Smallest Compact Formulation for the Permutahedron

Noname manuscript No.
(will be inserted by the editor)

Smallest Compact Formulation for the Permutahedron

Michel X. Goemans

Dedicated to Alberto Caprara who tragically died in a mountaineering accident in April 2012. The question answered here was asked by Alberto at the Oberwolfach meeting in Combinatorial Optimization in November 2008.

the date of receipt and acceptance should be inserted later

Abstract In this note, we consider the permutahedron, the convex hull of all permutations of $\{1, 2 \cdots, n\}$. We show how to obtain an extended formulation for this polytope from any sorting network. By using the optimal Ajtai-Komlós-Szemerédi (AKS) sorting network, this extended formulation has $\Theta(n \log n)$ variables and inequalities. Furthermore, from basic polyhedral arguments, we show that this is best possible (up to a multiplicative constant) since any extended formulation has at least $\Omega(n \log n)$ inequalities. The results easily extend to the generalized permutahedron.

Keywords Permutahedron, extended formulations, compact formulations, sorting networks

Mathematics Subject Classification (2010) 90C10

1 Introduction

A classical and fundamental approach to combinatorial optimization is the polyhedral approach based on linear programming and the characterization of the convex hull of feasible solutions in terms of linear inequalities. However, the polytopes corresponding to many (even computationally easy) combinatorial optimization problems require an exponential number of facets or inequalities in the natural space, as it is the case for example for the spanning tree, arborescence or matching polytopes. However, it is sometimes possible that, for a given polytope P in $\mathbb{R}^n$, there exists a polytope Q in a higher-dimensional space, say in $\mathbb{R}^{n+k}$, which projects onto P and which has a much smaller number of facets than P does. For the purpose of optimizing over P, one can simply optimize over Q, and therefore such extended formulations with fewer facets is important. The *extension complexity* of P is defined as the smallest possible number of facets of any extended formulation for P. As examples, for both the spanning tree polytope or the arborescence polytope on a graph with n vertices, it is possible to find such an *extended formulation* of P with polynomially many (in n) facets, showing that their extension complexity is polynomially bounded. It is still unknown whether the same result holds for the matching polytope, although Yannakakis in a fundamental paper [12] has shown that no compact *symmetric* extended formulation exists for the matching polytope; we refer the reader to [12] for a precise definition of symmetric.

Supported by NSF contracts CCF-0829878 and 1115849, and by ONR grant N00014-05-1-0148.

MIT, Deparment of Mathematics, Cambridge, MA, USA. E-mail: goemans@math.mit.edu.

In this note, we consider a question posed by Alberto Caprara in November 2008 regarding the smallest possible extended formulation of the permutahedron, the convex hull of all permutations of $\{1, 2, \cdots, n\}$. It is easy to obtain a compact formulation for the permutahedron with $O(n^2)$ inequalities (or facets) but Caprara asked whether $o(n^2)$ inequalities would suffice. To answer this question, we provide a construction which given a sorting network N provides an extended formulation for the permutahedron. Using the Ajtai-Komlós-Szemerédi (AKS) sorting network [1], this gives an extended formulation with $\Theta(n \log n)$ variables and inequalities. Furthermore, we show that any extended formulation for the permutahedron has at least $\Omega(n \log n)$ inequalities. We also show that any extended formulation with $O(n \log n)$ inequalities requires $\Theta(n \log n)$ dimensions. The only other tight results we are aware of show that both the hypercube and the Birkhoff polytope admit no smaller extended formulations than themselves, see Fiorini et al. [4].

The study of compact extended formulations and their existence or non-existence has been the focus of much renewed interest in the last few years, see for example the surveys [3,6] and this special issue. In the last section of this note, we briefly discuss some of the developments that have occurred since the writing of this note and the presentation of these results at ISMP in 2009.

2 Compact Formulation for the Permutahedron

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

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

Whenever the number of facets of Q is polynomial in n, we say that the extended formulation is *compact*. The extension complexity of P is the smallest t such that there exists an extended formulation for P with t facets.

For any integer n, the permutahedron P_n is defined as the convex hull of all permutations of the set of numbers $[n] := \{1, \cdots, n\}$. It follows from Rado [9] that the permutahedron can be described in terms of linear inequalities by:

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

where

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

This is the base polytope (see eg. [11]) 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:

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

This is the description we are using.

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 (faces of dimension one smaller than the dimension of P).

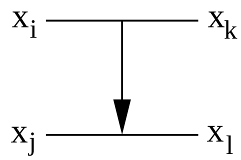

Fig. 1 A single comparator with inputs x_i and x_j and outputs $x_k = \min(x_i, x_j)$ and $x_l = \max(x_i, x_j)$. The arrow indicates which output corresponds to the maximum value.

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

The same result applies to the *generalized permutahedron* obtained by taking the convex hull of all permutations of the entries in a vector $a \in \mathbb{R}^n$, provided that these entries are distinct (to ensure that the number of vertices is indeed $n!$).

We 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 in terms of its extension complexity.

A *sorting network* N has n inputs and n outputs, and k 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. In a sorting network, see Figure 2 for an example, the comparators can be ordered in such a way that any input of a comparator is either an original input or an output of an earlier comparator. To be a sorting network, one needs to have the property that for any set of numbers on the n inputs, the n outputs are sorted in nondecreasing fashion. There exist 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 [2, Chapter 27]. In a major breakthrough, Ajtai, Komlós and Szemerédi [1] (see also [8]) 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 $x_{2k+1} \leq x_{2k+2} \leq \cdots \leq x_{2k+n}$.

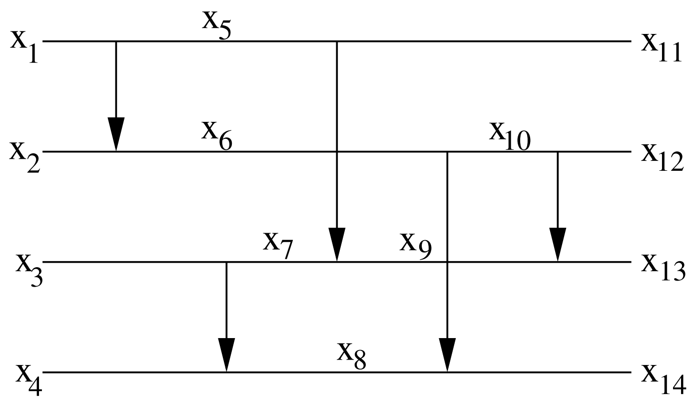

Fig. 2 A sorting network for $n = 4$ inputs and $k = 5$ comparators. The $2k + n = 14$ wires are labelled with variables so that the inputs are $x_1, \cdots, x_4$ and the outputs are $x_{11}, \cdots, x_{14}$.

To any sorting network with k comparators (and thus $2k + n$ wires), we construct a relaxation of it, and this corresponds to a polyhedron $Q \subset \mathbb{R}^{2k+n}$ 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:

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

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

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

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

Theorem 2 *Given any sorting network N with n inputs and k comparators, the polyhedron $Q(N)$ 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.

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, thereby establishing Corollary 1. The same construction applies verbatim to the generalized permutahedron, showing that its extension complexity is also $O(n \log n)$.

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 $n+1 \le j \le 2k+n$ 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 $\mathbf{a}$ be the non-decreasing sorting of a, i.e. there exists a permutation σ such that $\mathbf{a}_i = a_{\sigma(i)}$ for $i \in [n]$ and $\mathbf{a}_i \le \mathbf{a}_j$ for $i \le j$. 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. $\sum_{i\in[j]} \mathbf{a}_i \geq \sum_{i\in[j]} \mathbf{b}_i$ 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 $\mathbf{a}$ and $\mathbf{b}$ and not at all on the permutations transforming a into $\mathbf{a}$ and b into $\mathbf{b}$.

In the sorting network, 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 $\mathrm{proj}_n(Q(N)) \subseteq P_n$. 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 $\sum_{i\in S} x_i \geq \binom{|S|+1}{2}$ for all $S \subset [n]$); indeed, P_n could have been defined as the set of vectors majorizing $(1, 2, \cdots, n)$. 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 stays the same when going from $y^{(m-1)}$ to $y^{(m)}$ or decreases. $\square$

This result shows that no extended formulation for the permutahedron can have $o(n \log n)$ facets. Furthermore, we show that this construction is essentially optimal in the sense that the number of variables cannot be significantly decreased without significantly increasing the number of facets.

Theorem 3 *Given any c, let $Q \subset \mathbb{R}^{n+k}$ be an extended formulation of the permutahedron with $m \leq cn \ln n$ facets. Then $n + k = \Omega(n \ln n)$.*

Proof Let $n + k = dn \ln n$. We have that the number of vertices of Q is upper bounded by

$$\binom{m}{n+k} \leq \left(\frac{me}{n+k}\right)^{n+k} \leq \left(\frac{ce}{d}\right)^{dn \ln n}.$$

But Q has at least as many vertices as P does, and therefore

$$n! \leq \left(\frac{ce}{d}\right)^{dn \ln n}.$$

This implies that

$$\left(\frac{ce}{d}\right)^{d} \geq e - o(1),$$

or

$$d \ln(ce) - d \ln d \geq 1 - o(1).$$

This can be shown to imply that $d = \Omega(1)$ where the constant depends on c. $\square$

3 Subsequent work

In his pioneering paper on extended formulations [12], Yannakakis characterizes the extension complexity of a polytope in terms of the nonnegative rank of its slack matrix and also shows that no *symmetric* compact formulation exists for the matching polytope or the traveling salesman polytope; see [12] for definitions. In conclusion, we would like to mention a few of the many results that have been obtained recently. In view of Yannakakis' result and in response to this compact formulation of the permutahedron, Pashkovich [7] has shown that any *symmetric* extended formulation for the permutahedron has $\Omega(n^2)$ facets, showing that asymmetry really helps here. Recently, Rothvoß [10] has shown unconditionally that there exists polytopes in n dimensions having extension complexity $2^{n/2(1-o(1))}$ and this was followed by a proof by Fiorini et al. [5] that no compact formulation exists for the traveling salesman problem. The case of the matching polytope is still a major open question.

Acknowledgements In addition to Alberto Caprara, I would also like to thank Michele Conforti for a stimulating discussion at Oberwolfach, and Thomas Rothvoß for correcting calculations related to Theorem 3.

References

1. M. Ajtai, J. Komlós, and E. Szemerédi, "Sorting in $C \log N$ parallel steps", *Combinatorica*, **3**, 1–19, 1983.
2. T.H. Cormen, C.E. Leiserson, R.L. Rivest and C. Stein, "Introduction to Algorithms", 2001.
3. M. Conforti, G. Cornuéjols and G. Zambelli, "Extended formulations in combinatorial optimization", *4OR*, **8**, 1–48, 2010.
4. S. Fiorini, V. Kaibel, K. Pashkovich and D.O. Theis, "Combinatorial bounds on nonnegative rank and extended formulations", arXiv:1111.0444, 2012.
5. S. Fiorini, S. Massar, S. Pokutta, H.R. Tiwary and R. de Wolf, "Linear vs. semidefinite extended formulations: Exponential separation and strong lower bounds", *Proceedings of the 44th ACM Symposium on Theory of Computing (STOC 2012)*, 2012.
6. V. Kaibel, "Extended formulations in combinatorial optimization", *Optima*, **85**, 2–7, 2011.
7. K. Pashkovich, "Tight lower bounds on the sizes of symmetric extensions of permutahedra and similar results", arxiv:0912.3446, 2009.
8. M.S. Paterson, "Improved sorting networks with $O(\log N)$ depth", *Algorithmica*, **5**, 75–92, 1990.
9. R. Rado, "An inequality", *Journal of the London Mathematical Society*, **27**, 1–6, 1952.
10. T. Rothvoß, "Some 0/1 polytopes need exponential size extended formulations", *Mathematical Programming*, 2012.
11. A. Schrijver, "Combinatorial Optimization: Polyhedra and Efficiency", Springer, 2003.
12. M. Yannakakis, "Expressing combinatorial optimization problems by linear programs", *Journal of Computer and System Sciences*, **43**, 441–466, 1991.