

Balas formulation for the union of polytopes is optimal

Michele Conforti*, Marco Di Summa*, Yuri Faenza†

November 6, 2017

Abstract

A celebrated theorem of Balas gives a linear mixed-integer formulation for the union of two nonempty polytopes whose relaxation gives the convex hull of this union. The number of inequalities in Balas formulation is linear in the number of inequalities that describe the two polytopes and the number of variables is doubled.

In this paper we show that this is best possible: in every dimension there exist two nonempty polytopes such that if a formulation for the convex hull of their union has a number of inequalities that is polynomial in the number of inequalities that describe the two polytopes, then the number of additional variables is at least linear in the dimension of the polytopes.

We then show that this result essentially carries over if one wants to approximate the convex hull of the union of two polytopes and also in the more restrictive setting of lift-and-project.

1 Introduction

Linear extensions are a powerful tool in linear optimization, since they allow to reduce an optimization problem over a polyhedron P to an analogous problem over a second polyhedron Q , that may be describable with a smaller system of linear constraints. For this reason, a number of recent studies (e.g. [4, 8, 14, 17]) focus on proving upper and lower bounds on the *extension complexity* of a polytope P , i.e., on the minimum number of linear inequalities needed to describe a linear extension of P . Bounds on the extension complexity hence guarantee (or disprove) the theoretical efficiency of linear programming methods for certain optimization problems.

For practical purposes, the number of additional variables used in a linear extension is also an important parameter, see e.g. [21]. This paper studies the minimum number of variables needed to obtain a linear extension for the convex hull of the union of two polytopes, where the number of inequalities describing the linear extension is polynomially bounded with respect to the number of inequalities in the descriptions of the two polytopes.

1.1 Balas formulation for the union of polytopes

For any set $X \subseteq \mathbb{R}^d$, we denote by $\text{conv}(X)$ the convex hull of X . We recall the following theorem of Balas [1].

Theorem 1. *Let $P_1 := \{x \in \mathbb{R}^d : A_1x \leq b_1\}$ and $P_2 := \{x \in \mathbb{R}^d : A_2x \leq b_2\}$ be nonempty polytopes. Then*

$$\begin{aligned} \text{conv}(P_1 \cup P_2) = \{x \in \mathbb{R}^d : \exists (x_1, x_2, \lambda) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \text{ s.t.} \\ x = x_1 + x_2; Ax_1 \leq \lambda b_1; Ax_2 \leq (1 - \lambda)b_2; 0 \leq \lambda \leq 1\}. \end{aligned} \quad (1)$$

*Dipartimento di Matematica “Tullio Levi-Civita”, Università degli studi di Padova, Italy. Supported by the grant SID 2016.

†IEOR, Columbia University. Supported by a SEAS-IEOR starting grant and by the Swiss National Science Foundation.

The above result can be seen as follows. By definition of the convex hull operator, we have that

$$\text{conv}(P_1 \cup P_2) = \{x : \exists (y_1, y_2, \lambda) \text{ s.t. } x = \lambda y_1 + (1 - \lambda)y_2; Ay_1 \leq b_1; Ay_2 \leq b_2; 0 \leq \lambda \leq 1\}. \quad (2)$$

Since P_1 and P_2 are nonempty polytopes, $\{x \in \mathbb{R}^d : A_1x \leq 0\} = \{x \in \mathbb{R}^d : A_2x \leq 0\} = \{0\}$. Therefore it is easy to argue that the system in (2) can be linearized by substituting $x_1 := \lambda y_1$ and $x_2 := (1 - \lambda)y_2$ to obtain the system in (1). We refer to [6, Theorem 4.39] for the case in which P_1 and P_2 are (possibly empty) polyhedra. Note that as $\text{conv}(P_1 \cup \dots \cup P_k) = \text{conv}(\text{conv}(P_1 \cup \dots \cup P_{k-1}) \cup P_k)$, restricting to the case $k = 2$ is with no loss of generality.

Theorem 1 is fundamental for the geometric approach to Integer Programming, where, given a polytope $P \subseteq \mathbb{R}^d$, tighter and tighter polyhedral relaxations of the set $S := P \cap \mathbb{Z}^d$ are obtained as convex hulls of union of polytopes. An example of this paradigm is as follows. Given $(\pi, \pi_0) \in \mathbb{Z}^d \times \mathbb{Z}$, let

$$P_0 := \{x \in P : \pi x \leq \pi_0\} \text{ and } P_1 := \{x \in P : \pi x \geq \pi_0 + 1\}. \quad (3)$$

Then $\text{conv}(P_0 \cup P_1) \cap \mathbb{Z}^d = S$ and $\text{conv}(P_0 \cup P_1) \subseteq P$. This second containment is strict if and only if $\pi_0 < \pi v < \pi_0 + 1$ for some vertex v of P . In this case $\text{conv}(P_0 \cup P_1)$ is a tighter polyhedral relaxation for S . The *split cuts* used in Integer Programming, see e.g. Chapter 5 in [6], are the linear inequalities that are valid for $\text{conv}(P_0 \cup P_1)$, for some $(\pi, \pi_0) \in \mathbb{Z}^d \times \mathbb{Z}$.

Given polytope $P \subseteq \mathbb{R}^d$, polytope $Q \subseteq \mathbb{R}^{d+m}$ is a *linear extension* (with m additional variables) of P if there exists an affine map $\psi : \mathbb{R}^{d+m} \rightarrow \mathbb{R}^d$ such that $P = \psi(Q)$. We allow $Q = P$. In this paper, the only affine maps we consider are orthogonal projections: i.e., Q is a linear extension of P when $P = \text{proj}_x(Q) := \{x \in \mathbb{R}^d : \exists y \in \mathbb{R}^m \text{ s.t. } (x, y) \in Q\}$. Note that restricting to orthogonal projections is with no loss of generality.

A system of inequalities describing a linear extension Q of P is a *formulation* of P whose *size* is the number of inequalities. The *extension complexity* of P is the minimum size of a formulation of P . Therefore, if we let

$$Q := \{(x, x_1, \lambda) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} : Ax_1 \leq \lambda b_1; A(x - x_1) \leq (1 - \lambda)b_2; 0 \leq \lambda \leq 1\}, \quad (4)$$

then Theorem 1 says that $\text{proj}_x(Q) = \text{conv}(P_1 \cup P_2)$. Furthermore, given formulations of size f_1 and f_2 of nonempty polytopes P_1 and P_2 , we see that the formulation of Q given in (4) has size $f_1 + f_2 + 2$ and has $2d + 1$ variables. Hence the number of constraints is linear in $f_1 + f_2$ and the number of *additional* variables is $d + 1$. (Weltge [20, proof of Proposition 3.1.1] observed that the inequalities $0 \leq \lambda \leq 1$ can be omitted from (4) if both P_1 and P_2 have dimension at least 1.)

The fact that the formulation of Q has size $f_1 + f_2 + 2$ has been exploited by several authors to construct small size linear extensions of polytopes that can be seen as the convex hull of the union of a polynomial number of polytopes with few inequalities. These results are surveyed e.g. in [5], [13].

While most of the literature focuses on the smallest number of inequalities defining a linear extension of a given polytope, in this paper we focus on the minimum number of additional variables needed in a linear extension. More specifically, we address the following question:

Given formulations of nonempty polytopes $P_1, P_2 \subseteq \mathbb{R}^d$ with sizes f_1, f_2 respectively, let Q be a linear extension of $\text{conv}(P_1 \cup P_2)$ whose formulation has $\text{poly}(f_1 + f_2)$ inequalities. What is the minimum number of additional variables that Q must have?

We stress that in the above question, the property $\text{proj}_x(Q) = \text{conv}(P_1 \cup P_2)$ must be satisfied for every choice of nonempty polytopes P_1, P_2 . If P_1, P_2 have specific properties, then few additional variables may suffice. For instance, Kaibel and Pashkovich [14] show that when P_2 is the reflection of P_1 with respect to a hyperplane that leaves P_1 on one side, then $\text{conv}(P_1 \cup P_2)$ admits a linear extension with only $f_1 + 2$ inequalities and one additional variable.

1.2 Our contribution

Our main result shows that if one constructs a formulation of $\text{conv}(P_1 \cup P_2)$ whose size is polynomially bounded in the sizes of the descriptions of P_1 and P_2 , then $\Omega(d)$ additional variables are needed. In other words, the construction of Balas is optimal in this respect. More specifically, we have the following:

Theorem 2. *Fix a polynomial σ . For each odd $d \in \mathbb{N}$, there exist formulations of nonempty polytopes $P_1, P_2 \subseteq \mathbb{R}^d$ of size f_1 and f_2 respectively, such that any formulation of $\text{conv}(P_1 \cup P_2)$ of size at most $\sigma(f_1 + f_2)$ has $\Omega(d)$ additional variables.*

We then turn to polytopes whose orthogonal projection gives an outer approximation of $\text{conv}(P_1 \cup P_2)$. Given $\varepsilon \geq 0$, we say that a polytope $P' \subseteq \mathbb{R}^d$ is an ε -approximation of a nonempty polytope $P \subseteq \mathbb{R}^d$ if $P \subseteq P'$ and for all $c \in \mathbb{R}^d$ we have

$$\max_{x \in P'} cx - \min_{x \in P'} cx \leq (1 + \varepsilon) \left(\max_{x \in P} cx - \min_{x \in P} cx \right). \quad (5)$$

In particular, P and P' have that same affine hull. So the only ε -approximation of a point is the point itself.

Our second result can be seen as an ε -approximate version of Theorem 2.

Theorem 3. *Fix $\varepsilon > 0$ and a polynomial σ . For each $d \in \mathbb{N}$, there exist formulations of nonempty polytopes $P_1, P_2 \subseteq \mathbb{R}^d$ of size f_1 and f_2 respectively, such that any ε -approximation of $\text{conv}(P_1 \cup P_2)$ of size at most $\sigma(f_1 + f_2)$ has $\Omega(d/\log d)$ additional variables.*

1.3 MIP representations and the convex hull property

A set $S \subseteq \mathbb{R}^d$ is MIP (mixed-integer programming) representable if there exist matrices A, B, C and a vector d such that $S = \text{proj}_x(Q)$, where $Q := \{(x, y, z) : Ax + By + Cz \leq d, z \text{ integral}\}$. Under the condition that matrices A, B, C and vector d be rational, the MIP representable sets were characterized by Jeroslow and Lowe [11], see also Basu et al. [3] for a different characterization. We refer to the recent survey of Vielma [18] on MIP representability.

If P_1 and P_2 are nonempty polytopes, then $P_1 \cup P_2$ is a MIP representable set. Indeed a MIP representation of this set can be obtained by imposing integrality on variable λ in the system in (4); see [1]. This is not the only representation of $P_1 \cup P_2$: the famous big- M method gives a representation with $f_1 + f_2 + 2$ inequalities and only $d + 1$ variables, where f_1 and f_2 are the sizes of formulations of P_1, P_2 . So this representation is more compact than the one given by Balas.

If $Q := \{(x, y, z) : Ax + By + Cz \leq d, z \text{ integral}\}$ is a MIP representation of $P_1 \cup P_2$, then

$$\text{conv}(P_1 \cup P_2) \subseteq \text{proj}_x(\{(x, y, z) : Ax + By + Cz \leq d\}). \quad (6)$$

We say that a MIP representation of $P_1 \cup P_2$ has the *convex hull property* if the two sets in (6) coincide. It follows from Theorem 1 that the MIP representation obtained by imposing integrality on λ in (4) has the convex hull property and it is immediate to check that the one given by the big- M method does not.

The following is a consequence of Theorem 2.

Theorem 4. *Fix a polynomial σ . For each odd $d \in \mathbb{N}$, there exist formulations of nonempty polytopes $P_1, P_2 \subseteq \mathbb{R}^d$ of size f_1 and f_2 respectively, such that any MIP representation of $P_1 \cup P_2$ with at most $\sigma(f_1 + f_2)$ inequalities that has the convex hull property has $\Omega(d)$ additional variables.*

This paper is organized as follows. In Section 2 we describe the main idea of our approach, which relies on a counting argument and on the existence of a polytope $P \subseteq \mathbb{R}^d$ that has $d^{\Omega(d)}$ facets and is the convex hull of two polytopes with polynomially (in d) many facets. In Section 3 we develop some geometric tools for the construction of P , which is then obtained in Section 4 via a construction using the Cayley embedding. In Section 5 we prove Theorem 3, while in Section 6 we investigate some implications of Theorem 2 for the technique of lift-and-project (see Theorem 20).

2 An outline of the proof

We assume familiarity with polyhedral theory (see e.g. [6, 22]). Given $S \subseteq \mathbb{R}^d$, we denote with $\text{conv}(S)$, $\text{aff}(S)$ and $\text{cone}(S)$ its convex hull, affine hull and conical hull. We also let $\text{int}(S)$, $\text{relint}(S)$, $\text{dim}(S) := \text{dim}(\text{aff}(S))$ denote the interior, relative interior, and affine dimension of S . A d -polytope is a polytope of dimension d , and a k -face of a polytope P is a face of P of dimension k .

Our approach to proving Theorem 2 is based on the following lemma relating the number of facets of a linear extension of a given polytope to the number of additional variables. Given a $P \subseteq \mathbb{R}^d$, we let $f(P)$ denote the number of facets of P .

Lemma 5. *Let $Q \subseteq \mathbb{R}^{d+m}$ be a $(d+m)$ -polytope which is a linear extension of a d -polytope $P \subseteq \mathbb{R}^d$. Then $m = \Omega\left(\frac{\log f(P)}{\log f(Q)}\right)$.*

Proof. For $0 \leq k \leq d+m-1$, every k -face of Q is the intersection of $d+m-k$ facets of Q (this choice may not be unique). Then, by the binomial theorem, the number of proper faces of Q of dimension at least $d-1$ is at most

$$\sum_{j=1}^{m+1} \binom{f(Q)}{j} \leq \sum_{j=1}^{m+1} (f(Q))^j \cdot 1^{m+1-j} \leq (f(Q) + 1)^{m+1}.$$

Let F be a facet of P . Then F is the projection of a face F_Q of Q and $\text{dim}(F_Q) \geq \text{dim}(F) = d-1$. Therefore the number of facets of P is bounded by the number of proper faces of Q of dimension at least $d-1$ and by the above argument we have that $m = \Omega\left(\frac{\log f(P)}{\log f(Q)}\right)$. \square

We will show later (Theorem 15) that for every odd $d \geq 3$ there exists a d -polytope P having $d^{\Omega(d)}$ facets which is the convex hull of two polytopes P_1 and P_2 , each having $(d-1)^2$ facets. Let Q be any linear extension of P with $f(Q) = \sigma(f(P_1) + f(P_2))$. It is well-known that Q can be assumed to be full-dimensional. Since $f(P_1) = f(P_2) = (d-1)^2$, we have that $f(Q)$ is bounded by a polynomial in d . Let $d+m$ be the dimension of Q . Then by Lemma 5, the number of additional variables is $m = \Omega\left(\frac{\log f(P)}{\log f(Q)}\right) = \Omega\left(\frac{d \log d}{\log d}\right) = \Omega(d)$. This proves Theorem 2. Similar counting arguments (with different polytopes) settle Theorem 3 and Theorem 20.

The next two sections are devoting to proving Theorem 15, which is the missing ingredient to complete proof of Theorem 2.

3 Some tools

3.1 Optimality Cones

We let $\mathcal{F}(P)$ be the set of the nonempty faces of a polytope P .

Let $P \subseteq \mathbb{R}^d$ be a polytope and let $F \in \mathcal{F}(P)$ be a nonempty face of P . An inequality $cx \leq \delta$ defines F if

$$P \subseteq \{x \in \mathbb{R}^d : cx \leq \delta\} \text{ and } F = P \cap \{x \in \mathbb{R}^d : cx = \delta\}.$$

The *optimality cone* of F is the set

$$C_P(F) := \{c \in \mathbb{R}^d : \exists \delta \text{ s.t. } cx \leq \delta \text{ defines } F\}.$$

We also consider the set

$$\overline{C}_P(F) := \bigcup_{F' \in \mathcal{F}(P), F' \supseteq F} C_P(F').$$

Remark 6. *Let $P \subseteq \mathbb{R}^d$ be a nonempty polytope. The following hold:*

1. $\bigcup_{F \in \mathcal{F}(P)} C_P(F) = \mathbb{R}^d$ and $C_P(F_i) \cap C_P(F_j) = \emptyset$ for every $F_i, F_j \in \mathcal{F}(P)$ with $F_i \neq F_j$.

2. $C_P(P)$ is the subspace $\{c \in \mathbb{R}^d : \exists \delta \text{ s.t. } cx = \delta \forall x \in P\}$. So $\dim(P) = d$ if and only if $C_P(P) = \{0\}$.

3. For every $F \in \mathcal{F}(P)$, $\overline{C}_P(F)$ is the polyhedral cone generated by

$$\{c \in \mathbb{R}^d : \exists \delta \text{ s.t. } cx \leq \delta \text{ defines } P \text{ or a facet containing } F\}.$$

4. For every $F \in \mathcal{F}(P)$, we have $C_P(F) = \text{relint}(\overline{C}_P(F))$.

5. For every $F \in \mathcal{F}(P)$, $\dim(F) + \dim(C_P(F)) = d$.

Lemma 7. Let P be a d -polytope. For every dimension k , $0 \leq k \leq d$, there exists a linear subspace $V \subseteq \mathbb{R}^d$ such that $\dim(V) = k$ and $V \cap \text{aff}(\overline{C}_P(F)) = \{0\}$ for every k -face F of P .

Proof. Define

$$\mathcal{A} := \bigcup_{F \text{ is a } k\text{-face of } P} \text{aff}(\overline{C}_P(F)).$$

We iteratively construct subspaces $\{0\} =: V_0 \subset \dots \subset V_k$ such that $\dim(V_i) = i$ and $V_i \cap \mathcal{A} = \{0\}$, $i = 0, \dots, k$.

Assume $\dim(V_i) = i$ and $V_i \cap \mathcal{A} = \{0\}$ for some $i < k$. For every k -face F of P we have that $\text{aff}(\overline{C}_P(F) \cup V_i)$ is a linear space of dimension $d - k + i < d$. Since the number of k -faces of P is finite, the set

$$\mathcal{S} := \bigcup_{F \text{ is a } k\text{-face of } P} \text{aff}(\overline{C}_P(F) \cup V_i)$$

has Lebesgue measure 0. Therefore $\mathbb{R}^d \setminus \mathcal{S}$ contains a nonzero vector v . Let V_{i+1} be the linear space generated by $V_i \cup \{v\}$. Then $\dim(V_{i+1}) = i + 1$ and $V_{i+1} \cap \mathcal{A} = \{0\}$. \square

3.2 The polar of a cyclic polytope

The *moment curve* in \mathbb{R}^d is defined as

$$t \mapsto x(t) := \begin{pmatrix} t^1 \\ t^2 \\ \vdots \\ t^d \end{pmatrix} \in \mathbb{R}^d.$$

Given pairwise distinct real numbers t_1, \dots, t_k , the *cyclic polytope* $P^{Cy}(d, t_1, \dots, t_k)$ is $\text{conv}(x(t_1), \dots, x(t_k))$. It is well-known that, for d and k fixed, the combinatorial structure of a cyclic polytope does not depend on the choice of t_1, \dots, t_k . So we denote such a polytope by $P^{Cy}(d, k)$. In particular (see [9, Section 4.7]):

Lemma 8. For $k \geq d + 1$, $P^{Cy}(d, k)$ is a d -polytope with k vertices which is simplicial (i.e., all of its proper faces are simplices). For every subset S of vertices with $|S| \leq \frac{d}{2}$, $\text{conv}(S)$ is a $(|S| - 1)$ -face. So for $h \leq \frac{d}{2} - 1$, $P^{Cy}(d, k)$ has $\binom{k}{h+1}$ h -faces.

Given $P^{Cy}(d, k)$, $k \geq d + 1$, apply a translation so that $0 \in \text{int}(P^{Cy}(d, k))$. Let $D^{Cy}(d, k)$ be the polar of this translated polytope. Then by the above lemma and [22, Corollary 2.14] we obtain:

Remark 9. For $k \geq d + 1$, $0 \in \text{int}(D^{Cy}(d, k))$ and $D^{Cy}(d, k)$ is a d -polytope with k facets that is simple (i.e., every h -face, with $0 \leq h \leq d - 1$, is the intersection of exactly $d - h$ facets). For every subset S of facets with $|S| \leq \frac{d}{2}$, the intersection of the facets in S is a $(d - |S|)$ -face of P . So for $h \leq \frac{d}{2}$, P has $\binom{k}{h}$ $(d - h)$ -faces.

3.3 A perturbation of the polar of a cyclic polytope

Since for $k \geq d + 1$, $0 \in \text{int}(D^{Cy}(d, k))$, every valid inequality for $D^{Cy}(d, k)$ can be written in the form $ax \leq 1$. Assume now d even and $k = d^2$. Hence $D^{Cy}(d, d^2)$ has d^2 facets, and we arbitrarily partition the normals to its facets into $d/2$ color classes of size $2d$, so that $D^{Cy}(d, d^2)$ can be written as

$$\{x \in \mathbb{R}^d : a_j^i x \leq 1, i = 1, \dots, 2d, j = 1, \dots, d/2\},$$

where for every j the vectors a_j^1, \dots, a_j^{2d} are the normals that belong to color class j .

By Lemma 7, there exists a linear subspace $V \subseteq \mathbb{R}^d$ such that $\dim(V) = d/2$ and

$$V \cap \text{aff}(\overline{C}_{D^{Cy}(d, d^2)}(F)) = \{0\}$$

for each $(d/2)$ -face F of $D^{Cy}(d, d^2)$. Let $u_1, \dots, u_{d/2} \in V$ be such that $\text{conv}(u_1, \dots, u_{d/2})$ is a $(d/2 - 1)$ -simplex and $0 \in \text{reint}(\text{conv}(u_1, \dots, u_{d/2}))$. Note that the norms of vectors u_i can be arbitrarily small. Consider the following polytope:

$$Q^{Cy}(d, d^2) := \{x \in \mathbb{R}^d : (a_j^i + u_j)x \leq 1, i = 1 \dots, 2d, j = 1, \dots, d/2\}.$$

Remark 10. Since, by Remark 9, $D^{Cy}(d, d^2)$ is a simple d -polytope, $u_1, \dots, u_{d/2}$ can be scaled so that the combinatorial structures of $D^{Cy}(d, d^2)$ and $Q^{Cy}(d, d^2)$ coincide (see Section 2.5 in [22]). So the properties of Remark 9 hold for $Q^{Cy}(d, d^2)$ as well, and there is an isomorphism between the face lattices of $D^{Cy}(d, d^2)$ and $Q^{Cy}(d, d^2)$.

Call a $(d/2)$ -face of $D^{Cy}(d, d^2)$ colorful if it is the intersection of $d/2$ facets, no two of them having the same color. More precisely, a face F is colorful if there exist indices $i_j, j = 1, \dots, d/2$, such that:

$$F = \{x \in D^{Cy}(d, d^2) : a_j^{i_j} x = 1, j = 1, \dots, d/2\}. \quad (7)$$

Given a colorful face F described as above, let

$$F' := \{x \in Q^{Cy}(d, d^2) : (a_j^{i_j} + u_j)x = 1, j = 1, \dots, d/2\} \quad (8)$$

be the corresponding colorful face of $Q^{Cy}(d, d^2)$. Because of Remark 10, F' has dimension $d/2$.

Lemma 11. Let F and F' be corresponding colorful faces of $D^{Cy}(d, d^2)$ and $Q^{Cy}(d, d^2)$ respectively. Then

$$C_{D^{Cy}(d, d^2)}(F) \cap C_{Q^{Cy}(d, d^2)}(F') = \{\lambda r, \lambda > 0\}$$

for some $r \in \mathbb{R}^d \setminus \{0\}$.

Proof. Assume that F is given as in (7). Since $D^{Cy}(d, d^2)$ and $Q^{Cy}(d, d^2)$ are d -polytopes by Remarks 9 and 10, by Remark 6 we have that

$$\overline{C}_{D^{Cy}(d, d^2)}(F) = \text{cone}(a_j^{i_j}, j = 1, \dots, d/2), \quad \overline{C}_{Q^{Cy}(d, d^2)}(F') = \text{cone}(a_j^{i_j} + u_j, j = 1, \dots, d/2). \quad (9)$$

Again by Remark 6, $\overline{C}_{D^{Cy}(d, d^2)}(F)$ and $\overline{C}_{Q^{Cy}(d, d^2)}(F')$ are pointed cones. This shows that each point from $(\overline{C}_{D^{Cy}(d, d^2)}(F) \cap \overline{C}_{Q^{Cy}(d, d^2)}(F')) \setminus \{0\}$ corresponds to a solution to the system

$$\sum_{j=1}^{d/2} a_j^{i_j} \mu_j = \sum_{j=1}^{d/2} (a_j^{i_j} + u_j) \nu_j, \mu_j \geq 0, \nu_j \geq 0, j = 1 \dots, d/2 \quad (10)$$

where the μ_j 's are not all equal to 0 and the ν_j 's are not all equal to 0.

Since $u_1, \dots, u_{d/2}$ belong to V and $V \cap \text{aff}(\overline{C}_P(F)) = \{0\}$ by Lemma 7, every solution to the system $\sum_{j=1}^{d/2} a_j^{i_j} \mu_j = \sum_{j=1}^{d/2} (a_j^{i_j} + u_j) \nu_j$ must satisfy

$$\sum_{j=1}^{d/2} u_j \nu_j = 0.$$

Since by construction $\text{conv}(u_1, \dots, u_{d/2})$ is a $(d/2 - 1)$ -simplex and $0 \in \text{relint}(\text{conv}(u_1, \dots, u_{d/2}))$, the system

$$\sum_{j=1}^{d/2} u_j \nu_j = 0, \quad \nu_j \geq 0, \quad j = 1 \dots, d/2$$

admits a unique (up to scaling) nonzero solution $\bar{\nu}_j$, and furthermore $\bar{\nu}_j > 0$, $j = 1 \dots, d/2$.

Therefore the system (10) admits a unique (again, up to scaling) solution $\bar{\mu}, \bar{\nu}$, and this solution satisfies $\bar{\mu}_j = \bar{\nu}_j > 0$, $j = 1 \dots, d/2$.

By Remark 6, we have that

$$C_{D^{C_y}(d,d^2)}(F) = \text{relint}(\overline{C}_{D^{C_y}(d,d^2)}(F)) \quad \text{and} \quad C_{Q^{C_y}(d,d^2)}(F') = \text{relint}(\overline{C}_{Q^{C_y}(d,d^2)}(F')). \quad (11)$$

Let $r := \sum_{j=1}^{d/2} a_j^{i_j} \bar{\mu}_j$. Since $\bar{\mu}_j > 0$ for $j = 1 \dots, d/2$, we have that $r \in \text{relint}(\overline{C}_{D^{C_y}(d,d^2)}(F)) \cap \text{relint}(\overline{C}_{Q^{C_y}(d,d^2)}(F'))$. Then by (11) we have that $C_{D^{C_y}(d,d^2)}(F) \cap C_{Q^{C_y}(d,d^2)}(F') = \{\lambda r, \lambda > 0\}$. \square

4 A polyhedral construction

Let $P_0, P_1 \subseteq \mathbb{R}^{d-1}$ be $(d-1)$ -polytopes. The *Cayley embedding* [10] of P_0 and P_1 is the d -polytope

$$\text{conv} \left(\begin{pmatrix} P_0 \\ 0 \end{pmatrix} \cup \begin{pmatrix} P_1 \\ 1 \end{pmatrix} \right) \subseteq \mathbb{R}^d,$$

where for a set $S \subseteq \mathbb{R}^{d-1}$ we define notation $\begin{pmatrix} S \\ \alpha \end{pmatrix} := \left\{ \begin{pmatrix} x \\ \alpha \end{pmatrix} : x \in S \right\}$.

Note that given $x \in \mathbb{R}^{d-1}$, the point $(x, 1/2)$ belongs to the Cayley embedding of P_0 and P_1 if and only if $x \in \frac{1}{2}P_0 + \frac{1}{2}P_1$. Some extremal properties of the facial structure of the Minkowski sum of polytopes have been investigated, e.g., in [15, 19]. However, to the best of our knowledge the construction below is new.

Remark 12. Let $P_0, P_1 \subseteq \mathbb{R}^{d-1}$ be $(d-1)$ -polytopes and P be the Cayley embedding of P_0 and P_1 . Given $F_0 \in \mathcal{F}(P_0)$ and $F_1 \in \mathcal{F}(P_1)$, let F be the Cayley embedding of F_0 and F_1 . Then, given $x \in \mathbb{R}^d$, we have that $x \in F$ if and only if $0 \leq x_d \leq 1$ and there exist $x^0 \in \begin{pmatrix} F_0 \\ 0 \end{pmatrix}$ and $x^1 \in \begin{pmatrix} F_1 \\ 1 \end{pmatrix}$ such that $x = (1 - x_d)x^0 + x_d x^1$ (where x_d is the last component of x).

Lemma 13. Let $P_0, P_1 \subseteq \mathbb{R}^{d-1}$ be $(d-1)$ -polytopes and P be the Cayley embedding of P_0 and P_1 . Given $F_0 \in \mathcal{F}(P_0)$ and $F_1 \in \mathcal{F}(P_1)$, let F be the Cayley embedding of F_0 and F_1 . Then F is a face of P if and only if $C_{P_0}(F_0) \cap C_{P_1}(F_1) \neq \emptyset$. Furthermore, in this case, given $(r, \alpha) \in \mathbb{R}^{d-1} \times \mathbb{R}$ we have that $(r, \alpha) \in C_P(F)$ if and only if $r \in C_{P_0}(F_0) \cap C_{P_1}(F_1)$ and $\alpha = \max\{rx : x \in P_0\} - \max\{rx : x \in P_1\}$.

Proof. By Remark 12 we have that given $x \in \mathbb{R}^d$, $x \in P$ if and only if $0 \leq x_d \leq 1$ and there exist $x^0 \in \begin{pmatrix} P_0 \\ 0 \end{pmatrix}$ and $x^1 \in \begin{pmatrix} P_1 \\ 1 \end{pmatrix}$ such that $x = (1 - x_d)x^0 + x_d x^1$. Therefore, given $(r, \gamma) \in \mathbb{R}^{d-1} \times \mathbb{R}$ and $x \in P$, we have that

$$(r, \gamma)x = (r, \gamma)((1 - x_d)x^0 + x_d x^1) = (1 - x_d)(r, 0)x^0 + x_d((r, 0)x^1 + \gamma) \leq (1 - x_d)\alpha_0 + x_d(\alpha_1 + \gamma), \quad (12)$$

where $\beta_0 := \max\{rx : x \in P_0\}$, $\beta_1 := \max\{rx : x \in P_1\}$.

Assume now $r \in C_{P_0}(F_0) \cap C_{P_1}(F_1)$, i.e., $rx \leq \beta_0$ defines F_0 and $rx \leq \beta_1$ defines F_1 . Then if we let $\gamma = \alpha = \beta_0 - \beta_1$, we have that by (12) and Remark 12, the inequality $(r, \gamma)x \leq \beta_0$ is valid for P and is satisfied at equality if and only if $x \in F$. Therefore F is a face of P and $(r, \alpha) \in C_P(F)$. This proves the “if” direction of both equivalences in the statement.

Assume now that F is a face of P . Take $(r, \alpha) \in C_P(F)$ and let β be such that $(r, \alpha)x \leq \beta$ defines F . Then

$$\beta \geq \max\{rx : x \in P_0\} \quad \text{and} \quad \beta - \alpha \geq \max\{rx : x \in P_1\}.$$

Furthermore since $(r, \alpha) \in C_P(F)$, F is the Cayley embedding of F_0, F_1 , and F_0, F_1 are both nonempty, the above two inequalities are satisfied at equality. This shows $\alpha = \max\{rx : x \in P_0\} - \max\{rx : x \in P_1\}$.

We finally show $r \in C_{P_0}(F_0) \cap C_{P_1}(F_1)$. Let F_0^*, F_1^* be the faces of P_0, P_1 such that $r \in C_{P_0}(F_0^*) \cap C_{P_1}(F_1^*)$ (the existence of F_0^*, F_1^* is guaranteed by 1. of Remark 6). Assume $F_0 \neq F_0^*$ or $F_1 \neq F_1^*$, and let F^* be the Cayley embedding of F_0^*, F_1^* . Then $F^* \neq F$ and by the ‘‘if’’ part of the lemma, $(r, \alpha) \in C_P(F^*)$. Therefore $(r, \alpha) \in C_P(F) \cap C_P(F^*)$, a contradiction to 1. of Remark 6, and this concludes the proof of ‘‘only if’’ part. \square

Now Remark 6 and Lemma 13 imply the following:

Corollary 14. *Let $P_0, P_1 \subseteq \mathbb{R}^{d-1}$ be $(d-1)$ -polytopes and P be the Cayley embedding of P_0 and P_1 . Given $F_0 \in \mathcal{F}(P_0)$ and $F_1 \in \mathcal{F}(P_1)$, let F be the Cayley embedding of F_0 and F_1 . Then F is a facet of P if and only if $C_{P_0}(F_0) \cap C_{P_1}(F_1) = \{\lambda r, \lambda > 0\}$ for some $r \in \mathbb{R}^d \setminus \{0\}$.*

We now can provide a constructive proof of the following:

Theorem 15. *For every even $d \geq 2$ there exists a $(d+1)$ -polytope having $d^{\Omega(d)}$ facets which is the Cayley embedding of d -polytopes P_1 and P_2 , each having d^2 facets.*

Proof. Let $d \geq 2$ be even and fix a coloring of the facets of $D^{Cy}(d, d^2)$. Let F be a colorful face of $D^{Cy}(d, d^2)$ and F' be the corresponding face of $Q^{Cy}(d, d^2)$. By Lemma 11 we have that $C_{D^{Cy}(d, d^2)}(F) \cap C_{Q^{Cy}(d, d^2)}(F') = \{\lambda r, \lambda > 0\}$ for some $r \in \mathbb{R}^d \setminus \{0\}$. By Corollary 14, the Cayley embedding of F, F' is a facet of the Cayley embedding of $D^{Cy}(d, d^2)$ and $Q^{Cy}(d, d^2)$.

By Remark 9, the intersection of every $d/2$ facets of $D^{Cy}(d, d^2)$ forms a distinct face. By definition, the number of colorful $(d/2)$ -faces of $D^{Cy}(d, d^2)$ is $(2d)^{d/2} = d^{\Omega(d)}$. Therefore the Cayley embedding of $P_1 := D^{Cy}(d, d^2)$ and $P_2 := Q^{Cy}(d, d^2)$ has $d^{\Omega(d)}$ facets. Since P_1 and P_2 have d^2 facets each, this proves the theorem. \square

As shown in Section 2, the above theorem implies Theorem 2.

5 Proof of Theorem 3

The d -dimensional *cross-polytope* is (see e.g. [22]):

$$Q_d^\Delta := \left\{ x \in \mathbb{R}^d : \sum_{i \in I} x_i - \sum_{i \in [d] \setminus I} x_i \leq 1 : \forall I \subseteq [d] \right\},$$

with $[d] := \{1, \dots, d\}$. Q_d^Δ has 2^d facets, as every inequality in the above description defines a facet.

Consider the following $(d-1)$ -simplices:

$$P_1 := \left\{ x \in [0, 1]^d : \sum_{i=1}^d x_i = 1 \right\}, \quad P_{-1} := \left\{ x \in [-1, 0]^d : \sum_{i=1}^d x_i = -1 \right\}.$$

Since Q_d^Δ has $2d$ vertices, namely $\pm e_i$ for $i = 1, \dots, d$ (the unit vectors and their negatives), it follows that $Q_d^\Delta = \text{conv}(P_1 \cup P_{-1})$. Therefore Q_d^Δ is a d -polytope with 2^d facets which is the convex hull of two of its facets that are $(d-1)$ -simplices. (This choice is not unique: any two parallel facets of Q_d^Δ will do.)

We show that, for every constant $\varepsilon > 0$, any ε -approximation of the cross-polytope must still have an exponential number of facets. We will then invoke Lemma 5 to conclude the proof of Theorem 3.

The following observation allows us to focus on ε -approximations that only use facet-defining inequalities.

Lemma 16. *Let $Q \subseteq \mathbb{R}^d$ be an ε -approximation of a d -polytope $P \subseteq \mathbb{R}^d$ for some $\varepsilon > 0$. Then there exists an ε -approximation of P with at most $d \cdot f(Q)$ inequalities, each of which defines a facet of P .*

Proof. Let $cx \leq \delta$ be any facet-defining inequality for Q . Then $cx \leq \delta$ is valid for P and without loss of generality we may assume that $cx \leq \delta$ is supporting for P . Hence, by Caratheodory's theorem, it is a conic combination of at most d facet-defining inequalities for P . Hence, we can replace $cx \leq \delta$ with the facet-defining inequalities for P that define it and obtain a polytope P' such that $P \supseteq P' \supseteq Q$. By repeating the procedure for all facet-defining inequalities for Q , we obtain the claimed result. \square

Lemma 17. *Given $\varepsilon > 0$, there exists $\kappa > 1$ such that every ε -approximation of Q_d^Δ has $\Omega(\kappa^d)$ facets.*

Proof. We exhibit a set \mathcal{S} of points that *cannot* belong to any ε -approximation of Q_d^Δ , but such that any facet-defining inequality for Q_d^Δ can cut off *at most* t of them. Hence by Lemma 16, the number of inequalities needed to describe an ε -approximation of Q_d^Δ is at least $|\mathcal{S}|/(dt)$. (Our proof approach can be interpreted as an extension of those in [7, 16].)

Let $\delta > 2\varepsilon$ be fixed. Consider the family $\mathcal{S} \subseteq \mathbb{R}^d$ of 2^d points having coordinates equal to $\pm(1 + \delta)/d$.

Claim 18. *Let $x^* \in \mathcal{S}$. Then x^* does not belong to any ε -approximation of P .*

Proof of claim. Let c be the objective function with $c_i = 1$ if $x_i^* > 0$ and $c_i = -1$ if $x_i^* < 0$. Then for any polytope P' that contains $Q_d^\Delta \cup \{x^*\}$ we have:

$$\max_{x \in P'} cx - \min_{x \in P'} cx \geq (1 + \delta) - (-1) = 2 + \delta > 2(1 + \varepsilon) = (1 + \varepsilon) \left(\max_{x \in Q_d^\Delta} cx - \min_{x \in Q_d^\Delta} cx \right).$$

Therefore P' is not an ε -approximation to Q_d^Δ . \diamond

Claim 19. *Let $I \subseteq [d]$. There exists $\bar{\kappa} < 2$ such that $\sum_{i \in I} x_i - \sum_{i \in [d] \setminus I} x_i \leq 1$ is violated by at most $\bar{\kappa}^d$ points from \mathcal{S} .*

Proof of claim. By the symmetry of Q_d^Δ , it suffices to prove the statement for the inequality $\sum_{i \in [d]} x_i \leq 1$. Fix $x^* \in \mathcal{S}$ and suppose t of its components are positive. Then

$$\sum_{i \in [d]} x_i^* = t \frac{1 + \delta}{d} - (d - t) \frac{1 + \delta}{d} = 2t \frac{1 + \delta}{d} - 1 - \delta,$$

hence the inequality is violated if and only if

$$2t \frac{1 + \delta}{d} - 1 - \delta > 1 \quad \Leftrightarrow \quad t > \frac{d}{2} \cdot \frac{2 + \delta}{1 + \delta}.$$

Define $\gamma := \frac{1}{2} \cdot \frac{2 + \delta}{1 + \delta} > \frac{1}{2}$. Then a point in \mathcal{S} violates $\sum_{i \in [d]} x_i \leq 1$ if and only if it has more than γd positive entries. The number of points with this property is upper bounded by

$$\sum_{j=\lceil \gamma d \rceil}^d \binom{d}{j} = \sum_{j=0}^{\lfloor (1-\gamma)d \rfloor} \binom{d}{j} \leq 2^{dH(1-\gamma)},$$

where we used the well-known bound $\sum_{j=0}^k \binom{n}{j} \leq 2^{nH(k/n)}$ that is valid for $k \leq \frac{n}{2}$ and uses the entropy function $H(p) = -x \log_2(p) - (1 - p) \log_2(1 - p)$ (see e.g. [12]). It is well-known that $H(p) < H(1/2) = 1$ for all $0 \leq p < 1/2$. Since ε and δ are fixed, $1 - \gamma$ is a constant strictly less than $1/2$, and thus we conclude that $2^{dH(1-\gamma)} \leq \bar{\kappa}^d$ for some $\bar{\kappa} < 2$. \diamond

Putting everything together, the number of inequalities needed to describe an ε -approximation of Q_d^Δ is at least

$$\frac{|\mathcal{S}|}{d\bar{\kappa}^d} = \frac{1}{d} \left(\frac{2}{\bar{\kappa}} \right)^d = \Omega(\kappa^d)$$

for some $\kappa > 1$, as required. \square

Proof of Theorem 3. Fix $\varepsilon > 0$, and let P' be an ε -approximation of Q_d^Δ . Then P' has $\Omega(\kappa^d)$ facets for some $\kappa > 1$ by Lemma 17. Recall that Q_d^Δ is the convex hull of two polytopes with $d + 1$ facets each. By Lemma 5, every linear extension of P' with a number of facets polynomial in d has

$$\Omega\left(\frac{\log \kappa^d}{\log d^t}\right) = \Omega(d/\log d)$$

additional variables, as required. \square

6 A consequence for lift-and-project

Given $P \subseteq [0, 1]^d$, the lift-and-project method of Balas, Ceria and Cornuéjols [2] iteratively constructs polyhedral relaxations of $P \cap \mathbb{Z}^n$ that are the convex hull of the two faces of P defined by $x_j \geq 0$ and $x_j \leq 1$, for some $j = 1, \dots, d$.

We show that even in this restrictive setting, Theorem 2 is the best possible. More precisely, we prove the following:

Theorem 20. *Fix a polynomial σ . For each odd $d \in \mathbb{N}$, there exists a formulation of a nonempty polytope $P \subseteq [0, 1]^d$ of size f , such that any formulation of*

$$\text{conv}((P \cap \{x \in \mathbb{R}^d : x_d = 0\}) \cup (P \cap \{x \in \mathbb{R}^d : x_d = 1\}))$$

of size at most $\sigma(f)$ has $\Omega(d)$ additional variables.

Given polytope $Q \subseteq \mathbb{R}^d$ and $x^* \notin \text{aff}(Q)$, let C be the polyhedral cone generated by the vectors $\{x - x^* : x \in Q\}$. The polyhedron $\text{hom}(Q, x^*) := x^* + C$ is the *homogenization* of Q with respect to x^* . Note that F is a facet of Q if and only if $\text{hom}(F, x^*)$ is a facet of $\text{hom}(Q, x^*)$ and all facets of $\text{hom}(Q, x^*)$ arise in this way.

Remark 21. *Let $P_0, P_1 \in \mathbb{R}^{d-1}$ be $(d-1)$ -polytopes and pick x^0, x^1 in the interior of P_0, P_1 respectively. There exists $\varepsilon > 0$ such that*

$$H_0 := \text{hom}\left(\begin{pmatrix} P_0 \\ 0 \end{pmatrix}, \begin{pmatrix} x^0 \\ -\varepsilon \end{pmatrix}\right) \text{ contains } \begin{pmatrix} P_1 \\ 1 \end{pmatrix} \text{ in its interior, and}$$

$$H_1 := \text{hom}\left(\begin{pmatrix} P_1 \\ 1 \end{pmatrix}, \begin{pmatrix} x^1 \\ 1 + \varepsilon \end{pmatrix}\right) \text{ contains } \begin{pmatrix} P_0 \\ 0 \end{pmatrix} \text{ in its interior.}$$

In particular, if \bar{x} is a vertex of $H_0 \cap H_1$, then $\bar{x} = \begin{pmatrix} x^0 \\ -\varepsilon \end{pmatrix}$ or $\bar{x} = \begin{pmatrix} x^1 \\ 1 + \varepsilon \end{pmatrix}$ or $0 < \bar{x}_d < 1$.

Given $P_0 := D^{Cy}(d-1, (d-1)^2)$ and $P_1 := Q^{Cy}(d-1, (d-1)^2) \subseteq \mathbb{R}^{d-1}$, let $\varepsilon > 0$, H_0 and H_1 be as in Remark 21. By possibly scaling the first $d-1$ coordinates, we may assume that the polytope $P := H_0 \cap H_1 \cap \{x \in \mathbb{R}^d : 0 \leq x_d \leq 1\}$ is contained in $[0, 1]^d$. Note that $\begin{pmatrix} P_0 \\ 0 \end{pmatrix}$ is the facet of P defined by the inequality $x_d \geq 0$ and $\begin{pmatrix} P_1 \\ 1 \end{pmatrix}$ is the facet of P defined by the inequality $x_d \leq 1$.

Proof of Theorem 20. Let P be the polytope defined as above. By Remark 21, P has $2(d-1)^2 + 2$ facets. The proof of Theorem 15 shows that the polytope

$$\text{conv}((P \cap \{x \in \mathbb{R}^d : x_d = 0\}) \cup (P \cap \{x \in \mathbb{R}^d : x_d = 1\}))$$

has $d^{\Omega(d)}$ facets. By Lemma 5, any formulation of the above polytope has $\Omega(d)$ additional variables. \square

Acknowledgments We are grateful to Volker Kaibel and Stefan Weltge for helpful discussions and suggestions regarding the present work.

References

- [1] E. Balas. Disjunctive programming: properties of the convex hull of feasible points. *Discrete Applied Mathematics* 89, pp. 3–44, 1998.
- [2] E. Balas, S. Ceria, and G. Cornuéjols. A lift-and-project cutting plane algorithm for mixed 0-1 programs. *Mathematical Programming* 58, pp. 295–324, 1993.
- [3] A. Basu, K. Martin, C. Ryan, and G. Wang. Mixed-integer linear representability, disjunctions, and variable elimination. *Proceedings of 19th IPCO*, pp. 51–62, 2017.
- [4] S.O. Chan, J.R. Lee, P. Raghavendra, D. Steurer. Approximate constraint satisfaction requires large LP relaxations. *Journal of the ACM* 63, pp. 34:1–34:22, 2016.
- [5] M. Conforti, G. Cornuéjols, G. Zambelli. Extended formulations in combinatorial optimization. *4OR* 8, pp. 1–48, 2010.
- [6] M. Conforti, G. Cornuéjols, G. Zambelli. Integer Programming. *Springer*, 2014.
- [7] Y. Faenza and L. Sanità. On the existence of compact epsilon-approximation for the knapsack polytope in the original space. *Operations Research Letters* 43, pp. 339–342, 2015.
- [8] S. Fiorini, S. Massar, S. Pokutta, H.R. Tiwary, R. de Wolf. Exponential lower bounds for polytopes in combinatorial optimization. *Journal of the ACM* 62, pp. 17:1–17:23, 2015.
- [9] B. Grünbaum. Convex Polytopes. Springer, 1967.
- [10] B. Huber, J. Rambau, and F. Santos. The Cayley trick, lifting subdivisions, and the Bohne-Dress theorem on zonotopal tilings. *Journal of the European Mathematical Society* 2 pp. 179–198, 2000.
- [11] R.G. Jeroslow and J.K. Lowe. Modelling with integer variables. *Mathematical Programming Studies* 22, pp. 167–184, 1984.
- [12] S. Jukna. Extremal Combinatorics - With Applications in Computer Science. Springer, 2011.
- [13] V. Kaibel. Extended formulations in combinatorial optimization. *Optima* 85, pp. 2–7, 2011.
- [14] V. Kaibel and K. Pashkovich. Constructing extended formulations from reflection relations. *Proceedings of 15th IPCO*, pp. 287–300, 2011.
- [15] K. Adiprasito and R. Sanyal. Relative Stanley–Reisner theory and upper bound theorems for Minkowski sums. *Publications mathématiques de l’IHS* 124, pp. 99–163, 2016.
- [16] V. Kaibel and S. Weltge. Lower bounds on the sizes of integer programs without additional variables. *Mathematical Programming* 154, pp. 407–425, 2015.
- [17] T. Rothvoss. The matching polytope has exponential extension complexity. *Proceedings of 46th STOC*, pp. 263–272, 2014.
- [18] J.P. Vielma. Mixed integer linear programming formulation techniques. *SIAM Review* 57, pp. 3–57 2015.
- [19] C. Weibel. Minkowski Sums of Polytopes: Combinatorics and Computation. Ph.D. dissertation, École polytechnique fédérale de Lausanne, 2007.
- [20] S. Weltge. Sizes of Linear Descriptions in Combinatorial Optimization. Ph.D. dissertation, Otto-von-Guericke-Universität Magdeburg, 2016.
- [21] M. Van Vyve and L. Wolsey. Approximate extended formulations. *Mathematical Programming* 105, pp. 501–522, 2006.
- [22] G. Ziegler. Lectures on Polytopes. Springer, 1998.