Smallest Compact Formulation for the Permutahedron

Michel X. Goemans*

Abstract

We consider the permutahedron, the convex hull of all permutations of $\{1, 2 \cdots, n\}$. We show how to obtain an extended formulation for the permutahedron from any sorting network. By using the optimal Ajtai-Komlós-Szemerédi (AKS) sorting network, this extended formulation has $\Theta(n \log n)$ variables and constraints. Furthermore, from basic polyhedral arguments, we show that any extended formulation has at least $\Omega(n \log n)$ constraints.

For any integer n, the permutahedron P_n is defined as the convex hull of all permutations of the set of numbers $[n] := \{1, \cdots, n\}$. In terms of a system of linear inequalities, it can be described by:

$$\begin{aligned} P_n \quad = \quad \{x \in \mathbb{R}^n : \quad & x([n]) = g(n) \\ & x(S) \le g(|S|) \quad \forall S : \emptyset \neq S \subset [n]\}, \end{aligned}$$

where

$$g(k) = \sum_{j=n+1-k}^{n} j = \binom{n+1}{2} - \binom{n+1-k}{2}.$$

Indeed, this is the base polyhedron corresponding to the submodular function f given by $f(S) = g(|S|)$, and therefore all vertices are obtained by taking a permutation σ of $[n]$ and defining $x_{\sigma(i)} = g(i) - g(i-1) = n+1-i$ for $i \in [n]$. The permutahedron P_n has $n!$ vertices and $2^n - 2$ facet-defining inequalities. By using the equality $x([n]) = g(n)$, one can also rewrite P_n as:

$$\begin{aligned} P_n \quad = \quad \{x \in \mathbb{R}^n : \quad & x([n]) = g(n) \\ & x(S) \ge \tbinom{|S|+1}{2} \quad \forall S : \emptyset \neq S \subset [n]\}. \end{aligned}$$

This is the description we are using.

Given a polyhedron $P \subseteq \mathbb{R}^n$, we say that a polyhedron $Q \subseteq R^{n+q}$ is an *extended formulation* for P if

$$P = \mathrm{proj}_n(Q) := \{x \in \mathbb{R}^n | \exists y \in \mathbb{R}^q, (x, y) \in Q\}.$$

Furthermore, we say that the extended formulation is compact if both q and the number of facets of Q is polynomial in n. The definition of extended formulations is often in terms of systems of linear inequalities and not directly stated in terms of the corresponding polyhedron; however, one can simply consider a minimal description of Q.

Finding compact extended formulations is of great importance for integer programming. Compact extended formulations are known for several combinatorial optimization polytopes, such as

*MIT Department of Mathematics. `goemans@math.mit.edu`. Supported by NSF contract CCF-0829878 and by ONR grant N00014-05-1-0148.

the spanning tree polytope or the arborescence polytope. We refer the reader to Conforti et al. [2] for a recent survey of techniques and results for extended formulations. For the matching polytope, Yannakakis [5] has shown that no compact *symmetric* extended formulation exists, and we refer the reader to [5] for a precise definition of symmetric. The author is not aware of any combinatorial optimization polytope for which a tight result on the number of facets necessary and sufficient for an extended formulation is known.

We first show an elementary lower bound on the number of facets of any extended formulation Q for any polyhedron P which, somewhat surprisingly, does not appear to be known. For any polyhedron P, let $v(P)$ denote its number of vertices, $f(P)$ its number of faces (of any dimension), and $t(P)$ its number of facets (largest dimensional faces).

Theorem 1. *Let P be any polyhedron in $\mathbb{R}^n$ with $v(P)$ vertices. Then the number of facets $t(Q)$ of any extended formulation Q of P satisfies*

$$t(Q) \geq \log_2(v(P)).$$

Proof. Assume that $Q \subseteq \mathbb{R}^{n+q}$ is an extended formulation of P. Consider any face F_P of P, and let $F_Q = \{(x, y) \in Q : x \in F_P\}$. It is easy to argue that F_Q is a face of Q. Indeed, if F_P corresponds to the valid inequality $c^T x \leq b$ for P then F_Q is the face defined by the valid inequality $c^T x + 0^T y \leq b$ for Q. Therefore, the number of faces of P is at most the number $f(Q)$ of faces of Q, i.e. $f(P) \leq f(Q)$. This implies that $v(P) \leq f(P) \leq f(Q)$. Every face of a polyhedron Q is the intersection of a subset of the facets of Q. Thus, we get that

$$f(Q) \leq 2^{t(Q)}.$$

Therefore,

$$v(P) \leq f(P) \leq f(Q) \leq 2^{t(Q)},$$

implying that $t(Q) \geq \log_2(v(P))$. □

For the permutahedron P_n, the fact that $v(P_n) = n! = 2^{\Theta(n \log n)}$ therefore implies:

Corollary 2. *Any extended formulation Q of the permutahedron P_n has at least $\Omega(n \log n)$ facets.*

We will now describe an extended formulation in $\mathbb{R}^{n+q}$ for the permutahedron P_n based on sorting networks which has $q = O(n \log n)$ and $O(n \log n)$ facets, and thus this provides an optimum (up to constant factor) extended formulation with regard to the number of facets.

Sorting networks have been introduced to formalize and describe efficient parallel algorithms for sorting n numbers. The building blocks in a sorting network are *comparators*. Each comparator takes 2 numbers a and b as inputs, and outputs $\max(a, b)$ and $\min(a, b)$ as outputs, see Figure 1. A sorting network N for sorting n numbers has n inputs, n outputs, and a number, say k, or comparators. Any input of a comparator could be either an original input or an output of a previous comparator, see Figure 2 for an example of a sorting network. Comparators are arranged in such a way that for any set of numbers on the n inputs, the n outputs are sorted in nondecreasing fashion. This is the key requirement of a sorting network. It is trivial to construct a sorting network with $\binom{n}{2}$ comparators, but there exists a variety of (fairly simple) sorting networks with $k = O(n \log^2 n)$ comparators, including Batcher's bitonic sorting network or Shell sorting network, see for example [3, Chapter 27]. In a major breakthrough more than 25 years ago, Ajtai, Komlós and Szemerédi

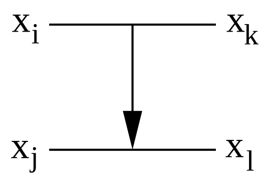

Figure 1: A single comparator with inputs x_i and x_j and outputs $x_k = \min(x_i, x_j)$ and $x_l = \max(x_i, x_j)$. The arrow indicates which output corresponds to the maximum value.

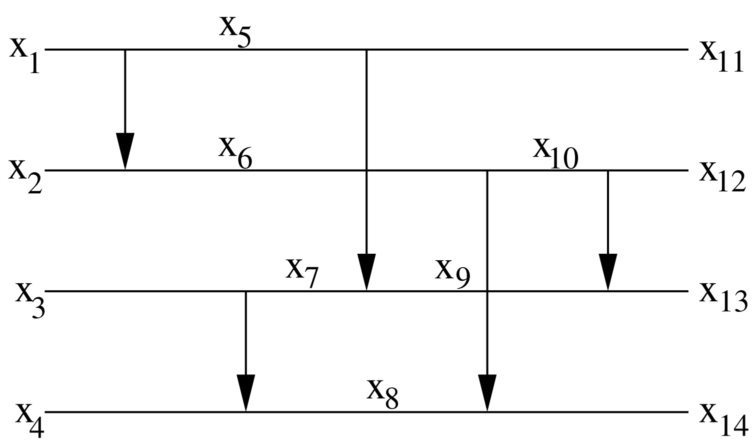

Figure 2: A sorting network for $n = 4$ inputs and $k = 5$ comparators. The $2k + n = 14$ wires are labelled with variables so that the inputs are $x_1, \cdots, x_4$ and the outputs are $x_{11}, \cdots, x_{14}$. For example, the third comparator takes x_5 and x_7 as inputs and sets x_9 to be their maximum and x_{11} their minimum.

[1] (see also [4]) have constructed a sorting network (known as an AKS sorting network) with $O(n \log n)$ comparators (although the constant in the $O(\cdot)$ notation is fairly large).

In a sorting network with k comparators, we have $2k + n$ wires, n of which are inputs, n are outputs and $2k - n$ are simultaneously an output of a comparator and an input of another. We denote by x_i for $1 \leq i \leq 2k + n$ the value on these wires, where the indexing is such that the n inputs are $x_1, x_2, \cdots, x_n$ and the n outputs are (in this order) $x_{2k+1}, x_{2k+2}, \cdots, x_{2k+n}$. See Figure 2. By construction, a sorting network is such that, for any inputs $x_1, \cdots, x_n$, the outputs satisfy $x_{2k+1} \leq x_{2k+2} \leq \cdots \leq x_{2k+n}$.

To any sorting network N with k comparators (and thus $2k+n$ wires), we construct a relaxation of it, and this corresponds to a polyhedron $Q(N) \subset \mathbb{R}^{2k+n}$ in the following way. We first impose that the ith output is equal to i, i.e.

$$x_{2k+i} = i \qquad\qquad i \in [n]. \tag{1}$$

Furthermore, for comparator m with inputs $x_{i(m)}$ and $x_{j(m)}$ and outputs $x_{k(m)} = \min(x_{i(m)}, x_{j(m)})$ and $x_{l(m)} = \max(x_{i(m)}, x_{j(m)})$, we relax these min and max constraints to linear constraints in the following way:

$$x_{i(m)} + x_{j(m)} = x_{k(m)} + x_{l(m)} \tag{2}$$

$$x_{k(m)} \leq x_{i(m)} \tag{3}$$

$$x_{k(m)} \leq x_{j(m)}. \tag{4}$$

This implies that $x_{k(m)} \leq \min(x_{i(m)}, x_{j(m)})$ and $x_{l(m)} \geq \max(x_{i(m)}, x_{j(m)})$. We claim that, for any sorting network, this relaxation provides an extended formulation of P_n.

Theorem 3. *Given any sorting network N with n inputs and k comparators, the polyhedron $Q(N) \subset \mathbb{R}^{n+2k}$ defined by the equations (1) for $i \in [n]$, the equations (2) and the inequalities (3) and (4) for $m \in [k]$ satisfies:*

$$\text{proj}_n(Q(N)) = P_n.$$

Thus, $Q(N)$ is an extended formulation for P_n with $k+n$ equalities and $2k$ inequalities in dimension $2k+n$.

By using an AKS sorting network N, we obtain an extended formulation for the permutahedron with dimension $\Theta(n \log n)$ and with $\Theta(n \log n)$ facets.

Proof. First, it is clear that $P_n \subseteq \text{proj}_n(Q(N))$. Indeed, by definition of the sorting network, if we set the x_i's for $i \in [n]$ to be any permutation of $[n]$ then we can find values x_j's for $n+1 \leq j \leq 2k+n$ such that $x \in Q(N)$. Indeed, it suffices to set $x_{k(m)} = \min(x_{i(m)}, x_{j(m)})$ and $x_{l(m)} = \max(x_{i(m)}, x_{j(m)})$ for each comparator m.

Before proving the converse, we need some notations. Given $a \in \mathbb{R}^n$, we let $\vec{a}$ be the non-decreasing sorting of a, i.e. $\vec{a}$ is such that there exists a permutation σ with $\vec{a}_i = a_{\sigma(i)}$ for $i \in [n]$ and $\vec{a}_i \leq \vec{a}_j$ for $i \leq j$. For $a, b \in \mathbb{R}^n$, we say that a majorizes b or $a \succeq b$ if

1. $\sum_{i \in [n]} a_i = \sum_{i \in [n]} b_i$,
2. $\sum_{i=1}^{j} \vec{a}_i \geq \sum_{i=1}^{j} \vec{b}_i$ for all $j \in [n]$.

Majorization is a partial order, so that if $a \succeq b$ and $b \succeq c$ then $a \succeq c$. Observe also that $a \succeq b$ depends only on $\vec{a}$ and $\vec{b}$ and not at all on the permutations transforming a into $\vec{a}$ and b into $\vec{b}$.

In the sorting network N, one can order the comparators linearly, say from 1 to k, such that an input of comparator m cannot be an output of a later comparator $m' > m$. Given this ordering, for any $0 \leq m \leq k$, let $y^{(m)} \in \mathbb{R}^n$ denote the values on the n outputs of a truncated sorting network with only the comparators with index $\leq m$. In other words, we have that $y^{(0)} = (x_1, x_2, \cdots, x_n)$ are the n inputs of the sorting network, and $y^{(m)}$ (for $1 \leq m \leq k$) can be obtained from $y^{(m-1)}$ by replacing $x_{i(m)}$ and $x_{j(m)}$ by $x_{k(m)}$ and $x_{l(m)}$. Observe that $y^{(k)} = (x_{2k+1}, \cdots, x_{2k+n})$.

We are now ready to prove that $\text{proj}_n(Q(N)) \subseteq P_n$. Consider any $x \in Q(N)$, and define $y^{(m)}$ as above for $0 \leq m \leq k$. We claim that

$$y^{(m-1)} \succeq y^{(m)},$$

for $m \in [k]$. Assuming this claim, this implies that

$$y^{(0)} = (x_1, x_2, \cdots, x_n) \succeq y^{(k)} = (1, 2, \cdots, n).$$

But this means that $(x_1, x_2, \cdots, x_n)$ satisfies all constraints defining P_n (namely, $\sum_{i\in[n]} x_i = \binom{n+1}{2}$ and $\sum_{i\in S} x_i = \binom{|S|+1}{2}$ for all $S \subset [n]$) and thus we have $\mathrm{proj}_n(Q(N)) \subseteq P_n$.

To prove the claim, observe the implications of replacing $x_{i(m)}$ and $x_{j(m)}$ by $x_{k(m)}$ and $x_{l(m)}$ satisfying (2)–(4). Clearly condition 1. of the definition of majorization will be satisfied (because of (2)) while condition 2. holds since the sum of the j smallest entries either stay the same when going from $y^{(m-1)}$ to $y^{(m)}$ or decrease. □

This work shows that no extended formulation for the permutahedron can have $o(n \log n)$ facets. However, it leaves open the question of the smallest dimension for which an extended formulation exists with $O(n \log n)$ facets (or even polynomially many). The AKS construction shows that $O(n \log n)$ dimensions are sufficient, but is it necessary? An argument similar to Theorem 1 shows that $t(Q)^{n+q} \geq n!$ for any extended formulation Q of the permutahedron, but this is too weak. In [5], Yannakakis provides a very nice characterization of the sum of the dimension and the number of facets needed for an extended formulation of any polyhedron P in terms of the *positive rank* of a slack matrix. This, however, does not provide any lower bound on just the dimension required for a compact extended formulation.

Acknowledgments

I would like to thank Alberto Caprara for mentioning this problem at an Oberwolfach meeting, and Michele Conforti for a stimulating discussion.

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi, "Sorting in $C \log N$ parallel steps", *Combinatorica*, **3**, 1–19, 1983.

[2] M. Conforti, G. Cornuéjols and G. Zambelli, "Extended formulations in Combinatorial Optimization", 2009.

[3] T.H. Cormen, C.E. Leiserson, R.L. Rivest and C. Stein, "Introduction to Algorithms", 2001.

[4] M.S. Paterson, "Improved sorting networks with $O(\log N)$ depth", *Algorithmica*, **5**, 75–92, 1990.

[5] M. Yannakakis, "Expressing Combinatorial Optimization Problems by Linear Programs", *Journal of Computer and System Sciences*, **43**, 441–466, 1991.