APPROXIMATE CONE FACTORIZATIONS AND LIFTS OF POLYTOPES

JOÃO GOUVEIA, PABLO A. PARRILO, AND REKHA R. THOMAS

Abstract. In this paper we show how to construct inner and outer convex approximations of a polytope from an approximate cone factorization of its slack matrix. This provides a robust generalization of the famous result of Yannakakis that polyhedral lifts of a polytope are controlled by (exact) nonnegative factorizations of its slack matrix. Our approximations behave well under polarity and have efficient representations using second order cones. We establish a direct relationship between the quality of the factorization and the quality of the approximations, and our results extend to generalized slack matrices that arise from a polytope contained in a polyhedron.

1. Introduction

A well-known idea in optimization to represent a complicated convex set $C \subset \mathbb{R}^n$ is to describe it as the linear image of a simpler convex set in a higher dimensional space, called a *lift* or *extended formulation* of C . The standard way to express such a lift is as an affine slice of some closed convex cone K, called a K-*lift* of C, and the usual examples of K are nonnegative orthants $\mathbb{R}^m_+$ and the cones of real symmetric positive semidefinite matrices $\mathcal{S}^m_+$. More precisely, C has a K-lift, where $K \subset \mathbb{R}^m$, if there exists an affine subspace $L \subset \mathbb{R}^m$ and a linear map $\pi : \mathbb{R}^m \to \mathbb{R}^n$ such that $C = \pi(K \cap L)$.

Given a nonnegative matrix $M \in \mathbb{R}^{p\times q}_+$ and a closed convex cone $K \subset \mathbb{R}^m$ with dual cone $K^* \subset (\mathbb{R}^m)^*$, a K-factorization of M is a collection of elements $a_1, \dots, a_p \in K^*$ and $b_1, \dots, b_q \in K$ such that $M_{ij} = \langle a_i, b_j\rangle$ for all i, j. In particular, a $\mathbb{R}^m_+$-factorization of M, also called a *nonnegative factorization* of M of size m, is typically expressed as $M = A^T B$ where A has columns $a_1, \dots, a_p \in (\mathbb{R}^m)^*_+$ and B has columns $b_1, \dots, b_q \in \mathbb{R}^m_+$. In [16], Yannakakis laid the foundations of polyhedral lifts of polytopes by showing the following.

Theorem 1.1. [16] *A polytope $P \subset \mathbb{R}^n$ has a $\mathbb{R}^m_+$-lift if and only if the* slack matrix *of P has a $\mathbb{R}^m_+$-factorization.*

This theorem was extended in [7] from $\mathbb{R}^m_+$-lifts of polytopes to K-lifts of convex sets $C \subset \mathbb{R}^n$, where K is any closed convex cone, via K-*factorizations* of the *slack operator* of C.

The above results rely on exact cone factorizations of the slack matrix or operator of the given convex set, and do not offer any suggestions for constructing lifts of the set in the absence of exact factorizations. In many cases, one only has access to approximate factorizations of the slack matrix, typically via numerical algorithms. In this paper we show how to take an approximate K-factorization of the slack matrix of a polytope and construct from it an inner and outer convex approximation of the polytope. Our approximations behave well under polarity and admit efficient representations via second order cones. Further, we

Gouveia was supported by the Centre for Mathematics at the University of Coimbra and Fundação para a Ciência e a Tecnologia, through the European program COMPETE/FEDER. Parrilo was supported by AFOSR FA9550-11-1-0305, and Thomas by the U.S. National Science Foundation grant DMS-1115293.

show that the quality of our approximations can be bounded by the error in the corresponding approximate factorization of the slack matrix.

Let $P := \{x \in \mathbb{R}^n : H^T x \leq \mathbb{1}\}$ be a full-dimensional polytope in $\mathbb{R}^n$ with the origin in its interior, and vertices $p_1, \ldots, p_v$. We may assume without loss of generality that each inequality $\langle h_i, x\rangle \leq 1$ in $H^T x \leq \mathbb{1}$ defines a facet of P. If H has size $n \times f$, then the *slack matrix* of P is the $f \times v$ nonnegative matrix S whose (i,j)-entry is $S_{ij} = 1 - \langle h_i, p_j\rangle$, the *slack* of the jth vertex in the ith inequality of P. Given an $\mathbb{R}^m_+$-factorization of S, i.e., two nonnegative matrices A and B such that $S = A^T B$, an $\mathbb{R}^m_+$-lift of P is obtained as

$$P = \left\{x \in \mathbb{R}^n : \exists y \in \mathbb{R}^m_+ \text{ s.t. } H^T x + A^T y = \mathbb{1}\right\}.$$

Notice that this lift is highly non-robust, and small perturbations of A make the right hand side empty, since the linear system $H^T x + A^T y = \mathbb{1}$ is in general highly overdetermined. The same sensitivity holds for all K-factorizations and lifts. Hence, it becomes important to have a more robust, yet still efficient, way of expressing P (at least approximately) from approximate K-factorizations of S. Also, the quality of the approximations of P and their lifts must reflect the quality of the factorization, and specialize to the Yannakakis setting when the factorization is exact. The results in this paper carry out this program and contain several examples, special cases, and connections to the recent literature.

1.1. **Organization of the paper.** In Section 2 we establish how an approximate K-factorization of the slack matrix of a polytope $P \subset \mathbb{R}^n$ yields a pair of inner and outer convex approximations of P which we denote as $\mathrm{Inn}_P(A)$ and $\mathrm{Out}_P(B)$ where A and B are the two "factors" in the approximate K-factorization. These convex sets arise naturally from two simple inner and outer second order cone approximations of the nonnegative orthant. We prove that these approximations behave well under polarity, in the sense that

$$\mathrm{Out}_{P^\circ}(A) = (\mathrm{Inn}_P(A))^\circ \quad \text{and} \quad \mathrm{Inn}_{P^\circ}(B) = (\mathrm{Out}_P(B))^\circ$$

where P° is the polar polytope of P. Given $P \subset \mathbb{R}^n$ and $K \subset \mathbb{R}^m$, our approximations admit efficient representations via slices and projections of $K \times \mathrm{SOC}_{n+m+2}$ where SOC_k is a second order cone of dimension k. We show that an ε-error in the K-factorization makes $\frac{1}{1+\varepsilon}P \subseteq \mathrm{Inn}_P(A)$ and $\mathrm{Out}_P(B) \subseteq (1+\varepsilon)P$, thus establishing a simple link between the error in the factorization and the gap between P and its approximations. In the presence of an exact K-factorization of the slack matrix, our results specialize to the Yannakakis setting.

In Section 3 we discuss two connections between our approximations and well-known constructions in the literature. In the first part we show that our inner approximation, $\mathrm{Inn}_P(A)$, always contains the Dikin ellipsoid used in interior point methods. Next we examine the closest rank one approximation of the slack matrix obtained via a singular value decomposition and the approximations of the polytope produced by it.

In Section 4 we extend our results to the case of generalized slack matrices that arise from a polytope contained in a polyhedron. We also show how an approximation of P with a K-lift produces an approximate K-factorization of the slack matrix of P. It was shown in [4] that the max clique problem does not admit polyhedral approximations with small polyhedral lifts. We show that this negative result continues to hold even for the larger class of convex approximations considered in this paper.

2. From approximate factorizations to approximate lifts

In this section we show how to construct inner and outer approximations of a polytope P from approximate K-factorizations of the slack matrix of P, and establish the basic properties of these approximations.

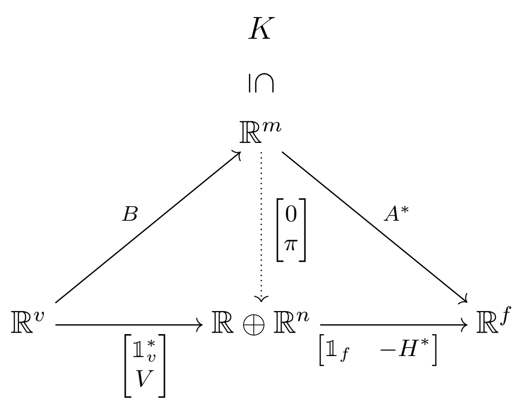
2.1. **K-factorizations and linear maps.** Let $P = \{x \in \mathbb{R}^n : H^T x \leq \mathbb{1}\}$ be a full-dimensional polytope with the origin in its interior. The vertices of the polytope are $p_1, \dots, p_v$, and each inequality $\langle h_i, x\rangle \leq 1$ for $i = 1, \dots, f$ in $H^T x \leq \mathbb{1}$ defines a facet of P. The *slack matrix* S of P is the $f \times v$ matrix with entries $S_{ij} = 1 - \langle h_i, p_j\rangle$. In matrix form, letting $H = [h_1 \dots h_f]$ and $V = [p_1 \dots p_v]$, we have the expression $S = \mathbb{1}_{f\times v} - H^T V$. We assume $K \subset \mathbb{R}^m$ is a closed convex cone, with dual cone $K^* = \{y \in (\mathbb{R}^m)^* : \langle y, x\rangle \geq 0 \quad \forall x \in K\}$.

Definition 2.1. ([7]) A *K-factorization* of the slack matrix S of the polytope P is given by $a_1, \dots, a_f \in K^*$, $b_1, \dots, b_v \in K$ such that $1 - \langle h_i, p_j\rangle = \langle a_i, b_j\rangle$ for $i = 1, \dots, f$ and $j = 1, \dots, v$. In matrix form, this is the factorization

$$S = \mathbb{1}_{f\times v} - H^T V = A^T B$$

where $A = [a_1 \dots a_f]$ and $B = [b_1 \dots b_v]$.

It is convenient to interpret a K-factorization as a composition of linear maps as follows. Consider B as a linear map from $\mathbb{R}^v \to \mathbb{R}^m$, verifying $B(\mathbb{R}^v_+) \subseteq K$. Similarly, think of A as a linear map from $(\mathbb{R}^f)^* \to (\mathbb{R}^m)^*$ verifying $A((\mathbb{R}^f)^*_+) \subseteq K^*$. Then, for the adjoint operators, $B^*(K^*) \subseteq (\mathbb{R}^v)^*_+$ and $A^*(K) \subseteq \mathbb{R}^f_+$. Furthermore, we can think of the slack matrix S as an affine map from $\mathbb{R}^v$ to $\mathbb{R}^f$, and the matrix factorization in Definition 2.1 suggests to define the slack operator, $S : \mathbb{R}^v \to \mathbb{R}^f$, as $S(x) = (\mathbb{1}_{f\times v} - H^* \circ V)(x)$, where $V : \mathbb{R}^v \to \mathbb{R}^n$ and $H : (\mathbb{R}^f)^* \to (\mathbb{R}^n)^*$.



We define a *nonnegative K-map* from $\mathbb{R}^v \to \mathbb{R}^m$ (where $K \subseteq \mathbb{R}^m$) to be any linear map F such that $F(\mathbb{R}^v_+) \subseteq K$. In other words, a nonnegative K-map from $\mathbb{R}^v \to \mathbb{R}^m$ is the linear map induced by an assignment of an element $b_i \in K$ to each unit vector $e_i \in \mathbb{R}^v_+$. In this language, a K-factorization of S corresponds to a nonnegative K^*-map $A : (\mathbb{R}^f)^* \to (\mathbb{R}^m)^*$ and a nonnegative K-map $B : \mathbb{R}^v \to \mathbb{R}^m$ such that $S(x) = (A^* \circ B)(x)$ for all $x \in \mathbb{R}^v$. As a consequence, we have the correspondence $a_i := A(e_i^*)$ for $i = 1, \dots, f$ and $b_j := B(e_j)$ for $j = 1, \dots, v$.

2.2. **Approximations of the nonnegative orthant.** In this section we introduce two canonical second order cone approximations to the nonnegative orthant, which will play a crucial role in our developments. In what follows, $\|\cdot\|$ will always denote the standard Euclidean norm in $\mathbb{R}^n$, i.e., $\|x\| = (\sum_{i=1}^n x_i^2)^{\frac{1}{2}}$.

Definition 2.2. Let $\mathcal{O}_{in}^n$, $\mathcal{O}_{out}^n$ be the cones

$$\mathcal{O}_{in}^n := \{x \in \mathbb{R}^n \,:\, \sqrt{n-1}\cdot \|x\| \leq \mathbb{1}^T x\}$$
$$\mathcal{O}_{out}^n := \{x \in \mathbb{R}^n \,:\, \|x\| \leq \mathbb{1}^T x\}.$$

If the dimension n is unimportant or obvious from the context, we may drop the superscript and just refer to them as $\mathcal{O}_{in}$ and $\mathcal{O}_{out}$.

As the following lemma shows, the cones $\mathcal{O}_{in}$ and $\mathcal{O}_{out}$ provide inner and outer approximations of the nonnegative orthant, which are dual to each other and can be described using second-order cone programming ([1, 9]).

Lemma 2.3. *The cones $\mathcal{O}_{in}$ and $\mathcal{O}_{out}$ are proper cones (i.e., convex, closed, pointed and solid) in $\mathbb{R}^n$ that satisfy*

$$\mathcal{O}_{in} \subseteq \mathbb{R}^n_+ \subseteq \mathcal{O}_{out},$$

and furthermore, $\mathcal{O}_{in}^* = \mathcal{O}_{out}$, and $\mathcal{O}_{out}^* = \mathcal{O}_{in}$.

The cones $\mathcal{O}_{in}$ and $\mathcal{O}_{out}$ are in fact the "best" second-order cone approximations of the nonnegative orthant, in the sense that they are the largest/smallest permutation-invariant cones with these containment properties; see also Remark 2.5. Lemma 2.3 is a direct consequence of the following more general result about (scaled) second-order cones:

Lemma 2.4. *Given $\omega \in \mathbb{R}^n$ with $\|\omega\| = 1$ and $0 < a < 1$, consider the set*

$$K_a := \{x \in \mathbb{R}^n \,:\, a\|x\| \leq \omega^T x\}.$$

Then, K_a is a proper cone, and $K_a^* = K_b$, where b satisfies $a^2 + b^2 = 1$, $b > 0$.

Proof: The set K_a is clearly invariant under nonnegative scalings, so it is a cone. Closedness and convexity of K_a follow directly from the fact that (for $a \geq 0$) the function $x \mapsto a\|x\| - \omega^T x$ is convex. The vector ω is an interior point (since $a\|\omega\| - \omega^T\omega = a - 1 < 0$), and thus K_a is solid. For pointedness, notice that if both x and $-x$ are in K_a, then adding the corresponding inequalities we obtain $2a\|x\| \leq 0$, and thus (since $a > 0$) it follows that $x = 0$.

The duality statement $K_a^* = K_b$ is perhaps geometrically obvious, since K_a and K_b are spherical cones with "center" ω and half-angles θ_a and θ_b, respectively, with $\cos\theta_a = a$, $\cos\theta_b = b$, and $\theta_a + \theta_b = \pi/2$. For completeness, however, a proof follows. We first prove that $K_b \subseteq K_a^*$. Consider $x \in K_a$ and $y \in K_b$, which we take to have unit norm without loss of generality. Let α, β, γ be the angles between (x, ω), (y, ω) and (x, y), respectively. The triangle inequality in spherical geometry gives $\gamma \leq \alpha + \beta$. Then,

$$\cos\gamma \geq \cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$$

or equivalently

$$x^T y \geq (\omega^T x)(\omega^T y) - \sqrt{1-(\omega^T x)^2}\sqrt{1-(\omega^T y)^2} \geq ab - \sqrt{1-a^2}\sqrt{1-b^2} = 0.$$

To prove the other direction ($K_a^* \subseteq K_b$) we use its contrapositive, and show that if $y \notin K_b$, then $y \notin K_a^*$. Concretely, given a y (of unit norm) such that $b\|y\| > \omega^T y$, we will construct an $x \in K_a$ such that $y^T x < 0$ (and thus, $y \notin K_a^*$). For this, define $x := a\omega - b\hat{\omega}$, where

$\hat{\omega} := (y-(\omega^T y)\omega)/\sqrt{1-(\omega^T y)^2}$ (notice that $\|\hat{\omega}\| = 1$ and $\omega^T\hat{\omega} = 0$). It can be easily verified that $\omega^T x = a$ and $\|x\|^2 = a^2 + b^2 = 1$, and thus $x \in K_a$. However, we have

$$y^T x = a(\omega^T y) - b\sqrt{1-(\omega^T y)^2} < ab - b\sqrt{1-b^2} = 0,$$

which proves that $y \notin K_a^*$. ☐

Proof: [of Lemma 2.3] Choosing $\omega = \mathbb{1}/\sqrt{n}$, $a = \sqrt{(n-1)/n}$ and $b = \sqrt{1/n}$ in Lemma 2.4, we have $\mathcal{O}_{in} = K_a$ and $\mathcal{O}_{out} = K_b$, so the duality statement follows. Since $x = \sum_i x_i e_i$, with $x_i \geq 0$, we have

$$\|x\| \leq \sum_i x_i \|e_i\| = \sum_i x_i = \mathbb{1}^T x,$$

and thus $\mathbb{R}^n_+ \subseteq \mathcal{O}_{out}$. Dualizing this expression, and using self-duality of the nonnegative orthant, we obtain the remaining containment $\mathcal{O}_{in} = \mathcal{O}_{out}^* \subseteq (\mathbb{R}^n_+)^* = \mathbb{R}^n_+$. ☐

Remark 2.5. Notice that $(\mathbb{1}^T x)^2 - \|x\|^2 = 2\sigma_2(x)$, where $\sigma_2(x)$ is the second elementary symmetric function in the variables $x_1, \ldots, x_n$. Thus, the containment relations in Lemma 2.3 also follow directly from the fact that the cone $\mathcal{O}_{out}$ is the $(n-2)$ derivative cone (or *Renegar derivative*) of the nonnegative orthant; see e.g. [12] for background and definitions and [13] for their semidefinite representability.

Remark 2.6. The following alternative description of $\mathcal{O}_{in}$ is often convenient:

$$\mathcal{O}_{in} = \{x \in \mathbb{R}^n \,:\, \exists t \in \mathbb{R} \quad \text{s.t. } \|t\mathbb{1} - x\| \leq t\}. \tag{1}$$

The equivalence is easy to see, since the condition above requires the existence of $t \geq 0$ such that $t^2(n-1) - 2t\mathbb{1}^T x + \|x\|^2 \leq 0$. Eliminating the variable t immediately yields $\sqrt{n-1} \cdot \|x\| \leq \mathbb{1}^T x$. The containment $\mathcal{O}_{in} \subseteq \mathbb{R}^n_+$ is now obvious from this representation, since $t - x_i \leq \|t\mathbb{1} - x\| \leq t$, and thus $x_i \geq 0$.

2.3. **From orthants to polytopes.** The cones $\mathcal{O}_{in}$ and $\mathcal{O}_{out}$ provide "simple" approximations to the nonnegative orthant. As we will see next, we can leverage these to produce inner/outer approximations of a polytope from any approximate factorizations of its slack matrix. The constructions below will use *arbitrary* nonnegative K^* and K-maps A and B (of suitable dimensions) to produce approximations of the polytope P (though of course, for these approximations to be useful, further conditions will be required).

Definition 2.7. Given a polytope P as before, a nonnegative K^*-map $A : (\mathbb{R}^f)^* \to (\mathbb{R}^m)^*$ and a nonnegative K-map $B : \mathbb{R}^v \to \mathbb{R}^m$, we define the following two sets:

$$\text{Inn}_P(A) := \left\{x \in \mathbb{R}^n \,:\, \exists y \in K, \text{ s.t. } \mathbb{1} - H^*(x) - A^*(y) \in \mathcal{O}_{in}^f\right\},$$
$$\text{Out}_P(B) := \left\{V(z) \,:\, z \in \mathcal{O}_{out}^v, \quad \mathbb{1}^T z \leq 1, \quad B(z) \in K\right\}.$$

By construction, these sets are convex, and the first observation is that the notation makes sense as the sets indeed define an inner and an outer approximation of P.

Proposition 2.8. *Let P be a polytope as before and A and B be nonnegative K^* and K-maps respectively. Then* $\text{Inn}_P(A) \subseteq P \subseteq \text{Out}_P(B)$.

Proof: Let $x \in \text{Inn}_P(A)$. Then, $\mathbb{1} - H^*(x) - A^*(y) \in \mathcal{O}_{in}^f \subseteq \mathbb{R}^f_+$, which implies $H^*(x) + A^*(y) \leq \mathbb{1}$. Since $A^*(y) \geq 0$ for $y \in K$, we have $H^*(x) \leq \mathbb{1}$ and thus $x \in P$.

For the second inclusion, by the convexity of $\text{Out}_P(B)$ it is enough to show that the vertices of P belong to this set. Any vertex p of P can be written as $p = V(e_i)$ for some canonical basis vector e_i. Furthermore, $B(e_i) \in K$ since B is a nonnegative K-map, $e_i \in \mathbb{R}^v_+ \subseteq \mathcal{O}^v_{out}$, $\mathbb{1}^T e_i \leq 1$, and so $p \in \text{Out}_P(B)$ as intended. ∎

If A and B came from a true K-factorization of S, then P has a K-lift and $S = A^* \circ B$. Then, the subset of $\text{Inn}_P(A)$ given by $\{x \in \mathbb{R}^n : \exists y \in K \text{ s.t. } H^*(x) + A^*(y) = \mathbb{1}\}$ contains P, since it contains every vertex $p_i = V(e_i)$ of P. This can be seen by taking $y = B(e_i) \in K$ and checking that $\mathbb{1} - H^*(p_i) = (\mathbb{1}_{f\times v} - H^* \circ V)(e_i) = (A^* \circ B)(e_i) = A^*(y)$. From Proposition 2.8 it then follows that $\text{Inn}_P(A) = P$. The definition of $\text{Out}_P(B)$ can be similarly motivated. An alternative derivation is through polarity, as we will see in Theorem 2.11.

Remark 2.9. Since the origin is in P, the set $\text{Out}_P(B)$ can also be defined with the inequality $\mathbb{1}^T z \leq 1$ replaced by the corresponding equation:

$$\text{Out}_P(B) = \left\{V(z) \,:\, z \in \mathcal{O}^v_{out}, \quad \mathbb{1}^T z = 1, \quad B(z) \in K\right\}.$$

To see this, suppose $q := V(z)$ such that $z \in \mathcal{O}^v_{out}$, $\mathbb{1}^T z \leq 1$ and $B(z) \in K$. Then there exists $s \geq 0$ such that $\mathbb{1}^T z + s = 1$. Since $0 \in P$, there exists $0 \leq \lambda_i$ with $\sum \lambda_i = 1$ such that $0 = \sum \lambda_i p_i$ where p_i are the vertices of P. Let $\tilde{z} := s\lambda + z \in \mathbb{R}^v$ where $\lambda = (\lambda_i)$. Then $\tilde{z} \in \mathcal{O}^v_{out}$ since $s \geq 0$, $\lambda \in \mathbb{R}^v_+ \subseteq \mathcal{O}^v_{out}$ and $z \in \mathcal{O}^v_{out}$. Further, $\mathbb{1}^T \tilde{z} = s\mathbb{1}^T\lambda + \mathbb{1}^T z = 1$. We also have that $B(\tilde{z}) = B(s\lambda + z) = sB(\lambda) + B(z) \in K$ since each component is in K (note that $B(\lambda) \in K$ since $\lambda \in \mathbb{R}^v_+$). Therefore, we can write $q = V(\tilde{z})$ with $\tilde{z} \in \mathcal{O}^v_{out}$, $\mathbb{1}^T \tilde{z} = 1$ and $B(\tilde{z}) \in K$ which proves our claim. This alternate formulation of $\text{Out}_P(B)$ will be useful in Section 4. However, Definition 2.7 is more natural for the polarity results in Section 2.4.

Example 2.10. Let P be the n-dimensional simplex given by the inequalities:

$$P = \{x \in \mathbb{R}^n \,:\, 1 + x_1 \geq 0, \ \ldots, \ 1 + x_n \geq 0, \ 1 - \textstyle\sum_i x_i \geq 0\}$$

with vertices $(n, -1, \cdots, -1), (-1, n, -1, \ldots, -1), \ldots, (-1, \ldots, -1, n), (-1, \ldots, -1)$. The slack matrix of this polytope is the $(n + 1) \times (n + 1)$ diagonal matrix with all diagonal entries equal to $n + 1$. Choosing A to be the zero map, for any cone K we have

$$\begin{aligned}\text{Inn}_P(0) &= \left\{x \in \mathbb{R}^n \,:\, n\left(\sum_i (1 + x_i)^2 + (1 - \sum_i x_i)^2\right) \leq (n + 1)^2\right\}\\ &= \left\{x \in \mathbb{R}^n \,:\, \sum_i x_i^2 + \left(\sum_i x_i\right)^2 \leq \frac{(n + 1)}{n}\right\}.\end{aligned}$$

For the case of $n = 2$ we have:

$$\text{Inn}_P(0) = \left\{(x_1, x_2) \,:\, x_1^2 + x_2^2 + (x_1 + x_2)^2 \leq \frac{3}{2}\right\} = \left\{(x_1, x_2) \,:\, 3(x_1 + x_2)^2 + (x_1 - x_2)^2 \leq 3\right\}.$$

For the outer approximation, if we choose $B = 0$ then we obtain the body

$$\text{Out}_P(0) = \left\{\left(-\sum_{i=1}^{n+1} z_i + (n + 1)z_j, \ \ j = 1, \ldots, n\right), \ \ \|z\| \leq \sum_{i=1}^{n+1} z_i \leq 1\right\}.$$

For $n = 2$, $\text{Out}_P(0) = \{(2z_1 - z_2 - z_3, -z_1 + 2z_2 - z_3) \,:\, \|z\| \leq z_1 + z_2 + z_3 \leq 1\}$, and eliminating variables we get

$$\text{Out}_P(0) = \left\{(x_1, x_2) \,:\, 3(x_1 + x_2)^2 + (x_1 - x_2)^2 \leq 12\right\}.$$

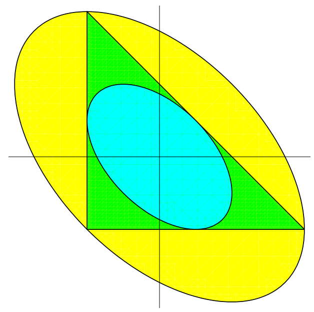


FIGURE 1. $\mathrm{Inn}_P(0)$ and $\mathrm{Out}_P(0)$ for a triangle centered at the origin.

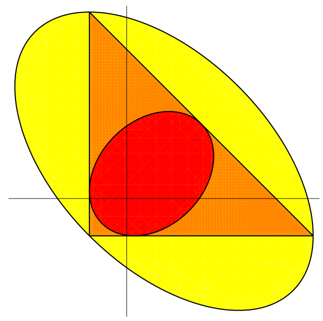


FIGURE 2. $\mathrm{Inn}_Q(0)$ and $\mathrm{Out}_Q(0)$ for a triangle not centered at the origin.

The simplex and its approximations can be seen in Figure 1.

Note that the bodies $\mathrm{Inn}_P(0)$ and $\mathrm{Out}_P(0)$ do not depend on the choice of a cone K and are hence canonical convex sets associated to the given representation of the polytope P. However, while $\mathrm{Out}_P(0)$ is invariant under translations of P (provided the origin remains in the interior), $\mathrm{Inn}_P(0)$ is sensitive to translation, i.e., to the position of the origin in the polytope P. To illustrate this, we translate the simplex in the above example by adding $(\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2})$ to it and denote the resulting simplex by Q. Then

$$Q = \left\{ x \in \mathbb{R}^n \,:\, 1 + 2x_1 \geq 0, \ldots, 1 + 2x_n \geq 0,\ 1 - \sum \frac{2}{n+2} x_i \geq 0 \right\}$$

and its vertices are $(n + \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \ldots, -\frac{1}{2}), \ldots, (-\frac{1}{2}, -\frac{1}{2}, \ldots, -\frac{1}{2}, n + \frac{1}{2}), (-\frac{1}{2}, -\frac{1}{2}, \ldots, -\frac{1}{2})$.

For $n = 2$, plugging into the formula for the inner approximation we get

$$\mathrm{Inn}_Q(0) = \{(x, y) : (3(x + y) - 2)^2 + 16(x - y)^2 \leq 16\},$$

while doing it for the outer approximation yields

$$\mathrm{Out}_Q(0) = \{(x, y) : 3(x + y - 1)^2 + (x - y)^2 \leq 12\}.$$

So we can see that while $\mathrm{Out}_Q(0)$ is simply a translation of the previous one, $\mathrm{Inn}_Q(0)$ has changed considerably as can be seen in Figure 2.

2.4. **Polarity.** We now show that our approximations behave nicely under polarity. Recall that the polar of the polytope $P \subset \mathbb{R}^n$ is the polytope $P^\circ = \{y \in (\mathbb{R}^n)^* \,:\, \langle y, x \rangle \leq 1,\ \forall\, x \in P\}$. The face lattices of P and P° are anti-isomorphic. In particular, if $P = \{x \in \mathbb{R}^n : H^T x \leq \mathbb{1}\}$ with vertices $p_1, \ldots, p_v$ as before, then the facet inequalities of P° are $\langle y, p_i \rangle \leq 1$ for all $i = 1, \ldots, v$ and the vertices are $h_1, \ldots, h_f$. Therefore, given a nonnegative K^*-map

$A : (\mathbb{R}^f)^* \to (\mathbb{R}^m)^*$ and a nonnegative K-map $B : \mathbb{R}^v \to \mathbb{R}^m$ we can use them, as above, to define approximations for P°, $\mathrm{Inn}_{P^\circ}(B)$ and $\mathrm{Out}_{P^\circ}(A)$, since facets and vertices of P and P° exchange roles. By applying polarity again, we potentially get new (different) outer and inner approximations of P via Proposition 2.8. We now prove that in fact, we recover the same approximations, and so in a sense, the approximations are dual to each other.

Theorem 2.11. *Let P be a polytope, A a nonnegative K^*-map and B a nonnegative K-map as before. Then,*

$$\mathrm{Out}_{P^\circ}(A) = (\mathrm{Inn}_P(A))^\circ,$$

and equivalently,

$$\mathrm{Inn}_{P^\circ}(B) = (\mathrm{Out}_P(B))^\circ.$$

Proof: Note that the two equalities are equivalent as one can be obtained from the other by polarity. Therefore, we will just prove the first statement. For notational convenience, we identify $\mathbb{R}^n$ and its dual. By definition,

$$(\mathrm{Inn}_P(A))^\circ = \{z \in \mathbb{R}^n \,:\, z^T x \leq 1 \;\forall\, x \in \mathrm{Inn}_P(A)\}.$$

Consider then the problem

$$\max\,\{z^T x \,:\, x \in \mathrm{Inn}_P(A)\} = \max\,\{z^T x \,:\, \mathbb{1} - H^*(x) - A^*(y) \in \mathcal{O}^f_{in},\; y \in K\}.$$

This equals the problem

$$\max\nolimits_{x, y\in K} \min\nolimits_{p \in (\mathcal{O}^f_{in})^*} \{z^T x + p^T(\mathbb{1} - H^*(x) - A^*(y))\}.$$

Strong duality [3] holds since the original problem $\max\,\{z^T x \,:\, x \in \mathrm{Inn}_P(A)\}$ is a convex optimization problem and $\mathrm{Inn}_P(A)$ has an interior as we will see in Proposition 2.16. This allows us to switch the order of min and max, to obtain

$$\min\nolimits_{p \in (\mathcal{O}^f_{in})^*} \max\nolimits_{x, y \in K} \{\mathbb{1}^T p + x^T(z - H(p)) - A(p)^T y\}.$$

For this max to be bounded above, we need $z = H(p)$ since x is unrestricted, and $A(p) \in K^*$ since $y \in K$. Therefore, using $(\mathcal{O}^f_{in})^* = \mathcal{O}^f_{out}$, we are left with

$$\min\nolimits_p \{\mathbb{1}^T p \,:\, p \in \mathcal{O}^f_{out},\; z = H(p),\; A(p) \in K^*\}.$$

Looking back at the definition of $(\mathrm{Inn}_P(A))^\circ$, we get

$$(\mathrm{Inn}_P(A))^\circ = \{H(p) \,:\, p \in \mathcal{O}^f_{out},\; \mathbb{1}^T p \leq 1,\; A(p) \in K^*\}$$

which is precisely $\mathrm{Out}_{P^\circ}(A)$. □

2.5. **Efficient representation.** There would be no point in defining approximations of P if they could not be described in a computationally efficient manner. Remarkably, the orthogonal invariance of the 2-norm constraints in the definition of $\mathrm{Inn}_P(\cdot)$ and $\mathrm{Out}_P(\cdot)$ will allow us to compactly represent the approximations via second order cones (SOC), with a problem size that depends only on the conic rank m (the affine dimension of K), and *not* on the number of vertices/facets of P. This feature is specific to our approximation and is of key importance, since the dimensions of the codomain of A^* and the domain of B can be exponentially large.

We refer to any set defined by a $k \times k$ positive semidefinite matrix $Q \succeq 0$ and a vector $a \in \mathbb{R}^k$ of the form

$$\{z \in \mathbb{R}^k \,:\, \|Q^{\frac{1}{2}} z\| \leq a^T z\}$$

as a k-second order cone or SOC_k. Here $Q^{\frac{1}{2}}$ refers to a matrix square root of $Q \succeq 0$. Suppose $M \in \mathbb{R}^{p\times n}$ is a matrix with $p \gg n$. Then $\|Mx\|^2 = x^T M^T M x = x^T Q x = \|Q^{\frac{1}{2}} x\|^2$ where $Q = M^T M$. Since Q has rank at most n, $Q^{\frac{1}{2}}$ is a $k \times n$ matrix where $k \leq n$ and $Q^{\frac{1}{2}}x$ is a vector of length k. Therefore, the expression $\|Mx\|$ (which corresponds to a 2-norm condition on $\mathbb{R}^p$) is equivalent to a 2-norm condition on $\mathbb{R}^k$ (and $k \ll p$). Notice that such a property is not true, for instance, for any ℓ_p norm for $p \neq 2$.

We use this key orthogonal invariance property of 2-norms to represent our approximations efficiently via second-order cones. Recall from the introduction that a convex set $C \subset \mathbb{R}^n$ has a K-lift, where $K \subset \mathbb{R}^m$ is a closed convex cone, if there is some affine subspace $L \subset \mathbb{R}^m$ and a linear map $\pi : \mathbb{R}^m \to \mathbb{R}^n$ such that $C = \pi(K \cap L)$. A set of the form $K' = \{z \in \mathbb{R}^k : \|Q^{\frac{1}{2}} z\|_2 \leq a_0 + a^T z\}$ where the right-hand side is affine can be gotten by slicing its homogenized second order cone

$$\left\{(z_0, z) \in \mathbb{R}^{k+1} \,:\, \left\| \begin{pmatrix} 0 & 0 \\ 0 & Q^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} z_0 \\ z \end{pmatrix} \right\| \leq a_0 z_0 + a^T z \right\}$$

with the affine hyperplane $\{(z_0, z) \in \mathbb{R}^{1+k} : z_0 = 1\}$ and projecting onto the z variables. In other words, K' has a SOC_{k+1}-lift.

Theorem 2.12. *Let $P \subseteq \mathbb{R}^n$ be a polytope, $K \subseteq \mathbb{R}^m$ a closed convex cone and A and B nonnegative K^* and K-maps respectively. Then the convex sets $\mathrm{Inn}_P(A)$ and $\mathrm{Out}_P(B)$ have $K \times \mathrm{SOC}_{(n+m+2)}$-lifts.*

Proof: Set M to be the $f \times (n+m+1)$ concatenated matrix $[H^* \ A^* \ -\mathbb{1}]$. Then, using Definition 2.7, the characterization in Remark 2.6 and $Q := M^T M$, we have

$$\begin{aligned} \mathrm{Inn}_P(A) &= \left\{ x \in \mathbb{R}^n \,:\, \exists y \in K, \xi \in \mathbb{R} \text{ s.t. } \left\| M \left(x^T \ y^T \ \xi \right)^T \right\| \leq 1 - \xi \right\} \\ &= \left\{ x \in \mathbb{R}^n \,:\, \exists y \in K, \xi \in \mathbb{R} \text{ s.t. } \|Q^{\frac{1}{2}} \left(x^T \ y^T \ \xi \right)^T \| \leq 1 - \xi \right\}. \end{aligned}$$

From the discussion above, it follows that $\mathrm{Inn}_P(A)$ is the projection onto the x variables of points $(z; x_0, x, y, \xi) \in K \times \mathrm{SOC}_{n+m+2}$ such that $z = y$ and $x_0 = 1$ and so $\mathrm{Inn}_P(A)$ has a $K \times \mathrm{SOC}_{n+m+2}$-lift.

The statement for $\mathrm{Out}_P(B)$ follows from Theorem 2.11 and the fact that if a convex set $C \subset \mathbb{R}^n$ with the origin in its interior has a K-lift, then its polar has a K^*-lift [7]. We also use the fact that the dual of a SOC_k is another SOC_k cone. □

Note that the lifting dimension is *independent* of f, the number of facets of P, which could be very large compared to n. A slightly modified argument can be used to improve this result by 1, and show that a $K \times \mathrm{SOC}_{(n+m+1)}$-lift always exists.

2.6. **Approximation quality.** The final question we will address in this section is how good an approximation $\mathrm{Inn}_P(A)$ and $\mathrm{Out}_P(B)$ are of the polytope P. One would like to prove that if we start with a good approximate K-factorization of the slack matrix S of P, we would get good approximate lifts of P from our definitions. This is indeed the case as we will show. Recall that our approximations $\mathrm{Inn}_P(A)$ and $\mathrm{Out}_P(B)$ each depend on only one of the factors A or B. For this reason, ideally, the goodness of these approximations should be quantified using only the relevant factor. Our next result presents a bound in this spirit.

Proposition 2.13. *For $x \in \mathbb{R}^f$, let $\mu(x) := \min\{t : x + t\mathbb{1} \in \mathcal{O}^f_{in} + A^*(K)\}$ where $A^*(K) \subset \mathbb{R}^f$ is the image of K under A^*. Then $\frac{P}{1+\varepsilon} \subseteq \mathrm{Inn}_P(A)$ for $\varepsilon = \max_i(\mu(S(e_i)))$.*

Proof: It suffices to show that $\frac{p}{1+\varepsilon} \in \text{Inn}_P(A)$ where $p = V(e_i)$ is a vertex of P. By the definition of ε, we have that $\varepsilon\mathbb{1} + S(e_i) \in \mathcal{O}_{in} + A^*(K)$ and hence, $\varepsilon\mathbb{1} + \mathbb{1} - H^*V(e_i) \in \mathcal{O}_{in} + A^*(K)$. Therefore, $\mathbb{1} - H^*(\frac{p}{1+\varepsilon}) \in \mathcal{O}_{in} + A^*(K)$ and hence, $\frac{p}{1+\varepsilon} \in \text{Inn}_P(A)$. □

From this proposition one sees that the approximation factor improves as $A^*(K)$ gets bigger. While geometrically appealing, this bound is not very convenient from a computational viewpoint. Therefore, we now write down a simple, but potentially weaker, result based on an alternative error measure $\xi(\cdot)$ defined as follows.

Lemma 2.14. *For $0 \neq x \in \mathbb{R}^f$, $f \geq 2$, let*

$$\xi(x) := \|x\| \cdot \alpha\left(\frac{\mathbb{1}^T x}{\sqrt{f}\|x\|}\right),$$

where $\alpha(s) := (\sqrt{(f-1)(1-s^2)} - s)/\sqrt{f}$. Then $\xi(x) = \min\{t : x + t\mathbb{1} \in \mathcal{O}_{in}^f\}$ and $\xi(x) \leq \|x\|$.

Proof: Notice that $\xi(x)$ is well-defined since $-1 \leq \frac{\mathbb{1}^T x}{\sqrt{f}\|x\|} \leq 1$ by Cauchy-Schwarz. For $t = \xi(x)$, one can verify that $\sqrt{f-1}\|x + t\mathbb{1}\| = \mathbb{1}^T(x + t\mathbb{1})$ (and it is the largest such value), and thus $x + t\mathbb{1}$ is on the boundary of $\mathcal{O}_{in}$. Since $\mathbb{1} \in \mathcal{O}_{in}^f$, increasing the value of t will keep $x + t\mathbb{1}$ in $\mathcal{O}_{in}^f$. For the second claim, an easy calculation shows that $\alpha(s) \leq 1$ for all $s \in [-1, 1]$ (since it is concave, and achieves this maximum at $s = -\frac{1}{\sqrt{f}}$), so the result follows. □

Using $\xi(\cdot)$ we can provide a sufficient condition for Proposition 2.13 based on the factorization error and not just on the factor A.

Lemma 2.15. *Let $A : (\mathbb{R}^f)^* \to (\mathbb{R}^m)^*$ be a nonnegative K^*-map and $B : \mathbb{R}^v \to \mathbb{R}^m$ be a nonnegative K-map. Let $\Delta = S - A^* \circ B$. Then, $\frac{1}{1+\varepsilon}P \subseteq \text{Inn}_P(A)$, for $\varepsilon = \max_i \xi(\Delta(e_i))$.*

Proof: Note that

$$\mu(S(e_i)) \leq \min\{t : \underbrace{S(e_i) - A^*(B(e_i))}_{\Delta(e_i)} + t\mathbb{1} \in \mathcal{O}_{in}\} \leq \xi(\Delta(e_i)),$$

where the first inequality follows since $B(e_i) \in K$ and therefore, $A^*(B(e_i)) \in A^*(K)$, and the second from the definition of $\xi(\cdot)$. □

We immediately get our main result establishing the connection between the quality of the factorization and the quality of the approximations. For simplicity, we state it using the simplified upper bound $\xi(x) \leq \|x\|$.

Proposition 2.16. *Let $A : (\mathbb{R}^f)^* \to (\mathbb{R}^m)^*$ be a nonnegative K^*-map and $B : \mathbb{R}^v \to \mathbb{R}^m$ be a nonnegative K-map. Let $\Delta := S - A^* \circ B$ be the factorization error. Then,*

(1) $\frac{1}{1+\varepsilon}P \subseteq \text{Inn}_P(A)$, *for* $\varepsilon = \|\Delta\|_{1,2}$*;*
(2) $\text{Out}_P(B) \subseteq (1 + \varepsilon)P$, *for* $\varepsilon = \|\Delta\|_{\infty,2}$*,*

where $\|\Delta\|_{1,2} = \max_i \|\Delta(e_i)\|$ is the induced ℓ_1, ℓ_2 norm and $\|\Delta\|_{\infty,2} = \max_i \|\Delta^*(e_i^*)\|$ is the induced ℓ_∞, ℓ_2 norm of the factorization error.

Proof: By Theorem 2.11, the two statements are equivalent. The proof now follows from Lemma 2.15. □

This means that if we start with A and B forming a K-factorization of a nonnegative S' which is close to the true slack matrix S, we get a $(1 + \|S - S'\|_{1,2})$-approximate inner approximation of P, as well as a $(1 + \|S - S'\|_{\infty,2})$-approximate outer approximation of P. Thus, good approximate factorizations of S do indeed lead to good approximate lifts of P.

Example 2.17. Consider the square given by the inequalities $1 \pm x \geq 0$ and $1 \pm y \geq 0$ with vertices $(\pm 1, \pm 1)$. By suitably ordering facets and vertices, this square has slack matrix

$$S = \begin{pmatrix} 2 & 2 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 2 & 2 \\ 2 & 0 & 0 & 2 \end{pmatrix}.$$

Let $A : (\mathbb{R}^4)^* \to (\mathbb{R}^2)^*$ and $B : \mathbb{R}^4 \to \mathbb{R}^2$ be the nonnegative maps given by the matrices

$$A = \begin{pmatrix} 4/3 & 4/3 & 0 & 0 \\ 0 & 0 & 4/3 & 4/3 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix}.$$

Then A and B are nonnegative $\mathbb{R}^2_+ = (\mathbb{R}^2_+)^*$ maps and

$$A^T B = \begin{pmatrix} 4/3 & 4/3 & 4/3 & 0 \\ 4/3 & 4/3 & 4/3 & 0 \\ 4/3 & 0 & 4/3 & 4/3 \\ 4/3 & 0 & 4/3 & 4/3 \end{pmatrix}.$$

It is easy to check that $\|S - A^T B\|_{1,2} = \frac{2}{3}\sqrt{10}$ while $\|S - A^T B\|_{\infty,2} = 2\sqrt{\frac{2}{3}}$. So by Proposition 2.16 this implies that

$$\frac{1}{1 + \frac{2}{3}\sqrt{10}} P \subseteq \mathrm{Inn}_P(A) \subseteq P \subseteq \mathrm{Out}_P(B) \subseteq (1 + 2\sqrt{\tfrac{2}{3}})P.$$

If we use instead Lemma 2.15, we can get the slightly better quality bounds

$$\frac{1}{\frac{4}{3} + \sqrt{3}} P \subseteq \mathrm{Inn}_P(A) \subseteq P \subseteq \mathrm{Out}_P(B) \subseteq (1 + \sqrt{2})P.$$

Finally, if we use directly Proposition 2.13, it is possible in this case to compute explicitly the true bounds

$$\frac{1}{\sqrt{3}} P \subseteq \mathrm{Inn}_P(A) \subseteq P \subseteq \mathrm{Out}_P(B) \subseteq \sqrt{3}P.$$

In Figure 3 we can see the relative quality of all these bounds.

3. Special Cases

In this section we relate our inner approximation to *Dikin's ellipsoid*, a well known inner approximation to a polytope that arises in the context of interior point methods. We also compute $\mathrm{Inn}_P(A)$ and $\mathrm{Out}_P(B)$ from a nonnegative factorization of the closest rank one approximation of the slack matrix of P. As we will see, these are the two simplest possible approximations of a polytope, and correspond to specific choices for the factors A and B.

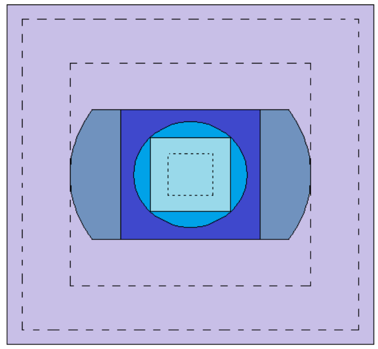


FIGURE 3. $\mathrm{Inn}_P(A)$, P and $\mathrm{Out}_P(B)$, as well as all the guaranteed approximations.

3.1. **Dikin's ellipsoid.** Recall that $\mathrm{Inn}_P(0)$ is an ellipsoidal inner approximation of P that is intrinsic to P and does not depend on any approximate factorization of the slack matrix of P through any cone. A commonly used inner approximation of a polytope is Dikin's ellipsoid defined as follows. Let $P = \{x \in \mathbb{R}^n : H^T x \leq \mathbb{1}\}$ be a polytope as before and let $h_1, \ldots, h_f$ be the columns of H. Define

$$\Gamma_P(x) := \sum_{i=1}^{f} \frac{h_i h_i^T}{(1 - \langle h_i, x \rangle)^2},$$

which is the Hessian of the standard logarithmic barrier function $\psi(x) := -\sum_{i=1}^{f} \log(1 - \langle h_i, x\rangle)$ associated to P. If x_0 is in the interior of P, then the Dikin ellipsoid centered at x_0 and of radius r, is the set

$$D_{x_0}^r := \{x \in \mathbb{R}^n : (x - x_0)^T \Gamma_P(x_0)(x - x_0) \leq r^2\}.$$

It can be checked that $D_{x_0}^1 \subseteq P$ (see Theorem 2.1.1 in [10]). Further, since Dikin's ellipsoid is invariant under translations, we may assume that $x_0 = 0$, and then

$$D_0^1 = \left\{ x \in \mathbb{R}^n : x^T \left(\sum_{i=1}^{f} h_i h_i^T \right) x \leq 1 \right\} = \left\{ x \in \mathbb{R}^n : \|H^T x\| \leq 1 \right\}.$$

Recall that

$$\mathrm{Inn}_P(0) = \left\{ x \in \mathbb{R}^n : \mathbb{1} - H^T x \in \mathcal{O}_{in}^f \right\} = \left\{ x \in \mathbb{R}^n : \exists t \in \mathbb{R} \text{ s.t. } \|(t-1)\mathbb{1} + H^T x\| \leq t \right\},$$

where we used the characterization of $\mathcal{O}_{in}^f$ given in Remark 2.6. Choosing $t = 1$, we see that $D_0^1 \subseteq \mathrm{Inn}_P(0) \subseteq P$. This inclusion is implicit in the work of Sturm and Zhang [15].

If the origin is also the analytic center of P (i.e., the minimizer of $\psi(x)$), then we have that $P \subseteq \sqrt{f(f-1)} D_0^1$ where f is the number of inequalities in the description of P; see [14] and [3, Section 8.5.3]. Also, in this situation, the first order optimality condition on $\psi(x)$ gives that $\sum_{i=1}^{f} h_i = 0$, and as a consequence, $f = \sum_{i=1}^{f}(1 - \langle h_i, p\rangle)$ for all vertices p of P. In other words, every column of the slack matrix sums to f. This implies an analogous (slightly stronger) containment result for $\mathrm{Inn}_P(0)$.

Corollary 3.1. *If the origin is the analytic center of the polytope P, then*

$$\mathrm{Inn}_P(0) \subseteq P \subseteq (f-1)\, \mathrm{Inn}_P(0),$$

furthermore, if P is centrally symmetric

$$\mathrm{Inn}_P(0) \subseteq P \subseteq \sqrt{f-1}\,\mathrm{Inn}_P(0).$$

Proof: This follows from Lemma 2.15 by using the fact that if S is the slack matrix of P then $\|S(e_i)\| \leq \mathbb{1}^T S(e_i) = f$. In this case, for $w = S(e_i)$, we have $\frac{\mathbb{1}^T w}{\sqrt{f}\|w\|} \geq \frac{1}{\sqrt{f}}$ and thus, since $\alpha(s)$ is decreasing for $s \geq -\frac{1}{\sqrt{f}}$, we get $\xi(w) = \|w\|\alpha(\frac{\mathbb{1}^T w}{\sqrt{f}\|w\|}) \leq \|w\|\alpha(\frac{1}{\sqrt{f}}) \leq f \cdot \frac{f-2}{f} = f-2$, from where the first result follows.

For the second result, note that if P is centrally symmetric, for every facet inequality $1 - \langle h_j, x\rangle \geq 0$ we have also the facet inequality $1 + \langle h_j, x\rangle \geq 0$, which implies

$$\|S(e_i)\| = \sqrt{\frac{1}{2}\sum_{j=1}^{f} [(1-\langle h_j, p_i\rangle)^2 + (1+\langle h_j, p_i\rangle)^2]} = \sqrt{f + \|H^T p_i\|^2}.$$

Using this fact together with $\mathbb{1}^T S(e_i) = f$, we get that $\xi(S(e_i)) = \sqrt{\frac{f-1}{f}}\|H^T p_i\| - 1$. To conclude just notice that since $\mathbb{1} - H^T p_i$ and $\mathbb{1} + H^T p_i$ are both valid, all entries in $H^T p_i$ are smaller than one in absolute value, hence $\|H^T p_i\| \leq \sqrt{f}$ concluding the proof. □

3.2. **Singular value decomposition.** Let P be a polytope as before and suppose $S = U\Sigma V^T$ is a singular value decomposition of the slack matrix S of P with U and V orthogonal matrices and Σ a diagonal matrix with the singular values of S on the diagonal. By the Perron-Frobenius theorem, the leading singular vectors of a nonnegative matrix can be chosen to be nonnegative. Therefore, if σ is the largest singular value of S, and u is the first column of U and v is the first column of V then $A = \sqrt{\sigma}u^T$ and $B = \sqrt{\sigma}v^T$ are two nonnegative matrices. By the Eckart-Young theorem, the matrix $A^T B = \sigma u v^T$ is the closest rank one matrix (in Frobenius norm) to S. We can also view $A^T B$ as an approximate $\mathbb{R}_+$-factorization of S and thus look at the approximations $\mathrm{Inn}_P(A) =: \mathrm{Inn}_P^{\mathrm{sing}}$ and $\mathrm{Out}_P(B) =: \mathrm{Out}_P^{\mathrm{sing}}$ and hope that in some cases they may offer good compact approximations to P. Naturally, these are not as practical as $\mathrm{Inn}_P(0)$ and $\mathrm{Out}_P(0)$ since to compute them we need to have access to a complete slack matrix of P and its leading singular vectors. We illustrate these approximations on an example.

Example 3.2. Consider the quadrilateral with vertices $(1,0), (0,2), (-1,0)$ and $(0,-1/2)$. This polygon has slack matrix

$$S = \begin{bmatrix} 0 & 5 & 2 & 0 \\ 0 & 0 & 2 & 5/4 \\ 2 & 0 & 0 & 5/4 \\ 2 & 5 & 0 & 0 \end{bmatrix},$$

and by computing a singular value decomposition and proceeding as outlined above, we obtain the 1×4 matrices

$$A = [1.9130, 0.1621, 0.1621, 1.9130] \text{ and } B = [0.5630, 2.5951, 0.5630, 0.0550],$$

verifying

$$S' = A^T B = \begin{bmatrix} 1.0770 & 4.9644 & 1.0770 & 0.1051 \\ 0.0912 & 0.4206 & 0.0912 & 0.0089 \\ 0.0912 & 0.4206 & 0.0912 & 0.0089 \\ 1.0770 & 4.9644 & 1.0770 & 0.1051 \end{bmatrix}.$$

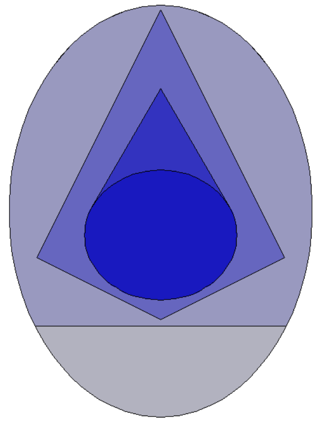


FIGURE 4. $\mathrm{Inn}_P(0) \subseteq \mathrm{Inn}_P^{\mathrm{sing}} \subseteq P \subseteq \mathrm{Out}_P^{\mathrm{sing}} \subseteq \mathrm{Out}_P(0)$.

Computing $\mathrm{Inn}_P^{\mathrm{sing}}$ and $\mathrm{Out}_P^{\mathrm{sing}}$ we get the inclusions illustrated in Figure 4. Notice how these rank one approximations from leading singular vectors use the extra information to improve on the trivial approximations $\mathrm{Inn}_P(0)$ and $\mathrm{Out}_P(0)$. Notice also that since

$$\mathrm{Out}_P^{\mathrm{sing}} = \mathrm{Out}_P(B) = \{V(z) \,:\, z \in \mathbb{R}^4,\, \|z\| \leq \mathbb{1}^T z \leq 1,\, Bz \geq 0\},$$

and $Bz \geq 0$ is a single linear inequality, $\mathrm{Out}_P^{\mathrm{sing}}$ is obtained from $\mathrm{Out}_P(0)$ by imposing a new linear inequality. By polarity, $\mathrm{Inn}_P^{\mathrm{sing}}$ is the convex hull of $\mathrm{Inn}_P(0)$ with a new point.

4. Nested polyhedra and Generalized Slack Matrices

In this section, we consider approximate factorizations of generalized slack matrices and what they imply in terms of approximations to a pair of nested polyhedra. The results here can be seen as generalizations of results in Section 2.

Let $P \subset \mathbb{R}^n$ be a polytope with vertices $p_1, \ldots, p_v$ and the origin in its interior, and $Q \subset \mathbb{R}^n$ be a polyhedron with facet inequalities $\langle h_j, x\rangle \leq 1$, for $j = 1, \ldots, f$, such that $P \subseteq Q$. The *slack matrix* of the pair P, Q is the $f \times v$ matrix $S_{P,Q}$ whose (i,j)-entry is $1 - \langle h_i, p_j\rangle$. In the language of operators from Section 2, $S_{P,Q}$ can be thought of as an operator from $\mathbb{R}^v \to \mathbb{R}^f$ defined as $S_{P,Q}(x) = (\mathbb{1}_{f\times v} - H_Q^* \circ V_P)(x)$ where V_P is the vertex operator of P and H_Q is the facet operator of Q. Every nonnegative matrix can be interpreted as such a generalized slack matrix after a suitable rescaling of its rows and columns, possibly with some extra rows and columns representing redundant points in P and redundant inequalities for Q.

4.1. Generalized slack matrices and lifts.

Yannakakis' theorem about $\mathbb{R}_+^m$-lifts of polytopes can be extended to show that $S_{P,Q}$ has a K-factorization (where $K \subset \mathbb{R}^m$ is a closed convex cone), if and only if there exists a convex set C with a K-lift such that $P \subseteq C \subseteq Q$. (For a proof in the case of $K = \mathbb{R}_+^m$, see [4, Theorem 1], and for $K = \mathcal{S}_+^m$ see [8]. Related formulations in the polyhedral situation also appear in [6, 11].)

Two natural convex sets with K-lifts that are nested between P and Q can be obtained as follows. For a nonnegative K^*-map $A : (\mathbb{R}^f)^* \to (\mathbb{R}^m)^*$ and a nonnegative K-map

$B : \mathbb{R}^v \to \mathbb{R}^m$, define the sets:

$$C_A := \left\{x \in \mathbb{R}^n \,:\, \exists\, y \in K \text{ s.t. } \mathbb{1} - H_Q^*(x) - A^*(y) = 0\right\},$$
$$C_B := \{V_P(z) \,:\, \mathbb{1}^T z = 1,\ B(z) \in K\}.$$

The sets C_A and C_B have K-lifts since they are obtained by imposing affine conditions on the cone K. From the definitions and Remark 2.9, it immediately follows that these containments hold:

$$P \subseteq \mathrm{Out}_P(B) \subseteq C_B \quad \text{and} \quad C_A \subseteq \mathrm{Inn}_Q(A) \subseteq Q.$$

As discussed in the introduction, the set C_A could potentially be empty for an arbitrary choice of A. On the other hand, since $P \subseteq C_B$, C_B is never empty. Thus in general, it is not true that C_B is contained in C_A. However, in the presence of an exact factorization we get the following chain of containments.

Proposition 4.1. *When $S_{P,Q} = A^* \circ B$, we get $P \subseteq \mathrm{Out}_P(B) \subseteq C_B \subseteq C_A \subseteq \mathrm{Inn}_Q(A) \subseteq Q$.*

Proof: We only need to show that $C_B \subseteq C_A$. Let $V_P(z) \in C_B$, with $\mathbb{1}_v^T z = 1$ and $B(z) \in K$. Then choosing $y = B(z)$, we have that $\mathbb{1}_f - H_Q^*(V_P(z)) - A^*(B(z)) = \mathbb{1}_f \mathbb{1}_v{}^T z - H_Q^*(V_P(z)) - A^*(B(z)) = S_{P,Q}(z) - A^*(B(z)) = 0$ since $S_{P,Q} = A^* \circ B$. Therefore, $C_B \subseteq C_A$ proving the result. □

Example 4.2. To illustrate these inclusions, consider the quadrilateral P with vertices $(1/2, 0), (0, 1), (-1/2, 0), (0, -1)$, and the quadrilateral Q defined by the inequalities $1 - x \geq 0$, $1 - x/2 - y/2 \geq 0$, $1 + x \geq 0$ and $1 + x/2 + y/2 \geq 0$. We have $P \subseteq Q$, and

$$S_{P,Q} = \frac{1}{4}\begin{pmatrix} 2 & 4 & 6 & 4 \\ 3 & 2 & 5 & 6 \\ 6 & 4 & 2 & 4 \\ 5 & 6 & 3 & 2 \end{pmatrix}.$$

For $K = \mathbb{R}^3_+$ we can find an exact factorization for this matrix, such as the one given by

$$S_{P,Q} = A^T B = \begin{pmatrix} 0 & 2 & 1 \\ 0 & 1 & \frac{3}{2} \\ 2 & 0 & 1 \\ 2 & 1 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 1 \end{pmatrix}.$$

In this example,

$$C_A = C_B = \{(x_1, x_2) \,:\, 1 - x_2 \geq 0, 1 + 2x_1 + x_2 \geq 0, 1 - 2x_1 + x_2 \geq 0\}.$$

By computing $\mathrm{Inn}_Q(A)$ and $\mathrm{Out}_P(B)$ we can see as in Figure 5 that

$$P \subsetneq \mathrm{Out}_P(B) \subsetneq C_B = C_A \subsetneq \mathrm{Inn}_Q(A) \subsetneq Q.$$

Note that if instead we pick the factorization

$$S_{P,Q} = A^T B = \begin{pmatrix} 0 & 2 & 1 & 1 \\ 0 & 1 & \frac{3}{2} & 0 \\ 2 & 0 & 1 & 1 \\ 2 & 1 & \frac{1}{2} & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

then $C_A \subsetneq C_B$ and we get the inclusions in Figure 6.

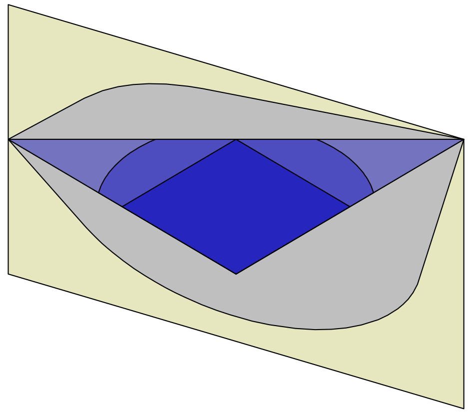


FIGURE 5. $P \subsetneq \mathrm{Out}_P(B) \subsetneq C_B = C_A \subsetneq \mathrm{Inn}_Q(A) \subsetneq Q$.

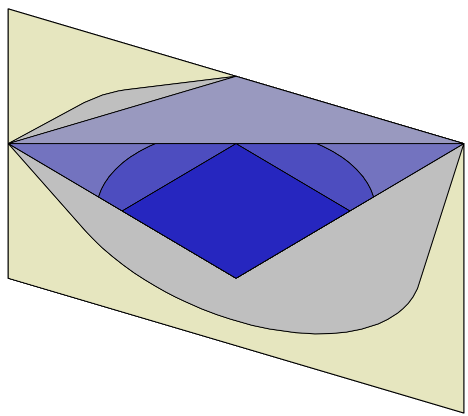


FIGURE 6. $P \subsetneq \mathrm{Out}_P(B) \subsetneq C_B \subsetneq C_A \subsetneq \mathrm{Inn}_Q(A) \subsetneq Q$.

4.2. **Approximate Factorizations of $S_{P,Q}$.** We now consider approximate factorizations of the generalized slack matrix $S_{P,Q}$ and provide three results about this situation. First, we generalize Proposition 2.16 to show that the quality of the approximations to P and Q is directly linked to the factorization error. Next we show how approximations to a pair of nested polytopes yield approximate K-factorizations of $S_{P,Q}$. We close with an observation on the recent inapproximability results in [4] in the context of this paper.

When we only have an approximate factorization of $S_{P,Q}$, we obtain a strict generalization of Proposition 2.16, via the same proof.

Proposition 4.3. *Let $A : (\mathbb{R}^f)^* \to (\mathbb{R}^m)^*$ a nonnegative K^*-map and $B : \mathbb{R}^v \to \mathbb{R}^m$ be a nonnegative K-map. Let $\Delta := S_{P,Q} - A^* \circ B$ be the factorization error. Then,*

(1) $\frac{1}{1+\varepsilon} P \subseteq \mathrm{Inn}_Q(A) \subseteq Q$, *for* $\varepsilon = \|\Delta\|_{1,2}$.
(2) $P \subseteq \mathrm{Out}_P(B) \subseteq (1+\varepsilon)Q$, *for* $\varepsilon = \|\Delta\|_{\infty,2}$.

So far we discussed how an approximate factorization of the slack matrix can be used to yield approximations of a polytope or a pair of nested polyhedra. It is also possible to go in the opposite direction in the sense that approximations to a pair of nested polyhedra $P \subseteq Q$ yield approximate factorizations of the slack matrix $S_{P,Q}$ as we now show.

Proposition 4.4. *Let $P \subseteq Q$ be a pair of polyhedra as before and $S_{P,Q}$ be its slack matrix. Suppose there exists a convex set C with a K-lift such that $\alpha P \subseteq C \subseteq \beta Q$ for some $0 < \alpha \leq 1$ and $\beta \geq 1$. Then there exists a K-factorizable nonnegative matrix S' such that $|(S' - S_{P,Q})_{ij}| \leq \frac{\beta}{\alpha} - 1$.*

Proof: Assume that the facet inequalities of Q are of the form $\langle h_j, x\rangle \leq 1$. Since C has a K-lift, it follows from [7] that we can assign an element $b_x \in K$ to each point x in C, and an element $a_y \in K^*$ to every valid inequality $y_0 - \langle y, x\rangle \geq 0$ for C, such that $\langle a_y, b_x\rangle = y_0 - \langle y, x\rangle$.

If $\alpha = \beta = 1$, then $P \subseteq C \subseteq Q$, and since C has a K-lift, $S_{P,Q}$ has a K-factorization and $S' = S_{P,Q}$ gives the result. So we may assume that $\frac{\beta}{\alpha} > 1$ and define $\eta > 0$ such that $(1+\eta)\alpha = \beta$. If $1 - \langle h, x\rangle \geq 0$ defines a facet of Q, then $1 + \eta - \frac{1}{\alpha}\langle h, x\rangle \geq 0$ is a valid inequality for $(1+\eta)\alpha Q = \beta Q$, and is therefore a valid inequality for C. Hence, we can pick $a_1, \ldots, a_f$ in K^*, one for each such inequality as mentioned above. Similarly, if v is a vertex of P, then αv belongs to C and we can pick $b_1, \ldots, b_v$ in K, one for each αv. Then,

$$\langle a_i, b_j\rangle = 1 + \eta - \frac{1}{\alpha}\langle h_i, \alpha v_j\rangle = \eta + (S_{P,Q})_{ij}.$$

The matrix S' defined by $S'_{ij} := \langle a_i, b_j\rangle$ yields the result. □

Note that Proposition 4.4 is not a true converse of Proposition 4.3. Proposition 4.3 says that approximate K-factorizations of the slack matrix give approximate $K \times$ SOC-lifts of the polytope or pair of polytopes, while Proposition 4.4 says that an approximation of a polytope with a K-lift gives an approximate K-factorization of its slack matrix. We have not ruled out the existence of a polytope with no good K-liftable approximations but whose slack matrix has a good approximate K-factorization.

A recent result on the inapproximability of a polytope by polytopes with small lifts, come from the max clique problem, as seen in [4] (and strengthened in [5]). In [4], the authors prove that for $P(n) = \mathrm{COR}(n) = \{bb^T \mid b \in \{0,1\}^n\}$ and

$$Q(n) = \{x \in \mathbb{R}^{n\times n} \mid \left\langle 2\mathrm{diag}(a) - aa^T, x\right\rangle \leq 1, a \in \{0,1\}^n\},$$

and for any $\rho > 1$, if $P(n) \subseteq C(n) \subseteq \rho Q(n)$ then any $\mathbb{R}^m_+$-lift of $C(n)$ has $m = 2^{\Omega(n)}$. Therefore, one cannot approximate $P(n)$ within a factor of ρ by any polytope with a small linear lift. This in turn says that the max clique problem on a graph with n vertices cannot be approximated well by polytopes with small polyhedral lifts. In fact they also prove that even if $\rho = O(n^\beta)$, for some $\beta \leq 1/2$, the size of m grows exponentially.

The above result was proven in [4] by showing that the *nonnegative rank* of the slack matrix $S_{P(n),\rho Q(n)}$, denoted as $\mathrm{rank}_+(S_{P(n),\rho Q(n)})$, has order $2^{\Omega(n)}$. The matrix $S_{P(n),\rho Q(n)}$ is a very particular perturbation of $S_{P(n),Q(n)}$. It is not hard to see that the proof of Theorem 5 in [4] in fact shows that all nonnegative matrices in a small neighborhood of $S_{P(n),Q(n)}$ have high nonnegative rank.

Proposition 4.5. *Let $P(n)$ and $Q(n)$ be as above, and $0 < \eta < 1/2$. For any nonnegative matrix $S'(n)$ such that $\|S_{P(n),Q(n)} - S'(n)\|_\infty \leq \eta$, $\mathrm{rank}_+(S'(n)) = 2^{\Omega(n)}$.*

Since $\|S_{P(n),Q(n)} - S_{P(n),\rho Q(n)}\|_\infty = \rho - 1$, Proposition 4.5 implies a version of the first part of the result in [4], i.e., that the nonnegative rank of $S_{P(n),\rho Q(n)}$ is exponential in n (if only for $1 < \rho < 3/2$). Further, Proposition 4.5 also says that even if we allowed approximations of $P(n)$ in the sense of Proposition 4.3 (i,e., approximations with a $\mathbb{R}^m_+ \times$ SOC-lift that do not truly nest between $P(n)$ and $Q(n)$), m would not be small. Therefore, the result in [4] on the outer inapproximability of $P(n)$ by small polyhedra is robust in a very strong sense.

Acknowledgements: We thank Anirudha Majumdar for the reference to [15].

References

[1] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Mathematical programming*, 95(1):3–51, 2003.
[2] G. Blekherman, P. A. Parrilo, and R. R. Thomas, editors. *Semidefinite optimization and convex algebraic geometry*, volume 13 of *MOS-SIAM Series on Optimization*. SIAM, 2012.
[3] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
[4] G. Braun, S. Fiorini, S. Pokutta, and D. Steurer. Approximation limits of linear programs (beyond hierarchies). In *Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on*, pages 480–489. IEEE, 2012.
[5] M. Braverman and A. Moitra. An information complexity approach to extended formulations. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 19, 2012.
[6] N. Gillis and F. Glineur. On the geometric interpretation of the nonnegative rank. *Linear Algebra and its Applications*, 437:2685–2712, 2012.
[7] J. Gouveia, P. A. Parrilo, and R. R. Thomas. Lifts of convex sets and cone factorizations. *Mathematics of Operations Research*, 38(2):248–264, 2013.
[8] J. Gouveia, R. Z. Robinson, and R. R. Thomas. Worst-case results for positive semidefinite rank. arXiv:1305.4600, 2013.
[9] M. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra and its Applications*, 284:193–228, 1998.
[10] Y. E. Nesterov and A. Nemirovski. *Interior point polynomial methods in convex programming*, volume 13 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, 1994.
[11] K. Pashkovich. *Extended Formulations for Combinatorial Polytopes*. PhD thesis, Magdeburg Universität, 2012.
[12] J. Renegar. Hyperbolic programs, and their derivative relaxations. *Found. Comput. Math.*, 6(1):59–79, 2006.
[13] J. Saunderson and P. A. Parrilo. Polynomial-sized semidefinite representations of derivative relaxations of spectrahedral cones. arXiv:1208.1443, 2012.
[14] Gy. Sonnevend. An "analytical centre" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. In *System modelling and optimization*, pages 866–875. Springer, 1986.
[15] J. F. Sturm and S. Zhang. An $O(\sqrt{nL})$ iteration bound primal-dual cone affine scaling algorithm for linear programming. *Mathematical Programming*, 72(2):177–194, 1996.
[16] M. Yannakakis. Expressing combinatorial optimization problems by linear programs. *J. Comput. System Sci.*, 43(3):441–466, 1991.

CMUC, Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal
E-mail address: jgouveia@mat.uc.pt

Department of Electrical Engineering and Computer Science, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
E-mail address: parrilo@mit.edu

Department of Mathematics, University of Washington, Box 354350, Seattle, WA 98195, USA
E-mail address: rrthomas@uw.edu