisfying $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$, the relation $s(\boldsymbol{x}) \in \mathbf{Q}$ holds.

Definition 2.16 We say that the inequality system $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ is *$\mathbf{Q}$-enforcing for s* if and only if for all $\boldsymbol{x} \in \mathbb{R}^n$ satisfying $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ it follows that $s(\boldsymbol{x}) \in \mathbf{Q}$.

Theorem 2.17 *Let polyhedra $\mathbf{P} \subseteq \mathbb{R}^n$ and $\mathbf{Q} \subseteq \mathbb{R}^d$ be given. Moreover, let a set of linear inequalities $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ with $\boldsymbol{A} \in \mathbb{Q}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{Q}^m$ be given that are valid for $\mathbf{P}$. Let $\sigma : \mathbb{R}^d \to \mathbb{R}^n$ be a linear projection such that $\sigma(\mathbf{Q}) \subseteq \mathbf{P}$ and let the map $s : \mathbb{R}^n \to \mathbb{R}^d$ be a σ-section. If the inequalities $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ are $\mathbf{Q}$-enforcing for s, then $\mathbf{P} = \{\boldsymbol{x} \in \mathbb{R}^m \mid \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\}$.*

Proof. Figure 2.2 illustrates the idea. Let $\mathbf{P}'$ denote the set $\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\}$. Clearly, $\mathbf{P} \subseteq \mathbf{P}'$, because the set of inequalities is valid for $\mathbf{P}$. So it remains to show that $\mathbf{P}' \subseteq \mathbf{P}$. Let $\boldsymbol{x} \in \mathbf{P}'$. As the inequalities are $\mathbf{Q}$-enforcing for s, $s(\boldsymbol{x}) \in \mathbf{Q}$. However, $\sigma(\mathbf{Q}) \subseteq \mathbf{P}$, and as s is a σ-section, $\sigma(s(\boldsymbol{x})) = \boldsymbol{x}$ holds. Hence, $\boldsymbol{x}$ must be in $\sigma(\mathbf{Q})$ and therefore in $\mathbf{P}$. $\square$

As will be shown in the following, the application of this method does not necessarily require an educated "guess" at the linear description of $\mathbf{P}$. Moreover, we remark that map s does not have to be defined in every point in $\mathbb{R}^n$. Instead, it

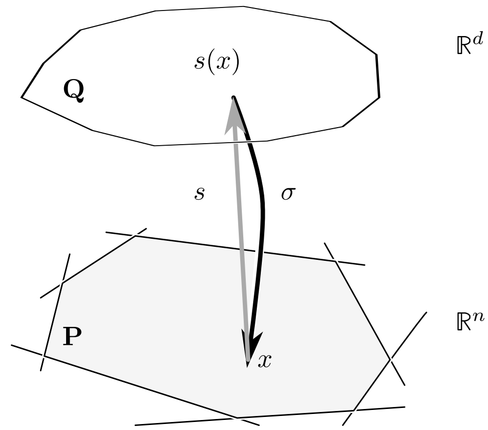

Figure 2.2: Projection σ and σ-sectioning s.

suffices for s to be defined for all $\boldsymbol{x} \in \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\}$, since it can be extended arbitrarily to $\mathbb{R}^n$.

In what follows, we will present some instructive examples for the application of faithful sectionings.

2.4.1 Applications and Examples

2.4.1.1 The Clique Polytope with Clique Size 2

Definition 2.18 (Clique polytopes) Let $G = (\mathcal{V}, \mathcal{E})$ be a graph and let $\mathcal{C}^k$ be a k-clique in G. The incidence vector $\boldsymbol{x}[C^k] \in \{0,1\}^{\mathcal{V}}$ induced by $\mathcal{C}^k$ has entries $x_v = 1$ if and only if $v \in \mathcal{C}^k$. For $k = 0$, we define $\boldsymbol{x}[\mathcal{C}^0] := \mathbb{0}$. Denoting by

$$\mathcal{X}^k := \{\boldsymbol{x} \in \{0,1\}^{\mathcal{V}} \mid \exists k\text{-clique } \mathcal{C}^k \text{ in } G \text{ such that } \boldsymbol{x} = \boldsymbol{x}[\mathcal{C}^k]\},$$

we define the *k-clique polytope* as

$$\mathbf{P}^k(G) := \operatorname{conv}(\mathcal{X}^k),$$

the *clique polytope with bounded clique size k* as

$$\mathbf{P}^{\leqslant k}(G) := \operatorname{conv}(\bigcup_{i=0}^{k} \mathcal{X}^i),$$

The clique number $\omega(G)$ of graph G denotes the size of a largest clique in G. The polytope $\mathbf{P}^{\leqslant \omega(G)}$ is called the *clique polytope*.

As the cliques of graph G correspond to stable sets in the complement $\overline{G}$ of G, polytope $\mathbf{P}^k(G)$ is isomorphic to the convex hull of stable sets of size k in $\overline{G}$ for any k. However, finding an independent set of maximal size in G (or, equivalently, finding a k-clique of maximal size in G) is an $\mathcal{NP}$-hard problem in general graphs ([46]). In the sense of the results of Papdimitriou and Karp (see page 3, there is no hope of obtaining a "nice" complete linear description for $\mathbf{P}^k(G)$ or $\mathbf{P}^{\leqslant k}(G)$.

But even for graph classes with known efficient algorithms for Max Independent Set — as for instance apple (pan) free graphs (which include the claw free graphs, see [22]) or $2K_2$-free graphs (see [79]) — there is no description for stable set polytopes available. However, there is a lot of development in this field. A complete description of $\mathbf{P}^{\leqslant \omega(G)}(G)$ was formulated some years ago by Eisenbrand

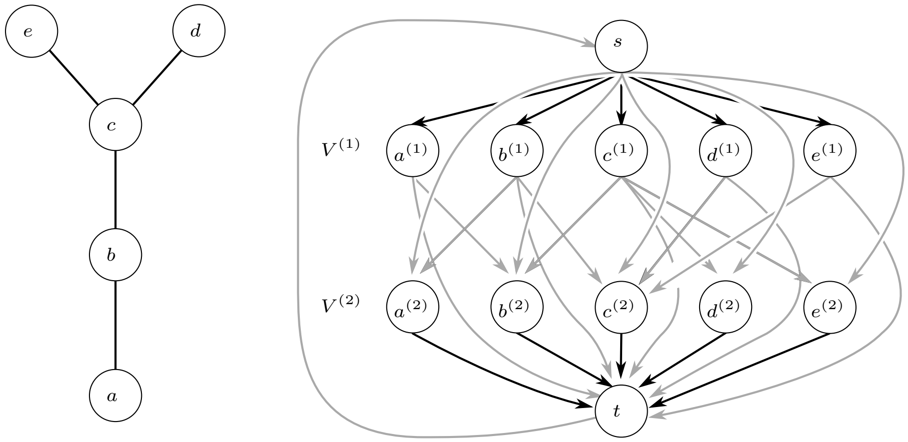

Figure 2.3: A graph G and its associated digraph D used in the construction of the extended formulation for $\mathbf{P}^{\leqslant 2}$. Gray arcs — except arc (t, s) — have lower capacity bound 0 and upper capacity bound $+\infty$.

et al. for quasi-line graphs ([33]). Ventura et al. found a linear description (by extended formulation) for geared (fuzzy) line graphs ([45, 44]) and Oriolo et al. recently obtained a description for claw free graphs with stability number at least 4 and containing neither homogeneous pairs nor 1-joins ([96]. The latter proof is especially interesting for us as it relies on graph decomposition and is related to branched polyhedral systems (see chapter 5; in fact, Oriolo et al. are using the strip-decomposition of graphs proposed by Chudnovsky and Seymour, see [25]).

Much less is known about the linear description of polytopes $\mathbf{P}^{\leqslant k}$, although for fixed k, optimization can be done in polynomial time. (Note that checking every subgraph of size at most k to be a clique costs time of $\mathcal{O}(n^k k^2)$ in a graph with n nodes.)

Janssen and Kilakos found a linear description of the polytope $\mathbf{P}^{\leqslant 2}(G)$ for general graphs G, formulated from the stable set point of view ([59]). Their proof is based on the work of Cook and Shepherd ([108]) and the use of sequential liftings ([99]). However, it has slight flaws concerning the characterization of facets (see 2.27).

Our own experiments show that for $\mathbf{P}^{\leqslant k}(G)$, things become much more involved if $k \geqslant 3$ and G is a general graph. For graphs G with bipartite complement, i.e. for graphs with a vertex set that can be partitioned into two subsets each inducing a complete subgraph, Janssen and Kilakos gave a linear description of $\mathbf{P}^{\leqslant 3}(G)$.

In what follows, we will use the technique of faithful sectioning to (re)prove a linear description of $\mathbf{P}^{\leqslant 2}$ for some given graph $G = (\mathcal{V}, \mathcal{E})$. The main idea is the use of an extended formulation based on the cycle polytope of a digraph D associated with G. While a complete "nice" description of the cycle polytope of general digraphs seems out of reach, it can nevertheless be formulated for special digraphs. In our case, we will derive the cycle polytope for a special digraph from the circulation polytope of D using Hoffman's theorem describing circulations in arbitrary digraphs.

Theorem 2.19 (Circulation theorem, [57]) *Let $D = (\mathcal{W}, \mathcal{A})$ be some digraph with upper capacity bound $u_a \in \mathbb{R} \cup \{-\infty, +\infty\}$ and lower capacity bound $\ell_a \in \mathbb{R} \cup \{-\infty, +\infty\}$ on each arc $a \in \mathcal{A}$. Then a circulation exists if and only if*

$$\sum_{a \in \delta_D^{\text{in}}(\mathcal{S})} \ell_a \leqslant \sum_{a \in \delta_D^{\text{out}}(\mathcal{S})} u_a \quad \forall \mathcal{S} \subset \mathcal{W}.$$

Moreover, if there is a feasible circulation in D and if $\boldsymbol{\ell}$ and $\boldsymbol{u}$ are both integral, then there exists an integer-valued feasible circulation.

DEFINITION 2.20 We call the set of all circulations in some digraph $D = (\mathcal{W}, \mathcal{A})$

$$\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}, \boldsymbol{\ell}) := \{\boldsymbol{y} \in \mathbb{R}^{\mathcal{A}} \mid \sum_{a \in \delta_D^{\text{in}}(w)} y_a - \sum_{a \in \delta_D^{\text{out}}(w)} y_a = 0 \ \forall w \in \mathcal{W},\ \ell_a \leqslant y_a \leqslant u_a \ \forall a \in \mathcal{A}\}$$

the *circulation polyhedron* of D. Polyhedron $\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}, \boldsymbol{\ell})$ is *integral* if $\boldsymbol{u}$ and $\boldsymbol{\ell}$ are integral. (This follows from the total unimodularity of the constraints matrix, since the $\mathcal{V}$-$\mathcal{E}$-incidence matrix of any digraph is total unimodular.)

Now let us return to $\mathbf{P}^{\leqslant 2}(G)$ on graph $G = (\mathcal{V}, \mathcal{E})$. We define a digraph $D = (\mathcal{W}, \mathcal{A})$ with node set $\mathcal{W}$ defined as

$$\mathcal{W} := (\mathcal{V} \times \{1, 2\}) \uplus \{s, t\}.$$

So $\mathcal{W}$ is the union of two copies $\mathcal{V}^1$ and $\mathcal{V}^2$ of the node set $\mathcal{V}$ of G and two additional nodes s and t. For a node $v \in \mathcal{V}$ or subset of nodes $\mathcal{U} \subseteq \mathcal{V}$, we denote by $v^i \in \mathcal{V}^i$ or $\mathcal{U}^i \subseteq \mathcal{V}^i$, $i \in \{1, 2\}$, their respective copies. The arc set $\mathcal{A}$ of D is defined as the set containing

- the arc (t, s),
- all arcs pointing from s to $\mathcal{V}^1 \cup \mathcal{V}^2$,
- all arcs pointing from from $\mathcal{V}^1 \cup \mathcal{V}^2$ to t, and
- for any edge $\{v, w\} \in \mathcal{E}$ one arc (v^1, w^2) and one arc (w^1, v^2).

Figure 2.3 shows an example for a graph G and its associated digraph D.

Defining upper and lower capacity bounds $\boldsymbol{u}^\star, \boldsymbol{\ell}^\star \in \mathbb{R}^{\mathcal{A}}$ as follows

$$\begin{aligned} \ell_a^\star &:= 0 && \text{for all } a \in \mathcal{A} \\ u_a^\star &:= 1 && \text{for } a = (t, s) \\ u_a^\star &:= +\infty && \text{for all } a \in \mathcal{A} \setminus \{(t, s)\}, \end{aligned}$$

we obtain a circulation polytope $\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^\star, \boldsymbol{\ell}^\star)$ on D, which is bounded because of the bounds on arc (t, s).

We are now ready to fix the ingredients for the faithful sectioning.

DEFINITION 2.21 We provide a projection σ as follows:

$$\sigma : \mathbb{R}^{\mathcal{A}} \to \mathbb{R}^{\mathcal{V}} \quad \text{with} \quad \sigma(\boldsymbol{y})_v = y_{(s,v^1)} + y_{(v^2,t)}.$$

Moreover, we define a map $s : \mathbb{R}_+^{\mathcal{V}} \to \mathbb{R}^{\mathcal{A}}$ which is mapping any point $\boldsymbol{x} \in \mathbb{R}_+^{\mathcal{V}}$ to an arbitrary point in the circulation polytope $\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^{\boldsymbol{x}}, \boldsymbol{\ell}^{\boldsymbol{x}})$, if the latter is not empty, and to an arbitrary point in $\mathbb{R}^{\mathcal{A}}$ otherwise. Upper and lower capacity bounds $\boldsymbol{u}^{\boldsymbol{x}}$ and $\boldsymbol{\ell}^{\boldsymbol{x}}$ are defined as follows:

$$\begin{aligned} &\left.\begin{aligned} \ell^{\boldsymbol{x}}_{(s,v^1)} &:= \tfrac{x_v}{2} \\ \ell^{\boldsymbol{x}}_{(v^2,t)} &:= \tfrac{x_v}{2} \end{aligned}\right\} \text{ for all } v \in \mathcal{V} \\ &\ell^{\boldsymbol{x}}_a := 0 \qquad \text{for all other } a \in \mathcal{A} \\ &\left.\begin{aligned} u^{\boldsymbol{x}}_{(s,v^1)} &:= \tfrac{x_v}{2} \\ u^{\boldsymbol{x}}_{(v^2,t)} &:= \tfrac{x_v}{2} \end{aligned}\right\} \text{ for all } v \in \mathcal{V} \\ &u^{\boldsymbol{x}}_{(t,s)} := 1 \qquad \text{for } a = (t, s) \\ &u^{\boldsymbol{x}}_a := +\infty \qquad \text{for all other } a \in \mathcal{A}, \end{aligned}$$

In figure 2.3, any gray arc except arc (t, s) gets lower capacity bound 0 and upper capacity bound $+\infty$. All black arcs get lower and upper bound $\frac{x_v}{2}$, where v^1 is end node or v^2 is start node of the arc, respectively.

Lemma 2.22 $\sigma(\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^\star, \boldsymbol{\ell}^\star)) \subseteq \mathbf{P}^{\leqslant 2}(G)$.

Proof. For any circulation $\boldsymbol{y}$ in $D = (\mathcal{W}, \mathcal{A})$ it holds that the flow on *each* arc in D is implicitly bounded from above by 1 by construction of D and the upper bound on arc (t, s). Moreover, $y_a \geqslant 0$ for any $a \in \mathcal{A}$. By Theorem 2.19 and Definition 2.20, we get that each vertex $\boldsymbol{y}$ of $\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^\star, \boldsymbol{\ell}^\star)$ is integral, which means that in particular $\boldsymbol{y}_a \in \{0, 1\}$ for each $a \in \mathcal{A}$. Hence, $\operatorname{supp}^+(\boldsymbol{y})$ is a simple cycle in D for each vertex $\boldsymbol{y}$, and by construction of D, $\sigma(\boldsymbol{y})$ is a vertex of $\mathbf{P}^{\leqslant 2}(G)$. As σ is linear, $\sigma(\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^\star, \boldsymbol{\ell}^\star)) \subseteq \mathbf{P}^{\leqslant 2}(G)$. □

Lemma 2.23 *If the polytope* $\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^{\boldsymbol{x}}, \boldsymbol{\ell}^{\boldsymbol{x}})$ *is nonempty, then* $\sigma(s(\boldsymbol{x})) = \boldsymbol{x}$.

Proof. If polytope $\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^{\boldsymbol{x}}, \boldsymbol{\ell}^{\boldsymbol{x}})$ is nonempty, then there is a feasible circulation in D satisfying upper and lower bounds. This circulation must send $\frac{x_v}{2}$ units of flow over arcs (s, v^1) and (v^2, t) for all $v \in \mathcal{V}$. Hence, $\sigma(s(\boldsymbol{x})) = \boldsymbol{x}$. □

The crucial observation is that for any $\boldsymbol{x} \in \mathbb{R}^{\mathcal{V}}$,

$$\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^{\boldsymbol{x}}, \boldsymbol{\ell}^{\boldsymbol{x}}) \subset \mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^\star, \boldsymbol{\ell}^\star).$$

Hence, if polytope $\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^{\boldsymbol{x}}, \boldsymbol{\ell}^{\boldsymbol{x}})$ is nonempty for some $\boldsymbol{x} \in \mathbb{R}^{\mathcal{V}}$, then $s(\boldsymbol{x}) \in \mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^\star, \boldsymbol{\ell}^\star)$. This means: if we ensure by some set of inequalities that for any $\boldsymbol{x} \in \mathbb{R}^{\mathcal{V}}$ satisfying these inequalities, polytope $\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^{\boldsymbol{x}}, \boldsymbol{\ell}^{\boldsymbol{x}})$ is nonempty, then this inequality system is at the same time $\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^\star, \boldsymbol{\ell}^\star)$-enforcing for s. However, this wanted inequality system has in general already been formulated in the main statement in Hoffman's circulation theorem.

Lemma 2.24 *The following set of inequalities is* $\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^\star, \boldsymbol{\ell}^\star)$*-enforcing for* s*:*

$$2\sum_{v \in \mathcal{T}} x_v \;+ \sum_{v \in \mathcal{V} \smallsetminus (\mathcal{T} \cup \mathrm{N}_G(\mathcal{T}))} x_v \leqslant 2 \quad \forall \mathcal{T} \subseteq \mathcal{V} \text{ stable in } G \tag{2.4}$$

$$x_v \geqslant 0 \quad \forall v \in \mathcal{V} \tag{2.5}$$

Proof. By Theorem 2.19, there is a circulation in $\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^{\boldsymbol{x}}, \boldsymbol{\ell}^{\boldsymbol{x}})$ if and only if

$$\sum_{a \in \delta_D^{\mathrm{in}}(\mathcal{S})} \ell_a^{\boldsymbol{x}} \leqslant \sum_{a \in \delta_D^{\mathrm{out}}(\mathcal{S})} u_a^{\boldsymbol{x}} \quad \forall \mathcal{S} \subset \mathcal{W}. \tag{*}$$

However, we observe:

- If $s \in \mathcal{S}$ and there is a node $v \in \mathcal{V}^2$ with $v \notin \mathcal{S}$, then the right-hand side of (*) becomes $+\infty$, and
- if $\{s\} \cup \mathcal{V}^2 \subseteq \mathcal{S}$, then the left-hand side of (*) becomes 0, as there is no arc a entering $\mathcal{S}$ with $\ell_a > 0$.
- Similarly, if $t \notin \mathcal{S}$ and $\mathcal{S} \cap \mathcal{V}^1 \neq \varnothing$, then the right-hand side of (*) becomes $+\infty$ and
- if $t \notin \mathcal{S}$ and $\mathcal{S} \cap \mathcal{V}^1 = \varnothing$, then the left-hand side of (*) becomes 0.
- If there is an arc $a \in \mathcal{V}^1 \times \mathcal{V}^2$ with $a \in \delta^{\mathrm{out}}(\mathcal{S})$, then the right-hand side of (*) becomes $+\infty$.

In all these cases, inequality (*) is trivially satisfied. Hence, we can in the following assume that

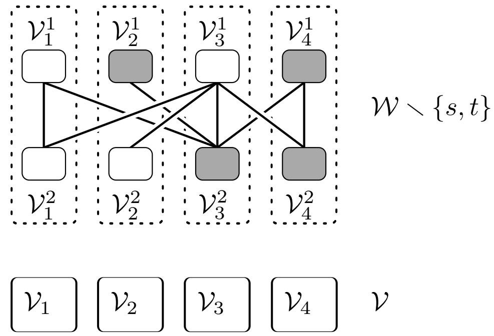

Figure 2.4: Partitioning node set $\mathcal{V}^1 \cup \mathcal{V}^2$. If $\mathcal{V}_i^1$ and $\mathcal{V}_j^1$ are not connected with a line, then we can assume that there is no arc in $\mathcal{V}_i^1 : \mathcal{V}_j^1$, as described in the proof. Subsets of $\mathcal{S}$ are gray. The partition of $\mathcal{W}$ induces a partition of $\mathcal{V}$ (dashed sets).

(i) $s \notin \mathcal{S}$, that
(ii) $t \in \mathcal{S}$ and that
(iii) $(\mathcal{S} \cap \mathcal{V}^1) : (\mathcal{S} \cap \mathcal{V}^2) = \varnothing$.

Now we partition $\mathcal{W}$ into sets $\mathcal{V}_1^i$ through $\mathcal{V}_4^i$, $i \in \{1,2\}$, such that

$$\begin{aligned} &\mathcal{V}_1^1 \cap \mathcal{S} = \varnothing \text{ and } \mathcal{V}_1^2 \cap \mathcal{S} = \varnothing && \mathcal{V}_2^1 \subseteq \mathcal{S} \text{ and } \mathcal{V}_2^2 \cap \mathcal{S} = \varnothing \\ &\mathcal{V}_3^1 \cap \mathcal{S} = \varnothing \text{ and } \mathcal{V}_3^2 \subseteq \mathcal{S} && \mathcal{V}_4^1 \subseteq \mathcal{S} \text{ and } \mathcal{V}_4^2 \subseteq \mathcal{S}. \end{aligned}$$

Note that this partitions at the same time set $\mathcal{V}$ into sets $\mathcal{V}_1$ through $\mathcal{V}_4$; see figure 2.4.

So, we can write the left-hand side of (*) as follows:

$$\begin{aligned} &\sum_{a \in \{s\}:\mathcal{V}_2^1} \ell_a^{\boldsymbol{x}} + \sum_{a \in \{s\}:\mathcal{V}_4^1} \ell_a^{\boldsymbol{x}} + \underbrace{\sum_{a \in \{s\}:\mathcal{V}_3^2} \ell_a^{\boldsymbol{x}}}_{=0} + \underbrace{\sum_{a \in \{s\}:\mathcal{V}_4^2} \ell_a^{\boldsymbol{x}}}_{=0} + \\ &\qquad \underbrace{\sum_{a \in \mathcal{V}_1^1:\{t\}} \ell_a^{\boldsymbol{x}}}_{=0} + \underbrace{\sum_{a \in \mathcal{V}_3^1:\{t\}} \ell_a^{\boldsymbol{x}}}_{=0} + \sum_{a \in \mathcal{V}_1^2:\{t\}} \ell_a^{\boldsymbol{x}} + \sum_{a \in \mathcal{V}_2^1:\{t\}} \ell_a^{\boldsymbol{x}} = \\ &\qquad\qquad = \frac{1}{2} \sum_{v \in \mathcal{V}_1} x_v + \sum_{v \in \mathcal{V}_2} x_v + \frac{1}{2} \sum_{v \in \mathcal{V}_4} x_v. \end{aligned}$$

Because of (iii), $\mathcal{V}_i^1 : \mathcal{V}_1^2 = \varnothing$ and $\mathcal{V}_i^1 : \mathcal{V}_2^2 = \varnothing$ for $i \in \{2,4\}$. So the right-hand side of $(*)$ becomes $u_{(t,s)} = 1$. We obtain the following inequality:

$$\sum_{v \in \mathcal{V}_1} x_v + 2 \sum_{v \in \mathcal{V}_2} x_v + \sum_{v \in \mathcal{V}_4} x_v \leqslant 2 \tag{2.6}$$

Observation (iii) also implies that $\mathcal{V}_2$ is a stable set in G. Moreover, there are no edges between $\mathcal{V}_1 \cup \mathcal{V}_4$ and $\mathcal{V}_2$ in G. Hence $\mathcal{V}_1 \cup \mathcal{V}_4 \subseteq \mathcal{V} \smallsetminus (\mathcal{V}_2 \cup \mathrm{N}_G(\mathcal{V}_2))$. In particular, every inequality from (2.6) is dominated by an inequality from (2.4) with $\mathcal{V}_2 = \mathcal{T}$ and $\mathcal{V}_1 \cup \mathcal{V}_4 \subseteq \mathcal{V} \smallsetminus (\mathcal{T} \cup \mathrm{N}_G(\mathcal{T}))$, since we can additionally assume that $x_v \geqslant 0$ for all $v \in \mathcal{V}$. Therefore, it suffices to use inequalities (2.4) in connection with (2.5). □

Observation 2.25

(a) The dimension of $\mathbf{P}^{\leqslant 2}(G) \subseteq \mathbb{R}^{\mathcal{V}}$ is $|\mathcal{V}|$ because it contains $\mathbb{0}$ and standard unit vectors $\mathbb{e}^v$ for all $v \in \mathcal{V}$.

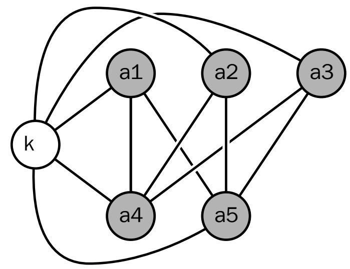

Figure 2.5: Inequality $\sum_{v\in\mathcal{A}} x_v + 2\sum_{v\in\mathcal{K}} x_v \leqslant 2$ is not facet defining in this graph.

(b) For the considerations below, the dimension of $\mathbf{P}^2(G)$ is also of interest. This dimension is clearly the rank of the node-edge incidence matrix of $G = (\mathcal{V}, \mathcal{E})$ (i.e., the dimension of the linear subspace spanned by $\mathbf{P}^2(G)$) minus one (since $\mathbf{P}^2(G)$ is contained in the hyperplane defined by $\sum_{v\in\mathcal{V}} x_v = 2$). As the rank of the node-edge incidence matrix of a connected graph with n nodes equals $n-1$ if the graph is bipartite and n otherwise (see e.g. [23]), we get that $\dim(\mathbf{P}^2(G)) = |\mathcal{V}| - 1 - \beta(G)$, with $\beta(G)$ the number of bipartite components of G.

Proposition 2.26 *For any graph $G = (\mathcal{V}, \mathcal{E})$, the following set of inequalities is a complete non-redundant linear description for $\mathbf{P}^{\leqslant 2}(G)$:*

$$2\sum_{v\in\mathcal{S}} x_v \;+ \sum_{v\in\mathcal{V}\smallsetminus(\mathcal{S}\cup \mathrm{N}_G(\mathcal{S}))} x_v \;\leqslant 2 \tag{2.7}$$

$$x_v \geqslant 0 \quad \forall v \in \mathcal{V} \tag{2.8}$$

for all $\mathcal{S} \subseteq \mathcal{V}$ that are (possibly empty) stable sets in G, where $G[\mathcal{V}\smallsetminus(\mathcal{S}\cup \mathrm{N}_G(\mathcal{S}))]$ does not contain any bipartite component.

Proof. The set of inequalities (2.4) and (2.5) provides a linear description of $\mathbf{P}^{\leqslant 2}(G)$ as all prerequisites for Theorem 2.17 are satisfied:

- The set of inequalities (2.4) and (2.5) is obviously valid for $\mathbf{P}^{\leqslant 2}(G)$.
- By Lemma 2.22, $\sigma(\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^\star, \boldsymbol{\ell}^\star)) \subseteq \mathbf{P}^{\leqslant 2}(G)$.
- By Lemmas 2.23 and 2.24, s is a σ-section for any $\boldsymbol{x}$ satisfying inequalities (2.4) and (2.5).
- Finally, by Lemma 2.24, the inequalities are $\mathbf{Q}_{\circlearrowleft}(D, \boldsymbol{u}^\star, \boldsymbol{\ell}^\star)$-enforcing.

If $\boldsymbol{x} \in \{0,1\}^{\mathcal{V}}$ is active for an inequality (2.4), then $\boldsymbol{x}$ is

- either the incidence vector of a 2-clique in $G[\mathcal{V}\smallsetminus(\mathcal{S}\cup \mathrm{N}_G(\mathcal{S}))]$ or
- $\mathbb{e}^v$ for some $v \in \mathcal{S}$ or
- the incidence vector of a 2-clique that intersects with $\mathcal{S}$.

From this and Observation 2.25 (b) follows that the affine dimension of the face defined by the inequality is $|\mathcal{V}|-1-\beta$, where β is the number of bipartite components in $\mathcal{V}\smallsetminus(\mathcal{S}\cup \mathrm{N}_G(\mathcal{S}))$. Hence, an inequality (2.4) is facet defining if and only if there are no bipartite components in $\mathcal{V}\smallsetminus(\mathcal{S}\cup \mathrm{N}_G(\mathcal{S}))$.

Consider now the active vertex set for inequality $x_v \geqslant 0$. It contains all vertices induced by 1- and 2-cliques $\mathcal{C}$ with $v \notin \mathcal{C}$ together with the 0-clique. Thus, its convex hull is the same as $\mathrm{PC}^{\leq 2}(G')$ for graph G' arising from G by dropping v and all edges adjacent to v. The dimension of $\mathrm{PC}^{\leq 2}(G')$ is $|\mathcal{V}| - 1$.

□

Remark 2.27 This corrects the results of Janssen and Kilakos, who erroneously characterize inequalities as facet defining which in fact are not, as example graph G in figure 2.5 shows.

The blank node $\mathcal{K}$ in the graph represents a clique which is adjacent to all nodes in the set $\mathcal{A} = \{a_1, a_2, a_3, a_4, a_5\}$ of (gray) nodes. So the following conditions are satisfied:

- Every maximal stable set in $G[\mathcal{A}]$ has size at least 2.
- $G[\mathcal{A}]$ has a stable set of size at least 3.
- $\mathcal{K}$ is maximal clique in $\tilde{N}(\mathcal{A}) := \bigcap_{v\in\mathcal{A}} \mathrm{N}_G(v) \smallsetminus \mathcal{A}$, where $\mathrm{N}_G(v)$ denotes the neighbourhood of v in G.

However,

$$\sum_{v\in\mathcal{A}} x_v + 2\sum_{v\in\mathcal{K}} x_v \leqslant 2$$

is not facet defining as Janssen and Kilakos claim, because it is the sum of inequalities

$$\sum_{v\in\mathcal{K}} x_v + x_{a_4} + \frac{1}{2}(x_{a_1} + x_{a_2} + x_{a_3}) \leqslant 1 \text{ and}$$
$$\sum_{v\in\mathcal{K}} x_v + x_{a_5} + \frac{1}{2}(x_{a_1} + x_{a_2} + x_{a_3}) \leqslant 1.$$

We close this section by characterizing $\mathbf{P}^2(G)$ linearly. This polytope is a face of $\mathbf{P}^{\leqslant 2}(G)$, namely the intersection of $\mathbf{P}^{\leqslant 2}(G)$ with equation $\sum_{v\in\mathcal{V}} x_v = 2$. Subtracting this equation from inequalities (2.4), we get that the following inequality set

$$\sum_{v\in\mathcal{T}} x_v - \sum_{v\in\mathrm{N}_G(\mathcal{T})} x_v \leqslant 0 \quad \forall \mathcal{T} \subseteq \mathcal{V} \text{ stable in } G \tag{*}$$

holds for any point in $\mathbf{P}^2(G)$. Any face of $\mathbf{P}^2(G)$ defined by an inequality from (*) is isomorphic to $\mathbf{P}^2(G')$ with G' the graph arising from G by deleting all edges with one end in $\mathrm{N}_G(\mathcal{T})$ and the other end not in $\mathcal{T}$. Similarly, for every $v \in \mathcal{V}$, the face of $\mathbf{P}^2(G)$ defined by $x_v \geqslant 0$ is isomorphic to $\mathbf{P}^2(G'')$, where G'' arises from G by removing v.

This establishes the following

Proposition 2.28 *For any graph $G = (\mathcal{V}, \mathcal{E})$, the following set of inequalities provides a complete non-redundant linear description for $\mathbf{P}^2(G)$:*

$$\begin{aligned}
\sum_{v\in\mathcal{V}} x_v &= 2 && \\
\sum_{v\in\mathcal{T}} x_v - \sum_{v\in\mathrm{N}_G(\mathcal{T})} x_v &= 0 && \textit{for all } \mathcal{T} \in \mathcal{B} \\
\sum_{v\in\mathcal{T}} x_v - \sum_{v\in\mathrm{N}_G(\mathcal{T})} x_v &\leqslant 0 && \textit{for all } \mathcal{T} \in \mathcal{S} \\
x_v &= 0 && \textit{for all isolated nodes } v \in \mathcal{V} \\
x_v &\geqslant 0 && \textit{for all } v \in \widetilde{\mathcal{V}}
\end{aligned}$$

where

- *$\mathcal{B}$ is the family of subsets of $\mathcal{V}$ containing from each bipartite component of G exactly one of the two shores,*
- *$\mathcal{S}$ is the family of stable sets $\mathcal{T} \subseteq \mathcal{V}$ of G with the property that removing from G all edges in $\mathcal{T} : (\mathcal{V} \smallsetminus \mathrm{N}_G(\mathcal{T}))$ increases the number of bipartite components by exactly one, and*
- *$\widetilde{\mathcal{V}}$ contains all nodes of G that can be removed without changing the number of bipartite components of G.*

Digression: the Edge Expansion of the Graph of $\mathbf{P}^2(G)$

Let $\mathbf{P}$ be some polytope with graph $H = (\mathcal{V}_H, \mathcal{E}_H)$. (For the definition of the graph of a polytope, see page 15.) The *edge expansion* of $\mathbf{P}$ is then defined as

$$\chi_e(\mathbf{P}) := \min_{\substack{0 \subset \mathcal{S} \subset \mathcal{V}_H \\ |\mathcal{S}| \leqslant \frac{|\mathcal{V}_H|}{2}}} \left(\frac{|\delta(\mathcal{S})|}{|\mathcal{S}|} \right)$$

The *vertex expansion* of the polytope is defined as

$$\chi_v(\mathbf{P}) = \min_{\substack{0 \subset \mathcal{S} \subset \mathcal{V}_H \\ |\mathcal{S}| \leqslant \frac{|\mathcal{V}_H|}{2}}} \left(\frac{|\,\mathrm{N}_H(\mathcal{S})|}{|\mathcal{S}|} \right),$$

where $\mathrm{N}_H(\mathcal{S})$ is the set of neighbours of $\mathcal{S}$ in $\mathcal{V}_H \setminus \mathcal{S}$. It is an open question whether the edge expansion of every 0/1-polytope is equal or larger than one (this has been conjectured by Mihail and Vazirani, see e.g. [37] and [91]). Clearly, the edge expansion is bounded from below by the vertex expansion. Unfortunately, 0/1-polytopes with (arbitrary) small vertex expansions do exist; they can be generated by means of probabilistic methods (see eg. [49]).

In contrast to this, we will show in Example 2.30 that $\mathbf{P}^2(G)$ can be used for a deterministic construction of polytopes with vertex expansion < 1. We will also show that $\mathbf{P}^2(G)$ has always edge expansion equal or larger than one.

We start with the characterization of adjacency in the graph of $P^2(G)$. We will use the fact that two vertices of the polytope are adjacent in the graph of the polytope if and only if there is a linear cost functional maximized by exactly these two vertices.

Lemma 2.29 *Let $e := \{u, v\}$ and $f := \{s, t\}$ be edges in $G = (\mathcal{V}, \mathcal{E})$. Then vertices $x[e]$ and $x[f]$ of $\mathbf{P}^2(G)$ are adjacent in the graph of the polytope if and only if e and f are not opposites in a four-cycle in G.*

Proof. $\Leftarrow$ Let e and f be no opposites in every 4-cycle. Let $\mathcal{K}$ the edge set of the complete graph on nodes u, v, s, t except of edges e and f. We will construct a cost vector $\boldsymbol{c}$ which is maximized by $x[e]$ and $x[f]$ only. We have to distinguish the following cases:

(i) e and f are incident in G, w.l.o.g. at node $v = s$. Define vector $\boldsymbol{c}$ with $c_u := 1$, $c_v := 2$, $c_t := 1$, components 0 otherwise. The maximum cost value of 3 is only achieved by $x[e]$ and $x[f]$.

(ii) e and f are not incident and $\mathcal{E} \cap \mathcal{K} = \varnothing$. Use then vector $\boldsymbol{c}$ with $c_u := 1$, $c_v := 1$, $c_s := 1$, $c_t := 1$ and components 0 otherwise.

(iii) e and f are not incident and $|\mathcal{E} \cap \mathcal{K}| = 1$. We can w.l.o.g. assume that edge $\{u, s\}$ exists. (Other constellations are isomorphic.) Use $c_v := 2$, $c_t := 2$, $c_u := 1$, $c_s := 1$ and components 0 otherwise.

(iv) e and f are not incident and $|\mathcal{E} \cap \mathcal{K}| = 2$. We can w.l.o.g. assume that edges $\{u, s\}$ and $\{v, s\}$ exist. (Since e and f are not in a four-cycle by assumption, all other constellations are isomorphic.) Use $c_u := 2$, $c_v := 2$, $c_s := 1$, $c_t := 3$ and components 0 otherwise.

$\Rightarrow$ Let now e and f be opposites in a four-cycle. W.l.o.g, the other edges in the cycle are $\{u, s\}$ and $\{v, t\}$. Assume we could find a cost vector $\boldsymbol{c}$ maximized by $x[e]$ and $x[f]$ only. Its components have to satisfy the following system of equations and inequalities:

$$\begin{aligned} c_u + c_v &= \zeta & \qquad c_s + c_t &= \zeta \\ c_u + c_s &< \zeta & \qquad c_v + c_t &< \zeta \end{aligned}$$

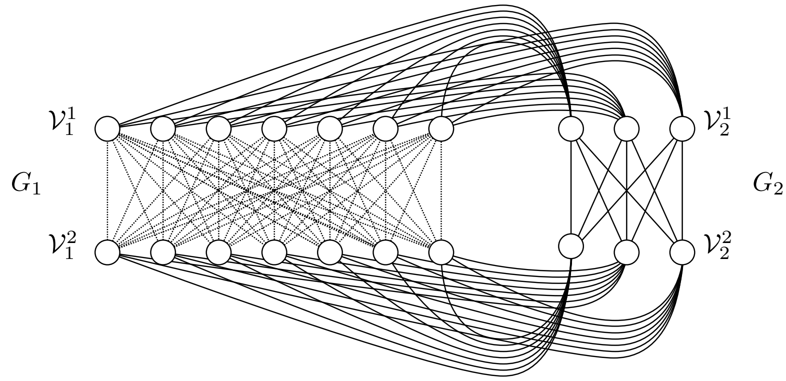

Figure 2.6: Example for graph $G^{7,3}$ inducing polytope $\mathbf{P}^2(G^{7,3})$ with small vertex expansion. Dotted edges are "red", solid edges are "blue".

So

$$\underbrace{c_u + c_s}_{<\zeta} + \underbrace{c_v + c_t}_{<\zeta} \stackrel{!}{=} 2\zeta.$$

Contradiction. □

Example 2.30 The vertex expansion of the graph of $\mathbf{P}^2(G)$ can be as small as $2(\sqrt{2}-1)+\epsilon$ for some $\epsilon > 0$, as the following example shows.

Let $G_1 := K_{m,m}$ and $G_2 := K_{n,n}$ be two complete bipartite graphs. Graph G_j has vertex set $\mathcal{V}_j := \mathcal{V}_j^1 \uplus \mathcal{V}_j^2$ and edge set $\mathcal{E}_j = \mathcal{V}_j^1 \times \mathcal{V}_j^2$ for $j \in \{1, 2\}$.

We construct the graph $G^{m,n}$ from G_1 and G_2 by adding all edges $\{u, v\}$ with $u \in V_1^j$ and $v \in V_2^j$, $j \in \{1, 2\}$. (See fig. 2.6 for an example with $m = 7$ and $n = 3$.)

Now we color all edges inside G_1 *red* and all other edges *blue*. This induces sets of red and blue nodes in the graph of polytope $\mathbf{P}^2(G^{m,n})$. We will study the adjacency of those red and blue nodes.

As long as $n^2 + 2mn \geqslant m^2$, there are at least as many blue than red nodes. There can be no connection between blue nodes coming from edges in the graph G_2 and any red node, since the corresponding edges in $G^{m,n}$ are opposites in 4-cycles. On the other hand, any other blue node is connected to a red node. Hence, the red nodes have $2mn$ blue neighbours. (In the example, there are 49 red edges and 51 blue ones, and the blue neighbourhood of the red node set has cardinality 42.)

For $m, n > 0$ with $n^2 + 2mn \geqslant m^2$, the vertex expansion of the graph of $\mathbf{P}^2(G^{m,n})$ is therefore bounded from above by $\frac{2n}{m}$. This value becomes smallest for m as large as possible, which is the case if $n^2 + 2mn - m^2 = 0$. This implies a bound of $2(\sqrt{2} - 1) \approx 0.82843$.

Proposition 2.31 *$P^2(G)$ has an edge expansion of at least* 1.

Proof. Let $G = (\mathcal{V}, \mathcal{E})$ and let $H := (\widetilde{\mathcal{V}}, \widetilde{\mathcal{E}})$ be the graph of polytope $\mathbf{P}^2(G)$. To distinguish the edges in G from the edges in H, we will refer to the latter as *meta-edges*. We color an arbitrary nonempty set of edges $\mathcal{E}_R \subset \mathcal{E}$ "red" and the complement $\mathcal{E}_B := \mathcal{E} \setminus \mathcal{E}_R$ "blue", where $|\mathcal{E}_R| \leqslant |\mathcal{E}_B|$. This induces sets $\widetilde{\mathcal{V}}_R \subset \widetilde{\mathcal{V}}$ and $\widetilde{\mathcal{V}}_B \subset \widetilde{\mathcal{V}}$ of red and blue nodes in H. Moreover, we define a blue subgraph $B := G[\mathcal{E}_B]$, a red subgraph $R := G[\mathcal{E}_R]$, and a set $\mathcal{V}_B := \{v \in \mathcal{V} \mid \deg_B(v) \geqslant 1\}$ containing all vertices in $\mathcal{V}$ incident with at least one blue edge.

The aim is to show that $|\delta_H(\widetilde{\mathcal{V}}_R)| \geqslant |\widetilde{\mathcal{V}}_R| = |\mathcal{E}_R|$.

Let $v \in V_B$. For any red edge $r \in \mathcal{E}_R$ incident with v, there are meta-edges between r and all edges in $\delta_B(v)$, because r cannot be opposite to these edges in a

4-cycle. This implies that

$$|\delta_H(\widetilde{\mathcal{V}}_R)| \geqslant \sum_{v \in \mathcal{V}_B} \deg_R(v) \deg_B(v).$$

If for all $v \in \mathcal{V}_B$, it holds that $\deg_R(v) \geqslant 1$, then

$$|\delta_H(\widetilde{\mathcal{V}}_R)| \geqslant \sum_{v \in V_B} \deg_B(v) = 2|\mathcal{E}_B| \geqslant 2|\mathcal{E}_R|.$$

But if $\deg_R(v) = 0$ for some $v \in V_B$, then v is not incident with any edge in $\mathcal{E}_R$. This implies that for every edge $r \in \mathcal{E}_R$, there exists at least one edge in $\mathcal{E}_B$ which is not opposite to r. (Assume that for some $r \in \mathcal{E}_R$, all edges in $\mathcal{E}_B$ are opposite. Then there must be in particular a 4-cycle containing v (since $\mathcal{V}_B$ does not contain isolated nodes). This cycle contains edge $b \in \mathcal{E}_B$ leading between one endpoint of r and v which cannot be opposite to r. Contradiction.) So for any edge in $\mathcal{E}_R$, there exists at least one meta-edge between $\widetilde{\mathcal{V}}_R$ and $\widetilde{\mathcal{V}}_B$, which implies $|\delta_H(\widetilde{\mathcal{V}}_R)| \geqslant |\mathcal{E}_R|$. □

2.4.1.2 The Path Set Polytope for Acyclic Digraphs

Our second example for the use of faithful sectionings concerns the linear description of the s-t-path set polytope for acyclic digraphs (see Definition 2.9), which has already been derived by Vande Vate ([112]) by means of a lift and project approach in combination with Bender's decomposition.

Let $D = (\mathcal{V}, \mathcal{A})$ be some acyclic digraph with unique source node s and unique sink node t. The s-t-path set polytope in D will be denoted by $\mathbf{P}^{s,t}(D)$.

Just like in the first example, we give an extended formulation based on a directed graph $\widetilde{D} = (\widetilde{\mathcal{V}}, \widetilde{\mathcal{A}})$. The construction is standard: node set $\widetilde{\mathcal{V}}$ is defined as

$$\widetilde{\mathcal{V}} := (\mathcal{V} \times \{\mathrm{in}, \mathrm{out}\}),$$

that is, we split each node $v \in \mathcal{V}$ into two clone nodes v^{in} and v^{out}. Again, for any subset of nodes $\mathcal{U} \subseteq \mathcal{V}$ we denote by $\mathcal{U}^{\mathrm{in}} \subseteq \mathcal{V}^{\mathrm{in}}$ or $\mathcal{U}^{\mathrm{out}} \subseteq \mathcal{V}^{\mathrm{out}}$ the sets of their respective copies in $\widetilde{\mathcal{V}}$.

The arc set $\widetilde{\mathcal{A}}$ is defined as

$$\widetilde{\mathcal{A}} := \{(v^{\mathrm{out}}, w^{\mathrm{in}}) \mid (v, w \in \mathcal{A})\} \,\cup\, \{(v^{\mathrm{in}}, v^{\mathrm{out}}) \mid v \in \mathcal{V}\} \,\cup\, (t^{\mathrm{out}}, s^{\mathrm{in}}).$$

Arcs in $\{(v^{\mathrm{out}}, w^{\mathrm{in}}) \mid (v, w \in \mathcal{A})\}$ will be referred to as *real arcs* while arcs in $\{(v^{\mathrm{in}}, v^{\mathrm{out}}) \mid v \in \mathcal{V}\}$ will be called *split arcs*; see figure 2.7 for an example. The construction of the faithful sectioning is as follows.

- We consider the circulation polytope $\mathbf{Q}_{\circlearrowleft}(\widetilde{D}, \boldsymbol{\ell}^\star, \boldsymbol{u}^\star)$ on the digraph $\widetilde{D}$ with the following bounds on the capacity:

$$\left.\begin{aligned} \ell^\star_{(v^{\mathrm{out}}, w^{\mathrm{in}})} &:= -\infty \\ u^\star_{(v^{\mathrm{out}}, w^{\mathrm{in}})} &:= +\infty \end{aligned}\right\} \text{for all } (v, w) \in \mathcal{A} \text{ (real arcs)}$$
$$\left.\begin{aligned} \ell^\star_{(v^{\mathrm{in}}, v^{\mathrm{out}})} &:= 0 \\ u^\star_{(v^{\mathrm{in}}, w^{\mathrm{out}})} &:= +\infty \end{aligned}\right\} \text{for all } v \in \mathcal{V} \text{ (split arcs)}$$
$$\begin{aligned} \ell_{(t^{\mathrm{out}}, s^{\mathrm{in}})} &:= 1 \\ u_{(t^{\mathrm{out}}, s^{\mathrm{in}})} &:= 1 \end{aligned}$$

  Any vertex $\boldsymbol{y}$ of $\mathbf{Q}_{\circlearrowleft}(\widetilde{D}, \boldsymbol{\ell}^\star, \boldsymbol{u}^\star)$ describes a (nonempty) cycle in $\widetilde{D}$.
- Projection σ is defined as:

$$\sigma : \mathbb{R}^{\widetilde{\mathcal{A}}} \to \mathbb{R}^{\mathcal{V}}, \quad \sigma(\boldsymbol{y})_v = y_{(v^{\mathrm{in}}, v^{\mathrm{out}})}$$

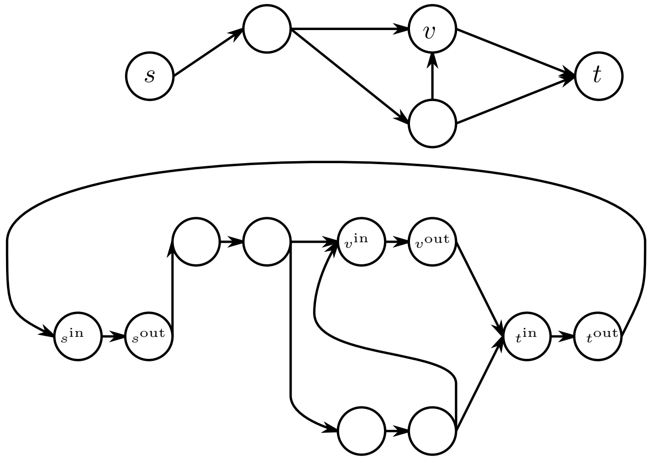

Figure 2.7: Example digraph $\widetilde{D}$ (bottom) obtained from acyclic digraph D (top).

- Map s is mapping any point in $\mathbb{R}_+$ to a circulation in $\mathbf{Q}_{\circlearrowleft}(\widetilde{D}, \boldsymbol{\ell}^{\boldsymbol{x}}, \boldsymbol{u}^{\boldsymbol{x}})$ (as long as it is not empty), which is defined by bounds

$$\left.\begin{aligned} \ell^{\boldsymbol{x}}_{(v^{\text{out}},w^{\text{in}})} &:= -\infty \\ u^{\boldsymbol{x}}_{(v^{\text{out}},w^{\text{in}})} &:= +\infty \end{aligned}\right\} \text{for all } (v,w) \in \mathcal{A}$$

$$\left.\begin{aligned} \ell^{\boldsymbol{x}}_{(v^{\text{in}},v^{\text{out}})} &:= x_v \\ u^{\boldsymbol{x}}_{(v^{\text{in}},v^{\text{out}})} &:= x_v \end{aligned}\right\} \text{for all } v \in \mathcal{V}$$

$$\begin{aligned} \ell_{(t^{\text{out}},s^{\text{in}})} &:= 1 \\ u_{(t^{\text{out}},s^{\text{in}})} &:= 1 \end{aligned}$$

It is easy to see that map s is a σ-section for those $\boldsymbol{x} \in \mathbb{R}^{\mathcal{V}}$ where it is defined, that $\sigma(\mathbf{Q}_{\circlearrowleft}(\tilde{D}, \boldsymbol{u}^{\star}, \boldsymbol{\ell}^{\star})) \subseteq \mathbf{P}^{s,t}(D)$ and that $\mathbf{Q}_{\circlearrowleft}(\tilde{D}, \boldsymbol{u}^{\boldsymbol{x}}, \boldsymbol{\ell}^{\boldsymbol{x}}) \subseteq \mathbf{Q}_{\circlearrowleft}(\tilde{D}, \boldsymbol{u}^{\star}, \boldsymbol{\ell}^{\star})$.

It remains to ensure by some set of inequalities, that for each $\boldsymbol{x} \in \mathbb{R}^{\mathcal{V}}$ satisfying these inequalities, polytope $\mathbf{Q}_{\circlearrowleft}(\widetilde{D}, \boldsymbol{\ell}^{\boldsymbol{x}}, \boldsymbol{u}^{\boldsymbol{x}})$ becomes nonempty, since in this case, the inequality system is $\mathbf{Q}_{\circlearrowleft}(\widetilde{D}, \boldsymbol{\ell}^{\star}, \boldsymbol{u}^{\star})$-enforcing for s.

Proposition 2.32 *Let $D = (\mathcal{V}, \mathcal{A})$ be an acyclic digraph with unique source s and sink t. The following system of inequalities provides a linear description for the s-t-path set polytope:*

$$\sum_{v \in \mathcal{T}} x_v - \sum_{v \in \operatorname{succ}_D(\mathcal{T})} x_v \leqslant 0 \qquad \forall \mathcal{T} \subset \mathcal{V} \setminus \{t\} \tag{*}$$

$$x_v \geqslant 0 \qquad \forall v \in \mathcal{V}$$

Proof. By theorem 2.19, $\mathbf{Q}_{\circlearrowleft}(\tilde{D}, \boldsymbol{u}^{\boldsymbol{x}}, \boldsymbol{\ell}^{\boldsymbol{x}})$ is nonempty if and only if

$$\sum_{a \in \delta^{\text{in}}_{\widetilde{D}}(\widetilde{\mathcal{S}})} \ell^{\boldsymbol{x}}_a \leqslant \sum_{a \in \delta^{\text{out}}_{\widetilde{D}}(\widetilde{\mathcal{S}})} u^{\boldsymbol{x}}_a \quad \forall \widetilde{\mathcal{S}} \subseteq \widetilde{\mathcal{V}}. \tag{**}$$

Let $\widetilde{\mathcal{S}}$ be an arbitrary subset of $\widetilde{\mathcal{V}}$. From $\widetilde{\mathcal{S}}$, we derive three subsets of $\mathcal{V}$:

$$\begin{aligned} \mathcal{S}^{\text{in}} &:= \{v \in \mathcal{V} \mid v^{\text{in}} \in \widetilde{\mathcal{S}} \text{ and } v^{\text{out}} \notin \widetilde{\mathcal{S}}\}, \\ \mathcal{S}^{\text{out}} &:= \{v \in \mathcal{V} \mid v^{\text{in}} \notin \widetilde{\mathcal{S}} \text{ and } v^{\text{out}} \in \widetilde{\mathcal{S}}\}, \text{ and} \\ \mathcal{S}^{\text{in out}} &:= \{v \in \mathcal{V} \mid v^{\text{in}} \in \widetilde{\mathcal{S}} \text{ and } v^{\text{out}} \in \widetilde{\mathcal{S}}\}. \end{aligned}$$

(i) Note that if $\delta^{\text{out}}_{\tilde{D}}(\widetilde{\mathcal{S}})$ or $\delta^{\text{in}}_{\tilde{D}}(\widetilde{\mathcal{S}})$ contain any real arc, then (**) is trivially satisfied. So, instead of (**), it suffices to ensure that

$$\alpha + \sum_{v\in\mathcal{S}^{\text{in}}} x_v \leqslant \sum_{v\in\mathcal{S}^{\text{out}}} x_v, \tag{***}$$

where α is set to 1 if $\delta^{\text{in}}_{\tilde{D}}(\widetilde{\mathcal{S}})$ contains arc $(t^{\text{out}}, s^{\text{in}})$; otherwise, we set $\alpha = 0$.

(ii) Assume now that $t \in \mathcal{S}^{\text{in}}$. We remove t^{out} from $\widetilde{\mathcal{S}}$. If $s^{\text{in}} \in \widetilde{\mathcal{S}}$, then this does not change the values of the left and right hand side in (***); if $s^{\text{in}} \notin \widetilde{\mathcal{S}}$, then we subtracted 1 from both sides.

(iii) $\alpha = 1$ implies that $s^{\text{in}} \in \widetilde{\mathcal{S}}$. By removing s^{in} from $\widetilde{\mathcal{S}}$, we either subtract 1 from both sides of inequality (***) (if $s^{\text{out}} \notin \widetilde{\mathcal{S}}$) or leave both sides as they are (if $s^{\text{out}} \in \widetilde{\mathcal{S}}$).

Hence, (***) is satisfied for $\widetilde{\mathcal{S}}$ if it is satisfied for a set $\widetilde{\mathcal{S}}'$ with $t \notin \mathcal{S}^{\text{in}}$ and $\alpha = 0$, which will be assumed in the following. Moreover, we can add $\sum_{v\in\mathcal{S}^{\text{in out}}} x_v$ to both sides of (***) obtaining

$$\sum_{v\in(\mathcal{S}^{\text{in}}\cup\mathcal{S}^{\text{in out}})} x_v \leqslant \sum_{v\in(\mathcal{S}^{\text{out}}\cup\mathcal{S}^{\text{in out}})} x_v, \tag{***'}$$

Last, we observe that because of (i), we can take it as given that

$$\text{succ}_D(\mathcal{S}^{\text{in out}} \cup \mathcal{S}^{\text{out}}) \subseteq \mathcal{S}^{\text{in out}} \cup \mathcal{S}^{\text{in}}. \tag{****}$$

Putting this together, we get

$$\sum_{v\in(\mathcal{S}^{\text{in}}\cup\mathcal{S}^{\text{in out}})} x_v \overset{(*)}{\leqslant} \sum_{v\in\text{succ}_D(\mathcal{S}^{\text{in}}\cup\mathcal{S}^{\text{in out}})} x_v \overset{(****)}{\leqslant} \sum_{v\in\mathcal{S}^{\text{in}}\cup\mathcal{S}^{\text{in out}}} x_v,$$

by using $\mathcal{T} = \mathcal{S}^{\text{in}} \cup \mathcal{S}^{\text{in out}}$ in (*) and exploiting that $x_v \geqslant 0$ for all $v \in \mathcal{V}$ to establish (***').

Hence, inequality system (***) together with $\boldsymbol{x} \geqslant \mathbb{0}$ is $\mathbf{Q}_{\circlearrowleft}(\tilde{D}, \boldsymbol{u}^\star, \boldsymbol{\ell}^\star)$-enforcing for s. As $\sigma(\mathbf{Q}_{\circlearrowleft}(\tilde{D}, \boldsymbol{u}^\star, \boldsymbol{\ell}^\star)) \subseteq \mathbf{P}^{s,t}(D)$ and the inequality system is valid for $\mathbf{P}^{s,t}(D)$, we obtain the statement.

□

2.4.1.3 The Matching Polytope of an Arbitrary Graph

As said at the beginning of chapter 1, the *perfect matching polytope* of a general graph is given by inequalities (1.1). Schrijver ([107]) derives a linear description of the *matching polytope* for general graphs by a reduction (as he calls it) from the linear description of the perfect matching polytope. In essence, he is using a faithful sectioning. We will rephrase his proof in terms of faithful sectionings.

Proposition 2.33 *The matching polytope $\mathbf{P}^M(G)$ of a general graph $G = (\mathcal{V}, \mathcal{E})$ is given by the following set of inequalities:*

$$x_e \geqslant 0 \qquad \forall\, e \in \mathcal{E} \tag{2.9}$$

$$\sum_{e\in\delta_G(v)} x_e \leqslant 1 \qquad \forall\, v \in \mathcal{V} \tag{2.10}$$

$$\sum_{e\in\mathcal{E}[\mathcal{U}]} x_e \leqslant \left\lfloor \frac{1}{2}|\mathcal{U}| \right\rfloor \qquad \forall\, \mathcal{U} \subseteq \mathcal{V} \text{ odd}, \tag{2.11}$$

Proof. Define graph $\widetilde{G} := G \times K_2$. We denote by v' the copy of node $v \in \mathcal{V}$ and by e' the copy of $e \in \mathcal{E}$. Moreover, we write $\mathcal{V}' := \{v' \mid v \in \mathcal{V}\}$ and $\mathcal{E}' := \{e' \mid e \in \mathcal{E}\}$. Hence, graph $\widetilde{G} = (\mathcal{V} \cup \mathcal{V}', \widetilde{\mathcal{E}})$ consists of two identical subgraphs $\mathcal{G}$ and $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$ and a set of edges $\{\{v, v'\} \mid v \in \mathcal{V}\}$ connecting both.

The crucial observation is that if $\widetilde{\mathcal{M}}$ is a perfect matching in graph $\widetilde{G}$, then $\widetilde{\mathcal{M}} \cap \mathcal{E}$ is a matching in graph G, and, on the other hand, supposed we found a matching $\mathcal{M}$ in G and define sets $\mathcal{M}' := \{e' \in \mathcal{E}' \mid e \in \mathcal{M}\}$ and $\overline{\mathcal{M}} := \Big\{\{v, v'\} \,\Big|\, v \notin \mathcal{V}[\mathcal{M}]\Big\}$, then

$$\widetilde{\mathcal{M}} := \mathcal{M} \,\cup\, \mathcal{M}' \,\cup\, \overline{\mathcal{M}}$$

is a perfect matching in $\widetilde{G}$. (See figure 2.8 left for an example.)

First, we construct an extended formulation for $\mathbf{P}^M(G)$ from the perfect matching polytope $\mathbf{P}^{PM}(\widetilde{G})$ and the orthogonal projection σ which maps any point $\boldsymbol{y} \in \mathbb{R}^{\widetilde{\mathcal{E}}}$ to $\boldsymbol{x} = \boldsymbol{y}_{\mathcal{E}}$. Therefore, σ maps in particular the incidence vector $\boldsymbol{y}$ of a perfect matching in $\widetilde{G}$ to the incidence vector $\boldsymbol{x}$ of a matching in G.

Next, we define a map $s : \mathbb{R}^{\mathcal{E}} \to \mathbb{R}^{\widetilde{\mathcal{E}}}$ as follows:

$$\begin{aligned}
y_e &:= x_e && \forall e \in \mathcal{E} \\
y_{e'} &:= x_e && \forall e' \in \mathcal{E}' \\
y_{\{v,v'\}} &:= 1 - \sum_{e \in \delta_G(v)} x_e && \forall \{v, v'\} \in \overline{\mathcal{E}}
\end{aligned}$$

We observe:

- The polytope $\mathbf{P}$ described by inequalities (2.9) through (2.11) contains $\mathbf{P}^M(G)$.
- $\sigma(\mathbf{P}^{PM}(\widetilde{G})) \subseteq \mathbf{P}^M(G)$.
- $\sigma(\boldsymbol{y}) = \boldsymbol{x}$ for all $\boldsymbol{y} = s(\boldsymbol{x})$; hence, s is a σ-section.

Therefore, by Theorem 2.17, it remains to show that inequalities (2.9) through (2.11) are $\mathbf{P}^{PM}(\widetilde{G})$-enforcing for s. In other words: we have to show that inequalities (1.1) are valid for any point $s(\boldsymbol{x})$ if inequalities (2.9) through (2.11) are valid for $\boldsymbol{x}$.

- Equations $\sum_{e \in \delta_{\widetilde{G}}(v)} y_e = 1$ follow from the definition of map s.
- Inequalities $y_e \geqslant 0$ for all $e \in \mathcal{E}$ and $y_{e'} \geqslant 0$ for all $e' \in \mathcal{E}'$ follow readily from inequalities (2.9).
- Inequalities $y_{\{v,v'\}} \geqslant 0$ for all $v \in \mathcal{V}$ follow from inequalities (2.10).
- Finally, we have to ensure inequalities $\sum_{e \in \delta_{\widetilde{G}}(\mathcal{U})} y_e \geqslant 1$ for all odd sets $\mathcal{U} \subseteq \widetilde{\mathcal{V}}$. Let $\mathcal{X} := \mathcal{U} \cap \mathcal{V}$ and $\mathcal{Y}' := \mathcal{U} \cap \mathcal{V}'$ and let $\mathcal{U}^{\varnothing} \subseteq \widetilde{\mathcal{V}}$ arise from $\mathcal{U}$ by deleting all nodes in $\mathcal{X} \cap \mathcal{Y}$ and $\mathcal{X}' \cap \mathcal{Y}'$ from $\mathcal{U}$. The symmetric difference $\delta_{\widetilde{G}}(\mathcal{U}^{\varnothing}) \,\triangle\, \delta_{\widetilde{G}}(\mathcal{U})$ can be partitioned into six subsets.

set...	contains all edges in $\widetilde{\mathcal{E}}[\mathcal{U}]$ with...
$\mathcal{A}_1$	...one node in $\mathcal{X} \setminus (\mathcal{X} \cap \mathcal{Y})$ and one node in $\mathcal{X} \cap \mathcal{Y}$
$\mathcal{A}_1'$	...one node in $\mathcal{X}' \setminus \mathcal{Y}'$ and one node in $\mathcal{X}' \cap \mathcal{Y}'$
$\mathcal{A}_2$	...one node in $\mathcal{Y}' \setminus (\mathcal{X}' \cap \mathcal{Y}')$ and one node in $\mathcal{X}' \cap \mathcal{Y}'$
$\mathcal{A}_2'$	...one node in $\mathcal{Y} \setminus \mathcal{X}$ and one node in $\mathcal{X} \cap \mathcal{Y}$
$\mathcal{A}_3$	...one node in $\mathcal{X} \cap \mathcal{Y}$ and one node in $\mathcal{V} \setminus (\mathcal{X} \cup \mathcal{Y}))$
$\mathcal{A}_3'$	...one node in $\mathcal{X}' \cap \mathcal{Y}'$ and one node in $\mathcal{V}' \setminus (\mathcal{X}' \cup \mathcal{Y}'))$

(See figure 2.8 right.) Because of the symmetry of $\widetilde{G}$, there is a one-to-one-correspondence between edge set $\mathcal{A}_i$ and edge set $\mathcal{A}_i'$ for any $i \in [3]$. Moreover, edges in $\mathcal{A}_1$ and $\mathcal{A}_2'$ are in $\delta_{\widetilde{G}}(\mathcal{U}^{\varnothing})$ but not in $\delta_{\widetilde{G}}(\mathcal{U})$, and, vice versa, edges

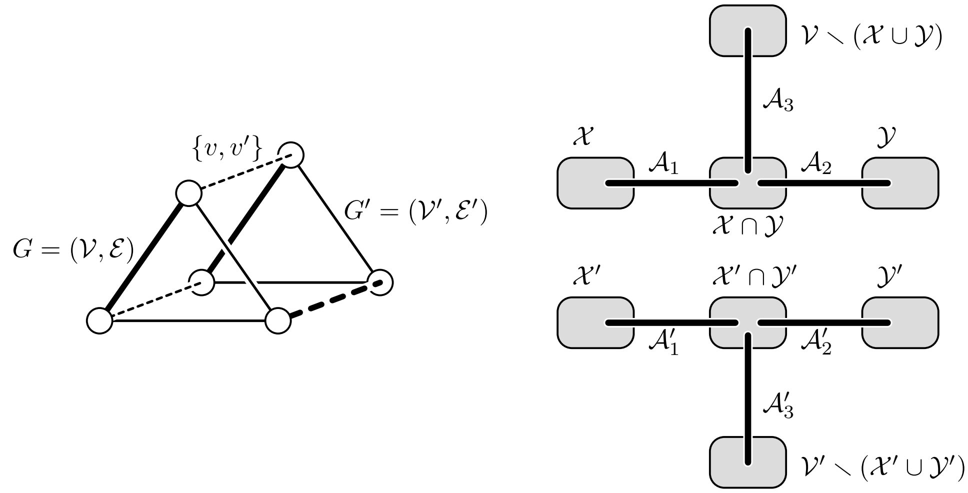

Figure 2.8: Left: example of graph $\widetilde{G} \simeq G \times K_2$. Edges $\{v, v'\}$ are dashed, and bold edges make a perfect matching in $\widetilde{G}$ obtained from a matching in G. Right: partitioning $\delta_{\widetilde{G}}(\mathcal{U}^{\varnothing})\Delta\delta_{\widetilde{G}}(\mathcal{U})$.

in $\mathcal{A}_1'$ and $\mathcal{A}_2$ are in $\delta_{\widetilde{G}}(\mathcal{U})$ but not in $\delta_{\widetilde{G}}(\mathcal{U}^{\varnothing})$. The edges in $\mathcal{A}_3 \cup \mathcal{A}_3'$ are in $\delta_{\widetilde{G}}(\mathcal{U})$ but not in $\delta_{\widetilde{G}}(\mathcal{U}^{\varnothing})$. As $y_e \geqslant 0$ for all $e \in \widetilde{\mathcal{E}}$, we obtain that

$$\sum_{e \in \delta_{\widetilde{G}}(\mathcal{U})} y_e \geqslant \sum_{e \in \delta_{\widetilde{G}}(\mathcal{U}^{\varnothing})} y_e$$

Hence, we can assume that $\mathcal{X} \cap \mathcal{Y} = \varnothing$. So, either $\mathcal{X}$ or $\mathcal{Y}$ must be odd. W.l.o.g., we can assume that $\mathcal{X}$ is odd; again, w.l.o.g, we can therefore assume that $\mathcal{Y} = \varnothing$.
Now,

$$\sum_{e \in \delta_{\widetilde{G}}(\mathcal{X})} y_e + 2 \sum_{e \in \mathcal{E}[\mathcal{X}]} y_e = \sum_{v \in \mathcal{X}} \delta_{\widetilde{G}}(v) = |\mathcal{X}|$$

because of equations $\sum_{e \in \delta_{\widetilde{G}}(v)} y_e = 1$. Rearranging gives

$$\sum_{e \in \delta_{\widetilde{G}}(\mathcal{X})} y_e = |\mathcal{X}| - 2 \sum_{e \in \mathcal{E}[\mathcal{X}]} y_e \overset{(*)}{\geqslant} |\mathcal{X}| - 2\left\lfloor \frac{1}{2}|\mathcal{X}| \right\rfloor = 1$$

where (*) holds because of inequalities (2.11).
So, in fact, inequalities (2.9) through (2.11) are $\mathbf{P}^{PM}(\widetilde{G})$-enforcing for s and we obtain the result by Theorem 2.17. □

Remark 2.34 Schrijver uses faithful sectionings also for other linear descriptions.

- He derives with faithful sectioning the linear description of several classes of polytopes related to the $\boldsymbol{b}$-matching polytope, in particular the $\boldsymbol{b}$-matching polytope (theorem 31.2 in [107]) and the $\boldsymbol{c}$-capacitated $\boldsymbol{b}$-matching polytope (theorem 32.2 in [107]). In the first case, the extended formulation is based on the matching polytope, in the second case, it is based on the $\boldsymbol{b}$-matching polytope.
- He also shows that the linear description of the $\boldsymbol{b}$-edge cover polyhedron can be obtained by faithful sectioning, using the description of the edge cover polytope (theorem 34.2 in [107]).

Chapter 3

Mapping the Terrain

3.1 Definition of Orbitopes

In 2006, Kaibel and Pfetsch introduced orbitopes[1] as the convex hull of all 0/1-matrices with lexicographically ordered columns. We will now specify the different classes of orbitopes in detail, largely following [65].

Throughout the following, we denote by $\mathcal{M}_{p,q}$ the set of 0/1-matrices with p rows and q columns and by $\boldsymbol{M}$ an element of $\mathcal{M}_{p,q}$. We let some subgroup G of the symmetric group $\mathfrak{S}_q$ on q elements act on $\mathcal{M}_{p,q}$:

$$G \times \mathcal{M}_{p,q} \to \mathcal{M}_{p,q}, \qquad\qquad (g, \boldsymbol{M}) \mapsto g(\boldsymbol{M}),$$

such that any group element $g \in G$ permutes the columns of matrix $\boldsymbol{M}$. So, G generates an *orbit* from each matrix $\boldsymbol{M}$:

$$G(\boldsymbol{M}) := \{g(\boldsymbol{M}) \mid g \in G\} \subseteq \mathcal{M}_{p,q}$$

and these orbits partition $\mathcal{M}_{p,q}$.

To pick out some unique representative from each orbit, we define an order on the elements in $\mathcal{M}_{p,q}$ by comparing their entries columnwise (see figure 3.1.); we will use the lexicographic ordering (see page 5) and compare the entries of vertices columnwise.

More precisely, we define an order on the entries by means of the following function τ mapping each element in $[pq]$ to a tupel $(i, j) \in [p] \times [q]$:

$$\tau := \begin{cases} [pq] & \to [p] \times [q] \\ \ell & \mapsto \big(((\ell-1) \bmod p) + 1, \lfloor (\ell-1)/p \rfloor + 1\big) \end{cases}$$

For two elements $\boldsymbol{M}, \boldsymbol{N} \in \mathcal{M}_{p,q}$ with entries $m_{i,j}$ and $n_{i,j}$, respectively, we say that $\boldsymbol{M}$ is *lexicographically larger* than $\boldsymbol{N}$ (or, equivalently, $\boldsymbol{M}$ and $\boldsymbol{N}$ are in *lexicographic order*, in short $\boldsymbol{M} \succ \boldsymbol{N}$), if there is some $\ell \in [pq]$ such that $m_{\tau(\ell)} > n_{\tau(\ell)}$ and $m_{\tau(i)} = n_{\tau(i)}$ for all $i \in [pq]$ with $i \in [\ell - 1]$.

Definition 3.1 We define $\mathcal{M}^{\max}_{p,q}(G)$ as the set of those matrices in $\mathcal{M}_{p,q}$ that are lexicographically maximal within their orbits under action of G.

Depending on the number $k \in [q]_0$ of 1-entries in the rows of the matrices in $\mathcal{M}_{p,q}$, we define now the following subsets of $\mathcal{M}_{p,q}$:

[1] Not to be confused with orbitopes in the sense of Sanyal et al., see [105]. They understand orbitopes as the convex hull of an orbit of a compact group acting linearly on a vector space.

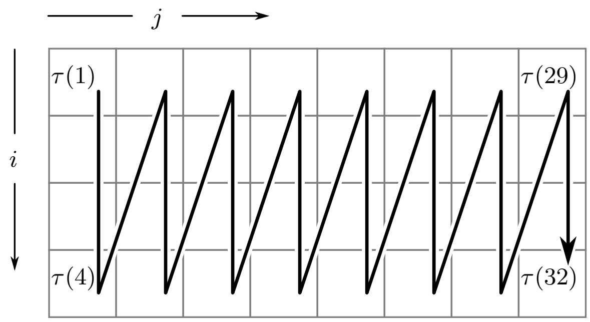

Figure 3.1: When comparing two matrices, the significance of the entries decreases columnwise. Top left entry with indices $(1,1)$ is the most significant entry. In this example, $p = 4$ and $q = 8$.

$$\mathcal{M}_{p,q}^{\leqslant k} := \{\boldsymbol{M} \in \mathcal{M}_{p,q} \mid \sum_{j\in[q]} m_{i,j} \leqslant k \quad \forall i \in [p]\}$$
$$\mathcal{M}_{p,q}^{= k} := \{\boldsymbol{M} \in \mathcal{M}_{p,q} \mid \sum_{j\in[q]} m_{i,j} = k \quad \forall i \in [p]\}$$
$$\mathcal{M}_{p,q}^{\geqslant k} := \{\boldsymbol{M} \in \mathcal{M}_{p,q} \mid \sum_{j\in[q]} m_{i,j} \geqslant k \quad \forall i \in [p]\}$$

Generalizing this concept, we let $\boldsymbol{k} \in \mathbb{N}^p$ be some list of integers. Entry k_i determines the number of 1-entries in row i. We define

$$\mathcal{M}_{p,q}^{=\boldsymbol{k}} := \{\boldsymbol{M} \in \mathcal{M}_{p,q} \mid \sum_{j\in[q]} m_{i,j} = k_i \quad \forall i \in [p]\}$$

Definition 3.2 For a given group $G \subseteq \mathfrak{S}_q$, the *full orbitope* is

$$\mathbf{O}_{p,q}(G) := \operatorname{conv} \mathcal{M}_{p,q}^{\max}(G),$$

the *k-packing orbitope* is

$$\mathrm{O}_{p,q}^{\leqslant k}(G) := \operatorname{conv}(\mathcal{M}_{p,q}^{\max}(G) \cap \mathcal{M}_{p,q}^{\leqslant k}),$$

the *k-partitioning orbitope* is

$$\mathrm{O}_{p,q}^{= k}(G) := \operatorname{conv}(\mathcal{M}_{p,q}^{\max}(G) \cap \mathcal{M}_{p,q}^{= k}),$$

and the *k-covering orbitope* is

$$\mathrm{O}_{p,q}^{\geqslant k}(G) := \operatorname{conv}(\mathcal{M}_{p,q}^{\max}(G) \cap \mathcal{M}_{p,q}^{\geqslant k}).$$

For an arbitrary list $\boldsymbol{k} \in \mathbb{N}^p$, we call

$$\mathrm{O}_{p,q}^{=\boldsymbol{k}}(G) := \operatorname{conv}(\mathcal{M}_{p,q}^{\max}(G) \cap \mathcal{M}_{p,q}^{=\boldsymbol{k}})$$

the *fixed row sum orbitope* over group G.

Note that we drop the group G if the context is clear.

Furthermore, we always drop k if $k = 1$ and speak in this case of *packing-* ($\mathrm{O}_{p,q}^{\leqslant}(G)$), *partitioning-* ($\mathrm{O}_{p,q}^{=}(G)$), and *covering-orbitopes* ($\mathrm{O}_{p,q}^{\geqslant}(G)$) over group G.

A majority of the following work concentrates on orbitopes over the full symmetric group and $q = 2$ columns. Only to a limited extend, we will study orbitopes over the cyclic and other groups, as well as orbitopes with more than two columns.

DEFINITION 3.3 A full orbitope with $q = 2$ columns and p rows over the full symmetric group $\mathfrak{S}_2$ is called an *orbisack*. We shorty write $\mathbf{O}_{p,2}$.

Clearly, in the case of orbisacks, the group G could just as well be the cyclic group. For general orbitopes however, the choice of group G definitely matters.

Remark 3.4 We remark that there is a relationship between orbitopes and *revlex-initial polytopes*. A set $\mathcal{X} \subseteq \{0,1\}^p$ is called revlex-initial, if $\mathcal{X}$ contains with any $x \in \mathcal{X}$ also all points in $\{0,1\}^p$ that are lexicographic smaller than $\boldsymbol{x}$. If set $\mathcal{X}$ is revlex-initial, then the convex hull of $\mathcal{X}$ is called a revlex-initial polytope $\mathbf{P}(\mathcal{X})$. Gillman and Kaibel gave a complete facial description of revlex-initial orbitopes, see [50].
Choose now an arbitrary vector $\boldsymbol{c} \in \{0,1\}^{[p]}$. From $\boldsymbol{c}$, we derive a vector $\tilde{\boldsymbol{c}} \in \mathbb{R}^{[p]\times[2]}$ by setting

$$\begin{aligned} \tilde{c}_{i,1} &= 1, && \text{if } c_i = 1, \\ \tilde{c}_{i,1} &= -1, && \text{if } c_i = 0, \text{ and} \\ \tilde{c}_{i,2} &= 0 && \text{for all } i \in [p]. \end{aligned}$$

Furthermore, we define $b := \langle \boldsymbol{c}_{*,1}, \tilde{\boldsymbol{c}} \rangle$. It is easy to see that the inequality $\tilde{\boldsymbol{c}}\boldsymbol{x} \leqslant b$ is valid for the orbisack. Moreover, the set $\{\boldsymbol{x} \in \{0,1\}^{[p]\times[2]} \mid \langle \tilde{\boldsymbol{c}}, \boldsymbol{x} \rangle = b\}$ contains all 0/1-vectors that have first column $\boldsymbol{c}$. Hence,

$$\mathbf{O}_{p,2} \cap \{\boldsymbol{x} \in \mathbb{R}^{[p]\times[2]} \mid \langle \tilde{\boldsymbol{c}}, \boldsymbol{x} \rangle \leqslant b\}$$

is a face of the orbisack. All faces that arise from fixing the first column in the vertices of an orbisack to a certain 0/1-pattern are reflex initial polytopes.

We close this section with a small remark concerning our definition of lexicographic order.

Remark 3.5 The lexicographic order defined above differs from the one used in [65]. This does not become apparent as long as one sticks to packing and partitioning orbitopes as Kaibel and Pfetsch did. However, with arbitrary groups (e.g. the cyclic group) and more than one 1 per row, $\mathcal{M}^{\max}_{p,q}(G)$ starts to look different. For example, using the ordering from [65], $\mathcal{M}^{\max}_{2,3}(\mathfrak{C}_3)$ contains the following matrix:

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

which is apparently not a representative according to our notion of lexicographic ordering. We chose our ordering, because it seems to fit better with column permutations.

3.2 What can be done and what cannot?

Immediately, several questions pop up. What do the vertices of the orbitopes look like, depending on the choice of group G and on the restriction on the number of 1s in each row? What is the complexity status for optimization over the different orbitopes and what do optimization algorithms look like? As outlined in the introduction, particularly the complexity status is interesting, as it indicates whether there are chances to find "nice" complete linear descriptions for the actual orbitopes (see page 3).

The results are quite encouraging for orbitopes over the full symmetric group (section 3.2.1) or over products of symmetric groups (section 3.2.5). Figure 3.2 gives a short overview over the known facts concerning orbitopes over the full symmetric group. The details will be presented in the following.

	complexity of optimization		complete lin. descr.
covering	$\mathcal{NP}$-hard	p. 61	—
packing	polynomial time	[65]	✓
partitioning	polynomial time	[65]	✓
k-packing	$\mathcal{NP}$-hard even for $k=2$	p. 61	—
k-partitioning	$\mathcal{NP}$-hard even for $k=2$	p. 61	—
fixed row sum	$\mathcal{NP}$-hard even for $k=2$	p. 61	—
full, $q>2$	polynomial time	Theorem 3.33	open
full, $q=2$	polynomial time	Theorem 3.33	✓

Figure 3.2: Complexity of optimization and the state of knowledge on linear descriptions for full orbitopes over the full symmetric group.

However, if the cyclic group is operating on the columns, things seem to become much more complicated. Apart from the results of Kaibel et al. ([65]) concerning partitioning and packing orbitopes over cyclic groups, a lot of questions are open. It is unclear how to optimize over full orbitopes when the columns can be cyclically permuted, and there is even no elementary description of the vertices of these orbitopes known.

3.2.1 Full Symmetric Group

Let $\boldsymbol{M} \in \mathcal{M}_{p,q}$. For the remainder of this section, the full symmetric group $\mathfrak{S}_q$ is operating on the columns of $\boldsymbol{M}$. Hence, each orbit $G(\boldsymbol{M})$ contains by definition all matrices arising from arbitrary permutations of the columns of $\boldsymbol{M}$. The representative of $G(\boldsymbol{M})$ is therefore the matrix with all columns in non-increasing order. In other words: in any vertex $\boldsymbol{x}$ of an orbitope — no matter which restrictions on the number of 1s per row exist —, any two columns $\boldsymbol{x}_{*,j}$ and $\boldsymbol{x}_{*,k}$, $j<k$, are either identical or there is some row index $i^\star \in [p]$ such that $x_{i,j} = x_{i,k}$ for all $i \in [i^\star - 1]$, and $x_{i^\star,j} = 1$ and $x_{i^\star,k} = 0$. This follows from the definition.

Example 3.6 The following is an example of two lexicographic descending ordered vectors with length $p = 6$. As there are up to two 1-entries per row, this could be the vertex of a packing orbitope $\mathrm{O}^{\leqslant 2}_{p,2} = \mathbf{O}_{p,2}$.

0	0
1	1
1	0
0	1
1	1
1	0

Roughly speaking, the gray row in this example "decides" that the columns of the 0/1-vector are in lexicographic descending order. This row will become particularly important when we are studying orbisacks. Therefore, we define for any vector in $\mathcal{M}_{p,2}$:

Definition 3.7 (Critical row) Let $\boldsymbol{v} \in \mathcal{M}_{p,q}$. We define a set $\mathcal{I}(\boldsymbol{v}) \subseteq [p]$ as follows:

$$i \in \mathcal{I}(\boldsymbol{v}) \;\Leftrightarrow\; \boldsymbol{v}_{i,*} = (1,0).$$

We call

$$\operatorname{crit}(\boldsymbol{v}) := \begin{cases} \min(\mathcal{I}(\boldsymbol{v}) \cup \{p+1\}), & \text{if } \boldsymbol{v} \in \mathcal{M}^{\max}_{p,2} \\ \text{undefined}, & \text{otherwise.} \end{cases} \quad \in [p+1]$$

the *critical row*. Note that if the two columns of $\boldsymbol{v}$ are equal, then the critical row is $p+1$, and if $\boldsymbol{v}_{*,1} \prec \boldsymbol{v}_{*,2}$, then $\operatorname{crit}(\boldsymbol{v})$ is undefined.

Using this concept, we define for an arbitrary number of columns:

Definition 3.8 (Split) Let $\boldsymbol{v} \in \mathcal{M}_{p,q}$. The *split pattern* $\boldsymbol{\sigma} \in [p+1]^{q-1}$ is a vector containing as entry σ_j, $j \in [q-1]$, the critical row of the submatrix $\boldsymbol{v}_{*,[j..j+1]}$ (that is the critical row of neighbouring columns $\boldsymbol{v}_{*,j}$ and $\boldsymbol{v}_{*,j+1}$). If $\sigma_j = i$, we call σ_j a *split in row* i.

Next, we make some small observations concerning split patterns of vectors in $\mathcal{M}^{\max}_{p,q}$.

Observation 3.9
- For any $j \in [q-2]$, it holds that if $\sigma_j = i$, then $\sigma_{j+1} \neq i$.
- There are at most $\min(q-1, 2^{i-1})$ splits in row i.
- There are at most $\mathcal{O}(p^{q-1})$ split patterns possible, since we can choose for each pair of neighbouring column $p+1$ possible splits.

Last, we will define a map that will be useful in several situations.

Definition 3.10 (Flipping a vector) Let $\boldsymbol{v} \in \mathcal{M}_{p,q}$ and define the following two affine transformations f to invert the order of columns and g to flip 1s and 0s in $\boldsymbol{v}$, that is, $f : \mathcal{M}_{p,q} \to \mathcal{M}_{p,q}$ with $f(\boldsymbol{x}) = \boldsymbol{y}$ defined by $y_{i,j} = x_{i,q-j+1}$ for all $(i,j) \in [p] \times [q]$, and $g : \mathcal{M}_{p,q} \to \mathcal{M}_{p,q}$ with $g(\boldsymbol{x}) = \mathbb{1} - \boldsymbol{x}$.

Observation 3.11 $f \circ g$ is an affine transformation with a nice property: $f \circ g(\boldsymbol{v})$ is in $\mathcal{M}^{\max}_{p,q}$ if and only if $\boldsymbol{v} \in \mathcal{M}^{\max}_{p,q}$, i.e. if the columns of $\boldsymbol{v}$ are in non-increasing lexicographic order. Proof: $\Rightarrow$ Transformation f inverts the order of the columns, and it is easy to see that any pair of neighbouring columns in $g(\boldsymbol{v})$ is ordered in lexicographic non-decreasing order if they were in lexicographic non-increasing order. Hence, if $\boldsymbol{v}$ is in lexicographic order, $f \circ g(\boldsymbol{v})$ is, too. $\Leftarrow$ $f \circ g$ is self-inverse.

This has the following implications:

Observation 3.12
- The partitioning orbitope $\mathrm{O}^{=k}_{p,q}$ is isomorphic to the partitioning orbitope $\mathrm{O}^{=q-k}_{p,q}$, since because of the fact that any vertex $\boldsymbol{v}$ of $\mathrm{O}^{=k}_{p,q}$ has $q-k$ 0-entries in each row, any vertex $f \circ g(\boldsymbol{v})$ has $q-k$ 1-entries in each row.
- Similarly, $\mathrm{O}^{\leqslant k}_{p,q}$ is isomorphic to $\mathrm{O}^{\geqslant q-k}_{p,q}$.

After these general observations and definitions, we will now consider special orbitopes in detail.

3.2.1.1 General Properties and Characterization of Vertices

3.2.1.1.1 Full Orbitopes

The vertices of the full orbitope $\mathbf{O}_{p,q}$ can have an arbitrary number of 1-entries per row. From the definition of splits, it follows that vector $\boldsymbol{v}$ is a vertex of $\mathbf{O}_{p,q}$ if and only if σ_j is defined for all $j \in [q-1]$. Moreover, there is at most one split in the first row (Observation 3.9).

Hence, we can characterize the vertices of the full orbitope as follows.

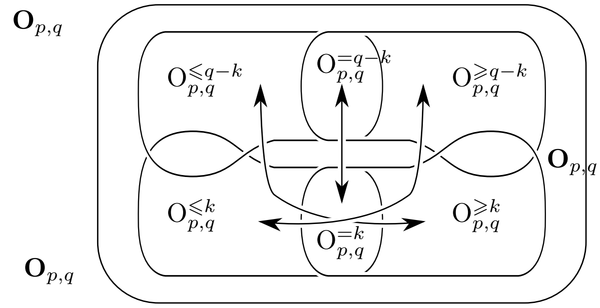

Figure 3.3: The vertex set of a full orbitope $\mathbf{O}_{p,q}$ contains in particular the vertex sets of $\mathrm{O}_{p,q}^{\leqslant k}$, $\mathrm{O}_{p,q}^{=k}$ and $\mathrm{O}_{p,q}^{\geqslant k}$ for any $k \in [q]_0$. Packing and covering orbitopes are isomorphic to each other by transformation $f \circ g$. (See definition 3.10 and observations 3.11 and 3.12.)

Lemma 3.13

(i) The first row of any vertex $\boldsymbol{x}$ of the full orbitope $\mathbf{O}_{p,q}$ over the full symmetric group consists of k 1-entries, with $k \in [q]_0$, followed by $q-k$ 0-entries. In particular, $(x_{1,j}, x_{1,j+1}) = (0,1)$ is forbidden for all $1 \leqslant j < q$.

(ii) For $1 \leqslant k \leqslant q-1$, both subvectors $\boldsymbol{x}_{[2..p]\times[k]}$ and $\boldsymbol{x}_{[2..p]\times[k+1..q]}$ have lexicographically ordered columns and are therefore vertices of $\mathbf{O}_{p-1,k}(\mathfrak{S}_k)$ and $\mathbf{O}_{p-1,q-k}(\mathfrak{S}_{q-k})$, respectively.

Proof. (i) follows from Observation 3.9. For (ii), see the example below. We consider first the left submatrix $\boldsymbol{x}_{[2..p]\times[k]}$. The statement is trivial for $k=0$ or $k=1$. So assume that $k>1$. Any column in $\boldsymbol{x}_{*,[k]}$ has first entry 1. Consider now the pattern of splits $\boldsymbol{\sigma} \in [i]^{[p-1]}$ for $\boldsymbol{x}$. By definition of splits, it holds that for any $j \in [k-1]$, the value $\sigma_j \in [2..p]$, since $\boldsymbol{v} \in \mathbf{O}_{p,q}$. This implies that the columns of $\boldsymbol{x}_{[2..p]\times[k]}$ are lexicographically ordered, and since there are no restrictions on the number of 1-entries per row, we obtain that $\boldsymbol{x}_{[2..p]\times[k]} \in \mathbf{O}_{p-1,k}$. The argumentation is analogous for columns in $\boldsymbol{x}_{[2..p]\times[k+1..q]}$. ☐

We will take advantage of this lemma when later formulating an algorithm for linear optimization over full orbitopes.

Example 3.14 This is an example of a vertex of the full orbitope $\mathbf{O}_{4,8}$. Note that both the left and the right gray submatrix are vertices of an orbitope of smaller dimension.

1	1	1	0	0	0	0	0
1	1	0	1	1	1	1	0
1	0	1	1	1	0	0	0
1	1	1	1	1	0	0	0

Observation 3.15 In any vertex $\boldsymbol{v}$ of $\mathbf{O}_{p,q}$, one can read the columns of $\boldsymbol{v}$ as the binary expansion of numbers in $[2^p-1]_0$. Ordering the columns lexicographically is equivalent to ordering them according to the value of their decimal representations.

This observation leads to the following characterization of the vertices of orbisacks $\mathbf{O}_{p,2}$: a vector $\boldsymbol{x} \in \{0,1\}^{[p]\times[2]}$ is a vertex of the orbisack if and only if it satisfies the following inequality:

$$\sum_{i=1}^{p} 2^{p-i}(x_{i,1} - x_{i,2}) \geqslant 0 \tag{3.1}$$

Lemma 3.16 *Orbisacks are knapsack polytopes.*

Proof. The knapsack inequality is (3.1). □

Define now the following affine transformation $\varphi : \mathbb{R}^{[p]\times[2]} \to \mathbb{R}^{[p]\times[2]}$, where $\varphi(\boldsymbol{x}) = \boldsymbol{y}$ is defined by

$$\begin{aligned} y_{i,1} &= 1 - x_{p-i+1,1} \\ y_{i,2} &= x_{p-i+1,2} \end{aligned} \qquad \forall i \in [p].$$

Since φ is one-to-one, $\boldsymbol{x} \in \{0,1\}^{[p]\times[2]}$ satisfies inequality (3.1) if and only if $\varphi(\boldsymbol{x})$ satisfies inequality

$$\sum_{i=1}^{p} 2^{i-1}(y_{i,1} + y_{i,2}) \leqslant 2^p - 1. \tag{3.2}$$

Therefore, the orbisack is in particular isomorphic to a special 0/1-knapsack in $\mathbb{R}^{[p]\times[2]}$, called the *sequential knapsack polytope*. This will be explained in detail (and exploited) in section 4.2.

Apart from the orbisacks, there are more knapsack polytopes among the full orbitopes: the orbitope $\mathbf{O}_{1,q}$ is a knapsack polytope for any $q \in \mathbb{N}_{>}$. This follows from the fact that $\mathbf{O}_{1,q}$ is isomorphic to a q-dimensional simplex $\widetilde{\triangle}_q$, which is a knapsack polytope (see page 14).

Lemma 3.17 *Orbitope $\mathbf{O}_{1,q}$ is isomorphic to the (non-standard) q-dimensional simplex $\widetilde{\triangle}_q$.*

Proof. The vertices of $\mathbf{O}_{1,q}$ make a set of $q+1$ affinely independent points. □

In fact, orbisacks and $\mathbf{O}_{p,1}$ are the only two classes of orbitopes over the full symmetric group that contain knapsack polytopes.

Proposition 3.18 *Any full orbitope over the full symmetric group with $q > 2$ columns and $p > 1$ rows is not a knapsack.*

Proof. Assume, there is a vector $\boldsymbol{c} \in \mathbb{R}^{[p]\times[q]}$ and a right-hand side $b \in \mathbb{R}$ defining a hyperplane $\langle \boldsymbol{c}, \boldsymbol{x}\rangle \leqslant b$ separating the vertices of the orbitope $\mathbf{O}_{p,q}$ from the remaining vertices of $\mathcal{M}_{p,q}$, i.e.

$$\langle \boldsymbol{c}, \boldsymbol{x}\rangle \leqslant b \quad \forall \boldsymbol{x} \in \mathcal{M}_{p,q} \cap \mathbf{O}_{p,q}$$

and

$$\langle \boldsymbol{c}, \boldsymbol{x}\rangle > b \quad \forall \boldsymbol{x} \in \mathcal{M}_{p,q} \setminus \mathbf{O}_{p,q}.$$

However, because of the lexicographic ordering, $\mathbb{e}^{1,1} + \mathbb{e}^{1,2}$ and $\mathbb{e}^{2,1} + \mathbb{e}^{2,2} + \mathbb{e}^{2,3}$ are vertices of $\mathbf{O}_{p,q}$, but $\mathbb{e}^{2,1} + \mathbb{e}^{2,3}$ and $\mathbb{e}^{1,1} + \mathbb{e}^{1,2} + \mathbb{e}^{2,2}$ are not (see figure 3.4). Therefore, the following inequalities must hold:

$$c_{1,1} + c_{1,2} \leqslant b \tag{3.3a}$$
$$c_{2,1} + c_{2,2} + c_{2,3} \leqslant b \tag{3.3b}$$
$$c_{2,1} + c_{2,3} > b \tag{3.4a}$$
$$c_{1,1} + c_{1,2} + c_{2,2} > b \tag{3.4b}$$

Multiplication of (3.3a) by -1 and addition to (3.4b) gives

$$c_{2,2} > 0,$$

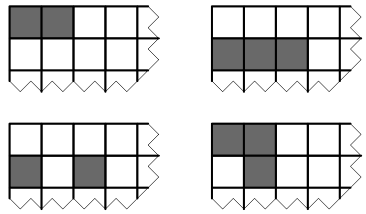

Figure 3.4: Vertices of the full orbitope (top row) and vertices of the cube that are not in the orbitope (bottom row). We show the upper left corner of each vertex; all other entries are 0.

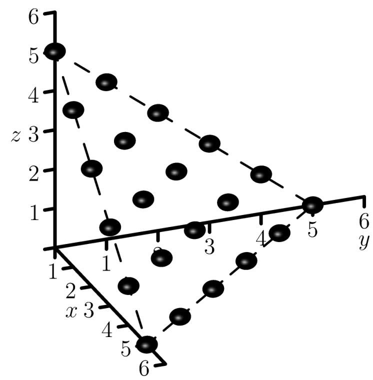

Figure 3.5: Visualization of triangular number $T(2,6) = 21$.

while multiplication of (3.4a) by -1 and addition to (3.3b) yields

$$c_{2,2} < 0.$$

Contradiction. □

We close this section by counting the number of vertices in full orbitopes. For this purpose, we use the *mth k-simplex number* $T(k,m)$, which can be defined as the number of integer points in the set

$$\mathcal{T}(k,m) := \{\boldsymbol{x} \in \mathbb{R}_+^{k+1} \mid \sum_{i=1}^{k+1} x_i = m-1\},$$

for $k,m \in \mathbb{N}_>$. Note that $T(k,2)$ is the number of integral points in the k-dimensional standard simplex and that $T(2,m)$ is also known as the *mth triangular number* (see figure 3.5 for an example). It is easy to see that

$$\sum_{i=1}^{m} T(k-1,i) = T(k,m)$$

holds, using inductively that there are $T(k-1, m-i)$ different integer points in $\mathcal{T}(k,m)$ with first component i for all $i \in [m-1]_0$.

Proposition 3.19 *The full orbitope $\mathbf{O}_{p,q}$ has $T(q, 2^p)$ vertices, where $T(k,m)$ denotes the mth k-simplex number.*

Proof. If $\boldsymbol{x} \in [m]^k$ is a k-tupel $\boldsymbol{x} = (x_1, \dots, x_k)$ with entries $x_1 \geqslant x_2 \geqslant \dots \geqslant x_k$, we refer to $\boldsymbol{x}$ as an *ordered* (k, m)*-tupel.* Because of Observation 3.15, the number of ordered $(q, 2^p)$-tupels equals the number of vertices of the full orbitope. (Note that each entry of an ordered $(q, 2^p)$-tupel lies between 1 and 2^p while the columns of the orbitope have decimal representations between 0 and $2^p - 1$.)

We count the number of ordered (k, m)-tupels inductively. For $k = 1$, there are

$$T(1, m) = \sum_{\ell=1}^{m} 1 = m$$

ordered $(1, m)$-tupels.

For $k > 1$, we observe that if $\boldsymbol{x}$ is an ordered (k, m)-tupel and entry x_1 has value ℓ, then the remaining entries $x_2, \dots, x_k$ form an ordered $(k-1, \ell)$-tupel. Hence, there are by induction $T(k-1, \ell)$ different (k, m)-tupels with entry $x_1 = \ell$ possible. Therefore, we get a number of

$$\sum_{\ell=1}^{m} T(k-1, \ell) = T(k, m)$$

ordered (k, m)-tupels.

□

$T(k, m)$ can be computed by

$$T(k, m) = \binom{m + k - 1}{k} = \frac{1}{k!} \prod_{i=1}^{k} (m + i - 1),$$

(which can for instance be shown using the properties of Pascal's triangle). Moreover, it is easy to see that $\binom{k}{\ell} \geqslant \left(\frac{k}{\ell}\right)^\ell$ using induction and the identity $\binom{k}{\ell} = \frac{k}{\ell}\binom{k-1}{\ell-1}$. This yields the following corollary.

Corollary 3.20 *The number of vertices of the full orbitope grows exponentially in* p *and* q*.*

For orbisacks, the identity from Proposition 3.19 can be simplified.

Corollary 3.21 *The orbisack* $\mathbf{O}_{p,2}$ *has*

$$\sum_{k=1}^{2^p} k = 2^{p-1}(1 + 2^p)$$

vertices.

3.2.1.1.2 (k)-Packing and (k)-Partitioning Orbitopes

In [65], Kaibel and Pfetsch study packing or partitioning orbitopes, i.e. k-packing and k-partitioning orbitopes with $k = 1$. They observe for example that for each vertex $\boldsymbol{v}$ of $\mathrm{O}^{\leqslant}_{p,q}$ or $\mathrm{O}^{=}_{p,q}$, the "upper right triangle" is fixed to 0:

Observation 3.22 For any vertex $\boldsymbol{v}$ of a packing or a partitioning orbitope, it holds that

$$v_{i,j} = 0 \text{ for each } (i, j) \in [p] \times [q] \text{ with } i < j.$$

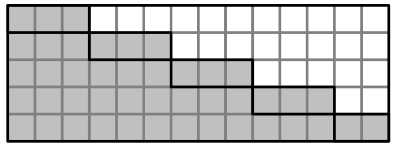

Figure 3.6: Example for upper triangle fixing for k-partitioning and k-packing. Here, $k = 3$, $p = 5$ and $q = 14$. Any white entry is fixed to 0.

This observation implies that for partitioning orbitopes, $v_{1,1} = 1$ must hold, and similarly, $v_{\ell,1} = 1$ must hold for packing orbitopes, where ℓ is the first row containing a nonzero entry. We will generalize these results to k-packing and k-partitioning orbitopes using basically the same argument as Kaibel and Pfetsch.

Proposition 3.23 *Let $k \geqslant 1$ and let $\boldsymbol{x}$ be a vertex of packing orbitope $\mathrm{O}_{p,q}^{\leqslant k}$ or of partitioning orbitope $\mathrm{O}_{p,q}^{=k}$. Then the following equations hold for $\boldsymbol{x}$:*

$$x_{i,j} = 0 \text{ for all } i \in [p] \text{ and } j \in [ik + 1..q].$$

(See figure 3.6.)

Proof. Assume the statement is false. Then there is an index (s, t) minimal with respect to the total order on $[p] \times [q]$ such that $x_{s,t} = 1$ and $t > sk$.

Since there are not more than k 1-entries in row s allowed, this implies that there must be a column ℓ among the k indices $[(s - 1)k + 1..sk]$ such that $v_{s,\ell} = 0$. Moreover, because of the minimality of (s, t), we know that $v_{s,\ell} = 0$ even for all $i \in [s]$. We move column t before column ℓ, i.e. we shift all columns $[\ell..t]$ cyclically by one position to the right:

$$(1, \ldots, \ell, \ldots, t, \ldots, q) \mapsto (1, \ldots, t, \ell, \ldots, t - 1, t + 1 \ldots, q)$$

This gives a vector lexicographic larger than $\boldsymbol{x}$. Contradiction. □

Proposition 3.24 *For partitioning orbitopes, the following equations hold additionally to the equations from Proposition 3.23:*

$$\sum_{j=1}^{q} x_{i,j} = k$$

for all $i \in [p]$.

Proof. Follows from the definition of partitioning orbitopes. □

Kaibel and Pfetsch also give an inductive method to generate vertices of packing and partitioning row by row. This can as well be extended to k-partitioning and k-packing orbitopes in general.

We will construct row ℓ of some vertex $\boldsymbol{v}$. For $\ell = 1$ and packing orbitopes, the first row can either contain only 0s or we fill the first row with up to k 1s on positions $v_{1,j}$, $j \in [k]$, starting with $v_{1,1}$. For partitioning orbitopes, $v_{1,j} = 1$ for all $j \in [k]$ is the only possibility. Note that this automatically fixes for partitioning orbitopes the split $\sigma_k = 1$, and for packing orbitopes there is a column $j \in [k]$ such that $\sigma_j = 1$ if the first row is not $\mathbb{0}$.

$\ell - 1 \rightsquigarrow \ell$. Note that for packing orbitopes, it is possible that all entries in the list of splits $\boldsymbol{\sigma}$ are yet undefined. We choose now a family $\mathcal{T}$ of pairwise disjoint ranges $[s..t] \subseteq [q]$ with the following properties:

(i) If range $[s..t] \in \mathcal{T}$, then there is no range in $\mathcal{T}$ starting at $t+1$.

(ii) For any range $[s..t] \in \mathcal{T}$, either $s = 1$ or there is a split defined in $\boldsymbol{\sigma}$ with $\sigma_j = s - 1$.

(iii) The total length $|\bigcup_{\mathcal{I} \in \mathcal{T}} \mathcal{I}|$ equals k for partitioning orbitopes and is less or equal k for packing orbitopes.

We set $v_{\ell,j} := 1$ for all $j \in \bigcup_{\mathcal{I} \in \mathcal{T}} \mathcal{I}$ and $v_{\ell,j} := 0$ otherwise. Moreover, for any $[s..t] \in \mathcal{T}$, if σ_t is undefined so far, we set $\sigma_t := \ell$.
Last, after defining the entries in row p, we set $\sigma_j := p + 1$ for any σ_j that is undefined so far.

Remark 3.25 Note that the vertices of the packing orbitope $\mathrm{O}^{\leqslant}_{p-1,q-1}$ can also be constructed from the vertices of the partitioning orbitope $\mathrm{O}^{=}_{p,q}$ by simply deleting the first column and the first row in each vertex. On the other hand, one can construct from each vertex of $\mathrm{O}^{\leqslant}_{p-1,q-1}$ a vertex of $\mathrm{O}^{=}_{p,q}$ by means of linear map $\phi : \mathbb{R}^{[p-1]\times[q-1]} \to \mathbb{R}^{[p]\times[q]}$ defined by

$$\phi(\boldsymbol{x})_{i,j} = \begin{cases} 1 - \sum_{i=1}^{q-1} x_{i,j}, & \text{if } j = 1 \\ x_{i-1,j-1}, & \text{otherwise} \end{cases} \quad \text{for all } (i,j) \in [p] \times [q].$$

Therefore, $\mathrm{O}^{=}_{p,q}$ and $\mathrm{O}^{\leqslant}_{p-1,q-1}$ are affinely isomorphic (see [65]).
Similarly, we can construct any vertex of a k-packing orbitope $\mathrm{O}^{\leqslant k}_{p-k,q-k}$ from a vertex of $\mathrm{O}^{=k}_{p,q}$ by deleting the first row and the first k columns. However, for $k > 1$, $\mathrm{O}^{=k}_{p,q}$ and $\mathrm{O}^{\leqslant k}_{p-1,q-k}$ are not isomorphic anymore.

We make some observations additional to Observation 3.12 which concern special cases of k-partitioning and k-packing orbitopes.

Observation 3.26

- For $p = 1$, the partitioning orbitope $\mathrm{O}^{=k}_{1,q}$, and for arbitrary p, the partitioning orbitope $\mathrm{O}^{=k}_{p,k}$ consist both of one vertex each.
- The packing orbitope $\mathrm{O}^{\leqslant k}_{p,q}$ is for $p = 1$ isomorphic to the full orbitope $\mathbf{O}_{1,k}$.
- For an arbitrary number of rows p, $\mathrm{O}^{\leqslant k}_{p,k}$ is isomorphic to $\mathbf{O}_{p,k}$.
- From Observation 3.12 it follows in particular that $\mathrm{O}^{=k}_{p,k+1}$ is isomorphic to $\mathrm{O}^{=}_{p,k+1}$.

Proposition 3.27 *For $k > 2$ and $p > 1$, k-packing $\mathrm{O}^{\leqslant k}_{p,q}$ orbitopes and for $k > 2$, $p > 1$, and $q > k + 1$, k-partitioning orbitopes are not isomorphic to knapsack orbitopes.*

Proof. For k-packing orbitopes, the proof is identical to the proof for full orbitopes (Proposition 3.18).
For k-partitioning orbitopes, we observe that the 0/1-vectors shown in the top row of figure 3.7 are in $\mathrm{O}^{=k}_{p,q}$ and the 0/1-vectors in the bottom row of figure 3.7 are not. Assuming there is a knapsack inequality $\langle \boldsymbol{c}, \boldsymbol{x} \rangle \leqslant b$ separating the allowed vectors from the forbidden ones, we get inequalities

$$\sum_{i=1}^{p} \sum_{j=1}^{k} c_{i,j} \leqslant b \tag{3.5a}$$

$$\sum_{i=1}^{p} \sum_{j=1}^{k} c_{i,j} - c_{2,k-1} - c_{2,k} + c_{2,k+1} + c_{2,k+2} \leqslant b \tag{3.5b}$$

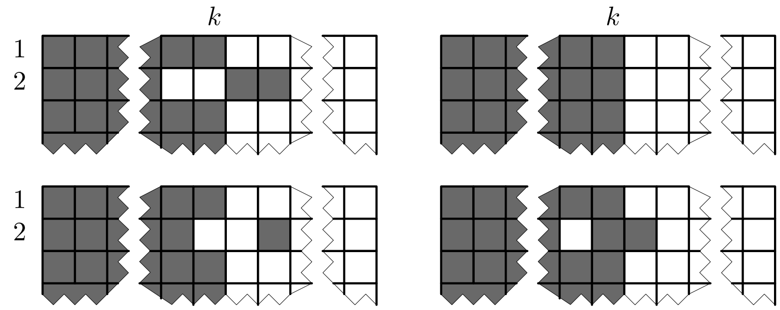

Figure 3.7: Vertices of the partitioning orbitope (top row) and vertices of the cube that are not in the orbitope (bottom row).

and

$$\sum_{i=1}^{p}\sum_{j=1}^{k} c_{i,j} - c_{2,k-1} + c_{2,k+1} > b \tag{3.5c}$$

$$\sum_{i=1}^{p}\sum_{j=1}^{k} c_{i,j} - c_{2,k} + c_{2,k+2} > b. \tag{3.5d}$$

Multiplying inequality (3.5a) with -1 and adding it to inequality (3.5c), we obtain that

$$c_{2,k+1} > c_{2,k-1}.$$

On the other hand, we can multiply inequality (3.5b) with -1 and add it to (3.5d) to get

$$c_{2,k-1} > c_{2,k+1}.$$

Contradiction. □

3.2.1.1.3 (k)-Covering Orbitopes

As has been observed in the section on k-packing and k-partitioning orbitopes above, $\mathrm{O}_{p,q}^{\leqslant k}$ is isomorphic to $\mathrm{O}_{p,q}^{\geqslant q-k}$. Therefore, most of the facts concerning k-partitioning and k-packing orbitopes can be transferred and adapted to k-covering orbitopes.

In particular, for $q = 2$ and $k = 1$, $\mathrm{O}_{p,2}^{\geqslant}$ and $\mathrm{O}_{p,2}^{\leqslant}$ are isomorphic. This opens the way to a full facial description of covering orbitope $\mathrm{O}_{p,2}^{\geqslant 1}$ via [65]. (See page 70, in particular Proposition 3.61.)

It is clear that for $k > 1$, the method of generation of vertices described for k-packing and k-partitioning can also be applied to k-covering orbitopes; in this case, one has to ensure that at least the first k entries in the first row are set to 1 and $|\bigcup_{\mathcal{I}\in\mathcal{T}} \mathcal{I}| \geqslant k$ for each row below.

We immediately obtain that all vertices of $\mathrm{O}_{p,q}^{\geqslant k}$ share the following property:

Proposition 3.28 *Let $k \geqslant 1$. For any vertex $\boldsymbol{x}$ of $\mathrm{O}_{p,q}^{\geqslant k}$, equations*

$$x_{1,j} = 1$$

hold for all $j \in [k]$.

Moreover, we obtain from Observation 3.26 and Proposition 3.27 the following proposition.

Proposition 3.29 *For $k < q-2$ and $p > 1$, k-covering orbitopes $\mathrm{O}^{\geqslant k}_{p,q}$ are not knapsack orbitopes.*

3.2.1.2 Optimization and Complexity

3.2.1.2.4 Full Orbitopes

Lemma 3.13 shows how to generate all vertices of orbitope $\mathbf{O}_{p,q}$. Roughly speaking, this can be done by "glueing" together an arbitrary vertex of $\mathbf{O}_{p-1,k}$ and another one of $\mathbf{O}_{p-1,q-k}$, and topping both with a row filled with k 1-entries followed by $q-k$ 0-entries. Note that $0 \leqslant k \leqslant q$, so in particular $k = 0$ or $q-k = 0$ are allowed.

This idea is the cornerstone of the construction of an algorithm for optimization over full orbitopes. For a formal description of this algorithm, we introduce the notion of *bricks*.

Definition 3.30 (Bricks) Let $\mathcal{I} := \{[s..t] \mid [s..t] \subseteq [q]\}$ be the family of all ranges in $[q]$. We define the following sets of tupels

$$\mathcal{B}^{\bullet}_{p,q} := [p] \times \mathcal{I} \times \{\text{black}\} \quad \text{and} \quad \mathcal{B}^{\circ}_{p,q} := [p] \times \mathcal{I} \times \{\text{white}\}$$

as well as set

$$\mathcal{B}_{p,q} := \mathcal{B}^{\bullet}_{p,q} \cup \mathcal{B}^{\circ}_{p,q}.$$

We denote by $b^{\bullet}_{k,[s..t]}$ the elements of $\mathcal{B}^{\bullet}_{p,q}$ and by $b^{\circ}_{k,[s..t]}$ the elements of $\mathcal{B}^{\circ}_{p,q}$. Furthermore, we define a map

$$\mathrm{M}^{p,q} : \mathcal{B}_{p,q} \to \mathcal{M}_{p,q}$$

that maps any element $b_{k,[s..t]} \in \mathcal{B}_{p,q}$ to a 0/1-matrix with p rows and q columns and is defined by

$$\left(\mathrm{M}^{p,q}(b^{\bullet}_{k,[s..t]})\right)_{i,j} = \begin{cases} 1, & \text{if } i = k \text{ and } j \in [s..t] \\ 0, & \text{otherwise} \end{cases} \qquad \text{for all } (i,j) \in [p] \times [q].$$

and

$$\mathrm{M}^{p,q}(b^{\circ}_{k,[s..t]}) = \mathbb{0}_{p,q}$$

for all $k \in [p]$ and $[s..t] \subseteq [q]$. The matrices in $\mathrm{M}^{p,q}(\mathcal{B}_{p,q})$ are referred to as *bricks*. Note that $\mathrm{M}^{p,q}$ is defined on $\mathcal{B}^{\circ}_{p,q}$ only for sake of completeness; for the following considerations, these bricks are not needed.

For the construction of the vertices of $\mathbf{O}_{p,q}(\mathfrak{S}_q)$, we use two liftings $\phi^{\text{left},q} : \mathbb{R}^{p-1,k} \to \mathbb{R}^{p,q}$ defined by

$$\phi^{\text{left},q}(\boldsymbol{x})_{i,j} = \begin{cases} 0, & \text{if } i = 1 \text{ or } j > k \\ x_{i-1,j}, & \text{otherwise,} \end{cases}$$

and $\phi^{\text{right},q} : \mathbb{R}^{p-1,k} \to \mathbb{R}^{p,k}$ defined by

$$\phi^{\text{right},q}(\boldsymbol{x})_{i,j} = \begin{cases} 0, & \text{if } i = 1 \text{ or } j < q-k \\ x_{i-1,j-(q-k)}, & \text{otherwise,} \end{cases}$$

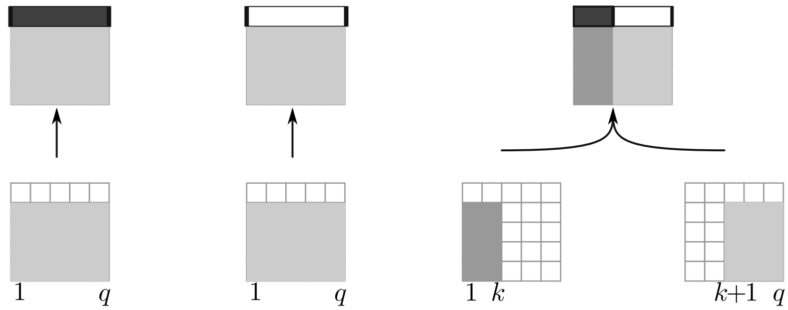

Figure 3.8: Vertices of a full orbitope over the full symmetric group can be generated recursively row by row. There are basically three possibilities to put together a vertex.

where $k \leqslant q$ in both cases. (Note that for $k = q$, it holds that $\phi^{\text{left},q} = \phi^{\text{right},q}$.) Lemma 3.13 shows that any vertex $\boldsymbol{v}$ of orbitope $\mathbf{O}_{p,q}$ has one of the following three forms. Either,

$$\boldsymbol{v} = \mathrm{M}^{p,q}(b^{\bullet}_{1,[1..q]}) + \phi^{\text{left},q}(\boldsymbol{v}') \tag{3.6a}$$

or

$$\boldsymbol{v} = \phi^{\text{left},q}(\boldsymbol{v}') \tag{3.6b}$$

with $\boldsymbol{v}'$ some vertex of $\mathbf{O}_{p-1,q}$ in both cases, or

$$\boldsymbol{v} = \mathrm{M}^{p,q}(b^{\bullet}_{1,[1..k]}) + \phi^{\text{left},q}(\boldsymbol{v}') + \phi^{\text{right},q}(\boldsymbol{v}'') \tag{3.6c}$$

with $\boldsymbol{v}'$ some vertex of $\mathbf{O}_{p-1,k}$, $\boldsymbol{v}''$ some vertex of $\mathbf{O}_{p,q-k}$, and $0 < k < q$. (For an illustration of these three possibilities, see figure 3.8.)

Next, we tackle the problem $\mathcal{P}$ of linear optimization (maximization) over full orbitope $\mathbf{O}_{p,q}$. Any instance $\mathcal{P}(p, q, \boldsymbol{c})$ of this problem is described by some cost vector $\boldsymbol{c} \in \mathbb{Q}^{[p]\times[q]}$, and we aim to find a vector $\boldsymbol{x}^\star \in \mathbf{O}_{p,q}$ such that

$$\langle \boldsymbol{x}^\star, \boldsymbol{c} \rangle = \max\{\langle \boldsymbol{c}, \boldsymbol{x} \rangle \mid \boldsymbol{x} \in \mathbf{O}_{p,q}\}. \tag{3.7}$$

Let

$$\zeta_{\mathcal{P}(p,q,\boldsymbol{c})} := \langle \boldsymbol{x}^\star, \boldsymbol{c} \rangle$$

denote the objective value of the optimal solution.

Lemma 3.31 *Let $\boldsymbol{c} \in \mathbb{R}^{[p]\times[q]}$ be some cost vector and let $\boldsymbol{x}^\star$ be an optimal solution to instance $\mathcal{P}(p, q, \boldsymbol{c})$. Then*

$$\begin{aligned} &\langle \boldsymbol{c}, \boldsymbol{x}^\star \rangle = \\ &= \max\Big(\Big\{\sum_{j=1}^{q} c_{1,j} + \zeta_{\mathcal{P}(p-1,q,\boldsymbol{c}_{[2..p]\times[q]})}\Big\} \cup \\ &\qquad \Big\{\zeta_{\mathcal{P}(p-1,q,\boldsymbol{c}_{[2..p]\times[q]})}\Big\} \cup \\ &\qquad \Big\{\sum_{j=1}^{k} c_{1,j} + \zeta_{\mathcal{P}(p-1,k,\boldsymbol{c}_{[2..p]\times[1..k]})} + \zeta_{\mathcal{P}(p-1,q-k,\boldsymbol{c}_{[2..p]\times[k+1..q]})} \,\Big|\, k \in [q-1]\Big\}\Big), \end{aligned} \tag{3.8}$$

where $\mathbf{c}_{[i..j]\times[k..\ell]}$ denotes the submatrix of $\mathbf{c}$ defined by rows i through j and columns k through ℓ.

Proof. We build $\boldsymbol{x}^\star$ according to Lemma 3.13 either from the combination of two optimal solutions to subproblems $\mathcal{P}(p-1, k, \boldsymbol{c}_{[2..p]\times[1..k]})$ and $\mathcal{P}(p-1, q-k, \boldsymbol{c}_{[2..p]\times[k+1..q]})$ or from the optimal solution to subproblem $\mathcal{P}(p-1, q, \boldsymbol{c}_{[2..p]\times[q]})$. □

Note that equation (3.8) is the Bellman equation for a dynamic programming algorithm to linearly optimize over $\mathbf{O}_{p,q}(\mathfrak{S}_q)$.

For a description of this algorithm, we first extend the set $\mathcal{B}_{p,q}$ by adding an element $b_{0,[1..q]}$ which will serve as final state in the algorithm; that is, we define set

$$\mathcal{B}^0_{p,q} := \mathcal{B}_{p,q} \cup \{b_{0,[1..q]}\}.$$

Next, we define a dynamic programming table $\mathcal{L}$ containing one entry $L(b_{i,[j..k]})$ for each element in $\mathcal{B}^0_{p,q}$.

- For element $b^\bullet_{k,[s..t]}$, $1 \leqslant k < p$ and $[s..t] \subseteq [1..q]$, entry $L(b^\bullet_{k,[s..t]})$ stores the value
$$\zeta_{\mathcal{P}(p-k,t-s+1,\boldsymbol{c}_{[k+1..p]\times[s..t]})} + \sum_{i=s}^{t} c_{k,i}.$$
- For element $b^\circ_{k,[s..t]}$, $1 \leqslant k < p$ and $[s..t] \subseteq [1..q]$, entry $L(b^\circ_{k,[s..t]})$ stores the value
$$\zeta_{\mathcal{P}(p-k,t-s+1,\boldsymbol{c}_{[k+1..p]\times[s..t]})}.$$
- For elements $b^\bullet_{p,[s..t]}$ in row p, $[s..t] \subseteq [1..q]$, we set
$$L(b^\bullet_{p,[s..t]}) := \sum_{j=s}^{t} c_{p,j},$$
- and for elements $b^\circ_{p,[s..t]}$ in row p, $[s..t] \subseteq [1..q]$, we define
$$L(b^\circ_{p,[s..t]}) := 0.$$
- Last, $L(b_{0,[1..q]})$ stores the optimal objective value to $\mathcal{P}(p, q, \boldsymbol{c})$.

The complete algorithm is shown in figure 3.9. Note that in the form shown there, only the optimal objective value is computed. However, if one stores in each state which predecessor node(s) are used to derive the respective partial solution, then one can easily reconstruct the set of nodes $\mathcal{S} \subset \mathcal{B}^0_{p,q}$ that take part in establishing the optimal objective value. The optimal solution is then given by

$$\boldsymbol{x}^\star = \sum_{b^\bullet_{k,[s..t]} \in \mathcal{S}\cap\mathcal{B}^\bullet_{p,q}} \mathrm{M}^{p,q}(b^\bullet_{k,[s..t]}).$$

Theorem 3.32 *The algorithm in figure 3.9 works correctly.*

Proof. Follows from Lemma 3.31. □

Theorem 3.33 *The algorithm in figure 3.9 has running time $\mathcal{O}(pq^3)$.*

Proof. There are $\mathcal{O}(pq^2)$ entries in $\mathcal{L}$. For each entry, the computation needs time of $\mathcal{O}(q)$. Hence over all, we need time $\mathcal{O}(pq^3)$. □

The algorithm for optimization over full orbitopes is based on dynamic programming. Hence, we can define a directed hypergraph associated with this algorithm (see section 2.3). Any vertex of the full orbitope can be mapped to a hyperpath in

Data: cost vector $\boldsymbol{c} \in \mathbb{R}^{[p]\times[q]}$
Result: optimum $\max\{\langle \boldsymbol{c}, \boldsymbol{x}\rangle \mid \boldsymbol{x} \in \mathbf{O}_{p,q}\}$
`// Initialize`
foreach $[s..t] \in [q]$ **do**
    $L(b^{\bullet}_{p,[s..t]}) \leftarrow \sum_{j=s}^{t} c_{p,j}$;
    $L(b^{\circ}_{p,[s..t]}) \leftarrow 0$;
end
`// Build up` $\mathcal{L}$ `row by row`
$k \leftarrow p-1$;
while $k \neq 0$ **do**
    foreach $[s..t] \in [q]$ **do**

$$L(b^{\bullet}_{k,[s..t]}) \leftarrow \sum_{j\in[s..t]} c_{k,j} + \max\big(\{L(b^{\bullet}_{k+1,[s..t]}), L(b^{\circ}_{k+1,[s..t]})\} \quad \cup \quad \{L(b^{\bullet}_{k+1,[s..\ell]}) + L(b^{\circ}_{k+1,[\ell+1..t]}) \mid s \leqslant \ell < t\}\big);$$

$$L(b^{\circ}_{k,[s..t]}) \leftarrow \max\big(\{L(b^{\bullet}_{k+1,[s..t]}), L(b^{\circ}_{k+1,[s..t]})\} \quad \cup \quad \{L(b^{\bullet}_{k+1,[s..\ell]}) + L(b^{\circ}_{k+1,[\ell+1..t]}) \mid s \leqslant \ell < t\}\big)$$

    end
    $k \leftarrow k-1$;
end
`// Global solution`

$$L(b_{0,[1..q]}) \leftarrow \max\big(\{L(b^{\bullet}_{1,[1..q]}), L(b^{\circ}_{1,[1..q]})\} \quad \cup \quad \{L(b^{\bullet}_{1,[1..\ell]}) + L(b^{\circ}_{1,[\ell+1..q]}) \mid 1 \leqslant \ell < q\}\big)$$

return $L(b_{0,[1..q]})$;

Figure 3.9: Algorithm for finding $\max\{\langle \boldsymbol{c}, \boldsymbol{x}\rangle \mid \boldsymbol{x} \in \mathbf{O}_{p,q}(\mathfrak{S}_q)\}$. The algorithm can straightforwardly be adjusted such that it also allows to reconstruct the optimal solution recursively.

this hypergraph. In what follows, we will briefly characterize the DP-hypergraph $H = (\mathcal{B}, \mathcal{A})$ with vertex set $\mathcal{B}$ and hyperarc set $\mathcal{A}$ that is associated with the DP-algorithm from above (in the following referred to as the *orbitope hypergraph*) and append some polyhedral results.

The vertex set $\mathcal{B}$ is partitioned into subsets $\mathcal{B}_0 \uplus \mathcal{B}_1 \uplus \ldots \uplus \mathcal{B}_p$ defined as follows:

$$\mathcal{B}_i := \begin{cases} \{b_{i,[1..q]}\}, & \text{if } i = 0 \\ \{b^{\bullet}_{i,[1..s]} \mid 1 \leqslant s \leqslant q\} \qquad \cup \{b^{\circ}_{i,[s..q]} \mid 1 \leqslant s \leqslant q\}, & \text{if } i = 1 \\ \{b^{\bullet}_{i,[s..t]} \mid 1 \leqslant s \leqslant t \leqslant q\} \cup \{b^{\circ}_{i,[s..t]} \mid 1 \leqslant s \leqslant t \leqslant q\}, & \text{otherwise} \end{cases}$$

for all $i \in [p]_0$. We refer to nodes corresponding to elements in $\mathcal{B}^{\bullet}_{p,q}$ as *black* nodes and to nodes associated with elements in $\mathcal{B}^{\circ}_{p,q}$ by *white* nodes.

The set of hyperarcs $\mathcal{A}$ is defined as

$$\Big\{\big(\{b^{\bullet}_{i,[s..\ell]}, b^{\circ}_{i,[\ell+1..t]}\},\ b_{i-1,[s..t]}\big) \,\Big|\, b^{\bullet}_{i,[s..\ell]}, b^{\circ}_{i,[\ell+1..t]} \in \mathcal{B}_i,\ b_{i-1,[s..t]} \in \mathcal{B}_{i-1}\ \forall i \in [p]\Big\} \cup \\ \Big\{\big(b_{i,[s..t]},\ b_{i-1,[s..t]}\big) \,\Big|\, b_{i,[s..t]} \in \mathcal{B}_i,\ b_{i-1,[s..t]} \in \mathcal{B}_{i-1}\ \forall i \in [p]\Big\}.$$

We denote in the following the final state $b_{0,[1..q]}$ also by b_t.

Observation 3.34

(a) It is easy to see that $H = (\mathcal{B}, \mathcal{A})$ is in fact a DP-hypergraph: it has a unique sink b_t and it is an acyclic B-hypergraph. The set of sources is $\mathcal{B}_p$. Obviously, the map $\mathcal{R} : b_{i,[j..k]} \mapsto [j..k]$ constitutes a reference system $\mathcal{R}$ satisfying alternative criterion [dph3'].

(b) Each B_p-b_t-hyperpath in this DP-hypergraph corresponds uniquely to a vertex of $\mathbf{O}_{p,q}$ and vice versa. This follows from the construction of vertices as described on page 53 ff. An example for a composition of a vertex is shown in figure 3.10. The nodes of the hypergraph are in this illustration ordered by color and by sets $\mathcal{B}_i$ they are contained in, so brick $b^{\bullet}_{i,[j..k]}$ can be found among the black nodes in $\mathcal{B}_i$ in row j and column k.

(c) The number of arcs in H is

$$|\mathcal{A}| = 1 + \frac{2}{3}q - 3q^2 + 2pq^2 + \frac{5}{3}pq + \frac{1}{3}pq^3 - \frac{2}{3}q^3 = \mathcal{O}(pq^3)$$

Proof. Let H associated with $\mathbf{O}_{p,q}$ and denote by $h_{q,i}$ ($\tilde{h}_{q,i}$) the number of hyperarcs with two (one) vertices in the tail, respectively, that are leaving from set $\mathcal{B}_i$ and ending in set $\mathcal{B}_{i-1}$.

- For $i = 1$, $h_{q,1} = q - 1$ and $\tilde{h}_{q,1} = 2$.
- For $i = 2$, it holds that for any hyperarc a with $|\operatorname{tail}(a)| = 2$ either $b^{\bullet}_{i,[1..s]} \in \operatorname{tail}(a)$ or $b^{\circ}_{i,[s..q]} \in \operatorname{tail}(a)$ (or both) for some $s \in [q]$. Moreover, the first kind of arcs has a black node in the head, the second one a white node. Hence, $h_{q,2} = q(q-1)$; on the other hand, $\tilde{h}_{q,2} = 4q$.
- Let $i \in [3..p]$. Clearly, $h_{2,i} = 2$. We proceed from $q-1$ to q. Then $h_{q,i} = h_{q-1,i} + h'_{q,i}$, where $h'_{q,i}$ counts those hyperarcs a with $|\operatorname{tail}(a)| = 2$ and $b^{\circ}_{i,[s+1..q]} \in \operatorname{tail}(a)$ for $s \in [1..q-1]$. For each value of s, these are $2s$ hyperarcs, namely the hyperarcs $(\{b^{\bullet}_{i,[\ell..s]}, b^{\circ}_{i,[s+1..q]}\}, b_{i-1,[\ell..q]})$ for $\ell \in [1..s]$. Hence

$$h_{q,i} = h_{q-1,i} + q(q-1) = q(q-1)(q+1)/3$$

On the other hand, $\tilde{h}_{q,i} = 2q(q+1)$.

□

(d) The number of nodes of H is

$$|\mathcal{B}| = 1 + 2q + \frac{1}{2}(p-1)q(q+1)$$

Now, we will focus on the polyhedral aspects.

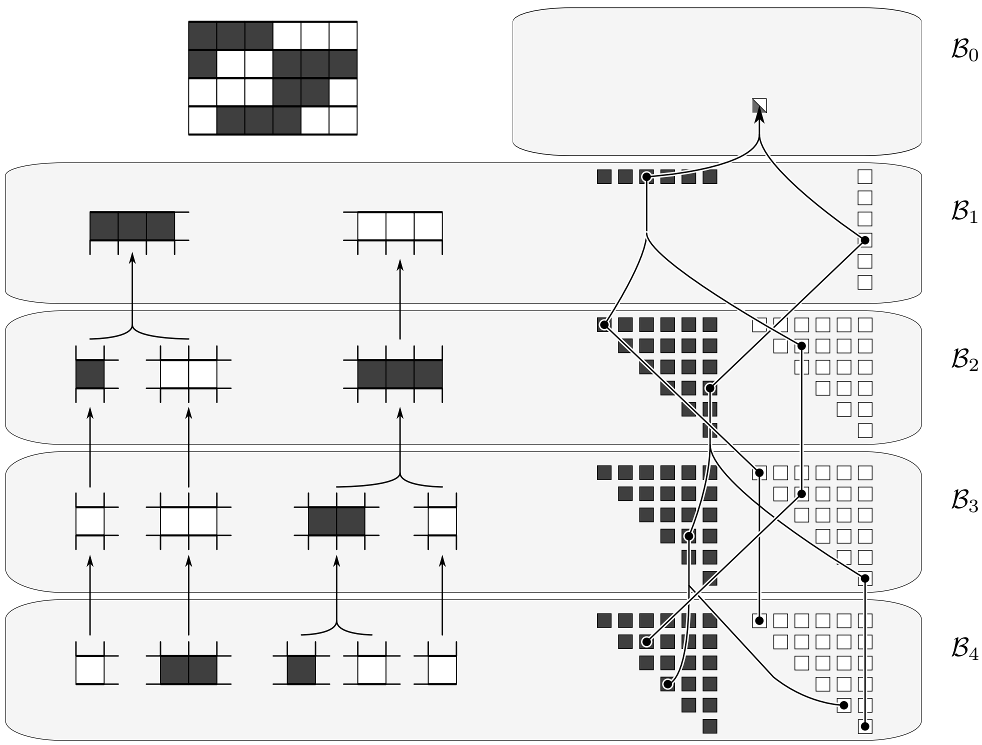

Figure 3.10: The composition of a vertex of the full orbitope ($\mathbf{O}_{4,6}$, in this case) by bricks. The complete vertex is shown on the top left. On the left side, one can see the relevant entries of the bricks in use. On the right side, the corresponding $\mathcal{B}_4$-b_t-hyperpath is shown, together with the complete set of nodes $\mathcal{B}$ of the DP hypergraph associated with orbitopes. Black nodes correspond to elements in $\mathcal{B}^{\bullet}_{p,q}$ and white nodes to elements in $\mathcal{B}^{\circ}_{p,q}$. The node set is partitioned into subsets $\mathcal{B}_0$ through $\mathcal{B}_4$ as described in the text. The nodes are ordered in arrays such that $b_{k,[s..t]}$ can be found in $\mathcal{B}_k$ in row s and column t.

Definition 3.35 We denote by $\mathbf{P}^{\mathrm{arc}}(H)$ the $\mathcal{B}_p$-b_t-hyperpath polytope in arc variables and by $\mathbf{P}^{\mathrm{node}}(H)$ the $\mathcal{B}_p$-b_t-hyperpath set polytope.

It is easy to obtain a complete linear description of $\mathbf{P}^{\mathrm{arc}}(H)$ by means of Theorem 2.14 due to Martin et al..

Corollary 3.36 *Denoting by $\boldsymbol{y} \in \{0,1\}^{\mathcal{A}}$ the incidence vector of a hyperpath in arc variables, the following inequalities provide a full description of $\mathbf{P}^{\mathrm{arc}}(H)$:*

$$\sum_{a \in \delta^{\mathrm{in}}_H(b_t)} y_a = 1 \tag{3.9}$$

$$\sum_{a \in \delta^{\mathrm{in}}_H(b_{i,[s..t]})} y_a - \sum_{a \in \delta^{\mathrm{out}}_H(b_{i,[s..t]})} y_a = 0 \qquad \forall b_{i,[s..t]} \in \bigcup_{i=1}^{p-1} \mathcal{B}_i \tag{3.10}$$

$$y_a \geqslant 0 \qquad \forall a \in \mathcal{A} \tag{3.11}$$

As a consequence, one can derive from this linear description an extended formulation for the orbitope $\mathbf{O}_{p,q}$ using a linear transformation which is based on the following observation:

Observation 3.37 If the hyperpath enters some node $b_{i,[j..k]} \in \mathcal{B}$, then this corresponds to fixing entries $x_{i,j}$ through $x_{i,k}$ in the associated vertex $\boldsymbol{x}$ of the orbitope to either 1 or 0 (depending on the color of the node). On the other hand, each entry $x_{i,j}$ of an orbitope vertex is determined by the choice of one node $b_{i,[s..t]} \in \mathcal{B}$ with $1 \leqslant s \leqslant j \leqslant t \leqslant q$. In particular, if $x_{i,j} = 1$, then the hyperpath corresponding to $\boldsymbol{x}$ must have been using one black node in $\{b^{\bullet}_{i,[s..t]} \mid 1 \leqslant s \leqslant j \leqslant t \leqslant q\}$, and this means at the same time that one arc in the outstars of these nodes has been used.

Lemma 3.38 *There is a linear projection of $\mathbf{P}^{\mathrm{arc}}(H)$ to $\mathbf{O}_{p,q}$ that establishes together with the linear description from Corollary 3.36 an extended formulation for the orbitope.*

Proof. Let linear projection $\vartheta : \mathbf{P}^{\mathrm{arc}}(H) \to \mathbf{O}_{p,q}$ be defined as follows:

$$\vartheta : \mathbb{R}^{\mathcal{A}} \to \mathbb{R}^{[p]\times[q]}, \quad x_{i,j} = \sum_{1 \leqslant s \leqslant j \leqslant t \leqslant q} \; \sum_{a \in \delta^{\mathrm{out}}_H(b^{\bullet}_{i,[s..t]})} y_a$$

$\subseteq$ $\vartheta(\mathbf{P}^{\mathrm{arc}}) \subseteq \mathbf{O}_{p,q}$. $\supseteq$ On the other hand, the DP-algorithm allows to generate for each vertex of the orbitope one B_p-b_t-path in H, so $\mathbf{O}_{p,q} \subseteq \vartheta(\mathbf{P}^{\mathrm{arc}})$. □

Remark 3.39 Of course, we can similarly project $\mathbf{P}^{\mathrm{node}}(H)$ to $\mathbf{O}_{p,q}$, by means of a projection $\widetilde{\vartheta} : \mathbf{P}^{\mathrm{node}}(H) \to \mathbf{O}_{p,q}$ defined by

$$x_{i,j} = \sum_{1 \leqslant s \leqslant j \leqslant t \leqslant q} u_{b_{i,[s..t]}},$$

where $\boldsymbol{u} \in \mathbf{P}^{\mathrm{node}}(H)$.

It is obvious that the algorithm from figure 3.9 can be seen as the computation of the length of a $\mathcal{B}_p$-b_t-hyperpath in H of maximal length.

3.2.1.2.5 Packing and Partitioning Orbitopes

The first optimization algorithm over orbitopes with symmetric groups has been given by Kaibel and Pfetsch for packing and partitioning orbitopes in [65]. In its

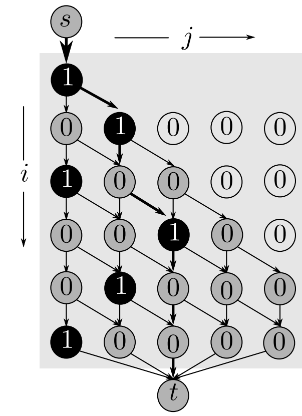

Figure 3.11: Digraph giving an extended formulation for partitioning orbitopes. The nodes inside the gray area correspond to the entries of the vertex of the partitioning orbitope; entries in the upper right triangle are fixed to 0 and are therefore disregarded for the digraph (see Observation 3.22). In each row, one can either put the 1-entry within the columns used so far or one can extend the column range by one. For each row, the s-t-path encodes the maximal allowed column. Modifying the setting to packing orbitopes is simply done by making column $j = 1$ encode whether a 0-row is used or not.

initial form, it had time complexity of $\mathcal{O}(p^2 q)$ in a unit-cost model. Later, it has been improved by Faenza and Kaibel [35] to time $\mathcal{O}(pq)$. On the way, they developed an extended formulation which relies on the fact that any vertex of the packing and partitioning orbitope induces an s-t-path in an acyclic network. We will briefly plot their idea in the partitioning setting which can easily be adapted to packing orbitopes.

Assuming $p \geqslant q \geqslant 1$, we start by defining the vertex set of a DAG $D = (\mathcal{V}, \mathcal{A})$ as

$$\mathcal{V} := \{(i,j) \in [p] \times [q] \mid i \leqslant j\} \uplus \{s,t\}.$$

For the arc set, we set

$$\mathcal{A}^{\downarrow} := \{\big((i,j),(i+1,j)\big) \mid (i,j) \in V \text{ with } i < p\}$$

and

$$\mathcal{A}^{\searrow} := \{\big((i,j),(i+1,j+1)\big) \mid (i,j) \in V \text{ with } i < p \text{ and } j < q\}.$$

Then the whole arc set $\mathcal{A}$ is defined as

$$\mathcal{A} := \mathcal{A}^{\searrow} \cup \mathcal{A}^{\downarrow} \cup \{\big(s,(1,1)\big)\} \cup \{\big((p,j),t\big) \mid j \in [q]\}.$$

(See figure 3.11 for an example.)

The main idea is that for optimization, it is not necessary to construct the whole vertex of a partitioning orbitope: it suffices to decide in row ℓ whether one stays within the columns used so far or uses a "new" column.

Let the cost functional be $\langle \boldsymbol{c}, \boldsymbol{x} \rangle$ with given cost vector $\boldsymbol{c} \in \mathbb{Q}^{[p]\times[q]}$ and assume we aim to find

$$\zeta := \max(\{\sum_{(i,j)\in[p]\times[q]} c_{i,j} x_{i,j} \mid \boldsymbol{x} \in \mathrm{O}^{=}_{p,q}\}).$$

This can be done by fixing the following arc weights in D:

- arc $(s,(1,1)$ gets weight $c_{1,1}$,
- any arc $\big((i,j),(i+1,j)\big) \in \mathcal{A}^{\downarrow}$ gets weight $\max(\{c_{i+1,k} \mid k \in [j]\})$,
- any arc $\big((i,j),(i+1,j+1)\big) \in \mathcal{A}^{\searrow}$ gets weight $c_{i+1,j+1}$, and

- all other arcs get weight 0.

Obviously, the length of a longest s-t-path in D is exactly the optimum value and ζ can be found in time $\mathcal{O}(pq)$.

The isomorphy between $\mathrm{O}^{\leqslant}_{p,q}$ and $\mathrm{O}^{=}_{p+1,q+1}$ (remark 3.25) makes it easy to adapt the algorithm to the packing case.

3.2.1.2.6 k-Packing, k-Partitioning, k-Covering and related Orbitopes

Optimization becomes in general hard as soon as the upper, exact or lower bounds on the number of 1s in each row of the solution become larger than 1. More precisely, we will show in the following that optimization over k-covering orbitopes is already $\mathcal{NP}$-hard for $k = 1$, while k-packing and k-partitioning become both $\mathcal{NP}$-hard with $k \geqslant 2$, as long as the respective orbitopes have sufficiently many rows and columns.

However, optimization over k-covering, k-packing, k-partitioning and fixed row sum orbitopes can all be done in polynomial time for constant number of rows or columns. (See propositions 3.53 and 3.52.)

Before proving these facts, we will formulate the decision problems associated with the optimization problems above.

Problem 3.40 (Orbitopal k-Packing) Given a number of rows p, a number of columns q, an objective vector $\boldsymbol{c} \in \mathbb{R}^{[p]\times[q]}$, an integer $k \geqslant 2$ and a bound B: is there a vertex $\boldsymbol{v}^\star \in \mathrm{O}^{\leqslant k}_{p,q}$ such that $\langle \boldsymbol{v}^\star, \boldsymbol{c}\rangle \leqslant B$?

Problem 3.41 (Orbitopal k-Partitioning) Given a number of rows p, a number of columns q, an objective vector $\boldsymbol{c} \in \mathbb{R}^{[p]\times[q]}$, an integer $k \geqslant 2$ and a bound B: is there a vertex $\boldsymbol{v}^\star \in \mathrm{O}^{=k}_{p,q}$ such that $\langle \boldsymbol{v}^\star, \boldsymbol{c}\rangle \leqslant B$?

Problem 3.42 (Orbitopal k-Covering) Given a number of rows p, a number of columns q, an objective vector $\boldsymbol{c} \in \mathbb{R}^{[p]\times[q]}$, and a bound B: is there a vertex $\boldsymbol{v}^\star \in \mathrm{O}^{\geqslant k}_{p,q}$ such that $\langle \boldsymbol{v}^\star, \boldsymbol{c}\rangle \leqslant B$?

Problem 3.43 (Orbitopal Fixed Row Sum Partitioning) Given a number of rows p, a number of columns q, an objective vector $\boldsymbol{c} \in \mathbb{R}^{[p]\times[q]}$, and a bound B: is there a vertex $\boldsymbol{v}^\star \in \mathrm{O}^{=\boldsymbol{k}}_{p,q}$ such that $\langle \boldsymbol{v}^\star, \boldsymbol{c}\rangle \leqslant B$? (Note that in contrast to Orbitopal k-Partitioning, $\boldsymbol{k}$ is a vector here; its entries are fixing the row sums independently from each other.)

We will reduce Minimum Exact Cover (exact set cover) to Orbitopal k-Partitioning.

Problem 3.44 (Minimum Exact Cover) Given is a family $\mathcal{C} = \{\mathcal{S}_1, \ldots, \mathcal{S}_n\}$ of sets $\mathcal{S}_i$ and a bound B. Let $\mathcal{S} := \cup_{\mathcal{S}_i \in \mathcal{C}} \mathcal{S}_i$. Is there a subset $\mathcal{C}^\star \subseteq \mathcal{C}$ with $|\mathcal{C}^\star| \leqslant B$, $\bigcup_{\mathcal{S}_i \in \mathcal{C}^\star} \mathcal{S}_i = \mathcal{S}$ and $\mathcal{C}^\star$ partitioning $\mathcal{S}$?

Remark 3.45 Minimum Cover is the slighty weaker problem than Minimum Exact Cover; here, $\mathcal{C}^\star$ is required to contain *at least* one set $\mathcal{S}_i$ for each element x in $\bigcup_{\mathcal{S}_i \in \mathcal{C}} \mathcal{S}_i$ $\mathcal{C}^\star$ such that $x \in \mathcal{S}_i$.

However, Minimum Cover is already strongly $\mathcal{NP}$-hard and hard to approximate. Feige [38] showed that there cannot be a $(1-\epsilon)\ln(|\mathcal{S}|)$ approximation algorithm for any $\epsilon > 0$, unless $\mathcal{NP}$ contains problems only solvable with certain super-polynomial time algorithms. Lund and Yannakakis ([80]) show that Minimum Exact Cover is at least as hard to approximate than Minimum Cover.

To reduce Minimum Exact Cover to Orbitopal k-Partitioning, we construct for $k \geqslant 2$ an objective vector $\boldsymbol{c}(\mathcal{C}, k)$ depending on the given family of sets $\mathcal{C} = \{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_n\}$ and on the bound k on the number of 1-entries in each row of the orbitope vertex. Any optimal solution to the optimization problem

$$\min(\{\langle \boldsymbol{c}(\mathcal{C}, k), \boldsymbol{x}\rangle \mid \boldsymbol{x} \in \mathbf{O}^{=k}_{p,q}\}) \tag{3.12}$$

will yield an optimal exact set cover of set

$$\mathcal{S} := \bigcup_{i=1}^{n} \mathcal{S}_i$$

and vice versa.

We define an ordered family $\widetilde{\mathcal{C}} = \mathcal{C} \times \{1,2\}$ as

$$\widetilde{\mathcal{C}} := \{(\mathcal{S}_1,1),(\mathcal{S}_1,2),(\mathcal{S}_2,1),(\mathcal{S}_2,2),(\mathcal{S}_3,1),(\mathcal{S}_3,2),\ldots\}.$$

Moreover, let α and β be two weights in $\mathbb{Q}^+$ that will later be fixed to appropriate values. We compose the objective $\boldsymbol{c}(\mathcal{C},k)$ from the following eight submatrices $\boldsymbol{M}^{1,1}$ through $\boldsymbol{M}^{4,2}$:

$$\boldsymbol{M}^{1,1} := \{-\alpha\}^{[1]\times[k]}, \qquad \boldsymbol{M}^{1,2} := \{0\}^{[1]\times\widetilde{\mathcal{C}}},$$

$$\boldsymbol{M}^{2,1} \in \{0,-\alpha\}^{\mathcal{C}\times[k]}, \quad \text{with } m^{2,1}_{i,j} := \begin{cases} -\alpha, & \text{for } j < k-1 \\ 0, & \text{otherwise,} \end{cases}$$

$$\boldsymbol{M}^{2,2} \in \{-\alpha,\alpha\}^{\mathcal{C}\times\widetilde{\mathcal{C}}}, \quad \text{with } m^{2,2}_{i,j} := \begin{cases} -\alpha, & \text{for } (i,j) \in \big\{\big(\mathcal{S}_\ell,(\mathcal{S}_\ell,1)\big),\big(\mathcal{S}_\ell,(\mathcal{S}_\ell,2)\big)\big\},\ \ell \in |\mathcal{C}| \\ \alpha, & \text{otherwise,} \end{cases}$$

$$\boldsymbol{M}^{3,1} \in \{-\alpha,0\}^{\mathcal{C}\times[k]}, \quad \text{with } m^{3,1}_{i,j} := \begin{cases} -\alpha, & \text{for } j < k \\ 0, & \text{otherwise,} \end{cases}$$

$$\boldsymbol{M}^{3,2} \in \{\beta/2,\alpha\}^{\mathcal{C}\times\widetilde{\mathcal{C}}}, \quad \text{with } m^{3,2}_{i,j} := \begin{cases} \beta/2, & \text{for } (i,j) = \big(\mathcal{S}_\ell,(\mathcal{S}_\ell,1)\big) \\ \alpha, & \text{otherwise,} \end{cases}$$

$$\boldsymbol{M}^{4,1} \in \{-\alpha,\alpha\}^{\mathcal{S}\times[k]}, \quad \text{with } m^{4,1}_{i,j} := \begin{cases} -\alpha, & \text{for } j < k \\ \alpha, & \text{otherwise,} \end{cases}$$

$$\boldsymbol{M}^{4,2} \in \{0,\alpha\}^{\mathcal{S}\times\widetilde{\mathcal{C}}}, \quad \text{with } m^{4,2}_{i,j} := \begin{cases} -\beta, & \text{for } (i,j) = \big(s,(\mathcal{S}_\ell,2)\big) \text{ with } s \in \mathcal{S}_\ell \\ \alpha, & \text{otherwise.} \end{cases}$$

Cost vector $\boldsymbol{c}(\mathcal{C},k)$ is then a composition of these submatrices:

$$\boldsymbol{c}(\mathcal{C},k) := \begin{pmatrix} \boldsymbol{M}^{1,1} & \boldsymbol{M}^{1,2} \\ \boldsymbol{M}^{2,1} & \boldsymbol{M}^{2,2} \\ \boldsymbol{M}^{3,1} & \boldsymbol{M}^{3,2} \\ \boldsymbol{M}^{4,1} & \boldsymbol{M}^{4,2} \end{pmatrix} \tag{3.13}$$

See figures 3.12 and 3.13 for an example with $\mathcal{S} = \{a,b,c,d,e,f\}$, set system $\mathcal{C} = \big\{\{a\},\{b\},\{c,e\},\{a,b,c\},\{a,d\},\{d,e\}\big\}$ and $k=3$. Note that in each solution $\boldsymbol{v} \in \mathrm{O}^{=k}_{|\mathcal{S}|+2|\mathcal{C}|+1,k+2|\mathcal{C}|}$ subject to a cost vector (3.12), we will interpret a 1-entry in the part of the solution corresponding to submatrix $\boldsymbol{M}^{4,2}$ as a covering of the element corresponding to that row. On the other hand, a 1-entry in the part of the solution that corresponds to $\boldsymbol{M}^{3,2}$ indicates which subset is responsible for this covering.

Proposition 3.46 Minimum Exact Cover *can be reduced to* Orbitopal k-Partitioning *for $k \geqslant 2$, $p \geqslant 4$ and $q \geqslant k+2$ in polynomial time.*

Proof. We choose α big enough such that any $\boldsymbol{c}(\mathcal{C},k)$-minimal vertex $\boldsymbol{v}^\star \in \mathrm{O}^{=k}_{p,q}$ has 1 at $-\alpha$-positions and 0 at α-positions, and define $\beta := 1$. Matrix $\boldsymbol{c}(\mathcal{C},k)$ as defined

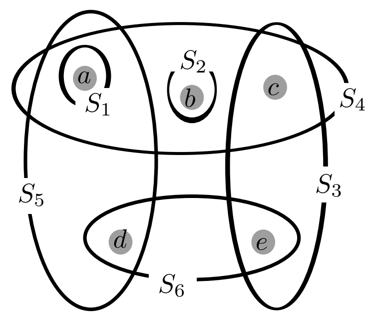

Figure 3.12: Example instance for minimum set cover with sets S_1 through S_6 over a ground set $S = \{a, b, c, d, e\}$.

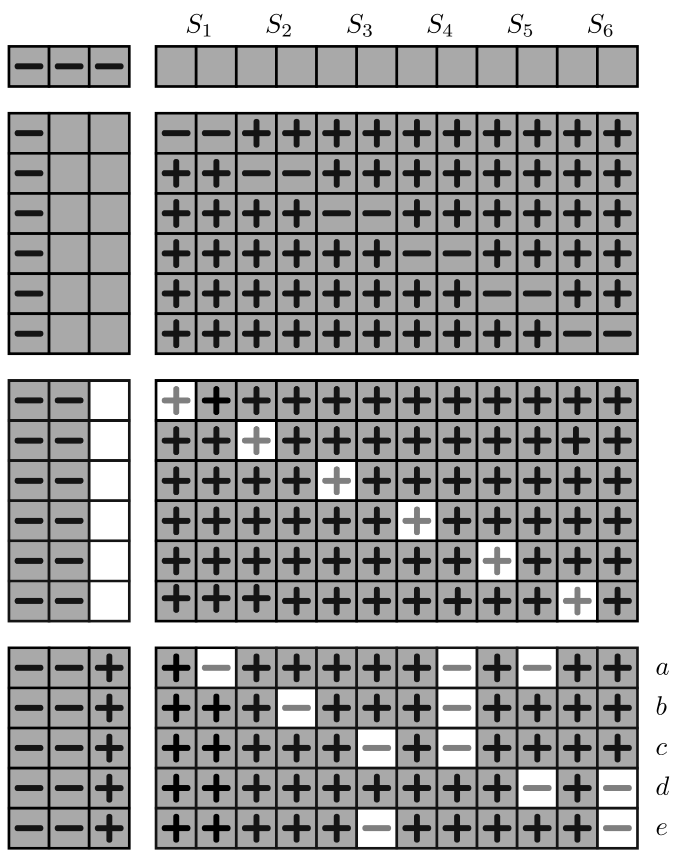

Figure 3.13: Objective vector associated with the instance of a minimum cover problem from figure 3.12. We are minimizing over $\mathrm{O}_{p,q}^{=3}$ here. Black "+" correspond to components $+\alpha$, black "−" correspond to $-\alpha$, gray "+" correspond to $+\beta/2$, and gray "−" correspond to $-\beta$. White cells contain 0-entries. α is chosen such that all components of a solution corresponding to cells with gray background are a priori fixed to 0 and 1, respectively, as described in the proof of Proposition 3.46.

in (3.13) can obviously be generated in time polynomial in the size of the instances of MINIMUM EXACT COVER.

Hence, in any of the last $|\mathcal{S}|$ rows of $\boldsymbol{v}^\star$, there remains exactly one 1 to be set to one of the -1-positions. However, this requires a 1 at the respective $+1/2$-position in $\boldsymbol{M}^{3,2}$.

A solution $\boldsymbol{v}^\star$ constructed in such way has value

$$\langle \boldsymbol{c}(\mathcal{C},k), \boldsymbol{v}^\star \rangle = \tfrac{K}{2} - |\mathcal{S}|(k\alpha - \alpha + 1) - |\mathcal{C}|\alpha(2k-1) - k\alpha$$

where K is the number of 1s in the part of the solution vector corresponding to submatrix $\boldsymbol{M}^{3,2}$; hence, K is the number of chosen subsets. The minimality of $\langle \boldsymbol{v}^\star, \boldsymbol{c}(\mathcal{C},k)\rangle$ guarantees the minimality of the set cover and the constant row sum ensures that all elements are covered by exactly one subset.

On the other hand, if one has a solution for an instance of MINIMUM EXACT COVER, it is easy to construct a solution for the associated instance of ORBITOPAL k-PARTITIONING from it.

Hence, we can cover $\mathcal{S}$ with K or less sets from $\mathcal{C}$ if and only if the associated optimization problem (3.13) has value $\frac{K}{2} - |\mathcal{S}|(k\alpha - \alpha + 1) - |\mathcal{C}|\alpha(2k-1) - k\alpha$ or smaller. □

Note that problems 3.40, 3.41, 3.42, and 3.43 are all in $\mathcal{NP}$. This yields

Corollary 3.47 ORBITOPAL k-PARTITIONING *is $\mathcal{NP}$-complete for all problems with $k \geqslant 2$, $p \geqslant 4$ and $q \geqslant k+2$.*

The same cost function $\boldsymbol{c}(\mathcal{C},k)$ as defined above for k-partitioning orbitopes can be used to reduce MINIMUM COVER to ORBITOPAL k-PACKING, similarly to above. We obtain

Proposition 3.48 ORBITOPAL k-PACKING *is $\mathcal{NP}$-complete for all problems with $k \geqslant 2$, $p \geqslant 4$ and $q \geqslant k+2$.*

Proof. To construct an optimal solution $\boldsymbol{v}^\star$ for cost vector $\boldsymbol{c}(\mathcal{C},k)$, we proceed as described in the proof for the partitioning case. Since $\langle \boldsymbol{v}^\star, \boldsymbol{c}(\mathcal{C},k)\rangle$ has to be as small as possible, we obtain that in each of the last $|\mathcal{S}|$ rows, exactly k entries must be set to 1. □

Exploiting the isomorphy between $\mathrm{O}_{p,q}^{\leqslant k}$ and $\mathrm{O}_{p,q}^{\geqslant q-k}$ (Observation 3.12), we obtain

Corollary 3.49 ORBITOPAL k-COVERING *is $\mathcal{NP}$-complete for all problems with $k \geqslant 2$, $p \geqslant 4$ and $q \geqslant k+2$.*

Remark 3.50 It is also possible to reduce WEIGHTED SET COVER to ORBITOPAL k-PACKING, ORBITOPAL k-PARTITIONING, and ORBITOPAL k-COVERING. Differing from MINIMUM (EXACT) COVER, in WEIGHTED (EXACT) SET COVER, there are weights $w_i \in \mathbb{Q}^+$ associated with the sets $\mathcal{S}_i \in \mathcal{C}$. The task is to find a set cover with minimal total weight. Hence, WEIGHTED SET COVER with weights $w_i = 1$ for all $i \in n$ becomes MINIMUM (EXACT) COVER. For WEIGHTED (EXACT) SET COVER, one can use essentially the same reductions as above with the following adaptions in the respective cost vector:

- In $\boldsymbol{M}^{3,2}$, entry $\beta/2$ is replaced by weight $w_i/2$ associated with the respective set $\mathcal{S}_i$.
- In $\boldsymbol{M}^{4,2}$, the entry $-\beta$ is chosen as $-\max(\{w_i \mid i \in [n]\})$.

Remark 3.51 The complexity results concerning partitioning orbitopes imply that ORBITOPAL FIXED ROW SUM PARTITIONING is also $\mathcal{NP}$-hard, as long as $k_i \geqslant 2$ for all $k_i \in \boldsymbol{k}$, $i \in [p]$. This can also be shown directly, since it is not too difficult to modify the cost vector as defined in (3.13) such that in each row of submatrix $\boldsymbol{M}^{3,2}$, one 1-entry can be positioned, and in each row of matrix $\boldsymbol{M}^{4,2}$, at least one 1 can be positioned in each row.

In contrast to the results obtained so far, optimization over all the aforementioned orbitopes can be done in polynomial time as long as the number of columns or rows are not part of the input.

Let us start with a fixed number of columns. The number of possible split patterns is bounded by $\mathcal{O}(p^q)$ (see Observation 3.9). Therefore, if we can show that it is possible to compute in a time that is polynomial in p for a given split pattern $\boldsymbol{\sigma} \in [p+1]^{q-1}$ and an objective vector $\boldsymbol{c} \in \mathbb{Q}^{[p]\times[q]}$ a vertex of one of the orbitopes in focus that produces an objective value which is minimal among all vertices with the same split pattern, we are finished.

How can this solution be constructed? The main idea is to define for each row a family of ranges $\mathcal{T}$ similar to the family used in the construction of vertices from page 50.

For this, we define for row ℓ a sequence (s_j) as follows: $s_1 := 1$ and, for $t > 1$,

$$s_t = \min(\{j \in [s_{t-1}+1..q-1] \mid \sigma_j \leqslant \ell\})$$

We add all ranges $[s_j..s_{j+1}]$ to $\mathcal{T}$. For each range, we have to distinguish three cases:

(i) $\sigma_{s_{j+1}} = \ell$; then we set all entries $v_{\ell,k}$ with $k \in [s_j..s_{j+1}]$ to 1.
(ii) $\sigma_{s_j-1} = \ell$; then we set all entries $v_{\ell,k}$ with $k \in [s_j..s_{j+1}]$ to 0.
(iii) Otherwise, we can either set the entries $v_{\ell,k}$ with $k \in [s_j..s_{j+1}]$ either all to 1 or all to 0.

For full orbitopes with no restriction on the number of 1s per row, we would choose $v_{\ell,k} = 1$ for all $k \in [s_j..s_{j+1}]$ if and only if $\sum_{k\in[s_j..s_{j+1}]} c_{\ell,k} < 0$. This is different for k-packing, k-partitioning and k-covering orbitopes. Here, the ranges in the third case require the solution of a subset sum or a knapsack problem for row ℓ. However, these subproblems can be solved by enumeration in time $\mathcal{O}(2^q)$ which is constant in the size of the input. Hence, we obtain the following proposition:

Proposition 3.52 *Optimization over k-partitioning, k-packing, k-covering and fixed row sum orbitopes can be done in polynomial time if the number of columns is fixed.*

Next, we assume that the number of rows is part of the input. In this case, we will generate vertex $\boldsymbol{v}$ of a k-packing, k-partitioning, k-covering or a fixed row sum orbitope column by column. There are 2^p possibilities to choose from when building column $\boldsymbol{v}_{*,j}$. However, we have to exclude those columns that are lexicographically larger than column $\boldsymbol{v}_{*,j-1}$, as well as those columns that violate the row sum conditions.

For this purpose, we define an acyclic digraph $D = (\mathcal{V}, \mathcal{A})$. In what follows, we will show what this digraph looks like in the case of fixed row sum orbitope $\mathrm{O}^{=\boldsymbol{k}}_{p,q}$. (The construction can easily be adapted to other orbitopes.)

The node set is defined as

$$\mathcal{V} := [q] \times \{0,1\}^p \times [k]_0^p \ \cup\ \{s,t\},$$

where $k := \max(\{k_i \mid i \in p\})$ is the maximal entry in the vector $\boldsymbol{k}$ defining the row sums. Node $(j, \boldsymbol{y}, \boldsymbol{z})$ will be identified with the choice of vector $\boldsymbol{y}$ as column $\boldsymbol{v}_{*,j}$ in case of $\sum_{\ell=1}^{j} v_{i,\ell} = z_i$ for all $i \in [p]$.

The arc set $\mathcal{A}$ contains the following arcs:

$$\begin{array}{ll} \big(s,(1,\boldsymbol{y},\boldsymbol{y})\big) & \text{for all } \boldsymbol{y} \in \{0,1\}^p \text{ and} \\ \big((j,\boldsymbol{y},\boldsymbol{z}),(j+1,\boldsymbol{y}',\boldsymbol{z}+\boldsymbol{y}')\big) & \text{for all } j \in [q-1],\ \boldsymbol{y},\boldsymbol{y}' \in \{0,1\}^p \text{ with } \boldsymbol{y} \succcurlyeq \boldsymbol{y}', \\ & \text{and } \boldsymbol{z} \in [k]_0^p. \end{array}$$

Additionally, there exist arcs that are specific for the orbitope. For the fixed row sum orbitope $\mathrm{O}_{p,q}^{=\boldsymbol{k}}$, these are the following ones:

$$\big((q,\boldsymbol{y},\boldsymbol{z}),t\big) \quad \text{for all } \boldsymbol{y} \in \{0,1\}^p \text{ and } \boldsymbol{z} = \boldsymbol{k}$$

It is obvious that there is a one-to-one-correspondence between the s-t-paths in D and the vertices of $\mathrm{O}_{p,q}^{=\boldsymbol{k}}$. Let now $\boldsymbol{c} \in \mathbb{R}^{[p]\times[q]}$ be the given cost vector. We will solve the minimization problem

$$\min(\{\langle \boldsymbol{x},\boldsymbol{c}\rangle \mid \boldsymbol{x} \in \mathrm{O}_{p,q}^{=\boldsymbol{k}}\}) \tag{*}$$

by assigning weights

$$\begin{array}{ll} \langle c_{*,1},\boldsymbol{y}\rangle & \text{to each arc } \big(s,(1,\boldsymbol{y},\boldsymbol{y})\big) \in \mathcal{A}, \\ \langle c_{*,j+1},\boldsymbol{y}'\rangle & \text{to each arc } \big((j,\boldsymbol{y},\boldsymbol{z}),(j+1,\boldsymbol{y}',\boldsymbol{z}+\boldsymbol{y}')\big) \in \mathcal{A}, \text{ and} \\ 0, & \text{to each arc } \big((j,\boldsymbol{y},\boldsymbol{z}),t\big) \in \mathcal{A}. \end{array}$$

A shortest s-t-path in D corresponds then to a minimal solution for (*).

The cardinality of $\mathcal{V}$ is within $\mathcal{O}(q2^pk^p)$, and the cardinality of $\mathcal{E}$ is within $\mathcal{O}(q2^{2p}k^{2p})$. It costs $\mathcal{O}(|\mathcal{V}|+|\mathcal{E}|)$ to bring $\mathcal{V}$ into topological order (which is possible since D is acyclic); a shortest path can then be found in time $\mathcal{O}(|\mathcal{E}|)$.

Since k-partitioning, k-packing and k-covering orbitopes can be treated similarly, we obtain the following proposition:

Proposition 3.53 *If the number of rows is not part of the input, one can optimize over k-partitioning $\mathrm{O}_{p,q}^{=k}$, k-packing $\mathrm{O}_{p,q}^{\leqslant k}$, k-covering $\mathrm{O}_{p,q}^{\geqslant k}$ and fixed row sum orbitopes $\mathrm{O}_{p,q}^{=\boldsymbol{k}}$ in linear time $\mathcal{O}(q)$.*

3.2.1.3 Dimension

Before we focus on the facial structure of the orbitopes, we briefly study their dimension. As the characterization of vertices (see page 45) already indicated, only full orbitopes are full dimensional.

Lemma 3.54 *The full orbitope $\mathbf{O}_{p,q}$ is full dimensional, i.e. it has dimension pq.*

Proof. For any $(i,j) \in [p]\times[q]$, we define a vector $\boldsymbol{v}^{i,[j]}$ with components

$$\boldsymbol{v}^{i,[j]}_{s,t} := \begin{cases} 1, & \text{if } s = i \text{ and } t \in [j] \\ 0, & \text{otherwise.} \end{cases}$$

Together with $\mathbb{0}_{p,q}$, this gives a set of $pq+1$ affinely independent points that are vertices of the orbitope. ☐

Lemma 3.55 *Writing q as $mk+n$ with $m,n \in \mathbb{N}$ and $n < k$, the packing orbitope $\mathrm{O}_{p,q}^{\leqslant k}$ has dimension*

$$pq - \frac{m(m-1)}{2}k - mn.$$

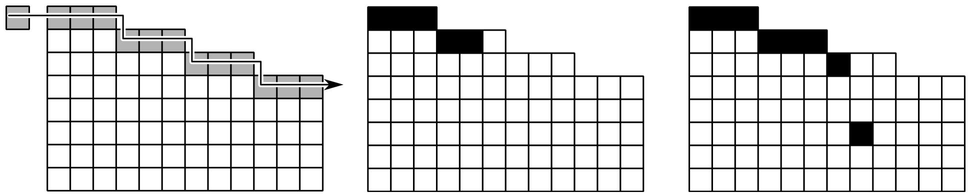

Figure 3.14: Examples concerning the packing orbitope $\mathrm{O}_{p,q}^{\leqslant 3}$. Set $\mathcal{I}$ is marked gray in the left image. The middle image shows vertex $\boldsymbol{u}(2,5)$, while the right image shows vertex $\boldsymbol{u}(6,8)$.

Proof. Let $\mathcal{S} := \{(i,j) \in [p] \times [q] \mid 1 \leqslant j \leqslant \min(\{q, ik\})\}$. In each vertex $\boldsymbol{v}$ of $\mathrm{O}_{p,q}^{\leqslant k}$, the upper right stepped triangle is fixed to 0, that is:

$$v_{i,j} = 0 \qquad \text{for all } (i,j) \in ([p] \times [q]) \setminus \mathcal{S}. \tag{*}$$

This makes $\frac{m(m-1)}{2}k + mn$ equations.

We define now the set

$$\mathcal{I} := \{(i,j) \in [p] \times [q] \mid j \in [((i-1)k+1)..ik]\} \,\cup\, \{(0,0)\}$$

(see figure 3.14 left). Since there is for each column in $[q]$ at most one element in $\mathcal{I}$, we can order set $\mathcal{I}$ columnwise. This order is denoted by $<_{\mathcal{I}}$.

- For each $(s,t) \in \mathcal{I}$, we define a vector $\boldsymbol{u}(s,t) \in [p] \times [q]$ with components $u_{i,j}(s,t) = 1$ for all $(i,j) \leqslant_{\mathcal{I}} (s,t)$, $(i,j) \in \mathcal{I} \setminus \{(0,0)\}$. Otherwise, vector $\boldsymbol{u}(s,t)$ has only 0-entries. (See figure 3.14 middle.) Note that $\boldsymbol{u}(0,0) = \mathbb{0}_{p,q}$ by this definition.
- For each $(s,t) \in \mathcal{S} \setminus \mathcal{I}$, we define vector $\boldsymbol{u}(s,t)$ as $\mathbb{e}^{s,t} + \boldsymbol{u}(s', j-1)$, where s' is choosen such that $(s', j-1) \in \mathcal{I}$. (See figure 3.14 right.)

We obtain by this a set of $pq - \frac{m(m-1)}{2}k - mn + 1$ affinely independent vertices $\boldsymbol{u}(i,j)$ of $\mathrm{O}_{p,q}^{\leqslant k}$. □

For the vertices of the *k-partitioning orbitope*, the equation set (*) is also valid. However, we find more valid equations.

- In the first row of each vertex $\boldsymbol{v}$, the first k entries are fixed to 1:

$$x_{1,j} = 1 \text{ for all } j \in [k]. \tag{**}$$

- In the second row, it holds that

$$x_{2,k-j+1} + x_{2,k+j} = 1 \tag{***}$$

  for all $j \in [k]$.
- For every row $i \in [p]$, it holds that $\sum_{t \in [q]} x_{i,t} = k$. (In fact, this follows for rows 1 and 2 already from equations (*), (**), and (***)).

Therefore, at least $\frac{m(m-1)k}{2} + mn + 2k + p - 2$ equations are holding for $\mathrm{O}_{p,q}^{=k}$, where m and n are defined as in Lemma 3.55. This leads to the following observation:

Observation 3.56 Let $q = mk + n$ with $m, n \in \mathbb{N}$ and $n < k$. Then the dimension of the partitioning orbitope $\mathrm{O}_{p,q}^{=k}$ is bounded from above by $pq - \frac{m(m-1)k}{2} - mn - 2k - p + 2$.

It is an open question whether this bound is tight or not.

For the *covering orbitope* $\mathrm{O}_{p,q}^{\geqslant k}$, the first k entries in the first row are fixed to 1. This gives the following set of equations:

$$x_{1,j} = 1 \text{ for all } j \in [k].$$

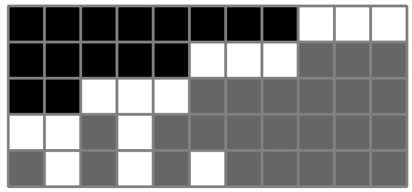

Figure 3.15: For this example concerning covering orbitopes, $p = 5$, $q = 11$ and $k = 8$. Therefore, $k = 2(q - k) + 2$, $m = 2$ and $n = 2$ and the black entries in rows 1 through $m + 1 = 3$ are fixed to 1, while the first entry in the fourth row can be 0. Gray positions are not necessarily fixed to 1.

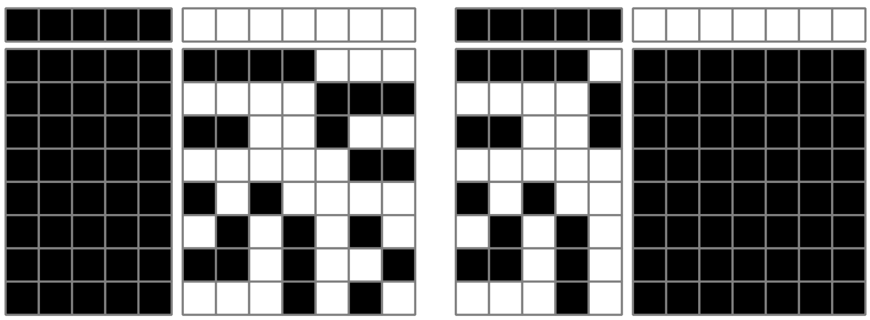

Figure 3.16: Constructing affinely independent vertices of the covering orbitope.

For the second row, one has to distinguish two cases: Either, $2k \leqslant q$ or $2k > q$. If $2k \leqslant q$, then we can place k 1-entries into columns $[k + 1..2k]$, which means that columns $[k]$ may contain 0-entries. However, if $2k > q$, we are forced to place into columns $[k]$ at least $2k - q$ 1-entries, starting from column 1. This implies that in this case, the equations

$$x_{2,j} = 1 \text{ for all } j \in [2k - q]$$

are holding. This can be iterated; if for the ith row it holds that $ik > (i - 1)q$, then we are forced to place $ik - (i - 1)q$ 1-entries into the first columns. Figure 3.15 shows an example for $k = 8$ and $q = 11$.

Hence, we obtain the following observation:

Observation 3.57 If $q \geqslant 2k$, then we get k equations. If $q < 2k$, then we decompose k into $m(q - k) + n$ with $m, n \in \mathbb{N}$ and $n < q - k$. In this case, there are at least $\frac{m(m+1)}{2}(q - k) + (m + 1)n$ equations valid for the covering orbitope $\mathrm{O}^{\geqslant k}_{p,q}$.

Lemma 3.58 *If $2k \leqslant q$, then the dimension of $\mathrm{O}^{\geqslant k}_{p,q}$ is $pq - k$.*

Proof. Define a vector

$$\boldsymbol{w}(\boldsymbol{v}) := \begin{pmatrix} \mathbb{1}^{1,k}_{v} & \mathbb{0}^{1,q-k}_{v} \\ \mathbb{1}^{p-1,k}_{v} & \boldsymbol{v} \end{pmatrix}$$

in $\{0,1\}^{[p]\times[q]}$ that is composed from a row vector $\mathbb{1}_{1,k}$ filled with k 1-entries, a row vector $\mathbb{0}_{1,q-k}$ filled with $q - k$ 0-entries, a vector $\mathbb{1}_{p-1,k}$ with $p - 1$ rows and k columns filled with 1s only, and an arbitrary vertex $\boldsymbol{v}$ of the full orbitope $\mathbf{O}_{p-1,q-k}$. Then $\boldsymbol{w}(\boldsymbol{v})$ is a vertex of $\mathrm{O}^{\geqslant k}_{p,q}$. Similarly,

$$\widetilde{\boldsymbol{w}}(\boldsymbol{u}) := \begin{pmatrix} \mathbb{1}^{1,k}_{v} & \mathbb{0}^{1,q-k}_{v} \\ \boldsymbol{u} & \mathbb{1}_{p-1,q-k} \end{pmatrix}$$

is also a vertex of $\mathrm{O}_{p,q}^{\geqslant k}$, where $\boldsymbol{u} \in \mathbf{O}_{p-1,k}$. (See figure 3.16 for examples.)

To construct a set of affinely independent vertices of $\mathrm{O}_{p,q}^{\geqslant k}$, we will now adopt the construction of vertices of the full orbitope from the proof of lemma 3.54. Consider the vertices $\boldsymbol{w}(\boldsymbol{v}^{i,[j]})$ and $\widetilde{\boldsymbol{w}}(\boldsymbol{u}^{i',[j']})$ with $(i,j) \in [p-1] \times [q-k]$ and $(i',j') \in [p-1] \times [k]$, and add to these for $j \in [q-k]_0$ the set of vertices

$$\begin{pmatrix} \mathbb{1}_{1,k} & \boldsymbol{v}^{1,[j]} \\ \mathbb{1}_{p-1,k} & \mathbb{0}_{p-1,q-k} \end{pmatrix},$$

where $\boldsymbol{v}^{1,[j]} \in \mathbf{O}_{1,q-k}$ and $\boldsymbol{v}^{1,[0]}$ is identified with $\mathbb{0}_{1,q-k}$. This gives a set of

$$(p-1)k + (p-1)(q-k) + q - k + 1 = pq - k + 1$$

vertices of $\mathrm{O}_{p,q}^{\geqslant k}$ which are obviously affinely independent. □

The dimension of $\mathrm{O}_{p,q}^{\geqslant k}$ for $2k > q$ is an open question.

3.2.1.4 Facial Structure

3.2.1.4.7 (k)Packing and (k)Partitioning Orbitopes

For $k = 1$, Kaibel and Pfetsch gave a complete linear description for k-packing and k-partitioning orbitopes. For this, they introduced a class of inequalities valid for packing and partitioning orbitopes.

Definition 3.59 (Shifted column inequalities) Let $(\alpha, \zeta) \in [p] \times [q]$ with $\alpha \leqslant \zeta$ and let $\mathcal{C}$ be a set of $\alpha - \zeta + 1$ column indices $c_i \in [q]$ with the property that $1 \leqslant c_1 \leqslant c_2 \leqslant \ldots \leqslant c_{\alpha-\zeta+1} < \zeta$. Moreover, let k some column such that $\zeta \leqslant k \leqslant \min(\{q, \zeta\})$. Then

$$\sum_{j=k}^{\min(\{q,\alpha\})} x_{\alpha,j} - \sum_{i=1}^{\alpha-\zeta+1} x_{i+c_i-1,c_i} \leqslant 0$$

is called a *shifted column inequality.*

Proposition 3.60 ([65]) *Let $k = 1$. The nonnegativity constraints, the shifted column inequalities and the row sum inequalities*

$$\sum_{j=1}^{q} x_{\ell,j} \leqslant 1 \text{ for all } \ell \in [p]$$

provide a complete description for packing orbitopes.
Replacing the row sum inequalities by equations, we obtain a complete description for partitioning orbitopes.

For $k \geqslant 2$, no "nice" linear description can be expected, since optimization is $\mathcal{NP}$-hard.

However, it is clear that if $\boldsymbol{a}\boldsymbol{x} \leqslant b$ defines a facet of $\mathrm{O}_{p,q}^{\leqslant k}$, then reading the columns of $\boldsymbol{a}$ backwards and multiplying them by -1 must give the normal of a facet of $\mathrm{O}_{p,q}^{\geqslant q-k}$; this follows from the isomorphy of both. Similarly, any facet of $\mathrm{O}_{p,q}^{=k}$ corresponds to a facet of $\mathrm{O}_{p,q}^{=q-k}$.

3.2.1.4.8 (k)Covering Orbitopes

Optimization is $\mathcal{NP}$-hard for $k \geqslant 2$, $p \geqslant 4$, and $q \geqslant k+2$. So we cannot expect a "nice" linear description for general p and q. However, if $q = k+1$, the covering orbitope $\mathrm{O}^{\geqslant k}_{p,k+1}$ becomes isomorphic to the packing orbitope $\mathrm{O}^{\leqslant}_{p,k+1}$, as Observation 3.11 shows. Therefore, one can from Proposition 3.60 derive a linear description for $\mathrm{O}^{\geqslant k}_{p,k+1}$. This is in particular possible for $k=1$.

Proposition 3.61

1. *The following inequalities provide a complete description for* $\mathrm{O}^{\geqslant}_{p,2}$.
   (a) Trivial (in)equalities:

$$\begin{aligned} x_{i,j} &\leqslant 1 && \textit{for all } 2 \leqslant i \leqslant p, 1 \leqslant j \leqslant 2, \\ -x_{1,2} &\leqslant 0 && \textit{and} \\ x_{1,1} &= 1 \end{aligned}$$

   (b) Row-sum inequalities:

$$-x_{i,1} - x_{i,2} \leqslant -1 \quad \textit{for all } 2 \leqslant i \leqslant p$$

   (c) Column-sum inequalities:

$$-x_{k,1} + \sum_{i=1}^{k-1} x_{i,2} \leqslant k-2 \quad \textit{for all } 2 \leqslant k \leqslant p$$

2. *All of these inequalities, but not the equation $x_{1,1} = 1$, define facets.*

Proof. Follows from Proposition 3.60 (and can also be shown by total unimodularity of the constraints matrix.) For the characterization of facets (in the packing case), see [65]. □

Note that all valid equations for the orbisack are also valid for the covering orbitope $\mathrm{O}^{\geqslant}_{p,2}$.

3.2.1.4.9 Full Orbitopes

Not much is known about the facial structure of the full orbitope. This is astounding, because it is possible to optimize polynomially over full orbitopes; therefore, a "nice" linear description seems not to be a priori excluded. However, our computer experiments using the software package `polymake` ([47]) exhibit a rather complicated facial structure.

As examples, we show the computed facets of $\mathbf{O}_{3,4}$, $\mathbf{O}_{4,4}$ and of $\mathbf{O}_{5,3}$, in figure 3.17, 3.18, and 3.19, respectively. All inequalities are in the form $\boldsymbol{ax} \leqslant b$ with right-hand side b at the top right corner of each facet; red entries mark negative coefficients, blue coefficients are positive.

As one can see, the linear description of $\mathbf{O}_{p,q}$ contains all lifted facets from $\mathbf{O}_{s,q}$, $s < p$. This is not by chance.

Lemma 3.62 *Let for $p > s \geqslant 1$ the lifting $\phi : \mathbb{R}^{s,q} \to \mathbb{R}^{p,q}$ be defined by*

$$\phi(\boldsymbol{x})_{i,j} = \begin{cases} x_{i,j}, & \textit{if } (i,j) \in [s] \times [q] \\ 0, & \textit{otherwise} \end{cases} \qquad \textit{for all } (i,j) \in [p] \times [q]$$

Then inequality $\boldsymbol{ax} \leqslant b$ with $\boldsymbol{a} \in \mathbb{R}^{[s]\times[q]}$ is facet defining for $\mathbf{O}_{s,q}$, if and only if inequality $\phi(\boldsymbol{a})\boldsymbol{x} \leqslant b$ is facet defining for $\mathbf{O}_{p,q}$.

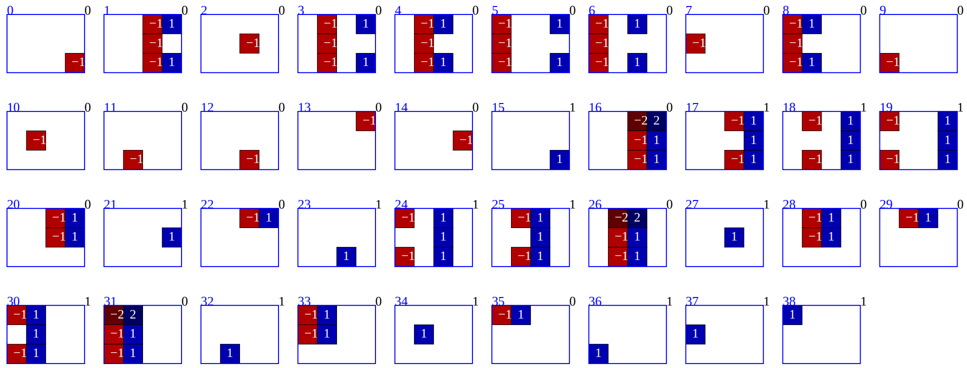

Figure 3.17: The facets of $\mathbf{O}_{3,4}(\mathfrak{S}_4)$.

Proof. $\textcircled{\Rightarrow}$ Since $\boldsymbol{ax} \leqslant b$ is facet defining, there exists a subset $\mathcal{C}$ of vertices of $\mathbf{O}_{s,q}$ containing sq affinely independent vectors active for $\boldsymbol{ax} \leqslant b$. From $\mathcal{C}$, we derive a subset $\mathcal{C}'$ of vertices of $\mathbf{O}_{p,q}$ as follows: For each $\boldsymbol{v} \in \mathcal{C}$, $\mathcal{C}'$ contains vector $\phi(\boldsymbol{v})$. Moreover, pick out some arbitrary vector $\boldsymbol{v}'^{\star} \in \mathcal{C}'$. For any $(i,j) \in [s+1..p] \times [q]$, we add to $\mathcal{C}'$ the modification of $\boldsymbol{v}'^{\star}$ in all components $\widetilde{\boldsymbol{v}}'^{\star}_{i,\ell} := 1$, $\ell \in [j]$. Obviously, any of these modifications is a vertex of $\mathbf{O}_{p,q}$.
Hence, $\mathcal{C}'$ contains $qs + (p-s)q = pq$ affinely independent vertices that are active for inequality $\phi(\boldsymbol{a})\boldsymbol{x} \leqslant b$.
$\textcircled{\Leftarrow}$ We use the orthogonal projection $\psi : \mathbb{R}^{[p]\times[q]} \longrightarrow \mathbb{R}^{[s]\times[q]}$ that maps $\boldsymbol{x}$ to $\boldsymbol{x}_{[s]\times[q]}$. $\psi(\mathtt{F})$ is the face of $\mathbf{O}_{s,q}$ defined by $\boldsymbol{ax} \leqslant b$. We have to show that $\psi(\mathtt{F})$ is facet defining. The face lattice of $\mathbf{O}_{p,q}$ and the sub-face lattice of $\mathbf{O}_{s,q}$ consisting of the ψ-compatible faces are isomorphic (see for instance [61]). Since $\mathtt{F}$ is obviously ψ-compatible, $\mathtt{F}$ is either a facet of $\mathbf{O}_{p,q}$ or $\mathbf{O}_{p,q} \subseteq \mathtt{F}$; however, the latter is not possible since $\mathbf{O}_{p,q}$ is full dimensional (Lemma 3.54). □

Corollary 3.63 *The linear description of orbitope $\mathbf{O}_{p,q}$ contains the lifted facets of $\mathbf{O}_{s,q}$ for any $1 \leqslant s < p$. In particular, the facet set of the orbisack $\mathbf{O}_{p,2}$ contains all lifted facets of any smaller orbisack $\mathbf{O}_{s,2}$, $1 \leqslant s < p$.*

Observation 3.64 Similarly as for k-packing and k-covering orbitopes (observation 3.11), if $\boldsymbol{ax} \leqslant b$ defines a facet of $\mathbf{O}_{p,q}$, then reading the columns of $\boldsymbol{a}$ backwards and multiplying all entries by -1 gives the normal of another facet of $\mathbf{O}_{p,q}$. This follows readily from the affine transformation from Definition 3.10.

Observation 3.65 It is also possible to lift facets by inserting empty columns instead of rows as in Lemma 3.62. The lifted facets are clearly valid, but do not have to be facet defining, as the computational experiments show.

3.2.2 Cyclic group

For this subsection, the full cyclic group $\mathfrak{C}_q$ is operating on the columns of the vertices of each considered orbitope. As we will see, much more questions than for the symmetric group are open. We will collect the few known facts.

In the case of two columns, the operation of cyclic and full symmetric group becomes the same. Therefore, everything that can been said about orbitopes with two columns over full symmetric group also holds for the cyclic group. This is in particular true for the orbisack (see chapter 4). We will in the following assume that $q > 2$.

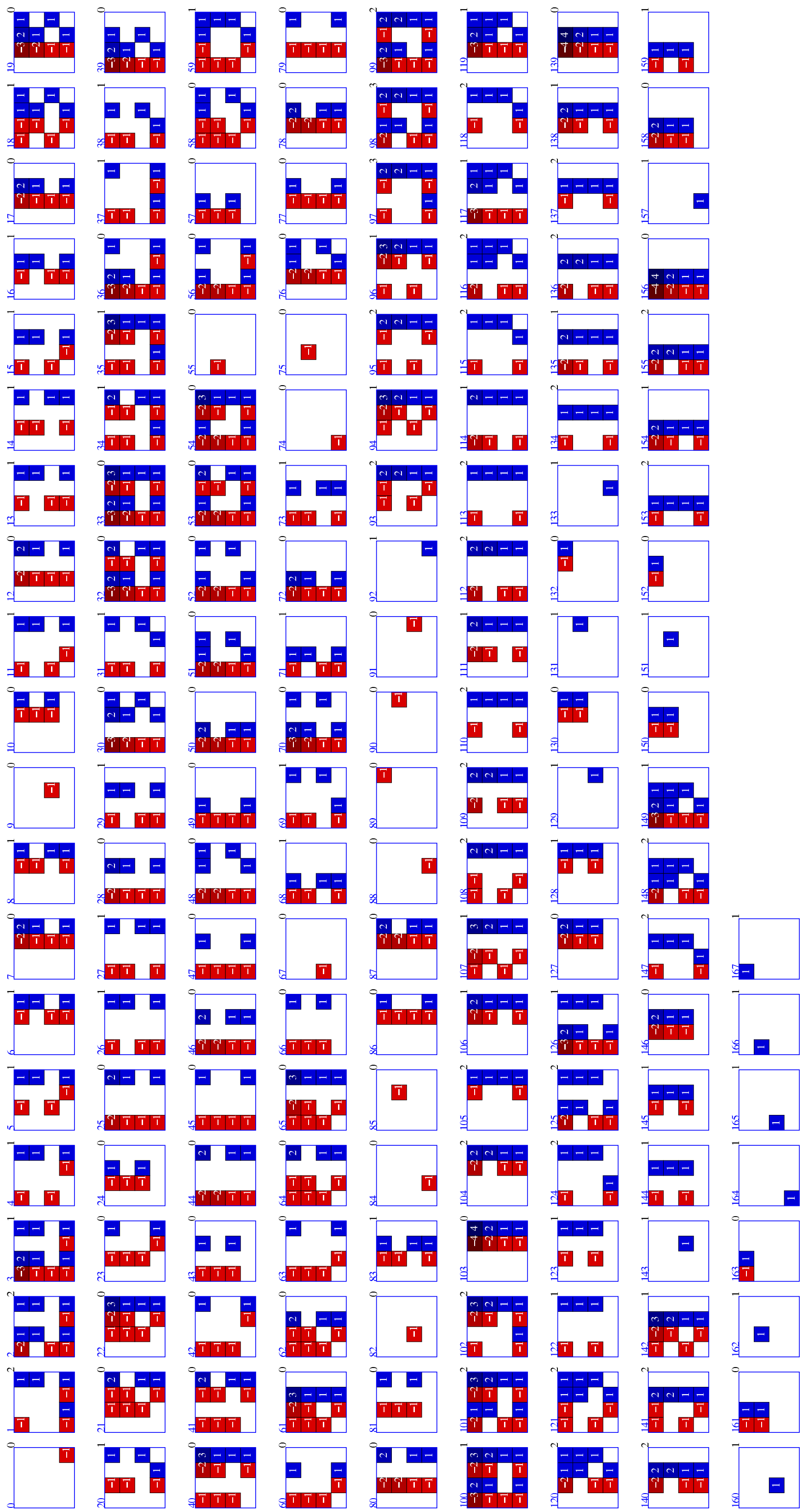

Figure 3.18: The facets of $\mathbf{O}_{4,4}(\mathfrak{S}_4)$.

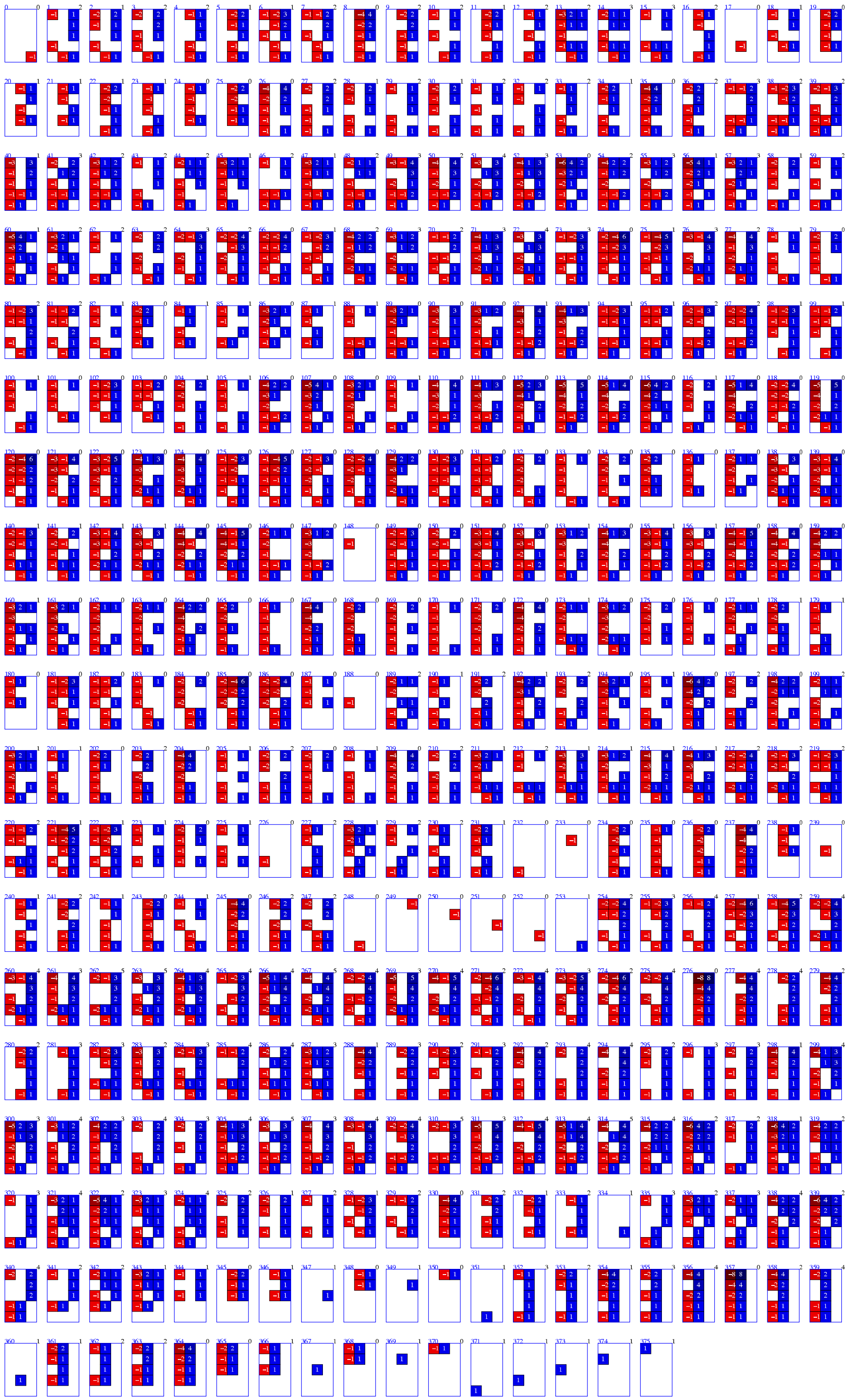

Figure 3.19: The facets of $\mathbf{O}_{5,3}(\mathfrak{S}_3)$.

3.2.2.1 Characterization of Vertices

For the cyclic group and $k = 1$, the vertices of the packing orbitope $\mathrm{O}^{\leqslant}_{p,q}(\mathfrak{C}_q)$ and the partitioning orbitope $\mathrm{O}^{=}_{p,q}(\mathfrak{C}_q)$ can be easily described. For partitioning orbitopes, the lexicographically maximal column in each vertex is unique, and for the packing case it is either unique or the matrix is $\mathbb{0}_{[p]\times[q]}$, i.e. all columns are equal $\mathbb{0}_p$ then. A vector in $\mathcal{M}_{p,q}$ is therefore vertex of the packing (partitioning) orbitope with respect to the cyclic group, if and only if

(i) its lexicographic largest column is the first one, and
(ii) every row contains at most (exactly) one 1-entry.

In particular, the vertices of $\mathrm{O}^{=}_{p,q}(\mathfrak{C}_q)$ are exactly the vectors $\boldsymbol{v}$ with $v_{1,1} = 1$ and row sum 1 for each row.

However, other (or no) restrictions on the number of 1s in each row seem to make things much more complicated. For the full orbitope $\mathbf{O}_{1,q}(\mathfrak{C}_q)$ with cyclic column permutations, the vertex set is already obviously quite intricate. In the following, we collect observations about these vertices to get nearer to a characterization.

Observation 3.66 (i) If a vertex $\boldsymbol{v} \in \mathbf{O}_{p,q}(\mathfrak{C})$ has last column $\mathbb{1}_p$, then all columns are $\mathbb{1}_p$. (ii) Similarly, if $\boldsymbol{v}_{*,1} = \mathbb{0}_p$, then $\boldsymbol{v}_{*,j} = \mathbb{0}_p$ for all $j \in [q]$.

Proof. (i) $\boldsymbol{v}_{*,q}$ must be lexicographically smaller or equal than $\boldsymbol{v}_{*,1}$, so $\boldsymbol{v}_{*,1} = \mathbb{1}_p$. But $\boldsymbol{v}_{*,1}$ must also be lexicographically smaller or equal than $\boldsymbol{v}_{*,2}$; otherwise $(\boldsymbol{v}_{*,q}, \boldsymbol{v}_{*,1}, \ldots, \boldsymbol{v}_{*,q-1})$ would be a lexicographically larger permutation of $\boldsymbol{v}$. Therefore $\boldsymbol{v}_{*,2} = \mathbb{1}_p$. By induction follows that $\boldsymbol{v}_{*,j} = \mathbb{1}_p$ for all $j \in [q]$. The case (ii) is analogous. □

It follows that in particular for $p = 1$, each vertex $\boldsymbol{v}$ of $\mathbf{O}_{1,q}$ is either $\mathbb{1}_q$, or it is $\mathbb{0}_q$, or it has first entry 1 and last entry 0.

Definition 3.67 All vertices $\boldsymbol{v} \in \mathbf{O}_{1,q}$ with $v_{1,1} = 1$ and $v_{1,q} = 0$ will in the following be referred to as *non-trivial* vertices.

Observation 3.68 Each non-trivial vertex of $\mathbf{O}_{1,q}$ can be identified with a sequence of ℓ tupels (α_i^1, α_i^0), $1 \leqslant i \leqslant \ell$ where α_i^1 and α_i^0 denote the length of the ith sequence of 1s and 0s, respectively, and

$$q = \sum_{i=1}^{\ell} (\alpha_i^1 + \alpha_i^0).$$

The lexicographic order induces an order on the set of tupels as follows:

$$(\alpha_i^1, \alpha_i^0) > (\alpha_j^1, \alpha_j^0) \text{ if and only if either } (\alpha_i^1 > \alpha_j^1) \text{ or } (\alpha_i^1 = \alpha_j^1 \text{ and } \alpha_i^0 < \alpha_j^0).$$

3.2.3 Optimization

Of course, it is possible to decide in polynomial time whether a given vector in $\mathcal{M}_{p,q}$ is in $\mathbf{O}_{p,q}(\mathfrak{C}_q)$ or not. However, it is unclear whether it is possible to optimize in polynomial time over $\mathbf{O}_{p,q}(\mathfrak{C}_q)$; the observation above does not seem to lead to an algorithm for optimization. Consequently, the complexity status for optimization over $\mathbf{O}_{p,q}(\mathfrak{C})$ (even for $p = 1$) is open for the full case as well as for all restricted cases (except $\leqslant 1$, $= 1$ and $k = \text{const}$).

However, even if there are no restrictions on the number of 1s per row, optimization over $\mathbf{O}_{1,q}(\mathfrak{C}_q)$ is at least as hard as optimization over $\mathbf{O}_{p,q}(\mathfrak{C}_q)$, as the following considerations show.

Let $m, n \in \mathbb{N}_>$ with $1 \leqslant m < n$. We define a projection $\phi_n : \mathcal{M}_{1,nq} \to \mathcal{M}_{n,q}$ that splits 0/1-vector $\boldsymbol{v} \in \mathcal{M}_{1,nq}$ into segments each of length n and interprets the jth segment as the jth column of the image vector, that is:

$$\phi_n\big(\boldsymbol{v} = (v_1, v_2, \dots, v_{nq})\big) = \begin{pmatrix} v_1 & v_{n+1} & \cdots & v_{(q-1)n+1} \\ \vdots & \vdots & & \vdots \\ v_n & v_{2n} & \cdots & v_{qn} \end{pmatrix}$$

Moreover, let $\psi_m : \mathcal{M}_{n,q} \to \mathcal{M}_{n-m-1,q}$ be the orthogonal projection that maps any vector $\boldsymbol{v}$ to $\boldsymbol{v}_{[m+1..n-1]\times[q]}$. In other words, ψ_m deletes the first m and the last row in $\boldsymbol{v} \in \mathcal{M}_{n\times q}$.

Lemma 3.69 *Let $p, q \in \mathbb{N}_>$. For any $\ell > p$, we set $s(\ell) := (2\ell + p + 1)q$ and define*

$$\tau := \psi_{2\ell} \circ \phi_{s(\ell)},$$

where ϕ and ψ are defined as described above. Moreover, let the set $\mathcal{B}_\ell \subseteq \mathcal{M}_{1,s(\ell)}$ contain all vectors in $\{0,1\}^{[1]\times[s(\ell)]}$ that are of the following form:

$$\boldsymbol{v} := (\mathbb{1}_\ell, \mathbb{0}_\ell, \boldsymbol{v}_1, 0,\ \mathbb{1}_\ell, \mathbb{0}_\ell, \boldsymbol{v}_2, 0, \dots, \mathbb{1}_\ell, \mathbb{0}_\ell, \boldsymbol{v}_q, 0), \tag{3.14}$$

where $\boldsymbol{v}_i \in \mathcal{M}_{1,p}$ for all $i \in [q]$. Then the following holds:

(i) If $\boldsymbol{v} \in \mathcal{B}_\ell \cap \mathbf{O}_{1,s(\ell)}(\mathfrak{C}_{s(\ell)})$, then $\tau(\boldsymbol{v}) \in \mathbf{O}_{p,q}(\mathfrak{C}_q)$.

(ii) For any vertex $\boldsymbol{w} \in \mathbf{O}_{p,q}(\mathfrak{C}_q)$ and $\ell > p$, there is a vector $\boldsymbol{v}' \in \mathbf{O}_{1,s(\ell)}(\mathfrak{C}_{s(\ell)}) \cap \mathcal{B}_\ell$ such that $\tau(\boldsymbol{v}') = \boldsymbol{w}$.

Proof. (i) Since $\boldsymbol{v} \in \mathcal{B}_\ell$, we know that

$$\tau(\boldsymbol{v}) = (\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_q).$$

Assume that $\tau(\boldsymbol{v}) \notin \mathbf{O}_{p,q}(\mathfrak{C}_q)$. Then there is a cyclic permutation $\pi : [q] \to [q]$ such that

$$\big(\boldsymbol{v}_{\pi(1)}, \boldsymbol{v}_{\pi(2)}, \dots, \boldsymbol{v}_{\pi(q)}\big)$$

is lexicographically larger than $\tau(\boldsymbol{v})$. However, this implies that

$$(\mathbb{1}_\ell, \mathbb{0}_\ell, \boldsymbol{v}_{\pi(1)}, 0,\ \mathbb{1}_\ell, \mathbb{0}_\ell, \boldsymbol{v}_{\pi(2)}, 0,\ \mathbb{1}_\ell, \mathbb{0}_\ell, \boldsymbol{v}_{\pi(3)}, 0,\ \dots,\ \mathbb{1}_\ell, \mathbb{0}_\ell, \boldsymbol{v}_{\pi(q)}, 0)$$

is lexicographically larger than $\boldsymbol{v}$. Contradiction.

(ii) It suffices to show that for any vertex $\boldsymbol{w}$ of $\mathbf{O}_{p,q}$, the vector

$$\boldsymbol{v}' = (\mathbb{1}_\ell, \mathbb{0}_\ell, \boldsymbol{w}_{*,1}, 0,\ \mathbb{1}_\ell, \mathbb{0}_\ell, \boldsymbol{w}_{*,2}, 0,\ \mathbb{1}_\ell, \mathbb{0}_\ell, \boldsymbol{w}_{*,3}, 0, \dots, \mathbb{1}_\ell, \mathbb{0}_\ell, \boldsymbol{w}_{*,q}, 0)$$

is in $\mathbf{O}_{1,s(\ell)}$ for an arbitrary $\ell > p$. Obviously, only those cyclic permutations $\pi : [s(\ell)] \to [s(\ell)]$ have to be considered that shift $(\mathbb{1}_\ell, \mathbb{0}_\ell)$-segments to the first position. Assume there is a π that gives a lexicographically larger vector than $\boldsymbol{v}'$, i.e. $\boldsymbol{v}' \notin \mathbf{O}_{1,s(\ell)}$. Then π induces a cyclic permutation $\pi' : [q] \to [q]$ such that

$$\big(\boldsymbol{w}_{*,\pi'(1)}, \boldsymbol{w}_{*,\pi'(2)}, \boldsymbol{w}_{*,\pi'(3)}, \dots, \boldsymbol{w}_{*,\pi'(q)}\big)$$

is lexicographically larger than $\boldsymbol{w}$. Contradiction.

□

Corollary 3.70 *Optimization over $\mathbf{O}_{1,q}$ is at least as hard as optimization over $\mathbf{O}_{p,q}$ for general p.*

Proof. From some given cost vector $\boldsymbol{c} \in \mathbb{R}^{[p]\times[q]}$, we construct the following cost vector

$$\boldsymbol{c}' := (\zeta\mathbb{1}_\ell,\ -\zeta\mathbb{1}_\ell,\ \boldsymbol{c}_{*,1},\ -\zeta,\ \zeta\mathbb{1}_\ell,\ -\zeta\mathbb{1}_\ell,\ \boldsymbol{c}_{*,2},\ -\zeta,\ldots,\zeta\mathbb{1}_\ell,\ -\zeta\mathbb{1}_\ell,\ \boldsymbol{c}_{*,q},\ -\zeta),$$

with $\zeta \in \mathbb{R}$ sufficiently large (i.e. $\zeta > \sum_{(i,j)\in\mathrm{supp}^+(\boldsymbol{c})} c_{i,j}$). Any optimal solution for $\max(\{\langle \boldsymbol{c}', \boldsymbol{x}\rangle \mid \boldsymbol{x} \in \mathbf{O}_{1,s(\ell)}\})$ induces by projection τ an optimal solution for $\max(\{\langle \boldsymbol{c}, \boldsymbol{x}\rangle \mid \boldsymbol{x} \in \mathbf{O}_{p,q}\})$. ☐

3.2.4 Facial Description of Packing and Partitioning Orbitopes

Kaibel and Pfetsch gave a linear description for packing and partitioning orbitopes also for the cyclic case, see [65]. The facial description for the partition orbitope $\mathrm{O}^{=}_{p,q}(\mathfrak{C}_q)$ is

$$\begin{aligned} x_{1,1} &= 1 \\ x_{1,j} &= 0 \quad \forall\, 2 \leqslant j \leqslant q \\ x_{i,j} &\geqslant 0 \quad \forall\, 2 \leqslant i \leqslant p \text{ and } j \in [q] \\ \sum_{j\in[q]} x_{i,j} &= 1 \quad \forall\, 2 \leqslant i \leqslant p, \end{aligned}$$

while the packing orbitope $\mathrm{O}^{\leqslant}_{p,q}(\mathfrak{C}_q)$ is fully and non-redundantly described by

$$\begin{aligned} x_{1,1} &\geqslant 0 \\ x_{1,j} &= 0 \quad \forall\, 2 \leqslant j \leqslant q \\ x_{i,j} &\geqslant 0 \quad \forall\, 2 \leqslant i \leqslant p \text{ and } j \in [q] \\ \sum_{j\in[q]} x_{i,j} &\leqslant 1 \quad \forall\, 2 \leqslant i \leqslant p \\ \sum_{j=2}^{q} x_{i,j} - \sum_{k=1}^{i-1} x_{k,1} &\leqslant 0 \quad \forall\, 2 \leqslant i \leqslant p. \end{aligned}$$

Apart from these special cases (and two-columned orbitopes), no facial description is known for orbitopes over the cyclic group. Moreover, our computer experiments seem to indicate that the facial structure of full orbitopes over the cyclic group is even more complicated than the facial structure of full orbitopes over the full symmetric group. For instance, we computed the facial structure of $\mathbf{O}_{3,4}(\mathfrak{C}_4)$ using the software package `polymake` ([47]; for a visualization of the facets, see the figure 3.20). The results do not seem very encouraging. (Compare with figure 3.17: for $p = 3$ and $q = 4$, there are 361 facets when the cyclic group is operating on the columns versus 39 facets with the symmetric group operating. However, what counts more is the fact that recurring patterns seem to become much less obvious.)

3.2.5 Other Groups Operating on the Columns

Partition the column indices $[q]$ into k subsets $\mathcal{G}_i$. If it is possible to simultaneously decompose the group G operating on the columns into a product of $G_1 \times \ldots \times G_k$ of subgroups G_i such that subgroup G_i is operating on subset $\mathcal{G}_i$ for all $i \in [k]$, then the representatives of $\mathbf{O}_{p,q}(G)$ can be characterized as follows.

Proposition 3.71 *Let $[q] = \dot{\bigcup}_{i=1}^{k} \mathcal{G}_i$ and $G \simeq G_1 \times \cdots \times G_k$. Let the action of G be defined by the actions of the G_i on $\mathcal{G}_i$. Denote moreover by $\mathcal{V}_i$ the vertex set of $\mathbf{O}_{p,\mathcal{G}_i}(G_i)$. Then $\boldsymbol{x}$ is a vertex of $\mathbf{O}_{p,q}(G)$ if and only if $\boldsymbol{x}_{*,\mathcal{G}_i} \in \mathcal{V}_i$ for all $i \in [k]$.*

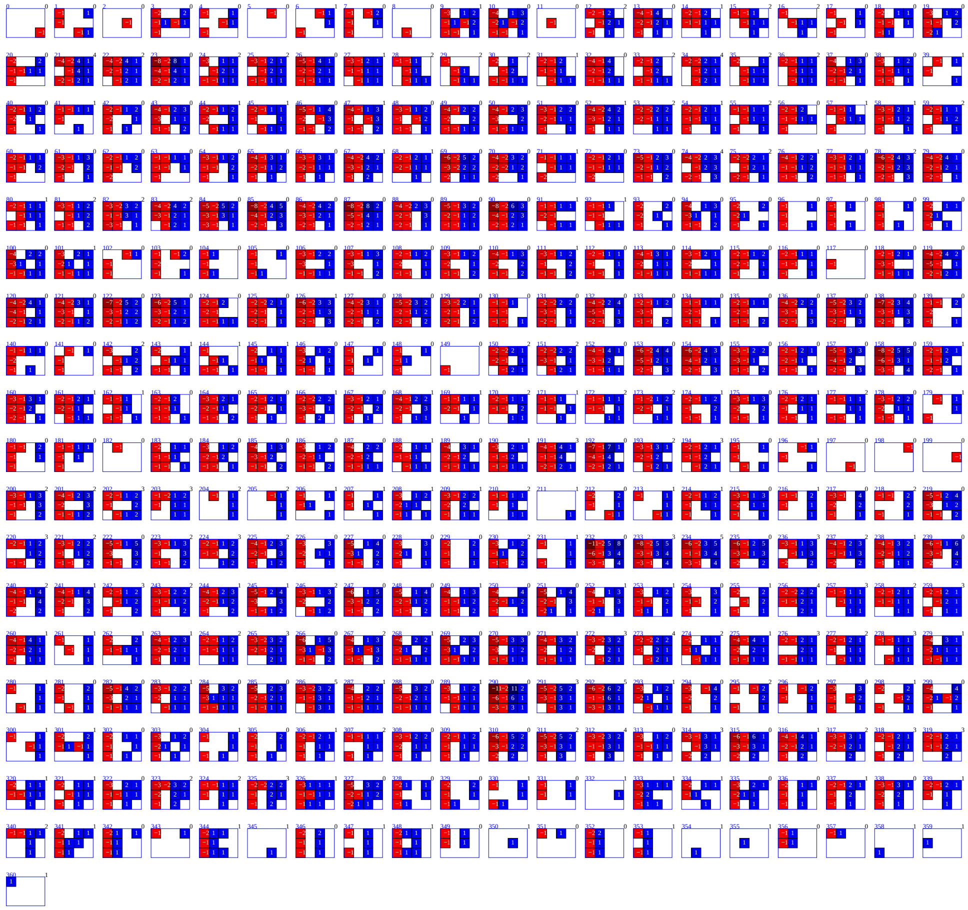

Figure 3.20: The facets of $\mathbf{O}_{3,4}(\mathfrak{C}_4)$.

Proof. ($\Rightarrow$) Assume $\boldsymbol{v}$ is a vertex of $\mathbf{O}_{p,q}(G)$ and there is an $i \in [k]$ such that $\boldsymbol{v}_{*,\mathcal{G}_i}$ is not vertex of $\mathbf{O}_{p,\mathcal{G}_i}(G_i)$. Then there is an element $g \in G_i$ such that $\boldsymbol{v}_{*,g(\mathcal{G}_i)} \succ \boldsymbol{v}_{*,\mathcal{G}_i}$. Therefore, applying $(\mathrm{id}, \dots, \mathrm{id}, g, \mathrm{id}, \dots, \mathrm{id}) \in G$ on the columns of $\boldsymbol{v}$ yields a lexicographically larger vector than $\boldsymbol{v}$. (id denotes here the identity in groups G_k, $k \neq i$.) Contradiction. ($\Leftarrow$) Assume that for some $\boldsymbol{v} \in \mathbf{O}_{p,q}(G)$, $\boldsymbol{v}_{*,\mathcal{G}_i}$ is a vertex of $\mathbf{O}_{p,\mathcal{G}_i}(G_i)$ for each $i \in [k]$, but there is an element $g \in G$ such that $\boldsymbol{v}_{*,g([q])}$ is lexicographically larger than $\boldsymbol{v}$. We compare $\boldsymbol{v}_{*,g([q])}$ and $\boldsymbol{v}_{*,q}$ column by column. Let j the index of the first column such that $\boldsymbol{v}_{*,j} \prec \boldsymbol{v}_{*,g(j)}$. Then by definition of G, there must be some subset $\mathcal{G}_i \subseteq [q]$ such that both $j, g(j) \in \mathcal{G}_i$. Moreover, the restriction $\widetilde{g}$ of g on the elements of $\mathcal{G}_i$ must be an element of G_i. But then, operating with $(\mathrm{id}, \dots, \mathrm{id}, \widetilde{g}, \mathrm{id}, \dots, \mathrm{id})$ on the columns of $\boldsymbol{v}$ gives a lexicographically larger vector than $\boldsymbol{v}$. Contradiction. $\square$

For $G_i \simeq \mathfrak{S}_{\mathcal{G}_i}$ for all $i \in [k]$, this fact has been observed by Faenza. It is clear that in full orbitopes over products of full symmetric groups, Proposition 3.71 implies that one can optimize over $\mathbf{O}_{p,q}(G)$ in polynomial time, by optimizing separately over each column partition $\mathcal{G}_i$. Note that an optimal vertex of $\mathbf{O}_{p,q}(G)$ can definitely have undefined splits, meaning: the columns of the vertex do not have to be in lexicographic order.

Proposition 3.71 has also implications for the linear description of $\mathbf{O}_{p,q}(G)$. If a linear description of $\mathbf{O}_{p,\mathcal{G}_i}(G_i)$ is available for any $i \in [k]$, one can obtain a linear description for $\mathbf{O}_{p,q}(G)$ by lifting the inequalities appropriately, independently of the type of groups G_i, since $\mathbf{O}_{p,q}(G_1 \times \cdots \times G_k) = \mathbf{O}_{p,\mathcal{G}_1}(G_1) \times \cdots \times \mathbf{O}_{p,\mathcal{G}_k}(G_k)$. In other words: assuming that for any $i \in [k]$, the linear description of $\mathbf{O}_{p,\mathcal{G}_i}(G_i)$ is given by

$$\mathbf{O}_{p,\mathcal{G}_i}(G_i) = \{\boldsymbol{y} \in \mathbb{R}^{[p]\times\mathcal{G}_i} \mid \boldsymbol{A}_i\boldsymbol{y} \leqslant \boldsymbol{b}_i\},$$

with $\boldsymbol{A}_i \in \mathbb{Q}^{[k_i]\times([p]\times\mathcal{G}_i)}$, $\boldsymbol{y} \in \{0,1\}^{[p]\times\mathcal{G}_i}$ and $\boldsymbol{b}_i \in \mathbb{Q}^{[k_i]}$ and defining

$$\mathbf{P} := \{\boldsymbol{x} \in \mathbb{R}^{[p]\times[q]} \mid \boldsymbol{A}_i\boldsymbol{x}_{*,\mathcal{G}_i} \leqslant \boldsymbol{b}_i \ \forall i \in [k]\},$$

then $\mathbf{P} = \mathbf{O}_{p,q}(G)$ holds.

For other groups, for instance alternating groups, we have neither experimental nor analytical results.

Chapter 4

Orbisacks

Orbisacks $\mathbf{O}_{p,2}$ are full orbitopes over the symmetric group, and, as we will see, they are – apart from some other special cases like packing and partitioning orbitopes – the best understood class of orbitopes. The reason is that orbisacks are special in many respects. As has been shown in Lemma 3.16, the orbisack is a knapsack polytope. Moreover, each orbitope vertex has exactly one split, the critical row, which opens the way to nice extended formulations.

But, because of their very special properties, orbisacks are a kind of laboratory mice. To us, there is no real world application known which could profit from the linear description of orbisacks. Our interest in orbisacks is more theoretical. From studying them, we tried to learn how to describe full orbitopes with more than two columns linearly. However, it turned out that all the properties that make orbisacks so nice get lost as soon as we consider more than two rows, as indicated in the introduction.

In this chapter, we will present in the first main part three different ways to derive a linear description of the orbisack $\mathbf{O}_{p,2}$.

- The first proof shows in a direct way that if inequality $\boldsymbol{a}\boldsymbol{x} \leqslant b$ is defining a facet, then the entries of normal vector $\boldsymbol{a}$ show a certain sign pattern, and all inequalities based on the same sign pattern are dominated by one with certain absolute values (the so called *valued block inequality*), which will prove to be facet defining.
- The second proof relies on the fact that the orbisack is a special knapsack polytope: its weight set is ordered such that the ratio of two subsequent weights is integral. This is what is generally known as a *sequential knapsack polytope* (SKP). In [117], Weismantel and Pochet gave a complete linear description of all SKPs which gives us a second approach for a linear description of orbisacks.
- Our third proof relies on an extended formulation for orbisacks (see page 19) and a two-step application of faithful sectioning (see Theorem 2.17).

The second main part of the chapter is mainly dedicated to the graph of the orbisack.

4.1 Facial Description of Orbisacks I (Combinatorial Proof)

For the following proofs, the term "critical row" will become of great importance (see Definition 3.7). Additionally, we will give now some more definitions that will prove to be useful.

Let $\mathcal{V}_{\mathbf{O}_{p,2}}$ denote the set of vertices of the orbisack.

Definition 4.1 (Layers) The position of the critical row can be used to partition the vertex set of the orbisack. For $k \in [p+1]$, we define sets

$$\mathcal{L}_k := \{\boldsymbol{v} \in \mathcal{V}_{\mathbf{O}_{p,2}} \mid \operatorname{crit}(\boldsymbol{v}) = k\}.$$

The equivalence classes $\mathcal{L}_k$ are called *layers*.

Definition 4.2 (Maximizing vertex set) For a given vector $\boldsymbol{a} \in \mathbb{R}^{[p]\times[2]}$, we define the set

$$\mathcal{V}[\boldsymbol{a}] := \operatorname{argmax}_{\boldsymbol{v}}(\{\langle \boldsymbol{v}, \boldsymbol{a}\rangle \mid \boldsymbol{v} \in \mathcal{V}_{\mathbf{O}_{p,2}}\})$$

of orbisack vertices maximizing cost functional $\langle \boldsymbol{x}, \boldsymbol{a}\rangle$. Sometimes, it will be convenient to also partition set $\mathcal{V}[\boldsymbol{a}]$ into subsets

$$\mathcal{V}_k[\boldsymbol{a}] := \mathcal{V}[\boldsymbol{a}] \cap \mathcal{L}_k, \quad k \in [p+1].$$

A key tool for our proofs will be the *modification of a vector*, that is: a change in selected components of some vertex (see page 5). For a given vector $\boldsymbol{v} \in \mathbb{R}^{[p]\times[2]}$, a modification $\widetilde{\boldsymbol{v}} \in \mathbb{R}^{[p]\times[2]}$ of $\boldsymbol{v}$ in component $\widetilde{v}_{k,\ell} := s$ is the vector

$$\widetilde{v}_{i,j} = \begin{cases} s, & \text{if } (i,j) = (k,\ell) \\ v_{i,j}, & \text{otherwise.} \end{cases} \qquad \text{for all } (i,j) \in [p] \times [2].$$

Before we look at the linear description in detail, we need two more definitions:

Definition 4.3 (n-rows, p-rows, np-rows) For some vector $\boldsymbol{a} \in \mathbb{R}^{[p]\times[2]}$,

- a *negative* row (in short *n-row*) has the form $(-\alpha, 0)$,
- a *positive* row (in short *p-row*) has the form $(0, \beta)$, and
- a *negative-positive* row (in short np-row) has the form $(-\alpha, \beta)$,

where $\alpha, \beta > 0$. If the row contains only 0s, we call it *empty*.

Definition 4.4 (Basement) Let $\boldsymbol{ax} \leqslant b$ be an inequality with $\boldsymbol{a} \in \mathbb{Q}^{[p]\times[2]}$ and $\boldsymbol{a} \neq \mathbb{0}$. We call row

$$\operatorname{base}(\boldsymbol{a}) := \max(\{i \in [p] \mid \boldsymbol{a}_{i,*} \text{ is not } (0,0) \text{ (empty)}\})$$

the *basement* of $\boldsymbol{a}$ (or of the inequality $\boldsymbol{ax} \leqslant b$, respectively).

Remark 4.5 From lemmas 3.54 and 3.62 it follows that any facet defining inequality with basement $s < p$ is a lifted facet from $\mathbf{O}_{s,2}$. Hence, if not otherwise stated, we can (and will) for the following proofs w.l.o.g. assume that any considered inequality $\boldsymbol{ax} \leqslant b$ has full extension, i.e. $\operatorname{base}(\boldsymbol{a}) = p$.

We will obtain the full description in four steps:

(i) We give for some candidate $\boldsymbol{ax} \leqslant b$ necessary conditions on the sign pattern of $\boldsymbol{a}$ and b.
(ii) We define the class of so called *valued block inequalities* (VBI). These show the sign-pattern of a facet defining inequality and dominate all inequalities with same sign pattern.
(iii) We prove that all VBI in fact define facets.
(iv) Last, we briefly consider trivial inequalities.

4.1.1 Sign Pattern of Non-Trivial Facet Defining Inequalities

If the facet $\mathtt{F}$ is defined by inequality $\boldsymbol{ax} \leqslant b$, then it is denoted by $\mathtt{F}(\boldsymbol{ax} \leqslant b)$.

Lemma 4.6 *Let facet* $\mathtt{F}(\boldsymbol{av} \leqslant b)$ *be non-trivial. For all* $i \in [p]$ *and* $j \in [2]$*, there is at least one vertex* $\boldsymbol{v}$ *with* $v_{i,j} = 1$ *and at least one vertex* $\boldsymbol{v}'$ *with* $v'_{i,j} = 0$ *both active for facet* $\mathtt{F}$*.*

Proof. If all vertices on facet $\mathtt{F}$ have component $v_{i,j} = 1$ in common, then

$$\mathtt{F} \subset \{\boldsymbol{x} \in \mathbf{O}_{p,2} \mid x_{i,j} = 1\},$$

i.e. $\mathtt{F}$ is contained in a trivial face of $\mathbf{O}_{p,2}$. But $\mathtt{F}$ is not trivial by assumption. Contradiction. Similarly with $v'_{i,j} = 0$. ☐

Lemma 4.7 *For any non-trivial facet* $\mathtt{F}(\boldsymbol{av} \leqslant b)$*, it holds that* $a_{i,1} \leqslant 0$ *and* $a_{i,2} \geqslant 0$ *for all* $i \in [p]$*.*

Proof. Assume $a_{i,1} > 0$. There is some vertex $\boldsymbol{v}$ in $\mathtt{F}$ with $v_{i,1} = 0$ (Lemma 4.6). We modify $\boldsymbol{v}$ in component $\widetilde{v}_{i,1} := 1$. Then $\widetilde{\boldsymbol{v}} \in \mathbf{O}_{p,2}$, but it violates $\boldsymbol{ax} \leqslant b$. Contradiction.
For the assertion $a_{i,2} \geqslant 0$, we modify some vertex $\boldsymbol{v}$ with $v_{i,2} = 1$ in component $\widetilde{v}_{i,2} := 0$. ☐

Lemma 4.8 *Let* $\mathtt{F}(\boldsymbol{av} \leqslant b)$ *be a non-trivial facet. Then there is no empty row* r *with* $1 \leqslant r < p$*.*

Proof. We assume that $\text{base}(\boldsymbol{a}) = p$ (Remark 4.5). This implies that either $a_{p,1} < 0$ or $a_{p,2} > 0$ or both (Lemma 4.7); let's say that $a_{p,1} < 0$. (The other case can be treated similarly).
Assume row $r < p$ was an empty row in $\boldsymbol{a}$. Let $\boldsymbol{v}$ be some vertex on $\mathtt{F}$ with $v_{p,1} = 1$ (Lemma 4.6). Then $\text{crit}(\boldsymbol{v}) \not< p$, because we could then modify $\boldsymbol{v}$ in $\widetilde{v}_{p,1} := 0$, obtaining $\widetilde{\boldsymbol{v}} \in \mathbf{O}_{p,2}$ with $\langle \boldsymbol{a}, \widetilde{\boldsymbol{v}} \rangle > \langle \boldsymbol{a}, \boldsymbol{v} \rangle = b$. So $\boldsymbol{v}_{p,*}$ is either $(1,1)$ or $(1,0)$, and we can find another modification of $\boldsymbol{v}$ in components $\widetilde{v}_{r,1} := 1$, $\widetilde{v}_{r,2} := 0$ and $\widetilde{v}_{p,1} := 0$. Row r is then critical row in $\widetilde{\boldsymbol{v}}$. $\widetilde{\boldsymbol{v}}$ is also vertex of $\mathbf{O}_{p,2}$, but $\langle \boldsymbol{a}, \widetilde{\boldsymbol{v}} \rangle > \langle \boldsymbol{a}, \boldsymbol{v} \rangle$. Contradiction. ☐

DEFINITION 4.9 We call $\boldsymbol{a}_{i,*}$ a *balanced* row, if it is an n-p-row and its row sum $a_{i,1} + a_{i,2} = 0$.

Lemma 4.10 *Let* $\mathtt{F}(\boldsymbol{av} \leqslant b)$ *be a non-trivial facet. Then row* $\boldsymbol{a}_{p,*}$ *is balanced.*

Proof. We assume $\text{base}(\boldsymbol{a}) = p$ (Remark 4.5). We assume $a_{p,2} > 0$ (the case $a_{p,1} < 0$ is similar).
There is a vertex $\boldsymbol{v}$ with $v_{p,1} = 1$ that lies on facet $\mathtt{F}$ (Lemma 4.6). Then $v_{p,2}$ must be 1, because if not, then we could modify $\boldsymbol{v}$ in component $\widetilde{v}_{p,2} := 1$ without leaving the orbisack, but $\langle \boldsymbol{a}, \widetilde{\boldsymbol{v}} \rangle > \langle \boldsymbol{a}, \boldsymbol{v} \rangle = b$. We can modify $\boldsymbol{v}$ also by $\widetilde{v}_{p,1} := 0$ and $\widetilde{v}_{p,2} := 0$. $\widetilde{\boldsymbol{v}}$ is a vertex of the orbisack, so $\boldsymbol{a}\widetilde{\boldsymbol{v}} \leqslant b$ and $\langle \boldsymbol{a}, \boldsymbol{v} \rangle - \langle \boldsymbol{a}, \widetilde{\boldsymbol{v}} \rangle \geqslant 0$. So we get that $a_{p,1} + a_{p,2} \geqslant 0$ must hold to keep $\boldsymbol{ax} \leqslant b$ valid.
But there is also a vertex $\boldsymbol{u}$ with $u_{p,2} = 0$ (Lemma 4.6) and $u_{p,1} = 0$. By modifying $\boldsymbol{u}$ by $\widetilde{u}_{p,1} := 1$ and $\widetilde{u}_{p,2} := 1$, we obtain $a_{p,1} + a_{p,2} \leqslant 0$. Hence $a_{p,1} + a_{p,2} = 0$. ☐

Lemma 4.11 *Let* $\mathtt{F}(\boldsymbol{av} \leqslant b)$ *be a non-trivial facet. Then row* $\boldsymbol{a}_{1,*}$ *is balanced.*

Proof. Assume first that $a_{1,1} + a_{1,2} < 0$. There is a vertex $\boldsymbol{v}$ with $v_{1,2} = 1$ (Lemma 4.6). But then $v_{1,1} = 1$ because of the lexicographic order, and we could enlarge $\langle \boldsymbol{a}, \boldsymbol{v} \rangle$ by modifying the first row of $\boldsymbol{v}$ to $(0, 0)$. Contradiction. If $a_{1,1}+a_{1,2} > 0$,

then it's the same situation with $\boldsymbol{v}_{1,*} = (0,0)$. Contradiction.
So $a_{1,1} + a_{1,2} = 0$. As the first row cannot be an empty row $(0,0)$ because of Lemma 4.8, the first row must be balanced. ☐

4.1.2 Block-Inequalities

We will now focus on the sign structure of $\boldsymbol{a}$ of non-trivial facet defining inequalities $\boldsymbol{ax} \leqslant b$.

DEFINITION 4.12 (Sign-pattern, block types) A vector $\sigma \in \{0,+,-\}^{[p]\times[2]}$ is called a *sign-pattern*. We are especially interested in those sign-patterns that can be segmented into certain substructures, called *blocks*. A block is a set of subsequent rows of σ with the following properties:

- The first row has sign-pattern $(-,+)$.
- All the following rows inside the block have sign pattern $(0,+)$ or $(-,0)$.

A sign-pattern that can be divided into a set of blocks is said to be of *general block type*. If the lowest block – that is, the block with a first row with largest row index among all blocks – additionally consists of one row only, the sign pattern is of *special block type*.

DEFINITION 4.13 (Function sign) We define function sign : $\mathbb{R}^{[p]\times[2]} \to \{0,+,-\}^{[p]\times[2]}$ by

$$\operatorname{sign}(\boldsymbol{a})_{i,j} = \begin{cases} +, & \text{if } a_{i,j} > 0 \\ -, & \text{if } a_{i,j} < 0 \\ 0, & \text{if } a_{i,j} = 0 \end{cases} \qquad \text{for all } (i,j) \in [p] \times [2].$$

The following observation is crucial for the further considerations.

Observation 4.14 Lemmas 4.6 through 4.11 show that if inequality $\boldsymbol{ax} \leqslant b$ defines a non-trivial facet, then sign pattern $\operatorname{sign}(\boldsymbol{a})$ is of special block type.

We will further narrow the set of candidates by considering a certain subset of these inequalities, the valued block-inequalities.

DEFINITION 4.15 (Valued block inequality (VBI)) We call an inequality $\boldsymbol{ax} \leqslant b$ a *valued block inequality (VBI)*, if the following holds:

- $\operatorname{sign}(\boldsymbol{a})$ is of special block type.
- Inside each block, every nonzero entry has the same absolute value. We will call it the *value of the block*.
- The values of the blocks are powers of 2 and ordered such that if $B_n, \ldots, B_0$ are the blocks from top to bottom with block B_0 the basement, then the value of B_0 is 2^0, and the value of B_i is 2^{i-1} for all $i \in [n]$.
- Last,

$$b := \sum_{\substack{i=1 \\ i \text{ is p-row}}}^{p} a_{i,2}$$

Fig. 4.1 shows two examples of inequalities. The left one is a VBI.

Lemma 4.16 *Let $\boldsymbol{a} \in \mathbb{R}^{[p]\times[2]}$. Then for every $\boldsymbol{a}$-maximizing vertex $\boldsymbol{v} \in \mathcal{V}_k[\boldsymbol{a}]$ with*

Block B_3	-4	4	$\leqslant 9$	-4	4	$\leqslant 10$
	-4	0		-4	0	
	0	4		0	2	
	-4	0		-3	0	
	-4	0		-4	0	
B_2	-2	2		-2	2	
	-2	0		-3	0	
	0	2		0	2	
	-2	0		-2	0	
	0	2		0	2	
B_1	-1	1		-1	1	
	0	1		0	1	
B_0	-1	1		-1	0	

Figure 4.1: The sign pattern of the left inequality is of special block type. Moreover, by choice of the values on the entries and of the right-hand side, the inequality is a VBI. The inequality to the right has a sign-pattern of general block type, and it is no VBI.

$k \in [p+1]$, the following holds:

$$\boldsymbol{v}_{i,*} \in \begin{cases} \{(1,1),(0,0)\}, & \text{if } a_{i,1}+a_{i,2}=0 \\ \{(1,1)\}, & \text{if } a_{i,1}+a_{i,2}>0 \\ \{(0,0)\}, & \text{if } a_{i,1}+a_{i,2}<0 \end{cases} \qquad \forall\, 1 \leqslant i < k$$

$$v_{i,\ell} \in \begin{cases} \{1,0\}, & \text{if } a_{i,\ell}=0 \\ \{1\}, & \text{if } a_{i,\ell}>0 \\ \{0\}, & \text{if } a_{i,\ell}<0 \end{cases} \qquad \forall\, k < i \leqslant p \text{ and } \ell \in [2]$$

Proof. Above the critical row, vertex $\boldsymbol{v}$ contains only rows $(1,1)$ or $(0,0)$. So if the row sum of row i is negative, we will choose row type $(0,0)$ when constructing an $\boldsymbol{a}$-maximizing vertex $\boldsymbol{v}$, and row type $(1,1)$ if the row sum is positive.
Below the critical row, $\boldsymbol{v}_{[k+1\ldots p],*}$ forms a cube. □

Note that the statement of Lemma 4.16 is in particular true if $\operatorname{sign}(\boldsymbol{a})$ is of special block type.

Corollary 4.17 *Let $\boldsymbol{a}, \boldsymbol{a}' \in \mathbb{R}^{[p]\times[2]}$ be two vectors with $\operatorname{sign}(\boldsymbol{a}) = \operatorname{sign}(\boldsymbol{a}')$ of special block type, and let moreover $\boldsymbol{a}'\boldsymbol{x} \leqslant b'$ be a valued block inequality. Then $\mathcal{V}[\boldsymbol{a}] \subseteq \mathcal{V}[\boldsymbol{a}']$.*

Proof. If $\mathcal{V}_k \neq \varnothing$ for $k \in [p+1]$, each vector $\boldsymbol{v}$ in $\mathcal{V}_k[\boldsymbol{a}']$ has the following shape.

$$\boldsymbol{v}_{i,*} \in \begin{cases} \{(1,1),(0,0)\}, & \text{if } \boldsymbol{a}'_{i,*} \text{ is np-row} \\ \{(1,1)\}, & \text{if } \boldsymbol{a}'_{i,*} \text{ is p-row} \\ \{(0,0)\}, & \text{if } \boldsymbol{a}'_{i,*} \text{ is n-row} \end{cases} \qquad \forall\, 1 \leqslant i < k$$

$$v_{i,\ell} \in \begin{cases} \{1,0\}, & \text{if } \boldsymbol{a}'_{i,*} \text{ is np-row} \\ \{1\}, & \text{if } \boldsymbol{a}'_{i,*} \text{ is p-row} \\ \{0\}, & \text{if } \boldsymbol{a}'_{i,*} \text{ is n-row} \end{cases} \qquad \forall\, k < i \leqslant p \text{ and } \ell \in [2]$$

In particular, $\mathcal{V}[\boldsymbol{a}'] \neq \varnothing$. Comparing with the vertices in $\mathcal{V}_k[\boldsymbol{a}]$ (proof of Lemma 4.16), we get that $\mathcal{V}[\boldsymbol{a}] \subseteq \mathcal{V}[\boldsymbol{a}']$. $\square$

Thus, showing that every VBI defines a facet finishes the characterization of the linear description of orbisacks. This problem will be tackled in the following section.

4.1.3 Block-Inequalities are Facet Defining Inequalities

For this and the following section, we drop the assumption that $\operatorname{base}(\boldsymbol{a}) = p$ for any VBI $\boldsymbol{ax} \leqslant b$, i.e. we also consider valued block inequalities that are lifted.

Lemma 4.18 *Valued block inequalities are valid.*

Proof. Let $\boldsymbol{ax} \leqslant b$ be a valued block inequality and let vertex $\boldsymbol{v} \in \mathcal{V}_k[\boldsymbol{a}]$. Let α be the value of the block where critical row k lives in. Then

$$\langle \boldsymbol{a}_{k,*}, \boldsymbol{v}_{k,*} \rangle = \begin{cases} -\alpha, & \text{if } a_{k,*} \text{ is n-row or np-row} \\ 0, & \text{if } a_{k,*} \text{ is p-row} \end{cases}$$

So there are two possibilities:

(i) Row k is p-row. Then $\boldsymbol{v}$ is collecting at most
- the positive entries in all p-rows of $\boldsymbol{a}$, except for entry $+\alpha$ in row k, and
- all positive entries in np-rows below row k.

(ii) Row k is np-row or n-row. Then $\boldsymbol{v}$ is collecting at most
- the positive entries in *all* p-rows of $\boldsymbol{a}$,
- all positive entries in np-rows below row k, and
- the negative entry $-\alpha$ in row k.

Thus, vector $\boldsymbol{v}$ collects in both cases at most the same value

$$\langle \boldsymbol{a}, \boldsymbol{v} \rangle = \sum_{\substack{i \in [p] \\ i \text{ p-row}}} a_{i,2} + \underbrace{\sum_{\substack{i \in [k+1..p] \\ i \text{ np-row}}} a_{i,2}}_{\overset{*}{=}\alpha} - \alpha = \sum_{\substack{i \in [p] \\ i \text{ p-row}}} a_{i,2} \overset{*}{=} b, \tag{4.1}$$

where both identities $(*)$ hold by definition of valued block inequalities 4.15. $\square$

The following almost immediately follows from the lemma above.

Observation 4.19 For any valued block inequality $\boldsymbol{ax} \leqslant b$,

$$\mathcal{V}_k[\boldsymbol{a}] \neq \varnothing \text{ for all } k \in [p+1] \setminus \{\operatorname{base}(\boldsymbol{a})\}.$$

Proof. Any vertex $\boldsymbol{v}$ of $\mathbf{O}_{p,2}$ with critical row in the basement $\operatorname{base}(\boldsymbol{a})$ that is supposed to maximize $\langle \boldsymbol{a}, \boldsymbol{x} \rangle$ collects the entries in all p-rows above the basement, as well as entry $a_{\operatorname{base}(\boldsymbol{a}),1} = -1$ in the basement. Therefore,

$$\langle \boldsymbol{a}, \boldsymbol{v} \rangle = \sum_{\substack{i \in [p] \\ i \text{ p-row}}} a_{i,2} - 1 \overset{*}{<} b,$$

where $(*)$ holds because of Definition 4.15. For $k \neq \operatorname{base}(\boldsymbol{a})$, the proof of Lemma 4.18 shows that $\langle \boldsymbol{a}, \boldsymbol{v} \rangle$ must equal b (see equation (4.1)), which implies that $\mathcal{V}_k[\boldsymbol{a}] \neq \varnothing$ then. $\square$

Proposition 4.20 *All valued block inequalities define facets.*

Proof. As it is clear that all trivial inequalities are valid (see Lemma 4.21), we know so far that the trivial and valued block inequalities together yield a complete description of the orbisack. Therefore, it suffices here to show that for any valued block inequality $\boldsymbol{ax} \leqslant b$, the set of vertices $\mathcal{V}[\boldsymbol{ax} \leqslant b]$ is not completely contained in the set of vertices on any other (block or trivial) inequality.

First, let $\mathtt{F}(\boldsymbol{ax} \leqslant b)$, $\mathtt{F}'(\boldsymbol{a}'\boldsymbol{x} \leqslant b')$ be the faces defined by $\boldsymbol{ax} \leqslant b$ and $\boldsymbol{a}'\boldsymbol{x} \leqslant b'$, respectively, and suppose that $\mathcal{V}[\boldsymbol{a}] \subseteq \mathcal{V}[\boldsymbol{a}']$.

From Observation 4.19 follows that $\operatorname{base}(\boldsymbol{a}') = \operatorname{base}(\boldsymbol{a})$ must hold; otherwise, $\mathtt{F}(\boldsymbol{ax} \leqslant b) \nsubseteq \mathtt{F}(\boldsymbol{a}'\boldsymbol{x} \leqslant b')$.

Let now row $1 < i < \operatorname{base}(\boldsymbol{a})$ be a row with differing sign pattern in $\boldsymbol{a}$ and $\boldsymbol{a}'$.

From Observation 4.19 follows that $\mathcal{V}_1[\boldsymbol{a}] \neq \varnothing$ and $\mathcal{V}_{p+1}[\boldsymbol{a}] \neq \varnothing$. In particular, depending on the shape of $\boldsymbol{a}_{i,*}$ and $\boldsymbol{a}'_{i,*}$, we can choose a vector $\boldsymbol{v}$ from $\mathcal{V}_1[\boldsymbol{a}]$ or from $\mathcal{V}_{p+1}[\boldsymbol{a}]$, respectively, with the following entries in row $\boldsymbol{v}_{i,*}$:

$\operatorname{sign}(\boldsymbol{a}_{i,*})$ \ $\operatorname{sign}(\boldsymbol{a}'_{i,*})$	$(-,0)$	$(0,+)$	$(-,+)$
$(-,0)$	—	$(0,0)$	$(0,0)$
$(0,+)$	$(1,1)$	—	$(1,1)$

$\boldsymbol{v} \in \mathcal{V}_1[\boldsymbol{a}]$

$\operatorname{sign}(\boldsymbol{a}_{i,*})$ \ $\operatorname{sign}(\boldsymbol{a}'_{i,*})$	$(-,0)$	$(0,+)$	$(-,+)$
$(-,+)$	$(1,1)$	$(0,0)$	—

$\boldsymbol{v} \in \mathcal{V}_{p+1}[\boldsymbol{a}]$

By this choice, $\boldsymbol{v} \in \mathcal{V}[\boldsymbol{a}]$ and $\boldsymbol{v} \notin \mathcal{V}[\boldsymbol{a}']$. Contradiction.

Last, suppose that either trivial inequality $x_{i,j} \leqslant 1$ or $x_{i,j} \geqslant 0$ contains $\mathtt{F}(\boldsymbol{ax} \leqslant b)$. However, this would imply that

$$\mathcal{V}[\boldsymbol{a}] \subseteq \{\boldsymbol{x} \in \mathbf{O}_{p,2} \mid x_{i,j} = \zeta\},$$

with $\zeta = 1$ in the first case and $\zeta = 0$ in the second. But Lemma 4.6 shows that this is not true. Contradiction. $\square$

4.1.4 Trivial Facet Defining Inequalities

Last, we consider the trivial inequalities.

Lemma 4.21 *Inequalities*

$$x_{i,j} \geqslant 0 \text{ and } x_{i,j} \leqslant 1 \quad \forall i \in [p] \text{ and } j \in [2]$$

are valid for the orbisack.

Proof. Obviously true. □

Proposition 4.22 *The following trivial inequalities are defining facets:*

(i) $x_{i,j} \leqslant 1$ *for* $(i,j) \in ([p] \times [2]) \setminus \{(1,2)\}$

(ii) $-x_{i,j} \leqslant 0$ *for* $(i,j) \in ([p] \times [2]) \setminus \{(1,1)\}$

Proof. We start with the exceptions.

$x_{1,2} \leqslant 1$: Any vertex $\boldsymbol{v}$ on the face $\mathtt{F}(x_{1,2} \leqslant 1)$ must have first row $\boldsymbol{v}_{1,*} = (1,1)$ because of the lexicographic order of the columns of $\boldsymbol{v}$. But then $\boldsymbol{v} \in \mathtt{F}(-x_{1,1} + x_{1,2} \leqslant 0)$. Therefore, $x_{1,2} \leqslant 1$ cannot be a facet defining inequality.

$-x_{1,1} \leqslant 0$: Similarly, face $\mathtt{F}(-x_{1,1} \leqslant 0)$ contains only vertices $\boldsymbol{v}$ with first row $\boldsymbol{v}_{1,*} = (0,0)$. Hence also $\mathtt{F}(-x_{1,1} \leqslant 0) \subseteq \mathtt{F}(-x_{1,1} + x_{1,2} \leqslant 0)$ and $-x_{1,1} \leqslant 0$ cannot be a facet defining inequality.

Now consider inequality $x_{i,j} \leqslant 1$ with $(i,j) \neq (1,2)$. The set of vertices on face $\mathtt{F}(x_{i,j} \leqslant 1)$ is clearly

$$\mathcal{V}[x_{i,j} \leqslant 1] = \{\boldsymbol{x} \in \mathcal{V}_{\mathbf{O}_{p,2}} \mid x_{i,j} = 1\}.$$

Since $(i,j) \neq (1,2)$, there is for each $k \in [p]$ a vector in $\mathcal{V}[x_{i,j} \leqslant 1]$ with critical row k; in particular, there is at least one vector $\boldsymbol{v}$ with $\mathrm{crit}(\boldsymbol{v}) = 1$. One can modify this vector $\boldsymbol{v}$ at any position $(s,t) \neq (i,j)$, $t > 1$ by component $\widetilde{v}_{s,t} := 1 - v_{s,t}$ and obtains a vector that is also in $\mathcal{V}[x_{i,j} \leqslant 1]$. The vector $\boldsymbol{w} := \mathbb{1}_{p,2}$ is also in $\mathcal{V}[x_{i,j} \leqslant 1]$, and if $(i,j) \neq (1,1)$, then we can additionally modify $\boldsymbol{w}$ by components $\widetilde{w}_{1,*} := (0,0)$ to obtain another vector in $\mathcal{V}[x_{i,j} \leqslant 1]$. Hence, we can find for every position $(s,t) \neq (i,j)$ at least two vectors $\boldsymbol{u}, \boldsymbol{u}'$ in $\mathcal{V}[x_{i,j} \leqslant 1]$ with $u_{s,t} = 1$ and $u'_{s,t} = 0$. Therefore, $\mathtt{F}(x_{i,j} \leqslant 1)$ cannot be contained in another trivial inequality.

For any valued block inequality $\boldsymbol{ax} \leqslant b$, it holds that $\mathcal{V}[\boldsymbol{ax} \leqslant b]$ does not contain a vector $\boldsymbol{v}$ with $\mathrm{crit}(\boldsymbol{v}) = \mathrm{base}(\boldsymbol{a})$ (Observation 4.19). So it follows from above that $\mathtt{F}(x_{i,j} \leqslant 1)$ can also not be contained in a valued block inequality.

Case (ii) is analogously. □

4.2 Facial Description of Orbisacks II (Sequential Knapsack)

A classic in combinatorial optimization is the *integer knapsack problem*:

$$\max \sum_{i=1}^{n} p_i x_i$$
$$\text{such that } \sum_{i=1}^{n} w_i x_i \leqslant c \ \text{ and } \ x_i \in \mathbb{N} \ \ \forall i \in [n],$$

where $w_i \in \mathbb{N}$ is the weight of item i, $p_i \in \mathbb{R}$ is the profit associated with item i and $c \in \mathbb{N}$ is the capacity of the knapsack. If for all $i \in [n]$, the values of x_i are additionally required to be in $[s_i]_0 \subset \mathbb{N}$ with some $s_i \in \mathbb{N}$, then the knapsack problem is called *bounded*. If $s_i = 1$ for all $i \in [n]$, we speak of a 0/1-*knapsack problem*. An integer knapsack problem is called *sequential* if one can order (after possible renumbering of components) the weights such that

$$0 < w_1 \leqslant w_2 \leqslant \ldots \leqslant w_m$$

and for any pair of successive weights w_{i-1}, w_i, it holds that $\frac{w_i}{w_{i-1}} \in \mathbb{N}$. Note that after possibly scaling weights and capacity, we can for sequential knapsack problems w.l.o.g. assume that $w_1 = 1$. The convex hull of the feasible solutions of a (sequential and) bounded integer knapsack problem is called the *(sequential) knapsack polytope*.

As has been observed in Lemma 3.16, the orbisack is isomorphic to the following 0/1-knapsack polytope

$$\mathbf{P}_{\mathbf{O}_{p,2}} := \operatorname{conv}\left(\left\{\boldsymbol{y} \in \mathcal{M}_{p,q} \,\middle|\, \sum_{i=1}^{p} 2^{i-1}(y_{i,1} + y_{i,2}) \leqslant 2^p - 1\right\}\right),$$

determined by knapsack inequality (3.2). The weights $w_{i,j} = 2^{i-1}$ as well as $c = 2^p - 1$ are in $\mathbb{N}$ and $y_{i,j} \in [1]_0$ for all $(i,j) \in [p] \times [2]$. Moreover, after ordering the weights as follows:

$$0 < \underbrace{w_{1,1}}_{=2^0} = \underbrace{w_{1,2}}_{2^0} \leqslant \underbrace{w_{2,1}}_{=2^1} = \underbrace{w_{2,2}}_{=2^1} \leqslant \ldots \leqslant \underbrace{w_{p,1}}_{=2^{p-1}} = \underbrace{w_{p,2}}_{=2^{p-1}},$$

it is clear that $\frac{w_{i,2}}{w_{i,1}} = 1 \in \mathbb{N}$ and $\frac{w_{i,1}}{w_{i-1,2}} = 2 \in \mathbb{N}$ for all $i \in [p]$. Hence, $\mathbf{P}_{\mathbf{O}_{p,2}}$ is in fact a sequential 0/1-knapsack polytope.

Observation 4.23 Using the isomorphism defined by $y_{i,1} = 1 - x_{p-i+1,1}$ and $y_{i,2} = x_{p-i+1,2}$, it is easy to see that any inequality $\boldsymbol{ax} \leqslant b$ valid for $\mathbf{O}_{p,2}$ transforms into an inequality $\boldsymbol{a'y} \leqslant b'$ valid for $\mathbf{P}_{\mathbf{O}_{p,2}}$ by setting

$$a'_{i,j} := \begin{cases} -a_{p-i+1,j}, & \text{if } j = 1 \\ a_{p-i+1,j}, & \text{if } j = 2 \end{cases}$$

for all $(i,j) \in [p] \times [2]$ and fixing right-hand side

$$b' := b - \sum_{i=1}^{p} a_{i,1}.$$

In particular, trivial inequalities transform as follows:

$$\begin{array}{rcrcrcr} x_{i,1} \leqslant 1 & \rightsquigarrow & -y_{i,1} \leqslant 0 & \qquad & x_{i,2} \leqslant 1 & \rightsquigarrow & y_{i,2} \leqslant 1 \\ -x_{i,1} \leqslant 0 & \rightsquigarrow & y_{i,1} \leqslant 1 & & -x_{i,2} \leqslant 0 & \rightsquigarrow & -y_{i,2} \leqslant 0 \end{array}$$

Example 4.24 From Proposition 4.20, we get that the inequality to the left is facet defining for orbisack $\mathbf{O}_{5,2}$. To the right, there is the corresponding facet defining inequality for $\mathbf{P}_{\mathbf{O}_{5,2}}$.

$$\begin{array}{|r|r|}\hline -4 & 4 \\ \hline 0 & 4 \\ \hline -2 & 2 \\ \hline -1 & 1 \\ \hline -1 & 1 \\ \hline \end{array} \leqslant 4 \qquad \begin{array}{|r|r|}\hline 1 & 1 \\ \hline 1 & 1 \\ \hline 2 & 2 \\ \hline 0 & 4 \\ \hline 4 & 4 \\ \hline \end{array} \leqslant 12$$

To the left, there is a vertex of $\mathbf{O}_{5,2}$ and to the right the corresponding vertex of $\mathbf{P}_{\mathbf{O}_{5,2}}$. It is easy to check that both are contained in the respective facets from above.

$$\begin{array}{|c|c|}\hline 0 & 0 \\ \hline 1 & 1 \\ \hline 1 & 0 \\ \hline 0 & 1 \\ \hline 0 & 1 \\ \hline \end{array} \qquad \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 1 & 1 \\ \hline 0 & 0 \\ \hline 0 & 1 \\ \hline 1 & 0 \\ \hline \end{array}$$

In 1998, Robert Weismantel and Yves Pochet gave an inductive scheme to compute the linear description of a sequential integer knapsack polytope $\mathbf{P}$ in general ([117]). We will in the following briefly present their main ideas.

Let $\mathbf{P} \subset \mathbb{R}^n$. For some objective vector $\boldsymbol{p} \in \mathbb{R}^n$, Weismantel and Pochet transform polytope $\mathbf{P}$ into a polytope $\widetilde{\mathbf{P}}_{\boldsymbol{p}}$ by combining certain subsets of items, called (item) blocks[1]. The transformed polytope has two nice properties:

(a) Optimization over $\mathbf{P}$ with respect to profit vector $\boldsymbol{p}$ is equivalent to optimization over $\widetilde{\mathbf{P}}_{\boldsymbol{p}}$ with respect to a modified profit vector $\widetilde{\boldsymbol{p}}$; i.e. there is an optimal solution $\boldsymbol{y}^\star$ to the original problem if and only if there is an optimal solution $\boldsymbol{z}^\star$ to the modified problem and $\langle \boldsymbol{y}^\star, \boldsymbol{p}\rangle = \langle \boldsymbol{z}^\star, \widetilde{\boldsymbol{p}}\rangle$.

(b) Using the standard dynamic programming scheme for knapsacks developed by Bellman and Dantzig ([14, 28, 15], overview for instance in [71]), one can characterize the optimal solutions over $\widetilde{\mathbf{P}}_{\boldsymbol{p}}$ with respect to profit vector $\widetilde{\boldsymbol{p}}$. Weismantel and Pochet show that it is also possible — parallel to the run of the dynamic programming algorithm — to recursively construct an inequality valid for $\widetilde{\mathbf{P}}_{\boldsymbol{p}}$ that is satisfied at equality by all optimal solutions to the problem of maximizing $\langle \boldsymbol{z}, \widetilde{\boldsymbol{p}}\rangle$ over $\widetilde{\mathbf{P}}_{\boldsymbol{p}}$.

Using (a), one can then transform the obtained inequality for $\widetilde{\mathbf{P}}_{\boldsymbol{p}}$ into an inequality valid for $\mathbf{P}$. Moreover, it turns out that the profit vectors can be partitioned in equivalence classes with respect to the inequalities; any pair of profit vectors from the same class yields the same inequality. So, one ends up with a finite set of inequalities. However, this set is pretty big: it contains at least $2^n n!$ inequalities.

To show that this set provides a full description of $\mathbf{P}$, Weismantel and Pochet show that any inequality transformed originating from $\widetilde{\mathbf{P}}_{\boldsymbol{p}}$ as described above is satisfied at equality by all optimal solutions to the problem of optimizing $\langle \boldsymbol{y}, \boldsymbol{p}\rangle$ over polytope $\mathbf{P}$. As this is true for arbitrary profit vectors $\boldsymbol{p}$, this implies that the set of inequalities valid for $\mathbf{P}$ must contain in particular all facet defining inequalities and therefore provides a complete linear description (this latter idea is due to Lovász, see [77]).

We will in the following reproduce the main definitions together with some results concerning orbisacks as far as they are needed to compute the linear description of the sequential knapsack polytope $\mathbf{P}_{\mathbf{O}_{p,2}}$ associated with the orbisack $\mathbf{O}_{p,2}$. In this special case, the knapsack inequality is $\langle \boldsymbol{w}, \boldsymbol{y}\rangle \leqslant c$ with weights $w_{i,j} = 2^{i-1}$ for all $(i,j) \in [p] \times [2]$ and capacity $c = 2^p - 1$. Note that for each $i \in [p]$, both items $(i,1)$ and $(i,2)$ have the same weight.

The main concept is the combination of items into *item blocks*.

Definition 4.25 Let $\mathcal{P}$ be the sequential knapsack problem

$$\max \sum_{i=1}^{n} p_i y_i \qquad \text{such that } \sum_{i=1}^{n} w_i y_i \leqslant c$$
$$y_i \in [s_i]_0 \quad \forall i \in [n]$$

Let $\mathcal{B} \subseteq [n]$ and $b_{\mathcal{B}} := \min(\mathcal{B})$. Set $\mathcal{B}$ is called an *item block* if for every $k \in \mathcal{B} \smallsetminus \{b_{\mathcal{B}}\}$, it holds that

$$w_k \leqslant w_{b_{\mathcal{B}}} + \sum_{\substack{i \in \mathcal{B} \\ i<k}} s_i w_i. \tag{*}$$

Lemma 4.26 *For orbisacks, a subset $\mathcal{B} \subseteq [p] \times [2]$ is not an item block if and only if there is a $k \in [2..p-1]$ such that $\mathcal{B}$ can be partitioned into two nonempty subsets*

[1] In fact, Weismantel and Pochet speak of *blocks* only. We use the term *item block* instead to avoid confusion with the blocks from block inequalities, see Definition 4.12.

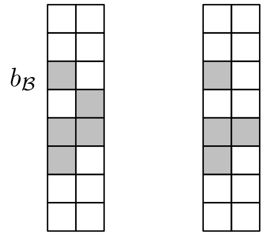

Figure 4.2: Example for an item block in the case of orbisacks (left). The gray marked indices are in the block. The example to the right is no item block.

$\mathcal{B}^< \subseteq [k-1] \times [2]$ and $\mathcal{B}^> \subseteq [k+1] \times [2]$. In other words: any subset $\mathcal{B}$ of $[p] \times [2]$ that does not "skip" both items $(k, 1)$ and $(k, 2)$ in some row k is an item block. (See figure 4.2 for an example.)

Proof. (i) Let $b^<$ and $b^>$ be the smallest elements in $\mathcal{B}^<$ and $\mathcal{B}^>$, respectively. Assume, set $\mathcal{B} = \mathcal{B}^< \cup \mathcal{B}^>$ and $\mathcal{B}^<, \mathcal{B}^> \neq \varnothing$. As $\mathcal{B}$ is a block,

$$w_{b^<} + \sum_{b \in \mathcal{B}^<} w_b \geqslant w_{b^>}$$

must hold. But the left-hand side is maximally $2^k - 1$, while the right-hand side is at least 2^k. Contradiction.

(ii) Let $\mathcal{B} \subseteq [p] \times [2]$. Let (s, t) be the smallest item in $\mathcal{B}$ and let $(k, \ell) > (s, t)$ be some other item in $\mathcal{B}$. If $k = s$, then $w_{s,t} = w_{k,\ell}$ and inequality (*) is satisfied. So we can assume that $k > s$. Let $\mathcal{W} := \{w_{i,j} \mid (i, j) < (k, \ell),\ (i, j) \in \mathcal{B}\}$. From (i), we get that the smallest possible set $\mathcal{W}$ is

$$\mathcal{W} = \{2^{i-1} \mid i \in [s..k-1]\}.$$

Hence, the right-hand side of inequality (*) is at least

$$2^{s-1} + \sum_{i=s}^{k-1} 2^{i-1} = 2^{k-1}.$$

On the other hand, $w_k = 2^{k-1}$. Hence, inequality (*) is satisfied.

□

4.2.1 Computation of Inequalities for $\mathbf{P}_{\mathbf{O}_{p,2}}$

Since we are mainly interested in the computation of coefficients and the right-hand side of these inequalities, we skip the details of transformation and generation of inequalities and refer for the details to the paper of Weismantel and Pochet.

Instead, we proceed with the definition of the set of variables which is needed to construct an algorithm to generate the inequalities for $\mathbf{P}$. (Note that the paper of Weismantel and Pochet contains some typos which are corrected here.)

We will in the following run through a family of item blocks $\mathcal{B}$. For each item block $\mathcal{B}_i$ in $\mathcal{B}$, we define:

Definition 4.27 Let $\ell := \operatorname{argmin}_i(\{w_i \mid i \in \mathcal{B}_j\})$. We define the *weight* of item block $\mathcal{B}_j$ by $\widetilde{w}_j := w_\ell$ and the *normalized weight* of $\mathcal{B}_j$ by $\widetilde{w}'_j := \frac{\widetilde{w}_j}{\widetilde{w}_1}$. Moreover, the *multiplicity* of item block $\mathcal{B}_j$ is denoted by $\widetilde{u}_j := \frac{1}{\widetilde{w}_j} \sum_{i \in \mathcal{B}} s_i w_i$. The *profit* of a item block is $\widetilde{p}_j := p_\ell$.

The inequalities to be computed are of the form

$$\sum_{j=1}^{m} \frac{d_j}{\widetilde{w}_j} \sum_{i \in \mathcal{B}_j} w_i y_i \leqslant g_m(N),$$

with a right-hand side depending on some $N \in \mathbb{N}$ that will be specified later.

- For $j = 1$ and any $\gamma \in \mathbb{N}$, we define

$$d_1 := 1 \qquad\qquad g_1(\gamma) := \min(\{\widetilde{u}_1, \gamma\}).$$

- For $j \in [2..m]$ and any $\gamma \in \mathbb{N}$, we define

$$\lambda_j^N := N \bmod \widetilde{w}'_j$$

and

$$N_j := \begin{cases} \lambda_j^N, & \text{if } \sum_{i \in \Delta_j} \widetilde{u}_i \widetilde{w}'_i < \lambda_j^N \\ \max\left(\left\{k \in \mathbb{N} \,\middle|\, k \leqslant \sum_{i \in \Delta_j} \widetilde{u}_i \widetilde{w}'_i \text{ and } k \bmod \widetilde{w}_j = \lambda_j^N \right\}\right), & \text{if } \sum_{i \in \Delta_j} \widetilde{u}_i \widetilde{w}'_i \geqslant \lambda_j^N. \end{cases}$$

The values of $g_j(\gamma)$ are defined by

$$g_j(\gamma) := \begin{cases} g_{j-1}(N_j) + \mu_j(\gamma) d_j, & \text{if } 0 \leqslant \mu_j(\gamma) \leqslant \widetilde{u}_j \\ g_{j-1}(\gamma), & \text{if } \mu_j(\gamma) < 0 \\ g_{j-1}(\gamma - \widetilde{w}'_j \widetilde{u}_j) + \widetilde{u}_j d_j, & \text{if } \mu_j(\gamma) > \widetilde{u}_j. \end{cases}$$

and are based on the value of

$$\mu_j(\gamma) := \frac{1}{\widetilde{w}'_j}(\gamma - N_j).$$

Last, the coefficients d_j in the inequality are computed recursively by

$$d_j := g_{j-1}(N_j + \widetilde{w}'_j) - g_{j-1}(N_j).$$

These definitions also rely on values Δ_j. In fact, this is where the objective $\boldsymbol{p}$ comes in, as Weismantel and Pochet define a first version of Δ_j by

$$\Delta_j := \left\{ i \in [j-1] \,\middle|\, \frac{\widetilde{p}_i}{\widetilde{w}'_i} > \frac{\widetilde{p}_j}{\widetilde{w}'_j} \right\} \text{ for all } j \in [m].$$

It is easy to see that the size of the coefficients of $\boldsymbol{p}$ has an immediate influence only on the *order* of indices. As any possible profit vector is considered when constructing the linear hull of $\mathbf{P}$, it is therefore possible to replace this definition by another one based on permutations π of $[m]$. This leads to the following definition:

Definition 4.28 Set Δ_j can independently of $\boldsymbol{p}$ be defined by

$$\Delta_j := \{i \in [j-1] \mid \pi(i) < \pi(j)\} \text{ for all } j \in [m],$$

where π is some permutation of $[m]$.

Now, the main result of Weismantel and Pochet can be formulated as follows:

Theorem 4.29 ([117]) *The following set of inequalities provides a complete linear description for polytope* $\mathbf{P}$*:*

$$y_k \geqslant 0 \quad \forall k \in [n]$$
$$\sum_{j=1}^{m} \frac{d_j}{\widetilde{w}_j} \sum_{i \in \mathcal{B}_j} w_i y_i \leqslant g_m(N),$$

where $N := (\lfloor \frac{c}{\widetilde{w}_1} \rfloor)$*, family* $\mathcal{B}$ *is any family of item blocks* $\mathcal{B}_i$ *partitioning any subset of items* $\mathcal{W} \subseteq [n]$*, and* g_i *and* d_i*,* $i \in [m]$*, are defined on the basis of any permutation* π *of* $[m]$*.*

Corollary 4.30 *The following set of inequalities provides a complete linear description for polytope* $\mathbf{P}_{\mathbf{O}_{p,2}}$*:*

$$y_k \geqslant 0 \quad \forall k \in [n]$$
$$\sum_{k=1}^{m} \frac{d_k}{\widetilde{w}_k} \sum_{(i,j) \in \mathcal{B}_k} w_{i,j} y_{i,j} \leqslant g_m(N), \tag{*}$$

where $w_{i,j} = 2^{i-1}$*,* $N := (\lfloor \frac{2^p-1}{\widetilde{w}_1} \rfloor)$*, family* $\mathcal{B}$ *is any family of item blocks* $\mathcal{B}_i$ *as described in Lemma 4.26 partitioning any subset of items* $\mathcal{W} \subseteq [p] \times [2]$*, and* g_i *and* d_i *are defined for* $i \in [m]$ *on the basis of any permutation* π *of* $[m]$*.*

It is obvious that all trivial inequalities

$$y_{i,j} \leqslant 1$$

with $(i,j) \in [p] \times [2]$ can be easily obtained by choosing item block system $\mathcal{B} = \{\mathcal{B}_1\}$ containing one single item block $\mathcal{B}_1$ with $\mathcal{B}_1 = \{(i,j)\}$.

It requires a bit more work to see that the valued block inequalities are among the inequalities (*). This will be done in the following.

Assume, $\boldsymbol{a}\boldsymbol{y} \leqslant b$ is a valued block inequality for $\mathbf{O}_{p,2}$ with basement $\beta := \mathrm{base}(\boldsymbol{a})$, and let $\boldsymbol{a}'\boldsymbol{y} \leqslant b'$ be the corresponding inequality for $\mathbf{P}_{\mathbf{O}_{p,2}}$ as described in Observation 4.23. We choose

$$\begin{aligned}
\pi[j] &:= \beta - j + 1 \\
\mathcal{W} &:= \mathrm{supp}(\boldsymbol{a}') \\
\mathcal{B}_j &:= \begin{cases} \{(p-\beta+j,1)\} & \text{if } \boldsymbol{a}_{\beta-j+1,*} \text{ is } n\text{-row} \\ \{(p-\beta+j,2)\} & \text{if } \boldsymbol{a}_{\beta-j+1,*} \text{ is } p\text{-row} \\ \{(p-\beta+j,1),(p-\beta+i,2)\} & \text{if } \boldsymbol{a}_{\beta-j+1,*} \text{ is } n\text{-}p\text{-row,} \end{cases}
\end{aligned}$$

for all $j \in [\beta]$. This choice implies the following:

$$\begin{aligned}
\Delta_j &= \varnothing \\
\widetilde{w}_j &= 2^{p-\beta+j-1} \\
\widetilde{w}'_j &= 2^{j-1} \\
N &= \Big\lfloor \frac{c}{\widetilde{w}_1} \Big\rfloor = \Big\lfloor \frac{2^p-1}{2^{p-\beta}} \Big\rfloor = 2^\beta - 1 \\
\widetilde{u}_j &= \frac{1}{\widetilde{w}_j} \sum_{(k,\ell) \in \mathcal{B}_j} 2^{k-1} = |\mathcal{B}_j| = \begin{cases} 2, & \text{if } \boldsymbol{a}_{k,*} \text{ is n-p-row} \\ 1, & \text{otherwise} \end{cases} \\
\sum_{i \in \Delta_j} \widetilde{u}_i \widetilde{w}'_i &= 0,
\end{aligned}$$

for all $j \in [\beta]$.

Observation 4.31

- By choice of $\mathcal{B}$, $\mathcal{W}$ is decomposed into β item blocks.
- By choice of $\mathcal{B}$, block $|\mathcal{B}_1| = 2$ and block $|\mathcal{B}_\beta| = 2$.
- For any element (k, ℓ) in block $\mathcal{B}_j$, $i \in [\beta]$, it holds that $\frac{w_{k,\ell}}{\widetilde{w}_j} = 1$. Therefore, inequality (*) can be simplified to

$$\sum_{j=1}^{m} d_j \sum_{(k,\ell)\in\mathcal{B}_j} y_{k,\ell} \leqslant g_m(2^\beta - 1).$$

Example 4.32 In this example, $p = 7$ and $\beta = 5$.

$\boldsymbol{a}\boldsymbol{x} \leqslant b$

-2	2	$\leqslant 2$
-2	0	
0	2	
-1	1	
-1	1	
0	0	
0	0	

$\boldsymbol{a}'\boldsymbol{x} \leqslant b'$

0	0	$\leqslant 8$
0	0	
1	1	
1	1	
0	2	
2	0	
2	2	

$\boldsymbol{w}$

1	1	
2	2	
4	4	$\mathcal{B}_1 = \{(3,1),(3,2)\}$
8	8	$\mathcal{B}_2 = \{(4,1),(4,2)\}$
16	16	$\mathcal{B}_3 = \{(5,2)\}$
32	32	$\mathcal{B}_4 = \{(6,1)\}$
64	64	$\mathcal{B}_5 = \{(7,1),(7,2)\}$

We will now compute d_j for all $j \in [\beta]$ as well as $g_\beta(N)$. For $j > 1$, the coefficients d_j depend on N_j. Hence, we have to compute this value first.

Lemma 4.33 *If $N = 2^\beta - 1$, then $N_j = 2^{j-1} - 1$ holds for any $j \in [2..\beta]$.*

Proof. With $N = 2^\beta - 1$, we obtain for any $j \in [2..\beta]$

$$\lambda_j^N = N \bmod \widetilde{w}'_j = 2^\beta - 1 \bmod 2^{j-1} = 2^{j-1} - 1.$$

On the other hand, $\sum_{i\in\Delta_j} \widetilde{u}_i\widetilde{w}_i = 0 < \lambda_j^N$ for all $j \in [\beta]$. Therefore, $N_j = \lambda_j^N$ by definition. □

Observation 4.34 In fact, Lemma 4.33 is independent of the choice of permutation π. Assume there is a permutation π' inducing sets Δ'_j and a index $j^\star$ such that

$$\sum_{i\in\Delta'_{j^\star}} \widetilde{u}_i\widetilde{w}_i \geqslant \lambda_{j^\star}^N$$

This implies by definition that

$$N_{j^\star} = \max\left(\left\{k \in \mathbb{N} \,\middle|\, k \leqslant \sum_{i\in\Delta'_{j^\star}} \widetilde{u}_i\widetilde{w}'_i \text{ and } k \bmod \widetilde{w}_{j^\star} = \lambda_{j^\star}^N\right\}\right).$$

However,

$$\sum_{i\in\Delta'_{j^\star}} \widetilde{u}_i\widetilde{w}'_i \leqslant \sum_{i=1}^{j^\star-1} \widetilde{u}_i\widetilde{w}'_i \leqslant 2\sum_{i=1}^{j^\star-1} \widetilde{w}'_i = 2\sum_{i=1}^{j^\star-1} 2^{i-1} = 2^{j^\star} - 2 <$$
$$< 2^{j^\star-1} + 2^{j^\star-1} - 1 \leqslant 2^{p-\beta+j^\star-1} + 2^{j^\star-1} - 1 \leqslant \widetilde{w}_{j^\star} + \lambda_{j^\star}^N$$

Hence, $k < \widetilde{w}_{j^\star} + \lambda_{j^\star}^N$, which implies that $N_{j^\star} = \lambda_{j^\star}^N$.

Lemma 4.35 *Let $N = 2^\beta - 1$ and define for $\beta \geqslant i \geqslant j > 1$ the scalar*

$$\gamma_{i,j} := \begin{cases} N_i + \widetilde{w}'_i - \sum_{k=j}^{i-1} \widetilde{w}'_k \widetilde{u}_k & \text{if } i > j \\ N_i + \widetilde{w}'_i & \text{if } i = j. \end{cases}$$

Then

$$g_{j-1}(\gamma_{i,j}) := \begin{cases} g_{i-2}(\gamma_{i,j-1}) + \widetilde{u}_{j-1} d_{j-1} & \text{if } j > 2 \\ 2 & \text{if } j = 2. \end{cases}$$

Proof. First let $j = 2$. From the definition of $g_1(\gamma)$ follows that

$$g_1(\gamma_{i,2}) = \min(\{\widetilde{u}_1,\ \gamma_{i,2}\}).$$

But

$$\gamma_{i,2} = N_i + \widetilde{w}'_i - \sum_{k=2}^{i-1} \widetilde{w}_k u_k = 2^{i-1} - 1 + 2^{i-1} - \sum_{k=2}^{i-1} 2^{k-1} |\mathcal{B}_k| \geqslant$$
$$\geqslant 2^i - 1 - \sum_{k=2}^{i-1} 2^k = 2^i - 1 - (2^i - 1 - 3) = 3,$$

which gives the statement for $j = 2$.

For $j > 2$, we have to show that $\mu_{j-1}(\gamma_{i,j}) > \widetilde{u}_{j-1}$ for any $\beta \geqslant i \geqslant j > 2$. This can be done using Lemma 4.33:

$$\mu_{j-1}(\gamma_{i,j}) = \frac{1}{\widetilde{w}'_{j-1}}(\gamma_{i,j} - N_{j-1}) =$$
$$= \frac{1}{2^{j-2}}(2^{i-1} - 1 + 2^{i-1} - \sum_{k=j}^{i-1} 2^{k-1}|\mathcal{B}_k| - 2^{j-2} + 1) \geqslant$$
$$\geqslant 2^{2-j}(2^i - 2\sum_{k=j}^{i-1} 2^{k-1} - 2^{j-2}) = 2^{2-j}(2^i - 2(2^{i-1} - 2^{j-1}) - 2^{j-2}) =$$
$$= \frac{3}{4} 2^2 = 3 > 2 \geqslant \widetilde{u}_{j-1}.$$

Now, the statement follows from the definition of $g_j(\gamma)$. □

Corollary 4.36 *We can now recursively compute $g_{j-1}(\gamma_{j,j})$ for any $\beta \geqslant j > 1$:*

$$g_{j-1}(\gamma_{j,j}) = g_{j-2}(\gamma_{j,j-1}) + \widetilde{u}_{j-1} d_{j-1} = \ldots =$$
$$= g_1(\gamma_{j,2}) + \sum_{k=2}^{j-1} \widetilde{u}_k d_k = 2 + \sum_{k=2}^{j-1} |\mathcal{B}_k| d_k = \sum_{k=1}^{j-1} |\mathcal{B}_k| d_k,$$

since $|\mathcal{B}_1| = 2$ by Observation 4.31.

Lemma 4.37 *Let $N = 2^\beta - 1$. Then*

$$g_{j-1}(N_j) := \begin{cases} g_{i-2}(N_{j-1}) + d_{j-1} & \text{if } j > 2 \\ 1 & \text{if } j = 2. \end{cases}$$

Proof. For $j = 2$, $N_2 = 2^1 - 1 = 1$. Hence, $g_1(N_2) = \min(\{2,\ N_2\}) = 1$.
For $j > 2$, we will show that $\mu_{j-1} = 1$:

$$\mu_{j-1} = \frac{1}{\widetilde{w}_{j-1}}(N_j - N_{j-1}) = 2^{2-j}(2^{j-1} - 1 - 2^{j-2} + 1) = 2^{2-j}(2^{j-2}) = 1.$$

The statement follows then from the definition of $g_i(\gamma)$. □

Corollary 4.38 *Lemma 4.37 enables us to recursively compute $g_{j-1}(N_j)$ for any $\beta \geqslant j > 1$:*

$$g_{j-1}(N_j) = g_{j-2}(N_{j-1}) + d_{j-1} = \ldots = g_1(N_2) + \sum_{k=2}^{j-1} d_k = 1 + \sum_{k=2}^{j-1} d_k = \sum_{k=1}^{j-1} d_k,$$

since $d_1 = 1$ by definition.

Corollaries 4.36 and 4.38 together make it now possible to compute d_j for $\beta \geqslant j > 1$ using the definition for d_j:

$$d_j = g_{j-1}(N_j + \widetilde{w}_j) - g_{j-1}(N_j) = \sum_{k=1}^{j-1} d_k(|\mathcal{B}_k| - 1).$$

It is easy to see that this implies that the coefficients of inequality (**) are those of a transformed facet defining inequality of the orbisack. For $j = 2$, we get that

$$d_2 = d_1(|\mathcal{B}_1| - 1) = 1$$

and for $j > 2$, we obtain

$$\begin{aligned} d_j = \sum_{k=1}^{j-1} d_k(|\mathcal{B}_k| - 1) = \\ = d_{j-1}(|\mathcal{B}_{j-1}| - 1) + \sum_{k=1}^{j-2} d_k(|\mathcal{B}_k| - 1) = d_{j-1}(|\mathcal{B}_{j-1}| - 1) + d_{j-1} = \\ = d_{j-1}|\mathcal{B}_{j-1}|. \end{aligned}$$

So, for $j > 2$, $d_j = 2d_{j-1}$ holds if $|\mathcal{B}_{j-1}|$ corresponds to a n-p-row; otherwise, $d_j = d_{j-1}$. It remains to compute the right-hand side $g_\beta(N)$ of inequality (**).

Lemma 4.39 *Let $N = 2^\beta - 1$. Then*

$$g_\beta(N) = g_{\beta-1}(N_j) + d_\beta.$$

Proof. Again, we use Lemma 4.33 to obtain that $\mu_\beta(N)$ is computed as

$$0 < \mu_\beta(N) = \frac{1}{\widetilde{w}'_\beta}(N - N_\beta) = 2^{1-\beta}(2^\beta - 1 - (2^{\beta-1} - 1)) = 2^0 = 1 \leqslant \widetilde{u}_j.$$

□

Lemmas 4.39 and 4.37 enable us to recursively compute the right-hand side $g_\beta(N)$ as follows:

$$g_\beta(N) = g_{\beta-1}(N_\beta) + d_\beta = \ldots = g_1(N_2) + \sum_{i=2}^{\beta} d_i = \sum_{i=1}^{\beta} d_i.$$

So far, we know that the set of inequalities from Theorem 4.29 contains all facet defining inequalities for the orbisack. However, the set contains more inequalities. We close this section with some examples. For all examples, $p = 6$.

Example 4.40

Item blocks: yield the following inequalities:

$$\begin{aligned} \mathcal{B}_1 &:= \{(1,1),\ldots,(6,1)\} \\ \mathcal{B}_2 &:= \{(1,2),\ldots,(6,2)\} \\ \pi(i) &:= m-i+1 \quad \forall i \in [m] \end{aligned} \qquad \begin{array}{|c|c|}\hline 1 & 1 \\\hline 2 & 2 \\\hline 4 & 4 \\\hline 8 & 8 \\\hline 16 & 16 \\\hline 32 & 32 \\\hline \end{array} \leqslant 63$$

$$\begin{aligned} \mathcal{B}_1 &:= \{(1,1),(2,1)\} \\ \mathcal{B}_2 &:= \{(1,2),(2,2)\} \\ \mathcal{B}_3 &:= \{(3,1),(4,1)\} \\ \mathcal{B}_4 &:= \{(3,2),(4,2)\} \\ \mathcal{B}_5 &:= \{(5,1),(6,1)\} \\ \mathcal{B}_6 &:= \{(5,2),(6,2)\} \\ \pi(i) &:= m-i+1 \quad \forall i \in [m] \end{aligned} \qquad \begin{array}{|c|c|}\hline 1 & 1 \\\hline 2 & 2 \\\hline 3 & 3 \\\hline 6 & 6 \\\hline 12 & 12 \\\hline 24 & 24 \\\hline \end{array} \leqslant 48$$

$$\begin{aligned} \mathcal{B}_i &\quad \text{same as above} \\ \pi(i) &:= i \quad \forall i \in [m] \end{aligned} \qquad \begin{array}{|c|c|}\hline 1 & 0 \\\hline 2 & 0 \\\hline 0 & 0 \\\hline 0 & 0 \\\hline 0 & 0 \\\hline 0 & 0 \\\hline \end{array} \leqslant 3$$

$$\begin{aligned} \mathcal{B}_1 &:= \{(3,1),(3,2)\} \\ \mathcal{B}_2 &:= \{\qquad\quad (4,2)\} \\ \mathcal{B}_3 &:= \{(5,1),(5,2)\} \\ \pi(i) &:= m-i+1 \quad \forall i \in [m] \end{aligned} \qquad \begin{array}{|c|c|}\hline 0 & 0 \\\hline 0 & 0 \\\hline 1 & 1 \\\hline 0 & 1 \\\hline 1 & 1 \\\hline 0 & 0 \\\hline \end{array} \leqslant 5$$

Obviously, one can easily generate inequalities different from the transformed facet defining inequalities for the orbisack.

It is an open question whether it is possible to describe the set of inequalities from Corollary 4.30 in a more appealing way than is done there. Moreover, the second and third inequality in Example 4.40 indicate that it is not possible to read off the item block system associated to an inequality from the coefficients of the inequality. Note that the set of inequalities from Corollary 4.30 is in $\mathcal{NP}$.

4.3 Complete Description of Orbisacks III (Proof by Faithful Sectioning)

4.3.1 Extended Formulations for Orbitopes

In the following, we will define three different extended formulations for orbisacks. The first two formulations are specific for orbisacks, i.e. restricted to $q = 2$ columns. The third formulation is based on the dynamic programming algorithms for orbitopes in general, i.e. it relies on the hyperpath polytope and the description of vertices of the orbitope as hypergraphs. So in this case, we obtain (also) an extended formulation for orbitopes with more than two columns.

At the end of the section, we will give a short overview of the extended formulations developed so far and the relationships between them.

4.3.1.1 The $\mathbf{P}^{x,y}$-Formulation

For this extended formulation, we append to each vertex $\boldsymbol{x}$ of an orbisack some 0/1-vector storing information about the position of the critical row of $\boldsymbol{x}$. More precisely, we define for each vertex $\boldsymbol{x}$ of the orbisack $\mathbf{O}_{p,2}$ some vector $\boldsymbol{y}(\boldsymbol{x}) \in \{0,1\}^{[p]}$ by

$$\boldsymbol{y}(\boldsymbol{x}) := \begin{cases} \mathbb{e}^{\operatorname{crit}(\boldsymbol{x})}, & \text{if } \operatorname{crit}(\boldsymbol{x}) < p+1 \\ \mathbb{0}, & \text{if } \operatorname{crit}(\boldsymbol{x}) = p+1 \end{cases}$$

Polytope $\mathbf{P}^{x,y}$ is defined as

$$\mathbf{P}^{x,y} := \operatorname{conv}(\{(\boldsymbol{x}, \boldsymbol{y}(\boldsymbol{x})) \in \mathbb{R}^{[p]\times[2]} \times \mathbb{R}^{[p]} \mid \boldsymbol{x} \text{ vertex of } \mathbf{O}_{p,2}\})$$

(Simplifying notation, we allow to write $(\boldsymbol{x}, \boldsymbol{y})$ instead of $\big(\boldsymbol{x}, \boldsymbol{y}(\boldsymbol{x})\big)$.)

Lemma 4.41 *A linear description of $\mathbf{P}^{x,y}$ together with an appropriate linear projection provides an extended formulation for the orbisack.*

Proof. The projection from $\mathbf{P}^{x,y}$ to $\mathbf{O}_{p,2}$ is simply the orthogonal projection

$$\sigma^{x,y} : \mathbb{R}^{[p]\times[2]} \times \mathbb{R}^{[p]} \to \mathbb{R}^{[p]\times[2]}, \quad (\boldsymbol{x}, \boldsymbol{y}) \mapsto \boldsymbol{x}$$

to the space of x-variables. □

This extended formulation will be later used for deriving the linear description of orbisacks with the means of faithful sectioning.

4.3.1.2 The $\mathbf{P}^{\tilde{x},y,z}$-Formulation

Our second approach for an extended formulation for the orbisack uses y-variables just like the $\mathbf{P}^{x,y}$-formulation above. However, instead of recycling $\boldsymbol{x}$, the x-variables are split into two new classes of variables, namely $\tilde{x}$ and z. For each vertex $\boldsymbol{x}$ of the orbisack, the $\tilde{x}$-variables display information about the entries of $\boldsymbol{x}$ *below* $\operatorname{crit}(\boldsymbol{x})$ and the z-variables store information about the type of the rows *above* $\operatorname{crit}(\boldsymbol{x})$.

In detail, we define for any vertex $\boldsymbol{x}$ of an orbisack $\mathbf{O}_{p,2}$ a vector $\big(\tilde{\boldsymbol{x}}(\boldsymbol{x}), \boldsymbol{y}(\boldsymbol{x}), \boldsymbol{z}(\boldsymbol{x})\big)$ in $\mathbb{R}^{[p]\times[2]} \times \mathbb{R}^{[p]} \times \mathbb{R}^{[p]}$ with

$$(\tilde{x}_{i,1}, \tilde{x}_{i,2}) = \begin{cases} (x_{i,1}, x_{i,2}), & \text{if } i > \operatorname{crit}(\boldsymbol{x}) \\ (0,0), & \text{otherwise} \end{cases} \qquad \text{for all } i \in [p]$$

and

$$z_i = \begin{cases} x_{i,1}, & \text{if } i < \mathrm{crit}(\boldsymbol{x}) \\ 0, & \text{otherwise} \end{cases} \qquad \text{for all } i \in [p]$$

and $\boldsymbol{y}(\boldsymbol{x}) = \boldsymbol{y}$ as defined for $\mathbf{P}^{x,y}$ above.

The corresponding polytope is defined as

$$\mathbf{P}^{\tilde{x},y,z} := \mathrm{conv}(\{(\tilde{\boldsymbol{x}}(\boldsymbol{x}), \boldsymbol{y}(\boldsymbol{x}), \boldsymbol{z}(\boldsymbol{x})) \in \mathbb{R}^{[p]\times[2]} \times \mathbb{R}^{[p]} \times \mathbb{R}^{[p]} \mid \boldsymbol{x} \text{ vertex of } \mathbf{O}_{p,2}\})$$

Again, we shorten notation by writing $(\tilde{\boldsymbol{x}}, \boldsymbol{y}, \boldsymbol{z})$ instead of $(\tilde{\boldsymbol{x}}(\boldsymbol{x}), \boldsymbol{y}(\boldsymbol{x}), \boldsymbol{z}(\boldsymbol{x}))$ in the following.

Lemma 4.42 *A linear description of $\mathbf{P}^{\tilde{x},y,z}$ together with an appropriate linear projection provides an extended formulation for $\mathbf{P}^{x,y}$.*

Proof. The projection $\sigma^{\tilde{x},y,z}$ from $\mathbf{P}^{\tilde{x},y,z}$ to $\mathbf{P}^{x,y}$ is mapping point $(\tilde{\boldsymbol{x}}, \boldsymbol{y}, \boldsymbol{z}) \in \mathbb{R}^{[p]\times[2]} \times \mathbb{R}^{[p]} \times \mathbb{R}^{[p]}$ to point $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^{[p]\times[2]} \times \mathbb{R}^{[p]}$ with

$$x_{i,1} = \tilde{x}_{i,1} + y_i + z_i \tag{4.2}$$
$$x_{i,2} = \tilde{x}_{i,2} + z_i \tag{4.3}$$

for all $i \in [p]$. The projection $\sigma^{\tilde{x},y,z}$ is linear. □

The following proposition gives a linear description of $\mathbf{P}^{\tilde{x},y,z}$. This provides the basis for the proof of the linear description via faithful sectioning(s).

Proposition 4.43 *Polytope $\mathbf{P}^{\tilde{x},y,z}$ is completely described by the following set of inequalities*

$$\tilde{x}_{i,1} - \sum_{k=1}^{i-1} y_k \leqslant 0 \qquad \forall i \in [p] \setminus \{1\} \tag{4.4}$$
$$\tilde{x}_{i,2} - \sum_{k=1}^{i-1} y_k \leqslant 0 \qquad \forall i \in [p] \setminus \{1\} \tag{4.5}$$
$$\sum_{k=1}^{i} y_k + z_i \leqslant 1 \qquad \forall i \in [p] \tag{4.6}$$
$$\tilde{x}_{i,j}, y_i, z_i \geqslant 0 \qquad \forall i \in [p] \textit{ and } j \in [2] \tag{4.7}$$
$$\tilde{x}_{1,1} = 0 \tag{4.8}$$
$$\tilde{x}_{1,2} = 0 \tag{4.9}$$

Proof. The system is totally unimodular, because the constraints matrix is an interval matrix (with attached unit matrices).

Obviously, the vertices $(\tilde{x}, y, z)$ satisfy all inequalities (4.4) through (4.7). On the other hand, (4.6) makes sure that no 0/1-vector $(\tilde{x}, y, z)$ can have more than one 1 in y, and if $y_k = 1$, then $z_i = 0$ for $k \leqslant i \leqslant p$. Inequalities (4.4), (4.5), and the equations (4.8) and (4.9) make sure that $\tilde{x}_{i,*} = 0$ for all rows i with $1 \leqslant i \leqslant k$. So the system in fact describes $\mathbf{P}^{\tilde{x},y,z}$. □

4.3.1.3 Extended Formulation associated with Dynamic Programming

As we have shown on page 55, one can, based on the DP algorithm for orbitopes, build a DP-hypergraph for orbitopes. There is a one-to-one-correspondence between the hyperpaths in this hypergraphs leading from the set of initial states $\mathcal{B}_p$ to the final state b_t and the vertices of the orbitope; the associated hyperpath polytope in arc variables leads to an extended formulation for orbitopes (see Lemma 3.38). (Note that we will in the following reuse notation from that section.)

Unfortunately, this extended formulation is not of much help for the linear description of the orbitope, as it is unclear how to describe the set of extreme rays of the projection cone. An alternative could be the use of node variables instead of arc variables. This would be interesting also for another reason: the number of nodes in $\mathcal{H}$ is of order $\mathcal{O}(pq^2)$ and the number of arcs is of order $\mathcal{O}(pq^3)$. Therefore, a linear description of the hyperpath set polytope would need fewer variables than the descriptions in arc variables does. Hence, a description in node variables would be nice to have. However, computer experiments are discouraging: they show that the description of the hyperpath polytope in node variables seems to be quite involved in general.

Nevertheless, for $q = 2$, a linear description for the $\mathcal{B}_p$-b_t-hyperpath set polytope $\mathbf{P}^{\text{node}}(H)$ is accessible. We will show that in fact, $\mathbf{P}^{\text{node}}(H)$ is isomorphic to polytope $\mathbf{P}^{\tilde{x},y,z}$ as defined above. Denoting by $\boldsymbol{w} \in \{0,1\}^{\mathcal{B}}$ the incidence vector of $\mathcal{B}[\mathcal{L}] \subseteq \mathcal{B}$ induced by a $\mathcal{B}_p$-b_t-hyperpath $\mathcal{L}$, we define for $q = 2$ the following transformation.

- Transformation $\tau : \mathbf{P}^{\text{node}}(H) \to \mathbf{P}^{\tilde{x},y,z}$.

$$\begin{aligned}
z_i &= w^{\bullet}_{i,[1..2]} && \forall i \in [p]\\
y_1 &= w^{\bullet}_{1,[1..1]} &&\\
y_2 &= w^{\bullet}_{2,[1..1]} + w^{\circ}_{2,[1..1]} - w^{\bullet}_{1,[1..1]} &&\\
y_i &= w^{\bullet}_{i,[1..1]} + w^{\circ}_{i,[1..1]} - w^{\bullet}_{i-1,[1..1]} - w^{\circ}_{i-1,[1..1]} && \forall i \in [3..p]\\
\tilde{x}_{i,1} &= w^{\bullet}_{i-1,[1..1]} + w^{\circ}_{i-1,[1..1]} - w^{\circ}_{i,[1..1]} && \forall i \in [2..p]\\
\tilde{x}_{i,2} &= w^{\bullet}_{i,[2..2]} && \forall i \in [2..p]\\
\tilde{x}_{1,j} &= 0 && \forall j \in [2]
\end{aligned}$$

- Transformation $\tau^{-1} : \mathbf{P}^{\tilde{x},y,z} \to \mathbf{P}^{\text{node}}(H)$.

$$\begin{aligned}
w^{\bullet}_{i,[1..2]} &= z_i && \forall i \in [p]\\
w^{\circ}_{i,[1..2]} &= 1 - (\textstyle\sum_{k=1}^{i} y_k + z_i) && \forall i \in [p]\\
w^{\bullet}_{i,[1..1]} &= \tilde{x}_{i,1} + y_i && \forall i \in [p]\\
w^{\circ}_{i,[1..1]} &= \textstyle\sum_{k=1}^{i-1} y_k - \tilde{x}_{i,1} && \forall i \in [2..p]\\
w^{\bullet}_{i,[2..2]} &= \tilde{x}_{i,2} && \forall i \in [2..p]\\
w^{\circ}_{i,[2..2]} &= \textstyle\sum_{k=1}^{i} y_k - \tilde{x}_{i,2} && \forall i \in [p]
\end{aligned}$$

We are using here and in the following the notation from page 55 ff..

Proposition 4.44 *For $q = 2$, the complete linear description of $\mathbf{P}^{\text{node}}(H)$ is given* *by the following set of inequalities:*

$$w^{\bullet}_{i,[2..2]} - w^{\bullet}_{i-1,[2..2]} - w^{\circ}_{i-1,[2..2]} \leqslant 0 \qquad \forall i \in [p] \setminus \{1\} \tag{4.10}$$
$$w^{\circ}_{i,[1..1]} - w^{\bullet}_{i-1,[1..1]} - w^{\circ}_{i-1,[1..1]} \leqslant 0 \qquad \forall i \in [p] \setminus \{1\} \tag{4.11}$$
$$w^{\bullet}_{i,[1..2]} + w^{\circ}_{i,[1..2]} - w^{\bullet}_{i-1,[1..2]} - w^{\circ}_{i-1,[1..2]} \leqslant 0 \qquad \forall i \in [p] \setminus \{1\} \tag{4.12}$$
$$w^{\bullet}_{i,[1..2]}, w^{\circ}_{i,[1..2]}, w^{\bullet}_{i,[1..1]}, w^{\circ}_{i,[1..1]}, w^{\bullet}_{i,[2..2]}, w^{\circ}_{i,[2..2]} \geqslant 0 \qquad \forall i \in [p] \tag{4.13}$$
$$w^{\circ}_{1,[1..1]} = 0 \tag{4.14}$$
$$w^{\bullet}_{1,[2..2]} = 0 \tag{4.15}$$
$$w_{0,[1..2]} = 1 \tag{4.16}$$
$$w^{\circ}_{i,[1..2]} + w^{\bullet}_{i,[1..2]} + w^{\bullet}_{i,[1..1]} + w^{\circ}_{i,[1..1]} = 1 \qquad \forall i \in [p] \tag{4.17}$$
$$w^{\circ}_{i,[1..2]} + w^{\bullet}_{i,[1..2]} + w^{\bullet}_{i,[2..2]} + w^{\circ}_{i,[2..2]} = 1 \qquad \forall i \in [p] \tag{4.18}$$

Proof. Follows directly from Proposition 4.43 and application of τ^{-1}. □

Despite the fact that the hypergraph point of view gives us a completely new set of variables (with new meaning), from the polyhedral point of view, it does not reveal more information than $\mathbf{P}^{\tilde{x},y,z}$ does.

One can also map $\mathbf{P}^{\mathrm{arc}}(H)$ to $\mathbf{P}^{\tilde{x},y,z}$ by means of the following projection. We denote the arc variables by u_a here.

- Projection $\gamma : \mathbf{P}^{\mathrm{arc}}(H) \to \mathbf{P}^{\tilde{x},y,z}$.

$$\begin{aligned}
\tilde{x}_{1,j} &= 0 && \forall j \in [2]\\
y_1 &= u_{(\{b^{\bullet}_{1,[1..1]}, b^{\circ}_{1,[2..2]}\}, b_{0,[1..2]})} &&\\
z_1 &= u_{(b^{\bullet}_{1,[1..2]}, b^{\bullet}_{0,[1..2]})} &&\\
y_i &= u_{(\{b^{\bullet}_{i,[1..1]}, b^{\circ}_{i,[2..2]}\}, b^{\bullet}_{i-1,[1..2]})} + u_{(\{b^{\bullet}_{i,[1..1]}, b^{\circ}_{i,[2..2]}\}, b^{\circ}_{i-1,[1..2]})} && \forall i \in [2..p]\\
x_{i,1} &= u_{(b^{\bullet}_{i,[1..1]}, b^{\bullet}_{i-1,[1..1]})} + u_{(b^{\bullet}_{i,[1..1]}, b^{\circ}_{i-1,[1..1]})} && \forall i \in [2..p]\\
x_{i,2} &= u_{(b^{\bullet}_{i,[2..2]}, b^{\bullet}_{i-1,[2..2]})} + u_{(b^{\bullet}_{i,[2..2]}, b^{\circ}_{i-1,[2..2]})} && \forall i \in [2..p]\\
z_i &= u_{(b^{\bullet}_{i,[1..2]}, b^{\bullet}_{i-1,[1..2]})} + u_{(b^{\bullet}_{i,[1..2]}, b^{\circ}_{i-1,[1..2]})} && \forall i \in [2..p]
\end{aligned}$$

Together with projection $\tau^{-1} : \mathbf{P}^{\tilde{x},y,z} \to \mathbf{P}^{\mathrm{node}}(H)$ from above, we get therefore also a projection $\tau^{-1} \circ \gamma : \mathbf{P}^{\mathrm{arc}}(H) \to \mathbf{P}^{\mathrm{node}}(H)$.

4.3.1.4 Overview of the Extended Formulations

Focusing on the orbisack, the extended formulations that have been defined so far can be arranged in a hierarchical structure, as the following illustration shows. Each of these extended formulations is set together from a polytope and a linear transformation mapping this polytope to another one. Projection ϑ has been defined in Lemma 3.38. Note that so far, for $q = 2$ columns, we have only linear descriptions for $\mathbf{P}^{\mathrm{node}}(H)$, $\mathbf{P}^{\mathrm{arc}}(H)$ and $\mathbf{P}^{\tilde{x},y,z}$. (Marked by boxes in the illustration.) We will undertake the task of formulating the linear description for the remaining polytopes in the following sections.

$$
\begin{tikzcd}
\mathbf{P}^{\mathrm{arc}}(H) \arrow[r, bend left, "\gamma"] \arrow[rr, bend left=40, "\tau^{-1}\circ\gamma"] \arrow[ddr, "\vartheta"] & \mathbf{P}^{\tilde{x},y,z} \arrow[r, shift left, "\tau^{-1}"] \arrow[d, "\sigma^{\tilde{x},y,z}"] & \mathbf{P}^{\mathrm{node}}(H) \arrow[l, shift left, "\tau"] \arrow[dl, "\sigma^{\tilde{x},y,z}\circ\tau"] \\
& \mathbf{P}^{x,y} \arrow[d, "\sigma^{x,y}"] & \\
& \mathbf{O}_{p,2} &
\end{tikzcd}
$$

4.3.2 Linear Description of $\mathrm{P}^{x,y}$ via Faithful Sectioning

We are now ready to derive a linear description of $\mathbf{P}^{x,y}$ via faithful sectioning, using the linear description of $\mathbf{P}^{\tilde{x},y,z}$ from Proposition 4.43.

We will start collecting the ingredients for the faithful sectioning. First, it is clear that the linear projection σ is as in Lemma 4.42:

$$\sigma := \sigma^{\tilde{x},y,z}$$

Next, we have to determine the map $s : \mathbb{R}^{[p]\times[3]} \to \mathbb{R}^{[p]\times[4]}, (\boldsymbol{x}, \boldsymbol{y}) \mapsto (\tilde{\boldsymbol{x}}, \boldsymbol{y}, \boldsymbol{z})$.

Observation 4.45 (a) Let s be some σ-section and $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ a linear description for $\mathbf{P}^{x,y}$ that is $\mathbf{P}^{\tilde{x},y,z}$-faithful for s. This implies that (at least) for any

$$(\boldsymbol{x}, \boldsymbol{y}) \in \{(\boldsymbol{x}, \boldsymbol{y}) \mid \boldsymbol{A}(\boldsymbol{x}, \boldsymbol{y}) \leqslant \boldsymbol{b}\},$$

(i) $s\big((\boldsymbol{x}, \boldsymbol{y})\big) \in \mathbf{P}^{\tilde{x},y,z}$ and
(ii) $\sigma\big(s\big((\boldsymbol{x}, \boldsymbol{y})\big)\big) = (\boldsymbol{x}, \boldsymbol{y})$.

From (i) follows that any point $(\tilde{\boldsymbol{x}}, \boldsymbol{y}, \boldsymbol{z}) = s\big((\boldsymbol{x}, \boldsymbol{y})\big)$ lifted from $\{(\boldsymbol{x}, \boldsymbol{y}) \mid \boldsymbol{A}(\boldsymbol{x}, \boldsymbol{y}) \leqslant \boldsymbol{b}\}$ must satisfy the inequalities from Proposition 4.43, in particular inequalities (4.4) and (4.5). From (ii) follows that we can use equations (4.2) and (4.3) to replace $\tilde{x}_{i,1}$ and $\tilde{x}_{i,2}$, respectively, to obtain the following conditions

$$x_{i,1} - \sum_{k=1}^{i} y_k \leqslant z_i \qquad \forall i \in [p] \setminus \{1\} \text{ and} \tag{*}$$

$$x_{i,2} - \sum_{k=1}^{i-1} y_k \leqslant z_i \qquad \forall i \in [p] \setminus \{1\}, \tag{**}$$

for any $(\boldsymbol{x}, \boldsymbol{y}) \in \{(\boldsymbol{x}, \boldsymbol{y}) \mid \boldsymbol{A}(\boldsymbol{x}, \boldsymbol{y}) \leqslant \boldsymbol{b}\}$. Moreover, from (4.7) follows that $z_i \geqslant 0$. Any other inequality in the linear description of $\mathbf{P}^{\tilde{x},y,z}$ bounds z_i either from above or not at all. So, defining component z_i of $s\big((\boldsymbol{x}, \boldsymbol{y})\big)$ as

$$z_i := \max\Big(\Big\{x_{i,1} - \sum_{k=1}^{i} y_k,\ x_{i,2} - \sum_{k=1}^{i-1} y_k,\ 0\Big\}\Big),$$

is not a contradiction to the fact that s is a σ-section and $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ is $\mathbf{P}^{\tilde{x},y,z}$-faithful.

(b) For $i = 1$, we obtain from equations (4.8) and (4.9) and the properties of σ that

$$z_1 = x_{1,1} - y_1 = x_{1,2}. \tag{4.19}$$

Therefore, equation $x_{1,1} - x_{i,2} - y_1 = 0$ must hold for $P^{x,y}$.

So, we define $s : (\boldsymbol{x}, \boldsymbol{y}) \mapsto (\tilde{\boldsymbol{x}}, \boldsymbol{y}, \boldsymbol{z})$ component wise as follows:

$$\left.\begin{array}{lcl} \tilde{x}_{i,1} & := & x_{i,1} - y_i - z_i \\ \tilde{x}_{i,2} & := & x_{i,2} - z_i \\ y_i & := & y_i \end{array}\right\} \text{ for all } i \in [p]$$

$$\begin{array}{lcll} z_1 & := & \max(\{x_{1,1} - y_1,\ x_{1,2},\ 0\}) & \\ z_i & := & \max(\{x_{i,1} - \sum_{k=1}^{i} y_k,\ x_{i,2} - \sum_{k=1}^{i-1} y_k,\ 0\}) & \text{for all } i \in [2..p]. \end{array}$$

Note that $i = 1$ and $i > 1$ can be considered simultaneously, if we keep in mind that equation (4.19) holds and thus identify y_0 with 0.

Lemma 4.46 *Map s is a σ-section.*

Proof. Easy computation shows that $\sigma\big(s\big((\boldsymbol{x}, \boldsymbol{y})\big)\big) = (\boldsymbol{x}, \boldsymbol{y})$ for all $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^{[p]\times[2]} \times \mathbb{R}^{[p]}$.

□

Lemma 4.47 *The following set of linear inequalities is $\mathbf{P}^{\tilde{x},y,z}$-enforcing for s.*

$$x_{i,1} - y_i \geqslant 0 \tag{4.20}$$

$$x_{i,2} + y_i \leqslant 1 \tag{4.21}$$

$$x_{i,1} - x_{i,2} - y_i + \sum_{k=1}^{i-1} y_k \geqslant 0 \tag{4.22}$$

$$-x_{i,1} + x_{i,2} + \sum_{k=1}^{i} y_k \geqslant 0 \tag{4.23}$$

$$\sum_{k=1}^{p} y_k \leqslant 1 \tag{4.24}$$

$$y_i \geqslant 0 \tag{4.25}$$

$$x_{i,1} \leqslant 1 \tag{4.26}$$

$$x_{i,2} \geqslant 0 \tag{4.27}$$

for all $i \in [p]$, together with equation

$$x_{1,1} - x_{1,2} - y_1 = 0 \tag{4.28}$$

Proof. We need to show that for all possibilities of z_i, $s(\boldsymbol{x}, \boldsymbol{y})$ is satisfying the inequality system from Proposition 4.43 determining $\mathbf{P}^{\tilde{x},y,z}$. (Remaining inequalities are satisfied by definition of s, see Observation 4.45.)

In the following table, we list in the first column the inequalities from Proposition 4.43, in the second column the possible cases for z_i, in the third column the inequalities which are necessary to satisfy the inequality from column one, and in the fourth and fifth column the respective numbers from the inequalities Lemma 4.47. So it can be read as follows: Inequality … is for $z_i = \ldots$ satisfied if inequality … is holding; this latter inequality has number … in Lemma 4.47 or it is dominated by inequality number … in the same Lemma.

inequality	z_i	ineq. satisfied if		dom. by
$\sum_{k=1}^{i} y_k + z_i \leqslant 1$	$x_{i,1} - \sum_{k=1}^{i} y_k$	$-x_{i,1} \geqslant -1$	(4.26)	
	$x_{i,2} - \sum_{k=1}^{i-1} y_k$	$-x_{i,2} - y_i \geqslant -1$	(4.21)	
	0	$-\sum_{k=1}^{i} y_k \geqslant -1$	(4.24)	
$\tilde{x}_{i,1} \geqslant 0$	$x_{i,1} - \sum_{k=1}^{i} y_k$	$\sum_{k=1}^{i-1} y_k \geqslant 0$		(4.25)
	$x_{i,2} - \sum_{k=1}^{i-1} y_k$	$x_{i,1} - x_{i,2} - y_i + \sum_{k=1}^{i-1} y_k \geqslant 0$	(4.22)	

	0	$x_{i,1} - y_i \geqslant 0$	(4.20)	
$\tilde{x}_{i,2} \geqslant 0$	$x_{i,1} - \sum_{k=1}^{i} y_k$	$-x_{i,1} + x_{i,2} + \sum_{k=1}^{i} y_k \geqslant 0$	(4.23)	
	$x_{i,2} - \sum_{k=1}^{i-1} y_k$	$\sum_{k=1}^{i-1} y_k \geqslant 0$		(4.25)
	0	$x_{i,2} \geqslant 0$	(4.27)	
$y_i \geqslant 0$	$x_{i,1} - \sum_{k=1}^{i} y_k$	$y_i \geqslant 0$	(4.25)	
	$x_{i,2} - \sum_{k=1}^{i-1} y_k$	$y_i \geqslant 0$	(4.25)	
	0	$y_i \geqslant 0$	(4.25)	

□

Corollary 4.48 *The inequality system from Lemma 4.47 provides a complete linear description of* $\mathbf{P}^{x,y}$.

Proof. Obviously, the inequalities are valid for $\mathbf{P}^{x,y}$. Moreover, $\sigma(\mathbf{P}^{\tilde{x},y,z}) \subseteq \mathbf{P}^{x,y}$, because σ and $\mathbf{P}^{\tilde{x},y,z}$ provide an extended formulation for $\mathbf{P}^{x,y}$ (Lemma 4.42). Map s is a σ-section because of Lemma 4.46. Lemma 4.47 shows that inequalities (4.20) through (4.27) and equation (4.28) are $\mathbf{P}^{\tilde{x},y,z}$-enforcing. Hence, the prerequisites for Theorem 2.17 are given. □

4.3.3 Linear Description of $\mathrm{O}_{p,2}$ via Faithful Sectioning

From the linear description of $\mathbf{P}^{x,y}$, we will now derive a linear description for the orbisack. We will proceed similarly to above.

We define $\sigma := \sigma^{x,y}$ by $(\boldsymbol{x}, \boldsymbol{y}) \mapsto \boldsymbol{x}$ (Lemma 4.41).

Observation 4.49 Consider some point $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbf{P}^{x,y}$. It satisfies inequalities (4.20) through (4.27) as well as equation (4.28). For $i > 1$, inequalities (4.20), (4.21), (4.22), and (4.24) define upper bounds for y_i. Hence, for any point $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbf{P}^{x,y}$, it holds that

$$y_i \leqslant \min\left(\left\{x_{i,1},\ 1 - x_{i,2},\ x_{i,1} - x_{i,2} + \sum_{k=1}^{i-1} y_k,\ 1 - \sum_{k=1}^{i-1} y_k\right\}\right) \quad \text{for all } i > 1$$

and

$$y_1 = x_{1,1} - x_{1,2} \quad \text{for } i = 1$$

Therefore, map s is defined by

$$s : \begin{cases} \mathbb{R}^{[p]\times[2]} & \to \mathbb{R}^{[p]\times[2]} \times \mathbb{R}^{[p]} \\ \boldsymbol{x} & \mapsto (\boldsymbol{x}, \boldsymbol{y}) \end{cases},$$

where

$$y_i = \min\left(\left\{x_{i,1},\ 1 - x_{i,2},\ x_{i,1} - x_{i,2} + \sum_{k=1}^{i-1} y_k,\ 1 - \sum_{k=1}^{i-1} y_k\right\}\right) \quad \text{for all } i > 1 \qquad (*)$$

and

$$y_1 = x_{1,1} - x_{1,2} \quad \text{for } i = 1.$$

Lemma 4.50 *s is a σ-section for all $\boldsymbol{x} \in \mathbb{R}^{[p]\times[2]}$.*

Proof. Obviously, $\sigma(s(\boldsymbol{x})) = \boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^{[p]\times[2]}$. □

Lemma 4.51 *The valued block inequalities (see Definition 4.15) and the cube inequalities form a set of $\mathbf{P}^{x,y}$-enforcing inequalities for s.*

Proof. We proceed by induction on i. For $i = 1$, equation $y_1 = x_{1,1} - x_{1,2}$ is satisfied by any point $s(\boldsymbol{x})$. The question is: which properties must $\boldsymbol{x}$ have such that $s(\boldsymbol{x})$ also satisfies inequalities (4.20) through (4.27) for $i = 1$?

inequality in description of $\mathbf{P}^{x,y}$	is satisfied if
$x_{1,1} - y_1 \geqslant 0$	$x_{1,2} \geqslant 0$
$x_{1,2} + y_1 \leqslant 1$	$x_{1,1} \leqslant 1$
$x_{1,1} - x_{1,2} - y_1 \geqslant 0$	—
$-x_{1,1} + x_{1,2} + y_1 \geqslant 0$	—
$y_1 \leqslant 1$	$x_{1,1} \leqslant 1$ and $x_{1,2} \geqslant 0$
$y_1 \geqslant 0$	$-x_{1,1} + x_{1,2} \leqslant 0$
$x_{1,1} \leqslant 1$	$x_{1,1} \leqslant 1$
$x_{1,2} \geqslant 0$	$x_{1,2} \geqslant 0$

Note that the block inequality $-x_{1,1} + x_{1,2} \leqslant 0$ and $x_{1,1} \leqslant 1$ imply that $x_{1,2} \leqslant 1$. Moreover, inequalities $-x_{1,1} + x_{1,2} \leqslant 0$ and $x_{1,2} \geqslant 0$ imply that $x_{1,1} \geqslant 0$. So, for $i = 1$, we add inequalities $x_{1,j} \leqslant 1$, $x_{1,j} \geqslant 0$, $j \in [2]$, as well as the block inequality $-x_{1,1} + x_{1,2} \leqslant 0$ to the set of $\mathbf{P}^{x,y}$-enforcing inequalities.

For induction step $i \rightsquigarrow i + 1$, we can assume that

$$y_k \geqslant 0 \text{ for all } k \in [i-1] \text{ and} \tag{**}$$

$$\sum_{k=1}^{i-1} y_k \leqslant 1 \tag{***}$$

Moreover, by choice of s, $s(\boldsymbol{x})$ already satisfies inequalities (4.20), (4.21), (4.22), and (4.24). It remains to identify those inequalities in $\boldsymbol{x}$ that ensure that $s(\boldsymbol{x})$ satisfies the remaining inequalities (4.23), (4.25), (4.26), and (4.27).

inequality in description of $\mathbf{P}^{x,y}$	y_i	ineq. satisfied by
$-x_{i,1} + x_{i,2} + \sum_{k=1}^{i} y_k \geqslant 0$	$x_{i,1}$	(**) and $x_{i,2} \geqslant 0$
	$1 - x_{i,2}$	(**) and $x_{i,1} \leqslant 1$
	$x_{i,1} - x_{i,2} + \sum_{k=1}^{i-1} y_k$	(**)
	$1 - \sum_{k=1}^{i-1} y_k$	$x_{i,1} \leqslant 1$ and $x_{i,2} \geqslant 0$
$y_i \geqslant 0$	$x_{i,1}$	$x_{i,1} \geqslant 0$
	$1 - x_{i,2}$	$x_{i,2} \leqslant 1$
	$x_{i,1} - x_{i,2} + \sum_{k=1}^{i-1} y_k$	(****)
	$1 - \sum_{k=1}^{i-1} y_k$	(***)
$x_{i,1} \leqslant 1$	$x_{i,1}$	$x_{i,1} \leqslant 1$
	$1 - x_{i,2}$	$x_{i,1} \leqslant 1$
	$x_{i,1} - x_{i,2} + \sum_{k=1}^{i-1} y_k$	$x_{i,1} \leqslant 1$
	$1 - \sum_{k=1}^{i-1} y_k$	$x_{i,1} \leqslant 1$
$x_{i,2} \geqslant 0$	$x_{i,1}$	$x_{i,2} \geqslant 0$
	$1 - x_{i,2}$	$x_{i,2} \geqslant 0$
	$x_{i,1} - x_{i,2} + \sum_{k=1}^{i-1} y_k$	$x_{i,2} \geqslant 0$
	$1 - \sum_{k=1}^{i-1} y_k$	$x_{i,2} \geqslant 0$

The inequalities obtained so far imply that $\boldsymbol{x} \in [0,1]^{[p]\times[2]}$. It remains to ensure that any $\boldsymbol{x} \in [0,1]^{[p]\times[2]}$ satisfies inequalities

$$x_{i,1} - x_{i,2} + \sum_{k=1}^{i-1} y_k \geqslant 0 \tag{****}$$

for all $i > 1$. Note that $y_1 = x_{1,1} - x_{1,2}$.

Example 4.52 Let $i = 7$. If y_k, $k \in [2..6]$, equals the second, fourth, and three times the third option in (*), we obtain:

$x_{1,1} - x_{1,2}$		$-(x_{1,1} - x_{1,2})$	0	0	0
	$1 - x_{2,2}$	$-(1 - x_{2,2})$	0	0	0
		1	1	2	4
			$x_{4,1} - x_{4,2}$	$x_{4,1} - x_{4,2}$	$2(x_{4,1} - x_{4,2})$
				$x_{5,1} - x_{5,2}$	$x_{5,1} - x_{5,2}$
					$x_{6,1} - x_{6,2}$
y_1	y_2	y_3	y_4	y_5	y_6

To satisfy the corresponding inequality (***) is equivalent to ensuring that the sum over all entries of this table (except the last row, of course), plus $x_{7,1} - x_{7,2}$, is larger or equal than 0; that is:

0	0	$x_{1,1}$	$x_{1,2}$	$\leqslant 8$
0	0	$x_{2,1}$	$x_{2,2}$	
0	0	$x_{3,1}$	$x_{3,2}$	
-4	$+4$	$x_{4,1}$	$x_{4,2}$	
-2	$+2$	$x_{5,1}$	$x_{5,2}$	
-1	$+1$	$x_{6,1}$	$x_{6,2}$	
-1	$+1$	$x_{7,1}$	$x_{7,2}$	

Note that this inequality is the sum of trivial inequalities $-x_{i,1} \leqslant 0$ and $x_{i,2} \leqslant 1$ for all $i \in [4..7]$.

On the other hand, we obtain with the first, second and three times the third option:

$x_{1,1} - x_{1,2}$			$x_{1,1} - x_{1,2}$	$2(x_{1,1} - x_{1,2})$	$4(x_{1,1} - x_{1,2})$
	$x_{2,1}$		$x_{2,1}$	$2x_{2,1}$	$4x_{2,1}$
		$1 - x_{3,2}$	$1 - x_{3,2}$	$2(1 - x_{3,2})$	$4(1 - x_{3,2})$
			$x_{4,1} - x_{4,2}$	$x_{4,1} - x_{4,2}$	$2(x_{4,1} - x_{4,2})$
				$x_{5,1} - x_{5,2}$	$x_{5,1} - x_{5,2}$
					$x_{6,1} - x_{6,2}$
y_1	y_2	y_3	y_4	y_5	y_6

Therefore, $\boldsymbol{x}$ must satisfy inequality

-8	$+8$	$x_{1,1}$	$x_{1,2}$	$\leqslant 8$
-8		$x_{2,1}$	$x_{2,2}$	
	$+8$	$x_{3,1}$	$x_{3,2}$	
-4	$+4$	$x_{4,1}$	$x_{4,2}$	
-2	$+2$	$x_{5,1}$	$x_{5,2}$	
-1	$+1$	$x_{6,1}$	$x_{6,2}$	
-1	$+1$	$x_{7,1}$	$x_{7,2}$	

Note that this inequality is a valued block inequality.

Observation 4.53 Obviously, one can define vectors $\boldsymbol{a}_{*,1}, \boldsymbol{a}_{*,2}$ and $\boldsymbol{b}$ all in $\mathbb{R}^{[p]}$ such that

$$x_{i,1} - x_{i,2} + \sum_{k=1}^{i-1} y_k = \sum_{k=1}^{p} (a_{k,1}x_{k,1} + a_{k,2}x_{k,2} + b_k)$$

simply by reading off the coefficients. Inequality (****) can then be rewritten as

$$-\langle \boldsymbol{a}, \boldsymbol{x} \rangle \leqslant \sum_{k=1}^{p} b_k,$$

where $\boldsymbol{a} = (\boldsymbol{a}_{*,1}, \boldsymbol{a}_{*,2})$ and $\boldsymbol{x} = (\boldsymbol{x}_{*,1}, \boldsymbol{x}_{*,2})$.

We will in the following describe the coefficients of $\boldsymbol{a}_{*,1}, \boldsymbol{a}_{*,2}$ and $\boldsymbol{b}$. Clearly, $a_{i,1} = -1$, $a_{i,2} = +1$, and $a_{k,j} = 0$ for $k > i$ and $j \in [2]$. Assume entries y_k have a value from (*) for $k \in [i-2]$. What influence has the value of y_{i-1} on the components of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$?

Observation 4.54

- If y_{i-1} is option 1 or 2 in (*), then for all $k < i-1$, $a_{k,1}$, $a_{k,2}$ and b_k keep unchanged, and either $a_{i-1,1} = -1$, $a_{i-1,2} = 0$ and $b_{i-1} = 0$ (option 1) or $a_{i-1,1} = 0$, $a_{i-1,2} = 1$ and $b_{i-1} = 1$ (option 2).
- If y_{i-1} is option 3 in (*), then for all $k < i-1$, $a_{k,1}$, $a_{k,2}$ and b_k are each multiplied by factor 2. Moreover, $a_{i-1,1} = -1$, $a_{i-1,2} = 1$ and $b_{i-1} = 0$.
- If y_{i-1} is option 4 in (*), then for all $k < i-1$, $a_{k,1} = 0$, $a_{k,2} = 0$ and $b_k = 0$. Moreover, $a_{i-1,1} = 0$, $a_{i-1,2} = 0$ and $b_{i-1} = 1$.
- By construction of the map s, $y_1 = x_{1,1} - x_{1,2}$. Together with the above-mentioned, this implies that $a_{1,1} = -\alpha$ and $a_{1,2} = \alpha$ with $\alpha \in \{2^\mu \mid \mu \in \mathbb{N}\} \cup \{0\}$.

From these observations, it follows that if y_k is equal to one of the first three options in (*) for all $k \in [2..i-1]$, then inequality (****) is a valued block inequality. If, on the other hand, not every y_k comes from the first three options, the resulting inequality $\boldsymbol{ax} \leqslant b$ looks like a block inequality with basement i, where rows 1 through $k^\star$ are replaced by empty rows for some $1 \leqslant k^\star \leqslant i-1$. The right-hand side is the sum of all p-rows in the modified vector $\boldsymbol{a}$ plus the value of the block that contains row $k^\star$; so the right-hand side equals the sum of all positive entries in $\boldsymbol{a}$. Therefore, the inequality $\boldsymbol{ax} \leqslant b$ is a block inequality plus the sum of trivial inequalities. (See the first of examples 4.52.)

Therefore, for any $i > 1$, if $\boldsymbol{x}$ satisfies all valued block inequalities with basement i, $s(\boldsymbol{x})$ does not violate inequality (****). □

Proposition 4.55 *The set of valued block inequalities and the cube inequalities provide a complete description of the orbisack.*

$$\boldsymbol{a} = \begin{pmatrix} \vdots & \vdots \\ -1 & 1 \\ -1 & 1 \\ 0 & 0 \\ \vdots & \vdots \end{pmatrix} \quad \rightsquigarrow \quad \underset{\text{case (A)}}{\begin{pmatrix} \vdots & \vdots \\ -1 & 1 \\ 0 & 1 \\ -1 & 1 \\ \vdots & \vdots \end{pmatrix}} \text{ or } \underset{\text{case (B)}}{\begin{pmatrix} \vdots & \vdots \\ -1 & 1 \\ -1 & 0 \\ -1 & 1 \\ \vdots & \vdots \end{pmatrix}} \text{ or } \underset{\text{case (C)}}{\begin{pmatrix} \vdots & \vdots \\ -2 & 2 \\ -1 & 1 \\ -1 & 1 \\ \vdots & \vdots \end{pmatrix}}$$

Figure 4.3: The idea behind the separation algorithm as listed in figure 4.4: row by row, we shift the basement down and modify vector $\boldsymbol{a}$ and right-hand side b appropriately such that $\boldsymbol{ax} \leqslant b$ stays a valued block inequality and $\langle \boldsymbol{ay} \rangle$ becomes maximal.

Proof. The valued block and cube inequalities are valid for the orbisack. Map s is a σ-section (Lemma 4.50). Moreover, Lemma 4.51 shows that the valued block inequalities and the cube inequalities make are $\mathbf{P}^{x,y}$-enforcing inequalities for s. So we obtain the statement by Theorem 2.17. $\square$

4.4 Selected Properties of the Orbisack

4.4.1 Number of Facets

It is easy to enumerate the facets of the orbisack.

Observation 4.56 The orbisack $\mathbf{O}_{p,2}$ has $\Theta(3^p)$ facets.

Proof. There are $2(2p-1)$ trivial facets. Each non-trivial facet $\boldsymbol{ax} \leqslant b$ is fixed by the sign pattern of the components of $\boldsymbol{a}$. There is one sign pattern (and therefore one non-trivial facet) with basement in row 1 and one with basement in row 2. For basement $k \geqslant 3$, each row between first row and the basement can be filled with one of three sign patterns. For $p > 2$, we get therefore a total number of facets of

$$2(2p-1) + 2 + \sum_{k=3}^{p} 3^{k-2} = 4p + \frac{3}{2}(3^{p-2} - 1),$$

using the geometric sum. $\square$

4.4.2 The Separation Problem for Orbisacks

The separation problem is to decide whether a given vector $\boldsymbol{y} \in \mathbb{R}^{[p]\times[2]}$ lies inside a polytope (here: an orbisack $\mathbf{O}_{p,2}$) or not and, if not, to find a valid inequality for the polytope that is violated by $\boldsymbol{y}$. This problem can be solved in time $\mathcal{O}(p^2)$ as the algorithm in fig. 4.4 shows. The main loop is executed not more than p-times, but in the worst case, it may be necessary to cycle through $\boldsymbol{a}$ in each run to multiply each entry by factor 2.

Lemma 4.57 *Algorithm 4.4 works correctly.*

Proof. The algorithm obviously terminates. Correctness can be shown inductively. For basements $\ell = 1$ and $\ell = 2$, the algorithm considers all existing VBIs. $\ell \rightsquigarrow \ell+1$: In the main loop, the algorithm varies the sign pattern of the inequality obtained in the preceding step (only) in the last row and appends a new basement. Let $\boldsymbol{ax} \leqslant b$

be the inequality obtained in the preceding step with $\text{base}(a) = \ell$. The algorithm constructs $\boldsymbol{a}'$ and b' from $\boldsymbol{a}$ and b such that $\boldsymbol{a}'\boldsymbol{y} - b'$ becomes

$$\begin{aligned}\max\big(\{\ &\langle \boldsymbol{a}, \boldsymbol{y}\rangle + y_{\ell,1} - y_{\ell+1,1} + y_{\ell+1,2} - b,\\ &\langle \boldsymbol{a}, \boldsymbol{y}\rangle - y_{\ell,2} - y_{\ell+1,1} + y_{\ell+1,2} - b - 1,\\ 2&\langle \boldsymbol{a}, \boldsymbol{y}\rangle + y_{\ell,1} - y_{\ell,2} - y_{\ell+1,1} + y_{\ell+1,2} - 2b\ \}\big).\end{aligned}$$

If there is an inequality $\tilde{\boldsymbol{a}}'\boldsymbol{x} \leqslant \tilde{b}'$ with $\text{base}(\tilde{\boldsymbol{a}}') = \ell + 1$ and $\tilde{\boldsymbol{a}}'\boldsymbol{y} - \tilde{b}' > \boldsymbol{a}'\boldsymbol{y} - b'$, then this implies that there exists an inequality $\tilde{\boldsymbol{a}}\boldsymbol{x} \leqslant \tilde{b}$ with basement ℓ and $\tilde{\boldsymbol{a}}\boldsymbol{y} - \tilde{b} > \boldsymbol{a}\boldsymbol{y} - b$ contradicting the induction hypothesis. □

4.4.3 The Graph of the Orbisack

The graph (sometimes also phrased as the "1-skeleton") $G_{\mathbf{P}} = (\mathcal{V}, \mathcal{E})$ of a polytope $\mathbf{P}$ has vertex set $\mathcal{V} = \mathcal{V}_{\mathbf{P}}$ and edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ with $\{\boldsymbol{v}, \boldsymbol{w}\} \in \mathcal{E}$ if and only if polytope $\mathbf{P}$ has an edge $\mathtt{F}$ such that $\boldsymbol{v}, \boldsymbol{w} \in \mathtt{F}$. As it is usual, $\boldsymbol{v}$ and $\boldsymbol{w}$ are called *neighbours* then.

The graph of a polytope reflects a lot of information about the polytope itself (for a survey, see [66]).

In this section, we will give a complete characterization of the graph $G_{\mathbf{O}_{p,2}}$ of the orbisack $\mathbf{O}_{p,2}$; this will allow us to compute the average degree of each vertex. Our aim was to prove that the graph of the orbisack has edge expansion at least 1 by inductively constructing the graph of $\mathbf{O}_{p,2}$ from graphs of $\mathbf{O}_{k,2}$, $k < p$. This would be a further support for the conjecture of Mihail and Vazirani (see page 34). However, we did not succeed.

First, we will show necessary conditions for adjacency, followed by a proof that these conditions are sufficient. The section is closed by the computation of the number of edges and the average degree.

4.4.3.1 Characterization of Adjacency

The main ingredient of the following proofs is a standard observation.

Observation 4.58 It is clear that two vertices $\boldsymbol{v}, \boldsymbol{w} \in \mathbf{P}$ cannot be adjacent if there are two other vertices $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{P}$ and scalars $0 \leqslant \mu, \lambda \leqslant 1$ such that

$$\mu\boldsymbol{x} + (1-\mu)\boldsymbol{y} = \lambda\boldsymbol{v} + (1-\lambda)\boldsymbol{w},$$

i.e. if $\text{conv}(\{\boldsymbol{v}, \boldsymbol{w}\}) \cap \text{conv}(\{\boldsymbol{x}, \boldsymbol{y}\}) \neq \varnothing$. Restricting to $\lambda = \mu = \frac{1}{2}$, we get that in particular, $\boldsymbol{v}$ and $\boldsymbol{w}$ cannot be adjacent if there are $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{P}$ such that

$$\boldsymbol{v} + \boldsymbol{w} = \boldsymbol{x} + \boldsymbol{y}. \tag{4.29}$$

For certain classes of 0/1-polytopes, the latter condition is also necessary for non-adjacency; so, for these polytopes, two vertices $\boldsymbol{v}$ and $\boldsymbol{w}$ are neighbours if and only if there is no other pair of vertices $\boldsymbol{x}, \boldsymbol{y}$ such that $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{v} + \boldsymbol{w}$. Among these polytopes are e.g. matching, matroid, and stable set polytopes. Naddef and Pulleyblank called these polytopes *combinatorial* (see [92]).

As we will see, orbisacks are combinatorial in this sense. Naddef and Pulleyblank showed that the graphs of combinatorial polytopes are either hypercubes or Hamilton connected. We will see that the latter is the case for the orbisack, i.e. any pair of nodes in the graph of orbisack is joined by a Hamilton path.

Data: $\boldsymbol{y} \in \mathbb{R}^{[p]\times[2]}$
Result: (false) or (true and violated inequality)
if $\boldsymbol{y} \notin [0,1]^{[p]\times[2]}$ **then**
  return *true;* $\boldsymbol{y}$ *violates trivial inequality*;
end
if $-y_{1,1} + y_{1,2} > 0$ **then**
  return *true;* $\boldsymbol{y}$ *violates inequality* $-x_{1,1} + x_{1,2} \leqslant 0$;
end
if $p > 1$ **then**
  $\boldsymbol{a} \leftarrow \begin{pmatrix} -1 & 1 \\ -1 & 1 \\ 0 & 0 \\ \vdots & \vdots \end{pmatrix}$ and $b \leftarrow 0$;
  while $\mathrm{base}(\boldsymbol{a}) \leqslant p$ **do**
    if $\langle \boldsymbol{a}, \boldsymbol{y} \rangle > b$ **then**
      return *true;* $\boldsymbol{y}$ *violates inequality* $\boldsymbol{ax} \leqslant b$*;*
    end
    if $\mathrm{base}(\boldsymbol{a}) < p$ **then**
      `// basement becomes p-row`
      $A \leftarrow \langle \boldsymbol{a}, \boldsymbol{y} \rangle + y_{\mathrm{base}(\boldsymbol{a}),1} - y_{\mathrm{base}(\boldsymbol{a})+1,1} + y_{\mathrm{base}(\boldsymbol{a})+1,2} - b - 1$;
      `// basement becomes n-row`
      $B \leftarrow \langle \boldsymbol{a}, \boldsymbol{y} \rangle - y_{\mathrm{base}(\boldsymbol{a}),2} - y_{\mathrm{base}(\boldsymbol{a})+1,1} + y_{\mathrm{base}(\boldsymbol{a})+1,2} - b$;
      `// basement becomes np-row`
      $C \leftarrow 2\langle \boldsymbol{a}, \boldsymbol{y} \rangle + y_{\mathrm{base}(\boldsymbol{a}),1} - y_{\mathrm{base}(\boldsymbol{a}),2} - y_{\mathrm{base}(\boldsymbol{a})+1,1} + y_{\mathrm{base}(\boldsymbol{a})+1,2} - 2b$;
      switch $\max(\{A, B, C\})$ **do**
        case A
          $b \leftarrow b + 1$;
          $\boldsymbol{a} \leftarrow \boldsymbol{a} + a_{\mathrm{base}(a),1} - a_{\mathrm{base}(a)+1,1} + a_{\mathrm{base}(a)+1,2}$;
        endsw
        case B
          $\boldsymbol{a} \leftarrow \boldsymbol{a} - a_{\mathrm{base}(a),2} - a_{\mathrm{base}(a)+1,1} + a_{\mathrm{base}(a)+1,2}$;
        endsw
        case C
          $b \leftarrow 2b$;
          $\boldsymbol{a} \leftarrow 2\boldsymbol{a} + a_{\mathrm{base}(a),1} - a_{\mathrm{base}(a),2} - a_{\mathrm{base}(a)+1,1} + a_{\mathrm{base}(a)+1,2}$;
        endsw
      endsw
    end
  end
end
return *false*;

Figure 4.4: Separation algorithm for orbisacks.

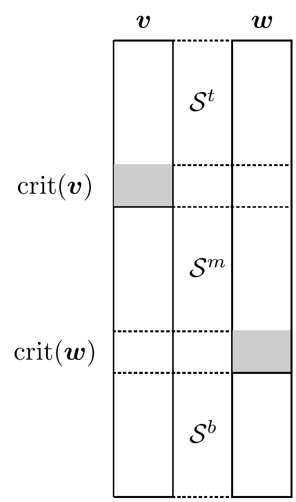

Figure 4.5: Partitioning the rows of two vertices $\boldsymbol{v}$ and $\boldsymbol{w}$ into segments $\mathcal{S}^t$, $\mathcal{S}^m$, and $\mathcal{S}^b$. Critical rows are marked gray.

Definition 4.59 (Segments) Let $\boldsymbol{v}, \boldsymbol{w}$ be a pair of vertices of $\mathbf{O}_{p,2}$. We define three intervals $\mathcal{S}^t, \mathcal{S}^m, \mathcal{S}^b \subseteq [p]$ of row indices and refer to them as *segments*. The range of each segment depends on the positions of the critical rows crit($\boldsymbol{v}$) and crit($\boldsymbol{w}$):

$$\begin{aligned}\mathcal{S}^t &:= \{i \in [p] \mid i < \min(\{\mathrm{crit}(\boldsymbol{v}), \mathrm{crit}(\boldsymbol{w})\})\},\\ \mathcal{S}^m &:= \{i \in [p] \mid i > \min(\{\mathrm{crit}(\boldsymbol{v}), \mathrm{crit}(\boldsymbol{w})\}) \text{ and } i < \max(\{\mathrm{crit}(\boldsymbol{v}), \mathrm{crit}(\boldsymbol{w})\}\}), \text{ and}\\ \mathcal{S}^b &:= \{i \in [p] \mid i > \max(\{\mathrm{crit}(\boldsymbol{v}), \mathrm{crit}(\boldsymbol{w})\})\}.\end{aligned}$$

Note that $\mathcal{S}^m = \varnothing$ if $\mathrm{crit}(\boldsymbol{v}) = \mathrm{crit}(\boldsymbol{w})$.
Moreover, $\mathcal{S}^b = \varnothing$ if $\max(\mathrm{crit}(\boldsymbol{v}), \mathrm{crit}(\boldsymbol{w})) = p + 1$.

Now we can start with a characterization of non-adjacent pairs of vertices.

Lemma 4.60 *Two vertices $\boldsymbol{v}$ and $\boldsymbol{w}$ with the same critical row cannot be adjacent if they differ*
(a) in more than one row in segment $\mathcal{S}^t$ or
(b) in more than one position (i, j) in segment $\mathcal{S}^b$ or
(c) in at least one row in segment $\mathcal{S}^t$ and at least one position (i, j) in segment $\mathcal{S}^b$.

Proof. (a) Suppose, the vertices differ in rows $i, k \in \mathcal{S}^t$, $i \neq k$. Modify $\boldsymbol{v}$ in $\widetilde{\boldsymbol{v}}_{i,*} := \mathbb{1} - \boldsymbol{v}_{i,*}$ and $\boldsymbol{w}$ by $\widetilde{\boldsymbol{w}}_{i,*} := \mathbb{1} - \boldsymbol{w}_{i,*}$. Then $\widetilde{\boldsymbol{v}} \notin \{\boldsymbol{v}, \boldsymbol{w}\}$ and $\widetilde{\boldsymbol{w}} \notin \{\boldsymbol{v}, \boldsymbol{w}\}$, but $\boldsymbol{v} + \boldsymbol{w} = \widetilde{\boldsymbol{v}} + \widetilde{\boldsymbol{w}}$ and condition (4.29) for non-adjacency is satisfied.
(b) Similarly, if $\boldsymbol{v}$ and $\boldsymbol{w}$ differ in positions (i, j) and (k, ℓ) with $i, k \in \mathcal{S}^b$, we modify $\boldsymbol{v}$ in $\widetilde{v}_{i,j} := 1 - v_{i,j}$ and $\boldsymbol{w}$ in $\widetilde{w}_{i,j} := 1 - w_{i,j}$.
(c) Analogous to (a) and (b).

□

Lemma 4.61 *Let $\boldsymbol{v}$ and $\boldsymbol{w}$ be two vertices with different critical rows. (W.l.o.g., we can assume that* $\mathrm{crit}(\boldsymbol{v}) < \mathrm{crit}(\boldsymbol{w})$*.) The vertices cannot be adjacent, if any of the following holds:*
(a) $\boldsymbol{v}_{\mathcal{S}^t,*} \neq \boldsymbol{w}_{\mathcal{S}^t,*}$,
(b) $\boldsymbol{v}_{\mathcal{S}^b,*} \neq \boldsymbol{w}_{\mathcal{S}^b,*}$,

(c) there is some row $i \in \mathcal{S}^m$ such that $\boldsymbol{v}_{i,*}, \boldsymbol{w}_{i,*} \in \{(1,1),(0,0)\}$ and $\boldsymbol{v}_{i,*} \neq \boldsymbol{w}_{i,*}$,
(d) there is some row $i \in \mathcal{S}^m$ such that $\boldsymbol{v}_{i,*} = (1,0)$,
(e) $\boldsymbol{v}_{\mathrm{crit}(\boldsymbol{w}),*} \in \{(1,1),(0,0),(0,1)\}$ and $\boldsymbol{v}_{\mathcal{S}^b,*}$ is lexicographically ordered.
Note that if for case (e), $\mathcal{S}^b = \varnothing$, then $\boldsymbol{v}_{\mathcal{S}^b,*}$ can be seen as lexicographically ordered.

Proof. (a) Assume, $\boldsymbol{v}_{i,*} \neq \boldsymbol{w}_{i,*}$ for $i \in \mathcal{S}^t$. Then modify $\boldsymbol{v}$ in components $\widetilde{\boldsymbol{v}}_{i,*} := \mathbb{1} - \boldsymbol{v}_{i,*}$ and $\boldsymbol{w}$ in components $\widetilde{\boldsymbol{w}}_{i,*} := \mathbb{1} - \boldsymbol{w}_{i,*}$. Because of $\mathrm{crit}(\boldsymbol{v}) \neq \mathrm{crit}(\boldsymbol{w})$, we get that $\boldsymbol{v}_{\mathrm{crit}(\boldsymbol{w},*)} \neq (1,0)$. So $\widetilde{\boldsymbol{v}} \notin \{\boldsymbol{v},\boldsymbol{w}\}$ and $\widetilde{\boldsymbol{w}} \notin \{\boldsymbol{v},\boldsymbol{w}\}$, but $\boldsymbol{v}+\boldsymbol{w} = \widetilde{\boldsymbol{v}} + \widetilde{\boldsymbol{w}}$.
(b) Analogous to (a).
(c) We swap the rows between $\boldsymbol{v}$ an $\boldsymbol{w}$: modify $\boldsymbol{w}$ in components $\widetilde{\boldsymbol{w}}_{i,*} := \boldsymbol{v}_{i,*}$ and $\boldsymbol{v}$ by $\widetilde{\boldsymbol{v}}_{i,*} := \boldsymbol{w}_{i,*}$.
(d) Analogous to (c).
(e) Analogous to (c). If $\boldsymbol{v}_{\mathrm{crit}(\boldsymbol{w}),*} \in \{(1,1),(0,0)\}$ and $\boldsymbol{w}_{\mathcal{S}^b,*}$ is in lexicographic order, swapping rows $\boldsymbol{v}_{\mathrm{crit}(\boldsymbol{w},*)}$ and $\boldsymbol{v}_{\mathrm{crit}(\boldsymbol{v},*)}$ leads to modified vectors $\widetilde{\boldsymbol{v}}$ and $\widetilde{\boldsymbol{w}}$ with $\boldsymbol{v}+\boldsymbol{w} = \widetilde{\boldsymbol{v}} + \widetilde{\boldsymbol{w}}$. If $\boldsymbol{v}_{\mathrm{crit}(\boldsymbol{w}),*} = (0,1)$ and $\boldsymbol{w}_{\mathcal{S}^b,*}$ is in lexicographic order, then modify $\boldsymbol{w}$ in components $\widetilde{\boldsymbol{w}}_{i,*} := (1,1)$ and $\boldsymbol{v}$ in components $\widetilde{\boldsymbol{v}}_{i,*} := (0,0)$. ☐

We summarize:

- To have a chance that two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ are adjacent, they must differ either in one row above $\mathrm{crit}(v)$ or in one entry below $\mathrm{crit}(\boldsymbol{v})$ but not both, as long as $\mathrm{crit}(\boldsymbol{v}) = \mathrm{crit}(\boldsymbol{w})$.
- If $\boldsymbol{v}$ and $\boldsymbol{w}$ are supposed to be adjacent and $\mathrm{crit}(\boldsymbol{v}) < \mathrm{crit}(\boldsymbol{w})$, then both vectors must be identical above $\mathrm{crit}(\boldsymbol{v})$ and below $\mathrm{crit}(\boldsymbol{w})$. In rows i between $\mathrm{crit}(\boldsymbol{v})$ and $\mathrm{crit}(\boldsymbol{w})$, the entries are either identical in $\boldsymbol{v}$ and $\boldsymbol{w}$ or row $\boldsymbol{v}_{i,*} = (0,1)$. Last, if $\boldsymbol{v}_{\mathcal{S}^b,*}$ is lexicographic ordered, then row $\boldsymbol{v}_{\mathrm{crit}(\boldsymbol{w}),*} = (1,0)$.

Proposition 4.62 *The conditions from Lemma 4.60 and 4.61 completely characterize non-adjacency, i.e.: If two vertices are not adjacent, then at least one of the conditions from Lemma 4.60 and 4.61 is violated.*

Proof. For two vertices $\boldsymbol{v}$ and $\boldsymbol{w}$ that satisfy the conditions from lemmas 4.60 and 4.61, we will construct a cost vector $\boldsymbol{c} \in \mathbb{R}^{[p]\times[2]}$ such that $\mathcal{V}[\boldsymbol{c}] = \{\boldsymbol{v},\boldsymbol{w}\}$. For this purpose, we define sets $\mathcal{I}^\alpha := \{(i,j) \in [p]\times[2] \mid v_{i,j} = w_{i,j} = \alpha\}$ for $\alpha \in \{0,1\}$. The cost vector will be in general defined as

$$\boldsymbol{c}(\boldsymbol{a}) := \underbrace{\sum_{(i,j)\in\mathcal{I}^1} \mathbb{e}^{i,j} - \sum_{(i,j)\in\mathcal{I}^0} \mathbb{e}^{i,j}}_{(*)} + \boldsymbol{a},$$

depending on some vector $\boldsymbol{a} \in \mathbb{R}^{[p]\times[2]}$ that will be fixed later. The unit vectors $(*)$ already fix for any vector $\boldsymbol{x} \in \mathcal{V}[\boldsymbol{c}(\boldsymbol{a})]$ entry $x_{i,j}$ to value $\alpha \in \{0,1\}$ if $v_{i,j} = \alpha$ and $w_{i,j} = \alpha$. (The solutions lie in the intersection of the appropriate trivial faces.) Thus we have to care only about entries where $\boldsymbol{v}$ and $\boldsymbol{w}$ differ from each other.

(i) First, we consider the case $\mathrm{crit}(v) = \mathrm{crit}(w)$. We have to show that if $\boldsymbol{v}$ and $\boldsymbol{w}$ do not violate conditions (a) through (c) from Lemma 4.60, then they are adjacent. Note that therefore, $\mathcal{I}^1 \cup \mathcal{I}^0$ covers all positions $(i,j) \in [p]\times[2]$ except either one row above $\mathrm{crit}(\boldsymbol{v})$ or one entry below $\mathrm{crit}(\boldsymbol{v})$. So, we have to distinguish only two cases:
 (i.1) If $\boldsymbol{v}$ and $\boldsymbol{w}$ differ in one position (i,j) below $\mathrm{crit}(\boldsymbol{v})$, we set $\boldsymbol{a} := \mathbb{0}$. Then, a choice is only possible for position (i,j), and therefore $\mathcal{V}[\boldsymbol{c}(\boldsymbol{a})] = \{\boldsymbol{v},\boldsymbol{w}\}$.

(i.2) If, on the other hand, $\boldsymbol{v}$ and $\boldsymbol{w}$ differ in one row k above $\operatorname{crit}(\boldsymbol{v})$, then we must ensure by choice of $\boldsymbol{a}$ that no vector in $\mathcal{V}[\boldsymbol{c}(\boldsymbol{a})]$ has critical row k. So we define vector $\boldsymbol{a}$ as the vector with the following properties

- there exists $b \in \mathbb{R}$ such that $\boldsymbol{ax} \leqslant b$ is a VBI,
- $\operatorname{base}(\boldsymbol{a}) = k$, and
- each nonempty row of $\boldsymbol{a}$ has sign pattern $\operatorname{sign}(\boldsymbol{a}_{i,*}) = (-, +)$.

Now the vectors in $\mathcal{V}[\boldsymbol{c}(\boldsymbol{a})]$ differ only in entries in row k. However, Observation 4.19 shows that there is no vector in $\mathcal{V}[\boldsymbol{a}]$ (and hence no vector in $\mathcal{V}[\boldsymbol{c}(\boldsymbol{a})]$) with critical row in row k. Therefore, again $\mathcal{V}[\boldsymbol{c}(\boldsymbol{a})] = \{\boldsymbol{v}, \boldsymbol{w}\}$.

(ii) Consider now the case $\operatorname{crit}(v) < \operatorname{crit}(w)$. We will show that if $\boldsymbol{v}$ and $\boldsymbol{w}$ do not violate Lemma 4.61 (a) through (e), then they must be adjacent. For the construction of appropriate additional cost vectors $\boldsymbol{a}$, we have to distinguish four sub cases:

	rows $(0,1)$ in $\mathcal{S}^m$?	$\boldsymbol{v}_{\operatorname{crit}(\boldsymbol{w}),*} = (1,0)$?
(1)	—	✓
(2)	✓	✓
(3)	✓	—
(4)	—	—

(ii.1) In this case, vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ are equal in all components except either position $(\operatorname{crit}(v), 1)$ or position $(\operatorname{crit}(v), 2)$. So we can choose $\boldsymbol{a} = \mathbb{0}$ and the situation is similar to (i.1).

(ii.2) First, we define set

$$\mathcal{T} := \{i \in \mathcal{S}^m \mid \exists\, k \in \mathcal{S}^m \text{ with } k \geqslant i \text{ and } \boldsymbol{v}_{k,*} = (0,1)\}.$$

This set contains all rows in $\mathcal{S}^m$ down to the last row of type $(0,1)$. We then define $\boldsymbol{a}$ as follows:

- There is a $b \in \mathbb{R}$ such that $\boldsymbol{ax} \leqslant b$ is a valued block inequality
- $\operatorname{base}(\boldsymbol{a}) = \max(\mathcal{T})$.
- The sign pattern of $\boldsymbol{a}$ is chosen as follows:

$$\operatorname{sign}(\boldsymbol{a}_{i,*}) := \begin{cases} (-,+), & \text{if } i \in \mathcal{S}^t \\ (-,+), & \text{if } i = \operatorname{crit}(\boldsymbol{v}) \\ (-,+), & \text{if } i \in \mathcal{T} \text{ and } \boldsymbol{v}_{i,*} = (0,1) \\ (-,0), & \text{if } i \in \mathcal{T} \text{ and } \boldsymbol{v}_{i,*} = (0,0) \\ (0,+), & \text{if } i \in \mathcal{T} \text{ and } \boldsymbol{v}_{i,*} = (1,1). \end{cases}$$

(See the left picture in figure 4.6.)
Let now $\boldsymbol{x} \in \mathcal{V}[\boldsymbol{c}(\boldsymbol{a})]$ be some $\boldsymbol{c}(\boldsymbol{a})$-maximizing vertex of the orbisack. By definition of $(*)$, $\boldsymbol{x}$ is not completely fixed in row $\operatorname{crit}(\boldsymbol{v})$ and in rows $i \in \mathcal{S}^m$ with $\boldsymbol{v}_{i,*} = (0,1)$. However, $(*)$ ensures that vertex $\boldsymbol{x}$ doesn't have its critical row in any of the latter rows. Hence, there are two possibilities:

- $\operatorname{crit}(\boldsymbol{x}) = \operatorname{crit}(\boldsymbol{v})$. Then all rows $i \in \mathcal{S}^m$ with $\boldsymbol{v}_{i,*} = (0,1)$ are set to $\boldsymbol{x}_{i,*} = (0,1)$ as cost vector $\boldsymbol{a}$ has np-rows there and thus $\boldsymbol{x}$ equals $\boldsymbol{v}$.
- $\operatorname{crit}(\boldsymbol{x}) = \operatorname{crit}(\boldsymbol{w})$. Then in any row above $\operatorname{crit}(\boldsymbol{x})$ that is not completely fixed, one component is fixed by choice of $(*)$. Hence, these rows will be set as in $\boldsymbol{w}$, hence $\boldsymbol{x} = \boldsymbol{w}$.

So, $\mathcal{V}[\boldsymbol{a}(\boldsymbol{c})] = \{\boldsymbol{v}, \boldsymbol{w}\}$.

(ii.3) We define now:

$$\mathcal{T} := \mathcal{S}^m \cup \{i \in \mathcal{S}^b \mid \nexists\, k \in \mathcal{S}^b \text{ with } \boldsymbol{v}_{k,*} = (0,1) \text{ and } k < i\}$$

$\boldsymbol{v}$		$\boldsymbol{w}$		sign($\boldsymbol{a}$)		
1	1	1	1	$-$	$+$	
0	0	0	0	$-$	$+$	
1	0	1	1	$-$	$+$	
1	1	1	1	0	$+$	
0	0	0	0	$-$	0	
0	1	1	1	$-$	$+$	$\mathcal{T}$
0	1	0	0	$-$	$+$	
1	1	1	1	0	0	
1	0	1	0	0	0	
1	1	1	1	0	0	
1	0	1	0	0	0	
0	0	0	0	0	0	
1	1	1	1	0	0	

$\boldsymbol{v}$		$\boldsymbol{w}$		sign($\boldsymbol{a}$)		
1	1	1	1	$-$	$+$	
0	0	0	0	$-$	$+$	
1	0	1	1	$-$	$+$	
1	1	1	1	0	$+$	
0	0	0	0	$-$	0	
0	1	1	1	$-$	$+$	
0	1	0	0	$-$	$+$	
1	1	1	1	0	$+$	$\mathcal{T}$
0	1	1	0	$-$	$+$	
1	1	1	1	0	$+$	
0	1	0	1	$-$	$+$	
0	0	0	0	0	0	
1	1	1	1	0	0	

Figure 4.6: Examples for the choice of additional cost vectors $\boldsymbol{a}$ in case (ii.2) (left) and case (ii.3) (right) for the proof of Proposition 4.62. Critical rows crit($\boldsymbol{v}$) and crit($\boldsymbol{w}$) are drawn gray. In both cases, rows $\boldsymbol{v}_{i,*} = (0,1)$ are allowed in $\mathcal{S}^m$. However, in case (ii.2), $\boldsymbol{v}_{\mathrm{crit}(w),*} = (1,0)$ holds, while in (ii.3), $\boldsymbol{v}_{\mathrm{crit}(()\boldsymbol{w}),*} \in \{(0,1),(1,1),(0,0)\}$ here. This implies that in case (ii.3), $\boldsymbol{v}_{\mathcal{S}^b,*}$ must be in reverse lexicographic order.

and $\boldsymbol{a}$ as in case (ii.2). (See right picture in figure 4.6.)
Set $\mathcal{T}$ now contains all rows in $\mathcal{S}^m$ and the rows in $\mathcal{S}^b$ down to the first row of type $(0,1)$ (which must exist because $\boldsymbol{v}_{\mathcal{S}^b,*}$ must be in lexicographically reverse order not to violate Lemma 4.61 (e)).
Note that by the new definition of set $\mathcal{T}$, base($\boldsymbol{a}$) lies now in set $\mathcal{S}^b$.
The construction of $\boldsymbol{c}(\boldsymbol{a})$-maximizing vertices is basically the same as for case (ii.2).

(ii.4) This case is completely analogous to case (ii.3).

□

4.4.3.2 Number of Edges and Average Degree

To count the edges of the orbisack, we will break up the edge set $\mathcal{E}$ into partitions, depending on the endpoints of the edges. For this purpose, we will split up the set of vertices into classes containing all vertices that are identical in a certain subset of rows. Additionally, we partition the vertices by the position of their critical row into layers (see definition 4.1).

Definition 4.63 Let $\mathcal{R} \subseteq [p]$ be a (possibly empty) set of row indices. We define an equivalence relation $\sim_{\mathcal{R}}$ on $\mathcal{V}$ as follows:

- If $\mathcal{R} = \varnothing$, then $\boldsymbol{v} \sim_{\mathcal{R}} \boldsymbol{w}$ holds for all $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.
- If $\mathcal{R} \neq \varnothing$, then for $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$, it holds that

$$\boldsymbol{v} \sim_{\mathcal{R}} \boldsymbol{w} \quad \Leftrightarrow \quad \boldsymbol{v}_{i,*} = \boldsymbol{w}_{i,*} \; \forall\, i \in \mathcal{R}.$$

So, if $\mathcal{R} = \varnothing$, then set $\mathcal{V}$ is the only equivalence class; if $\mathcal{R} = [p]$, then $\mathcal{V}$ decomposes into $|\mathcal{V}|$ different equivalence classes.

We will now use this equivalence class to partition the set of edges $\mathcal{E}$ into four subsets $\mathcal{E}_1$ through $\mathcal{E}_4$. Subsets $\mathcal{E}_1$ and $\mathcal{E}_2$ contain edges between vertices with same critical row, i.e. in the same layer; subsets $\mathcal{E}_3$ and $\mathcal{E}_4$ contain edges between vertices with different critical rows.

(E1) $\{\boldsymbol{v}, \boldsymbol{w}\} \in \mathcal{E}_1 \Leftrightarrow \{\boldsymbol{v}, \boldsymbol{w}\} \in \mathcal{E}$ and $(\boldsymbol{v}, \boldsymbol{w} \in \mathcal{L}_k$ and $\boldsymbol{v} \sim_{[1..k-1]} \boldsymbol{w})$ for $k \in [1..p-1]$

(E2) $\{\boldsymbol{v}, \boldsymbol{w}\} \in \mathcal{E}_2 \Leftrightarrow \{\boldsymbol{v}, \boldsymbol{w}\} \in \mathcal{E}$ and $(\boldsymbol{v}, \boldsymbol{w} \in \mathcal{L}_k$ and $\boldsymbol{v} \sim_{[k+1..p]} \boldsymbol{w})$ for $k \in [2..p+1]$

(E3) $\{\boldsymbol{v}, \boldsymbol{w}\} \in \mathcal{E}_3 \Leftrightarrow \{\boldsymbol{v}, \boldsymbol{w}\} \in \mathcal{E}$ and $\boldsymbol{v} \in \mathcal{L}_k, \boldsymbol{w} \in \mathcal{L}_\ell,$ with $1 \leqslant k < \ell \leqslant p$

(E4) $\{\boldsymbol{v}, \boldsymbol{w}\} \in \mathcal{E}_4 \Leftrightarrow \{\boldsymbol{v}, \boldsymbol{w}\} \in \mathcal{E}$ and $\boldsymbol{v} \in \mathcal{L}_k$ with $k \in [p], \boldsymbol{w} \in \mathcal{L}_{p+1}$

Proposition 4.64 *The graph $G_{\mathbf{O}_{p,2}}$ of the orbisack $\mathbf{O}_{p,2}$ has*

$$|\mathcal{E}| = 2^p \left(\tfrac{7}{4}\left(1 - 2^p\right) + p(1 + 9 \cdot 2^{p-3})\right)$$

edges.

Proof.

(E1) $\{\boldsymbol{v}, \boldsymbol{w}\} \in \mathcal{E}_1$ implies that $\boldsymbol{v}$ and $\boldsymbol{w}$ share the same rows above and in the critical row. Because of Lemma 4.60 and Proposition 4.62, there are $2(p-k)$ edges in $\mathcal{E}_1$ containing some vector $\boldsymbol{v} \in \mathcal{L}_k$ for $1 \leqslant k \leqslant p-1$. There are $2^{k-1}2^{2(p-k)}$ vertices in layer $\mathcal{L}_k$ for $1 \leqslant k \leqslant p-1$. Introducing factor $\frac{1}{2}$ because of counting each edge twice, we obtain:

$$\begin{aligned}
|\mathcal{E}_1| &= \sum_{k=1}^{p-1} 2^{k-1}2^{2(p-k)}(p-k) = \\
&= 2^{2p-1}\left(p\sum_{k=1}^{p-1}2^{-k} - \sum_{k=1}^{p-1}k\,2^{-k}\right) = \qquad (4.30)\\
&= 2^{2p-1}\left(p\left(1-2^{-p+1}\right) - 2^{-p+1}(2^p - p - 1)\right) = \\
&= 2^p\left(1 - 2^p\left(1 - p\,2^{-1}\right)\right).
\end{aligned}$$

(E2) Two vertices $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{L}_k$ with $\boldsymbol{v} \sim_{[k+1..p]} \boldsymbol{w}$ are identical in and below the critical row k for $2 \leqslant k \leqslant p$. They are neighbours if and only if they differ in exactly one row above the critical row. For $k = p+1$, two vertices $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{L}_k$ are also neighbours if and only they differ in exactly one row (Lemma 4.60, Proposition 4.62, and Definition 4.63).
Hence, for $2 \leqslant k \leqslant p+1$, each vertex in layer $\mathcal{L}_k$ has $k-1$ neighbours (sharing the same rows in and below the critical row if $k < p+1$). Layer $\mathcal{L}_k$ contains 2^{2p-k-1} vertices for $1 \leqslant k \leqslant p$ and 2^p vertices for $k = p+1$.

Summing up over all relevant layers gives

$$
\begin{aligned}
|\mathcal{E}_2| &= p\,\frac{1}{2}2^p + \sum_{k=2}^{p} \frac{1}{2}(k-1)2^{2p-k-1} = \\
&= p\,2^{p-1} + 2^{2p-3}\sum_{k=1}^{p-1} k\,2^{-k} = \\
&= p\,2^{p-1} + 2^{2p-3}\,2^{-(p-1)}\big(2^p - (p-1) - 2\big) = \\
&= 2^{p-2}\big(2^p + p - 1\big).
\end{aligned} \tag{4.31}
$$

(E3) To count the edges between different layers, we work "bottom up". That means: we count the edges that start from some vertex $\boldsymbol{v} \in \mathcal{L}_k$ with $1 < k \leqslant p$ and end at any vertex $\boldsymbol{w} \in \mathcal{L}_{\ell<k}$.
Let us at first fix some vertex $\boldsymbol{v} \in \mathcal{L}_k$. If $\boldsymbol{v}$ is adjacent to some vertex $\boldsymbol{w} \in \mathcal{L}_{\ell<k}$, then $\boldsymbol{v}_{[1..\ell-1],*} = \boldsymbol{w}_{[1..\ell-1],*}$ and $\boldsymbol{v}_{[k+1..p],*} = \boldsymbol{w}_{[k+1..p],*}$ must hold.
However, each row $\boldsymbol{w}_{i,*}$ with $\ell < i < k$ can be chosen either to be the same as in $\boldsymbol{v}$ or to $(0,1)$. This gives $2^{k-\ell-1}$ possibilities.
For row k of $\boldsymbol{w}$, there is either one row type possible, namely

$$\boldsymbol{v}_{k,*} \in \{(1,0)\}, \text{ if } \boldsymbol{v}_{[k+1..p],1} \succcurlyeq \boldsymbol{v}_{[k+1..p],2}$$

or there is one of four row types possible, namely

$$\boldsymbol{v}_{k,*} \in \{(1,0),(0,1),(0,0),(1,1)\}, \text{ if } \boldsymbol{v}_{[k+1..p],1} \prec \boldsymbol{v}_{[k+1..p],2}.$$

That means: vertex $\boldsymbol{v}$ has either $2^{k-\ell-1}$ or $4\cdot 2^{k-\ell-1}$ neighbours in any layer $\mathcal{L}_{\ell<k}$, depending on whether $\boldsymbol{v}_{[k+1..p],*}$ is lexicographically ordered or not. Thus in all layers $\mathcal{L}_{\ell<k}$, the number of neighbours of vertex $\boldsymbol{v}$ sums up to

$$
\begin{array}{rll}
\sum_{\ell=1}^{k-1} 2^{k-\ell-1} & = 2^{k-1}-1, & \text{if } \boldsymbol{v}_{[k+1,p],1} \succcurlyeq \boldsymbol{v}_{[k+1,p],2} \\
2^2\sum_{\ell=1}^{k-1} 2^{k-\ell-1} & = 2^2(2^{k-1}-1), & \text{if } \boldsymbol{v}_{[k+1,p],1} \prec \boldsymbol{v}_{[k+1,p],2}.
\end{array}
$$

In each set $\mathcal{L}_k^{\boldsymbol{m}}$, we find

$$
\begin{array}{lll}
2^{p-k-1}(2^{p-k}+1) & \text{vertices with} & \boldsymbol{v}_{[k+1,p],1} \succcurlyeq \boldsymbol{v}_{[k+1,p],2} \\
2^{p-k-1}(2^{p-k}-1) & \text{vertices with} & \boldsymbol{v}_{[k+1,p],1} \prec \boldsymbol{v}_{[k+1,p],2},
\end{array}
$$

which can be calculated with the help of Proposition 3.19. As there are 2^{k-1} different sets $\mathcal{L}_k^{\boldsymbol{m}}$, we get

$$
\begin{array}{lll}
2^{p-2}(2^{p-k}+1) & \text{vertices with} & \boldsymbol{v}_{[k+1,p],1} \succcurlyeq \boldsymbol{v}_{[k+1,p],2} \\
2^{p-2}(2^{p-k}-1) & \text{vertices with} & \boldsymbol{v}_{[k+1,p],1} \prec \boldsymbol{v}_{[k+1,p],2}
\end{array}
$$

in the entire layer $\mathcal{L}_k$. That means that there are

$$
\begin{aligned}
&\underbrace{2^{p-2}(2^{p-k}+1)\;(2^{k-1}-1)}_{\text{lex. ordered}} + \underbrace{2^{p-2}(2^{p-k}-1)\;2^2(2^{k-1}-1)}_{\text{lex. rev. ordered}} = \\
&\qquad = 2^{p-2}(2^{k-1}-1)(5\cdot 2^{p-k}-3)
\end{aligned} \tag{4.32}
$$

edges between all vertices in layer $\mathcal{L}_k$ and all vertices in layers $\mathcal{L}_{\ell<k}$. We have to sum (4.32) up over all layers $1 < k \leqslant p$. Note that (4.32) equals 0 for $k = 1$, hence we can start the sum from $k = 1$:

$$\begin{aligned} |\mathcal{E}_3| &= \sum_{k=1}^{p} 2^{p-2}(2^{k-1}-1)(5\cdot 2^{p-k}-3) = \\ &= 2^{p-2}\sum_{k=1}^{p}(5\cdot 2^{p-1} - 3\cdot 2^{k-1} - 5\cdot 2^{p-k} + 3) = \\ &= 2^{p-2}\big(5p\cdot 2^{p-1} + 3p - 8(2^p-1)\big). \end{aligned} \tag{4.33}$$

(E4) There are 2^p vertices in layer $\mathcal{L}_{p+1}$. For each $\boldsymbol{v} \in \mathcal{L}_{p+1}$, we can find 2^{p-k} neighbours $\boldsymbol{w} \in \mathcal{L}_{k<p+1}$, as

$$\boldsymbol{w}_{[1..k-1],*} = \boldsymbol{v}_{[1..k-1],*}$$

and

$$\boldsymbol{w}_{i,*} \in \{\boldsymbol{v}_{i,*}, (0,1)\} \text{ for all } i \in [k+1..p]$$

must hold (Lemma 4.60 and Proposition 4.62). Thus all vertices in layer $p+1$ have together 2^{2p-k} neighbours in any layer $\mathcal{L}_{k<p+1}$. Summing up over all $k \in [p]$ gives:

$$|\mathcal{E}_4| = \sum_{k=1}^{p} 2^{2p-k} = 2^p\ (2^p - 1). \tag{4.34}$$

Summing up equations (4.30), (4.31), (4.33), and (4.34), we get in total

$$|\mathcal{E}| = 2^p\ \left(\tfrac{7}{4}\ (1-2^p) + p(1 + 9\cdot 2^{p-3})\right).$$

□

This result immediately gives us the possibility to compute the average degree of a vertex of the orbisack.

Corollary 4.65 (Average vertex degree) *As there are $2^{p-1}(2^p+1)$ vertices in orbisack $\mathbf{O}_{p,2}$, for the average vertex degree $D(p)$, it holds that $D(p) \leqslant 2.25 \cdot p$.*

Proof. It is easy to see that $\lim_{p\to\infty} \frac{D(p)}{p} = \frac{9}{4}$ and that $\frac{D(p)}{p}$ is bounded from above by $\frac{9}{4}$. □

Corollary 4.66 *The graph of the orbisack is Hamilton connected.*

Proof. The orbisack is a combinatorial polytope, so the graph can be either a hypercube or Hamilton connected ([92]). In a d-hypercube, all vertices have same degree d. However, it is obvious that this is not the case for the graph of the orbisack. For instance, a vertex $\boldsymbol{v} \in \mathcal{L}_{p+1}$ has p neighbours inside $\mathcal{L}_{p+1}$ (E3) and

$$\sum_{k=0}^{p-1} 2^k = 2^p - 1$$

into layers $\mathcal{L}_k$ with $k < p+1$ (E5). However, node $\boldsymbol{w} \in \mathcal{L}_k^{\boldsymbol{m}}$ has at least $2^{2(p-k)}$ neighbours in $\mathcal{L}_k^{\boldsymbol{m}}$, and therefore generally (i.e. for $p > 1$), $\boldsymbol{w}$ has a higher degree than $\boldsymbol{v}$. □

Chapter 5

Branched Polyhedral Systems

A large class of dynamic programming algorithms can be associated with a certain class of directed hypergraphs, the DP hypergraphs (see definition 2.11). The solutions of these problems can then be identified with paths or hyperpaths in the hypergraph. Describing these hyperpaths in arc variables, the convex hull of the corresponding incidence vectors has been described by Martin et al. ([86]), see Theorem 2.14. However, things become much more involved as soon as one chooses node variables.

For motivation, we will briefly sketch the general idea for this framework coming from dynamic programming. Let $H = (\mathcal{V}, \widetilde{\mathcal{A}})$ be a DP-hypergraph with source nodes $\mathcal{W}$ and final state $\boldsymbol{t}$, let $\mathcal{L} \subseteq \widetilde{\mathcal{A}}$ be a $\mathcal{W}$-$\boldsymbol{t}$-hyperpath in H, and let the relaxation of H be $D = (\mathcal{V}, \mathcal{A})$. For each $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$, the hypergraph defines a family of sets

$$\mathcal{F}^{\boldsymbol{v}} := \{\mathcal{S} \subset \mathcal{V} \mid (\mathcal{S}, \boldsymbol{v}) \in \widetilde{\mathcal{A}}\}.$$

Hence, $\mathcal{F}^{\boldsymbol{v}}$ contains those sets of states that can be combined and directly translated into state $\boldsymbol{v}$ while running the DP-algorithm.

Based on these families, we can now define a nonempty 0/1-polytope $\mathbf{P}^{\boldsymbol{v}}$ for each $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$ as the convex hull of feasible combinations of in-neighbours in D, i.e.

$$\mathbf{P}^{\boldsymbol{v}} := \operatorname{conv}\left(\left\{\boldsymbol{x} \in \{0,1\}^{\mathrm{N}_D^{\mathrm{in}}(\boldsymbol{v})} \,\middle|\, \boldsymbol{x} = \boldsymbol{x}[\mathcal{S}] \text{ for some } \mathcal{S} \in \mathcal{F}^{\boldsymbol{v}}\right\}\right). \tag{*}$$

Let $\boldsymbol{x} = \boldsymbol{x}\big[\mathcal{V}[\mathcal{L}]\big] \in \{0,1\}^{\mathcal{V}}$ be the incidence vector of the nodes that hyperpath $\mathcal{L}$ is using. From the properties of a hyperpath follows that $\boldsymbol{x}$ must satisfy three properties:

- $x_{\boldsymbol{t}} = 1$
- If component $x_{\boldsymbol{v}}$ is set to 1 and $\boldsymbol{v} \notin \mathcal{W}$, then there must be a set $\mathcal{S} \in \mathcal{F}^{\boldsymbol{v}}$ of in-neighbours of $\boldsymbol{v}$ in D such that $x_{\boldsymbol{w}} = 1$ for all $\boldsymbol{w} \in \mathcal{S}$ and $x_w = 0$ for all $w \in \mathrm{N}_D^{\mathrm{in}}(v) \setminus \mathcal{S}$.
- If component $x_{\boldsymbol{v}} = 1$ and $\boldsymbol{v} \neq \boldsymbol{t}$, then there must be a node $\boldsymbol{w} \in \mathrm{N}_D^{\mathrm{out}}(\boldsymbol{v})$ such that $x_{\boldsymbol{w}} = 1$.

We will refer to these properties as *path conditions*.
In particular, the incidence vector of the $\mathcal{W}$-$\boldsymbol{t}$-hyperpath restricted to the in-neighbours of some node $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$ must be a vertex of $\mathbf{P}^{\boldsymbol{v}}$. These vertices correspond to partial solutions to subproblems of the problem the DP algorithm is solving, and they are arranged in a directed tree (arborescence) structure.

It suggests itself to generalize this approach by replacing polytopes $\mathbf{P}^{\boldsymbol{v}}$ with general polyhedra $\mathbf{Q}^{\boldsymbol{v}}$ in $\mathbb{R}^{\mathrm{N}_D^{\mathrm{in}}(\boldsymbol{v})}$ and adapting the path conditions appropriately. (We will in the following section define more precisely what we mean by that.)
The main goal is to describe the linear hull of all feasible combinations of points

of polyhedra $\mathbf{Q}^{\boldsymbol{v}}$ that belong to hyperpaths by using the linear descriptions of the polyhedra $\mathbf{Q}^{\boldsymbol{v}}$ themselves. We will derive a description of the hyperpath polytope in arc and node variables.

However, in general, a linear description purely in node variables is still pending.

The material presented here is joint work with Volker Kaibel and has been published in [63]. However, that paper goes a bit further than we will in the following chapter: It also studies other applications of branched polyhedral systems, for instance to generalize Balas' unions of polyhedra or to study certain stable set polytopes. In contrast to this, we will focus here exclusively on branched polyhedral systems in the context of orbitopes. Note that different from [63], we direct in this chapter the digraph D that is underlying the branched polyhedral systems from its sources $\mathcal{W}$ to the final state t. This makes the construction consistent with the definitions from the previous chapters, in particular with DP hypergraphs.

5.1 Branched Polyhedral Systems (BPS)

Definition 5.1 (Branched Polyhedral System) Let $D = (\mathcal{V}, \mathcal{A})$ be an acyclic digraph with unique sink $\boldsymbol{t}$ and set of sources $\mathcal{W}$. For each node $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$, a polyhedron $\mathbf{P}^{\boldsymbol{v}} \subseteq \mathbb{R}^{\mathrm{N}^{\mathrm{in}}_D(\boldsymbol{v})}$ is given by

$$\mathbf{P}^{\boldsymbol{v}} := \operatorname{conv}(\widehat{\mathcal{G}}^{\boldsymbol{v}}_{\diamond}) + \operatorname{cone}(\widehat{\mathcal{G}}^{\boldsymbol{v}}_{\heartsuit}),$$

generated by finite sets $\varnothing \neq \widehat{\mathcal{G}}^{\boldsymbol{v}}_{\diamond} \subset \mathbb{R}^{\mathrm{N}^{\mathrm{in}}_D(\boldsymbol{v})}$ and $\widehat{\mathcal{G}}^{\boldsymbol{v}}_{\heartsuit} \subseteq \mathbb{R}^{\mathrm{N}^{\mathrm{in}}_D(\boldsymbol{v})}$, with the additional properties that any point $\boldsymbol{x} \in \widehat{\mathcal{G}}^{\boldsymbol{v}}_{\diamond} \cup \widehat{\mathcal{G}}^{\boldsymbol{v}}_{\heartsuit}$ satisfies

$$r^{\mathrm{in}}_D(\boldsymbol{w}) \cap r^{\mathrm{in}}_D(\boldsymbol{w}') = \varnothing \text{ for } \boldsymbol{w}, \boldsymbol{w}' \in \operatorname{supp}(\boldsymbol{x}) \text{ and } \boldsymbol{w} \neq \boldsymbol{w}'. \tag{*}$$

and that for each $\boldsymbol{x} \in \widehat{\mathcal{G}}^{\boldsymbol{v}}_{\diamond} \cup \widehat{\mathcal{G}}^{\boldsymbol{v}}_{\heartsuit}$, it holds that $x_{\boldsymbol{w}} > 0$ for all $\boldsymbol{w} \in \operatorname{supp}(\boldsymbol{x}) \setminus \mathcal{W}$.

Let $\mathcal{P} := \{\mathbf{P}^{\boldsymbol{v}} \mid \boldsymbol{v} \in \mathcal{V}\}$. We call the pair $(D, \mathcal{P})$ a *branched polyhedral system (BPS)*.

Let $\mathcal{C} = (D, \mathcal{P})$ be some BPS. Pick some $\boldsymbol{v}' \in \mathcal{V} \setminus \mathcal{W}$ and set $\mathcal{V}' := r^{\mathrm{in}}_D(\boldsymbol{v}')$ as well as $\mathcal{P}' := \{\mathbf{P}^{\boldsymbol{v}} \mid \boldsymbol{v} \in \mathcal{V}'\}$. Then the pair $(D[\mathcal{V}'], \mathcal{P}')$ on the subgraph of D induced by $\mathcal{V}'$ is also a BPS. We call it the *truncation of the BPS* $(D, \mathcal{P})$ *at node* $\boldsymbol{v}'$, denoted by $\mathcal{C}^{\boldsymbol{v}'}$.

We proceed with some small remarks.

Remark 5.2

- If polyhedra $\mathbf{P}^{\boldsymbol{v}}$ are pointed, it suffices for the linear description of $\mathbf{P}^{\boldsymbol{v}}$ that the set $\widehat{\mathcal{G}}^{\boldsymbol{v}}_{\diamond}$ only contains all vertices of $\mathbf{P}^{\boldsymbol{v}}$. Similarly, $\widehat{\mathcal{G}}^{\boldsymbol{v}}_{\heartsuit}$ only has to contain all generators of extreme rays of $\mathbf{P}^{\boldsymbol{v}}$. However, we do not assume that a priori.
- As we can identify the in-neighbourhood $\mathrm{N}^{\mathrm{in}}_D(\boldsymbol{v})$ with the in-star $\delta^{\mathrm{in}}_D(\boldsymbol{v})$, we will also consider the polyhedra $\mathbf{P}^{\boldsymbol{v}}$ as subsets of $\mathbb{R}^{\delta^{\mathrm{in}}_D(\boldsymbol{v})}$.

Let from now on $\mathcal{C} = (D, \mathcal{P})$ be some BPS based on some choice of sets $\widehat{\mathcal{G}}^{\boldsymbol{v}}_{\diamond}$ and $\widehat{\mathcal{G}}^{\boldsymbol{v}}_{\heartsuit}$ for all $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$.

Definition 5.3 (Polyhedron associated with BPS) We define set $\mathcal{G}^{\mathcal{C}}_{\diamond}$ as the set of points $\boldsymbol{x} \in \mathbb{R}^{\mathcal{V}}$ with the following properties:

(V.1) $x_{\boldsymbol{t}} = 1$

(V.2) For each $\boldsymbol{v} \in \operatorname{supp}(\boldsymbol{x}) \setminus \mathcal{W}$ it holds that $\frac{1}{x_{\boldsymbol{v}}} \boldsymbol{x}_{\mathrm{N}^{\mathrm{in}}_D(\boldsymbol{v})} \in \widehat{\mathcal{G}}^{\boldsymbol{v}}_{\diamond}$.

(V.3) For each $\boldsymbol{v} \in \operatorname{supp}(\boldsymbol{x}) \setminus \{t\}$, it holds that $\boldsymbol{x}_{\mathrm{N}^{\mathrm{out}}_D(\boldsymbol{v})} \neq \mathbb{0}_{\mathrm{N}^{\mathrm{out}}_D(\boldsymbol{v})}$.

The set $\mathcal{G}^{\mathcal{C}}_{\heartsuit}$ is defined as the set of all points $\boldsymbol{x} \in \mathbb{R}^{\mathcal{V}}$ for which there exists a node $\widetilde{\boldsymbol{v}} \in \mathcal{V} \setminus \mathcal{W}$ with the following properties:

(R.1) $\boldsymbol{x}_{\mathrm{N}^{\mathrm{in}}_D(\widetilde{\boldsymbol{v}})} \in \widehat{\mathcal{G}}^{\widetilde{\boldsymbol{v}}}_{\heartsuit}$

(R.2) Let $\widetilde{\mathcal{V}}_{\boldsymbol{x}} := \mathrm{N}_D^{\mathrm{in}}(\widetilde{\boldsymbol{v}}) \cap \operatorname{supp}(\boldsymbol{x})$. For any $\boldsymbol{w} \in \mathcal{V} \setminus \bigcup_{\boldsymbol{v} \in \widetilde{\mathcal{V}}_{\boldsymbol{x}}} r_D^{\mathrm{in}}(\boldsymbol{v})$, it holds that $x_{\boldsymbol{w}} = 0$. (In particular, $x_{\widetilde{\boldsymbol{v}}} = 0$.)

(R.3) $\frac{1}{x_{\boldsymbol{w}}} \boldsymbol{x}_{r_D^{\mathrm{in}}(\boldsymbol{w})} \in \mathcal{G}_{\diamond}^{\mathcal{C}^{\boldsymbol{w}}}$ for all $\boldsymbol{w} \in \mathrm{N}_D^{\mathrm{in}}(\widetilde{\boldsymbol{v}}) \cap \operatorname{supp}(\boldsymbol{x})$.

The polyhedron associated with BPS $\mathcal{C}$ is then defined as

$$\mathbf{P}^{\mathcal{C}} := \operatorname{conv}(\mathcal{G}_{\diamond}^{\mathcal{C}}) + \operatorname{cone}(\mathcal{G}_{\heartsuit}^{\mathcal{C}})$$

Remark 5.4

- As $\mathbf{P}^{\boldsymbol{v}} \neq \varnothing$ for all $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$, we get that $\mathcal{G}_{\diamond}^{\mathcal{C}} \neq \varnothing$.
- From the finiteness of $\widehat{\mathcal{G}}_{\diamond}^{\boldsymbol{v}}$ for all $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$ follows that $\mathcal{G}_{\diamond}^{\mathcal{C}}$ is finite, as can be seen by applying (V.1) and (V.2) in topological order of $\mathcal{V}$.
- Similarly, one obtains from the finiteness of $\mathcal{G}_{\diamond}^{\mathcal{C}}$ and (R.1) through (R.3) that $\mathcal{G}_{\heartsuit}^{\mathcal{C}}$ is also finite.
- From (V.1), (V.2), (R.1) through (R.3) follows that if $\mathbf{P}^{\boldsymbol{v}} \subseteq \mathbb{R}_+^{\mathrm{N}_D^{\mathrm{in}}(\boldsymbol{v})}$ for all $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$, then $\mathbf{P}^{\mathcal{C}} \subseteq \mathbb{R}_+^{\mathcal{V} \setminus \mathcal{W}} \times \mathbb{R}^{\mathcal{W}}$.
- If $\mathbf{P}^{\boldsymbol{v}}$ is a pointed integral polyhedron for all $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$, then $\mathbf{P}^{\mathcal{C}}$ is also a pointed integral polyhedron.

Remark 5.5 It is also possible to assign to each node $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$ a combinatorial problem. Let $\mathcal{S}^{\boldsymbol{v}}$ be the set of solutions to the combinatorial problem connected with node $\boldsymbol{v}$, and denote by $\mathcal{S}$ the family of all sets $\mathcal{S}^{\boldsymbol{v}}$. Then the pair $(D, \mathcal{S})$ is called a *branched combinatorial system*. It is obvious that by describing the solutions in $\mathcal{S}^{\boldsymbol{v}}$ by incidence vectors, one can derive from each branched combinatorial system a branched polyhedral system, where all $\mathbf{P}^{\boldsymbol{v}} \in \mathcal{P}$ are 0/1-polytopes, namely the convex hulls of the incidence vectors of the feasible solutions.

5.1.1 Linear Description

We start with a small lemma.

Lemma 5.6 *Vector $\boldsymbol{x} \in \mathcal{G}_{\diamond}^{\mathcal{C}}$ if and only if*

(i) $x_{\boldsymbol{t}} = 1$,

(ii) $\frac{1}{x_v} \boldsymbol{x}_{r_D^{\mathrm{in}}(\boldsymbol{v})} \in \mathcal{G}_{\diamond}^{\mathcal{C}^{\boldsymbol{v}}}$ *for each* $\boldsymbol{v} \in \operatorname{supp}(\boldsymbol{x}) \setminus \mathcal{W}$, *and*

(iii) $\operatorname{supp}(\boldsymbol{x})$ *induces an arborescence in* D *rooted at* $\boldsymbol{t}$.

Proof. ⇒ Let $\boldsymbol{x}$ satisfy (i) through (iii). Clearly, (i) implies (V.1), (ii) implies (V.2) and (iii) implies (V.3). Hence, $\boldsymbol{x} \in \mathcal{G}_{\diamond}^{\mathcal{C}}$. ⇐ Let $\boldsymbol{x} \in \mathcal{G}_{\diamond}^{\mathcal{C}}$. Then (i) is satisfied by (V.1). This implies that $\boldsymbol{t} \in \operatorname{supp}(\boldsymbol{x})$. Let now $D' = D[\operatorname{supp}(\boldsymbol{x})]$ be the graph induced by $\operatorname{supp}(\boldsymbol{x})$. Proceeding in topological order starting with $\boldsymbol{t}$, we obtain by inductively applying (V.3) that from any node $\boldsymbol{w} \in D'$, there is a path leading to $\boldsymbol{t}$. Moreover, this path is unique: assume there are two different paths Γ_1 and Γ_2 leading from some $\boldsymbol{u} \in \operatorname{supp}(\boldsymbol{x})$ to $\boldsymbol{t}$. There must be a node $\boldsymbol{v}$ where Γ_1 and Γ_2 intersect with $(\delta_D^{\mathrm{in}}(\boldsymbol{v}) \cap \Gamma_1) \neq (\delta_D^{\mathrm{in}}(\boldsymbol{v}) \cap \Gamma_2)$, contradicting (*). This establishes (iii).
To show (ii), let $\boldsymbol{v} \in \operatorname{supp}(\boldsymbol{x}) \setminus \mathcal{W}$ and let $\boldsymbol{x}' := \frac{1}{x_v} \boldsymbol{x}_{r_D^{\mathrm{in}}(v)}$. (V.1) is trivially satisfied for $\boldsymbol{x}'$, since $\frac{x_v}{x_v} = 1$. (V.2) is satisfied for $\boldsymbol{x}'$, since for any $\boldsymbol{w} \in \operatorname{supp}(\boldsymbol{x}) \cap (r_D^{\mathrm{in}}(\boldsymbol{v}) \setminus \mathcal{W})$, it holds that

$$\frac{1}{x'_w} \boldsymbol{x}'_{\mathrm{N}_D^{\mathrm{in}}(w)} = \frac{1}{\frac{x_w}{x_v}} \frac{1}{x_v} \boldsymbol{x}_{\mathrm{N}_D^{\mathrm{in}}(w)} = \frac{1}{x_w} \boldsymbol{x}_{\mathrm{N}_D^{\mathrm{in}}(w)},$$

which is in $\widehat{\mathcal{G}}_{\diamond}^{\boldsymbol{w}}$, since (V.2) holds for $\boldsymbol{x}$. Last, let $\boldsymbol{w} \in r_D^{\mathrm{in}}(\boldsymbol{v})$. Since $\boldsymbol{x}$ satisfies (V.3), $|\,\mathrm{N}_D^{\mathrm{out}}(\boldsymbol{w}) \cap \operatorname{supp}(\boldsymbol{x})| \geqslant 1$. In fact, the inequality cannot be strict, since then D' could not be an arborescence. □

In our following considerations, we will use the *homogenization* of a polyhedron. Let $\varnothing \neq \mathbf{Q} \subseteq \mathbb{R}^n$ be some nonempty polyhedron. The recession cone of $\mathbf{Q}$ is defined as

$$\operatorname{rec}(\mathbf{Q}) := \{\boldsymbol{w} \in \mathbb{R}^n \mid \boldsymbol{v} + \boldsymbol{w} \in \mathbf{Q} \text{ for all } \boldsymbol{v} \in \mathbf{Q}\}$$

and the homogenization of $\mathbf{Q}$ is defined as

$$\hom(\mathbf{Q}) := \operatorname{cone}(\{(\boldsymbol{v}, 1) \mid \boldsymbol{v} \in \mathbf{Q}\}) + \{(\boldsymbol{v}, 0) \mid \boldsymbol{v} \in \operatorname{rec} \mathbf{Q}\}.$$

Using the linear description $\mathbf{Q} = \{\boldsymbol{x} \in \mathbb{R}^{\boldsymbol{n}} \mid \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\}$ with $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$ for $\mathbf{Q}$, we obtain that

$$\hom(\mathbf{Q}) = \{(\boldsymbol{x}, \lambda) \in \mathbb{R}^n \times \mathbb{R} \mid \boldsymbol{A}\boldsymbol{x} - \lambda \boldsymbol{b} \leqslant \mathbb{0}, \lambda \geqslant 0\}$$

(For details, see for instance section 1.5 in [120].)

Theorem 5.7 *Let $\mathcal{C} = (D, \mathcal{P})$ be a polyhedral branching system, where $D = (V, A)$ is an acyclic digraph with unique sink node $\boldsymbol{t} \in \mathcal{V}$ and set of sources $\mathcal{W} \subseteq \boldsymbol{V}$. Moreover, there is an extended formulation given for each $\mathbf{P}^{\boldsymbol{v}} \in \mathcal{P}$, that is: for each $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$, we have given a polyhedron $\mathbf{Q}^{\boldsymbol{v}} \subseteq \mathbb{R}^{d(\boldsymbol{v})}$ and a linear map $\pi^{\boldsymbol{v}} : \mathbb{R}^{d(\boldsymbol{v})} \to \mathbb{R}^{\delta^{\text{in}}_D(\boldsymbol{v})}$ such that $\mathbf{P}^{\boldsymbol{v}} = \pi^{\boldsymbol{v}}(Q^{\boldsymbol{v}})$.*

Then, the polyhedron $\mathcal{Q}^{\mathcal{B}} \subseteq \mathbb{R}^{\mathcal{V}} \times \mathbb{R}^{\mathcal{A}} \times \bigtimes_{\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}} \mathbb{R}^{d(\boldsymbol{v})}$ defined by the system

$$x_{\boldsymbol{t}} = 1 \tag{5.1}$$

$$x_{\boldsymbol{v}} - \sum_{a \in \delta^{\text{out}}_D(\boldsymbol{v})} y_a = 0 \qquad \text{for all } \boldsymbol{v} \in \mathcal{V} \setminus \{t\} \tag{5.2}$$

$$\boldsymbol{y}_{\delta^{\text{in}}_D(\boldsymbol{v})} - \pi^{\boldsymbol{v}}(\boldsymbol{z}) = 0 \qquad \text{for all } \boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W} \tag{5.3}$$

$$(\boldsymbol{z}^{\boldsymbol{v}}, x_v) \in \hom(\mathbf{Q}^{\boldsymbol{v}}) \qquad \text{for all } \boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W} \tag{5.4}$$

together with the orthogonal projection $\pi : \mathbb{R}^{\mathcal{V}} \times \mathbb{R}^{\mathcal{A}} \times \bigtimes_{\boldsymbol{v} \in \mathcal{C} \setminus \mathcal{W}} \mathbb{R}^{d(v)} \to \mathbb{R}^{\mathcal{V}}$ provide an extended formulation to $\mathcal{P}^{\mathcal{C}}$, that is:

$$\pi(\mathbf{Q}^{\mathcal{C}}) = \mathbf{P}^{\mathcal{C}} = \operatorname{conv}(\mathcal{G}^{\mathcal{C}}_{\diamond}) + \operatorname{cone}(\mathcal{G}^{\mathcal{C}}_{\triangledown}),$$

where $\mathcal{G}^{\mathcal{C}}_{\diamond}$ and $\mathcal{G}^{\mathcal{C}}_{\triangledown}$ are defined with respect to any choice of sets $\widehat{\mathcal{G}}^{\boldsymbol{v}}_{\diamond}$ and $\widehat{\mathcal{G}}^{\boldsymbol{v}}_{\triangledown}$) defining polyhedra $\mathbf{P}^{\boldsymbol{v}}$ for all $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$.

Proof. ⊆ We show that $\mathbf{P}^{\mathcal{C}} \subseteq \pi(\mathbf{Q}^{\mathcal{C}})$ by constructing from each vector $\boldsymbol{x} \in \mathbf{P}^{\mathcal{C}}$ a vector $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ that satisfies (5.1) through (5.4). Let $\boldsymbol{x} \in \mathcal{G}^{\mathcal{C}}_{\diamond}$.

(i) For all $\boldsymbol{v} \in (\mathcal{V} \setminus \mathcal{W}) \setminus \operatorname{supp}(\boldsymbol{x})$, we set $y_{(\boldsymbol{w},\boldsymbol{v})} := 0$ for all $\boldsymbol{w} \in \mathrm{N}^{\text{in}}_D(\boldsymbol{v})$, as well as $\boldsymbol{z}^{\boldsymbol{v}} := \mathbb{0}$.

(ii) For all $\boldsymbol{v} \in (\mathcal{V} \setminus \mathcal{W}) \cap \operatorname{supp}(\boldsymbol{x})$, we set $y_{(\boldsymbol{w},\boldsymbol{v})} := x_{\boldsymbol{w}}$ for all $\boldsymbol{w} \in \mathrm{N}^{\text{in}}_D(\boldsymbol{v})$.

(iii) Since for $\boldsymbol{v} \in (\mathcal{V} \setminus \mathcal{W}) \cap \operatorname{supp}(\boldsymbol{x})$,

$$\mathbf{P}^{\boldsymbol{v}} \overset{(\text{V.2})}{\ni} \frac{1}{x_{\boldsymbol{v}}} \boldsymbol{x}_{\mathrm{N}^{\text{in}}_D(\boldsymbol{v})} \overset{(*)}{=} \frac{1}{x_{\boldsymbol{v}}} \boldsymbol{y}_{\delta^{\text{in}}_D(\boldsymbol{v})},$$

where equality (*) holds because of our definition of $\boldsymbol{y}$ from above, we get that there is a point $\widetilde{\boldsymbol{z}}^{\boldsymbol{v}} \in \mathbf{Q}^{\boldsymbol{v}}$ such that $\pi^{\boldsymbol{v}}(\widetilde{\boldsymbol{z}}^{\boldsymbol{v}}) = \frac{1}{x_{\boldsymbol{v}}} \boldsymbol{y}_{\delta^{\text{in}}_D(\boldsymbol{v})}$ for each $\boldsymbol{v} \in (\mathcal{V} \setminus \mathcal{W}) \cap \operatorname{supp}(\boldsymbol{x})$. This allows us to set $\boldsymbol{z}^{\boldsymbol{v}} := x_{\boldsymbol{v}} \widetilde{\boldsymbol{z}}^{\boldsymbol{v}}$ for each such $\boldsymbol{v}$.

By (V.1), $\boldsymbol{x}_{\boldsymbol{t}} = 1$, so (5.1) is satisfied.
From Lemma 5.6 follows that there is exactly one node $\boldsymbol{w} \in \mathrm{N}^{\text{out}}_D(\boldsymbol{v}) \cap \operatorname{supp}(\boldsymbol{x})$ for each $\boldsymbol{v} \in \operatorname{supp}(\boldsymbol{x})$. From (i) and (ii) follows that $y_{(\boldsymbol{v},\boldsymbol{w})} = x_{\boldsymbol{v}}$ and $y_{(\boldsymbol{v},\boldsymbol{w}')} = 0$ for all $\boldsymbol{w} \neq \boldsymbol{w}' \in \mathrm{N}^{\text{out}}_D(\boldsymbol{v}) \setminus \operatorname{supp}(\boldsymbol{x})$. On the other hand, (i) and (ii) ensure for $\boldsymbol{v} \notin \operatorname{supp}(\boldsymbol{x})$ that $y_a = 0$ for all $a \in \delta^{\text{out}}_D(\boldsymbol{v})$. Hence, (5.2) is satisfied.
By choice of $\boldsymbol{z}^{\boldsymbol{v}}$ in (i) and (iii), $\pi^{\boldsymbol{v}}(\boldsymbol{z}^{\boldsymbol{v}}) = \boldsymbol{y}_{\delta^{\text{in}}_D(\boldsymbol{v})}$ for all $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$. Hence, $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$

satisfies (5.3).
As $x_{\boldsymbol{v}} \geqslant 0$ and by choice of $\boldsymbol{z}$, we get that $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ satisfies (5.4).
Therefore, $\operatorname{conv}(\mathcal{G}_{\diamond}^{\mathcal{C}}) \subseteq \pi(\mathbf{Q}^{\mathcal{C}})$.
Next, we will prove that $\operatorname{cone}(\mathcal{G}_{\heartsuit}^{\mathcal{C}}) \subseteq \pi(\operatorname{rec}(\mathbf{Q}^{\mathcal{C}}))$ by showing that $\lambda \boldsymbol{x}' \in \pi(\operatorname{rec}(\mathbf{Q}^{\mathcal{C}}))$ for any $\boldsymbol{x}' \in \mathcal{G}_{\heartsuit}^{\mathcal{C}}$ and positive scalar $\lambda > 0$.
For $\boldsymbol{x} := \lambda \boldsymbol{x}'$, we define vectors $\boldsymbol{y}$ and $\boldsymbol{z}$ as above except for the root node $\boldsymbol{v}_{\boldsymbol{x}'}$ of $\boldsymbol{x}'$. Here, we define $\boldsymbol{y}_{\delta_D^{\text{in}}(\boldsymbol{v}_{\boldsymbol{x}'})} := \boldsymbol{x}_{\mathrm{N}_D^{\text{in}}(\boldsymbol{v}_{\boldsymbol{x}'})}$. Moreover, we choose a vector $\widetilde{\boldsymbol{z}}^{\boldsymbol{v}_{\boldsymbol{x}'}} \in \operatorname{rec}(\mathbf{Q}^{\boldsymbol{v}_{\boldsymbol{x}'}})$ with $\pi^{\boldsymbol{v}_{\boldsymbol{x}'}}(\widetilde{\boldsymbol{z}}^{\boldsymbol{v}_{\boldsymbol{x}'}}) = \frac{1}{\lambda}\boldsymbol{y}_{\delta_D^{\text{in}}(\boldsymbol{v}_{\boldsymbol{x}'})}$, which must exist, since

$$\pi^{\boldsymbol{v}_{\boldsymbol{x}'}}(\operatorname{rec}(\mathbf{Q}^{\boldsymbol{v}_{\boldsymbol{x}'}})) \overset{(*)}{=} \operatorname{rec}(\mathbf{P}^{\boldsymbol{v}_{\boldsymbol{x}}}) \overset{\text{(R.1)}}{\ni} \boldsymbol{x}'_{\mathrm{N}_D^{\text{in}}(\widetilde{\boldsymbol{v}})} \overset{(**)}{=} \frac{1}{\lambda}\boldsymbol{y}_{\delta_D^{\text{in}}(\widetilde{\boldsymbol{v}})},$$

where (*) holds because $\mathbf{P}^{\boldsymbol{v}_{\boldsymbol{x}'}}$ is the projection of $\mathbf{Q}^{\boldsymbol{v}_{\boldsymbol{x}'}}$, and (**) holds because of our choice of $\boldsymbol{y}$. With the means of $\widetilde{\boldsymbol{v}^{\boldsymbol{x}'}}$, we define now $\boldsymbol{z}^{\boldsymbol{v}_{\boldsymbol{x}'}} := \lambda \widetilde{\boldsymbol{z}}^{\boldsymbol{v}_{\boldsymbol{x}'}}$.
Hence, $\boldsymbol{z}^{\boldsymbol{v}_{\boldsymbol{x}'}} \in \operatorname{rec}(\mathbf{Q}^{\boldsymbol{v}_{\boldsymbol{x}'}})$, which implies that $(\boldsymbol{z}^{\boldsymbol{v}_{\boldsymbol{x}'}}, 0) \in \operatorname{hom}(\mathbf{Q}^{\boldsymbol{v}_{\boldsymbol{x}'}})$. So, for $\boldsymbol{v}_{\boldsymbol{x}'}$ $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ satisfies (5.4).
From (R.2) and (R.3) follows that $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ also satisfies (5.4) for all other $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$, as well as all equations (5.1) through (5.3).

$\supseteq$ We will show that $\pi(\mathbf{Q}^{\mathcal{C}}) \subseteq \mathbf{P}^{\mathcal{C}}$ by proving that for all $\boldsymbol{c} \in \mathbb{R}^{\mathcal{V}}$ and

$$\omega := \max(\{\langle \boldsymbol{c}, \boldsymbol{x}\rangle \mid (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \in \mathbf{Q}^{\mathcal{C}}\}),$$

there exist

- $\boldsymbol{x}^\star \in \mathcal{G}_{\diamond}^{\mathcal{C}}$ with $\langle \boldsymbol{x}^\star, \boldsymbol{c}\rangle = \omega$, if $\omega < \infty$ and
- $\boldsymbol{x}^\star \in \mathcal{G}_{\heartsuit}^{\mathcal{C}}$ with $\langle \boldsymbol{x}^\star, \boldsymbol{c}\rangle > 0$, if $\omega = \infty$.

Let $\mathbf{Q}^{\boldsymbol{v}} = \{\boldsymbol{z}^{\boldsymbol{v}} \in \mathbb{R}^{d(\boldsymbol{v})} \mid \boldsymbol{A}^{\boldsymbol{v}}\boldsymbol{x} \leqslant \boldsymbol{b}^{\boldsymbol{v}}\}$ be the $\mathcal{H}$-representation of $\mathbf{Q}^{\boldsymbol{v}}$ and let $\boldsymbol{T}^{\boldsymbol{v}} \in \mathbb{R}^{\delta_D^{\text{in}}(\boldsymbol{v}) \times [d(\boldsymbol{v})]}$ be the matrix associated with linear projection $\pi^{\boldsymbol{v}}$. Then the inequality system describing $\mathbf{Q}^{\mathcal{C}}$ from above can be reformulated as follows:

$$x_{\boldsymbol{t}} = 1 \tag{5.5}$$

$$x_{\boldsymbol{v}} - \sum_{a \in \delta_D^{\text{out}}(\boldsymbol{v})} \boldsymbol{y}_a = 0 \qquad \forall \boldsymbol{v} \in \mathcal{V} \setminus \{t\} \tag{5.6}$$

$$\boldsymbol{y}_{\delta_D^{\text{in}}(\boldsymbol{v})} - \boldsymbol{T}^{\boldsymbol{v}}\boldsymbol{z}^{\boldsymbol{v}} = 0 \qquad \forall \boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W} \tag{5.7}$$

$$\boldsymbol{A}^{\boldsymbol{v}}\boldsymbol{z}^{\boldsymbol{v}} - x_{\boldsymbol{v}}\boldsymbol{b}^{\boldsymbol{v}} \leqslant 0 \qquad \forall \boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W} \tag{5.8}$$

$$x_{\boldsymbol{v}} \geqslant 0 \qquad \forall \boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W} \tag{5.9}$$

We will construct vector $\boldsymbol{x}^\star$ along the topological order from sources to sink on the node set $\mathcal{V} \setminus \mathcal{W}$.

Since the node set of D is by definition finite, the algorithm terminates, either at some node $\boldsymbol{w} \neq \boldsymbol{t}$ or at sink $\boldsymbol{t}$. The solution $\boldsymbol{v}^\star$ is in $\mathcal{G}_{\diamond}^{\mathcal{C}} \cup \mathcal{G}_{\heartsuit}^{\mathcal{C}}$, which can be shown by induction: the statement is obviously true for all nodes in $\mathcal{W}$. Now let node $\boldsymbol{v}$ some processed node in $\mathcal{V} \setminus \mathcal{W}$. If $\zeta^{\boldsymbol{w}} < \infty$, then $\boldsymbol{x}^{\boldsymbol{v}} \in \mathcal{G}_{\diamond}^{\mathcal{C}^{\boldsymbol{v}}}$ and $c_{\boldsymbol{v}}^\star = \langle \boldsymbol{c}, \boldsymbol{x}^{\boldsymbol{v}}\rangle$; if $\zeta^{\boldsymbol{w}} = \infty$, then $\boldsymbol{x}^{\boldsymbol{v}} \in \mathcal{G}_{\heartsuit}^{\mathcal{C}^{\boldsymbol{v}}}$ and $\langle \boldsymbol{c}, \boldsymbol{x}^{\boldsymbol{v}}\rangle > 0$.
We define now dual variables additionally to the $\boldsymbol{\lambda}$ associated with inequalities (5.8) already defined in the algorithm.

- For inequalities (5.5) and (5.6), we define dual variables $\nu_{\boldsymbol{v}} := c_{\boldsymbol{v}}^\star$ for all $\boldsymbol{v} \in \mathcal{V}$.
- For inequalities (5.7), we define dual variables $\boldsymbol{\mu}^{\boldsymbol{v}} := -\boldsymbol{c}_{\mathrm{N}_D^{\text{in}}(\boldsymbol{v})}$ for all $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$.

The objective value of vector $(\boldsymbol{\nu}, \boldsymbol{\mu}, \boldsymbol{\lambda})$ is obviously $\nu_{\boldsymbol{t}} = c_{\boldsymbol{t}}^\star$.
Moreover, $(\boldsymbol{\nu}, \boldsymbol{\mu}, \boldsymbol{\lambda})$ is dual feasible: for all constraints associated with $\boldsymbol{z}$-variables, we get that

$$(\lambda^{\boldsymbol{v}})^\intercal \boldsymbol{A}^{\boldsymbol{v}} - (\mu^{\boldsymbol{v}})^\intercal \boldsymbol{T}^{\boldsymbol{v}} = ((\boldsymbol{T}^{\boldsymbol{v}})^\intercal \boldsymbol{c}^\star_{\mathrm{N}_D^{\text{in}}(\boldsymbol{v})})^\intercal - (\boldsymbol{c}^\star_{\mathrm{N}_D^{\text{in}}(\boldsymbol{v})})^\intercal \boldsymbol{T}^{\boldsymbol{v}} = \mathbb{0},$$

```
Data: c
Result: x*
// Initialization
foreach v ∈ 𝒱 ∖ 𝒲 do
 | x^v ← 𝟘 ∈ ℝ^𝒱 ;                                  // x^v ∈ ℝ^𝒱
end
foreach v ∈ 𝒲 do
 | x^v ← 𝕖^v ∈ ℝ^𝒱 ;                                 // x^v ∈ ℝ^𝒱
end
c*_𝒲 ← c_𝒲 ;                         // auxiliary cost vector c* ∈ ℝ^𝒱
foreach v ∈ 𝒱 ∖ 𝒲 along topological order do
 | let ζ^v ← max({⟨c*_{N^in_D(v)}, x̃⟩ | x̃ ∈ P^v});
 | if ζ^v = ∞ then
 | | choose x̃ ∈ Ĝ^v_◇ with ⟨c*_{N^in_D(v)}, x̃⟩ > 0;
 | | set x* ← Σ_{w∈supp(x̃)} x̃_w x^v;
 | | break;
 | else
 | | choose x̃ ∈ Ĝ^v_◇ with ⟨c*_{N^in_D(v)}, x̃⟩ = ζ^v;
 | | compute optimal solution λ^v to dual problem
 | | max({⟨(T^v)ᵀ c*_{N^in_D(v)}, z̃⟩ | A^v z̃ ⩽ b^v});
 | | set x^v ← 𝕖^v + Σ_{w∈supp(x̃)} x̃_w x^w;
 | | set c*_v ← c_v + ζ^v;
 | end
end
return x*
```

for all $v \in V \setminus V_T$. For the constraints associated with the $\boldsymbol{y}$-variables, we find that

$$-\mu^{\boldsymbol{v}}_{\boldsymbol{w}} - \nu_{\boldsymbol{w}} = c^\star_{\boldsymbol{w}} - c^\star_{\boldsymbol{w}} = 0$$

for all $(\boldsymbol{v}, \boldsymbol{w}) \in \mathcal{A}$. And for the constraints associated with $\boldsymbol{x}$-variables, we get that

$$-(\lambda^{\boldsymbol{v}})^\intercal b^{\boldsymbol{v}} + \nu_{\boldsymbol{v}} = -\zeta^{\boldsymbol{v}} + c^\star_v = 0.$$

for all $\boldsymbol{v} \in \mathcal{V}$, even for all $\boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W}$. This implies that the nonnegativity constraints (5.9) are redundant. □

The description becomes simpler if the polyhedra $\mathbf{P}^{\boldsymbol{v}}$ are not described by extended formulations.

Corollary 5.8 *Let $\mathcal{C} = (D, \mathcal{P})$ be a BPS based on a digraph $D = (\mathcal{V}, \mathcal{A})$ with sink node $\boldsymbol{t} \in \mathcal{V}$ and set of sources $\mathcal{W} \subseteq \mathcal{V}$. Then polyhedron $\mathbf{Q}^{\mathcal{C}} \subseteq \mathbb{R}^{\mathcal{V}} \times \mathbb{R}^{\mathcal{A}}$ defined by system*

$$x_{\boldsymbol{t}} = 1 \tag{5.10}$$

$$x_{\boldsymbol{v}} - \sum_{a \in \delta^{\text{out}}_D(\boldsymbol{v})} y_a = 0 \qquad \forall \boldsymbol{v} \in \mathcal{V} \setminus \{\boldsymbol{t}\} \tag{5.11}$$

$$(\boldsymbol{y}_{\delta^{\text{in}}_D(\boldsymbol{v})}, x_{\boldsymbol{v}}) \in \text{hom}(\mathbf{P}^{\boldsymbol{v}}) \qquad \forall \boldsymbol{v} \in \mathcal{V} \setminus \mathcal{W} \tag{5.12}$$

and the orthogonal projection $\pi : \mathbb{R}^{\mathcal{V}} \times \mathbb{R}^{\mathcal{A}} \to \mathbb{R}^{\mathcal{V}}$ provide an extended formulation for polyhedron $\mathbf{P}^{\mathcal{C}}$.

5.1.2 BPS and Orbitopes

We will use the branched polyhedral systems now to develop an extended formulation for full orbitopes over the full symmetric group.

The underlying digraph $D = (\mathcal{B}, \mathcal{A})$ is the relaxation of the DP-hypergraph $H = (\mathcal{B}, \mathcal{A})$ associated with orbitopes (see page 55).

Define for each node $b_{i,[s..t]} \in \mathcal{B}$, $i \in [p-1]$, the following set

$$\mathcal{S}^{b_{i,[s..t]}} := \{b^{\bullet}_{i+1,[s..t]}\} \,\cup\, \{b^{\circ}_{i+1,[s..t]}\} \,\cup\, \Big\{\{b^{\bullet}_{i+1,[s..\ell]}, b^{\circ}_{i+1,[\ell+1..t]}\} \,\Big|\, \ell \in [s..t-1]\Big\}$$

and for $b_{0,[1..q]} \in \mathcal{B}$ the set

$$\mathcal{S}^{b_{0,[1..q]}} := \{b^{\bullet}_{1,[1..q]}\} \,\cup\, \{b^{\circ}_{1,[1..q]}\} \,\cup\, \Big\{\{b^{\bullet}_{1,[1..\ell]}, b^{\circ}_{1,[\ell+1..q]}\} \,\Big|\, \ell \in [1..q-1]\Big\}.$$

Denote by $\mathcal{S} := \{\mathcal{S}^{\boldsymbol{v}} \mid \boldsymbol{v} \in \mathcal{V} \setminus \mathcal{B}_p\}$. Then $(D, \mathcal{S})$ is a branched combinatorial system which induces a branched polyhedral system $(D, \mathcal{P})$ over a family of polytopes $\mathbf{P}^{b_{i,[s..t]}}$, $b_{i,[s..t]} \in \mathcal{B} \setminus \mathcal{B}_p$, where each polytope $\mathbf{P}^{b_{i,[s..t]}}$ is defined as $\operatorname{conv}(\{\boldsymbol{x}[\boldsymbol{s}] \in \{0,1\}^{[s..t]} \mid \boldsymbol{s} \in \mathcal{S}^{b_{i,[s..t]}}\})$, that is: the convex hull over all incidence vectors of solutions in $\mathcal{S}^{b_{i,[s..t]}}$.

However, these polytopes can be linearly described, since for each $b_{i,[s..t]} \in \mathcal{B} \setminus \mathcal{B}_p$, $\mathbf{P}^{b_{i,[s..t]}}$ is isomorphic to

$$\operatorname{conv}\big(\{\mathbb{e}^1, \mathbb{e}^2\} \cup \{\mathbb{e}^{2\ell-1} + \mathbb{e}^{2\ell} \mid \ell \in [2..n]\}\big), \tag{*}$$

with $n = t - s + 1$. It is obvious that the polytope (*) can be linearly described by

$$\begin{aligned} u_{2\ell-1} + u_{2\ell} &= 0 \quad \forall \ell \in [2..n] \\ 2u_1 + 2u_2 + \textstyle\sum_{\ell=3}^{2n} u_\ell &= 2 \\ u_i &\geqslant 0 \quad \forall i \in [2n] \end{aligned}$$

Thus, we obtain the following proposition:

Proposition 5.9 *The full orbitope $\mathbf{O}_{p,q}$ over the full symmetric group is a projection of the polyhedron defined by*

$$x_{b_{i,[s..t]}} - \sum_{a \in \delta^{\text{out}}_D(b_{i,[s..t]})} y_a = 0 \;\; \forall b_{i,[s..t]} \in \mathcal{B} \setminus \{b_t\}$$

$$y_{(b^{\bullet}_{1,[1..\ell]}, b_t)} - y_{(b^{\circ}_{1,[\ell+1..q]}, b_t)} = 0 \;\; \forall \ell \in [q-1]$$

$$\begin{aligned} &2\big(y_{(b^{\bullet}_{1,[1..q]}, b_t)} + y_{(b^{\circ}_{1,[1..q]}, b_t)}\big) + \\ &+ \sum_{\ell=1}^{q-1} \big(y_{(b^{\bullet}_{1,[1..\ell]}, b_t)} + y_{(b^{\circ}_{1,[\ell+1..q]}, b_t)}\big) = 2 \end{aligned}$$

$$y_{(b^{\bullet}_{i+1,[s..\ell]}, b_{i,[s..t]})} - y_{(b^{\circ}_{i+1,[\ell+1..t]}, b_{i,[s..t]})} = 0 \;\; \forall \ell \in [s..t-1] \text{ and } b_{i,[s..t]} \in \mathcal{B} \setminus \mathcal{B}_p$$

$$\begin{aligned} &2\big(y_{(b^{\bullet}_{i+1,[s..t]}, b_{i,[s..t]})} + y_{(b^{\circ}_{i+1,[s..t]}, b_{i,[s..t]})}\big) + \\ &+ \sum_{\ell=1}^{t-1} \big(y_{(b^{\bullet}_{i+1,[s..\ell]}, b_{i,[s..t]})} + y_{(b^{\circ}_{i+1,[\ell+1..t]}, b_{i,[s..t]})}\big) - 2x_{b_{i,[s..t]}} = 2 \;\; \forall b_{i,[s..t]} \in \mathcal{B} \setminus \mathcal{B}_p \end{aligned}$$

$$y_a \geqslant 0 \;\; \forall a \in \mathcal{A}$$

where $b_t = b_{0,[1..q]}$ and

Proof. First project the polyhedron orthogonally to the $\boldsymbol{x}$-coordinates, then use projection $\widetilde{\vartheta}$ from remark 3.39. $\square$

Note that this extended formulation is isomorphic to the one from Theorem 2.14. It remains an open question whether this extended formulation can be useful for obtaining the linear description of full orbitopes; it is for instance unclear how the projection cone for this polyhedron can be generated.

Chapter 6

Conclusions

One main purpose of this thesis is to mark off for which classes of orbitopes "nice" inequality descriptions can be expected. We showed that if the full symmetric group is operating on the columns, then in general, the full orbitopes and the orbitopes isomorphic to packing and partitioning orbitopes are the only orbitopes that allow such a linear description, at least as long as $\mathcal{NP} \neq$ co-$\mathcal{NP}$. The reason is that in general, optimization is $\mathcal{NP}$-hard over any orbitope over the full symmetric group apart from the above mentioned. On the other hand, the dynamic programming algorithm for optimization over full orbitopes which runs in time $\mathcal{O}(pq^3)$ still gives hope that a linear description for these orbitopes can be found.

A series of computer experiments indicated nevertheless that the facial structure of the full orbitope seems to be quite involved, as soon as the number of columns exceeds two. There seems to be no obvious set of rules that would allow to guess how to generate the set of facets in general.

Therefore, for "warming up" and deriving a better understanding of full orbitopes in general, the full orbitope with two columns (the so-called orbisack) has been studied in detail. We present in this thesis a set of tools and methods that allow the derivation of the linear description of orbisacks in three different ways. Two of these approaches, namely the faithful sectioning and the direct combinatorial proof, even yield a non-redundant description. Since the application of faithful sectioning requires an appropriate extended formulation, we studied branched polyhedral systems (BPS) in order to develop extended formulations for full orbitopes. However, we obtained only an extended formulation isomorphic to the one based on the work of Kipp Martin et al.

The framework of BPS can be seen as a tool of very general possibilities; it generalizes for instance Balas' disjunctive programming. However, it has been treated in this thesis only very briefly, in so far it concerns orbitopes. It would be a research field in itself to study BPS in detail. However, this study could prove to be interesting, since BPS may in future play in particular a role in context with the polyhedral description of the solution set of combinatorial algorithms on graphs that can be decomposed in certain ways. (There are several ideas of how to decompose graphs. For tree decompositions, see [20] or the works of Arnborg and Bodlaender, e.g. [4, 19], for the strip-decomposition of claw-free graphs [24], for other decompositions e.g. [60, 27, 116]). There is evidence that graph decomposition can open new possibilities for polyhedral descriptions associated with combinatorial problems (see for instance Margot's thesis [81]). A concrete example is the linear description of the stable set polytope for a subclass of claw-free graphs due to Oriolo et al. ([96]), which is tightly connected to branched polyhedral systems.

As has been shown in examples in chapter 2, the tools developed to obtain the linear description of orbisacks are not restricted to use with orbisacks; they can definitely be applied to other polytopes. We used them in particular to prove linear descriptions of clique polyhedra with clique size less or equal 2 and exactly 2 and the path set polytope of a directed acyclic digraph.

Nevertheless, the tools can apparently not easily be extended to orbitopes in general, although they fit well with orbisacks. The main reason for this — apart from the lack of an idea of the facial structure of orbitopes we mentioned before — is that for the known extended formulations for orbitopes $\mathbf{O}_{p,q>2}$, which are based on the path polytope in arc variables and on branched polyhedral systems (BPS), no constructions of faithful sectionings are known. Since full orbitopes are in general no knapsack polytopes and the combinatorial proof seems also not to be extendable to more than two columns, neither of the three approaches presented in chapter 4 can be directly adjusted to orbitopes in general.

For the same reasons, we are also skeptical that a linear description of the full orbitope $\mathbf{O}_{p,q}$ with $q > 2$ can be obtained by means of a modification of one of the three approaches from chapter 4. However, another extended formulation based on BPS could perhaps open a new direction. One such extended formulation in node and arc variables is given in chapter 5; it is unclear how its projection to node variables looks like. (This would give the hyperpath set polytope on the DP hypergraph associated with full orbitopes). For the linear description of orbitopes, these extended formulations can possibly be useful in connection with faithful sectionings.

A different technique is based on an idea of Lovász ([77]). As Weismantel and Pochet have shown ([117]), this approach proves to work for sequential knapsack polytopes, mainly because of the existence of a dynamic programming algorithm for these polytopes. Using the dynamic programming algorithm from section 3.2.1.2.4, their strategy could therefore perhaps be adapted to orbitopes: For any cost vector $\boldsymbol{c} \in \mathbb{Q}^{[p]\times[q]}$, one would formulate an inequality which is valid for $\mathbf{O}_{p,q}$ and which defines a face containing all vertices of $\mathbf{O}_{p,q}$ maximizing the functional $\langle \boldsymbol{x}, \boldsymbol{c} \rangle$. The aim is to make the set finite by a suitable construction of these inequalities, and to identify all facet defining inequalities in it. However, it is currently unclear how these steps could be done in detail. Hence, there is still a lot of work to be done.

Much time has been spent on attempts to show that the edge expansion of the orbisack is bounded by 1 from below, which would further confirm the conjecture of Mihail and Vazirani ([91]). (We conjecture that the bound of 1 is tight in case of the orbisack.) For these purposes, the adjacency structure of the orbisack has been worked out. However, it turned out that the inductive approaches which have been taken probably do not sufficiently take into account the structure of the graph. We think that approaches which incorporate the details of the adjacency structure are more promising.

Last, not much is known if groups different from the full symmetric group or products of full symmetric groups are permuting the columns, in particular if the cyclic group acts on the columns. For these orbitopes, possible applications could possibly be found in scheduling problems, for instance for the development of cyclic timetables. However, our computer experiments indicate that the facial structure of full orbitopes over cyclic groups is even more intricate than for the symmetric group. An interesting next step would therefore be to prove or disprove that optimization over full orbitopes over full cyclic groups is $\mathcal{NP}$-complete.

Bibliography

[1] Karen Aardal and Stan van Hoesel, *Polyhedral Techniques in Combinatorial Optimization — I (Theory) and II (Computations)*, Statistica Nederlandica (1995).

[2] Karen Aardal and Robert Weismantel, *Polyhedral combinatorics: An annotated bibliography*, Annotated bibliographies in combinatorial Optimization (M. Dell'Amico, F. Maffioli, and Silvano Martello, eds.), 1998.

[3] Stefan Arnborg, *Efficient algorithms for combinatorial problems on graphs with bounded decomposability — (Survey)*, BIT Numerical Mathematics **25** (1985), no. 1, 1–23.

[4] Stefan Arnborg, Jens Lagergren, and Detlef Seese, *Easy problems for tree-decomposable graphs*, Journal of Algorithms **12** (1991), no. 2, 308–340.

[5] Giorgio Ausiello, Paolo G. Franciosa, and Daniele Frigioni, *Directed Hypergraphs: Problems, Algorithmic Results, and a Novel Decremental Approach*, Lecture Notes in Computer Science (2001), 312–327.

[6] Giorgio Ausiello and Guiseppe F. Italiano, *On-line algorithms for polynomially solvable satisfiability problems*, Journal of Logic Programming **10** (1990), no. 1, 90.

[7] Sven Axsäter, *Inventory control*, Springer, 2006.

[8] Egon Balas, *Disjunctive programming: properties of the convex hull of feasible points*, Discrete Applied Mathematics **89** (1998), no. 1-3, 3–44.

[9] ———, *Projection and lifting in combinatorial optimization*, Lecture Notes in Computer Science (2001), 26–56.

[10] ———, *Projection and lifting and extended formulation in integer and combinatorial optimization*, Annals of Operations Research (2005), no. 140, 125–161.

[11] Egon Balas and William R. Pulleyblank, *The perfectly matchable subgraph polytope of a bipartite graph*, Networks **13** (1983), 495–516.

[12] ———, *The perfectly matchable subgraph polytope of an arbitrary graph*, Combinatorica **9** (1989), no. 4, 321–337.

[13] Jørgen Bang-Jensen and Gregory Gutin, *Digraphs: Theory, algorithms and applications*, Springer, 2006.

[14] Richard E. Bellman, *Notes on the theory of dynamic programming IV — maximization over discrete sets*, Naval research Logistics Quaterly (1956), no. 3, 67–70.

[15] ______, *Comment on Dantzig's paper on discrete variable extremum problems*, Operations Research (1957), no. 5, 723–724.

[16] ______, *Dynamic programming*, Princeton University Press, 1957.

[17] Dimitri P. Bertsekas, *Dynamic programming and optimal control*, vol. I, Athena Scientific, Belmont, Massachusetts, 2000.

[18] ______, *Dynamic programming and optimal control*, vol. II, Athena Scientific, Belmont, Massachusetts, 2001.

[19] Hans L. Bodlaender, *Treewidth: Algorithmic techniques and results*, Mathematical Foundations of Computer Science 1997 (1997), 19–36.

[20] Hans L. Bodlaender and Arie M. C. A. Koster, *Combinatorial Optimization on Graphs of Bounded Treewidth*, The Computer Journal (2007), 1–15.

[21] John Adrian Bondy and Uppaluri Siva Ramachandra Murty, *Graph theory*, Graduate Texts in Mathematics, vol. 244, Springer, 2008.

[22] Andreas Brandstädt, Thilo Klembt, Vadim V. Lozin, and Raffaele Mosca, *Independent sets of maximum weight in apple-free graphs*, LNCS 5369 (Seok-Hee Hong, Hiroshi Nagamochi, and Takuro Fukunaga, eds.), Springer, 2008, pp. 848–858.

[23] Richard A. Brualdi and Herbert John Ryser, *Combinatorial matrix theory*, Cambridge University Press, 1991.

[24] Maria Chudnovsky and Paul D. Seymour, *Surveys in combinatorics*, London Mathematical Society Lecture Note Series, vol. 327, ch. The structure of claw-free graphs, pp. 153–172, Cambridge University Press, 2005.

[25] ______, *Claw-free graphs IV. global structure.*, Journal of Combinatorial Theory, Series B **98** (2008), 1373–1410.

[26] Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli, *Extended Formulations in Combinatorial Optimization*, Tech. report, 2009.

[27] Bruno Courcelle and Stephan Olariu, *Clique width: A graph complexity measure — preliminary results and open problems*, Proceedings of the 5th International Workshop on Graph Grammars and Their Application to Computer Science, 1994, pp. 263–270.

[28] George B. Dantzig, *Discrete variable extremum problems*, Operations Research (1957), no. 5, 266–277.

[29] Babette Lucie Elisabeth DeFluiter, *Algorithms for graphs with small treewidth*, Ph.D. thesis, Universiteit Utrecht, 1997.

[30] Stuart E. Dreyfus and Robert A. Wagner, *The Steiner problem in graphs*, Networks **1** (1972), no. 3, 195–207.

[31] Jack Edmonds, *Maximum matching and a polyhedron with* $0,1$ *vertices*, Journal of Research of the National Bureau of Standards **69 B** (1965), 125–130.

[32] Andreas Eisenblätter, *Frequency assignment in GSM networks: Models, heuristics, and lower bounds*, Ph.D. thesis, Technical University Berlin, 2001.

[33] Friedrich Eisenbrand, Gianpaolo Oriolo, Gautier Stauffer, and Paolo Ventura, *The stable set polytope of quasi-line graphs*, Combinatorica **28** (2008), no. 1, 45–67.

[34] Cees Elzinga, Sven Rahmann, and Hui Wang, *Algorithms for subsequence combinatorics*, Theoretical Computer Science **409** (2008), 394–404.

[35] Yuri Faenza and Volker Kaibel, *Extended formulations for packing and partitioning orbitopes*, Mathematics of Operations Research **34** (2009), no. 3, 686–697.

[36] Thorsten Fahle, Stefan Schamberger, and Meinolf Sellmann, *Symmetry breaking*, vol. Lecture Notes in Computer Science 2239, pp. 93–107, Springer, 2001.

[37] Tomás Feder and Milena Mihail, *Balanced matroids*, STOC '92: Proceedings of the twenty-fourth annual ACM symposium on Theory of computing (New York, NY, USA), ACM Press, 1992, pp. 26–38.

[38] Uriel Feige, *A threshold of* $\ln n$ *for approximating set cover*, Journal of the ACM **45** (1998), no. 4, 634–652.

[39] Sàndor P. Fekete and Jörg Schepers, *A combinatorial characterization of higher-dimensional orthogonal packing*, Mathematics of Operations Research **29** (2004), 353–368.

[40] Gert Fischer, *Lineare Algebra*, 9 ed., Vieweg, Braunschweig, 1989.

[41] Eric J. Friedman, *Fundamental domains for integer programs with symmetries*, COCOA'07: Proceedings of the 1st international conference on Combinatorial optimization and applications (Andreas Dress, Yinfeng Xu, and Binhai Zhu, eds.), Springer-Verlag, 2007, pp. 146–153.

[42] Giorgio Gallo, Giustino Longo, and Stefano Pallottino, *Directed hypergraphs and applications*, Discrete Applied Mathematics **42** (1993), no. 2, 177–201.

[43] Giorgio Gallo and Maria Grazia Scutellà, *Directed hypergraphs as a modelling paradigm*, Decisions in Economics and Finance **21** (1998), no. 1, 97–123.

[44] Anna Gallucio, Claudio Gentile, and Paolo Ventura, *Gear composition and the stable set polytope*, Operations Research Letters **36** (2008), no. 4, 419–423.

[45] ———, *On the stable set polytope of claw-free graphs*, COCOA 2008 LNCS 5165 (B. Yang, D.-Z. Du, and W.A. Wang, eds.), Springer, 2008, pp. 339–350.

[46] Michael R. Garey and David S. Johnson, *Computers and intractability: A guide to the theory of NP-completeness*, W. H. Freeman & Co, 1979.

[47] Ewgenij Gawrilow and Michael Joswig, *polymake: a framework for analyzing convex polytopes*, Polytopes — Combinatorics and Computation (Gil Kalai and Günter M. Ziegler, eds.), Birkhäuser, 2000, pp. 43–74.

[48] Ian P. Gent, Karen E. Petrie, and Jean-François Puget, *Symmetry in constraint programming*, Handbook of Constraint Programming (F. Rossi, P. van Beck, and Toby Walsh, eds.), 2006.

[49] Rafael Gillmann, 0/1-*polytopes: Typical and extremal properties*, Ph.D. thesis, Technical University Berlin, 2007.

[50] Rafael Gillmann and Volker Kaibel, *Revlex-initial 0/1-polytopes*, Journal of Combinatorial Theory **A 113** (2006), no. 5, 799–821.

[51] Stefania Gnesi, Ugo Montanari, and Alberto Martelli, *Dynamic programming as graph searching: An algebraic approach*, Journal of the ACM **28** (1981), no. 4, 737–751.

[52] Martin Grötschel, László Lovász, and Alexander Schrijver, *The ellipsoid method and its consequences in combinatorial optimization.*, Combinatorica **1** (1981), no. 2, 169–197.

[53] Branko Grünbaum, *Convex polytopes*, Graduate Texts in Mathematics, Springer, 2003.

[54] Warwick Harvey, *Symmetry breaking and the social golfer problem*, Proceedings of Symmetry in Constraints (SymCon-01) (2001).

[55] Katrin Herr and Richard Bödi, *Symmetries in integer programs*, arxiv (2009), no. 0908.3331v1.

[56] ———, *Symmetries in linear and integer programs*, arxiv (2009), no. 0908.3329v1.

[57] Alan J. Hoffman, *Some recent applications of the theory of linear inequations to extremal combinatorial analysis*, Proceedings of Symposia in Applied Mathematics, American Mathematical Society, Providence (Richard E. Bellman and Marshall Hall, eds.), vol. 10, 1960, pp. 113–127.

[58] Liang Huang, *Advanced dynamic programming in semiring and hypergraph frameworks*, Coling 2008, Manchester, UK (2008).

[59] Jeannette Janssen and Kyriakos Kilakos, *Bounded stable sets: Polytopes and colorings*, Siam Journal of Discrete Mathematics **12** (1999), no. 2, 262–275.

[60] Öjvind Johansson, *Graph decomposition using node labels*, Ph.D. thesis, Royal Insitute of Technology, Department of Numerical Analysis and Computer Science, Stockholm, 2001.

[61] Volker Kaibel, *Polyhedral combinatorics of the quadratic assignment problem*, Ph.D. thesis, University of Köln, 1997.

[62] ———, *Two theorems on projections of polyhedra*, manuscript, 2009.

[63] Volker Kaibel and Andreas Loos, *Branched polyhedral systems*, Integer Programming and Combinatorial Optimization (Friedrich Eisenbrand and F. Bruce Shepherd, eds.), Lecture Notes in Computer Science 6080, vol. Proceedings of IPCO XIV, Springer, 2010, pp. 177–190.

[64] Volker Kaibel, Matthias Peinhardt, and Marc E. Pfetsch, *Orbitopal Fixing*, IPCO (Matteo Fischetti and David P. Williamson, eds.), Lecture Notes in Computer Science, vol. 4513, Springer, 2007, pp. 74–88.

[65] Volker Kaibel and Marc E. Pfetsch, *Packing and partitioning orbitopes*, Mathematical Programming **114** (2008), no. 1, 1–36.

[66] Gil Kalai, *Polytope skeletons and paths*, pp. 331–344, CRC Press, Inc., Boca Raton, FL, USA, 1997.

[67] Richard M. Karp and Christos H. Papadimitriou, *On linear characterizations of combinatorial optimization problems*, Siam Journal on Computing (1982), no. 11, 620–632.

[68] Donald E. Knuth, *The Art of Computer Programming. Vol. 3, Sorting and Searching*, Addison-Wesley Reading, MA, 1973.

[69] Matthias Köppe, Robert Weismantel, and Utz-Uwe Haus, *A primal all-integer algorithm based on irreducible solutions*, Mathematical Programming Series B **96** (2003), 205–246.

[70] Matthias Köppe, Robert Weismantel, Quentin Louveaux, and Laurence A. Wolsey, *Extended formulations for gomory corner polyhedra*, Discrete Optimization **1** (2004), 141–165.

[71] Bernhard Korte and Jens Vygen, *Combinatorial optimization (theory and algorithms)*, 4 ed., Springer, 2008.

[72] Joseph B. Kruskal and David Sankoff, *String edits and macromolecules: The theory and practice of sequence comparision*, Addison-Wesley Reading, MA, 1983.

[73] A.J.D. (Fred) Lambert, *Disassembly sequencing: a survey*, International Journal of Production Research **41** (2003), no. 16, 3721–3759.

[74] Art Lew and Holger Mauch, *Dynamic programming (a computational tool)*, Springer, 2007.

[75] Falko Lorenz, *Lineare Algebra I*, vol. 1, B.I. Wissenschaftsverlag, Mannheim, 1988.

[76] ———, *Lineare Algebra II*, vol. 2, B.I. Wissenschaftsverlag, Mannheim, 1989.

[77] László Lovász, *Graph theory and integer programming*, Annals of Discrete Mathematics (1979), no. 4, 141–158.

[78] László Lovász and Alexander Schrijver, *Matrix cones, projection representations and stable set polyhedra*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science **1** (1990), 1–17.

[79] Vadim V. Lozin and Raffaele Mosca, *Independent sets in extensions of $2K_2$-free graphs*, Discrete Applied Mathematics **146** (2005), 74–80.

[80] Carsten Lund and Mihalis Yannakakis, *On the hardness of approximating minimization problems*, Journal of the ACM **41** (1994), no. 5, 960–981.

[81] François Margot, *Composition de polytopes combinatoires: Une approche par projection*, Ph.D. thesis, École Polytechnique Fédérale de Lausanne, 1994.

[82] ———, *Pruning by isomorphism in branch-and-cut*, Mathematical Programming **A** (2002), no. 94, 71–90.

[83] ———, *Exploiting orbits in symmetric ILP*, Mathematical Programming **B** (2003), no. 98, 3–21.

[84] ———, *Symmetric ILP: Coloring and small integers*, Discrete Optimization **4** (2007), 40–62.

[85] ———, *Symmetry in Integer Linear Programming*, Tepper Working Paper **E-37** (2008), Carnegie Mellon University.

[86] R. Kipp Martin, Ronald L. Rardin, and Brian A. Campbell, *Polyhedral characterization of discrete dynamic programming*, Operations Research **38** (1990), no. 1, 127–138.

[87] D. McKay, *Isomorphim-free exhaustive generation*, Journal of Algorithms **26** (1998), 306–324.

[88] Isabel Méndéz-Díaz and Paul Zabala, *A polyhedral approach for graph coloring*, Electronics Notes in Discrete Mathematics **7** (2001).

[89] ______, *A cutting plane algorithm for graph coloring*, Discrete Applied Mathematics **156** (2008), 159–179.

[90] Isabel Méndéz-Díaz, Paul Zabala, Javier Marenco, and Pablo Coll, *Facets for the graph coloring polytope*, Annals of Operations Research **116** (2002), 79–90.

[91] Milena Mihail, *On the expansion of combinatorial polytopes*, MFCS '92: Proceedings of the 17th International Symposium on Mathematical Foundations of Computer Science, Springer-Verlag, London, UK, 1992, pp. 37–49.

[92] Denis Naddef and William R. Pulleyblank, *Hamiltonicity and combinatorial polyhedra*, Journal of Combinatorial Theory **31** (1981), no. B, 297–312.

[93] Saul B. Needleman and Christian D. Wunsch, *A general method applicable to the search of similarities in the amino-acid sequence of two proteins*, Journal for Molecular Biology (1970), no. 48, 443–453.

[94] George L. Nemhauser and Laurence A. Wolsey, *Integer and combinatorial optimization*, John Wiley & Sons, New York, 1988.

[95] Sang Nguyen and Stefano Pallottino, *Hyperpaths and shortest hyperpaths*, COMO '86: Lectures given at the third session of the Centro Internazionale Matematico Estivo (C.I.M.E.) on Combinatorial optimization (New York, NY, USA), Springer-Verlag New York, Inc., 1989, pp. 258–271.

[96] Gianpaolo Oriolo, Yuri Faenza, and Gautier Stauffer, *The hidden matching structure of the composition of strips: A polyhedral perspective*, to appear, 2010.

[97] James Ostrowski, Jeff Linderoth, Fabrizio Rossi, and Stefano Smriglio, *Orbital Branching*, IPCO (Matteo Fischetti and David P. Williamson, eds.), Lecture Notes in Computer Science, vol. 4513, Springer, 2007, pp. 104–118.

[98] ______, *Constraint orbital branching*, IPCO (Andrea Lodi, Alessandro Panconesi, and Giovanni Rinaldi, eds.), Lecture Notes in Computer Science, vol. 5035, Springer, 2008, pp. 225–239.

[99] Manfred Padberg, *On the facial structure of set packing polyhedra*, Mathematical Programming (1973), no. 5, 199–215.

[100] Kyungchul Park, Kyungsik Lee, and Sungsoo Park, *An extended formulation to the edge-weighted maximal clique problem*, European Journal of Operational Research **95** (1996), 671–682.

[101] Daniele Pretolani and Lars Relund Nielsen, *A remark on the definition of b-hyperpath*, 2001.

[102] William R. Pulleyblank and F. Bruce Shepherd, *Formulations for the stable set polytope*, Proceeding of IPCO III, Springer, 1993, pp. 267–279.

[103] Ronald C. Read, *Every one a winner*, Annals of Discrete Mathematics (1978), no. 2, 107–120.

[104] Ed Rothberg, *Using cuts to remove symmetries*, Presentation ISMP 2000, 2000.

[105] Raman Sanyal, Bernd Sturmfels, and Frank Sottile, *Orbitopes*, arxiv 0911.5436v3 (2010).

[106] Alexander Schrijver, *Theory of linear and integer programming*, John Wiley & Sons, 1986.

[107] ______, *Combinatorial optimization (polyhedra and efficiency)*, vol. A-C, Springer, 2004.

[108] F. Bruce Shepherd, *Near perfection and stable set polyhedra*, Ph.D. thesis, University of Waterloo, 1990.

[109] Hanif D. Sherali and Warren P. Adams, *A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems*, Siam Journal of Discrete Mathematics **3** (1990), no. 3, 411–430.

[110] Barbara M. Smith, *Reducing symmetry in a combinatorial design problem*, CPAIOR'01: Proceedings of the Third International Workshop on Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems, 2001, pp. 351–359.

[111] Rüdiger Stephan, *Polyhedral aspects of cardinality constrained combinatorial optimization problems*, Ph.D. thesis, Technical University Berlin, 2009.

[112] John H. Vande Vate, *The path set polytope of an acyclic, directed graph with an application to machine sequencing*, Networks **19** (1989), no. 5, 607–614.

[113] Robert J. Vanderbei, *Linear programming – foundations and extensions*, Springer, 2008.

[114] Rolf Walter, *Lineare Algebra und analytische Geometrie*, Vieweg, Braunschweig, 1985.

[115] ______, *Einführung in die lineare Algebra*, 2 ed., Vieweg, Braunschweig, 1986.

[116] Egon Wanke and Frank Gurski, *k-nlc graphs and polynomial algorithms*, Discrete Applied Mathematics **54** (1994), no. 2-3, 251–266.

[117] Robert Weismantel and Yves Pochet, *The sequential knapsack polytope*, Siam Journal of Optimization **8** (1998), no. 1, 248–264.

[118] Laurence A. Wolsey, *Strong formulations for mixed integer programs: valid inequalities and extended formulations*, Mathematical Programming **97** (2003), no. 1–2, 423–447.

[119] Laurence A. Wolsey and Yves Pochet, *Production planning by mixed integer programming*, Springer Series in Operations Research and Financial Engineering ed., Springer, 2006.

[120] Günter M. Ziegler, *Lectures on polytopes*, Springer, 1994.

[121] ______, *Lectures on 0/1-polytopes*, Polytopes — Combinatorics and Computation (Gil Kalai and Günter M. Ziegler, eds.), DMV Seminar 29, Birkhäuser, 2000, pp. 1–42.

Schriftliche Erklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Insbesondere habe ich nicht die Hilfe einer kommerziellen Promotionsberatung in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form als Dissertation, Diplom- oder ähnliche Prüfungsarbeit eingereicht und ist als Ganzes auch noch nicht veröffentlicht.

Berlin, 8. Juli 2011