Facial Reduction and Partial Polyhedrality

Bruno F. Lourenço [*] Masakazu Muramatsu[†] Takashi Tsuchiya [‡]

November 2015 (Revised: October 2016, April 2017, April 2018)

Abstract

We present FRA-Poly, a facial reduction algorithm (FRA) for conic linear programs that is sensitive to the presence of polyhedral faces in the cone. The main goals of FRA and FRA-Poly are the same, i.e., finding the minimal face containing the feasible region and detecting infeasibility, but FRA-Poly treats polyhedral constraints separately. This reduces the number of iterations drastically when there are many linear inequality constraints. The worst case number of iterations for FRA-Poly is written in the terms of a "distance to polyhedrality" quantity and provides better bounds than FRA under mild conditions. In particular, in the case of the doubly nonnegative cone, FRA-Poly gives a worst case bound of n whereas the classical FRA is $\mathcal{O}(n^2)$. Of possible independent interest, we prove a variant of Gordan-Stiemke's Theorem and a proper separation theorem that takes into account partial polyhedrality. We provide a discussion on the optimal facial reduction strategy and an instance that forces FRAs to perform many steps. We also present a few applications. In particular, we will use FRA-Poly to improve the bounds recently obtained by Liu and Pataki on the dimension of certain affine subspaces which appear in weakly infeasible problems.

1 Introduction

Consider the following pair of primal and dual conic linear programs (CLPs):

$$\begin{array}{lll} \inf\limits_{x} & \langle c, x\rangle & \\ \text{subject to} & \mathcal{A}x = b & \\ & x \in \mathcal{K}^* & \end{array} \qquad \text{(P)} \qquad\qquad \begin{array}{ll} \sup\limits_{y} & \langle b, y\rangle \\ \text{subject to} & c - \mathcal{A}^*y \in \mathcal{K}, \end{array} \qquad \text{(D)}$$

where $\mathcal{K} \subseteq \mathbb{R}^n$ is a closed convex cone and $\mathcal{K}^*$ is the dual cone $\{s \in \mathbb{R}^n \mid \langle s, x\rangle \geq 0, \forall x \in \mathcal{K}\}$. We have that $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ and $\mathcal{A}^*$ denotes the adjoint map. We also have $\mathcal{A}^*y = \sum_{i=1}^m \mathcal{A}_i y_i$, for certain elements $\mathcal{A}_i \in \mathbb{R}^n$. The inner product is denoted by $\langle \cdot, \cdot \rangle$. We will use θ_P and θ_D to denote the primal and dual optimal value, respectively. It is understood that $\theta_P = +\infty$ if (P) is infeasible and $\theta_D = -\infty$ if (D) is infeasible.

In the absence of either a primal relative interior feasible solution or a dual relative interior slack, it is possible that $\theta_P \neq \theta_D$. A possible way of correcting that is to let $\mathcal{F}_{\min}^D$ be the minimal face of $\mathcal{K}$ which contains the feasible slacks $\mathcal{F}_{\mathrm{D}}^s = \{c - \mathcal{A}^*y \in \mathcal{K} \mid y \in \mathbb{R}^m\}$, then we substitute $\mathcal{K}$ by $\mathcal{F}_{\min}^D$ and $\mathcal{K}^*$ by $(\mathcal{F}_{\min}^D)^*$.

[*]Department of Mathematical Informatics, Graduate School of Information Science & Technology, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan. (Email: lourenco@mist.i.u-tokyo.ac.jp)

[†]Department of Computer and Network Engineering, The University of Electro-Communications 1-5-1 Chofugaoka, Chofu-shi, Tokyo, 182-8585 Japan. (E-mail: MasakazuMuramatsu@uec.ac.jp)

[‡]National Graduate Institute for Policy Studies 7-22-1 Roppongi, Minato-ku, Tokyo 106-8677, Japan. (E-mail: tsuchiya@grips.ac.jp)

M. Muramatsu and T. Tsuchiya are supported in part with Grant-in-Aid for Scientific Research (B)24310112 and (C) 26330025. M. Muramatsu is also partially supported by the Grant-in-Aid for Scientific Research (B)26280005. T. Tsuchiya is also partially supported by the Grant-in-Aid for Scientific Research (B)15H02968.

	FRA	FRA-Poly
$\mathcal{K}$	$\ell_{\mathcal{K}}$	$1 + \ell_{\text{poly}}(\mathcal{K})$
$\mathcal{K}^1 \times \ldots \times \mathcal{K}^r$	$1 + \sum_{i=1}^r (\ell_{\mathcal{K}^i} - 1)$	$1 + \sum_{i=1}^r \ell_{\text{poly}}(\mathcal{K}^i)$
$\mathcal{Q}^{t_1} \times \ldots \times \mathcal{Q}^{t_{r_1}} \times \mathcal{S}_+^{n_1} \times \ldots \times \mathcal{S}_+^{n_{r_2}}$	$1 + 2r_1 + \sum_{j=1}^{r_2} n_j$	$1 + r_1 + \sum_{j=1}^{r_2} (n_j - 1)$
$\mathcal{D}^n$	$1 + \frac{n(n+1)}{2}$	n

Table 1: Summary of the worst case number of reduction steps predicted by the classical FRA analysis and by FRA-Poly.

With that, the new primal optimal value $\theta_P{}'$ will satisfy $\theta_P{}' = \theta_D$. This is precisely what facial reduction [4, 22, 29] approaches do.

In this paper, we analyze how to take advantage of the presence of polyhedral faces in $\mathcal{K}$ when doing Facial Reduction. To do that, we introduce FRA-Poly, which is a facial reduction algorithm (FRA) that, in many cases, provides a better worst case complexity than the usual approach, especially when $\mathcal{K}$ is a direct product of several cones. The idea behind it is as follows. Facial reduction algorithms work by successively identifying what is called "reducing directions" $\{d_1, \ldots, d_\ell\}$. Starting with $\mathcal{F}_1 = \mathcal{K}$, these directions define faces of $\mathcal{K}$ by the relation $\mathcal{F}_{i+1} = \mathcal{F}_i \cap \{d_i\}^\perp$. For feasible problems, d_i must be such that $\mathcal{F}_{i+1}$ is a face of $\mathcal{K}$ containing $\mathcal{F}^s_{\text{D}}$. We then obtain a sequence $\mathcal{F}_1 \supseteq \ldots \supseteq \mathcal{F}_\ell$ of faces of $\mathcal{K}$ such that $\mathcal{F}_i \supseteq \mathcal{F}^s_{\text{D}}$ for every i. Usually, a FRA proceeds until $\mathcal{F}^D_{\min}$ is found.

A key observation is that as soon as we reach a polyhedral face $\mathcal{F}_i$, we can jump to the minimal face $\mathcal{F}^D_{\min}$ in a single facial reduction step. In addition, when $\mathcal{K}$ is a direct product $\mathcal{K} = \mathcal{K}^1 \times \ldots \times \mathcal{K}^r$, each $\mathcal{F}_i$ is also a direct product $\mathcal{F}^1_i \times \ldots \times \mathcal{F}^r_i$. In this case an even weaker condition is sufficient to jump to $\mathcal{F}^D_{\min}$, namely, if every block $\mathcal{F}^j_i$ is polyhedral or it is already equal to j-th block of the minimal face.

Our proposed algorithm FRA-Poly works in two phases. In Phase 1, it proceeds until a face $\mathcal{F}_i$ satisfying the condition above is reached or until a certificate of infeasibility is found. In Phase 2, $\mathcal{F}^D_{\min}$ is obtained with single facial reduction step. One interesting point is that even if $\mathcal{F}_i \neq \mathcal{F}^D_{\min}$, if we substitute $\mathcal{K}$ for $\mathcal{F}_i$ in (D), then strong duality will hold. The theoretical backing for that is given by Proposition 2, which is a generalization of the classical strong duality theorem. In Section 4, we will give a generalization of the Gordan-Stiemke Theorem for the case when $\mathcal{K}$ is the direct product of a closed convex cone and a polyhedral cone, see Theorem 5. We also prove a proper separation theorem that will be the engine behind FRA-Poly, see Theorem 4.

In order to analyze the number of facial reduction steps, we introduce a quantity called *distance to polyhedrality* $\ell_{\text{poly}}(\mathcal{K})$. This is the length *minus one* of the longest strictly ascending chain of nonempty faces $\mathcal{F}_1 \subsetneq \ldots \subsetneq \mathcal{F}_\ell$ for which $\mathcal{F}_1$ is polyhedral and $\mathcal{F}_i$ is not polyhedral for all $i > 1$. If $\mathcal{K}$ is a direct product of cones $\mathcal{K}^1 \times \ldots \times \mathcal{K}^r$, we prove that FRA-Poly stops in at most $1 + \sum_{i=1}^r \ell_{\text{poly}}(\mathcal{K}^i)$ steps. This is no worse than the bound given by classical FRA and, provided that at least two of the cones $\mathcal{K}^i$ are not subspaces, it is strictly smaller. We also discuss whether our bounds are achieved by some problem instance, see Section 5.4 and Proposition 24 (Appendix B).

As an application, we give a nontrivial bound for the singularity degree of CLPs over cones that are intersections of two other cones. In particular, for the case of the doubly nonnegative cone $\mathcal{D}^n$, we show that the longest chain of nonempty faces of $\mathcal{D}^n$ has length $1 + \frac{n(n+1)}{2}$. Therefore, the classical analysis gives the upper bound $\frac{n(n+1)}{2}$ for the singularity degree of feasible problems over $\mathcal{D}^n$. On the other hand, using our technique, we show that the singularity degree of any problem over $\mathcal{D}^n$ is at most n. We also use FRA-Poly to improve bounds obtained by Liu and Pataki in Corollary 1 of [12] on the dimension of certain subspaces connected to weakly infeasible problems.

Table 1 contains a summary of the bounds predicted by FRA and FRA-Poly for several cases. The notation $\ell_{\mathcal{K}}$ indicates the length of the longest strictly ascending chain of nonempty faces of $\mathcal{K}$. The first line corresponds to a single cone, the second to a product of r arbitrary closed convex cones and the third to the product of r_1 Lorentz cones and r_2 positive semidefinite cones, respectively. These results follow from Theorem 10 and Example 1. The last line contains the bounds for the doubly nonnegative cone, which follows from Proposition 21 and Corollary 20.

This work is divided as follows. In Section 2 we give some background on related notions. In Section 3, we review facial reduction. In Section 4 we prove versions of two classical theorems taking into account partial polyhedrality. In Section 5 we analyze FRA-Poly and in Section 6 we discuss two applications. Appendix A contains the proof of a strong duality criterion. Appendix B illustrates FRA-Poly and contains an example which generalizes an earlier worst case SDP instance by Tunçel.

2 Notation and Preliminaries

In this section, we will define the notation used throughout this article and review a few concepts. More details can be found in [25, 21]. For C a closed convex set, we will denote by $\operatorname{ri} C$ and $\operatorname{cl} C$ the relative interior and the closure of C, respectively. If $\mathcal{U}$ is an arbitrary set, we denote by $\mathcal{U}^\perp$ the subspace which contains the elements orthogonal to it. We will denote by $\mathcal{S}^n_+$ the cone of $n \times n$ symmetric positive semidefinite matrices and by $\mathcal{Q}^n$ the Lorentz cone $\{(x_0, \overline{x}) \in \mathbb{R} \times \mathbb{R}^{n-1} \mid x_0 \geq \|\overline{x}\|_2\}$, where $\|.\|_2$ is the usual Euclidean norm. The nonnegative orthant will be denoted by $\mathbb{R}^n_+$.

If $\mathcal{K}$ is a closed convex cone, we write $\mathcal{K}^*$ for its dual cone. We write $\operatorname{span} \mathcal{K}$ for its linear span and $\operatorname{lin} \mathcal{K}$ for its lineality space, which is $\mathcal{K} \cap -\mathcal{K}$. $\mathcal{K}$ is said to be *pointed* if $\operatorname{lin} \mathcal{K} = \{0\}$. We have $\operatorname{lin}(\mathcal{K}^*) = \mathcal{K}^\perp$, see Theorem 14.6 in [25]. Also, if we select $e \in \operatorname{ri} \mathcal{K}$ and $x \in \mathcal{K}^*$, then $x \in \mathcal{K}^\perp$ if and only if $\langle e, x \rangle = 0$[1].

The conic linear program (D) can be in four different feasibility statuses: i) strongly feasible if $\mathcal{F}^s_{\mathrm{D}} \cap \operatorname{ri} \mathcal{K} \neq \emptyset$; ii) weakly feasible if $\mathcal{F}^s_{\mathrm{D}} \neq \emptyset$ but $\mathcal{F}^s_{\mathrm{D}} \cap \operatorname{ri} \mathcal{K} = \emptyset$; iii) weakly infeasible if $\mathcal{F}^s_{\mathrm{D}} = \emptyset$ but $\operatorname{dist}(c + \operatorname{range} \mathcal{A}^*, \mathcal{K}) = 0$; iv) strongly infeasible if $\operatorname{dist}(c + \operatorname{range} \mathcal{A}^*, \mathcal{K}) > 0$. Note that (P) admits analogous definitions.

The strong duality theorem states that if (D) is strongly feasible and $\theta_D < +\infty$, then $\theta_P = \theta_D$ and θ_P is attained. On the other hand, if (P) is strongly feasible and $\theta_P > -\infty$, then $\theta_P = \theta_D$ and θ_D is attained. We will also need a special version of the strong duality theorem. First, we need the following definition.

Definition 1 (Partial Polyhedral Slater's condition)**.** *Let $\mathcal{K} = \mathcal{K}^1 \times \mathcal{K}^2$, where $\mathcal{K}^1 \subseteq \mathbb{R}^{n_1}, \mathcal{K}^2 \subseteq \mathbb{R}^{n_2}$ are closed convex cones such that $\mathcal{K}^2$ is polyhedral. We say that* (D) *satisfies the Partial Polyhedral Slater's (PPS) condition if there is a slack $(s_1, s_2) = c - \mathcal{A}^* y$, such that $s_1 \in \operatorname{ri} \mathcal{K}^1$ and $s_2 \in \mathcal{K}^2$. Similarly, we say that* (P) *satisfies the PPS condition, if there is a primal feasible solution $x = (x_1, x_2)$ for which $x_1 \in \operatorname{ri} (\mathcal{K}^1)^*$.*

The following is a strong duality theorem based on the PPS condition. As we could not find a precise reference for it, we give a proof in the Appendix A.

Proposition 2 (PPS-Strong Duality)**.** *Let $\mathcal{K} = \mathcal{K}^1 \times \mathcal{K}^2$, where $\mathcal{K}^1 \subseteq \mathbb{R}^{n_1}, \mathcal{K}^2 \subseteq \mathbb{R}^{n_2}$ are closed convex cones such that $\mathcal{K}^2$ is polyhedral.*

(i) If θ_P is finite and (P) *satisfies the PPS condition, then $\theta_P = \theta_D$ and the dual optimal value is attained.*

(ii) If θ_D is finite and (D) *satisfies the PPS condition, then $\theta_P = \theta_D$ and the primal optimal value is attained.*

3 Facial Reduction

Facial Reduction was developed by Borwein and Wolkowicz to restore strong duality in convex optimization [3, 4]. Descriptions for the conic linear programming case have appeared, for instance, in Pataki [22] and in Waki and Muramatsu [29].

Here, we will suppose that our main interest is in the dual problem (D). Facial Reduction hinges on the fact that strong feasibility fails if and only if there is $d \in \mathcal{K}^*$ such that $\mathcal{A}d = 0$ and one of the two alternatives holds: (i) $\langle c, d \rangle = 0$ and $d \notin \mathcal{K}^\perp$; or (ii) $\langle d, c \rangle < 0$, see Lemma 3.2 in [29]. If alternative (i) holds, $\mathcal{F} = \mathcal{K} \cap \{d\}^\perp$ is a proper face of $\mathcal{K}$ containing $\mathcal{F}^s_{\mathrm{D}}$. We then substitute $\mathcal{K}$ by $\mathcal{F}$ and repeat. If (ii) holds, (D) is infeasible. We write below a generic facial reduction algorithm similar to the one described in [29].

[1]Suppose $\langle e, x \rangle = 0$. Then, given $y \in \mathcal{K}$ we have $\alpha e + (1 - \alpha) y \in \mathcal{K}$ for some $\alpha > 1$, due to Theorem 6.4 in [25]. Taking the inner product with x, we see that $\langle y, x \rangle$ must be zero.

[**Generic Facial Reduction**]

Input: (D)

Output: A set of reducing directions $\{d_1, \ldots, d_\ell\}$ and $\mathcal{F}^D_{\min}$.

1. $\mathcal{F}_1 \leftarrow \mathcal{K}, i \leftarrow 1$
2. Let d_i be an element in $\mathcal{F}_i^* \cap \ker \mathcal{A}$ such that either: *i)* $d_i \notin \mathcal{F}_i^\perp$ and $\langle c, d_i \rangle = 0$; or *ii)* $\langle c, d_i \rangle < 0$. If no such d_i exists, let $\mathcal{F}^D_{\min} \leftarrow \mathcal{F}_i$ and stop.
3. If $\langle c, d_i \rangle < 0$, let $\mathcal{F}^D_{\min} \leftarrow \emptyset$ and stop.
4. If $\langle c, d_i \rangle = 0$, let $\mathcal{F}_{i+1} \leftarrow \mathcal{F}_i \cap \{d_i\}^\perp, i \leftarrow i + 1$ and return to 2).

We will refer to the directions satisfying $d_i \in \mathcal{F}_i^* \cap \ker \mathcal{A}$ and $\langle c, d_i \rangle \leq 0$ as *reducing directions*, so that the d_i in Step 2. are indeed reducing directions. An important issue when doing facial reduction is how to model the search for the reducing directions. It is sometimes said that doing facial reduction can be as hard as solving the original problem. However, an important difference is that the search for the d_i can be cast as a pair of primal and dual problems which are always strongly feasible. This was shown in the work by Cheung, Schurr and Wolkowicz [5] and in our previous work [13]. Recently, Permenter, Friberg and Andersen showed that d_i can also be obtained as by-products of self-dual homogeneous methods [23]. There are also approximate approaches such as the one described by Permenter and Parrilo [24], where the search for the d_i is conducted in a more tractable cone at the cost of, perhaps, failing to identify $\mathcal{F}^D_{\min}$, but still simplifying the problem nonetheless. See also the article by Frieberg [9], where conic constraints are dropped when searching for the reducing directions, making it easier to find the d_i but introducing representational issues.

In this article, we will search for reducing directions by considering the pair ($P_\mathcal{K}$) and ($D_\mathcal{K}$) introduced in [13], which are parametrized by $\mathcal{A}$, c, $\mathcal{K}, e, e^*$. In Phase 1 of FRA-Poly, we will always select e and e^* according to Lemma 3. Different choices will be discussed/used in Phase 2 of FRA-Poly and on Sections 5.2 and 5.4.

$$
\begin{array}{lllr}
\underset{x,t,w}{\text{minimize}} & t & & (P_\mathcal{K}) \\
\text{subject to} & -\langle c, x - te^* \rangle + t - w & = 0 & (1) \\
& \langle e, x \rangle + w & = 1 & (2) \\
& \mathcal{A}x - t\mathcal{A}e^* & = 0 & (3) \\
& (x, t, w) \in \mathcal{K}^* \times \mathbb{R}_+ \times \mathbb{R}_+ & &
\end{array}
$$

$$
\begin{array}{llr}
\underset{y_1,y_2,y_3}{\text{maximize}} & y_2 & (D_\mathcal{K}) \\
\text{subject to} & cy_1 - ey_2 - \mathcal{A}^*y_3 \in \mathcal{K} & (4) \\
& 1 - y_1(1 + \langle c, e^* \rangle) + \langle e^*, \mathcal{A}^*y_3 \rangle \geq 0 & (5) \\
& y_1 - y_2 \geq 0 & (6) \\
& (y_1, y_2, y_3) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^m. &
\end{array}
$$

Recall that our goal is to find a point $x \in \ker \mathcal{A} \cap \mathcal{K}$ satisfying $\langle c, x \rangle \leq 0$. The idea behind ($P_\mathcal{K}$) is to shift the problem by $-te^*$ (Equations (1) and (3)) and add constraints to ensure the x stays in a bounded region (Equation (2)). These changes ensure that ($P_\mathcal{K}$) and ($D_\mathcal{K}$) satisfy the PPS condition when the parameters e, e^* are chosen appropriately as in the next section.

4 Partial Polyhedrality Theorems

We are now in position to present a choice of e, e^* for ($P_\mathcal{K}$) and ($D_\mathcal{K}$) taking into account the PPS condition.

Lemma 3. *Let $\mathcal{K} = \mathcal{K}^1 \times \mathcal{K}^2$ be a closed convex cone, such that $\mathcal{K}^2$ is polyhedral. Consider the pair* ($P_\mathcal{K}$) *and* ($D_\mathcal{K}$) *with e and e^* such that $e = (e_1, 0) \in (\operatorname{ri} \mathcal{K}^1) \times \{0\}$ and $e^* \in \operatorname{ri} \mathcal{K}^*$. The following properties hold.*

(i) *The following are solutions to* ($P_{\mathcal{K}}$) *and* ($D_{\mathcal{K}}$) *that satisfy the PPS condition:*

$$(x,t,w) = \left(\frac{e^*}{\langle e, e^*\rangle + 1}, \frac{1}{\langle e, e^*\rangle + 1}, \frac{1}{\langle e, e^*\rangle + 1}\right)$$
$$(y_1, y_2, y_3) = (0, -1, 0).$$

In particular, $\theta_{P_{\mathcal{K}}} = \theta_{D_{\mathcal{K}}}$.

Let (x^*, t^*, w^*) *be any primal optimal solution and* (y_1^*, y_2^*, y_3^*) *be any dual optimal solution.*

(ii) $\theta_{P_{\mathcal{K}}} = \theta_{D_{\mathcal{K}}} = 0$ *if and only if one of the two alternatives below holds:*

(a) $\langle c, x^*\rangle < 0$ *and* $\mathcal{F}^s_D = \mathcal{F}^D_{\min} = \emptyset$, *or*

(b) $\langle c, x^*\rangle = 0$, $\mathcal{F}^s_D \subseteq \mathcal{K} \cap \{x^*\}^\perp \subsetneq \mathcal{K}$ *and* $x_1^* \notin (\mathcal{K}^1)^\perp = \operatorname{lin}((\mathcal{K}^1)^*)$.

(iii) $\theta_{P_{\mathcal{K}}} = \theta_{D_{\mathcal{K}}} > 0$ *if and only if the PPS condition is satisfied for* (D). *In this case, we have* $c - \mathcal{A}^*\frac{y_3^*}{y_1^*} \in (\operatorname{ri}\mathcal{K}^1) \times \{0\}$. *(In particular, it is feasible for* (D)*)*

Proof. (i) Due to choice of e and e^*, it is clear that the solutions meet the PPS condition.

(ii) $\boxed{\theta_{P_{\mathcal{K}}} = 0 \Rightarrow \text{(a) or (b) holds.}}$ Suppose that $\theta_{P_{\mathcal{K}}} = 0$ and let $(x^*, 0, w^*)$ be an optimal solution for ($P_{\mathcal{K}}$). Due to the constraints in ($P_{\mathcal{K}}$), we have

$$\mathcal{A}x^* = 0, \qquad -\langle c, x^*\rangle = w^* \geq 0, \qquad x^* \in \mathcal{K}^*.$$

Note that if $\langle c, x^*\rangle < 0$, then $\mathcal{F}^D_{\min} = \mathcal{F}^s_{\mathrm{D}} = \emptyset$. This is alternative (a).
If $\langle c, x^*\rangle = 0$, then ((1)) implies $w^* = 0$. Since $\mathcal{A}x^* = 0$, we obtain

$$\mathcal{F}^s_{\mathrm{D}} \subseteq \mathcal{K} \cap \{x^*\}^\perp.$$

By (2), we have $\langle e, x^*\rangle = 1$, so that $x^* = (x_1^*, x_2^*)$ satisfies $x_1^* \notin (\mathcal{K}^1)^\perp$, due to the choice of e. Therefore, the inclusion $\mathcal{K} \cap \{x^*\}^\perp \subseteq \mathcal{K}$ is strict, that is, $\mathcal{K} \cap \{x^*\}^\perp \subsetneq \mathcal{K}$. This is alternative (b).

$\boxed{\text{(a) or (b) holds} \Rightarrow \theta_{P_{\mathcal{K}}} = 0.}$ First recall that t is constrained to be nonnegative, therefore $\theta_{P_{\mathcal{K}}} \geq 0$. Next, we will prove the following inequality

$$\langle e, x^*\rangle - \langle c, x^*\rangle > 0. \tag{7}$$

Since $e \in \mathcal{K}, x^* \in \mathcal{K}^*$, we have $\langle e, x^*\rangle \geq 0$. If (a) holds, then $-\langle c, x^*\rangle > 0$. In particular, (7) holds. Now, we consider the case where (b) holds. In this case, we have $\langle c, x^*\rangle = 0$. If $\langle e, x^*\rangle$ is also zero, then $\langle e_1, x_1^*\rangle$ is zero and since $e_1 \in \operatorname{ri}\mathcal{K}^1$, x_1^* must belong to $(\mathcal{K}^1)^\perp$. However, this is not allowed since we assumed that $x_1^* \notin (\mathcal{K}^1)^\perp$. Therefore, we have $\langle e, x^*\rangle > 0$ and, in particular, (7) also holds.
Let

$$\alpha = \frac{1}{\langle e, x^*\rangle - \langle c, x^*\rangle}.$$

Then $\alpha > 0$ and $(\alpha x^*, 0, -\alpha\langle c, x^*\rangle)$ is a solution to ($P_{\mathcal{K}}$) with value 0, so that $\theta_{P_{\mathcal{K}}} = 0$.

(iii) $\boxed{\theta_{D_{\mathcal{K}}} > 0 \Rightarrow \text{PPS holds for (D)}}$ From (4), we have

$$cy_1^* - ey_2^* - \mathcal{A}^*y_3^* \in \mathcal{K}.$$

Since $y_2^* = \theta_{D_{\mathcal{K}}} > 0$, we have, due to (6), $y_1^* > 0$. Hence,

$$c - e\frac{y_2^*}{y_1^*} - \mathcal{A}^*\frac{y_3^*}{y_1^*} \in \mathcal{K}.$$

We conclude that $c - \mathcal{A}^* y_3^*/y_1^* \in (\mathcal{K}^1 + y_2^*/y_1^* e_1) \times \{0\} \subseteq (\operatorname{ri} \mathcal{K}^1) \times \{0\}$, due to the choice of e and the fact that $y_2^*/y_1^* > 0$.

$\boxed{\theta_{D_\mathcal{K}} = 0 \Rightarrow \text{PPS does not hold for (D)}}$ If $t^* = 0$ and $(x^*, 0, w^*)$ is an optimal solution for ($P_\mathcal{K}$), then either (a) or (b) of item (ii) is satisfied. If (a) is satisfied, then (D) is infeasible and thus the PPS condition cannot hold. If (b) is satisfied, then $\langle c, x^* \rangle = 0$, $\mathcal{A}x^* = 0$ and $x_1^* \notin (\mathcal{K}^1)^\perp$. If (s_1, s_2) is a feasible slack for (D), we have $\langle s_1, x_1^* \rangle + \langle s_2, x_2^* \rangle = 0$, so that $\langle s_1, x_1^* \rangle = 0$. As $x_1^* \notin (\mathcal{K}^1)^\perp$, we have that $s_1 \notin \operatorname{ri} \mathcal{K}^1$. This means that there is no feasible solution for (D) satisfying the PPS condition.

□

We now prove a theorem that dualizes the criterion in Proposition 2.

Theorem 4. *Let $c \in \mathbb{R}^n$, $\mathcal{L} \subseteq \mathbb{R}^n$ be a subspace and $\mathcal{K} = \mathcal{K}^1 \times \mathcal{K}^2$ be a closed convex cone, such that $\mathcal{K}^2$ is polyhedral. Then $(\mathcal{L} + c) \cap ((\operatorname{ri} \mathcal{K}^1) \times \mathcal{K}^2) = \emptyset$ if and only if one of the conditions below holds:*

(a) there exists $x \in \mathcal{K}^* \cap \mathcal{L}^\perp$ such that $\langle c, x \rangle < 0$;

(b) there exists $x = (x_1, x_2) \in \mathcal{K}^* \cap \mathcal{L}^\perp \cap \{c\}^\perp$ such that $x_1 \notin (\mathcal{K}^1)^\perp$.

Proof. Select a linear map $\mathcal{A}$ such that $\mathcal{L} = \operatorname{range} \mathcal{A}^*$ (therefore, $\mathcal{L}^\perp = \ker \mathcal{A}$) and consider the problem (D) and the pair of problems ($P_\mathcal{K}$) and ($D_\mathcal{K}$). Note that $(\mathcal{L} + c) \cap ((\operatorname{ri} \mathcal{K}^1) \times \mathcal{K}^2) = \emptyset$ if and only if the PPS condition is *not* satisfied for (D). The result then follows from Lemma 3. □

One of the points of doing facial reduction is to solve problems that would not be solvable directly if, say, we fed them to an interior point method based solver. Therefore, it is natural to consider whether the problems ($P_\mathcal{K}$), ($D_\mathcal{K}$) are themselves solvable, with the choice of e, e^* provided by Lemma 3. We note that due to item (i) of Lemma 3, the pair ($P_\mathcal{K}$), ($D_\mathcal{K}$) can be solved by infeasible interior-point methods in the case of semidefinite and second order cone programming, even though they might fail to be strongly feasible. This is because the convergence theory relies on the existence of optimal solutions affording zero duality gap, rather than strong feasibility. See, for instance, item 2. of Theorem 11 in the work by Nesterov, Todd and Ye [20].

We remark that Theorem 4 implies a version of the Gordan-Stiemke's Theorem that takes into account partial polyhedrality. It contains as a special case the classical version described in Corollary 2 in Luo, Sturm and Zhang [17].

Corollary 5 (Partial Polyhedral Gordan-Stiemke's Theorem)**.** *Let $\mathcal{L}$ be a subspace and $\mathcal{K} = \mathcal{K}^1 \times \mathcal{K}^2$ be a closed convex cone, such that $\mathcal{K}^2$ is polyhedral. Then:*

$$\mathcal{K}^* \cap \mathcal{L}^\perp \subseteq (\operatorname{lin}((\mathcal{K}^1)^*)) \times (\mathcal{K}^2)^* \Leftrightarrow \big((\operatorname{ri}(\mathcal{K}^1)) \times (\mathcal{K}^2)\big) \cap \mathcal{L} \neq \emptyset.$$

Proof. Take $c = 0$ in Theorem 4 and recall that $\operatorname{lin}((\mathcal{K}^1)^*) = (\mathcal{K}^1)^\perp$. □

For more results taking into account partial polyhedrality see Chapter 20 of [25] and Propositions 1 and 2 of [15].

5 Distance to polyhedrality, FRA-Poly and tightness

Here we will discuss FRA-Poly, which is a facial reduction algorithm divided into two phases. The first detects infeasibility and restores strong duality, while the second finds the minimal face. For an example illustrating FRA-Poly, see Appendix B.

The idea behind the classical FRA is that whenever strong feasibility fails, we can obtain reducing directions until strong feasibility is satisfied again. Similarly, Phase 1 of FRA-Poly is based on the fact that whenever the PPS condition in Proposition 2 fails, we may also obtain reducing directions until the PPS condition is satisfied, thanks to Theorem 4. After that, a single extra facial reduction step is enough to go to the minimal face. As the PPS condition is weaker than full-on strong feasibility, FRA-Poly has better worst case bounds in many cases.

We now present a disclaimer of sorts. The theoretical results presented in this section and the next stand whether FRA-Poly is doable or not for a given $\mathcal{K}$. If we wish to do facial reduction concretely (even if it is by hand!), we need to make a few assumptions on our computational capabilities and on our knowledge on the lattice of faces of $\mathcal{K}$. First of all, we must be able to solve problems over faces of $\mathcal{K}$ such that both the primal and the dual satisfy the PPS condition and we must also be able to do basic linear algebraic operations. Also, for each face $\mathcal{F}$ of $\mathcal{K}$ we must know:

1. $\operatorname{span} \mathcal{F}$,
2. at least one point $e \in \operatorname{ri} \mathcal{F}$,
3. at least one point $e^* \in \operatorname{ri} \mathcal{F}^*$,
4. whether $\mathcal{F}$ is polyhedral or not.

We remark that apart from knowledge about the polyhedral faces, our assumptions are not very different from what it is usually assumed *implicitly* in the FRA literature. For symmetric cones, which include direct products of $\mathcal{S}^n_+$, $\mathcal{Q}^n$ and $\mathbb{R}^n_+$, they are reasonable since their lattice of faces is well-understood and every face is again a symmetric cone. So, for instance, e can be taken as the identity element for the corresponding Jordan algebra. On the other hand, if $\mathcal{K}$ is, say, the copositive cone $\mathcal{C}^n$, we might have some trouble fulfilling the requirements, inasmuch as our knowledge of the faces of $\mathcal{C}^n$ is still lacking.

5.1 Distance to Polyhedrality

Here we introduce the notion of *distance to polyhedrality*. In what follows, if we have a chain of faces $\mathcal{F}_1 \subsetneq \ldots \subsetneq \mathcal{F}_\ell$, the length of the chain is defined to be ℓ.

Definition 6. *Let $\mathcal{K}$ be a nonempty closed convex cone. The* distance to polyhedrality $\ell_{poly}(\mathcal{K})$ *is the length* minus one *of the longest strictly ascending chain of nonempty faces $\mathcal{F}_1 \subsetneq \ldots \subsetneq \mathcal{F}_\ell$ which satisfies:*

1. *$\mathcal{F}_1$ is polyhedral;*
2. *$\mathcal{F}_j$ is not polyhedral for $j > 1$.*

The distance to polyhedrality is a well-defined concept, because the lineality space of $\mathcal{K}$ is always a polyhedral face of $\mathcal{K}$. Moreover, $\ell_{\text{poly}}(\mathcal{K})$ counts the maximum number of facial reduction steps that can be taken before we reach a polyhedral face.

Example 1. *See section 2 and examples 2.5 and 2.6 in [21] for more details on the facial structure of $\mathcal{S}^n_+$ and $\mathcal{Q}^n$. For the positive semidefinite cone $\mathcal{S}^n_+$, we have $\ell_{poly}(\mathcal{S}^n_+) = n-1$. Liu and Pataki defined in [12]* smooth cones *as full-dimensional, pointed cones (i.e., $\mathcal{K} \cap -\mathcal{K} = \{0\}$) such that any face that is not $\{0\}$ nor $\mathcal{K}$ must be a half-line. For those cones we have $\ell_{poly}(\mathcal{K}) = 1$, when the dimension of $\mathcal{K}$ is greater than 2. Examples of smooth cones include the Lorentz cone $\mathcal{Q}^n$ and the p-cones, for $1 < p < \infty$. For comparison, the longest chain of nonempty faces of $\mathcal{S}^n_+$ has length $n + 1$ and the one for any smooth cone with dimension greater than 2 has length 3. Note that $\mathcal{S}^n_+$ and $\mathcal{Q}^n$ are examples of* symmetric cones *and we mention in passing that a discussion about the longest chain of faces of a general symmetric cone can be found in Section 5.3 of [10].*

5.2 Strict complementarity in ($P_{\mathcal{K}}$) and ($D_{\mathcal{K}}$)

The last ingredient we need is a discussion on the cases where jumping to $\mathcal{F}^D_{\min}$ with a single facial reduction step is possible. Let x^*, y^* be any pair of optimal solutions to (P) and (D). Recall that if $\mathcal{K}$ is $\mathbb{R}^n_+, \mathcal{S}^n_+$ or $\mathcal{Q}^n$, then x^*, y^* are said to be *strictly complementary* if the following equivalent conditions hold:

$$s^* \in \operatorname{ri}(\mathcal{K} \cap \{x^*\}^\perp) \Leftrightarrow x^* \in \operatorname{ri}(\mathcal{K}^* \cap \{s^*\}^\perp) \Leftrightarrow \langle x^*, s^* \rangle = 0 \text{ and } x^* + s^* \in \operatorname{ri} \mathcal{K},$$

where $s^* = c - \mathcal{A}^* y^*$. For general $\mathcal{K}$, these equivalencies may not hold and we might need to distinguish between primal and dual strict complementarity, see for instance, Definition 3.4 and Remark 3.6 in the chapter by Pataki [21] and Equation (2.6) in Section 2 of the work by Tunçel and Wolkowicz [28]. Based on those references, we will say that ($P_{\mathcal{K}}$) and ($D_{\mathcal{K}}$) satisfies *(dual) strict complementarity* if $\theta_{P_{\mathcal{K}}} = \theta_{D_{\mathcal{K}}}$ and there are optimal solutions $(x^*, t^*, w^*), (y_1^*, y_2^*, y_3^*)$ such that

$$cy_1^* - ey_2^* - \mathcal{A}^* y_3^* \in \operatorname{ri}(\mathcal{K} \cap \{x^*\}^\perp) \tag{8}$$
$$t^* + 1 - y_1^*(1 + \langle c, e^* \rangle) + \langle e^*, \mathcal{A}^* y_3^* \rangle > 0 \tag{9}$$
$$w^* + y_1^* - y_2^* > 0. \tag{10}$$

Proposition 7. *Suppose $\theta_{P_{\mathcal{K}}} = \theta_{D_{\mathcal{K}}} = 0$ and that we have optimal solutions to ($P_{\mathcal{K}}$) and ($D_{\mathcal{K}}$) satisfying dual strict complementarity. If $w^* = 0$, then $\mathcal{F}^D_{\min} = \mathcal{K} \cap \{x^*\}^\perp$.*

Proof. Since $w^* = 0$, (10) implies that $y_1^* > y_2^*$. Then, $\theta_{P_{\mathcal{K}}} = \theta_{D_{\mathcal{K}}}$ and $\theta_{P_{\mathcal{K}}} \geq 0$ implies that $y_2^* \geq 0$, so that $y_1^* > 0$ as well. Therefore, from (8) we obtain $c - \mathcal{A}^* \frac{y_3^*}{y_1^*} \in \operatorname{ri}(\mathcal{K} \cap \{x^*\}^\perp)$. □

Therefore, under strict complementarity, we can find $\mathcal{F}^D_{\min}$ with a single facial reduction step. Note that, here, we do not care about the choice of e, e^*. For semidefinite programming, a similar observation was made in Theorem 12.28 of [5], where reducing directions are found through an auxiliary problem (AP). There, the authors show that a single direction is needed if and only if their AP satisfy strict complementarity. Another characterization of when one direction is enough can be found in Theorem 4.1 of [7]. One small advantage of ($P_{\mathcal{K}}$) and ($D_{\mathcal{K}}$) is that only linear constraints are used in addition to the conic constraints induced by $\mathcal{K}$. In contrast, AP also adds quadratic constraints.

5.3 FRA-Poly

Henceforth, we will assume that $\mathcal{K}$ is the product of r cones and we will write $\mathcal{K} = \mathcal{K}^1 \times \ldots \times \mathcal{K}^r$. Then, recall that if $\mathcal{F}$ is face of $\mathcal{K}$, we can write $\mathcal{F} = \mathcal{F}^1 \times \ldots \times \mathcal{F}^r$, where $\mathcal{F}^i$ is a face of $\mathcal{K}^i$ for every i.

Consider the following FRA variant, which we call FRA-Poly.

[**Facial Reduction Poly - Phase 1**]

Input: (D)

Output: A set of reducing directions $\{d_1, \ldots, d_\ell\}$. If (D) is feasible, it outputs some face $\mathcal{F} \subseteq \mathcal{K}$ for which the PPS condition holds, together with a dual slack s' for which $s'_j \in \operatorname{ri} \mathcal{F}^j$ for every j such that $\mathcal{F}^j$ is nonpolyhedral. If (D) is infeasible, the directions form a certificate of infeasibility.

1. $\mathcal{F}_1 \leftarrow \mathcal{K}, i \leftarrow 1$
2. Let (x^*, t^*, w^*) and (y_1^*, y_2^*, y_3^*) be any pair of optimal solutions to ($P_{\mathcal{K}}$) and ($D_{\mathcal{K}}$) with
   - $\mathcal{F}_i$ in place of $\mathcal{K}$, where $\mathcal{F}_i = \mathcal{F}_i^1 \times \ldots \times \mathcal{F}_i^r$,
   - any $e^* \in \operatorname{ri} \mathcal{F}_i^*$,
   - any e such that $e_j = 0$ if $\mathcal{F}_i^j$ is polyhedral and $e_j \in \operatorname{ri} \mathcal{F}_i^j$, otherwise.
3. If $t^* = 0$ and $\langle c, x^* \rangle < 0$, let $\mathcal{F}^D_{\min} \leftarrow \emptyset$ and stop. (D) is infeasible.
4. If $t^* = 0$ and $\langle c, x^* \rangle = 0$, let $d_i \leftarrow x^*, \mathcal{F}_{i+1} \leftarrow \mathcal{F}_i \cap \{d_i\}^\perp, i \leftarrow i + 1$ and return to 2).
5. If $t^* > 0$, $s' \leftarrow c - \mathcal{A}^* \frac{y_3^*}{y_1^*}$, $\mathcal{F} \leftarrow \mathcal{F}_i$ and stop. PPS condition is satisfied.

Note that Phase 1 of FRA-Poly might not end at the minimal face, but still, due to Proposition 2, strong duality will be satisfied. First, we will prove the correctness of Phase 1, which essentially follows from Lemma 3.

Proposition 8. *The following hold.*

(i) if (D) *is feasible, then the output face $\mathcal{F}$ satisfies $\mathcal{F}^D_{\min} \subseteq \mathcal{F}$. Moreover, s' is a dual feasible slack such that $s'_j \in \operatorname{ri} \mathcal{F}^j$ for every j such that $\mathcal{F}^j$ is nonpolyhedral, i.e., the PPS condition is satisfied for $\mathcal{F}$.*

In this case, Phase 1 stops after finding at most $\sum_{i=1}^r \ell_{poly}(\mathcal{K}^i)$ directions.

(ii) (D) *is infeasible if and only if Step 3 is reached. In this case, Phase 1 stops after finding at most $1 + \sum_{i=1}^r \ell_{poly}(\mathcal{K}^i)$ directions.*

Proof. We will focus on the statements about the bounds, since the other statements are direct consequences of Lemma 3. Note that whenever Step 4 is reached, we have $\mathcal{F}_{i+1} \subsetneq \mathcal{F}_i$, since $x^*_j \notin (\mathcal{F}^j_i)^\perp$ for at least one nonpolyhedral cone $\mathcal{F}^j_i$, due to item (ii)-(b) of Lemma 3. Therefore, whenever a new (proper) face is found, it is because we are making progress towards a polyhedral face for at least one nonpolyhedral cone.

By definition, after finding $\hat{\ell} = \sum_{i=1}^r(\ell_{\text{poly}}(\mathcal{K}^i))$ directions, $\mathcal{F}_{\hat{\ell}+1}$ is polyhedral. We now consider what happens if the algorithm has not stopped after all these directions were found. In this case, when it is time to build $(P_\mathcal{K})$ and $(D_\mathcal{K})$ with $\mathcal{F}_{l+1}$ in place of $\mathcal{K}$ at Step 2, Phase 1 selects $e = 0$ and $e^* \in (\mathcal{F}_{\hat{\ell}+1})^*$.

First, suppose that (D) is feasible and let y be such that $c - \mathcal{A}^* y \in \mathcal{F}^s_{\text{D}}$. Since $e = 0$, $(\alpha, \alpha, \alpha y)$ is a feasible solution for $(D_\mathcal{K})$, when $\alpha > 0$ is sufficiently small. It follows that $\theta_{D_\mathcal{K}} > 0$ and that Phase 1 eventually reaches Step 5. This gives item (i).

Finally, suppose that (D) is infeasible. By Lemma 3, $\theta_{D_\mathcal{K}} = 0$. Since $e = 0$, (2) implies that every optimal solution of $(P_\mathcal{K})$ will be a triple of the form $(x^*, 0, 1)$, which implies that Step 3 will be reached and a single new direction will be added. This gives item (ii). ☐

Remark. *Let ℓ be the number of directions found in Phase 1. When* (D) *is feasible, $\ell+1$ is the total number of times the pair of problems $(P_\mathcal{K})$, $(D_\mathcal{K})$ are solved during Phase 1. After solving $(P_\mathcal{K})$, $(D_\mathcal{K})$ ℓ times, a face $\mathcal{F}$ of $\mathcal{K}$ satisfying the PPS condition is computed. However, it is necessary to solve $(P_\mathcal{K})$, $(D_\mathcal{K})$ one extra time to reach the stopping criteria in Step 5 and obtain s', which is a certificate that $\mathcal{F}$ indeed satisfies the PPS condition. When* (D) *is infeasible, ℓ coincides with the number of times problems $(P_\mathcal{K})$, $(D_\mathcal{K})$ are solved during Phase 1.*

If we merely want a face of $\mathcal{K}$ satisfying the PPS condition, we can stop at Phase 1. In any case, we will now show that is possible to jump directly to the minimal face in a single facial reduction step.

[**Facial Reduction Poly - Phase 2**]
Input: (D), the output of Phase 1: $\mathcal{F}$ and s', with $\mathcal{F} \neq \emptyset$.
Output: $\mathcal{F}^D_{\min}$, a dual feasible slack $\hat{s} \in \operatorname{ri} \mathcal{F}^D_{\min}$ and, perhaps, an extra reducing direction d.

1. Let $\hat{\mathcal{K}} = \hat{\mathcal{K}}^1 \times \ldots \times \hat{\mathcal{K}}^r$ such that $\hat{\mathcal{K}}^j = \mathcal{F}^j$ if $\mathcal{F}^j$ is polyhedral and $\hat{\mathcal{K}}^j = \operatorname{span} \mathcal{F}^j$ otherwise.
2. Let (x^*, t^*, w^*) and (y^*_1, y^*_2, y^*_3) be any pair of *strictly complementary* optimal solutions to $(P_\mathcal{K})$ and $(D_\mathcal{K})$ with
   - $\hat{\mathcal{K}}$ in place of $\mathcal{K}$,
   - any $e \in \operatorname{ri} \hat{\mathcal{K}}$,
   - any $e^* \in \operatorname{ri} \hat{\mathcal{K}}^*$.
3. If $t^* = 0$, let $d \leftarrow x^*$, $\mathcal{F}^D_{\min} \leftarrow \mathcal{F} \cap \{x^*\}^\perp$. Let $\tilde{s}$ be $c - \mathcal{A}^* \frac{y^*_3}{y^*_1}$. Then, we let $\hat{s}$ be a convex combination of $\tilde{s}$ and s' such that $\hat{s} \in \operatorname{ri} \mathcal{F}^D_{\min}$ and stop.[2]
4. If $t^* > 0$, $\mathcal{F}^D_{\min} \leftarrow \mathcal{F}$. Let $\tilde{s}$ be $c - \mathcal{A}^* \frac{y^*_3}{y^*_1}$. Then, we let $\hat{s}$ be a convex combination of $\tilde{s}$ and s' such that $\hat{s} \in \operatorname{ri} \mathcal{F}^D_{\min}$ and stop.[2]

[2] In Proposition 9 it is shown that if $z_\beta = \beta s' + (1-\beta)\tilde{s}$ is such that $\beta \in (0, 1)$ and β is sufficiently close to 1 then $z_\beta \in \operatorname{ri} \mathcal{F}^D_{\min}$. Therefore, after determining $\mathcal{F}^D_{\min}$, $\hat{s}$ can be found through a simple search procedure, e.g., the bisection method.

Note that in Phase 2, the cone $\hat{\mathcal{K}}$ is polyhedral, therefore, both ($D_{\mathcal{K}}$) and ($P_{\mathcal{K}}$) are polyhedral problems. Therefore, strictly complementary solutions are ensured to exist, which is a consequence of Goldman-Tucker Theorem and also follows from the results of McLinden [18] and Akgül [1]. We also remark that a strictly complementary solution of a polyhedral problem can be found by solving a single linear program, see, for instance, the article by Freund, Roundy and Todd [8] and the related work by Mehrotra and Ye [19].

We now try to motivate the next proposition. At Phase 2, a single facial reduction iteration is performed. In the usual facial reduction approach, we would build the problems ($D_{\mathcal{K}}$) and ($P_{\mathcal{K}}$) using $\mathcal{K} = \mathcal{F}$ and seek a reducing direction belonging to $\mathcal{F}^*$. The subtle point in Phase 2 is that we use $\hat{\mathcal{K}}$ in place of $\mathcal{F}$, which is potentially larger since the nonpolyhedral blocks were relaxed to their span. This restricts our search for reducing directions to $\hat{\mathcal{K}}^*$, which is potentially smaller than $\mathcal{F}^*$. However, the proof in Proposition 9 will show that, at this stage, any reducing direction must be already confined to $\hat{\mathcal{K}}^*$.

Proposition 9. *The output face of Phase 2 is $\mathcal{F}^D_{\min}$ and there exists $\hat{s}$ as in Steps 3. and 4.*

Proof. Suppose that the output face $\mathcal{F}$ of Phase 1 satisfies $\mathcal{F} \neq \mathcal{F}^D_{\min}$. By Lemma 3.2 in [29], there is a reducing direction x such that $x \in \mathcal{F}^* \cap \ker \mathcal{A} \cap \{c\}^\perp$ and $x \notin \mathcal{F}^\perp$. Due to Proposition 8, any such reducing direction x satisfies $\langle x, s' \rangle = 0$, which implies that $x_j \in (\mathcal{F}^j)^\perp = (\operatorname{span} \mathcal{F}^j)^\perp$ for every j such that $\mathcal{F}^j$ is not polyhedral, since $s'_j \in \operatorname{ri} \mathcal{F}^j$ for those j. Therefore, the possible reducing directions are confined to the polyhedral cone $\hat{\mathcal{K}}^*$, where $\hat{\mathcal{K}}$ is the cone in Step 1. of Phase 2.

Since $\hat{\mathcal{K}}$ is polyhedral, the problems ($P_{\mathcal{K}}$) and ($D_{\mathcal{K}}$) are polyhedral and they admit strictly complementary optimal solutions (x^*, t^*, w^*), (y_1^*, y_2^*, y_3^*). The fact that $x \notin \mathcal{F}^\perp$ implies that $\langle e, x \rangle \neq 0$ so that $(x/\langle e, x \rangle, 0, 0)$ is an optimal solution to ($P_{\mathcal{K}}$). Therefore, $t^* = y_2^* = 0$. Moreover, since (D) is feasible, we have $w^* = 0$. By Proposition 7, we have $c - \mathcal{A}^* \frac{y_3^*}{y_1^*} \in \operatorname{ri}(\hat{\mathcal{K}} \cap \{x^*\}^\perp)$.

Let $\tilde{s} = c - \mathcal{A}^* \frac{y_3^*}{y_1^*}$. Note that $\mathcal{F} \cap \{x^*\}^\perp$ is a face of $\mathcal{F}$ containing $\mathcal{F}^s_{\mathrm{D}}$, since we argued that x^* must be a reducing direction. We will prove that $\mathcal{F}^D_{\min} = \mathcal{F} \cap \{x^*\}^\perp$ by showing that some convex combination of s' and $\tilde{s}$ lies in $\operatorname{ri}(\mathcal{F} \cap \{x^*\}^\perp)$.

Let $z_\beta = \beta s' + (1 - \beta)\tilde{s}$. For all $\beta \in (0, 1)$ and all j such that $\mathcal{F}^j$ is polyhedral, we have $(z_\beta)_j \in \operatorname{ri}(\mathcal{F}^j \cap \{x_j^*\}^\perp)$, because $\tilde{s}_j \in \operatorname{ri}(\mathcal{F}^j \cap \{x_j\}^\perp)$ and s' is feasible. If $\mathcal{F}^j$ is not polyhedral, then $\mathcal{F}^j \cap \{x_j^*\}^\perp = \mathcal{F}^j$, since $x_j \in (\mathcal{F}^j)^\perp$. Because $\tilde{s}_j \in \operatorname{span} \mathcal{F}^j$ and $s'_j \in \operatorname{ri} \mathcal{F}^j$, for all β sufficiently close to 1 we have $(z_\beta)_j \in \operatorname{ri} \mathcal{F}^j$. Therefore, it is possible to select $\beta \in (0, 1)$ such that $(z_\beta)_j \in \operatorname{ri}(\mathcal{F}^j \cap \{x_j\}^\perp)$ for all j. This shows that $\mathcal{F}^D_{\min} = \mathcal{F} \cap \{x^*\}^\perp$.

If $\mathcal{F}$ was already the minimal face to begin with, then $t^* > 0$. We can then proceed in a similar fashion. The only difference is that due to (4), we will have that $\tilde{s} = c - \mathcal{A}^* \frac{y_3^*}{y_1^*}$ satisfies $\tilde{s}_j \in \operatorname{ri}(\mathcal{F}^j)$ for every j such that $\mathcal{F}^j$ is polyhedral. And as before, we can select a convex combination of s' and $\tilde{s}$ belonging to the relative interior of $\mathcal{F}^D_{\min}$. ☐

We then arrive at the main result of this section.

Theorem 10. *Let $\mathcal{K} = \mathcal{K}^1 \times \ldots \times \mathcal{K}^r$. The minimum face $\mathcal{F}^D_{\min}$ that contains the feasible region of* (D) *can be found in no more than $1 + \sum_{i=1}^r \ell_{poly}(\mathcal{K}^i)$ facial reduction steps.*

Proof. If (D) is infeasible, then $\mathcal{F}^D_{\min} = \emptyset$ and the result follows from Proposition 8. So suppose now that (D) is feasible. Then Phase 1 ends after finding at most $\sum_{i=1}^r \ell_{\mathrm{poly}}(\mathcal{K}^i)$ directions. Due to Proposition 9, at most one extra direction is needed to jump to the minimal face. ☐

Recall that $\ell_{\mathcal{K}}$ is the length of the longest chain of strictly ascending nonempty faces of $\mathcal{K}$. If one uses the "classical" facial reduction approach, it takes no more than $\ell_{\mathcal{K}} - 1$ facial reduction steps to find the minimal face, when (D) is feasible. See, for instance, Theorem 1 in [22] or Corollary 3.1 in [29]. When $\mathcal{K}$ is a direct product of several cones, we have $\ell_{\mathcal{K}} = 1 + \sum_{i=1}^r (\ell_{\mathcal{K}^i} - 1)$. We will end this subsection by showing that, under the relatively weak hypothesis that $\mathcal{K}^i$ is not a subspace, we have $\ell_{\mathrm{poly}}(\mathcal{K}^i) < \ell_{\mathcal{K}^i} - 1$. This means that FRA-Poly compares favorably to the classical FRA analysis and the difference between the two bounds grows at least linearly with the number of cones.

Theorem 11. *If $\mathcal{K}$ is not a subspace then $1 + \ell_{poly}(\mathcal{K}) \leq \ell_{\mathcal{K}} - 1$. In particular, if $\mathcal{K}$ is the direct product of r closed convex cones that are not subspaces we have:*

$$1 + r + \sum_{i=1}^{r} \ell_{poly}(\mathcal{K}^i) \leq 1 + \sum_{i=1}^{r} (\ell_{\mathcal{K}^i} - 1).$$

Proof. Let $U = \operatorname{lin} \mathcal{K}$. Then we have $\mathcal{K} = (\mathcal{K} \cap U^\perp) + U$. If we let $\hat{\mathcal{K}} = \mathcal{K} \cap (U^\perp)$, we have that $\operatorname{lin}(\hat{\mathcal{K}}) = \{0\}$ so that $\hat{\mathcal{K}}$ is pointed and $\hat{\mathcal{K}} \neq \{0\}$ if $\mathcal{K}$ is not a subspace. Recall that the minimal nonzero face of any nonzero pointed cone must be an extreme ray, i.e., an one dimensional face. Therefore, the first two faces of any longest chain of faces of $\hat{\mathcal{K}}$ must be $\{0\}$ and some extreme ray. Therefore, we have $1 + \ell_{\text{poly}}(\hat{\mathcal{K}}) \leq \ell_{\hat{\mathcal{K}}} - 1$.

Note that there is a bijection between the faces of $\mathcal{K}$ and the set $\{\mathcal{F} + U \mid \mathcal{F} \text{ is a face of } \hat{\mathcal{K}}\}$. A similar correspondence holds between the polyhedral faces of $\mathcal{K}$ and the set $\{\mathcal{F} + U \mid \mathcal{F} \text{ is a polyhedral face of } \hat{\mathcal{K}}\}$. Therefore, $\ell_{\mathcal{K}} = \ell_{\hat{\mathcal{K}}}$ and $\ell_{\text{poly}}(\mathcal{K}) = \ell_{\text{poly}}(\hat{\mathcal{K}})$. This shows that $1 + \ell_{\text{poly}}(\mathcal{K}) \leq \ell_{\mathcal{K}} - 1$. To conclude, note that if $\mathcal{K}$ is a direct product of r cones then $\ell_{\mathcal{K}} = 1 + \sum_{i=1}^{r} (\ell_{\mathcal{K}^i} - 1)$, so the result follows from applying what we have done so far to each $\mathcal{K}^i$. ☐

5.4 Tightness of the bound

It is reasonable to consider whether there are instances that actually need the amount of steps predicted by Theorem 10. In this section we will take a look at this issue. The following notion will be helpful.

Definition 12 (Singularity degree)**.** *Consider the set of possible outputs $\{d_1, \ldots, d_\ell\}$ of the Generic Facial Reduction algorithm in Section 3. The singularity degree of* (D) *is the minimum ℓ among all the possible outputs and is denoted by $d(D)$.*

That is, the singularity degree is the minimum number of facial reduction steps before $\mathcal{F}_{\min}^D$ is found. In the recent work by Liu and Pataki [12], there is also an equivalent definition of singularity degree for feasible problems, see Definition 6 therein. As far as we know, the expression "singularity degree" in this context is due to Sturm in [26], where he showed the connection between the singularity degree of a positive semidefinite program and error bounds, see also [16].

The singularity degree of (D) is a quantity that depends on $c, \mathcal{A}$ and $\mathcal{K}$. The classical facial reduction strategy gives the bounds $d(D) \leq \ell_{\mathcal{K}} - 1$ when (D) is feasible and $d(D) \leq \ell_{\mathcal{K}}$ when (D) is infeasible. Theorem 10 readily implies that $d(D) \leq 1 + \sum_{i=1}^{r} \ell_{\text{poly}}(\mathcal{K}^i)$, no matter whether (D) is feasible or not. Due to Theorem 11, this bound is likely to compare favorably to $\ell_{\mathcal{K}} - 1 = \sum_{i=1}^{r} (\ell_{\mathcal{K}^i} - 1)$.

Tunçel constructed an SDP instance with singularity degree $d(D) = n - 1 = \ell_{\mathcal{S}_+^n} - 2$, see Section 2.6 in [27] or the section "Worst case instance" in [5]. Now, let $\mathcal{K} = \mathcal{Q}^{t_1} \times \ldots \times \mathcal{Q}^{t_{r_1}} \times \mathcal{S}_+^{n_1} \times \ldots \times \mathcal{S}_+^{n_{r_2}}$ be the direct product of r_1 second order (Lorentz) cones and r_2 positive semidefinite cones. In this case, Theorem 10 implies that $d(D) \leq 1 + r_1 + \sum_{j=1}^{r_2} (n_j - 1)$. In Appendix B we extend Tunçel's example and show that for every such $\mathcal{K}$ there is a feasible instance for which $d(D) = r_1 + \sum_{j=1}^{r_2} (n_j - 1)$, thus showing the worst case bound in Theorem 10 is off by at most one.

This type of $\mathcal{K}$ was also studied by Luo and Sturm in [16], where they discussed a regularization procedure which ends in at most $r_1 + \sum_{j=1}^{r_2} (n_j - 1)$ steps, see Theorem 7.4.1 therein. However, their definition of regularity does not imply strong feasibility, so similarly to Phase 1 of FRA-Poly, an additional step is necessary (akin to a facial reduction step) before the minimal face is reached, see Lemma 7.3.3. In total we get the same bound predicted by Theorem 10.

We remark that we were unable to construct a feasible instance with singularity degree $1 + r_1 + \sum_{j=1}^{r_2} (n_j - 1)$. Note that if $\mathcal{K} = \mathcal{S}_+^n$, since each facial reduction step reduces the possible ranks of feasible matrices, if we need n steps it is because $\mathcal{F}_{\text{D}}^s = \{0\}$. But if $\mathcal{F}_{\text{D}}^s = \{0\}$, then $c \in \operatorname{range} \mathcal{A}^*$ and Gordan-Stiemke's Theorem implies the existence of $d \in (\operatorname{ri} \mathcal{K}^*) \cap \ker \mathcal{A}$. Therefore, we can go to $\mathcal{F}_{\min}^D$ with a single step, since $\mathcal{K} \cap \{d\}^\perp = \{0\}$. So, in fact, we never need more than $n - 1$ steps for feasible SDPs and Tunçel's example is indeed the worse that could happen in this case. A similar argument holds when $\mathcal{K} = \mathcal{Q}^n$, where we never need more than a single step if we select the directions optimally. But when we have direct products, the possible interactions between the blocks makes it hard to argue that the +1 is unnecessary, although the partial polyhedral Gordan-Stiemke theorem (Corollary 5) helps rule out a few cases.

We will now take a look at what could be done to ensure that a facial reduction algorithm never takes more steps than the necessary to find $\mathcal{F}^D_{\min}$. Consider the Generic Facial Reduction algorithm in Section 3. All the directions, with the possible exception of the last, belong to $\mathcal{F}_i^* \cap \ker \mathcal{A} \cap \{c\}^\perp$. In particular, the FRAs considered in [26, 16] and the FRA-CE variant in [29] always select the most interior direction possible. In our context, this means that whenever step 2.i) is reached the following choice is made:

$$d_i \in \operatorname{ri}(\mathcal{F}_i^* \cap \ker \mathcal{A} \cap \{c\}^\perp). \tag{11}$$

In fact, the singularity degree was originally defined not as in Definition 12, but as the number of steps that their particular algorithms take to find the minimal face[3]. Although intuitive, it is not entirely obvious that the choice in (11) minimizes the number of directions, so let us take a look at this issue.

Proposition 13. *Suppose that* (D) *is feasible and that at each step of the Generic Facial Reduction algorithm* d_i *is selected as in* (11)*. Then, the algorithm will output exactly* $d(D)$ *directions.*

Proof. Suppose $d(D) > 0$ and let $(d_1, \ldots, d_\ell)$ be a sequence of reducing directions such that $\ell = d(D)$ and the last face is $\mathcal{F}^D_{\min}$. Let $d_1^* \in \operatorname{ri}(\mathcal{K}^* \cap \ker \mathcal{A} \cap \{c\}^\perp)$.

Since d_1^* is a relative interior point, there is $\alpha > 1$ such that $\alpha d_1^* + (1-\alpha) d_1 \in \mathcal{K}^* \cap \ker \mathcal{A} \cap \{c\}^\perp$, see Theorem 6.4 in [25]. Now, let $x \in \mathcal{K} \cap \{d_1^*\}^\perp$. We must have

$$\langle x, \alpha d_1^* + (1-\alpha) d_1 \rangle \geq 0.$$

Since $\langle x, d_1^* \rangle = 0$ and $(1-\alpha) < 0$, we have $\langle x, d_1 \rangle = 0$ as well. That is, we have

$$\mathcal{K} \cap \{d_1^*\}^\perp \subseteq \mathcal{K} \cap \{d_1\}^\perp. \tag{12}$$

(Note that this shows that if $d_1 \notin \mathcal{K}^\perp$ then $d_1^* \notin \mathcal{K}^\perp$ as well.) Since taking the dual cone inverts the containment, we have

$$(\mathcal{K} \cap \{d_1^*\}^\perp \cap \cdots \cap \{d_i\}^\perp)^* \supseteq (\mathcal{K} \cap \{d_1\}^\perp \cap \cdots \cap \{d_i\}^\perp)^*,$$

for every i. Therefore, $(d_1^*, \ldots, d_\ell)$ is still a valid sequence of reducing directions for (D) and the corresponding chain of faces still ends in the minimal face, due to (12). Likewise, we substitute d_2 by d_2^* following (11) with $\mathcal{F}_2 = \mathcal{K} \cap \{d_1^*\}^\perp$ and proceed inductively. This shows that selecting according to (11) does indeed produce the least number of directions. □

We remark that the argument that leads to (12) also shows that if d_1 was already chosen according to (11), we would have in fact $\mathcal{K} \cap \{d_1^*\}^\perp = \mathcal{K} \cap \{d_1\}^\perp$. So that if we use the choice in (11) the resulting chain of faces is unique even if the directions themselves are not.

For some cases, we can expect to implement the choice in (11). If (D) and (P) are both strongly feasible and $\mathcal{K} = \mathcal{S}^n_+$, then it is known that the central path converges to a solution that is a relative interior point of the set of optimal solutions [6] and the facial reduction approach in [23] uses this fact in an essential way. Therefore, the choice in (11) might be implementable in the context of interior point methods although it is not known whether for other algorithms, say augmented Lagrangian methods, a similar property holds. Still, as interior point methods are very revelant to conic linear programming, one of the referees prompted us to prove the following.

Proposition 14. *Let* $e \in \mathcal{K}, e^* \in \mathcal{K}$ *and let* Ω *denote the optimal solution set of* $(P_\mathcal{K})$*. Let* $(x^*, t^*, w^*) \in \operatorname{ri} \Omega$*. If* $t^* = w^* = 0$*, then*

$$x^* \in \operatorname{ri}(\mathcal{K}^* \cap \ker \mathcal{A} \cap \{c\}^\perp).$$

[3] There is a minor incompatibility between the two definitions. Sturm considered that a problem with $\mathcal{F}^s_{\mathrm{D}} = \{0\}$ has singularity degree 0, see Step 1 in Procedure 1 in [26]. According to the definition in [12] and our own, such a problem would have singularity degree 1.

Proof. Let P_x be the linear map that takes $(x, t, w) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$ to x. Since at optimality we have $t^* = w^* = 0^*$, Equation (2) implies that we have

$$\Omega_x = \mathcal{K}^* \cap \ker \mathcal{A} \cap \{c\}^\perp \cap H,$$

where $\Omega_x = P_x(\Omega)$ and $H = \{x \in \mathbb{R}^n \mid \langle e, x\rangle = 1\}$. As P_x is linear, we have $P_x(\operatorname{ri}\Omega) = \operatorname{ri}\Omega_x$, see Theorem 6.6 in [25]. Therefore, $(x^*, t^*, w^*) \in \operatorname{ri}\Omega$ implies $x^* \in \operatorname{ri}\Omega_x$.

Let $C = \mathcal{K}^* \cap \ker \mathcal{A} \cap \{c\}^\perp$, so that $\Omega_x = C \cap H$. Note that the proposition will be proved if we show that $\operatorname{ri}\Omega_x = (\operatorname{ri}(C)) \cap H$.

First, observe that since H is an affine space, we have $\operatorname{ri} H = H$. Then, by Theorem 6.5 in [25], a sufficient condition for $(\operatorname{ri}(C)) \cap H = \operatorname{ri}(C \cap H)$ to hold is that $(\operatorname{ri}(C)) \cap H \neq \emptyset$. We will now construct a point in $(\operatorname{ri}(C)) \cap H$.

Let $z \in \operatorname{ri}(C)$. Note that $x^* \in C$ as well, so there is $\alpha > 1$ such that $\alpha z + (1-\alpha)x^* \in C$, by Theorem 6.4 in [25]. Then, $\langle e, \alpha z + (1-\alpha)x^*\rangle \geq 0$ together with $(1-\alpha) < 0$ and $\langle e, x^*\rangle = 1$ implies that $\langle e, z\rangle > 0$. Therefore, $\frac{z}{\langle e, z\rangle} \in (\operatorname{ri}(C)) \cap H$. This shows that $\operatorname{ri}\Omega_x = (\operatorname{ri}(C)) \cap H$. □

6 Applications of FRA-Poly

In this section, we discuss applications of FRA-Poly. In the first one, we sharpen a result proven by Liu and Pataki [12] on the geometry of weakly infeasible problems. In the second, we show that the singularity degree of problems over the doubly nonnegative cone is at most n.

As mentioned before, the singularity degree only depends on $c, \mathcal{A}$ and $\mathcal{K}$. Finding the minimal face $\mathcal{F}^D_{\min}$ ensures that no matter which b we select, as long as θ_D is finite, there will be zero duality gap and primal attainment. This suggests the following definition that also depends on b and, thus, produce a less conservative quantity.

Definition 15 (Distance to strong duality)**.** *The distance to strong duality $d_{str}(D)$ is the minimum number of facial reduction steps (at (D)) needed to ensure $\theta_{\hat{P}} = \theta_D$, where $(\hat{P})$ is the problem $\inf\{\langle c, x\rangle \mid \mathcal{A}x = b, x \in \mathcal{F}^*_{\ell+1}\}$ and $\mathcal{F}_{\ell+1}$ is a face obtained after a sequence of ℓ facial reduction steps. If $-\infty < \theta_D < +\infty$, we also require attainment of $\theta_{\hat{P}}$.*

Similarly, we define $d_{str}(P)$ as the minimum number of facial reduction steps needed to ensure that $\theta_P = \theta_{\hat{D}}$ and that $\theta_{\hat{D}}$ is attained when $-\infty < \theta_P < +\infty$, where $(\hat{D})$ is the problem in dual standard form arising after some sequence of facial reduction steps is done at (P).

Clearly, we have $d_{\mathrm{str}}(D) \leq d(D)$. However, since Phase 1 of FRA-Poly restores strong duality in the sense of Definition 15, we obtain the nontrivial bound $d_{\mathrm{str}}(D) \leq \sum_{i=1}^r \ell_{\mathrm{poly}}(\mathcal{K}^i)$.

6.1 Weak infeasibility

Let $\mathcal{V}$ denote the affine space $c + \operatorname{range}\mathcal{A}^*$ and let the tuple $(\mathcal{V}, \mathcal{K})$ denote the feasibility problem of seeking an element in the intersection $\mathcal{V} \cap \mathcal{K} = \mathcal{F}^s_{\mathrm{D}}$. In [14], we showed that if $\mathcal{K} = \mathcal{S}^n_+$ and (D) is weakly infeasible, then there is an affine subspace $\mathcal{V}'$ contained in $\mathcal{V}$ of dimension at most $n-1$ such that $(\mathcal{V}', \mathcal{K})$ is also weakly infeasible. This can be interpreted as saying that "we need at most $n-1$ directions to approach the positive semidefinite cone". In [12], Liu and Pataki generalized this result and proved that those affine spaces always exist and $\ell_{\mathcal{K}^*} - 1$ is an upper bound for the dimension of $\mathcal{V}'$, see Corollary 1 therein. We proved a bound of r for the direct product of r Lorentz cones [15], which is tighter than the one in [12]. Here we will refine these results. Consider the following pair of problems.

$$\begin{array}{lll} \inf\limits_{x} & \langle c, x\rangle & \qquad (\mathrm{P_{feas}}) \\ \text{subject to} & \mathcal{A}x = 0 & \\ & x \in \mathcal{K}^* & \end{array} \qquad\qquad \begin{array}{lll} \sup\limits_{y} & 0 & \qquad (\mathrm{D_{feas}}) \\ \text{subject to} & c - \mathcal{A}^*y \in \mathcal{K}. & \end{array}$$

Recall that strong infeasibility of (D) is equivalent to the existence of x such that $x \in \mathcal{K}^* \cap \ker \mathcal{A}$ and $\langle c, x\rangle < 0$, see Lemma 5 in [17]. Therefore, $\theta_{\mathrm{P_{feas}}} = -\infty$ if and only if (D) is strongly infeasible. It follows that (D) is weakly infeasible if and only if $\theta_{\mathrm{P_{feas}}}$ is zero and $\theta_{\mathrm{D_{feas}}} = -\infty$.

When $\theta_{\mathrm{P_{feas}}} = 0$ and we restore strong duality to ($\mathrm{P_{feas}}$) in the sense of Definition 15, a feasible solution will appear at the dual side. Even if that solution is not feasible for the original problem (D), it will give us some information about (D) and this is the motivation behind Theorem 17 below.

We first need an auxiliary result that shows that if (D) is strongly infeasible and we try to regularize ($\mathrm{P_{feas}}$), then ($\mathrm{D_{feas}}$) (and, therefore, (D)) will stay strongly infeasible.

Lemma 16. *Let d be a reducing direction for ($\mathrm{P_{feas}}$), i.e., $d \in (\operatorname{range} \mathcal{A}^*) \cap \mathcal{K}$. Let $\hat{\mathcal{K}} = (\mathcal{K}^* \cap \{d\}^\perp)^*$. Then ($\mathrm{D_{feas}}$) is strongly infeasible if and only if ($\hat{D}_{feas}$) is strongly infeasible, where ($\hat{D}_{feas}$) is the problem with $\hat{\mathcal{K}}$ in place of $\mathcal{K}$.*

Proof. ($\Leftarrow$) This part is clear, since $\mathcal{K} \subseteq \hat{\mathcal{K}}$.

($\Rightarrow$) Strong infeasibility of ($\mathrm{D_{feas}}$) is equivalent to the existence of x such that $x \in \mathcal{K}^* \cap \ker \mathcal{A}$ and $\langle c, x\rangle < 0$. Since $d \in \operatorname{range} \mathcal{A}^*$, we have $\langle x, d\rangle = 0$. By the same principle, x induces strong infeasibility for ($\hat{\mathrm{D}}_{\mathrm{feas}}$) as well. ☐

Theorem 17.

(i) (D) is not strongly infeasible if and only if there are:

(a) a sequence of reducing directions $\{d_1, \ldots, d_\ell\}$ for ($\mathrm{P_{feas}}$) restoring strong duality in the sense of Definition 15 with $\ell = d_{str}(P_{feas})$ and

(b) $\hat{y}$ such that $c - \mathcal{A}^*\hat{y} \in (\mathcal{K}^* \cap \{d_1\}^\perp \cap \cdots \cap \{d_\ell\}^\perp)^*$.

(ii) If (D) is not strongly infeasible, there is an affine subspace $\mathcal{V}' \subseteq c - \operatorname{range} \mathcal{A}^*$ such that $(\mathcal{V}', \mathcal{K})$ is not strongly infeasible and the dimension of $\mathcal{V}'$ satisfies

$$\dim (\mathcal{V}') \leq d_{str}(P') \leq \sum_{i=1}^{r} \ell_{poly}((\mathcal{K}^i)^*).$$

In particular, if (D) is weakly infeasible, then $(\mathcal{V}', \mathcal{K})$ is weakly infeasible.

Proof. (i) ($\Rightarrow$) Due to the assumption that (D) is not strongly infeasible, we have $\theta_{\mathrm{P_{feas}}} = 0$. Now, let $\{d_1, \ldots, d_\ell\}$ be a sequence of reducing directions for ($\mathrm{P_{feas}}$) that restores strong duality in the sense of Definition 15 with $\ell = d_{\mathrm{str}}(\mathrm{P_{feas}})$. Let $\hat{\mathcal{F}}_{\ell+1} = \mathcal{K}^* \cap \{d_1\}^\perp \cap \cdots \cap \{d_\ell\}^\perp$.

Now, ($\mathrm{P_{feas}}$) shares the same feasible region with the problem

$$\inf \{\langle c, x\rangle \mid \mathcal{A}x = 0, x \in \hat{\mathcal{F}}_{\ell+1}\}. \tag{$\hat{\mathrm{P}}$}$$

Since facial reduction preserves the optimal value, we have $\theta_{\hat{P}} = \theta_{\mathrm{P_{feas}}} = 0$. Because the reducing directions restore strong duality, we have $\theta_{\hat{D}} = \{0 \mid c - \mathcal{A}^* y \in \hat{\mathcal{F}}^*_{\ell+1}\}$ and $\theta_{\hat{D}}$ is attained. In particular, there is $\hat{y}$ such that $c - \mathcal{A}^*\hat{y} \in \mathcal{F}^*_{\ell+1}$.

($\Leftarrow$) By (a), $\theta_{\mathrm{P_{feas}}} = \theta_{\hat{D}} = \{0 \mid c - \mathcal{A}^* y \in \hat{\mathcal{F}}^*_{\ell+1}\}$, where $\hat{\mathcal{F}}_{\ell+1} = \mathcal{K}^* \cap \{d_1\}^\perp \cap \cdots \cap \{d_\ell\}^\perp$. Due to (b), $\theta_{\mathrm{P_{feas}}} = \theta_{\hat{D}} = 0$. Therefore, (D) is not strongly infeasible.

(ii) Let $\{d_1, \ldots, d_\ell\}$, $\hat{y}$ and $\hat{\mathcal{F}}_{\ell+1} = \mathcal{K}^* \cap \{d_1\}^\perp \cap \cdots \cap \{d_\ell\}^\perp$ be as in item (i). Let $\mathcal{V}'$ be the affine space $\hat{s} + \mathcal{L}'$, where $\mathcal{L}'$ is spanned by the directions $\{d_1, \ldots, d_\ell\}$ and $\hat{s} = c - \mathcal{A}^*\hat{y}$. Since $\ell = d_{\mathrm{str}}(\mathrm{P_{feas}})$, we have $\dim \mathcal{V}' \leq d_{\mathrm{str}}(\mathrm{P_{feas}})$. Suppose for the sake of contradiction that $(\mathcal{V}', \mathcal{K})$ is strongly infeasible. Then, we can use $\{d_1, \ldots, d_\ell\}$ as reducing directions for $\inf\{\langle \hat{s}, x\rangle \mid x \in \mathcal{L}'^\perp, x \in \mathcal{K}^*\}$. However, Lemma 16 implies that $\sup\{0 \mid s \in (\hat{s} + \mathcal{L}') \cap \hat{\mathcal{F}}^*_{\ell+1}\}$ is strongly infeasible, which contradicts the fact that $\hat{s}$ is a feasible solution.

Since the number steps required for Phase 1 of FRA-Poly gives an upper bound for $d_{\mathrm{str}}(\mathrm{P_{feas}})$, we obtain $d_{\mathrm{str}}(\mathrm{P_{feas}}) \leq \sum_{i=1}^{r} \ell_{\mathrm{poly}}((\mathcal{K}^i)^*)$.

When (D) is weakly infeasible, since $\mathcal{V}' \subseteq c - \operatorname{range} \mathcal{A}^*$ and $(\mathcal{V}', \mathcal{K})$ is not strongly infeasible, it must be the case that $(\mathcal{V}', \mathcal{K})$ is weakly infeasible.

□

Due to Theorem 11, the bound in Theorem 17 will usually compare favorably to $\ell_{\mathcal{K}^*} - 1$. Moreover, it also recovers the bounds described in [14, 15].

6.2 An application to the intersection of cones

In this subsection, we discuss the case where $\mathcal{K} = \mathcal{K}^1 \cap \mathcal{K}^2$. We can rewrite (D) as a problem over $\mathcal{K}^1 \times \mathcal{K}^2$ by duplicating the entries.

$$\begin{aligned} \inf_{x^1,x^2} \quad & \langle c, x^1 + x^2 \rangle & (P_{\mathrm{dup}}) \\ \text{subject to} \quad & \mathcal{A}(x^1 + x^2) = 0 \\ & (x^1, x^2) \in {\mathcal{K}^1}^* \times {\mathcal{K}^2}^* \end{aligned}$$

$$\begin{aligned} \sup_{y} \quad & \langle b, y \rangle & (D_{\mathrm{dup}}) \\ \text{subject to} \quad & (c - \mathcal{A}^*y, c - \mathcal{A}^*y) \in \mathcal{K}^1 \times \mathcal{K}^2 \end{aligned}$$

If we apply FRA-Poly to (D_{dup}), we will obtain a face $\mathcal{F}^1 \times \mathcal{F}^2$ of $\mathcal{K}^1 \times \mathcal{K}^2$, so that $\mathcal{F}^1 \cap \mathcal{F}^2$ will be a face of $\mathcal{K}$ containing $\mathcal{F}^s_{\mathrm{D}}$. Doing facial reduction using the formulation (D_{dup}) might be more convenient, since we need to search for reducing directions in $(\mathcal{K}^1)^* \times (\mathcal{K}^2)^*$ instead of $\operatorname{cl}((\mathcal{K}^1)^* + (\mathcal{K}^2))^*$ and deciding membership in $(\mathcal{K}^1)^* \times (\mathcal{K}^2)^*$ could be more straightforward than doing the same for $\operatorname{cl}((\mathcal{K}^1)^* + (\mathcal{K}^2))^*$.

Before we proceed we need an auxiliary result. If $\mathcal{K} = \mathcal{K}^1 \cap \mathcal{K}^2$, it is always true that the intersection of a face of $\mathcal{K}^1$ with a face of $\mathcal{K}^2$ results in a face of $\mathcal{K}$. However, it is not obvious that every face of $\mathcal{K}$ arises as an intersection of faces of $\mathcal{K}^1$ and $\mathcal{K}^2$, so we remark that as a proposition although it is probably a well-known result.

Proposition 18. *Let $\mathcal{F}$ be a nonempty face of $\mathcal{K}^1 \cap \mathcal{K}^2$. Let $\mathcal{F}^1$ and $\mathcal{F}^2$ be the minimal faces of $\mathcal{K}^1$ and $\mathcal{K}^2$, respectively, containing $\mathcal{F}$. Then $\mathcal{F} = \mathcal{F}^1 \cap \mathcal{F}^2$ and $\mathcal{F}^* = (\mathcal{F}^1)^* + (\mathcal{F}^2)^*$.*

Proof. Since $\mathcal{F}^1 \cap \mathcal{F}^2$ is a face of $\mathcal{K}^1 \cap \mathcal{K}^2$, in order to prove that $\mathcal{F}^1 \cap \mathcal{F}^2 = \mathcal{F}$ it is enough to show that their relative interiors intersect, which we will do next. By the choice of $\mathcal{F}^1$ and $\mathcal{F}^2$, we have $\operatorname{ri}(\mathcal{F}) \subseteq \operatorname{ri}(\mathcal{F}^1)$ and $\operatorname{ri}(\mathcal{F}) \subseteq \operatorname{ri}(\mathcal{F}^2)$, see item (*iii*) of Proposition 2.2 in [21][4]. In particular, this implies that $\operatorname{ri}(\mathcal{F}^1) \cap \operatorname{ri}(\mathcal{F}^2) \neq \emptyset$. Therefore, $\operatorname{ri}(\mathcal{F}^1 \cap \mathcal{F}^2) = \operatorname{ri}(\mathcal{F}^1) \cap \operatorname{ri}(\mathcal{F}^2)$, by Theorem 6.5 in [25]. We conclude that $\operatorname{ri}(\mathcal{F}) \cap \operatorname{ri}(\mathcal{F}^1 \cap \mathcal{F}^2) = \operatorname{ri}(\mathcal{F}) \cap \operatorname{ri}(\mathcal{F}^1) \cap \operatorname{ri}(\mathcal{F}^2) \neq \emptyset$. It follows that $\mathcal{F} = \mathcal{F}^1 \cap \mathcal{F}^2$.

Because $\operatorname{ri}(\mathcal{F}^1) \cap \operatorname{ri}(\mathcal{F}^2) \neq \emptyset$, the sum $(\mathcal{F}^1)^* + (\mathcal{F}^2)^*$ is closed (see Corollary 16.4.2 in [25]), so that $\mathcal{F}^* = \operatorname{cl}((\mathcal{F}^1)^* + (\mathcal{F}^2)^*) = (\mathcal{F}^1)^* + (\mathcal{F}^2)^*$. □

In what follows, we will need two well-known equalities. Let $\mathcal{K}^1, \mathcal{K}^2$ be closed convex cones and $d^1 \in (\mathcal{K}^1)^*, d^2 \in (\mathcal{K}^2)^*$. Then:

$$(\mathcal{K}^1 \times \mathcal{K}^2) \cap \{(d^1, d^2)\}^\perp = \left(\mathcal{K}^1 \cap \{d^1\}^\perp\right) \times \left(\mathcal{K}^2 \cap \{d^2\}^\perp\right) \tag{13}$$

$$\mathcal{K}^1 \cap \mathcal{K}^2 \cap \{d^1_1\}^\perp \cap \{d^2_1\}^\perp = \mathcal{K}^1 \cap \mathcal{K}^2 \cap \{d^1_1 + d^2_1\}^\perp. \tag{14}$$

Theorem 19. *Let $\mathcal{K} = \mathcal{K}^1 \cap \mathcal{K}^2$.*

(i) Let $\mathcal{F}^{(D_{\mathrm{dup}})}_{\mathrm{min}} = \mathcal{F}^1 \times \mathcal{F}^2$ be the minimal face of $\mathcal{K}^1 \times \mathcal{K}^2$ containing the set of feasible slacks of (D_{dup}). Then, $\mathcal{F}^D_{\mathrm{min}} = \mathcal{F}^1 \cap \mathcal{F}^2$.

(ii) The singularity degree and the distance to strong duality of (D) satisfy

$$\begin{aligned} d(D) \leq & d(D_{dup}) \leq & 1 + \ell_{poly}(\mathcal{K}^1) + \ell_{poly}(\mathcal{K}^2) \\ d_{str}(D) \leq & d_{str}(D_{dup}) \leq & \ell_{poly}(\mathcal{K}^1) + \ell_{poly}(\mathcal{K}^2) \end{aligned}$$

[4] Proposition 2.2 ensures that $\operatorname{ri}(\mathcal{F})$ intersects $\operatorname{ri}(\mathcal{F}^1)$. However, given that $\mathcal{F}_1$ contains $\mathcal{F}$, this is enough for the containment $\operatorname{ri}(\mathcal{F}) \subseteq \operatorname{ri}(\mathcal{F}^1)$, due to Corollary 6.5.2 in [25]. The same goes for $\mathcal{F}_2$.

Proof. (i) $\mathcal{F}^1$ must be the minimal face of $\mathcal{K}^1$ containing $\mathcal{F}_{\mathrm{D}}^s = \{c - \mathcal{A}^* y \in \mathcal{K}\}$. Otherwise, if some proper face $\widetilde{\mathcal{F}}$ of $\mathcal{F}^1$ is minimal, then $\widetilde{\mathcal{F}} \times \mathcal{F}^2$ contains the feasible slacks of (D_{dup}), which contradicts the minimality of $\hat{\mathcal{F}}$. The same must hold for $\mathcal{F}^2$. Then Proposition 18 implies $\mathcal{F}_{\min}^D = \mathcal{F}^1 \cap \mathcal{F}^2$.

(ii) To prove the bounds we first show that given ℓ reducing directions for (D_{dup}) we can also construct ℓ reducing directions for (D) and that there are relations between the faces defined by both sets of directions.

Suppose that $\{(d_1^1, d_1^2), \ldots, (d_\ell^1, d_\ell^2)\}$ are reducing directions for (D_{dup}). Define

$$\mathcal{F}_1^1 \times \mathcal{F}_1^2 := \mathcal{K}^1 \times \mathcal{K}^2 \tag{15}$$

$$\mathcal{F}_{i+1}^1 \times \mathcal{F}_{i+1}^2 := (\mathcal{F}_i^1 \times \mathcal{F}_i^2) \cap \{(d_i^1, d_i^2)\}^\perp \qquad 1 < i \leq \ell. \tag{16}$$

We will show by induction on i that $\{d_1^1 + d_1^2, \ldots, d_\ell^1 + d_\ell^2\}$ are reducing directions for (D) and that

$$\mathcal{F}_{i+1}^1 \cap \mathcal{F}_{i+1}^2 = \mathcal{K}^1 \cap \mathcal{K}^2 \cap \{d_1^1 + d_1^2\}^\perp \cap \cdots \cap \{d_i^1 + d_i^2\}^\perp. \tag{17}$$

If $\boxed{i = 1}$, because (d_1^1, d_1^2) is a reducing direction for (D_{dup}), it satisfies

$$\mathcal{A}(d_1^1 + d_1^2) = 0, \qquad \langle c, d_1^1 + d_1^2 \rangle \leq 0, \qquad (d_1^1, d_1^2) \in {\mathcal{K}^1}^* \times {\mathcal{K}^2}^*. \tag{18}$$

Our goal is to show that $d_1^1 + d_2^1$ is a reducing direction for (D), that is

$$\mathcal{A}(d_1^1 + d_1^2) = 0, \qquad \langle c, d_1^1 + d_1^2 \rangle \leq 0, \qquad d_1^1 + d_1^2 \in \mathrm{cl}\,({\mathcal{K}^1}^* + {\mathcal{K}^2}^*) \tag{19}$$

and that (17) holds when $i = 1$. Note that (19) follows directly from (18). From (15), (16) and (13) we have:

$$\mathcal{F}_2^1 \times \mathcal{F}_2^2 = (\mathcal{K}^1 \times \mathcal{K}^2) \cap \{(d_1^1, d_1^2)\}^\perp = \left(\mathcal{K}^1 \cap \{d_1^1\}^\perp\right) \times \left(\mathcal{K}^2 \cap \{d_1^2\}^\perp\right).$$

By (14), we conclude that $\mathcal{F}_2^1 \cap \mathcal{F}_2^2 = \mathcal{K}^1 \cap \mathcal{K}^2 \cap \{d_1^1\}^\perp \cap \{d_1^2\}^\perp = \mathcal{K}^1 \cap \mathcal{K}^2 \cap \{d_1^1 + d_1^2\}^\perp$.

If $\boxed{i > 1}$, by induction, we have that (17) holds up to $i-1$. By hypothesis, (d_i^1, d_i^2) is a reducing direction, so that

$$\mathcal{A}(d_i^1 + d_i^2) = 0, \qquad \langle c, d_i^1 + d_i^2 \rangle \leq 0, \qquad (d_i^1, d_i^2) \in {\mathcal{F}_i^1}^* \times {\mathcal{F}_i^2}^*. \tag{20}$$

We have to show that (17) holds and that

$$\mathcal{A}(d_i^1 + d_i^2) = 0, \qquad \langle c, d_i^1 + d_i^2 \rangle \leq 0, \qquad d_i^1 + d_i^2 \in \mathrm{cl}\,({\mathcal{F}_i^1}^* + {\mathcal{F}_i^2}^*). \tag{21}$$

As before, (21) follows directly from (20). From (16), (20) and (13) we obtain

$$\mathcal{F}_{i+1}^1 \times \mathcal{F}_{i+1}^2 = (\mathcal{F}_i^1 \cap \{d_i^1\}^\perp) \times (\mathcal{F}_i^2 \cap \{d_i^2\}^\perp).$$

We conclude that:

$$\begin{aligned} \mathcal{F}_{i+1}^1 \cap \mathcal{F}_{i+1}^2 &= \mathcal{F}_i^1 \cap \mathcal{F}_i^2 \cap \{d_i^1 + d_i^2\}^\perp \\ &= \mathcal{K}^1 \cap \mathcal{K}^2 \cap \{d_1^1 + d_1^2\}^\perp \cap \cdots \cap \{d_i^1 + d_i^2\}^\perp, \end{aligned}$$

where the first equality follows from (14) and the second follows from the induction hypothesis. This concludes the induction.

Finally, if $\mathcal{F}_{\min}^{(D_{\mathrm{dup}})} = \mathcal{F}_{\ell+1}^1 \times \mathcal{F}_{\ell+1}^2$, then (i) implies $\mathcal{F}_{\min}^D = \mathcal{F}_{\ell+1}^1 \cap \mathcal{F}_{\ell+1}^2$. Similarly, if substituting $\mathcal{K}^1 \times \mathcal{K}^2$ in (D_{dup}) for $\mathcal{F}_{\ell+1}^1 \times \mathcal{F}_{\ell+1}^2$ restores strong duality, then the same holds for (D) if we substitute $\mathcal{K}$ for $\mathcal{F}_{\ell+1}^1 \cap \mathcal{F}_{\ell+1}^2$. This shows that $d(D) \leq d(D_{\mathrm{dup}})$ and $d_{\mathrm{str}}(D) \leq d_{\mathrm{str}}(D_{\mathrm{dup}})$, respectively. The other bounds follow from Propositions 8 and 9. □

We now consider the case where $\mathcal{K}$ is the doubly nonnegative cone $\mathcal{D}^n = \mathcal{S}_+^n \cap \mathcal{N}^n$, where $\mathcal{N}^n$ is the cone of $n \times n$ symmetric matrices with nonnegative entries. This cone is important because it can be used as a relatively tractable relaxation for the cone of completely positive matrices, see [30, 11, 2].

Corollary 20. *When $\mathcal{K} = \mathcal{D}^n$, we have $d(D) \leq n$ and $d_{str}(D) \leq n-1$.*

Proof. Follows from Theorem 19 since $\ell_{\text{poly}}(\mathcal{S}^n_+) = n-1$ and $\ell_{\text{poly}}(\mathcal{N}^n) = 0$. $\square$

We will compare the bound in Corollary 20 with the one predicted by the classical FRA. To do that, we need to compute $\ell_{\mathcal{D}^n}$.

Proposition 21. *The longest chain of nonempty faces in $\mathcal{D}^n$ has length $\frac{n(n+1)}{2}+1$, which is the maximum possible for a cone contained in $\mathcal{S}^n$.*

Proof. Maximality follows from the fact that the dimension of $\mathcal{S}^n$ is $\frac{n(n+1)}{2}$ and that if we have two faces such that $\mathcal{F} \subsetneq \hat{\mathcal{F}}$ then $\dim(\mathcal{F}) < \dim(\hat{\mathcal{F}})$.

Let $\mathcal{G}$ be any set of tuples (i,j) with $i,j \in \{1,\ldots,n\}$ and let $\mathcal{N}^n(\mathcal{G})$ be the face of $\mathcal{N}^n$ which corresponds to the matrices x such that the only entries $x_{i,j}$ that are allowed to be nonzero are the ones for which either $(i,j) \in \mathcal{G}$ or $(j,i) \in \mathcal{G}$. We will first define two chains of faces of $\mathcal{N}^n$. First, let $\mathcal{G}_0 = \emptyset$ and define $\mathcal{G}_i = \mathcal{G}_{i-1} \cup \{(i,i)\}$ for $i \in \{1,\ldots,n\}$. We now consider the following construction written in pseudocode.

$k \leftarrow 1$, $\mathcal{H}_0 \leftarrow \mathcal{G}_n$

For $i \leftarrow 1, i \leq n$ **do**

For $j \leftarrow 1$, $j < i$ **do**

$\mathcal{H}_k \leftarrow \mathcal{H}_{k-1} \cup \{i,j\}$
$k \leftarrow k+1$
$j \leftarrow j+1.$

$i \leftarrow i+1.$

The idea is to add one non-diagonal entry per iteration, so that $\mathcal{N}^n(\mathcal{H}_k) \subsetneq \mathcal{N}^n(\mathcal{H}_{k+1})$. First $(2,1)$ will be added, then $(3,1),(3,2)$ and so on. We have

$$\mathcal{S}^n_+ \cap \mathcal{N}^n(\mathcal{G}_0) \subsetneq \ldots \subsetneq \mathcal{S}^n_+ \cap \mathcal{N}^n(\mathcal{G}_n) \subsetneq \mathcal{S}^n_+ \cap \mathcal{N}^n(\mathcal{H}_1) \subsetneq \ldots \subsetneq \mathcal{S}^n_+ \cap \mathcal{N}^n(\mathcal{H}_{\frac{n(n-1)}{2}})$$

and all inclusions are indeed strict. The first n inclusions are strict because $\mathcal{S}^n_+ \cap \mathcal{N}^n(\mathcal{G}_i) = \mathcal{N}^n(\mathcal{G}_i)$ and it is clear that $\mathcal{N}^n(\mathcal{G}_i) \subsetneq \mathcal{N}^n(\mathcal{G}_{i+1})$. Now, let $\mathcal{I}_n$ denote the $n \times n$ identity matrix. If $k > 0$ and $x \in \operatorname{ri} \mathcal{N}^n(\mathcal{H}_k)$ then $x_{i,j} > 0$ for some (i,j) entry such that neither (i,j) nor (j,i) belong to $\mathcal{H}_{k-1}$. For $\alpha > 0$ sufficiently large, we have $x + \alpha\mathcal{I}_n \in \mathcal{S}^n_+ \cap \mathcal{N}^n(\mathcal{H}_k)$ and $x + \alpha\mathcal{I}_n \notin \mathcal{S}^n_+ \cap \mathcal{N}^n(\mathcal{H}_{k-1})$. This shows the remainder of the containments and concludes the proof, since the chain has length $\frac{n(n+1)}{2}+1$. $\square$

For feasible problems, the classical FRA analysis gives either the bound $\ell_{\mathcal{D}^n} - 1 = \frac{n(n+1)}{2}$ or, using Theorem 19, the bound $\ell_{\mathcal{S}^n_+} - 1 + \ell_{\mathcal{N}^n} - 1 = n + \frac{n(n+1)}{2}$. Both bounds are quadratic in n in opposition to the linear bound obtained in Corollary 20.

Acknowledgements

We thank the referees and the associate editor for the detailed feedback we were given. The article benefited greatly from it and, in particular, the 2nd referee motivated the discussion in Section 5.4 and suggested the first instance in Appendix B, for which we are grateful.

References

[1] Mustafa Akgül. On polyhedral extension of some LP theorems. *Mathematical Programming*, 30(1):112–120, 1984.

[2] Naohiko Arima, Sunyoung Kim, Masakazu Kojima, and Kim-Chuan Toh. Lagrangian-conic relaxations, part i: A unified framework and its applications to quadratic optimization problems. Technical Report B-475, Tokyo Institute of Technology, 2014.

[3] Jonathan M. Borwein and Henry Wolkowicz. Facial reduction for a cone-convex programming problem. *Journal of the Australian Mathematical Society (Series A)*, 30(03):369–380, 1981.

[4] Jonathan M. Borwein and Henry Wolkowicz. Regularizing the abstract convex program. *Journal of Mathematical Analysis and Applications*, 83(2):495 – 530, 1981.

[5] Yuen-Lam Cheung, Simon Schurr, and Henry Wolkowicz. Preprocessing and regularization for degenerate semidefinite programs. In *Computational and Analytical Mathematics*, volume 50 of *Springer Proceedings in Mathematics & Statistics*, pages 251–303. Springer New York, 2013.

[6] E. de Klerk, C. Roos, and T. Terlaky. Initialization in semidefinite programming via a self-dual skew-symmetric embedding. *Operations Research Letters*, 20(5):213–221, 1997.

[7] Dmitriy Drusvyatskiy, Gábor Pataki, and Henry Wolkowicz. Coordinate shadows of semidefinite and euclidean distance matrices. *SIAM Journal on Optimization*, 25(2):1160–1178, 2015.

[8] Robert M. Freund, Robin Roundy, and Michael Todd. Identifying the set of always-active constraints in a system of linear inequalities by a single linear program. Working papers 1674-85., Massachusetts Institute of Technology (MIT), Sloan School of Management, 1985. URL: http://EconPapers.repec.org/RePEc:mit:sloanp:2111.

[9] Henrik A. Friberg. A relaxed-certificate facial reduction algorithm based on subspace intersection. *Operations Research Letters*, 44(6):718 – 722, 2016.

[10] Masaru Ito and Bruno F. Lourenço. A bound on the Carathéodory number. *Linear Algebra and its Applications*, 532:347 – 363, 2017.

[11] Sunyoung Kim, Masakazu Kojima, and Kim-Chuan Toh. A Lagrangian–DNN relaxation: a fast method for computing tight lower bounds for a class of quadratic optimization problems. *Mathematical Programming*, 156(1):161–187, Mar 2016.

[12] Minghui Liu and Gábor Pataki. Exact duals and short certificates of infeasibility and weak infeasibility in conic linear programming. *Mathematical Programming*, 2017. doi:10.1007/s10107-017-1136-5.

[13] Bruno F. Lourenço, Masakazu Muramatsu, and Takashi Tsuchiya. Solving SDP completely with an interior point oracle, July 2015. arXiv:1507.08065.

[14] Bruno F. Lourenço, Masakazu Muramatsu, and Takashi Tsuchiya. A structural geometrical analysis of weakly infeasible SDPs. *Journal of the Operations Research Society of Japan*, 59(3):241–257, 2016.

[15] Bruno F. Lourenço, Masakazu Muramatsu, and Takashi Tsuchiya. Weak infeasibility in second order cone programming. *Optimization Letters*, 10(8):1743–1755, 2016.

[16] Z. Luo and J. F. Sturm. Error analysis. In Henry Wolkowicz, Romesh Saigal, and Lieven Vandenberghe, editors, *Handbook of semidefinite programming: theory, algorithms, and applications*. Kluwer Academic Publishers, 2000.

[17] Z. Luo, J. F. Sturm, and S. Zhang. Duality results for conic convex programming. Technical report, Econometric Institute, Erasmus University Rotterdam, The Netherlands, 1997.

[18] L. McLinden. Polyhedral extensions of some theorems of linear programming. *Mathematical Programming*, 24(1):162–176, 1982.

[19] Sanjay Mehrotra and Yinyu Ye. Finding an interior point in the optimal face of linear programs. *Mathematical Programming*, 62(1-3):497–515, 1993.

[20] Y. Nesterov, M. J. Todd, and Y. Ye. Infeasible-start primal-dual methods and infeasibility detectors for nonlinear programming problems. *Mathematical Programming*, 84(2):227–267, 1999. doi:10.1007/s10107980009a.

[21] Gábor Pataki. The geometry of semidefinite programming. In Henry Wolkowicz, Romesh Saigal, and Lieven Vandenberghe, editors, *Handbook of semidefinite programming: theory, algorithms, and applications*. Kluwer Academic Publishers, online version at http://www.unc.edu/~pataki/papers/chapter.pdf, 2000.

[22] Gábor Pataki. Strong duality in conic linear programming: Facial reduction and extended duals. In *Computational and Analytical Mathematics*, volume 50, pages 613–634. Springer New York, 2013.

[23] Frank Permenter, Henrik A. Friberg, and Erling D. Andersen. Solving conic optimization problems via self-dual embedding and facial reduction: A unified approach. *SIAM Journal on Optimization*, 27(3):1257–1282, 2017.

[24] Frank Permenter and Pablo Parrilo. Partial facial reduction: simplified, equivalent SDPs via approximations of the PSD cone. *Mathematical Programming*, Jun 2017. doi:10.1007/s10107-017-1169-9.

[25] R. T. Rockafellar. *Convex Analysis* . Princeton University Press, 1997.

[26] Jos F. Sturm. Error bounds for linear matrix inequalities. *SIAM Journal on Optimization*, 10(4):1228–1248, January 2000. doi:10.1137/S1052623498338606.

[27] Levent Tunçel. *Polyhedral and Semidefinite Programming Methods in Combinatorial Optimization*. American Mathematical Society, Providence, R.I. : Toronto, Ont, November 2010.

[28] Levent Tunçel and Henry Wolkowicz. Strong duality and minimal representations for cone optimization. *Computational Optimization and Applications*, 53(2):619–648, 2012.

[29] Hayato Waki and Masakazu Muramatsu. Facial reduction algorithms for conic optimization problems. *Journal of Optimization Theory and Applications*, 158(1):188–215, 2013.

[30] A. Yoshise and Y. Matsukawa. On optimization over the doubly nonnegative cone. In *Computer-Aided Control System Design (CACSD), 2010 IEEE International Symposium on*, pages 13–18, Sept 2010.

A Partial Polyhedrality and Slater's condition

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, +\infty\}$ be a convex function. We denote the domain of f by $\operatorname{dom} f = \{x \in \mathbb{R}^n \mid f(x) < \infty\}$. If $\operatorname{dom} f \neq \emptyset$ and f is never $-\infty$, then f is said to be proper. Its conjugate will be denoted by f^* and it satisfies $f^*(s) = \sup_x \langle s, x\rangle - f(x)$. If the epigraph of f is a polyhedral set, then f is said to be a polyhedral function. We recall Theorem 20.1 from [25].

Theorem 22 (Rockafellar)**.** *Let $f_1, \ldots, f_m$ be proper convex functions and let $f_{k+1}, \ldots, f_m$ be polyhedral functions. Suppose also that*

$$\operatorname{ri}(\operatorname{dom} f_1) \cap \cdots \cap \operatorname{ri}(\operatorname{dom} f_k) \cap \operatorname{dom} f_{k+1} \cap \cdots \cap \operatorname{dom} f_m \neq \emptyset.$$

Then the following holds:

$$(f_1 + \cdots + f_m)^*(s) = \inf\{f_1^*(s_1) + \cdots + f_m^*(s_m) \mid s_1 + \cdots + s_m = s\},$$

where for each s the infimum is attained whenever it is finite.

Proposition 23. *Let $\mathcal{K} = \mathcal{K}^1 \times \mathcal{K}^2$, where $\mathcal{K}^1 \subseteq \mathbb{R}^{n_1}, \mathcal{K}^2 \subseteq \mathbb{R}^{n_2}$ are closed convex cones such that $\mathcal{K}^2$ is polyhedral.*

(i) If θ_P is finite and (P) *satisfies the PPS condition, then $\theta_P = \theta_D$ and the dual optimal value is attained.*

(ii) If θ_D is finite and (D) *satisfies the PPS condition, then $\theta_P = \theta_D$ and the primal optimal value is attained.*

Proof. We will prove (i) first. Let f_1 be such that $f_1(x) = \langle c, x\rangle$ if $Ax = b$ and $+\infty$ otherwise. Let f_2 be the indicator function of $\mathbb{R}^{n_1} \times (\mathcal{K}^2)^*$ and f_3 be the indicator function of $(\mathcal{K}^1)^* \times \mathbb{R}^{n_2}$. Since there is a primal feasible solution $x = (x_1, x_2)$ such that $x_1 \in \operatorname{ri}(\mathcal{K}^1)^*$, we have that $\operatorname{dom} f_1 \cap \operatorname{dom} f_2 \cap \operatorname{ri}(\operatorname{dom} f_3)$ is nonempty. In addition, f_1 and f_2 are polyhedral functions. Let us now observe that:

$$f_1^*(s) = \begin{cases} \langle b, y\rangle & \text{if there is } y \text{ with } s - c = \mathcal{A}^* y \\ +\infty & \text{otherwise} \end{cases}$$

Note that, due to feasibility, for fixed s, $\langle b, y\rangle$ does not depend on the choice of y, as long as $c + \mathcal{A}^* y = s$. This is because since there is x such that $Ax = b$, we have $\langle b, y\rangle = \langle x, s - c\rangle$. The conjugate f_2^* is the indicator function of $-\{0\} \times \mathcal{K}^2$ and f_3^* is the indicator function of $-\mathcal{K}^1 \times \{0\}$. Applying Theorem 22 with $s = 0$, we have:

$$\begin{aligned}(f_1 + f_2 + f_3)^*(0) &= \inf\left\{\langle b, y\rangle \mid c + \mathcal{A}^* y = s_1, s_1 - (s_3, s_2) = 0, s_2 \in \mathcal{K}^2, s_3 \in \mathcal{K}^1\right\} \\ &= \inf\left\{\langle b, y\rangle \mid c + \mathcal{A}^* y = s_1, s_1 \in \mathcal{K}^1 \times \mathcal{K}^2\right\} \\ &= -\sup\left\{\langle b, y\rangle \mid c - \mathcal{A}^* y = s_1, s_1 \in \mathcal{K}^1 \times \mathcal{K}^2\right\},\end{aligned}$$

where the sup in the last equation is attained. So, there is some dual feasible y such that $(f_1+f_2+f_3)^*(0) = \langle b, y\rangle$. However, using the definition of conjugate, we also have:

$$(f_1+f_2+f_3)^*(0) = -\inf\{\langle c, x\rangle \mid Ax = b, x \in \mathcal{K}^1 \times \mathcal{K}^2\} = -\theta_P.$$

It follows that $\theta_P = \theta_D$ and the dual is attained at y. To prove (ii), let $g_1 = f_1^*$, and let g_2 and g_3 be the indicator functions of $\mathbb{R}^{n_1} \times \mathcal{K}^2$ and $\mathcal{K}^1 \times \mathbb{R}^{n_2}$, respectively. Again, it is enough to compute $(g_1+g_2+g_3)^*(0)$ using both the definition of conjugate function and using Theorem 22. □

B Examples

Example 2. *We will apply FRA-Poly to the following problem.*

$$\begin{aligned} &\textit{find} \quad y \\ &\textit{subject to} \quad (y_1, -y_1) \times \begin{pmatrix} y_1 & y_2 \\ y_2 & y_3 \end{pmatrix} \in \mathbb{R}^2_+ \times \mathcal{S}^2_+. \end{aligned}$$

At the first step we have $\mathcal{F}_1 = \mathbb{R}^2_+ \times \mathcal{S}^2_+$ and we build ($P_\mathcal{K}$) and ($D_\mathcal{K}$) using $e = 0 \times I_2$ and $e^* = (1,1) \times I_2$, where I_2 is the 2×2 identity matrix. Solving ($P_\mathcal{K}$) and ($D_\mathcal{K}$), suppose that we have found the reducing direction

$$d_1 = (1,2) \times \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then $\mathcal{F}_2 = \mathcal{K} \cap \{d_1\}^\perp = \{0\} \times \hat{\mathcal{F}}$, where the matrices in $\hat{\mathcal{F}}$ have zeros in all entries except in the $(2,2)$ entry. Note that $\mathcal{F}_2$ is polyhedral, so when it is time to build ($P_\mathcal{K}$) and ($D_\mathcal{K}$) again, we will take $e = 0$ and since the PPS condition is satisfied, we will have $\theta_{P_\mathcal{K}} > 0$. Note that this is a case where Phase 1 ends at the minimal face and there is no need to proceed further, although we might want to solve the problem in Phase 2 in order to obtain a point in $\mathrm{ri}\,\mathcal{F}^D_{\min}$.

Note that if the first block of d_1 were $(0,1)$ instead of $(1,2)$, then we would have $\mathcal{F}_2 = \mathbb{R}_+ \times \{0\} \times \hat{\mathcal{F}}$, so we would need a Phase 2 iteration to find a reducing direction d_2 such that $\mathcal{F}^D_{\min} = \mathcal{F}_2 \cap \{d_2\}^\perp$. Still, if we are able to implement the choice in (11) we will never need to go for a second direction, since we will always take the most interior direction possible.

Let $\mathcal{K} = \mathcal{Q}^{t_1} \times \ldots \times \mathcal{Q}^{t_{r_1}} \times \mathcal{S}^{n_1}_+ \times \ldots \times \mathcal{S}^{n_{r_2}}_+$. We will assume that $r_1 + r_2 > 0$, $t_j \geq 3$ and $n_j \geq 3$ for every j. Given x, we will use $x^j_{i,k}$ to denote the (i,k) entry of the j-th matrix block and x^j_i to denote the i-entry of the j-th vector block. We will also use the same notation to single out a few special elements. For $j \in [1, r_2]$, $a^j_{i,k} \in \mathcal{K}$ is such that all its blocks are zero except for the block corresponding to $\mathcal{S}^{n_j}_+$. In that block $a^j_{i,k}$ contains the $n_j \times n_j$ matrix that has one at the (i,k) and (k,i) entries and zero elsewhere. Similarly, for $j \in [1, r_1]$, $a^j_i \in \mathcal{K}$ is such that all its blocks are zero except for the block corresponding to $\mathcal{Q}^{t_j}$, where a^j_i corresponds to the i-th unit vector.

Recall that if $d = (d_0, \overline{d})$ and $x = (x_0, \overline{x})$ are points of $\mathcal{Q}^n$ with $d_0, x_0 \in \mathbb{R}$ and $\overline{d}, \overline{x} \in \mathbb{R}^{n-1}$, then $\langle d, x\rangle = 0$ implies $d_0\overline{x} + x_0\overline{d} = 0$. In particular, if d is a nonzero boundary point of $\mathcal{Q}^n$, then the face $\mathcal{Q}^n \cap \{d\}^\perp$ is equal to the half-line $h_{d'} = \{\alpha d' \mid \alpha \geq 0\}$ where $d' = (d_0, -\overline{d})$. We also have $h^*_{d'} = \{x \mid \langle x, d'\rangle \geq 0\}$.

Let $\mathcal{L}^\perp$ be the space spanned by the following vectors:

1. $a^1_1 + a^1_2$ and $\{a^{j-1}_3 + a^j_1 + a^j_2 \mid 1 < j \leq r_1\}$,
2. $a^{r_1}_3 + a^1_{1,1}$ and $\{a^1_{i,i} + a^1_{i-1,i+1} \mid 1 < i < n_1\}$ (if $r_1 = 0$, use $a^1_{1,1}$ instead of $a^{r_1}_3 + a^1_{1,1}$),
3. $a^{j-1}_{n_j, n_{j-1}} + a^j_{1,1}$ and $\{a^j_{i,i} + a^j_{i-1,i+1} \mid 1 < i < n_j\}$, for $1 < j \leq r_2$.

Remark. *It will be helpful to keep in mind the case where $r_1 = r_2 = 2$, $n_1 = n_2 = t_1 = t_2 = 3$. In this case, $\mathcal{L}^\perp$ is spanned by elements having the following format*

$$\begin{pmatrix} y_1 \\ y_1 \\ y_2 \end{pmatrix} \times \begin{pmatrix} y_2 \\ y_2 \\ y_3 \end{pmatrix} \times \begin{pmatrix} y_3 & 0 & y_4 \\ 0 & y_4 & y_5 \\ y_4 & y_5 & 0 \end{pmatrix} \times \begin{pmatrix} y_5 & 0 & y_6 \\ 0 & y_6 & 0 \\ y_6 & 0 & 0 \end{pmatrix}.$$

Proposition 24. *Consider the problem (D) with $c = 0$ and $\mathcal{A}$ such that $\mathrm{range}\,\mathcal{A}^* = \mathcal{L}$, where $\mathcal{L}^\perp$ is the subspace constructed above. Then $d(D) = r_1 + \sum_{j=1}^{r_2}(n_j - 1)$.*

Proof. First, suppose that $r_1 > 0$. The first reducing direction must be some $x \in (\mathcal{K} \cap \mathcal{L}^\perp) \setminus \mathcal{K}^\perp$. However, if $x \in \mathcal{L}^\perp$, then $x_1^1 = x_2^1$. Then, because $x^1 \in \mathcal{Q}^{t_1}$, we have $x_i^1 = 0$ for all $i \geq 3$. Therefore, the coefficient of $a_3^1 + a_1^2 + a_2^2$ appearing in x must be zero as well. It follows that all blocks of x are zero, except for x^1. We conclude that x must be a positive multiple of $a_1^1 + a_2^1$. So let $d_1 = a_1^1 + a_2^1$, we then have

$$\mathcal{F}_2 = \mathcal{K} \cap \{d_1\}^\perp = h_{d_1'} \times \mathcal{Q}^{t_2} \times \ldots \times \mathcal{Q}^{t_{r_1}} \times \mathcal{S}_+^{n_1} \times \ldots \times \mathcal{S}_+^{n_{r_2}},$$

where $h_{d_1'}$ is contained in $\mathcal{Q}^{t_1}$ and is the half-line along the direction defined by the nonzero part of $a_1^1 - a_2^1$. At the next step, it turns out that only the positive multiples of $a_3^1 + a_1^2 + a_2^2$ belong to $(\mathcal{F}_2^* \cap \mathcal{L}^\perp) \setminus \mathcal{F}_2^\perp$. This means that facial reduction must proceed by successively selecting positive multiples of:

1. $d_1 = a_1^1 + a_2^1$ and $d_j = a_3^{j-1} + a_1^j + a_2^j$, for $1 < j \leq r_1$.

After r_1 steps, all the Lorentz cone blocks will be transformed to half-lines and we will have $\mathcal{F}_{r_1+1} = h_{d_1'} \times \ldots \times h_{d_{r_1}'} \times \mathcal{S}_+^{n_1} \times \ldots \times \mathcal{S}_+^{n_{r_2}}$, where for every $1 < j \leq r_1$, $h_{d_j'}$ is the half-line in $\mathcal{Q}^{t_j}$ along the direction defined by the nonzero part of $a_1^j - a_2^j$. If $r_2 = 0$, we are done. Otherwise, we have $(\mathcal{F}_{r_1+1}^* \cap \mathcal{L}^\perp) \setminus \mathcal{F}_{r_1+1}^\perp = \{t(a_3^{r_1} + a_{1,1}^1) \mid t > 0\}$. Again, we must proceed "one row at time" and select positive multiples of $a_{i,i}^1 + a_{i-1,i+1}^1$ for $1 < i < n_1$ as the reducing directions. In total, we find $n_1 - 1$ directions before we can move to the next block. For each block $n_j - 1$ directions will be found, so in total we obtain $r_1 + \sum_{j=1}^{r_2}(n_j - 1)$ directions. The case $r_1 = 0$ follows similarly. $\square$