Facial reduction heuristics and the motivational example of mixed-integer conic optimization

Henrik A. Friberg *

February 24, 2016

Abstract

Facial reduction heuristics are developed in the interest of added performance and reliability in methods for mixed-integer conic optimization. Specifically, the process of branch-and-bound is shown to spawn subproblems for which the conic relaxations are difficult to solve, and the objective bounds of linear relaxations are arbitrarily weak. While facial reduction algorithms already exist to deal with these issues, heuristic variants represent a very potent supplement due to their inherent speed and accuracy. The paper covers a family of heuristics based on linear optimization, subgradient matching, single-cone analysis, and cone factorization.

1 Introduction

For $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$, consider the following primal-dual pair of conic optimization problems over the nonempty, closed, convex cone $\mathcal{K} \subseteq \mathbb{R}^n$ and its dual cone $\mathcal{K}^* \subseteq \mathbb{R}^n$:

$$(P): \quad \theta_P = \inf_x \{c^T x : Ax = b,\, x \in \mathcal{K}\}, \qquad (D): \quad \theta_D = \sup_{s,y} \{b^T y : c - A^T y = s,\, s \in \mathcal{K}^*\}. \qquad (1)$$

By careful construction, the primal-dual pair can always be formulated such that the problem of interest, (P) or (D), is either *strongly feasible*[1] or *strongly infeasible*[2] [29, 41]. These properties serve in the interest of a successful solve in both theory and practice (see, e.g., [40, 20, 42]), and problems not satisfying either are denoted *ill-posed* following Renegar [35, 16].

Careful construction is unfortunately a false premise in many cases, such as when problem formulations are the product of preprocessing, cut generation or other automated modifications. As a result, ill-posed formulations may occur naturally in some of these cases (albeit perhaps rarely), and optimization software needs robust countermeasures for dealing with them to function reliably. This is the first reason for looking at the example of mixed-integer conic optimization, namely because it serves as a particularly great motivation for these countermeasures. Specifically, any solver would potentially fail from time to time if neglecting them, due to the ill-posed conic relaxations constructed by its own branch-and-bound procedure, as elaborated in Section 2.

Facial reduction algorithms [8, 32, 41] automate the process of careful construction by iteratively

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

[1] Existence of a feasible point in the relative interior of all non-polyhedral cones (i.e., the generalized Slater's condition) qualifying use of the usual KKT-type optimality conditions [9].

[2] Existence of a dual improving ray qualifying use of the usual Farkas-type theorem of alternatives [29]; if this property is not satisfied for an infeasible problem, the infimum distance to feasibility is zero!

applying reformulations until either strong feasibility or strong infeasibility is satisfied. At every iteration until this happens, there exist so-called facial reduction certificates to verify and drive the process of these reformulations [41]. At termination, the problem is said to be *regularized*. Even so, the algorithm can still be continued in dual space as one way of dealing with unattainment [34, 28]. In any case, and for any of this to work in practice, core decisions have to be made regarding the generation of the facial reduction certificates.

First of all, the immediate conic formulation of the feasible set of certificates may itself be ill-posed and hence problematic to solve. Alternatives are given in [13] and [28], where the feasible set is lifted (for strong feasibility) by a nonnegative artificial variable which is then minimized. More recently, [34] shows that the self-dual embedding of (1) also contains the feasible set of certificates. Specifically, these certificates make up the optimal set if and only if (1) has no primal-dual optimal solutions or improving rays. Regularization and optimization can thus be interleaved with expected computational advantages (yet to be verified in practice), and the appealing property that facial reduction-induced reformulation is used only as needed.

This leads to the second point. If the certificates are generated by solving the conic optimization problems described above (from [13], [28] or [34]) using floating-point-based algorithms, the applied reformulations are only approximately valid. In fact, given the property of backwards stability shown in [13], this corresponds to reformulating a perturbed primal-dual pair (1). Given that there are many sources of inaccuracy in conic optimization software (see, e.g., [45, 4]), it may occur that the perturbation becomes too big for a continuation of the facial reduction algorithm to make sense. The very process of solving conic optimization problems—the size of the original problem—in each iteration of the facial reduction algorithm can, of course, also be very time consuming.

These concerns may to some extent be resolved by *facial reduction heuristics* attempting to construct accurate facial reduction certificates (or even exact, in rational arithmetic) within a short period of time. Specifically, with the right choice of heuristics (e.g., inspired by the application-specific reduction techniques of [14, 27, 11, 5, 47, 43]), heuristics alone may turn out to be sufficient in many cases. Partial regularization can also be useful, however, as shown by the example of mixed-integer conic optimization. Specifically, as elaborated in Section 2, objective bounds computed from linear relaxations can be arbitrarily weak if the conic relaxation is ill-posed, and only rapid regularization techniques make sense to include in such simple bound computations.

This paper introduces the relevance of facial reduction in mixed-integer conic optimization, and extends upon previous work on facial reduction heuristics. A family of heuristics based on *linear optimization*, inspired by [33], is presented in Section 4, along with strengthened versions of heuristics from [21, 13], denoted *subgradient matching* and *cone factorization*. This section also introduces a new heuristic based on *single-cone analysis*, inspired by the ability to solve single second-order cone problems analytically [19].

2 Motivational example: Mixed-integer conic optimization

Presolve, cut generation and branching are standard elements of mixed-integer optimization [1], and all act to modify either the feasible set of a problem or its formulation. This poses a risk for the procedure, as the conic relaxations solved in the nodes of the branch-and-bound search tree may become ill-posed at any time during the branch-and-bound algorithm.

This sudden occurrence of ill-posedness is illustrated to occur in [23] for the conic representation of the conditional constraint, $x \geq 4$ if $z = 0$, where many solvers are shown to give the wrong answer or produce errors. The conic relaxations in the search tree for this example are also unattained, however, motivating the following somewhat "cleaner" example where primal and dual optimal

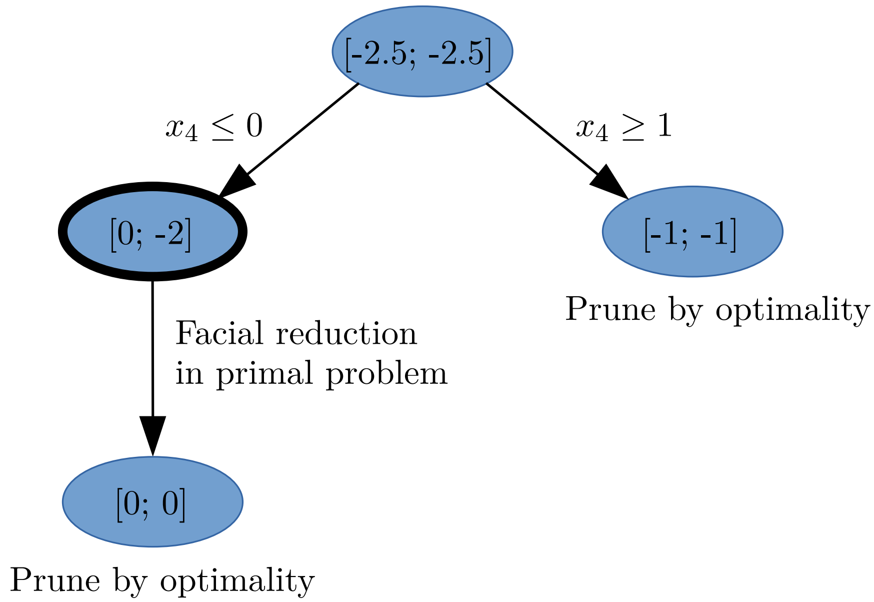

Figure 1: The branch-and-bound search tree for the problem of Example 1. Each node shows the range from the primal optimal value of the continuous relaxation to its dual optimal value.

values are attained throughout the search tree. The paper returns to analyze the conditional constraint of [23] in Section 4.3.

Example 1. *Consider the mixed-integer second-order cone optimization problem,*

$$\begin{array}{rcll} \theta_{MIP} & = & \inf\limits_{x} & 2x_3 + 2x_4 - x_5 \\ & & \text{s.t.} & x_1 + x_2 - x_4 \leq 0, \\ & & & 4x_4 - x_5 \geq 0, \\ & & & x_3 \geq -1, \\ & & & x_5 \leq 1, \\ & & & x \in \mathcal{Q}^3 \times \mathbb{R}^2_+ \\ & & & x_4 \in \mathbb{Z}, \end{array} \tag{2}$$

where $\mathcal{Q}^n := \{x \in \mathbb{R}^n : x_1 \geq \|x_{2:n}\|_2\}$ *is the quadratic cone. Solving this problem gives* $\theta_{MIP} = -1$ *with the branch-and-bound search tree of Figure 1. The solutions to the node relaxations of this search tree can be found analytically as follows.*

root node *Assume the relaxation can be solved without the constraint* $x_1 + x_2 - x_4 \leq 0$*. Omitting it, two separate subproblems are obtained which can be solved by inspection, namely*

$$\begin{array}{ll} \inf\limits_{x_1,x_2,x_3} & 2x_3 \\ \text{s.t.} & x_3 \geq -1, \\ & (x_1,x_2,x_3) \in \mathcal{Q}^3, \end{array} \quad \text{and} \quad \begin{array}{ll} \inf\limits_{x_4,x_5} & 2x_4 - x_5 \\ \text{s.t.} & 4x_4 - x_5 \geq 0, \\ & x_5 \leq 1, \\ & x_4, x_5 \geq 0. \end{array}$$

Solutions are, respectively, $(x_1, x_2, x_3) = (3, -2.75, -1)$ *and* $(x_4, x_5) = (0.25, 1)$*. Given that the omitted constraint is not violated by these values, the assumption was correct, and an optimal solution of the root node relaxation is given by* $x = (x_1, x_2, x_3, x_4, x_5)$ *with value* -2.5*. This is also the dual optimal value, since strong feasibility is certified by the identified solution, noting* $(x_1, x_2, x_3) \in \text{relint}\, \mathcal{Q}^3$*, and implies zero duality gap [29].*

$[x_4 \geq 1]$**-node** *The same analysis as before applies with $(x_4, x_5) = (1, 1)$ being the new optimal solution of the second subproblem. Hence, since the omitted constraint is not violated, an optimal solution is given by $x = (3, -2.75, -1, 1, 1)$ with value -1, and this is also the dual optimal value.*

$[x_4 \leq 0]$**-node** *The till now omitted constraint can no longer be satisfied by optimal solutions of the separate subproblems. Instead, ocular inspection of (2) reveals the fixations $x_4 = x_5 = 0$, whereby this node relaxation reduces to the analytically solvable primal-dual pair,*

$$\begin{array}{rclll} \theta_P & = & \inf\limits_{x_1,x_2,x_3} & 2x_3 \\ & & \text{s.t.} & x_1 + x_2 \leq 0, \\ & & & x_3 \geq -1, \\ & & & (x_1, x_2, x_3) \in \mathcal{Q}^3, \end{array} \qquad \begin{array}{rclll} \theta_D & = & \sup\limits_{s,y} & -y_2 \\ & & \text{s.t.} & y_1 = -s_1, \\ & & & y_1 = -s_2, \\ & & & y_2 - 2 = -s_3, \\ & & & y_1 \leq 0,\ y_2 \geq 0,\ s \in \mathcal{Q}^3. \end{array} \tag{3}$$

Given $x_1 + x_2 \leq 0$, let $x_2 = -x_1 - t$ for some $t \geq 0$ such that the conic constraint $(x_1, x_2, x_3) \in \mathcal{Q}^3$ becomes $x_1^2 \geq (-x_1 - t)^2 + x_3^2$ for $x_1 \geq 0$. This shows that $t = 0$ and $x_3 = 0$. Hence, the primal optimal value is $\theta_P = 0$ as attained, e.g., by $(x_1, x_2, x_3) = (0, 0, 0)$. Similarly, the conic constraint $s \in \mathcal{Q}^3$ becomes $y_1^2 \geq y_1^2 + (y_2 - 2)^2$ from which $y_2 = 2$ follows. The dual optimal value is hence $\theta_D = -2$ as attained, e.g., by $s = (0, 0, 0)$ and $y = (0, 2)$.

The $[x_4 \leq 0]$-node relaxation has positive duality gap and is therefore ill-posed. This ill-posedness is single-handedly caused by $x_1 + x_2 \leq 0$, as shown by $x_1 + x_2 \geq 0$ being a supporting halfspace [38] of $(x_1, x_2, x_3) \in \mathcal{Q}^3$ (compare to the unsatisfied definition of strong feasibility, requiring a feasible relative interior point of the cone). Hence, the conic relaxation is problematic to optimize without regularization and the method of facial reduction is now exemplified.

$[x_4 \leq 0]$**-node using facial reduction** *The previous analysis of this node concluded $x_4 = x_5 = 0$, as well as $x_1 + x_2 = t = 0$ and $x_3 = 0$ for the primal-dual pair (3). We can regularize the primal problem of (3) using either one of the latter two equations, albeit only the first one corresponds to a facial reduction. Indeed, taking $x_1 + x_2 = 0$ as example, it is possible to reformulate the conic constraint $\left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right) \in \mathcal{Q}^3$ as*

$$\left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right) \in \mathcal{Q}^3 \cap \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right)^{\perp}, \tag{4}$$

because $\left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right) \in \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right)^{\perp}$ is just another way of writing $\left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right)^T \left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right) = x_1 + x_2 = 0$. In turn, given $\mathcal{Q}^3 \cap \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right)^{\perp} = \left(\begin{smallmatrix} 1 \\ -1 \\ 0 \end{smallmatrix}\right) \mathbb{R}_+$ (see [17]), this reformulation actually changes the conic constraint into the linear constraints $x_1 \geq 0$, $x_2 = -x_1$ and $x_3 = 0$. That is, the regularized conic relaxation becomes a linear relaxation for which strong duality always holds, in this case at value $\theta_P = 0$.

Example 1 shows that branching operations can cause ill-posed relaxations to occur in the search tree of a branch-and-bound algorithm. In particular, by branching, the property of strong feasibility was lost when the feasible set was reduced in a way that kept the relaxation feasible without intersecting the relative interior of non-polyhedral cones. This specific type of reduction could as well have been caused by presolving or cut generation. Moreover, it is not even necessary for the non-polyhedral cone in question to be part of the original problem formulation, as it could be added later on as a conic cut [6, 25].

The implications of not being able to solve a relaxation is, in general, that there is a part of the search tree which cannot be pruned and hence prevent conclusions such as infeasibility or optimality for the mixed-integer problem. This fact motivates the integration of facial reduction techniques in the solution process of conic relaxations. Of course, turning attention to linear relaxations, these will not require regularization for solvability given that they can be solved symbolically assuming rational coefficients [26]. It will now be shown, however, that linear relaxations may need regularization for another reason, namely to strengthen the objective bounds they compute.

2.1 Objective bounds from linear relaxations

Cutting plane methods iteratively refining and resolving a linear relaxation of (P), formed by outer approximation of the non-polyhedral cones (e.g., [31]), converge with refinement towards the optimal value of (D)—not (P)!—and vice versa. That is, unless strong duality fails for the linear relaxation. Formally:

Proposition 1. *Let $\theta_P, \theta_D \in \mathbb{R} \cup \{-\infty, +\infty\}$ be the respective (and possibly unattained) optimal values of the primal-dual pair* (1)*, denoted (P) and (D). Suppose $\mathcal{C} \subseteq \mathbb{R}^n$ is a polyhedral set replacing $\mathcal{K}$ in (P) and consider the new primal-dual pair:*

$$\widehat{\theta}_P = \inf_x \{c^T x : Ax = b,\, x \in \mathcal{C}\}, \qquad \widehat{\theta}_D = \sup_{s,y} \{b^T y : c - A^T y = s,\, s \in \mathcal{C}^*\}. \tag{5}$$

Strong duality either fails for the primal-dual pair (5)*, or the following statements hold:*

1. *If $\mathcal{C} \supseteq \mathcal{K}$, then $\theta_P \geq \theta_D \geq \widehat{\theta}_D = \widehat{\theta}_P$;*
2. *If $\mathcal{C}^* \supseteq \mathcal{K}^*$, then $\widehat{\theta}_D = \widehat{\theta}_P \geq \theta_P \geq \theta_D$,*

Proof. If strong duality holds for the linear relaxation ($\widehat{\theta}_P = \widehat{\theta}_D$), the statements follow by weak duality in conic optimization ($\theta_P \geq \theta_D$). In particular, restricting the dual feasible set in Statement 1 (i.e., $\mathcal{C}^* \subseteq \mathcal{K}^*$) implies $\theta_D \geq \widehat{\theta}_D$, contra restricting the primal feasible set in Statement 2 which implies $\widehat{\theta}_P \geq \theta_P$. □

Remarkably, the subtlety of Proposition 1 is real as strong duality may fail for the primal-dual pair (5) even if it satisfied for the primal-dual pair (1). This is seen, e.g., by [34, Example 1] when the domain $x \in \mathbb{R}^2_+$ is taken to be an outer approximation of $x \in \{0\}^2$. Just as remarkable is the opposite case when strong duality holds for the linear relaxation. In particular, independently from the level of refinement used to construct the linear relaxation, the approximation error (as measured in terms of the optimal value) remains bounded by the duality gap of the conic relaxation.

Corollary 1. *Consider the primal-dual pairs* (1) *and* (5) *and suppose strong duality holds for* (5)*. The following statements hold:*

1. *If $\mathcal{C} \supseteq \mathcal{K}$, then $\theta_P - \widehat{\theta}_P \geq \theta_P - \theta_D$;*
2. *If $\mathcal{C}^* \supseteq \mathcal{K}^*$, then $\widehat{\theta}_D - \theta_D \geq \theta_P - \theta_D$,*

where $\theta_P - \theta_D \geq 0$ is the duality gap of the primal-dual pair (1)*.*

As the duality gap can be arbitrarily large for ill-posed conic relaxations (see, e.g., [40]), the objective bounds computed from linear relaxations can be arbitrarily weak. For branch-and-bound algorithms guided solely by the value of linear relaxations, this may cause a far greater number of nodes to be explored than actually needed. One remedy is to strengthen the objective bounds by using facial reduction algorithms. The expense alone, however, of checking for ill-posedness (by solving the conic optimization problems from [13], [28] or [34]), totally ruins the advantage of the linear relaxations which can be solved efficiently. This motivates the use of facial reduction heuristics to heuristically detect and fix ill-posedness and therethrough strengthen the objective bounds computed from linear relaxations.

3 Background

3.1 Cones, faces and facial reduction

Cones are subsets $\mathcal{C} \subseteq \mathbb{R}^n$ closed under positive scaling, i.e., $\lambda x \in \mathcal{C}$ holds for any $\lambda > 0$ and $x \in \mathcal{C}$. One such cone is the *dual cone* of any subset $\mathcal{C} \subseteq \mathbb{R}^n$, defined as $\mathcal{C}^* := \{y \in \mathbb{R}^n : y^T x \geq 0, \;\; \forall x \in \mathcal{C}\}$. This paper only concerns nonempty, closed, convex cones $\mathcal{C}$. In this case, the dual cone $\mathcal{C}^*$ is also a nonempty, closed, convex cone, $\mathcal{C}$ equals $(\mathcal{C}^*)^*$ and the origin is contained [36].

The conical hull of any set $S \subseteq \mathbb{R}^n$, denoted $\operatorname{cone}(S)$, is defined as the union of the origin and all finite conical combinations, $\lambda_1 s_1 + \ldots + \lambda_k s_k$ for $\lambda \in \mathbb{R}^k_{++}$, of points $s_1, \ldots, s_k \in S$. By definition, it is a nonempty, convex cone [36].

Polyhedral cones are conical hulls, $\operatorname{cone}(S)$, of any finite set of extreme rays S. By definition, it is a nonempty, closed, convex cone [36, Theorem 19.1].

Proper cones are nonempty, closed and convex, besides being *solid* and *pointed* [10]. Specifically, *solid cones* have nonempty interior, $\operatorname{int} \mathcal{C} \neq \emptyset$, or, equivalently, are full-dimensional in the sense that $\operatorname{span} \mathcal{C} = \mathbb{R}^n$ and $\mathcal{C}^\perp = (\operatorname{span} \mathcal{C})^\perp = \{0\}^n$. *Pointed cones*, on the other hand, contain no lines, i.e., $\mathcal{C} \cap (-\mathcal{C}) = \{0\}^n$.

Self-dual cones satisfy $\mathcal{C} = \mathcal{C}^*$. The self-dual and proper cones used in the examples of this paper are given by the nonnegative orthant $\mathbb{R}^n_+$, the quadratic cone $\mathcal{Q}^n := \{x \in \mathbb{R}^n : x_1 \geq \|x_{2:n}\|_2\}$, the rotated quadratic cone $\mathcal{Q}^n_r := \{(x_1, x_2, x_{3:n}) \in \mathbb{R}^2_+ \times \mathbb{R}^{n-2} \mid 2x_1 x_2 \geq \|x_{3:n}\|_2^2\}$, and the semidefinite cone $\mathcal{S}^N_+ := \left\{VV^T : V \in \mathbb{R}^{N \times N}\right\}$ (see, e.g., [44] for more properties).

Faces of a set $S \subseteq \mathbb{R}^N$ are subsets $\mathcal{F} \subseteq \mathcal{S}$ for which any line segment in $\mathcal{S}$, with a midpoint in $\mathcal{F}$, has both endpoints in $\mathcal{F}$ [36]. A *proper face* of $\mathcal{S}$ is a face which is nonempty and not equal to $\mathcal{S}$. These definitions generalize extreme points and other faces from polyhedra, and will now be related to the nonempty, closed, convex cones $\mathcal{C}$.

Let $z^\perp := \{x \in \mathbb{R}^n : x^T z = 0\}$. For $z \in \mathcal{C}^*$, the intersection $\mathcal{C} \cap z^\perp$ contains the origin and is a face of $\mathcal{C}$ as it maximizes $-z^T x$ over $x \in \mathcal{C}$ [36]. Hence, if $z \in \mathcal{C}^* \setminus \mathcal{C}^\perp$, the intersection $\mathcal{C} \cap z^\perp$ cannot equal $\mathcal{C}$ and hence represent a proper face of $\mathcal{C}$.

In these terms, a *facial reduction* can finally be defined as a valid problem reformulation in which a cone $\mathcal{C}$ is replaced by one of its proper faces, e.g., $\mathcal{C} \cap z^\perp$ for some $z \in \mathcal{C}^* \setminus \mathcal{C}^\perp$.

3.2 Facial reduction certificates

Consider again the primal-dual pair (1), restated here for the convenience of the reader:

$$(P): \quad \theta_P = \inf_x \{c^T x : Ax = b, \, x \in \mathcal{K}\}, \qquad (D): \quad \theta_D = \sup_{s,y} \{b^T y : c - A^T y = s, \, s \in \mathcal{K}^*\}.$$

where $\mathcal{K}, \mathcal{K}^* \subseteq \mathbb{R}^n$ are nonempty, closed, convex cones. Any equation of the form $z^T x = 0$, valid in (P) for some $z \in \mathcal{K}^* \setminus \mathcal{K}^\perp$, justifies the facial reduction from cone $\mathcal{K}$ to its proper face $\mathcal{K} \cap z^\perp$. If the equation is implied by the equation system $Ax = b$, formally denoted $z^\perp \supseteq \{x \in \mathbb{R}^n : Ax = b\}$, the exposing vector z is called a *facial reduction certificate* for (P) [34]. Similarly, the exposing vector $z \in \mathcal{K} \setminus (\mathcal{K}^*)^\perp$ of a facial reduction in (D), is called a facial reduction certificate for (D) if the required equation $z^T s = 0$ is implied by the equation system $c - A^T y = s$. The immediate conic formulation of the feasible set of facial reduction certificates for (P) and (D), respectively, can thus be stated as follows.

Proposition 2 ([34]). *The following statements hold:*

1. *$z \in \mathbb{R}^n$ is a facial reduction certificate for (P) if the exists a $\omega \in \mathbb{R}^m$ for which*

$$b^T\omega = 0, \qquad z = -A^T\omega, \qquad z \in \mathcal{K}^* \setminus \mathcal{K}^\perp,$$

and all facial reduction certificates are of this form if $\{x \in \mathbb{R}^n : Ax = b\} \neq \emptyset$.

2. *$z \in \mathbb{R}^n$ is a facial reduction certificate for (D) if and only if*

$$c^Tz = 0, \qquad Az = 0, \qquad z \in \mathcal{K} \setminus (\mathcal{K}^*)^\perp.$$

The conic formulations in statement 1 and 2 of Proposition 2 are called *auxiliary problems* of the primal-dual pair (1). Note that a constraint of the form $x \in \mathcal{C} \setminus (\mathcal{C}^*)^\perp$ is satisfied if and only if $x \in \mathcal{C}$ and x has nonzero inner-product with any point in $\operatorname{relint} \mathcal{C}^*$. Hence, $x \in \mathcal{C} \setminus (\mathcal{C}^*)^\perp$ can be reformulated as $x \in \mathcal{C}$ and $\hat{p}^Tx = 1$ for $\hat{p} \in \operatorname{relint} \mathcal{C}^*$ in the statements of Proposition 2, thereby normalizing the solutions to the homogeneous auxiliary problems.

3.3 Subgradient-based outer approximation

Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a nonempty, closed, convex cone with membership indicator function

$$\chi_\mathcal{C}(x) = \begin{cases} 0 & \text{if } x \in \mathcal{C}, \\ +\infty & \text{otherwise.} \end{cases}$$

The indicator function is convex, since the set $\mathcal{C}$ is convex [36], and offers the useful description $C = \{x \in \mathbb{R}^n : \chi_\mathcal{C}(x) \leq 0\}$. In particular, valid inequalities of $\{x \in \mathbb{R}^n : \chi_\mathcal{C}(x) \leq 0\}$ are readily provided by the subgradient inequality

$$\chi_\mathcal{C}(\hat{x}) + \overline{\xi}^T(x - \hat{x}) \ \leq \ 0, \tag{6}$$

holding for any $\hat{x} \in \mathbb{R}^n$, and all subgradients $\overline{\xi}$ of the corresponding subdifferential set $\partial\chi_\mathcal{C}(\hat{x}) \subseteq \mathbb{R}^n$. This is a direct consequence of the inequality $\chi_\mathcal{C}(x) \leq 0$, combined with the fact that subgradients $\overline{\xi} \in \partial\chi_\mathcal{C}(\hat{x})$ are defined (e.g., in [36]) as solutions to

$$\chi_\mathcal{C}(\hat{x}) + \overline{\xi}^T(x - \hat{x}) \ \leq \ \chi_\mathcal{C}(x) \text{ for all } x \in \mathbb{R}^n. \tag{7}$$

For the subset of points $\hat{x} \notin \mathcal{C}$, the system (7) has no solution as $\chi_\mathcal{C}(\hat{x}) = +\infty$. For all other points, $\hat{x} \in \mathcal{C}$, the system is solved by $\overline{\xi} \in -\mathcal{C}^* \cap \hat{x}^\perp$ [36, Corollary 23.5.4]. This allows us to characterize the family of subgradient inequalities (6) in a much simpler way.

Theorem 1. *For nonempty, closed, convex cones $\mathcal{C} \subseteq \mathbb{R}^n$, the set obtained by intersecting all subgradient inequalities of the form (6), is concisely described by*

$$\xi^Tx \geq 0, \qquad \textit{for } \xi \in \Omega,$$

where $\Omega \subseteq \mathbb{R}^n$ is any set satisfying $\operatorname{cone}(\Omega) = \mathcal{C}^*$.

Proof. The subdifferential set is the solutions to (7) given by

$$\partial\chi_\mathcal{C}(\hat{x}) = \begin{cases} \emptyset & \text{if } \hat{x} \notin \mathcal{C}, \\ -\mathcal{C}^* \cap \hat{x}^\perp & \text{otherwise,} \end{cases}$$

as argued. Hence, to satisfy $\bar{\xi} \in \partial\chi_{\mathcal{C}}(\hat{x})$ as needed in (6), both $\chi_{\mathcal{C}}(\hat{x}) = 0$ and $\bar{\xi}^T\hat{x} = 0$ is required. This simplifies (6) to $\bar{\xi}^T x \leq 0$, holding for any $\hat{x} \in \mathcal{C}$ and all $\bar{\xi} \in -\mathcal{C}^* \cap \hat{x}^\perp$. The claim is hence shown for $\Omega = \cup_{\hat{x}\in\mathcal{C}}(\mathcal{C}^* \cap \hat{x}^\perp) = \mathcal{C}^*$ (noting that $\mathcal{C}$ contains the origin), in terms of the negated subgradient $\xi = -\bar{\xi}$. Finally, if $\text{cone}(\Omega) = \mathcal{C}^*$, then any $\xi \in \mathcal{C}^*$ has a finite conical factorization $\xi = \sum_{j=1}^k \lambda_j \xi_j^T$ for $\lambda \in \mathbb{R}^k_{++}$ and $\xi_j \in \Omega$. This shows $\xi^T x = \lambda_1\xi_1^T x + \ldots + \lambda_k\xi_k^T x \geq 0$ to be redundant given $\xi_j^T x \geq 0$ for $j = 1, \ldots, k$. □

Needless to say, an outer approximation of $\mathcal{C}$ is obtained by Theorem 1 from any finite subset of $\Omega \subseteq \mathbb{R}^n$, or less pedantic, from any finite subset of $\mathcal{C}^*$.

Corollary 2. *A subgradient-based outer approximation of a nonempty, closed, convex cone $\mathcal{C} \subseteq \mathbb{R}^n$ is given by $\hat{\Omega}^* = \{x : \xi^T x \geq 0 \text{ for } \xi \in \hat{\Omega}\}$, that is $\hat{\Omega}^* \supseteq \mathcal{C}$, for any finite subset $\hat{\Omega} \subseteq \mathcal{C}^*$.*

The rather trivial corollary above does not exploit the distinction between Ω and $\mathcal{C}^*$ in Theorem 1, and may hence lead to formulations with redundancies. What Theorem 1 actually suggests is that Ω can be chosen as any minimal set of conically independent points generating $\mathcal{C}^*$ in the sense that $\text{cone}(\Omega) = \mathcal{C}^*$. This notion can be clarified in case $\mathcal{C}^*$ is a pointed, nonempty, closed, convex cone not equal to $\{0\}^n$. Specifically, in this case, a necessary and sufficient condition for $\text{cone}(\Omega) = \mathcal{C}^*$ is that Ω contains a relative interior point from all one-dimensional faces of $\mathcal{C}^*$. This follows by [36, Corollary 18.5.2.], noting that *extreme rays* are used to denote the set of half-line faces, i.e., the set of one-dimensional faces for pointed, nonempty, closed, convex cones. This leads to the following stricter definition of an outer approximation obtained by Theorem 1.

Corollary 3. *Suppose Ω is any minimal set of conically independent points generating $\mathcal{C}^*$ in the sense that* $\text{cone}(\Omega) = \mathcal{C}^*$*. A subgradient-based outer approximation of a nonempty, closed, convex cone $\mathcal{C} \subseteq \mathbb{R}^n$ is then given by $\hat{\Omega}^* = \{x : \xi^T x \geq 0 \text{ for } \xi \in \hat{\Omega}\}$ for any finite subset $\hat{\Omega} \subseteq \Omega$.*

Outer approximations on the form of Corollary 3 appears for second-order cones in [7], and for semidefinite cones in [30, 33]. These outer approximations are constructed either statically as in [33], or dynamically as in [7, 30] by iteratively separating violated points. This indicates that the implications of Theorem 1 are folklore, although never formalized to the author's knowledge. The paper returns to show the static outer approximations of [33] in the context of facial reduction heuristics in Section 4.1.

4 Facial reduction heuristics

The auxiliary problems of Proposition 2 describe the feasible set of facial reduction certificates for (P) and (D), respectively. Solving simplifications of these sets are hence a straightforward way to design facial reduction heuristics. This is done for the family of heuristics based on linear optimization in Section 4.1, and for the heuristics based on cone factorization in Section 4.4. Alternatively, one may search for equations of the form $z^T x = 0$ in (P) (or $z^T s = 0$ in (D)) and detect whenever the exposing vector $z \in \mathbb{R}^n$ exposes a proper face of the respective cone. This is done in the subgradient matching heuristic of Section 4.2 and in the single-cone analysis of Section 4.3.

4.1 A family of heuristics based on linear optimization

This family of facial reduction heuristics have in common that they solve the auxiliary problems of Proposition 2 using linear inner approximations of the conic variable domains. This allows the partially represented set of facial reduction certificates to be explored fairly efficient and to high accuracy using simplex methods. In fact, the use of simplex methods allow *exact certificates* in rational arithmetic to be computed for this family of heuristics, within reasonable computational efforts [26]. This motivates the use of heuristics based on linear optimization.

To begin, note that a partial set of facial reduction certificates is clearly obtained through inner approximation of the auxiliary problems in Proposition 2. One example of this is given in [33], where inner approximations of the conic variable domain are obtained from span-invariant outer approximations of the cone considered for facial reduction. Taking the auxiliary problem of Proposition 2-(1) as example, an inner approximation of the conic variable domain,

$$\mathcal{C}^* \setminus \mathcal{C}^\perp \subseteq \mathcal{K}^* \setminus \mathcal{K}^\perp,$$

clearly holds if $\mathcal{C}^* \subseteq \mathcal{K}^*$ and $\mathcal{C}^\perp \supseteq \mathcal{K}^\perp$ is satisfied. The former containment shows $\mathcal{C}$ to be an outer approximation of $\mathcal{K}$ by its contrapositive, $\mathcal{C} \supseteq \mathcal{K}$, and the latter containment establish span-invariance. In particular, the latter containment and its contrapositive, $\operatorname{span} \mathcal{C} \subseteq \operatorname{span} \mathcal{K}$, must hold with equality given $\mathcal{C} \supseteq \mathcal{K}$. This characterization leads to the following family of inner approximated auxiliary problems.

Proposition 3. *A partial set of facial reduction certificates, $z \in \mathbb{R}^n$, is obtained from any span-invariant outer approximation of the cone from the considered problem. Specifically:*

1. *Given $\mathcal{K} \subseteq \mathcal{C} \subseteq \mathbb{R}^n$ where $\operatorname{span} \mathcal{K} = \operatorname{span} \mathcal{C}$, a partial description of the facial reduction certificates for (P) is given by*

$$b^T \omega = 0, \qquad z = -A^T \omega, \qquad z \in \mathcal{C}^* \setminus \mathcal{C}^\perp.$$

2. *Given $\mathcal{K}^* \subseteq \mathcal{C}^* \subseteq \mathbb{R}^n$ where $\operatorname{span} \mathcal{K}^* = \operatorname{span} \mathcal{C}^*$, a partial description of the facial reduction certificates for (D) is given by*

$$c^T z = 0, \qquad Az = 0, \qquad z \in \mathcal{C} \setminus (\mathcal{C}^*)^\perp.$$

Restricting attention to primal-dual pairs (1) where the cone considered for facial reduction is solid, the span-invariance of Proposition 3 is automatically satisfied. This makes it possible to apply the subgradient-based outer approximation of Corollary 2 directly. A partial description of the facial reduction certificates for (P) is thus obtained from Proposition 3-(1), using

$$\mathcal{C} = \mathcal{C}_1 \times \cdots \times \mathcal{C}_k, \text{ where } \mathcal{C}_j = \{x \in \mathbb{R}^n : \xi^T x \geq 0, \ \forall \xi \in \hat{\Omega}_j\}, \tag{8}$$

given any finite subset of negated subgradients, $\hat{\Omega}_j \subseteq \mathcal{K}_j^*$, for all factors of the Cartesian product $\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_k$. It is worth noting that $C_j = \hat{\Omega}_j^*$ by definition and hence $\mathcal{C}_j^* = \operatorname{cone}(\hat{\Omega}_j)$ by the bipolar theorem [36]. Two examples of this inner approximation are now given, satisfying the stricter requirements of Corollary 3 on the selection of $\hat{\Omega}_j$.

Example 2. *The semidefinite cone $\mathcal{S}_+^N$ is outer approximated in accordance with Corollary 3, by a finite subset $\hat{\Omega}$ of rank-one matrices of the form $\omega\omega^T$ for $\omega \in \mathbb{R}^N$, representing the set of extreme rays for $\mathcal{S}_+^N$ [24]. Two examples from [33] are given by:*

1. *The non-negative diagonal approximation chooses ω as any permutation of $(1, 0, \ldots, 0)^T$. Hence, $\mathcal{C} = \{X \in \mathcal{S}^n : X_{ii} \geq 0\}$ and $\mathcal{C}^* = \{X \in \mathcal{S}^n : X_{ii} \geq 0, \ X_{ij} = 0 \text{ for } i \neq j\}$.*

2. *The diagonally-dominant approximation chooses* ω *as any permutation of either* $(1, 0, \ldots, 0)^T$, $(1, 1, 0, \ldots, 0)^T$ *or* $(1, -1, 0, \ldots, 0)^T$. *Hence,* $\mathcal{C} = \{X \in \mathcal{S}^n : X_{ii} \geq 0, X_{ii} + X_{jj} \geq 2|X_{ij}|\}$ *and* $\mathcal{C}^* = \{X \in \mathcal{S}^n : X_{ii} \geq \sum_{j \neq i} |X_{ij}|\}$.

These realizations of $\mathcal{C}$ *give rise to inner approximated auxiliary problems through Proposition 3, that are shown effective in [33] for a wide range of ill-posed semidefinite optimization instances.*

With the right choice of $\hat{\Omega}_j$ in Corollary 2, any facial reduction certificate can be found using the subgradient-based inner approximated auxiliary problems of Proposition 3. The only catch is that the right choice of $\hat{\Omega}_j$ literally have to guess which facial reductions are possible, in order for this family of heuristics to recognize them.

Proposition 4. *Let* $z = (z_1, \ldots, z_k)$ *be a facial reduction certificate for* (P), *where* $\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_k$ *is a solid cone. Suppose further that this* z *is feasible in the auxiliary problem of Proposition 3-*(1) *for a cone* $\mathcal{C} = \mathcal{C}_1 \times \cdots \times \mathcal{C}_k$ *defined as in* (8). *Then* $\mathcal{C}^* = \operatorname{cone}(\hat{\Omega}_1) \times \cdots \times \operatorname{cone}(\hat{\Omega}_k)$ *for finite subsets* $\hat{\Omega}_j \subseteq \mathcal{K}_j^*$, *giving rise to finite conical factorizations* $z = (\sum_{i=1}^{p_1} \lambda_1^i \xi_1^i, \ldots, \sum_{i=1}^{p_k} \lambda_k^i \xi_k^i)$ *for* $\lambda_j^i > 0$ *and nonzero* $\xi_j^i \in \hat{\Omega}_j$. *It now follows that the reformulation from* $\mathcal{K}_j$ *to* $\mathcal{K}_j \cap (\xi_j^i)^\perp$ *is a valid facial reduction in* (P) *for all* $j = 1, \ldots, k$ *and* $i = 1, \ldots, p_j$.

Proof. A facial reduction certificate for (P) satisfy $z \in \mathcal{K}^*$ by Proposition 2. Hence,

$$\mathcal{K} \cap z^\perp = (\mathcal{K}_1 \cap z_1^\perp) \times \cdots \times (\mathcal{K}_k \cap z_k^\perp)$$

as seen, e.g., by [17, Corollary 1]. In turn, each Cartesian factor can be rewritten as

$$\mathcal{K}_j \cap z_j^\perp = \mathcal{K}_j \cap (\textstyle\sum_{i=1}^{p_j} \lambda_j^i \xi_j^i)^\perp = \mathcal{K}_j \bigcap_{i=1}^{p_j} (\lambda_j^i \xi_j^i)^\perp = \mathcal{K}_j \bigcap_{i=1}^{p_j} (\xi_j^i)^\perp,$$

where the first equality follows by definition. The second equality is a consequence of [17, Proposition 3], given $\lambda_j^i \xi_j^i \in \mathcal{K}_j^*$ as implied by $\lambda_j^i > 0$ and $\xi_j^i \in \hat{\Omega}_j \subseteq \mathcal{K}_j^*$. The last equality is from invariance of the orthogonal complement to positive scaling. The claim hence follows from $\mathcal{K}_j \cap (\xi_j^i)^\perp$ being a relaxation of $\mathcal{K} \cap z^\perp$, and from noting that $\mathcal{K}_j \cap (\xi_j^i)^\perp$ defines a proper face of the solid cone $\mathcal{K}_j$. □

The equivalent statement for the subgradient-based inner approximated auxiliary problem of Proposition 3-(2) is shown similarly. There is another and possibly more intuitive way to think of this. In particular, note that the partial set of certificates for the primal-dual pair (1), is the complete set of certificates for the linear primal-dual pair:

$$\widehat{\theta}_P = \inf_x \{c^T x : Ax = b,\, x \in \mathcal{C}\}, \qquad \widehat{\theta}_D = \sup_{s,y} \{b^T y : c - A^T y = s,\, s \in \mathcal{C}^*\}, \tag{9}$$

as seen by comparing Proposition 2 with Proposition 3. Hence, you have to guess the facial reductions in order to find them in (9), since all facial reductions of linear problems are inequalities holding as implied equalities [13] and the only inequalities of (9) are those with normal vectors from $\hat{\Omega}_j$. As example, if $\mathcal{C}$ is defined by the non-negative diagonal approximation of Example 2-(1), the only facial reductions that can be found are from diagonal entries fixed to zero. Simpler variable bound analysis may also yield these conclusions and hence might enjoy much of the success reported for the non-negative diagonal approximation in [33]. This motivates subgradient matching.

4.2 Subgradient matching

In contrast to the previous family of heuristics that require one to solve an optimization problem, albeit linear, the subgradient matching technique integrates with domain propagation to provide a faster facial reduction heuristic. This integration has synergistic effects too, even when no facial reductions are identified, as subgradient matching acts to strengthen the variable and constraint activity bounds derived through domain propagation as will be shown.

Consider the activity bounds of an affine expression $a^T x$ for some $a \in \mathbb{R}^n$, given by

$$L_{\min} = \sum_{j:a_j>0} a_j l_j + \sum_{j:a_j<0} a_j u_j \quad \text{and} \quad L_{\max} = \sum_{j:a_j>0} a_j u_j + \sum_{j:a_j<0} a_j l_j, \tag{10}$$

where $l, u \in (\mathbb{R} \cup \{-\infty, +\infty\})^n$ holds the domain propagated lower and upper variable bounds. These are called the simplest, but also the weakest, activity bounds by Savelsbergh [37]. A simple advancement of these bounds is to include subgradient information from the conic variable domain $x \in \mathcal{K}$ as shown in Algorithm 1. The correctness of this algorithm is proven in Proposition 5.

Algorithm 1: Computing activity bounds of $a^T x$ using subgradient matching.

Data: Suppose $x = (x^1, \ldots, x^k) \in \mathcal{K}_1 \times \cdots \times \mathcal{K}_k$, dictating similar partitioning of $a = (a^1, \ldots, a^k)$, $l = (l^1, \ldots, l^k)$ and $u = (u^1, \ldots, u^k)$.

1 **for all** $i \in \{1, \ldots, k\}$ **do**

2 Compute the simple activity bounds of $(a^i)^T x^i$:

$$L^i_{\min} \leftarrow \sum_{j:a^i_j>0} a^i_j l^i_j + \sum_{j:a^i_j<0} a^i_j u^i_j \quad \text{and} \quad L^i_{\max} \leftarrow \sum_{j:a^i_j>0} a^i_j u^i_j + \sum_{j:a^i_j<0} a^i_j l^i_j.$$

3 **if** $L^i_{\min} \leq 0$ **and** $a^i \in \mathcal{K}^*_i$ **then**

4 Update $L^i_{\min} \leftarrow 0$.

5 **end**

6 **if** $L^i_{\max} \geq 0$ **and** $a^i \in -\mathcal{K}^*_i$ **then**

7 Update $L^i_{\max} \leftarrow 0$.

8 **end**

9 **end**

10 **return** $L_{\min} \leftarrow \sum_{i=1}^k L^i_{\min}$ and $L_{\max} \leftarrow \sum_{i=1}^k L^i_{\max}$.

Proposition 5. *The activity bounds computed by Algorithm 1 are valid and stronger than those computed by* (10)*.*

Proof. If line 4 (resp. line 7) of Algorithm 1 never executes, then $L_{\min}$ (resp. $L_{\max}$) equals the value computed by (10). Otherwise, when line 4 executes, then $(a^i)^T x^i \geq 0$ holds by definition of dual cones and strengthens the current lower bound if $L^i_{\min} < 0$. Similarly, when line 7 executes, then $(a^i)^T x^i \leq 0$ holds and strengthens the current upper bound if $L^i_{\max} > 0$. □

The attentive reader may question the use of non-strict inequalities on line 4 and 7 of Algorithm 1, as the updating of $L^i_{\min}$ and $L^i_{\max}$ is needless if they are already zero. This is done intentionally, however, to stress that one should pay attention to all cases in which the subgradient matching bound is the tightest bound computed. In particular, whenever this holds, valid facial reductions can be identified from *forcing* constraints on the affine expression $a^T x$.

Definition 1. *Suppose $\sum_{i=1}^p L^i = b$. A constraint is considered forcing if*

1. $\sum_{i=1}^p (a^i)^T x^i \leq b$ *for lower bounded terms* $(a^i)^T x^i \geq L^i$ *for all* $i \in \{1, \ldots, p\}$;
2. $\sum_{i=1}^p (a^i)^T x^i \geq b$ *for upper bounded terms* $(a^i)^T x^i \leq L^i$ *for all* $i \in \{1, \ldots, p\}$;
3. $\sum_{i=1}^p (a^i)^T x^i = b$ *for lower bounded terms (resp. upper bounded terms) as above,*

whereby $(a^i)^T x^i = L^i$ *for all* $i \in \{1, \ldots, p\}$ *is implied.*

Forcing constraints appear already in [3], but the idea that they may be used to identify valid facial reductions is new. This idea is formalized in the following proposition.

Proposition 6. *Consider a forcing constraint from Definition 1. If any of the implied equations* $(a^i)^T x^i = L^i$ *holds with* $L^i = 0$ *and* $a^i \in \mathcal{K}^* \setminus \mathcal{K}^\perp$ *for some* $i \in \{1, \ldots, p\}$*, the facial reduction from* $x^i \in \mathcal{K}$ *to* $x^i \in \mathcal{K} \cap (a^i)^\perp$ *is justified.*

Comparing Proposition 6 to Algorithm 1, it can now be verified that line 4 and 7 captures the sought conditions for a facial reduction whenever $a^i \notin \mathcal{K}_i^\perp$. Note that this is equivalent to $a^i \neq 0$ for solid cones. In an attempt to evaluate the usefulness of subgradient matching without an implementation, the reader is invited to verify that it is capable of regularizing the $[x_4 \leq 0]$-node of Example 1. It can furthermore be shown to recognize all primal facial reductions of all examples in [34], including the following less trivial reduction.

Example 3. *Consider the primal-dual pair originating with [2], given by*

$$
\begin{array}{llcll}
\inf\limits_{x} & x_3 & \qquad & \sup\limits_{s,y} & y_2 \\
\text{s.t.} & x_1 + x_2 + x_4 + x_5 = 0, & & \text{s.t.} & \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} y_1 \\ y_1 \\ -y_2 \\ y_1 + y_2 \\ y_1 \end{pmatrix} = s, \\
& -x_3 + x_4 = 1, & & & \\
& x \in \mathcal{Q}^3 \times \mathcal{Q}^2, & & & s \in \mathcal{Q}^3 \times \mathcal{Q}^2.
\end{array}
\tag{11}
$$

Simple activity bounds for the equation $x_1 + x_2 + x_4 + x_5 = 0$ *leads to* $L_{\min} = -\infty$ *and* $L_{\max} = \infty$ *since* x_2 *and* x_5 *are free variables. This equation is nevertheless forcing according to Definition 1, as shown by the following partition into Cartesian factors,*

$$a^T x = (a^1)^T x^1 + (a^2)^T x^2 = (x_1 + x_2) + (x_4 + x_5) = 0,$$

where $a^1 = (1, 1, 0)^T$ *and* $a^2 = (1, 1)^T$*. In particular, subgradient matching finds* $L_{\min}^1 = 0$ *since* $a^1 \in \mathcal{Q}^3$*, as well as* $L_{\min}^2 = 0$ *since* $a^2 \in \mathcal{Q}^2$*, to conclude* $L_{\min} = 0$ *which matches the value* $a^T x$ *is constrained to take. Hence,* $x_1 + x_2 = 0$ *and* $x_4 + x_5 = 0$ *by Proposition 6, exposing two facial reductions. That is, from* $x^1 \in \mathcal{Q}^3$ *to* $x^1 \in \mathcal{Q}^3 \cap (a^1)^\perp$ *and from* $x^2 \in \mathcal{Q}^2$ *to* $x^2 \in \mathcal{Q}^2 \cap (a^2)^\perp$.

Finally, note that simpler forms of subgradient matching have already appeared in literature. Gruber et al. [21] (and latter Cheung et al. [13]) notice that the existence of an equation $a^T x = 0$ in the considered problem, for a semidefinite coefficient and variable $a, x \in \mathcal{S}_+^n$, implies validity of the facial reduction from $x \in \mathcal{S}_+^n$ to $x \in \mathcal{S}_+^n \cap a^\perp$. This special case of subgradient matching was shown useful for the side chain positioning problem in [12].

Another far less apparent appearance of subgradient matching is given by the widely known facial reduction made possible by diagonal entries of a semidefinite cone fixed to zero (see, e.g., [39, page 535]). We elaborate on this specific usecase in the following paragraph.

Subgradient matching on variable bounds For a conic variable domain $x^i \in \mathcal{K}$, consider the set of variables selected by $a^T x$ for all permutations of $a = (1, 0, \dots, 0)$ belonging to $\mathcal{K}^* \setminus \mathcal{K}^\perp$ (resp. to $-\mathcal{K}^* \setminus \mathcal{K}^\perp$). The variables of this set all have lower (resp. upper) bounds of zero by definition of dual cones. Suppose then that an upper bound $x_j \leq 0$ (resp. a lower bound $x_j \geq 0$) was derived for a variable of this set, e.g., using domain propagation. This would then show that the variable was forced to zero (as the execution of Algorithm 1 on this bound constraint would also reveal), whereby the facial reduction of Proposition 6 can be used on the conic variable domain.

Subgradient matching with general conic constraints One may confirm that Definition 1 and Proposition 6 can be generalized from expressions $\sum_{i=1}^p (a^i)^T x^i$ following the variable partition $x = (x^1, \dots, x^p)$, to any finite list of variable-overlapping expressions of the form $\sum_{i=1}^p (a^i)^T x$. As a consequence, Algorithm 1 can be extended to include subgradient information from general conic constraints, $Dx - d \in \mathcal{K}$ for some $D \in \mathbb{R}^{p \times n}$ and $b \in \mathbb{R}^p$, using the partitioning strategy

$$a^T x = (a^1)^T x + (a^2)^T x = \Big(a^T x - \lambda^T (Dx - d)\Big) + \lambda^T (Dx - d),$$

for any $\lambda \in \mathbb{R}^p$. The term $(a^2)^T x = \lambda^T (Dx - d)$ is nonnegative and invariant to the computation of $L_{\min}$ if $\lambda \in \mathcal{K}^*$, or nonpositive and invariant to the computation of $L_{\max}$ if $\lambda \in -\mathcal{K}^*$. Hence, by restricting the domain of λ, one can disregard $(a^2)^T x$ in the respective bound computation and actively choose λ to strengthen the $L_{\min}$ (resp. $L_{\max}$) bound of $(a^1)^T x$, e.g., by eliminating unbounded variables. If the constraint on $a^T x$ is shown forcing, the expression $(a^2)^T x$ should of course be used to check for facial reductions in Proposition 6. In particular, if $\lambda \in \mathcal{K} \setminus (\mathcal{K}^*)^\perp$ (resp. $\lambda \in -\mathcal{K} \setminus (\mathcal{K}^*)^\perp$) holds, then the facial reduction from $Dx - d \in \mathcal{K}$ to $Dx - d \in \mathcal{K} \cap \lambda^\perp$ is justified. An exact strategy for subgradient matching with general conic constraints is left unexplored.

4.3 Single-cone analysis

As opposed to subgradient matching, which integrates information from conic constraints into the analysis of a single linear constraint, single-cone analysis acts to integrate information from linear constraints into the analysis of a single conic constraint. This is realized as a facial reduction heuristic for the second-order cones in this section. In motivation of this heuristic, the following example from [23] is now presented.

Example 4. *The non-standard mixed-integer optimization problem,*

$$\begin{array}{cl} \inf\limits_{x,t} & x^2 + t \\ \text{s.t.} & x - 4 \geq 0 \text{ if } t = 0, \\ & x \in \mathbb{R}_+, \\ & t \in \{0, 1\}, \end{array}$$

is solved by $(x, t) = (0, 1)$ and representable as a conic optimization problem by

$$\begin{array}{cl} \inf\limits_{x,t,\omega,y,\gamma} & \omega + t \\ \text{s.t.} & x - y - 4 \geq 0, \\ & \begin{pmatrix} 1/2 \\ \omega \\ x \end{pmatrix} \in \mathcal{Q}_r^3, \quad \begin{pmatrix} \gamma + t \\ \gamma - t \\ 2y \end{pmatrix} \in \mathcal{Q}^3, \\ & x \in \mathbb{R}_+, \\ & t \in \{0, 1\}. \end{array}$$

To give a feeling for this conic formulation, the first conic constraint models $\omega \geq x^2$ to represent the squared objective contribution. The second conic constraint models $y = 0$ if $t = 0$ and $\sqrt{\gamma} \geq |y|$ if $t = 1$ (for unbounded γ) to represent the conditional constraint (see [23] for more details).

Branching on $t = 0$, the relaxation is ill-posed as there is no way to satisfy $\begin{pmatrix} \gamma+t \\ \gamma-t \\ 2y \end{pmatrix} \in \operatorname{relint} \mathcal{Q}^3$ as needed for strong feasibility. In this case, however, there is also a linear dependency exposed by $z^T \begin{pmatrix} \gamma+t \\ \gamma-t \\ 2y \end{pmatrix} = 0$ for $z = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$. That is, the conic constraint $\begin{pmatrix} \gamma+t \\ \gamma-t \\ 2y \end{pmatrix} \in \mathcal{Q}^3$ can be restated as

$$\begin{pmatrix} \gamma+t \\ \gamma-t \\ 2y \end{pmatrix} \in \mathcal{Q}^3 \cap \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}^{\perp} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \mathbb{R}_+ \text{ (see [17]).}$$

This simplifies to $\gamma \geq 0$ and $2y = 0$, and regularizes the relaxation as the reader may confirm by verifying the property of strong feasibility.

In order to systematize the search, exemplified above, for linear dependencies between the entries of conic constraints, let $D \in \mathbb{R}^{p \times n}$, $d \in \mathbb{R}^p$ and $\mathcal{K} \subseteq \mathbb{R}^p$ and consider

$$Dx - d = \left(D,\, d \right) \left(\begin{smallmatrix} x \\ -1 \end{smallmatrix} \right) \in \mathcal{K}.$$

Any linear dependency $z^T \left(D,\, d \right) = 0$, for nonzero $z \in \mathbb{R}^p$, justifies reformulation to the subspace intersected conic constraint $Dx - d \in \mathcal{K} \cap z^{\perp}$. For z to expose a proper face of a $\mathcal{K}$, as needed for a facial reduction, we require a solution to the system

$$z^T \left(D,\, d \right) = 0, \quad z \in \mathcal{K}^* \setminus \mathcal{K}^{\perp},$$

which is comparable to Proposition 2-(2) and equivalent to the system

$$\left(D,\, d \right)^T z = 0, \quad \hat{p}^T z = 1, \quad z \in \mathcal{K}^*, \tag{12}$$

for some relative interior point $\hat{p} \in \operatorname{relint} \mathcal{K}$ [33]. Finally, if $\mathcal{K}^*$ is the image under linear mapping of another set, i.e., $\mathcal{K}^* = H\hat{\mathcal{K}}^*$, then (12) with $z = H\hat{z}$ is equivalent to the system

$$\left(H^T D,\, H^T d \right)^T \hat{z} = 0, \quad (H^T \hat{p})^T \hat{z} = 1, \quad \hat{z} \in \hat{\mathcal{K}}^*. \tag{13}$$

These systems are now shown analytically solvable for the second-order cones. The first step in this direction is to rewrite (12) as a least-norm optimization problem for the quadratic cone.

Proposition 7. *Let $D = \begin{pmatrix} \alpha^T \\ A \end{pmatrix} \in \mathbb{R}^{p \times n}$ and $d = \begin{pmatrix} \beta \\ b \end{pmatrix} \in \mathbb{R}^p$. The system* (12)*, with $\mathcal{K} = \mathcal{Q}^p$ and the relative interior point $\hat{p} = (1, 0, \ldots, 0)^T \in \operatorname{relint} \mathcal{Q}^p$, is feasible if and only if $\theta_P \leq 1$ for the least-norm optimization problem*

$$\begin{array}{rcll} \theta_P & = & \inf\limits_{\lambda} & \|\lambda\|_2 \\ & & \text{s.t.} & \left(A,\, b \right)^T \lambda = -\left(\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix} \right), \\ & & & \lambda \in \mathbb{R}^{p-1}. \end{array} \tag{14}$$

Specifically, a feasible point of (12) *is given by $z = (1, \hat{\lambda}^T)^T$ for any feasible point $\hat{\lambda}$ of* (14) *with an objective value less than or equal to one.*

Proof. Let $z \in \mathbb{R}^p$ satisfy (12) such that $\hat{p}^T z = z_1 = 1$. Then $z = (1, \lambda^T)^T$ for some $\lambda \in \mathbb{R}^{p-1}$. The claim follows by $z \in \mathcal{Q}^p \iff \|\lambda\|_2 \leq 1$ and $z^T \left(D \; d \right) = \left(\alpha^T \; \beta \right) + \lambda^T \left(A \; b \right)$. □

Algorithm 2: Single-cone analysis for a second-order cone.

Data: Let $D \in \mathbb{R}^{p \times n}$, $b \in \mathbb{R}^p$ and $\mathcal{K} \subseteq \mathbb{R}^p$, such that $Dx - d \in \mathcal{K}$ is a second-order cone constraint (or variable domain) of the considered problem.

while $Dx - d$ contains a singleton variable **do**

  Substitute the singleton variable in $Dx - d$ with its unique definition in the equation system of the considered problem.

end

return a facial reduction certificate, if one is found following the instructions of Corollary 4.

The least-norm optimization problem of Proposition 7 is widely documented elsewhere and can, although alternatives exists, be solved to a fair balance between speed and accuracy by means of QR-decomposition [22]. For completeness, this approach is now established.

Proposition 8. *Consider the least-norm optimization problem of Proposition 7, and the QR-decomposition with pivoting* $(Q_1, Q_2) \left(\begin{smallmatrix} R_1 \\ 0 \end{smallmatrix} \right) = \left(\begin{smallmatrix} A, & b \end{smallmatrix} \right) P$ *for full row rank* $R_1 \in \mathbb{R}^{r \times p}$. *Then* $\lambda = Q \left(\begin{smallmatrix} \lambda' \\ 0 \end{smallmatrix} \right)$ *is a solution to* (14) *if and only if* λ' *is a solution to*

$$(R_1)^T \lambda' = -P^T \left(\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix} \right), \tag{15}$$

which can be determined by forward substitution.

Proof. The system (15) is equivalent to $PR^T \left(\begin{smallmatrix} \lambda' \\ 0 \end{smallmatrix} \right) = P(R_1)^T \lambda' = -\left(\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix} \right)$. Hence, by definition of λ and $Q^T Q = \mathrm{I}$, the system implies feasibility of λ in (14) as shown by $PR^T Q^T \lambda = (A, b)^T \lambda = -\left(\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix} \right)$. To show that λ is the unique optimal solution of (14), note that any other feasible point $\hat{\lambda}$ of (14) satisfy $(\hat{\lambda} - \lambda)^T \lambda = (\hat{\lambda} - \lambda)^T Q \left(\begin{smallmatrix} \lambda' \\ 0 \end{smallmatrix} \right) = (\hat{\lambda} - \lambda)^T Q R P^T \omega = ((A, b)^T \hat{\lambda} - (A, b)^T \lambda)^T \omega = 0$ where $R(P^T \omega) = \left(\begin{smallmatrix} \lambda' \\ 0 \end{smallmatrix} \right)$ is solvable because R_1 has full row rank. Thus, $\|\hat{\lambda}\|_2^2 = \|\hat{\lambda} - \lambda + \lambda\|_2^2 = \|\hat{\lambda} - \lambda\|_2^2 + \|\lambda\|_2^2 + 2(\hat{\lambda} - \lambda)^T \lambda = \|\hat{\lambda} - \lambda\|_2^2 + \|\lambda\|_2^2 > \|\lambda\|_2^2$ for $\lambda \neq \hat{\lambda}$. □

These results are finally be combined in Corollary 4 and realized in the facial reduction heuristic of Algorithm 2, allowing heuristic regularization of the second-order cones in (P) and (D).

Corollary 4. *Let* $D \in \mathbb{R}^{p \times n}$ *and* $b \in \mathbb{R}^p$ *such that* $Dx - d \in \mathcal{K}$ *is a second-order cone constraint. Any linear dependency between the conic entries, exposing a facial reduction, can be found analytically.*

1. *If* $\mathcal{K} = \mathcal{Q}^p$, *the system* (12) *can be solved using* $\hat{p} = (1, 0, \ldots, 0)^T \in \operatorname{relint} \mathcal{Q}^p$ *as shown by Proposition 7 and Proposition 8.*

2. *If* $\mathcal{K} = \mathcal{Q}_r^p$, *the system* (13) *can be solved as in statement 1, using* $H = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ 0 & 0 & \mathrm{I} \end{pmatrix}$, *such that* $(\mathcal{Q}_r^p)^* = H(\mathcal{Q}^p)^*$, *and* $\hat{p} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0, \ldots, 0)^T \in \operatorname{relint} \mathcal{Q}_r^p$, *such that* $H^T \hat{p} = (1, 0, \ldots, 0)$.

Note that the substitution mechanism of Algorithm 2 is rather unsophisticated, simply purging singleton variables, and improvements could possibly be made in this direction. The merits of additional substitutions are hard to quantify without executing the computations of Corollary 4 repeatedly, however, and the idea is hence left unexplored.

4.4 Heuristics based on cone factorization

Heuristics based on cone factorization exploits the fact that the equations of the auxiliary problems of Proposition 2 are homogeneous. In particular, if a cone can be put on product-form, then inner approximations of the auxiliary problems are simply obtained by dropping factors of that product-form. A simple analogy of this relation is that $ax = 0$ implies $axy = 0$ for variables $x, y \in \mathbb{R}$. Taking the semidefinite cone as example, this approach leads to the facial reduction heuristic from [13, Algorithm 1.0.2].

Example 5. *The semidefinite cone from Section 3.1 was defined in its product-form as*

$$\mathcal{S}_+^N := \left\{ VV^T : V \in \mathbb{R}^{N\times N} \right\} \subseteq \mathbb{R}^{N\times N}. \tag{16}$$

Substituting this definition into the auxiliary problem from Proposition 2-(2), and taking advantage of $VV^T \neq 0 \iff V \neq 0$, one obtain in matrix notation

$$\langle C, VV^T\rangle = \langle CV, V\rangle = 0, \qquad \begin{pmatrix} \langle A_1, VV^T\rangle \\ \vdots \\ \langle A_m, VV^T\rangle \end{pmatrix} = \begin{pmatrix} \langle A_1 V, V\rangle \\ \vdots \\ \langle A_m V, V\rangle \end{pmatrix} = 0, \qquad V \in \mathbb{R}^{N\times N} \setminus \{0\},$$

where $\langle X, Y\rangle = \mathrm{Tr}(Y^T X)$ is the trace inner product. An inner approximation is thus obtained by

$$CV = 0, \qquad \begin{pmatrix} A_1 V \\ \vdots \\ A_m V \end{pmatrix} = 0, \qquad V \in \mathbb{R}^{N\times N} \setminus \{0\},$$

from which facial reduction certificates for (D) can be computed as VV^T, for nonzero matrices V composed of columns from the common nullspace of C and A_i for $i = 1, \ldots, m$.

More generally, one may consider all proper cones factorisable into a bilinear matrix-vector product of the form

$$\mathcal{K} = \{L(v)v : v \in \mathbb{R}^n\} \subseteq \mathbb{R}^n. \tag{17}$$

The semidefinite cone from Example 5 is included in vectorized form by this characterization as will become clear, and a facial reduction heuristic for (D), based on the auxiliary problem from Proposition 2-(2), can similarly be derived.

Theorem 2. *Consider the product-form (17) of a Cartesian product of proper cones,*

$$\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_k = \left\{ L(v)v = \begin{pmatrix} L_1(v_1) & & \\ & \ddots & \\ & & L_k(v_k) \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix} : v \in \mathbb{R}^n = \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_k} \right\} \subseteq \mathbb{R}^n.$$

Substituting this definition into the auxiliary problem from Proposition 2-(2), an inner approximation can be obtained, as proven, by

$$c^T \begin{pmatrix} L_1(v^1) & & \\ & \ddots & \\ & & L_k(v^k) \end{pmatrix} = 0, \qquad A \begin{pmatrix} L_1(v^1) & & \\ & \ddots & \\ & & L_k(v^k) \end{pmatrix} = 0, \qquad v \in \mathbb{R}^n \setminus \{0\}.$$

Let $c^T = (c_1^T, \ldots, c_k^T)$ and $A = (A_1, \ldots, A_k)$ according to the Cartesian product. Facial reduction certificates for (D) can then be computed as vectors $z = L(v)v$, for nonzero $v = (v^1, \ldots, v^k) \in \mathbb{R}^n$ satisfying $c_i^T L_i(v^i) = 0$ and $A_i L_i(v^i) = 0$ for $i = 1, \ldots, k$.

Proof. The only difficult part is

$$\mathcal{K} \setminus (\mathcal{K}^*)^\perp = \mathcal{K} \setminus \{0\} = \{L(v)v : v \in \mathbb{R}^n \setminus \{0\}\}.$$

The first step holds since $\mathcal{K}^*$ is solid, as shown by the closure of $\mathcal{K}$ being pointed [10]. The second step holds since $L(v)v = 0$ only if $v = 0$, as can be shown by $\mathcal{K}$ being solid. In particular, assume that $L(\hat{v})\hat{v} = 0$ for some $\hat{v} \neq 0$. Then $L(\lambda\hat{v})(\lambda\hat{v}) = \lambda^2 L(\hat{v})\hat{v} = 0$ for all $\lambda \in \mathbb{R}$ by bilinearity. Hence, $\operatorname{span} \mathcal{K} \neq \mathbb{R}^n$ contradicting $\mathcal{K}$ being solid. □

Heuristics for (D) based on the cone factorization approach of Theorem 2 can be realized for the class of symmetric cones [44]. Formally:

Proposition 9. *The product-form* (17) *is achieved by the following cones:*

1. *The nonnegative orthant* $\mathbb{R}_+$, *representable by* $L(v) = v$;
2. *The quadratic cone* $\mathcal{Q}^n$, *representable by* $L(v) = \begin{pmatrix} v_1 & v_{2:n}^T \\ v_{2:n} & v_1 I \end{pmatrix}$;
3. *The semidefinite cone* $\operatorname{svec}(\mathcal{S}_+^N)$, *representable by* $L(v) = I \otimes_s \operatorname{smat}(v)$, *where* $v \in \mathbb{R}^{N(N+1)/2}$;
4. *The semidefinite cone* $\operatorname{vec}(\mathcal{S}_+^n)$, *representable by* $L(v) = I \otimes \operatorname{mat}(v)^T$, *where* $v \in \mathbb{R}^{N^2}$.

Proof. Statements 1-3 follow from the fact that all symmetric cones have the product-form,

$$\mathcal{K} = \{v \circ v : v \in \mathbb{R}^n\} \subseteq \mathbb{R}^n,$$

for some Euclidean Jordan algebra with associated Jordan product '$\circ$'. The Jordan product is bilinear such that $v \circ w = L(v)w$ for some symmetric linear mapping $L(v) : \mathbb{R}^n \to \mathcal{S}^{n\times n}$, and the definitions of $L(v)$ for the cones above are well known [15, 44]. The symmetric Kronecker product '$\otimes_s$' as found, e.g., in [44], satisfies $L(v)w = \operatorname{svec}(\operatorname{smat}(v)\operatorname{smat}(w) + \operatorname{smat}(w)\operatorname{smat}(v))/2$.

Statement 4 follows from the product-form (16) and the Kronecker product '$\otimes$' for which $L(v)w = \operatorname{vec}(\operatorname{mat}(v)^T \operatorname{mat}(w))$. □

It is unclear whether Theorem 2 represents any advancement from the special case of Example 5. First of all, Statement 1 is an idempotent as $A_i L(v)v = 0 \iff A_i L(v) = 0$ in Theorem 2 for all $L(v) = \lambda v$ where $\lambda > 0$. Secondly, Statement 2 and Statement 3 are restricted in usability by symmetry of $L(v)$, and Statement 4 leads to the characterization shown in Example 5. Finally, the author was unable to use the product-form (17) to construct an inner approximation for Proposition 2-(1) that was not trivially infeasible. Similar heuristics for (P), based on cone factorization, might thus not be possible.

5 Conclusion

Facial reduction is theoretically established as a useful countermeasure against ill-posedness in mixed-integer optimization which affects both conic and linear relaxations. Speedy and/or accurate facial reduction heuristics are further motivated by the slow and inaccurate alternative of having to solve conic optimization problems in each iteration of the facial reduction algorithm. This led to the development of the heuristics based on linear optimization, subgradient matching and single-cone analysis, which were shown useful in various scenarios. A fourth type of heuristic based on cone factorization may also prove useful, although it seems to be limited in applicability to dual facial reductions of semidefinite cones.

Further work is needed to test how these heuristics compare and performs in practice, such as on the instances of CBLIB [18]. Also, it remains to be investigated how these heuristics should be integrated with the brand-and-bound algorithm for greatest accuracy, greatest speed, or a balance thereof. In this regard, it is likely that local information on branching decisions, presolve changes and generated cuts can be used to guide the employment of heuristics. In a sense, subgradient matching already achieves this by integrating with domain propagation to capture facial reductions only from the changes propagated out.

Acknowledgments The author, Henrik A. Friberg, owe a special thanks to Erling D. Andersen and Mathias Stolpe who supervised the work on this paper. The author was funded by MOSEK ApS and the Danish Ministry of Higher Education and Science through the Industrial PhD project "Combinatorial Optimization over Second-Order Cones and Industrial Applications".

References

[1] T. Achterberg. *Constraint Integer Programming*. PhD thesis, Technische Universität Berlin, 2008.

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

[3] E. D. Andersen and K. D. Andersen. Presolving in linear programming. *Mathematical Programming*, 71(2):221–245, 1995.

[4] M. F. Anjos and S. Burer. On handling free variables in interior-point methods for conic linear optimization. *SIAM Journal on Optimization*, 18(4):1310–1325, 2007.

[5] M. F. Anjos and H. Wolkowicz. Strengthened semidefinite relaxations via a second lifting for the Max-Cut problem. *Discrete Applied Mathematics*, 119(1-2):79–106, 2002.

[6] P. Belotti, J. C. Goez, I. Pólik, T. K. Ralphs, and T. Terlaky. A conic representation of the convex hull of disjunctive sets and conic cuts for integer second order cone optimization. In *Numerical Analysis and Optimization*, volume 134 of *Springer Proceedings in Mathematics & Statistics*, pages 1–35. Springer International Publishing, 2015.

[7] H. Y. Benson and Ü. Sağlam. Mixed-integer second-order cone programming: A survey. In *Tutorials in Operations Research*, INFORMS, Minneapolis, 2013.

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

[9] J. M. Borwein and A. S. Lewis. Partially finite convex programming, Part II: Explicit lattice models. *Mathematical Programming*, 57(1-3):49–83, 1992.

[10] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[11] F. Burkowski, Y.-L. Cheung, and H. Wolkowicz. Efficient Use of Semidefinite Programming for Selection of Rotamers in Protein Conformations. *INFORMS Journal on Computing*, 26(4):748–766, 2014.

[12] Y.-L. Cheung. *Preprocessing and Reduction for Semidefinite Programming via Facial Reduction: Theory and Practice*. PhD thesis, University of Waterloo, 2013. URL `http://orion.uwaterloo.ca/~hwolkowi/henry/reports/thesisvrisnov13.pdf`.

[13] Y.-L. Cheung, S. Schurr, and H. Wolkowicz. Preprocessing and Regularization for Degenerate Semidefinite Programs. In *Computational and Analytical Mathematics*, volume 50 of *Springer Proceedings in Mathematics & Statistics*, pages 251–303. Springer New York, 2013.

[14] D. Drusvyatskiy, G. Pataki, and H. Wolkowicz. Coordinate shadows of semi-definite and euclidean distance matrices. *arXiv*, 2015. URL `http://arxiv.org/abs/1405.2037`.

[15] J. Faraut and A. Korányi. *Analysis on symmetric cones*. 1994.

[16] R. M. Freund and J. R. Vera. Some characterizations and properties of the "distance to ill-posedness" and the condition measure of a conic linear system. *Mathematical Programming*, 86(2):225–260, 1999.

[17] H. A. Friberg. A relaxed-certificate facial reduction algorithm based on subspace intersection. *Optimization Online*, 2016. URL `http://www.optimization-online.org/DB_HTML/2016/01/5283.html`.

[18] H. A. Friberg. CBLIB 2014: a benchmark library for conic mixed-integer and continuous optimization. *Mathematical Programming Computation*, Online 2015.

[19] F. Glineur. Second-order cone optimization with a single second-order cone. In *18th International Symposium on Mathematical Programming (ISMP)*, 2003.

[20] G. Gruber and F. Rendl. Computational experience with ill-posed problems in semidefinite programming. *Computational Optimization and Applications*, 21(2):201–212, 2002.

[21] G. Gruber, S. Kruk, F. Rendl, and H. Wolkowicz. Presolving for Semidefinite Programs Without Constraint Qualifications. Technical Report CORR 98-32, University of Waterloo, Waterloo, Ontario, 1998.

[22] N. J. Higham. Underdetermined Systems. In *Accuracy and Stability of Numerical Algorithms.* Society for Industrial and Applied Mathematics, 2nd edition, 2002.

[23] H. Hijazi and L. Liberti. Constraint Qualification Failure in Second-Order Cone Formulations of Unbounded Disjunctions. Technical report, NICTA, Canberra ACT Australia, 2014.

[24] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, 1985.

[25] F. Kılınç-Karzan and S. Yıldız. Two-Term Disjunctions on the Second-Order Cone. In *Integer Programming and Combinatorial Optimization*, volume 8494 of *Lecture Notes in Computer Science*, pages 345–356. Springer International Publishing, 2014.

[26] T. Koch. The final NETLIB-LP results. *Operations Research Letters*, 32(2):138–142, 2004.

[27] N. Krislock and H. Wolkowicz. Explicit sensor network localization using semidefinite representations and clique reductions. *SIAM Journal on Optimization*, 20(5):2679–2708, 2010.

[28] B. F. Lourenço, M. Muramatsu, and T. Tsuchiya. Solving SDP Completely with an Interior Point Oracle. *Optimization Online*, 2015. URL `http://www.optimization-online.org/DB_HTML/2015/06/4982.html`.

[29] Z.-Q. Luo, J. F. Sturm, and S. Zhang. Duality Results for Conic Convex Programming. Technical Report 9719/A, Econometric Institute, Erasmus University Rotterdam, The Netherlands, 1997.

[30] S. Mars and L. Schewe. An SDP-Package for SCIP. Technical Report August, TU Darmstadt, 2012. URL `http://www.opt.tu-darmstadt.de/~smars/scip_sdp.html`.

[31] A. Nemirovski. Efficient methods in convex programming. Technion - The Israel Institute Of Technology Faculty Of Industrial Engineering & Management, 1994.

[32] G. Pataki. Strong duality in conic linear programming: facial reduction and extended duals. In *Springer Proceedings in Mathematics and Statistics*, volume 50, pages 613–634. Springer New York, 2013.

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

[34] F. Permenter, H. A. Friberg, and E. D. Andersen. Solving conic optimization problems via self-dual embedding and facial reduction : a unified approach. Technical report, 2015. `http://www.optimization-online.org/DB_HTML/2015/09/5104.html`.

[35] J. Renegar. Incorporating Condition Measures into the Complexity Theory of Linear Programming. *SIAM Journal on Optimization*, 5(3):506–524, 1995.

[36] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

[37] M. W. P. Savelsbergh. Preprocessing and probing techniques for mixed integer programming problems. *ORSA Journal on Computing*, 6(4):445–454, 1994.

[38] J. Stoer and C. Witzgall. *Convexity and Optimization in Finite Dimensions I.* Springer Verlag, Berlin, 1970.

[39] M. J. Todd. Semidefinite optimization. *Acta Numerica*, 10(C):515–560, 2001.

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

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

[42] 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, 2011.

[43] H. Wolkowicz and Q. Zhao. Semidefinite programming relaxations for the graph partitioning problem. *Discrete Applied Mathematics*, 96-97:461–479, 1999.

[44] H. Wolkowicz, R. Saigal, and L. Vandenberghe. Handbook of Semidefinite Programming. 2000.

[45] M. Yamashita, K. Fujisawa, K. Nakata, M. Nakata, M. Fukuda, K. Kobayashi, and K. Goto. A high-performance software package for semidefinite programs: SDPA 7. Technical report, 2010.

[46] Y. Ye, M. J. Todd, and S. Mizuno. An $\mathcal{O}(\sqrt{n}L)$-Iteration Homogeneous and Self-Dual Linear Programming Algorithm, 1994.

[47] Q. Zhao, S. E. Karisch, F. Rendl, and H. Wolkowicz. Semidefinite Programming Relaxations for the Quadratic Assignment Problem. *Journal of Combinatorial Optimization*, 2(1):71–109, 1998.