EXTENDED FORMULATIONS FOR PACKING AND PARTITIONING ORBITOPES

YURI FAENZA AND VOLKER KAIBEL

ABSTRACT. We give compact extended formulations for the packing and partitioning orbitopes (with respect to the full symmetric group) described and analyzed in [6]. These polytopes are the convex hulls of all 0/1-matrices with lexicographically sorted columns and at most, resp. exactly, one 1-entry per row. They are important objects for symmetry reduction in certain integer programs. Using the extended formulations, we also derive a rather simple proof of the fact [6] that basically shifted-column inequalities suffice in order to describe those orbitopes linearly.

1. INTRODUCTION

Exploitation of symmetries is crucial for many very difficult integer programming models. Over the last few years, significant progress has been achieved with respect to general techniques for dealing with symmetries within branch-and-cut algorithms. Very nice and effective procedures have been devised, like isomorphism pruning [10, 12, 11, 13] and orbital branching [7, 8]. There has also been progress in understanding linear inequalities to be added to certain integer programs in order to remove symmetry. Towards this end, orbitopes have been introduced in [6].

The *packing orbitope* $O_{p,q}^{\leq}$ and the *partitioning orbitope* $O_{p,q}^{=}$ are the convex hulls of all 0/1-matrices of size $p \times q$ whose columns are in lexicographically decreasing order having at most or exactly, respectively, one 1-entry per row. In [6], complete descriptions with linear inequalities have been derived for these polytopes (see Thm. 16 and 17 in [6]). Knowledge on orbitopes turns out to be quite useful in practical symmetry reduction for certain integer programming models. For instance, in a well-known formulation of the graph partitioning problem (for graphs having p nodes to be partitioned into q parts) the symmetry on the 0/1-variables x_{ij} indicating whether node i is put into part j of the partitioning the symmetry arising from permuting the parts can be removed by requiring $x \in O_{p,q}^{=}$. We refer to [5] and [6] for a more detailed discussion of the practical use of orbitopes.

The topic of this paper are extended formulations for these orbitopes, i.e., (simple) linear descriptions of higher dimensional polytopes which can be projected to $O_{p,q}^{\leq}$ and $O_{p,q}^{=}$. In fact, such extended formulations play important roles in polyhedral combinatorics and integer programming in general, because rather than solving a linear optimization problem over a polyhedron in the original space, one may solve it over a (hopefully simpler described) polyhedron of which the first one is a linear projection. For instance, the lift-and-project approach [1] and other general reformulation schemes (e.g., the ones due to Lovász and Schrijver [9] as well as Adams and Sherali [15]) are based on extended formulations. Recent examples include work on mixed integer programming for duals of network matrices [3, 4, 2]. More classical is the general theory on extended formulations obtained from (certain) dynamic programming algorithms [14]. The results of this paper are much in the spirit of the latter work.

In order to give an overview on the contributions of this paper let us first recall a few facts on orbitopes. As no 0/1-matrix with at most one 1-entry per row and lexicographically (decreasing) sorted columns has a one above its main diagonal, we may assume without loss of generality that $O_{p,q}^{=} \subseteq O_{p,q}^{\leq} \subseteq \mathbb{R}^{\mathcal{I}_{p,q}}$ with $\mathcal{I}_{p,q} = \{(i,j) \in [p] \times [q] \ : \ i \geq j\}$ (where $[n] = \{1, 2, \ldots, n\}$). In fact, $O_{p,q}^{=}$ is the face of $O_{p,q}^{\leq}$ defined by requiring that all *row-sum inequalities* $x(\mathrm{row}_i) \leq 1$ for $i \in [p]$ are satisfied with equality, where $\mathrm{row}_i = \{(i,j) \in \mathcal{I}_{p,q} \ : \ j \in [q]\}$.

The main result of [6] is a complete description of $O_{p,q}^{\leq}$ and $O_{p,q}^{=}$ by means of linear equations and inequalities. This system of constraints (the *SCI-system*) consists, next to nonnegativity constraints and row-sum inequalities (or row-sum equations for $O_{p,q}^{=}$), of the exponentially large class of *shifted column inequalities* (*SCI*) that will be defined at the end of Sect. 2. In [6] it is also proved that, up to a few exceptions, these exponentially many SCIs define facets of the orbitopes. The proof given in [6] of the fact that the SCI-system completely describes these orbitopes is rather lengthy and somewhat technical. Extending over pages 18 to 27, it hardly leaves (not only) the reader with a good idea of the reasons for the SCI-system being sufficient to describe the orbitopes.

In contrast to this, the contributions of the present work are the following: We provide a quite simple extended formulation for $O_{p,q}^{\leq}$ (along with a rather short proof establishing this) and, moreover, we show by some simple and natural (not technical) arguments that the SCI-system describes the projection of the feasible region of that extended formulation to the original space, thus providing a new proof showing that the SCI-system describes $O_{p,q}^{\leq}$.

This latter proof is much shorter than the original one, and it seems to provide much better insight into the reasons for the SCI-system to describe the orbitopes. Clearly, as $O_{p,q}^{=}$ is a face of $O_{p,q}^{\leq}$, the results for the latter polytope immediately yield corresponding results for the first one. However, besides leading to that simpler proof, we believe that our extended formulation for $O_{p,q}^{\leq}$ is interesting itself. It provides a description of a quite natural polytope (the orbitope $O_{p,q}^{\leq}$) by

a system of constraints in a space whose dimension is roughly twice the original dimension $|\mathcal{I}_{p,q}|$ with only linearly (in $|\mathcal{I}_{p,q}|$) many nonzero coefficients, while every linear description of the orbitope in the original space requires exponentially many inequalities. This may also turn out to be computationally attractive.

The basic idea of our extended formulation is to assign to each vertex of $O_{p,q}^{\leq}$ a directed path in a certain acyclic digraph. The additional variables in our extended formulations are used to suitably express these paths. The digraph we work with is set up in Sect. 2, where we also fix some notations and define SCIs. In Sect. 3 we then describe the extended formulations for $O_{p,q}^{\leq}$ and $O_{p,q}^{=}$ (Thm. 6 and Cor. 8). The main work is done in Sect. 3.1, where the extended formulation for $O_{p,q}^{\leq}$ with additional variables encoding the paths mentioned above is introduced and proved to define an integral polyhedron (Thm. 4). From this it is easy to conclude that the formulation indeed defines a polytope that projects down to $O_{p,q}^{\leq}$ (Thm. 6). Both the extension of such results to the partitioning case $O_{p,q}^{=}$ (Cor. 8 in Sect. 3.2), and the transformations of the systems in order to reduce the numbers of variables and nonzero coefficients (Thm. 10 in Sect. 3.3) are obtained without much work. On the way, we also derive linear (in $|\mathcal{I}_{p,q}|$) time algorithms for optimizing linear objective functions over $O_{p,q}^{\leq}$ and $O_{p,q}^{=}$ (Cor. 7 and 9). In Sect. 4 we finally prove that the projection of the feasible region defined by the extended formulation contains the polytope defined by the SCI-system (Thm. 12), thus providing the new proof of the fact (Thm. 11) that the latter polytope equals $O_{p,q}^{\leq}$. We conclude with a few remarks and acknowledgements in Sect. 5.

2. THE SETUP

Let us assume $p \geq q \geq 1$ throughout the paper. We define a directed acyclic graph $D_{p,q} = (V_{p,q}, A_{p,q})$ with node set

$$V_{p,q} = \mathcal{I}_{p,q} \uplus ([p]_0 \times \{0\}) \uplus \{s\} \uplus \{t\}$$

(where $[n]_0 = [n] \cup \{0\}$ and $\uplus$ means disjoint union). Using the notation $q(i) = \min\{i, q\}$, the set of arcs of $D_{p,q}$ is

$$A_{p,q} = A_{p,q}^{\downarrow} \cup A_{p,q}^{\searrow} \cup \{(s, (0,0))\} \cup \{((p,j),t) \,:\, j \in [q]_0\},$$

where

$$A_{p,q}^{\downarrow} = \{((i,j),(i+1,j)) \,:\, i \in [p-1]_0, j \in [q(i)]_0\}$$

is the set of *vertical arcs* that are denoted by $(i,j)^{\downarrow} = ((i,j),(i+1,j))$, and

$$A_{p,q}^{\searrow} = \{((i,j),(i+1,j+1)) \,:\, i \in [p-1]_0, j \in [q(i+1)-1]_0\}$$

is the set of *diagonal arcs* that are denoted by $(i,j)^{\searrow} = ((i,j),(i+1,j+1))$. The crucial property of $D_{p,q}$ is that every vertex of $O_{p,q}^{\leq}$ induces an s-t-path in $D_{p,q}$ as indicated in Fig. 1. Note that different vertices may induce the same path.
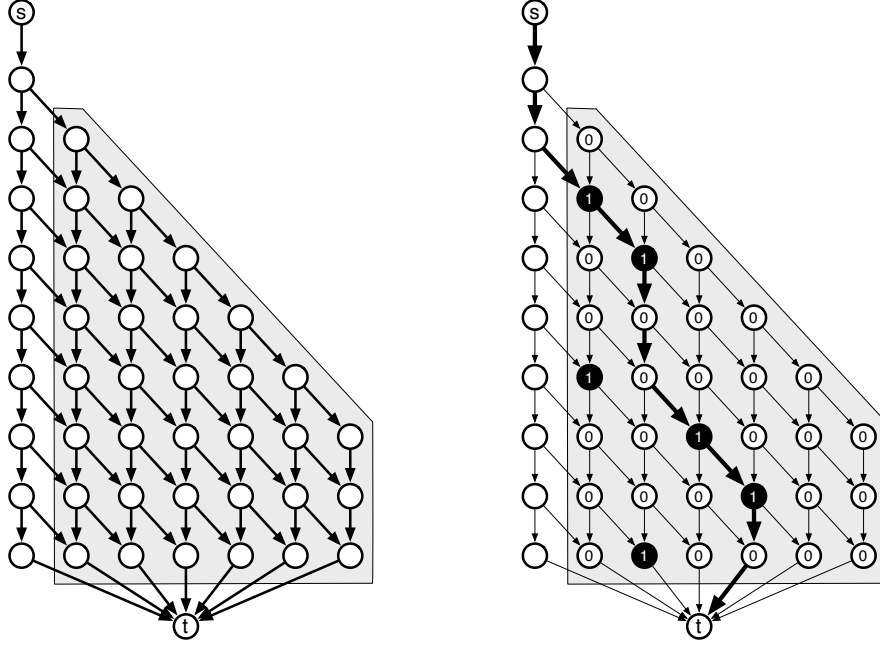
FIGURE 1. The digraph $D_{8,6}$ and a vertex of $O_{8,6}^{\leq}$ along with its s-t-path.

For a subset $W \subseteq V_{p,q}$ we use the following notation:

$$
\begin{aligned}
\text{out}(W) &= \{(w,u) \in A_{p,q} : w \in W, u \notin W\} \\
\text{out}^{\downarrow}(W) &= \{(w,u) \in A_{p,q}^{\downarrow} : w \in W, u \notin W\} \\
\text{in}(W) &= \{(u,w) \in A_{p,q} : w \in W, u \notin W\} \\
\text{in}^{\downarrow}(W) &= \{(u,w) \in A_{p,q}^{\downarrow} : w \in W, u \notin W\} \\
\text{in}^{\searrow}(W) &= \{(u,w) \in A_{p,q}^{\searrow} : w \in W, u \notin W\}
\end{aligned}
$$

For a directed path Γ in $D_{p,q}$, we denote by

$V(\Gamma) \subseteq V_{p,q}$ the set of nodes on the path Γ,

$S(\Gamma) \subseteq V(\Gamma)$ the set of nodes on Γ not entered by Γ via diagonal arcs, and

$T(\Gamma) \subseteq V(\Gamma)$ the set of nodes on Γ left by Γ via diagonal arcs

(see Fig. 2). Note that $S(\Gamma)$ always contains the start node of Γ, and $T(\Gamma)$ always excludes the end node of Γ.

Remark 1. *For every directed path Γ in $D_{p,q}$ with end node $(i,j) \in \mathcal{I}_{i,j}$, we have*

$$
\text{in}^{\searrow}(V(\Gamma)) = \text{in}^{\searrow}(S(\Gamma)) \quad and \quad \text{out}^{\downarrow}(V(\Gamma) \setminus \{(i,j)\}) = \text{out}^{\downarrow}(T(\Gamma)) .
$$

A subset $S \subseteq \mathcal{I}_{p,q}$ is a *shifted column* if and only if $S = S(\Gamma)$ for some (ℓ,ℓ)-$(i-1, j-1)$-path Γ in $D_{p,q}$ with $i \in [p] \setminus \{1\}$, $j \in [q] \setminus \{1\}$, and $\ell \in [q]$. The associated *shifted-column inequality* is $x(\text{bar}_{i,j}) \leq x(S)$, where $\text{bar}_{i,j} = \{(i,\ell) \in \mathcal{I}_{p,q} : \ell \geq j\}$
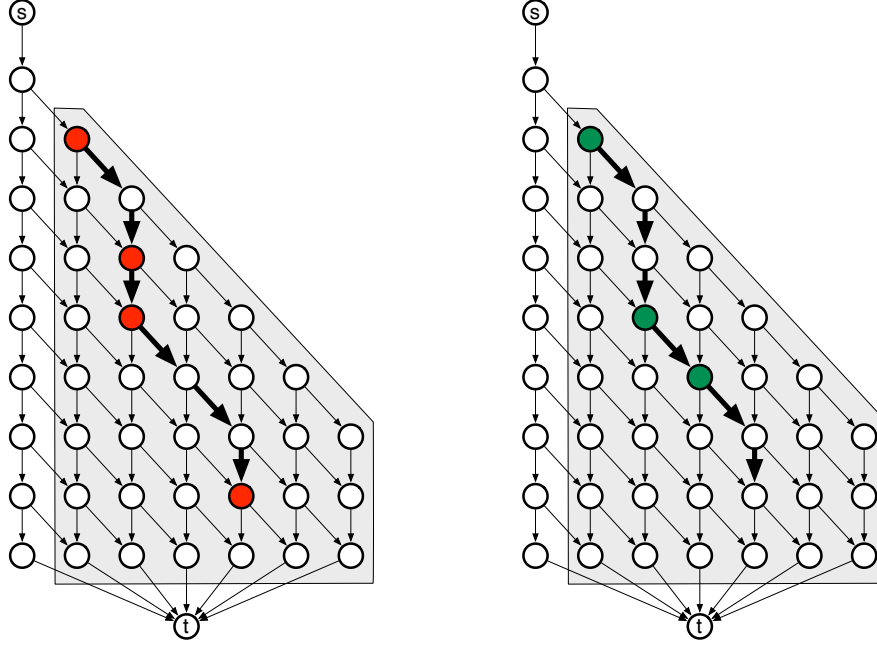
FIGURE 2. A path Γ along with the sets $S(\Gamma)$ (left) and $T(\Gamma)$ (right).

and, as usual, we write $z(N) = \sum_{e \in N} z_e$ for some vector $z \in \mathbb{R}^M$ and a subset $N \subseteq M$ (see Fig. 3).

3. Extended formulations

3.1. **The packing case.** Denote by $F_{p,q} \subseteq \mathbb{R}^{A_{p,q}}$ the set of all s-t-flows (without any capacity restrictions) in $D_{p,q}$ with flow value one. Clearly, $F_{p,q}$ is an integral polytope. Since $D_{p,q}$ is acyclic, the vertices of $F_{p,q}$ are the incidence vectors of the directed s-t-paths (viewed as subsets of arcs) in $D_{p,q}$.

For a flow $y \in F_{p,q}$ and a node $(i,j) \in \mathcal{I}_{p,q}$, we denote by

$$y(i,j) = y(\text{in}(i,j)) = y(\text{out}(i,j))$$

the amount of flow passing node (i,j). For a subset $W \subseteq \text{row}_i$ of nodes in the same row, $y(W) = \sum_{w \in W} y(w)$ is the total amount of flow entering W (or, equivalently, leaving W).

Lemma 2. *For a directed (k,ℓ)-(i,j)-path Γ in $D_{p,q}$ with $(i,j) \in \mathcal{I}_{p,q}$ the following statements hold for all $y \in F_{p,q}$ (see Fig. 4):*

(1) *If $k = \ell \geq 1$ then*

$$y(\text{in}^\searrow(S(\Gamma))) - y(\text{out}^\downarrow(T(\Gamma))) = y(\text{bar}_{i,j}).$$

(2) *If $\ell = 0$ then*

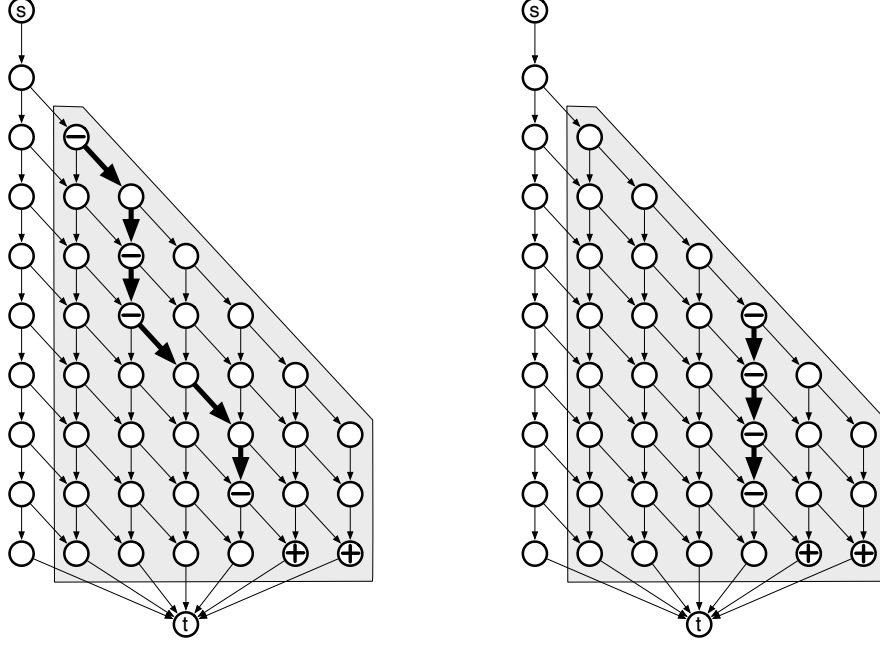$$1 - y(\text{out}^\downarrow(T(\Gamma))) = y(\text{bar}_{i,j}).$$

FIGURE 3. Coefficient vectors of two SCIs with the same bar $\mathrm{bar}_{8,5}$.

Proof. For both cases, let

$$W = \{(a', b) \in \mathrm{V}_{p,q} \setminus \{s, t\} \,:\, a' \le a \text{ for some } (a, b) \in \mathrm{V}(\Gamma) \cup \mathrm{bar}_{i,j}\}$$

be the set of all nodes (different from s and t) in or above $\mathrm{V}(\Gamma) \cup \mathrm{bar}_{i,j}$.

For case (1), we start by observing

$$(1) \qquad\qquad\qquad y(\mathrm{in}(W)) = y(\mathrm{out}(W))$$

(as $y \in \mathrm{F}_{p,q}$ is an s-t-flow). Because of $(0,0) \notin W$ (thus s being not adjacent to W) we have $\mathrm{in}(W) = \mathrm{in}^{\searrow}(\mathrm{V}(\Gamma))$, which according to Remark 1 equals $\mathrm{in}^{\searrow}(\mathrm{S}(\Gamma))$, yielding

$$(2) \qquad\qquad\qquad y(\mathrm{in}(W)) = y(\mathrm{in}^{\searrow}(\mathrm{S}(\Gamma)))\,.$$

Similarly, we have $\mathrm{out}(W) = \mathrm{out}^{\downarrow}(\mathrm{V}(\Gamma) \setminus \{(i, j)\}) \uplus \mathrm{out}(\mathrm{bar}_{i,j})$, where Remark 1 gives $\mathrm{out}^{\downarrow}(\mathrm{V}(\Gamma) \setminus \{(i, j)\}) = \mathrm{out}^{\downarrow}(\mathrm{T}(\Gamma))$. Thus, we obtain

$$(3) \qquad\qquad y(\mathrm{out}(W)) = y(\mathrm{out}^{\downarrow}(\mathrm{T}(\Gamma))) + y(\mathrm{out}(\mathrm{bar}_{i,j}))\,.$$

Equations (1), (2), and (3) imply the statement on case (1).

For case (2), we exploit the fact that the s-t-flow $y \in \mathrm{F}_{p,q}$ of value one satisfies

$$(4) \qquad\qquad\qquad 1 = y(\mathrm{out}(W \cup \{s\})) - y(\mathrm{in}(W \cup \{s\}))\,.$$

We have $\mathrm{in}(W \cup \{s\}) = \varnothing$ and $\mathrm{out}(W \cup \{s\}) = \mathrm{out}(W)$ (due to $(0,0) \in W$ in this case). From

$$\mathrm{out}(W) = \mathrm{out}^{\downarrow}(\mathrm{V}(\Gamma) \setminus \{(i, j)\}) \uplus \mathrm{out}(\mathrm{bar}_{i,j})$$
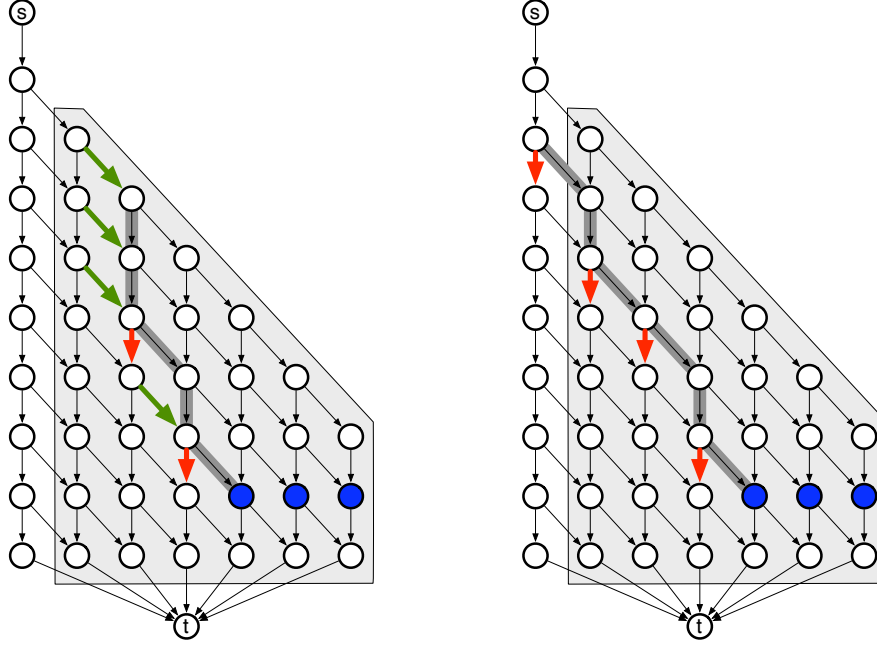
FIGURE 4. Illustration of part (1) (left) and part (2) (right) of Lemma 2 with $(i, j) = (7, 4)$.

and $\mathrm{out}^\downarrow(\mathrm{V}(\Gamma) \setminus \{(i, j)\}) = \mathrm{out}^\downarrow(\mathrm{T}(\Gamma))$ (see Remark 1), we thus derive the statement on case (2) from (4). $\square$

For $(i, j) \in \mathcal{I}_{p,q}$ we denote by

$$\mathrm{col}_{i,j} = \{(k, j) \,:\, j \le k \le i\}$$

the upper part of the j-th column from (j, j) down to (i, j), including both nodes. For the directed path Γ with $\mathrm{V}(\Gamma) = \mathrm{col}_{i,j}$, we have $\mathrm{S}(\Gamma) = \mathrm{col}_{i,j}$ and $\mathrm{T}(\Gamma) = \varnothing$. Thus, part (1) of Lemma 2 implies the following.

Remark 3. *For all $y \in \mathrm{F}_{p,q}$ and $(i, j) \in \mathcal{I}_{p,q}$, we have*

$$y(\mathrm{in}^{\searrow}(\mathrm{col}_{i,j})) = y(\mathrm{bar}_{i,j}) \,.$$

The central object of study of this paper is the polytope

$$\mathrm{P}_{p,q} = \{(x, y) \in \mathbb{R}^{\mathcal{I}_{p,q}} \times \mathbb{R}^{\mathrm{A}_{p,q}} \,:\, y \in \mathrm{F}_{p,q} \text{ and } (x, y) \text{ satisfies (5) and (6) below}\}$$

with

(5) $$y_{(i-1,j-1)^{\searrow}} \le x_{ij} \quad \text{for all } (i, j) \in \mathcal{I}_{p,q}$$

and

(6) $$x(\mathrm{bar}_{i,j}) \le y(\mathrm{in}^{\searrow}(\mathrm{col}_{i,j})) \quad \text{for all } (i, j) \in \mathcal{I}_{p,q}$$

(see Fig. 5). Since $\mathrm{F}_{p,q}$ is the set of all s-t-flows of value one, all $(x, y) \in \mathrm{P}_{p,q}$
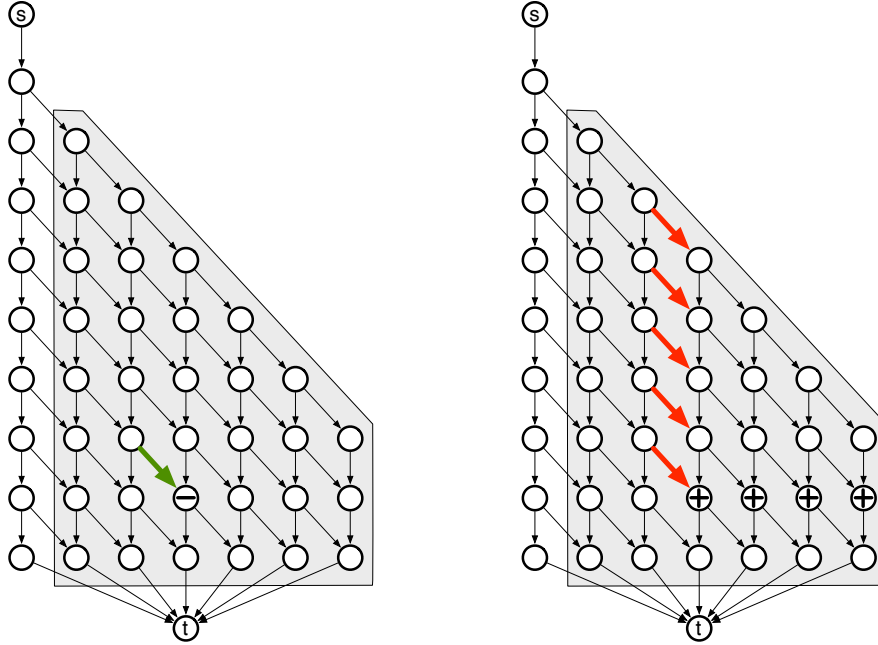
FIGURE 5. Coefficient vectors of inequalities (5) (left) and (6) (right).

satisfy $\mathbf{0} \leq y \leq \mathbf{1}$, and thus $\mathbf{0} \leq x \leq \mathbf{1}$ due to (5) and (6). Furthermore, inequalities (6) imply the row-sum inequalities $x(\mathrm{row}_i) \leq 1$ for all $(x, y) \in \mathrm{P}_{p,q}$.

Theorem 4. *The polytope $\mathrm{P}_{p,q}$ is integral.*

For the proof of this theorem, we need the following result.

Lemma 5. *Let $x_1, \ldots, x_n \geq 0$ and $y_1, \ldots, y_n \in \mathbb{R}$ with*

$$\sum_{\ell=j}^{n} x_\ell \leq \sum_{\ell=j}^{n} y_\ell \quad \text{for all } 1 \leq j \leq n\,.$$

For all numbers $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ and $0 \leq \beta_1 \leq \beta_2 \leq \cdots \leq \beta_n$ with

$$\alpha_j \leq \beta_j \quad \text{for all } 1 \leq j \leq n$$

the inequality

$$\sum_{j=1}^{n} \alpha_j x_j \leq \sum_{j=1}^{n} \beta_j y_j$$

holds.

Proof of Lemma 5. We prove the claim by induction on n. The case $n = 1$ is trivial, thus let $n \geq 1$. Ignoring index 1 and decreasing the remaining α_j and β_j

by β_1, from the induction hypothesis we obtain

$$\sum_{j=2}^{n}(\alpha_j - \beta_1)x_j \leq \sum_{j=2}^{n}(\beta_j - \beta_1)y_j \, .$$

Due to $\alpha_1 \leq \beta_1$ and $x_1 \geq 0$ we have $(\alpha_1 - \beta_1)x_1 \leq 0$, thus we deduce

$$\sum_{j=1}^{n}(\alpha_j - \beta_1)x_j \leq \sum_{j=1}^{n}(\beta_j - \beta_1)y_j \, .$$

Hence we have

$$\sum_{j=1}^{n}\alpha_j x_j - \sum_{j=1}^{n}\beta_j y_j \leq \beta_1 \cdot \left(\sum_{j=1}^{n} x_j - \sum_{j=1}^{n} y_j \right),$$

with a nonpositive right-hand side due to $\beta_1 \geq 0$ and $\sum_{j=1}^{n} x_j \leq \sum_{j=1}^{n} y_j$. $\square$

Proof of Theorem 4. In order to show that $P_{p,q}$ is integral, we show that for an arbitrary objective function vector $c \in \mathbb{R}^{\mathcal{I}_{p,q}} \times \mathbb{R}^{A_{p,q}}$ the optimization problem

$$(7) \qquad \max\{\langle c, (x,y)\rangle \, : \, (x,y) \in P_{p,q}\}$$

has an optimal solution with $0/1$-components.

   We define two vectors $c^{(1)}, c^{(2)} \in \mathbb{R}^{\mathcal{I}_{p,q}} \times \mathbb{R}^{A_{p,q}}$ which are zero in all components with the following exceptions:

$$\begin{aligned}
c^{(1)}_{(i-1,j-1)^{\searrow}} &= c_{(i,j)} & \text{for all } (i,j) \in \mathcal{I}_{p,q} \\
c^{(2)}_{(i-1,j)^{\downarrow}} &= \max\{0, c_{(i,1)}, \ldots, c_{(i,j)}\} & \text{for all } (i,j) \in \mathcal{I}_{p,q}
\end{aligned}$$

We are going to establish the following two claims:
   (1) For each $(x,y) \in P_{p,q}$ we have

$$\langle c, (x,y)\rangle \leq \langle c + c^{(1)} + c^{(2)}, (\mathbf{0}, y)\rangle \, .$$

   (2) For each s-t-flow $y \in F_{p,q} \cap \{0,1\}^{A_{p,q}}$ there is some $x \in \{0,1\}^{\mathcal{I}_{p,q}}$ with

$$(x,y) \in P_{p,q} \quad \text{and} \quad \langle c + c^{(1)} + c^{(2)}, (\mathbf{0}, y)\rangle = \langle c, (x,y)\rangle \, .$$

With these two claims, the existence of an integral optimal solution to (7) can be established as follows: Let $\tilde{c} \in \mathbb{R}^{A_{p,q}}$ be the y-part of $c + c^{(1)} + c^{(2)}$. As $F_{p,q}$ is a $0/1$-polytope, there is a $0/1$-flow $y^{\star} \in F_{p,q} \cap \{0,1\}^{A_{p,q}}$ with

$$\langle \tilde{c}, y^{\star}\rangle = \max\{\langle \tilde{c}, y\rangle \, : \, y \in F_{p,q}\} \, .$$

Due to $\langle c + c^{(1)} + c^{(2)}, (\mathbf{0}, y)\rangle = \langle \tilde{c}, y\rangle$ for all $y \in F_{p,q}$, claim (1) implies that the optimal value of (7) is at most $\langle \tilde{c}, y^{\star}\rangle$. On the other hand, claim (2) ensures that there is some $x^{\star} \in \{0,1\}^{\mathcal{I}_{p,q}}$ with $(x^{\star}, y^{\star}) \in P_{p,q}$ and

$$\langle c, (x^{\star}, y^{\star})\rangle = \langle c + c^{(1)} + c^{(2)}, (\mathbf{0}, y^{\star})\rangle = \langle \tilde{c}, y^{\star}\rangle \, .$$

Thus, $(x^{\star}, y^{\star})$ is an integral optimal solution to (7).

In order to prove claim (2), let $y \in F_{p,q} \cap \{0,1\}^{A_{p,q}}$, i.e., y is the incidence vector of an s-t-path in $D_{p,q}$. For the construction of some $x \in \{0,1\}^{\mathcal{I}_{p,q}}$ as required we start by initializing $x = \mathbf{0}$. For each $(i,j) \in V_{p,q}$ with $y_{(i,j)^{\searrow}} = 1$ set $x_{i+1,j+1} = 1$. For every $(i,j) \in V_{p,q}$ with $j \geq 1$ and $y_{(i,j)^{\downarrow}} = 1$ choose $\ell \in [j]$ with

$$c_{i+1,\ell} = \max\{c_{(i+1),1}, \ldots, c_{(i+1),j}\},$$

and set $x_{i+1,\ell} = 1$ if $c_{i+1,\ell} \geq 0$.

For claim (1) let $(x,y) \in P_{p,q}$. Define $x' \in \mathbb{R}^{\mathcal{I}_{p,q}}$ via

$$x'_{ij} = x_{ij} - y_{(i-1,j-1)^{\searrow}}$$

for all $(i,j) \in \mathcal{I}_{p,q}$. As (x,y) satisfies (5), $x' \geq \mathbf{0}$ holds. Furthermore, we have

(8) $$\langle c^{(1)}, (\mathbf{0}, y) \rangle = \langle c, (x - x', \mathbf{0}) \rangle.$$

Therefore, it suffices to show

(9) $$\langle c^{(2)}, (\mathbf{0}, y) \rangle \geq \langle c, (x', \mathbf{0}) \rangle,$$

because (8) and (9) yield

$$
\begin{aligned}
\langle c + c^{(1)} + c^{(2)}, (\mathbf{0}, y) \rangle &= \langle c, (\mathbf{0}, y) \rangle + \langle c^{(1)}, (\mathbf{0}, y) \rangle + \langle c^{(2)}, (\mathbf{0}, y) \rangle \\
&\geq \langle c, (\mathbf{0}, y) \rangle + \langle c, (x - x', \mathbf{0}) \rangle + \langle c, (x', \mathbf{0}) \rangle \\
&= \langle c, (x, y) \rangle.
\end{aligned}
$$

In order to establish (9), we prove for every $i \in [p]$

(10) $$\sum_{j=1}^{q(i)} c_{ij} x'_{ij} \leq \sum_{j=1}^{q(i-1)} c^{(2)}_{(i-1,j)^{\downarrow}} y_{(i-1,j)^{\downarrow}},$$

which by summation over $i \in [p]$ yields (9). To see (10) for some $i \in [p]$, observe that, for every $j \in [q(i)]$, we have

$$x'(\mathrm{bar}_{i,j}) = x(\mathrm{bar}_{i,j}) - y(\mathrm{in}^{\searrow}(\mathrm{bar}_{i,j})) = x(\mathrm{bar}_{i,j}) - y(\mathrm{bar}_{i,j}) + y(\mathrm{in}^{\downarrow}(\mathrm{bar}_{i,j})).$$

Due to Remark 3, and as (x,y) satisfies (6), this implies

$$x'(\mathrm{bar}_{i,j}) \leq y(\mathrm{in}^{\downarrow}(\mathrm{bar}_{i,j})).$$

Defining $y_{(i,q(i))^{\downarrow}} = 0$ in case of $q(i-1) < q(i)$, we thus have

$$\sum_{\ell=j}^{q(i)} x'_{i\ell} \leq \sum_{\ell=j}^{q(i)} y_{(i-1,\ell)^{\downarrow}}$$

for every $j \in [q(i)]$. Setting $c^{(2)}_{(i-1,q(i))^{\downarrow}}$ to the biggest component of c in case of $q(i-1) < q(i)$, we furthermore have

$$0 \leq c^{(2)}_{(i-1,1)^{\downarrow}} \leq \cdots \leq c^{(2)}_{(i-1,q(i))^{\downarrow}}$$

and $c_{ij} \leq c^{(2)}_{((i-1),j)^{\downarrow}}$ for all $j \in [q(i)]$. Thus we can use Lemma 5 (with $n = q(i)$, $x_j = x'_{ij} \geq 0$, $y_j = y_{(i-1,j)^{\downarrow}}$, $\alpha_j = c_{ij}$, and $\beta_j = c^{(2)}_{(i-1,j)^{\downarrow}}$) to deduce

$$\sum_{j=1}^{q(i)} c_{ij} x'_{ij} \leq \sum_{j=1}^{q(i)} c^{(2)}_{(i-1,j)^{\downarrow}} y_{(i-1,j)^{\downarrow}} \,,$$

which yields (10). $\square$

From Theorem 4 one obtains that $P_{p,q}$ is an extended formulation for $O^{\leq}_{p,q}$.

Theorem 6. *The orbitope $O^{\leq}_{p,q} \subseteq \mathbb{R}^{\mathcal{I}}_{p,q}$ is the orthogonal projection of the polytope $P_{p,q} \subseteq \mathbb{R}^{\mathcal{I}_{p,q}} \times \mathbb{R}^{A_{p,q}}$ to the space $\mathbb{R}^{\mathcal{I}_{p,q}}$.*

Proof. Let $x \in O^{\leq}_{p,q} \cap \{0,1\}^{\mathcal{I}_{p,q}}$ be an arbitrary vertex of $O^{\leq}_{p,q}$. The incidence vector $y \in \{0,1\}^{A_{p,q}}$ of the unique s-t-path using all arcs $(i-1, j-1)^{\searrow}$ with $(i,j) \in \mathcal{I}_{p,q}$ and $x_{ij} = 1$ satisfies $(x,y) \in P_{p,q}$. Thus, $O^{\leq}_{p,q}$ is contained in the projection of $P_{p,q}$.

To see that vice versa the projection of $P_{p,q}$ is contained in $O^{\leq}_{p,q}$, by Theorem 4 it suffices to observe that every 0/1-point $(x,y) \in P_{p,q}$ is contained in $O^{\leq}_{p,q}$. Clearly, for such a point x has at most one one-entry per row (since the row-sum inequalities are implied by the fact $(x,y) \in P_{p,q}$). Furthermore, if the j-th column of x was lexicographically larger than the $(j-1)$-st column of x with i being minimal such that $x_{ij} = 1$ holds, then one would find that $y(\mathrm{col}_{i,j-1}) = 0$ holds (because of (5)), contradicting (6) for $(i, j-1)$. $\square$

From the proof of Theorem 4, we derive a combinatorial algorithm for the linear optimization problem

$$(11) \qquad\qquad \max\{\langle d, x \rangle \ : \ x \in O^{\leq}_{p,q}\}$$

with $d \in \mathcal{I}_{p,q}$. Indeed, with $c = (d, \mathbf{0}) \in \mathbb{R}^{\mathcal{I}_{p,q}} \times \mathbb{R}^{A_{p,q}}$ every optimal solution $(x^{\star}, y^{\star})$ to

$$(12) \qquad\qquad \max\{\langle c, (x,y) \rangle \ : \ (x,y) \in P_{p,q}\}$$

yields an optimal solution $x^{\star}$ to (11). From the proof of Theorem 4 we know that we can compute an optimal solution $(x^{\star}, y^{\star})$ to (12) by first computing the incidence vector $y^{\star} \in \{0,1\}^{A_{p,q}}$ of a longest s-t-path in the digraph $D_{p,q}$ with respect to arc length given by $c^{(1)} + c^{(2)}$ (which can be done in linear time since $D_{p,q}$ is acyclic) and then setting $x^{\star} \in \{0,1\}^{\mathcal{I}_{p,q}}$ as described in the proof of claim (2) (in the proof of Theorem 4).

Corollary 7. *Linear optimization over $O^{\leq}_{p,q}$ can be solved in time $O(pq)$.*

3.2. **The partitioning case.** The previous results can be easily extended to the partitioning case. Since the row-sum inequalities $x(\text{row}_i) \leq 1$ are valid for $\text{P}_{p,q}$,

$$\text{P}_{p,q}^{=} = \{(x,y) \in \text{P}_{p,q} \ : \ x(\text{row}_i) = 1 \text{ for all } i \in [p]\}$$

is a face of $\text{P}_{p,q}$. Clearly, due to Theorem 6 this face maps to the face (see Sect. 1)

$$\{x \in \text{O}_{p,q}^{\leq} \ : \ x(\text{row}_i) = 1 \text{ for all } i \in [p]\} = \text{O}_{p,q}^{=}$$

of $\text{O}_{p,q}^{\leq}$ via the orthogonal projection onto the x-space $\mathbb{R}^{\mathcal{I}_{p,q}}$.

Corollary 8. $\text{P}_{p,q}^{=}$ *is an extended formulation for* $\text{O}_{p,q}^{=}$.

Suppose we want to solve

(13) $$\max\{\langle d, x \rangle \ : \ x \in \text{O}_{p,q}^{=}\}$$

for some $d \in \mathbb{R}^{\mathcal{I}_{p,q}}$. As all points $x \in \text{O}_{p,q}^{=}$ satisfy the row-sum equations $x(\text{row}_i) = 1$ for all $i \in [p]$, we may add, for each i, an arbitrary constant to the objective function coefficients of the variables belonging to row_i without changing the optimal solutions to (13) (though, of course, changing the objective function values of the solutions). Therefore, we may assume that d has only positive components. But then

$$\max\{\langle d, x \rangle \ : \ x \in \text{O}_{p,q}^{=}\} = \max\{\langle d, x \rangle \ : \ x \in \text{O}_{p,q}^{\leq}\},$$

and all optimal solutions to the optimization problem over $\text{O}_{p,q}^{\leq}$ are points in $\text{O}_{p,q}^{=}$. Thus we derive the following from Corollary 7.

Corollary 9. *Linear optimization over* $\text{O}_{p,q}^{=}$ *can be solved in time* $\text{O}(pq)$.

3.3. **Reducing the number of variables and nonzero elements.** Let us manipulate the defining system of $\text{P}_{p,q}$ in order to decrease the number of variables and nonzero coefficients. This may be advantageous for practical purposes. It furthermore emphasizes the simplicity of the extended formulation. For the sake of readability, we define $\text{bar}_{i,j} = \varnothing$ and $\text{col}_{i,j} = \varnothing$ whenever $j > q(i)$.

Since every $y \in \text{F}_{p,q}$ satisfies $y_{(i,j)^{\downarrow}} = y(\text{bar}_{i,j}) - y(\text{bar}_{i+1,j+1})$, we deduce from Remark 3 that for all $(x,y) \in \text{P}_{p,q}$

$$y_{(i,j)^{\downarrow}} = y(\text{in}^{\searrow}(\text{col}_{i,j})) - y(\text{in}^{\searrow}(\text{col}_{i+1,j+1}))$$

holds for all vertical arcs $(i,j)^{\downarrow}$ with $(i,j) \in \mathcal{I}_{p,q}$ (and $i < p$) as well as

$$y_{((p,j),t)} = y(\text{in}^{\searrow}(\text{col}_{p,j})) - y(\text{in}^{\searrow}(\text{col}_{p,j+1}))$$

for all arcs $((p,j),t)$ with $j \in [q]_0$, where we defined $y(\text{in}^{\searrow}(\text{col}_{p,q+1})) = 0$. Similarly to the derivation of Remark 3, one furthermore deduces that every $(x,y) \in \text{P}_{p,q}$ satisfies

$$y_{(i,0)^{\downarrow}} = 1 - y(\text{in}^{\searrow}(\text{col}_{i+1,1}))$$

for all $i \in [p-1]_0$. Finally, every $(x,y) \in \text{P}_{p,q}$ clearly satisfies $y_{(s,(0,0))} = 1$. Therefore, we can eliminate from the system describing $\text{P}_{p,q}$ all arc variables except for the ones corresponding to diagonal arcs.

We finally apply the linear transformation defined by

$$z_{ij} = x(\text{bar}_{i,j}) \quad \text{and} \quad w_{ij} = y(\text{in}^{\searrow}(\text{col}_{i,j})) \quad \text{for all } (i,j) \in \mathcal{I}_{p,q},$$

to $\mathbb{R}^{\mathcal{I}_{p,q}} \times \mathbb{R}^{A^{\searrow}_{p,q}}$, whose inverse is given by

$$x_{ij} = z_{i,j} - z_{i,j+1} \quad \text{for all } (i,j) \in \mathcal{I}_{p,q}.$$

(defining $z_{i,q(i)+1} = 0$, for all $i \in [p]$) and

$$y_{(i,j)^{\searrow}} = w_{i+1,j+1} - w_{i,j+1} \quad \text{for all } i \in [p-1]_0, j \in [q(i+1)-1]_0.$$

Few calculations are needed to check that the previous transformation (bijectively) maps $P_{p,q}$ onto the polytope $P_{p,q}^{\text{comp}} \subseteq \mathbb{R}^{\mathcal{I}_{p,q}} \times \mathbb{R}^{\mathcal{I}_{p,q}}$ defined by the following "very compact" set of constraints:

$$\begin{align}
(14) && w_{i+1,j+1} - w_{i,j+1} &\geq 0 && \text{for } i \in [p-1]_0, j \in [q(i+1)-1]_0 \\
(15) && w_{i,j} - w_{i+1,j+1} &\geq 0 && \text{for } (i,j) \in \mathcal{I}_{p,q}, i < p \\
(16) && w_{p,1} &\leq 1 \\
(17) && w_{i,j} - w_{i-1,j} - z_{ij} + z_{i,j+1} &\leq 0 && \text{for } (i,j) \in \mathcal{I}_{p,q} \\
(18) && z_{i,j} - w_{i,j} &\leq 0 && \text{for } (i,j) \in \mathcal{I}_{p,q} \\
(19) && w_{i,q(i)} &\geq 0 && \text{for } i \in [p]
\end{align}$$

Here, (14) represent the nonnegativity constraints on the diagonal arcs. Nonnegativity on the vertical arcs $(i,j)^{\downarrow}$ with $i \in [p-1]_0$ is reflected by (15) for $j \in [q(i)]$ and by (16) (together with the nonnegativity of w, which is implied by (19) and (15)) for $j = 0$. Finally, equations (5) and (6) translate to (17) and (18), respectively.

Ignoring the nonnegativity constraints, system (14)–(19) has less than $2pq$ variables and $4pq$ constraints, for a total number of nonzero coefficients that is smaller than $10pq$.

Note that $w_{1,1} \leq 1$ is a valid inequality for $P_{p,q}^{\text{comp}}$. The face of $P_{p,q}^{\text{comp}}$ defined by $w_{1,1} = 1$ is the image of the face $P_{p,q}^{=}$ of $P_{p,q}$. Thus, adding $w_{1,1} = 1$ to the system (14)–(19) one arrives at another extended formulation of $O_{p,q}^{=}$.

We summarize the results of this subsection.

Theorem 10. *The polytope* $P_{p,q}^{comp} \subseteq \mathbb{R}^{\mathcal{I}_{p,q}} \times \mathbb{R}^{\mathcal{I}_{p,q}}$ *defined by* (14)–(19) *is an extended formulation of* $O_{p,q}^{\leq}$. *The face of* $P_{p,q}^{comp}$ *defined by* $w_{1,1} = 1$ *is an extended formulation of* $O_{p,q}^{=}$.

4. THE PROJECTION

Let $Q_{p,q} \subseteq \mathbb{R}^{\mathcal{I}_{p,q}}$ be the polytope defined by the nonnegativity constraints $x \geq \mathbf{0}$, the row-sum inequalities $x(\text{row}_i) \leq 1$ for all $i \in [p]$ and all shifted-column inequalities. By checking the vertices (0/1-vectors) of $O_{p,q}^{\leq}$ it is easy to see that $O_{p,q}^{\leq} \subseteq Q_{p,q}$ holds. Thus, in order to prove

Theorem 11. $O_{p,q}^{\leq} = Q_{p,q}$

(which is Prop. 13 in [6]) it suffices (due to Theorem 6) to show the following:

Theorem 12. *For each $x \in Q_{p,q}$ there is some $y \in F_{p,q}$ with $(x,y) \in P_{p,q}$.*

Proof. For $x \in Q_{p,q}$ consider the network $D_{p,q}$ with

$$\text{capacity } x_{ij} \text{ on the diagonal arc } (i-1, j-1)^\searrow$$

for each $(i,j) \in \mathcal{I}_{p,q}$ and infinite capacities on all other arcs. In this network, we construct a feasible flow $y \in F_{p,q}$ of value one with the property

$$(20) \qquad\qquad y_{(i-1,j-1)^\downarrow} > 0 \quad \Rightarrow \quad y_{(i-1,j-1)^\searrow} = x_{ij}$$

for all $(i,j) \in \mathcal{I}_{p,q}$. Phrased verbally, the flow y uses a vertical arc only if the diagonal arc emanating from its tail is saturated. Such a flow can easily be constructed in the following way: start by sending one unit of flow from s to t along the vertical path in column zero. At each step, if the flow y constructed so far violates (20) for some $(i,j) \in \mathcal{I}_{p,q}$, choose such a pair (i,j) with minimal j, breaking ties by choosing i minimally as well. With

$$\vartheta = \min\{y_{(i-1,j-1)^\downarrow}, x_{ij} - y_{(i-1,j-1)^\searrow}\}$$

(i.e., the minimum of the flow on the vertical arc and the residual capacity on the diagonal arc starting at $(i-1, j-1)$) reroute ϑ units of the flow currently travelling on the vertical arc $(i-1, j-1)^\downarrow$ along the path starting with the diagonal arc $(i-1, j-1)^\searrow$ and then using the vertical arcs in column j. Note that this affects only arcs leaving nodes (k, ℓ) with $k \geq i$ and $\ell \geq j$. After this rerouting, (20) holds for (i,j). The minimality requirements in the choice of (i,j) ensure that the flow on the two arcs leaving (i,j) is not changed again afterwards. Thus, (20) will always be satisfied for (i,j) in the future. Therefore, the procedure eventually ends with a flow as required.

As (x,y) satisfies (5) by construction, it suffices to show (6) in order to prove $(x,y) \in P_{p,q}$. To this end, let $(i,j) \in \mathcal{I}_{p,q}$. Due to Remark 3, we only need to prove

$$(21) \qquad\qquad y(\mathrm{bar}_{i,j}) \geq x(\mathrm{bar}_{i,j}).$$

We construct a directed (k,ℓ)-(i,j)-path Γ in the residual network with respect to the flow y (containing only those arcs of $D_{p,q}$ that are not saturated by y) with $k = \ell$ or $\ell = 0$ in the following way: Starting from the trivial (length zero) w-(i,j)-path with $w = (i,j)$, in each step we extend the path at its current start node w by the diagonal arc entering w if this arc is part of the residual network, and by the vertical arc entering w otherwise. As the residual network contains all vertical arcs, we clearly can proceed this way until the start node of the current path is some node (k, ℓ) with $k = \ell$ or with $\ell = 0$.

Since Γ is a path in the residual network and due to (20), we have

$$(22) \qquad\qquad y(\mathrm{out}^\downarrow(\mathrm{T}(\Gamma))) = 0.$$

If $\ell = 0$, part (2) of Lemma 2 together with (22) yields $y(\mathrm{bar}_{i,j}) = 1$, from which (21) follows since x satisfies the row-sum inequality $x(\mathrm{row}_i) \leq 1$ and the nonnegativity constraints.

If $\ell \neq 0$, then $k = \ell \geq 1$. Thus, according to part (1) of Lemma 2 and due to (22), we have

$$(23) \qquad\qquad y(\mathrm{in}^{\searchangle}(\mathrm{S}(\Gamma))) = y(\mathrm{bar}_{i,j}) \,.$$

Since we preferred diagonal arcs from the residual network in our backwards construction of Γ, we find that all arcs from $\mathrm{in}^{\searchangle}(\mathrm{S}(\Gamma))$ are saturated by the flow y. Therefore, for the shifted column $S = \mathrm{S}(\Gamma)$, we have (using (23))

$$(24) \qquad\qquad x(S) = y(\mathrm{in}^{\searchangle}(\mathrm{S}(\Gamma)) = y(\mathrm{bar}_{i,j}) \,.$$

Let $a \in A_{p,q}$ be the arc in Γ entering (i,j), and denote by Γ' the path arising from Γ by removing a.

If a is diagonal, then the (ℓ,ℓ)-$(i-1, j-1)$-path Γ' satisfies $\mathrm{S}(\Gamma') = S$, and the shifted-column inequality $x(\mathrm{bar}_{i,j}) \leq x(S)$ (satisfied by x) establishes (21) via (24).

If a is vertical, by construction of Γ, we have $y_{(i-1,j-1)^{\searchangle}} = x_{ij}$. Furthermore, the (ℓ,ℓ)-$(i-1, j)$-path Γ' satisfies $\mathrm{S}(\Gamma') = S \setminus \{(i,j\})$. Thus using the shifted-column inequality $x(\mathrm{bar}_{i,j+1}) \leq x(S \setminus \{(i,j)\})$ one obtains from (24) the inequality

$$y(\mathrm{bar}_{i,j}) = x(S) = x(S \setminus \{(i,j)\}) + x_{ij} \geq x(\mathrm{bar}_{i,j+1}) + x_{ij} = x(\mathrm{bar}_{i,j}) \,.$$

Thus, (21) is established also in this case, which finally proves Theorem 12. $\square$

5. REMARKS

In our view, the extended formulations for the orbitopes $O_{p,q}^{\leq}$ and $O_{p,q}^{=}$ presented in this paper once more demonstrate the power that lies in the concept of extended formulations. Not only do the extended formulations provide a very compact way of describing the orbitopes, but also do they allow to derive rather simple proofs of the fact that nonnegativity constraints, row-sum inequalities/equations, and SCIs suffice in order to linearly describe $O_{p,q}^{\leq}$ and $O_{p,q}^{=}$.

To us it seems that these proofs better reveal the reason why SCIs are necessary and (basically) sufficient in these descriptions. The construction of the flow in the proof of Thm. 12 is quite natural. The rest of the proof (i.e., the backwards construction of the path Γ) one may also have done without knowing the SCIs in advance. Thus, knowing the extended formulation, one possibly could also have *detected* SCIs on the way trying to do this proof.

An interesting practical question is whether the very sparse and compact extended formulations for orbitopes lead to performance gains in branch-and-cut algorithms compared to versions that dynamically add SCIs via the linear time separation algorithm (described in [6]).

References

1. Egon Balas, Sebastián Ceria, and Gérard Cornuéjols, *A lift-and-project cutting plane algorithm for mixed 0-1 programs*, Math. Programming **58** (1993), no. 3, Ser. A, 295–324.
2. Michele Conforti, Marco Di Summa, Friedrich Eisenbrand, and Laurence A. Wolsey, *Network formulations of mixed integer programs*, CORE Discussion Paper 2006/117.
3. Michele Conforti, Marco Di Summa, and Laurence A. Wolsey, *The mixing set with flows*, SIAM J. Discrete Math. **21** (2007), no. 2, 396–407.
4. Michele Conforti, Bert Gerards, and Giacomo Zambelli, *Mixed-integer vertex covers on bipartite graphs*, Proceedings of IPCO XII (Matteo Fischetti and David Williamson, eds.), LNCS, vol. 4513, Springer-Verlag, 2007, pp. 324–336.
5. Volker Kaibel, Matthias Peinhardt, and Marc E. Pfetsch, *Orbitopal fixing*, Proceedings of IPCO XII (Matteo Fischetti and David Williamson, eds.), LNCS, vol. 4513, Springer-Verlag, 2007, pp. 74–88.
6. Volker Kaibel and Marc E. Pfetsch, *Packing and partitioning orbitopes*, Math. Programming, Ser. A **114** (2008), no. 1, 1–36.
7. Jeff Linderoth, James Ostrowski, Fabrizio Rossi, and Stefano Smriglio, *Orbital branching*, Proceedings of IPCO XII (Matteo Fischetti and David Williamson, eds.), LNCS, vol. 4513, Springer-Verlag, 2007, pp. 106–120.
8. _____ , *Constraint orbital branching*, Proceedings of IPCO XIII (Andrea Lodi and Giovanni Rinaldi, eds.), LNCS, Springer-Verlag, 2008, to appear.
9. L. Lovász and A. Schrijver, *Cones of matrices and set-functions and 0-1 optimization*, SIAM J. Optim. **1** (1991), no. 2, 166–190.
10. Fránćois Margot, *Pruning by isomorphism in branch-and-cut*, Math. Program. **94** (2002), no. 1, 71–90.
11. _____ , *Exploiting orbits in symmetric ILP*, Math. Program. **98** (2003), no. 1–3, 3–21.
12. _____ , *Small covering designs by branch-and-cut*, Math. Program. **94** (2003), no. 2–3, 207–220.
13. _____ , *Symmetric ILP: Coloring and small integers*, Discrete Opt. **4** (2007), no. 1, 40–62.
14. R. Kipp Martin, Ronald L. Rardin, and Brian A. Campbell, *Polyhedral characterization of discrete dynamic programming*, Oper. Res. **38** (1990), no. 1, 127–138.
15. Hanif D. Sherali and Warren P. Adams, *A hierarchy of relaxations and convex hull characterizations for mixed-integer zero-one programming problems*, Discrete Appl. Math. **52** (1994), no. 1, 83–106.

(Faenza) Dipartimento di Ingegneria dell'Impresa, Università di Roma "Tor Vergata", Rome, Italy

E-mail address, Faenza: `faenza@disp.uniroma2.it`

(Kaibel) Otto-von-Guericke Universität Magdeburg, Fakultät für Mathematik, Universitätsplatz 2, 39106 Magdeburg, Germany

E-mail address, Kaibel: `kaibel@ovgu.de`