PACKING AND PARTITIONING ORBITOPES

VOLKER KAIBEL AND MARC E. PFETSCH

Abstract. We introduce *orbitopes* as the convex hulls of 0/1-matrices that are lexicographically maximal subject to a group acting on the columns. Special cases are packing and partitioning orbitopes, which arise from restrictions to matrices with at most or exactly one 1-entry in each row, respectively. The goal of investigating these polytopes is to gain insight into ways of breaking certain symmetries in integer programs by adding constraints, e.g., for a well-known formulation of the graph coloring problem.

We provide a thorough polyhedral investigation of packing and partitioning orbitopes for the cases in which the group acting on the columns is the cyclic group or the symmetric group. Our main results are complete linear inequality descriptions of these polytopes by facet-defining inequalities. For the cyclic group case, the descriptions turn out to be totally unimodular, while for the symmetric group case, both the description and the proof are more involved. The associated separation problems can be solved in linear time.

1. Introduction

Symmetries are ubiquitous in discrete mathematics and geometry. They are often responsible for the tractability of algorithmic problems and for the beauty of both the investigated structures and the developed methods. It is common knowledge, however, that the presence of symmetries in integer programs may severely harm the ability to solve them. The reasons for this are twofold. First, the use of branch-and-bound methods usually leads to an unnecessarily large search tree, because equivalent solutions are found again and again. Second, the quality of LP relaxations of such programs typically is extremely poor.

A classical approach to "break" such symmetries is to add constraints that cut off equivalent copies of solutions, in hope to resolve these problems. There are numerous examples of this in the literature; we will give a few references for the special case of graph coloring below. Another approach was developed by Margot [11, 12]. He studies a branch-and-cut method that ensures to investigate only one representative of each class of equivalent solutions by employing methods from computational group theory. Furthermore, the symmetries are also used to devise cutting planes. Methods for symmetry breaking in the context of constraint programming have been developed, for instance, by Fahle, Schamberger, and Sellmann [7] and Puget [16].

Date: 11/22/2006.

2000 *Mathematics Subject Classification.* Primary 90C10; Secondary 90C57, 52B12.

Key words and phrases. integer programming, symmetry breaking, lexicographic representatives.

Supported by the DFG Research Center Matheon in Berlin.

The main goal of this paper is to start an investigation of the polytopes that are associated with certain symmetry breaking inequalities. In order to clarify the background, we first discuss the example of a well-known integer programming (IP) formulation for the graph coloring problem.

Let $G = (V, E)$ be a loopless undirected graph without isolated nodes. A *(vertex) coloring* of G using at most C colors is an assignment of colors $\{1, \dots, C\}$ to the nodes such that no two adjacent nodes receive the same color. The *graph coloring* problem is to find a vertex coloring with as few colors as possible. This is one of the classical NP-hard problems [9]. It is widely believed to be among the hardest problems in combinatorial optimization. In the following classical IP formulation, $V = \{1, \dots, n\}$ are the nodes of G and C is some upper bound on the number of colors needed.

$$\begin{aligned} \min \quad & \sum_{j=1}^{C} y_j & & \\ & x_{ij} + x_{kj} \le y_j & \{i,k\} \in E,\ j \in \{1,\dots,C\} \quad & \text{(i)} \\ & \sum_{j=1}^{C} x_{ij} = 1 & i \in V \quad & \text{(ii)} \\ & x_{ij} \in \{0,1\} & i \in V,\ j \in \{1,\dots,C\} \quad & \text{(iii)} \\ & y_j \in \{0,1\} & j \in \{1,\dots,C\} \quad & \text{(iv)} \end{aligned} \tag{1}$$

In this model, variable x_{ij} is 1 if and only if color j is assigned to node i and variable y_j is 1 if color j is used. Constraints (i) ensure that color j is assigned to at most one of the two adjacent nodes i and k; it also enforces that y_j is 1 if color j is used, because there are no isolated nodes. Constraints (ii) guarantee that each node receives exactly one color.

It is well known that this formulation exhibits symmetry: Given a solution (x, y), any permutation of the colors, i.e., the columns of x (viewed as an $n \times C$-matrix) and the components of y, results in a valid solution with the same objective function value. Viewed abstractly, the symmetric group of order C acts on the solutions (x, y) (by permuting the columns of x and the components of y) in such a way that the objective function is constant along every orbit of the group action. Each orbit corresponds to a symmetry class of feasible colorings of the graph. Note that "symmetry" here always refers to the symmetry of permuting colors, not to symmetries of the graph.

The weakness of the LP-bound mentioned above is due to the fact that the point $(x^\star, y^\star)$ with $x^\star_{ij} = 1/C$ and $y^\star_j = 2/C$ is feasible for the LP relaxation with objective function value 2. The symmetry is responsible for the feasibility of $(x^\star, y^\star)$, since $x^\star$ is the barycenter of the orbit of an arbitrary $x \in \{0,1\}^{n \times C}$ satisfying (ii) in (1).

It turned out that the symmetries make the above IP-formulation for the graph coloring problem difficult to solve. One solution is to develop different formulations for the graph coloring problem. This line has been pursued, e.g., by Mehrotra and Trick [13], who devised a column generation approach. See Figueiredo, Barbosa, Maculan, and de Souza [8] and Cornaz [5] for alternative models.

Another solution is to enhance the IP-model by additional inequalities that cut off as large parts of the orbits as possible, keeping at least one element of each orbit in the feasible region. Méndez-Díaz and Zabala [15] showed that a branch-and-cut algorithm using this kind of symmetry breaking inequalities performs well in practice. The polytope corresponding to (1) was investigated by Campêlo, Corrêa, and Frota [3] and Coll, Marenco, Méndez-Díaz, and Zabala [4]. Ramani, Aloul, Markov, and Sakallah [17] studied symmetry breaking in connection with SAT-solving techniques to solve the graph coloring problem.

The strongest symmetry breaking constraints that Méndez-Díaz and Zabala [14, 15] introduced are the inequalities

$$x_{ij} - \sum_{k=1}^{i-1} x_{k,j-1} \leq 0, \qquad \text{for all } i \text{ and } j \geq 2. \tag{2}$$

From each orbit, they cut off all points except for one representative that is the maximal point in the orbit with respect to a lexicographic ordering. A solution (x, y) of the above IP-model is such a representative if and only if the columns of x are in decreasing lexicographic order. We introduce a generalization and strengthening of Inequalities (2) in Section 4.1.

Breaking symmetries by adding inequalities like (2) does not depend on the special structure of the graph coloring problem. These inequalities single out the lexicographic maximal representative from each orbit (with respect to the symmetric group acting on the columns) of the whole set of all 0/1-matrices with exactly one 1-entry per row. The goal of this paper is to investigate the structure of general "symmetry breaking polytopes" like the convex hull of these representatives. We call these polytopes *orbitopes*. The idea is that general knowledge on orbitopes (i.e., valid inequalities) can be utilized for different symmetric IPs in order to address both the difficulties arising from the many equivalent solutions and from the poor LP-bounds. In particular with respect to the second goal, for concrete applications it will be desirable to combine the general knowledge on orbitopes with concrete polyhedral knowledge on the problem under investigation in oder to derive strengthened inequalities. For the example of graph coloring, we indicate that (and how) this can be done in Section 5. Figure 1 illustrates the geometric situation.

The case of a symmetric group acting on the columns is quite important. It does not only appear in IP-formulations for the graph coloring problem, but also in many other contexts like, e.g., block partitioning of matrices [1], k-partitioning in the context of frequency assignment [6], or line-planning in public transport [2]. However, other groups are interesting as well. For instance, in the context of timetabling in public transport systems [19], cyclic groups play an important role.

We thus propose to study different types of orbitopes, depending on the group acting on the columns of the variable-matrix and on further restrictions like the number of 1-entries per row being exactly one (*partitioning*), at most one (*packing*), at least one (*covering*), or arbitrary (*full*).

The main results of this paper are complete and irredundant linear descriptions of packing and partitioning orbitopes for both the symmetric group and for the cyclic group acting on the columns of the variable-matrix. We also provide (linear time) separation algorithms for the corresponding sets of inequalities. While this work lays the theoretical foundations on orbitopes,

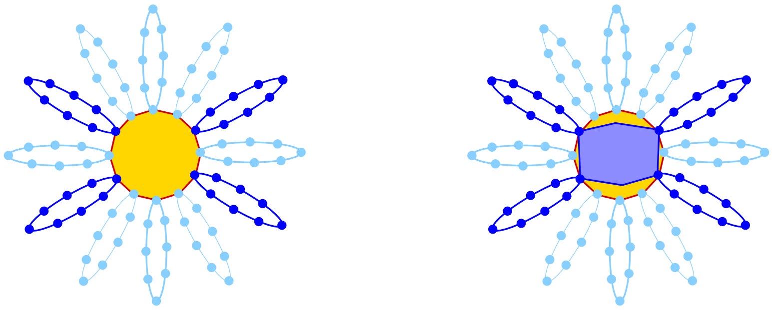


Figure 1: Breaking symmetries by orbitopes. The left figure illustrates an orbitope, i.e., the convex hull of the representatives of a large system of orbits. For a concrete problem, like graph coloring, only a subset of the orbits are feasible (the dark orbits). Combining a (symmetric) IP-formulation for the concrete problem with the orbitope removes the symmetry from the formulation (right figure).

a thorough computational investigation of the practical usefulness of the results will be the subject of further studies (see also the remarks in Section 5).

The outline of the paper is as follows. In Section 2, we introduce some basic notations and define orbitopes. In Section 2.1 we show that optimization over packing and partitioning orbitopes for symmetric and cyclic groups can be done in polynomial time. In Section 3 we give complete (totally unimodular) linear descriptions of packing and partitioning orbitopes for cyclic groups. Section 4 deals with packing and partitioning orbitopes for symmetric groups, which turn out to be more complicated than their counterparts for cyclic groups. Here, besides (strengthenings of) Inequalities (2), one needs exponentially many additional inequalities, the "shifted column inequalities", which are introduced in Section 4.2. We show that the corresponding separation problem can be solved in linear time, see Section 4.3. Section 4.4 gives a complete linear description, and Section 4.5 investigates the facets of the polytopes. We summarize the results for symmetric groups in Section 4.6 for easier reference. Finally, we close with some remarks in Section 5.

2. Orbitopes: General Definitions and Basic Facts

We first introduce some basic notation. For a positive integer n, we define $[n] := \{1, 2, \dots, n\}$. We denote by $\mathbf{0}$ the 0-matrix or 0-vector of appropriate sizes. Throughout the paper let p and q be positive integers. For $x \in \mathbb{R}^{[p]\times[q]}$ and $S \subseteq [p] \times [q]$, we write

$$x(S) := \sum_{(i,j)\in S} x_{ij}.$$

For convenience, we use $S - (i, j)$ for $S \setminus \{(i, j)\}$ and $S + (i, j)$ for $S \cup \{(i, j)\}$, where $S \subseteq [p] \times [q]$ and $(i, j) \in [p] \times [q]$. If p and q are clear from the context, then $\text{row}_i := \{(i, 1), (i, 2), \dots, (i, q)\}$ are the entries of the ith row.

Let $\mathcal{M}_{p,q} := \{0, 1\}^{[p]\times[q]}$ be the set of 0/1-matrices of size $p \times q$. We define

- $\mathcal{M}^{\leq}_{p,q} := \{x \in \mathcal{M}_{p,q} : x(\text{row}_i) \leq 1 \text{ for all } i\}$
- $\mathcal{M}^{=}_{p,q} := \{x \in \mathcal{M}_{p,q} : x(\text{row}_i) = 1 \text{ for all } i\}$
- $\mathcal{M}^{\geq}_{p,q} := \{x \in \mathcal{M}_{p,q} : x(\text{row}_i) \geq 1 \text{ for all } i\}$.

Let $\prec$ be the lexicographic ordering of $\mathcal{M}_{p,q}$ with respect to the ordering

$$(1,1) < (1,2) < \cdots < (1,q) < (2,1) < (2,2) < \cdots < (2,q) < \cdots < (p,q)$$

of matrix positions, i.e., $A \prec B$ with $A = (a_{ij}), B = (b_{ij}) \in \mathcal{M}_{p,q}$ if and only if $a_{k\ell} < b_{k\ell}$, where (k,ℓ) is the first position (with respect to the ordering above) where A and B differ.

Let $\mathfrak{S}_n$ be the group of all permutations of $[n]$ (*symmetric group*) and let G be a subgroup of $\mathfrak{S}_q$, acting on $\mathcal{M}_{p,q}$ by permuting columns. Let $\mathcal{M}^{\max}_{p,q}(G)$ be the set of matrices of $\mathcal{M}_{p,q}$ that are $\prec$-maximal within their orbits under the group action G.

We can now define the basic objects of this paper.

Definition 1 (Orbitopes)**.**
(1) *The* full orbitope *associated with the group G is*

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

(2) *We associate with the group G the following restricted orbitopes:*

$$\begin{aligned} \mathrm{O}^{\leq}_{p,q}(G) &:= \mathrm{conv}(\mathcal{M}^{\max}_{p,q}(G) \cap \mathcal{M}^{\leq}_{p,q}) \quad \text{(packing orbitope)}\\ \mathrm{O}^{=}_{p,q}(G) &:= \mathrm{conv}(\mathcal{M}^{\max}_{p,q}(G) \cap \mathcal{M}^{=}_{p,q}) \quad \text{(partitioning orbitope)}\\ \mathrm{O}^{\geq}_{p,q}(G) &:= \mathrm{conv}(\mathcal{M}^{\max}_{p,q}(G) \cap \mathcal{M}^{\geq}_{p,q}) \quad \text{(covering orbitope)} \end{aligned}$$

Remark. By definition, $\mathrm{O}^{=}_{p,q}(G)$ is a face of both $\mathrm{O}^{\leq}_{p,q}(G)$ and $\mathrm{O}^{\geq}_{p,q}(G)$.

In this paper, we will be only concerned with the cases of G being the *cyclic group* $\mathfrak{C}_q$ containing all q cyclic permutations of $[q]$ (Section 3) or the symmetric group $\mathfrak{S}_q$ (Section 4). Furthermore, we will restrict attention to packing and partitioning orbitopes. For these, we have the following convenient characterizations of vertices:

Observation 1.
(1) *A matrix of $\mathcal{M}_{p,q}$ is contained in $\mathcal{M}^{\max}_{p,q}(\mathfrak{S}_q)$ if and only if its columns are in non-increasing lexicographic order (with respect to the order $\prec$ defined above).*
(2) *A matrix of $\mathcal{M}^{\leq}_{p,q}$ is contained in $\mathcal{M}^{\max}_{p,q}(\mathfrak{C}_q)$ if and only if its first column is lexicographically not smaller than the remaining ones (with respect to the order $\prec$).*
(3) *In particular, a matrix of $\mathcal{M}^{=}_{p,q}$ is contained in $\mathcal{M}^{\max}_{p,q}(\mathfrak{C}_q)$ if and only if it has a 1-entry at position $(1,1)$.*

2.1. Optimizing over Orbitopes

The main aim of this paper is to provide complete descriptions of $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$, $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$, $\mathrm{O}^{=}_{p,q}(\mathfrak{C}_q)$, and $\mathrm{O}^{\leq}_{p,q}(\mathfrak{C}_q)$ by systems of linear equations and linear inequalities. If these orbitopes admit "useful" linear descriptions then the corresponding linear optimization problems should be solvable efficiently, due to the equivalence of optimization and separation, see Grötschel, Lovász, and Schrijver [10].

We start with the cyclic group operation, since the optimization problem is particularly easy in this case.

Theorem 1. *Both the linear optimization problem over* $\mathcal{M}^{\max}_{p,q}(\mathfrak{C}_q) \cap \mathcal{M}^{\leq}_{p,q}$ *and over* $\mathcal{M}^{\max}_{p,q}(\mathfrak{C}_q) \cap \mathcal{M}^{=}_{p,q}$ *can be solved in time* $O(pq)$.

Proof. We first give the proof for the packing case.

For a vector $c \in \mathbb{Q}^{[p]\times[q]}$, we consider the linear objective function

$$\langle c, x\rangle := \sum_{i=1}^{p}\sum_{j=1}^{q} c_{ij}\, x_{ij}.$$

The goal is to find a matrix $A^\star \in \mathcal{M}^{\max}_{p,q}(\mathfrak{C}_q) \cap \mathcal{M}^{\leq}_{p,q}$ such that $\langle c, A^\star\rangle$ is maximal. Let $A^\star$ be such a c-maximal matrix, and let $a^\star \in \{0,1\}^p$ be its first column. If $a^\star = \mathbf{0}$, then $A^\star = \mathbf{0}$ by Part (2) of Observation 1. By the same observation it follows that if $a^\star \neq \mathbf{0}$ and $i^\star \in [p]$ is the minimum row-index i with $a^\star_i = 1$, then $A^\star$ has only zero entries in its first $i^\star$ rows, except for the 1-entry at position $(i^\star, 1)$ (there is at most one 1-entry in each row). Furthermore, each row $i > i^\star$ of $A^\star$ either has no 1-entry or it has its (unique) 1-entry at some position where c is maximal in row i.

Thus, we can compute an optimal solution as follows: (1) For each $i \in [p]$ determine a vector $b^i \in \{0,1\}^q$ that is the zero vector if c does not have any positive entries in row i and otherwise is the j-th standard unit vector, where $j \in [q]$ is chosen such that $c_{ij} = \max\{c_{i\ell} : \ell \in [q]\}$; set $\sigma_i := 0$ in the first case and $\sigma_i := c_{ij}$ in the second. (2) Compute the values $s_p := \sigma_p$ and $s_i := \sigma_i + s_{i+1}$ for all $i = p-1, p-2, \dots, 1$. (3) Determine $i^\star$ such that $c_{i^\star,1} + s_{i^\star+1}$ is maximal among $\{c_{i,1} + s_{i+1} : i \in [p]\}$. (4) If $c_{i^\star,1} + s_{i^\star+1} \leq 0$, then $\mathbf{0}$ is an optimal solution. Otherwise, the matrix whose i-th row equals b^i for $i \in \{i^\star+1, \dots, p\}$ and which is all-zero in the first $i^\star$ rows, except for a 1-entry at position $(i^\star, 1)$, is optimal.

From the description of the algorithm it is easy to see that its running time is bounded by $O(pq)$ (in the unit-cost model).

The partitioning case is then straightforward and even becomes easier due to Part (3) of Observation 1. □

Theorem 2. *Both the linear optimization problem over* $\mathcal{M}^{\max}_{p,q}(\mathfrak{S}_q) \cap \mathcal{M}^{\leq}_{p,q}$ *and over* $\mathcal{M}^{\max}_{p,q}(\mathfrak{S}_q) \cap \mathcal{M}^{=}_{p,q}$ *can be solved in time* $O(p^2 q)$.

Proof. We give the proof for the partitioning case, indicating the necessary modifications for the packing case at the relevant points.

As in the proof of Theorem 1, we maximize the linear objective function given by $\langle c, x\rangle$ for $c \in \mathbb{Q}^{[p]\times[q]}$. We describe a two-step approach.

In the first step, for $i_1, i_2 \in [p]$ with $i_1 \leq i_2$ and $j \in [q]$, we let $M(i_1, i_2, j)$ be c-maximal among the matrices in $\{0,1\}^{\{i_1, i_1+1, \dots, i_2\}\times[j]}$ with exactly (in the packing case: at most) one 1-entry in every row. Denote by $\mu(i_1, i_2, j)$ the c-value of $M(i_1, i_2, j)$, i.e.,

$$\mu(i_1, i_2, j) = \sum_{k=i_1}^{i_2}\sum_{\ell=1}^{j} c_{k\ell}\, M(i_1, i_2, j)_{k\ell}\,.$$

The values $\mu(i_1, i_2, j)$ can be computed in time $O(p^2 q)$ as follows. First, we compute all numbers $\lambda(i, j) = \max\{c_{i\ell} : \ell \in [j]\}$ (in the packing case: $\lambda(i, j) = \max(0, \{c_{i\ell} : \ell \in [j]\})$) for all $i \in [p]$ and $j \in [q]$. This can clearly be done in $O(pq)$ steps by using the recursions $\lambda(i, j) = \max\{\lambda(i, j-1), c_{ij}\}$

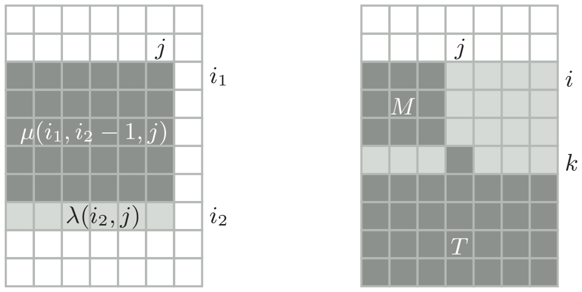


Figure 2: Illustration of the proof of Theorem 2. *Left:* Computation of $\mu(i_1, i_2, j)$. *Right:* Computation of $\tau(i, j)$ via the dynamic programming relation (3). Indicated are the matrix $M(i, k-1, j-1)$ and corresponding term $\mu(i, k-1, j-1)$ and matrix $T(k+1, j+1)$ with corresponding term $\tau(k+1, j+1)$.

for $j \geq 2$. Then, after initializing $\mu(i, i, j) = \lambda(i, j)$ for all $i \in [p]$ and $j \in [q]$, one computes $\mu(i_1, i_2, j) = \mu(i_1, i_2 - 1, j) + \lambda(i_2, j)$ for all $j \in [q]$, $i_1 = 1, 2, \dots, p$, and $i_2 = i_1 + 1, i_1 + 2, \dots, q$; see Figure 2.

In the second step, for $i \in [p]$ and $j \in [q]$, let $T(i, j)$ be c-maximal among the matrices in $\{0, 1\}^{\{i, i+1, \dots, p\} \times [q]}$ with exactly (in the packing case: at most) one 1-entry in every row and with columns $j, j + 1, \dots, q$ being in non-increasing lexicographic order. Thus, by Part (1) of Observation 1, $T(1, 1)$ is an optimal solution to our linear optimization problem. Denote by $\tau(i, j)$ the c-value of $T(i, j)$, i.e.,

$$\tau(i, j) = \sum_{k=i}^{p} \sum_{\ell=1}^{q} c_{k\ell}\, T(i, j)_{k\ell}.$$

Let $k \in \{i, i+1, \dots, p+1\}$ be the index of the first row, where $T(i, j)$ has a 1-entry in column j (with $k = p+1$ if there is no such 1-entry); see Figure 2. Then $T(i, j)$ has a c-maximal matrix T in rows $k+1, \dots, p$ with exactly (in the packing case: at most) one 1-entry per row and lexicographically sorted columns $j+1, \dots, q$ (contributing $\tau(k+1, j+1)$). In row k, there is a single 1-entry at position (k, j) (contributing c_{kj}). And in rows $i, \dots, k - 1$, we have a c-maximal matrix M with exactly (in the packing case: at most) one 1-entry per row in the first $j-1$ columns (contributing $\mu(i, k-1, j-1)$) and zeroes in the remaining columns. Therefore, we obtain

$$\tau(i, j) = \mu(i, k - 1, j - 1) + c_{kj} + \tau(k + 1, j + 1).$$

Hence, considering all possibilities for k, we have

$$\tau(i, j) = \max \{\, \mu(i, k - 1, j - 1) + c_{kj} + \tau(k + 1, j + 1) \;: \; k \in \{i, i + 1, \dots, p + 1\} \}, \tag{3}$$

for all $i \in [p]$ and $j \in [q]$. For convenience we define $\mu(k_1, k_2, 0) = 0$ for $k_1, k_2 \in [p]$ with $k_1 \leq k_2$ and $\mu(k, k - 1, \ell) = 0$ for all $k \in [p]$ and $\ell \in \{0, 1, \dots, q\}$. Furthermore, we set $c_{p+1,\ell} = 0$ for all $\ell \in [q]$. Finally, we define $\tau(p+2, \ell) = \tau(p+1, \ell) = \tau(k, q+1) = 0$ for all $k \in [p]$ and $\ell \in [q+1]$.

Thus, by dynamic programming, we can compute the table $\tau(i, j)$ via Equation (3) in the order $i = p, p - 1, \dots, 1$, $j = q, q - 1, \dots, 1$. For each pair (i, j) the evaluation of (3) requires no more than $O(p)$ steps, yielding a total running time bound of $O(p^2 q)$.

Furthermore, if during these computations for each (i,j) we store a maximizer $k(i,j)$ for k in (3), then we can easily reconstruct the optimal solution $T(1,1)$ from the k-table without increasing the running time asymptotically: For $i \in [p]$, $j \in [q]$ the matrix $T(i,j)$ is composed of $M(i,k(i,j)-1,j-1)$ (if $k(i,j) \geq i+1$ and $j \geq 2$), $T(k(i,j)+1,j+1)$ (if $k(i,j) \leq p-1$ and $j \leq q-1$), and having 0-entries everywhere else, except for a 1-entry at position $(k(i,j),j)$ (if $k(i,j) \leq p$). Each single matrix $M(i_1,i_2,j)$ can be computed in $O((i_2-i_1)j)$ steps. Furthermore, for the matrices $M(i_1,i_2,j)$ needed during the recursive reconstruction of $T(1,1)$, the sets $\{i_1,\dots,i_2\}\times[j]$ are pairwise disjoint (see Figure 2). Thus, these matrices all together can be computed in time $O(pq)$. At the end there might be a single $T(k,q+1)$ to be constructed, which trivially can be done in $O(pq)$ steps. □

Thus, with respect to complexity theory there are no "obstructions" to finding complete linear descriptions of packing and partitioning orbitopes for both the cyclic and the symmetric group action. In fact, for cyclic group actions we will provide such a description in Theorem 3 and Theorem 4 for the partitioning and packing case, respectively. For symmetric group actions we will provide such a description for partitioning orbitopes in Theorems 16 and for packing orbitopes in Theorem 17. The algorithm used in the proof of Theorem 1 (for cyclic groups) is trivial, while the one described in the proof of Theorem 2 (for symmetric groups) is a bit more complicated. This is due to the simpler characterization of the cyclic case in Observation 1 and is reflected by the fact that the proofs of Theorems 16 and 17 (for symmetric groups) need much more work than the ones of Theorems 3 and 4 (for cyclic groups).

The algorithms described in the above two proofs heavily rely on the fact that we are considering only matrices with at most one 1-entry per row. For cyclic group operations, the case of matrices with more ones per row becomes more involved, because we do not have a simple characterization (like the one given in parts 2 and 3 of Observation 1) of the matrices in $\mathcal{M}^{\max}_{p,q}(\mathfrak{C}_q)$ anymore. For the action of the symmetric group, though we still have the characterization provided by Part (1) of Observation 1, the dynamic programming approach used in the proof of Theorem 2 cannot be adapted straight-forwardly without resulting in an exponentially large dynamic programming table (unless q is fixed). These difficulties apparently are reflected in the structures of the corresponding orbitopes (see the remarks in Section 5).

3. Packing and Partitioning Orbitopes for Cyclic Groups

From the characterization of the vertices in parts (2) and (3) of Observation 1 one can easily derive IP-formulations of both the partitioning orbitope $\mathrm{O}^{=}_{p,q}(\mathfrak{C}_q)$ and the packing orbitope $\mathrm{O}^{\leq}_{p,q}(\mathfrak{C}_q)$ for the cyclic group $\mathfrak{C}_q$. In fact, it turns out that these formulations do already provide linear descriptions of the two polytopes, i.e., they are totally unimodular. We refer the reader to Schrijver [18, Chap. 19] for more information on total unimodularity.

It is easy to see that for the descriptions given in Theorems 3 and 4 below, the separation problem can be solved in time $O(pq)$.

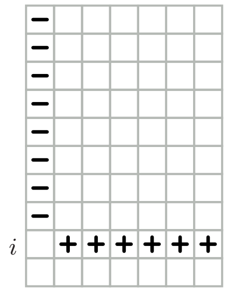


Figure 3: Example of the coefficient vector for an inequality of type (4); "−" stands for a −1, "+" for a +1.

Theorem 3. *The partitioning orbitope* $\mathrm{O}^{=}_{p,q}(\mathfrak{C}_q)$ *for the cyclic group* $\mathfrak{C}_q$ *equals the set of all* $x \in \mathbb{R}^{[p]\times[q]}$ *that satisfy the following linear constraints:*

- *the equations* $x_{11} = 1$ *and* $x_{1j} = 0$ *for all* $2 \leq j \leq q$,
- *the nonnegativity constraints* $x_{ij} \geq 0$ *for all* $2 \leq i \leq p$ *and* $j \in [q]$,
- *the row-sum equations* $x(\mathrm{row}_i) = 1$ *for all* $2 \leq i \leq p$.

This system of constraints is non-redundant.

Proof. The constraints $x(\mathrm{row}_i) = 1$ for $i \in [p]$ and $x_{ij} \geq 0$ for $i \in [p], j \in [q]$ define an integral polyhedron, since they describe a transshipment problem (and thus, the coefficient matrix is totally unimodular). Hence, the constraint system given in the statement of the theorem describes an integer polyhedron, because it defines a face of the corresponding transshipment polytope.

By Part (3) of Observation 1, the set of integer points satisfying this constraint system is $\mathcal{M}^{=}_{p,q} \cap \mathcal{M}^{\max}_{p,q}(\mathfrak{C}_q)$. Hence the given constraints completely describe $\mathrm{O}^{=}_{p,q}(\mathfrak{C}_q)$. The non-redundancy follows from the fact that dropping any of the constraints enlarges the set of feasible integer solutions. □

Theorem 4. *The packing orbitope* $\mathrm{O}^{\leq}_{p,q}(\mathfrak{C}_q)$ *for the cyclic group* $\mathfrak{C}_q$ *equals the set of all* $x \in \mathbb{R}^{[p]\times[q]}$ *that satisfy the following linear constraints:*

- *the constraints* $0 \leq x_{11} \leq 1$ *and* $x_{1j} = 0$ *for all* $2 \leq j \leq q$,
- *the nonnegativity constraints* $x_{ij} \geq 0$ *for all* $2 \leq i \leq p$ *and* $j \in [q]$,
- *the row-sum inequalities* $x(\mathrm{row}_i) \leq 1$ *for all* $2 \leq i \leq p$,
- *the inequalities*
$$\sum_{j=2}^{q} x_{ij} - \sum_{k=1}^{i-1} x_{k1} \leq 0 \tag{4}$$
for all $2 \leq i \leq p$ *(see Figure 3 for an example).*

This system of constraints is non-redundant.

Proof. From Part (2) of Observation 1 it follows that an integer point is contained in $\mathrm{O}^{\leq}_{p,q}(\mathfrak{C}_q)$ if and only if it satisfies the constraints described in the statement, where Inequalities (4) ensure that the first column of x is lexicographically not smaller than the other ones (note that we have at most one 1-entry in each row of x). Dropping any of the constraints enlarges

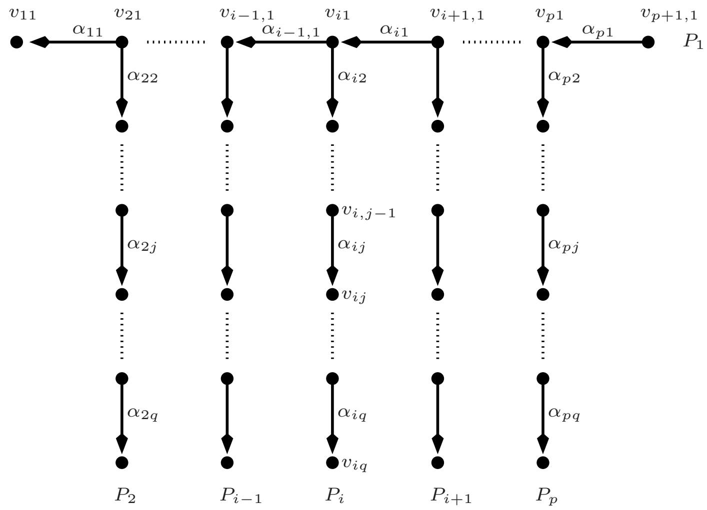


Figure 4: The network matrix constructed in the proof of Theorem 4.

the set of integer solutions, which proves the statement on non-redundancy. Thus, as in the proof of the previous theorem, it remains to show that the polyhedron defined by the constraints is integral. We prove this by showing that the coefficient matrix A of the row-sum inequalities $x(\text{row}_i) \leq 1$ (for $2 \leq i \leq p$) and Inequalities (4) (for all $2 \leq i \leq p$) is a network matrix (and thus, totally unimodular). Adding the nonnegativity constraints amounts to adding an identity matrix and preserves total unimodularity, which also holds for the inclusion of $x_{11} \leq 1$ into the system.

In order to establish the claim on the network structure of A, we will identify a directed tree T, whose arcs are in bijection with $[p] \times [q]$ (the set of indices of the columns of A), such that there are pairs of nodes (v_r, w_r) of T in bijection with the row indices $r \in [2(p-1)]$ of A with the following property. The matrix A has a $(+1)$-entry in row r and column (i,j), if the unique path π_r from node v_r to node w_r in the tree T uses arc (i,j) in its direction from i to j, a (-1)-entry, if π_r uses (i,j) in its reverse direction, and a 0-entry, if π_r does not use (i,j).

For the construction of the tree T, we take a directed path P_1 of length p on nodes $\{v_{11}, v_{21}, \dots, v_{p+1,1}\}$ with arcs $\alpha_{i1} := (v_{i+1,1}, v_{i1})$ for $i \in [p]$; see Figure 4. For each $2 \leq i \leq p$, we append a directed path P_i of length $q-1$ to node v_{i1}, where P_i has node set $\{v_{i1}, v_{i2}, \dots, v_{iq}\}$ and arcs $\alpha_{ij} := (v_{i,j-1}, v_{ij})$ for $2 \leq j \leq q$. Choosing the pair $(v_{i+1,1}, v_{iq})$ for the i-th row sum-inequality and the pair (v_{11}, v_{iq}) for the i-th Inequality (4), finishes the proof (using the bijection between the arcs of T and the columns of A indicated by the notation α_{ij}). $\square$

4. Packing and Partitioning Orbitopes for Symmetric Groups

For packing orbitopes $\text{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ and partitioning orbitopes $\text{O}^{=}_{p,q}(\mathfrak{S}_q)$ with respect to the symmetric group it follows readily from the characterization in Part (1) of Observation 1 that the equations

$$x_{ij} = 0 \qquad \text{for all } i < j \tag{5}$$

are valid. Thus, we may drop all variables corresponding to components in the upper right triangle from the formulation and consider

$$\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q),\ \mathrm{O}^{=}_{p,q}(\mathfrak{S}_q) \subset \mathbb{R}^{\mathcal{I}_{p,q}} \qquad \text{with} \quad \mathcal{I}_{p,q} := \{(i,j) \in [p] \times [q] \,:\, i \geq j\}.$$

We also adjust the definition of

$$\mathrm{row}_i := \{(i,1),(i,2),\dots,(i,\min\{i,q\})\} \qquad \text{for } i \in [p]$$

and define the jth column for $j \in [q]$ as

$$\mathrm{col}_j := \{(j,j),(j+1,j),\dots,(p,j)\}.$$

Furthermore, we restrict ourselves to the case

$$p \geq q \geq 2$$

in this context. Because of (5), the case of $q > p$ can be reduced to the case $p = q$ and the case of $q = 1$ is of no interest.

The next result shows a very close relationship between packing and partitioning orbitopes for the case of symmetric group actions.

Proposition 5. *The polytopes* $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$ *and* $\mathrm{O}^{\leq}_{p-1,q-1}(\mathfrak{S}_{q-1})$ *are affinely isomorphic via orthogonal projection of* $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$ *onto the space*

$$\mathcal{L} := \{x \in \mathbb{R}^{\mathcal{I}_{p,q}} \,:\, x_{i1} = 0 \text{ for all } i \in [p]\}$$

(*and the canonical identification of this space with* $\mathbb{R}^{\mathcal{I}_{p-1,q-1}}$).

Proof. The affine subspace

$$\mathcal{A} := \{x \in \mathbb{R}^{\mathcal{I}_{p,q}} \,:\, x(\mathrm{row}_i) = 1 \text{ for all } i\}$$

of $\mathbb{R}^{\mathcal{I}_{p,q}}$ clearly contains $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$. Let $\pi : \mathcal{A} \to \mathbb{R}^{\mathcal{I}_{p-1,q-1}}$ be the orthogonal projection mentioned in the statement (identifying $\mathcal{L}$ in the canonical way with $\mathbb{R}^{\mathcal{I}_{p-1,q-1}}$); note that the first row is removed since it only contains the element $(1,1)$. Consider the linear map $\phi : \mathbb{R}^{\mathcal{I}_{p-1,q-1}} \to \mathbb{R}^{\mathcal{I}_{p,q}}$ defined by

$$\phi(y)_{ij} = \begin{cases} 1 - y(\mathrm{row}_{i-1}) & \text{if } j = 1 \\ y_{i-1,j-1} & \text{otherwise} \end{cases} \qquad \text{for } (i,j) \in \mathcal{I}_{p,q}$$

(where $\mathrm{row}_0 = \varnothing$ and $y(\varnothing) = 0$). This is the inverse of π, showing that π is an affine isomorphism. As we have $\pi(\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)) = \mathrm{O}^{\leq}_{p-1,q-1}(\mathfrak{S}_{q-1})$, this finishes the proof. □

It will be convenient to address the elements in $\mathcal{I}_{p,q}$ via a different "system of coordinates":

$$\langle \eta, j \rangle := (j + \eta - 1, j) \qquad \text{for } j \in [q],\ 1 \leq \eta \leq p - j + 1.$$

Thus (as before) i and j denote the row and the columns, respectively, while η is the index of the diagonal (counted from above) containing the respective element; see Figure 5 (a) for an example. For $(k,j) = \langle \eta, j \rangle$ and $x \in \mathbb{R}^{\mathcal{I}_{p,q}}$, we write $x_{\langle \eta,j \rangle} := x_{(k,j)} := x_{kj}$.

For $x \in \{0,1\}^{\mathcal{I}_{p,q}}$ we denote by $I^x := \{(i,j) \in \mathcal{I}_{p,q} \,:\, x_{ij} = 1\}$ the set of all coordinates (positions in the matrix), where x has a 1-entry. Conversely, for $I \subseteq \mathcal{I}_{p,q}$, we use $\chi^I \in \{0,1\}^{\mathcal{I}_{p,q}}$ for the 0/1-point with $\chi^I_{ij} = 1$ if and only if $(i,j) \in I$.

For $(i, j) \in \mathcal{I}_{p,q}$, we define the *column*

$$\mathrm{col}(i, j) = \{(j, j), (j + 1, j), \ldots, (i - 1, j), (i, j)\} \subseteq \mathcal{I}_{p,q},$$

and for $(i, j) = \langle \eta, j\rangle$ we write $\mathrm{col}\langle \eta, j\rangle := \mathrm{col}(i, j)$. Of course, we have $\mathrm{col}\langle \eta, j\rangle = \{\langle 1, j\rangle, \langle 2, j\rangle, \ldots, \langle \eta, j\rangle\}$.

The rest of this section is organized as follows. First, in Section 4.1, we deal with basic facts about integer points in packing and partitioning orbitopes for the symmetric group. To derive a linear description of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ and $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$ that only contains integer vertices, we need additional inequalities, the *shifted column inequalities*, which are introduced in Section 4.2. We then show that the corresponding separation problem can be solved in linear time (Section 4.3). Section 4.4 proves the completeness of the linear description and Section 4.5 investigates the facets of the polytopes.

4.1. Characterization of Integer Points

We first derive a crucial property of the vertices of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$.

Lemma 6. *Let x be a vertex of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ with $\langle \eta, j\rangle \in I^x$ $(j \geq 2)$. Then we have $I^x \cap \mathrm{col}\langle \eta, j - 1\rangle \neq \varnothing$.*

Proof. With $\langle \eta, j\rangle = (i, j)$ we have $x_{ij} = 1$, which implies $x_{i,j-1} = 0$ (since x has at most one 1-entry in row i). Thus, $I^x \cap \mathrm{col}\langle \eta, j - 1\rangle = \varnothing$ would yield $x_{k,j-1} = 0$ for all $k \leq i$, contradicting the lexicographic order of the columns of x (see Part (1) of Observation 1). □

Definition 2 (Column inequality)**.** *For $(i, j) \in \mathcal{I}_{p,q}$ and the set $B = \{(i, j),$ $(i, j + 1), \ldots, (i, \min\{i, q\})\}$, we call*

$$x(B) - x(\mathrm{col}(i - 1, j - 1)) \leq 0$$

a column inequality*; see Figure 5 (b) for an example with $(i, j) = (9, 5)$.*

The column inequalities are strengthenings of the symmetry breaking inequalities

$$x_{ij} - x(\mathrm{col}(i - 1, j - 1)) \leq 0, \qquad (6)$$

introduced by Méndez-Díaz and Zabala [14] in the context of vertex-coloring (see (2) in the introduction).

Proposition 7. *A point $x \in \{0, 1\}^{\mathcal{I}_{p,q}}$ is contained in $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ ($\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$) if and only if x satisfies the row-sum constraints $x(\mathrm{row}(i)) \leq 1$ ($x(\mathrm{row}(i)) = 1$) for all $i \in [p]$ and all column inequalities.*

Proof. By Lemma 6, Inequalities (6) are valid for $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ (and thus, for its face $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$ as well). Because of the row-sum constraints, all column inequalities are valid as well. Therefore, it suffices to show that a point $x \in \{0, 1\}^{\mathcal{I}_{p,q}}$ that satisfies the row-sum constraints $x(\mathrm{row}(i)) \leq 1$ and all column inequalities is contained in $\mathcal{M}^{\max}_{p,q}(\mathfrak{S}_q)$.

Suppose, this was not the case. Then, by Part (1) of Observation 1, there must be some $j \in [q]$ such that the $(j-1)$-st column of x is lexicographically smaller than the jth column. Let i be minimal with $x_{ij} = 1$ (note that column j cannot be all-zero). Thus, $x_{k,j-1} = 0$ for all $k < i$. This implies $x(\mathrm{col}(i - 1, j - 1)) = 0 < 1 = x_{ij}$, showing that the column inequality

$x(B) - x(\mathrm{col}(i-1, j-1)) \leq 0$ is violated by the point x for the bar $B = \{(i,j), (i,j+1), \dots, (i, \min\{i,q\})\}$. □

4.2. Shifted Column Inequalities

Proposition 7 provides a characterization of the vertices of the packing- and partitioning orbitopes for symmetric groups among the integer points. Different from the situation for cyclic groups (see Theorems 3 and 4), however, the inequalities in this characterization do not yield complete descriptions of these orbitopes. In fact, we need to generalize the concept of a column inequality in order to arrive at complete descriptions. This will yield exponentially many additional facets (see Proposition 14).

Definition 3 (Shifted columns). *A set* $S = \{\langle 1, c_1\rangle, \langle 2, c_2\rangle, \dots, \langle \eta, c_\eta\rangle\} \subset \mathcal{I}_{p,q}$ *with* $\eta \geq 1$ *and* $c_1 \leq c_2 \leq \dots \leq c_\eta$ *is called a* shifted column. *It is a* shifting *of each of the columns*

$$\mathrm{col}\langle \eta, c_\eta\rangle, \mathrm{col}\langle \eta, c_\eta + 1\rangle, \dots, \mathrm{col}\langle \eta, q\rangle.$$

Remark.
- As a special case we have column $\mathrm{col}(i,j)$, which is the shifted column $\{\langle 1, j\rangle, \langle 2, j\rangle, \dots, \langle \eta, j\rangle\}$ for $\langle \eta, j\rangle = (i,j)$.
- By definition, if $S = \{\langle 1, c_1\rangle, \langle 2, c_2\rangle, \dots, \langle \eta, c_\eta\rangle\} \subset \mathcal{I}_{p,q}$ is a shifted column, then so is $\{\langle 1, c_1\rangle, \langle 2, c_2\rangle, \dots, \langle \eta', c_{\eta'}\rangle\}$ for every $1 \leq \eta' \leq \eta$.

Lemma 8. *Let x be a vertex of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ with $\langle \eta, j\rangle \in I^x$ ($j \geq 2$). Then we have $I^x \cap S \neq \varnothing$ for all shiftings S of $\mathrm{col}\langle \eta, j-1\rangle$.*

Proof. We proceed by induction on j. The case $j = 2$ follows from Lemma 6, because the only shifting of $\mathrm{col}\langle \eta, 1\rangle$ is $\mathrm{col}\langle \eta, 1\rangle$ itself. Therefore, let $j \geq 3$, and let $S = \{\langle 1, c_1\rangle, \langle 2, c_2\rangle, \dots, \langle \eta, c_\eta\rangle\}$ be a shifting of $\mathrm{col}\langle \eta, j-1\rangle$ (hence, $c_1 \leq c_2 \leq \dots \leq c_\eta \leq j-1$). Since by assumption $\langle \eta, j\rangle \in I^x$, Lemma 6 yields that there is some $\eta' \leq \eta$ with $\langle \eta', j-1\rangle \in I^x$. If $\langle \eta', j-1\rangle \in S$, then we are done. Otherwise, $c_{\eta'} < j-1$ holds. Hence, $\{\langle 1, c_1\rangle, \langle 2, c_2\rangle, \dots, \langle \eta', c_{\eta'}\rangle\}$ is a shifting of ($\mathrm{col}\langle \eta', c_{\eta'}\rangle$ and hence of) $\mathrm{col}\langle \eta', j-2\rangle$, which, by the inductive hypothesis, must intersect I^x. □

Definition 4 (Shifted column inequalities). *For $(i,j) = \langle \eta, j\rangle \in \mathcal{I}_{p,q}$, $B = \{(i,j), (i,j+1), \dots, (i, \min\{i,q\})\}$, and a shifting S of $\mathrm{col}\langle \eta, j-1\rangle$, we call*

$$x(B) - x(S) \leq 0$$

a shifted column inequality (SCI). *The set B is the* bar *of the SCI, and (i,j) is the* leader *of (the bar of) the SCI. The set S is the* shifted column (SC) *of the SCI. See Figure 5 for examples.*

In particular, all column inequalities are shifted column inequalities. The class of shifted column inequalities, however, is substantially richer: It contains exponentially many inequalities (in q).

Proposition 9. *Shifted column inequalities are valid both for the packing orbitopes $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ and for the partitioning orbitopes $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$.*

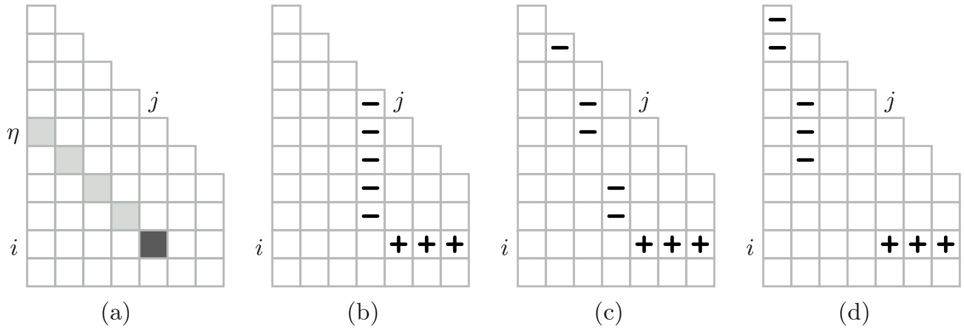


Figure 5: (a) Example for coordinates $(9,5) = \langle 5,5\rangle$. (b)–(d) Shifted column inequalities with leader $\langle 5,5\rangle$, see Definition 4. All SCI inequalities are $\leq$-inequalities with right-hand sides zero and "$-$" stands for a (-1)-coefficient, "+" for a $(+1)$ coefficient. The shifted column of (c) is $\{\langle 1,2\rangle, \langle 2,3\rangle, \langle 3,3\rangle, \langle 4,4\rangle, \langle 5,4\rangle\}$.

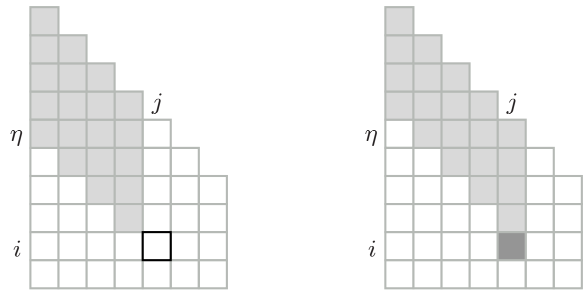


Figure 6: The two cases arising in the dynamic programming algorithm of Section 4.3.

Proof. As $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$ is a face of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$, it is enough to prove the proposition for packing orbitopes $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$. Therefore, let $(i,j) = \langle \eta, j\rangle \in \mathcal{I}_{p,q}$, with $j \geq 2$, and let $S = \{\langle 1,c_1\rangle, \langle 2,c_2\rangle, \dots, \langle \eta, c_\eta\rangle\}$ be a shifting of $\mathrm{col}\langle \eta, j-1\rangle$. Denote by B the bar of the corresponding SCI.

Let $x \in \{0,1\}^{\mathcal{I}_{p,q}}$ be a vertex of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$. If $B \cap I^x = \varnothing$, then clearly $x(B) - x(S) = 0 - x(S) \leq 0$ holds. Otherwise, there is a unique element $(i,j') = \langle \eta', j'\rangle \in B \cap I^x$. As $j' \geq j$, we have $\eta' \leq \eta$. Therefore $S' = \{\langle 1,c_1\rangle, \langle 2,c_2\rangle, \dots, \langle \eta', c_{\eta'}\rangle\} \subseteq S$ is a shifting of $\mathrm{col}\langle \eta', j'-1\rangle$. Thus, by Lemma 8, we have $S' \cap I^x \neq \varnothing$. This shows $x(S) \geq x(S') \geq 1$, implying $x(B) - x(S) \leq 1 - 1 = 0$. □

4.3. A Linear Time Separation Algorithm for SCIs

In order to devise an efficient separation algorithm for SCIs, we need a method to compute minimal shifted columns with respect to a given weight vector $w \in \mathbb{Q}^{\mathcal{I}_{p,q}}$. The crucial observation is the following. Let $S = \{\langle 1,c_1\rangle, \langle 2,c_2\rangle, \dots, \langle \eta, c_\eta\rangle\}$ with $1 \leq c_1 \leq c_2 \leq \cdots \leq c_\eta \leq j$ be a shifting of $\mathrm{col}\langle \eta, j\rangle$ for $\langle \eta, j\rangle \in \mathcal{I}_{p,q}$ with $\eta > 1$. If $c_\eta < j$, then S is a shifting of $\mathrm{col}\langle \eta, j-1\rangle$ (*Case 1*). If $c_\eta = j$, then

$$S - \langle \eta, j\rangle = \{\langle 1,c_1\rangle, \langle 2,c_2\rangle, \dots, \langle \eta-1, c_{\eta-1}\rangle\}$$

is a shifting of $\mathrm{col}\langle \eta-1, j\rangle$ (*Case 2*); see Figure 6.

For all $\langle \eta, j\rangle \in \mathcal{I}_{p,q}$, let $\omega\langle \eta, j\rangle$ be the weight of a w-minimal shifting of $\mathrm{col}\langle \eta, j\rangle$. The table $(\omega\langle \eta, j\rangle)$ can be computed by dynamic programming as follows; we also compute a table of values $\tau\langle \eta, j\rangle \in \{1, 2\}$, for each $\langle \eta, j\rangle$, which are needed later to reconstruct the corresponding shifted columns:

(1) For $j = 1, 2, \ldots, q$, initialize $\omega\langle 1, j\rangle := \min\{w_{\langle 1,\ell\rangle} : \ell \in [j]\}$.
(2) For $\eta = 2, 3, \ldots, p$, initialize $\omega\langle \eta, 1\rangle := \omega\langle \eta - 1, 1\rangle + w_{\langle \eta,1\rangle}$.
(3) For $\eta = 2, 3, \ldots, p$, $j = 2, 3, \ldots, q$ (with $\langle \eta, j\rangle \in \mathcal{I}_{p,q}$): Compute

$$\omega_1 := \omega\langle \eta, j - 1\rangle \quad \text{and} \quad \omega_2 := \omega\langle \eta - 1, j\rangle + w_{\langle \eta,j\rangle}$$

corresponding to Cases 1 and 2, respectively. Then set

$$\omega\langle \eta, j\rangle = \min\{\omega_1,\ \omega_2\} \quad \text{and} \quad \tau\langle \eta, j\rangle = \begin{cases} 1 & \text{if } \omega_1 \le \omega_2 \\ 2 & \text{otherwise.} \end{cases}$$

Thus, the tables $(\omega\langle \eta, j\rangle)$ and $(\tau\langle \eta, j\rangle)$ can be computed in time $O(pq)$. Furthermore, for a given $\langle \eta, j\rangle \in \mathcal{I}_{p,q}$, we can compute a w-minimal shifting $S\langle \eta, j\rangle$ of $\mathrm{col}\langle \eta, j\rangle$ in time $O(\eta)$ from the table $(\tau\langle \eta, j\rangle)$: We have $S\langle 1, j\rangle = \{\langle 1, j\rangle\}$ for all $j \in [q]$, $S\langle \eta, 1\rangle = \mathrm{col}\langle \eta, 1\rangle$ for all $\eta \in [p]$, and

$$S\langle \eta, j\rangle = \begin{cases} S\langle \eta, j - 1\rangle & \text{if } \tau\langle \eta, j\rangle = 1 \\ S\langle \eta - 1, j\rangle \cup \{\langle \eta, j\rangle\} & \text{if } \tau\langle \eta, j\rangle = 2 \end{cases}$$

for all other $\langle \eta, j\rangle$. This proves the following result.

Theorem 10. *Let $w \in \mathbb{Q}^{\mathcal{I}_{p,q}}$ be a given weight vector. There is an $O(pq)$ time algorithm that simultaneously computes the weights of w-minimal shiftings of $\mathrm{col}\langle \eta, j\rangle$ for all $\langle \eta, j\rangle \in \mathcal{I}_{p,q}$ and a data structure that afterwards, for a given $\langle \eta, j\rangle$, allows to determine a corresponding shifted column in $O(\eta)$ steps.*

In particular, we obtain the following:

Corollary 11. *The separation problem for shifted column inequalities can be solved in linear time $O(pq)$.*

Proof. Let a point $x^\star \in \mathbb{Q}^{\mathcal{I}_{p,q}}$ be given. We can compute the $x^\star$-values $\beta(i, j) := x^\star(B(i, j))$ of all bars $B(i, j) = \{(i, j), (i, j + 1), \ldots, (i, \min\{i, q\})\}$ in linear time in the following way: First, we initialize $\beta(i, \ell) = x^\star_{i\ell}$ for all $i \in [p]$ and $\ell = \min\{i, q\}$. Then, for each $i \in [p]$, we calculate the value $\beta(i, j) = x^\star_{ij} + \beta(i, j + 1)$ for $j = \min\{i, q\} - 1, \min\{i, q\} - 2, \ldots, 1$.

Using Theorem 10 (and the notations introduced in the paragraphs preceeding it), we compute the table $(\omega\langle \eta, j\rangle)$ and the mentioned data structure in time $O(pq)$. Then in time $O(pq)$ we check whether there exists an $(i, j) = \langle \eta, j\rangle \in \mathcal{I}_{p,q}$ with $j \ge 2$ and $\omega\langle \eta, j - 1\rangle < \beta(i, j)$. If there exists such an $\langle \eta, j\rangle$, we compute the corresponding shifted column $S\langle \eta, j - 1\rangle$ (in additional time $O(\eta) \subseteq O(p)$), yielding an SCI that is violated by $x^\star$. Otherwise $x^\star$ satisfies all SCIs. □

Of course, the procedure described in the proof of the corollary can be modified to find a maximally violated SCI if $x^\star$ does not satisfy all SCIs.

4.4. Complete Inequality Descriptions

In this section we prove that nonnegativity constraints, row-sum equations, and SCIs suffice to describe partitioning and packing orbitopes for symmetric groups. The proof will be somewhat more involved than in the case of cyclic groups. In particular, the coefficient matrices are not totally unimodular anymore. In order to see this, consider the three column inequalities

$$x_{3,3} - x_{2,2} \leq 0, \quad x_{4,3} + x_{4,4} - x_{2,2} - x_{3,2} \leq 0, \quad \text{and}$$
$$x_{5,4} + x_{5,5} - x_{3,3} - x_{4,3} \leq 0.$$

The submatrix of the coefficient matrix belonging to these three rows and the columns corresponding to $(2,2)$, $(3,3)$, and $(4,3)$ is the matrix

$$\begin{pmatrix} -1 & +1 & 0 \\ -1 & 0 & +1 \\ 0 & -1 & -1 \end{pmatrix},$$

whose determinant equals -2. Note that the above three inequalities define facets both of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ and $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$ for $p \geq q \geq 5$ (see Propositions 14 and 15, respectively).

Proposition 12. *The partitioning orbitope $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$ is completely described by the nonnegativity constraints, the row-sum equations, and the shifted column inequalities:*

$$\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q) = \{\, x \in \mathbb{R}^{\mathcal{I}_{p,q}} \;:\; x \geq \mathbf{0},\ x(\mathrm{row}_i) = 1 \text{ for } i = 1, \dots, p,$$
$$x(B) - x(S) \leq 0 \text{ for all SCIs with SC } S \text{ and bar } B \,\}.$$

Proof. Let P be the polyhedron on the right-hand side of the statement above. From Propositions 7 and 9 we know already that

$$P \cap \mathbb{Z}^{\mathcal{I}_{p,q}} = \mathrm{O}^{=}_{p,q}(\mathfrak{S}_q) \cap \mathbb{Z}^{\mathcal{I}_{p,q}}$$

holds. Thus, it suffices to show that P is an integral polytope (as $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$ is by definition). In the following, we first describe the strategy of the proof.

For the rest of the proof, fix an arbitrary vertex $x^\star$ of P. A *basis* $\mathcal{B}$ of $x^\star$ is a cardinality $|\mathcal{I}_{p,q}|$ subset of the constraints describing P that are satisfied with equality by $x^\star$ with the property that the $|\mathcal{I}_{p,q}| \times |\mathcal{I}_{p,q}|$-coefficient matrix of the left-hand sides of the constraints in $\mathcal{B}$ is non-singular. Thus, the equation system obtained from the constraints in $\mathcal{B}$ has $x^\star$ as its unique solution.

We will show that there exists a basis $\mathcal{B}^\star$ of $x^\star$ that does not contain any SCI. Thus, $\mathcal{B}^\star$ contains a subset of the p row-sum equations and at least $|\mathcal{I}_{p,q}| - p$ nonnegativity constraints. This shows that $x^\star$ has at most p nonzero entries and, since $x^\star$ satisfies the row-sum equations, it has a nonzero entry in every row. Therefore, $\mathcal{B}^\star$ contains all p row-sum equations, and all p nonzero entries must in fact be 1. Hence, $x^\star$ is a $0/1$-point. So the existence of such a basis proves the proposition.

The *weight* of a shifted column $S = \{\langle 1, c_1\rangle, \langle 2, c_2, ,\rangle \dots, \langle \eta, c_\eta\rangle\}$ with $1 \leq c_1 \leq c_2 \leq \cdots \leq c_\eta < q$ (we will not need shifted columns with $c_\eta = q$

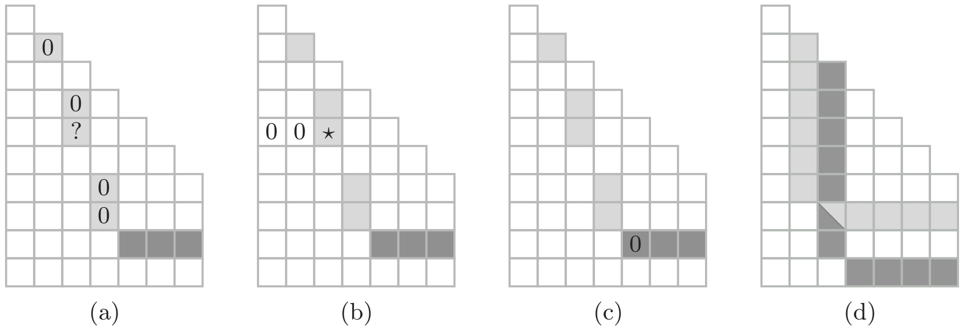


Figure 7: Illustration of trivial SCIs and of the three types of configurations not present in reduced bases of minimal weight, see Claim 3. Bars are shown in dark gray, shifted columns in light gray. Figure (a) shows trivial SCIs ("?" refers to a 0 or 1). Figures (b), (c), and (d) refer to parts (1), (2), and (3) of Claim 3, respectively ("$\star$" indicates any nonzero number).

here, as they do not appear in SCIs) is

$$\text{weight}(S) := \sum_{i=1}^{\eta} c_i\, q^i.$$

In particular, if S_1 and S_2 are two shifted columns with $|S_1| < |S_2|$, then we have $\text{weight}(S_1) < \text{weight}(S_2)$. The *weight* of an SCI is the weight of its shifted column, and the *weight* of a basis $\mathcal{B}$ is the sum of the weights of the SCIs contained in $\mathcal{B}$ (note that a shifted column can appear in several SCIs).

A basis of $x^\star$ that contains all row-sum equations and all nonnegativity constraints corresponding to 0-entries of $x^\star$ is called *reduced*. As the coefficient vectors (of the left-hand sides) of these constraints are linearly independent, some reduced basis of $x^\star$ exists. Hence, there is also a reduced basis $\mathcal{B}^\star$ of $x^\star$ of minimal weight.

To prove the proposition, it thus suffices to establish the following claim.

Claim 1. *A reduced basis of $x^\star$ of minimal weight does not contain any SCI.*

The proof of Claim 1 consists of three parts:

(1) We show that a reduced basis of $x^\star$ does not contain any "trivial SCIs" (Claim 2).
(2) We prove that a reduced basis of $x^\star$ of minimal weight satisfies three structural conditions on its (potential) SCIs (Claim 3).
(3) Finally, assuming that a reduced basis of $x^\star$ with minimal weight contains at least one SCI, we will derive a contradiction by constructing a different solution $\tilde{x} \neq x^\star$ of the corresponding equation system.

We are now ready to start with Part 1. We call an SCI with shifted column S *trivial* if $x^\star(S) = 0$ holds or if we have $x^\star(S) = 1$ and $x^\star_{k\ell} = 0$ for all $(k,\ell) \in S - (i,j)$ for some $(i,j) \in S$ (thus satisfying $x^\star_{ij} = 1$) (see Figure 7 (a)).

Claim 2. *A reduced basis $\mathcal{B}$ of $x^\star$ does not contain any trivial SCIs.*

Proof. Let S be the shifted column S and B be the bar of some SCI that is satisfied with equality by $x^\star$.

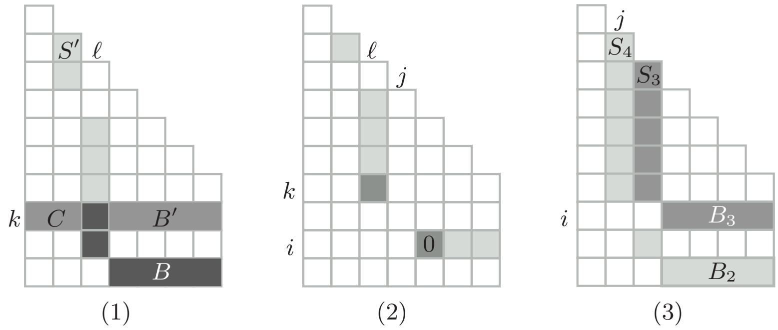


Figure 8: Illustration of the proof of Claim 3, parts (1) to (3).

If $x^\star(S) = 0$, then the coefficient vector of the SCI is a linear combination of the coefficient vectors of the inequalities $x_{ij} \geq 0$ for $(i, j) \in S \cup B$, which all are contained in $\mathcal{B}$ (due to $x^\star(B) = x^\star(S) = 0$). Since the coefficient vectors of the inequalities in $\mathcal{B}$ form a non-singular matrix, the SCI can not be in $\mathcal{B}$. (By "coefficient vector" we always mean the vector formed by the coefficients of the left-hand side of a constraint.)

If S contains exactly one entry $(k, \ell) \in S$ with $x^\star_{k\ell} = 1$, then we have $x^\star(S) = x^\star(B) = 1$. Let i be the index of the row that contains the bar B. The nonnegativity constraints $x_{rs} \geq 0$ for $(r, s) \in S - (k, \ell)$, $x_{ks} \geq 0$ for $(k, s) \in \text{row}_k - (k, \ell)$, and $x_{is} \geq 0$ for $(i, s) \in \text{row}_i \setminus B$ are contained in $\mathcal{B}$.

Since the coefficient vector of the considered SCI can linearly be combined from the coefficient vectors of these nonnegativity constraints and of the row-sum equations $x(\text{row}_k) = 1$ and $x(\text{row}_i) = 1$, this SCI cannot be contained in $\mathcal{B}$. □

Claim 3. *A minimal weight reduced basis $\mathcal{B}$ of $x^\star$ satisfies the following three conditions:*

(1) *If (k, ℓ) is contained in the shifted column of some SCI in $\mathcal{B}$, then there exists some $s < \ell$ with $x^\star_{ks} > 0$.*
(2) *If (i, j) is the leader of an SCI in $\mathcal{B}$, then $x^\star_{ij} > 0$ holds.*
(3) *If (i, j) is the leader of an SCI in $\mathcal{B}$, then there is no SCI in $\mathcal{B}$ whose shifted column contains (i, j).*

See Figure 7, (b)–(d) for an illustration of the three conditions.

Proof. Part (1): Assume there exists an SCI in $\mathcal{B}$ with shifted column S and bar B that contains the first nonzero entry of a row k, i.e., there is $(k, \ell) \in S$ with $x^\star_{k\ell} > 0$ and $x^\star_{ks} = 0$ for all $s < \ell$. Let $S' := S \cap \mathcal{I}_{k-1,q}$ be the entries of S above row k. Let $C = \{(k, 1), (k, 2), \ldots, (k, \ell - 1)\}$ and $B' = \text{row}_k \setminus (C + (k, \ell))$. See Figure 8 (1) for an illustration.

Because S' is a shifting of $\text{col}(k - 1, \ell)$, $x(B') - x(S') \leq 0$ is an SCI and hence satisfied by $x^\star$. Since we have $|S'| < |S|$ (thus, $\text{weight}(S') < \text{weight}(S)$), it suffices to show that replacing the original SCI $x(B) - x(S) \leq 0$ by $x(B') - x(S') \leq 0$ gives another basis $\mathcal{B}'$ of $x^\star$ (which also is reduced), contradicting the minimality of the weight of $\mathcal{B}$.

Due to $x^\star(\text{row}_k) = 1$, $x^\star(C) = 0$, $x^\star(B') - x^\star(S') \leq 0$, and $S' + (k,\ell) \subseteq S$ we have

$$1 = x^\star_{k\ell} + x^\star(B') \leq x^\star_{k\ell} + x^\star(S') \leq x^\star(S) = x^\star(B) \leq 1. \tag{7}$$

Therefore, equality must hold throughout this chain. In particular, this shows $x^\star(B') - x^\star(S') = 0$. Thus, its suffices to show that the coefficient matrix of the equation system obtained from $\mathcal{B}'$ is non-singular, which can be seen as follows.

Since $x^\star(S' + (k,\ell)) = 1 = x^\star(S)$ (see (7)), we know that all nonnegativity constraints $x_{rs} \geq 0$ with $(r,s) \in S \setminus (S' + (k,\ell))$ are contained in $\mathcal{B}$ and $\mathcal{B}'$. The same holds for $x_{ks} \geq 0$ with $(k,s) \in C$ and for $x_{is} \geq 0$ with $(i,s) \in \text{row}_i \setminus B$, where row i contains bar B (since $x^\star(B) = 1$ by (7)). Thus, we can linearly combine the coefficient vector of $x(B) - x(S) \leq 0$ from the coefficient vectors of the constraints $x(B') - x(S') \leq 0$, $x(\text{row}_k) = 1$, $x(\text{row}_i) = 1$, and the nonnegativity constraints mentioned above. Since all these constraints are contained in $\mathcal{B}'$, this shows that the coefficient matrix of $\mathcal{B}'$ has the same row-span as that of $\mathcal{B}$, thus proving that it is non-singular as well.

Part (2): Assume that there exists an SCI in $\mathcal{B}$ with leader (i,j), bar B, and shifted column S such that $x^\star_{ij} = 0$. If $S = \{\langle 1, c_1\rangle, \langle 2, c_2\rangle, \dots, \langle \eta, c_\eta\rangle\}$, then we have $(i,j) = \langle \eta, j\rangle$. Define $B' := B - (i,j)$, $S' := S - \langle \eta, c_\eta\rangle$, and observe that $B' \neq \varnothing$, $S' \neq \varnothing$, i.e., $|B| > 1$ and $|S| > 1$, because a reduced basis does not contain trivial SCIs by Claim 2; see Figure 8 (2). Hence, $x(B') - x(S') \leq 0$ is an SCI. We therefore have:

$$0 = x^\star(B) - x^\star(S) = x^\star(B') - x^\star(S) \leq x^\star(B') - x^\star(S') \leq 0, \tag{8}$$

where the first equation holds because $x(B) - x(S) \leq 0$ is satisfied with equality by $x^\star$ and the second equation follows from $x^\star_{ij} = 0$. Hence, we know that $x^\star(B') - x^\star(S') = 0$. Since we have $|S'| < |S|$ (and consequently $\text{weight}(S') < \text{weight}(S)$), again it remains to show that the coefficient vector of $x(B) - x(S) \leq 0$ can be linearly combined from the coefficient vector of $x(B') - x(S') \leq 0$ and some coefficient vectors of nonnegativity constraints in $\mathcal{B}$ and $\mathcal{B}'$. But this is clear, as we have $x^\star_{ij} = 0$ and $x^\star_{\langle \eta, c_\eta\rangle} = 0$, where the latter follows from (8).

Part (3): Assume that in $\mathcal{B}$ there exists an SCI

$$x(B_1) - x(S_1) \leq 0 \tag{9}$$

with leader $(i,j) = \langle \eta, j\rangle$, bar B_1, and shifted column

$$S_1 = \{\langle 1, c_1\rangle, \langle 2, c_2\rangle, \dots, \langle \eta, c_\eta\rangle\}$$

(in particular: $c_\eta < j$) and another SCI

$$x(B_2) - x(S_2) \leq 0 \tag{10}$$

with bar B_2 and shifted column

$$S_2 = \{\langle 1, d_1\rangle, \langle 2, d_2\rangle, \dots, \langle \eta, j\rangle, \langle \eta+1, d_{\eta+1}\rangle, \dots, \langle \tau, d_\tau\rangle\}.$$

Hence, we have $(i,j) = \langle \eta, j\rangle \in S_2$. Define

$$S_3 := \{\langle 1, d_1\rangle, \langle 2, d_2\rangle, \dots, \langle \eta - 1, d_{\eta-1}\rangle\}$$

(i.e, the part of S_2 lying strictly above row i) and

$$S_4 := \{\langle 1, c_1\rangle, \dots, \langle \eta, c_\eta\rangle, \langle \eta+1, d_{\eta+1}\rangle, \dots, \langle \tau, d_\tau\rangle\}$$

(i.e, S_1 together with the part of S_2 strictly below row i). Clearly, S_3 is a shifting of $\mathrm{col}\langle \eta-1, j\rangle = \mathrm{col}(i-1, j)$, and S_4 is a shifted column as well (due to $c_\eta < j \le d_{\eta+1}$). Thus, with $B_3 = B_1 - (i, j)$, we obtain the SCIs

$$x(B_3) - x(S_3) \le 0 \tag{11}$$

$$x(B_2) - x(S_4) \le 0 \tag{12}$$

(see Figure 8 (3)).

Since (9) and (10) are contained in $\mathcal{B}$, we have $x^\star(B_1) - x^\star(S_1) = 0$ and $x^\star(B_2) - x^\star(S_2) = 0$. Adding these two equations yields

$$\big(x^\star(B_3) - x^\star(S_3)\big) + \big(x^\star(B_2) - x^\star(S_4)\big) = 0, \tag{13}$$

because $x^\star_{ij}$ cancels due to $(i, j) \in B_1 \cap S_2$. Since $x^\star$ satisfies the SCIs (11) and (12), Equation (13) shows that in fact we have $x^\star(B_3) - x^\star(S_3) = 0$ and $x^\star(B_2) - x^\star(S_4) = 0$.

It is not clear, however, that we can simply replace (9) and (10) by (11) and (12) in order to obtain a new basis of $x^\star$. Nevertheless, if v_1, v_2, v_3, and v_4 are the coefficient vectors of (9), (10), (11), and (12), respectively, we have $v_1 + v_2 = v_3 + v_4$, which implies

$$v_2 = v_3 + v_4 - v_1. \tag{14}$$

Let $V \subset \mathbb{R}^{\mathcal{I}_{p,q}}$ be the subspace of $\mathbb{R}^{\mathcal{I}_{p,q}}$ that is spanned by the coefficient vectors of the constraints different from (10) in $\mathcal{B}$. Thus, the linear span of $V \cup \{v_2\}$ is the whole space $\mathbb{R}^{\mathcal{I}_{p,q}}$. Due to (14), the same holds for $V \cup \{v_3, v_4\}$ (since $v_1 \in V$). Therefore, there is $\alpha \in \{3, 4\}$ such that $V \cup \{v_\alpha\}$ spans $\mathbb{R}^{\mathcal{I}_{p,q}}$. Let (a) be the corresponding SCI from $\{(11), (12)\}$. Hence, $\mathcal{B}' := \mathcal{B} \setminus \{(10)\} \cup \{(a)\}$ is a (reduced) basis of $x^\star$ as well.

Since we have $|S_3| < |S_2|$ and $\mathrm{weight}(S_4) < \mathrm{weight}(S_2)$ (due to $c_\eta < j$), the weight of $\mathcal{B}'$ is smaller than that of $\mathcal{B}$, contradicting the minimality of the weight of $\mathcal{B}$. ☐

Before we finish the proof of the proposition by establishing Claim 1, we need one more structural result on the SCIs in a reduced basis of $x^\star$. Let $S = \{\langle 1, c_1\rangle, \langle 2, c_2\rangle, \dots, \langle \eta, c_\eta\rangle\}$ be any shifted column with $x^\star_{\langle \gamma, c_\gamma\rangle} > 0$ for some $\gamma \in [\eta]$. We call $\langle \gamma, c_\gamma\rangle$ the *first nonzero element* of S if

$$x^\star_{\langle 1, c_1\rangle} = \cdots = x^\star_{\langle \gamma-1, c_{\gamma-1}\rangle} = 0$$

holds. Similarly, $\langle \gamma, c_\gamma\rangle$ is called the *last nonzero element* of S if we have

$$x^\star_{\langle \gamma+1, c_{\gamma+1}\rangle} = \cdots = x^\star_{\langle \eta, c_\eta\rangle} = 0.$$

Claim 4. *Let $\mathcal{B}$ be a reduced basis of $x^\star$, and let S_1, S_2 be the shifted columns of some SCIs in $\mathcal{B}$ ($S_1 = S_2$ is allowed).*

(1) *If (i, j) is the first nonzero element of S_1 and $(i, j) \in S_2$, then (i, j) is also the first nonzero element of S_2.*
(2) *If (i, j) is the last nonzero element of S_1 with $x^\star(S_1) = 1$ and $(i, j) \in S_2$, then (i, j) is also the last nonzero element of S_2 and $x^\star(S_2) = 1$.*
(3) *If (i, j) is the last nonzero element of S_1 with $x^\star(S_1) = 1$, then (i, j) is not the first nonzero element of S_2.*

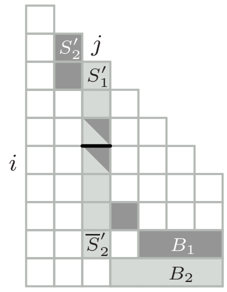


Figure 9: Illustration of sets used in the proof of Claim 4.

Proof. Let

$$S_1 = \{\langle 1, c_1\rangle, \langle 2, c_2\rangle, \ldots, \langle \eta, c_\eta\rangle\} \quad \text{and} \quad S_2 = \{\langle 1, d_1\rangle, \langle 2, d_2\rangle, \ldots, \langle \tau, d_\tau\rangle\}$$

be two shifted columns of SCIs with bars B_1 and B_2, respectively, in the reduced basis $\mathcal{B}$ of $x^\star$. Suppose that $(i, j) = \langle \gamma, j\rangle \in S_1 \cap S_2$, i.e., $c_\gamma = j = d_\gamma$ holds. Define

$$S_1' := \{\langle 1, c_1\rangle, \langle 2, c_2\rangle, \ldots, \langle \gamma - 1, c_{\gamma-1}\rangle\},$$
$$S_2' := \{\langle 1, d_1\rangle, \langle 2, d_2\rangle, \ldots, \langle \gamma - 1, d_{\gamma-1}\rangle\},$$

and $\overline{S}_2' := S_2 \setminus S_2'$, see Figure 9. Since $\langle \gamma, j\rangle \in S_1 \cap S_2$ holds, $S_1' \cup \overline{S_2'}$ is a shifted column and $x(B_2) - x(S_1' \cup \overline{S_2'}) \leq 0$ is an SCI. Thus, we obtain

$$x^\star(B_2) - x^\star(S_1') - x^\star(\overline{S}_2') \leq 0. \tag{15}$$

Furthermore, since $x(B_2) - x(S_2) \leq 0$ is contained in the basis $\mathcal{B}$ of $x^\star$, we have

$$x^\star(B_2) - x^\star(S_2') - x^\star(\overline{S}_2') = 0. \tag{16}$$

Subtracting (16) from (15) yields $x^\star(S_2') - x^\star(S_1') \leq 0$. We thus conclude

$$x^\star(S_2') \leq x^\star(S_1') \quad \text{and} \quad x^\star(S_1') \leq x^\star(S_2') \tag{17}$$

(where the second inequality follows by exchanging the roles of S_1 and S_2 in the argument).

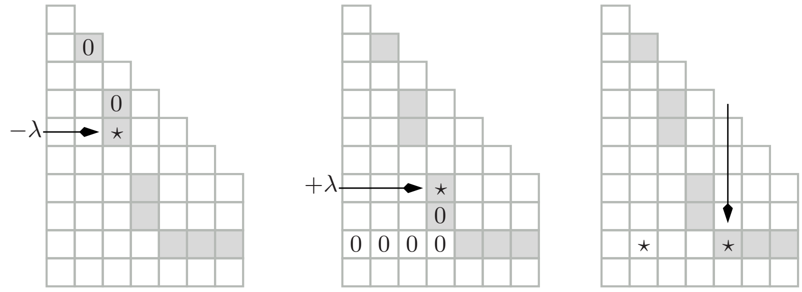
Part (1): If (i, j) is the first nonzero element of S_1, then we have $x^\star(S_1') = 0$. Thus, the first inequality of (17) implies $x^\star(S_2') = 0$, showing that (i, j) is the first nonzero element of S_2.

Part (2): If (i, j) is the last nonzero element of S_1 and $x^\star(S_1) = 1$ holds, then we have $x^\star(S_1' + (i, j)) = 1$. With the second inequality of (17) we obtain:

$$1 = x^\star(S_1' + (i, j)) \leq x^\star(S_2' + (i, j)) \leq x^\star(S_2) = x^\star(B_2) \leq 1,$$

where the last equation holds because $x(B_2) - x(S_2) \leq 0$ is contained in $\mathcal{B}$. It follows that $x^\star(S_2) = 1$ and (i, j) is the last nonzero element of S_2.

Part (3): This follows from the first two parts of the claim, since $\mathcal{B}$ does not contain any trivial SCIs by Claim 2. □



(a) (b) (c)

Figure 10: Illustration of the construction of $\tilde{x}$, Steps (1) to (3).

We will now proceed with the proof of Claim 1. Thus, assume that $\mathcal{B}^\star$ is a reduced basis of $x^\star$ of minimal weight and suppose that $\mathcal{B}^\star$ contains at least one SCI. We are going to construct a point $\tilde{x} \neq x^\star$ that satisfies the equation system obtained from $\mathcal{B}^\star$, contradicting the fact the $x^\star$ is the unique solution to this system of equations.

At the beginning, we set $\tilde{x} = x^\star$, and let $\lambda > 0$ be an arbitrary positive number. Then we perform the following four steps (see Figure 10 for illustrations of the first three).

(1) For every (i, j) that is the first nonzero element of the shifted column of at least one SCI in $\mathcal{B}^\star$, we reduce $\tilde{x}_{ij}$ by λ.
(2) For every (i, j) that is the last nonzero element of the shifted column S of at least one SCI in $\mathcal{B}^\star$ with $x^\star(S) = 1$, we increase $\tilde{x}_{ij}$ by λ.
(3) For each $i \in [p]$ and for all $j = \min\{i, q\}, \min\{i, q\} - 1, \ldots, 1$ (in this order): If (i, j) is the leader of some SCI in $\mathcal{B}^\star$, we adjust $\tilde{x}_{ij}$ such that, with $B = \{(i, j), (i, j + 1), \ldots, (i, \min\{i, q\})\}$,
$$\tilde{x}(B) = \begin{cases} 1 & \text{if } x^\star(B) = 1 \\ x^\star(B) - \lambda & \text{otherwise} \end{cases}$$
holds.
(4) For each $i \in [p]$, adjust $\tilde{x}_{ij}$ in order to achieve $\tilde{x}(\text{row}_i) = 1$, where $j = \min\{\ell : x^\star_{i\ell} > 0\}$.

The reason for treating the case $x^\star(S) = 1$ separately in Step 2 will become evident in the proof of Claim 8 below.

The following four claims will yield that $\tilde{x}$ is a solution of the equation system corresponding to $\mathcal{B}^\star$.

Claim 5. *After Step 2, for each shifted column S of some SCI in $\mathcal{B}^\star$ we have*
$$\tilde{x}(S) = \begin{cases} 1 & \textit{if } x^\star(S) = 1 \\ x^\star(S) - \lambda & \textit{otherwise.} \end{cases}$$

Proof. Let S be the shifted column of some SCI in $\mathcal{B}^\star$. It follows from Part (1) of Claim 4 that the first nonzero element (i, j) of S is the only element in S whose $\tilde{x}$-component is changed (reduced by λ) in Step 1. Thus, after Step 1 we have $\tilde{x}(S) = x^\star(S) - \lambda$.

If $x^\star(S) < 1$, then, by Part (2) of Claim 4, $\tilde{x}(S)$ is not changed in Step 2. Otherwise, $x^\star(S) = 1$, and $\tilde{x}_{k\ell}$ is increased by λ in Step 2, where (k, ℓ) is the last nonzero element of S. According to Part (2) of Claim 4, no other component of $\tilde{x}$ belonging to some element in S is changed in Step 2. Thus, in both cases the claim holds. □

Claim 6. *No component of $\tilde{x}$ belonging to the shifted column of some SCI in $\mathcal{B}^\star$ is changed in Step 3.*

Proof. Let S be the shifted column of some SCI in $\mathcal{B}^\star$. According to Part (3) of Claim 3, S does not contain the leader of any SCI in $\mathcal{B}^\star$, since $\mathcal{B}^\star$ is a reduced basis of minimal weight. □

Claim 7. *After Step 3, for each SCI in $\mathcal{B}^\star$ with shifted column S and bar B we have $\tilde{x}(S) = \tilde{x}(B)$.*

Proof. For an SCI in $\mathcal{B}^\star$ with shifted column S and bar B, we have $x^\star(S) = x^\star(B)$. Thus, from Claims 5 and 6 it follows that $\tilde{x}(S) = \tilde{x}(B)$ holds after Step 3. □

Claim 8. *Step 4 does not change any component of $\tilde{x}$ that belongs to the shifted column or the bar of some SCI in $\mathcal{B}^\star$.*

Proof. Let (i, j) be such that $x^\star_{i\ell} = 0$ for all $\ell < j$ and $x^\star_{ij} > 0$. By Part (1) of Claim 3, (i, j) is not contained in any shifted column of an SCI in $\mathcal{B}^\star$. If (i, j) is contained in the bar B of some SCI in $\mathcal{B}^\star$, then clearly $x^\star(B) = 1$ holds. Thus, after Step 3, we have $\tilde{x}(\mathrm{row}_i) = \tilde{x}(B) = 1$, which shows that $\tilde{x}_{ij}$ is not changed in Step 4. □

We can now finish the proof of the proposition. Claims 7 and 8 show that $\tilde{x}$ satisfies all SCIs contained in $\mathcal{B}^\star$ with equality. Furthermore, in all steps of the procedure only components $\tilde{x}_{ij}$ with $x^\star_{ij} > 0$ are changed (this is clear for Steps 1, 2, and 4; for Step 3 it follows from Part (2) of Claim 3). Since after Step 4, $\tilde{x}$ satisfies all row-sum equations, this proves that $\tilde{x}$ is a solution to the equation system obtained from $\mathcal{B}^\star$.

We assumed that $\mathcal{B}^\star$ contains at least one SCI. Let S be the shifted column of one of these. We know $x^\star(S) > 0$ by Claim 2. Thus, let (i, j) be the first nonzero element of S. Hence, after Step 1, we have $\tilde{x}_{ij} = x^\star_{ij} - \lambda$. By Part (3) of Claim 4, this still holds after Step 2. As $\tilde{x}_{ij}$ is also not changed in Steps 3 and 4 (see Claims 6 and 8), we deduce $\tilde{x} \neq x^\star$, contradicting the fact that $x^\star$ is the unique solution to the equation system belonging to $\mathcal{B}^\star$.

This concludes the proof of Proposition 12. □

We hope that reading this proof was somewhat enjoyable. Anyway, at least it also gives us a linear description of the packing orbitopes for symmetric groups almost for free.

Proposition 13. *The packing orbitope $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ is completely described by the nonnegativity constraints, the row-sum inequalities, and the shifted column inequalities:*

$$\begin{aligned}\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q) = \{\, x \in \mathbb{R}^{\mathcal{I}_{p,q}} \;:\; & x \geq \mathbf{0},\ x(\mathrm{row}_i) \leq 1 \text{ for } i = 1, \dots, p, \\ & x(B) - x(S) \leq 0 \text{ for all SCIs with SC } S \text{ and bar } B \,\}.\end{aligned}$$

Proof. Let $Q \subset \mathbb{R}^{\mathcal{I}_{p,q}}$ be the polyhedron on the right-hand side of the statement. We define $\mathcal{A} := \{x \in \mathbb{R}^{\mathcal{I}_{p+1,q+1}} : x(\mathrm{row}_i) = 1 \text{ for all } i \in [p+1]\}$.

The proof of Proposition 12 in fact shows that its statement remains true if we drop all SCIs with shifted column S and $S \cap \mathrm{col}_1 \neq \varnothing$ from the linear description. This follows from the fact that, due to $x^\star_{11} = 1$ and Claim 2, no such SCI can be contained in any reduced basis of $x^\star$ (using the notations from the proof of Proposition 12). Thus we obtain

$$\mathrm{O}^{=}_{p+1,q+1}(\mathfrak{S}_{q+1}) = \mathcal{A} \cap \tilde{Q}, \tag{18}$$

with

$$\begin{aligned}\tilde{Q} = \{x \in \mathbb{R}^{\mathcal{I}_{p+1,q+1}} :\ & x(B) - x(S) \leq 0 \text{ for all SCIs with bar } B\\ & \text{and shifted column } S \text{ with } S \cap \mathrm{col}_1 = \varnothing,\\ & x_{ij} \geq 0 \text{ for all } (i,j) \in \mathcal{I}_{p+1,q+1} \setminus \mathrm{col}_1,\\ & x(\mathrm{row}_i - (i,1)) \leq 1 \text{ for all } i = 2, \dots, p+1\},\end{aligned}$$

where the last inequalities are equivalent (with respect to $\mathrm{O}^{=}_{p+1,q+1}(\mathfrak{S}_{q+1})$) to the nonnegativity constraints associated with the elements of col_1 by addition of row-sum equations.

Define $\mathcal{L} := \{x \in \mathbb{R}^{\mathcal{I}_{p+1,q+1}} : x_{i1} = 0 \text{ for all } i \in [p+1]\}$, and denote by $\tilde{\pi} : \mathbb{R}^{\mathcal{I}_{p+1,q+1}} \to \mathcal{L}$ the orthogonal projection. Since none of the inequalities defining $\tilde{Q}$ has a nonzero coefficient in col_1, we have $\tilde{\pi}^{-1}(\tilde{Q} \cap \mathcal{L}) = \tilde{Q}$, hence $\tilde{Q} \cap \mathcal{L} = \tilde{\pi}(\tilde{Q})$. This yields $\tilde{\pi}(\mathcal{A} \cap \tilde{Q}) = \tilde{\pi}(\mathcal{A}) \cap \tilde{\pi}(\tilde{Q})$, which, due to $\tilde{\pi}(\mathcal{A}) = \mathcal{L}$, implies $\tilde{\pi}(\mathcal{A} \cap \tilde{Q}) = \tilde{Q} \cap \mathcal{L}$. Thus, we obtain

$$\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q) = \tilde{\pi}(\mathrm{O}^{=}_{p+1,q+1}(\mathfrak{S}_{q+1})) = \tilde{\pi}(\mathcal{A} \cap \tilde{Q}) = \tilde{Q} \cap \mathcal{L} = Q,$$

where the first equation is due to Proposition 5, the second equation follows from (18), and the final arises from identifying $\mathcal{L}$ with $\mathbb{R}^{\mathcal{I}_{p,q}}$. □

4.5. Facets

In this section, we investigate which of the constraints from the linear descriptions of $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$ and $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ given in Propositions 12 and 13, respectively, define facets. This will also yield non-redundant descriptions.

It seems to be more convenient to settle the packing case first and then to carry over the results to the partitioning case. Recall that we assume $2 \leq p \leq q$.

Proposition 14.

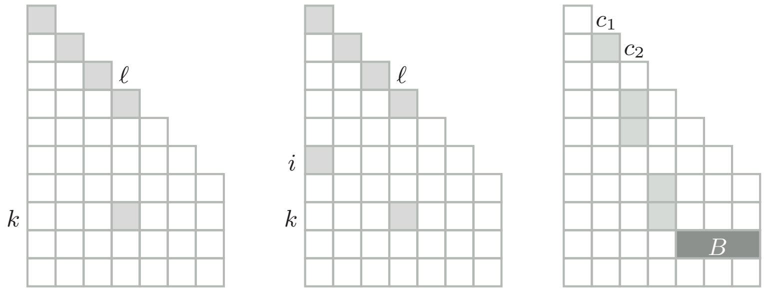
(1) *The packing orbitope $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q) \subset \mathbb{R}^{\mathcal{I}_{p,q}}$ is full dimensional:*

$$\dim(\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)) = |\mathcal{I}_{p,q}| = pq - \tfrac{q(q-1)}{2} = \big(p - \tfrac{q-1}{2}\big)q.$$

(2) *A nonnegativity constraint $x_{ij} \geq 0$, $(i,j) \in \mathcal{I}_{p,q}$, defines a facet of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$, unless $i = j < q$ holds. The faces defined by $x_{jj} \geq 0$ with $j < q$ are contained in the facet defined by $x_{qq} \geq 0$.*

(3) *Every row-sum constraint $x(\mathrm{row}_i) \leq 1$ for $i \in [p]$ defines a facet of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$.*

(4) *A shifted column inequality $x(B) - x(S) \leq 0$ with bar B and shifted column $S = \{\langle 1, c_1\rangle, \langle 2, c_2\rangle, \dots, \langle \eta, c_\eta\rangle\}$ defines a facet of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$, unless $\eta \geq 2$ and $c_1 < c_2$ (exception I) or $\eta = 1$ and $B \neq \{\langle 1, c_1 + 1\rangle\}$*



(a) matrix $V^{k\ell}$ (b) matrix $\hat{V}^{k\ell}$ (c)

Figure 11: (a)–(b): Illustration of the matrices used in the proof of parts (1) and (3) of Proposition 14. (c): Example of an SCI that does not define a facet; see the proof of Part (4) of Proposition 14.

(exception II) hold. In case of exception I, the corresponding face is contained in the facet defined by the SCI with bar B and shifted column $\{\langle 1, c_2\rangle, \langle 2, c_2\rangle, \dots, \langle \eta, c_\eta\rangle\}$. In case of exception II, the face is contained in the facet defined by the SCI $x_{\langle 1,c_1+1\rangle} - x_{\langle 1,c_1\rangle} \leq 0$.

Proof. Part (1): For all $(k, \ell) \in \mathcal{I}_{p,q}$, we define $V^{k\ell} = (v^{k\ell}_{ij}) \in \mathbb{R}^{\mathcal{I}_{p,q}}$ by

$$v^{k\ell}_{ij} = \begin{cases} 1 & \text{if } \big(i = j \leq \ell \text{ and } j < q\big) \text{ or } (i, j) = (k, \ell) \\ 0 & \text{otherwise} \end{cases} \qquad \text{for } (i, j) \in \mathcal{I}_{p,q},$$

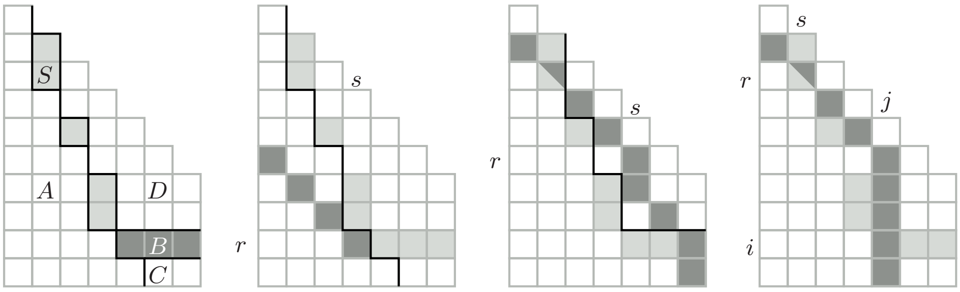
that is, $V^{k\ell}$ has 1-entries at position (k, ℓ) and on the main diagonal up to column ℓ, except that $v^{k\ell}_{qq} = 0$ unless $(k, \ell) = (q, q)$; see Figure 11 (a). The columns of each $V^{k\ell}$ are in non-increasing lexicographic order. Hence, by Part (1) of Observation 1, each $V^{k\ell}$ is a vertex of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$.

In order to show that these vectors are linearly independent, we fix an arbitrary ordering of the $V^{k\ell}$ that starts with $V^{11}, V^{22}, \dots, V^{q-1,q-1}$. For each $(k, \ell) \in \mathcal{I}_{p,q}$, all points V^{rs} preceding $V^{k\ell}$ have a 0-entry at position (k, ℓ), while $v^{k\ell}_{k\ell} = 1$. This shows that these $|\mathcal{I}_{p,q}|$ vertices of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ are linearly independent. Together with $\mathbf{0}$ this gives $|\mathcal{I}_{p,q}|+1$ affinely independent points contained in $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$, proving that $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ is full dimensional. The calculations in the statement are straightforward.

Part (2): For $(i, j) \in \mathcal{I}_{p,q} \setminus \{(j, j) \,:\, j < q\}$ all points $V^{k\ell}$ with $(k, \ell) \neq (i, j)$ are contained in the face defined by $x_{ij} \geq 0$. Since this is also true for $\mathbf{0}$, the face defined by $x_{ij} \geq 0$ contains $|\mathcal{I}_{p,q}|$ affinely independent points (see the proof of Part (1)), i.e., it is a facet of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$.

For every vertex $x^\star \in \mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ contained in the face defined by $x_{jj} \geq 0$ for some $j < q$, we have $x^\star_{\ell\ell} = 0$ for all $\ell \geq j$ (because otherwise the columns of $x^\star$ would not be in non-increasing lexicographic order). This shows that $x^\star$ is contained in the facet defined by $x_{qq} \geq 0$.

Part (3): In order to show that $x(\mathrm{row}_i) \leq 1$ defines a facet of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ for $i \in [p]$, we construct points $\hat{V}^{k\ell}$ (depending on i) from the points $V^{k\ell}$ defined in Part (1) by adding a 1 at position $(i, 1)$ if $V^{k\ell}(\mathrm{row}_i) = 0$ (see



(a) All cases (b) Case A, W^{rs} (c) Case D, U^{rs} (d) Case S

Figure 12: Illustration of the constructions in the proof of Part (4) of Proposition 14.

Figure 11 (b)). The $(|\mathcal{I}_{p,q}| - 1)$ points $\hat{V}^{k\ell}$ for all $(k, \ell) \in \mathcal{I}_{p,q} - (i, 1)$, and the unit vector E^{i1} (with a single 1 in position $(i, 1)$) satisfy $x(\text{row}_i) = 1$. Furthermore, they are affinely independent, since subtracting E^{i1} from all vectors $\hat{V}^{k\ell}$ yields vectors $\tilde{V}^{k\ell}$, which can be shown to be linearly independent similarly to Part (1); here, we need $(k, \ell) \neq (i, 1)$.

Part (4): Let $x(B) - x(S) \leq 0$ be an SCI with bar B, leader $(i, j) = \langle \eta, j \rangle$, and shifted column $S = \{\langle 1, c_1 \rangle, \langle 2, c_2 \rangle, \ldots, \langle \eta, c_\eta \rangle\}$.

If $\eta \geq 2$ and $c_1 < c_2$ hold (exception I), then the SCI is the sum of the SCI

$$x_{\langle 1, c_1+1 \rangle} - x_{\langle 1, c_1 \rangle} \leq 0$$

and the SCI with bar B and shifted column $\{\langle 1, c_1 + 1 \rangle, \langle 2, c_2 \rangle, \ldots, \langle \eta, c_\eta \rangle\}$; see Figure 11 (c). Repeating this argument $(c_2 - c_1 - 1)$ times proves the second statement of Part (4) for exception I.

If $\eta = 1$ and $B = \{\langle 1, j \rangle\}$ with $j > c_1 + 1$ hold (exception II), then the SCI is the sum of the SCIs $x_{\langle 1, c_1+1 \rangle} - x_{\langle 1, c_1 \rangle} \leq 0, \ldots, x_{\langle 1, j \rangle} - x_{\langle 1, j-1 \rangle} \leq 0$. This proves the second statement of Part (4) for exception II.

Otherwise, let $\mathcal{V}$ be the set of vertices of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ that satisfy the SCI with equality, and let $\mathcal{L} = \text{lin}(\mathcal{V} \cup \{E^{ij}\})$ be the linear span of $\mathcal{V}$ and the unit vector E^{ij}. We will show that $\mathcal{L} = \mathbb{R}^{\mathcal{I}_{p,q}}$, which proves $\dim(\text{aff}(\mathcal{V})) = |\mathcal{I}_{p,q}| - 1$ (since $\mathbf{0} \in \mathcal{V}$). Hence, the SCI defines a facet of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$.

To show that $\mathcal{L} = \mathbb{R}^{\mathcal{I}_{p,q}}$, we prove that $E^{rs} \in \mathcal{L}$ for all $(r, s) \in \mathcal{I}_{p,q}$. We partition the set $\mathcal{I}_{p,q} \setminus (B \cup S)$ into three parts (see Figure 12 (a)):

$$\begin{aligned} A :=& \{\langle \rho, s \rangle \in \mathcal{I}_{p,q} : (\rho \leq \eta \text{ and } s < c_\rho) \text{ or } \rho > \eta\}, \\ C :=& \{\langle \rho, s \rangle = (r, s) \in \mathcal{I}_{p,q} : \rho \leq \eta \text{ and } r > i\}, \text{ and} \\ D :=& \{\langle \rho, s \rangle = (r, s) \in \mathcal{I}_{p,q} : \rho < \eta,\ s > c_\rho, \text{ and } r < i\}. \end{aligned}$$

For $(r, s) = \langle \rho, s \rangle$, denote by $\text{diag}^{\leq}(r, s) = \{\langle \rho, 1 \rangle, \langle \rho, 2 \rangle, \ldots, \langle \rho, s \rangle\}$ the diagonal starting at $\langle \rho, 1 \rangle = (r - s + 1, 1)$ and ending at $\langle \rho, s \rangle = (r, s)$. Similarly, denote by $\text{diag}^{\geq}(r, s) = \{\langle \rho, s \rangle, \langle \rho, s + 1 \rangle, \ldots\} \cap \mathcal{I}_{p,q}$ the diagonal starting at (r, s) and ending in col_q or in row_p.

Claim 9. *For all $(r, s) = \langle \rho, s \rangle \in A \cup C$ we have $E^{rs} \in \mathcal{L}$.*

Proof. Denote the incidence vector of $\text{diag}^{\leq}(r, s)$ by $W^{rs} = \chi^{\text{diag}^{\leq}(r,s)}$ (see Figure 12 (b)). Both W^{rs} and $W^{rs} - E^{rs}$ are vertices of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$. We have

$\text{diag}^{\leq}(r,s) \cap (B \cup S) = \varnothing$ for $(r,s) \in A$. Furthermore

$$|\text{diag}^{\leq}(r,s) \cap B| = 1 = |\text{diag}^{\leq}(r,s) \cap S|$$

for $(r,s) \in C$. Hence, these two vertices satisfy the SCI with equality and we obtain $E^{rs} = W^{rs} - (W^{rs} - E^{rs}) \in \mathcal{L}$. □

Claim 10. *For all* $(r,s) = \langle \rho, s \rangle \in D$ *we have* $E^{rs} \in \mathcal{L}$.

Proof. Define the set

$$U(r,s) := \text{diag}^{\leq}(r,s) \cup \text{diag}^{\geq}(r+1,s) \cup \big(\{\langle \rho+1, q\rangle, \langle \rho+2, q\rangle, \dots\} \cap \mathcal{I}_{p,q}\big),$$

see Figure 12 (c). Let $U^{rs} := \chi^{U(r,s)}$. By construction, the three points U^{rs}, $U^{rs} - E^{rs}$, and $U^{rs} - E^{r+1,s}$ are vertices of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$.

If $\rho = 1$, we have $|U(r,s) \cap B| = 1$ and $|U(r,s) \cap S| = 1$, where we need $c_1 = c_2$ in case of $s = c_1 + 1$ (notice that in case of $\eta = 1$ we have $D = \varnothing$). Due to $(r,s) \notin B \cup S$, both U^{rs} and $U^{rs} - E^{rs}$ satisfy the SCI with equality. This yields $E^{rs} = U^{rs} - (U^{rs} - E^{rs}) \in \mathcal{L}$.

If $\rho > 1$, then $|U(r,s) \cap S| = 1$ does not hold in all cases (e.g., if $s = c_{\rho+1}$, we have $(r+1,s) \in S$). However, since $\rho > 1$, $U(r-1,s)$ is well-defined and

$$|U(r-1,s) \cap B| = 1 \qquad \text{and} \qquad |U(r-1,s) \cap S| = 1$$

hold. Hence the vertices $U^{r-1,s}$ and $U^{r-1,s} - E^{rs}$ satisfy the SCI with equality, giving $E^{rs} = U^{r-1,s} - (U^{r-1,s} - E^{rs}) \in \mathcal{L}$. □

Claim 11. *For all* $(r,s) = \langle \rho, s \rangle \in S$ *we have* $E^{rs} \in \mathcal{L}$.

Proof. Define the set

$$T(r,s) := \text{diag}^{\leq}(r+j-s, j) \cup \big(\{\langle \rho+1, j\rangle, \langle \rho+2, j\rangle, \dots\} \cap \mathcal{I}_{p,q}\big),$$

see Figure 12 (d). The incidence vector $T^{rs} := \chi^{T(r,s)}$ is a vertex of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$, which, due to $T(r,s) \cap S = \{(r,s)\}$ and $T(r,s) \cap B = \{(i,j)\}$ satisfies the SCI with equality. Thus, from

$$E^{rs} = T^{rs} - E^{ij} - \sum_{(k,\ell) \in T(r,s) \cap A} E^{k\ell} - \sum_{(k,\ell) \in T(r,s) \cap C} E^{k\ell} - \sum_{(k,\ell) \in T(r,s) \cap D} E^{k\ell}$$

we conclude $E^{rs} \in \mathcal{L}$, since $E^{ij} \in \mathcal{L}$ by definition of $\mathcal{L}$, and $E^{k\ell} \in \mathcal{L}$ for all $(k,\ell) \in A \cup C \cup D$ by Claims 9 and 10. □

Claim 12. *For all* $(i,s) = \langle \rho, s \rangle \in B$ *we have* $E^{rs} \in \mathcal{L}$.

Proof. The vector $W^{is} := \chi^{\text{diag}^{\leq}(i,s)}$ is a vertex of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ that satisfies the SCI with equality. Furthermore, we have

$$E^{is} = W^{is} - E^{rc_\rho} - \sum_{(k,\ell) \in \text{diag}^{\leq}(i,s) \cap A} E^{k\ell} - \sum_{(k,\ell) \in \text{diag}^{\leq}(i,s) \cap D} E^{k\ell},$$

where $(r, c_\rho) := \langle \rho, c_\rho \rangle \in S$. Thus, we conclude $E^{is} \in \mathcal{L}$, since $E^{k\ell} \in \mathcal{L}$ for all $(k,\ell) \in A \cup D \cup S$ by Claims 9, 10, and 11. □

Claims 9 to 12 show $E^{rs} \in \mathcal{L}$ for all $(r,s) \in \mathcal{I}_{p,q}$. This proves that the SCI defines a facet of $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ (unless exception I or II hold). □

Finally, we carry the results of Proposition 14 over to partitioning orbitopes.

Proposition 15.

(1) *The partitioning orbitope* $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q) \subset \mathbb{R}^{\mathcal{I}_{p,q}}$ *has dimension*

$$\dim(\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)) = |\mathcal{I}_{p-1,q-1}| = |\mathcal{I}_{p,q}| - p = \big(p - \tfrac{q}{2}\big)(q-1).$$

The constraints $x(\mathrm{row}_i) = 1$ *form a complete and non-redundant linear description of* $\mathrm{aff}(\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q))$.

(2) *A nonnegativity constraint* $x_{ij} \geq 0$, $(i,j) \in \mathcal{I}_{p,q}$, *defines a facet of* $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$, *unless* $i = j < q$ *holds. The faces defined by* $x_{jj} \geq 0$ *with* $j < q$ *are contained in the facet defined by* $x_{qq} \geq 0$.

(3) *A shifted column inequality* $x(B) - x(S) \leq 0$ *with bar* B *and shifted column* $S = \{\langle 1, c_1\rangle, \langle 2, c_2\rangle, \dots, \langle \eta, c_\eta\rangle\}$ *defines a facet of* $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$, *unless* $c_1 = 1$ *(Exception I) or* $\eta \geq 2$ *and* $c_1 < c_2$ *(Exception II) or* $\eta = 1$ *and* $B \neq \{\langle 1, c_1 + 1\rangle\}$ *(Exception III). In case of Exception I, the corresponding face is contained in the facet defined by* $x_{i1} \geq 0$, *where* i *is the index of the row containing* B. *In case of Exception II, the face is contained in the facet defined by the SCI with bar* B *and shifted column* $\{\langle 1, c_2\rangle, \langle 2, c_2\rangle, \dots, \langle \eta, c_\eta\rangle\}$. *In case of Exception III, the face is contained in the facet defined by the SCI* $x_{\langle 1,c_1+1\rangle} - x_{\langle 1,c_1\rangle} \leq 0$.

Proof. According to Proposition 5, $\mathrm{O}^{\leq}_{p-1,q-1}(\mathfrak{S}_{q-1})$ is isomorphic to $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$ via the orthogonal projection of the latter polytope to the space

$$\mathcal{L} := \{x \in \mathbb{R}^{\mathcal{I}_{p,q}} \,:\, x_{i1} = 0 \text{ for all } i \in [p]\}$$

(and via the canonical identification of $\mathcal{L}$ and $\mathbb{R}^{\mathcal{I}_{p-1,q-1}}$). This shows the statement on the dimension of $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$; the calculations and the claim on the non-redundancy of the equation system are straightforward.

Furthermore, this projection (which is one-to-one on $\mathrm{aff}(\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q))$) maps every face of $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$ that is defined by some inequality

$$\langle a, x\rangle := \sum_{(i,j)\in\mathcal{I}_{p,q}} a_{ij}\, x_{ij} \leq a_0,$$

with $a \in \mathbb{R}^{\mathcal{I}_{p,q}}$, $a_0 \in \mathbb{R}$, and $a_{i1} = 0$ for all $i \in [p]$ to a face of $\mathrm{O}^{\leq}_{p-1,q-1}(\mathfrak{S}_{q-1})$ of the same dimension defined by

$$\sum_{(i,j)\in\mathcal{I}_{p-1,q-1}} a_{i+1,j+1}\, x_{ij} \leq a_0.$$

Conversely, if $\langle \tilde{a}, x\rangle \leq \tilde{a}_0$ defines a face of $\mathrm{O}^{\leq}_{p-1,q-1}(\mathfrak{S}_{q-1})$ for $\tilde{a} \in \mathbb{R}^{\mathcal{I}_{p-1,q-1}}$ and $\tilde{a}_0 \in \mathbb{R}$, then the inequality

$$\sum_{(i,j)\in\mathcal{I}_{p,q}} \tilde{a}_{ij}\, x_{i+1,j+1} \leq \tilde{a}_0$$

defines a face of $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$ of the same dimension.

Due to parts (2) and (3) of Proposition 14, this proves Part (2) of the proposition, where we use the fact that the inequalities $x_{i1} \geq 0$ are equivalent to $x\big(\mathrm{row}_i - (i,1)\big) \leq 1$ with respect to $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$.

Furthermore, due to Part (4) of Proposition 14, the above arguments also imply the statements of Part (3) for $c_1 \geq 2$ (including Exception II and III).

Finally, we consider the case $c_1 = 1$ (Exception I). Since we have $x_{1,1} = 1$ for all $x \in \mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$, the equation $x(B) - x(S) = 0$ implies

$$1 \geq x(B) = x(S) \geq x_{1,1} = 1,$$

and hence $x_{i,1} = 0$ (using the row-sum equation for row i containing B). This concludes the proof. □

4.6. Summary of Results on the Symmetric Group

We collect the results on the packing- and partitioning orbitopes for symmetric groups.

Theorem 16. *The partitioning orbitope* $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$ *(for* $p \geq q \geq 2$*) with respect to the symmetric group* $\mathfrak{S}_q$ *equals the set of all* $x \in \mathbb{R}^{\mathcal{I}_{p,q}}$ *that satisfy the following linear constraints:*

- *the row-sum equations* $x(\mathrm{row}_i) = 1$ *for all* $i \in [p]$*,*
- *the nonnegativity constraints* $x_{ij} \geq 0$ *for all* $(i,j) \in \mathcal{I}_{p,q} \setminus \{(j,j) : j < q\}$*,*
- *the shifted column inequalities* $x(B) - x(S) \leq 0$ *for all bars*

  $$B = \{(i,j),(i,j+1),\dots,(i,\min\{i,q\})\}$$

  with $(i,j) = \langle \eta, j\rangle \in \mathcal{I}_{p,q}$*,* $j \geq 2$*, and shifted columns*

  $$S = \{\langle 1,c_1\rangle,\langle 2,c_2\rangle,\dots,\langle \eta,c_\eta\rangle\} \text{ with } 2 \leq c_1 = c_2 \leq \cdots \leq c_\eta \leq j-1,$$

  where in case of $\eta = 1$ *the last condition reduces to* $2 \leq c_1$ *and we additionally require* $j = c_1 + 1$*.*

This system of constraints is non-redundant. The corresponding separation problem can be solved in time $O(pq)$*.*

For the result on the completeness of the description, see Proposition 12, for the question of redundancy see Proposition 15, and for the separation algorithm see Corollary 11. Note that the SCI with shifted column $\{(1,1)\}$ and bar $\{(2,2)\}$ defines the same facet of $\mathrm{O}^{=}_{p,q}(\mathfrak{S}_q)$ as the nonnegativity constraint $x_{2,1} \geq 0$.

Theorem 17. *The packing orbitope* $\mathrm{O}^{\leq}_{p,q}(\mathfrak{S}_q)$ *(for* $p \geq q \geq 2$*) with respect to the symmetric group* $\mathfrak{S}_q$ *equals the set of all* $x \in \mathbb{R}^{\mathcal{I}_{p,q}}$ *that satisfy the following linear constraints:*

- *the row-sum inequalities* $x(\mathrm{row}_i) \leq 1$ *for all* $i \in [p]$*,*
- *the nonnegativity constraints* $x_{ij} \geq 0$ *for all* $(i,j) \in \mathcal{I}_{p,q} \setminus \{(j,j) : j < q\}$*,*
- *the shifted column inequalities* $x(B) - x(S) \leq 0$ *for all bars*

  $$B = \{(i,j),(i,j+1),\dots,(i,\min\{i,q\})\}$$

  with $(i,j) = \langle \eta, j\rangle \in \mathcal{I}_{p,q}$*,* $j \geq 2$*, and shifted columns*

  $$S = \{\langle 1,c_1\rangle,\langle 2,c_2\rangle,\dots,\langle \eta,c_\eta\rangle\} \text{ with } c_1 = c_2 \leq \cdots \leq c_\eta \leq j-1,$$

  where in case of $\eta = 1$ *we additionally require* $j = c_1 + 1$*.*

This system of constraints is non-redundant. The corresponding separation problem can be solved in time $O(pq)$*.*

For the result on the completeness of the description, see Proposition 13, for the question of redundancy see Proposition 14, and for the separation algorithm see Corollary 11.

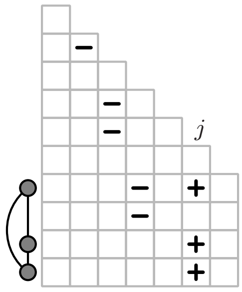


Figure 13: Combination of a clique inequality and an SCI.

5. Concluding Remarks

We close with some remarks on the technique used in the proof of Proposition 12, on the combination of SCIs and clique-inequalities for the graph-coloring problem, and on full and covering orbitopes.

The Proof Technique. Our technique to prove Proposition 12 can be summarized as follows. Assume a polytope $Q \subset \mathbb{R}^n$ is described by some (finite) system $\mathcal{Q}$ of linear equations and inequalities. Suppose that $\mathcal{Q}'$ is a subsystem of $\mathcal{Q}$ for which it is known that $\mathcal{Q}'$ defines an integral polytope $Q' \supseteq Q$. One can prove that Q is integral by showing that every vertex $x^\star$ of Q is a vertex of Q' in the following way. Here we call a basis (with respect to $\mathcal{Q}$) of $x^\star$ *reduced* if it contains as many constraints from $\mathcal{Q}'$ as possible:

(1) Starting from an arbitrary reduced basis $\mathcal{B}$ of $x^\star$, construct iteratively a reduced basis $\mathcal{B}^\star$ of $x^\star$ that satisfies some properties that are useful for the second step.
(2) Under the assumption that $\mathcal{B}^\star \not\subseteq \mathcal{Q}'$, modify $x^\star$ to some $\tilde{x} \neq x^\star$ that also satisfies the equation system corresponding to $\mathcal{B}^\star$ (contradicting the fact that $\mathcal{B}^\star$ is a basis).

(In our proof of Proposition 12, Step (1) was done by showing that a reduced basis of "minimal weight" has the desired properties.)

Such a proof is conceivable for every 0/1-polytope Q by choosing $Q' = [0,1]^n$ as the whole 0/1-cube and $\mathcal{Q}'$ as the set of the $2n$ trivial inequalities $0 \leq x_i \leq 1$, for $i = 1, \ldots, n$ (if necessary, modifying $\mathcal{Q}$ in order to contain them all).

We do not know whether this kind of integrality proof has been used in the literature. It may well be that one can interpret some of the classical integrality proofs in this setting. Anyway, it seems to us that the technique might be useful for other polytopes as well.

The Graph-Coloring Problem. As mentioned in the introduction, for concrete applications like the graph coloring problem one can (and probably has to) combine the polyhedral knowledge on orbitopes with the knowledge on problem specific polyhedra. We illustrate this by the example of clique inequalities for the graph coloring model (1) described in the introduction.

Fix a color index $j \in [C]$. If $W \subseteq V$ is a clique in the graph $G = (V, E)$, then clearly the inequality $\sum_{i \in W} x_{ij} \leq 1$ is valid. In fact, the strengthened inequalities $\sum_{i \in W} x_{ij} \leq y_j$ are known to be facet-defining for the convex hull of the solutions to (1), see [4]. Suppose that $S \subset \mathcal{I}_{|V|,C}$ is a shifted column and that we have $\eta \leq |S|$ for all $\langle \eta, j \rangle = (i, j)$ with $i \in W$. Then the inequality

$$\sum_{i \in W} x_{ij} - x(S) \leq 0$$

is valid for all solutions to the model obtained from (1) by adding inequalities (2) (which are all "column inequalities" in terms of orbitopes), see Figure 13. The details and a computational study will be the subject of a follow-up paper.

Full and Covering Orbitopes. As soon as one starts to consider 0/1-matrices that may have more than one 1-entry per row, things seem to become more complicated.

With respect to cyclic group actions, we loose the simplicity of the characterizations in Observation 1. The reason is that the matrices under investigation may have several equal nonzero columns. In particular, the lexicographically maximal column may not be unique.

With respect to the action of the symmetric group, we still have the characterization of the representatives as the matrices whose columns are in non-increasing lexicographic order (see Part 1 of Observation 1). The structures of the respective full and covering orbitopes, however, become much more complicated. In particular, we know from computer experiments that several powers of two arise as coefficients in the facet-defining inequalities. This increase in complexity is reflected by the fact that optimization of linear functionals over these orbitopes seems to be more difficult than over packing and partitioning orbitopes (see the remarks at the end of Section 2.1).

Let us close with a comment on our choice of the set of representatives as the maximal elements with respect to a lexicographic ordering (referring to the row-wise ordering of the components of the matrices). It might be that the difficulties for full and covering orbitopes mentioned in the previous paragraph can be overcome by the choice of a different system of representatives. The choice of representatives considered in this paper, however, seems to be appropriate for the packing and partitioning cases.

Whether the results presented in this paper are useful in practice will turn out in the future. In any case, we hope that the reader shares our view that orbitopes are neat mathematical objects. It seems that symmetry strikes back by its own beauty, even when mathematicians start to fight it.

Acknowledgment. We thank the referees for their work. In particular, we are indepted to one of them for several insightful and constructive remarks, including the proof of Theorem 4 that we have in this final version. We furthermore thank Yuri Faenza, Andreas Loos, and Matthias Peinhardt for helpful comments.

References

[1] R. Borndörfer, C. E. Ferreira, and A. Martin, *Decomposing matrices into blocks*, SIAM J. Optim. **9**, no. 1 (1998), pp. 236–269.

[2] R. Borndörfer, M. Grötschel, and M. E. Pfetsch, *A column-generation approach for line planning in public transport*, Transportation Sci. (2006, to appear).

[3] M. Campêlo, R. Corrêa, and Y. Frota, *Cliques, holes and the vertex coloring polytope*, Inform. Process. Lett. **89**, no. 4 (2004), pp. 159–164.
[4] P. Coll, J. Marenco, I. Méndez Díaz, and P. Zabala, *Facets of the graph coloring polytope*, Ann. Oper. Res. **116** (2002), pp. 79–90.
[5] D. Cornaz, *On forests, stable sets and polyhedras associated with clique partitions.* Preprint, 2006. Available at `www.optimization-online.org`.
[6] A. Eisenblätter, *Frequency Assignment in GSM Networks: Models, Heuristics, and Lower Bounds*, PhD thesis, TU Berlin, 2001.
[7] T. Fahle, S. Schamberger, and M. Sellmann, *Symmetry breaking*, in Principles and Practice of Constraint Programming – CP 2001: 7th International Conference, T. Walsh, ed., LNCS 2239, Springer-Verlag, Berlin/Heidelberg, 2001, pp. 93–107.
[8] R. Figueiredo, V. Barbosa, N. Maculan, and C. de Souza, *New 0-1 integer formulations of the graph coloring problem*, in Proceedings of XI CLAIO, 2002.
[9] M. R. Garey and D. S. Johnson, *Computers and Intractability. A Guide to the Theory of NP-Completeness*, W. H. Freeman and Company, New York, 1979.
[10] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Algorithms and Combinatorics 2, Springer-Verlag, Heidelberg, 2nd ed., 1993.
[11] F. Margot, *Pruning by isomorphism in branch-and-cut*, Math. Programming **94**, no. 1 (2002), pp. 71–90.
[12] F. Margot, *Small covering designs by branch-and-cut*, Math. Programming **94**, no. 2–3 (2003), pp. 207–220.
[13] A. Mehrotra and M. A. Trick, *A column generation approach for graph coloring*, INFORMS J. Comput. **8**, no. 4 (1996), pp. 344–354.
[14] I. Méndez-Díaz and P. Zabala, *A polyhedral approach for graph coloring*, Electron. Notes Discrete Math. **7** (2001).
[15] I. Méndez-Díaz and P. Zabala, *A branch-and-cut algorithm for graph coloring*, Discrete Appl. Math. **154**, no. 5 (2006), pp. 826–847.
[16] J.-F. Puget, *Symmetry breaking revisited*, Constraints **10**, no. 1 (2005), pp. 23–46.
[17] A. Ramani, F. A. Aloul, I. L. Markov, and K. A. Sakallah, *Breaking instance-independent symmetries in exact graph coloring*, in Design Automation and Test in Europe Conference, 2004, pp. 324–329.
[18] A. Schrijver, *Theory of linear and integer programming*, John Wiley & Sons, Chichester, 1986. Reprint 1998.
[19] P. Serafini and W. Ukovich, *A mathematical model for periodic scheduling problems.*, SIAM J. Discrete Math. **2**, no. 4 (1989), pp. 550–581.

Zuse Institute Berlin, Takustr. 7, 14195 Berlin, Germany
E-mail address: `[kaibel,pfetsch]@zib.de`