

Extended Formulations in Combinatorial Optimization

Michele Conforti *
Università di Padova, conforti@math.unipd.it

Gérard Cornuéjols †
Carnegie Mellon University
gc0v@andrew.cmu.edu

Giacomo Zambelli
London School of Economics and Political Science
g.zambelli@lse.ac.uk

October 2012

Abstract

This survey is concerned with the size of perfect formulations for combinatorial optimization problems. By “perfect formulation”, we mean a system of linear inequalities that describes the convex hull of feasible solutions, viewed as vectors. Natural perfect formulations often have a number of inequalities that is exponential in the size of the data needed to describe the problem. Here we are particularly interested in situations where the addition of a polynomial number of extra variables allows a formulation with a polynomial number of inequalities. Such formulations are called “compact extended formulations”. We survey various tools for deriving and studying extended formulations, such as Fourier’s procedure for projection, Minkowski-Weyl’s theorem, Balas’ theorem for the union of polyhedra, Yannakakis’ theorem on the size of an extended formulation, dynamic programming, and variable discretization. For each tool that we introduce, we present one or several examples of how this tool is applied. In particular, we present compact extended formulations for several graph problems involving cuts, trees, cycles and matchings, and for the mixing set, and we present the proof of Fiorini, Massar, Pokutta, Tiwary and de Wolf of an exponential lower bound for the cut polytope. We also present Bienstock’s approximate compact extended formulation for the knapsack problem, Goemans’ result on the size of an extended formulation for the permutahedron, and the Faenza-Kaibel extended formulation for orbitopes.

An earlier version of this survey appeared in *4OR - Quarterly Journal of Operations Research* 8 (2010) 1-48.

*Supported by the Progetto di Eccellenza 2008-2009 of the Fondazione Cassa di Risparmio di Padova e Rovigo

†Supported by NSF grant CMMI0653419, ONR grant N00014-03-1-0188 and ANR grant BLAN06-1-138894.

1 Introduction

Many combinatorial optimization problems can be written in the following form: Optimize a linear function of a vector x of feasible solutions, where the feasible set is given by linear constraints (inequalities and/or equalities) and the restriction that some (or all) components of the vector x are integral. A “perfect” formulation is one where the linear constraints describe the convex hull of feasible solutions. The interest in perfect formulations comes from the fact that the corresponding combinatorial optimization problems can then be solved as linear programs. The number of linear constraints needed in a perfect formulation is often exponential in the size of the data used to describe the problem. In this survey, we study situations where the addition of a polynomial number of new variables allows a formulation with polynomially many linear constraints. When this is possible, we say that the problem has a “compact extended formulation”. Such formulations are important in integer programming and combinatorial optimization, both theory and computations. In this paper, we survey various tools for deriving and studying extended formulations. For each tool that we introduce, we present one or several examples of how this tool is applied.

Projection is the most basic tool for relating extended formulations to formulations in the original space of variables. There are two classical methods for projecting, one uses the notion of projection cone and the other is Fourier’s procedure. We introduce these concepts in Sections 2, and show how Fourier’s method can be used to derive Barahona’s compact extended formulation of the cycle relaxation of the cut polytope. In Section 4 we illustrate two ways in which the projection cone can be used: to obtain the perfect formulation of a polytope in the original space once an extended formulation is known, which we illustrate with Balas-Pulleyblank’s formulation of the perfectly matchable subgraphs of a graph; or to prove the correctness of an extended formulation if the perfect formulation in the original space is known. We illustrate this for several graph problems (*st*-cut dominant, arborescences, trees, cuts, cycle cone).

In Section 3 we present Minkowski-Weyl’s theorem and apply it to derive a compact extended formulation for the mixing set.

In Section 5 we present another very useful tool to construct extended formulations: Balas’ theorem for the union of polyhedra. We apply it to several problems: All even subsets, cut dominant, an approximate compact extended formulation for the knapsack problem, and the continuous mixing set.

A striking theorem of Yannakakis, presented in Section 6, characterizes the size of the smallest extended formulation for a polytope. Yannakakis’ result is the main component in the recent proof of Fiorini et al. [27] that the cut polytope does not admit a compact extended formulation; this is the first and most crucial step in showing that the Travelling Salesman polytope does not admit a compact extended formulation. We give several other applications of Yannakakis’ theorem to extended formulations for spanning trees, matchings, stable sets, and the permutahedron.

The next tool that we present, in Section 7, is dynamic programming. We show how it implies extended formulations for the knapsack problem, stable sets in distance claw-free graphs, and packing and partitioning orbitopes.

Finally, in Section 8 we discuss variable discretization, a tool that has proved extremely useful when dealing with mixed integer linear sets. This approach has been particularly

successful in dealing with problems arising in lot-sizing.

We now give a formal presentation of the results. Given a set $S \subset \mathbb{R}^n$ and a linear subspace L of \mathbb{R}^n , the *orthogonal projection of S onto L* is the set of all points $u \in L$ such that there exists a vector $v \in \mathbb{R}^n$ orthogonal to L such that $u + v \in S$.

We are interested in orthogonal projections onto the space $L = \{x \in \mathbb{R}^n : x_i = 0, i \in N \setminus M\}$ where $N = \{1, \dots, n\}$ and M is a subset of N . We denote the set $\{x \in \mathbb{R}^M : \exists z \in \mathbb{R}^{N \setminus M} \text{ s.t. } \begin{pmatrix} x \\ z \end{pmatrix} \in S\}$ by $\text{proj}_x(S)$.

Our interest comes from the fact that the two programs

$$\max\{f(x) : x \in \text{proj}_x(S)\} \text{ and } \max\{f(x) + \mathbf{0}z : \begin{pmatrix} x \\ z \end{pmatrix} \in S\}$$

are equivalent. However, sometimes the set S is easier to describe than its projection, so looking at the second problem is easier. In this survey, we study sets of the type:

$$X = \{x \in \mathbb{R}^N : A'x \leq b', x_i \text{ integer}, i \in I \subseteq N\} \quad (1)$$

that arise in combinatorial optimization and integer programming, pure (when $I = N$) or mixed (when I is a proper subset of the set of variable indices).

Let $\text{conv}(X)$ denote the convex hull of the vectors in X . Namely, $\text{conv}(X) = \{\lambda_1 x_1 + \dots + \lambda_k x_k : k \geq 1, x_1, \dots, x_k \in X, \lambda_1 + \dots + \lambda_k = 1, \lambda_1, \dots, \lambda_k \geq 0\}$. A theorem of Meyer [47] states that $\text{conv}(X)$ is a rational polyhedron whenever A', b' have rational entries. We aim to find an external description of the polyhedron $\text{conv}(X)$ as the projection of a polyhedron in a higher dimensional space.

That is, we wish to find a polyhedron $Q = \{\begin{pmatrix} x \\ z \end{pmatrix} \in \mathbb{R}^n \times \mathbb{R}^p : Ax + Bz \leq b\}$ such that:

- $\text{proj}_x(Q) = \text{conv}(X)$. When this happens, we say that the system $Ax + Bz \leq b$ that defines Q provides an *extended formulation* for the set X .
- The linear program $\max\{hx + gz : Ax + Bz \leq b\}$ is easy to solve.
This is certainly the case when the size of the system $A'x \leq b'$ that defines X is small (e.g., a polynomial number of constraints) and the size of the matrix $(A|B|b)$ is polynomially bounded in the size of the matrix $(A'|b')$. When this happens, we say that the system $Ax + Bz \leq b$ provides a *compact* extended formulation for X .
Another case is when the separation problem over $\{(x, z) : Ax + Bz \leq b\}$ can be solved in polynomial time.

2 Projections

2.1 The projection cone

Given a polyhedron $Q = \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^p : Ax + Bz \leq b\}$, consider a valid inequality $ax \leq b$ for Q . It follows from the definition of projection that $ax \leq b$ is also valid for $\text{proj}_x(Q)$. Since, for every vector u satisfying $u \geq 0, uB = 0$, the inequality $uAx \leq ub$ is

valid for Q , this observation shows that $uAx \leq ub$ is also valid for $\text{proj}_x(Q)$. The next theorem states that the converse also holds. We define the *projection cone* of a polyhedron $Q = \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^p : Ax + Bz \leq b\}$ as

$$C_Q = \{u \in \mathbb{R}^m : uB = 0, u \geq 0\}.$$

Theorem 2.1. *Given a polyhedron $Q = \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^p : Ax + Bz \leq b\}$, its projection onto the x -space is $\text{proj}_x(Q) = \{x \in \mathbb{R}^n : uAx \leq ub \text{ for all } u \in C_Q\}$.*

Proof. Given $\bar{x} \in \mathbb{R}^n$, note that \bar{x} is in $\text{proj}_x(Q)$ if and only if the polyhedron $\{z \in \mathbb{R}^p : Bz \leq b - A\bar{x}\}$ is nonempty. In other words \bar{x} does not belong to $\text{proj}_x(Q)$ if and only if the system $Bz \leq b - A\bar{x}$ is infeasible. By Farkas' Lemma, the system $Bz \leq b - A\bar{x}$ is infeasible if and only if there exists a vector $u \geq 0$ such that $uB = 0$ and $uA\bar{x} > ub$. The inequality $uAx \leq ub$ is valid for $\text{proj}_x(Q)$ and violated by \bar{x} . \square

This theorem has a number of equivalent variants depending on the form in which the polyhedron Q is given. As an example, we state the result when Q has equality constraints and the variables to be projected out are nonnegative.

Corollary 2.2. *Given a polyhedron $Q = \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^p : Ax + Bz = b, z \geq 0\}$, its projection is $\text{proj}_x(Q) = \{x \in \mathbb{R}^n : uAx \leq ub \text{ for all } u \text{ s.t. } uB \geq 0\}$.*

2.2 Fourier's method for computing projections of polyhedra

Next we describe Fourier's elimination procedure to compute projections of polyhedra [29]. It performs row operations on a system of linear inequalities to eliminate one variable at a time, in the same spirit as Gauss' method to solve systems of linear equations. Here we describe an iteration of Fourier's method.

Given a system $Ax + cz \leq b$, (unknowns $x \in \mathbb{R}^n, z \in \mathbb{R}$) with inequalities indexed by M , define

$$\begin{aligned} I^+ &= \{i \in M : c_i > 0\} \\ I^- &= \{i \in M : c_i < 0\} \\ I^0 &= \{i \in M : c_i = 0\}. \end{aligned}$$

Multiplying the rows by appropriate positive numbers we assume that the entries of c are $0, \pm 1$. The system $Ax + cz \leq b$ can be rewritten as:

$$\begin{aligned} \sum_{j=1}^n a_{ij}x_j + z &\leq b_i, \quad i \in I^+ \\ \sum_{j=1}^n a_{ij}x_j - z &\leq b_i, \quad i \in I^- \\ \sum_{j=1}^n a_{ij}x_j &\leq b_i, \quad i \in I^0. \end{aligned}$$

For every pair of indices $i \in I^+$ and $k \in I^-$, construct the inequality:

$$\sum_{j=1}^n a_{ij}x_j + \sum_{j=1}^n a_{kj}x_j \leq b_i + b_k. \quad (2)$$

Let $A'x \leq b'$ be the system consisting of the inequalities in I^0 and the $|I^+| \times |I^-|$ inequalities (2).

Theorem 2.3. Let $Q = \{(x, z) \in \mathbb{R}^n \times \mathbb{R} : Ax + cz \leq b\}$ and $P = \{x \in \mathbb{R}^n : A'x \leq b'\}$. Then $P = \text{proj}_x(Q)$.

Proof. By construction, all the inequalities in the system $A'x \leq b'$ are valid for both Q and $\text{proj}_x(Q)$, therefore $\text{proj}_x(Q) \subseteq P$.

To prove that $P \subseteq \text{proj}_x(Q)$, we show that if x' satisfies $A'x' \leq b'$, then there exists $z' \in \mathbb{R}$ such that $(x', z') \in Q$. That is, the system $cz \leq b - Ax'$ is feasible. Again we may assume that the entries of c are $0, \pm 1$.

Let $u = \min_{i \in I^+} \{b_i - \sum_{j=1}^n a_{ij}x'_j\}$ (where $u = +\infty$ if $I^+ = \emptyset$), and $l = \max_{i \in I^-} \{\sum_{j=1}^n a_{ij}x'_j - b_i\}$ (where $l = -\infty$ if $I^- = \emptyset$). Since $A'x' \leq b'$, the construction of $A'x' \leq b'$ implies $l \leq u$, therefore, given z' such that $l \leq z' \leq u$, $(x', z') \in Q$. \square

Remark 2.4. At each iteration, Fourier's method generates a finite number of inequalities. Therefore it provides an algorithmic proof that the projection of a polyhedron is a polyhedron.

Remark 2.5. At each iteration, Fourier's method removes $|I^+| + |I^-|$ inequalities and adds $|I^+| \times |I^-|$ inequalities, hence the number of inequalities has an order which may be squared at each iteration. So the number of inequalities produced by eliminating p variables may be exponential in p . In general, not all the inequalities produced by the procedure are necessary to define the projected polyhedron.

Fourier's method can be used to solve systems of inequalities: Project away all the variables until a polyhedron with one variable is left: Notice that when $n = 1$, the system $A'x \leq b'$ after scaling consists of inequalities of the type $x_1 \leq u_i$ and $x_1 \geq l_j$ (one of the two sets may be empty). Let $u_{\min} = \min\{u_i\}$ and $l_{\max} = \max\{l_j\}$. After removing redundant inequalities, we get the equivalent system: $l_{\max} \leq x_1 \leq u_{\min}$. The above procedure constructs the inequality $0x_1 \leq u_{\min} - l_{\max}$ which is feasible if and only if $u_{\min} - l_{\max} \geq 0$.

2.3 An application of Fourier's method: The cycle relaxation of the cut polytope

Let $G = (V, E)$ be an undirected graph, possibly with loops and parallel edges. For any set of nodes $S \subseteq V$, let $\delta(S) = \{uv \in E : u \in S, v \notin S\}$. A *cut* of G is a set of the form $\delta(S)$, for some $S \subseteq V$. Note that, according to this definition, the empty set is a cut. The *cut polytope* of G $P^{\text{cut}}(G)$ is the convex hull of the incidence vectors of the cuts of G .

Since the cuts of G are the subsets of E having even intersection with every cycle of G , a vector in $\{0, 1\}^E$ is the incidence vector of a cut of G if and only if it satisfies the following linear constraints:

$$\begin{aligned} x(F) - x(C \setminus F) &\leq |F| - 1, & C \in \mathcal{C}(G), F \subseteq C, |F| \text{ odd} \\ 0 &\leq x_e \leq 1, & e \in E \end{aligned} \tag{3}$$

where $\mathcal{C}(G)$ denotes the family of cycles of G . Above, and throughout this paper, we use the notation $x(D)$ to represent $\sum_{e \in D} x_e$.

Let $R(G)$ denote the polytope in \mathbb{R}^E defined by (3). In general, the cut polytope $P^{\text{cut}}(G)$ is strictly contained in $R(G)$. However the two polytopes coincide when G is planar (this will be proven in Section 6.6) and more generally when G is not contractible to K_5 [8]. But there are other graphs G for which $P^{\text{cut}}(G)$ coincides with $R(G)$. An important result of Guenin [34] characterizes these graphs.

Theorem 2.6. (Barahona [8]) $R(G)$ has a compact extended formulation.

We give a proof of this result.

Lemma 2.7. Let $G = (V, E)$ and $G' = (V, E')$ be two graphs on the same nodeset such that $E \subset E'$. Then $R(G)$ is the projection of $R(G')$ onto \mathbb{R}^E .

Proof. It is sufficient to prove the statement when $E' = E \cup \{e'\}$. We use Fourier's method to project out the variable $x_{e'}$. The inequalities defining $R(G')$ in which $x_{e'}$ appears with nonzero coefficient are

- (a) $x_{e'} \leq 1$;
- (b) $x_{e'} + x(F \setminus \{e'\}) - x(C \setminus F) \leq |F| - 1$ where $C \in \mathcal{C}(G')$, $e' \in F \subseteq C$, $|F|$ odd;
- (c) $-x_{e'} \leq 0$;
- (d) $-x_{e'} + x(F) - x(C \setminus (F \cup \{e'\})) \leq |F| - 1$ where $C \in \mathcal{C}(G')$, $e' \in C \setminus F$, $|F|$ odd;

Fourier's method sums each inequality of type (a) or (b) with an inequality of type (c) or (d) to eliminate $x_{e'}$. It is easy to see that the only inequalities that are not redundant are obtained by combining an inequality of type (b) with an inequality of type (d). Let C' and C'' be cycles of G' containing e' , let $F' \subseteq C'$, $|F'|$ odd, such that $e' \in F'$, and $F'' \subseteq C''$, $|F''|$ odd, such that $e' \notin F''$. Let C_1, \dots, C_k be disjoint cycles whose union is $C' \Delta C''$. Let $F = F'' \Delta (F' \setminus \{e'\})$. Note that $|F|$ is odd. The inequality obtained by summing the two inequalities determined by C', F' and C'', F'' is implied by the following inequalities, valid for $R(G)$:

$$\begin{aligned} x(C_i \cap F) - x(C_i \setminus F) &\leq |F \cap C_i| - 1 \text{ if } |F \cap C_i| \text{ is odd} \\ 0 \leq x_e &\leq 1 \quad e \in (C' \cup C'') \setminus \{e'\}. \end{aligned}$$

Therefore the only irredundant inequalities produced by Fourier's method are the ones in (3). \square

Lemma 2.8. Let $C \in \mathcal{C}(G)$ be a cycle with a chord $e = uv$ and $F \subseteq C$ be a set of odd cardinality. Then the inequality $x(F) - x(C \setminus F) \leq |F| - 1$ is implied by the other inequalities in (3).

Proof. Let P_1 and P_2 be the two distinct paths in C between u and v , and let C_1, C_2 be the cycles defined by $P_1 \cup \{e\}$ and $P_2 \cup \{e\}$, respectively. By symmetry, we may assume $|F \cap P_1|$ is odd and $|F \cap P_2|$ is even. Let $F_1 = F \cap P_1$ and $F_2 = (F \cap P_2) \cup \{e\}$. Then $x(F) - x(C \setminus F) \leq |F| - 1$ is the sum of the two inequalities $x(F_1) - x(C_1 \setminus F_1) \leq |F_1| - 1$ and $x(F_2) - x(C_2 \setminus F_2) \leq |F_2| - 1$. \square

Let $G' = (V, E')$ be the graph with the same nodeset as G such that $E' = E \cup \{uv : uv \notin E, u \neq v\}$. By Lemma 2.7, $R(G)$ is the projection onto \mathbb{R}^E of $R(G')$. By Lemma 2.8, $R(G')$

is defined by the following system of inequalities

$$\left. \begin{array}{r} x_{e_1} + x_{e_2} + x_{e_3} \leq 2 \\ x_{e_1} - x_{e_2} - x_{e_3} \leq 0 \\ -x_{e_1} + x_{e_2} - x_{e_3} \leq 0 \\ -x_{e_1} - x_{e_2} + x_{e_3} \leq 0 \end{array} \right\} \begin{array}{l} \{e_1, e_2, e_3\} \text{ triangle of } G' \\ \\ \\ \end{array}$$

$$\begin{array}{l} x_e = 0 \\ x_{e_1} = x_{e_2} \end{array} \begin{array}{l} e \text{ loop of } G' \\ e_1, e_2 \text{ parallel edges of } G'. \end{array}$$

Indeed, the only chordless cycles of G' are loops, parallel edges and triangles. One can easily see that the system (3) implies $x_e = 0$ for every loop e of G and $x_{e_1} = x_{e_2}$ for every pair of parallel edges e_1, e_2 of G . Furthermore, since every nonloop edge $e \in E'$ is contained in a triangle of G' , the inequalities $0 \leq x_e \leq 1$ are easily seen to be redundant for the above system.

Since $x_{e_1} = x_{e_2}$ for each pair of parallel edges, we only need to consider the $O(|V|^3)$ inequalities relative to triangles. Thus the extended formulation has $O(|V|^2)$ variables and $O(|V|^3)$ constraints.

3 The Theorem of Minkowski-Weyl

We first present the Minkowski-Weyl theorem for cones.

3.1 Minkowski-Weyl for Cones

Theorem 3.1. *For a set $C \subseteq \mathbb{R}^n$, the following two conditions are equivalent:*

1. *There exists a matrix A such that $C = \{x \in \mathbb{R}^n : Ax \leq 0\}$.*
2. *There exists a matrix R such that $C = \{x \in \mathbb{R}^n : x = R\mu \text{ for some } \mu \geq 0\}$.*

The next lemma shows that, in order to prove the equivalence between Conditions 1 and 2, it suffices to prove one of the two directions. If matrix A satisfies Condition 1 and matrix R satisfies Condition 2, then the number of columns of A equals the number of rows of R . Furthermore $AR\mu \leq 0$ for every $\mu \geq 0$, hence all entries of AR are nonpositive.

Lemma 3.2. $\{x : Ax \leq 0\} = \{x : x = R\mu \text{ for some } \mu \geq 0\}$ if and only if $\{y : R^T y \leq 0\} = \{y : y = A^T \nu \text{ for some } \nu \geq 0\}$.

Proof. We only need to show that the first equality implies the second. By Corollary 2.2, the set $\{x : x = R\mu \text{ for some } \mu \geq 0\}$ coincides with the set of vectors x for which $x^T y \leq 0$ holds for every y satisfying $R^T y \leq 0$. Therefore

$$\{x : Ax \leq 0\} = \{x : x^T y \leq 0 \text{ for every } y \text{ such that } R^T y \leq 0\}. \quad (4)$$

We need to show that $\{y : R^T y \leq 0\} = \{y : y = A^T \nu \text{ for some } \nu \geq 0\}$.

We first show $\{y : R^T y \leq 0\} \subseteq \{y : y = A^T \nu \text{ for some } \nu \geq 0\}$. Let \bar{y} such that $R^T \bar{y} \leq 0$. By (4), inequality $\bar{y}^T x \leq 0$ is valid for $Ax \leq 0$. By Farkas' Lemma, it follows that there exists a vector ν such that $\bar{y} = A^T \nu$ and $\nu \geq 0$.

We show the reverse inclusion. Given \bar{y} such that $\bar{y} = A^T \nu$ for some $\nu \geq 0$, $R^T \bar{y} = R^T A^T \nu \leq 0$, where the inequality follows from the property that AR has nonpositive entries. \square

Proof of Theorem 3.1

By Lemma 3.2, it suffices to prove that Condition 2 of the theorem implies 1. Consider the cone $C_\mu = \{(x, \mu) \in \mathbb{R}^n \times \mathbb{R}^q : x = R\mu, \mu \geq 0\}$. Then the implication states that $\text{proj}_x(C_\mu) = \{x \in \mathbb{R}^n : Ax \leq 0\}$. By Remark 2.4, the procedure of Fourier applied to the system $x = R\mu, \mu \geq 0$ that defines C_μ computes a matrix A such that $\text{proj}_x(C_\mu) = \{x \in \mathbb{R}^n : Ax \leq 0\}$. \square

3.2 Minkowski-Weyl for Polyhedra

Given a set $S \subseteq \mathbb{R}^n$, let $\text{cone}(S) = \{\mu_1 x_1 + \dots + \mu_k x_k : k \geq 1, x_1, \dots, x_k \in S, \mu_1, \dots, \mu_k \geq 0\}$ be the cone generated by the elements in S . We adopt the convention that $\text{cone}(\emptyset) = \{0\}$ and $\text{conv}(\emptyset) = \emptyset$. The *Minkowski sum* $A + B$ of sets $A, B \subseteq \mathbb{R}^n$ is the set $\{a + b : a \in A, b \in B\}$. If A or B is empty, $A + B$ is also empty.

Theorem 3.3. *For a subset Q of \mathbb{R}^n , the following two conditions are equivalent:*

1. Q is a polyhedron, i.e., there is a matrix A and a vector b such that $Q = \{x \in \mathbb{R}^n : Ax \leq b\}$.
2. There exist $v^1, \dots, v^p \in \mathbb{R}^n, r^1, \dots, r^q \in \mathbb{R}^n$ such that $Q = \text{conv}\{v^1, \dots, v^p\} + \text{cone}\{r^1, \dots, r^q\}$.

Proof. Given cone $H_Q := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : Ax - by \leq 0, y \geq 0\}$, then $Q = \{x : (x, 1) \in H_Q\}$. By Theorem 3.1, the cone H_Q is finitely generated. Since $y \geq 0$ for every vector $(x, y) \in H_Q$, we can assume that $y = 0$ or 1 for all the rays that generate H_Q . That is,

$$H_Q = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^n \times \mathbb{R} : \begin{pmatrix} x \\ y \end{pmatrix} \in \text{cone} \left\{ \begin{pmatrix} v^1 \\ 1 \end{pmatrix}, \dots, \begin{pmatrix} v^k \\ 1 \end{pmatrix}, \begin{pmatrix} r^1 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} r^q \\ 0 \end{pmatrix} \right\} \right\}.$$

Since $Q = \{x : (x, 1) \in H_Q\}$, this shows that

$$Q = \{x \in \mathbb{R}^n : x = v + r \text{ for some } v \in \text{conv}\{v^1, \dots, v^k\} \text{ and } r \in \text{cone}\{r^1, \dots, r^q\}\}.$$

This equivalence also proves the converse statement. \square

We say that $Ax \leq b$ is an *external description* of Q , and $v^1, \dots, v^p, r^1, \dots, r^q$ is an *internal description* of Q . Note that an internal description of a polyhedron Q yields an extended formulation in the variables $(x, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q$:

$$Q = \text{proj}_x \left\{ (x, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q : \begin{array}{l} x = \lambda_1 v^1 + \dots + \lambda_p v^p + \mu_1 r^1 + \dots + \mu_q r^q \\ \lambda_1 + \dots + \lambda_p = 1, \lambda \geq 0, \mu \geq 0 \end{array} \right\}.$$

The cone $\{r \in \mathbb{R}^n : r = \mu_1 r^1 + \dots + \mu_q r^q, \mu \geq 0\} = \{r \in \mathbb{R}^n : Ar \leq 0\}$ is the *recession cone* of Q .

Remark 3.4. *It follows from polyhedral theory (see, e.g., [14]) that:*

- A nonredundant external description of a polyhedron $Q = \text{conv}(V) + \text{cone}(R)$ is of the form: $Q = \{x \in \mathbb{R}^n : A^-x = b^-, A^<x \leq b^<\}$ where $A^-x = b^-$ is a system of linearly independent equations that defines the smallest affine subspace of \mathbb{R}^n that contains Q and every inequality in $A^<x \leq b^<$ defines a distinct facet of Q .

- If $Q = \{x \in \mathbb{R}^n : Ax \leq b\}$ is pointed, then Q has a unique minimal internal description $Q = \text{conv}(V) + \text{cone}(R)$ where V is the set of vertices of Q and R is the set of extreme rays. We recall that Q is pointed if and only if A has rank n .
- Given cone $C = \{x \in \mathbb{R}^n : Ax \leq 0\}$, let r^1, \dots, r^k be a set of generators of the linear space $L = \{x \in \mathbb{R}^n : Ax = 0\}$. A minimal internal description of C is obtained by taking the rays r^1, \dots, r^k , the rays $-r^1 - \dots - r^k$ and one ray in every face F containing L that is minimal with this property.

Theorem 3.3 can be used to prove the following theorem of Meyer [47] (see also [14]).

Theorem 3.5. *Given rational matrices A , G and a rational vector b , let $P = \{(x, y) : Ax + Gy \leq b\}$ and $S = \{(x, y) \in P : x \text{ integral}\}$. Then $\text{conv}(S)$ is a polyhedron and, if $S \neq \emptyset$, the recession cone of $\text{conv}(S)$ is equal to the recession cone of P .*

For a mixed integer set X of the form (1), the internal description of $\text{conv}(X)$ typically involves an exponential number of vertices and extreme rays. Column generation methods solve mixed integer programs by working with a compact list that involves a subset of vertices and extreme rays. The list is dynamically updated by solving an optimization problem that produces a profitable new vertex or ray.

For some selected mixed integer sets X , the internal description of $\text{conv}(X)$ provides a compact extended formulation. We present such an example.

3.3 An application: The mixing set

The *mixing set* MIX is defined as follows

$$MIX = \{(x_0, \dots, x_n) \in \mathbb{R} \times \mathbb{Z}^n : x_0 + x_t \geq b_t, t = 1, \dots, n, x_0 \geq 0\}.$$

The mixing set was introduced by Günlük and Pochet [35] and it can be traced back to work of Pochet and Wolsey [50] on a lot-sizing problem. A minimal external description of the polyhedron $P^{\text{mix}} = \text{conv}(MIX)$ was characterized by Günlük and Pochet [35]. Such a description needs exponentially many inequalities whose coefficients may be large (the “mixing inequalities” of Pochet and Wolsey [50]). However Pochet and Wolsey give a compact extended formulation for P^{mix} .

Let $f_t = b_t - \lfloor b_t \rfloor$. For simplicity of exposition, we assume here that $0 < f_1 < \dots < f_n$. Define $f_0 = 0$.

The vertices and extreme rays of P^{mix} are easy to characterize.

For $t = 0, \dots, n$, let $v^t \in \mathbb{R} \times \mathbb{R}^n$ be defined by

$$v_i^t = \begin{cases} f_t & \text{if } i = 0 \\ \lfloor b_i \rfloor & \text{if } 1 \leq i \leq t \\ \lfloor b_i \rfloor & \text{if } t + 1 \leq i \leq n. \end{cases}$$

One can verify that $v^t \in MIX$.

Let $r^0 \in \mathbb{R} \times \mathbb{R}^n$ be defined by $r_0^0 = 1$, $r_i^0 = -1$, $i = 1, \dots, n$. For $t = 1, \dots, n$, let $r^t \in \mathbb{R} \times \mathbb{R}^n$ be defined by $r_t^t = 1$, $r_i^t = 0$ for $i \neq t$, $0 \leq i \leq n$. The vectors r^0, \dots, r^n are rays of P^{mix} .

Theorem 3.6. P^{mix} is the projection onto the space of x -variables of the polyhedron

$$\{(x, \lambda, \mu) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} : x = \sum_{t=0}^n \lambda_t v^t + \sum_{t=0}^n \mu_t r^t, \sum_{t=0}^n \lambda_t = 1, \lambda \geq 0, \mu \geq 0\}.$$

Proof. Note that, since the system of inequalities defining MIX has full column rank, the polyhedron P^{mix} is pointed. By Remark 3.4 we only need to show that v^0, \dots, v^n are all the vertices of P^{mix} and r^0, \dots, r^n are all its extreme rays. By Theorem 3.5, the recession cone of P^{mix} is

$$\{x \in \mathbb{R}^{n+1} : x_0 + x_t \geq 0, t = 1, \dots, n; x_0 \geq 0\}.$$

The extreme rays of this cone are the rays of the cone satisfying n linearly independent inequalities at equality. Since the system defining the recession cone has $n + 1$ inequalities, and they are linearly independent, the $n + 1$ extreme rays are easily seen to be the vectors r^0, \dots, r^n .

Claim. Let \bar{x} be a vertex of P^{mix} . Either $\bar{x}_0 = 0$ or $\bar{x}_0 = f_t$ for some $t = 1, \dots, n$.

We first show $\bar{x}_0 < 1$. Suppose not. Then P^{mix} contains both points $\bar{x} + r^0$ and $\bar{x} - r^0$. Since $\bar{x} = \frac{1}{2}((\bar{x} + r^0) + (\bar{x} - r^0))$, \bar{x} is not a vertex, a contradiction. Suppose now that $\bar{x}_0 \neq 0$ and $\bar{x}_0 \neq f_t, 1 \leq t \leq n$. Since $\bar{x}_1, \dots, \bar{x}_n$ are integer, \bar{x} does not satisfy at equality any of the inequalities $x_0 + x_t \geq b_t, 1 \leq t \leq n, x_0 \geq 0$. Therefore $(\bar{x}_0 \pm \epsilon, \bar{x}_1, \dots, \bar{x}_n) \in P^{\text{mix}}$ for $\epsilon > 0$ sufficiently small, thus \bar{x} is not a vertex. This completes the proof of the claim.

A similar argument shows that $\bar{x}_t = \lceil b_t - \bar{x}_0 \rceil, t = 1, \dots, n$. Hence $\bar{x} = v^t$ for some $t \in \{0, \dots, n\}$. \square

Note that the above extended formulation has $3n + 3$ variables and $3n + 4$ constraints. It is therefore a compact formulation for P^{mix} .

4 Finiteness of the projection cone

Given a polyhedron $Q = \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^p : Ax + Bz \leq b\}$, we recall from Section 2.1 that its projection cone is $C_Q = \{u \in \mathbb{R}^m : uB = 0, u \geq 0\}$. Since C_Q is contained in the nonnegative orthant, C_Q is pointed. By Theorem 3.1 and Remark 3.4,

$$C_Q = \{u \in \mathbb{R}^m : u = r^1 \mu_1 + \dots + r^q \mu_q, \mu \geq 0\}$$

where r^1, \dots, r^q are the extreme rays of C_Q . Theorem 2.1 implies the following.

Theorem 4.1. Given a polyhedron $Q = \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^p : Ax + Bz \leq b\}$, let r^1, \dots, r^q be the extreme rays of C_Q . Then $\text{proj}_x(Q) = \{x \in \mathbb{R}^n : r^t Ax \leq r^t b, 1 \leq t \leq q\}$. Therefore $\text{proj}_x(Q)$ is defined by a finite number of inequalities, i.e., it is a polyhedron.

Remark 4.2. Theorem 4.1 shows that the inequalities $r^t Ax \leq r^t b, 1 \leq t \leq q$ are sufficient to describe $\text{proj}_x(Q)$. However, some of these inequalities might be redundant even though r^1, \dots, r^q are extreme rays.

Avis and Fukuda [3] show that one can enumerate all feasible bases of a linear program in standard form using Bland's pivoting rule. By normalizing the sum of the components of vectors in C_Q , one can generate all extreme rays using this method.

4.1 Perfectly matchable subgraphs of a bipartite graph

The results presented in this section are due to Balas and Pulleyblank [6, 7]. Let $G = (V, E)$ be an undirected graph. The *perfectly matchable subgraph polytope* of G , denoted by $P^{\text{matchable}}$, is the convex hull of the incidence vectors of all subsets V' of V such that the graph $G[V']$ induced by V' has a perfect matching. ($G[V'] = (V', E')$ denotes the graph where $E' \subseteq E$ is the subset of edges that have both endnodes in V' .)

A vector $(x, z), x \in \mathbb{R}^V, z \in \mathbb{R}^E$ is the *extended incidence vector of a perfectly matchable subgraph* of G induced by V' if x is the incidence vector of V' and z is the incidence vector of a matching that saturates the nodes in V' and leaves exposed the nodes in $V \setminus V'$. Note that $(x, z) \in \mathbb{R}^V \times \mathbb{R}^E$ is the extended incidence vector of a perfectly matchable subgraph of G if and only if it satisfies

$$\begin{aligned} \sum_{e \in \delta(v)} z_e &= x_v, & v \in V \\ 0 \leq x_v &\leq 1, & v \in V \\ z_e &\geq 0, \text{ integer}, & e \in E. \end{aligned} \tag{5}$$

Let $Q^{\text{matchable}}$ be the convex hull of extended incidence vectors of perfectly matchable subgraphs of G .

Theorem 4.3. *Let $G = (V, E)$ be a bipartite graph. Then $Q^{\text{matchable}}$ is the polytope*

$$\left\{ (x, z) \in \mathbb{R}^V \times \mathbb{R}^E : \begin{array}{ll} \sum_{e \in \delta(v)} z_e = x_v, & v \in V \\ 0 \leq x_v \leq 1, & v \in V \\ z_e \geq 0, & e \in E \end{array} \right\}. \tag{6}$$

Proof. It is straightforward to check that any 0,1-vector (x, z) in the polytope in (6) is an extended incidence vector of a perfectly matchable subgraph of G . On the other hand, since G is a bipartite graph, the constraint matrix that defines the above system of inequalities is totally unimodular. Since the right-hand side is integral, it follows from a classical theorem of Hoffman and Kruskal [37] that the polytope defined in (6) has 0,1 vertices, which by the above argument are the incidence vectors of the members of $Q^{\text{matchable}}$. \square

Given $U \subseteq V$, let $N(U)$ be the subset of nodes in $V \setminus U$ having at least one neighbor in U .

Theorem 4.4. *Let $G = (V, E)$ be a bipartite graph with bipartition V_1, V_2 . The following is an external description of $P^{\text{matchable}}$*

$$\left\{ x \in \mathbb{R}^V : \begin{array}{ll} \sum_{u \in V_1} x_u = \sum_{v \in V_2} x_v, & U \subseteq V_1 \\ \sum_{u \in U} x_u \leq \sum_{v \in N(U)} x_v, & u \in V \\ 0 \leq x_u \leq 1, & \end{array} \right\}. \tag{7}$$

Proof. Since $P^{\text{matchable}} = \text{proj}_x(Q^{\text{matchable}})$ and by Theorem 4.3 the polytope $Q^{\text{matchable}}$ has the external description (6), the relevant projection cone is

$$C = \{y \in \mathbb{R}^E : y_u + y_v \geq 0, uv \in E\}.$$

We may assume without loss of generality that G is connected. Since the edge-node incidence matrix of a connected bipartite graph has rank $|V| - 1$ (this is well known and easy to check

directly) and the lineality space L of C is the space $\{y \in \mathbb{R}^E : y_u + y_v = 0, uv \in E\}$, then L has dimension 1. Thus L is generated by the vector \bar{y} defined by $\bar{y}_u = -1, u \in V_1, \bar{y}_v = 1, v \in V_2$. By Remark 3.4, an internal description of C contains rays \bar{y} and $-\bar{y}$ and both produce the equation $\sum_{u \in V_1} x_u = \sum_{v \in V_2} x_v$.

Let F be a face of C that is minimal with the property that $F \supset L$, and let $y^* \in F \setminus L$. Since the lineality space of C is generated by \bar{y} , then $F = \{\lambda y^* + \mu \bar{y} : \lambda, \mu \in \mathbb{R}, \lambda \geq 0\}$, hence we may choose y^* so that $y_v^* \leq 0$ for all $v \in V_1$ and $y_{v^*}^* = 0$ for some $v^* \in V_1$.

Since y^* is a ray in a minimal face of C , then y^* satisfies at equality $n - 2$ linearly independent inequalities defining C . Since G is bipartite, it is known that any set $E^=$ of edges corresponding to linearly independent inequalities is acyclic. Since $|E^=| = |V| - 2$, then the graph $(V, E^=)$ is a forest with exactly two connected components, say F_1, F_2 . One can easily see that $y_u^* = -y_v^*$ for all $uv \in E(F_i), i = 1, 2$. By symmetry, we may assume $v^* \in V(F_2)$. This shows that $y_u^* = 0, u \in V(F_2)$, and we can assume $y_u^* = -1, u \in v(F_1) \cap V_1, y_v^* = 1, v \in v(F_1) \cap V_2$.

Define $U = V(F_1) \cap V_1$. Since $y^* \in C$, then for every edge $uv \in E$ with $u \in U, y_u^* + y_v^* \geq 0$, hence $y_v^* \geq 1$, which implies that $v \in V(F_1)$. This shows that $N(U) \subseteq V(F_1) \cap V_2$. Hence $y_u^* = -1, u \in U, y_v^* = 1, v \in N(U), y_w^* = 0, w \in V \setminus (U \cup N(U))$, and the corresponding inequality in the projection is $\sum_{u \in U} x_u - \sum_{v \in N(U)} x_v \leq 0$. \square

The above theorem can be extended to the case where G is non bipartite as follows. In (5), substitute the variable $s_v = 1 - x_v, v \in V$, obtaining

$$\begin{aligned} \sum_{e \in \delta(v)} z_e + s_v &= 1, & v \in V; \\ s_v &\geq 0, & v \in V; \\ z_e &\geq 0, \text{ integer}, & e \in E. \end{aligned}$$

Thus s_v is just the slack variable of the degree inequality relative to $v \in V$ in the system

$$\begin{aligned} \sum_{e \in \delta(v)} z_e &\leq 1, & v \in V; \\ z_e &\geq 0, \text{ integer}, & e \in E. \end{aligned}$$

Since the feasible solutions of the latter system are the incidence vectors of matchings in G , by the famous matching polytope theorem of Edmonds we obtain the following.

Theorem 4.5. *The polytope $Q^{\text{matchable}}$ is the set*

$$\left\{ (x, z) \in \mathbb{R}^V \times \mathbb{R}^E : \begin{array}{ll} \sum_{e \in \delta(v)} z_e = x_v, & v \in V \\ \sum_{e \in E(U)} z_e \leq \frac{|U|-1}{2}, & U \subseteq V, |U| \text{ odd} \\ 0 \leq x_v \leq 1, & v \in V \\ z_e \geq 0, & e \in E \end{array} \right\}. \quad (8)$$

Notice, however, that the above extended formulation is not compact, since there is an *odd cut inequality* for each $U \subseteq V$ of odd cardinality. Nonetheless, since odd cut inequalities can be separated in polynomial time [49], the problem of optimizing a linear function over $Q^{\text{matchable}}$, and thus over $P^{\text{matchable}}$, can be solved in polynomial time.

Balas and Pulleyblank [7] also project the above extended formulation onto the x -space by characterizing the extreme rays of the projection cone, thus giving an explicit external description of $P^{\text{matchable}}$.

4.2 The st -cut dominant

A polyhedron $Q \subseteq \mathbb{R}_+^n$ is a *dominant polyhedron* if the recession cone of Q is the nonnegative orthant \mathbb{R}_+^n . So Q is a dominant polyhedron if and only if, for every $c \in \mathbb{R}^n$, the linear program $\min\{cx : x \in Q\}$ has a finite optimum if and only if $c \geq 0$.

Given any polytope $P \subset \mathbb{R}^n$, the *dominant* of P is the polyhedron $P_+ = P + \mathbb{R}_+^n$. Note that, if $c \in \mathbb{R}^n$ is nonnegative, then $\min\{cx : x \in P\} = \min\{cx : x \in P_+\}$.

Let $D = (V, A)$ be a directed graph, and $|A| = m$, $|V| = n$. Given $S \subseteq V$, the set $\delta^+(S) = \{(u, v) \in A : u \in S, v \in V \setminus S\}$ is a *cut* of D . Given $s, t \in V$, a cut $\delta^+(S)$ is an *st -cut* if S separates s from t , that is $s \in S$, $t \notin S$. We define $\delta^-(S) = \delta^+(V \setminus S)$. Let $P_+^{\text{st-cut}}(D)$ be the convex hull of incidence vectors of all the st -cuts of D and let $P_+^{\text{st-cut}}(D)$ be its dominant.

Minimizing a linear function over $P_+^{\text{st-cut}}(D)$ is *NP-hard* [44], therefore it is unlikely that a tractable external formulation exists for $P_+^{\text{st-cut}}(D)$ (in the original space or extended). On the other hand, minimizing a nonnegative linear function over $P_+^{\text{st-cut}}(D)$ is the same as minimizing it over $P_+^{\text{st-cut}}(D)$, which is the problem of computing a minimum capacity cut, and can be done in polynomial time. The external description of $P_+^{\text{st-cut}}(D)$ in the original space is well known, and is due to Fulkerson [30] (see [55] Section 13.1).

Let \mathcal{P}_{st} be the collection of all the directed paths from s to t .

Theorem 4.6.

$$P_+^{\text{st-cut}}(D) = \left\{ x \in \mathbb{R}^A : \sum_{a \in P} x_a \geq 1, P \in \mathcal{P}_{st}; x_a \geq 0, a \in A \right\}.$$

Proof. Let $R = \{x \in \mathbb{R}^A : \sum_{a \in P} x_a \geq 1, P \in \mathcal{P}_{st}; x_a \geq 0, a \in A\}$. Clearly $P_+^{\text{st-cut}}(D) \subseteq R$ and the recession cone of R is \mathbb{R}_+^A . Thus we only need to show that every vertex of R is the incidence vector of an st -cut. To prove this, we show that given a nonnegative vector $c \in \mathbb{R}^A$, the linear program $\min\{\sum_{a \in A} c_a x_a : x \in R\}$ has an optimal solution which is the incidence vector of an st -cut. Let $\delta^+(S)$ be an st -cut such that $c(\delta^+(S))$ is smallest possible, and let $c^* = c(\delta^+(S))$. By the Max-flow Min-cut theorem, c^* is also the maximum value of an st -flow. By the path decomposition of flows (see for example [55]), there exists a set of nonnegative multipliers v_P , $P \in \mathcal{P}_{st}$ such that:

- $\sum_{P \ni a} v_P \leq c_a, a \in A;$
- $\sum_{P \in \mathcal{P}_{st}} v_P = c^*.$

Thus $(v_P)_{P \in \mathcal{P}}$ is a feasible solution for the dual of $\min\{\sum_{a \in A} c_a x_a : \sum_{a \in P} x_a \geq 1, P \in \mathcal{P}_{st}; x_a \geq 0, a \in A\}$ with value c^* . This shows that the incidence vector of $\delta^+(S)$ is an optimal solution of the above problem. \square

Clearly the previous description of $P_+^{\text{st-cut}}(D)$ is not compact, since it has an inequality for each st -path, and their number might be exponential. Next we describe a compact extended formulation. Add the arc (t, s) to A (note that this does not change the st -cut polytope).

Consider the polyhedron

$$Q^{\text{st-cut}} = \left\{ (x, z) \in \mathbb{R}^{m+n} : \begin{array}{l} z_v - z_u + x_{(u,v)} \geq 0, \quad (u, v) \in A \setminus \{(t, s)\} \\ z_s - z_t + x_{(t,s)} \geq 1, \\ x_a \geq 0, \quad a \in A \\ x_{(t,s)} = 0 \end{array} \right\}. \quad (9)$$

Theorem 4.7.

$$P_+^{\text{st-cut}}(D) = \text{proj}_x(Q^{\text{st-cut}})$$

Proof. The projection cone associated to the external description (9) is

$$C = \{f \in \mathbb{R}^A : \sum_{a \in \delta^-(v)} f_a - \sum_{a \in \delta^+(v)} f_a = 0, v \in V; f_a \geq 0, a \in A\}.$$

Therefore a vector f is in C if and only if f is a circulation of D . Since C is a pointed cone, only extreme rays of C can produce facets of $\text{proj}_x(Q^{\text{st-cut}})$. It is well known and easy to prove that any extreme ray of C is the incidence vector of some directed cycle F of D .

This shows that an irredundant inequality for $\text{proj}_x(Q^{\text{st-cut}})$ is either a nonnegativity constraint $x_a \geq 0$ or an inequality $\sum_{a \in F} x_a \geq \alpha$ for some directed cycle F of D , where $\alpha = 1$ if $(t, s) \in F$, $\alpha = 0$ otherwise.

If $\alpha = 0$ the inequality $\sum_{a \in F} x_a \geq \alpha$ is just the sum of the constraints $x_a \geq 0$, $a \in F$, and is therefore redundant. When $\alpha = 1$, let $P = F \setminus \{(t, s)\}$. Since $x_{(t,s)} = 0$, the inequality $\sum_{a \in F} x_a \geq 1$ is just $\sum_{a \in P} x_a \geq 1$. Hence $\text{proj}_x(Q^{\text{st-cut}}) = \{x \in \mathbb{R}^A : \sum_{a \in P} x_a \geq 1, P \in \mathcal{P}_{st}; x_a \geq 0, a \in A\}$, and the statement follows from Theorem 4.6. \square

Let $G = (V, E)$ be an undirected graph. An st -cut in G is a cut $\delta(S)$ such that $s \in S$, $t \notin S$. The st -cut polytope $P^{\text{st-cut}}(G)$ is the convex hull of incidence vectors of st -cuts of G . Let $P_+^{\text{st-cut}}(G)$ be its dominant.

Let $D = (V, A)$ be the digraph obtained from G by substituting every edge $e = uv$ with the pair of opposite arcs (u, v) , (v, u) . Then

$$P_+^{\text{st-cut}}(G) = \{x \in \mathbb{R}^E : \exists y \in P_+^{\text{st-cut}}(D) \text{ s.t. } x_{uv} = y_{(u,v)} + y_{(v,u)}, uv \in E\}.$$

Therefore the compact extended formulation for $P_+^{\text{st-cut}}(D)$ yields a compact extended formulation for $P_+^{\text{st-cut}}(G)$.

4.3 Arborescences and Trees

Let $D = (V, A)$ be a digraph and let $r \in V$. An r -arborescence T in D is a subset of arcs such that no arc in T enters r and, for every node $v \in V \setminus \{r\}$, there is a unique arc of T entering v and there exists a directed path from r to v in (V, T) . An r -cut in D is a set $\delta^+(S)$ where $r \in S$ and $V \setminus S \neq \emptyset$.

Let P_+^{arb} be the dominant of the convex hull of incidence vectors of r -arborescences. By a well know theorem of Edmonds [23],

$$P_+^{\text{arb}} = \{y \in \mathbb{R}^A : y(C) \geq 1 \text{ for all } r\text{-cuts } C; y \geq 0\}. \quad (10)$$

By the Max-flow Min-cut theorem, P_+^{arb} can be expressed as follows:

$$P_+^{\text{arb}} = \{y \in \mathbb{R}^A : \forall v \in V \setminus \{r\} \exists rv\text{-flow } z^v \text{ of value 1 s.t. } z^v \leq y\}$$

Indeed, for the inclusion “ \subseteq ”, let y be the incidence vector of an r -arborescence T . For each arc $e \in A$ and node $v \in V \setminus \{r\}$, let z_e^v

$$z_e^v = \begin{cases} 1 & \text{if } e \text{ is on the path from } r \text{ to } v \text{ in } T \\ 0 & \text{otherwise.} \end{cases}$$

Then, for every $v \in V \setminus \{r\}$, z^v is an rv -flow of value 1 such that $z^v \leq y$.

We show “ \supseteq ”. Let $y \in \mathbb{R}^A$ be such that, for every $v \in V \setminus \{r\}$, there exists an rv -flow z^v of value 1 such that $z^v \leq y$. Suppose by contradiction $y \notin P_+^{\text{arb}}$. Then, by Edmonds’ theorem, there exists an r -cut $C = \delta^+(S)$ such that $y(C) < 1$. But then, given $v \in V \setminus S$, there is no flow of value 1 from r to v satisfying the capacities y_e , $e \in A$, a contradiction.

Thus P_+^{arb} is the projection onto the y -space of the set of feasible solutions of the system

$$\begin{aligned} 0 &\leq z_e^v \leq y_e && v \in V \setminus \{r\}, e \in A \\ \sum_{e \in \delta^+(r)} z_e^v &= 1 && v \in V \setminus \{r\} \\ \sum_{e \in \delta^+(u)} z_e^v - \sum_{e \in \delta^-(u)} z_e^v &= 0 && u, v \in V \setminus \{r\}, u \neq v. \end{aligned}$$

Let $G = (V, E)$ be an undirected graph. Let Π be the set of partitions of V . For every $\mathcal{P} \in \Pi$, let $|\mathcal{P}|$ denote the number of classes in the partition, and $\delta(\mathcal{P})$ be the set of edges with endnodes in distinct classes of the partition. Edmonds [21] showed that the dominant of the convex hull of incidence vectors of spanning trees P_+^{tree} is

$$P_+^{\text{tree}} = \{x \in \mathbb{R}^E : \sum_{e \in \delta(\mathcal{P})} x_e \geq |\mathcal{P}| - 1, \mathcal{P} \in \Pi; x \geq 0\}.$$

Let $D = (V, A)$ be the digraph obtained from G by substituting every edge $e = uv$ with the arcs (u, v) , (v, u) . Then

$$P_+^{\text{tree}} = \{x \in \mathbb{R}^E : \exists y \in P_+^{\text{arb}} \text{ s.t. } x_{uv} = y_{(u,v)} + y_{(v,u)}, uv \in E\}.$$

Therefore the compact extended formulation for P_+^{arb} yields a compact extended formulation for P_+^{tree} .

4.4 Cuts

Let $D = (V, A)$ be a digraph and $r \in V$. Here we observe how the compact extended formulation for P_+^{arb} yields an extended formulation for the dominant of the convex hull of r -cuts, P_+^{cut} . This follows from the next result.

Lemma 4.8. *Let $R = \{y \in \mathbb{R}^n : \exists z \text{ s.t. } Ay + Cz \geq b\}$ and $P = \{x \in \mathbb{R}^n : y^T x \geq 1 \text{ for all } y \in R\}$. Then $P = \text{proj}_x(Q)$, where $Q = \{u : b^T u \geq 1, A^T u = x, C^T u = 0, u \geq 0\}$.*

Proof. Given $\bar{x} \in \mathbb{R}^n$, \bar{x} is in P if and only if $1 \leq \min\{\bar{x}^T y : y \in R\} = \min\{\bar{x}^T y : Ay + Cz \geq b\}$. By strong duality, this is equivalent to $1 \leq \max\{b^T u : A^T u = \bar{x}, C^T u = 0, u \geq 0\}$. \square

Edmonds' theorem [23] states that

$$P_+^{\text{rcut}} = \{x \in \mathbb{R}^A : x^T y \geq 1 \text{ for all } y \in P_+^{\text{arb}}; x \geq 0\}.$$

Since P_+^{arb} has a compact extended formulation, Lemma 4.8 shows how to derive an extended formulation for P_+^{rcut} .

Let $G = (V, E)$ be an undirected graph and let $R^{\text{cut}}(G)$ be the convex hull of all incidence vectors of cuts of G of the form $\delta(S)$, $S, V \setminus S \neq \emptyset$. Note that the cut polytope $P^{\text{cut}}(G)$ defined in Section 2.3 is the convex hull of $R^{\text{cut}}(G)$ union the origin. The polyhedron $R_+^{\text{cut}}(G)$ which is the dominant of $R^{\text{cut}}(G)$ is the *cut dominant*. Let $D = (V, A)$ be the digraph obtained from G by substituting every edge $e = uv$ with the arcs (u, v) , (v, u) . Then

$$R_+^{\text{cut}} = \{x \in \mathbb{R}^E : \exists y \in P_+^{\text{rcut}} \text{ s.t. } x_{uv} = y_{(u,v)} + y_{(v,u)} \quad uv \in E\}.$$

Therefore the compact extended formulation for P_+^{rcut} yields a compact extended formulation for R_+^{cut} .

4.5 Cycle cone

Given an undirected graph G , the *cycle cone* of G , denoted by C^{cycle} , is the cone generated by the incidence vectors of cycles of G . Seymour [56] shows that

$$C^{\text{cycle}} = \{x \in \mathbb{R}^E : x \geq 0, \quad x(\delta(S)) - 2x_e \geq 0 \text{ for all } e \in E \text{ and } S \subset V \text{ s.t. } e \in \delta(S)\}. \quad (11)$$

Barahona [9] constructs a compact extended formulation for C^{cycle} by observing the following. Let $D = (V, A)$ be the digraph obtained from G by substituting every edge $e = uv$ with the arcs (u, v) , (v, u) .

Remark 4.9. A vector $x \in \mathbb{R}^E$ is in C^{cycle} if and only if, for every $uv \in E$, there exists a uv -flow y^{uv} in D of value at least $2x_{uv}$ such that $0 \leq y^{uv} \leq x$.

Proof. Given $uv \in E$, by the Max-flow Min-cut theorem a nonnegative vector $x \in \mathbb{R}^E$ satisfies $x(\delta(S)) - 2x_{uv} \geq 0$ for all $S \subset V$ s.t. $uv \in \delta(S)$ if and only if there exists a uv -flow y^{uv} in D of value at least $2x_{uv}$ such that $0 \leq y^{uv} \leq x$. Hence the statement follows from (11). \square

The above remark shows that the following system of inequalities is an extended formulation for C^{cycle} .

$$\begin{aligned} 0 &\leq y_e^{uv} \leq x_e && uv \in E, e \in A \\ \sum_{e \in \delta^+(u)} y_e^{uv} &\geq 2x_{uv} && uv \in E \\ \sum_{e \in \delta^+(w)} y_e^{uv} - \sum_{e \in \delta^-(w)} y_e^{uv} &= 0 && uv \in E, w \neq V \setminus \{u, v\}. \end{aligned} \quad (12)$$

Note that the vertices of the polytope $C^{\text{cycle}} \cap \{x \in \mathbb{R}^E : \sum_{e \in E} x_e = 1\}$ are of the form $\frac{\chi^C}{|C|}$ for some cycle C , where χ^C denotes the incidence vector of C . Therefore, given $w \in \mathbb{R}^E$, the problem of finding a cycle of *minimum mean weight*, that is $\min\{w(C)/|C| : C \text{ cycle of } D\}$ can be reduced to solving the polynomial size linear program

$$\min\left\{\sum_{e \in E} w_e x_e : \sum_{e \in E} x_e = 1 \text{ and } \exists y \text{ s.t. } (x, y) \text{ satisfies (12)}\right\}.$$

Barahona [9] shows that computing a maximum weight matching can be reduced to a sequence of $O(|E|^2 \log |V|)$ minimum mean weight cycles computations. Therefore, although no compact extended formulation for the matching polytope is known, one can still efficiently solve the matching problem with linear programming.

5 Union of Polyhedra

In this section, we prove a result of Balas [4], [5] about the union of k polyhedra and we give several applications.

Theorem 5.1. (Balas [4], [5]) *Given k polyhedra $P^i := \{x \in \mathbb{R}^n : A^i x \leq b^i\}$, $i \in K$, let $r_1^i, \dots, r_{q_i}^i$ be such that for $i \in K$, the cone $C^i := \{x : A^i x \leq 0\} = \text{cone}\{r_1^i, \dots, r_{q_i}^i\}$. If $P^i \neq \emptyset$, let $v_1^i, \dots, v_{p_i}^i$ be such that $P^i = \text{conv}\{v_1^i, \dots, v_{p_i}^i\} + \text{cone}\{r_1^i, \dots, r_{q_i}^i\}$.*

Consider the polyhedron

$$P = \text{conv}\left(\bigcup_{i:P^i \neq \emptyset} \{v_1^i, \dots, v_{p_i}^i\}\right) + \text{cone}\left(\bigcup_{i \in K} \{r_1^i, \dots, r_{q_i}^i\}\right).$$

Then P is the projection onto the x -space of the polyhedron Y of points $(x, x^1, \dots, x^k, \delta) \in \mathbb{R}^n \times (\mathbb{R}^n)^k \times \mathbb{R}^k$ satisfying the system:

$$\begin{aligned} A^i x^i &\leq \delta^i b^i & i \in K \\ \sum_{i \in K} x^i &= x \\ \sum_{i \in K} \delta^i &= 1 \\ \delta^i &\geq 0 & i \in K. \end{aligned} \tag{13}$$

Proof. Given $\bar{x}^i \in P^i$, the vector $x^i = \bar{x}^i$, $x^j = 0$, $j \neq i$, $x = \bar{x}^i$, $\delta^i = 1$, $\delta^j = 0$, $j \neq i$ is in Y . Therefore if $Y = \emptyset$, then $\cup_{i \in K} P^i = \emptyset$.

In the system defining Y , constraint $\sum_{i \in K} \delta^i = 1$ forces at least one of δ^i to be positive. Therefore if the system defining Y is feasible, at least one of the systems $A^i x^i \leq b^i$ must be feasible. This shows that $Y = \emptyset$ if and only if $\cup_{i \in K} P^i = \emptyset$.

Since $P = \emptyset$ whenever $\cup_{i \in K} P^i = \emptyset$, this shows that $P = \text{proj}_x(Y)$ whenever P is empty.

We now assume that P is nonempty, i.e., the index set K can be partitioned into a nonempty set $K^N = \{i : P^i \neq \emptyset\}$ and $K^E = K \setminus K^N$.

Let $x \in P$. Then there exist points $v^i \in \text{conv}\{v_1^i, \dots, v_{p_i}^i\}$, $i \in K^N$ and rays $r^i \in \text{cone}\{r_1^i, \dots, r_{q_i}^i\}$, $i \in K$ such that

$$x = \sum_{i \in K^N} \delta^i v^i + \sum_{i \in K^N} r^i + \sum_{i \in K^E} r^i \text{ for some } \sum_{i \in K^N} \delta^i = 1, \delta^i \geq 0, i \in K^N.$$

For $i \in K^N$, define $x^i = \delta^i v^i + r^i$ and for $i \in K^E$ define $x^i = 0 + r^i$ and $\delta^i = 0$. Since r^i is a ray of C^i , then $A^i x^i \leq \delta^i b^i$ for all $i \in K$. Therefore every $x \in P$ can be completed with vectors x^i , $i \in K$ and scalars δ^i , $i \in K$ such that $(x, x^i, \delta^i, i \in K) \in Y$. This shows $P \subseteq \text{proj}_x(Y)$.

Let $(x, x^1, \dots, x^k, \delta)$ be a vector in Y . Define $K^P = \{i \in K : \delta^i > 0\}$ and let $z^i = \frac{x^i}{\delta^i}$, $i \in K^P$. Then $A^i z^i \leq b^i$. So $P^i \neq \emptyset$ in this case and $z^i \in \text{conv}\{v_1^i, \dots, v_{p_i}^i\} + \text{cone}\{r_1^i, \dots, r_{q_i}^i\}$.

For $i \in K \setminus K^P$, $A^i x^i \leq 0$, and therefore $x^i \in \text{cone}\{r_1^i, \dots, r_{q_i}^i\}$. Since $x = \sum_{i \in K^P} \delta^i z^i + \sum_{i \in K \setminus K^P} x^i$ and $\sum_{i \in K^P} \delta^i = 1, \delta^i \geq 0$, then $x \in \text{conv}(\bigcup_{i \in K^P} \{v_1^i, \dots, v_{p_i}^i\}) + \text{cone}(\bigcup_{i \in K} \{r_1^i, \dots, r_{q_i}^i\})$ and therefore $x \in P$. This shows that $\text{proj}_x(Y) \subseteq P$ and the proof is complete. \square

Remark 5.2. *Theorem 5.1 shows that the system of inequalities defining Y gives an extended formulation of the polyhedron P that uses $O(k(n+1))$ variables and the size of its formulation is approximately the sum of the sizes of the formulations that define the polyhedra P^i . So if k is small and the formulations defining the polyhedra P^i are compact, the formulation defining Y is also compact.*

We now give some examples: We will not always give an explicit description of the system Y . We only give the compact external descriptions of a small number k of polyhedra P^i and show that the internal descriptions of the polyhedra P^i and P satisfy the condition of Theorem 5.1.

5.1 All even subsets

We consider here the set:

$$EVEN^n = \{x \in \{0, 1\}^n : x \text{ has an even number of } 1^s\}$$

Jeroslow [39] proves the following:

Theorem 5.3. *Let \mathcal{S} be the family of subsets of $N = \{1, \dots, n\}$ having odd cardinality. Then*

$$\text{conv}(EVEN^n) = \left\{ x \in \mathbb{R}^n : \begin{array}{l} \sum_{i \in \mathcal{S}} x_i - \sum_{i \in N \setminus \mathcal{S}} x_i \leq |\mathcal{S}| - 1, \quad \mathcal{S} \in \mathcal{S} \\ 0 \leq x_i \leq 1, \quad i \in N \end{array} \right\}.$$

Proof. Let x^E be an optimal vector for the program:

$$\max \sum_{i \in N} c_i x_i : x \in \text{conv}(EVEN^n) \quad (14)$$

and let E be the even subset of N represented by x^E . Consider the pair of dual linear programs:

$$\begin{array}{ll} \max \sum_{i \in N} c_i x_i & \\ \sum_{i \in \mathcal{S}} x_i - \sum_{i \in N \setminus \mathcal{S}} x_i \leq |\mathcal{S}| - 1 & \mathcal{S} \in \mathcal{S} \\ x_i \leq 1 & i \in N \\ x_i \geq 0 & i \in N \end{array} \quad (P)$$

$$\begin{array}{ll} \min \sum_{\mathcal{S} \in \mathcal{S}} (|\mathcal{S}| - 1) y_{\mathcal{S}} + \sum_{i \in N} z_i & \\ \sum_{\mathcal{S} \in \mathcal{S}, \mathcal{S} \ni i} y_{\mathcal{S}} - \sum_{\mathcal{S} \in \mathcal{S}, \mathcal{S} \not\ni i} y_{\mathcal{S}} + z_i \geq c_i & i \in N \\ z_i \geq 0 & i \in N \\ y_{\mathcal{S}} \geq 0 & \mathcal{S} \in \mathcal{S}. \end{array} \quad (D)$$

We show that (D) admits a feasible solution (y^*, z^*) having value $\sum_{i \in N} c_i x_i^E$. Since x^E is feasible for (P), this shows that x^E and (y^*, z^*) are optimal solutions for (P) and (D).

Since x^E is optimal for the program (14), it is easy to see that E satisfies one of the following three cases:

Case 1: $c_i \geq 0, i \in E$ and $c_i \leq 0, i \in N \setminus E$.

The vector (y^*, z^*) satisfying the above requirement is:

$$y_S^* = 0, S \in \mathcal{S}, \quad z_i^* = \max\{c_i, 0\}, i \in N.$$

Case 2: There is an element $i^* \in E$ such that $c_{i^*} < 0$ and $c_i + c_{i^*} \geq 0, i \in E \setminus \{i^*\}$, $c_i \geq c_{i^*}, i \in N \setminus E$.

Let $S^* = E \setminus \{i^*\}$. The vector (y^*, z^*) satisfying the above requirement is:

$$y_{S^*}^* = -c_{i^*}, \quad y_S^* = 0, S \in \mathcal{S} \setminus \{S^*\}, \quad z_i^* = \max\{c_i + c_{i^*}, 0\}, i \in N.$$

Case 3: There is an element $i^* \in N \setminus E$ such that $c_{i^*} > 0$ and $c_i + c_{i^*} \leq 0, i \in N \setminus (E \cup \{i^*\})$, $c_i \geq c_{i^*}, i \in E$. Let $S^* = E \cup \{i^*\}$. The vector (y^*, z^*) satisfying the above requirement is:

$$y_{S^*}^* = c_{i^*}, \quad y_S^* = 0, S \in \mathcal{S} \setminus \{S^*\}, \quad z_i^* = \max\{c_i - c_{i^*}, 0\}, i \in N.$$

□

The formulation of $\text{conv}(EVEN^n)$ given in Theorem 5.3 is not compact. However, Theorem 5.1 gives us the means of obtaining a compact extended formulation. We present it next.

Let $S_k = \{x \in \{0, 1\}^n : x \text{ has } k \text{ } 1^s\}$. It is straightforward to see that $\text{conv}(S_k) = \{x \in \mathbb{R}^n : 0 \leq x_i \leq 1, i \in N; \sum_{i \in N} x_i = k\}$ (Easy direct proof, total unimodularity).

Let $N^{ev} = \{0 \leq k \leq n, k \text{ even}\}$. Since $EVEN^n = \bigcup_{k \in N^{ev}} S_k$, Theorem 5.1 implies that

$$\text{conv}(EVEN^n) = \text{proj}_x(Q)$$

where Q is the polytope defined by the following system

$$\begin{aligned} x_i - \sum_{k \in N^{ev}} x_i^k &= 0 & i \in N \\ \sum_{i \in N} x_i^k &= k\lambda^k & k \in N^{ev} \\ \sum_{k \in N^{ev}} \lambda^k &= 1 \\ x_i^k &\leq \lambda^k & i \in N, k \in N^{ev} \\ x_i^k &\geq 0 & i \in N, k \in N^{ev} \\ \lambda^k &\geq 0 & k \in N^{ev}. \end{aligned}$$

5.2 Cut Dominant

Given an undirected graph $G = (V, E)$, the cut dominant polyhedron $R_+^{\text{cut}}(G)$ has been defined in Section 4.4. We provide another extended formulation for $R_+^{\text{cut}}(G)$ using Theorem 5.1.

Let V^{cut} be the set of incidence vectors of cuts of G of the form $\delta(S)$, $S, V \setminus S \neq \emptyset$. Let $s \in V$. For every $t \in V \setminus \{s\}$, let us denote by $V^{st\text{-cut}}$ the set of incidence vectors of st -cuts of G . Then

$$V^{\text{cut}} = \bigcup_{t \in V \setminus \{s\}} V^{st\text{-cut}}$$

therefore

$$R_+^{\text{cut}}(G) = \text{conv}\left(\bigcup_{t \in V \setminus \{s\}} P_+^{st\text{-cut}}(G)\right).$$

Since $P_+^{st\text{-cut}}(G)$ has a compact extended formulation, discussed in Section 4.2, Theorem 5.1 gives a compact extended formulation for $R_+^{\text{cut}}(G)$. Using Remark 5.2, we observe that this extended formulation has $O(|V||E|)$ variables and $O(|V||E|)$ constraints.

To the best of our knowledge, the most compact extended formulation for the cut dominant in a dense graph is currently the one given by Carr et al. [13], and it uses $O(|V|^2)$ variables and $O(|V|^3)$ constraints.

An external description of $R_+^{\text{cut}}(G)$ in the original space is not known. Describing the facets of $R_+^{\text{cut}}(G)$ is equivalent to describing the vertices of the subtour relaxation of the Graphical Traveling Salesman polytope, which is the blocking polyhedron [30].

5.3 An approximate formulation for the knapsack set

Consider the 0 – 1 knapsack set $K(n, b)$ with n items and capacity b :

$$K(n, b) = \{x \in \{0, 1\}^n : \sum_{i=1}^n a_i x_i \leq b\}.$$

Given a vector $c \in \mathbb{R}^n$, let

$$W^* = \max\left\{\sum_{i=1}^n c_i x_i : x \in K(n, b)\right\}.$$

Ibarra and Kim [38], and Lawler [45] gave a fully polynomial-time approximation scheme for the knapsack problem, i.e., an algorithm that, for any given $\varepsilon > 0$, returns a feasible solution whose value is at least $(1 - \varepsilon)W^*$ in time polynomial in n and ε^{-1} . This prompts a question, posed by Van Vyve and Wolsey [58], of whether there exists a system of inequalities $Ax + Bz \leq d$, of size polynomial in n and ε^{-1} , such that

- (i) $\text{conv}(K(n, b)) \subseteq \{x \in \mathbb{R}^n : \exists z \text{ s.t. } Ax + Bz \leq d\}$,
- (ii) for every $c \in \mathbb{R}^n$, $W^* \geq (1 - \varepsilon) \max\{\sum_{i=1}^n c_i x_i : Ax + Bz \leq d\}$.

Note that (i) and (ii) are equivalent to saying that, given $Q = \{x \in \mathbb{R}^n : \exists z \text{ s.t. } Ax + Bz \leq d\}$, $(1 - \varepsilon)Q \subseteq \text{conv}(K(n, b)) \subseteq Q$.

The above question is still open. However, Bienstock [11] proved that there exists a system $Ax + Bz \leq d$ satisfying (i) and (ii) whose size is polynomial in n if ε is fixed. More precisely, the formulation has $O(\varepsilon^{-1}n^{1+\lceil\varepsilon^{-1}\rceil})$ variables and constraints.

The formulation is based on Theorem 5.1. The idea is to give a family \mathcal{F} of $O(n^{\lceil\varepsilon^{-1}\rceil})$ polytopes inside the unit cube such that $K(n, b) \subseteq \cup_{Q \in \mathcal{F}} Q$ and, for every $c \in \mathbb{R}^n$, $W^* \geq (1 - \varepsilon) \max\{\sum_{i=1}^n c_i x_i : x \in Q\}$ for every $Q \in \mathcal{F}$.

Let us denote $H = \lceil\varepsilon^{-1}\rceil$. The family \mathcal{F} has a member Q^S for each set $S \subseteq \{1, \dots, n\}$ of cardinality at most H such that $\sum_{i \in S} a_i \leq b$. Note that the number of these sets S is $O(Hn^H)$.

The polytope Q^S is defined as follows:

- If $|S| < H$, $Q^S = \{x \in \mathbb{R}^n : x_j = 1, j \in S; x_j = 0, j \notin S\}$,
- If $|S| = H$,

$$Q^S = \left\{ x \in \mathbb{R}^n : \begin{array}{ll} \sum_{j=1}^n a_j x_j \leq b & \\ 0 \leq x_j \leq 1 & j = 1, \dots, n \\ x_j = 1 & j \in S \\ x_j = 0 & j \notin S \text{ s.t. } a_j > \min_{i \in S} a_i \end{array} \right\}.$$

It is clear that $K(n, b)$ is contained in the union of the Q^S . Indeed, given $\bar{x} \in K(n, b)$, let $\text{supt}(x) = \{i : x_i = 1\}$. If $|\text{supt}(\bar{x})| < H$, then if $S = \text{supt}(\bar{x})$, $Q^S = \{\bar{x}\}$. If $|\text{supt}(\bar{x})| \geq H$, let S be the set of the indices in $\text{supt}(\bar{x})$ relative to the H largest elements a_j , $j \in \text{supt}(\bar{x})$. Then $\bar{x} \in Q^S$.

Let $P = \text{conv}(\cup_{Q^S \in \mathcal{F}} Q^S)$. Note that, for every $Q^S \in \mathcal{F}$, Q^S is a polytope expressed by $O(n)$ linear inequalities. Thus, by Balas' Theorem 5.1, the system of inequalities (13) gives an extended formulation for P with $n + |\mathcal{F}|n + |\mathcal{F}|$ variables and $|\mathcal{F}|O(n) + n + 1 + |\mathcal{F}|$ constraints. Both numbers are of order $O(\varepsilon^{-1}n^{1+\lceil\varepsilon^{-1}\rceil})$.

Theorem 5.4. For every $c \in \mathbb{R}^n$,

$$\max\left\{\sum_{i=1}^n c_i x_i : x \in K(n, b)\right\} \geq (1 - \varepsilon) \max\left\{\sum_{i=1}^n c_i x_i : x \in P\right\}.$$

Proof. Let $c \in \mathbb{R}^n$ and let $W^* = \max\{\sum_{i=1}^n c_i x_i : x \in K(n, b)\}$.

We only need to show that, given $S \subseteq \{1, \dots, n\}$ of cardinality at most H such that $\sum_{i \in S} a_i \leq b$, we have $W^* \geq (1 - \varepsilon) \max\{\sum_{i=1}^n c_i x_i : x \in Q^S\}$.

If $|S| < H$, then the result holds trivially since Q^S consists of the point \bar{x} such that $\text{supt}(\bar{x}) = S$, and $\bar{x} \in K(n, b)$.

Suppose now $|S| = H$. We may assume that, among all such sets, S is chosen so that $\max\{\sum_{i=1}^n c_i x_i : x \in Q^S\}$ is largest possible. Let \tilde{x} be an optimal basic solution for the above problem. We may assume \tilde{x} is not integral, otherwise the statement follows.

Thus there exists exactly one index q such that $0 < \tilde{x}_q < 1$. Since $\tilde{x}_q > 0$ and $\tilde{x} \in Q^S$, it follows that $a_q \leq \min_{i \in S} a_i$. We next show that $c_q \leq \min_{i \in S} c_i$. If not, then let $p \in S$

such that $c_p < c_q$. Let \hat{x} be the vector defined by $\hat{x}_p = \tilde{x}_q$, $\hat{x}_q = 1$, and $\hat{x}_j = \tilde{x}_j$, $1 \leq j \leq n$, $j \neq p, q$. Let $S' = S \setminus \{p\} \cup \{q\}$. Since $a_p \geq a_q$, it follows $\hat{x} \in Q^{S'}$. Furthermore, since $c_p < c_q$, $\sum_{i=1}^n c_i \tilde{x}_i < \sum_{i=1}^n c_i \hat{x}_i$, contradicting our choice of S .

We define the vector \bar{x} by $\bar{x}_j = \lfloor \tilde{x}_j \rfloor$, $j = 1, \dots, n$. Clearly $\bar{x} \in K(n, b)$. Furthermore, since $c_q \leq \min_{i \in S} c_i$, we have

$$\frac{\sum_{i=1}^n c_i \tilde{x}_i - \sum_{i=1}^n c_i \bar{x}_i}{\sum_{i=1}^n c_i \tilde{x}_i} \leq \frac{c_q}{Hc_q} \leq \varepsilon$$

and therefore $(1 - \varepsilon) \sum_{i=1}^n c_i \tilde{x}_i \leq \sum_{i=1}^n c_i \bar{x}_i \leq W^*$. \square

Bienstock and McClosky [12] give further approximate extended formulations for knapsack and fixed charge network flow problems based on disjunctions. Van Vyve and Wolsey [58] give extended approximate formulations for several lot-sizing problems.

5.4 Continuous mixing set

The set of vertices of some selected mixed-integer sets can sometimes be partitioned into few subsets in which the continuous variables take specified fractional values. The convex hull of each of the subsets can then be found by applying the theory of (pure) Integer Programming.

We illustrate this by considering the *continuous mixing set*, defined as the following mixed-integer set:

$$X^{CMIX} = \{(s, y, x) \in \mathbb{R}_+ \times \mathbb{R}_+^n \times \mathbb{Z}_+^n : s + y_t + x_t \geq b_t, 1 \leq t \leq n\}.$$

Miller and Wolsey [48] gave a compact extended formulation for the polyhedron $\text{conv}(X^{CMIX})$ and characterized the vertices and rays. It follows from their work that the internal description of $\text{conv}(X^{CMIX})$ has exponential size. Van Vyve [57] has provided a new more compact extended formulation that only involves $O(n)$ additional variables, and has shown that the separation problem in the original space can be solved by flow techniques.

Remark 5.5. *Let (s^*, y^*, x^*) be a vertex of $\text{conv}(X^{CMIX})$. Then (s^*, y^*) is a vertex of the polyhedron $Q = \{(s, y) \in \mathbb{R}_+ \times \mathbb{R}_+^n : s + y_t \geq b_t - x_t^*, 1 \leq t \leq n\}$.*

Define $b_0 = 0$.

Lemma 5.6. *Let (s^*, y^*, x^*) be a vertex of $\text{conv}(X^{CMIX})$. Then*

1. $s^* \equiv b_t \pmod{1}$ for some $0 \leq t \leq n$.
2. For $1 \leq t \leq n$, either $y_t^* = 0$ or $y_t^* \equiv b_t - s^* \pmod{1}$.
3. The $2n + 1$ extreme rays of $\text{conv}(X^{CMIX})$ are: $(1, 0, 0)$, $(0, e_j, 0)$ and $(0, 0, e_j)$.

Proof. Assume first that $s^* > 0$. Let T be the set of indices of the inequalities $s + y_t + x_t \geq b_t$ that are satisfied at equality by (s^*, y^*, x^*) .

We claim that T is nonempty and $y_t^* = 0$ for at least one $t \in T$. If not, let e_T be the incidence vector of the set T . Then for some $\epsilon > 0$, the polyhedron Q defined in Remark 5.5 contains both $(s^*, y^*) \pm \epsilon(1, e_T)$, showing that (s^*, r^*) is not a vertex of Q , a contradiction to Remark

5.5 and the claim is proved.

Let t^* be such that $s^* + y_{t^*}^* + x_{t^*}^* = b_{t^*}$ and $y_{t^*}^* = 0$, then $s^* \equiv b_{t^*} \pmod{1}$ and Statement 1 is proved.

Statement 2 follows immediately from Remark 5.5 and the statement about extreme rays is immediate. \square

For $0 \leq i \leq n$, define $X_i = \{(s, y, x) \in X^{CMIX} : s \equiv b_i \pmod{1}\}$. By Lemma 5.6, the set of vertices of polyhedron $\text{conv}(X^{CMIX})$ is the union of the sets of vertices of polyhedra $\text{conv}(X_i)$ and the recession cones of $\text{conv}(X^{CMIX})$ and $\text{conv}(X_i)$ coincide. Therefore a compact formulation for $\text{conv}(X_i)$ can be used to derive a compact extended formulation for $\text{conv}(X^{CMIX})$.

For $1 \leq t \leq n$ and $0 \leq i \leq n$, let $f_{ti} = (b_t - f_i) - \lfloor b_t - f_i \rfloor$.

Theorem 5.7. *The following set of inequalities gives a formulation for $\text{conv}(X_i)$:*

$$\begin{aligned} s &\geq f_i \\ s + y_t + x_t &\geq b_t, \quad 1 \leq t \leq n \\ y_t + f_{ti}(x_t + s) &\geq f_{ti}(\lfloor b_t - f_i \rfloor + f_i), \quad 1 \leq t \leq n \\ s &\in \mathbb{R}_+, y \in \mathbb{R}_+^n, x \in \mathbb{R}_+^n. \end{aligned}$$

Proof. We model the condition $s \equiv f_i \pmod{1}$ with $s = \sigma + f_i$, $\sigma \in \mathbb{Z}_+$. Substituting for s in the set of inequalities defining X^{CMIX} , we obtain:

$$\begin{aligned} \sigma + y_t + x_t &\geq b_t - f_i \text{ for } t = 1, \dots, n \\ \sigma &\in \mathbb{Z}_+, y \in \mathbb{R}_+^n, x \in \mathbb{Z}_+^n. \end{aligned}$$

By Lemma 5.6, in a vertex of the convex hull of this set, either $y_t \equiv 0 \pmod{1}$ or $y_t \equiv f_{ti} \pmod{1}$. This leads us to write $y_t = \mu_t + f_{ti}\delta_t$ with $\mu_t \in \mathbb{Z}_+$, $\delta_t \in \{0, 1\}$. Substituting for y_t in the above system, we obtain :

$$\begin{aligned} \sigma + \mu_t + f_{ti}\delta_t + x_t &\geq b_t - f_i \text{ for } t = 1, \dots, n \\ \sigma &\in \mathbb{Z}_+, \mu \in \mathbb{Z}_+^n, \delta \in \{0, 1\}^n, x \in \mathbb{Z}_+^n. \end{aligned}$$

Applying Chvátal-Gomory rounding to the above system, an equivalent, but tighter, set of inequalities is:

$$\begin{aligned} \sigma + \mu_t + \lceil f_{ti} \rceil \delta_t + x_t &\geq \lceil b_t - f_i \rceil, \quad 1 \leq t \leq n \\ \sigma &\in \mathbb{Z}_+, \delta \in \{0, 1\}^n, \mu \in \mathbb{Z}_+^n, x \in \mathbb{Z}_+^n. \end{aligned}$$

Observe now that the matrix defining the above system is a totally unimodular matrix, and the requirements vector and bounds are integer. It follows from the theorem of Hoffman and Kruskal [37] that we can replace the integrality requirements with variable bounds, and obtain a formulation for the above set. This yields the following extended formulation for $\text{conv}(X^i)$:

$$\begin{aligned} s &= \sigma + f_i \\ y_t &= \mu_t + f_{ti}\delta_t \text{ for } t = 1, \dots, n \\ \sigma + \mu_t + \lceil f_{ti} \rceil \delta_t + x_t &\geq \lceil b_t - f_i \rceil, \quad 1 \leq t \leq n \\ \sigma &\in \mathbb{R}_+, \mu \in \mathbb{R}_+^n, \delta \in [0, 1]^n, x \in \mathbb{R}_+^n. \end{aligned}$$

Projecting back into the original s, r, y space using Fourier's method, eliminating first σ , then μ and finally δ , it is easily checked that one obtains the set of inequalities in the statement of the theorem. \square

The above theorem is a simplified version of a result that appears in [17].

6 The size of an extended formulation

6.1 Yannakakis's Theorem

Given a polyhedron $P \subseteq \mathbb{R}^n$, what is the smallest number of variables and constraints needed in an extended formulation? Yannakakis [61] gave an elegant answer to this question.

The *nonnegative rank* of a nonnegative matrix S is the smallest integer t such that $S = FW$ where F and W are nonnegative matrices of size $m \times t$ and $t \times p$.

Given a polytope $P \subseteq \mathbb{R}^n$, let v^1, \dots, v^p be its vertices and let $Ax \leq b$ be a system of linear inequalities such that $P = \{x \in \mathbb{R}^n : Ax \leq b\}$. Let $a^i x \leq b_i$, $i = 1, \dots, m$, denote the inequalities in $Ax \leq b$. Define the *slack matrix* of $Ax \leq b$ to be the $m \times p$ matrix S where the entry

$$s_{ij} := b_i - a^i v^j$$

is the slack of vertex v^j in the i th inequality. Note that S is a nonnegative matrix.

We first want to show that the nonnegative rank of the slack matrix does not depend on the particular system of inequalities defining P , but it is a property of the polytope itself. To do this we need the following observation, which strengthens a basic well known fact in polyhedral theory.

Lemma 6.1. *If $P = \{x : Ax \leq b\}$ is a polytope of dimension at least one, and $\alpha x \leq \beta$ is a valid inequality for P , then there exists $\lambda \geq 0$ such that $\lambda A = \alpha$ and $\lambda d = \beta$.*

Proof. Since P is a polytope and $\dim(P) > 1$, $\delta_1 := \max\{\alpha x : Ax \leq d\}$ and $\delta_2 := \max\{-\alpha x : Ax \leq d\}$ both exist finite and $d_1 + d_2 > 0$. By LP duality there exist $\lambda^1 \geq 0$, $\lambda^2 \geq 0$ s.t. $\lambda^1 A = \alpha$, $\lambda^1 d = d_1$, $\lambda^2 A = -\alpha$, $\lambda^2 d = d_2$. Since $\alpha x \leq \beta$, $d_1 \leq \beta$. It follows that the vector $\lambda := \lambda^1 + (\lambda^1 + \lambda^2)(\beta - d_1)(d_1 + d_2)^{-1}$ satisfies the statement of the lemma. \square

Lemma 6.2. *Let $Ax \leq b$ and $Cx \leq d$ be systems defining polytope P and let S, S' be the corresponding slack matrices. Then the nonnegative ranks of S and S' coincide.*

Proof. The statement is obvious if P consists of only one point, thus we assume $\dim(P) \geq 1$. By symmetry it suffices to show that the nonnegative rank of S is at most the nonnegative rank of S' . Because every inequality of $Ax \leq b$ is valid for P , it follows from Lemma 6.1 that there exists a nonnegative matrix L such that $A = LC$ and $b = Ld$. This implies that $S = LS'$. Given any factorization $S' = F'W$ where F', W are nonnegative matrices, if we let $F = LF'$, we have that F is a nonnegative matrix and $S = FW$. \square

By the above remark, we may define the *nonnegative rank of the polytope P* as the nonnegative rank of the slack matrix S of any system defining P .

Theorem 6.3 (Yannakakis [61]). *Let $P \subset \mathbb{R}^n$ be a polytope, and let t be its nonnegative rank. Every extended formulation of P has at least t constraints, and there exists an extended formulation of P with $t + n$ variables and $O(t + n)$ constraints.*

Proof. The statement is obvious if P consist of only one point, thus we may assume that $\dim(P) \geq 1$. Let $Ax \leq b$ be a system defining P , and let S be the slack matrix of $Ax \leq b$.

We first prove that every extended formulation has at least t constraints. Indeed, let

$$Q = \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^k : Rx + Tz \leq d\}$$

be a polyhedron such that $P = \text{proj}_x(Q)$, where (R, T) is an $h \times (n + k)$ matrix. By Lemma 6.1, there exists a nonnegative vector f^i such that $f^i T = 0$, $f^i R = a^i$, $f^i d = b_i$, $i = 1, \dots, m$. Also, for any vertex v^j of P there exists $w^j \in \mathbb{R}^k$ such that $(v^j, w^j) \in Q$. Let $w^j := d - Rv^j - Tw^j$, $j = 1, \dots, p$. Note that w^j is nonnegative. For $i = 1, \dots, h$ and $j = 1, \dots, p$, we have $f^i w^j = f^i(d - Rv^j - Tw^j) = b_i - a^i v^j$. If we define the $m \times h$ matrix F with rows f^i , $i = 1, \dots, m$, and the $h \times p$ matrix W with columns w^j , $j = 1, \dots, p$, we have $S = FW$. Since F and W are nonnegative matrices and F has h columns, it follows that $t \leq h$.

Next we show that there exists an extended formulation with $t + n$ variables and $O(t + n)$ constraints. Indeed, let F and W be nonnegative matrices of size $m \times t$ and $t \times p$ such that $S = FW$. Let $f^1, \dots, f^m \in \mathbb{R}^t$ be the rows of F and $w^1, \dots, w^p \in \mathbb{R}^t$ be the columns of W . Define

$$Q := \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^t : Ax + Fz = b, z \geq 0\}. \quad (15)$$

We will show that $P = \text{proj}_x(Q)$. To prove the inclusion $\text{proj}_x(Q) \subseteq P$, note that for any $(x, z) \in Q$ we have that $Ax = b - Fz \leq b$, because $F \geq 0$ and $z \geq 0$, and therefore $x \in P$. For the inclusion $P \subseteq \text{proj}_x(Q)$, it suffices to show that every vertex v^j of P is contained in $\text{proj}_x(Q)$. This is the case since, for any vertex v^j of P , the point (v^j, w^j) is in Q because $w^j \geq 0$ and $f^i w^j = b_i - a^i v^j$ for $i = 1, \dots, m$.

We have shown that $P = \text{proj}_x(Q)$. Since $Q \subseteq \mathbb{R}^n \times \mathbb{R}^t$, at most $n + t$ of the equations $Ax + Fz = b$ are linearly independent. Thus Q can be described with $n + t$ variables and at most $n + 2t$ constraints. \square

Note that expressing the polytope P by a linear system in the original variables corresponds to the factorization $S = IS$ with the identity matrix I .

Yannakakis' result is stated for polytopes, but it can be extended in a natural way to pointed polyhedra. Indeed, let P be a polyhedron, and let points v^1, \dots, v^p and vectors r^1, \dots, r^q be a minimal internal description of P (thus $P := \text{conv}\{v^1, \dots, v^p\} + \text{cone}\{r^1, \dots, r^q\}$). A *tight external description* of P is a system $Ax \leq b$ such that $P = \{x : Ax \leq b\}$ and every inequality of $Ax \leq b$ defines a nonempty face of P .

Given any tight description $Ax \leq b$ of P , let m be the number of rows of A , and define the slack matrix of P to be the $m \times (p + q)$ matrix S whose entries are defined by

$$s_{s_{ij}} = \begin{cases} b_i - a^i v^j & \text{if } 1 \leq j \leq p; \\ -a^i r^{j-p} & \text{if } p + 1 \leq j \leq q. \end{cases}$$

It can be shown, analogously to Remark 6.2, that the nonnegative rank of the slack matrix is independent on the choice of the tight representation, and we can therefore define the nonnegative rank of P as the nonnegative rank of the slack matrix of any tight description. The statement of Theorem 6.3 now extends naturally to the case where P is a polyhedron (see [15]).

Given a factorization of S as FW , the proof of Yannakakis' theorem is constructive for obtaining an extended formulation. The catch is that factorizing S is a difficult task in general. Vavasis [59] shows that computing the nonnegative rank of a given matrix is NP-hard. On the other hand, Yannakakis' theorem is a powerful tool for proving lower bounds on the size of extended formulation. The most remarkable example of this is the proof, by Fiorini et al. [27], that the cut polytope, the TSP polytope, and the stable set polytope, do not admit compact extended formulations. We give the proof for the cut polytope in Section 6.3.

In Section 6.4 we will show examples where compact extended formulations can be constructed by finding a nonnegative factorization of the slack matrix of polynomial rank.

There are scant examples of families of polytopes for which the nonnegative rank is actually known. One such example is given by the permutahedron, whose nonnegative rank is of order $\Theta(n \log(n))$ (Goemans [33]). We will discuss this result in Section 6.7. Another example is the class of regular m -gons, whose nonnegative rank is $\Theta(\log m)$ (Ben-Tal and Nemirowski [10]). For general m -gons, Fiorini, Rothvoss, and Tiwary [28] prove a lower-bound of $\Omega(\sqrt{m})$.

6.2 Rectangle covering bound

A 0,1 matrix $R \in \mathbb{R}^{m \times n}$ is said to be a *rectangle matrix* if there exist sets $I \subseteq \{1, \dots, m\}$ and $J \subseteq \{1, \dots, n\}$ such that, for $i = 1, \dots, m$ and $j = 1, \dots, n$, $r_{ij} = 1$ if $i \in I$ and $j \in J$, while $r_{ij} = 0$ otherwise. Given a matrix $S \in \mathbb{R}^{m \times n}$, a *rectangle cover* of S is a family of rectangle matrices R^1, \dots, R^k such that, for every $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$, $s_{ij} = 0$ if and only if $r_{ij}^1, \dots, r_{ij}^k = 0$. The *rectangle covering number* of S is the smallest k such that a rectangle cover of S of cardinality k exists.

Lemma 6.4. *Let $P \subset \mathbb{R}^n$ be a polytope. The number of constraints in any extended formulation of P is at least the rectangle covering number of the slack matrix of any system defining P .*

Proof. Let $S \in \mathbb{R}^{m \times n}$ be the slack matrix of a system defining P , and let t be its nonnegative rank. Let $F \in \mathbb{R}^{m \times t}$ and $W \in \mathbb{R}^{t \times n}$ be nonnegative matrices such that $S = FW$. Let ϕ^1, \dots, ϕ^t be the columns of F , and $\omega^1, \dots, \omega^t$ be the rows of W . For $h = 1, \dots, t$, define the $m \times n$ matrix R^h as follows,

$$r_{ij}^h = \begin{cases} 1 & \text{if } \phi_i^h \omega_j^h > 0 \\ 0 & \text{if } \phi_i^h \omega_j^h = 0. \end{cases} \quad i = 1, \dots, m, j = 1, \dots, n.$$

One can readily verify that R^1, \dots, R^t is a rectangle cover of S , therefore the nonnegative rank of S is at least the rectangle covering number of S . The statement now follows from Theorem 6.3. \square

6.3 An exponential lower-bound for the cut polytope

Fiorini et al. [27] showed that TSP, max-cut and the stable set polytope cannot have a polynomial-size extended formulation. The core of their proof is in establishing the lower bound for the cut polytope. The lower bounds for the stable set and TSP polytopes then follow by reductions to the cut polytope in a way that is quite similar to NP completeness proofs.

Let P_n^{cut} denote the cut polytope of order n , that is, the convex hull of incidence vectors of all cuts of the complete graph $K_n = (V_n, E_n)$ on n vertices (i.e. $P_n^{\text{cut}} := P^{\text{cut}}(K_n)$).

The proof of the lower bound is not done directly on the cut polytope, but on a polytope that is linearly isomorphic to it. The *correlation polytope* P_n^{corr} of order n is the convex hull of all 2^n $n \times n$ matrices of the form bb^\top , $b \in \{0, 1\}^n$.

Two polytopes $P \subseteq \mathbb{R}^p$ and $Q \subseteq \mathbb{R}^q$ are *linearly isomorphic* if there exists a linear function $f : \mathbb{R}^p \rightarrow \mathbb{R}^q$ such that $f(P) = Q$ and for all $y \in Q$ there is a unique $x \in P$ for which $f(x) = y$.

Lemma 6.5. *For all n , P_n^{corr} and P_{n+1}^{cut} are linearly isomorphic.*

Proof. Let $f : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{E_{n+1}}$ be the function that associates to each $x \in \mathbb{R}^{n \times n}$ the element $y \in \mathbb{R}^{E_{n+1}}$ defined by

$$y_{ij} = \begin{cases} x_{ii} & \text{if } 1 \leq i \leq n, j = n+1, \\ x_{ii} + x_{jj} - 2x_{ij} & \text{if } 1 \leq i < j \leq n. \end{cases}$$

One can verify that f maps P_n^{corr} to P_{n+1}^{cut} and that, for every $y \in P_{n+1}^{\text{cut}}$, the only point x in P_n^{corr} such that $f(x) = y$ is the one defined by

$$x_{ij} = \begin{cases} y_{i,n+1} & \text{if } i = j, \\ \frac{1}{2}(y_{i,n+1} + y_{j,n+1} - y_{ij}) & \text{if } i \neq j, \end{cases} \quad i, j = 1, \dots, n.$$

□

Given a vector $a \in \mathbb{R}^n$, we denote by $\text{diag}(a)$ the $n \times n$ matrix with a on the diagonal and 0 everywhere else. We denote by $\langle \cdot, \cdot \rangle$ the *Frobenius product* of two matrices, that is $\langle A, B \rangle := \sum_i \sum_j a_{ij} b_{ij}$. We recall that the *trace* of a square matrix C , denoted by $\text{tr}(C)$, is the sum of its diagonal entries. It is well known that $\langle A, B \rangle = \text{tr}(A^\top B) = \text{tr}(BA^\top)$.

Lemma 6.6. *For all $a \in \{0, 1\}^n$, the inequality*

$$\langle 2\text{diag}(a) - aa^\top, y \rangle \leq 1 \tag{16}$$

is valid for P_n^{corr} . Furthermore, the slack of the vertex $y = bb^\top$, $b \in \{0, 1\}^n$, is $(1 - a^\top b)^2$.

Proof. For all $a, b \in \{0, 1\}^n$, we have

$$\begin{aligned} 1 - \langle 2\text{diag}(a) - aa^\top, bb^\top \rangle &= 1 - 2a^\top b + \text{tr}(aa^\top bb^\top) \\ &= 1 - 2a^\top b + \text{tr}(b^\top aa^\top b) \\ &= (1 - a^\top b)^2, \end{aligned}$$

where the first equality follows from the fact that $b \in \{0, 1\}^n$. Since $(1 - a^\top b)^2 \geq 0$, (16) is valid for P_n^{corr} . □

Theorem 6.7. *There exists a constant $C > 0$ such that, for all n , every extended formulation for P_n^{cut} has at least 2^{Cn} constraints.*

Proof. By Lemma 6.5, it suffices to prove that there exists some constant C such that every extended formulation for P_n^{corr} has at least 2^{Cn} constraints. Let $Ay \leq b$ be a system of inequalities such that $P_n^{\text{corr}} = \{y \in \mathbb{R}^{n \times n} : Ay \leq b\}$ and such that all the inequalities (16) are included in $Ay \leq b$. Let S be the slack matrix of $Ay \leq b$. Let M be the $2^n \times 2^n$ matrix whose rows and columns are indexed by the vectors in $\{0, 1\}^n$, defined by $M_{ab} = (1 - a^\top b)^2$ for all $a, b \in \{0, 1\}^n$. By Lemma 6.6, M is the row-submatrix of S relative to the 2^n constraints (16). De Wolf [60] showed that there exists some constant C such that the rectangle covering number of M is at least 2^{Cn} for all n . Since M is a submatrix of S , it follows that the rectangle covering number of S is at least 2^{Cn} . It follows by Lemma 6.4 that every extended formulation for P_n^{corr} has at least 2^{Cn} variables and constraints. \square

Fiorini et al. [27] show that, for every n , there exists a graph H with $O(n^2)$ vertices such that its stable set polytope $STAB(H)$ has a face that is an extended formulation for P_n^{cut} . This implies that there exists an infinite family of graphs $\{H_n\}_{n \in \mathbb{N}}$ such that every extended formulation for $STAB(H_n)$ has at least $2^{\sqrt{Cn}}$ variables and constraints. Finally, Yannakakis [61] showed that, for every graph G with p vertices, $STAB(G)$ is the projection of a face of a TSP polytope on $O(p^2)$ vertices. This implies that, for all n , every extended formulation for the TSP polytope on n vertices has at least $2^{(Cn)^{1/4}}$ constraints. This settles a question of Yannakakis [61], who had shown an exponential lower bound on the size of any *symmetric* extended formulations of the TSP, and had asked if a similar result could be proved in general.

Rothvoss [53] showed that there exist 0/1 polytopes that do not admit extended formulations with less than $2^{n/2(1-o(1))}$ constraints and variables. The proof employs a sophisticated rounding technique. The core of the proof is showing that, given a 0/1 polytope $P \subseteq \mathbb{R}^n$ with nonnegative rank t , and given the extended formulation $Ax + Fz = b, y \geq 0$ determined by a factorization $S = FW$ of the slack matrix as in (15), it is possible to select at most $n + t$ independent rows of (A, F, b) so that, by rounding appropriately, one obtains a matrix $(\bar{A}, \bar{F}, \bar{b})$ whose coefficients have encoding size that is polynomial in n and t , and such that a point $x \in \mathbb{R}^n$ belongs to P if and only if there exists $y \geq 0$ such that $\|\bar{A}x + \bar{F}y\|_\infty \leq \frac{1}{4(n+t)}$. Since the number of 0/1 polytopes in \mathbb{R}^n is doubly exponential in n , it follows that t cannot be polynomially bounded by n , otherwise there would not be enough matrices $(\bar{A}, \bar{F}, \bar{b})$ to represent all 0/1 polytopes.

Rothvoss' technique also implies that there are matroid polytopes which do not admit compact extended formulations. This follows from the same argument and the fact that the number of matroids on n elements is doubly exponential in n .

6.4 Compact extended formulations from the slack matrix factorization

Next we show two examples where Yannakakis's theorem can be used to produce a polynomial size extended formulation.

6.4.1 The spanning tree polytope

Given an undirected graph $G = (V, E)$, $|V| = n$, $|E| = m$, the *spanning tree* polytope $SP(G)$ is the convex hull of the incidence vectors of the spanning trees of G . For $S \subseteq V$, let $E(S)$ denote the set of edges with both ends in S . Edmonds [22] showed that

$$SP(G) = \left\{ x \in \mathbb{R}^m : \begin{array}{l} \sum_{e \in E} x_e = n - 1, \\ \sum_{e \in E(S)} x_e \leq |S| - 1, \quad S \subset V \\ x_e \geq 0, \quad e \in E \end{array} \right\}.$$

The slack matrix corresponding to the inequalities $\sum_{e \in E(S)} x_e \leq |S| - 1$ can be described as follows: Given a node set $S \subset V$ and a spanning tree T , the slack of the corresponding constraint is the number of connected components of the forest induced by S on T minus 1. An equivalent description is the following. Given a node $k \in V$, let A_k be the arborescence obtained by rooting T at k . Then one can readily verify that, if $k \in S$, the slack is the number of nodes of $S \setminus \{k\}$ whose father in A_k is not in S . So, if we let $\lambda_{kij} = 1$ if j is the father of i in A_k , 0 otherwise, $i, j, k \in V$, the slack s_{ST} for $S \subset V$ and spanning tree T is

$$s_{ST} = \sum_{i \in S, j \in V \setminus S} \lambda_{kij}.$$

This yields a factorization of the slack matrix $S^V = (s_{ST})$ into nonnegative matrices $S^V = FW$ as follows. For every $S \subset V$, choose an element $k_S \in S$, and let f^S be the vector with $V \times V \times V$ entries, where

$$f_{kij}^S = \begin{cases} 1 & \text{if } k = k_S, i \in S, j \in V \setminus S; \\ 0 & \text{otherwise.} \end{cases}$$

For every tree T , let w^T be the vector with $V \times V \times V$ entries, where $w_{kij}^T = \lambda_{kij}$. Let F be the matrix with rows f^S , $S \subset V$, and W be the matrix with columns w^T for every tree T of G . Then $S^V = FW$.

It can be shown that F and W yield the following extended formulation for the polytope $SP(G)$. Namely $SP(G) = \text{proj}_x(Q)$ where Q is the set of points (x, λ) satisfying

$$\begin{aligned} \sum_{i,j \in V} x_{ij} &= n - 1, \\ x_{ij} &\geq 0, & ij \in E, \\ x_{ij} &= \lambda_{kij} + \lambda_{kji}, & 1 \leq i, j, k \leq n \\ \sum_j \lambda_{kij} &= 1, & 1 \leq i, k \leq n, i \neq k \\ \lambda_{kij} &\geq 0, \lambda_{kkj} = 0, \lambda_{kii} = 0, & 1 \leq i, j, k \leq n. \end{aligned}$$

6.4.2 Stable sets

Given an undirected graph $G = (V, E)$ the stable set polytope $P^{\text{stab}}(G)$ is the convex hull of the incidence vectors of the stable sets of G . Fiorini et al. [27] showed that stable set polytopes do not admit, in general, a compact extended formulation. Next we show that stable sets of t -perfect graphs and comparability graphs always admit compact extended formulations.

Since optimizing a linear function over $P^{\text{stab}}(G)$ is an NP-hard problem [44], a complete external description of $P^{\text{stab}}(G)$ may not be obtainable. However several families of valid inequalities are known. We consider the following inequalities.

$$0 \leq x_v \leq 1 \quad v \in V \quad (17)$$

$$x_{v_i} + x_{v_j} \leq 1 \quad v_i v_j \in E \quad (18)$$

$$x(C) \leq \frac{|C| - 1}{2} \quad C \text{ is an odd cycle of } G \quad (19)$$

$$x(K) \leq 1 \quad K \text{ is a clique of } G \quad (20)$$

Let $P^{\text{oc}}(G) = \{x \in \mathbb{R}^V : x \text{ satisfies (17), (18) and (19)}\}$ and $P^K(G) = \{x \in \mathbb{R}^V : x \text{ satisfies (17) and (20)}\}$ (inequalities (18) are implied by (20) and (17)).

A graph G is *t-perfect* if $P^{\text{stab}}(G) = P^{\text{oc}}(G)$ and G is *perfect* if $P^{\text{stab}}(G) = P^K(G)$. The class of perfect graphs is extremely rich and includes bipartite graphs and comparability graphs among many others.

Consider the polytope

$$Q^{\text{oc}}(G) = \left\{ (x, o, e) \in \mathbb{R}^V \times \mathbb{R}^{V^2} \times \mathbb{R}^{V^2} : \begin{array}{ll} 0 \leq x_{v_i} \leq 1 & v_i \in V \\ 0 \leq o_{ij} \leq 1 - x_{v_i} - x_{v_j} & v_i v_j \in E \\ o_{ij} \leq o_{ik} + e_{kj} & v_i v_k \in E, v_j \in V \\ e_{ij} \leq o_{ik} + o_{kj} & v_i v_k \in E, v_j \in V \\ o_{ii} \geq 1 & v_i \in V \end{array} \right\}.$$

Theorem 6.8. (Yannakakis [61]) *Polytope $P^{\text{oc}}(G)$ is the projection of $Q^{\text{oc}}(G)$ onto the x -space.*

Proof. Assign length $1 - x_{v_i} - x_{v_j}$ to each edge $v_i v_j$ of G . If x satisfies (18), these lengths are nonnegative. Furthermore a vector x violates constraints (19) if and only if G contains an odd cycle C having length < 1 . Define o_{ij} and e_{ij} to be the shortest lengths of an odd and even walk, respectively, between v_i and v_j (multiple traversals of edges or nodes allowed). If no such walk exists, assign a high value to the corresponding variable o_{ij} or e_{ij} . It is easy to see that these definitions of o_{ij} and e_{ij} are consistent with the constraints defining $Q^{\text{oc}}(G)$, except possibly for $o_{ii} \geq 1$.

Furthermore G contains an odd cycle of length < 1 if and only if $o_{ii} < 1$ for some $v_i \in V$, but this contradicts $o_{ii} \geq 1$, $v_i \in V$. Therefore $P^{\text{oc}}(G) \subseteq \text{proj}_x(Q^{\text{oc}}(G))$.

It is easy to see that the constraints defining $Q^{\text{oc}}(G)$ force o_{ij} and e_{ij} to be lower bounds on the shortest lengths of an odd and even path between v_i and v_j respectively. Therefore $P^{\text{oc}}(G) \supseteq \text{proj}_x(Q^{\text{oc}}(G))$. \square

A graph $G = (V, E)$ is a *comparability graph* if there exists a partial order $>$ on $V = \{v_1, \dots, v_n\}$, such that $v_i v_j \in E$ if and only if $v_i > v_j$ or $v_j > v_i$. Cliques of G correspond to chains $v_i > v_j > \dots > v_k$ and stable sets of G correspond to sets of mutually incomparable elements.

It is well known that comparability graphs are perfect (see Schrijver [55] Corollary 66.2a), hence $P^{\text{stab}}(G) = P^K(G)$. We consider the slack matrix S^V of $P^K(G)$ when G is a comparability graph and we disregard constraints (17). Since the rows and columns of S^V correspond

to cliques and stable sets, then the entry of S^V corresponding to clique K and stable set I is 1 if $K \cap I = \emptyset$, 0 otherwise.

Given a clique $K = \{v_1, \dots, v_k\}$, where $v_1 > \dots > v_k$, consider the subsets

$$\{v_1\} \subset V, \{(v_i v_{i+1}), 1 \leq i \leq k-1\} \subset V \times V, \{v_k\} \subset V.$$

Consider the vector in $\{0, 1\}^{|V|} \times \{0, 1\}^{|V|^2} \times \{0, 1\}^{|V|}$ that contains in sequence the incidence vectors of the above three subsets and let f_K be its transpose. Let F be the matrix whose rows are the vectors f_K , for all cliques K of G .

We say that v precedes u if $v > u$. Given a stable set I , consider the subsets

$$\begin{aligned} &\{v \notin I : v \text{ does not precede any node of } I\} \subset V \\ &\{(v_i v_j) : v_i > v_j, v_i \text{ precedes some node of } I, v_j \notin I \text{ and does not precede any node of } I\} \\ &\subset V \times V \\ &\{v : v \text{ precedes some node of } I\} \subset V. \end{aligned}$$

Let w_I be the vector in $\{0, 1\}^{|V|} \times \{0, 1\}^{|V|^2} \times \{0, 1\}^{|V|}$ that contains the incidence vectors of the above three subsets in sequence and let W be the matrix whose columns are the vectors w_I , for all stable sets I of G .

Theorem 6.9. (Yannakakis[61]) *The stable set polytope of a comparability graph $G = (V, E)$ admits an extended formulation of size $o(|V|^2)$.*

Proof. Given clique $K = \{v_1, \dots, v_k\}$ where $v_1 > \dots > v_k$ and stable set I such that $K \cap I = \emptyset$, exactly one of the following alternatives occurs:

1. v_1 does not precede a node of I in the partial order $>$.
2. v_i precedes some node of I while v_{i+1} does not precede a node of I .
3. v_k precedes some node of I .

By construction, this shows that $f_k w_i = 1$ if $K_i \cap I_j = \emptyset$ and $f_k w_i = 0$ otherwise. Therefore $S^V = FW$, and F, W are 0,1-matrices. By Theorem 6.3, this gives a compact extended formulation for $P^{\text{stab}}(G)$ when G is a comparability graph. \square

Yannakakis also gives an extended formulation of size $|V|^{o(\log(|V|))}$ for $P^{\text{stab}}(G)$ when G is perfect. This is subexponential size.

6.5 Matchings

Let $K_n = (V, E)$ be the complete undirected graph on n nodes, where n is even. The *perfect matching polytope* P^{matching} of order n is the convex hull of perfect matchings of K_n . One of the fundamental results in polyhedral combinatorics is Edmonds' perfect matching polytope theorem [20], showing that P^{matching} is the set of solutions of the system

$$\begin{aligned} \sum_{e \in \delta(v)} x_e &= 1 & v \in V \\ \sum_{e \in \delta(U)} x_e &\geq 1 & U \subset V, |U| \text{ odd} \\ x_e &\geq 0 & e \in E. \end{aligned} \tag{21}$$

An outstanding open question is whether there exists a compact extended formulation for the perfect matching polytope. Yannakakis [61] gives a partial negative answer by showing that there is no subexponential size *symmetric* extended formulation for P^{matching} . Next we give the precise statement.

Let $V = \{1, \dots, n\}$. Given a permutation π of V and a vector $x \in \mathbb{R}^E$, let $\pi(x) \in \mathbb{R}^E$ be the vector defined by $(\pi(x))_{ij} = x_{\pi(i)\pi(j)}$, $1 \leq i < j \leq n$.

Let $Ax + By \leq c$ be an extended formulation for P^{matching} , where B is a matrix with q columns. Given a permutation σ of $\{1, \dots, q\}$ and $y \in \mathbb{R}^q$, we denote by $\sigma(y)$ the vector in \mathbb{R}^q defined by $(\sigma(y))_i = y_{\sigma(i)}$. The formulation $Ax + By \leq c$ is *symmetric* if, for every permutation π of V , there exists a permutation σ of $\{1, \dots, q\}$ such that, for every $x \in \mathbb{R}^E$ and $y \in \mathbb{R}^q$, $Ax + By \leq c$ if and only if $A\pi(x) + B\sigma(y) \leq c$.

Theorem 6.10. (Yannakakis [61]) *For every positive even integer n , let $f(n)$ denote the minimum number of variables and constraints of a symmetric extended formulation for the perfect matching polytope of order n . Then $f(n) \in \Omega\left(\frac{n}{4}\right)$.*

Yannakakis [61] shows that an analogous result holds for the *traveling salesman polytope*, that is the convex hull of all Hamiltonian cycles of K_n .

It is unclear how strong the symmetry assumption is. Yannakakis [61] conjectures that Theorem 6.10 should hold even without the symmetry assumption, stating that he feels that “asymmetry does not help much”. As seen in Section 6.3, Yannakakis’ insight is true for the TSP polytope: an exponential lower bound holds even for non-symmetric formulations. However, Kaibel, Pashkovich and Theis [41] gave examples where symmetry can indeed make a huge difference. Consider the family $\mathcal{M}^\ell(n)$ of matchings of K_n with exactly ℓ edges, and let $P_{\text{match}}^\ell(n)$ be the convex hull of incidence vectors of $\mathcal{M}^\ell(n)$. They show that, for $\ell = \lfloor \log n \rfloor$, there is no symmetric compact extended formulation for $P_{\text{match}}^\ell(n)$, while these polytopes have a compact extended formulation (which must therefore be asymmetric).

Theorem 6.11. (Kaibel et al. [41]) *For every n and every $1 \leq \ell \leq n/2$, the size of a symmetric extended formulation for $P_{\text{match}}^\ell(n)$ is at least $\Omega\left(\frac{n}{\lfloor (\ell-1)/2 \rfloor}\right)$.*

In particular, Theorem 6.11 implies that, for $\Omega(\log n) \leq \ell \leq n/2$, there is no compact symmetric extended formulation for $P_{\text{match}}^\ell(n)$. The proof follows Yannakakis’ method for proving Theorem 6.10. Conversely, they show the following.

Theorem 6.12. (Kaibel et al. [41]) *For every n and every $1 \leq \ell \leq n/2$, there exists an extended formulation for $P_{\text{match}}^\ell(n)$ of size bounded by $2^{O(\ell)}n^2 \log n$.*

In particular, Theorem 6.12 implies that, for $\ell \leq \lfloor \log n \rfloor$, there is a compact extended formulation for $P_{\text{match}}^\ell(n)$. The proof is based on the following lemma.

Lemma 6.13. (Alon et al. [2]) *There are maps $\phi_1, \dots, \phi_{q(n,r)} : [n] \rightarrow [r]$ with $q(n,r) \leq 2^{O(r)} \log n$ such that, for every $W \subseteq [n]$ with $|W| = r$, there is some $i \in [q(n,r)]$ for which the map ϕ_i is bijective on W .*

To prove Theorem 6.12, let $\phi_1, \dots, \phi_{q(n,r)}$ be as in Lemma 6.13, with $r = 2\ell$, and let $q = q(n,r)$. Let $\mathcal{M}_i = \{M \in \mathcal{M}^\ell(n) : \phi_i \text{ is bijective on } V(M)\}$ (where $V(M)$ denotes the

set of vertices covered by M), and let P_i be the convex hull of incidence vectors of \mathcal{M}_i . By the choice of $\phi_1, \dots, \phi_{q(n,r)}$, $\mathcal{M}^\ell(n) = \bigcup_{i=1}^q \mathcal{M}_i$, and thus $P_{\text{match}}^\ell(n) = \text{conv}(P_1 \cup \dots \cup P_q)$. Kaibel et al. then show that

$$P_i = \left\{ x \in \mathbb{R}_+^E : \begin{array}{ll} x_e = 0 & e \in E \setminus E_i \\ x(\delta(\phi^{-1}(t))) = 1 & t = 1, \dots, 2\ell \\ x(E_i(\phi^{-1}(S))) \leq (|S| - 1)/2 & S \subseteq [2\ell], |S| \text{ odd} \end{array} \right\},$$

where $E_i = E \setminus \bigcup_{j=1}^{2\ell} E(\phi_i^{-1}(j))$. We do not report the proof of this latter fact here.

Finally, since $P_{\text{match}}^\ell(n) = \text{conv}(P_1 \cup \dots \cup P_q)$, we may apply Balas' union of polyhedra (Theorem 5.1) to obtain an extended formulation for $P_{\text{match}}^\ell(n)$. Note that the number of variables and inequalities in the description of P_i is bounded by $2^{O(\ell)} + n^2$, thus the extended formulation given by Theorem 5.1 has at most $2^{O(\ell)} n^2 \log n$ variables and constraints.

Kaibel et al. [41] apply similar techniques to show that there is no compact symmetric extended formulation for the polytope of cycles with $\lfloor \log n \rfloor$ edges (i.e. the convex hull of incidence vectors of cycles of length $\lfloor \log n \rfloor$ in K_n), while there are compact extended formulations that are not symmetric.

The proof that the cut polytope (and hence the TSP polytope) do not admit an extended formulation, is based on the rectangle covering bound. The rectangle covering bound is, however, very weak for the matching polytope, as observed by Yannakakis [61]. Indeed, it is not difficult to find a rectangle cover with less than n^4 elements for the slack matrix of the system (21). The cover has a rectangle for every pair of edges, where the rows are defined by the odd sets U such that $\delta(U)$ contains both edges, and the columns are defined by the matchings that contain both edges.

6.6 Matchings in planar graphs

While Yannakakis [61] shows that no symmetric subexponential formulation for the matching polytope exists, Barahona [8] describes a compact formulation for the perfect matching polytope of planar graphs.

Let $G = (V, E)$ be an undirected graph and $T \subseteq V$ a set of even cardinality. A set $J \subseteq E$ is a T -join if T is the set of nodes of odd degree of the graph (V, J) . The T -join polytope $P^{TJ}(G)$ is the convex hull of incidence vectors of T -joins. An external description of the T -join polytope in the original space of variables is known (see, e.g., [55] Corollary 29.2e p. 491).

Note that the perfect matchings of G are the V -joins of cardinality $|V|/2$. Furthermore, any V -join of G has cardinality at least $|V|/2$. Therefore the perfect matching polytope of G is the face of $P^{VJ}(G)$ defined by the hyperplane $\{x \in \mathbb{R}^E : \sum_{e \in E} x_e = |V|/2\}$.

We describe Barahona's compact extended formulation for $P^{VJ}(G)$ when G is a planar graph.

A well-known fact about T -joins is that the symmetric difference of two T -joins is a \emptyset -join (that is, a disjoint union of cycles) and the symmetric difference of a T -join and a \emptyset -join is a T -join. This implies that, given a V -join J^* of G , any other V -join is the symmetric

difference of J^* and some \emptyset -join. Thus

$$P^{VJ}(G) = \left\{ y \in \mathbb{R}^E : \exists x \in P^{\emptyset J}(G) \text{ s.t. } \begin{array}{ll} y_e = x_e, & e \in E \setminus J^* \\ y_e = 1 - x_e, & e \in J^* \end{array} \right\}.$$

Thus a compact extended formulation for $P^{\emptyset J}(G)$ provides a compact extended formulation for $P^{VJ}(G)$. Edmonds and Johnson [24] showed that, for any graph G (not necessarily planar),

$$P^{\emptyset J}(G) = \left\{ x \in \mathbb{R}^E : \begin{array}{ll} x(F) - x(\delta(S) \setminus F) \leq |F| - 1, & S \subset V, F \subseteq \delta(S), |F| \text{ odd} \\ 0 \leq x_e \leq 1, & e \in E \end{array} \right\}. \quad (22)$$

Note that, in the above description, an inequality associated with a set $S \subset V$ is irredundant only if $\delta(S)$ is a minimal nonempty cut. That is, the graphs induced by S and $V \setminus S$ are both connected.

If G is planar, let $G^* = (V^*, E^*)$ be the dual of a plane representation of G . For each edge $e \in E$, let $\phi(e)$ be the edge of G^* joining the (possibly identical) nodes representing the two faces of the plane representation of G having e on their boundaries. Given $S \subset V$, the set $\{\phi(e) : e \in \delta(S)\}$ is an \emptyset -join of G^* , and viceversa, given an \emptyset -join J of G , $\{\phi(e) : e \in J\}$ is a cut of G^* . In particular, simple cycles of G^* are in one-to-one correspondence with minimal nonempty cuts of G .

For every $x \in \mathbb{R}^E$, we denote by $\phi(x)$ the vector in \mathbb{R}^{E^*} whose component indexed by $\phi(e)$ is x_e , for all $e \in E$. Therefore

$$P^{\emptyset J}(G) = \{x \in \mathbb{R}^E : \phi(x) \in P^{\text{cut}}(G^*)\}, \quad (23)$$

where $P^{\text{cut}}(G^*)$ denotes the cut polytope of G^* . Recall that the polytope $R(G^*)$, defined in Section 2.3, is the following:

$$\left\{ x \in \mathbb{R}^{E^*} : \begin{array}{ll} x(F) - x(C \setminus F) \leq |F| - 1, & C \in \mathcal{C}(G^*), F \subseteq C, |F| \text{ odd} \\ 0 \leq x_e \leq 1, & e \in E^* \end{array} \right\}$$

where $\mathcal{C}(G^*)$ denotes the family of cycles of G^* . By (22) and (23), we have that $P^{\text{cut}}(G^*) = R(G^*)$.

In Section 2.3 we described a compact extended formulation for $R(G^*)$. By the above discussion, this gives a compact extended formulation for the matching polytope of planar graphs.

Gerards [31] gives a compact extended formulation for the matching polytope in graphs with bounded genus.

6.7 The Permutahedron

The permutahedron $\Pi_n \subseteq \mathbb{R}^n$ is the convex hull of all vectors that can be obtained by permuting the coordinates of the vector $(1, 2, \dots, n)$. It is well known (see Ziegler [62] or Goemans [33]) that Π_n is the set of all $x \in \mathbb{R}^n$ satisfying

$$\begin{aligned} \sum_{i=1}^n x_i &= \binom{n+1}{2} \\ \sum_{i \in S} x_i &\geq \binom{|S|+1}{2} \quad S \subset \{1, \dots, n\}. \end{aligned}$$

It can be shown that the above is an irredundant external description of Π_n .

Goemans [33] shows that the size of a smallest extended formulation for Π_n is of order $\Theta(n \log n)$. We report here his elegant geometric argument showing the lower bound.

Proposition 6.14. *Any extended formulation for Π_n has at least $\log(n!)$ facets.*

Proof. Let Q be a polyhedron in $\mathbb{R}^n \times \mathbb{R}^q$ whose projection onto the space $\mathbb{R}^n \times \{0\}$ is Π_n , and let f be the number of facets of Q . Since each vertex of Π_n is the projection of some face of Q , the number of faces of Q is at least the number of vertices of Π_n , that is $n!$. Since each face of Q is uniquely determined as the intersection of a subset of the facets of Q , Q has at most 2^f faces. Hence $n! \leq 2^f$. \square

The above shows that any extended formulation must have $\Omega(n \log(n))$ constraints. Goemans [33] also gives an extended formulation for the permutahedron with $O(n \log(n))$ variables and constraints. The extended formulation is based on sorting networks. We recall that a *comparison network* is comprised of *wires* and *comparators*. Each wire lies on a horizontal line between an input and a output wire. There are n such lines, thus there are n input wires a_1, \dots, a_n and n output wires b_1, \dots, b_n . Each wire x carries a value $v(x)$. Each comparator is depicted as a vertical line and consists of two input wires x', x'' and two output wires y', y'' , and $v(y') = \min\{v(x'), v(x'')\}$ while $v(y'') = \max\{v(x'), v(x'')\}$. Note that the number of wires is n plus twice the number of comparators.

A *sorting network* is a comparison network such that, for any possible choice of $v(a_1), \dots, v(a_n)$, we have $v(b_1) \leq v(b_2) \leq \dots \leq v(b_n)$. We refer the reader to [19] for an introduction to comparison and sorting networks. Figure 1 depicts a sorting network for $n = 4$.

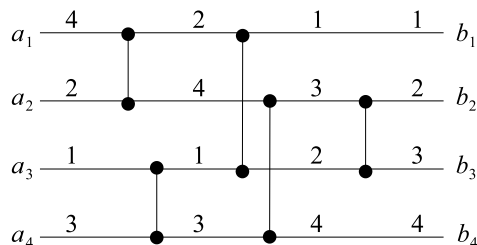


Figure 1: A sorting network with 4 lines, 5 comparators and 14 wires.

Given a comparison network N with n lines, we say that a permutation π of $\{1, \dots, n\}$ is *feasible* for N if, once we set $v(a_i) = \pi_i$ for $i = 1, \dots, n$, then $v(b_i) = i$ for $i = 1, \dots, n$. In particular, N is a sorting network if and only if all permutations of $\{1, \dots, n\}$ are feasible for N .

Let $P(N)$ be the convex hull of the vectors that are a feasible permutation for N of the components of the vector $(1, 2, \dots, n)$. When N is a sorting network, $P(N) = \Pi_n$.

We give an extended formulation for $P(N)$. The extended formulation has a variable for each wire and 5 constraints for each comparator. Ajtai, Komlós, and Szemerédi [1] show that there exist sorting networks with n input wires with $O(n \log n)$ comparators. This gives an extended formulation for Π_n with $O(n \log n)$ variables and constraints.

Let w_{i1}, \dots, w_{ih_i} be the wires in line i . Wires w_{i1} are the input wires, while wires w_{ih_i} are the output wires. We denote each comparator as a pair $\{(i, j), (i', j')\}$ whose input wires are w_{ij} and $w_{i'j'}$ and whose output wires are $w_{i,j+1}$ and $w_{i',j'+1}$.

For each wire w_{ij} , we have an auxiliary variable y_{ij} , $1 \leq i \leq n$, $1 \leq j \leq h_i$. For $i = 1, \dots, n$, we set

$$y_{ih_i} = i. \quad (24)$$

For each comparator $\{(i, j), (i', j')\}$, where $i < i'$, we have the following 5 constraints. (Only the first 3 are necessary).

$$\begin{aligned} y_{ij} + y_{i'j'} &= y_{i,j+1} + y_{i',j'+1} \\ y_{ij} &\geq y_{i,j+1} \\ y_{i'j'} &\geq y_{i',j'+1} \\ y_{ij} &\leq y_{i',j'+1} \\ y_{i'j'} &\leq y_{i,j+1}. \end{aligned} \quad (25)$$

Note that, if π is a feasible permutation for N , and we set $v(a_i) = \pi_i$, $i = 1, \dots, n$, then the solution defined by $y_{ij} = v(w_{ij})$, $1 \leq i \leq n$, $1 \leq j \leq h_i$ is feasible for the system defined by (24) for each i and (25) for each comparator $\{(i, j), (i', j')\}$, where $i < i'$.

Theorem 6.15. *Let N be a comparison network and $Q(N)$ be the polyhedron defined by (24) for $i = 1, \dots, n$ and (25) for each comparator $\{(i, j), (i', j')\}$, where $i < i'$. The polytope $P(N)$ is the projection of $Q(N)$ onto the space of variables y_{11}, \dots, y_{n1} . In particular, if N is a sorting network, Π_n is the projection of $Q(N)$ onto the space of variables y_{11}, \dots, y_{n1} .*

Proof. Let \bar{y} be a vertex of $Q(N)$. We will show that $(\bar{y}_{11}, \dots, \bar{y}_{n1})$ is a feasible permutation for N . We prove this by induction on the number c of comparators. The result is trivial if $c = 0$, since the identity is the unique feasible permutation and (24) implies $\bar{y}_{i1} = i$.

Let us assume that $c \geq 1$. By definition of a comparison network (since comparators correspond to vertical lines), there exists a comparator of the form $\{(h, 1)(k, 1)\}$, $h < k$. We first observe that either $\bar{y}_{h1} = \bar{y}_{h2}$ and $\bar{y}_{k1} = \bar{y}_{k2}$, or $\bar{y}_{h1} = \bar{y}_{k2}$ and $\bar{y}_{k1} = \bar{y}_{h2}$. If not, then $\bar{y}_{h1} > \bar{y}_{h2}$, $\bar{y}_{k1} < \bar{y}_{k2}$, $\bar{y}_{h1} < \bar{y}_{k2}$, and $\bar{y}_{k1} > \bar{y}_{h2}$, because $\bar{y}_{h1} + \bar{y}_{k1} = \bar{y}_{h2} + \bar{y}_{k2}$. Then, for $\varepsilon > 0$ small enough, the vectors y' and y'' defined next are in $Q(N)$ and $\bar{y} = \frac{1}{2}(y' + y'')$, contradicting that \bar{y} is a vertex of $Q(N)$:

$$y'_{ij} = \begin{cases} \bar{y}_{h1} - \varepsilon & \text{if } i = h, j = 1 \\ \bar{y}_{k1} + \varepsilon & \text{if } i = k, j = 1 \\ \bar{y}_{ij} & \text{otherwise.} \end{cases} \quad y''_{ij} = \begin{cases} \bar{y}_{h1} + \varepsilon & \text{if } i = h, j = 1 \\ \bar{y}_{k1} - \varepsilon & \text{if } i = k, j = 1 \\ \bar{y}_{ij} & \text{otherwise.} \end{cases}$$

Therefore, $\bar{y}_{h2} = \min\{\bar{y}_{h1}, \bar{y}_{k1}\}$ and $\bar{y}_{k2} = \max\{\bar{y}_{h1}, \bar{y}_{k1}\}$.

Let N' be the comparison network obtained from N by deleting the comparator $\{(h, 1)(k, 1)\}$ and wires w_{h1} and w_{k1} . Note that wires w_{h2} and w_{k2} become input wires of N' .

Let \tilde{y} be the vector obtained from \bar{y} by removing the components indexed by $(h, 1)$ and $(k, 1)$. Clearly $\tilde{y} \in Q(N')$. We show that \tilde{y} is a vertex of $Q(N')$. Suppose not, then there exist two points y' and y'' of $Q(N')$ distinct from \tilde{y} such that $\tilde{y} = \frac{1}{2}(y' + y'')$. Let \bar{y}' be

defined by $\bar{y}'_{ij} = y'_{ij}$ for every $(i, j) \neq (h, 1), (k, 1)$, $\bar{y}'_{h1} = y'_{h2}$ and $\bar{y}'_{k1} = y'_{k2}$ if $\bar{y}_{h1} = \bar{y}_{h2}$ and $\bar{y}_{k1} = \bar{y}_{k2}$, $\bar{y}'_{h1} = y'_{k2}$ and $\bar{y}'_{k1} = y'_{h2}$ if $\bar{y}_{h1} = \bar{y}_{k2}$ and $\bar{y}_{k1} = \bar{y}_{h2}$. If we define \bar{y}'' analogously, then $\bar{y} = \frac{1}{2}(\bar{y}' + \bar{y}'')$, contradicting the fact that \bar{y} is a vertex of $Q(N)$. Thus, by induction, $(\tilde{y}_{11}, \dots, \tilde{y}_{h-1,1}, \tilde{y}_{h2}, \tilde{y}_{h+1,1}, \dots, \tilde{y}_{k-1,1}, \tilde{y}_{k2}, \tilde{y}_{k+1,1}, \dots, \tilde{y}_{n1})$ is a feasible permutation for N' . Therefore $(\bar{y}_{11}, \dots, \bar{y}_{n1})$ is a feasible permutation for N . \square

A $O(n^2)$ extended formulation for Π_n is much simpler to describe. We introduce auxiliary binary variables δ_{ij} , $i = 1, \dots, n$, $j = 1, \dots, n$, such that $\delta_{ij} = 1$ if and only if $x_i = j$. Let Q_n be the polyhedron containing all points $(x, \delta) \in \mathbb{R}^n \times \mathbb{R}^{n \times n}$ satisfying the following system:

$$\begin{aligned} x_i &= \sum_{j=1}^n j \delta_{ij}, \quad 1 \leq i \leq n \\ \sum_{j=1}^n \delta_{ij} &= 1, \quad 1 \leq i \leq n \\ \sum_{i=1}^n \delta_{ij} &= 1, \quad 1 \leq j \leq n \\ \delta_{ij} &\geq 0, \quad 1 \leq i, j \leq n. \end{aligned} \tag{26}$$

Theorem 6.16. *The polytope Π_n is the projection of Q_n in the x -space.*

Proof. The first set of equations in the system (26) defines x . Let $A\delta \geq b$ be the system comprising all inequalities not involving x . The matrix A is the node-edge incidence matrix of a bipartite graph thus it is totally unimodular. Since b is an integral vector, the polytope Q_n is integral. This proves the theorem. \square

Kaibel and Pashkovic [42] generalise the above result to the convex hull of the union of the orbits of a given polytope under a finite reflection group. More formally, consider a finite reflection group G defined by hyperplanes H_1, \dots, H_k in \mathbb{R}^n passing through the origin. Given a polytope $P \in \mathbb{R}^n$, the G -permutahedron of P is the convex hull of the orbits of P under the action of G . Note that the permutahedron is the G -permutahedron of the 0-dimensional polytope $\{(1, 2, \dots, n)\}$, where G is the reflection group defined by the hyperplanes of the form $\{x \in \mathbb{R}^n : x_i - x_j = 0\}$, $1 \leq i < j \leq n$.

Kaibel and Pashkovich [42] show that, if P is contained in the closure of one of the connected regions of $\mathbb{R}^n \setminus (H_1 \cup \dots \cup H_k)$, there exists an extended formulation with $f + O(n \log m) + O(n \log n)$ variables and constraints, where f is the number of facets of P and m is a parameter depending on G , which we do not discuss here. When specialised to the permutahedron, the result gives a $O(n \log n)$ extended formulation. Note that any regular m -gon is the G -permutahedron of a point in \mathbb{R}^2 , where G is the reflection group defined by m lines passing through the origin and subdividing \mathbb{R}^2 into wedges of angle $2\pi/m$. For regular m -gons, Kaibel and Pashkovich's extended formulation matches the best possible bound of $O(\log m)$ of Ben-Tal and Nemirovski [10].

7 Dynamic Programming

It is sometimes the case where a polynomial time algorithm to solve a combinatorial optimization problem also offers a compact extended formulation for the corresponding polytope. A wide class of algorithms for which this is often the case is that of dynamic programming algorithms. Martin, Rardin and Campbell [46] provided sufficient conditions on the general

class of such algorithms that guarantee to yield extended formulations. A generalisation is given by Kaibel and Loos [40] within the framework of “branched polyhedral systems”, which encompass both the results of Martin et al. on dynamic programming [46] and of Balas [4, 5] on union of polyhedra. Below we give several examples of extended formulations derived from dynamic programming algorithms.

7.1 The knapsack problem

The 0–1 knapsack set $K(n, b)$ with n items and capacity b has been introduced in Section 5.3. The knapsack problem is $\max\{\sum_{i=1}^n c_i x_i : x \in K(n, b)\}$.

For $k \leq n$ and $0 \leq b' \leq b$, let $f(k, b')$ be the optimal value of the knapsack problem over the set $K(k, b')$ and using the first k items: Then $f(k, b')$ can be computed using the following recursion:

$$f(k, b') = \max\{f(k-1, b'), f(k-1, b'-a_k) + c_k\}.$$

Consider the acyclic network $N(V, A)$, where the vertices $v_{k,b'}$ represent all possible states $f(k, b')$, $0 \leq k \leq n$, $0 \leq b' \leq b$ and the edges are $v_{k-1,b'}v_{k,b'}$ having weight 0 and $v_{k-1,b'-a_k}v_{k,b'}$ having weight c_k . We call A_k the set of arcs with weight c_k . The computation of $f(n, b)$ by the above recursion is equivalent to the computation of a path P of minimum weight from $v_{0,0}$ to $v_{n,b}$ in the network $N(V, A)$. Note that $x_k = 1$ if $P \cap A_k \neq \emptyset$. Consider the polyhedron Q defined by the following inequalities:

$$\begin{aligned} x_k &= \sum_{a \in A_k} z_a, \quad 1 \leq k \leq n \\ \sum_{a \in \delta^-(v_{0,0})} z_a - \sum_{a \in \delta^+(v_{0,0})} z_a &= -1 \\ \sum_{a \in \delta^-(v_{n,b})} z_a - \sum_{a \in \delta^+(v_{n,b})} z_a &= 1 \\ \sum_{a \in \delta^-(v_{k,b'})} z_a - \sum_{a \in \delta^+(v_{k,b'})} z_a &= 0, \quad v_{k,b'} \in V \setminus \{v_{0,0}, v_{n,b}\} \\ z_a &\geq 0, \quad a \in A. \end{aligned}$$

Theorem 7.1. *Let $K(n, b) = \{x \in \{0, 1\}^n : \sum_{i=1}^n a_i x_i \leq b\}$. Then $\text{conv}(K(n, b)) = \text{proj}_x(Q)$.*

Proof. In the system of inequalities defining Q , disregard the first set of equations, that define x_k . The remaining system is the path polytope, since $D = (V, A)$ does not contain directed cycles. \square

7.2 Stable sets in Distance Claw-free graphs

The results in this section appear in [52]. Given an undirected graph $G = (V, E)$, we denote with $\alpha(G)$ the *stability number* of G , i.e., the maximum number of pairwise nonadjacent nodes of G . Given $v \in V$, let $N_i(v)$ be the set of nodes of G at distance *exactly* i from v and $G_i(v)$ the subgraph of G induced by $N_i(v)$. A graph G is *claw-free* if $\alpha(G_1(v)) \leq 2$ for every $v \in V$. This is equivalent to the fact that G does not contain a claw as induced subgraph. G is *distance claw-free* if $\alpha(G_1(v)) \leq 2$ and $\alpha(G_2(v)) \leq 2$ for every $v \in V$.

Lemma 7.2. *If G is a distance claw-free graph, then $\alpha(G_i(v)) \leq 2$ for every $v \in V$ and distance i .*

Proof. Assume not and choose v such that for a minimum i , $G_i(v)$ contains three pairwise nonadjacent nodes x_1, x_2, x_3 . Since G is distance claw-free, $i > 2$. So no pair x_i, x_j has a common neighbor t , in $N_{i-1}(v)$, else t would be the center of a claw. Let $y_1, y_2, y_3 \in N_{i-1}(v)$ be neighbors of x_1, x_2, x_3 respectively. As just observed, the nodes y_1, y_2, y_3 are distinct. By minimality of i , we may assume that $y_1 y_2 \in E$ and let $z \in N_{i-2}(v)$ be a neighbor of y_1 . Now $z y_2 \in E$, else y_1 is the center of a claw. This shows that z and y_3 are nonadjacent, else $\{x_1, x_2, x_3\} \in N_2(z)$, a contradiction to G being distance claw-free. This shows that y_3 is not adjacent to y_1 or y_2 , else y_1 or y_2 would be the center of a claw.

Let P be a shortest path between z and y_3 whose intermediate nodes lie in $\bigcup_{0 \leq k \leq i-2} N_k(v)$ and let w be the node at distance 2 from z in P . Then $\{x_1, x_2, w\} \in N_2(z)$, again a contradiction to G being distance claw-free. \square

We now introduce a Dynamic Programming recursion to compute a maximum weight stable set in a (general) graph G and use it to derive an extended formulation for the stable set polytope of G . If G is distance claw-free, we use the above lemma to show that the recursion runs in polynomial time and the associated extended formulation is compact.

Given a connected graph $G = (V, E)$ and $v \in V$, let d be the largest value i for which $N_i(v) \neq \emptyset$. Let \mathcal{S} the family of the stable sets of G and \mathcal{S}_i the family of stable sets of $G_i(v)$. Consider a directed acyclic graph D whose nodes are $\{v_S^i : S \in \mathcal{S}_i, 0 \leq i \leq d\}$ together with s and t .

The arcs are:

- sv_\emptyset^0 and sv_v^0
- $v_{S \cap N_i(v)}^i v_{S \cap N_{i+1}(v)}^{i+1}$ for each stable set S contained in $N_i(v) \cup N_{i+1}(v)$.
- $v_S^d t$ for each stable set S in \mathcal{S}_d .

The weight of an arc is the weight of the stable set represented by the head of the arc. It is now clear that there is a bijection between the family \mathcal{S} of stable sets in G and the set of directed paths from s to t in D that preserves the weight.

If G contains a node v such that $\alpha(G_i(v)) \leq k, 1 \leq i \leq d$, the size of the digraph D is polynomial in the size of G and therefore a stable set of maximum weight can be efficiently computed by shortest path techniques.

This is the case if G is distance claw-free since, by Lemma 7.2 for every node v , $\alpha(G_i(v)) \leq 2, 1 \leq i \leq d$.

An extended formulation can be derived in a manner that is similar to the one introduced for the knapsack problem: $x_u = 1$ whenever the optimal stable set contains node u . Define A_u to be the subset of arcs of D whose head represents a stable set containing u . This happens whenever the maximum weight path contains an arc $a \in A_u$. So by adding the equations $x_u = \sum_{a \in A_u} z_a, u \in V(G)$ to the path formulation one obtains an extended formulation. This formulation is compact whenever the size of D is polynomial in the size of G and this is the case for distance claw-free graphs.

Giles and Trotter [32] show that if G is a distance claw-free graph, the external description of the stable set polytope of G in the original space contains complicated inequalities.

No compact extended formulation is known in general for the stable set polytope of claw-free graphs. Finding one such formulation, or proving that none exists, is probably a very tall task, considering that the stable set polytope of claw free graphs encompasses the matching polytope of general graphs, as the line graph of any given graph is claw-free, and stable sets in the line graph correspond to matchings in the given graph. On the other hand, Faenza, Oriolo and Stauffer [26] provide an extended formulation for the stable set polytope of claw-free graphs that, albeit of exponential size in general, provides an efficient separation routine in the original space of variables.

7.3 Packing and partitioning orbitopes

Let m, n be two positive integers. The *packing orbitope* $O_{m,n}^{\leq}$ is the convex hull of all $m \times n$ 0, 1-matrices with at most a nonzero element in each row and whose columns are in lexicographic decreasing order. The *partitioning orbitope* $O_{m,n}^{\equiv}$ is the convex hull of all $m \times n$ 0, 1-matrices with exactly one nonzero element in each row and whose columns are in lexicographic decreasing order.

Packing and partitioning orbitopes have been introduced and studied by Kaibel and Pfetsch [43] who characterized the facet inducing inequalities. These inequalities can be added in the formulation of certain combinatorial optimization problems in order to break symmetries. An example is the formulation for graph coloring with binary variables x_{ij} assuming value 1 if and only if color j is assigned to node i .

Faenza and Kaibel [25] gave a compact extended formulation for packing and partitioning orbitopes. This formulation is based on a dynamic programming algorithm to maximize a linear function over the packing or partitioning orbitope. Here we restrict ourselves to packing orbitopes, since O_{mn}^{\equiv} is a face of O_{mn}^{\leq} .

We first illustrate this algorithm, which reduces the problem to a longest path computation in an auxiliary acyclic digraph. Given a cost matrix $c \in \mathbb{R}^{mn}$, we want to find a 0, 1 matrix $x \in \mathbb{R}^{mn}$ maximizing $cx = \sum_{i=1}^m \sum_{j=1}^n c_{ij}x_{ij}$ such that x has at most a 1 in each row and its columns are in lexicographic decreasing order.

We construct a digraph $D = (V, A)$ where $V = \{(i, j) : 0 \leq i \leq m, 0 \leq j \leq n\}$ and A is the union of the set A^{\downarrow} containing an arc $(i, j)^{\downarrow} = ((i, j), (i + 1, j))$ for every $0 \leq i \leq m - 1, 0 \leq j \leq n$, and the set A^{\searrow} containing an arc $(i, j)^{\searrow} = ((i, j), (i + 1, j + 1))$ for every $0 \leq i \leq m - 1, 0 \leq j \leq n - 1$. Note that D is acyclic. We assign lengths to the arcs as follows: $\ell_{(ij)^{\downarrow}} = \max\{0, c_{(i+1)1}, \dots, c_{(i+1)j}\}$ and $\ell_{(ij)^{\searrow}} = c_{(i+1)(j+1)}$.

Next we show that, the length of a longest path in D from $(0, 0)$ to $\{(m, 0), \dots, (m, n)\}$ is equal to the maximum value of cx , and that an optimal matrix x can be constructed from such a longest path. Since D is acyclic, a longest path can be computed in polynomial time.

Let P be the longest path from $(0, 0)$ to the set of nodes $\{(m, 0), \dots, (m, n)\}$. We construct an $m \times n$ 0, 1-matrix x as follows. For every arc of the form $(i - 1, j - 1)^{\searrow} \in P$, we set $x_{ij} = 1$. For every arc $(i - 1, j)^{\downarrow} \in P$, let $k = \arg \max\{c_{ih} : 1 \leq h \leq j\}$; if $c_{ik} > 0$, we set $x_{ik} = 1$. All other entries of x are set to 0. Note that, by construction of D , ℓ and x , x is a 0, 1 matrix with at most a 1 per row and whose columns are in lexicographic decreasing order. Furthermore, $cx = \sum_{e \in P} \ell_e$.

Viceversa, given a 0, 1 matrix x with at most a 1 per row and whose columns are in lexicographic decreasing order, we construct a path P in D from $(0, 0)$ to $\{(m, 0), \dots, (m, n)\}$

as follows. For every $1 \leq i \leq m$ and $1 \leq j \leq n$ such that $x_{ij} = 1$ and $x_{hj} = 0$ for every $1 \leq h \leq i-1$, P contains the arc $(i-1, j-1) \searrow$. All other arcs in P are in A^\downarrow . By construction of D , since the columns of x are in lexicographic decreasing order, there exists a unique such path P . By construction $cx \leq \sum_{e \in P} \ell_e$.

The extended formulation proposed by Faenza and Kaibel [25] is as follows. Let $F \subset \mathbb{R}^A$ be the convex hull of the incidence vectors of paths in D from $(0, 0)$ to $\{(m, 0), \dots, (m, n)\}$. Note that, since D is acyclic, F is the set of unit flows from $(0, 0)$ to $\{(m, 0), \dots, (m, n)\}$. Thus F is described by $m(n+1) - 1$ flow-conservation equations, $|A|$ nonnegativity bounds and by the equation $y_{(0,0)^\downarrow} + y_{(0,0)^\searrow} = 1$.

The formulations has a variable x_{ij} for $0 \leq i \leq m$ and $0 \leq j \leq n$ and a variable y_e for every $e \in A$. The variables x_{ij} , $1 \leq i \leq m$, $1 \leq j \leq n$, are the original variables. Variables x_{0j} , $j = 0, \dots, n$, x_{i0} , $i = 1, \dots, m$, are dummy variables, while the variables y_e , $e \in A$, are 0, 1-variables representing paths from $(0, 0)$ to $\{(m, 0), \dots, (m, n)\}$. Given (i, j) , $0 \leq i \leq m$, $0 \leq j \leq n$, let $S_{ij} = \{(i, h) : j \leq h \leq n\}$. The constraints are

$$\begin{aligned} x_{00} = 1, x_{0j} = 0 & \quad j = 1, \dots, n \\ x_{ij} \geq y_{(i-1, j-1)^\searrow} & \quad i = 1, \dots, m, j = 1, \dots, n, \\ \sum_{(i, h) \in S_{ij}} x_{ih} \leq \sum_{e \in \delta^-(S_{ij})} y_e & \quad i = 1, \dots, m, j = 0, \dots, n, \\ y \in F & \\ x \geq 0. & \end{aligned} \tag{27}$$

$$\tag{28}$$

Constraints (27) ensure that, whenever $(i-1, j-1) \searrow$ is in the path defined by y , $x_{ij} = 1$. Constraints (28) ensure that, for given (i, j) , whenever the path defined by y does not contain any node (i, h) , $j \leq h \leq n$, $x_{ih} = 0$ for all $h \geq j$. Furthermore, since $y \in F$, by construction of D we have $\sum_{e \in \delta^-(S_{i0})} y_e = 1$, hence constraints (28) relative to $(i, 0)$, $i = 1 \dots, m$ become $\sum_{h=0}^n x_{ih} \leq 1$.

It follows from the construction of $D = (V, A)$ and from the previous discussion that a 0, 1-vector $(x_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ is in O_{mn}^{\leq} if and only if it is the projection of a 0, 1-vector $(x, y) \in \mathbb{R}^{(m+1)(n+1)} \times \mathbb{R}^A$ satisfying the above constraints.

Theorem 7.3. (Faenza and Kaibel [25]) *Let $Q \subset \mathbb{R}^{(n+1)(m+1)} \times \mathbb{R}^A$ be the polyhedron defined by the above constraints. Then O_{mn}^{\leq} is the projection of Q onto the space of the variables x_{ij} , $1 \leq i \leq m$, $1 \leq j \leq n$.*

Proof. We will show that Q is integral. It suffices to prove that, given any vector $(c, d) \in \mathbb{R}^{(m+1)(n+1)} \times \mathbb{R}^A$, the problem $\max\{cx + dy : (x, y) \in Q\}$ has an optimal solution that is integral. Since $x_{00} = 1$, we may assume $c_{00} = 0$. We define two vectors c^\downarrow and c^\searrow in \mathbb{R}^A that are zero in all components except:

$$\begin{aligned} c_{(i-1, j-1)^\searrow}^\searrow &= c_{ij} & 1 \leq i \leq m, 1 \leq j \leq n, \\ c_{(i-1, j)^\downarrow}^\downarrow &= \max\{0, c_{i1}, \dots, c_{ij}\} & 1 \leq i \leq m, 0 \leq j \leq n. \end{aligned}$$

We show that the following hold:

- i) For every $(x, y) \in Q$, $cx + dy \leq (d + c^{\searrow} + c^{\downarrow})y$;
- ii) For every integral $y \in F$, there exists $x \in \{0, 1\}^{(m+1)(n+1)}$ such that $(x, y) \in Q$ and $cx + dy = (d + c^{\searrow} + c^{\downarrow})y$.

The above two properties imply the theorem. Indeed, since F is an integral polyhedron, there exists an optimal integral solution y^* for the problem $\max\{(d + c^{\searrow} + c^{\downarrow})y : y \in F\}$. By i), for every $(x, y) \in Q$, $cx + dy \leq (d + c^{\searrow} + c^{\downarrow})y \leq (d + c^{\searrow} + c^{\downarrow})y^*$. By ii), there exists $x^* \in \{0, 1\}^{(m+1)(n+1)}$ such that $(x^*, y^*) \in Q$ and $cx^* + dy^* = (d + c^{\searrow} + c^{\downarrow})y^*$. Thus (x^*, y^*) is an optimal integral solution for $\max\{cx + dy : (x, y) \in Q\}$.

We show ii). Let $y \in F$. We define $x \in \{0, 1\}^{(m+1)(n+1)}$ such that x is zero everywhere except:

- $x_{00} = 1$,
- For every (i, j) such that $y_{(i-1, j-1)\searrow} = 1$, $x_{ij} = 1$ ($1 \leq i \leq m$, $1 \leq j \leq n$),
- For every (i, j) such that $y_{(i-1, j)\downarrow} = 1$, $x_{ik} = 1$ where $k = \arg \max\{0, c_{i1}, \dots, c_{ij}\}$ ($1 \leq i \leq m$, $0 \leq j \leq n$).

One can readily verify that $(x, y) \in Q$ and $cx + dy = (d + c^{\searrow} + c^{\downarrow})y$.

We now prove i). Define $x' \in \mathbb{R}^{(m+1)(n+1)}$ by $x'_{ij} = x_{ij} - y_{(i-1, j-1)\searrow}$, $1 \leq i \leq m$, $1 \leq j \leq n$, and $x'_{ij} = x_{ij}$ if $i = 0$ or $j = 0$. By (27), $x' \geq 0$. By construction, $c(x - x') = c^{\searrow}y$.

It suffices to show $c^{\downarrow}y \geq cx'$, since this implies $(d + c^{\searrow} + c^{\downarrow})y \geq dy + c(x - x') + cx' = cx + dy$.

We prove that, for $i = 1, \dots, n$,

$$\sum_{j=0}^n c_{ij}x'_{ij} \leq \sum_{j=0}^n c_{(i-1, j)\downarrow}^{\downarrow} y_{(i-1, j)\downarrow}, \quad (29)$$

which implies $c^{\downarrow}y \geq cx'$. Observe first that, for $1 \leq i \leq m$ and $0 \leq j \leq n$

$$\sum_{h=j}^n x'_{ih} = \sum_{h=j}^n x_{ih} - \sum_{h=\max\{1, j\}}^n y_{(i-1, h-1)\searrow} = \sum_{(i, h) \in S_{ij}} x_{ih} - \sum_{e \in \delta^-(S_{ij})} y_e + \sum_{h=j}^n y_{(i-1, h)\downarrow} \leq \sum_{h=j}^n y_{(i-1, h)\downarrow}$$

where the last inequality follows from (28).

By construction, $0 \leq c_{(i-1, 0)\downarrow}^{\downarrow} \leq c_{(i-1, 1)\downarrow}^{\downarrow} \leq \dots \leq c_{(i-1, n)\downarrow}^{\downarrow}$ and $c_{ij} \leq c_{(i-1, j)\downarrow}^{\downarrow}$ for $1 \leq i \leq m$, $0 \leq j \leq n$. Now inequality (29) is implied by the following claim.

Let $w, z \in \mathbb{R}^q$ such that $w \geq 0$ and $\sum_{i=j}^q w_i \leq \sum_{i=j}^q z_i$ for $j = 1, \dots, q$. For every $\alpha, \beta \in \mathbb{R}^q$ such that $0 \leq \beta_1 \leq \beta_2 \leq \dots \leq \beta_q$ and $\alpha \leq \beta$, the inequality $\sum_{i=1}^q \alpha_i w_i \leq \sum_{i=1}^q \beta_i z_i$ holds.

The above claim is proved by induction on q , the case $q = 1$ being trivial. Suppose $q > 1$. By inductive hypothesis $\sum_{i=2}^q (\alpha_i - \beta_1)w_i \leq \sum_{i=2}^q (\beta_i - \beta_1)z_i$. Since $\alpha_1 \leq \beta_1$ and $w_1 \geq 0$, $\sum_{i=1}^q (\alpha_i - \beta_1)w_i \leq \sum_{i=2}^q (\beta_i - \beta_1)z_i$. Hence $\sum_{i=1}^q \alpha_i w_i - \sum_{i=1}^q \beta_i z_i \leq \beta_1 \sum_{i=1}^q (w_i - z_i) \leq 0$, where the last inequality follows from $\beta_1 \geq 0$ and $\sum_{i=1}^q w_i \leq \sum_{i=1}^q z_i$. \square

8 Variable discretization

Most of the results exposed in this survey deal with integer or 0/1-variables. Over the last decade, progress in characterizing the convex hull of structured mixed integer sets has

been achieved by “discretizing” the continuous variables. That is, continuous variables are expressed as a combination of few auxiliary integer variables, thus reducing the original problem to a pure integer one. This approach has been particularly successful in several problems arising in lot-sizing [50], [35], [48], [57], [16] (see [51] for a survey).

We describe a framework that unifies many of the problems studied in the papers referenced above. This framework was proposed in [16]. Here we follow the presentation given in [18]. The problem is an extension to the mixed-integer case of the classic vertex covering in bipartite graphs.

Given a bipartite graph $G = (U \cup V, E)$ a set $I \subseteq U \cup V$ and rational numbers b_{ij} , $ij \in E$, we define the set of *mixed integer vertex covers* as

$$S^{(G,I)} = \{x \in \mathbb{R}^{U \cup V} : x_i + x_j \geq b_{ij}, ij \in E; x_i \in \mathbb{Z}, i \in I\}.$$

Let k be a positive integer such that $b_{ij} = \lfloor b_{ij} \rfloor + \frac{h_{ij}}{k}$ for some integer h_{ij} between 0 and $k - 1$ for each $ij \in E$.

Next we present an extended formulation for $\text{conv}(S^{(G,I)})$ which is polynomial on $|V(G)|$, $|E|$ and k . Thus, when k is polynomial in the input size, the formulation is compact. The size of this formulation can be improved by studying the possible fractional parts taken by the continuous variables at points lying on a minimal face of $\text{conv}(S^{(G,I)})$ (see [16]). In particular, this formulation can be made compact when G is a tree, irrespective of k . Most problems studied in the literature are of this type.

Let $V(G) = U \cup V$ and $L = V(G) \setminus I$. For any vector $x \in \mathbb{R}^{V(G)}$, we denote by x_I and x_L its restrictions to the components indexed by I and L respectively.

Remark 8.1. *Every point of $S^{(G,I)}$ is a convex combination of points $x^1, \dots, x^h \in S^{(G,I)}$ such that kx^ℓ is integral for $\ell = 1, \dots, h$.*

Proof. The constraint matrix of the system $x_i + x_j \geq b_{ij}$, $ij \in E$ is the edge-node incidence matrix A of G . Since G is bipartite, A is totally unimodular. Let $\bar{x} \in S^{(G,I)}$. By definition of $S^{(G,I)}$, \bar{x}_I is integral. Furthermore, $\bar{z} = k\bar{x}_L$ is in the polyhedron Q defined by $A_L z \geq k(b - A_I \bar{x}_I)$, where A_L and A_I are the column submatrices of A indexed by L and I , respectively. Since A_L is totally unimodular and $k(b - A_I \bar{x}_I)$ is an integral vector, Q is an integral polyhedron. Hence \bar{z} is a convex combination of integral points z^1, \dots, z^h of Q . Let $x^i = (\bar{x}_I, k^{-1}z^i) \in S^{(G,I)}$, then \bar{x} is a convex combination of x^1, \dots, x^h . \square

Remark 8.2. *The polyhedron $\text{conv}(S^{(G,I)})$ is not pointed. Indeed the dimension of its lineality space is the number of connected components of G . A basis of the lineality space is given by the vectors, defined by each component C , of the form $x_i = 1$, $i \in U \cap C$, $x_i = -1$, $i \in V \cap C$, $x_i = 0$, $i \notin C$.*

Remark 8.1 shows that every minimal face of $\text{conv}(S^{(G,I)})$ contains a point x that is $1/k$ -integral, namely kx is an integral vector.

Consider a point \bar{x} of $\text{conv}(S^{(G,I)})$ with $k\bar{x}$ is integral. Thus $\bar{x}_i = \lfloor \bar{x}_i \rfloor + \frac{r_i}{k}$ for some integer r_i , $i \in U \cup V$. For every $i \in U \cup V$, define $\bar{\mu}_i^t = \lfloor \bar{x}_i \rfloor$ for $t = 0, \dots, k - r_i - 1$, $\bar{\mu}_i^t = \lceil \bar{x}_i \rceil$ for $t = k - r_i, \dots, k - 1$. Then $\bar{x}_i = \frac{1}{k}(\bar{\mu}_i^0 + \dots + \bar{\mu}_i^{k-1})$ where $\bar{\mu}_i^0, \dots, \bar{\mu}_i^{k-1} \in \mathbb{Z}$. Furthermore, if $i \in I$, then \bar{x}_i is integer and $\bar{\mu}_i^0 = \dots = \bar{\mu}_i^{k-1}$.

Therefore we associate k auxiliary integer variables $\mu_i^0, \dots, \mu_i^{k-1}$ to each variable x_i , $i \in V(G)$. This allows us to define $\text{conv}(S^{(G,I)})$ as the projection of a polyhedron in the (x, μ) space, as follows.

Theorem 8.3. *The polyhedron $\text{conv}(S^{(G,I)})$ is the projection onto the space of the x -variables of the polyhedron \mathcal{Q}_I defined as the set of points (x, μ) satisfying*

$$x_i - \frac{1}{k}(\mu_i^0 + \dots + \mu_i^{k-1}) = 0 \quad i \in U \cup V \quad (30)$$

$$\mu_i^t + \mu_j^{k-h_{ij}-1-t} \geq \lfloor b_{ij} \rfloor \quad t = 0, \dots, k - h_{ij} - 1 \quad ij \in E \quad (31)$$

$$\mu_i^t + \mu_j^{2k-h_{ij}-1-t} \geq \lceil b_{ij} \rceil \quad t = k - h_{ij}, \dots, k - 1 \quad ij \in E \quad (32)$$

$$\mu_i^{t-1} - \mu_i^t = 0 \quad t = 1, \dots, k - 1, i \in I. \quad (33)$$

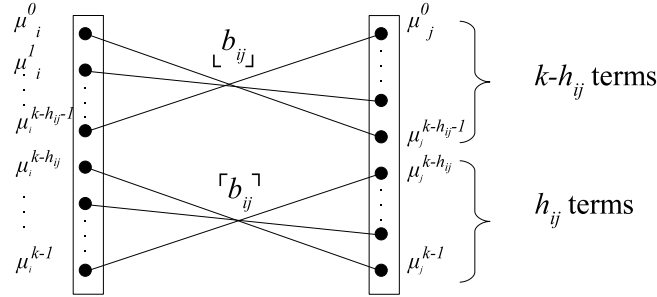


Figure 2: Representation of the extended formulation of a single inequality $x_i + x_j \geq b_{ij}$. Each edge joins the node corresponding to variable μ_i^t on the left to the node corresponding to variable $\mu_j^{q_{ijt}}$ on the right. Edges at the top group represent inequalities (31), while edges below represent inequalities (32).

We will need the following result of Heller and Tompkins.

Theorem 8.4 (Heller and Tompkins [36]). *Let A be a $0, \pm 1$ -matrix with at most two nonzero entries per row. The matrix A is totally unimodular if and only if the columns of A can be partitioned into two sets R and B such that the sum of the columns in R minus the sum of the columns in B is a $0, \pm 1$ -vector.*

For ease of notation, given an edge $ij \in E$ and an index t , $0 \leq t \leq k - 1$, we define

$$q_{ijt} = \begin{cases} k - h_{ij} - 1 - t, & t = 0, \dots, k - h_{ij} - 1; \\ 2k - h_{ij} - 1 - t, & t = k - h_{ij}, \dots, k - 1. \end{cases}$$

Proof of Theorem 8.3. We first show that the projection of \mathcal{Q}_I onto the x -space is contained in $\text{conv}(S^{(G,I)})$. Notice that, given $ij \in E$, summing all inequalities (31)-(32) relative to ij , dividing them by k and adding the equations $x_i - \sum_{t=0}^{k-1} \frac{\mu_i^t}{k} = 0$ and $x_j - \sum_{t=0}^{k-1} \frac{\mu_j^t}{k} = 0$, one obtains $x_i + x_j \geq b_{ij}$. Therefore $x_i + x_j \geq b_{ij}$ is valid for the projection of \mathcal{Q}_I .

Let M be the constraint matrix of the system defined by (31),(32),(33). Then M is a $0, \pm 1$ matrix with exactly two nonzero elements in each row, and the sum of the columns of

M corresponding to nodes in U minus the sum of the columns corresponding to nodes in V yields the vector of all zeroes. By Theorem 8.4, matrix M is totally unimodular.

Consider a point $(\bar{x}, \bar{\mu})$ of \mathcal{Q}_I . We wish to show that $\bar{x} \in \text{conv}(S^{(G,I)})$. Since the constraint matrix M of the system defined by (31),(32),(33) is totally unimodular, and the right-hand-side of such system is integral, $\bar{\mu}$ can be written as a convex combination of integral vectors μ^1, \dots, μ^h satisfying (31),(32),(33). Let x^1, \dots, x^h be the vector defined by μ^1, \dots, μ^h in the system of equations (30). By (33), x_j^1, \dots, x_j^h are integral vectors, therefore $x^1, \dots, x^h \in S^{(G,I)}$. Furthermore \bar{x} is a convex combination of x^1, \dots, x^h , thus $\bar{x} \in \text{conv}(S^{(G,I)})$.

Conversely, we show that $\text{conv}(S^{(G,I)})$ is contained in the projection of \mathcal{Q}_I onto the x -space. By Remark 8.1, given a point \bar{x} of $S^{(G,I)}$ such that $k\bar{x}$ is integral, we only need to show that there is a vector $\bar{\mu}$ such that $(\bar{x}, \bar{\mu}) \in \mathcal{Q}_I$. Since $k\bar{x}$ is integral, $\bar{x}_i = \lfloor \bar{x}_i \rfloor + \frac{r_i}{k}$ for some integer r_i , $i \in U \cup V$. Also, $\bar{x}_i \in \mathbb{Z}$ for every $i \in I$, thus $r_i = 0$ for every $i \in I$. For every $i \in U \cup V$, define $\bar{\mu}_i^t = \lfloor \bar{x}_i \rfloor$ for $t = 0, \dots, k - r_i - 1$, $\bar{\mu}_i^t = \lceil \bar{x}_i \rceil$ for $t = k - r_i, \dots, k - 1$. Clearly $(\bar{x}, \bar{\mu})$ satisfies (30) and (33). We now show that $\bar{\mu}$ satisfies (31),(32). In fact, given $ij \in E$ and an index t , $0 \leq t \leq k - 1$, observe that the vector $\bar{\mu}$ defined above satisfies

$$\begin{aligned} \frac{1}{k}(\bar{\mu}_i^0 + \dots + \bar{\mu}_i^{k-1}) + \frac{1}{k}(\bar{\mu}_j^0 + \dots + \bar{\mu}_j^{k-1}) &\geq b_{ij}; \\ \frac{1}{k}\bar{\mu}_i^t - \frac{1}{k}\bar{\mu}_i^\ell &\geq 0, \quad \ell = 0, \dots, t - 1; \\ \frac{1}{k}\bar{\mu}_i^t - \frac{1}{k}\bar{\mu}_i^\ell &\geq -\frac{1}{k}, \quad \ell = t + 1, \dots, k - 1; \\ \frac{1}{k}\bar{\mu}_j^{q_{ijt}} - \frac{1}{k}\bar{\mu}_j^\ell &\geq 0, \quad \ell = 0, \dots, q_{ijt} - 1; \\ \frac{1}{k}\bar{\mu}_j^{q_{ijt}} - \frac{1}{k}\bar{\mu}_j^\ell &\geq -\frac{1}{k}, \quad \ell = q_{ijt} + 1, \dots, k - 1. \end{aligned}$$

Summing all these inequalities, we obtain

$$\bar{\mu}_i^t + \bar{\mu}_j^{q_{ijt}} \geq b_{ij} - \frac{1}{k}(2k - 2 - t - q_{ijt}).$$

For $t = 0, \dots, k - h_{ij} - 1$, this gives $\bar{\mu}_i^t + \bar{\mu}_j^{q_{ijt}} \geq b_{ij} - 1 - \frac{1}{k}(h_{ij} - 1)$, that is $\bar{\mu}_i^t + \bar{\mu}_j^{q_{ijt}} \geq \lfloor b_{ij} \rfloor - 1 + \frac{1}{k}$. Since $\bar{\mu}$ is integral, it satisfies (31).

For $t = k - h_{ij}, \dots, k - 1$, this gives $\bar{\mu}_i^t + \bar{\mu}_j^{q_{ijt}} \geq b_{ij} - \frac{1}{k}(h_{ij} - 1)$, that is $\bar{\mu}_i^t + \bar{\mu}_j^{q_{ijt}} \geq \lfloor b_{ij} \rfloor + \frac{1}{k}$. Since $\bar{\mu}$ is integral, it satisfies (32). \square

Remark 8.5. Observe that constraints (33) are equivalent to $x_i = \mu_i^1 = \dots = \mu_i^{k-1}$ for $i \in I$. Therefore, for $i \in I$, the variables μ_i^t , $t = 0, \dots, k - 1$ can be eliminated in the system defining \mathcal{Q}_I by replacing them with variable x_i . The constraint matrix of the system obtained from (31),(32), $ij \in E$, by the above substitution is again totally unimodular.

Example. The mixing set [50] is the set

$$X_n^{MIX} = \{(x_0, x) \in \mathbb{R} \times \mathbb{Z}^n : x_0 + x_i \geq b_i, i = 1, \dots, n\}.$$

Let $b_i = \frac{h_i}{k}$, $i = 1, \dots, n$, where k and h_1, \dots, h_n are integer. After replacement of each variable μ_i^t , $i = 1, \dots, k$, $t = 0, \dots, k - 1$ by the variable x_i as explained in Remark 8.5, the extended formulation for $\text{conv}(X_n^{MIX})$ given in Theorem 8.3 becomes

$$\begin{aligned} kx_0 - \mu_0^0 - \dots - \mu_0^{k-1} &= 0, \\ \mu_0^t + x_i &\geq \lfloor b_i \rfloor, \quad t = 0, \dots, k - h_i - 1, \quad i = 1, \dots, n; \\ \mu_0^t + x_i &\geq \lceil b_i \rceil, \quad t = k - h_i, \dots, k - 1, \quad i = 1, \dots, n. \end{aligned}$$

Figure 3 depicts the constraints defining X_n^{MIX} and the corresponding extended formulation. ■

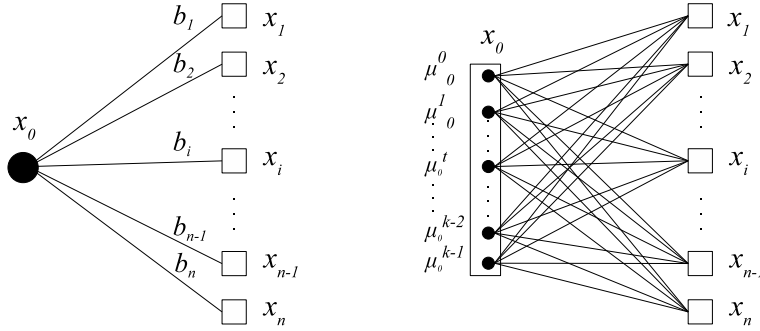


Figure 3: Mixing set and extended formulation.

Remark 8.6. *The extended formulation in Theorem 8.3 has $O(k|U \cup V|)$ variables and $O(k|E|)$ constraints. Therefore its size is pseudopolynomial, but in general not polynomial, in the size of the encoding of $S^{(G,I)}$. In fact one can construct counterexamples to polynomiality [16].*

The extended formulation in Theorem 8.3 can be turned into a formulation of polynomial size whenever we have the property that every point in $\text{conv}(S^{(G,I)})$ can be expressed as convex combination of points of $S^{(G,I)}$ in which the fractional parts of the coordinates of these points can take only “a small number” of possible values. More formally.

Remark 8.7. *Suppose we are given $\mathcal{L} \subset \{0, \dots, k - 1\}$ with the property that*

$$\text{conv}(S^{(G,I)}) = \text{conv}(S^{(G,I)} \cap \{x : k(x_i - \lfloor x_i \rfloor) \in \mathcal{L}\}).$$

(By Lemma 8.1, the set $\{0, \dots, k - 1\}$ has the above property.)

The extended formulation presented in Section 2 can be turned into an extended formulation of size polynomial in $|\mathcal{L}|$ by setting $\mu_i^t = \mu_i^{t-1}$ whenever $k - t \notin \mathcal{L}$ and then eliminating variables and duplicate constraints.

In particular, whenever such a set $|\mathcal{L}|$ is known whose size is polynomial in the input data, this yields a polynomial size extended formulation.

Conforti et al. [16] give bounds on the smallest size of a set \mathcal{L} satisfying the conditions of Remark 8.7. Given any such set \mathcal{L} , they also give an extended formulation for $\text{conv}(S^{(G,I)})$ that has $|\mathcal{L}||V(G)|$ additional variables. It can be seen that such an extended formulation is that given in Remark 8.7, with the additional constraints $\mu_i^t - \mu_i^{t-1} \geq 0$, $t = 1, \dots, k-1$, $\mu_i^0 - \mu_i^{k-1} \geq -1$.

9 Conclusion

This survey presented a number of tools for deriving and studying extended formulations: Projection, Minkowski-Weyl, union of polyhedra, dynamic programming, and variable discretization. In each case, we showed how to use the tool in one or several applications. A theorem of Yannakakis gives a lower bound on the size of an extended formulation. This result does not seem to be well known and it certainly deserves greater recognition.

Acknowledgments: We thank Marco DiSumma for his valuable comments.

References

- [1] M. Ajtai, J. Komlós, E. Szemerédi, An $O(n \log n)$ sorting network, *Proceedings of the 15th Annual ACM Symposium on Theory of Computing* (1983), 1–9.
- [2] N. Alon, R. Yuster, U. Zwick. Color-coding, *Journal of the ACM* 42 (1995), 844856.
- [3] D. Avis, K. Fukuda, A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra, *Discrete Computational Geometry* 8 (1992), 295-313.
- [4] E. Balas, Disjunctive programming: properties of the convex hull of feasible points, GSIA Management Science Research Report MSRR 348, Carnegie Mellon University (1974), published as invited paper in *Discrete Applied Mathematics* 89 (1998) 1–44.
- [5] E. Balas, Disjunctive programming and a hierarchy of relaxations for discrete optimization problems, *SIAM Journal on Algebraic and Discrete Methods* 6 (1985) 466–486.
- [6] E. Balas and W. R. Pulleyblank, The perfectly matchable subgraph polytope of a bipartite graph, *Networks* 13 (1983), 495-516.
- [7] E. Balas and W. R. Pulleyblank, The perfectly matchable subgraph polytope of an arbitrary graph, *Combinatorica* 9 (1989), 321-337.
- [8] F. Barahona, On cuts and matchings in planar graphs, *Mathematical Programming* 60 (1993), 53–68.
- [9] F. Barahona, Reducing matching to polynomial size linear programming, *SIAM Journal of Optimization* 3 (1993), 688–695.
- [10] A. Ben-Tal, A. Nemirovski, On polyhedral approximations of the second order cone, *Mathematics of Operations Research* 26 (2001) 193-205.
- [11] D. Bienstock, Approximate formulations for 0 – 1 knapsack sets, *Operations Research Letters* 36 (2008) 317–320.
- [12] D. Bienstock, B. McClosky, Tightening simple mixed-integer sets with guaranteed bounds, *Mathematical Programming* 133 (2012), 337-363.

- [13] R. Carr, G. Konjevod, G. Little, V. Natarajan, O. Parekh, Compacting cuts: a new linear formulation for minimum cut, *Proceeding of the Symposium on Discrete Algorithms* (2007), 43-52.
- [14] M. Conforti, G. Cornuéjols and G. Zambelli, Polyhedral Approaches to Mixed Integer Linear Programming, in *50 Years of Integer Programming 1958-2008*, Springer Berlin Heidelberg (2009) 343–385.
- [15] M. Conforti, G. Cornuéjols and G. Zambelli, Extended formulations in combinatorial optimization, *4OR: A Quarterly Journal of Operations Research* 8 (2010) 1-48.
- [16] M. Conforti, M. Di Summa, F. Eisenbrand, L.A. Wolsey, Network formulations of mixed-integer programs, *Mathematics of Operations Research* 34 (2009) 194-209.
- [17] M. Conforti and L. Wolsey, Compact formulations as a union of polyhedra, *Mathematical Programming* 114 (2008), 277-289.
- [18] M. Conforti, L. Wolsey and G. Zambelli, Projecting an extended formulation for mixed-integer covers on bipartite graphs, *Mathematics of Operations Research* 35 (2010), 603-623.
- [19] T.H. Cormen, C.E. Leiserson, R.L. Rivest, C. Stein, *Introduction to Algorithms*, MIT Press, 2001.
- [20] J. Edmonds, Maximum matching and a polyhedron with 0,1-vertices, *Journal of Research National Bureau of Standards Section B* 69 (1965) 125-130.
- [21] J. Edmonds, Submodular functions, matroids, and certain polyhedra, in *Combinatorial Structures and Their Applications* (Proceedings Calgary International Conference on Combinatorial Structures and Their Applications, Calgary, Alberta, 1969; R. Guy, H. Hanani, N. Sauer, J. Schönheim, eds.), Gordon and Breach, New York, 1970, 69–87.
- [22] J. Edmonds, Matroids and the greedy algorithm, *Mathematical Programming* 1 (1971) 127–136.
- [23] J. Edmonds, Edge-disjoint branchings, in *Combinatorial Algorithms* (Courant Computer Science Symposium 9, Monterey, California, 1972; R. Rustin, ed.), Algorithmic Press, New York, 1973, 91-96.
- [24] J. Edmonds, E.L. Johnson, Matching, Euler tours and the Chinese postman, *Mathematical Programming* 5 (1973) 88–124.
- [25] Y. Faenza, V. Kaibel, Extended Formulations for Packing and Partitioning Orbitopes, *Mathematics of Operations Research* 34 (2009) 686-697.
- [26] Y. Faenza, G. Oriolo, and G. Stauffer, Separating stable sets in claw-free graphs via Padberg-Rao and compact linear programs, *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2012)* (2012) 1298-1308.
- [27] S. Fiorini, s. Massar, S. Pokutta, H.R. Tiwary and R. de Wolf, Linear vs. Semidefinite Extended Formulations: Exponential Separation and Strong Lower Bounds, in *Proceedings of the 44th symposium on Theory of Computing* (H. J. Karloff, T. Pitassi eds.), ACM, 2012, 95-106.
- [28] S. Fiorini, T. Rothvoss, H. R. Tiwary, Extended formulations for polygons, 2012, Manuscript.

- [29] J.B.J. Fourier, Solution d'une question particulière du calcul des inégalités, *Nouveau Bulletin des Sciences par la Société Philomatique de Paris* (1826) 317–319.
- [30] D. R. Fulkerson, Blocking polyhedra, in: *Graph Theory and its Applications*, (Proceedings Advanced Seminar Madison, Wisconsin, (1969); B. Harris ed.), Academic Press, New York, (1970) 93-112.
- [31] B. Gerards, Compact systems for T-join and perfect matching polyhedra of graphs with bounded genus, *Operations Research Letters* 10 (1991) 377-382.
- [32] R. Giles and L. Trotter Jr., On the stable set polytope for $K_{1,3}$ -free graphs, *Journal of Combinatorial Theory B* 31 (1981) 313–326.
- [33] M.X. Goemans, Smallest compact formulation for the permutahedron, working paper, MIT Department of Mathematics, 2010.
- [34] B. Guenin, A characterization of weakly bipartite graphs, *Journal of Combinatorial Theory B* 83 (2001) 112–168.
- [35] O. Günlük and Y. Pochet, Mixing mixed-integer inequalities, *Mathematical Programming* 90 (2001) 429–458.
- [36] I. Heller and C.B. Tompkins, An extension of a theorem of Danzig's, in *Linear Inequalities and Related Systems* (H. W. Kuhn and A. W. Tucker, eds.) Princeton University Press, Princeton, N.J., 1956, 247-254.
- [37] A. J. Hoffman and J. B. Kruskal, Integral boundary points of polyhedra, in *Linear Inequalities and Related Systems* (H. W. Kuhn and A. W. Tucker eds.), Princeton University Press, Princeton, N.J., 1956, 223–246.
- [38] O.H. Ibarra, C.E. Kim, Fast approximation algorithms for the knapsack and sum of subset problems, *Journal of the ACM* 22 (1975) 463–468.
- [39] R. Jeroslow, On defining sets of vertices of the hypercube by linear inequalities, *Discrete Mathematics* 11 (1975) 119-124.
- [40] V. Kaibel, A. Loos, Branched Polyhedral Systems, in *Integer Programming and Combinatorial Optimization XIV, Lecture Notes in Computer Science 6080* (F. Eisenbrand and B. Shepherd eds.), Springer, 2010, 177-190.
- [41] V. Kaibel, K. Pashkovich, D.O. Theis, Symmetry matters for sizes of extended formulations, in *Integer Programming and Combinatorial Optimization XIV, Lecture Notes in Computer Science 6080* (F. Eisenbrand and B. Shepherd eds.), Springer, 2010, 135-148.
- [42] V. Kaibel, K. Pashkovich, Constructing Extended Formulations from Reflection Relations, in *Integer Programming and Combinatorial Optimization XV, Lecture Notes in Computer Science 6655* (O. Günlük, G. Woeginger eds.), Springer, 2011, 287-300.
- [43] V. Kaibel, M.E. Pfetsch, Packing and partitioning orbitopes, *Mathematical Programming A* 114 (2008) 1-36.
- [44] R.M. Karp, Reducibility among combinatorial problems, in: *Complexity of Computer Computations* (R.E. Miller and J.W. Thatcher eds.), Plenum Press, New York (1972) 85–103.
- [45] E.L. Lawler, Fast approximation algorithms for knapsack problems, *Symposium on Foundations of Computer Science* (1977) 206–213.

- [46] R. K. Martin, R. L. Rardin and B. A. Campbell, Polyhedral Characterization of Discrete Dynamic Programming, *Operations Research* 38 (1990) 127-38.
- [47] R.R. Meyer, On the Existence of Optimal Solutions to Integer and Mixed Integer Programming Problems, *Mathematical Programming* 7 (1974) 223-235.
- [48] A. Miller and L.A. Wolsey, Tight formulations for some simple MIPs and convex objective IPs, *Mathematical Programming B* 98 (2003) 73-88.
- [49] M.W. Padberg, M.R. Rao, Odd minimum cut-sets and b -matchings, *Mathematics of Operations Research* 7 (1982) 67-80.
- [50] Y. Pochet and L.A. Wolsey, Polyhedra for lot-sizing with Wagner-Whitin costs, *Mathematical Programming* 67 (1994) 297-324.
- [51] Y. Pochet and L.A. Wolsey, *Production Planning by Mixed Integer Programming*, Springer Series in Operations Research and Financial Engineering, New York, 2006.
- [52] W. R. Pulleyblank and F. B. Shepherd, Formulations for the stable set polytope, *Proceedings of IPCO III* (eds. G. Rinaldi, L.A. Wolsey), Springer, Berlin, 1993, pp. 267-279.
- [53] T. Rothvoss, Some 0/1 polytopes need exponential size extended formulations, *Mathematical Programming* (2012) DOI: 10.1007/s10107-012-0574-3.
- [54] A. Schrijver, *Theory of Linear and Integer Programming*, Wiley, New York (1986).
- [55] A. Schrijver, *Combinatorial Optimization, Polyhedra and Efficiency*, Springer, Berlin (2003).
- [56] P.D. Seymour, Sums of circuits, in *Graph Theory and Related Topics* (J.A. Bondy and U.S.R. Murty, eds), Academic Press, New York, 1979, 341-355.
- [57] M. Van Vyve, The continuous mixing polyhedron, *Mathematics of Operations Research* 30 (2005) 441-452.
- [58] M. Van Vyve, L.A. Wolsey, Approximate extended formulations, *Mathematical Programming* 105 (2006) 501-522.
- [59] S. Vavasis, On the complexity of nonnegative matrix factorization, *SIAM Journal on Optimization* 20 (2009) 1364-1377.
- [60] R. de Wolf, Nondeterministic quantum query and communication complexities, *SIAM Journal on Computing* 32 (2003) 681-699.
- [61] M. Yannakakis, Expressing combinatorial optimization problems by linear programs, *Journal of Computer and System Sciences* 43 (1991) 441-466.
- [62] G.M. Ziegler, *Lectures on Polytopes*, Springer, New York (1995).