DTU Library

Solving conic optimization problems via self-dual embedding and facial reduction: A unified approach

Permenter, Frank; Friberg, Henrik Alsing; Andersen, Erling D.

Link back to DTU Orbit

Solving conic optimization problems via self-dual embedding and facial reduction: a unified approach

Frank Permenter [*]   Henrik A. Friberg [§]   Erling D. Andersen [¶]

August 18, 2016

Abstract

We establish connections between the facial reduction algorithm of Borwein and Wolkowicz and the self-dual homogeneous model of Goldman and Tucker when applied to conic optimization problems. Specifically, we show the self-dual homogeneous model returns facial reduction certificates when it fails to return a primal-dual optimal solution or a certificate of infeasibility. Using this observation, we give an algorithm, based on facial reduction, for solving the primal problem that, in principle, always succeeds. (An analogous algorithm is easily stated for the dual problem.) This algorithm has the appealing property that it only performs facial reduction when it is required, not when it is possible; e.g., if a primal-dual optimal solution exists, it will be found in lieu of a facial reduction certificate even if Slater's condition fails. For the case of linear, second-order and semidefinite optimization, we show the algorithm can be implemented assuming oracle access to the central-path limit-point of an extended-embedding, a strictly feasible conic problem with strictly feasible dual. We then give numerical experiments illustrating barriers to practical implementation.

1 Introduction

Let $\mathcal{V}$ be a finite-dimensional inner-product space and $A : \mathcal{V} \to \mathbb{R}^m$ a linear map with adjoint $A^* : \mathbb{R}^m \to \mathcal{V}$. For $b \in \mathbb{R}^m$ and $c \in \mathcal{V}$, consider the following primal-dual pair of conic optimization problems over a closed, convex cone $\mathcal{K} \subseteq \mathcal{V}$ and its dual cone $\mathcal{K}^* := \{ s \in \mathcal{V} : \langle s, x \rangle \geq 0 \;\; \forall x \in \mathcal{K} \}$:

$$
\begin{array}{llll}
\text{minimize} & \langle c, x \rangle & \text{maximize} & \langle b, y \rangle \\
\text{subject to} & Ax = b & \text{subject to} & c - A^*y = s \\
& x \in \mathcal{K} & & s \in \mathcal{K}^*, \; y \in \mathbb{R}^m.
\end{array}
\tag{1}
$$

Here, $x \in \mathcal{V}$ is the decision variable of the primal problem and $(s, y) \in \mathcal{V} \times \mathbb{R}^m$ is the decision variable of the dual. The *self-dual embedding* technique [16, 7, 20, 26], originally due to Goldman and Tucker [12], solves (1) by finding solutions to the following *self-dual homogeneous model*:

$$
\begin{array}{rcl}
Ax - b\tau & = & 0, \\
-A^*y - s + c\tau & = & 0, \\
\langle b, y \rangle - \langle c, x \rangle - \kappa & = & 0, \\
(x, s, y, \tau, \kappa) \in \mathcal{K} \times \mathcal{K}^* \times \mathbb{R}^m \times \mathbb{R}_+ \times \mathbb{R}_+.
\end{array}
\tag{2}
$$

---

[*]F. Permenter is with the Laboratory for Information and Decision Systems, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge MA 02139, USA. Email: fperment@mit.edu.

[§]H. A. Friberg is with the Department of Wind Energy, Technical University of Denmark, DK-4000 Roskilde, Denmark, and employed at MOSEK ApS. E-mail: metronware@gmail.com.

[¶]E. D. Andersen is the founder and CEO of MOSEK ApS, DK-2100 Copenhagen O, Denmark. E-mail: e.d.andersen@mosek.com.

Any solution (x, s, y, τ, κ) of (2) satisfies the complementarity condition $\tau\kappa = 0$; hence, solutions fall into one of three categories. If $\tau > 0$, then $\frac{1}{\tau}(x, s, y)$ is a primal-dual optimal solution for (1) with duality gap $\frac{1}{\tau^2}\langle x, s\rangle$ equal to zero; in other words, $\frac{1}{\tau}(x, s, y)$ is a *complementary solution*. If $\kappa > 0$, then x and/or y are *improving rays* that certify dual and/or primal infeasibility. If, on the other hand, $\tau = \kappa = 0$, the solution (x, s, y, τ, κ) reveals nothing about the primal or dual optimal value—either could be finite or unbounded from above or below.[1] Further, if $\tau = \kappa = 0$ holds for all solutions, then complementary solutions and improving rays do not exist, and at least one problem—primal or dual—fails the Slater regularity condition [8].

In this paper, we reexamine this latter case. Specifically, we show relative interior solutions of (2) yield *facial reduction certificates* for (1) when $\tau = \kappa = 0$ holds for all solutions. (Restricting to the relative interior is inspired by the analysis of de Klerk et al. [7].) As we review, these certificates allow one to regularize the primal or dual in the manner proposed by Borwein and Wolkowicz [5] and resolve (2), repeating until a complementary solution or an improving ray is obtained. As we show, this idea leads to simple algorithms that solve *arbitrary* instances of the primal or dual, where solve means to find the optimal value and a solution if one exists. These algorithms, of course, rely on a subroutine that produces relative interior solutions to (2). As we show, such solutions are obtained from relative interior solutions to the *extended-embedding* of Ye et al. [35], a strictly feasible conic optimization problem with strictly feasible dual. Note if $\mathcal{K}$ is a symmetric cone, e.g., if (1) is a linear, second-order cone, or semidefinite program, the central path of the extended-embedding converges to the relative interior of its solution set [13, 27]. Hence, implementations of our algorithms are conceptually simple in this case, involving only basic linear algebra (for regularization) and repeated calls to an interior-point method (for tracking the central path to its limit point).

Though the aforementioned implementations are *conceptually* simple, they nevertheless face significant *practical* barriers. Numerical experiments will be given to illustrate the following three issues. First, interior-point methods only approximate the limit point of the central path; hence, in practice, only approximate facial reduction certificates are obtained. Second, regularization with approximate certificates can change the optimal value an arbitrary amount, e.g., by changing a feasible problem to an infeasible problem. Note this second issue is rooted in the fact at least one problem—primal or dual—is *ill-posed* in the sense of Renegar [28] when $\tau = \kappa = 0$ for all solutions;[2] that is, at least one problem has optimal value infinitely-sensitive to perturbations of A, b and c. Third, strict complementarity can fail badly, which can cause poor convergence of interior-point algorithms. Indeed, the extended-embedding never has a strictly complementary solution when $\tau = \kappa = 0$ holds for all solutions, since, by definition, strict complementarity requires $\tau\kappa = 0$ and $\tau + \kappa > 0$. We emphasize these issues currently preclude a practical implementation of the mentioned algorithms. Nevertheless, experiments indicate 'good' approximate facial reduction certificates can be provided in many cases, which may have practical value; indeed, in Section 5.4, we use approximate certificates to explain poor performance [18] of primal-dual solvers on a DIMACS library instance [23].

Existence of facial reduction certificates implies failure of Slater's condition, which, for feasible problems, is a generic property [9]. In addition, failure of Slater's condition is only a *necessary* condition for non-existence of complementary solutions and improving rays. Hence, one might assume instances (theoretically) solvable by our approach not (theoretically) solvable by a standard

---

[1]Though the solution reveals nothing, the asymptotic behavior of central-path-following techniques gives information in certain cases. See Luo et al. [16] and de Klerk et al. [8].

[2]In this case, no complementary solution or improving ray exists. At least one problem—primal or dual—is thus feasible and fails Slater's condition or infeasible and fails to have a dual improving ray. In other words, at least one problem is *weakly feasible* or *weakly infeasible* in the sense of [17] and hence ill-posed in the sense of Renegar [28].

Case	Interpretation
$\tau > 0, \kappa = 0$	Complementary solution
$\tau = 0, \kappa > 0$	Improving ray(s)
$\tau = 0, \kappa = 0$	Facial reduction certificate(s)

Table 1: Interpretation of a relative interior solution to the self-dual homogeneous model (2) in terms of the conic optimization problem (1). A main observation of this paper, given by Corollary 1, is summarized by the last row.

interior-point algorithm are never encountered unless they are artificially constructed. We offer two responses. First, the aforementioned instance from [23] suggests such instances *can* arise in applications (though they may be rare). Second, the homogeneous model (2) is the back-bone of widely-used solvers [30, 19, 21]; hence, careful examination of its solution set is warranted to (at the very least) clarify the behavior of solvers across the full spectrum of inputs.

Our paper also contributes to the theory of facial reduction in a few ways. To explain, we first recall the basic operation of facial reduction algorithms [5] [22] [33]: finding facial reduction certificates, which themselves are solutions to conic optimization problems (so-called auxiliary problems). The recent papers of Cheung et al. [6] and Lourenço et al. [15] propose methods for finding certificates, addressing issues of numerical robustness [6] and strict feasibility of auxiliary problems [6, 15], while other papers take a heuristic approach [24, 25, 11]. In this paper, we show by finding relative interior solutions to (2), certificates are found only when they are *needed* (complementary solutions and improving rays do not exist) as opposed to when they *exist* (e.g., Slater's condition fails for feasible problems). In contrast, the methods of [5, 22, 33, 6, 15] find a complete set of certificates for feasible problems and regularize until Slater's condition holds (which can be costly and unnecessary), and the heuristic methods of [24, 25, 11] may fail to find needed certificates. We also show solutions to (2) automatically identify which problem—primal or dual— needs regularization. In contrast, facial reduction procedures often only regularize the problem one is interested in solving. This is insufficient, for instance, to solve the primal problem if it has a finite unattained optimal value; in this case, dual regularization or regularization in the sense of Abrams [1] is required. As we will show, relative interior solutions to (2) always provide the necessary certificate for the required regularization. Indeed, the certificates provided by (2) allow one to handle all pathologies (duality gaps, unattainment, etc.) in a unified facial-reduction-based framework, where, in contrast, the method of [15] for semidefinite programming uses a combination of techniques.

This paper is organized as follows. Section 2 introduces notation and reviews facial reduction. Section 3 studies the solution set of the homogeneous model and the extended-embedding. Section 4 gives a theoretical algorithm for solving arbitrary instances of the primal problem. (An analogous algorithm for the dual is easily stated.) Section 5 gives numerical experiments, highlighting barriers to practical implementation.

2   Background on facial reduction

This section reviews the basic concepts underlying facial reduction algorithms [5, 22, 33]. These algorithms take as input either the primal problem or the dual problem of (1) and replace $\mathcal{K}$ or $\mathcal{K}^*$ with a face containing the primal feasible set or the set of dual feasible slacks. After this replacement, the new problem has optimal value equal to that of the input problem. Moreover, if

the input problem is feasible, the new problem satisfies Slater's condition, and if the input problem is infeasible, the new problem is *strongly* infeasible, meaning a strictly-separating hyperplane (i.e., an improving ray for a dual problem) exists proving infeasibility. For this reason, this replacement is called *regularization.*

These algorithms work by finding a finite sequence of hyperplanes that provably contain the feasible set, where the normal vectors of these hyperplanes are called *facial reduction certificates.* They terminate when facial reduction certificates no longer exist. The number of steps taken by these algorithms depends on the facial reduction certificates used. The minimum number of steps taken is called the *singularity degree* [31] and is intrinsic to the input problem. To explain these ideas in detail, we first review properties of convex cones and their faces.

2.1 Cones and faces

A subset $\mathcal{C}$ of $\mathcal{V}$ is called a *cone* if it is closed under positive scaling, i.e. $\lambda x \in \mathcal{C}$ for any $\lambda > 0$ when $x \in \mathcal{C}$. A *convex cone* is a cone that is convex. The *dual cone* $\mathcal{C}^*$ of any subset $\mathcal{C}$ of $\mathcal{V}$ is the convex cone $\{z \in \mathcal{V} : \langle z, x \rangle \geq 0, \quad \forall x \in \mathcal{C}\}$. In this paper, we are only concerned with convex cones that are also *closed* and *non-empty.* Note if $\mathcal{C}$ is a non-empty, closed, convex cone, then $\mathcal{C}$ contains the origin and $\mathcal{C}^{**} = \mathcal{C}$.

Let $\mathcal{C}$ be a non-empty, closed, convex cone. A *face* $\mathcal{F}$ of $\mathcal{C}$ is a closed convex subset for which $a, b \in \mathcal{C}$ and $\frac{a+b}{2} \in \mathcal{F}$ implies $a, b \in \mathcal{F}$. A face $\mathcal{F}$ of $\mathcal{C}$ is called a *proper face* if it is non-empty and not equal to $\mathcal{C}$. (Note this definition includes $\mathcal{C} \cap (-\mathcal{C})$ as a proper face, which some authors exclude.) Let $z^{\perp}$ denote the hyperplane $\{x \in \mathcal{V} : \langle x, z \rangle = 0\}$. For $z \in \mathcal{C}^*$, the set $\mathcal{C} \cap z^{\perp}$ is easily seen to be a non-empty face (said to be *exposed* by z). The face $\mathcal{C} \cap z^{\perp}$ is hence proper if and only if $z \in \mathcal{C}^* \setminus \mathcal{C}^{\perp}$, where $\mathcal{C}^* \setminus \mathcal{C}^{\perp}$ is the subset of $\mathcal{C}^*$ not contained in the orthogonal complement of the span of $\mathcal{C}$. Since the dual cone $\mathcal{C}^*$ and the proper faces of $\mathcal{C}$ are also nonempty, closed, convex cones, all of these concepts translate if $\mathcal{C}$ is replaced with a proper face $\mathcal{F}$, with the dual cone $\mathcal{C}^*$, or with one of the proper faces of the dual cone. For instance, $z \in \mathcal{F}^* \setminus \mathcal{F}^{\perp}$ exposes a proper face $\mathcal{F} \cap z^{\perp}$ of $\mathcal{F}$, just as $z \in \mathcal{C} \setminus (\mathcal{C}^*)^{\perp}$ exposes a proper face $\mathcal{C}^* \cap z^{\perp}$ of $\mathcal{C}^*$.

2.2 Primal and dual problems

It will be convenient to define primal and dual problems parametrized by a closed, convex cone $\mathcal{C}$:

Definition 1. *For a non-empty, closed, convex cone $\mathcal{C} \subseteq \mathcal{V}$, and the problem data $A : \mathcal{V} \to \mathbb{R}^m$, $b \in \mathbb{R}^m$ and $c \in \mathcal{V}$ of the primal-dual pair* (1)*, let $\mathbf{P}(\mathcal{C})$ denote the primal optimization problem*

$$\begin{array}{ll} \text{minimize} & \langle c, x \rangle \\ \text{subject to} & Ax = b \\ & x \in \mathcal{C}, \end{array}$$

and let $\mathbf{D}(\mathcal{C})$ denote its dual optimization problem

$$\begin{array}{ll} \text{maximize} & \langle b, y \rangle \\ \text{subject to} & c - A^*y = s \\ & s \in \mathcal{C}^*, \ y \in \mathbb{R}^m. \end{array}$$

Note with this notation, $\mathbf{P}(\mathcal{K})$ and $\mathbf{D}(\mathcal{K})$ denote the primal and dual problems of (1), respectively. We call $\inf\left\{\langle c, x \rangle : Ax = b, x \in \mathcal{C}\right\}$ the *optimal value of* $\mathbf{P}(\mathcal{C})$ and $\sup\left\{\langle b, y \rangle : c - A^*y \in \mathcal{C}^*, y \in \mathbb{R}^m\right\}$ the *optimal value of* $\mathbf{D}(\mathcal{C})$. Moreover, we say $\mathbf{P}(\mathcal{C})$ satisfies Slater's condition if $\{x \in \mathcal{V} : Ax =$

$b\}\cap\mathrm{relint}\,\mathcal{C}$ is non-empty. Similarly, $\mathbf{D}(\mathcal{C})$ satisfies Slater's condition if $\{c - A^*y : y \in \mathbb{R}^m\} \cap \mathrm{relint}\,\mathcal{C}^*$ is non-empty.

We also call $x \in \mathcal{C}$ an *improving ray* for $\mathbf{P}(\mathcal{C})$ if $Ax = 0$ and $\langle c, x \rangle < 0$. Similarly, we call $(s, y) \in \mathcal{C}^* \times \mathbb{R}^m$ an improving ray for $\mathbf{D}(\mathcal{C})$ if $\langle b, y \rangle > 0$ and $s = -A^*y$. Note existence of an improving ray for $\mathbf{P}(\mathcal{C})$ implies infeasibility of $\mathbf{D}(\mathcal{C})$, and vice-versa. Finally, we call $(x, s, y) \in \mathcal{C} \times \mathcal{C}^* \times \mathbb{R}^m$ a *complementary solution* for $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$ if x is primal feasible, (s, y) is dual feasible and they satisfy $\langle c, x \rangle = \langle b, y \rangle$, or equivalently, $\langle x, s \rangle = 0$. Note if (x, s, y) is a complementary solution, then x is primal optimal and (s, y) is dual optimal.

2.3 Facial reduction certificates

A facial reduction certificate is a normal vector to a particular type of hyperplane. It is defined in terms of a cone $\mathcal{C} \subseteq \mathcal{V}$ and either the primal problem $\mathbf{P}(\mathcal{C})$ or dual problem $\mathbf{D}(\mathcal{C})$.

Definition 2. *For a non-empty, closed, convex cone $\mathcal{C} \subseteq \mathcal{V}$, and the problem data $A : \mathcal{V} \to \mathbb{R}^m$, $b \in \mathbb{R}^m$ and $c \in \mathcal{V}$ of (1), define facial reduction certificates as follows:*

- *Call $s \in \mathcal{C}^*$ a facial reduction certificate for $\mathbf{P}(\mathcal{C})$ if the hyperplane $s^\perp$ contains the affine set $\{x \in \mathcal{V} : Ax = b\}$ and $\mathcal{C} \cap s^\perp \subseteq \mathcal{C}$ holds strictly.*

- *Call $x \in \mathcal{C}$ a facial reduction certificate for $\mathbf{D}(\mathcal{C})$ if the hyperplane $x^\perp$ contains the affine set $\{c - A^*y : y \in \mathbb{R}^m\}$ and $\mathcal{C}^* \cap x^\perp \subseteq \mathcal{C}^*$ holds strictly.*

A facial reduction certificate for $\mathbf{P}(\mathcal{C})$ exposes a proper face of $\mathcal{C}$ containing the feasible set of $\mathbf{P}(\mathcal{C})$. Similarly, a facial reduction certificate for $\mathbf{D}(\mathcal{C})$ exposes a proper face of $\mathcal{C}^*$ containing the set of dual feasible slacks for $\mathbf{D}(\mathcal{C})$. Hence, existence of these certificates imply failure of Slater's condition for $\mathbf{P}(\mathcal{C})$ or $\mathbf{D}(\mathcal{C})$. For feasible problems, the converse is also true: a facial reduction certificate exists if Slater's condition fails. Specifically, the following is well-known.

Proposition 1. *Let $\mathcal{C} \subseteq \mathcal{V}$ be a non-empty, closed, convex cone. The following statements hold.*

- *Suppose the primal problem $\mathbf{P}(\mathcal{C})$ is feasible. A facial reduction certificate for $\mathbf{P}(\mathcal{C})$ exists if and only if $\{x \in \mathcal{V} : Ax = b\} \cap \mathrm{relint}\,\mathcal{C}$ is empty.*

- *Suppose the dual problem $\mathbf{D}(\mathcal{C})$ is feasible. A facial reduction certificate for $\mathbf{D}(\mathcal{C})$ exists if and only if $\{c - A^*y : y \in \mathbb{R}^m\} \cap \mathrm{relint}\,\mathcal{C}^*$ is empty.*

(See, e.g., Lemma 2 of [25] for a proof, and Theorem 7.1 of [5], Lemma 12.6 of [6], or Lemma 1 of [22] for closely-related statements.)

Facial reduction certificates are also solutions to conic feasibility problems (so-called auxiliary problems). Indeed, a hyperplane contains a non-empty affine set if and only if it has a normal vector satisfying certain linear equations. Hence, the set of facial reduction certificates for $\mathbf{P}(\mathcal{C})$ or $\mathbf{D}(\mathcal{C})$ is defined by particular linear and conic constraints:

Proposition 2. *Let $\mathcal{C} \subseteq \mathcal{V}$ be a non-empty, closed, convex cone. The following statements hold.*

- *$s \in \mathcal{V}$ is a facial reduction certificate for $\mathbf{P}(\mathcal{C})$ if there exists $y \in \mathbb{R}^m$ for which*

$$\langle b, y \rangle = 0, \qquad s = -A^*y, \qquad s \in \mathcal{C}^* \setminus \mathcal{C}^\perp,$$

and all facial reduction certificates are of this form if $\{x \in \mathcal{V} : Ax = b\}$ is non-empty.

5

- $x \in \mathcal{V}$ *is a facial reduction certificate for the dual problem* $\mathbf{D}(\mathcal{C})$ *if*

$$\langle c, x \rangle = 0, \qquad Ax = 0, \qquad x \in \mathcal{C} \setminus (\mathcal{C}^*)^\perp,$$

and all facial reduction certificates are of this form (since $\{c - A^* y : y \in \mathbb{R}^m\}$ *is non-empty).*

Note the constraint $x \in \mathcal{C} \setminus (\mathcal{C}^*)^\perp$ is satisfied if and only if $x \in \mathcal{C}$ and x has non-zero inner-product with any point in relint $\mathcal{C}^*$, and similarly for $s \in \mathcal{C}^* \setminus \mathcal{C}^\perp$.

Optimal facial reduction certificates. Faces exposed by facial reduction certificates are partially-ordered by set inclusion. Thus, there is a natural notion of optimality for certificates. Formally:

Definition 3. *Let* $\mathcal{C} \subseteq \mathcal{V}$ *be a non-empty, closed, convex cone. Let* $Z_p \subseteq \mathcal{V}$ *denote the set of facial reduction certificates for* $\mathbf{P}(\mathcal{C})$ *and* $Z_d \subseteq \mathcal{V}$ *the set of facial reduction certificates for* $\mathbf{D}(\mathcal{C})$.

- $s \in Z_p$ *is an* optimal facial reduction certificate *for* $\mathbf{P}(\mathcal{C})$ *if* $\mathcal{C} \cap s^\perp$ *satisfies*

$$\mathcal{C} \cap s^\perp \subseteq \mathcal{C} \cap \hat{s}^\perp \quad \textit{for all } \hat{s} \in Z_p.$$

- $x \in Z_d$ *is an* optimal facial reduction certificate *for* $\mathbf{D}(\mathcal{C})$ *if* $\mathcal{C}^* \cap x^\perp$ *satisfies*

$$\mathcal{C}^* \cap x^\perp \subseteq \mathcal{C}^* \cap \hat{x}^\perp \quad \textit{for all } \hat{x} \in Z_d.$$

If facial reduction certificates exist, then so do optimal ones. Since facial reduction certificates are closed under addition, this follows easily from the following identities

$$\mathcal{C} \cap (s_1 + s_2)^\perp = \mathcal{C} \cap s_1^\perp \cap s_2^\perp, \qquad \mathcal{C}^* \cap (x_1 + x_2)^\perp = \mathcal{C}^* \cap x_1^\perp \cap x_2^\perp,$$

which hold for any $s_1, s_2 \in \mathcal{C}^*$ and $x_1, x_2 \in \mathcal{C}$. Hence, the sum of a maximal set of linearly independent certificates for $\mathbf{P}(\mathcal{C})$ is an optimal certificate for $\mathbf{P}(\mathcal{C})$, and similarly for $\mathbf{D}(\mathcal{C})$.

2.4 Facial reduction algorithms

A facial reduction algorithm regularizes the primal problem $\mathbf{P}(\mathcal{K})$ of (1) by finding a sequence of facial reduction certificates for $\mathbf{P}(\mathcal{F}_i)$, where each $\mathcal{F}_i$ is a face of $\mathcal{K}$. Similarly, it regularizes the dual problem $\mathbf{D}(\mathcal{K})$ using a sequence of facial reduction certificates for dual problems defined by faces of $\mathcal{K}^*$. We explain the basic idea using the primal problem, and then summarize how it extends to the dual. Additional details can be found in [5, 22, 33].

Facial reduction of the primal problem. Suppose we had a facial reduction certificate $z \in \mathcal{K}^*$ for $\mathbf{P}(\mathcal{K})$. Replacing $\mathcal{K}$ with the face $\mathcal{K} \cap z^\perp$ yields a new primal-dual pair

$$
\begin{array}{llll}
\text{minimize} & \langle c, x \rangle & \text{maximize} & \langle b, y \rangle \\
\text{subject to} & Ax = b & \text{subject to} & c - A^* y = s \\
& x \in \mathcal{K} \cap z^\perp & & s \in (\mathcal{K} \cap z^\perp)^*, \ y \in \mathbb{R}^m,
\end{array}
\tag{3}
$$

where the primal problem $\mathbf{P}(\mathcal{K} \cap z^\perp)$ and the original $\mathbf{P}(\mathcal{K})$ have the same feasible set and equal optimal values (since, by Definition 2, the hyperplane $z^\perp$ contains all solutions to $Ax = b$). We can, of course, repeat this process. For an integer $d_P > 0$, consider the recursion

$$\mathcal{F}_0 = \mathcal{K}, \quad \mathcal{F}_i = \mathcal{F}_{i-1} \cap z_i^\perp \quad i \in \{1, \ldots, d_P\},$$

where $z_i \in \mathcal{F}_{i-1}^*$ is a facial reduction certificate for $\mathbf{P}(\mathcal{F}_{i-1})$. If $\mathcal{C}$ is replaced with any face $\mathcal{F}_i$ in this recursion, the primal problem $\mathbf{P}(\mathcal{F}_i)$ and the original $\mathbf{P}(\mathcal{K})$ also have the same feasible set and equal optimal values, given that each hyperplane $z_i^\perp$ contains all solutions to $Ax = b$. We call replacement of $\mathcal{K}$ with one of these faces *primal regularization.*

Facial reduction algorithms compute the recursion $\mathcal{F}_i = \mathcal{F}_{i-1} \cap z_i^\perp$ and terminate when a facial reduction certificate for $\mathbf{P}(\mathcal{F}_i)$ does not exist. If optimal certificates z_i are used, the length d_P of the sequence $\mathcal{F}_0, \ldots, \mathcal{F}_{d_P}$ is unique and does not depend on the specific certificates z_i. This length is called the *singularity degree* of $\mathbf{P}(\mathcal{K})$. If $\mathbf{P}(\mathcal{K})$ is feasible, the last face $\mathcal{F}_{d_P}$ in the sequence is called the *minimal face* of $\mathbf{P}(\mathcal{K})$. One can show if $\mathbf{P}(\mathcal{K})$ is feasible, the regularized problem $\mathbf{P}(\mathcal{F}_{d_P})$ satisfies Slater's condition, and if $\mathbf{P}(\mathcal{K})$ is infeasible, the dual $\mathbf{D}(\mathcal{F}_{d_P})$ of the regularized primal problem has an improving ray.

Facial reduction of the dual problem. The dual problem $\mathbf{D}(\mathcal{K})$ of (1) is regularized in a similar way. Given a facial reduction certificate $z \in \mathcal{K}$ for $\mathbf{D}(\mathcal{K})$, one can reformulate (1) as:

$$
\begin{array}{llll}
\text{minimize} & \langle c, x \rangle & \text{maximize} & \langle b, y \rangle \\
\text{subject to} & Ax = b & \text{subject to} & c - A^*y = s \\
& x \in (\mathcal{K}^* \cap z^\perp)^* & & s \in \mathcal{K}^* \cap z^\perp, \; y \in \mathbb{R}^m,
\end{array}
\tag{4}
$$

where $\mathcal{K}^*$ has been replaced with the face $\mathcal{K}^* \cap z^\perp$. Since the hyperplane $z^\perp$ contains all vectors of the form $c - A^*y$, the dual problem $\mathbf{D}((\mathcal{K}^* \cap z^\perp)^*)$ and the original $\mathbf{D}(\mathcal{K})$ have the same feasible set and equal optimal values. As with $\mathbf{P}(\mathcal{K})$, we can repeat this process to identify a sequence of faces containing the set of dual feasible slacks. Replacement of $\mathcal{K}^*$ with one of these faces is called *dual regularization.* The singularity degree and minimal face of $\mathbf{D}(\mathcal{K})$ is defined similarly.

2.5 Primal-dual facial reduction asymmetry

While a facial reduction algorithm leaves the feasible set of the input problem unchanged, the same is not true for the corresponding Lagrangian dual problem—in other words, facial reduction is asymmetric with respect to duality. Compare the primal-dual pair (1) with the primal-regularized pair (3). While the *primal* feasible sets are the same, the *dual* feasible set of (3) is potentially larger. An analogous statement holds when comparing (1) with the dual-regularized pair (4); while the *dual* feasible sets are the same, the *primal* feasible set of (4) is potentially larger. Hence, by solving (3) or (4), one won't (generally) find solutions to both the primal $\mathbf{P}(\mathcal{K})$ and dual $\mathbf{D}(\mathcal{K})$.

2.6 Connections with our approach

In this paper, we show optimal facial reduction certificates can be obtained from the homogeneous model when complementary solutions or improving rays do not exist. This allows us to perform primal or dual regularization and resolve the homogeneous model, repeating until a complementary solution or improving ray is obtained. In addition, complementary solutions and improving rays will always be obtained when they exist, even if facial reduction certificates exist as well. As a consequence, we can find complementary solutions without having to first identify the minimal face—i.e., we do not have to regularize until Slater's condition holds. Towards making these statements precise, we now study homogeneous models in more detail.

3 Solutions to homogeneous models

In this section, we study the solution set of the homogeneous model:

Definition 4. *For a non-empty, closed, convex cone $\mathcal{C} \subseteq \mathcal{V}$, and the problem data $A : \mathcal{V} \to \mathbb{R}^m$, $b \in \mathbb{R}^m$ and $c \in \mathcal{V}$ of* (1), *define $\mathbf{H}(\mathcal{C})$ as the convex cone of solutions (x, s, y, τ, κ) to the system:*

$$
\begin{aligned}
Ax - b\tau &= 0, \\
-A^*y - s + c\tau &= 0, \\
\langle b, y \rangle - \langle c, x \rangle - \kappa &= 0, \\
(x, s, y, \tau, \kappa) \in \mathcal{C} \times \mathcal{C}^* \times \mathbb{R}^m \times \mathbb{R}_+ \times \mathbb{R}_+.
\end{aligned}
\tag{5}
$$

Note if $\mathcal{C} = \mathcal{K}$, then $\mathbf{H}(\mathcal{C})$ equals the solution set of the homogeneous model (2). Allowing $\mathcal{C}$ to differ from $\mathcal{K}$ is convenient for stating an algorithm in the next section.

Not every point in $\mathbf{H}(\mathcal{C})$ provides information about the primal-dual pair $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$; indeed, $\mathbf{H}(\mathcal{C})$ contains the zero vector. We therefore focus on two subsets of $\mathbf{H}(\mathcal{C})$ whose points all provide information. We classify points in these subsets based on τ and κ, leveraging the fact $\tau\kappa = 0$; that is, for $(x, s, y, \tau, \kappa) \in \mathbf{H}(\mathcal{C})$,

$$
0 \leq \langle x, s \rangle = \langle x, c\tau - A^*y \rangle = \tau(\langle c, x \rangle - \langle b, y \rangle) = -\tau\kappa \leq 0.
\tag{6}
$$

In other words, for each subset, we consider the three cases $\tau > 0$, $\kappa > 0$ and $\tau = \kappa = 0$, providing novel insight into this latter case.

The first subset of interest is the relative interior of $\mathbf{H}(\mathcal{C})$. Our main result (Theorem 1) implies the results of Table 1—specifically, it shows points in the relative interior yield complementary solutions when $\tau > 0$, improving rays when $\kappa > 0$, and optimal facial reduction certificates when $\tau = \kappa = 0$. Specialized to SDP, this expands analysis of [8, 7].

We next consider the subset $M \cap \mathbf{H}(\mathcal{C})$ where M is a distinguished type of hyperplane introduced in Ye et al. [35] and studied in Freund [10]. As with the relative interior, points in this subset yield complementary solutions when $\tau > 0$ and improving rays when $\kappa > 0$. When $\tau = \kappa = 0$, we obtain either an improving ray or a facial reduction certificate; in other words, we obtain a certificate Slater's condition has failed.

Finally, we show how to find points in both subsets by solving the extended-embedding of Ye et al. [35]—a strictly feasible conic problem with strictly feasible dual.

3.1 The relative interior

The following theorem classifies $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$ by the values of τ and κ. A corollary follows restating key statements in terms of complementary solutions, improving rays and facial reduction certificates for the primal-dual pair given by $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$. The key observation for the case of $\tau = \kappa = 0$ is the following: if $(x, s, y, 0, 0) \in \mathbf{H}(\mathcal{C})$, then, by inspection, so is at least one of the points $(0, s, y, 0, \langle b, y \rangle)$ or $(x, 0, 0, 0, -\langle c, x \rangle)$. This will imply $\langle c, x \rangle = \langle b, y \rangle = 0$ when $\tau = \kappa = 0$ holds for points in the relative interior.

Theorem 1. *Let $\mathcal{C} \subseteq \mathcal{V}$ be a non-empty, closed, convex cone and suppose $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$. Then, the complementarity condition $\tau\kappa = 0$ holds. In addition,*

1. *If $\tau > 0$, then $\frac{1}{\tau}(Ax) = b$ and $\frac{1}{\tau}(A^*y + s) = c$, and $\langle b, y \rangle = \langle c, x \rangle$.*

2. *If $\kappa > 0$, then $Ax = 0$, $A^*y + s = 0$ and $\langle b, y \rangle > \langle c, x \rangle$.*

3. *If $\tau = \kappa = 0$, then $\hat{\tau} = \hat{\kappa} = 0$ and $\langle c, \hat{x} \rangle = \langle b, \hat{y} \rangle = 0$ for all $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa}) \in \mathbf{H}(\mathcal{C})$. Further, letting $\mathcal{F}_p := \mathcal{C} \cap s^\perp$, $\mathcal{F}_d := \mathcal{C}^* \cap x^\perp$, $\mathcal{A}_p := \{x \in \mathcal{V} : Ax = b\}$ and $\mathcal{A}_d := \{c - A^*y : y \in \mathbb{R}^m\}$,*

    (a) *The hyperplane $s^\perp$ contains $\mathcal{A}_p$;*

8

(b) The hyperplane $x^\perp$ contains $\mathcal{A}_d$;

(c) The face $\mathcal{F}_p$ is proper if and only if $\mathcal{A}_p \cap \operatorname{relint} \mathcal{C}$ is empty;

(d) The face $\mathcal{F}_d$ is proper if and only if $\mathcal{A}_d \cap \operatorname{relint} \mathcal{C}^*$ is empty;

(e) At least one of the faces $\mathcal{F}_p$ or $\mathcal{F}_d$ is proper;

(f) The inclusion $\mathcal{F}_p \subseteq \mathcal{C} \cap \hat{s}^\perp$ holds for all $\hat{s} \in \mathcal{C}^*$ satisfying $\mathcal{A}_p \subseteq \hat{s}^\perp$;

(g) The inclusion $\mathcal{F}_d \subseteq \mathcal{C}^* \cap \hat{x}^\perp$ holds for all $\hat{x} \in \mathcal{C}$ satisfying $\mathcal{A}_d \subseteq \hat{x}^\perp$.

Proof. We only prove the third statement, noting the first two are well-known and follow easily from the complementarity condition $\tau\kappa = 0$ implied by (6). To begin, let $w := (x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$ and assume $\tau = \kappa = 0$. Now pick arbitrary $\hat{w} := (\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa}) \in \mathbf{H}(\mathcal{C})$. First note $\hat{\tau} = \hat{\kappa} = 0$; otherwise, we'd have $w - \alpha\hat{w} \notin \mathbf{H}(\mathcal{C})$ for every $\alpha > 0$, contradicting the fact $w \in \operatorname{relint}\mathbf{H}(\mathcal{C})$. Since $\hat{\kappa} = 0$, it follows $\langle c, \hat{x}\rangle - \langle b, \hat{y}\rangle = 0$, i.e., that $\langle c, \hat{x}\rangle = \langle b, \hat{y}\rangle$. We'll now tighten this to $\langle c, \hat{x}\rangle = \langle b, \hat{y}\rangle = 0$. To begin, let $\hat{\theta} := \langle c, \hat{x}\rangle = \langle b, \hat{y}\rangle$. By inspection, at least one point—$(0, \hat{s}, \hat{y}, 0, \hat{\theta})$ or $(\hat{x}, 0, 0, 0, -\hat{\theta})$—is in $\mathbf{H}(\mathcal{C})$. But as just argued, the κ-coordinate of any point in $\mathbf{H}(\mathcal{C})$ must be zero. We conclude $\hat{\theta} = 0$; hence $\langle c, \hat{x}\rangle = \langle b, \hat{y}\rangle = 0$, as desired. We now use this to show statements (3a)-(3b).

To see (3a)-(3b), first note that $\langle \hat{x}, s\rangle = -\langle \hat{x}, A^*y\rangle = -\langle b, y\rangle = 0$ for all solutions $\hat{x}$ of $Ax = b$. Hence, the solution set $\{x \in \mathcal{V} : Ax = b\}$ is contained in the hyperplane $s^\perp$. Likewise, $\langle x, \hat{s}\rangle = \langle c, x\rangle = 0$ for all $\hat{s} \in \{c - A^*y : y \in \mathbb{R}^m\}$; hence, $x^\perp$ contains $\{c - A^*y : y \in \mathbb{R}^m\}$.

We now show (3d). One direction is trivial; if $\mathcal{A}_d$ is contained in a proper face of $\mathcal{C}^*$, then $\mathcal{A}_d \cap \operatorname{relint} \mathcal{C}^*$ must be empty. For the converse direction, suppose $\mathcal{A}_d \cap \operatorname{relint} \mathcal{C}^*$ is empty. The main separation theorem (Theorem 11.3) of [29] states that a hyperplane exists *properly* separating these sets. Using Theorem 11.7 of [29], we can additionally assume this hyperplane passes through the origin since $\mathcal{C}^*$ is a cone. In other words, there exists $\hat{x} \in \mathcal{C}^{**} = \mathcal{C}$ satisfying

$$\langle \hat{x}, c - A^*y\rangle \leq 0, \quad \forall y \in \mathbb{R}^m,$$
$$\langle \hat{x}, z\rangle \neq 0 \text{ for some } z \in \{c - A^*y : y \in \mathbb{R}^m\} \cup \mathcal{C}^*.$$

It follows that $\langle c, \hat{x}\rangle \leq \langle A^*y, \hat{x}\rangle$ for arbitrary $y \in \mathbb{R}^m$, which implies $A\hat{x} = 0$ and $\langle c, \hat{x}\rangle \leq 0$. But $\langle c, \hat{x}\rangle = 0$, otherwise $\hat{x}$ is an improving ray for $\mathbf{P}(\mathcal{C})$, and $(\hat{x}, 0, 0, 0, -\langle c, \hat{x}\rangle) \in \mathbf{H}(\mathcal{C})$ with $\kappa > 0$. Hence, the hyperplane $\hat{x}^\perp$ contains $\{c - A^*y : y \in \mathbb{R}^m\}$ implying $\langle \hat{x}, z\rangle \neq 0$ for some $z \in \mathcal{C}^*$ given proper separation of the sets. That is, $\hat{x}$ exposes a proper face of $\mathcal{C}^*$. We now use $\hat{x}$ to show that x exposes a proper face of $\mathcal{C}^*$ as claimed. Clearly, $\hat{w} := (\hat{x}, 0, 0, 0, 0) \in \mathbf{H}(\mathcal{C})$. Since w is in the relative interior of $\mathbf{H}(\mathcal{C})$, it holds that $w \pm \alpha\hat{w} \in \mathbf{H}(\mathcal{C})$, and thus $x \pm \alpha\hat{x} \in \mathcal{C}$, for some $\alpha > 0$. Hence, for any $u \in \mathcal{C}^*$, the inequality $\langle u, x \pm \alpha\hat{x}\rangle \geq 0$ holds, which in turn implies $\langle u, \hat{x}\rangle = 0$ when $\langle u, x\rangle = 0$. In other words, $\mathcal{C}^* \cap x^\perp$ is contained in a proper face, i.e.,

$$\mathcal{C}^* \cap x^\perp \subseteq \mathcal{C}^* \cap \hat{x}^\perp, \tag{7}$$

and is hence proper.

Applying the argument of the previous paragraph to the set $\mathcal{A}_p \cap \operatorname{relint} \mathcal{C}$ shows (3c).

Statement (3e) follows if at least one of the sets—$\{x \in \mathcal{V} : Ax = b\} \cap \operatorname{relint} \mathcal{C}$ or $\{c - A^*y : y \in \mathbb{R}^m\} \cap \operatorname{relint} \mathcal{C}^*$—is empty. Suppose this weren't the case. Then, Slater's condition is satisfied for both $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$ showing existence of an optimal primal-dual solution with zero duality gap [4, Section 7.2.2]. Hence, there exists a point in $\mathbf{H}(\mathcal{C})$ with $\tau > 0$, contradicting the assumption that (x, s, y, τ, κ) is in the relative interior of $\mathbf{H}(\mathcal{C})$.

The same argument that shows the containment (7) shows (3f) and (3g). $\qquad\square$

The following corollary arises simply from definitions. Variants of the first two statements (and their converses) are well-known for semidefinite optimization [8, Theorem 5.3.2]. Note taking $\mathcal{C} = \mathcal{K}$ yields Table 1.

Corollary 1. *Let $\mathcal{C} \subseteq \mathcal{V}$ be a non-empty, closed, convex cone, and suppose $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$. The following statements hold for the primal-dual pair $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$.*

1. *If $\tau > 0$, then $\frac{1}{\tau}(x, s, y)$ is a complementary solution for $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$.*

2. *If $\kappa > 0$, then x is an improving ray for $\mathbf{P}(\mathcal{C})$ and/or (s, y) is an improving ray for $\mathbf{D}(\mathcal{C})$.*

3. *If $\tau = \kappa = 0$, then x and/or s are facial reduction certificates.*

Moreover, converses of the first two statements hold: if $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$ have a complementary solution, then $\tau > 0$; if $\mathbf{P}(\mathcal{C})$ and/or $\mathbf{D}(\mathcal{C})$ have an improving ray, then $\kappa > 0$.

We now strengthen the third statement of Corollary 1, showing facial reduction certificates are obtained for both problems if they both fail Slater's condition. We then illustrate the second statement does not have an analogous strengthening—that is, improving rays are not necessarily obtained for both problems even if they exist.

Facial reduction certificates for the primal and dual. The third statement of Corollary 1 does not guarantee s and x are *both* facial reduction certificates when such certificates exist for *both* primal and dual problems. Statements (3c)-(3d) of Theorem 1 provide this guarantee. Moreover, statements (3f)-(3g) imply theses certificates are optimal. Formally:

Corollary 2. *Let $\mathcal{C} \subseteq \mathcal{V}$ be a non-empty, closed, convex cone, and suppose $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$. If $\tau = \kappa = 0$, the following statements hold:*

- *The set $\{x \in \mathcal{V} : Ax = b\} \cap \operatorname{relint} \mathcal{C}$ is empty if and only if s is a facial reduction certificate for $\mathbf{P}(\mathcal{C})$.*

- *The set $\{c - A^*y : y \in \mathbb{R}^m\} \cap \operatorname{relint} \mathcal{C}^*$ is empty if and only if x is a facial reduction certificate for $\mathbf{D}(\mathcal{C})$.*

Moreover, if x (resp. s) is a facial reduction certificate, then x (resp. s) is optimal in the sense of Definition 3.

This corollary can be compared to Proposition 1 (the basis of the facial reduction algorithm [22]), which asserted that, for feasible problems, existence of facial reduction certificates is equivalent to failure of Slater's condition. Also note optimality of certificates is ensured by the restriction to $\operatorname{relint} \mathbf{H}(\mathcal{C})$; note that [31, Procedure 1] and [33, Algorithm 4.1] find optimal certificates using similar restrictions.

The next example illustrates the first two statements of Corollary 2. Here, the primal-dual pair has finite but non-zero duality gap. This implies $\tau = \kappa = 0$ for all points in $\operatorname{relint} \mathbf{H}(\mathcal{C})$ and that Slater's condition fails for both the primal and dual.

Example 1 (Example from [2] with positive duality gap.)**.** *Let $\mathcal{Q}^n := \left\{ x_1 \geq \sqrt{\sum_{i=2}^{n} x_i^2} \right\}$ denote the quadratic (Lorentz) cone. The following primal-dual pair has a positive duality gap of 1:*

$$
\begin{array}{ll}
\text{minimize} & x_3 \\
\text{subject to} & x_1 + x_2 + x_4 + x_5 = 0 \\
& -x_3 + x_4 = 1 \\
& x \in \mathcal{Q}^3 \times \mathcal{Q}^2
\end{array}
\qquad
\begin{array}{ll}
\text{maximize} & y_2 \\
\text{subject to} & \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} y_1 \\ y_1 \\ -y_2 \\ y_1+y_2 \\ y_1 \end{pmatrix} = s \\
& s \in \mathcal{Q}^3 \times \mathcal{Q}^2.
\end{array}
$$

Indeed, if $x \in \mathcal{Q}^3 \times \mathcal{Q}^2$, then $x_1 + x_2 \geq 0$ and $x_4 + x_5 \geq 0$; if, in addition, $x_1 + x_2 + x_4 + x_5 = 0$, then $x_1 + x_2 = 0$, implying $x_3 = 0$ if $(x_1, x_2, x_3) \in \mathcal{Q}^3$. On the other hand, dual feasible points satisfy $s_1 = s_2$, which in turn implies $s_3 = 0$, i.e., $y_2 = -1$. Since both problems are feasible, the duality gap is $0 - (-1) = 1$.

Since there is a duality gap of 1, any point in $\mathrm{relint}\, \mathbf{H}(\mathcal{Q}^3 \times \mathcal{Q}^2)$ satisfies $\tau = \kappa = 0$, e.g.,

$$\hat{x} = (1, -1, 0, 0, 0)^T, \quad \hat{s} = (1, 1, 0, 1, 1)^T, \quad \hat{y} = (-1, 0)^T, \quad \hat{\tau} = \hat{\kappa} = 0.$$

We see $\hat{s}$ and $\hat{x}$ are facial reduction certificates for the primal and dual, as predicted by Corollary 2.

Improving rays for the primal and dual. Statement two of Corollary 1 does not guarantee s and x are *both* improving rays when such rays exists for *both* primal and dual problems. Unfortunately, this statement cannot be strengthened—in the $\kappa > 0$ case, there are instances for which relative interior points do not yield improving rays for both problems, even if these rays exist. This is a known shortcoming of the homogeneous self-dual model that occurs even in the linear programming case (see, e.g., [35]). The following example illustrates this shortcoming:

Example 2. *Consider the following primal-dual pair of linear programs*

$$\begin{array}{ll} \text{minimize} & -x_1 \\ \text{subject to} & x_1 - x_2 = 0 \\ & -(x_1 - x_2) = 1 \\ & x \in \mathbb{R}^2_+ \end{array} \qquad \begin{array}{ll} \text{maximize} & y_2 \\ \text{subject to} & \begin{pmatrix} -1 \\ 0 \end{pmatrix} - \begin{pmatrix} y_1 - y_2 \\ -(y_1 - y_2) \end{pmatrix} = s \\ & s \in \mathbb{R}^2_+, \end{array}$$

where both the primal and dual problem are infeasible. Indeed, the point $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa}) \in \mathrm{relint}\, \mathbf{H}(\mathbb{R}^2_+)$ yields an improving ray $\hat{x}$ for the primal and an improving ray $\hat{y}$ for the dual, where

$$\hat{x} = (1, 1)^T, \quad \hat{s} = (0, 0)^T, \quad \hat{y} = (1, 1)^T, \quad \hat{\tau} = 0, \quad \hat{\kappa} = 2.$$

Nevertheless, the entire family of points:

$$\tilde{x} = (r, r)^T, \quad \tilde{s} = (0, 0)^T, \quad \tilde{y} = (t, t)^T, \quad \tilde{\tau} = 0, \quad \tilde{\kappa} = r + t, \quad \text{for } r > -t \geq 0,$$

are also in the relative interior of solutions to the homogeneous model, and only give improving rays for the primal problem.

Note this example illustrates one cannot in general decide infeasibility of the primal and dual from a single point in $\mathrm{relint}\, \mathbf{H}(\mathcal{C})$ with $\kappa > 0$. We will revisit this issue in Section 4.3, as it complicates a presented solution algorithm for finding optimal values.

3.2 The intersection with distinguished hyperplanes

The next subset of $\mathbf{H}(\mathcal{C})$ of interest is its intersection with a distinguished type of hyperplane. Such a hyperplane M is defined by $\mu > 0$ and a fixed point $(\hat{x}, \hat{s}, \hat{\tau}, \hat{\kappa}) \in \mathrm{relint}(\mathcal{C} \times \mathcal{C}^* \times \mathbb{R}_+ \times \mathbb{R}_+)$ via

$$M := \{(x, s, y, \tau, \kappa) : \langle \hat{s}, x \rangle + \langle \hat{x}, s \rangle + \hat{\tau}\kappa + \hat{\kappa}\tau = \mu\}.$$

For a particular μ, membership in M is an implicit constraint of the extended-embedding of [35]. Also, as shown in [10], it can be interpreted as a norm constraint on $(x, s, \tau, \kappa) \in \mathcal{C} \times \mathcal{C}^* \times \mathbb{R}_+ \times \mathbb{R}_+$ when $\mathcal{C}$ is a *proper cone*, i.e., a cone that is full-dimensional, pointed, convex, and closed.

As we now show, from $(x, s, y, \tau, \kappa) \in \mathbf{H}(\mathcal{C}) \cap M$, one always obtains one of the following objects: a complementary solution or a certificate that the primal or dual has failed Slater's condition. As with $\mathrm{relint}\, \mathbf{H}(\mathcal{C})$, complementary solutions are obtained when $\tau > 0$ and improving rays when $\kappa > 0$. The case of $\tau = \kappa = 0$, however, is now more delicate: we either obtain a facial reduction certificate *or* an improving ray. Facial reduction certificates are also not necessarily optimal.

Theorem 2. *Let $\mathcal{C} \subseteq \mathcal{V}$ be a non-empty closed, convex cone. For a scalar $\mu > 0$ and point $(\hat{x}, \hat{s}, \hat{\tau}, \hat{\kappa})$ in the relative interior of $\mathcal{C} \times \mathcal{C}^* \times \mathbb{R}_+ \times \mathbb{R}_+$ consider the hyperplane*

$$M := \{(x, s, y, \tau, \kappa) : \langle \hat{s}, x \rangle + \langle \hat{x}, s \rangle + \hat{\tau}\kappa + \hat{\kappa}\tau = \mu\}.$$

For $(x, s, y, \tau, \kappa) \in M \cap \mathbf{H}(\mathcal{C})$, the complementarity condition $\tau\kappa = 0$ holds. In addition,

1. *If $\tau > 0$, then $\frac{1}{\tau}(x, s, y)$ is a complementary solution of $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$.*

2. *If $\kappa > 0$, then x is an improving ray for $\mathbf{P}(\mathcal{C})$ and/or (s, y) is an improving ray for $\mathbf{D}(\mathcal{C})$.*

3. *If $\tau = \kappa = 0$, at least one of the following statements is true:*

    (a) *x is an improving ray for $\mathbf{P}(\mathcal{C})$*

    (b) *(s, y) is an improving ray for $\mathbf{D}(\mathcal{C})$*

    (c) *x is a facial reduction certificate for $\mathbf{D}(\mathcal{C})$, i.e., $\mathcal{C}^* \cap x^\perp \subseteq \mathcal{C}^*$ holds strictly and the hyperplane $x^\perp$ contains $\mathcal{A}_d := \{c - A^* y : y \in \mathbb{R}^m\}$.*

    (d) *s is a facial reduction certificate for $\mathbf{P}(\mathcal{C})$, i.e., $\mathcal{C} \cap s^\perp \subseteq \mathcal{C}$ holds strictly and the hyperplane $s^\perp$ contains $\mathcal{A}_p := \{x \in \mathcal{V} : Ax = b\}$.*

Proof. The first two statements are trivial from the complementarity condition $\tau\kappa = 0$. For the third statement, we note from $\kappa = 0$ that $\langle b, y \rangle = \langle c, x \rangle$ and consider the three cases: $\langle b, y \rangle = \langle c, x \rangle > 0$, $\langle b, y \rangle = \langle c, x \rangle < 0$ and $\langle c, x \rangle = \langle b, y \rangle = 0$. In the first case, its trivial to check that (s, y) is an improving ray and in the second that x is an improving ray. In the third case, where $\langle c, x \rangle = \langle b, y \rangle = 0$, it follows $x^\perp$ and $s^\perp$ contain $\mathcal{A}_d$ and $\mathcal{A}_p$, respectively; see the proof for statements (3a)-(3b) of Theorem 1. In addition, at least one exposes a proper face since $\tau = \kappa = 0$ implies $\langle \hat{s}, x \rangle + \langle \hat{x}, s \rangle = \mu > 0$; that is, the face exposed by $x^\perp$ doesn't contain $\hat{s}$ if $\langle \hat{s}, x \rangle > 0$ and the face exposed by $s^\perp$ doesn't contain $\hat{x}$ if $\langle \hat{x}, s \rangle > 0$. $\square$

We conclude with descriptions of relint($M \cap \mathbf{H}(\mathcal{C})$). We'll use this description in the next section to find points in relint $\mathbf{H}(\mathcal{C})$.

Theorem 3. *Let $\mathcal{C} \subseteq \mathcal{V}$ be a non-empty closed, convex cone. For a scalar $\mu > 0$ and point $(\hat{x}, \hat{s}, \hat{\tau}, \hat{\kappa})$ in the relative interior of $\mathcal{C} \times \mathcal{C}^* \times \mathbb{R}_+ \times \mathbb{R}_+$ consider the hyperplane*

$$M := \{(x, s, y, \tau, \kappa) : \langle \hat{s}, x \rangle + \langle \hat{x}, s \rangle + \hat{\tau}\kappa + \hat{\kappa}\tau = \mu\}.$$

The following statements hold:

1. $\text{relint}\left(M \cap \mathbf{H}(\mathcal{C})\right) = M \cap \text{relint}\,\mathbf{H}(\mathcal{C})$

2. *The conic hull of $M \cap \text{relint}\,\mathbf{H}(\mathcal{C})$ equals $\text{relint}\,\mathbf{H}(\mathcal{C})$.*

Proof. We first show the second statement. The inclusion $\subseteq$ is trivial: since $M \cap \text{relint}\,(\mathbf{H}(\mathcal{C}))$ is a subset of $\text{relint}\,\mathbf{H}(\mathcal{C})$, its conic hull is contained in $\text{relint}\,\mathbf{H}(\mathcal{C})$ since $\text{relint}\,\mathbf{H}(\mathcal{C})$ is a cone. For the reverse, let $(x, s, y, \tau, \kappa) \in \text{relint}\,\mathbf{H}(\mathcal{C})$. By Corollary 1, $\tau > 0$, $\kappa > 0$, $s \notin \mathcal{C}^\perp$ or $x \notin (\mathcal{C}^*)^\perp$. Hence, the sum $\langle \hat{x}, s \rangle + \langle x, \hat{s} \rangle + \hat{\tau}\kappa + \tau\hat{\kappa}$ equals some positive number, say, α. It follows $\lambda(x, s, y, \tau, \kappa) \in M \cap \text{relint}\,\mathbf{H}(\mathcal{C})$ for the strictly positive number $\lambda = \frac{\mu}{\alpha}$ by the fact $\text{relint}\,\mathbf{H}(\mathcal{C})$ is a cone.

The second statement implies $M \cap \text{relint}\,\mathbf{H}(\mathcal{C})$ is non-empty; hence, the first statement follows from Corollary 6.5.1 of [29]. $\square$

Remark 1. *Note the special structure of M is crucial to statements* (3c) *and* (3d) *of Theorem 2. If, for instance, we replaced M with an arbitrary hyperplane, the inclusions $\mathcal{C} \cap s^\perp \subseteq \mathcal{C}$ and $\mathcal{C}^* \cap x^\perp \subseteq \mathcal{C}^*$ wouldn't necessarily hold strictly. It is also crucial to Theorem 3: an arbitrary hyperplane doesn't necessarily intersect the relative interior of $\mathbf{H}(\mathcal{C})$.*

3.3 Finding solutions via extended-embeddings

We now discuss how to find points in relint $\mathbf{H}(\mathcal{C})$ and $M \cap \mathbf{H}(\mathcal{C})$, where M is a hyperplane of the distinguished type considered in Section 3.2. For convenience, we fix $(\hat{x}, \hat{s}, \hat{\tau}, \hat{\kappa}) \in \text{relint}(\mathcal{C} \times \mathcal{C}^* \times \mathbb{R}_+ \times \mathbb{R}_+)$ and define M as a function of μ:

$$M(\mu) := \{(x, s, y, \tau, \kappa) : \langle \hat{s}, x \rangle + \langle \hat{x}, s \rangle + \hat{\tau}\kappa + \hat{\kappa}\tau = \mu\}. \tag{8}$$

As shown in Ye et al. [35], points in $M(\langle \hat{x}, \hat{s} \rangle + \hat{\tau}\hat{\kappa}) \cap \mathbf{H}(\mathcal{C})$ are obtained by solving an *extended-embedding*. Hence, points in $M(\mu) \cap \mathbf{H}(\mathcal{C})$ for arbitrary $\mu > 0$ can be obtained by rescaling solutions. We'll review this result and then build on it, showing points in relint $\mathbf{H}(\mathcal{C})$ are obtained from relative interior solutions to the extended-embedding (which is not obvious).

Picking $\hat{y} \in \mathbb{R}^m$ and letting $\alpha = \langle \hat{s}, \hat{x} \rangle + \hat{\tau}\hat{\kappa}$, the extended-embedding takes the following form

$$
\begin{aligned}
\text{minimize} \quad & \alpha\theta \\
\text{subject to} \quad & Ax - b\tau = r_p\theta \\
& -A^*y - s + c\tau = r_d\theta \\
& \langle b, y \rangle - \langle c, x \rangle - \kappa = r_g\theta \\
& \langle r_p, y \rangle + \langle r_d, x \rangle + r_g\tau = -\alpha,
\end{aligned}
\tag{9}
$$

where $(x, s, y, \tau, \kappa, \theta) \in \mathcal{C} \times \mathcal{C}^* \times \mathbb{R}^m \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$ is the decision variable and r_p, r_d, and r_g are additional parameters defined by $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa})$ via

$$r_p = A\hat{x} - b\hat{\tau}, \qquad r_d = -A^*\hat{y} - \hat{s} + c\hat{\tau}, \qquad r_g = \langle b, \hat{y} \rangle - \langle c, \hat{x} \rangle - \hat{\kappa}.$$

As shown in [35], the point $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa}, 1)$ is strictly feasible; further, the optimal value is zero and it is attained, implying $\theta = 0$ at optimality. This latter property is easy to see from duality. Precisely, the dual problem is to maximize $-\alpha\bar{\theta}$ over dual variables $(\bar{x}, \bar{s}, \bar{y}, \bar{\tau}, \bar{\kappa}, \bar{\theta})$ satisfying identical constraints. Hence, an optimal solution, which exists by strict feasibility of (9) and its dual, can be viewed as a dual optimal solution. We conclude $\alpha\theta = -\alpha\theta$ at optimality since there is no duality gap (by strict feasibility). This shows $\theta = 0$ since $\alpha > 0$.

Note if $\theta = 0$, then the first three constraints of (9) reduce to the defining equations of $\mathbf{H}(\mathcal{C})$; i.e., a point $(x, s, y, \tau, \kappa, \theta)$ with $\theta = 0$ satisfies these constraints if and only if $(x, s, y, \tau, \kappa) \in \mathbf{H}(\mathcal{C})$. For points in $\mathbf{H}(\mathcal{C})$, the remaining constraint is also equivalent to membership in $M(\alpha)$:

$$
\begin{aligned}
-\alpha &= \langle r_p, y \rangle + \langle r_d, x \rangle + r_g\tau \\
&= \langle A\hat{x} - b\hat{\tau}, y \rangle + \langle -A^*\hat{y} - \hat{s} + c\hat{\tau}, x \rangle + (\langle b, \hat{y} \rangle - \langle c, \hat{x} \rangle - \hat{\kappa})\tau \\
&= \langle -\hat{x}, -A^*y + c\tau \rangle - \langle \hat{s}, x \rangle - \langle \hat{y}, Ax - b\tau \rangle - \hat{\tau}(\langle b, y \rangle - \langle c, x \rangle) - \hat{\kappa}\tau \\
&= -\langle \hat{x}, s \rangle - \langle \hat{s}, x \rangle - \hat{\tau}\kappa - \hat{\kappa}\tau.
\end{aligned}
$$

It follows a feasible point $(x, s, y, \tau, \kappa, \theta)$ is optimal if and only if $\theta = 0$ and $(x, s, y, \tau, \kappa) \in M(\alpha) \cap \mathbf{H}(\mathcal{C})$—a result originally due to [35]. We restate this as the first statement of the following theorem. The second statement—which to our knowledge is new—describes the relative interior of the optimal solution set of (9) in terms of relint $\mathbf{H}(\mathcal{C})$.

Theorem 4. *Let $\Omega \times \{0\} \subseteq \mathcal{C} \times \mathcal{C}^* \times \mathbb{R}^m \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$ denote the set of optimal solutions of* (9) *and $M(\mu)$ the hyperplane* (8). *For $\mu = \alpha$, the following statements hold:*

1. *$\Omega = M(\mu) \cap \mathbf{H}(\mathcal{C})$.*

2. *$\operatorname{relint} \Omega = M(\mu) \cap \operatorname{relint} \mathbf{H}(\mathcal{C})$; in particular, the conic hull of $\operatorname{relint} \Omega$ equals $\operatorname{relint} \mathbf{H}(\mathcal{C})$.*

Proof. The first statement is from [35]. The second follows immediately from the first statement and Theorem 2. $\qquad \square$

Remark 2. *Note the second statement is* not *a trivial consequence of the first; that is, for convex sets $\mathcal{C}_1$ and $\mathcal{C}_2$, the relative interior of $\mathcal{C}_1 \cap \mathcal{C}_2$ does not in general equal $\operatorname{relint} \mathcal{C}_1 \cap \operatorname{relint} \mathcal{C}_2$, even if one of these sets is affine.*

In conclusion, we obtain points in $M(\alpha) \cap \mathbf{H}(\mathcal{C})$ and $\operatorname{relint} \mathbf{H}(\mathcal{C})$ from points in $\Omega \times \{0\}$ and $\operatorname{relint} \Omega \times \{0\}$, respectively. We also note when $\mathcal{C}$ is a symmetric cone (and the range of A equals $\mathbb{R}^m$), the central path of the extended-embedding exists and converges to a point in $\operatorname{relint} \Omega \times \{0\}$ by strict feasibility and results of [13, 27]. Hence, for the cases of semidefinite, linear, and second-order cone optimization, we can approximate points in $\operatorname{relint} \Omega \times \{0\}$—and therefore points in $\operatorname{relint} \mathbf{H}(\mathcal{C})$—using interior-point methods. We'll investigate convergence behavior—focusing on cases where $\tau = \kappa = 0$ for all $(x, s, y, \tau, \kappa) \in \mathbf{H}(\mathcal{C})$ in Section 5. We next develop a solution algorithm assuming oracle access to $\mathbf{H}(\mathcal{C})$.

4 An algorithm based on homogeneous models

As we've shown, we always obtain complementary solutions, improving rays, or facial reduction certificates from two subsets of $\mathbf{H}(\mathcal{C})$—the relative interior of $\mathbf{H}(\mathcal{C})$ and, for a distinguished type of hyperplane M, the set $M \cap \mathbf{H}(\mathcal{C})$. We've also seen facial reduction certificates are obtained from $\operatorname{relint} \mathbf{H}(\mathcal{C})$ precisely when they are needed—i.e., when complementary solutions or improving rays don't exist (Corollary 1). These facts suggest an iterative procedure that finds a point in $\operatorname{relint} \mathbf{H}(\mathcal{C})$, regularizes the primal or dual if necessary, and repeats until an improving ray or complementary solution is obtained. In this section we develop such a procedure.

To design a useful procedure, one must make an upfront decision: should regularization leave the primal or the dual optimal value unchanged? (Recall from Section 2.5 we cannot guarantee both are unchanged.) In this section, we choose the former, stating an algorithm that finds a cone $\mathcal{C}$ such that $\mathbf{P}(\mathcal{C})$ and $\mathbf{P}(\mathcal{K})$ have equal optimal values and a point $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$ satisfying $\tau > 0$ or $\kappa > 0$. We then interpret (x, s, y, τ, κ) in terms of $\mathbf{P}(\mathcal{K})$, showing an optimal solution is obtained when $\tau > 0$ and a certificate of infeasibility when $\kappa > 0$ and $\langle b, y \rangle > 0$. In the remaining case where, $\kappa > 0$ and $\langle b, y \rangle \leq 0$, the primal problem $\mathbf{P}(\mathcal{K})$ is either unbounded or infeasible. A trivial extension (given in Section 4.3) resolves this ambiguity. The procedure appears in Algorithm 1.

4.1 Basic properties

As indicated, Algorithm 1 solves a sequence of homogeneous models, terminating when $\tau > 0$ or $\kappa > 0$. If $\tau = \kappa = 0$, primal regularization is performed if $s \notin \mathcal{C}^\perp$ and dual regularization is performed otherwise. Note when $\tau = \kappa = 0$, the condition $s \notin \mathcal{C}^\perp$ holds if and only if $\mathbf{P}(\mathcal{C})$ fails Slater's condition—a consequence of Theorem 1-(3c). Hence, if necessary, this algorithm performs primal regularization until $\mathbf{P}(\mathcal{C})$ satisfies Slater's condition, then switches to dual regularization—never switching back. Precise statements of this property and others follow.

14

Algorithm 1: Given $\mathcal{K}$, returns $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$ satisfying $\tau > 0$ or $\kappa > 0$ where $\mathbf{P}(\mathcal{C})$ and $\mathbf{P}(\mathcal{K})$ have equal optimal values.

$\mathcal{C} \leftarrow \mathcal{K}$
repeat
    Find (x, s, y, τ, κ) in the relative interior of $\mathbf{H}(\mathcal{C})$
    if $\tau > 0$ **or** $\kappa > 0$ **then**
        return (x, s, y, τ, κ)
    else
        if $s \notin \mathcal{C}^\perp$ **then**
            Regularize primal: $\mathcal{C} \leftarrow \mathcal{C} \cap s^\perp$
        else
            Regularize dual: $\mathcal{C} \leftarrow (\mathcal{C}^* \cap x^\perp)^*$
        end
    end
until algorithm returns;

Theorem 5. *Let $\mathcal{A}_p := \{x \in \mathcal{V} : Ax = b\}$, and let $\theta_p \in \mathbb{R} \cup \{\pm\infty\}$ denote the optimal value of $\mathbf{P}(\mathcal{K})$, i.e.,*

$$\theta_p := \inf \left\{ \langle c, x \rangle : x \in \mathcal{A}_p \cap \mathcal{K} \right\}.$$

Algorithm 1 has the following basic properties:

1. *Algorithm 1 terminates in finitely-many iterations. Further, it terminates after one iteration, with $\mathcal{C} = \mathcal{K}$, if and only if a complementary solution, primal improving ray or dual improving exists for the primal-dual pair $\mathbf{P}(\mathcal{K})$ and $\mathbf{D}(\mathcal{K})$.*

2. *At each iteration, the optimal values of the primal problem $\mathbf{P}(\mathcal{K})$ and the regularized problem $\mathbf{P}(\mathcal{C})$ are equal, i.e.,*

$$\theta_p = \inf \left\{ \langle c, x \rangle : x \in \mathcal{A}_p \cap \mathcal{C} \right\}.$$

3. *Suppose the dual regularization step $\mathcal{C} \leftarrow (\mathcal{C}^* \cap x^\perp)^*$ executes, and let $\mathcal{C}'$ and $\mathcal{C}''$ denote $\mathcal{C}$ just before and after execution. The following statements hold:*

   (a) *The primal problems $\mathbf{P}(\mathcal{C}')$ and $\mathbf{P}(\mathcal{C}'')$ satisfy Slater's condition, that is, $\mathcal{A}_p \cap \operatorname{relint} \mathcal{C}'$ and $\mathcal{A}_p \cap \operatorname{relint} \mathcal{C}''$ are both non-empty. Further, $\mathcal{A}_p \cap \operatorname{relint} \mathcal{C}' \subseteq \mathcal{A}_p \cap \operatorname{relint} \mathcal{C}''$.*

   (b) *The primal regularization step $\mathcal{C} \leftarrow \mathcal{C} \cap s^\perp$ is not executed at any following iteration.*

4. *The dual regularization step $\mathcal{C} \leftarrow (\mathcal{C}^* \cap x^\perp)^*$ executes if and only if one of the following statements hold:*

   (a) *The optimal value of $\mathbf{P}(\mathcal{K})$ is finite and unattained;*

   (b) *The optimal value of $\mathbf{P}(\mathcal{K})$ is unbounded below ($\theta_p = -\infty$) and the set of improving rays $\{x \in \mathcal{K} : Ax = 0, \langle c, x \rangle < 0\}$ is empty.*

5. *At termination, $\mathcal{A}_p \cap \mathcal{K} \subseteq \mathcal{A}_p \cap \mathcal{C}$, and with strict inclusion only if (4a) or (4b) holds.*

Proof. In arguments below, we call $\mathcal{C} \leftarrow \mathcal{C} \cap s^\perp$ the primal regularization step, and $\mathcal{C} \leftarrow (\mathcal{C}^* \cap x^\perp)^*$ the dual regularization step. When one of these steps executes, we let $\mathcal{C}'$ and $\mathcal{C}''$ denote the cone $\mathcal{C}$ before and after execution, respectively.

Statement 4. We will show $\mathbf{P}(\mathcal{C}')$ satisfies (4a) or (4b) when a dual regularization step executes. Using this, we then show $\mathbf{P}(\mathcal{K})$ also satisfies (4a) or (4b). We use the following facts from Theorem 1: when the dual regularization step executes, all points in $\mathbf{H}(\mathcal{C}')$ satisfy $\tau = \kappa = 0$; and, when the dual regularization step executes, no facial reduction certificate exists for $\mathbf{P}(\mathcal{C}')$ (since $s \in \mathcal{C}'^\perp$).

To begin, we first show $\mathbf{P}(\mathcal{C}')$ is feasible, and hence has finite optimal value or is unbounded below. Suppose $\mathbf{P}(\mathcal{C}')$ is infeasible. If $\{x \in \mathcal{V} : Ax = b\}$ is empty, then there exists $\hat{y}$ for which $\langle b, \hat{y} \rangle = 1$ and $A^*\hat{y} = 0$. Hence $(0, 0, \hat{y}, 0, \langle b, \hat{y} \rangle)$ is a point in $\mathbf{H}(\mathcal{C}')$ with $\kappa > 0$, which is a contradiction. On the other hand, if $\{x \in \mathcal{V} : Ax = b\}$ is non-empty, there exists a hyperplane properly seperating $\mathcal{A}_p := \{x \in \mathcal{V} : Ax = b\}$ from the relative interior of $\mathcal{C}'$. That is, there exists $\hat{s} \in \mathcal{C}'^*$ for which

$$\langle \hat{s}, x \rangle \leq 0, \quad \forall x \in x_0 + \text{null } A,$$

$$\langle \hat{s}, z \rangle \neq 0 \text{ for some } z \in (x_0 + \text{null } A) \cup \mathcal{C}',$$

where $x_0 \in \mathcal{A}_p$ and $\mathcal{A}_p = x_0 + \text{null } A$. This implies $\hat{s} \in (\text{null } A)^\perp = \text{range } A^*$. Hence, $\hat{s} = -A^*\hat{y}$ for some $\hat{y}$, where, evidently, $\langle \hat{s}, x \rangle = -\langle b, \hat{y} \rangle \leq 0$ for all $x \in \mathcal{A}_p$. If $\langle b, \hat{y} \rangle = 0$, then $\langle \hat{s}, z \rangle \neq 0$ for some $z \in \mathcal{C}'$ by proper separation of the sets. Hence, $\hat{s}$ is a facial reduction certificate which, as mentioned above, cannot exist. On the other hand, if $\langle b, \hat{y} \rangle > 0$, then $(0, -A^*\hat{y}, \hat{y}, 0, \langle b, \hat{y} \rangle)$ is a point in $\mathbf{H}(\mathcal{C}')$ with $\kappa > 0$, which is a contradiction. Hence, $\mathbf{P}(\mathcal{C}')$ must be feasible and either has finite optimal value or an optimal value that is unbounded below.

We have established that $\mathbf{P}(\mathcal{C}')$ is feasible and that no facial reduction certificate for $\mathbf{P}(\mathcal{C}')$ exists. Hence, by Proposition 1, $\mathbf{P}(\mathcal{C}')$ is strictly feasible. Now suppose that $\mathbf{P}(\mathcal{C}')$ has a finite optimal value. Then, by Slater's condition, the dual $\mathbf{D}(\mathcal{C}')$ of $\mathbf{P}(\mathcal{C}')$ has equal optimal value that is attained. Hence, if $\mathbf{P}(\mathcal{C}')$ attains its optimal value, $\mathbf{P}(\mathcal{C}')$ has a complementary solution (x, s, y) where $(x, s, y, 1, 0)$ is a point in $\mathbf{H}(\mathcal{C}')$ with $\tau > 0$; a contradiction. Suppose next the optimal value equals $-\infty$. If an improving ray $\hat{x}_{ray}$ exists, then $(\hat{x}_{ray}, 0, 0, 0, -\langle c, \hat{x} \rangle_{ray})$ is a point in $\mathbf{H}(\mathcal{C}')$ with $\kappa > 0$; a contradiction. Hence, $\mathbf{P}(\mathcal{C}')$ satisfies (4a) or (4b).

We now show $\mathbf{P}(\mathcal{K})$ also satisfies (4a) or (4b). Consider the first time the dual regularization step executes. Since the feasible sets of $\mathbf{P}(\mathcal{C}')$ and $\mathbf{P}(\mathcal{K})$ are equal, it trivially follows that $\mathbf{P}(\mathcal{K})$ satisfies (4a) if $\mathbf{P}(\mathcal{C}')$ does. If $\mathbf{P}(\mathcal{C}')$ satisfies (4b), then $\mathbf{P}(\mathcal{K})$ is clearly unbounded. Suppose then $\mathbf{P}(\mathcal{K})$ has an improving ray x_{ray}. Then, for any feasible point x_0 and facial reduction certificate s used by the primal regularization step,

$$0 = \langle s, x_0 + x_{ray} \rangle = \langle s, x \rangle_{ray}.$$

Hence, $x_{ray} \in \mathcal{C}'$ and is therefore an improving ray for $\mathbf{P}(\mathcal{C}')$, a contradiction.

For the converse direction, we will argue $\tau = \kappa = 0$ holds at each iteration unless the dual regularization step executes. Since the primal regularization step can execute only finitely many times (since $\mathcal{K}$ is finite-dimensional), the converse direction therefore follows. To begin, suppose the optimal value of $\mathbf{P}(\mathcal{K})$ is finite and unattained, i.e., suppose (4a) holds. Then $\tau = \kappa = 0$ for all $(x, s, y, \tau, \kappa) \in \text{relint } \mathbf{H}(\mathcal{K})$; otherwise, either an improving ray would exist, contradicting finiteness, or a complementary solution would exist, contradicting unattainment. Since the feasible sets of $\mathbf{P}(\mathcal{C}')$ and $\mathbf{P}(\mathcal{K})$ are equal unless the dual regularization step executes, repeating this argument shows $\tau = \kappa = 0$ unless the dual regularization step executes. A similar argument shows the claim assuming (4b).

Statement 3a. Strict feasibility of $\mathbf{P}(\mathcal{C}')$ was established in the proof of statement 4. This implies $\mathbf{P}(\mathcal{C}'')$ is strictly feasible as shown by

$$\mathcal{A}_p \cap \text{relint } \mathcal{C}' \subseteq \mathcal{A}_p \cap \text{relint}(\mathcal{C}' + \text{span } x) = \mathcal{A}_p \cap \text{relint}(\mathcal{C}'^* \cap x^\perp)^* = \mathcal{A}_p \cap \text{relint } \mathcal{C}''.$$

Statement 3b. By (3a), if dual regularization is performed, then $\mathbf{P}(\mathcal{C}')$ satisfies Slater's condition, and continues to satisfy Slater's condition at each ensuing iteration. Hence, a facial reduction certificate s cannot exist at any ensuing iteration by Proposition 1.

Statement 1. Since $\mathcal{K}$ is finite dimensional, it trivially follows, using (3b), that both regularization steps can execute only finitely many times. Hence, the algorithm must terminate. Corollary 1 implies termination in one iteration when complementary solutions or improving rays exist for $\mathbf{P}(\mathcal{K})$ and $\mathbf{D}(\mathcal{K})$.

Statement 2. The optimal values of $\mathbf{P}(\mathcal{C}')$ and $\mathbf{P}(\mathcal{C}'')$ are equal when the primal regularization step executes. Similarly, the optimal values of $\mathbf{D}(\mathcal{C}')$ and $\mathbf{D}(\mathcal{C}'')$ are equal when the dual regularization step executes. Moreover, (3a) and Slater's condition imply the optimal values of $\mathbf{P}(\mathcal{C}')$ and $\mathbf{D}(\mathcal{C}')$ are equal when the dual regularization step executes. Combining these facts with (3b) shows the optimal value of $\mathbf{P}(\mathcal{K})$ equals the optimal value of $\mathbf{P}(\mathcal{C})$ at every iteration.

Statement 5. When the primal regularization step executes, $\mathbf{P}(\mathcal{C}')$ and $\mathbf{P}(\mathcal{C}'')$ have equal feasible sets since s is a facial reduction certificate for $\mathbf{P}(\mathcal{C}')$. When the dual regularization step executes, the feasible set $\mathcal{A}_p \cap \mathcal{C}'$ of $\mathbf{P}(\mathcal{C}')$ and the feasible set $\mathcal{A}_p \cap \mathcal{C}''$ of $\mathbf{P}(\mathcal{C}'')$ satisfy

$$\mathcal{A}_p \cap \mathcal{C}' \subseteq \mathcal{A}_p \cap \overline{\mathcal{C}' + \operatorname{span} x} = \mathcal{A}_p \cap (\mathcal{C}'^* \cap x^\perp)^* = \mathcal{A}_p \cap \mathcal{C}''.$$

Combining these facts with (3b) shows $\mathcal{A}_p \cap \mathcal{K} \subseteq \mathcal{A}_p \cap \mathcal{C}$. Since dual regularization is performed if and only if (4a) or (4b) hold, the claim follows. $\qquad\square$

An illustration of the basic steps is given next on an infeasible problem with no improving ray.

Example 3 (Weak infeasibility). *Consider the following primal-dual pair, where $\mathcal{Q}_r^{2+k} = \{(r, x) \in \mathbb{R}_+^2 \times \mathbb{R}^k : 2r_1 r_2 \geq x^T x\}$ denotes a rotated Lorentz cone:*

$$
\begin{array}{ll}
\text{minimize} & 0 \\
\text{subject to} & x_1 = 0 \\
& x_3 = 1 \\
& x_4 = 1 \\
& x \in \mathcal{Q}_r^3 \times \mathbb{R}_+
\end{array}
\qquad
\begin{array}{ll}
\text{maximize} & y_2 + y_3 \\
\text{subject to} & \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} y_1 \\ 0 \\ y_2 \\ y_3 \end{pmatrix} = s \\
& s \in \mathcal{Q}_r^3 \times \mathbb{R}_+.
\end{array}
\qquad (10)
$$

Let $\mathcal{K} = \mathcal{Q}_r^3 \times \mathbb{R}_+$. *The primal problem is infeasible, but the dual problem has no improving ray.*

<u>*The first iteration.*</u> *At the first iteration, Algorithm 1 finds a point in* $\operatorname{relint} \mathbf{H}(\mathcal{K})$ *satisfying* $\tau = \kappa = 0$, *e.g.,* $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa})$, *where*

$$\hat{x} = (0, 1, 0, 0), \qquad \hat{s} = (1, 0, 0, 0), \qquad \hat{y} = (-1, 0, 0), \qquad \hat{\tau} = \hat{\kappa} = 0.$$

Since $\hat{s} \notin \mathcal{K}^\perp$, *Algorithm 1 regularizes the primal problem, reformulating it over* $\mathcal{K} \cap \hat{s}^\perp = \{0\} \times \mathbb{R}_+ \times \{0\} \times \mathbb{R}_+$. *This yields a new primal-dual pair* $\mathbf{P}(\mathcal{K} \cap \hat{s}^\perp)$ *and* $\mathbf{D}(\mathcal{K} \cap \hat{s}^\perp)$:

$$
\begin{array}{ll}
\text{minimize} & 0 \\
\text{subject to} & x_1 = 0 \\
& x_3 = 1 \\
& x_4 = 1 \\
& x \in \{0\} \times \mathbb{R}_+ \times \{0\} \times \mathbb{R}_+
\end{array}
\qquad
\begin{array}{ll}
\text{maximize} & y_2 + y_3 \\
\text{subject to} & \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} y_1 \\ 0 \\ y_2 \\ y_3 \end{pmatrix} = s \\
& s \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+.
\end{array}
$$

<u>*The second iteration.*</u> *At the next iteration, Algorithm 1 finds a point in* $\operatorname{relint} \mathbf{H}(\mathcal{K} \cap \hat{s}^\perp)$ *satisfying* $\kappa > 0$, *e.g.,* $(\tilde{x}, \tilde{s}, \tilde{y}, \tilde{\tau}, \tilde{\kappa})$, *where*

$$\tilde{x} = (0, 0, 0, 0), \quad \tilde{s} = (0, 0, -1, 0), \quad \tilde{y} = (0, 1, 0), \quad \tilde{\tau} = 0, \quad \tilde{\kappa} = 1.$$

Since $\kappa > 0$ the algorithm terminates. Note $(\tilde{s}, \tilde{y})$ is an improving ray for $\mathbf{D}(\mathcal{K} \cap \hat{s}^\perp)$ showing infeasibility of $\mathbf{P}(\mathcal{K} \cap \hat{s}^\perp)$ and hence of $\mathbf{P}(\mathcal{K})$ by Theorem 5, statement 5.

Theorem 5 states an iteration of dual regularization will be performed by Algorithm 1 iff the primal optimal value is finite but unattained or the primal problem is unbounded but has no improving ray. The next example illustrates this latter scenario.

Example 4 (An unbounded problem with no improving ray)**.** *Consider the following primal-dual pair also over a rotated Lorentz cone:*

$$
\begin{array}{ll}
\text{minimize} & x_3 \\
\text{subject to} & x_1 = 1 \\
& x \in \mathcal{Q}_r^3
\end{array}
\qquad
\begin{array}{ll}
\text{maximize} & y \\
\text{subject to} & \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} y \\ 0 \\ 0 \end{pmatrix} = s \\
& s \in \mathcal{Q}_r^3.
\end{array}
\qquad (11)
$$

The primal is unbounded but has no improving ray: if x and $x + d$ are feasible, then $d_1 = 0$; further, if $d \in \mathcal{Q}_r^3$, then $d_3 = 0$.

   The first iteration. Algorithm 1 finds a point in $\mathrm{relint}\,\mathbf{H}(\mathcal{Q}_r^3)$ *satisfying* $\tau = \kappa = 0$, *e.g.,* $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa})$, *where*

$$
\hat{x} = (0, 1, 0), \quad \hat{s} = (0, 0, 0), \quad \hat{y} = 0, \quad \hat{\tau} = \hat{\kappa} = 0.
$$

Since $\hat{s} \in (\mathcal{Q}_r^3)^\perp = \{0\}^3$, the dual problem is regularized by replacing $\mathcal{Q}_r^3$ with $\mathcal{Q}_r^3 \cap \hat{x}^\perp$:

$$
\begin{array}{ll}
\text{minimize} & x_3 \\
\text{subject to} & x_1 = 1 \\
& x \in \mathbb{R}_+ \times \mathbb{R}^2
\end{array}
\qquad
\begin{array}{ll}
\text{maximize} & y \\
\text{subject to} & \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} y \\ 0 \\ 0 \end{pmatrix} = s \\
& s \in \mathbb{R}_+ \times \{0\}^2.
\end{array}
\qquad (12)
$$

   The second iteration. A point in $\mathbf{H}(\mathbb{R}_+ \times \mathbb{R}^2)$ with $\kappa > 0$ is obtained, yielding a primal improving ray. Since $\mathbf{P}(\mathbb{R}_+ \times \mathbb{R}^2)$ is also feasible, we conclude it, and hence $\mathbf{P}(\mathcal{Q}_r^3)$, is unbounded.

4.2   Interpretation of outputs

Algorithm 1 returns a point $(x, s, y, \tau, \kappa) \in \mathrm{relint}\,\mathbf{H}(\mathcal{C})$ satisfying $\tau > 0$ or $\kappa > 0$, where the optimal value of the primal problem $\mathbf{P}(\mathcal{C})$ equals the optimal value of $\mathbf{P}(\mathcal{K})$, the problem of interest. By Corollary 1, we obtain a complementary solution or an improving ray for $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$ from the output. How, then, does one interpret this output in terms of $\mathbf{P}(\mathcal{K})$? We answer this with the following corollary of Theorem 5.

Corollary 3. *The following statements hold about the output (x, s, y, τ, κ) of Algorithm 1, where $\mathcal{A}_p := \{x \in \mathcal{V} : Ax = b\}$ and $\theta_p \in \mathbb{R} \cup \{\pm\infty\}$ denotes the optimal value of $\mathbf{P}(\mathcal{K})$, i.e.,*

$$
\theta_p = \inf\{\langle c, x \rangle : x \in \mathcal{A}_p \cap \mathcal{K}\}.
$$

1. *$\tau > 0$ holds if and only if θ_p is finite. Further, θ_p is finite and attained if and only if $\tau > 0$ and Algorithm 1 does not execute dual regularization steps.*

2. *$\kappa > 0$ holds if and only if $\mathbf{P}(\mathcal{K})$ is infeasible ($\theta_p = \infty$) or unbounded ($\theta_p = -\infty$).*

3. *If $\tau > 0$, then $\theta_p = \frac{1}{\tau}\langle c, x \rangle = \frac{1}{\tau}\langle b, y \rangle$. If θ_p is attained, then $\frac{1}{\tau}x$ is a solution of $\mathbf{P}(\mathcal{K})$.*

4. *If $\kappa > 0$ and $\langle b, y \rangle > 0$, then $\mathbf{P}(\mathcal{K})$ is infeasible ($\theta_p = \infty$).*

Note from this corollary, we can only conclude the optimal value of $\mathbf{P}(\mathcal{K})$ is not finite when $\kappa > 0$ and $\langle b, y \rangle \leq 0$; that is, in this situation, $\mathbf{P}(\mathcal{K})$ could be infeasible or it could be unbounded. Though it is tempting to assume $\langle b, y \rangle > 0$ whenever $\mathbf{P}(\mathcal{K})$ is infeasible—i.e., that (s, y) is a dual improving ray—Example 2 of Section 3.1 illustrated this is not necessarily the case. Hence, more must be done to distinguish unboundedness from infeasibility.

4.3 A complete algorithm

As just indicated, one cannot determine the optimal value of $\mathbf{P}(\mathcal{K})$ from the output (x, s, y, τ, κ) of Algorithm 1 if $\kappa > 0$ and the dual objective satisfies $\langle b, y \rangle \leq 0$; that is, $\mathbf{P}(\mathcal{K})$ could be unbounded or it could be infeasible. Fortunately, there is a simple remedy: reexecuting Algorithm 1 after setting the cost vector of $\mathbf{P}(\mathcal{K})$ to zero. That is, one can set $c = 0$ to obtain the following primal-dual pair:

$$
\begin{array}{ll}
\text{minimize} & 0 \\
\text{subject to} & Ax = b \\
& x \in \mathcal{K}
\end{array}
\qquad
\begin{array}{ll}
\text{maximize} & \langle b, y \rangle \\
\text{subject to} & 0 - A^*y = s \\
& s \in \mathcal{K}^*, \ y \in \mathbb{R}^m,
\end{array}
\tag{13}
$$

and then reexecute Algorithm 1. The returned solution $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa})$ satisfies $\hat{\kappa} > 0$ if and only if the primal problem of (13) is infeasible or unbounded (Corollary 3). However, as this problem cannot be unbounded (since $c = 0$), it follows $\hat{\kappa} > 0$ if and only if $\mathcal{K}$ contains no solution to $Ax = b$—i.e., if and only if the original problem $\mathbf{P}(\mathcal{K})$ is infeasible. On the other hand, if $\hat{\kappa} = 0$, then $\mathbf{P}(\mathcal{K})$ must have been unbounded.

Putting everything together, we get a complete algorithm for solving $\mathbf{P}(\mathcal{K})$:

Algorithm 2: Finds optimal value of $\mathbf{P}(\mathcal{K})$ and point attaining it if one exists.

Execute Algorithm 1 and let (x, s, y, τ, κ) denote the output
if $\tau > 0$ **or** ($\kappa > 0$ **and** $\langle b, y \rangle > 0$) **then**
   | **return** optimal value θ_p and point attaining θ_p (if one exists) using Corollary 3.
else
   Re-execute Algorithm 1 with $c = 0$ and let $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa})$ denote the output.
   return $\begin{cases} \infty & \text{if } \hat{\kappa} > 0 \text{ and } \hat{\tau} = 0; \ \mathbf{P}(\mathcal{K}) \text{ is infeasible.} \\ -\infty & \text{if } \hat{\kappa} = 0 \text{ and } \hat{\tau} > 0; \ \mathbf{P}(\mathcal{K}) \text{ is unbounded.} \end{cases}$
end

To conclude, we illustrate execution of Algorithm 2 with a simple example.

Example 5. *Consider the following primal-dual pair, where $\mathcal{Q}_r^{2+k} = \{(r, x) \in \mathbb{R}_+^2 \times \mathbb{R}^k \mid 2r_1 r_2 \geq x^T x\}$ denotes a rotated Lorentz cone:*

$$
\begin{array}{ll}
\text{minimize} & -x_2 \\
\text{subject to} & x_1 = 0 \\
& x_3 = 1 \\
& x_4 = 1 \\
& x \in \mathcal{Q}_r^3 \times \mathbb{R}_+
\end{array}
\qquad
\begin{array}{ll}
\text{maximize} & y_2 + y_3 \\
\text{subject to} & \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} y_1 \\ 0 \\ y_2 \\ y_3 \end{pmatrix} = s \\
& s \in \mathcal{Q}_r^3 \times \mathbb{R}_+.
\end{array}
\tag{14}
$$

Here, the primal and dual problem are both infeasible, but no dual improving ray exists. Indeed, all dual rays satisfy $s_2 = 0$ and $y_3 \leq 0$; since, in addition, $s_2 = 0$ implies $y_2 = 0$, we conclude $y_2 + y_3 > 0$ cannot hold. On the other hand, $x = (0, 1, 0, 0)$ is a primal improving ray.

We now illustrate execution of Algorithm 2. Since a primal improving ray exists and no dual improving ray exists, Algorithm 1 executes one iteration and returns (x, s, y, τ, κ) satisfying $\kappa > 0$ and $\langle b, y \rangle = 0$. Since $\langle b, y \rangle = 0$, the feasibility problem is constructed:

$$
\begin{array}{ll}
\text{minimize} & 0 \\
\text{subject to} & x_1 = 0 \\
& x_3 = 1 \\
& x_4 = 1 \\
& x \in \mathcal{Q}_r^3 \times \mathbb{R}_+
\end{array}
\qquad
\begin{array}{ll}
\text{maximize} & y_2 + y_3 \\
\text{subject to} & \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} y_1 \\ 0 \\ y_2 \\ y_3 \end{pmatrix} = s \\
& s \in \mathcal{Q}_r^3 \times \mathbb{R}_+,
\end{array}
$$

which matches the primal-dual pair of Example 3. Algorithm 1 reexecutes on this primal-dual pair (performing the steps illustrated in Example 3) and returns a point satisfying $\kappa > 0$. Since $\kappa > 0$, Algorithm 2 reports the primal of (14) *is infeasible.*

5 Numerical experiments

We conclude with numerical experiments. Precisely, we evaluate tracking of the central path (of the extended-embedding) to its limit point, which yields an element of relint $\mathbf{H}(\mathcal{K})$ by Theorem 4 and results of [13, 27]. We then verify the classification of $(x, s, y, \tau, \kappa) \in$ relint $\mathbf{H}(\mathcal{K})$ given by Corollary 2: if $\tau = \kappa = 0$ and the primal (resp. dual) fails Slater's condition, then s (resp. x) is a facial reduction certificate; we'll also confirm there certificates are optimal on a subset of examples. On the negative side, we show regularizing without significantly changing optimal values (a key step of an Algorithm 1 implementation) can be difficult. We also illustrate tracking the central path of the extended-embedding is hard for semidefinite programs described in Waki et al. [34]; we attribute this to high singularity degree and significant failure of strict complementarity. Finally, an approximate certificate is used to (informally) argue an instance of the DIMACS library [23] has an unattained optimal value. This illustrates approximate facial reduction certificates can provide useful insight, even if they cannot be used reliably for regularization.

To perform experiments, we have made trivial modifications to the interface of SeDuMi [30], which already tracks the central path of the extended-embedding (though without maintaining the slack variable s). We modify the interface to return a tuple (x, y, τ, κ); the point $(x, c\tau - A^*y, y, \tau, \kappa)$ then lies in relint $\mathbf{H}(\mathcal{K})$ to within tolerances achieved by SeDuMi. Each test case considered is a semidefinite program (SDP), i.e., $\mathcal{K}$ equals $\mathbb{S}_+^n$, the cone of psd matrices of order n. Finally, in reported results, $\|\cdot\|$ denotes the Frobenius norm when the argument is a matrix and the 2-norm otherwise; $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue of its argument.

5.1 Approximate facial reduction certificates

Weak infeasibility library The first test cases are SDPs taken from the URL documented in [14]. For each instance, c is strictly feasible for the dual and the primal is weakly infeasible (i.e., infeasible but with no dual improving ray). It follows $\tau = \kappa = 0$ for $(x, s, y, \tau, \kappa) \in$ relint $\mathbf{H}(\mathcal{K})$. Further, Corollary 2 implies s is facial reduction certificate and that $x = 0$. The tables below confirm that these implications hold.

set of instances	τ	κ	$\lambda_{\min}(s)$	$|\langle b, y \rangle|$	$\|s\|$	$\|x\|$
weak_clean_10_10	4.8e$-$9	2.0e$-$8	7.3e$-$11	3.0e$-$7	5.6	1.9e$-$7
weak_messy_10_10	7.5e$-$8	6.7e$-$8	7.2e$-$10	1.1e$-$6	7.6	8.6e$-$7
weak_clean_20_10	1.3e$-$9	1.6e$-$8	1.9e$-$11	2.6e$-$7	5.3	1.7e$-$7
weak_messy_20_10	7.0e$-$8	4.0e$-$7	7.9e$-$10	4.5e$-$6	8.8	3.6e$-$6

(a) means

set of instances	τ	κ	$\lambda_{\min}(s)$	$|\langle b, y \rangle|$	$\|s\|$	$\|x\|$
weak_clean_10_10	4.4e$-$9	2.0e$-$8	9.1e$-$11	2.5e$-$7	3.7	1.6e$-$7
weak_messy_10_10	8.0e$-$8	7.0e$-$8	7.6e$-$10	1.3e$-$6	4.6	1.1e$-$6
weak_clean_20_10	1.8e$-$9	1.8e$-$8	3.1e$-$11	2.8e$-$7	4.9	1.9e$-$7
weak_messy_20_10	8.4e$-$8	4.5e$-$7	8.8e$-$10	3.7e$-$6	4.8	3.0e$-$6

(b) standard deviations

Note in these tables, we report the mean and standard deviation taken over each set of instances, which each contain 100 SDPs.

Finite but non-zero duality gaps Using Algorithm 12.3 of [6], we generated instances with duality gap equal to 100. The other inputs to this algorithm are the number of equations m, the order of the matrices n, the rank r_1 of an optimal facial reduction certificate for the dual problem, and an additional parameter p specifying structure of the constraint matrices. If $(x, s, y, \tau, \kappa) \in$ relint $\mathbf{H}(\mathcal{K})$, then $\tau = \kappa = 0$ necessarily holds given the duality gap. Further, by Corollary 2, both x and s are optimal facial reduction certificates, since both primal and dual problems fail Slater's condition. Results below show (approximate) facial reduction certificates are indeed found:

(n, m, p, r_1)	τ	κ	$\lambda_{\min}(s)$	$|\langle b, y \rangle|$	$\|s\|$	$\lambda_{\min}(x)$	$|\langle c, x \rangle|$	$\|Ax\|$	$\|x\|$
$(10, 10, 5, 5)$	1.2e−6	2.3e−6	−8.1e−14	4.4e−6	5.8e2	−8.6e−14	2.2e−6	3.4e−4	1.7e3
$(20, 20, 10, 5)$	1.2e−6	1.3e−6	−2.6e−14	2.0e−6	7.8e2	−2.2e−13	6.9e−7	4.6e−4	1.3e3
$(40, 40, 10, 5)$	5.9e−7	5.2e−8	−4.0e−15	1.0e−7	6.3e1	−1.1e−13	4.9e−8	3.5e−4	8.0e2
$(40, 40, 10, 10)$	8.9e−7	3.1e−7	−7.4e−15	5.0e−7	3.0e2	−9.0e−14	1.9e−7	6.5e−4	1.0e3

Estimating rank as the number of eigenvalues larger than 1e−4 yields the following table, indicating s and x are (approximate) optimal certificates, i.e., rank $s = n - r_1 - 1$ and rank $x = r_1$:

(n, m, p, r_1)	rank s	rank x
$(10, 10, 5, 5)$	4	5
$(20, 20, 10, 5)$	14	5
$(40, 40, 10, 5)$	34	5
$(40, 40, 10, 10)$	29	10

Also note strict-complementarity fails modestly for these instances: rank s + rank $x = n - 1$. Section 5.3 will illustrate large numerical error on instances failing this condition more severely.

5.2 Regularization using approximate certificates

Given an approximate point $(x, s, y, \tau, \kappa) \in$ relint $\mathbf{H}(\mathcal{K})$, an implementation of Algorithm 1 must decide if $\tau > 0$, $\kappa > 0$ or $\tau = \kappa = 0$ for an actual point in relint $\mathbf{H}(\mathcal{K})$. For the case of SDP $(\mathcal{K} = \mathbb{S}_+^n)$, it also must estimate the kernel of an actual facial reduction certificate; indeed, for $s \in \mathcal{K}$, the face $\mathcal{K} \cap s^\perp$ equals the set of psd matrices with range contained in the kernel of s by [3].

Example 6 below shows the computed optimal value can be sensitive to errors in these estimates. In this example, we estimate τ and κ are actually zero if they're less than a threshold $T_{\tau, \kappa}$ and estimate eigenvalues of s are actually zero if they're less than a threshold T_λ. Results for different thresholds $T_{\tau, \kappa}$ and T_λ follow:

thresholds	computed optimal value	
$T_{\tau, \kappa} = 5\mathrm{e}{-5}$	3.148e−1	
$T_{\tau, \kappa} = 5\mathrm{e}{-9}, T_\lambda = 5\mathrm{e}{-4}$	1.0000	(agrees with actual value)
$T_{\tau, \kappa} = 5\mathrm{e}{-9}, T_\lambda = 5\mathrm{e}{-6}$	∞	(i.e., regularized problem is infeasible)

The mentioned example used to generate this table is now given.

Example 6. *We consider an SDP with optimal value 1 and a finite non-zero duality gap, given by*

$$\begin{array}{ll}
\text{minimize} & \langle c, x\rangle \\
\text{subject to} & \langle a_1, x\rangle = 1 \\
& \langle a_2, x\rangle = 0 \\
& x \in \mathbb{S}_+^3,
\end{array} \qquad (15)$$

where the data matrices c, a_1 and a_2 are defined in terms of an orthogonal matrix $q \in \mathbb{R}^{3\times 3}$ via:

$$q := \begin{bmatrix} 7/11 & 6/11 & 6/11 \\ 6/11 & -9/11 & 2/11 \\ -6/11 & -2/11 & 9/11 \end{bmatrix} \quad c := q \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} q^T, \ a_1 := q \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} q^T, \ a_2 := q \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} q^T.$$

The solver SeDuMi finds an approximate point in $\mathrm{relint}\, \mathbf{H}(\mathbb{S}_+^3)$, *given by* $\tau = 7.945\mathrm{e}{-5}$, $\kappa = 1.028\mathrm{e}{-5}$,

$$s = \begin{bmatrix} 1.995 & -2.993 & -6.652\mathrm{e}{-1} \\ -2.993 & 4.490 & 9.977\mathrm{e}{-1} \\ -6.652\mathrm{e}{-1} & 9.977\mathrm{e}{-1} & 2.217\mathrm{e}{-1} \end{bmatrix}, \quad x = \begin{bmatrix} 1.518 & 5.060\mathrm{e}{-1} & 2.277 \\ 5.060\mathrm{e}{-1} & 1.687\mathrm{e}{-1} & 7.590\mathrm{e}{-1} \\ 2.277 & 7.590\mathrm{e}{-1} & 3.416 \end{bmatrix}.$$

Picking a threshold $T_{\tau,\kappa} = 5\mathrm{e}{-5}$, we estimate $\tau > 0$ and $\kappa = 0$ and compute the optimal value:

$$\frac{1}{\tau}\langle c, x\rangle = 3.148\mathrm{e}{-1},$$

which significantly differs from the actual optimal value of (15).

On the other hand, taking $T_{\tau,\kappa} = 5\mathrm{e}{-9}$, we estimate $\tau = \kappa = 0$ and interpret s and x as facial reduction certificates. To regularize, we compute an eigenvalue decomposition $\sum_{i=1}^{3} \lambda_i v_i v_i^T$ of s:

$$\lambda = \begin{bmatrix} 4.416\mathrm{e}{-5} \\ -1.857\mathrm{e}{-10} \\ 6.707 \end{bmatrix}, \quad v_1 = \begin{bmatrix} 6.364\mathrm{e}{-1} \\ 5.455\mathrm{e}{-1} \\ -5.455\mathrm{e}{-1} \end{bmatrix} \ v_2 = \begin{bmatrix} 5.455\mathrm{e}{-1} \\ 1.818\mathrm{e}{-1} \\ 8.182\mathrm{e}{-1} \end{bmatrix} \ v_3 = \begin{bmatrix} -5.455\mathrm{e}{-1} \\ 8.182\mathrm{e}{-1} \\ 1.818\mathrm{e}{-1} \end{bmatrix}.$$

We then estimate a basis for the kernel of s by interpreting eigenvalues below a threshold T_λ as zero. For $T_\lambda = 5\mathrm{e}{-4}$, an estimated basis is $u = (v_1, v_2)$, leading to the regularized SDP:

$$\begin{array}{ll}
\text{minimize} & \langle c, u x u^T\rangle \\
\text{subject to} & \langle a_1, u x u^T\rangle = 1 \\
& \langle a_2, u x u^T\rangle = 0 \\
& x \in \mathbb{S}_+^2.
\end{array}$$

Solving with SeDuMi, we compute an optimal value of 1.0000, which agrees with the actual optimal value of (15). On the other hand, taking $T_\lambda = 5\mathrm{e}{-6}$ yields $u = (v_2)$ and the SDP

$$\begin{array}{ll}
\text{minimize} & \langle c, u x u^T\rangle \\
\text{subject to} & \langle a_1, u x u^T\rangle = 1 \\
& \langle a_2, u x u^T\rangle = 0 \\
& x \in \mathbb{S}_+^1.
\end{array}$$

Solving with SeDuMi, we obtain a dual improving ray; hence, we incorrectly conclude (15) is infeasible.

5.3 Error, singularity degree and strict complementarity

For the instances of Section 5.1, we were able to obtain 'good' approximations of facial reduction certificates. The next set of instances, taken from [34], illustrate this is not always the case. For the reported instances, the dual optimal value is finite but unattained. Hence, $\tau = \kappa = 0$ holds for all points in $\mathbf{H}(\mathbb{S}^n_+)$. As reported in Table 4, significant error in τ is observed on several instances:

Instance	τ	κ	k_p	n	rank x	rank s
unboundDim1R2	1.2e−5	6.0e−9	3	7	0	2
unboundDim1R3	4.5e−7	3.4e−9	5	10	0	2
unboundDim1R4	4.3e−7	4.5e−9	7	13	0	2
unboundDim1R5	1.1†	1.8e−9	9	16	0	2
unboundDim1R6	9.8e−1†	2.5e−9	11	19	0	2
unboundDim1R7	7.9e−1†	5.2e−9	13	22	0	2
unboundDim1R8	7.2e−1†	3.6e−9	15	25	0	2
unboundDim1R9	6.9e−1†	3.6e−9	17	28	0	2
unboundDim1R10	6.6e−1†	9.2e−10	19	31	0	2

Table 4: k_p denotes singularity degree of the primal; n the order of the semidefinite constraint; rank x and rank s are actual values for any $(x, s, y, \tau, \kappa) \in \text{relint } \mathbf{H}(\mathbb{S}^n_+)$. Large error is marked †.

We offer two explanations for this error based on this table. The first is the singularity degree k_p (see Section 2) of the primal problem (reported in [33]) is high. Note when singularity degree is high, distance to feasibility (forward error) can be large even when residuals (backwards errors) are small [31]. The other (related) explanation is the extent to which strict complementarity fails—that is, the extent to which rank x + rank s is less than n. For $(x, s, y, \tau, \kappa) \in \text{relint } \mathbf{H}(\mathbb{S}^n_+)$, we see rank x + rank $s = 2 < n$ for each instance. Indeed, x is the zero matrix by Corollary 2 and strict feasibility of the dual problem [34, Section 2] and the matrix s has rank two by [33, Section 5.2].

We remark singularity degree can be high and strict complementarity can fail even if $\mathbf{P}(\mathbb{S}^n_+)$ and $\mathbf{D}(\mathbb{S}^n_+)$ have complementary solutions or improving rays. In other words, the suspected causes of error can also occur when $\tau > 0$ or $\kappa > 0$ holds in the relative interior of $\mathbf{H}(\mathbb{S}^n_+)$.

5.4 Difficult instances from the DIMACS library

Mittelmann [18] makes the following remark about two instances of the DIMACS SDP library [23]:

> 'In the case of the `hinf12` and `hinf13` instances the results obtained by the various codes are so different, that we cannot be sure whether these problems have in fact ever been solved.'

We'll compile evidence the dual optimal value of `hinf12` is finite and unattained, explaining the difficulty of solving this instance. To make a more convincing argument, we obtain approximate points in relint $\mathbf{H}(\mathcal{K})$ using both the self-dual solver of SDPT3v4.0 [32] and SeDuMi. Properties of these points are summarized below:

Instance (solver)	τ	κ	$\lambda_{\min}(s)$	$|\langle b, y \rangle|$	$\|s\|$	$\|x\|$
`hinf12` (SeDuMi)	1.5e−6	2.9e−8	4.1e−9	−1.5e−7	5.5e2	1.2e−6
`hinf12` (SDPT3)	4.4e−11	1.4e−15	2.0e−15	−2.0e−15	2.5e1	3.6e−11

These approximate points suggest the following: by Corollary 1, no complementary solution or improving ray exists (since $\tau \approx 0$ and $\kappa \approx 0$); by Corollary 2, the primal fails Slater's condition

(since $\|s\| \neq 0$); by Corollary 2, the dual satisfies Slater's condition (since $\|x\| \approx 0$). Hence, these approximate points suggest either the primal is weakly infeasible or the dual optimal value is finite and unattained. But one can verify that $e_6 e_6^T \in \mathbb{S}_+^{24}$ is primal feasible (where e_i is a standard basis vector of $\mathbb{R}^{24}$). Hence, we suspect the dual optimal value is finite and unattained.

6 Conclusions

We've unified the facial reduction algorithm of Borwein and Wolkowicz with the self-dual embedding of Goldman and Tucker, bringing together both techniques to, in principle, solve arbitrary conic optimization problems. Implementing a suggested algorithm involves only conceptually-simple modifications to solvers that track the central path of extended-embeddings (such as SeDuMi [30]), and these modifications only affect solver execution when both complementary solutions and improving rays do not exist. Nevertheless, numerical experiments illustrate significant practical barriers: in practice, one obtains only approximate facial reduction certificates; optimal values can be infinitely-sensitive to inexact regularization; and tracking the central-path can be difficult due to loss of strict-complementarity. Further work is needed to better understand these barriers. We also used approximate facial reduction certificates to make an informed conjecture about an SDP instance from [23]. Clarifying the usefulness of such certificates is another topic for future research; we note approximate certificates are used in [6] with backwards-stability guarantees.

7 Acknowledgements

References

[1] R. A. Abrams. Projections of convex programs with unattained infima. *SIAM Journal on Control*, 13(3):706–718, 1975.

[2] E. Andersen, C. Roos, and T. Terlaky. Notes on Duality in Second Order and p-Order Cone Optimization. *Optimization*, 51(4):627–643, 2002.

[3] G. P. Barker and D. Carlson. Cones of diagonally dominant matrices. *Pacific Journal of Mathematics*, 57(1):15–32, 1975.

[4] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar. *Convex analysis and optimization*. Athena Scientific Belmont, 2003.

[5] J. Borwein and H. Wolkowicz. Regularizing the abstract convex program. *Journal of Mathematical Analysis and Applications*, 83(2):495–530, 1981.

[6] Y.-L. Cheung, S. Schurr, and H. Wolkowicz. Preprocessing and regularization for degenerate semidefinite programs. In *Computational and Analytical Mathematics*, pages 251–303. Springer, 2013.

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

[8] E. de Klerk, T. Terlaky, and K. Roos. *Self-Dual Embeddings*, pages 111–138. Springer US, Boston, MA, 2000. ISBN 978-1-4615-4381-7. doi: 10.1007/978-1-4615-4381-7_5. URL http://dx.doi.org/10.1007/978-1-4615-4381-7_5.

[9] M. Dür, B. Jargalsaikhan, and G. Still. The Slater condition is generic in linear conic programming. pages 1–15, 2012. URL http://www.optimization-online.org/DB_FILE/2012/11/3675.pdf.

[10] R. M. Freund. On the behavior of the homogeneous self-dual model for conic convex optimization. *Mathematical programming*, 106(3):527–545, 2006.

[11] H. A. Friberg. Facial reduction heuristics and the motivational example of mixed-integer conic optimization. *Optimization Online*, 2016. URL http://www.optimization-online.org/DB_HTML/2016/02/5324.html.

[12] A. J. Goldman and A. W. Tucker. Theory of linear programming. *Linear inequalities and related systems*, 38:53–97, 1956.

[13] M. Halická, E. de Klerk, and C. Roos. On the convergence of the central path in semidefinite optimization. *SIAM Journal on Optimization*, 12(4):1090–1099, 2002.

[14] M. Liu and G. Pataki. Exact duals and short certificates of infeasibility and weak infeasibility in conic linear programming. 2015.

[15] B. F. Lourenço, M. Muramatsu, and T. Tsuchiya. Solving SDP completely with an interior point oracle. 2015.

[16] Z.-Q. Luo, J. F. Sturm, and S. Zhang. Duality and self-duality for conic convex programming. Technical Report 9620/A, Econometric Institute, Erasmus University Rotterdam, 1996.

[17] Z.-Q. Luo, J. F. Sturm, and S. Zhang. Duality Results for Conic Convex Programming. Technical report, Erasmus University Rotterdam, Erasmus School of Economics (ESE), Econometric Institute, 1997.

[18] H. D. Mittelmann. An independent benchmarking of SDP and SOCP solvers. *Mathematical Programming*, 95(2):407–430, 2003.

[19] Mosek APS. The MOSEK optimization software. Online at http://www.mosek.com.

[20] Y. Nesterov. Infeasible start interior-point primal-dual methods in nonlinear programming. CORE Discussion Papers 1995067, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE), 1995.

[21] B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd. Operator splitting for conic optimization via homogeneous self-dual embedding. *arXiv preprint arXiv:1312.3039*, 2013.

[22] G. Pataki. Strong duality in conic linear programming: facial reduction and extended duals. *Computational and Analytical Mathematics*, pages 613–634, 2013.

[23] G. Pataki and S. Schmieta. The DIMACS library of semidefinite-quadratic-linear programs. Available at http://dimacs.rutgers.edu/Challenges/Seventh/Instances, 1999.

[24] F. Permenter and P. A. Parrilo. Basis selection for SOS programs via facial reduction and polyhedral approximations. In *Proceedings of the IEEE Conference on Decision and Control*, 2014.

[25] F. Permenter and P. A. Parrilo. Partial facial reduction: simplified, equivalent SDPs via inner approximations of the PSD cone. http://arxiv.org/abs/1408.4685, 2014.

[26] F. A. Potra and R. Sheng. On homogeneous interior-point algorithms for semidefinite programming. *Optimization Methods and Software*, 9(1-3):161–184, 1998.

[27] H. Ramírez and D. Sossa. On the central paths in symmetric cone programming. *Journal of Optimization Theory and Applications*, pages 1–20, 2016.

[28] J. Renegar. Incorporating Condition Measures into the Complexity Theory of Linear Programming. *SIAM Journal on Optimization*, 5(3):506–524, 1995. ISSN 1052-6234. doi: 10.1137/0805026.

[29] R. T. Rockafellar. *Convex analysis*, volume 28. Princeton University Press, 1997.

[30] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization methods and software*, 11(1-4):625–653, 1999.

[31] J. F. Sturm. Error bounds for linear matrix inequalities. *SIAM Journal on Optimization*, 10(4):1228–1248, 2000.

[32] K.-C. Toh, M. J. Todd, and R. Tütüncü. SDPT3 version 4.0: a MATLAB software for semidefinite-quadratic-linear programming. *http://www.math.nus.edu.sg/ mattohkc/sdpt3.html*, 2009.

[33] H. Waki and M. Muramatsu. Facial reduction algorithms for conic optimization problems. *Journal of Optimization Theory and Applications*, pages 188–215, 2013.

[34] H. Waki, M. Nakata, and M. Muramatsu. Strange behaviors of interior-point methods for solving semidefinite programming problems in polynomial optimization. *Computational Optimization and Applications*, 53(3):823–844, 2012.

[35] Y. Ye, M. J. Todd, and S. Mizuno. An $\mathcal{O}(\sqrt{nL})$-iteration homogeneous and self-dual linear programming algorithm. *Mathematics of Operations Research*, 19(1):53–67, 1994.