Math. Program., Ser. A (2013) 142:255–268
DOI 10.1007/s10107-012-0574-3

FULL LENGTH PAPER

Some 0/1 polytopes need exponential size extended formulations

Thomas Rothvoß

Received: 2 November 2011 / Accepted: 15 June 2012 / Published online: 8 July 2012

Abstract We prove that there are $0/1$ polytopes $P \subseteq \mathbb{R}^n$ that do not admit a compact LP formulation. More precisely we show that for every n there is a set $X \subseteq \{0,1\}^n$ such that $\text{conv}(X)$ must have extension complexity at least $2^{n/2\cdot(1-o(1))}$. In other words, every polyhedron Q that can be linearly projected on $\text{conv}(X)$ must have exponentially many facets. In fact, the same result also applies if $\text{conv}(X)$ is restricted to be a matroid polytope. Conditioning on $\mathbf{NP} \not\subseteq \mathbf{P}_{/\mathbf{poly}}$, our result rules out the existence of a compact formulation for any $\mathbf{NP}$-hard optimization problem even if the formulation may contain arbitrary real numbers.

Mathematics Subject Classification 90Cxx

1 Introduction

Combinatorial optimization deals with finding the best solution out of a finite number of choices $X \subseteq \{0,1\}^n$, e.g. finding the cheapest spanning tree in a graph. If possible one aims of course to design a polynomial time algorithm. However another popular way to study combinatorial problems is to express the convex hull $P = \text{conv}(X)$ by linear inequalities $Ax \leq b$, i.e. describing them as the solutions of a linear program. A drawback of this approach is that in general an exponential number of inequalities is needed. In principle one could use the Ellipsoid method to optimize these systems, if at least the separation problem can be solved in polynomial time. But in practice this method is considered to be not applicable. A more satisfactory approach is to allow polynomially many extra variables in order to reduce the number of necessary

T. Rothvoß (✉)
M.I.T., Cambridge, MA, USA
e-mail: rothvoss@math.mit.edu

inequalities to a polynomial. This is called a *compact formulation* $P = \{x \mid \exists y : Ax + Uy \leq b\}$. Such compact formulations exist for example for the spanning tree polytope [18], the parity polytope and the permutahedron (see [20] for an extensive account).

The advantages of such a compact formulation are that (1) one can now optimize any linear function over X in polynomial time; (2) one can solve the problem with a powerful general purpose LP solver, without the need to implement a custom-tailored algorithm.

This naturally leads to the question for which problems such a compact formulation does *not* exist. Yannakakis [23] showed that the TSP polytope P_{TSP} (the convex full of the characteristic vectors of all Hamiltonian cycles in the complete graph on n nodes) does not have a subexponential size *symmetric* formulation. Surprisingly the same result holds true for the matching polytope, though here a complete description of all facets is known due to Edmonds [8] and the problem itself as well as the separation problem are solvable in polynomial time. Kaibel, Pashkovich and Theis [17] demonstrate that symmetric formulations are in some cases more restricted by proving that there is a compact non-symmetric formulation for all $\log n$-size matchings, while symmetric formulations still need size $n^{\Omega(\log n)}$.

However, it remains a fundamental open problem to show that the matching polytope or the TSP polytope do not admit any non-symmetric compact formulation. In fact, it was even an open problem to prove that there *exists* any family of $0/1$ polytopes without a compact formulation.[1] In this paper we answer this question affirmatively.

Our idea is based on a counting argument similar to Shannon's theorem [21] (see also [1]) for lower bounds on circuit sizes: Let us assume for the sake of contradiction that all n-dimensional 0/1 polytopes have a compact formulation $P = \{x \mid \exists y \geq \mathbf{0} : Ax + Uy = b\}$ of polynomial size $r(n)$. Since there are doubly-exponentially many 0/1 polytopes, there must also be at least that many formulations of size $r(n)$. This would lead to a contradiction under the additional assumption that all coefficients in the system $Ax + Uy = b$ have polynomial encoding length. Unfortunately there is no known result which guarantees that the coefficients of U will even be rational and already a single real number can contain an infinite amount of information[2] ruling out a simple counting argument.

Our contribution

In our approach, we bypass these difficulties by selecting a linearly independent subsystem of $Ax + Uy = b$ which maximizes the volume of the spanned parallelepiped; then we discretize the entries of U. We thus obtain a subsystem $\bar{A}x + \bar{U}y = \bar{b}$ with the property that $x \in X$ if and only if there is a short certificate y such that $\bar{A}x + \bar{U}y \approx \bar{b}$ for the rounded system. Secondly, all numbers in $\bar{A}, \bar{U}, \bar{b}$ have an encoding length

[1] This was posed as an open problem by Volker Kaibel on the 1st Cargèse Workshop in Combinatorial Optimization.

[2] Note that the usual argument that a polytope with rational vertices admits rational inequalities and vice versa does not apply, since both, the vertices and the inequalities of the extension polyhedron might be irrational.

which is bounded by a polynomial in n. In other words, this construction defines an injective map, taking a set X as input and providing $(\bar{A}, \bar{U}, \bar{b})$. Since there are doubly-exponentially many sets $X \subseteq \{0, 1\}^n$ and by injectivity, the number of such systems $(\bar{A}, \bar{U}, \bar{b})$ must also be doubly-exponential, which then implies the result.

It is folklore, that if **NP** problems do not all have polynomial size circuits, then no **NP**-hard optimization problem admits a compact formulation in which the numbers are rationals with polynomial encoding length. We can argue that the latter condition can be omitted.

2 Related work

A formulation of size $O(n \log n)$ for the permutahedron was provided by Goemans [14]. In fact, [14] also showed that this is tight up to constant factors. The lower bound of [14] is based on the insight that the number of facets of any extension must be at least logarithmic in the number of vertices of the target polytope (which is $n!$ for the permutahedron). The perfect matching polytope for planar graphs and graphs with bounded genus does admit a compact formulation [4,13]. A useful tool to design such formulations is the Theorem of Balas [2,3], which describes the convex hull of the union of polyhedra. For **NP**-hard problems, one can of course not expect the existence of any *exact* compact formulation. Nevertheless, Bienstock [5] gave an approximate formulation of size $n^{O(1/\varepsilon)}$ for the Knapsack polytope. This means, optimizing any linear function over the approximate polytope will give the optimum Knapsack value, up to a $1 + \varepsilon$ factor (Pritchard [19] generalized this to multidimensional Knapsack). For a more detailed literature review, we refer to the surveys of Conforti, Cornuéjols and Zambelli [6] and of Kaibel [15].

3 Preliminaries

Let $P \subseteq \mathbb{R}^n$ be a polytope with non-redundant inequality representation $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$. An *extension* is a polyhedron $Q \subseteq \mathbb{R}^m$ together with a linear projection $p : \mathbb{R}^m \to \mathbb{R}^n$ such that $p(Q) = P$. An *extended formulation* is a description of Q with linear inequalities and equations $Q = \{z \in \mathbb{R}^m \mid Cz \leq c, \; Dz = d\}$ (together with p). The *size* of the extended formulation is the number of inequalities in the description, i.e. the number of rows in C. We do not need to account for the number of equations, since they can always be eliminated. Now we can define the *extension complexity* $\mathrm{xc}(P)$ as the smallest size of any extended formulation (see [15] for more details).

Let $X = \{x_1, \ldots, x_v\} \subseteq P$ be the *vertices* (or *extreme points*) of P and let f be the number of inequalities in the description $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$.

Then the *slack-matrix* $S \in \mathbb{R}^{f \times v}$ of P is defined by $S_{ij} = b_i - A_i x_j$. Recall that the *rank* of a matrix S is the smallest r such that one can factor $S = UV$, where U is a matrix with r columns and V is a matrix with r rows. A notion which is very important for studying extended formulations is the *non-negative rank* of a matrix:

$$\mathrm{rk}_+(S) = \min\{r \mid \exists U \in \mathbb{R}^{f \times r}_{\geq 0}, V \in \mathbb{R}^{r \times v}_{\geq 0} : S = UV\}$$

Note that given a matrix $A \subseteq \mathbb{Q}_{\geq}^{m\times n}$, deciding whether $\mathrm{rk}(A) = \mathrm{rk}_+(A)$ is **NP**-hard [22]. A basic theorem concerning extended formulations, is the insight of Yannakakis, that the non-negative factorization of the slack-matrix with minimum r gives the smallest extension:

Theorem 1 (Yannakakis [23]) *Let P be a polytope with vertices $X = \{x_1, \dots, x_v\}$, non-redundant inequality description $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ and corresponding slack matrix S. Then $xc(P) = rk_+(S)$. Moreover, for any factorization $S = UV$ with $U, V \geq \mathbf{0}$ one can write $P = \{x \in \mathbb{R}^n \mid \exists y \geq \mathbf{0} : Ax + Uy = b\}$ and for every $x_j \in X$ one has $Ax_j + U \cdot V^j = b$.*

In other words: Given a polytope $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$, the smallest extension can be found by factoring the slack matrix S into non-negative factors U and V with minimum number of columns/rows. Then the smallest extended formulation comprises of $Q = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^{\mathrm{xc}(P)} \mid Ax + Uy = b,\ y \geq \mathbf{0}\}$ together with the *projection on the x-variables* $\mathrm{proj}_x(Q) = \{x \in \mathbb{R}^n \mid \exists y : (x, y) \in Q\}$. While for a polytope P, the inequality description $Ax \leq b$ is not unique, Theorem 1 implies that the non-negative rank is the same for all these descriptions.

For any matrix A, we denote its ith row by A_i and the ith column by A^i. For linearly independent vectors $w_1, \dots, w_k \in \mathbb{R}^n$, we define $\mathrm{vol}(w_1, \dots, w_k)$ as the k-dimensional volume of the parallelepiped, spanned by $w_1, \dots, w_k$. Hence for $k = n$ one has $\mathrm{vol}(w_1, \dots, w_k) = |\det(B)|$ where B is a matrix, having $w_1, \dots, w_k$ as column vectors in an arbitrary order. Note that for any vector $w \in span(w_1, \dots, w_k)$, there are unique coefficients $\lambda \in \mathbb{R}^k$ such that $w = \sum_{i=1}^k \lambda_i w_i$ and by Cramer's rule

$$|\lambda_i| = \frac{\mathrm{vol}(w_1, \dots, w_{i-1}, w, w_{i+1}, \dots, w_k)}{\mathrm{vol}(w_1, \dots, w_k)}.$$

For $q \in \mathbb{R}$, let $q\mathbb{Z}_{\geq 0} = \{0, q, 2q, \dots\}$ denote all non-negative integer multiples of q.

4 A lower bound for general 0/1 polytopes

In the following we fix a set $X \subseteq \{0, 1\}^n$. It is well known, that one can choose a matrix A and a vector b with integral entries such that $P = \{x \in \mathbb{R}^n \mid Ax \leq b\} = \mathrm{conv}(X)$, while the absolute values of any entry in A and b are bounded by $\Delta := \Delta(n) := (\sqrt{n+1})^{n+1} \leq 2^{n \log(2n)}$ (see e.g. Cor. 26 in [24]). Let S be the corresponding slack-matrix, then S is non-negative by definition and integral, since A, b and all vertices are integral. More precisely $S_{ij} = b_j - A_i x_j \in \{0, \dots, (n+1)\Delta\}$. Let $S = UV$ be any non-negative factorization, i.e. $U \in \mathbb{R}_{\geq 0}^{f\times r}$ and $V \in \mathbb{R}_{\geq 0}^{r\times v}$. As already argued above, we cannot make any assumption on the rationality/encoding length of the coefficients of U and V. But what we can do is to bound their absolute values.

Observe that if we simultaneously scale a column ℓ of U by $\lambda > 0$ and row ℓ of V by $\frac{1}{\lambda}$, then the matrix product UV stays invariant. Thus we may scale the rows and columns such that $\|U^\ell\|_\infty = \|V_\ell\|_\infty$ (if $U^\ell = \mathbf{0}$, then we can just set $V_\ell := \mathbf{0}$ as well). We call such pairs of matrices *normalized*.

Lemma 2 *For normalized matrices, one has $\|U\|_\infty \leq \Delta$ and $\|V\|_\infty \leq \Delta$.*

Proof Assume for the sake of contradiction that $U_{i\ell} > \Delta$. Thus $\|V_\ell\|_\infty > \Delta$, hence there must be an entry $V_{\ell j} > \Delta$. Then $S_{ij} = U_i \cdot V^j \geq U_{i\ell} \cdot V_{\ell j} > \Delta^2 \geq (n+1)\Delta$, which is a contradiction. □

Recalling Theorem 1, we can write $\text{conv}(X) = \{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}_{\geq 0}^{\text{xc}(\text{conv}(X))} : Ax + Uy = b\}$. Our main technical ingredient is to select a linear independent subsystem $\bar{A}x + \bar{U}y = \bar{b}$ of $Ax + Uy = b$ such that the entries of $\bar{U}$ can be rounded to rational numbers with small encoding length and still $x \in X$ iff $\bar{A}x + \bar{U}y \approx \bar{b}$ for some y.

Theorem 3 *For any non-empty $X \subseteq \{0,1\}^n$, there are matrices $\bar{A} \in \mathbb{Z}^{(n+r)\times n}$, $\bar{U} \in (\frac{1}{4r(n+r)\Delta}\mathbb{Z}_{\geq 0})^{(n+r)\times r}$ and a vector $\bar{b} \in \mathbb{Z}^{n+r}$ with $\|\bar{A}\|_\infty$, $\|\bar{b}\|_\infty$ and $\|\bar{U}\|_\infty$ upper bounded by Δ such that*

$$X = \left\{x \in \{0,1\}^n \mid \exists y \in [0,\Delta]^r : \|\bar{A}x + \bar{U}y - \bar{b}\|_\infty \leq \frac{1}{4(n+r)}\right\}$$

Here is $r := xc(conv(X))$ and $\Delta := \Delta(n) := (\sqrt{n+1})^{n+1}$.

Proof Let $X = \{x_1, \ldots, x_v\}$ and let $Ax \leq b$ with $A \in \mathbb{Z}^{f\times n}$ and $b \in \mathbb{Z}^f$ be a non-redundant description of $\text{conv}(X)$ with $\|A\|_\infty, \|b\|_\infty \leq \Delta$. Furthermore let $S \in \mathbb{Z}_{\geq 0}^{f\times|X|}$ be the corresponding slack matrix.

By Yannakakis' Theorem 1, we can write

$$P = \text{conv}(X) = \{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^r : Ax + Uy = b, \quad y \geq \mathbf{0}\}$$

where U, V are the non-negative factorization of the slack-matrix, i.e. $S = UV$. By Lemma 2 we may assume that $\|U\|_\infty, \|V\|_\infty \leq \Delta$. Let $W = \text{span}(\{(A_i, U_i) \mid i = 1, \ldots, f\})$ be the span of the constraint matrix of the system $Ax + Uy = b$ and let $k = \dim(W)$ be its dimension. Choose $I \subseteq \{1, \ldots, f\}$ of size $|I| = k$ such that $\text{vol}(\{(A_i, U_i) \mid i \in I\})$ is maximized. Recall that U_I is the matrix U, restricted to the rows in I. Let U'_I be the matrix U_I where coefficients are rounded down to the nearest multiple of $\frac{1}{4r(n+r)\Delta}$. Our choice will be $\bar{A} := A_I, \bar{U} := U'_I, \bar{b} := b_I$, hence it remains to show that

$$X \overset{!}{=} \left\{x \in \{0,1\}^n \mid \exists y \in [0,\Delta]^r : \|A_I x + U'_I y - b_I\|_\infty \leq \frac{1}{4(n+r)}\right\} =: Y$$

Claim $X \subseteq Y$.

Proof of Claim Consider a vector $x_j \in X$. Using Yannakakis' Theorem 1, we can simply choose $y := V^j \geq \mathbf{0}$ and have $Ax_j + U \cdot y = b$. Due to normalization, $\|y\|_\infty \leq \|V\|_\infty \leq \Delta$. Note that $\|U - U'\|_\infty \leq \frac{1}{4r(n+r)\Delta}$. By the triangle inequality

$$\begin{aligned}\|A_I x_j + U_I' y - b_I\|_\infty &= \| \underbrace{A_I x_j + U_I y - b_I}_{=\mathbf{0}} + (U_I' - U_I) y\|_\infty \\ &\leq r \cdot \underbrace{\|U_I' - U_I\|_\infty}_{\leq \frac{1}{4r(n+r)\Delta}} \cdot \underbrace{\|y\|_\infty}_{\leq \Delta} \leq \frac{1}{4(n+r)}\end{aligned}$$

Thus $x_j \in Y$. □

Claim $X \supseteq Y$.

Proof of Claim We show that for $x \in \{0,1\}^n$ with $x \notin X$ one has $x \notin Y$. Since $x \notin X$, there must be a row ℓ with $A_\ell x > b_\ell$. Since A, b and x are integral, one even has $A_\ell x \geq b_\ell + 1$. Unfortunately ℓ is in general not among the selected constraints I. But there are unique coefficients $\lambda \in \mathbb{R}^k$ such that we can express constraint $A_\ell x + U_\ell y = b_\ell$ as a linear combination of those in I, i.e.

$$\left(A_\ell, U_\ell\right) = \sum_{i \in I} \lambda_i \left(A_i, U_i\right).$$

Note that automatically we have $\sum_{i \in I} \lambda_i b_i = b_\ell$, since otherwise the system $Ax + Uy = b$ could not have any solution (x, y) at all and $X = \emptyset$. The next step is to bound the coefficients λ_i. Here we recall that by Cramer's rule

$$|\lambda_i| = \frac{\mathrm{vol}\left(\left\{(A_{i'}, U_{i'}) \mid i' \in I \backslash \{i\} \cup \{\ell\}\right\}\right)}{\mathrm{vol}\left(\left\{(A_{i'}, U_{i'}) \mid i' \in I\right\}\right)} \leq 1$$

since we picked I such that $\mathrm{vol}(\{(A_{i'}, U_{i'}) \mid i' \in I\})$ is maximized. Fix an arbitrary $y \in [0, \Delta]^r$, then

$$\begin{aligned}1 \leq | \underbrace{A_\ell x - b_\ell}_{\geq 1} + \underbrace{U_\ell y}_{\geq 0} | &= \Big| \sum_{i \in I} \lambda_i (A_i x - b_i + U_i y) \Big| \\ &\leq \sum_{i \in I} \underbrace{|\lambda_i|}_{\leq 1} \cdot |A_i x - b_i + U_i y| \\ &\leq (n + r) \cdot \|A_I x - b_I + U_I y\|_\infty\end{aligned} \tag{1}$$

using the triangle inequality and the fact that $|I| \leq n + r$. Again making use of the triangle inequality yields

$$\begin{aligned}\|A_I x - b_I + U_I y\|_\infty &= \|A_I x - b_I + U_I' y + (U_I - U_I') y\|_\infty \\ &\leq \|A_I x - b_I + U_I' y\|_\infty + r \cdot \underbrace{\|U_I - U_I'\|_\infty}_{\leq \frac{1}{4r(n+r)\Delta}} \cdot \underbrace{\|y\|_\infty}_{\leq \Delta} \\ &\leq \|A_I x - b_I + U_I' y\|_\infty + \frac{1}{4(n+r)}\end{aligned} \tag{2}$$

Combining (1) and (2) gives $\|A_I x - b_I + U_I' y\|_\infty \geq \frac{1}{n+r} - \frac{1}{4(n+r)} \geq \frac{1}{2(n+r)}$ and consequently $x \notin Y$.

The assertion of the Theorem follows. Note that by padding empty rows, we can ensure that $\bar{A}, \bar{U}, \bar{b}$ have exactly $n + r$ rows. □

Theorem 4 *For any $n \in \mathbb{N}$, there exists a set $X \subseteq \{0, 1\}^n$ such that $xc(conv(X)) \geq \Omega(2^{n/2}/\sqrt{n \log(2n)})$.*

Proof Let $R := R(n)$ be the maximum value of $\mathrm{xc}(\mathrm{conv}(X))$ over all $X \subseteq \{0, 1\}^n$. Note that $R \leq 2^n$. The construction in Theorem 3 implicitly defines a function Φ which maps a set X to a system $(\bar{A}, \bar{U}, \bar{b})$.[3] The important observation is that due to Theorem 3, for a given system $(\bar{A}, \bar{U}, \bar{b})$, one can reconstruct the corresponding set X. In other words, the function Φ is injective. In fact, adding zero rows and columns to those matrices does not change the claim, hence we may assume that $\bar{A}$ is an $(n + R) \times n$ matrix and $\bar{U}$ is an $(n + R) \times R$ matrix. Every entry in $\bar{U}$ has absolute value at most Δ and is a multiple of $\frac{1}{4r(n+r)\Delta}$ for some $r \in \{1, \ldots, R\}$. In other words, the domain for each entry contains at most $\sum_{r=1}^{R}(2 \cdot 4r(n + r)\Delta \cdot \Delta + 1) \leq 18\Delta^5$ many possible values (here we use the generous estimates $R \leq 2^n \leq \Delta$ and $n \leq \Delta$). By injectivity of Φ, the number of sets X (which is $2^{2^n} - 1$) cannot be larger than the number of systems $(\bar{A}, \bar{U}, \bar{b})$. Thus

$$2^{2^n} - 1 \leq (18\Delta^5)^{(n+R+1)\cdot(n+R)} \leq 2^{C(n^4+n\log(2n)\cdot R^2)}$$

for some constant $C > 0$. Hence $R \geq C' \cdot 2^{n/2}/\sqrt{n \log(2n)}$ for some $C' > 0$. □

Observe that this proof also implies that *most* 0/1 polytopes will have large extension complexity.

Corollary 5 *Let $X^{(1)}, \ldots, X^{(M)} \subseteq \{0, 1\}^n$ be distinct subsets and let $0 < \delta < 1$ be a parameter such that $\delta M \geq 2^{n^4}$. Then for at least $(1 - \delta) \cdot M$ many indices $j \in \{1, \ldots, M\}$ one has $xc(conv(X^{(j)})) \geq \Omega\left(\sqrt{\frac{\log(\delta M)}{n \log(2n)}}\right)$.*

Proof Suppose that δM many non-empty polytopes $\mathrm{conv}(X^{(j)})$ have extension complexity at most R. By the same arguments as in the proof of Theorem 4,

$$\delta M \leq (18\Delta^5)^{(n+R+1)\cdot(n+R)} \leq 2^{\frac{n^4}{2}+Cn\log(2n)\cdot R^2}$$

for some constant $C > 0$ and n large enough. Rearranging this yields the claim. □

[3] The initial system $Ax \leq b$ describing $\mathrm{conv}(X)$ might not be unique, as well as index set I. For Φ to be well defined one can make an arbitrary canonical choice, like choosing $Ax \leq b$ and I lexicographical minimal.

5 A lower bound for matroid polytopes

The main drawback of our result is that it does not rule out compact formulations for any explicitly known polytope. However, we can extend the result to matroid polytopes. Recall that a pair $([n], \mathcal{I})$ is called a *matroid* with *ground set* $[n] = \{1, \dots, n\}$ and *independent sets* $\mathcal{I} \subseteq 2^{[n]}$, if (I) $I \in \mathcal{I}$, $J \subseteq I \Rightarrow J \in \mathcal{I}$ and (II) for all $I, J \in \mathcal{I}$ with $|I| < |J|$ there is a $z \in J \backslash I$ with $I + z \in \mathcal{I}$. Note that all non-trivial facet-defining inequalities for $\text{conv}(\chi(\mathcal{I}))$ are of the form $\sum_{i \in S} x_i \leq r_{\mathcal{I}}(S)$ with $S \subseteq [n]$, where $r_{\mathcal{I}}$ denotes the *rank function* of the matroid ($\chi(\mathcal{I})$ denotes the set of characteristic vectors of $\mathcal{I}$). Secondly, any linear objective function can be optimized over $\text{conv}(\chi(\mathcal{I}))$ using the greedy algorithm, which involves calling a membership oracle a polynomial number of times. See e.g. the textbook of Schrijver [20] for more details.

Nevertheless, it is well known that the number of matroids with ground set $\{1, \dots, n\}$ is at least $2^{\binom{n}{\lfloor n/2 \rfloor}/(2n)} \geq 2^{2^n/(10n^{3/2})}$ for n large enough [7]. In other words, there are doubly-exponentially many matroids. Applying Corollary 5 (say with $\delta := 1/2$) yields:

Corollary 6 *There exists a family $M_n = (\{1, \dots, n\}, \mathcal{I}_n)$ of matroids such that $xc(conv(\chi(\mathcal{I}_n))) = \Omega(2^{n/2}/(n^{5/4}\sqrt{\log(2n)}))$.*

Finally, we want to remark that the counting argument in this paper can also be applied to show lower bounds on the extension complexity of polygons (see [10]).

6 Approximating 0/1 polytopes

In this section, we want to extend the result of Theorem 3 such that any 0/1 polytope P can be arbitrarily well approximated as a projection of a polytope Q with $O(n + \text{xc}(P))$ facets but still small encoding length. See Fig. 1 for an illustration. In the following, for any $\varepsilon > 0$, let $P + \varepsilon = \{x + z \in \mathbb{R}^n \mid x \in P, \|z\|_2 \leq \varepsilon\}$.

Theorem 7 *For any non-empty 0/1 polytope $P = conv(X)$ ($X \subseteq \{0, 1\}^n$) and any $\varepsilon > 0$, there exists a polytope $Q = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^{xc(P)} \mid Bx + Cy \leq d\}$ such that $B \in \mathbb{Q}^{(4xc(P)+2n)\times n}$, $C \in \mathbb{Q}^{(4xc(P)+2n)\times xc(P)}$ and $b \in \mathbb{Q}^{4xc(P)+2n}$ have encoding length $poly(n, xc(P), \log(\frac{1}{\varepsilon}))$ and $P \subseteq proj_x(Q) \subseteq P + \varepsilon$.*

Furthermore for any objective function $c \in \mathbb{R}^n$, $\max\{c^T x \mid x \in proj_x(Q)\} - \max\{c^T x \mid x \in P\} \leq \varepsilon \cdot \|c\|_2$.

Proof W.l.o.g. assume that $\frac{1}{\varepsilon}$ is integral. Again let $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ be a non-redundant inequality description of P such that A and b have entries from $\{-\Delta, \dots, \Delta\}$. Abbreviate $r := \text{xc}(P)$. We again apply Theorem 3 to obtain a system A_I, U'_I, b_I. But this time, we round the entries in the matrix U_I down to the nearest multiple of $\frac{\delta}{4r(n+r)\Delta}$ (instead of $\frac{1}{4r(n+r)\Delta}$), for $\delta := \min\{\frac{1}{2(n\Delta)^{2n+2}}, \frac{\varepsilon}{n\cdot(n\Delta)^n}\}$. We choose

$$Q := \left\{(x, y) \mid \|A_I x + U'_I y - b_I\|_\infty \leq \frac{\delta}{4(n+r)}, y \in [0, \Delta]^r\right\}$$

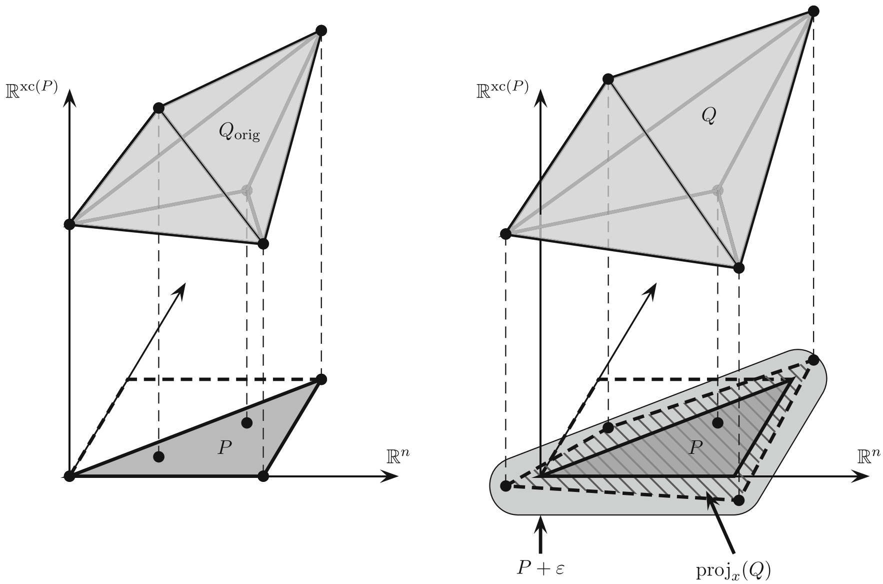


Fig. 1 Visualization of Theorem 7

Note that Q is in fact a polytope which can be written in the form $Q = \{(x, y) \mid Bx + Cy \leq d\}$ such that B, C, d are of the claimed format. Furthermore the encoding length of B, C, d is polynomial in n, $\mathrm{xc}(P)$ and $\log(1/\varepsilon)$.[4] In the remaining proof we show that $P \subseteq \mathrm{proj}_x(Q) \subseteq \{x \in \mathbb{R}^n \mid Ax \leq b + \delta\mathbf{1}\} \subseteq P + \varepsilon$.

Claim $P \subseteq \mathit{proj}_x(Q)$.

Proof of Claim As in Theorem 3, for any vertex $x_j \in P$, one has $(x_j, V^j) \in Q$ (since $\|A_I x_j + U_I' V^j - b_I\|_\infty \leq r \cdot \|U_I' - U_I\|_\infty \cdot \|V^j\|_\infty \leq \frac{\delta}{4(n+r)}$). Consequently $P \subseteq \mathrm{proj}_x(Q)$. □

Claim $\mathit{proj}_x(Q) \subseteq \{x \in \mathbb{R}^n \mid Ax \leq b + \delta\mathbf{1}\}$.

Proof of Claim Suppose for the sake of contradiction, that there is an $x^* \in \mathrm{proj}_x(Q)$ such that for some ℓ one has $A_\ell x^* > b_\ell + \delta$. Revisiting again Inequalities (1) and (2), we see that for any $y \in [0, \Delta]^r$ now

$$\begin{aligned}\delta &\overset{(1)}{\leq} (n+r) \cdot \|A_I x^* - b_I + U_I\|_\infty \\ &\overset{(2)}{\leq} (n+r) \cdot \Big(\|A_I x^* - b_I + U_I' y\|_\infty + r \cdot \underbrace{\|U_I - U_I'\|_\infty}_{\leq \delta/(4r(n+r)\Delta)} \cdot \underbrace{\|y\|_\infty}_{\leq \Delta}\Big) \\ &\leq (n+r) \cdot \|A_I x^* - b_I + U_I' y\|_\infty + \frac{\delta}{4}\end{aligned}$$

[4] This follows from the fact that all coefficients in B, C, d are products of $n, \mathrm{xc}(P), \delta, \varepsilon, \Delta$ (or their reciprocals) and $\log(\Delta) \leq O(n \cdot \log n)$, $\log(1/\delta) \leq \log(1/\varepsilon) + O(n^2 \log n)$.

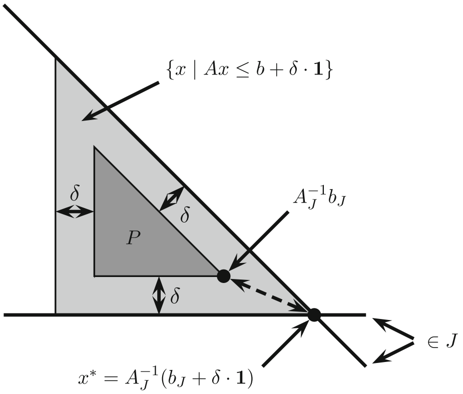

$\{x \mid Ax \leq b + \delta \cdot \mathbf{1}\}$
δ
δ
P
δ
$A_J^{-1} b_J$
$\in J$
$x^* = A_J^{-1}(b_J + \delta \cdot \mathbf{1})$

Fig. 2 We bound the distance of x^* to P by the distance to $A_J^{-1} b_J$ (see *dashed line*)

which implies that $\|A_I x^* - b_I + U_I' y\|_\infty \geq \frac{\delta}{n+r} - \frac{\delta}{4(n+r)} > \frac{\delta}{4r(n+r)}$ and consequently $x^* \notin \text{proj}_x(Q)$. This is a contradiction. □

Claim $\{x \in \mathbb{R}^n \mid Ax \leq b + \delta \mathbf{1}\} \subseteq P + \varepsilon$.

Proof of Claim It suffices to prove that every vertex x^* of $\{x \mid Ax \leq b + \delta \mathbf{1}\}$ has a distance of at most ε to P. There is a subsystem $A_J x \leq b_J + \delta \mathbf{1}$ of n constraints such that x^* is the unique solution of $A_J x = b_J + \delta \mathbf{1}$ or in other words $x^* = A_J^{-1}(b_J + \delta \mathbf{1})$. Since A has integral entries with absolute value at most Δ, we know that we can write $A_J^{-1} = (\frac{\alpha_{ij}}{\beta})_{i,j}$ with $\alpha_{ij}, \beta \in \{-(n\Delta)^n, \ldots, (n\Delta)^n\}$.[5]

Let us assume for the sake of contradiction that J was not a feasible basis for P, i.e. $A(A_J^{-1} b_J) \nleq b$. Well, then there is an index i with $A_i(A_J^{-1} b_J) > b_i$. In fact, even $A_i(A_J^{-1} b_J) \geq b_i + \frac{1}{\beta}$. But since we picked δ small enough, $|A_i x^* - A_i(A_J^{-1} b_J)| = |A_i A_J^{-1} \delta \mathbf{1}| \leq n^2 \cdot \Delta \cdot (n\Delta)^n \delta < \frac{1}{(n\Delta)^n} \leq \frac{1}{\beta}$, which is a contradiction.

Hence we may assume that J is indeed a feasible basis for P and we can bound the distance of x^* to P by the distance that the basic solution corresponding to basis J "moved" by shifting the hyperplanes by δ (see Fig. 2):

$$\|x^* - A_J^{-1} b_J\|_2 = \|A_J^{-1}(b_J + \delta \mathbf{1}) - A_J^{-1} b_J\|_2 = \|A_J^{-1} \delta \mathbf{1}\|_2 \leq n \cdot \delta \cdot (\Delta n)^n \leq \varepsilon.$$

Here we again used our choice of δ. □

Combining the proven claims yields $P \subseteq \text{proj}_x(Q) \subseteq P + \varepsilon$. □

7 Complexity theory considerations

After an informal publication of this result on ArXiv in April 2011, in November 2011, Fiorini, Massar, Pokutta, Tiwary and de Wolf [9] were able to prove that the extension

[5] By Cramer's rule, every entry (i, j) of the inverse of an $n \times n$ matrix M can be written as $\pm \frac{\det(M')}{\det(M)}$ for some submatrix M' of M. By the Hadamard bound, $|\det(M)| \leq \prod_{i=1}^n \|M^i\|_2 \leq (n\|M\|_\infty)^n$.

complexity of the *TSP polytope* is at least $2^{\Omega(n^{1/4})}$, where n is the number of nodes. Differently from this paper, their result is based on lower bounds for non-deterministic communication complexity. Moreover, their superpolynomial lower bound carries over to the *boolean quadric polytope*, the *cut polytope* and the *stable set polytope*.

The set of problems that admit compact formulations induce a non-uniform complexity class in a natural way. In the following, we want to briefly discuss, how this class relates to other, well studied classes (especially in view of [9]). For an up-to-date introduction into the topic of complexity theory, we recommend the textbook of [1]. Recall that $\{0,1\}^* = \bigcup_{n\geq 0}\{0,1\}^n$ is the set of all binary strings. By a slight abuse of notation we consider a $0/1$ string of length n also as a binary vector of dimension n. We want to define the class of problems that admit compact formulations as follows:

Definition 1 Let **CF** be the set of languages $L \subseteq \{0,1\}^*$ for which there exists a polynomial p such that for all $n \in \mathbb{N}$ there exist $A \in \mathbb{R}^{p(n)\times n}$, $B \in \mathbb{R}^{p(n)\times p(n)}$, $b \in \mathbb{R}^{p(n)}$ such that

$$\operatorname{conv}(L_n) = \{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^{p(n)} : Ax + By \leq b\},$$

where we abbreviate $L_n := \{x \in L : |x| = n\}$. By $\mathbf{CF}^{\text{enc}} \subseteq \mathbf{CF}$ we denote the subclass of languages, for which there exist integral matrices A, B and vectors b such that $\log(\max\{\|A\|_\infty, \|B\|_\infty, \|b\|_\infty\}) \leq p(n)$.

7.1 Properties of **CF**

Since any LP of polynomial size and encoding length can be solved in polynomial time, it is rather obvious that $\mathbf{CF}^{\text{enc}} \subseteq \mathbf{P}_{/\mathbf{poly}}$ (see also the remark of Yannakakis [23]). However, using Theorem 3, we can show that compact LPs can be solved in polynomial time exactly even if the coefficients are irrational.

Lemma 8 *Let $L \in$ **CF**. Then there is a non-uniform Turing machine (taking advice of size poly(n)) which on input $c \in \mathbb{Z}^n$, computes an optimum solution $x \in \{0,1\}^n$ to* $\max\{c^T x \mid x \in L_n\}$ *in time polynomial in n and* $\log \|c\|_\infty$.

Proof The first observation is that we can limit the encoding length of c by remembering the result of Frank and Tardos [12]:

> On input $c \in \mathbb{Z}^n$ and $N \in \mathbb{N}$ one can find in time polynomial in n, $\log\|c\|_\infty$ and $\log N$ a vector $c' \in \mathbb{Z}^n$ with $\|c'\|_\infty \leq 2^{4n^3} N^{n(n+2)}$ such that $\operatorname{sign}(c^T y) = \operatorname{sign}(c'^T y)$ for all $y \in \mathbb{Z}^n$ with $\|y\|_1 \leq N - 1$.[6]

In other words, choosing $N := n + 1$, for every objective function c, one can compute a $c' \in \mathbb{Z}^n$ with $\|c'\|_\infty \leq \alpha = 2^{O(n^3)}$ such that the set of optimum solutions for $\max\{c^T x \mid x \in L_n\}$ and $\max\{c'^T x \mid x \in L_n\}$ are identical. Next, define another vector $c'' \in \mathbb{Z}^n$ with $c''_i := 2^n c'_i + 2^{i-1}$. Then the optimum of $\max\{c''^T x \mid x \in L_n\}$

[6] As usual $\operatorname{sign}(r) \in \{-1, 0, 1\}$ gives the sign of a real number r.

is unique and it is also an optimum solution for $\max\{c'^T x \mid x \in L_n\}$ (the converse is not necessarily true).

Now, let Q be the approximation of $\text{conv}(L_n)$ with parameter $\varepsilon := \frac{1}{4n^2 \cdot 2^{2n}\alpha}$ as in Theorem 7. Note that the description length of Q is polynomial in n.

Next, we consider a non-uniform Turing machine which has the description of Q as advice and takes $c \in \mathbb{Z}^n$ as input. As described above, we then compute the vector c'' in polynomial time. Then we optimize c'' in polynomial time over Q (see [16]) and obtain an optimum solution x^*. Unfortunately, x^* does not need to be integral. However, the objective function value $c''x^*$ exceeds the value of the optimum integral solution by at most $\varepsilon\|c''\|_2 \leq \frac{1}{4}$ (see Theorem 7), thus $\gamma := \lfloor c''x^* \rfloor = \max\{c''x \mid x \in L_n\}$. Now the last n bits of γ encode the (unique) optimum integral solution, which we extract and return. □

Note that this implies also $\mathbf{CF} \subseteq \mathbf{P}_{/\mathbf{poly}}$, since given $x \in \{0,1\}^n$ testing whether $x \in Q$ is trivial. Secondly, Lemma 8 provides the following corollary:

Corollary 9 *Let $L \in \{0,1\}^*$ be a language such that the following problem is* **NP**-*hard:*

> *Given $c \in \mathbb{Z}^n$ and $k \in \mathbb{Z}$ as input. Decide whether there is an $x \in L_n$ with $c^T x \geq k$.*

Then $\mathbf{NP} \not\subseteq \mathbf{P}_{/\mathbf{poly}} \Rightarrow L \notin \mathbf{CF}$.

In our opinion this provides some evidence, that large coefficients might not be necessary for small extended formulations. We make the following conjecture:

Conjecture 10 $\mathbf{CF}^{\text{enc}} = \mathbf{CF}$.

For a graph $G = (V, E)$, the *stable set polytope* is $\text{conv}\{\chi(S) \mid S \subseteq V$ is a stable set in $G\}$, where we call S a *stable set* if it does not include adjacent vertices. Here $\chi(S) \in \{0,1\}^V$ denotes the characteristic vector of S. We will use a theorem of [9], which shows that there exists a sequence of graphs $\{G_n = ([n], E_n)\}_{n\in\mathbb{N}}$ such that the extension complexity of their stable set polytopes is at least $2^{\Omega(n^{1/2})}$. In other words, if we define the corresponding language

$$\texttt{STABLE} := \bigcup_{n\in\mathbb{N}} \{\chi(S) \in \mathbb{R}^n \mid S \subseteq [n] \text{ is stable set in } G_n\}$$

for this particular sequence of graphs, then $\texttt{STABLE} \notin \mathbf{CF}$.

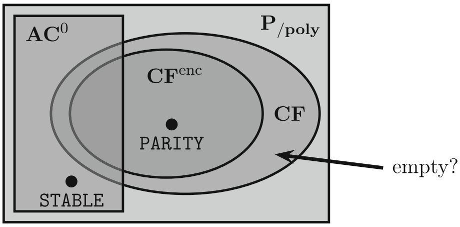
Corollary 11 $\mathbf{CF} \neq \mathbf{P}_{/\mathbf{poly}}$.

Proof This follows since $\texttt{STABLE} \notin \mathbf{CF}$ [9] and $\mathbf{CF} \subseteq \mathbf{P}_{/\mathbf{poly}}$, while $\texttt{STABLE} \in \mathbf{P}_{/\mathbf{poly}}$. □

Since optimizing over STABLE is **NP**-hard (while checking membership is easy) it comes to no suprise that there is no compact formulation for it.

However, at this point it is perfectly possible that there exists also a language $L \subseteq \{0,1\}^*$ with the following properties:

Fig. 3 Overview over complexity landscape



(A) There is an algorithm such that: On input $c \in \mathbb{Z}^n$, the algorithm computes $\max\{c^T x \mid x \in L_n\}$ in time polynomial in n and $\log \|c\|_\infty$.

(B) $L \notin \mathbf{CF}$.

In the opinion of the author, this is most likely the case.

7.2 **CF** versus $\mathbf{AC}^0$

It seems a challenging question, whether alternative characterizations exist for **CF**. However, it turns out that this class is incomparable to a well-studied subclass of $\mathbf{P}_{/\mathbf{poly}}$, which is called $\mathbf{AC}^0$. Recall that $\mathbf{AC}^0$ is the set of languages for which there are circuits with bounded depth and unbounded fan-in.

Recall that PARITY is the set of all $x \in \{0, 1\}^*$ such $\|x\|_1$ is odd. Then PARITY admits a compact formulation (with small integral coefficients; see e.g. [6]), thus PARITY $\in \mathbf{CF}^{\text{enc}}$. In a seminal result, Furst, Saxe and Sipser [11] showed that PARITY $\notin \mathbf{AC}^0$ and hence $\mathbf{CF} \not\subseteq \mathbf{AC}^0$ (in fact, even $\mathbf{CF}^{\text{enc}} \not\subseteq \mathbf{AC}^0$). However, the reverse is true as well (Fig. 3).

Theorem 12 $\mathbf{AC}^0 \not\subseteq \mathbf{CF}$.

Proof Recall that STABLE $\notin$ **CF**. On the other hand, it is not difficult to see that

$$\bigwedge_{\{u,v\} \in E_n} (\neg x_u \vee \neg x_v)$$

is a polynomial size, constant depth formula for STABLE, thus STABLE $\in \mathbf{AC}^0$. □

Acknowledgments The author is grateful to Samuel Fiorini for carefully reading a preliminary draft. Furthermore the author wants to thank Michel X. Goemans, Neil Olver and Rico Zenklusen as well as the anonymous referees for helpful comments. This study is supported by the Alexander von Humboldt Foundation within the Feodor Lynen program, by ONR grant N00014-11-1-0053 and by NSF contract CCF-0829878.

References

1. Arora, S., Barak, B.: Computational Complexity: A Modern Approach. Cambridge University Press, Cambridge (2009)
2. Balas, E.: Disjunctive programming and a hierarchy of relaxations for discrete optimization problems. SIAM J. Algebraic Discret. Methods **6**(3), 466–486 (1985)

3. Balas, E.: Disjunctive programming: properties of the convex hull of feasible points. Discret. Appl. Math. **89**(1–3), 3–44 (1998)
4. Barahona, F.: On cuts and matchings in planar graphs. Math. Program. **60**, 53–68 (1993). doi:10.1007/BF01580600
5. Bienstock, D.: Approximate formulations for 0–1 knapsack sets. Oper. Res. Lett. **36**(3), 317–320 (2008)
6. Conforti, M., Cornuéjols, G., Zambelli, G.: Extended formulations in combinatorial optimization. 4OR Q. J. Oper. Res. **8**, 1–48 (2010). doi:10.1007/s10288-010-0122-z
7. Dukes, W.M.B.: Bounds on the number of generalized partitions and some applications. Aust. J. Combin. **28**, 257–261 (2003)
8. Edmonds, J.: Maximum matching and a polyhedron with 0, 1-vertices. J. Res. Nat. Bur. Standards Sect. B **69**, 125–130 (1965)
9. Fiorini, S., Massar, S., Pokutta, S., Tiwary, H., de Wolf, R.: Linear vs. semidefinite extended formulations: exponential separation and strong lower bounds. CoRR abs/1111.0837, 2011. (To appear in STOC) (2012)
10. Fiorini, S., Rothvoß, T., Tiwary, H.: Extended formulations for polygons. Discret. Comput. Geom., pp. 1–11 (2012). doi:10.1007/s00454-012-9421-9
11. Furst, M., Saxe, J.B., Sipser, M.: Parity, circuits, and the polynomial-time hierarchy. Math. Syst. Theory **17**(1), 13–27 (1984)
12. Frank, A., Tardos, É.: An application of simultaneous diophantine approximation in combinatorial optimization. Combinatorica **7**(1), 49–65 (1987)
13. Gerards, A.M.H.: Compact systems for t-join and perfect matching polyhedra of graphs with bounded genus. Oper. Res. Lett. **10**(7), 377–382 (1991)
14. Goemans, M.: Smallest compact formulation for the permutahedron. Working paper. http://math.mit.edu/~goemans/PAPERS/permutahedron.pdf (2010)
15. Kaibel, V.: Extended Formulations in Combinatorial Optimization, ArXiv e-prints (2011)
16. Khachiyan, L.G.: A polynomial algorithm for linear programming. Soviet Math. Doklady **20**, 191–194, (Russian original in Doklady Akademiia Nauk SSSR, **244**, 1093–1096) (1979)
17. Kaibel, V., Pashkovich, K., Theis, D.O.: Symmetry matters for the sizes of extended formulations. In: IPCO, pp. 135–148 (2010)
18. Kipp Martin, R.: Using separation algorithms to generate mixed integer model reformulations. Oper. Res. Lett. **10**(3), 119–128 (1991)
19. Pritchard, D.: An lp with integrality gap 1+epsilon for multidimensional knapsack. CoRR, abs/1005.3324 (2010)
20. Schrijver, A.: Combinatorial Optimization. Polyhedra and Efficiency. vol. A,B,C, vol 24 of Algorithms and Combinatorics. Springer, Berlin (2003)
21. Shannon, C.E.: The synthesis of two-terminal switching circuits. Bell Syst. Tech. J. **28**, 59–98 (1949)
22. Vavasis, S.A.: On the complexity of nonnegative matrix factorization. SIAM J. Optim. **20**(3), 1364–1377 (2009)
23. Yannakakis, M.: Expressing combinatorial optimization problems by linear programs. J. Comput. Syst. Sci. **43**(3), 441–466 (1991)
24. Ziegler, G.M.: Lectures on 0/1-polytopes. In: Polytopes—combinatorics and computation (Oberwolfach, 1997), vol. 29 of DMV Sem., pp. 1–41. Birkhäuser, Basel (2000)