Solving SDP Completely with an Interior Point Oracle

Bruno F. Lourenço *  Masakazu Muramatsu†  Takashi Tsuchiya ‡

July 2015 (Revised: August 2019)

Abstract

We suppose the existence of an oracle which solves any semidefinite programming (SDP) problem satisfying Slater’s condition simultaneously at its primal and dual sides. We note that such an oracle might not be able to directly solve general SDPs even after certain regularization schemes are applied. In this work we fill this gap and show how to use such an oracle to “completely solve” an arbitrary SDP. Completely solving an SDP includes, for example, distinguishing between weak/strong feasibility/infeasibility and detecting when the optimal value is attained or not. We will employ several tools, including a variant of facial reduction where all auxiliary problems are ensured to satisfy Slater’s condition at all sides. Our main technical innovation, however, is an analysis of double facial reduction, which is the process of applying facial reduction twice: first to the original problem and then once more to the dual of the regularized problem obtained during the first run. Although our discussion is focused on semidefinite programming, the majority of the results are proved for general convex cones.

Keywords: double facial reduction, facial reduction, semidefinite programming, feasibility problem.

1 Introduction

Consider the following pair of primal and dual linear semidefinite programs (SDPs).

\[
\begin{align*}
\inf_x \langle c, x \rangle & \quad \text{(SDP-P)} \\
\text{subject to} \ A x = b & \quad \text{subject to} \ c - A^* y \in \mathcal{S}_+^n, \\
x \in \mathcal{S}_+^n & \quad \text{where } \mathcal{S}_+^n \text{ denotes the cone of } n \times n \text{ symmetric positive semidefinite matrices, which is contained in } \mathcal{S}^n \text{ (the space of } n \times n \text{ symmetric matrices). Here } A : \mathcal{S}^n \to \mathbb{R}^m \text{ is a linear map, } b \in \mathbb{R}^m, c \in \mathcal{S}^n \text{ and } A^* \text{ denotes the adjoint map of } A. \text{ In addition, we assume that both } \mathbb{R}^m \text{ and } \mathcal{S}^n \text{ are equipped with the usual Euclidean product and the trace inner product, respectively. We will use the same symbol } \langle \cdot, \cdot \rangle \text{ to express both inner products.}
\end{align*}
\]

We start with the following observation.

To the best of our knowledge, all (or, almost all) methods for solving SDP require some kind of assumption on the problems (SDP-P) and (SDP-D) in order for its convergence theory to work. In addition, there seems to be no method that can solve arbitrary SDP instances and distinguish between all the kinds of ill-behaviour that can happen in semidefinite programming.

*Department of Mathematical Informatics, Graduate School of Information Science & Technology, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan. (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: tuchiya@grips.ac.jp)

B. F. Lourenço is partially supported by the JSPS Grant-in-Aid for Young Scientists 19K20217. M. Muramatsu is partially supported by the JSPS Grant-in-Aid for Scientific Research (B)26280005 and (C)17K00031. T. Tsuchiya is partially supported by the JSPS Grant-in-Aid for Scientific Research (B)15H02068. M. Muramatsu and T. Tsuchiya are also supported in part by the JSPS Grant-in-Aid for Scientific Research (B)24310112 and (C) 2630025. B. F. Lourenço and T. Tsuchiya are also partially supported by the JSPS Grant-in-Aid for Scientific Research (B)18H03206.
The previous observation contains both surprising and non-surprising aspects. On one hand, it might seem almost obvious that some condition must be imposed on the pair (SDP-P) and (SDP-D) in order to get meaningful convergence results, which is the pattern established in classical nonlinear programming since the early days, where many convergence results required some constraint qualification to hold. On the other hand, for linear programming we have the simplex method, which, at least in theory, is able to solve any linear program and detect all possible outcomes including infeasibility and unboundedness. Given that semidefinite programming is widely used and is often considered one of the natural extensions of linear programming, it is somewhat disappointing that, as of this writing, we still cannot claim to be able to solve SDPs in the same thorough way.

However, there are indeed classes of SDPs that we can reasonably claim that are solvable by current methods. One of those classes consists of the SDPs for which the so-called Slater’s condition holds at both (SDP-P) and (SDP-D). They are solvable, for instance, by using interior point methods [25, 2]. In this paper, we aim to show the following result.

Suppose we have access to an oracle that can solve any SDP instance, provided that the instance is both primal and dual strongly feasible. Then, we can “completely solve” any SDP instance with polynomially (in n) many calls to this oracle.

Later in Section 1.2 we state our results more precisely including a suitable definition of “completely solve” but, for now, we give some background to our research and results.

1.1 Background and previous works

Interior point methods (IPMs) [25, 2] are one of the standard methods for solving semidefinite programs (SDP-P) and (SDP-D). However, in order to function properly, they require that the problem at hand satisfy certain regularity conditions which may fail to be satisfied and this leads to numerical difficulties. The usual requirement is that there is a primal feasible solution x such that x is positive definite and that there is a dual feasible solution y such that the corresponding slack s = c - A*y is positive definite. In other words, Slater’s condition must hold for both (SDP-P) and (SDP-D). We now discuss briefly how Slater’s condition is connected with some research trends in continuous optimization.

- **Interior point algorithms and software.** Most modern IPM softwares [39, 9, 41] including the commercial solver Mosek do not require explicit knowledge of an interior feasible point beforehand. SeDuMi [39], for instance, transforms a standard form problem into the so-called homogeneous self-dual formulation, which has a trivial starting point. SDPA [9] and SDPT3 [41] use an infeasible interior point method. The fact that these methods can work without explicit knowledge of an interior feasible point, does not mean that they do not require the existence of an interior feasible point. Quite the opposite, the absence of interior feasible points may introduce theoretical and numerical difficulties in recovering a solution for the original problem. Also, detection of infeasibility is a complicated task. Some interior point methods, such as the one discussed in [26] by Nesterov, Todd and Ye, are able to obtain a certificate of infeasibility if the problem is dual or primal strongly infeasible, but the situation is less clear in the presence of the so-called weak infeasibility [18].

- **Ramana’s extended dual.** When neither the primal nor the dual have interior feasible points, there could be a nonzero duality gap between the (SDP-P) and (SDP-D), i.e., the optimal values of (SDP-P) and (SDP-D) might differ. To address this, Ramana [36, 37] developed an alternative duality theory for (SDP-D) that requires no regularity assumptions. Remarkable features of Ramana’s dual include the fact that it can be written as an SDP, always affords zero duality gap and that the dual is always attained whenever the primal optimal value is finite. However, Ramana’s dual is not necessarily suitable to be used with IPMs due to the fact that it does not ensure the existence of interior feasible points at both sides.

- **Facial reduction.** Denote by \( \mathcal{F}_D^S \), the set of feasible slacks of (SDP-D), i.e.,

\[
\mathcal{F}_D^S = \{ s \in S_n^+ \mid \exists y, s = c - A^*y \}.
\]
Let $\mathcal{F}_\text{min}^D$ be the minimal face of $S^n_+$ which contains $\mathcal{F}_D^S$. If we substitute $S^n_+$ for $\mathcal{F}_\text{min}^D$ in (SDP-D), then the new (SDP-D) will satisfy Slater’s condition, because $\mathcal{F}_\text{min}^D$ is characterized as the unique face for which $\mathcal{F}_D^S$ intersects the relative interior of $\mathcal{F}_\text{min}^D$. The process of finding $\mathcal{F}_\text{min}^D$ is called *facial reduction* [45, 28] and was developed originally by Borwein and Wolkowicz [5, 4]. We will overview facial reduction in more detail in Section 3.

However, one important point is that facial reduction only guarantees that Slater’s condition is satisfied at one side of the problem. So, again, even this regularized problem might fail to have interior solutions at both primal and dual sides.

We remark that strong feasibility at only one of the sides of the problem can also be a source of numerical difficulties. In Section 2 of [46], Waki, Nakata and Muramatsu shows an instance that satisfies Slater’s condition at the primal side, but not at the dual side. Its optimal value is zero but both SDPA [9] and SeDuMi [39] output 1 instead.

- **Algebraic approaches.** Henrion, Naldi and Din described an algebraic approach to the problem of obtaining a feasible solution to (SDP-D), see [10, 11]. Interesting features of their algorithm include, among others, the fact that their algorithm is implementable in exact arithmetic (as opposed to floating point arithmetic) and that, as long as (SDP-D) satisfies certain genericity assumptions, the algorithm can find solutions even in degenerate cases when Slater’s condition is not satisfied. In addition, when a solution is found, a so-called rational parametrization is provided for it. A description of their package *Spectra* is given in [11]. Drawbacks, however, include that in most cases, only small instances can be solved, see Section 1 of [11]. Furthermore, optimization problems cannot be solved directly.

One of the reasons why general SDPs are hard is because there can be afflicted by all sorts of pathologies. Fortunately, there is a growing body of research aimed at understanding SDPs and conic linear programs having pathological behaviours such as nonzero duality gaps and weak infeasibility. Here we will mention a few of them. A problem is said to be *weakly infeasible* if there is no feasible solution but the distance between the underlying affine space and the cone under consideration is zero. Weak infeasibility is known to be very hard to detect numerically, see for instance Pólik and Terlaky [34]. In [44], Waki showed that weakly infeasible problems sometimes arise from polynomial optimization. There is also a discussion on weak infeasibility semidefinite programming and second order cone programming in [18] and [20], respectively. Some of the results in [18] were generalized to arbitrary closed convex cones by Liu and Pataki, see [15] for more details. See also [19], where some results of [15] on weakly infeasible problems are sharpened when the underlying cone has polyhedral faces.

It is hard to obtain finite certificates of infeasibility for SDPs, because there is no straightforward extension of Farkas’ Lemma for non-polyhedral cones. The first finite infeasibility certificate was obtained by Ramana in [36]. Since then, Sturm mentioned the possibility of obtaining a finite certificate for infeasibility by using the directions produced in his regularization procedure, see page 1243 of [40]. More recently, Liu and Pataki have also obtained finite certificates through elementary reformulations [14]. Interestingly, Klep and Schweighofer [13] also obtained certificates through a completely different approach using tools from real algebraic geometry. As we move from SDPs to conic linear programs over arbitrary cones, facial reduction seems to one of the few approaches that can provide finite certificates of infeasibility see, for example, [15].

In [46], Waki, Nakata and Muramatsu discussed SDP instances for which known solvers failed to obtain the correct answer and in one case, this happened even though the problem had an interior feasible point at the primal side. In [29, 30], Pataki gave a definition of “bad behaviour” and showed that all SDPs in that class can be put in the same form, after performing an elementary reformulation. A discussion on duality gaps and many interesting examples of pathological SDPs are given by Tunçel and Wolkowicz in [42]. Pataki has recently provided an extensive study of duality gaps in semidefinite programming in [31]. He showed, for instance, that all SDPs with positive duality gap and $m = 2$ (i.e., the dual problem has two variables) have a common reformulation, see Theorem 1 therein.
1.2 Summary of results

In general, SDPs can be plagued by all sorts of ill-behavior and this stands in sharp contrast to the relatively nice nature of SDPs where both the primal and dual satisfy Slater’s condition. In order to capture this distinction, we consider the following oracle.

**Oracle 1:** The interior point oracle $O_{\text{int}}$ for SDPs

**Input:** The problem data: $S^n_+, A, b, c$. Both (SDP-P) and (SDP-D) must satisfy Slater’s condition.

**Output:** A zero duality gap optimal solution pair $x^*, y^*$. That is, $x^*$ and $y^*$ satisfy

\[
\langle c, x^* \rangle = \langle b, y^* \rangle \\
c - A^* y^* \in S^n_+ \\
A x^* = b \\
x^* \in S^n_+.
\]

We will refer to the oracle in Oracle 1 as $O_{\text{int}}$. We now define the meaning of completely solving an SDP.

**Definition 1** (Completely solving an SDP instance). *An algorithm, procedure or a scheme is said to completely solve an SDP instance, if it receives as input $S^n_+, A, b, c$ and $\epsilon > 0$ and achieves the following goals.*

(a) It decides whether the (SDP-D) is feasible or not.

(b) When (SDP-D) is feasible, it computes the optimal value. If the optimal value is attained, it computes an optimal solution. If the optimal value is not attained, it computes an $\epsilon$-optimal solution.

(c) When (SDP-D) is infeasible, it correctly distinguishes between strong infeasibility and weak infeasibility. If the (SDP-D) is weakly infeasible, then it find a matrix that is arbitrarily close to feasibility (this will be made precise later).

Although we are focused on semidefinite programming, the majority of our results will be proved for general conic linear programs (CLPs). With that, our contributions in this paper are as follows.

1. We present an algorithm for completely solving general CLPs, provided that we can solve certain auxiliary problems that satisfy Slater’s condition, see Section 4 and Algorithm 4. In particular, we will show that an arbitrary SDP can be completely solved by $O(n)$ calls to $O_{\text{int}}$. This shows that even though an arbitrary SDP may have unfavourable properties, we can always completely solve it in the sense of Definition 1 if we assume that we are capable of solving instances satisfying Slater’s condition at both primal and dual sides. An important feature of our approach is that it is method agnostic and does not rely in any way on the inner working of $O_{\text{int}}$.

2. We present a detailed discussion of double facial reduction for general conic linear programs, which is the process of applying facial reduction twice: first to an CLP and then, to the dual of the regularized CLP obtained at the first step.

Through double facial reduction, whenever the optimal value of (SDP-D) is finite, we are ensured to obtain a new pair of primal and dual problems that satisfy Slater’s condition and whose common optimal value coincides with the optimal value of (SDP-D).

Although we cannot always recover optimal solutions for (SDP-D) from this new pair of problems (after all, (SDP-D) might not even have optimal solutions in the first place), we will show how it is possible to obtain feasible solutions that are arbitrarily close to optimality, for any desired accuracy, by using the directions that appear when applying facial reduction. See Section 4.1 and Algorithm 2 for more details. The discussion on obtaining almost optimal solution leads naturally to an approach for obtaining almost feasible solution for weakly infeasible problems and this is discussed in Section 4.2.
3. We present several technical results about facial reduction that we believe might be of independent interest. For example, we show how to perform facial reduction by solving auxiliary problems that are ensured to satisfy Slater’s condition at both primal and dual sides, see Lemma 10 and Algorithm 1. We also provide a technical result on how the feasibility properties of a problem might change when facial reduction is applied to its dual, see Proposition 14 and Theorem 15.

We remark that this paper is a thorough extension and reformulation of an earlier technical report [17], where the results were only proved for semidefinite programming by different techniques.

1.3 Limitations of this work

Although we stand by our theoretical results including the ones on facial reduction and double facial reduction, we must admit that the algorithm for completely solving conic linear programs (Algorithm 4) is somewhat hypothetical. This is because, except in very special cases [43, 47, 24], we cannot solve exactly an SDP even if primal and dual Slater’s conditions are simultaneously satisfied. Usually, the best we can do is to compute solutions that are approximately feasible and approximately optimal to some specified tolerance $\epsilon > 0$ or, under special circumstances, provide a rational parametrization to the solution set as in [10, 11]. So, strictly speaking, only an approximate version of the oracle $O_{\text{int}}$ might be practically implementable. Because of that, we believe our approach, in all likelihood, does not lead to a practical algorithm for general problems.

Missing from our analysis is how to deal with the case where there is some imprecision in the answer returned by $O_{\text{int}}$. This is a very complex issue because since regularity conditions might fail, small perturbations in the input data might lead to problems whose optimal values are vastly different. Furthermore, impreciseness whilst doing facial reduction might lead to a wrong face being computed and feasible solutions could be inadvertently removed.

We believe however, that the analysis of the exact case is an important stepping stone and we see a similar pattern in many subareas of optimization. For example, for augmented Lagrangian methods, understanding the behavior of the algorithm when subproblems are solved exactly seems to be quite important for getting the larger picture of the algorithm and its convergence analysis, even though, in practice, the subproblems are only approximately solved.

We remark that related approaches by de Klerk, Roos and Terlaky [7] and Permenter, Friberg and Andersen [32] also assume that exact solutions are obtainable. However, numerical experiments are provided in [32] to check how their approach fare under inexactness. We provide a detailed comparison between [7, 32] and our approach in Section 6.

1.4 Structure of this work

Our work is divided as follows. Section 2 discusses the notation used throughout the paper and contains a review of the necessary notions from convex analysis. Some technical aspects related to the interior point oracle $O_{\text{int}}$ are discussed in Section 2.3. Section 3 presents a facial reduction algorithm that is suitable to be used in conjunction with $O_{\text{int}}$. Section 4 discusses double facial reduction and how it can be used to obtain almost optimal solutions and analyze weak infeasibility. Section 5 contains the description of an algorithm for completely solving a general conic linear program which can be adapted to use $O_{\text{int}}$ when the underlying cone is $S^n_+$. Section 6 contains a discussion on related approaches. Section 7 concludes this work.

2 Preliminary discussion and review of relevant notions

Let $C \subseteq E$ be a closed convex set contained in a real finite dimensional space $E$. Its relative interior, closure, linear span and dimension are denoted by $\text{ri} \, C$, $\text{cl} \, C$, $\text{span} \, C$ and $\dim \, C$, respectively. We assume that $E$ is equipped with some inner product $\langle \cdot , \cdot \rangle$ and we will denote by $C^\perp$ the subspace of $E$ which contains the
elements orthogonal to \( C \) with respect to \( \langle \cdot, \cdot \rangle \). We will denote by \( \| \cdot \| \) the norm induced by \( \langle \cdot, \cdot \rangle \). For a pair of sets \( C, D \subseteq \mathcal{E} \) we define the distance between \( C \) and \( D \) as

\[
\text{dist} (C, D) := \inf \{ \| x - z \| \mid x \in C, z \in D \}.
\]

If \( x \in \mathcal{E} \), we will use \( \text{dist} (x, C) \) as a shorthand for \( \text{dist} \{ \{ x \} \}, C \).

If \( A \) is a linear map from \( \mathcal{E} \) to some \( \mathbb{R}^m \), we will denote the image, the kernel and the adjoint of \( A \) by \( \text{range} \, A, \ker \, A \) and \( A^* \), respectively.

For \( K \subseteq \mathcal{E} \) a closed convex cone, we denote by \( \text{lin} \, K \) the largest subspace contained in \( K \), i.e.,

\[
\text{lin} \, K := K \cap -K.
\]

We denote by \( K^* \) the dual cone of \( K \):

\[
K^* := \{ x \in \mathcal{E} \mid \langle s, x \rangle \geq 0, \forall s \in K \}.
\]

A closed convex subset \( F \) contained in \( K \) is said to be a face of \( K \) if

\[
s, \hat{s} \in K, \frac{s + \hat{s}}{2} \in F \Rightarrow s, \hat{s} \in F.
\]

The conjugate face of \( F \) is defined as

\[
F^\Delta := K^* \cap F^\perp.
\]

Given \( x \in K \), we write \( F(x, K) \) for the intersection of all faces of \( K \) containing \( x \). \( F(x, K) \) is the minimal (with respect to inclusion) face of \( K \) containing \( x \).

For a given \( x \in C \), we write \( \text{dir} (x, K) \) for the cone of feasible directions of \( C \) at \( x \). This is the set

\[
\text{dir} (x, K) := \{ z \in \mathbb{R}^n \mid \exists t > 0, x + tz \in C \}.
\]

The closure of \( \text{dir} (x, K) \) is the tangent cone of \( C \) at \( x \) and is denoted by \( \text{tan} (x, K) \). The tangent space of \( C \) at \( x \) is the lineality space of \( \text{tan} (x, K) \) and is denoted by \( T_x K \). In summary, we have

\[
\text{tan} (x, K) := \text{cl} \, \text{dir} (x, K),
\]

\[
T_x K := \text{tan} (x, K) \cap -\text{tan} (x, K).
\]

Some of the relationships between the sets defined so far will be summarized at Lemma 3.

Although our focus is on semidefinite programming, most of the results will be proved for general conic linear programming. We will denote by \( (K, A, b, c) \), the following pair of primal and dual conic linear programs.

\[
\inf_x \langle c, x \rangle \quad \text{subject to } Ax = b \quad \text{(Conic-P)} \quad \sup_y \langle b, y \rangle \quad \text{subject to } c - A^* y \in K, \quad \text{(Conic-D)}
\]

where \( A \) is a linear map between \( \mathcal{E} \) and some \( \mathbb{R}^m \), \( b \in \mathbb{R}^m \), \( c \in \mathcal{E} \). Semidefinite programming corresponds to the specific case where \( \mathcal{E} = \mathcal{S}^n \) and \( K = \mathcal{S}^n_+ \).

We will denote by \( \theta_P \) and \( \theta_D \), the optimal values of \( \text{(Conic-P)} \) and \( \text{(Conic-D)} \) respectively. It is understood that \( \theta_P = +\infty \) if \( \text{(Conic-P)} \) is infeasible and \( \theta_D = -\infty \) if \( \text{(Conic-D)} \) is infeasible. The pair \( \text{(Conic-P)} \) and \( \text{(Conic-D)} \) is said to have zero duality gap if \( \theta_P = \theta_D \). The primal and dual feasible regions are defined as follows:

\[
\mathcal{F}_P := \{ x \in K^* \mid Ax = b \},
\]

\[
\mathcal{F}_D := \{ y \in \mathbb{R}^m \mid c - A^* y \in K \},
\]

\[
\mathcal{F}_S := \{ s \in K \mid \exists y \in \mathbb{R}^m, s = c - A^* y \} = (c + \text{range} \, A^*) \cap K.
\]
An element $s \in F^S_D$ is called a dual slack. The dual optimal value $\theta_D$ is said to be attained if there is $y \in F^D_D$ such that $\langle b, y \rangle = \theta_D$. The notion of primal attainment is analogous. We recall the following basic constraint qualification.

**Proposition 2** (Slater). Consider the pair (Conic-P) and (Conic-D).

(i) If there exists $x \in (\text{ri} K^*) \cap F^P_P$ and $\theta_P$ is finite, then $\theta_P = \theta_D$ and the dual optimal value $\theta_D$ is attained.

(ii) If there exists $s \in (\text{ri} K) \cap F^S_S$ and $\theta_D$ is finite, then $\theta_P = \theta_D$ and the primal optimal value $\theta_P$ is attained.

For the reader’s convenience, before we proceed we recall a few basic facts from convex analysis. We provide short proofs for most items for the sake of completeness.

**Lemma 3.** Let $K \subset E$ be a closed convex cone, $e \in \text{ri} K$, $x \in K$ and $z \in K^*$.

(i) $K^{**} = K$.

(ii) $\text{lin} K = K^{*\perp}$, where $K^{*\perp}$ is a short-hand for $(K^*)^\perp$.

(iii) $K^\perp = \text{lin} (K^*)$.

(iv) $x + e \in \text{ri} K$.

(v) There exists $\alpha > 1$ such that $\alpha e + (1 - \alpha)x \in K$.

(vi) $z \in K^\perp$ if and only if $\langle e, z \rangle = 0$.

(vii) $x \in \text{ri} F(x, K)$.

(viii) $F(x, K)^\Delta = K^* \cap \{x\}^\perp$.

(ix) $(\tan(x, K))^* = F(x, K)^\Delta$.

(x) $T_x K = F(x, K)^{\perp\Delta}$.

(xi) If $w \in \tan(x, K)$ then $\lim_{t \to +\infty} \text{dist}(tx + w, K) = 0$.

**Proof.** (i) This is the bipolar theorem, see Theorem 14.1 of [38].

(ii) If $z \in \text{lin} K$, then $\langle z, y \rangle \geq 0$ and $\langle -z, y \rangle \geq 0$, for every $y \in K^*$. It follows that $z \in K^{*\perp}$. Reciprocally, if $z \in K^{*\perp}$, then $z \in K^{**} = K$, by item (i). Since $K^{*\perp}$ is a subspace, we have $K^{*\perp} \subseteq \text{lin} K$.

(iii) It follows from applying (ii) to $K^*$.

(iv) Since $e \in \text{ri} K$, for any $z \in K$ we have that all points in the relative interior of the line segment connecting $z$ and $e$ also belong to the relative interior of $K$, see Theorem 6.1 of [38]. Since $x + e = e\frac{1}{2} + (2x + e)\frac{1}{2}$, we have $x + e \in \text{ri} K$.

(v) See Theorem 6.4 of [38].

(vi) If $z \in K^\perp$, it is clear that $\langle e, z \rangle$ is zero. Now, suppose that $\langle e, z \rangle$ is zero. By item (iv), there is $\alpha > 1$ such that $u := \alpha e + (1 - \alpha)x \in K$.

On one hand, since $z \in K^*$, we have $\langle u, z \rangle \geq 0$. On the other, $\langle u, z \rangle = (1 - \alpha)\langle x, z \rangle \leq 0$. So, we must have $\langle x, z \rangle = 0$. As $x$ is an arbitrary element, it holds that $z \in K^\perp$.  

7
(vii) Suppose \( x \notin {\text{ri}} \mathcal{F}(x, \mathcal{K}) \). Then \( x \) and \( \mathcal{F}(x, \mathcal{K}) \) can be properly separated by an hyperplane \( H \) (see Theorem 11.3 in [38]), where we recall that proper separation means that \( x \) and \( \mathcal{F}(x, \mathcal{K}) \) belong to opposite closed half-spaces defined by \( H \) and at least one of them is not entirely contained in \( H \). Since \( x \in \mathcal{F}(x, \mathcal{K}) \), we must have \( x \in H \) and, therefore, \( \mathcal{F}(x, \mathcal{K}) \) is not entirely contained in \( H \).

The fact that \( \mathcal{F}(x, \mathcal{K}) \) is contained in one of the closed half-spaces defined by \( H \) together with \( x \in H \) imply that \( \mathcal{F}(x, \mathcal{K}) \cap H \) is a face of \( \mathcal{F}(x, \mathcal{K}) \) (and, therefore, a face of \( \mathcal{K} \)). We also have

\[
\mathcal{F}(x, \mathcal{K}) \cap H \subseteq \mathcal{F}(x, \mathcal{K}),
\]

because \( \mathcal{F}(x, \mathcal{K}) \) is not entirely contained in \( H \). This contradicts the minimality of \( \mathcal{F}(x, \mathcal{K}) \).

(viii) By definition, \( \mathcal{F}(x, \mathcal{K})^\Delta = \mathcal{K}^* \cap \mathcal{F}(x, \mathcal{K})^\perp \). Since \( x \in \mathcal{F}(x, \mathcal{K}) \), we have

\[
\mathcal{F}(x, \mathcal{K})^\Delta = \mathcal{K}^* \cap \mathcal{F}(x, \mathcal{K})^\perp \subseteq \mathcal{K}^* \cap \{x\}^\perp.
\]

For the converse, suppose that \( s \in \mathcal{K}^* \cap \{x\}^\perp \). We have \( s \in \mathcal{F}(x, \mathcal{K})^\perp \), since \( \mathcal{F}(x, \mathcal{K}) \subseteq \mathcal{K} \). By item (vii), \( x \in {\text{ri}} \mathcal{F}(x, \mathcal{K}) \). From item (vi) we conclude that \( s \in \mathcal{F}(x, \mathcal{K})^\perp \) and, therefore, \( s \in \mathcal{F}(x, \mathcal{K})^\Delta \).

(ix) First we show that \( \mathcal{F}(x, \mathcal{K})^\Delta \subseteq (\tan (x, \mathcal{K}))^* \). If \( s \in \mathcal{F}(x, \mathcal{K})^\Delta \) and \( z \in \text{dir}(x, \mathcal{K}) \), then, for some \( t > 0 \) we have

\[
\langle s, x + tz \rangle \geq 0.
\]

Because \( \langle s, x \rangle = 0 \), we must have \( \langle s, z \rangle \geq 0 \), which shows that \( s \in \text{dir}(x, \mathcal{C})^* = (\tan (x, \mathcal{K}))^* \), since a set and its closure have the same dual.

Conversely, suppose that \( s \in (\tan (x, \mathcal{K}))^* \). Because \( \mathcal{K} \subseteq \tan (x, \mathcal{K}) \), we have that \( s \in \mathcal{K}^* \). In addition, since both \( x \) and \(-x\) belong to \( \tan (x, \mathcal{K}) \), we have \( \langle s, x \rangle = 0 \). This implies that \( s \in \mathcal{K}^* \cap \{x\}^\perp \), which coincides with \( \mathcal{F}(x, \mathcal{K})^\Delta \), by item (viii).

(x) Follows from (iii) and (ix).

(xi) Since \( \mathcal{K} \) is a closed convex cone, we have \( \text{dist} (a+b, \mathcal{K}) \leq \text{dist} (a, \mathcal{K}) + \text{dist} (b, \mathcal{K}) \), for all \( a, b \in \mathcal{E} \). Now, for every \( \epsilon > 0 \), there exists \( w_\epsilon \in \text{dir}(x, \mathcal{K}) \) such that \( \text{dist} (w, w_\epsilon) < \epsilon \). Moreover, there exists \( t_\epsilon \) such that \( t_\epsilon x + w_\epsilon \in \mathcal{K} \). It follows that

\[
\text{dist} (tx + w, \mathcal{K}) \leq \text{dist} (tx + w_\epsilon, \mathcal{K}) + \text{dist} (w - w_\epsilon, \mathcal{K})
\]

\[
\leq \text{dist} (tx + w_\epsilon, \mathcal{K}) + \epsilon,
\]

where the last inequality follows from the fact that \( 0 \in \mathcal{K} \), so \( \text{dist} (w - w_\epsilon, \mathcal{K}) \leq \text{dist} (w - w_\epsilon, 0) \). However, since \( t_\epsilon x + w_\epsilon \in \mathcal{K} \), we must have \( \lim_{t \to +\infty} \text{dist} (tx + w_\epsilon, \mathcal{K}) = 0 \), since for \( t \) sufficiently large we have \( tx + w_\epsilon \in \mathcal{K} \). It follows that \( \lim_{t \to +\infty} \text{dist} (tx + w, \mathcal{K}) \leq \epsilon \). Since \( \epsilon \) is arbitrary, we conclude that (xi) must hold.

\[\square\]

2.1 Types of feasibility, almost optimality, almost feasibility

Here, we review the fact that a conic linear program can be in four different feasibility statuses. We say that (Conic-D) is

(i) strongly feasible if \( (\text{ri} \mathcal{K}) \cap (c + \text{range} \mathcal{A}^*) \neq \emptyset \) (i.e., Slater’s condition hold),

(ii) weakly feasible if it is feasible but not strongly feasible,

(iii) weakly infeasible if it is infeasible but \( \text{dist}(c + \text{range} \mathcal{A}^*, \mathcal{K}) = 0 \),

(iv) strongly infeasible if \( \text{dist}(c + \text{range} \mathcal{A}^*, \mathcal{K}) > 0 \).
Strong/weak feasibility/infeasibility of (Conic-P) is defined analogously by replacing \((c + \text{range} \mathcal{A}^*)\) by \(V := \{x \mid \mathcal{A}x = b\}\). As a matter of convention, if \(V = 0\), we will say that (Conic-P) is strongly infeasible. If a problem is either weakly infeasible or weakly feasible we will say that it is in weak status.

We have the following characterization of strong infeasibility, see Lemma 5 in [21].

**Proposition 4** (Characterization of strong infeasibility). The following hold.

(i) (Conic-P) is strongly infeasible if and only if there exists \(y\) such that
\[
(b, y) = 1 \quad \text{and} \quad -\mathcal{A}^* y \in \mathcal{K}.
\]

(ii) (Conic-D) is strongly infeasible if and only if there exists \(x\) such that
\[
(c, x) = -1 \quad \text{and} \quad x \in \mathcal{K}^* \cap \ker \mathcal{A}
\]

Moving on, let \(y \in \mathcal{E}\) and \(\epsilon > 0\). We say that \(y\) is an \(\epsilon\)-feasible solution to (Conic-D) if \(\text{dist} (c - \mathcal{A}^* y, \mathcal{K}) \leq \epsilon\). In addition, we say that \(y\) is an \(\epsilon\)-optimal solution to (Conic-D) if \(y\) is feasible for (SDP-D) and \((b, y) \geq \theta_D - \epsilon\). These notions will be used in Sections 4.1 and 4.2.

As a word of caution, we remark that, in general, even if \(s = c - \mathcal{A}^* y\) is such that \(\text{dist} (s, \mathcal{K})\) is small, there is no guarantee that \(\text{dist} (s, \mathcal{F}^D)\) will also be small. In this case, the quantities \(\text{dist} (s, \mathcal{K})\) and \(\text{dist} (s, \mathcal{F}^D)\) are sometimes called the backward error and forward error, respectively. The problem of bounding the forward error by the backward error is intrinsically connected with the notion of error bounds. See, for example, the fundamental work by Sturm [40] on error bounds for linear matrix inequalities, where he showed the importance of facial reduction in analyzing these questions. See also [16] for some generalizations of Sturm’s results to the so-called amenable cones.

### 2.2 Facial structure of \(\mathcal{S}_+^n\)

The cone of positive semidefinite symmetric matrices has a very special structure and every face of \(\mathcal{S}_+^n\) is linearly isomorphic to some \(\mathcal{S}_r^m\) for \(r \leq n\). This is a well-known fact which we register as a proposition for future reference. For a proof, see [27]. See also Section 6 of [3].

**Proposition 5.** Let \(\mathcal{F}\) be a nonempty face of \(\mathcal{S}_+^n\). There exists \(r \leq n\) and an orthogonal \(n \times n\) matrix \(Q\) such that
\[
Q^\top \mathcal{F} Q = \left\{ \begin{pmatrix} U & 0 \\ 0 & 0 \end{pmatrix} \mid U \in \mathcal{S}_r^m \right\}
\]

Using the fact that \(Q\) is orthogonal, we have that the dual of \(\mathcal{F}\) satisfies
\[
Q^\top \mathcal{F}^\ast Q = (Q^\top \mathcal{F} Q)^\ast = \left\{ \begin{pmatrix} U & V \\ V^\top & W \end{pmatrix} \in \mathcal{S}_n \mid U \in \mathcal{S}_r^m \right\}.
\]

### 2.3 The interior point oracle and related aspects

We can regard \(\mathcal{O}_\text{int}\) as a machine running an idealized version of either the homogeneous self-dual embedding [35, 7, 22], an infeasible interior point method [26], the ellipsoid method or even an augmented Lagrangian method. An important point is that no assumption is made on the inner workings of the oracle.

In the definition of \(\mathcal{O}_\text{int}\), the affine space is contained in the space of \(n \times n\) symmetric matrices and the optimization is carried over \(\mathcal{S}_+^n\). Note that \(n\) is the same for both \(\mathcal{S}_r^m\) and \(\mathcal{S}_n\). However, for fixed \(\mathcal{A}, b, c\) we might be interested in solving problems over a face of \(\mathcal{S}_n\) or over the dual of a face of \(\mathcal{S}_+^n\).

In this case, even if (Conic-D) and (Conic-P) are both primal and dual strongly feasible, it is not immediately clear how to use \(\mathcal{O}_\text{int}\) to solve (Conic-D) and (Conic-P), since they are not exactly SDPs. One possibility would be to consider a, a priori, stronger oracle that is also able to solve strongly feasible problems over faces of \(\mathcal{S}_+^n\).

We will shown in Appendix A that this is not necessary and, after an appropriate reformulation, we can still solve (Conic-P) and (Conic-D) using \(\mathcal{O}_\text{int}\). We register this fact as a proposition for future reference.
Proposition 6. Suppose that $\mathcal{K}$ is either a non-empty proper face of $\mathcal{S}_+^n$ or the dual of a non-empty proper face of $\mathcal{S}_+^n$. Suppose that (Conic-P) and (Conic-D) are strongly feasible. Then, (Conic-P) and (Conic-D) are solvable with a single call to $\mathcal{O}_{\text{int}}$.

3 Facial reduction with $\mathcal{O}_{\text{int}}$

A major obstacle for solving (SDP-D) with the oracle $\mathcal{O}_{\text{int}}$ is that, in general, Slater’s condition does not hold. Therefore, a critical step towards solving (SDP-D) is to reformulate (SDP-D) as an SDP instance for which Slater’s condition holds at both primal and dual sides. By using facial reduction, we are able to either detect infeasibility or to reformulate (SDP-D) as an SDP instance for which Slater’s condition holds at one side of the problem. This will be an important step towards completely solving (SDP-D).

In this section, we discuss facial reduction for general conic linear programs and how it can be carried out by solving auxiliary problems that are ensured to be strongly feasible (i.e., Slater’s condition holds) at both primal and dual sides. In particular, when the underlying cone $\mathcal{K}$ is $\mathcal{S}_+^n$, this will mean that facial reduction can be implemented through calls to $\mathcal{O}_{\text{int}}$. Although we will focus on problems formulated in the dual form (Conic-D), any analysis carried out for (Conic-D) can be translated back to (Conic-P).

Facial Reduction was developed by Borwein and Wolkowicz as a tool to restore strong duality in convex optimization [4, 5]. Descriptions for the conic linear programming case have appeared, for instance, in Pataki [28] and in Waki and Muramatsu [45]. Here, we will follow the version described in [45], which relies on the following key result.

Lemma 7 (The facial reduction lemma: Lemma 3.2 in [45]). (Conic-D) is not strongly feasible (i.e., Slater’s condition fails) if and only if there is $d \in \mathcal{K}^* \cap \ker \mathcal{A}$ such that:

(i) $(c, d) = 0$ and $d \notin \mathcal{K}^\perp$,

(ii) $(c, d) < 0$.

Remark. The primal counter part of Lemma 7 is obtained when $\ker \mathcal{A}, \mathcal{K}^*$ are substituted by range $\mathcal{A}^*$, $\mathcal{K}$ respectively and the condition on $c$ is replaced by a condition on $b$. To wit, (Conic-P) is not strongly feasible if and only if there exists $(f, y)$ such that $f = -\mathcal{A}^* y \in \mathcal{K}$ and either (i) $\langle b, y \rangle = 0$ and $f \notin \mathcal{K}^\perp = \text{lin} \mathcal{K}$ (item (iii) of Lemma 3) holds or (ii) $\langle b, y \rangle > 0$ holds.

Therefore, whenever (Conic-D) lacks a relative interior solution (i.e., $\mathcal{K} \cap (c + \text{range} \mathcal{A}^*) = \emptyset$), it is either because (Conic-D) is infeasible (alternative (ii) together with Proposition 4) or because the set of dual feasible slacks $\mathcal{F}^S_D$ is contained in $\mathcal{K} \cap \{d\}^\perp$ (alternative (i))\footnote{One must be careful that even if (Conic-D) is infeasible it might be the case that alternative (ii) is not satisfied at this stage. This happens, for instance, if (Conic-D) is weakly infeasible.}.

If alternative (i) holds, since $d \notin \mathcal{K}^\perp$, we have

$$\mathcal{K} \cap \{d\}^\perp \subsetneq \mathcal{K},$$

that is, the face $\mathcal{F}_2 := \mathcal{K} \cap \{d\}^\perp$ is properly contained in $\mathcal{K}$. We then substitute $\mathcal{K}$ for $\mathcal{F}_2$ and repeat. As long as $(\text{ri } \mathcal{F}_i) \cap (c + \text{range } \mathcal{A}^*) = \emptyset$, we can find a new direction $d$.

We recall that if $\mathcal{F}$ is a face of $\mathcal{K}$, then $\mathcal{F} \subseteq \mathcal{K}$ holds if and only if $\text{dim } \mathcal{F} < \text{dim } \mathcal{K}$. Therefore, (5) implies that after a finite number of iterations, we will either find some face $\mathcal{F}_i$ such that $(\text{ri } \mathcal{F}_i) \cap (c + \text{range } \mathcal{A}) \neq \emptyset$ or we will eventually find out that the problem is infeasible.

It turns out that $\mathcal{F}_i$ must be the smallest face $\mathcal{F}^D_{\text{min}}$ of $\mathcal{K}$ which contains $\mathcal{F}^S_D$. This process is called facial reduction and it aims at finding $\mathcal{F}^D_{\text{min}}$. The direction $d$ will be henceforth called a reducing direction. If there is no dual feasible solution, we have $\mathcal{F}^D_{\text{min}} = \emptyset$. For the sake of preciseness, we will state the following definition.

Definition 8 (Reducing directions). A reducing direction for (Conic-D) is an element $d \in \mathcal{K}^* \cap \ker \mathcal{A}$ such that $\langle c, d \rangle \leq 0$. A reducing direction for (Conic-P) is a pair $(f, y)$ such that $f \in \mathcal{K}$, $f = -\mathcal{A}^* y$ (i.e., $f \in \text{range } \mathcal{A}^*$) and $\langle b, y \rangle \geq 0$.
Next, \( \{d_1, \ldots, d_k\} \) is said to be a sequence of reducing directions for \((\text{Conic-D})\) if
\[
d_i \in (\mathcal{K} \cap \{d_1\}^\perp \cap \cdots \cap \{d_{i-1}\}^\perp)^* \cap \ker \mathcal{A} \cap \{c\}^\perp, \quad \text{for } i = 1, \ldots, \ell - 1 \tag{6}
\]
\[
d_i \in (\mathcal{K} \cap \{d_1\}^\perp \cap \cdots \cap \{d_{i-1}\}^\perp)^* \cap \ker \mathcal{A}, \quad (c, d_i) \leq 0. \tag{7}
\]

Analogously, \( \{(f_1, y_1), \ldots, (f_\ell, y_\ell)\} \) is said to be a sequence of reducing directions for \((\text{Conic-P})\) if
\[
f_i = -\mathcal{A}^* y_i, \quad y_i \in \{b\}^\perp, \quad f_i \in (\mathcal{K}^* \cap \{f_1\}^\perp \cap \cdots \cap \{f_{i-1}\}^\perp)^*, \quad \text{for } i = 1, \ldots, \ell - 1 \tag{8}
\]
\[
f_\ell = -\mathcal{A}^* y_\ell, \quad \langle b, y_\ell \rangle \geq 0, \quad f_\ell \in (\mathcal{K}^* \cap \{f_1\}^\perp \cap \cdots \cap \{f_{\ell-1}\}^\perp)^*. \tag{9}
\]

**Remark.** Liu and Pataki defined in [15] the notion of facial reduction cone, see Definition 2 therein. The \(k\)-th facial reduction cone of \(\mathcal{K}\) is given by
\[
\text{FR}_k(\mathcal{K}) = \{(d_1, \ldots, d_k) \mid d_1 \in \mathcal{K}^*, d_i \in (\mathcal{K} \cap \{d_1\}^\perp \cap \cdots \cap \{d_{i-1}\}^\perp)^*, i = 2, \ldots, k\}.
\]

With that, (6), (7) and (8), (9) imply that
\[
(d_1, \ldots, d_k) \in \text{FR}_k(\mathcal{K}), \quad (f_1, \ldots, f_\ell) \in \text{FR}_\ell(\mathcal{K}^*).
\]

The minimal face also has the following well-known characterization, see for instance, Proposition 3.2.2 in [27].

**Proposition 9** (Characterizations of the minimal face). Let \(\mathcal{F}\) be a face of \(\mathcal{K}\) containing \(\mathcal{F}^S_B\). Suppose \(\mathcal{F}\) and \(\mathcal{F}^S_B\) are both non-empty. Then the conditions below are equivalent.

1. \(\mathcal{F}^S_B \cap \text{ri} \mathcal{F} \neq \emptyset\).
2. \(\text{ri} \mathcal{F}^S_B \subseteq \text{ri} \mathcal{F}\).
3. \(\mathcal{F} = \mathcal{F}^S_B \cap \mathcal{F}^S_{\text{min}}\).

Facial reduction is a very powerful procedure and it can be used to solve feasibility problems over arbitrary closed convex cones. The computationally challenging part of doing facial reduction is computing \(d\). In general, it is necessary to solve another CLP in order to find \(d\). Therefore, at first glance, it might seem that we are again stuck solving an CLP that might also not satisfy Slater’s condition.

However, an important point is that it is possible to carefully formulate the problem of seeking \(d\) as a linear conic program having relative interior feasible points at both its primal and dual sides, as we will show shortly. In other words, even if the original CLP does not satisfy Slater’s condition, the problem of searching for \(d\) can always be done by solving problems that do indeed satisfy Slater’s condition.

The search for efficient ways of doing facial reduction and computing the reducing directions is an area of active research. Permenter, Friberg and Andersen have recently shown that reducing directions can be obtained naturally if we have access to relative interior solutions to a certain self-dual homogeneous model of \((\text{Conic-P})\) and \((\text{Conic-D})\), see Theorem 3.2 and Section 4 of [32].

It is also possible to relax the search criteria in order to make the problem of finding \(d\) more tractable by considering, for example, polyhedral approximations as in the Partial Facial Reduction approach of Permenter and Parrilo [33] or relaxing the definition of reducing direction as in the approach by Friberg [8] by removing the conic constraints. See also the work of Zhu, Pataki and Tran-Dinh for an heuristic facial reduction algorithm for SDPs in primal standard format [23]. In the case of [33] and [23], the drawback is that facial reduction might end with a face other than \(\mathcal{F}^S_{\text{min}}\), although their experiments show that many interesting instances become easier to solve nonetheless. For the approach in [8], there are some representability issues affecting the cones obtained by intersecting \(\mathcal{K}\) with the hyperplanes defined by the reducing directions, see Sections 4 and 6 therein.

The next lemma shows how to formulate the problem of finding a reducing direction as a problem that can be solved with \(\mathcal{O}\_m\), when \(\mathcal{K} = \mathcal{S}^n_T\). We remark that, for SDPs, Cheung, Schurr and Wolkowicz also discuss an auxiliary problem that is primal and dual strongly feasible, see the problem \((\text{AP})\) in [6]. However, \((\text{AP})\) uses an additional second order cone constraint. Since a second order cone constraint can be transformed to a semidefinite constraint, \((\text{AP})\) is also an SDP, when \(\mathcal{K} = \mathcal{S}^n_T\). However, \((\text{AP})\) is not readily generalizable to other families of cones that are not able to express second order cone constraints.
Lemma 10 (Finding a reducing direction through strongly feasible auxiliary problems). Let $e \in \text{ri} \mathcal{K}$, $e^* \in \text{ri} \mathcal{K}^*$ and consider the following pair of primal and dual problems.

\[
\begin{align*}
\inf_{x,t,w} & \quad t & \quad (P_{\mathcal{K}}) \\
\text{subject to} & \quad -\langle c, x - te^* \rangle + t - w = 0 & (10) \\
& \quad \langle e, x \rangle + w = 1 & (11) \\
& \quad Ax - tAe^* = 0 & (12) \\
& \quad (x,t,w) \in \mathcal{K}^* \times \mathbb{R}_+ \times \mathbb{R}_+ \\
\end{align*}
\]

\[
\begin{align*}
\sup_{y_1,y_2,y_3} & \quad y_2 & \quad (D_{\mathcal{K}}) \\
\text{subject to} & \quad cy_1 - ey_2 - A^*y_3 \in \mathcal{K} & (13) \\
& \quad 1 - y_1(1 + \langle c,e^* \rangle) + \langle e^*, A^*y_3 \rangle \geq 0 & (14) \\
& \quad y_1 - y_2 \geq 0 & (15) \\
\end{align*}
\]

The following properties hold.

(i) Both $(P_{\mathcal{K}})$ and $(D_{\mathcal{K}})$ are strongly feasible.

Let $(x^*, t^*, w^*)$ be an optimal solution to $(P_{\mathcal{K}})$ and $(y_1^*, y_2^*, y_3^*)$ be an optimal solution to $(D_{\mathcal{K}})$.

(ii) The primal optimal value $\theta_{P_{\mathcal{K}}}$ is zero if and only if $F_{\text{min}}^D \subseteq \mathcal{K}$. In this case, one of the two alternatives below must hold:

(a) $\langle c, x^* \rangle < 0$ and $(c + \text{range } A^*) \cap \mathcal{K} = \emptyset$ (i.e., $(\text{Conic-D})$ is infeasible), or

(b) $\langle c, x^* \rangle = 0$ and $(c + \text{range } A^*) \cap \mathcal{K} \subseteq \mathcal{K} \cap \{x^* \} \perp \subseteq \mathcal{K}$.

(iii) The primal optimal value $\theta_{P_{\mathcal{K}}}$ is positive if and only if $(\text{Conic-D})$ is strongly feasible, i.e., $F_{\text{min}}^D = \mathcal{K}$. In this case, we have

\[
e - A^* y_3^* y_1^* \in \text{ri } \mathcal{K}.
\]

Proof. (i) Let

\[
t := \frac{1}{\langle e, e^* \rangle + 1}, \quad w := \frac{1}{\langle e, e^* \rangle + 1}, \quad x := \frac{e^*}{\langle e, e^* \rangle + 1}.
\]

Then $(x,t,w)$ is a feasible solution to $(P_{\mathcal{K}})$ satisfying Slater’s condition, i.e.,

\[
(x,t,w) \in \text{ri } (\mathcal{K} \times \mathbb{R}_+ \times \mathbb{R}_+) = \text{ri } \mathcal{K} \times \text{ri } \mathbb{R}_+ \times \text{ri } \mathbb{R}_+.
\]

Next, we observe that $(y_1^*, y_2^*, y_3^*) := (0, -1, 0)$ is a feasible solution to $(D_{\mathcal{K}})$ such that $(14)$, $(15)$ are satisfied strictly and

\[
cy_1 - ey_2 - A^* y_3 = e \in \text{ri } \mathcal{K}.
\]

We have thus shown that both $(\text{Conic-P})$ and $(\text{Conic-D})$ are strongly feasible.

(ii) First, let $(x^*, t^*, w^*)$ be an optimal solution to $(\text{Conic-P})$ and suppose that $\theta_{P_{\mathcal{K}}}$ is zero. Since $\theta_{P_{\mathcal{K}}}$ is assumed to be zero, we have $t^* = 0$. Then, $(10)$ and $(12)$ together with $x^* \in \mathcal{K}^*$ and $w^* \geq 0$ imply that

\[
x^* \in \text{ker } A \cap \mathcal{K}^*, \quad \langle c, x^* \rangle \leq 0.
\]

Then, we have two possibilities.
(a) Suppose $\langle c, x^* \rangle < 0$. We will show that $(\text{Conic-D})$ must be infeasible. Let $s \in (c + \text{range } \mathcal{A}^*)$, then (16) implies $\langle s, x^* \rangle < 0$. Since $x^* \in K^*$, we conclude that $s$ cannot belong to $K$, because otherwise we would have $\langle x^*, s \rangle \geq 0$.

Therefore, $(\text{Conic-D})$ must be infeasible and $(c + \text{range } \mathcal{A}^*) \cap K = \emptyset$. In this case, we have $\mathcal{F}_\text{min}^D = \emptyset$ and, indeed, $\mathcal{F}_\text{min}^D \subseteq K$.

(b) Suppose $\langle c, x^* \rangle = 0$. This, together with (16) implies that

$$(c + \text{range } \mathcal{A}^*) \cap K = \mathcal{F}_\text{S}^D \subseteq K \cap \{x^* \}^\perp.$$ 

Next, we will check that the inclusion $K \cap \{x^* \}^\perp \subseteq K$ is indeed proper. We observe that since $t^* = 0$ and $\langle c, x^* \rangle = 0$, (10) implies that $w^* = 0$ too. Therefore, (11) implies that $\langle c, x^* \rangle = 1$. In particular, $x^*$ does not belong to $K^\perp$. In other words,

$$K \cap \{x^* \}^\perp \subseteq K.$$ 

Since $\mathcal{F}_\text{min}^D \subseteq K \cap \{x^* \}^\perp$, we also have $\mathcal{F}_\text{min}^D \subseteq K$.

Now, we will prove the converse. That is, we will suppose that $\mathcal{F}_\text{min}^D \subseteq K$ and we will show that $\theta_{P_K} = 0$. We start by observing that since the objective function of $(P_K)$ is “$t$” and $t$ is constrained to be nonnegative, if we exhibit a feasible solution for $(P_K)$ having $t = 0$ this would be enough to show that $\theta_{P_K} = 0$.

Since $\mathcal{F}_\text{min}^D \subseteq K$, $(\text{Conic-D})$ is not strongly feasible. By Lemma 7, there exists some $x \in K^* \cap \ker A$ such that either (a) $\langle c, x \rangle = 0$ and $x \not\in K^\perp$ or (b) $\langle c, x \rangle < 0$. Let us check each case.

(a) Suppose $\langle c, x \rangle = 0$ and $x \not\in K^\perp$. Then the condition $x \not\in K^\perp$ implies that $\langle c, x \rangle > 0$, by item (v) of Lemma 3. Then,

$$\left( \frac{x}{\langle e, x \rangle}, 0, 0 \right)$$

is a feasible solution for $(P_K)$, which shows that $\theta_{P_K} = 0$.

(b) Suppose that $\langle c, x \rangle < 0$. We define

$$\alpha := \frac{1}{\langle e, x \rangle - \langle c, x \rangle}$$

and this is well-defined because $-\langle c, x \rangle > 0$ and $\langle e, x \rangle \geq 0$. Then $(x\alpha, 0, -\alpha \langle c, x \rangle)$ is a feasible solution to $(P_K)$, which also shows that $\theta_{P_K} = 0$.

(iii) Since $(P_K)$ and $(D_K)$ both satisfy Slater’s condition, we have, in particular, that $\theta_{P_K} = \theta_{D_K}$ and there is an optimal solution to $(D_K)$ $(y_1^*, y_2^*, y_3^*)$ satisfying $y_2^* = \theta_{P_K}$. By item (ii), we have that $\theta_{P_K} = 0$ if and only if $\mathcal{F}_\text{min}^D \subseteq K$, which happens if and only if $(\text{Conic-D})$ is not strongly feasible, by Proposition 9. As $\theta_{P_K}$ is always nonnegative (because $t$ is constrained to be nonnegative), we conclude that $(\text{Conic-D})$ is strongly feasible if and only if $\theta_{P_K}$ is positive.

Next, suppose that $\theta_{P_K}$ is indeed positive. In this case we have that $y_2^* = \theta_{P_K}$ is positive and that

$$ey_3^* \in \text{ri } K,$$

since $e \in \text{ri } K$. This fact, together with (13) and item (iv) of Lemma 3, implies that $z$

$$ey_1^* - \mathcal{A}^* y_3^* \in \text{ri } K.$$ 

To conclude, we observe that (15) implies that $y_1^* \geq y_2^* > 0$. Therefore,

$$c - A y_3^* y_1^* \in \text{ri } K$$

Using Proposition 9, we conclude that indeed $\mathcal{F}_\text{min}^D = K$. 

}\]
Remark. Lemma 10 holds for any pair of $e, e^*$ satisfying $e \in \text{ri} \mathcal{K}, e^* \in \text{ri} \mathcal{K}^*$. When $\mathcal{K} = \mathcal{S}^n_+$, we may take $e$ and $e^*$ to be, for example, both equal to the $n \times n$ identity matrix. If $\mathcal{K}$ is some face of $\mathcal{S}^n_+$, Proposition 5 together with (3) and (4) indicate how to find $e$ and $e^*$ without much effort. One can take $e = e^*$ and let $e$ be such that $Q^T e Q = \left( \begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix} \right)$, where $I_r$ is the $r \times r$ identity matrix.

In [19], we proposed “FRA-Poly”, a two-phase facial reduction algorithm that takes into consideration the presence of polyhedral faces in the face lattice of $\mathcal{K}$. Instead of performing facial reduction until Slater’s condition is satisfied, Phase 1 of the algorithm in [19] regularizes the problem until the so-called partial polyhedral Slater’s condition is satisfied. Then, in Phase 2, the algorithm jumps directly to the minimal face. An extension of Lemma 10 appropriate for FRA-Poly is proved in Lemma 3 of [19].

With the aid of Lemma 10, we now are able to provide a facial reduction algorithm that can be easily adapted to use the oracle $\mathcal{O}_{\text{inst}}$, when $\mathcal{K} = \mathcal{S}^n_+$.

Algorithm 1: Facial reduction with strongly feasible auxiliary problems

**Input:** $\mathcal{K}, \mathcal{A}, c$

**Output:** Reducing directions $d_1, \ldots, d_i$ for (Conic-D) (Definition 8) together with Feasible or Infeasible. If Feasible, a pair $(s, y)$ is also returned so that $s = c - A^* y \in \text{ri} (\mathcal{K} \cap \{d_1\}^\perp \cap \cdots \cap \{d_i\}^\perp)$.

1. $\mathcal{F}_1 \leftarrow \mathcal{K}, i \leftarrow 1$.
2. Replace $\mathcal{K}, \mathcal{K}^*$ by $\mathcal{F}_i, \mathcal{F}_i^*$ in $(\mathcal{D}_\mathcal{K})$ and $(\mathcal{P}_\mathcal{K})$, respectively. Solve the resulting pair of problems and denote the obtained optimal solutions by $(x^*, t^*, w^*)$ and $(y_1^*, y_2^*, y_3^*)$.
3. if $t^* = 0$ then
   4. $d_i \leftarrow x^*$ /* Found a reducing direction */
   5. if $\langle c, x^* \rangle < 0$ then
   6. $\mathcal{F}_{\text{min}} \leftarrow \emptyset$ /* $\langle c, x^* \rangle < 0$ attests that (Conic-D) is infeasible */
   7. return Infeasible, $\mathcal{F}_{\text{min}}^D, d_1, \ldots, d_i$
   8. else
   9. $\mathcal{F}_{i+1} \leftarrow \mathcal{F}_i \cap \{d_i\}^\perp$ /* In this case we have $\langle c, x^* \rangle = 0$ */
   10. $i \leftarrow i + 1$
   11. go to line 2
12. end
13. else
14. $\mathcal{F}_{\text{min}}^D \leftarrow \mathcal{F}_i$, /* Found the minimal face */
15. $s \leftarrow c - A^* \frac{y_2^*}{y_1^*}$ /* $s \in \text{ri} \mathcal{F}_i$ */
16. return Feasible, $\mathcal{F}_{\text{min}}^D, d_1, \ldots, d_i, (s, \frac{y_2^*}{y_1^*})$
17. end

Proposition 11 (Algorithm 1 is correct). Algorithm 1 correctly detects whether (Conic-D) is feasible or not. If (Conic-D) is feasible, Algorithm 1 correctly identifies $\mathcal{F}_{\text{min}}^D$ and the pair $(s, y)$ returned by Algorithm 1 does indeed satisfy $s = c - A^* y \in \text{ri} \mathcal{F}_{\text{min}}^D$.

Proof. The correctness of Algorithm 1 is a consequence of Lemma 10 and we will now explain some of the details. We have several claims.

**Claim 1** For all $i$, $\mathcal{F}_i$ contains $(c + \text{range } A^*) \cap \mathcal{K}$ and $\mathcal{F}_{i+1}$ is strictly contained in $\mathcal{F}_i$.

This claim holds by induction. When Algorithm 1 starts, we have $\mathcal{F}_1 = \mathcal{K}$. Now, suppose that for some $i$ we have that $\mathcal{F}_i$ contains $(c + \text{range } A^*) \cap \mathcal{K}$. Given $\mathcal{F}_i$, we have that $\mathcal{F}_{i+1}$ is constructed by the relation $\mathcal{F}_{i+1} = \mathcal{F}_i \cap \{d_i\}^\perp$. 

14
However, $\mathcal{F}_{i+1}$ is only computed if the optimal value of $(P_K)$ is 0 and $\langle c, x^* \rangle = 0$, see Lines 5 and 9. In this case, item (ii)(b) of Lemma 10 ensures

$$
(c + \text{range } A^*) \cap \mathcal{F}_i \subseteq \mathcal{F}_{i+1} \subseteq \mathcal{F}_i.
$$

(17)

Since $\mathcal{F}_i$ is a face (and, therefore, a subset) of $\mathcal{K}$, the hypothesis that $\mathcal{F}_i$ contains $(c + \text{range } A^*) \cap \mathcal{K}$ implies that, in fact,

$$(c + \text{range } A^*) \cap \mathcal{F}_i = (c + \text{range } A^*) \cap \mathcal{K},$$

which, combined with (17), implies that $\mathcal{F}_{i+1}$ must also contain $(c + \text{range } A^*) \cap \mathcal{K}$. This concludes the proof of Claim 1.

**Claim 2** The minimal face of $\mathcal{F}_i$ containing $(c + \text{range } A^*) \cap \mathcal{K}$ coincides with $\mathcal{F}^D_{\text{min}}$.

**Claim 2** follows from Claim 1 and the fact that if $\mathcal{F}$ is a face of $\mathcal{K}$ and $\tilde{\mathcal{F}}$ is a face of $\mathcal{F}$, then $\tilde{\mathcal{F}}$ is a face of $\mathcal{K}$.

**Claim 3** For all $i$, $(c + \text{range } A^*) \cap \mathcal{F}_i = \emptyset$ holds if and only if (Conic-D) is infeasible. Claim 3 is a consequence of Claim 1.

Now, Claim 1 implies that whenever a new face $\mathcal{F}_{i+1}$ is computed, it must be strictly smaller than $\mathcal{F}_i$ and, therefore, the dimension must also be strictly smaller$^2$. Since we cannot have an infinite strictly descending of faces, at some point, the optimal value of $(P_K)$ must become positive or a certificate that $(c + \text{range } A^*) \cap \mathcal{F}_i = \emptyset$ will be found (see Lines 5 and 6). In the first case, Claim 2 together with item (iii) of Lemma 10 (applied to $\mathcal{F}_i$) implies that $\mathcal{F}^D_{\text{min}} = \mathcal{F}_i$ and that

$$s = c - A^* y \in \text{ri } \mathcal{F}^D_{\text{min}},$$

where $y = y^*_3/y^*_1$. In the second case, Claim 3 and item (ii)(a) of Lemma 10 ensures that, indeed, (Conic-D) must be infeasible.

Next, we examine the computational cost of Algorithm 1. The same iteration analysis that appear in other facial reduction approaches (e.g., [28, 45]) is also valid here, but, nevertheless, we will explain the details. When Algorithm 1 is invoked, a chain of faces of $\mathcal{K}$ is constructed as follows

$$\mathcal{K} = \mathcal{F}_1 \supseteq \cdots \supseteq \mathcal{F}_\ell.$$

We recall that if $\mathcal{F}, \tilde{\mathcal{F}}$ are faces of $\mathcal{K}$ such that $\mathcal{F} \subseteq \tilde{\mathcal{F}}$, then $\mathcal{F} \neq \tilde{\mathcal{F}}$ if and only if $\dim \mathcal{F} < \dim \tilde{\mathcal{F}}$. As $\mathcal{K}$ is finite dimensional, we conclude that at most $\dim \mathcal{K} + 1$ faces will be found when Algorithm 1 is invoked. This estimate can be sharpened in several different ways. For example, let $\ell_{\mathcal{K}}$ denote the longest chain of strictly decreasing non-empty faces of $\mathcal{K}$. Then, the number of non-empty faces that will be found when Algorithm 1 is invoked is bounded above by $\ell_{\mathcal{K}}$. In particular, when $\mathcal{K} = S^n_+$, we have

$$\dim \mathcal{K} = \frac{n(n+1)}{2}, \quad \ell_{S^n_+} = n + 1.$$

This shows that, in some cases, $\ell_{\mathcal{K}}$ can be a much better bound than $\dim \mathcal{K}$. For a proof that $\ell_{S^n_+} = n + 1$ see, for example, Theorem 14 in [12] where it is shown that that whenever $\mathcal{K}$ is a symmetric cone (homogeneous self-dual cone), we have $\ell_{\mathcal{K}} = \text{rank } \mathcal{K} + 1$, where $\text{rank } \mathcal{K}$ is the Jordan algebraic rank of $\mathcal{K}$. We summarize this discussion in the next proposition.

**Proposition 12** (Computational cost of Algorithm 1). The number of times that Algorithm 1 solves the pair $(P_K)$ and $(D_K)$ is bounded above by $\ell_{\mathcal{K}}$. In particular, when $\mathcal{K} = S^n_+$, Algorithm 1 can be implemented by invoking $O_{\text{int}}$ at most $n + 1$ times.

**Proof.** In the proof of Proposition 11, we have shown that Algorithm 1 constructs a strictly nondecreasing chain of faces as follows

$$\mathcal{K} = \mathcal{F}_1 \supseteq \cdots \supseteq \mathcal{F}_\ell.$$

(18)

$^2$The fact that $\mathcal{F}_i$ and $\mathcal{F}_{i+1}$ are faces is important, because, in general, $C_1 \subsetneq C_2$ does not imply $\dim C_1 < \dim C_2$. 

15
We divide the proof in two cases. Suppose first that (Conic-D) is feasible. Then, \( F_\ell = F^D_{\text{min}} \) by Proposition 11 and \( F^D_{\text{min}} \) is not empty. Finding a new face \( F_\ell \) in Algorithm 1 corresponds to solving the pair \((P_K)\) and \((D_K)\) once. So, after solving the pair \((P_K)\) and \((D_K)\) at most \( \ell_K - 1 \) times, Algorithm 1 will set \( F_\ell \) to \( F^D_{\text{min}} \). Then, \((P_K)\) and \((D_K)\) will be solved one extra time in order to check that \( F_\ell \) is indeed the minimal face and to obtain \( s \in \text{ri} \mathcal{K} \), as in Lines 14 and 15. In total, \((P_K)\) and \((D_K)\) is solved at most \( \ell_K \) times.

Next, suppose that (Conic-D) is infeasible. In this case, the last face \( F_\ell \) will be empty (see Line 6), but all faces up to \( \ell - 1 \) will be nonempty. Therefore, \( \ell - 1 \leq \ell_K \). As in the previous case, each face in the chain \((18)\) corresponds to solving the pair \((P_K)\) and \((D_K)\) once. In summary, after solving \((P_K)\) and \((D_K)\) at most \( \ell_K - 1 \) times, Algorithm 1 will find the last nonempty face \( F_{\ell-1} \) and, then, \((P_K)\) and \((D_K)\) will be solved once more in order to set \( F_\ell \) to “\( \emptyset \)”,

To conclude, we suppose that \( \mathcal{K} = S^m_{\mathcal{K}} \). Then, Algorithm 1 successively solves the problem \((P_K)\) and \((D_K)\) over \( S^m_{\mathcal{K}} \) and its faces at most \( \ell_{S^m_{\mathcal{K}}} = n + 1 \) times. By Lemma 10 and Proposition 6, these are strongly feasible problems that can be solved by invoking \( O_{\text{int}} \) a single time.

We mention in passing that the minimal number of facial reduction steps needed to find the minimal face of (Conic-D) is often called the singularity degree of (Conic-D). The singularity degree is also bounded by \( \ell_K \), but sharper estimates can be obtained by considering facial reduction strategies that take into account the existence of polyhedral faces of \( \mathcal{K} \) as in the case of the FRA-Poly algorithm in [19].

To conclude this section, we emphasize that what sets Algorithm 1 apart from other facial reduction approaches in the literature is that all auxiliary problems that must be solved are always primal and dual strongly feasible and that only linear constraints are added, in addition to conic constraints originating from the original problem.

## 4 Double facial reduction

In Section 3, we saw how to perform facial reduction by solving auxiliary problems that are always primal and dual strongly feasible. However, as we remarked previously, facial reduction only guarantees that one side of the problem will be strongly feasible, after reformulating the problem over the minimal face. In order to finally obtain a problem where both the primal and dual are strongly feasible, we need to do facial reduction twice. We call this double facial reduction.

In this section, we will discuss technical aspects related to double facial reduction and how it can be used to compute the optimal value of (Conic-D). Double facial reduction will also enable us to compute almost optimal solutions when the optimal value of (Conic-D) is not attained, as we will see in Section 4.1. We will also show how to obtain almost feasible solutions when (Conic-D) is weakly infeasible, see Section 4.2.

Now, suppose that (Conic-D) is feasible. Then, after applying facial reduction to (Conic-D) we obtain the following pair of CLPs:

\[
\begin{align*}
\inf_x \langle c, x \rangle & \quad \text{(P)} & \sup_y \langle b, y \rangle & \quad \text{(D)} \\
\text{subject to} \quad Ax = b & \quad \text{subject to} \quad c - A^* y \in F^D_{\text{min}},
\end{align*}
\]

Here, (D) is strongly feasible (i.e., Slater’s condition is satisfied), but it could still be the case that (P) does not satisfy Slater’s condition. Therefore, when \( \mathcal{K} = S^m_{\mathcal{K}} \) the pair (P) and (D) might still not be solvable with \( O_{\text{int}} \). To remedy this issue, it is reasonable to consider applying facial reduction to (P), which leads to the following pair of problems:

\[
\begin{align*}
\inf_x \langle c, x \rangle & \quad \text{(P*)} & \sup_y \langle b, y \rangle & \quad \text{(D*)} \\
\text{subject to} \quad Ax = b & \quad \text{subject to} \quad c - A^* y \in (F^D_{\text{min}})^*.
\end{align*}
\]
Here, \( F^0_{\text{min}} \) is the minimal face of \( (F^{D}_{\text{min}})^* \) which contains the feasible region of \((\hat{P})\). Now, both \((P^*)\) and \((\hat{D})\) are ensured to satisfy Slater’s condition, but it is not obvious at all whether \((D^*)\) still satisfies Slater’s condition, since \( F^{D}_{\text{min}} \subseteq (F^0_{\text{min}})^* \). In principle, it seems possible that \((D^*)\) might lose strong feasibility because, in general, \( C_1 \subseteq C_2 \) does not imply \( \ri C_1 \subseteq \ri C_2 \).

Nevertheless, we will show in this section that, in fact, if \((\text{Conic-D})\) and \((\text{Conic-P})\) are both feasible, then \((D^*)\) will still be strongly feasible. In essence, the question boils down to understanding the possible ways that the feasibility properties of \((\text{Conic-D})\) might change when a single facial reduction step is performed at \((\text{Conic-P})\) and \( K \) is replaced by \((K^* \cap \{ f \}^\perp)^* \) in \((\text{Conic-D})\), for some \( f \in K \cap \text{range} \ A^* \). First, we need a few auxiliary results.

**Lemma 13.** Let \( u \in \ri K, \ d \in K \) and \( v \in T_d K \), where \( T_d K \) is the tangent space of \( K \) at \( d \). Then, there is \( t > 0 \) such that \( u + v + td \in \ri K \).

The intuition for Lemma 13 is as follows. If \( v + td \) were a point in \( K \), then it would be clear that \( u + v + td \in \ri K \), by item \((iv)\) of Lemma 3. Unfortunately, this does not happen in general. However, as \( t \) increases, \( v + td \) gets closer and closer to \( K \), so adding \( u \) will eventually drag everything to the relative interior.

**Proof.** Let \( C = \{ u + v + td \mid t \geq 0 \} \).

To prove the lemma, it is enough to show that \( \ri C \cap \ri K \neq \emptyset \). Suppose, for the sake of obtaining a contradiction, that \( \ri C \cap \ri K = \emptyset \). This implies that there is a separating hyperplane

\[
H = \{ w \in E \mid \langle z, w \rangle = \theta \}
\]

such that \( H \) properly separates \( C \) and \( K \), see Theorem 11.3 in [38]. We recall that **proper separation** means that \( C \) and \( K \) lie in opposite closed half-spaces and \( H \) does not contain both at the same time. We have

\[
\langle u, z \rangle + \langle v, z \rangle + \langle td, z \rangle \leq \theta \leq \langle w, z \rangle
\]

for all \( t \geq 0 \) and all \( w \in K \). For the inequality above to hold, we must have \( z \in K^* \) and \( \theta \leq 0 \). Since \( d \in K \) and \( z \in K^* \) we must have

\[
\langle d, z \rangle = 0,
\]

since \( t \) can be taken to be any nonnegative number. Therefore, \( z \in F(d, K)^\Delta \). Because \( T_d K = F(d, K)^\Delta \) (item \((x)\) of Lemma 3), we obtain that \( \langle v, z \rangle = 0 \).

Gathering all we have, we conclude that \( \langle u, z \rangle = 0 \), so that \( C \subseteq H \) and \( \theta = 0 \). By item \((vi)\) of Lemma 3, we have \( K \subseteq H \) as well. This contradicts the properness of the separation. \( \square \)

We are now ready to state a result on conservation of feasibility after one facial reduction step. For what follows, we say that \((\text{Conic-D})\) is **in weak status** if it is weakly feasible or weakly infeasible. We also recall the following basic facts. A face \( F \) of \( K \) always satisfies \( F = K \cap \text{span} F \), therefore we have

\[
F^* = \text{cl}(K^* + F^\perp),
\]

(19)

Also, if \( C_1 \) and \( C_2 \) are two convex sets we have \( \ri (C_1 + C_2) = \ri C_1 + \ri C_2 \).

**Proposition 14** (Conservation of feasibility). Let \( f \in K \cap \text{range} A^* \) and let \( F := K^* \cap \{ f \}^\perp = F(K, f)^\Delta \) (see item \((viii)\) of Lemma 3). Let \((D')\) be the problem obtained by replacing \( K \) by \( F^* \) in \((\text{Conic-D})\), i.e.,

\[
\sup_y \langle b, y \rangle
\]

subject to \( c - A^* y \in F^* \).

We have the following relations:
(i) (Conic-D) is strongly feasible if and only if (D') is;
(ii) (Conic-D) is strongly infeasible if and only if (D') is;
(iii) (Conic-D) is in weak status if and only if (D') is.

**Proof.** (i) First, since $F^* = \text{cl}(K + F^\perp)$ and $\text{ri} F^\perp = F^\perp$, we have

$$\text{ri} F^* = \text{ri}(\text{cl}(K + F^\perp)) = (\text{ri} K) + F^\perp. \quad (20)$$

Now, suppose that (Conic-D) is strongly feasible. Since $\text{ri} K \subseteq F^\perp + \text{ri} K^*$, we conclude that (D') must be strongly feasible as well.

Conversely, suppose that $s = c - A^* y$ is a relative interior feasible solution for (D'), i.e., $s \in \text{ri} F^*$. By (20), we have

$$s = u + v,$$

where $u \in \text{ri} K$ and $v \in F^\perp$. By item (x) of Lemma 3, we have $F^\perp = T_f K$. Invoking Lemma 13 we conclude that there exists $t > 0$ such that $u + v + tf \in \text{ri} K^*$. As $f \in \text{range} A^*$, there exists $\hat{y}$ such that $f = -A^* \hat{y}$. We conclude that

$$s + tf = u + v + tf = c - A^*(y + \hat{y}) \in \text{ri} K,$$

thus showing that (Conic-D) is strongly feasible.

(ii) Since $K \subseteq F^*$, if (D') is strongly infeasible, (Conic-D) must be strongly infeasible as well.

Conversely, suppose that (Conic-D) is strongly infeasible. Then, Proposition 4 implies the existence of $x$ satisfying

$$\langle c, x \rangle = -1, \quad x \in K^* \cap \ker A.$$

Since $x \in \ker A$ and $f \in \text{range} A^*$, we have $\langle x, f \rangle = 0$. So, in fact, $x \in F$. Therefore, by Proposition 4, the same $x$ attests that (D') is strongly infeasible.

(iii) Follows by elimination.

We can now state and prove our main result on the preservation of feasibility status after facial reduction is performed on (Conic-P). Intuitively, Theorem 15 means following. Whenever facial reduction is applied to, say, (Conic-P), we obtain a new problem which is ensured to satisfy Slater’s condition, if (Conic-P) is feasible. This new problem will also have a dual problem whose feasibility properties might be different than the original dual problem (Conic-D). However, Theorem 15 says that no drastic changes are allowed, i.e., if (Conic-P) was strongly feasible to begin with, it will stay strongly feasible. The only possible room for change is that a weakly feasible/infeasible problem might become weakly infeasible/feasible. Theorem 15 also contains the relatively surprising fact that strong feasibility of the new dual implies strong feasibility of (Conic-D).

**Theorem 15** (Preservation of feasibility under facial reduction). Let $F^\text{P}_{\text{min}}$ denote the minimal face of $K^*$ that contains the feasible region of (Conic-P) and suppose that $F^\text{P}_{\text{min}} \neq \emptyset$. Consider the problem obtained by replacing $K$ by $(F^\text{P}_{\text{min}})^*$ in (Conic-D), i.e.,

$$\sup_y \langle b, y \rangle \quad \text{subject to} \quad c - A^* y \in (F^\text{P}_{\text{min}})^*,$$

where $(\text{Conic-D-FP})$

The following hold.

(i) (Conic-D) is strongly feasible if and only if (Conic-D-FP) is strongly feasible;
(ii) (Conic-D) is strongly infeasible if and only if (Conic-D-FP) is strongly infeasible;

(iii) (Conic-D) is in weak status if and only if (Conic-D-FP) is in weak status.

Proof. Applying facial reduction to (Conic-P) (e.g., Algorithm 1), we see that $\mathcal{F}_{\min}^P$ can be written as

$$\mathcal{F}_{\min}^P = \mathcal{K}^* \cap \{d_1\}^* \cap \cdots \cap \{d_\ell\}^*,$$

where each $d_i$ satisfies

$$d_i \in (\mathcal{K}^* \cap \{d_1\}^* \cap \cdots \{d_{i-1}\}^*)^* \cap \text{range} \mathcal{A}^*.$$

Now, denote by $(D_i)$ the problem obtained by replacing $\mathcal{K}$ by $(\mathcal{K}^* \cap \{d_1\}^* \cap \cdots \{d_{i-1}\}^*)^*$ in (Conic-D).

We observe the following:

1. $(D_1)$ and $(D_{\ell+1})$ are precisely (Conic-D) and (Conic-D-FP), respectively.

2. $d_i$ is a reducing direction for $(D_i)$, so Proposition 14 applies to $d_i$, $(D_i)$ and $(D_{i+1})$, for $i = 1, \ldots, \ell$.

By induction, we conclude that items (i), (ii) and (iii) hold.

We are now in position to state our main result on double facial reduction.

**Theorem 16 (Double facial reduction).** Suppose $\mathcal{F}_{\min}^D \neq \emptyset$ and consider the problems $(\hat{P})$ and $(\hat{D})$ above.

Let $\mathcal{F}_{\min}^\hat{P}$ be the minimal face of $(\mathcal{F}_{\min}^D)^*$ that contains the feasible region of $(\hat{P})$. Consider the pair of problems $(P^*)$ and $(D^*)$, which we repeat below for convenience.

$$\inf_x \langle c, x \rangle \quad \text{(P*)} \quad \sup_y \langle b, y \rangle \quad \text{(D*)}$$

subject to $\mathcal{A}x = b$

$$x \in \mathcal{F}_{\min}^\hat{P}$$

subject to $c - \mathcal{A}^* y \in (\mathcal{F}_{\min}^\hat{P})^*$.

The following hold.

(i) The optimal value of (Conic-D) ($\theta_D$) is finite if and only if $\mathcal{F}_{\min}^\hat{P} \neq \emptyset$. In this case, $(P^*)$ and $(D^*)$ are both strongly feasible and

$$\theta_D = \theta_{P^*} = \theta_{D^*}.$$

(ii) $\theta_D = +\infty$ if and only if $\mathcal{F}_{\min}^\hat{P} = \emptyset$.

Proof. (i) Suppose that the optimal value of (Conic-D) is finite. Then, the optimal value of $(\hat{P})$ must be equal to $\theta_D$, since $(\hat{D})$ is strongly feasible. In particular, $(\hat{P})$ must be feasible and, therefore, $\mathcal{F}_{\min}^\hat{P} \neq \emptyset$. Since $\mathcal{F}_{\min}^\hat{P}$ is the minimal face of $(\mathcal{F}_{\min}^D)^*$ that contains the feasible region of $(\hat{P})$, $(P^*)$ is strongly feasible and its optimal value must coincide with the optimal value of $(\hat{P})$, which is $\theta_D$.

Next, since $(P^*)$ is strongly feasible and has finite optimal value, $(D^*)$ must have the same optimal value. Therefore, as stated, we have

$$\theta_D = \theta_{P^*} = \theta_{D^*}.$$

By item (i) of Theorem 15, substituting $\mathcal{F}_{\min}^D$ by $(\mathcal{F}_{\min}^\hat{P})^*$ preserves strong feasibility, so $(P^*)$ and $(D^*)$ are both strongly feasible.

Conversely, suppose that $\mathcal{F}_{\min}^\hat{P} \neq \emptyset$. This means that $(P^*)$ is feasible. So $(\hat{P})$ must be feasible as well, because any feasible solution to $(P^*)$ must be a feasible solution to $(\hat{P})$. Since we are assuming that $\mathcal{F}_{\min}^\hat{P} \neq \emptyset$, $(D)$ must be feasible as well. Therefore, $(\hat{P})$ and $(\hat{D})$ are feasible primal and dual problems sharing the same optimal value, which must be finite. Since $(D)$ shares the same optimal value with (Conic-D), we conclude that $\theta_D$ is indeed finite.
The conclusion is that, when $\theta_D$ is finite, the pair of problems $(P^*)$ and $(D^*)$ both satisfy Slater’s condition, so when $K = S^n_+$, they can indeed be solved by $O_{\text{int}}$ in order to obtain $\theta_D$. At this stage, even though $\theta_D$ might have been unattained for (Conic-D), $(D^*)$ is never hindered by unattainment.

The problem, however, is that a feasible solution to $(D^*)$ might not be feasible to (Conic-D). And, indeed, if $\theta_D$ is finite but not attained, even though $(D^*)$ has an optimal solution, (Conic-D) will not have optimal solutions. When $\theta_D$ is finite but not attained, the best we can do is to compute some solution $y$, satisfying $\langle b, y \rangle \geq \theta_D - \epsilon$, for some arbitrary $\epsilon > 0$. We will discuss this issue in the next subsection.

**Remark.** Double facial reduction is related to the concept of double reformulation presented by Pataki in Section 5 of [31] for SDPs, but it is different. In Definition 1 of [31], Pataki defined that a reformulation of (SDP-D) corresponds to the SDP primal and dual pair obtained by applying certain elementary operations to (SDP-P) and (SDP-D). These elementary operations preserve the properties of the original problem such as duality gaps and whether the optimal value is attained or not. In simplified terms, Pataki showed in Theorem 4 of [31] that (SDP-D) can be “doubly reformulated” as

$$
\sup_y \langle b', y \rangle \quad \text{(SDP-Ref)}
$$

subject to 
$$
c' - \sum_{i=1}^m A'_i y_i \in S^n_+, 
$$

where $A_i \in S^n$ for all $i$ and in such a way that $c'$ belongs to $\text{ri} \mathcal{F}^D_{\text{min}}$. Furthermore, for some $\ell$, $(c', A'_1, \ldots, A'_\ell)$ can be used to obtain the minimal face associated to the so-called “homogeneous dual” of (SDP-Ref). So, this double reformulation, in a sense, reveal both the minimal face of (SDP-Ref) and the minimal face associated to a homogenized version of the corresponding dual problem of (SDP-Ref).

As far as we could see, the homogeneous dual of (SDP-Ref) is related but it is quite different from either $(\hat{P})$ or $(P^*)$ even when $K = S^n_+$. We note that the homogeneous dual in Theorem 4 of [31] is computed with respect the cone $S^n_+$ and not with respect the minimal face associated to (SDP-Ref). Therefore, the closest analogous of the double reformulation in our setting would be if we applied the second facial reduction to (Conic-P) instead of applying to $(\hat{P})$. In conclusion, it seems to us that double facial reduction and Pataki’s double reformulation serve different purposes.

Before we move on, we give an intuitive explanation of why unattainment disappears when doing facial reduction. In fact, there are a few ways of explaining intuitively why unattainment disappears. The “dual” explanation is that Slater’s condition is satisfied at $(P^*)$ (which is the Lagrangean dual of $(D^*)$), so, of course, $(D^*)$ must be attained when $\theta_D$ value is finite.

We now give an “primal” explanation of why unattainment disappears. Suppose that the optimal value $\theta_D$ of (Conic-D) is finite but not attained. Then, there exists a sequence $\{y^k\}$ such that $\langle b, y_k \rangle \to \theta_D$ and

$$
s^k := c - A^k y^k \in K, \quad \forall k.
$$

As we are assuming unattainment and $\mathcal{F}^S_D$ is closed, $\{y^k\}$ cannot be bounded. Passing to a subsequence if necessary, we may assume that $\|y^k\| \to \infty$ and $y^k/\|y^k\|$ converges to some $y \in \mathbb{R}^n$ and that $\theta_D \geq \langle b, y \rangle$ $\theta_D - 1$ for every $k$. Dividing $c - A^k y^k \in K$ and $\theta_D \geq \langle b, y^k \rangle \geq \theta_D - 1$ by $\|y^k\|$ and taking limits, we conclude that $f := -A^* y$ satisfies

$$
f \in K, \quad f \in \text{range} \ A^*, \quad \langle b, y \rangle = 0
$$

and $y$ must be nonzero. This shows that $(f, y)$ is a reducing direction for (Conic-P), see Definition 8. In fact, with more effort, we can show that there must be at least one pair $(f, y)$ as above satisfying $f \not\in \text{lin} K$, see the remark after Lemma 7. In other words, a necessary condition for unattainment of (Conic-D) is the existence of $(f, y)$ as above with $f \not\in \text{lin} K$. Informally speaking, $(f, y)$ acts as a “recession direction” for the problem (Conic-D). This suggests that one possible way of fixing unattainment is by preventing $f$ from
Lemma 17. Suppose that $s \in (c + \text{range } \mathcal{A}^\ast) \cap \text{ri } \mathcal{K}$ and let $\mathcal{F}_{\text{min}}^P$ be the minimal face of $\mathcal{K}^\ast$ that contains the feasible region of \((\text{Conic-P})\). If $\mathcal{F}_{\text{min}}^P \neq \emptyset$, then $s \in \text{ri } ((\mathcal{F}_{\text{min}}^P)^\ast)$. 

\begin{align}
(\mathcal{F} \cap \{f\}^\perp)^\ast &= (\mathcal{F}(f, \mathcal{F}^\ast)^\Delta)^\ast = \text{cl } (\mathcal{F}^\ast + \mathcal{F}(f, \mathcal{F}^\ast)^\Delta^\perp) = \text{cl } (\mathcal{F}^\ast + T_f \mathcal{F}^\ast) \\
&= (21)
\end{align}

becoming a recession direction. This is accomplished, for instance, by substituting $\mathcal{K}$ by $\text{cl } (\mathcal{K} + \text{span } \{f\})$, so that $f \in \text{lin } (\text{cl } (\mathcal{K} + \text{span } \{f\}))$. However, $\text{cl } (\mathcal{K} + \text{span } \{f\})$ is equal to $(\mathcal{K}^\ast \cap \{f\}^\perp)^\ast$, which corresponds to a single facial reduction step done at \((\text{Conic-P})\).

In other words, from the point of view of \((\text{Conic-D})\), facial reduction done at \((\text{Conic-P})\) removes recession directions that affect attainment. We remark that Abrams [1] also proposed a regularization procedure that removes recession directions, in order to fix unattainment in convex programming.

4.1 Obtaining feasible almost optimal solutions

The pair of problems \((\text{D}^\ast)\) and \((\text{P}^\ast)\) are strongly feasible and the common optimal value is $\theta_D$. However, an optimal solution to \((\text{D}^\ast)\) is unlikely to be feasible for \((\text{Conic-D})\). In fact, it may happen that $\theta^D$ is not attained, in which case \((\text{Conic-D})\) has no optimal solution at all.

Nevertheless, we will show how to construct feasible solutions that are almost optimal for \((\text{Conic-D})\) using the directions obtained when calling Algorithm 1 with \((\hat{\text{P}})\) as input. We will denote by $y^\ast$ an optimal solution to \((\text{D}^\ast)\) and by $s^\ast$ the corresponding slack $s^\ast = c - \mathcal{A}^\ast y^\ast$.

\begin{algorithm}[h]
\caption{Finding an $\epsilon$-optimal solution to \((\text{Conic-D})\)}
\begin{algorithmic}[1]
\Input{}
1. Reducing directions for \((\hat{\text{P}})\) (Definition 8): $(f_1, y_1), \ldots, (f_{\ell_2}, y_{\ell_2})$.
2. $\hat{y}$ such that $c - \mathcal{A}^\ast \hat{y} \in \text{ri } \mathcal{F}_{\text{min}}^D$.
3. an optimal solution $y^\ast$ to \((\text{D}^\ast)\).
4. $\epsilon > 0$
\Output{} A feasible solution $y^\epsilon$ to \((\text{Conic-D})\) satisfying $\langle b, y^\epsilon \rangle \geq \theta_D - \epsilon$.
\If{$\langle b, \hat{y} \rangle \geq \theta_D - \epsilon$}
\Return{$\hat{y}$}
\Else
\State $\beta \leftarrow \frac{\theta_D - \langle b, \hat{y} \rangle - \epsilon}{\theta_D - \langle b, \hat{y} \rangle}$
\State $w_{i+1} \leftarrow \beta y^\ast + (1 - \beta)\hat{y}$
\State $\mathcal{F}_{\ell_2+1} \leftarrow ((\mathcal{F}_{\text{min}}^D)^\ast \cap \{f_1\}^\perp \cap \{f_{\ell_2}\}^\perp)^\ast$
\For{$i \leftarrow \ell_2$ to 1}
\State $\mathcal{F}_i \leftarrow ((\mathcal{F}_{\text{min}}^D)^\ast \cap \{f_1\}^\perp \cap \{f_{i-1}\}^\perp)^\ast$
\State Find $\alpha_i$ positive such that $c - \mathcal{A}^\ast (w_{i+1} + \alpha_i y_i) = c - \mathcal{A}^\ast w_{i+1} + \alpha_i f_i \in \text{ri } \mathcal{F}_i^*$
\State $w_i \leftarrow w_{i+1} + \alpha_i y_i$
\EndFor
\Return{$w_1$}
\EndIf
\end{algorithmic}
\end{algorithm}

Note that the inner loop goes from $\ell_2$ to 1. This is because we start from a relative interior solution to $\mathcal{F}_{\ell_2+1}$ and we have to work our way until the bottom of the chain $\mathcal{F}_{\text{min}}^D$. Of course, the tricky part is ensuring that at each step there is indeed an $\alpha_i$ as in Line 9. If there is at least one $\alpha_i$, then any number larger than $\alpha_i$ will work. Therefore, it is enough to keep trying larger and larger numbers until the condition in Line 9 is met. We will now show that an appropriate $\alpha_i$ always exists and that Algorithm 2 is indeed correct. For that, we need a few auxiliary results. First, suppose that $f \in \mathcal{F}^*$, for $\mathcal{F}$ a closed convex cone. We have by items \((viii)\) and \((x)\) of Lemma 3 and \((19)\) that

\begin{align}
(\mathcal{F} \cap \{f\}^\perp)^\ast &= (\mathcal{F}(f, \mathcal{F}^\ast)^\Delta)^\ast = \text{cl } (\mathcal{F}^\ast + \mathcal{F}(f, \mathcal{F}^\ast)^\Delta^\perp) = \text{cl } (\mathcal{F}^\ast + T_f \mathcal{F}^\ast)
\end{align}

(21)
By construction, we have:
$$\mathcal{F}_{\min}^P = \mathcal{K}^* \cap \{f_1\}^+ \cap \cdots \cap \{f_s\}^+,$$
where each \( f_i \) satisfies
$$f_i \in (\mathcal{K}^* \cap \{f_1\}^+ \cap \cdots \cap \{f_{i-1}\}^+)^* \cap \text{range} \mathcal{A}^*.$$

Now, let \( \mathcal{F}_1 := (\mathcal{K}^* \cap \{f_1\}^+ \cap \cdots \cap \{f_{i-1}\}^+)^* \). We observe the following:
- \( \mathcal{F}_1 = \mathcal{K}^* \) and \( \mathcal{F}_{i+1} = \mathcal{F}_{\min}^P \).
- \( \mathcal{F}_i \leftarrow \mathcal{F}_{i-1} \cap \{f_{i-1}\}^+ \), for all \( i > 1 \).

By hypothesis, we have \( s \in \text{ri} \mathcal{F}_1^* \) and by (21) we have:
$$\mathcal{F}_2^* = \text{cl} (\mathcal{F}_1^* + T_{f_1} \mathcal{F}_1^*)$$
$$\text{ri} \mathcal{F}_2^* = (\text{ri} \mathcal{F}_1^*) + T_{f_1} \mathcal{F}_1^*.$$

Therefore, \( s \in \text{ri} \mathcal{F}_2^* \) as well. At the \( i \)-th step, we have:
$$\mathcal{F}_i^* = \text{cl} (\mathcal{F}_{i-1}^* + T_{f_{i-1}} \mathcal{F}_{i-1}^*)$$
$$\text{ri} \mathcal{F}_i^* = (\text{ri} \mathcal{F}_{i-1}^*) + T_{f_{i-1}} \mathcal{F}_{i-1}^*.$$

By induction, we conclude that \( s \in \text{ri} \mathcal{F}_i^* \) for every \( i \). In particular, \( s \in \text{ri} ((\mathcal{F}_{i+1})^*) = \text{ri} ((\mathcal{F}_{\min}^P)^*) \).

**Theorem 18.** Algorithm 2 is correct, that is, the output \( y_c \) is indeed a feasible solution to (Conic-D) satisfying \( \langle b, y_c \rangle \geq \theta_D - \epsilon \).

**Proof.** \( y_c \) is \( \epsilon \)-optimal. By construction, \( \langle b, w_{\ell_2+1} \rangle \geq \theta_D - \epsilon \). Moreover, all the \( y_i \) satisfy \( \langle b, y_i \rangle = 0 \). Therefore, \( \langle b, y_c \rangle > \theta_D - \epsilon \).

\( y_c \) is feasible for (Conic-D). If the algorithm stops before Line 5, \( y_c \) is feasible because \( \mathcal{F}_{\min}^D \subseteq \mathcal{K} \). So suppose that we have reached Line 5. Since \( \mathcal{F}_1^* = \mathcal{F}_{\min}^D \), if Line 9 is correct, then \( y_c \) is feasible for (Conic-D).

We now show that Line 9 is indeed correct.

Let
$$\hat{s} := c - \mathcal{A}^* \hat{y},$$
$$s_i := c - \mathcal{A}^* w_i, \text{ for } i = 1, \ldots, \ell_2 + 1.$$

We have \( \hat{s} \in \text{ri} \mathcal{F}_{\min}^D \) and, by Lemma 17, \( \hat{s} \in \text{ri} ((\mathcal{F}_{\min}^P)^*) \) as well. Note that \( s_{\ell_2+1} \) is a strict convex combination of \( c - \mathcal{A}^* y^* \) and \( \hat{s} \). These points belong to \( (\mathcal{F}_{\min}^P)^* \) and \( (\mathcal{F}_{\min}^P)^* \), respectively, so \( s_{\ell_2+1} \) must belong to \( \text{ri} ((\mathcal{F}_{\min}^P)^*) \) as well. In addition, \( s_{\ell_2+1} \) is a feasible slack for (D*).

Now suppose that we have shown that \( s_{i+1} \in \text{ri} \mathcal{F}_i^* \), for some \( i \). By (19), we have
$$\mathcal{F}_{i+1}^* = (\mathcal{F}_i \cap \{f_i\}^+)^* = \text{cl} (\mathcal{F}_i^* + T_{f_i} \mathcal{F}_i^*)$$
$$\text{ri} \mathcal{F}_{i+1}^* = (\text{ri} \mathcal{F}_i^*) + T_{f_i} \mathcal{F}_i^*.$$

Therefore, \( s_{i+1} = u_i + v_i \) for some \( u_i \in \text{ri} \mathcal{F}_i^* \) and \( v_i \in T_{f_i} \mathcal{F}_i^* \). We can apply Lemma 13 to \( u_i, v_i, f_i, \mathcal{F}_i^* \) and conclude the existence of positive \( \alpha_i \) such that \( s_{i+1} + \alpha f_i \) belongs to \( \text{ri} \mathcal{F}_i^* \). Therefore,
$$c - \mathcal{A}^* (w_{i+1} + \alpha y_i) \in \text{ri} \mathcal{F}_i^*.$$

In other words, \( s_i = c - \mathcal{A}^* w_i \in \text{ri} \mathcal{F}_i^* \).

By induction, we conclude that at each iteration it is possible to find \( \alpha_i \) as stated in Line 9. \( \square \)
4.1.1 Computational aspects of Algorithm 2

Having proved the correctness of Algorithm 2 in Theorem 13, we discuss the computation of \( \alpha_i \) in Line 9, which is the most computationally expensive part of the algorithm. As remarked previously, the existence of \( \alpha_i \) follows from Lemma 13. So, we will discuss the computation of \( t \) as in Lemma 13.

Let \( u, v, d \) be as in Lemma 13, i.e., \( u \in \mathrm{ri} \mathcal{K}, \ d \in \mathcal{K} \) and \( v \in T_d \mathcal{K} \). As we remarked before Lemma 17, if \( t > 0 \) is such that \( u + v + td \in \mathrm{ri} \mathcal{K} \), then any \( t \geq t \) will also work. So, the simplest algorithm for computing \( t \) starts with some arbitrary positive value and keeps doubling it, until \( u + v + td \in \mathrm{ri} \mathcal{K} \).

Still, for the sake of completeness, we show that we can also obtain \( t \) by solving a pair of primal and dual strongly feasible problems. First, we consider the following pair of problems:

\[
\begin{align*}
\inf_{x} & \quad \langle u + v, x \rangle \\
\text{subject to} & \quad \langle d, x \rangle = 1 \\
& \quad x \in \mathcal{K}^* \\
\end{align*}
\]

\[
\begin{align*}
\sup_{t} & \quad -t \\
\text{subject to} & \quad u + v + td \in \mathcal{K}. \\
\end{align*}
\]

Lemma 13 guarantees that \( (D_d) \) is strongly feasible, so we can apply Lemma 10 to the pair \( (P_d) \) and \( (D_d) \) by replacing \( b, c \) by \(-1\) and \( u + v \) respectively and \( \mathcal{A}^* \) by the map that takes \( t \) to \(-td\). From item (iii) of Lemma 10, if we solve the pair \( (P_K) \) and \( (D_K) \) we will obtain \( t \) such that \( u + v + td \in \mathrm{ri} \mathcal{K} \). If \( t \) turns out to be negative, we can just set it to zero.

Therefore, it is enough to solve the pair \( (P_K) \) and \( (D_K) \), in order to obtain \( t \). In particular, this can be done with \( O_{\text{int}} \), when \( \mathcal{K} \) is \( S^n_+ \) or a face of \( S^n_+ \). This leads immediately to the following result.

**Proposition 19.** When \( \mathcal{K} = S^n_+ \), Algorithm 2 can be implemented by invoking \( O_{\text{int}} \) at most \( \ell_2 \) times, where \( \ell_2 \) is the number of reducing directions (see input 1. of Algorithm 2). In particular, since the number of reducing directions is bounded above \( n + 1 \), we also have that \( n + 1 \) is an upper bound for the number of times \( O_{\text{int}} \) is invoked.

We remark, however, that Proposition 19 seems to be exceedingly conservative. If membership in the cone \( \mathcal{K} \) can be decided efficiently, the line search problem of finding \( t \) with \( u + v + td \in \mathrm{ri} \mathcal{K} \) seems to be much cheaper than solving a conic linear program over \( \mathcal{K} \).

We also notice the following curious feature of Algorithm 2. Except for the problem of finding \( \alpha_i \) in Line 9, the complexity of Algorithm 2 does not depend on \( \epsilon \). Decreasing \( \epsilon \), however, might lead to larger \( \alpha_i \) in Algorithm 2.

4.2 Handling infeasibility

When \( (\text{Conic-D}) \) is infeasible, it can be either strongly or weakly infeasible. Strong infeasibility is relatively straightforward to analyze. Indeed, by Proposition 4, if we wish to show that \( (\text{Conic-D}) \) is strongly infeasible, it is enough to exhibit some \( x \in \mathcal{K} \cap \ker \mathcal{A} \) such that \( \langle c, x \rangle = -1 \). Therefore, in order to prove that \( (\text{Conic-D}) \) is strongly infeasible we need to solve an CLP feasibility problem. In particular, when \( \mathcal{K} = S^n_+ \), this can be done in at most \( n + 1 \) calls to \( O_{\text{int}} \), by Proposition 12.

When \( (\text{Conic-D}) \) is weakly infeasible, the situation is far more complicated. In order to prove that \( (\text{Conic-D}) \) is weakly infeasible, we have to prove that \( (\text{Conic-D}) \) is infeasible (which can also be done by Algorithm 1) and that the feasibility problem associated to strong infeasibility is infeasible, i.e., we have to show that there is no solution to

\[
\text{find } x \in \mathcal{K}^* \cap \{ x \in \ker \mathcal{A} \mid \langle c, x \rangle = -1 \}.
\]

In this subsection, we will use the techniques of Section 4.1 to analyze weak infeasibility. This is not surprising because weak infeasibility and non-attainment of optimal solutions (Section 4.1) are closely related as we will see in a moment. In fact, let \( \epsilon \in \mathrm{ri} \mathcal{K} \) and consider the following problem and its primal counterpart.
Before, we proceed, we need two preliminary results.

**Proposition 20.** If \((c + \text{range} \mathcal{A}^*) \cap \text{span} \mathcal{K} \neq \emptyset\), then (D-Feas) is strongly feasible. If \((c + \text{range} \mathcal{A}^*) \cap \text{span} \mathcal{K} = \emptyset\), then (Conic-D) is strongly infeasible.

**Proof.** Suppose that \((c + \text{range} \mathcal{A}^*) \cap \text{span} \mathcal{K} \neq \emptyset\) and let \(y\) and \(s\) be such that

\[
\begin{align*}
\inf_{x} \langle c, x \rangle & \quad (P-\text{Feas}) \\
\text{subject to} & \quad Ax = 0 \\
& \quad \langle e, x \rangle = 1 \\
& \quad x \in \mathcal{K}^* \end{align*}
\]

Then, since \(e \in \text{ri} \mathcal{K}\), there exists \(\alpha > 0\) such that \(e + \alpha s \in \text{ri} \mathcal{K}\). Since \(\mathcal{K}\) is a cone, we have \(e/\alpha + s \in \text{ri} \mathcal{K}\). Therefore, \((t, y) := (-\alpha, y)\) is a solution for (P-Feas) for which the corresponding slack \(c - te - \mathcal{A}^* y\) belongs to \(\text{ri} \mathcal{K}\), thus showing that (D-Feas) is strongly feasible.

Next, suppose that \((c + \text{range} \mathcal{A}^*) \cap \text{span} \mathcal{K} = \emptyset\). Because \(c + \text{range} \mathcal{A}^*\) and \(\text{span} \mathcal{K}\) are polyhedral sets, this implies that

\[
dist(c + \text{range} \mathcal{A}^*, \text{span} \mathcal{K}) > 0,
\]

e.g., see Corollary 19.3.3 and Theorem 11.4 in [38]. In particular, we must have \(\text{dist}(c + \text{range} \mathcal{A}^*, \mathcal{K}) > 0\) as well, thus showing that (Conic-D) is strongly infeasible.

**Lemma 21.** Let \(\mathcal{L} \subseteq \mathcal{E}\) be a subspace, \(c \in \mathcal{E}\) and \(\mathcal{K}\) be a closed convex cone such that \(\text{dist}(\mathcal{L} + c, \mathcal{K}) = 0\). Then, \((\mathcal{L} + c) \cap \text{span} \mathcal{K} \neq \emptyset\) and

\[
dist((\mathcal{L} + c) \cap \text{span} \mathcal{K}, \mathcal{K}) = 0.
\]

**Proof.** Since \(\text{dist}(\mathcal{L} + c, \mathcal{K}) = 0\), we also have \(\text{dist}(\mathcal{L} + c, \text{span} \mathcal{K}) = 0\). However, because \(\mathcal{L} + c\) and \(\text{span} \mathcal{K}\) are polyhedral sets, this implies that \((\mathcal{L} + c) \cap \text{span} \mathcal{K} \neq \emptyset\), see Corollary 19.3.3 and Theorem 11.4 in [38]. So, let \(\hat{c} \in (\mathcal{L} + c) \cap \text{span} \mathcal{K}\). We have

\[
(\mathcal{L} + c) \cap \text{span} \mathcal{K} = (\mathcal{L} \cap \text{span} \mathcal{K}) + \hat{c}.
\]

For the sake of obtaining a contradiction, assume that \(\text{dist}((\mathcal{L} + c) \cap \text{span} \mathcal{K}, \mathcal{K}) > 0\). By item (ii) of Proposition 4, there exists \(x\) such that

\[
\langle \hat{c}, x \rangle = -1, \quad x \in \mathcal{K}^* \cap ((\mathcal{L} \cap \text{span} \mathcal{K})^\perp).
\]

Therefore, \(x\) satisfies \(x = u + v\), where \(u \in \mathcal{L}^\perp\) and \(v \in \mathcal{K}^\perp\). Recall that, since \(\hat{c} \in \mathcal{L} + c\), there exists \(l \in \mathcal{L}\) such that \(\hat{c} = l + c\). We have

\[
-1 = \langle \hat{c}, x \rangle = \langle l + c, u + v \rangle = \langle c, u \rangle,
\]

because \(l + c \in \text{span} \mathcal{K}\) and \(v \in \mathcal{K}^\perp\), so \(\langle l + c, v \rangle = 0\). Furthermore, \(u \in \mathcal{K}^*\), because \(u = x - v\) and \(\mathcal{K}^\perp \subseteq \mathcal{K}^*\). Gathering all we have shown, we obtain

\[
\langle c, u \rangle = -1, \quad u \in \mathcal{K}^* \cap \mathcal{L}^\perp.
\]

Again, by item (ii) of Proposition 4, we conclude that \(\text{dist}(\mathcal{L} + c, \mathcal{K}) > 0\), which contradicts our assumptions.

**Proposition 22.** Denote by \(\theta_{D-\text{Feas}}\) the optimal value of (D-Feas). Then,

(i) \(\theta_{D-\text{Feas}} > 0\) if and only if (Conic-D) is strongly feasible.

(ii) \(\theta_{D-\text{Feas}} = 0\) if and only if (Conic-D) is in weak status (i.e., either weakly infeasible or weakly feasible).
(iii) $\theta_{D-\text{Feas}} = 0$ and is not attained if and only if (Conic-D) is weakly infeasible.

(iv) $\theta_{D-\text{Feas}} < 0$ if and only if (Conic-D) is strongly infeasible.

Proof. (i) First, suppose that (Conic-D) is strongly feasible and let $s, y$ be such that

$$s = c - A^* y \in \text{ri} K.$$ 

By hypothesis, we have $e \in \text{ri} K$. By item (v) of Lemma 3, there exists $\alpha > 1$ such that

$$\alpha s + (1 - \alpha) e \in \text{ri} K.$$ 

Therefore,

$$c - te - A^* y \in K,$$

where $t = (\alpha - 1)/\alpha$. This shows that $\theta_{D-\text{Feas}} > 0$.

Conversely, if $\theta_{D-\text{Feas}} > 0$, there exists $(t, y)$ such that $c - te - A^* y \in K$ with $t > 0$. By item (iv) of Lemma 3, we have $c - A^* y \in \text{ri} K$.

(ii) Suppose that (Conic-D) is in weak status. If (Conic-D) is weakly feasible, then there is $y$ such that $(0, y)$ is feasible for (D-Feas). Therefore, $\theta_{D-\text{Feas}} \geq 0$. By item (i), we must have $\theta_{D-\text{Feas}} = 0$.

Next, we suppose that (Conic-D) is weakly infeasible. By Lemma 21, we have

$$\text{dist} \left( (c + \text{range} A^*) \cap \text{span} K, K \right) = 0.$$ 

Therefore, there exists $y$ such that $c - A^* y \in \text{span} K$. And, because $e \in \text{ri} K$, there exists $t < 0$ such that $c - te - A^* y \in \text{ri} K$. This, together with item (i), shows that (D-Feas) must be strongly feasible and that $\theta_{D-\text{Feas}} \leq 0$. Therefore, (P-Feas) must be attained and have finite optimal value equal to $\theta_{D-\text{Feas}} \leq 0$. That is, there exists $x \in K$ such that

$$\langle c, x \rangle = \theta_{D-\text{Feas}}, \quad \langle e, x \rangle = 1, \quad Ax = 0.$$ 

If $\theta_{D-\text{Feas}} < 0$, we would have that (Conic-D) is strongly infeasible, which contradicts our assumptions. Therefore, we must have $\theta_{D-\text{Feas}} = 0$.

This concludes the first half of item (ii). Now suppose that $\theta_{D-\text{Feas}} = 0$. Then, there are sequences $\{t_\epsilon\}, \{y_\epsilon\}$ such that $t_\epsilon \to 0$ and

$$c - t_\epsilon e - A^* y_\epsilon \in K, \quad \forall \epsilon.$$ 

We have

$$\text{dist} \left( (c + \text{range} A^*) \cap K, K \right) \leq \| (c - A^* y_\epsilon) - (c - t_\epsilon e - A^* y_\epsilon) \| \leq \| t_\epsilon e \|, \quad \forall \epsilon.$$ 

Since $t_\epsilon \to 0$, this shows that $\text{dist} (c + \text{range} A^*, K) = 0$ and, therefore, (Conic-D) is in weak status.

(iii) From item (ii), we know that (Conic-D) is in weak status if and only if $\theta_{D-\text{Feas}} = 0$. In particular, if (Conic-D) is weakly infeasible then the optimal value of (D-Feas) cannot be attained for, otherwise, we would obtain a feasible solution to (Conic-D). Conversely, if $\theta_{D-\text{Feas}} = 0$ and is not attained, then (Conic-D) is in weak status and cannot be weakly feasible, so it must be weakly infeasible.

(iv) Follows by elimination.

Remark. Item (iv) of Proposition 22 includes the possibility that $\theta_{D-\text{Feas}} = -\infty$, i.e., (D-Feas) might be infeasible. This happens, for example, when $K$ is a subspace and $c + \text{range} A^*$ does not intersect $K$. However, under the hypothesis that $K$ is full-dimensional (i.e., $\text{span} K = E$), (D-Feas) must always be feasible because $e \in \text{int} K$, see Proposition 20.
From Proposition 22 we see that if \((\text{Conic-D})\) is weakly infeasible, we can obtain almost feasible solutions to \((\text{Conic-D})\) by constructing almost optimal solution solutions to \((\text{D-Feas})\), which can be done through the discussion in Section 4.1 and Algorithm 2. For future reference, we register this fact as a proposition.

**Proposition 23** (From almost optimality to almost feasibility). Suppose that \(\epsilon > 0\) and that \((t_\epsilon, y_\epsilon)\) is a feasible solution to \((\text{D-Feas})\) satisfying \(0 \geq t_\epsilon \geq -\epsilon\). Then,

\[
\text{dist} (c - A^* y_\epsilon, K) \leq \epsilon \|e\|.
\]

**Proof.** Since \((t_\epsilon, y_\epsilon)\) is feasible for \((\text{D-Feas})\) we have \(c - t_\epsilon e - A^* y_\epsilon \in K\). Therefore,

\[
\text{dist} (c - A^* y_\epsilon, K) \leq \|c - A^* y_\epsilon - (c - t_\epsilon e - A^* y_\epsilon)\| \leq \epsilon \|e\|.
\]

To conclude this section, we present an algorithm for handling infeasibility, Algorithm 3. The algorithm is able to distinguish between weak and strong infeasibility and, for weakly infeasible problems, it returns almost feasible solutions. During the algorithm’s run, we will need to apply facial reduction to \((\text{P-Feas})\). For convenience, denote by \(\mathcal{F}_{\text{P-Feas}}\) the minimal face of \(K^*\) that contains the feasible region of \((\text{P-Feas})\). Applying facial reduction to \((\text{P-Feas})\) leads to the following pair of problems:

\[
\begin{align*}
\inf_x \langle c, x \rangle & \quad \text{(P-Feas)} \\
\text{subject to} \quad A x &= 0 \\
\quad \langle e, x \rangle &= 1 \\
\quad x &\in \mathcal{F}_{\text{P-Feas}}
\end{align*}
\]

\[
\begin{align*}
\sup_{t,y} t & \quad \text{(D-Feas)} \\
\text{subject to} \quad c - t e - A^* y &\in (\mathcal{F}_{\text{min}})\text{*.}
\end{align*}
\]

The pair \((\text{P-Feas})\) and \((\text{D-Feas})\) satisfy the following property.

**Proposition 24.** Suppose \((\text{D-Feas})\) is feasible and \(\mathcal{F}_{\text{min}}\) \neq \emptyset, then the pair \((\text{P-Feas})\), \((\text{D-Feas})\) are both strongly feasible and their common optimal value is equal to \(\theta_{\text{D-Feas}}\). Moreover, if \(\mathcal{F}_{\text{min}}\) \(= \emptyset\) then \((\text{Conic-D})\) is strongly feasible.

**Proof.** Under the assumption that \((\text{D-Feas})\) is feasible, \((\text{D-Feas})\) must be, in fact, strongly feasible, by Proposition 20. Therefore, the minimal face of \(K\) that contains the feasible region of \((\text{D-Feas})\) is \(K\) itself. That is,

\[
\mathcal{F}_{\text{D-Feas}} = K.
\]

Applying Theorem 16 to \((\text{D-Feas})\) we conclude that \(\theta_{\text{D-Feas}}\) is finite if and only if \(\mathcal{F}_{\text{min}}\) \(\neq \emptyset\). We also obtain from item (i) of Theorem 16 that, if indeed \(\mathcal{F}_{\text{min}}\) \(\neq \emptyset\) holds, then \((\text{P-Feas})\), \((\text{D-Feas})\) are both strongly feasible and their common optimal value must coincide with \(\theta_{\text{D-Feas}}\). Alternatively, from item (ii) of Theorem 16, we conclude that \(\mathcal{F}_{\text{min}}\) \(= \emptyset\) if and only if \(\theta_{\text{D-Feas}} = +\infty\), in which case \((\text{Conic-D})\) must be strongly feasible by Proposition 22. \(\square\)
We can now state Algorithm 3.

**Algorithm 3:** Handling infeasibility in (Conic-D)

**Input:** $K, A, b, c, \epsilon$ ((Conic-D) is assumed to be infeasible)

**Output:** Weakly Infeasible or Strongly Infeasible. If Weakly Infeasible then $y_\epsilon$ such that $\text{dist}(c - A^*y_\epsilon, K) \leq \epsilon$ is also returned.

1. If $(c + \text{range} A^*) \cap \text{span} K = \emptyset$ then
   2. return Strongly Infeasible

3. end

4. Apply Algorithm 1 to (P-Feas) and let $(f_1, y_1), \ldots, (f_\ell, y_\ell)$ be obtained corresponding reducing directions.

5. if Algorithm 1 returned Infeasible (i.e., $F_{\text{P-Feas}} - \text{Feas} \cap \emptyset$) then
   6. return Strongly Feasible

7. else
   8. Solve the pair $(\hat{P} \text{-Feas})$ and $(\hat{D} \text{-Feas})$ and denote the optimal value by $\theta$, and the optimal solution of $(\hat{D} \text{-Feas})$ by $(t^*, y^*)$.

   9. if $\theta < 0$ then
      10. return Strongly Infeasible

11. else if $\theta = 0$ then
      12. Let $\hat{t}, \hat{y}$ be such that $c - \hat{t}e - A^*\hat{y} \in \text{ri} K$.

      13. Apply Algorithm 2 to (D-Feas) using as input $(f_1, y_1), \ldots, (f_\ell, y_\ell), (\hat{t}, \hat{y}), (t^*, y^*), \epsilon/\|e\|$.

      14. let $y_\epsilon$ be the output of Algorithm 2.

      15. return $y_\epsilon$ and Weakly Infeasible

16. else if $\theta > 0$ then
      17. return Strongly Feasible

18. end

*By Proposition 20, $(\hat{t}, \hat{y})$ must exist because if we have reached this line, (Conic-D) is not strongly infeasible.

**Proposition 25** (Algorithm 3 is correct). The following hold.

(i) Assuming that (Conic-D) is infeasible, Algorithm 3 correctly identifies whether (Conic-D) is strongly or weakly infeasible.

(ii) When (Conic-D) is weakly infeasible, the output of Algorithm 3 is indeed an $\epsilon$-feasible solution.

(iii) When $K = S_n^+$, Algorithm 3 is implementable with $O(n)$ calls to $O_{\text{int}}$.

**Proof.** The correctness of Algorithm 3 follows from Proposition 24 and the correctness of Algorithm 2. We will now explain some details.

By Propositions 22 and 24, to distinguish between weak and strong infeasibility it is enough to check the following three items: whether (D-Feas) is feasible or not; whether $F_{P \text{-Feas}} - \text{Feas}$ is empty or not; whether the optimal value of the pair $(\hat{P} \text{-Feas})$ and $(\hat{D} \text{-Feas})$ is negative or zero. These three items are checked at Lines 1, 5, 9 and 11 of Algorithm 3.

At Line 1, if $(c + \text{range} A^*) \cap \text{span} K = \emptyset$, then the optimal value of (D-Feas) is $-\infty$ and (Conic-D) is strongly infeasible by Proposition 22.

However, if we progress until the check of Line 5, (D-Feas) must be strongly feasible, by Proposition 20. By this point, facial reduction is applied to (P-Feas) and if $F_{\text{P-Feas}} \cap \emptyset = \emptyset$, then Proposition 24 tells us that (Conic-D) is strongly feasible. As we are assuming that (Conic-D) is infeasible, this should not happen.

If the algorithm reaches Line 9 then (D-Feas) is feasible, $F_{\text{min}} \neq \emptyset$ and, therefore, Proposition 24 applies. Therefore, if $\theta < 0$ it must be, indeed the case that (Conic-D) is strongly infeasible. If $\theta = 0$ then (Conic-D) is weakly infeasible and the correctness of Algorithm 2 shows that $(t_\epsilon, y_\epsilon)$ is indeed an $\epsilon/\|e\|$.  

27
optimal solution to (D-Feas) which, by Proposition 23 implies that \( y \) must be an \( \epsilon \)-feasible solution to (Conic-D).

Finally, suppose that \( \mathcal{K} = S^n_+ \). The only lines where SDPs need to be solved are when Algorithm 1 is invoked and at Line 8. Algorithm 1 and Algorithm 2 require at most \( n + 1 \) calls to \( O_{\text{int}} \) each so, we only need to check that we can indeed solve the SDP at Line 8 with \( O_{\text{int}} \). We note that if we reach Line 8, then (D-Feas) is feasible and \( \mathcal{F}^{\text{P-Feas}}_{\text{min}} \neq \emptyset \) which, by Proposition 24 implies that the pair (P-Feas) and (D-Feas) are both strongly feasible and can indeed be solved by \( O_{\text{int}} \). Therefore, Algorithm 3 can be implemented with \( O(n) \) calls to \( O_{\text{int}} \).

To conclude, we note that, when \( \mathcal{K} = S^n_+ \), the problem at Line 12 can be solved by invoking \( O_{\text{int}} \) (which would not affect the \( O(n) \) complexity of Algorithm 3), but that is not necessary. Let \( \lambda_{\text{min}}(\cdot) \) denote the minimum eigenvalue function and recall \( \lambda_{\text{min}}(U + V) \geq \lambda_{\text{min}}(U) + \lambda_{\text{min}}(V) \) always holds for \( U, V \in S^n \). With that, if \( \alpha \) is positive then
\[
\alpha > \frac{\lambda_{\text{min}}(c - A^* \hat{y})}{\lambda_{\text{min}}(e)} \Rightarrow c - A^* \hat{y} + \alpha e \in S^n_+.
\]
So, with two minimum eigenvalue computations, we can solve the problem in Line 12 of Algorithm 3. As in Section 4.1.1, a strategy of starting with some negative \( t \) and doubling it at each step would also work.

5 Completely solving (Conic-D)

Using the techniques described in Sections 3 and 4, we can now present a general algorithm for completely solving (Conic-D), in the sense of Definition 1. In particular, when \( \mathcal{K} = S^n_+ \), we can completely solve (SDP-D) through polynomially many calls to \( O_{\text{int}} \). For ease of reference, we write down below again some of the auxiliary problems that are referenced in Algorithm 4 below.

\[
\inf_x \langle c, x \rangle \quad \text{(P)} \quad \sup_y \langle b, y \rangle \quad \text{(D)}
\]
\[
\text{subject to } A x = b \quad \text{subject to } c - A^* y \in \mathcal{F}^\text{D}_{\text{min}}.
\]
\[
\inf_x \langle c, x \rangle \quad \text{(P*)} \quad \sup_y \langle b, y \rangle \quad \text{(D*)}
\]
\[
\text{subject to } A x = b \quad \text{subject to } c - A^* y \in (\mathcal{F}^\text{D}_{\text{min}})^*.
\]

Here, we recall that \( \mathcal{F}^\text{D}_{\text{min}} \) is the minimal face of \( \mathcal{K} \) that contains \( (c + \text{range } A^*) \cap \mathcal{K} \) and \( \mathcal{F}^\text{D}_{\text{min}} \) is the minimal face of \( (\mathcal{F}^\text{D}_{\text{min}})^* \) that contains the feasible region of \( \text{(P)} \). We also recall that, by Theorem 16, \( \theta_D \) is finite if and only if (D-Feas) is feasible and \( \mathcal{F}^\text{D}_{\text{min}} \neq \emptyset \). In this case, (P*) and (D*) are both strongly feasible and, when \( \mathcal{K} = S^n_+ \), they can be solved by invoking \( O_{\text{int}} \). By doing so, we are able to obtain the dual optimal value \( \theta_D \). Checking whether \( \theta_D \) is attained can be done by solving the following feasibility problem.

\[
\text{find } y \quad \text{(D-OPT)}
\]
\[
\text{subject to } c - A^* y \in \mathcal{F}^\text{D}_{\text{min}}, \quad \langle b, y \rangle = \theta_D.
\]

Let \( L : \mathbb{R}^m \to \mathbb{R}^m \) be an affine map (that is, \( L - u \) is linear for some \( u \in \mathbb{R}^m \)) such that
\[
\text{range } L = \{ y | \langle b, y \rangle = \theta_D \}
\]
In particular, \( \langle b, L(\hat{y}) \rangle = \theta_D \) holds for every \( \hat{y} \). With that we can put (D-OPT) in “dual standard format” as follows

\[
\text{find } \hat{y} \quad \text{(D-OPT-STD)}
\]
\[
\text{subject to } c - A^*(L(\hat{y})) \in \mathcal{F}^\text{D}_{\text{min}}.
\]
We observe that \((D\text{-OPT})\) is feasible if and only if \((D\text{-OPT-STD})\) is feasible\(^3\). Once \((D\text{-OPT-STD})\) is solved (for example, with Algorithm 1) and a solution \(\hat{y}^*\) is obtained, a solution to \((D\text{-OPT})\) is obtained by letting \(y^* = L\hat{y}^*\). Of course, it might be the case \((D\text{-OPT})\) is not feasible in the first place. Nevertheless, we now have all pieces in place to state Algorithm 4.

**Algorithm 4: Completely Solving \((\text{Conic-D})\)**

\begin{verbatim}
Input: \(K, A, b, c, \epsilon\)
1 Apply Algorithm 1 to \((\text{Conic-D})\) and let \(d_1, \ldots, d_{\ell_1}\) be the corresponding reducing directions.
2 if Algorithm 1 returned Infeasible then
3     Invoke Algorithm 3, return its outputs.
4 else
5     Let \(\hat{s}\) be such that \(\hat{s} \in \text{ri} F\text{\_min}^D\) (see output 2. of Algorithm 1) and \(\hat{y}\) such that \(c - A^*\hat{y} = \hat{s}\).
6     Apply Algorithm 1 to \((\hat{P})\), obtain reducing directions \((f_1, y_1), \ldots, (f_{\ell_2}, y_{\ell_2})\).
7     if \(F\text{\_min}^{\hat{P}} = \emptyset\) then
8         return \(\hat{s}\) and Feasible Unbounded.
9     else
10    Solve \((P^*)\) and \((D^*)\) and obtain \(\theta_D\) and optimal solutions \(y^*, x^*\) to \((D^*)\) and \((P^*)\), respectively.
11       Apply Algorithm 1 to \((D\text{-OPT-STD})\).
12       if Algorithm 1 returned Infeasible then
13           Use Algorithm 2 with \(f_1, \ldots, f_{\ell_1}, \hat{y}, y^*, \epsilon\) as inputs and return \(y_\epsilon\) (the output of Algorithm 2), \(\theta_D\) and Feasible Unattained.
14           else
15             Let \((y, s)\) be the feasible solution returned by Algorithm 1.
16             return \(y, \theta_D\) and Feasible Attained.
17       end
18    end
19 end
\end{verbatim}

**Theorem 26** (Algorithm 4 is correct). Algorithm 4 completely solves \((\text{Conic-D})\). That is, it correctly determines whether \((\text{Conic-D})\) is feasible or not. If \((\text{Conic-D})\) is infeasible, Algorithm 4 distinguishes between weak and strong infeasibility and, in case of weak infeasibility, an \(\epsilon\)-feasible solution is returned. If \((\text{Conic-D})\) is feasible, Algorithm 4 computes the optimal value of \((\text{Conic-D})\). If the optimal value is finite and attained, an optimal solution is returned. If the optimal value is finite but not attained, an \(\epsilon\)-optimal solution is returned. If the optimal value is \(+\infty\), a feasible solution is returned.

In addition, if \(K = S^+_n\), then \((\text{Conic-D})\) can be implemented with \(O(n)\) calls to the oracle \(O_{\text{int}}\).

**Proof.** To prove the result we gather everything we have done so far. We consider the following cases.

1. \((\text{Conic-D})\) is infeasible. The correctness of Facial Reduction and Algorithm 1 implies that if \((\text{Conic-D})\) is infeasible, then this will be correctly detected after Line 1. Furthermore, the correctness of Algorithm 3 (Proposition 25) ensures that weak infeasibility and strong infeasibility will be correctly detected. And, in case of weak infeasibility, an \(\epsilon\)-feasible solution will be returned.

2. \((\text{Conic-D})\) is feasible but unbounded. If the algorithm advances until Line 7, it is because \((\text{Conic-D})\) is feasible and, in particular, \(F\text{\_min}^D\) is not empty. In this case, we are under the hypothesis of Theorem 16. By item \((\text{ii})\) of Theorem 16, we have \(\theta_D = +\infty\) if and only if \(F\text{\_min}^{\hat{P}} = \emptyset\).

3. \((\text{Conic-D})\) is feasible, \(\theta_D\) is finite but not attained. In that case, when Algorithm 1 is invoked at Line 11, it is correctly detected that \((D\text{-OPT-STD})\) is infeasible and Algorithm 2 correctly constructs an \(\epsilon\)-optimal solution.

\(^3\)If \(y\) is feasible for \((D\text{-OPT})\), then \(y \in \text{range} L\), so there exists \(\hat{y}\) such that \(L(\hat{y}) = y\). Reciprocally, if \(\hat{y}\) is feasible for \((D\text{-OPT-STD})\) then \(L\hat{y}\) is feasible for \((D\text{-OPT})\).
4. \((\text{Conic-D})\) is feasible, \(\theta_D\) is finite and attained. In that case, when Algorithm 1 is invoked at Line 11, a feasible solution to \((\text{D-OPT-STD})\) will be obtained, which corresponds to an optimal solution to \((\text{Conic-D})\).

For the last part of the proof, suppose that \(\mathcal{K} = S^n_+.\) Algorithm 4 directly invokes facial reduction (Algorithm 1) at most 3 times (Lines 1, 6 and 11). It also directly invokes Algorithm 2 and Algorithm 3 at most one time, each. The only other time where SDPs need to be solved is at Line 10 where we need to solve the SDPs \((P^*)\) and \((P^*)\), which are both strongly feasible (item (i) of Theorem 16) and therefore can be solved by a single call to \(\mathcal{O}_{\text{int}}.\) By Propositions 12, 19, 25, we have that Algorithms 1, 2 and 3 can also be implemented with \(O(n)\) calls to \(\mathcal{O}_{\text{int}}\), so the same must be true of Algorithm 4.

6 Comparison with other approaches

As far as we know, the discussion on using reducing directions to concretely construct almost optimal solutions and almost feasible solutions for general conic linear programs are new, although some of those ideas were present in our previous works on semidefinite programming [18] and second order cone programming [20]. Furthermore, technical results on facial reduction and double facial reduction such as Proposition 14, Theorems 15 and 16 seem to be novel as well. Nevertheless, the idea of completely solving a problem in some sense is definitely not new and, in this section, we compare our approach with two other proposals in the literature that had similar goals.

In section 5.10 of [7], de Klerk, Terlaky and Roos have described a possible sequence of steps to solve \((\text{SDP-D})\). Their tool of choice is a self-dual embedding strategy of the original pair \((\text{SDP-P})\) and \((\text{SDP-D})\).

As we mentioned before, in the absence of both primal and dual strong feasibility, the embedded problem might fail to reveal the optimal value of the original problem or detect infeasibility/nonattainment. To account for that, they go for a second step, where they consider an embedded problem using Ramana’s dual. The Ramana’s dual \((P_R)\) is a substitute for \((\text{SDP-P})\) and they consider the pair formed by \((P_R)\) and its dual \((D_{\text{cor}})\), which is a “corrected” version of \((\text{SDP-D})\). The pair \((P_R, D_{\text{cor}})\) can then be solved by their embedding strategy to find \(\theta_D.\) As the embedded problem is both primal and dual strongly feasible, it is possible to invoke \(\mathcal{O}_{\text{int}}\) to solve it. However, if the solution given by \(\mathcal{O}_{\text{int}}\) is not of maximum rank at both steps, their strategy might not work. We should mention that they do show in detail how to build an interior point method suitable for their approach. Our analysis, on the other hand, is completely agnostic to the inner workings of the interior point oracle and no assumption is made on the optimal solutions returned by \(\mathcal{O}_{\text{int}}\).

As our approach does not rely on Ramana’s dual, our analysis is easily generalizable to other classes of cones. Indeed, Algorithm 4 is valid and correct for any closed convex cone \(\mathcal{K}.\) We remark that although there is a strong connection between Ramana’s dual and facial reduction [37, 28], no similar construction is known for any other class of cones. For example, following Pataki’s approach in [28], one could formulate an alternative dual system for a second order cone programming problem. Such a system would have many of the properties that Ramana’s dual has, but it is not clear whether that system can be expressed via second order cone constraints.

Permenter, Friberg and Andersen present in [32] a very elegant approach for general conic linear programming based on self-dual embeddings and they are able to achieve most of the goals included in Definition 1. They showed that the relative interior of the set of solutions to a certain self-dual embedding of the pair \((\text{Conic-P})\) and \((\text{Conic-D})\) will reveal reducing directions for \((\text{Conic-P})\) and \((\text{Conic-D})\) under certain circumstances, see Corollary 3.3 in [32]. They used this property to present an algorithm for solving \((\text{Conic-P})\) and \((\text{Conic-D})\) while identifying several pathologies, see Algorithms 1 and 2 in [32].

We remark that even if \((\text{Conic-P})\) and/or \((\text{Conic-D})\) are not strongly feasible, it could still be the case that the duality gap is zero and both problems are attained. In this case, certain self-dual embeddings might recover optimal solutions to the pair \((\text{Conic-P}), (\text{Conic-D})\) even in the absence of strong feasibility. Indeed, a crucial advantage of the approach in [32] is that a facial reducing step is performed only if it strictly necessary in order to recover zero duality gap and attainment, see item 1. of Theorem 4.1. As such, the approach in [32] regularizes a problem only if needed.
However, the main drawback in [32] seems to be the fact that it requires a relative interior solution to their self-dual embedding, which is a stronger requirement than our assumption of having access to $\mathcal{O}_{\text{int}}$, since $\mathcal{O}_{\text{int}}$ is allowed to return any optimal solution. While relative interior optimal solutions might be obtainable via interior point algorithms, this is not necessarily true for other methods. Nevertheless, although our algorithm is more general, the approach in [32] seems to be more likely to lead to a practical implementation than ours, especially in conjunction with interior point methods. Indeed, the numerical experiments in Section 5 of [32] suggest that even when reducing directions are computed inexactly there are cases where they are still useful for analyzing the problem, although sometimes these approximate directions can also lead to incorrect conclusions.

7 Concluding remarks

In this paper, we have discussed how to use facial reduction and double facial reduction to completely solving (Definition 1) a general conic linear program, under the assumption that certain auxiliary problems can be solved, see Algorithm 4 and Theorem 26. When specialized to the particular case of semidefinite programming, these results imply that an arbitrary semidefinite program over $n \times n$ matrices can be completely solved by invoking at most $O(n)$ times an oracle that only return solutions to primal and dual strongly feasible SDPs. We also provided technical results on facial reduction and double facial reduction that might be of independent interest, see Sections 3 and 4.

For limitations, drawbacks and comparison to other approaches, see Sections 1.3 and 6. In particular, as discussed in Section 1.3, in our analysis we assumed that the oracle $\mathcal{O}_{\text{int}}$ returns an exact solution. An interesting topic of future research would be to consider the effects of impreciseness in the solutions returned by $\mathcal{O}_{\text{int}}$.

Acknowledgements

We thank Prof. Henry Wolkowicz for bringing to our attention the paper by Abrams [1]. The first author would also like to acknowledge the support and the interesting discussions with Prof. Mituhiro Fukuda at the Tokyo Institute of Technology.

References


A Reformulation of SDPs

In this section we discuss and prove Proposition 6, which we repeat below for convenience.

**Proposition.** Suppose that $K$ is either a non-empty proper face of $S^n_+$ or the dual of a non-empty proper face of $S^n_+$. Suppose that (Conic-P) and (Conic-D) are strongly feasible. Then, (Conic-P) and (Conic-D) are solvable with a single call to $O_{int}$.

In what follows, we suppose that $E = S^n$ and the inner product $\langle \cdot, \cdot \rangle$ is such that $\langle x, y \rangle$ is equal to the trace of $xy$, for all $x, y \in S^n$. In particular, $\langle x, y \rangle = \sum_{i,j} x_{ij} y_{ij}$, where $x_{ij}$ denotes the $(i,j)$ component of $x$.

Initially, suppose that $K$ is the dual of a non-empty face of $S^n_+$. From Proposition 5 there is an $n \times n$ orthogonal matrix $Q$ and $r \leq n$ such that

$$Q^T K Q = \left\{ \begin{pmatrix} U & V \\ V^T & W \end{pmatrix} \in S^n \mid U \in S^r_+ \right\}. \quad (22)$$

Using the fact that $Q$ is orthogonal, we also have

$$Q^T K^* Q = \left\{ \begin{pmatrix} U & 0 \\ 0 & 0 \end{pmatrix} \in S^n \mid U \in S^r_+ \right\}. \quad (23)$$

We now have to set up several auxiliary maps. Let $\pi_r$ denote the map that takes an $n \times n$ matrix and maps it to its upper left $r \times r$ block. We have

$$\pi_r \begin{pmatrix} U \\ V^T \\ W \end{pmatrix} = U.$$

Let $\psi : S^n \to S^r$ be the linear map that maps $x \in S^n$ to $\pi_r(Q^T x Q)$. Since $A : S^n \to \mathbb{R}^m$ is a linear map, there are $A_1, \ldots, A_m$ such that

$$c - A^* y = c - \sum_{i=1}^m A_i y_i, \quad \text{and} \quad Ax = b \Leftrightarrow \langle A_i, x \rangle = b_i, i = 1, \ldots, m.$$

We reformulate (Conic-P) and (Conic-D) as follows.

$$\inf_{\tilde{x}} \langle \psi(c), \tilde{x} \rangle \quad \text{(P)} \quad \sup_{\tilde{y}} \langle b, \tilde{y} \rangle \quad \text{(D)}$$

subject to $\langle \psi(A_i), \tilde{x} \rangle = b_i, \quad i = 1, \ldots, m$

subject to $\psi(c) - \psi(A^* \tilde{y}) \in S^r_+$.

**Proposition 27.** The pair (P) and (D) is strongly feasible if and only if (Conic-P) and (Conic-D) are strongly feasible. Furthermore,
(i) \( \hat{y}^* \) is an optimal solution to (D), if and only if \( \hat{y}^* \) is an optimal solution to (Conic-D).

(ii) \( \hat{x}^* \) is an optimal solution to (P) if and only if \( \psi^* \hat{x}^* \) is an optimal solution to (Conic-P).

Proof. Let \( S^n_++ \) denote the \( n \times n \) positive definite real symmetric matrices. Since (22) and (23) holds, we have

\[
Q^T (\text{ri } K) Q = \left\{ \begin{pmatrix} U & V \\ V & W \end{pmatrix} \in S^n \mid U \in S^n_+ \right\}, \quad Q^T \text{ri } (K^*) Q = \left\{ \begin{pmatrix} U & 0 \\ 0 & 0 \end{pmatrix} \in S^n \mid U \in S^n_+ \right\}. \tag{24}
\]

Therefore,

\[
c - \mathcal{A}^* \hat{y} \in \text{ri } K \Leftrightarrow \psi(c) - \psi(\mathcal{A}^* \hat{y}) \in S^n_+,
\]

which shows that (Conic-D) is strongly feasible if and only if (D) is strongly feasible.

Next, we first observe that the adjoint of \( \pi_r \) is the map that takes \( \hat{x} \in S^r \) and maps it to the matrix in \( S^n \) that has \( \hat{x} \) in its upper left \( r \times r \) block and zero elsewhere. With that in mind, the adjoint of \( \psi \) is such that

\[
\psi^* \hat{x} = Q \pi_r^* (\hat{x}) Q^T.
\]

This expression for \( \psi^* \), together with (24) and the fact that \( \langle \psi(\mathcal{A}_i), \hat{x} \rangle = \langle \mathcal{A}_i, \psi^* \hat{x} \rangle \), allow us to conclude that if \( \hat{x} \) is feasible for (P) and \( \hat{x} \in S^n_+ \), then \( \psi^* \hat{x} \) is feasible for (Conic-P) and belongs to ri \( K^* \). Conversely, suppose that \( x \) is feasible for (Conic-P) and belongs to \( \text{ri } K^* \). By (24), \( \psi(x) \in S^n_+ \). In addition, for every \( i = 1, \ldots, m \) we have

\[
b_i = \langle \mathcal{A}_i, x \rangle = \langle Q^T \mathcal{A}_i Q, Q^T x Q \rangle = \langle \pi_r (Q^T \mathcal{A}_i Q), \pi_r (Q^T x Q) \rangle = \langle \psi(\mathcal{A}_i), \psi(x) \rangle, \tag{25}
\]

where the second equality follows from (24) and the fact that \( \langle \cdot, \cdot \rangle \) is the trace inner product. The third inequality follows from the fact that the nonzero elements of \( Q^T x Q \) are all in the upper left \( r \times r \) block, by (23). The overall conclusion is that (P) is strongly feasible if and only if (Conic-P) is strongly feasible.

We now move on the proofs of items (i) and (ii). First, since

\[
c - \mathcal{A}^* \hat{y} \in K \Leftrightarrow \psi(c) - \psi(\mathcal{A}^* \hat{y}) \in S^n_+,
\]

we conclude that item (i) indeed holds.

For the proof of item (ii), we check two facts. First, if \( x \) is feasible for (Conic-P), then (23) and (25) imply that \( \psi(x) \) is also feasible for (P). In addition, analogous to (25), we have \( \langle c, x \rangle = \langle \psi(c), \psi(x) \rangle \). Conversely, if \( \hat{x} \) is feasible for (P), then (23) implies that \( \psi^* \hat{x} \) is feasible for (Conic-D). Furthermore, we have \( \langle \psi(c), \hat{x} \rangle = \langle c, \psi^* x \rangle \), so \( \hat{x} \) and and and \( \psi^* x \) have the same value.

From Proposition 27, we conclude that (Conic-D) and (Conic-P) can be solved via a single call to \( \mathcal{O}_{\text{int}} \), when \( K \) is the dual of a face of \( S^n_+ \).

If \( K \) is a face of \( S^n_+ \), reformulating (Conic-D) as an SDP in dual standard format gets more cumbersome because we need to include the (linear) constraints that all entries of \( Q^T (c - \mathcal{A}^* y) Q \) not belonging to the upper left \( r \times r \) block must be zero.

Alternatively, we can reformulate (Conic-D) as a problem in primal standard format as follows. Let \( \hat{\mathcal{A}} \) and \( \hat{b} \) be such that

\[
c + \text{range} \mathcal{A}^* = \{ s \in \mathcal{E} \mid \hat{\mathcal{A}} s = \hat{b} \}
\]

Also, let \( x_0 \in \mathcal{E} \) be such that \( A x_0 = b \). Here, it is not required that \( x_0 \) belongs to \( K^* \). With that, if \( s = c - \mathcal{A}^* y \), we have

\[
-\langle b, y \rangle = \langle x_0, -\mathcal{A}^* y \rangle = \langle x_0, s \rangle - \langle x_0, c \rangle.
\]

Therefore,

\[
\sup_{c - \mathcal{A}^* y \in K} \langle b, y \rangle = \langle x_0, c \rangle - \inf_{\hat{\mathcal{A}} s = \hat{b}, s \in K} \langle x_0, s \rangle
\]

In this way, we can reformulate (Conic-D) as a problem in primal standard format over a face \( K \). For this problem, the discussion leading to Proposition 27 is applicable. This shows that Proposition 6 is indeed true.