

Preprocessing and Reduction for Degenerate Semidefinite Programs

Yuen-Lam Cheung ^{*} Simon Schurr [†] Henry Wolkowicz [‡]

February 18, 2011

University of Waterloo
Department of Combinatorics & Optimization
Waterloo, Ontario N2L 3G1, Canada
Research Report CORR 2011-02

Key words and phrases: Preprocessing, semidefinite programming, degeneracy, strong duality.

AMS subject classifications: 90C46, 90C22, 90C25, 49K40, 65K10

Abstract

This paper presents a backward stable preprocessing technique for (nearly) ill-posed semidefinite programming, SDP, problems, i.e., programs for which Slater’s constraint qualification, existence of strictly feasible points, (nearly) fails.

Current popular algorithms for semidefinite programming rely on *primal-dual interior-point*, *p-d i-p* methods. These algorithms require Slater’s constraint qualification for both the primal and dual problems. This assumption guarantees the existence of Lagrange multipliers, well-posedness of the problem, and stability of algorithms. However, there are many instances of SDPs where Slater’s constraint qualification fails or *nearly* fails. Our backward stable preprocessing technique is based on finding a *rank-revealing* rotation of the problem followed by facial reduction. This results in a smaller, well-posed, *nearby* problem that can be solved by standard SDP solvers.

Contents

1	Introduction	3
1.1	Outline	4
1.2	Preliminary definitions	4

^{*}Department of Combinatorics and Optimization, University of Waterloo, Ontario N2L 3G1, Canada Research supported by TATA Consultancy Services.

[†]Research supported by The Natural Sciences and Engineering Research Council of Canada. Email: schurr@math.uwaterloo.ca

[‡]Research supported by The Natural Sciences and Engineering Research Council of Canada. Email: hwolkowicz@uwaterloo.ca

⁰ URL for paper: orion.math.uwaterloo.ca/~hwolkowi/henry/reports/ABSTRACTS.html

20	2 Theory	5
21	2.1 Strong duality for cone optimization	5
22	2.2 Theorems of the alternative	7
23	2.3 Stable auxiliary subproblem	9
24	2.4 Rank-revealing rotation and equivalent problems	12
25	2.5 LP: an example	16
26	3 Facial Reduction	17
27	4 Sensitivity Analysis	19
28	4.1 Rank-revealing rotation and the minimal face $QS_+^{n-1}Q^T$	19
29	4.2 Equivalent problems	22
30	4.3 Reduction to a smaller problem	24
31	4.3.1 The $\delta^* = 0$ case	24
32	4.3.2 The $\delta^* > 0$ case	24
33	5 Numerical Results	28
34	5.1 Worst case instance	28
35	5.2 Generating instances with finite nonzero duality gaps	28
36	5.3 Numerical results from SeDuMi	33
37	6 Conclusions and future work	34
38	A Proof of Corollary 4.2	34
39	Bibliography	36
40	Index	40
41	List of Algorithms	
42	3.1 One iteration of facial reduction	18
43	3.2 Solving (AP)	19
44	List of Tables	
45	1 Comparisons with/without facial reduction	33
46	List of Figures	
47	1 Minimal Face; $0 < \delta^* \ll 1$	24

48 **1 Introduction**

49 The aim of this paper is to develop a backward stable preprocessing technique to handle (nearly)
 50 ill-posed semidefinite programming, SDP, problems, i.e., programs for which Slater’s constraint
 51 qualification, existence of strictly feasible points, (nearly) fails. The technique is based on finding
 52 a *rank-revealing* rotation of the problem followed by *facial reduction*. This results in a smaller,
 53 well-posed, approximately equivalent problem that can be solved by standard SDP solvers.

In particular, we study SDPs of the following form

$$(P) \quad v_P := \sup\{b^T y : \mathcal{A}^* y \preceq c\}, \tag{1.1}$$

54 where the optimal value v_P is finite, $b \in \mathbb{R}^m$, $c \in \mathbb{S}^n$, and $\mathcal{A} : \mathbb{S}^n \rightarrow \mathbb{R}^m$ is an onto linear
 55 transformation from the space \mathbb{S}^n of $n \times n$ real symmetric matrices to \mathbb{R}^m . The adjoint of \mathcal{A} is
 56 $\mathcal{A}^* y = \sum_{i=1}^m y_i A_i$, where $A_i \in \mathbb{S}^n, i = 1, \dots, m$. The symbol \preceq denotes the Löwner partial order
 57 induced by the cone \mathbb{S}_+^n of positive semidefinite matrices, i.e., $\mathcal{A}^* y \preceq c \iff c - \mathcal{A}^* y \in \mathbb{S}_+^n$.¹ If (P)
 58 is *strictly feasible*, then one can use standard solution techniques; if (P) is *strongly infeasible*, then
 59 one can set $v_P = -\infty$, e.g., [20, 25, 38, 39]. If neither of these two feasibility conditions can be
 60 verified, then we apply our preprocessing technique that finds a rotation of the problem that is akin
 61 to *rank-revealing* matrix rotations. (See e.g., [34, 35] for equivalent matrix results.) This rotation
 62 finds an equivalent (nearly) block diagonal problem which allows for simple strong dualization by
 63 solving only the most significant block of (P), for which Slater’s condition (strict feasibility) holds.
 64 This is equivalent to restricting the original problem to a face of \mathbb{S}_+^n , i.e., the preprocessing can be
 65 considered as a *facial reduction* of (P). Moreover, it provides a *backward stable* approach for solving
 66 (P) when it is feasible and the Slater constraint qualification (CQ) fails; and it solves a nearby
 67 problem when (P) is *weakly infeasible*.

The Lagrangian dual to (1.1) is

$$(D) \quad v_D := \inf \{ \langle c, x \rangle : \mathcal{A}x = b, x \succeq 0 \}, \tag{1.2}$$

68 where $\langle c, x \rangle := \text{trace } cx = \sum_{ij} c_{ij} x_{ij}$ denotes the trace inner product of the symmetric matrices c
 69 and x ; and, $\mathcal{A}x = (\langle A_i, x \rangle) \in \mathbb{R}^m$. Weak duality $v_D \geq v_P$ follows easily, e.g., for any primal-dual
 70 feasible points y, x , we get $\langle c, x \rangle \geq \langle \mathcal{A}^* y, x \rangle = \langle y, \mathcal{A}x \rangle = \langle y, b \rangle$. The usual constraint qualification
 71 (CQ) used for (P) is the Slater condition, i.e., strict feasibility $\mathcal{A}^* y \prec c$ (or $c - \mathcal{A}^* y \in \mathbb{S}_{++}^n$, the
 72 cone of positive definite matrices). If we assume Slater’s CQ holds and the primal optimal value
 73 is finite, then strong duality holds, i.e., we have a zero duality gap and attainment of the dual
 74 optimal value. Strong duality results for (1.1) without any constraint qualification are given in
 75 [5, 6, 7, 43] and [27, 28], and more recently in [29, 40]. Related closure conditions appear in [23];
 76 and, properties of problems where strong duality fails appear in [22].

77 Many popular algorithms for (P) are based on Newton’s method and a *primal-dual interior-*
 78 *point, p-d i-p*, approach, e.g., the codes (latest at the URLs in the citations) CSDP, SeDuMi,
 79 SDPT3, SDPA [4, 36, 42, 47]; see also the
 80 SDP URL: www-user.tu-chemnitz.de/~helmberg/sdp_software.html.

81 To find the search direction, these algorithms apply symmetrization in combination with block
 82 elimination to find the Newton search direction. The symmetrization and elimination steps both

¹Note that the cone optimization problem (1.1) is commonly used as the dual problem in the SDP literature, though it is often the primal in the Linear Matrix Inequality (LMI) literature, e.g., [8].

83 result in ill-conditioned linear systems, even for well conditioned SDP problems, e.g., [9, 44]. And,
 84 these methods are very susceptible to numerical difficulties and high iteration counts in the case
 85 when Slater's CQ nearly fails, see e.g., [10, 11, 12, 13]. Our aim in this paper is to provide
 86 a regularization process for problems where strict feasibility (almost) fails. Related papers on
 87 regularization are e.g., [16, 21]; and papers on high accuracy solutions for algorithms SDPA-GMP,-
 88 QD,-DD are e.g., [48].

89 1.1 Outline

90 We continue in Section 1.2 with preliminary notation and results for cone programming. Section
 91 2 presents the theoretical background and tools needed for the facial reduction algorithm. This
 92 includes results on strong duality and theorems of the alternative in Section 2.1. A stable auxiliary
 93 problem for identifying faces containing the feasible set is presented and studied in Section 2.3. An
 94 outline of the facial reduction using a rank-revealing rotation process is given in Section 3.

95 Sensitivity analysis and backward stability results are presented in Section 4. Preliminary
 96 numerical tests, as well as a technique for generating instances with a finite duality gap useful
 97 for numerical tests, are given in Section 5. We conclude with remarks in Section 6. (An index is
 98 included to help the reader, see page 40.)

99 1.2 Preliminary definitions

Let K be a (closed) convex cone, i.e., $\lambda K \subseteq K, \forall \lambda \geq 0$, and $K + K \subseteq K$. K is *pointed* if
 $K \cap (-K) = \{0\}$; K is *proper* if K is pointed and $\text{int } K \neq \emptyset$; the polar or dual cone of K is
 $K^* = \{\phi : \langle \phi, k \rangle \geq 0, \forall k \in K\}$. We denote by \preceq_K the partial order with respect to K . That
 is, $x_1 \preceq_K x_2$ means that $x_2 - x_1 \in K$. We also write $x_1 \prec_K x_2$ to mean that $x_2 - x_1 \in \text{int } K$. In
 particular, $K = \mathbb{S}_+^n$ yields the partial order induced by the cone of positive semidefinite matrices
 in \mathbb{S}^n , i.e., the so-called Löwner partial order. We denote this simply with $x \preceq y$ for $y - x \in \mathbb{S}_+^n$.

$\text{cone}(S)$ denotes the convex cone generated by the set S . In particular, for any non-zero vector
 x , the *ray generated by* x is defined by $\text{cone}(x)$. The ray generated by $s \in K$ is called an *extreme*
ray if $0 \preceq_K u \preceq_K s$ implies that $u \in \text{cone}(s)$. The subset $F \subseteq K$ is a *face of the cone* K , denoted
 $F \trianglelefteq K$, if

$$(s \in F, 0 \preceq_K u \preceq_K s) \implies (\text{cone}(u) \subseteq F). \quad (1.3)$$

100 Equivalently, $F \trianglelefteq K$ if F is a cone and $(x, y \in K, \frac{1}{2}(x + y) \in F) \implies (\{x, y\} \subseteq F)$. If $F \trianglelefteq K$ but
 101 is not equal to K , we write $F \triangleleft K$. If $0 \neq F \triangleleft K$, then F is a *proper face* of K . (Similarly, $S_1 \subset S_2$
 102 denotes a proper subset, i.e., $S_1 \subseteq S_2, S_1 \neq S_2$.) For $S \subseteq K$, we let $\text{face}(S)$ denote the smallest face
 103 of K that contains S . A face $F \trianglelefteq K$ is an *exposed face* if it is the intersection of K with a hyperplane.
 104 The cone K is *facially exposed* if every face $F \trianglelefteq K$ is exposed. If $F \trianglelefteq K$, then the *conjugate face*
 105 is $F^c := K^* \cap \{F\}^\perp$. Note that the conjugate face F^c is *exposed* using any $s \in \text{relint } F$ (where
 106 $\text{relint } S$ denotes the *relative interior* of set S), i.e., $F^c = K^* \cap \{s\}^\perp, \forall s \in \text{relint } F$. In addition, note
 107 that \mathbb{S}_+^n is self-dual (self-polar) and it is facially exposed. For the general cone problem, the linear
 108 transformation \mathcal{A} maps between two Euclidean spaces $\mathcal{A} : \mathcal{V} \rightarrow \mathcal{W}$, and $K \subseteq \mathcal{V}$.

Denote the feasible solution and slack sets of (1.1) and (1.2) by $\mathcal{F}_P = \mathcal{F}_P^y = \{y : A^*y \preceq c\}$,
 $\mathcal{F}_P^z = \{z : z = c - A^*y \succeq 0\}$, and $\mathcal{F}_D = \{x : Ax = b, x \succeq 0\}$, respectively. The *minimal face* of
 (1.1) is the intersection of all faces of K containing the feasible slack vectors:

$$f_P := \cap \{H \trianglelefteq K : c - \mathcal{A}^*(\mathcal{F}_P) \subseteq H\} = \text{face}(c - \mathcal{A}^*(\mathcal{F}_P)).$$

109 Here, $\mathcal{A}^*(\mathcal{F}_P)$ is the linear image of the set \mathcal{F}_P under \mathcal{A}^* .

We denote the *triangular number* $t(n) := n(n+1)/2$, and for $S \in \mathbb{S}^n$, we define $\text{s2vec}(S) \in \mathbb{R}^{t(n)}$ to be the vector formed, columnwise, from the upper-triangular part of S where the strict upper-triangular part is multiplied by $\sqrt{2}$. Therefore, s2vec is an isometry from \mathbb{S}^n to $\mathbb{R}^{t(n)}$. The inverse (and adjoint) is $\text{s2Mat} = \text{s2vec}^* = \text{s2vec}^{-1}$. We let e_i denote the i -th unit vector and $E_{ij} := \frac{1}{\sqrt{2}}(e_i e_j^T + e_j e_i^T)$ are the unit matrices in \mathbb{S}^n . The linear transformation $\mathcal{A} : \mathbb{S}^n \rightarrow \mathbb{R}^m$ can be expressed as

$$\mathcal{A}x = (\text{trace } A_i x) = (\text{s2vec}(A_i)^T \text{s2vec}(x)) \in \mathbb{R}^m,$$

for specific $A_i \in \mathbb{S}^n, i = 1, \dots, m$. We let $\|\mathcal{A}\|_2$ denote the spectral norm of \mathcal{A} and define the Frobenius norm (Hilbert-Schmidt norm) of \mathcal{A} as $\|\mathcal{A}\|_F := \sqrt{\sum_{i=1}^m \|A_i\|_F^2}$. Alternatively, if $A^T = [\text{s2vec}(A_1) \ \dots \ \text{s2vec}(A_m)]$, i.e., A is the matrix with rows $\text{s2vec}(A_i)^T$, then

$$\|\mathcal{A}\|_2 = \sigma_{\max}(A), \quad \text{cond}(\mathcal{A}) = \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)}, \quad \|\mathcal{A}\|_F = \sqrt{\sum_{i=1}^m \sigma_i^2(A)} = \sqrt{\text{trace } A^T A},$$

where σ_i denotes the i -th singular value. Unless stated otherwise, all matrix norms in this paper are Frobenius norms. Then, e.g., [17, Chapter 5], for any $s \in \mathbb{S}^n$,

$$\|\mathcal{A}s\|_2 \leq \|\mathcal{A}\|_2 \|s\|_F \leq \|\mathcal{A}\|_F \|s\|_F. \quad (1.4)$$

110 All vector norms are assumed to be 2-norm, unless stated otherwise.

111 The following easily-proven result presents the effect of shifting the dual objective function
112 coefficients c by an element from $\mathcal{R}(\mathcal{A}^*)$. This allows us to shift so that $c \in \mathcal{N}(\mathcal{A}^*)$ (done below).

113 **Lemma 1.1.** *Suppose that the optimal value v_P for (1.1) is finite, and let $\bar{y} \in \mathbb{R}^m$. If c is replaced
114 by $\bar{c} = c - \mathcal{A}^*\bar{y}$, then the optimal value is shifted to $v_P - \langle b, \bar{y} \rangle$. Moreover, if v_P is attained in (1.1)
115 at the optimum y^* , then the optimum is shifted to $w^* = y^* - \bar{y}$. ■*

116

117 We summarize our assumptions in the following.

118 **Assumption 1.2.** $\mathcal{F}_P \neq \emptyset$; \mathcal{A} is surjective.

119 2 Theory

120 We now present the theoretical tools that are needed for the facial reduction algorithm. This
121 includes the well known results for strong duality, the theorems of the alternative to identify strict
122 feasibility, a stable subproblem to apply the theorems of the alternative, and the backward error
123 analysis for the output of the stable subproblem.

124 2.1 Strong duality for cone optimization

125 We first summarize some results on *strong duality* for the conic convex problem in the form (1.1).
126 Strong duality for (1.1) means that there is a *zero duality gap*, $v_P = v_D$, and the dual optimal
127 value v_D (1.2) is attained. However, it is easy to construct examples where strong duality fails, see
128 e.g., [22, 28, 45] and Section 5, below.

129 It is well known that, in finite dimensional linear optimization (LP), strong duality fails only
130 if the primal problem and/or its dual are infeasible. In fact, in LP both problems are feasible and
131 both of the optimal values are attained (and equal) if, and only if, the optimal value of one of
132 the problems is finite. In general (conic) convex optimization the situation is more complicated,
133 since the underlying cone in the primal and dual optimization problems need not be polyhedral.
134 Consequently, even if a primal problem and its dual are feasible, a nonzero duality gap and/or
135 non-attainment of the optimal values may ensue unless some *constraint qualification* holds; see
136 e.g., [3, 31]. More specific examples for our cone situations appear in e.g., [20], [30, Section 3.2],
137 and [37, Section 4]. Applications that exploit the failure of Slater’s CQ for SDP relaxations appear
138 in e.g., [1, 2, 19, 46, 51].

Failure of strong duality is problematic, since many classes of p-d i-p algorithms require not
only that a primal-dual pair of problems possess a zero duality gap, but also that the (generalized)
Slater constraint qualification holds for both primal and dual, i.e., that strict feasibility holds for
both problems. In [5, 6, 7], an equivalent *strongly dualized primal problem*

$$(SP) \quad v_P = v_{SP} := \sup\{\langle b, y \rangle : \mathcal{A}^*y \preceq_{f_P} c\} \quad (2.1)$$

is considered. Its Lagrangian dual is given by

$$(DSP) \quad v_{DSP} := \inf\{\langle c, x \rangle : \mathcal{A}x = b, x \succeq_{f_P^*} 0\}. \quad (2.2)$$

139 **Theorem 2.1** ([5]). *Suppose that the optimal value v_P is finite. Then strong duality holds for the*
140 *pair (2.1) and (2.2), or equivalently, for the pair (1.1) and (2.2); i.e., $v_P = v_{DSP}$ and the dual*
141 *optimal value v_{DSP} is attained. ■*

142

143 **Remark 2.2.** *We can also close the duality gap from the dual side by writing a strongly dualized*
144 *problem involving the minimal face of (D) in (1.2), and using the Lagrangian dual of this strongly*
145 *dualized problem. Alternatively, we can transform (D) into the form (P). This is best done using the*
146 *subspace formulations, see e.g., [40]. Define a one-one linear transformation \mathcal{K} so that the range*
147 *of \mathcal{K} and nullspace of \mathcal{A} satisfy $\mathcal{R}(\mathcal{K}) = \mathcal{N}(\mathcal{A})$; and, let \hat{x} be feasible for (D), $\mathcal{A}\hat{x} = b, \hat{x} \succeq_K 0$.*
148 *Then $(\mathcal{A}x = b, x \succeq_K 0)$ if, and only if, $0 \preceq_K x = \hat{x} + \mathcal{K}(y)$, for some y . Thus we can rewrite the*
149 *constraint in (D) in the form (P) as $-\mathcal{K}(y) \preceq_K \hat{x}$. Moreover, one can strongly dualize the problem*
150 *using a minimal representation for the linear transformation \mathcal{A} instead of (or together with) using*
151 *the minimal face of K , see [40].*

152 *An equivalent approach to the strongly dualized program (SP) in the SDP case is given using*
153 *an extended Lagrange dual (ELSD), [27, 28] and [29]. Similarly, there is an extended Lagrange*
154 *dual (ELSD) from the dual side. Whereas the strongly dualized primal (SP) and dual (DSD) are*
155 *Lagrangian duals of each other, this is not the case for (ELSD) and the primal (P) or (SP). As*
156 *a result, until this current work, it was not clear how to use interior-point methods in an effective*
157 *and efficient manner on (ELSD), e.g., [26].*

158 *In addition, we note that when applying the algorithm to construct the minimal face, in order to*
159 *ensure direct correspondence between (DSP) and (ELSD), points in the relative interior of various*
160 *faces of the dual cone K^* need to be computed. This was mentioned in [29, p. 650].*

161 2.2 Theorems of the alternative

162 Our reduction process is based on the following lemma that specializes [7, Theorem 7.1].²

163 **Lemma 2.3** ([7]). *Suppose that $\text{int}(K) \neq \emptyset$ and $\mathcal{F}_P \neq \emptyset$. Then exactly one of the following two*
 164 *systems is consistent:*

165 1. $\mathcal{A}d = 0, 0 \neq d \succeq_{K^*} 0$, and $\langle c, d \rangle = 0$.

166 2. $\mathcal{A}^*y \prec_K c$ (Slater constraint qualification).

167 *Proof.* We modify the proof from [7] for our special case. Suppose that $\hat{y} \in \mathcal{F}_P$. Define the
 168 vector-valued and real-valued functions $g(y) := \mathcal{A}^*y - c$ and $g_d(y) := \langle d, \mathcal{A}^*y - c \rangle$. Note that
 169 $\nabla_y g_d(y) = \mathcal{A}d$.

170 Suppose that d satisfies the system in Item 1. Then $\nabla_y g_d(\hat{y}) = \mathcal{A}d = 0$ and $g_d(\hat{y}) = \langle \mathcal{A}d, \hat{y} \rangle -$
 171 $\langle d, c \rangle = 0 \forall y$. Therefore \hat{y} is a global minimizer of the linear function $g_d(y)$, i.e., $\langle d, \mathcal{A}^*(\mathbb{R}^m) - c \rangle = 0$.
 172 Since $0 \neq d \succeq_{K^*} 0$, this implies that the Slater CQ fails.

Conversely, suppose that the Slater CQ in Item 2 fails. We have $\text{int} K \neq \emptyset$ and

$$0 \notin (\mathcal{A}^*(\mathbb{R}^m) - c) + \text{int} K.$$

Therefore, we can find $d \neq 0$ to separate the set from 0, i.e.,

$$\langle d, (\mathcal{A}^*(\mathbb{R}^m) - c) + \text{int} K \rangle \geq 0.$$

173 Therefore, \hat{y} is again a global minimizer of $g_d(y)$, and the optimality conditions imply that this d
 174 satisfies the conditions in Item 1. \square

175 We have an equivalent characterization for the generalized Slater condition for the dual problem.
 176 This can be used to extend our results to (D).

177 **Corollary 2.4.** *Suppose that $\text{int}(K^*) \neq \emptyset$ and $\mathcal{F}_D \neq \emptyset$. Then exactly one of the following two*
 178 *systems is consistent:*

179 1. $0 \neq \mathcal{A}^*v \succeq_K 0$, and $\langle b, v \rangle = 0$.

180 2. $\mathcal{A}x = b, x \succ_{K^*} 0$ (generalized Slater constraint qualification).

181 *Proof.* Let \mathcal{K} be a one-one linear transformation with range $\mathcal{R}(\mathcal{K}) = \mathcal{N}(\mathcal{A})$. Then, system 2 is
 182 consistent if, and only if, there exists \hat{u}, \hat{x} such that $\mathcal{A}\hat{x} = b$ and $x = \hat{x} + \mathcal{K}\hat{u} \succ_{K^*} 0$. This is
 183 equivalent to $-\mathcal{K}\hat{u} \prec_{K^*} \hat{x}$. Therefore, $-\mathcal{K}, \hat{x}$ play the roles of \mathcal{A}^*, c , respectively, in Lemma 2.3.
 184 Therefore, an alternative system is $-\mathcal{K}^*z = 0, 0 \neq z \succeq_K 0$, and $\langle \hat{x}, z \rangle = 0$. Since $\mathcal{N}(\mathcal{K}^*) = \mathcal{R}(\mathcal{A}^*)$,
 185 this is equivalent to $0 \neq z = \mathcal{A}^*v \succeq_K 0$, and $\langle \hat{x}, z \rangle = 0$, or $0 \neq \mathcal{A}^*v \succeq_K 0$, and $\langle b, v \rangle = 0$. \square

186 We can extend the above Lemma 2.3 to the cases where $\text{int} K = \emptyset$. We let \mathcal{T}^\dagger denote the
 187 Moore-Penrose generalized inverse of \mathcal{T} .

²The result in [7, Theorem 7.1] involves a feasible set of the form $\{x \in \Omega : g(x) \preceq_S 0\} \neq \emptyset$, where Ω is a convex set, S is a convex cone, and g is a S -convex function. Taking Ω to be the whole space \mathbb{R}^m , $S = \mathbb{S}_+^n$ and $g(y) = \mathcal{A}^*y - c$ yields the feasible set of (1.1). In addition, the objective function in [7] was minimized rather than maximized.

Corollary 2.5. *Suppose that $\text{int } K = \emptyset$, $\mathcal{F}_P \neq \emptyset$, and $c \in \text{span}(K)$. Then the linear manifold*

$$\mathbb{S}_y := \{y \in \mathcal{W} : c - \mathcal{A}^*y \in \text{span}(K)\}$$

is a subspace. Moreover, let \mathcal{P} be a one-one linear transformation with

$$\mathcal{R}(\mathcal{P}) = (\mathcal{A}^*)^\dagger(K - K).$$

188 *Then exactly one of the following two systems is consistent:*

189 1. $\mathcal{P}^* \mathcal{A}d = 0$, $0 \neq d \succeq_{K^*} 0$, $d \in \text{span}(K)$, and $\langle c, d \rangle = 0$.

190 2. $c - \mathcal{A}^*y \in \text{relin } K$ (*generalized Slater constraint qualification*).

191 *Proof.* Since $c \in \text{span}(K)$, we get that $0 \in \mathbb{S}_y$, i.e., \mathbb{S}_y is a subspace.

Let \mathcal{T} denote an onto linear transformation acting on \mathcal{V} such that the nullspace $\mathcal{N}(\mathcal{T}) = \text{span}(K)^\perp$, and \mathcal{T}^* is an orthogonal linear transformation, i.e., $\mathcal{T}^* = \mathcal{T}^\dagger$. Therefore, \mathcal{T} is one-to-one onto $\text{span}(K)$. Then

$$\begin{aligned} \mathcal{A}^*y \preceq_K c &\iff \mathcal{A}^*y \preceq_K c \text{ and } \mathcal{A}^*y \in K - K, && \text{since } c \in K - K \\ &\iff (\mathcal{A}^*\mathcal{P})w \preceq_K c, y = \mathcal{P}w, \text{ for some } w, && \text{by definition of } \mathcal{P} \\ &\iff (\mathcal{T}\mathcal{A}^*\mathcal{P})w \preceq_{\mathcal{T}(K)} \mathcal{T}c, y = \mathcal{L}w, \text{ for some } w, && \text{by definition of } \mathcal{T}, \end{aligned}$$

i.e., (1.1) is equivalent to

$$v_P := \sup\{\langle \mathcal{P}^*b, w \rangle : (\mathcal{T}\mathcal{A}^*\mathcal{P})w \preceq_{\mathcal{T}(K)} \mathcal{T}c\}.$$

The corresponding dual is

$$v_D := \inf\{\langle \mathcal{T}c, d \rangle : (\mathcal{P}^* \mathcal{A} \mathcal{T}^*)d = \mathcal{P}^*b, d \succeq_{(\mathcal{T}(K))^*} 0\}.$$

192 By construction, $\text{int } \mathcal{T}(K) \neq \emptyset$, so we may apply Lemma 2.3. We conclude that exactly one of
193 the following two systems is consistent:

194 1. $(\mathcal{P}^* \mathcal{A} \mathcal{T}^*)d = 0$, $0 \neq d \succeq_{(\mathcal{T}(K))^*} 0$, and $\langle \mathcal{T}c, d \rangle = 0$.

195 2. $(\mathcal{T}\mathcal{A}^*\mathcal{P})w \prec_{\mathcal{T}(K)} \mathcal{T}c$ (*Slater constraint qualification*).

196 The required result follows, since we can now identify \mathcal{T}^*d with $d \in \text{span}(K)$, and $\mathcal{T}c$ with c . \square

Definition 2.6. *The set of recession directions for (1.2) is*

$$\mathcal{R}_D := \{d : \mathcal{A}d = 0, d \succeq_{K^*} 0, \langle c, d \rangle = 0\}. \quad (2.3)$$

197 Of course \mathcal{R}_D consists of feasible directions along which the dual objective function is constant.
198 Note that $\mathcal{R}_D \setminus \{0\}$ is the set of points satisfying condition 1 in Lemma 2.3.³

Lemma 2.7. *Suppose that the feasible set $\mathcal{F}_P \neq \emptyset$ for (1.1), and let $0 \neq d \in \mathcal{R}_D$. Then the minimal face of (1.1) satisfies*

$$f_P \trianglelefteq K \cap \{d\}^\perp \triangleleft K.$$

³The set \mathcal{R}_D for $K = \mathbb{S}_+^n$ is denoted \mathcal{U}_1 in [29].

Proof. We have

$$0 = \langle c, d \rangle = \langle c, d \rangle - \langle \mathcal{F}_P, \mathcal{A}d \rangle = \langle c - \mathcal{A}^*(\mathcal{F}_P), d \rangle.$$

199 Hence $c - \mathcal{A}^*(\mathcal{F}_P) \in \{d\}^\perp$. It follows that $f_P \subset \{d\}^\perp$, or equivalently, $f_P \subset \{d\}^\perp \cap K$. The required
200 result now follows from the fact that f_P is (by definition) a face of K , and d is nonzero. \square

201 **Remark 2.8.** Ideally, we would like to find $\hat{d} \in \text{relint}(\mathcal{F}_P^z)^c = ((c + \mathcal{R}(\mathcal{A}^*)) \cap K)^c$, since then we
202 have found the minimal face $f_P = \{\hat{d}\}^\perp \cap K$. This is difficult to do numerically. Instead, Lemma
203 2.3 compromises and finds a point in a larger set $d \in (\mathcal{N}(\mathcal{A}) \cap \{c\}^\perp \cap K^*) \setminus \{0\}$. This allows for the
204 reduction of $K \leftarrow K \cap \{d\}^\perp$. Repeating to find another d is difficult without the subspace reduction
205 using \mathcal{P} in Corollary 2.5. This emphasizes the importance of the minimal subspace form reduction
206 as an aid to the minimal cone reduction, [40].

207 A similar argument applies to the regularization of the dual as given in Corollary 2.4. Let
208 $\mathcal{F}_D = \hat{x} + \mathcal{N}(\mathcal{A}) \cap K^*$, where $\mathcal{A}\hat{x} = b$. We note that a compromise to finding $\hat{z} \in \text{relint}(\mathcal{F}_P^z)^c =$
209 $((\hat{x} + \mathcal{N}(\mathcal{A})) \cap K^*)^c$, $f_D = \{\hat{z}\}^\perp \cap K^*$ is finding $z \in (\mathcal{R}(\mathcal{A}^*) \cap \{\hat{x}\}^\perp \cap K) \setminus \{0\}$, where $0 = \langle z, \hat{x} \rangle =$
210 $\langle \mathcal{A}^*v, \hat{x} \rangle = \langle v, b \rangle$.

211 2.3 Stable auxiliary subproblem

212 Each iteration of the facial reduction algorithm involves two steps. First, we apply Lemma 2.3 and
213 find a point d in the relative interior of \mathcal{R}_D . Then, we project onto the span of the conjugate face
214 $\{d\}^\perp \cap K \supseteq f_p$. This yields a smaller dimensional equivalent problem. The first step to find d is
215 well-suited for interior-point algorithms if we can formulate a suitable conic optimization problem.
216 We now formulate and present the properties of a stable auxiliary problem for finding d . The
217 following is well-known, e.g., [45].

218 **Theorem 2.9.** If the (generalized) Slater CQ holds for both primal and dual problems, then as the
219 barrier parameter $\mu \rightarrow 0^+$, the primal-dual central path converges to a point $(\hat{x}, \hat{y}, \hat{s})$ such that \hat{x} is
220 in the relative interior of the set of optimal solutions of (1.1) and (\hat{y}, \hat{s}) is in the relative interior
221 of the set of optimal solutions of (1.2). \blacksquare

222

223 Thus, if we can formulate a pair of auxiliary primal-dual cone optimization problems, each with
224 generalized Slater points such that the relative interior of \mathcal{R}_D coincides with the relative interior
225 of the optimal solution set of one of our auxiliary problems, then we can design an interior-point
226 algorithm for the auxiliary primal-dual pair, making sure that the iterates of our algorithm stay
227 close to their central path (as they approach the optimal solution set) and generate our desired
228 $x \in \text{relint}(\mathcal{R}_D)$. This is precisely what we accomplish next. Note that in the special case of
229 $K = \mathbb{S}_+^n$, this corresponds to finding maximum rank feasible solutions for the underlying auxiliary
230 SDPs, since the relative interiors of the faces are characterized by their maximal rank elements.

Define the linear transformation $\mathcal{A}_c : \mathbb{S}^n \rightarrow \mathbb{R}^{m+1}$ by

$$\mathcal{A}_c d = \begin{pmatrix} \mathcal{A}d \\ \langle c, d \rangle \end{pmatrix}, \quad \mathcal{A}_c = \begin{bmatrix} \mathcal{A} \\ c \end{bmatrix}.$$

231 This presents a homogenized form of the constraint of (1.1) and combines the two constraints in
232 Lemma 2.3, Item 1. We choose $e = \frac{1}{\sqrt{n}}I$, where I is the identity matrix of order n .

Consider the following conic optimization problem, which we shall henceforth refer to as the *auxiliary problem*.

$$(AP) \quad \begin{aligned} \text{val}_P^{aux} := \min_{d,\delta} \quad & \delta \\ \text{s.t.} \quad & \|\mathcal{A}_c d\| \leq \delta \\ & \langle e, d \rangle = 1 \\ & d \succeq 0. \end{aligned} \quad (2.4)$$

This auxiliary problem is related to the study of the distances to infeasibility in e.g., [24]. The Lagrangian dual of (2.4) is

$$\begin{aligned} & \sup_{w \succeq 0, \begin{pmatrix} \beta \\ z \end{pmatrix} \succeq_{\mathcal{Q}} 0} \inf_{d,\delta} \delta + \gamma(1 - \langle d, e \rangle) - \langle w, d \rangle - \left\langle \begin{pmatrix} \beta \\ z \end{pmatrix}, \begin{pmatrix} \delta \\ \mathcal{A}_c d \end{pmatrix} \right\rangle \\ & = \sup_{w \succeq 0, \begin{pmatrix} \beta \\ z \end{pmatrix} \succeq_{\mathcal{Q}} 0} \inf_{d,\delta} \delta(1 - \beta) - \langle d, \mathcal{A}_c^* z + \gamma e + w \rangle + \gamma, \end{aligned} \quad (2.5)$$

where \mathcal{Q} refers to the second order cone. Since the inner infimum of (2.5) is unconstrained, we get the following equivalent dual.

$$(DAP) \quad \begin{aligned} \text{val}_D^{aux} := \sup_{\gamma, w, z} \quad & \gamma \\ \text{s.t.} \quad & \mathcal{A}_c^* z + \gamma e + w = 0 \\ & w \succeq 0 \\ & \|z\| \leq 1. \end{aligned} \quad (2.6)$$

We define

$$\delta_e := \|\mathcal{A}_c e\|.$$

A strictly feasible primal-dual point is given by

$$d = e, \quad \delta > \delta_e, \quad \text{and} \quad z = 0, \quad \gamma = -1, \quad w = e, \quad (2.7)$$

233 showing that the generalized Slater CQ holds for the pair (2.4)–(2.5).

234 Observe that the complexity of solving (2.4) is essentially that of solving the original dual (1.2).
 235 Recalling that if a path-following interior point method is applied to solve (2.4), one arrives at
 236 a point in the relative interior of the set of optimal solutions, a primal optimal solution (δ^*, d^*)
 237 obtained is such that d^* is of maximum rank possible.

Interestingly, finding $0 \neq d \succeq 0$ that solves $\mathcal{A}_c d = 0$ is equivalent to the SDP

$$\begin{aligned} \inf_d \quad & \|d\| \\ \text{s.t.} \quad & \mathcal{A}_c d = 0, \quad \langle e, d \rangle = 1, \quad d \succeq 0, \end{aligned} \quad (2.8)$$

238 a program for which the Slater CQ generally fails. (See item 2 of Theorem 2.10 below.) This
 239 suggests that the problem of finding the recession direction $0 \neq d \succeq 0$ that certifies a failure for
 240 (1.1) to satisfy Slater condition may be a difficult problem, yet if such d exists, theoretically one
 241 should be able to obtain a maximal rank solution d safely using a path-following interior point
 242 method.

243 We now show how a solution of the auxiliary problem and its dual can be used to obtain useful
 244 information about the original pair of conic problems.

245 **Theorem 2.10.** Assume that the feasible set $\mathcal{F}_P \neq \emptyset$. Then both problems (2.4) and (2.5) satisfy
 246 the generalized Slater CQ, both have optimal solutions, and their (nonnegative) optimal values are
 247 equal. Moreover:

1. If $\delta^* = 0$ and $d^* \succ 0$, then Slater's CQ fails for (1.1) but the generalized Slater CQ holds for (1.2). In fact the primal minimal face and the unique primal optimal solution are

$$f_P = \{0\}, \quad y^* = (\mathcal{A}^*)^\dagger c.$$

2. If $\delta^* = 0$ and $d^* \not\succeq 0$, then Slater's CQ fails for (1.1) and the minimal face satisfies

$$f_P \trianglelefteq \mathbb{S}_+^n \cap \{d^*\}^\perp \triangleleft \mathbb{S}_+^n. \quad (2.9)$$

3. If $\delta^* > 0$, then Slater's CQ holds for (1.1). Let $z^* = (z_1^*, z_2^*, \dots, z_{m+1}^*) \in \mathbb{R}^{m+1}$ be an
 249 optimal solution of (2.6). If $z_{m+1}^* > 0$, then $c - \sum_{i=1}^m \frac{z_i^*}{z_{m+1}^*} A_i \succeq \delta^* e$. If $z_{m+1}^* = 0$, then
 250 $c - \lambda \sum_{i=1}^m z_i^* A_i \succ 0$ for sufficiently large $\lambda > 0$.

4. Moreover,

$$\delta^* > 0 \iff \mathcal{N}(\mathcal{A}_c) \cap \mathbb{S}_+^n = \{0\}. \quad (2.10)$$

251 *Proof.* A strictly feasible pair for (2.4)–(2.5) is given in (2.7). Hence both problems have equal
 252 optimal values and both values are attained.

1. Suppose that $\delta^* = 0$ and $d^* \succ 0$. It follows that $\mathcal{A}_c(d^*) = 0$ and $d^* \neq 0$, i.e., $d^* \in \mathcal{R}_D$, where \mathcal{R}_D was defined in (2.3). It follows from Lemma 2.7 that

$$f_P \trianglelefteq \mathbb{S}_+^n \cap \{d\}^\perp = \mathbb{S}_+^n \cap ((\mathbb{S}_+^n)^*)^\perp = \{0\}.$$

253 Hence all feasible points for (1.1) satisfy $c - \mathcal{A}^*y = 0$. Since \mathcal{A} is an onto mapping, we
 254 conclude that the unique solution of this linear system is $y = (\mathcal{A}^*)^\dagger c$.

255 Since \mathcal{A} is onto, there exists \bar{x} such that $\mathcal{A}\bar{x} = b$. Thus, for every $t \geq 0$, $\mathcal{A}(\bar{x} + td^*) = b$, and
 256 for t large enough, $\bar{x} + td^* \succ 0$. Therefore, the generalized Slater CQ holds for (1.2).

2. The result follows from Lemma 2.7.

3. If $\delta^* > 0$, then $\mathcal{R}_D = \{0\}$, where \mathcal{R}_D was defined in (2.3). It follows from Lemma 2.3 that
 259 the Slater CQ holds for (1.1).

260 For any feasible solution $(\gamma, w, z) \in \mathbb{R} \times \mathbb{S}_+^n \times \mathbb{R}^{m+1}$ of (2.6), we have $z_{m+1}c - \sum_{i=1}^m z_i A_i \succeq \gamma e$.
 261 By strong duality, the optimal values of (2.4) and (2.6) coincide, thus $z_{m+1}^*c - \sum_{i=1}^m z_i^* A_i \succeq$
 262 $\delta^* e \succ 0$. When $z_{m+1}^* = 0$, $c - \lambda \sum_{i=1}^m z_i^* A_i \succeq c - \lambda \delta^* e \succ 0$ for sufficiently large $\lambda > 0$.

4. This follows immediately from the definition of (AP) in (2.4).

264 □

265 Since $(\delta, d) = (\mathcal{A}_c e, e)$ is feasible for (2.4), the optimal solution of (2.4) satisfies $\delta^* \leq \delta_e =$
 266 $\|\mathcal{A}_c(\frac{1}{\sqrt{n}}I)\|$. Thus, in the case that $\delta_e = 0$, $d^* = e$ is optimal for (2.4). The next lemma pertains to
 267 the case that $\delta_e > 0$.

268 **Lemma 2.11.** *Let $\delta_e > 0$. Then, one of the following three cases must occur.*

- 269 1. *If $val_P^{aux} = val_D^{aux} = 0$, then Slater's CQ fails for (1.1).*
- 270 2. *If $0 < val_P^{aux} = val_D^{aux} < \delta_e$, then Slater's CQ holds for (1.1). All primal optimal solutions d^**
 271 *satisfy $\|\mathcal{A}_c d^*\| = val_P^{aux}$ and every optimal z^* is of the form $-\mathcal{A}_c d^* / \|\mathcal{A}_c d^*\|$ for some primal*
 272 *optimal d^* .*
3. *If $val_P^{aux} = val_D^{aux} = \delta_e$, then all results in the previous case hold, and an optimal solution of*
 (2.4) *is given by $d^* = e$. Moreover, this case occurs if, and only if, \mathcal{A} and c satisfy*

$$\mathcal{A}_c^*(\mathcal{A}_c e) = \|\mathcal{A}_c e\|^2 I. \quad (2.11)$$

273 *Proof.* 1. This case was discussed in the previous lemma.

274 2. It is clear that $\|\mathcal{A}_c d^*\| = \delta^* = val_P^{aux}$. Since a constraint qualification holds for (2.4)–(2.5), a
 275 necessary optimality condition for this pair is that the Jordan product of the vectors $(\delta^*, \mathcal{A}_c d^*)$
 276 and $(1, z^*)$ be zero, i.e., these vectors lie in conjugate faces of the second-order cone. It follows
 277 that $\delta^* z^* + \mathcal{A}_c d^* = 0$, giving $z^* = -\mathcal{A}_c d^* / \|\mathcal{A}_c d^*\|$.

278 3. In the case that $val_P^{aux} = val_D^{aux} = \delta_e$, clearly $d^* = e$ is optimal. Moreover, by complemen-
 279 tarity slackness, $w^* = 0$, so it follows from dual feasibility that $\mathcal{A}_c^* z^* + \delta^* I = 0$. Using the
 280 relation for z^* proved in the previous part, we obtain (2.11). □

282 The relation in (2.11) implies: *if the assumptions in Case 3 in Lemma 2.11 holds and $\langle c, e \rangle > 0$,*
 283 *then the point $y = -\frac{\mathcal{A}_c e}{\langle c, e \rangle}$ is strongly feasible for (1.1).*

Remark 2.12. *Theorem 2.10 shows that if the primal problem is feasible, then $\delta^* > 0$ if, and only if, Slater's CQ holds for the primal. Equivalently, Item 4 in Theorem 2.10 means that $\delta^* = 0$ if, and only if, \mathcal{A}_c has a right singular vector d such that $d \succeq 0$ and the corresponding singular value is zero, i.e we could replace (AP) with $\min \{\|\mathcal{A}_c d\| : \|d\| = 1, d \succeq 0\}$. Therefore, we could solve (AP) using a basis for the nullspace of \mathcal{A}_c , e.g., $\mathcal{R}(\mathcal{N}_A) = \mathcal{N}(\mathcal{A}_c)$, and an approach based on maximizing the smallest eigenvalue:*

$$\delta \approx \sup_y \{\lambda_{\min}(\mathcal{N}_A y) : \text{trace}(\mathcal{N}_A y) = 1, \|y\| \leq 1\},$$

284 *so, in the case when $\delta^* = 0$, both (AP) and (DAP) can be seen as a max-min eigenvalue problem*
 285 *(subject to a bound and a linear constraint).*

286 2.4 Rank-revealing rotation and equivalent problems

Following the results from Theorem 2.10, we focus on the case where $\delta^* = 0$ and $\mathcal{R}_D^n \cap \mathbb{S}_{++}^n = \emptyset$, where \mathcal{R}_D^n is the set of normalized recession directions of (1.2), that is,

$$\mathcal{R}_D^n := \{d \in \mathbb{S}^n : \mathcal{A}d = 0, \langle d, c \rangle = 0, \langle e, d \rangle = 1, d \succeq 0\} \neq \emptyset.$$

287 We have the following observation:

Proposition 2.13. Suppose $\mathcal{F}_P \neq \emptyset$, $\delta^* = 0$ and $\text{int}(\mathcal{R}_D^n) = \emptyset$. Let $d^* \in \text{ri}(\mathcal{R}_D^n)$ have the eigenvalue decomposition

$$d^* = \begin{bmatrix} P & Q \end{bmatrix} \begin{bmatrix} d_+ & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P^T \\ Q^T \end{bmatrix},$$

288 where $\begin{bmatrix} P & Q \end{bmatrix}$ is orthogonal, $Q \in \mathbb{R}^{n \times \bar{n}}$ and $d_+ \succ 0$. Then $f_P \preceq QS_+^{\bar{n}}Q^T$, and $\mathcal{R}_D^n \subseteq PS_+^{n-\bar{n}}P^T$.

289 *Proof.* The first claim follows directly from Lemma 2.7: $f_P \preceq S_+^n \cap \{d^*\}^\perp = QS_+^{\bar{n}}Q^T$.

290 To prove the second claim, recall that (Theorem 18.2 [31]) if $\text{face}(\mathcal{R}_D^n)$ is the minimal face
291 of S_+^n containing \mathcal{R}_D^n , then $\text{ri}(\mathcal{R}_D^n) \subseteq \text{ri}(\text{face}(\mathcal{R}_D^n))$. Hence $d^* = Pd_+P^T \in \text{ri}(\mathcal{R}_D^n)$ implies that
292 $\mathcal{R}_D^n \subseteq \text{face}(\mathcal{R}_D^n) = PS_+^{n-\bar{n}}P^T$. \square

293 Note in particular that the inclusions are independent of the choice of P, Q ; it depends only on
294 the linear subspaces $\mathcal{R}(P \cdot P^T)$ and $\mathcal{R}(Q \cdot Q^T)$.

295 Given the assumptions and notations in Proposition 2.13, we may reduce (1.1) to an equivalent
296 problem over a spectrahedron of lower dimension using Q . The reduction is independent of the
297 specific choice of Q by Proposition 2.13.

298 **Proposition 2.14.** Suppose that \mathcal{A} is surjective; and, let $\delta^* = 0$ and $d^* = \begin{bmatrix} P & Q \end{bmatrix} \begin{bmatrix} d_+ & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P^T \\ Q^T \end{bmatrix}$,

299 where $\begin{bmatrix} P & Q \end{bmatrix}$ is orthogonal, $Q \in \mathbb{R}^{n \times \bar{n}}$ and $d_+ \succ 0$. Then:

1.

$$0 = \min_{y \in \mathbb{R}^m, w \in S^{\bar{n}}} \|c - (\mathcal{A}^*y + QwQ^T)\|. \quad (2.12)$$

300 2. If $\mathcal{R}(Q \cdot Q^T) \cap \mathcal{R}(\mathcal{A}^*) = \{0\}$, then for any $y_1, y_2 \in \mathcal{F}_P$, $b^T y_1 = b^T y_2 = v_P$.

3. If $\mathcal{R}(Q \cdot Q^T) \cap \mathcal{R}(\mathcal{A}^*) \neq \{0\}$, let $\hat{\mathcal{A}} : S_+^{\bar{n}} \rightarrow \mathbb{R}^{\bar{m}}$ be a surjective map that satisfies

$$\mathcal{R}(\hat{\mathcal{A}}^*) = \mathcal{R}(\mathcal{A}^*) \cap \mathcal{R}(Q \cdot Q^T).$$

Let (y_Q, w_Q) be an optimal solution to the least squares problem (2.12), and let $c_Q := c - \mathcal{A}^*y_Q$.
Then

$$\mathcal{F}_P = y_Q + \left\{ y : c_Q - \hat{\mathcal{A}}^*v \succeq 0, \mathcal{A}^*y = \hat{\mathcal{A}}^*v \right\}.$$

301 *Proof.* 1. For any feasible solution $y \in \mathcal{F}_P$, we have $c - \mathcal{A}^*y \in \mathcal{R}(Q \cdot Q^T)$. Therefore, $(y, Q^T(c -$
302 $\mathcal{A}^*y)Q)$ is an optimal solution to the least squares problem; thus the problem attains an
303 objective value zero.

2. Since $c - \mathcal{A}^*y \in \mathcal{R}(Q \cdot Q^T), \forall y \in \mathcal{F}_P$, we get $\mathcal{A}^*(y_2 - y_1) = (c - \mathcal{A}^*y_1) - (c - \mathcal{A}^*y_2) \in$
 $\mathcal{R}(Q \cdot Q^T) \cap \mathcal{R}(\mathcal{A}^*) = \{0\}$. Given that \mathcal{A} is surjective, we get $b = \mathcal{A}\hat{x}$, for some $\hat{x} \in S^n$, and

$$b^T(y_2 - y_1) = \langle \hat{x}, \mathcal{A}^*(y_2 - y_1) \rangle = 0.$$

3. Any optimal solution (y_Q, w_Q) of (2.12) satisfies $c = \mathcal{A}^*y_Q + Qw_QQ^T$, by Item 1. Hence

$$\begin{aligned}\mathcal{F}_P &= \{y : c - \mathcal{A}^*y \in \mathcal{R}(Q \cdot Q^T) \cap \mathbb{S}_+^n\} \\ &= \{y : c_Q - \mathcal{A}^*(y - y_Q) \in \mathcal{R}(Q \cdot Q^T) \cap \mathbb{S}_+^n\} \\ &= y_Q + \{y : c_Q - \mathcal{A}^*y \in \mathcal{R}(Q \cdot Q^T) \cap \mathbb{S}_+^n\}.\end{aligned}$$

Since $c_Q \in \mathcal{R}(Q \cdot Q^T)$, we see that $c_Q - \mathcal{A}^*y \in \mathcal{R}(Q \cdot Q^T)$ iff $\mathcal{A}^*y \in \mathcal{R}(Q \cdot Q^T) \cap \mathcal{R}(\mathcal{A}^*)$, iff $\mathcal{A}^*y = \hat{\mathcal{A}}^*v$ for some $v \in \mathbb{R}^{\bar{m}}$. Consequently,

$$\begin{aligned}\mathcal{F}_P &= y_Q + \{y : c_Q - \hat{\mathcal{A}}^*v \in \mathcal{R}(Q \cdot Q^T) \cap \mathbb{S}_+^n, \mathcal{A}^*y = \hat{\mathcal{A}}^*v\} \\ &= y_Q + \{y : c_Q - \hat{\mathcal{A}}^*v \succeq 0, \mathcal{A}^*y = \hat{\mathcal{A}}^*v\}.\end{aligned}$$

304

□

Corollary 2.15. *Suppose that $\delta^* = 0$ and*

$$d^* = [P \quad Q] \begin{bmatrix} d_+ & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P^T \\ Q^T \end{bmatrix},$$

solve (AP), where $[P \quad Q]$ is orthogonal, $Q \in \mathbb{R}^{n \times \bar{n}}$, $d_+ \succ 0$, and $\mathcal{R}(Q \cdot Q^T) \cap \mathcal{R}(\mathcal{A}^*) \neq \{0\}$. Let $\hat{\mathcal{A}} : \mathbb{S}_+^n \rightarrow \mathbb{R}^{\bar{m}}$ be a surjective map that satisfies

$$\mathcal{R}(\hat{\mathcal{A}}^*) = \mathcal{R}(\mathcal{A}^*) \cap \mathcal{R}(Q \cdot Q^T).$$

Let (y_Q, w_Q) be an optimal solution to the least squares problem (2.12), and let $c_Q := c - \mathcal{A}^*y_Q$. Then

$$v_P = b^T y_Q + \sup \left\{ \left(\hat{\mathcal{A}}(\mathcal{A}^\dagger b) \right)^T v : Q^T (c_Q - \hat{\mathcal{A}}^*v) Q \succeq 0 \right\}.$$

Moreover, if $\hat{\mathcal{A}}^* = \mathcal{A}^* \mathcal{P}$ for some injective linear map $\mathcal{P} : \mathbb{R}^{\bar{m}} \rightarrow \mathbb{R}^m$, then

$$v_P = b^T y_Q + \sup \left\{ (\mathcal{P}^* b)^T v : Q^T (c_Q - \hat{\mathcal{A}}^*v) Q \succeq 0 \right\}, \quad (2.13)$$

305 and if v^* is an optimal solution of (2.13), then $y^* = y_Q + \mathcal{P}v^*$ is an optimal solution of (1.1).

Proof. Since \mathcal{A} is surjective, we get $b = \mathcal{A}\mathcal{A}^\dagger b$. From Proposition 2.14, we have that

$$\begin{aligned}v_P &= \sup \{b^T y : y \in \mathcal{F}_P\} \\ &= b^T y_Q + \sup \left\{ \left\langle \mathcal{A}^\dagger b, \mathcal{A}^*y \right\rangle : c_Q - \hat{\mathcal{A}}^*v \succeq 0, \mathcal{A}^*y = \hat{\mathcal{A}}^*v \right\} \\ &= b^T y_Q + \sup \left\{ \left\langle \mathcal{A}^\dagger b, \hat{\mathcal{A}}^*v \right\rangle : c_Q - \hat{\mathcal{A}}^*v \succeq 0, \mathcal{A}^*y = \hat{\mathcal{A}}^*v \right\} \\ &= b^T y_Q + \sup \left\{ \left(\hat{\mathcal{A}}\mathcal{A}^\dagger b \right)^T v : c_Q - \hat{\mathcal{A}}^*v \succeq 0 \right\}.\end{aligned}$$

Since $c_Q \in \mathcal{R}(Q \cdot Q^T)$ and $\mathcal{R}(\hat{\mathcal{A}}^*) \subseteq \mathcal{R}(Q \cdot Q^T)$, we have

$$v_P = b^T y_Q + \sup \left\{ \left(\hat{\mathcal{A}}\mathcal{A}^\dagger b \right)^T v : Q^T (c_Q - \hat{\mathcal{A}}^*v) Q \succeq 0 \right\}.$$

If $\hat{\mathcal{A}}^* = \mathcal{A}^* \mathcal{P}$ for some injective linear map $\mathcal{P} : \mathbb{R}^{\bar{m}} \rightarrow \mathbb{R}^m$, then $\mathcal{A}^* y = \hat{\mathcal{A}}^* v$ iff $y = \mathcal{P} v$ (as \mathcal{A}^* is injective), so

$$\mathcal{F}_P = y_Q + \{\mathcal{P} v : c_Q - \hat{\mathcal{A}}^* v \succeq 0\}. \quad (2.14)$$

Since $\hat{\mathcal{A}} \mathcal{A}^\dagger b = \mathcal{P}^* \mathcal{A} \mathcal{A}^\dagger b = \mathcal{P}^* b$, we have

$$v_P = b^T y_Q + \sup \left\{ (\mathcal{P}^* b)^T v : Q^T c_Q Q - Q^T (\hat{\mathcal{A}}^* v) Q \succeq 0 \right\}.$$

306 By (2.14), if v^* is an optimal solution of (2.13), then $y^* := y_Q + \mathcal{P} v^*$ is an optimal solution of
307 (1.1). \square

308 This gives a theoretical means of reducing a degenerate SDP to one of smaller dimension. In
309 practice, we would like to avoid the computation of $\hat{\mathcal{A}} x$ for some solution $\mathcal{A} x = b$, directly, as that
310 requires finding explicitly a particular solution x to the linear system $\mathcal{A} x = b$. Indeed, it is possible
311 to express this in terms of the original b without the knowledge of any particular solution x , as long
312 as we are able to obtain the operator \mathcal{P} mentioned in Corollary 2.15.

Proposition 2.16. *For any n -by- n orthogonal matrix $U = \begin{bmatrix} P & Q \end{bmatrix}$ and surjective linear operator $\mathcal{A} : \mathbb{S}^n \rightarrow \mathbb{R}^m$, there exist $\bar{m} \leq m$ and a one-one linear transformation $\mathcal{P} : \mathbb{R}^{\bar{m}} \rightarrow \mathbb{R}^m$ that satisfies*

$$\mathcal{R}(\mathcal{A}^* \mathcal{P}) = \mathcal{R}(Q \cdot Q^T) \cap \mathcal{R}(\mathcal{A}^*). \quad (2.15)$$

$$\text{and } \mathcal{R}(\mathcal{P}) = \mathcal{N}(P^T (\mathcal{A}^* \cdot) P) \cap \mathcal{N}(P^T (\mathcal{A}^* \cdot) Q), \quad (2.16)$$

Moreover, $\bar{\mathcal{A}} : \mathbb{S}^{\bar{n}} \rightarrow \mathbb{R}^{\bar{m}}$ defined by

$$\bar{\mathcal{A}}^*(\cdot) := Q^T (\mathcal{A}^* \mathcal{P}(\cdot)) Q$$

313 is surjective.

Proof. Recall that $X \in \mathcal{R}(Q \cdot Q^T)$ if, and only if, $P^T X P = 0$ and $P^T X Q = 0^4$. For any $y \in \mathbb{R}^m$, $\mathcal{A}^* y \in \mathcal{R}(Q \cdot Q^T)$ if, and only if,

$$\sum_{i=1}^m (P^T A_i P) y_i = 0 \quad \text{and} \quad \sum_{i=1}^m (P^T A_i Q) y_i = 0,$$

which holds if, and only if, $y \in \text{span}\{\beta\}$, where $\beta := \{y_1, \dots, y_{\bar{m}}\}$ is a basis of the linear subspace

$$\left\{ y : \sum_{i=1}^m (P^T A_i P) y_i = 0 \right\} \cap \left\{ y : \sum_{i=1}^m (P^T A_i Q) y_i = 0 \right\} = \mathcal{N}(P^T (\mathcal{A}^* \cdot) P) \cap \mathcal{N}(P^T (\mathcal{A}^* \cdot) Q).$$

Now define $\mathcal{P} : \mathbb{R}^{\bar{m}} \rightarrow \mathbb{R}^m$ by

$$\mathcal{P} \lambda = \sum_{i=1}^{\bar{m}} \lambda_i y_i \quad \text{for } \lambda \in \mathbb{R}^{\bar{m}}.$$

⁴Recall that for any matrix $X \in \mathbb{S}^n$,

$$X = P P^T X P P^T + P P^T X Q Q^T + Q Q^T X P P^T + Q Q^T X Q Q^T,$$

so $X \in \mathcal{R}(Q \cdot Q^T)$ if, and only if, $P^T X P = 0$ and $P^T X Q = 0$.

Then, by definition of \mathcal{P} , we have

$$\mathcal{R}(\mathcal{A}^*\mathcal{P}) = \mathcal{R}(Q \cdot Q^T) \cap \mathcal{R}(\mathcal{A}^*) \quad \text{and} \quad \mathcal{R}(\mathcal{P}) = \mathcal{N}(P^T(\mathcal{A}^*\cdot)P) \cap \mathcal{N}(P^T(\mathcal{A}^*\cdot)Q).$$

314 The onto property of $\bar{\mathcal{A}}$ follows from (2.15) and the fact that both $\mathcal{P}, \mathcal{A}^*$ are one-one⁵. □

315 2.5 LP: an example

316 In this subsection we consider the special case where (1.1) is a linear program (and, more generally,
317 the special case of optimizing over an arbitrary polyhedral cones, see e.g., [32, 33, 50, 49]). If the
318 original problem does not satisfy Slater's condition, then one single iteration of facial reduction
319 yields a reduced problem that satisfies Slater's condition.

320 **Theorem 2.17.** *Assume the hypotheses as in Corollary 2.15, with d^* being a maximal rank optimal*
321 *solution of (AP). If $A_i = \text{Diag}(a_i)$ for some $a_i \in \mathbb{R}^n$, for $i = 1, \dots, m$, and $c = \text{Diag}(\underline{c})$, for some*
322 *$\underline{c} \in \mathbb{R}^n$, then the reduced problem (2.13) satisfies Slater's condition.*

Proof. Since A_i, c are diagonal, without loss of generality we may assume that the solution of (AP) is diagonal (so is a vector) and satisfies

$$d^* = \begin{bmatrix} \bar{d}^* \\ 0 \end{bmatrix} \in \mathbb{R}^n,$$

where $\bar{d}^* \in \mathbb{R}^{n-\bar{n}}$ satisfies $\bar{d}^* > 0$. (Then Q is the submatrix formed by deleting the first $n - \bar{n}$ columns of I_n .) Let $A^T = [a_1, \dots, a_m]$. Then for all $\underline{z} = \underline{c} - A^T y \geq 0$, $\underline{z}_i = 0$ for $i = 1, \dots, n - \bar{n}$, and $\text{diag}(cQ) = \begin{bmatrix} 0 \\ \bar{c} \end{bmatrix}$ for some $\bar{c} \in \mathbb{R}^{\bar{n}}$. Let

$$A = [A_1, A_2],$$

where $A_2 \in \mathbb{R}^{m \times \bar{n}}$, and $L \in \mathbb{R}^{m \times \bar{m}}$ be a nullspace representation of A_1^T , i.e. $\mathcal{N}(A_1^T) = \mathcal{R}(L)$ and L has orthogonal columns. Also, noting that $\mathcal{R}(\mathcal{A}^*) \cap \mathcal{R}(Q \cdot Q^T)$ is a subset of $n \times n$ diagonal matrices,

$$\begin{aligned} \text{diag}(\mathcal{R}(\mathcal{A}^*) \cap \mathcal{R}(Q \cdot Q^T)) &= \left\{ \begin{bmatrix} A_1^T y \\ A_2^T y \end{bmatrix} : A_1^T y = 0 \right\} \\ &= \left\{ \begin{bmatrix} A_1^T y \\ A_2^T y \end{bmatrix} : y = Lv \text{ for some } v \in \mathbb{R}^{\bar{m}} \right\} \\ &= \{A^T L v : v \in \mathbb{R}^{\bar{m}}\}. \end{aligned}$$

From Corollary 2.15, (1.1) is equivalent to

$$\sup_v (L^T b)^T v \quad \text{s.t.} \quad \bar{c} - \bar{A}^T v \geq 0,$$

323 where $\bar{A}^T = A_2^T L$.

⁵If $\bar{\mathcal{A}}^* v = 0$, noting that $\mathcal{A}^* \mathcal{P} v = Q w Q^T$ for some $w \in \mathbb{S}^{\bar{n}}$ by (2.15), we have that $w = 0$ so $\mathcal{A}^* \mathcal{P} v = 0$. Since both \mathcal{A}^* and \mathcal{P} injective, we have that $v = 0$.

If there exists $\bar{d} \in \mathbb{R}^{\bar{n}}$ such that

$$L^T A_2 \bar{d} = \bar{A} \bar{d} = 0, \quad \bar{c}^T \bar{d} = 0, \quad \bar{e}^T \bar{d} = 1,$$

then

$$A_2 \bar{d} \in \mathcal{N}(L^T) = \mathcal{N}(A_1^T)^\perp = \mathcal{R}(A_1).$$

Therefore, there exists some $\tilde{d} \in \mathbb{R}^{n-\bar{n}}$ such that $A_1 \tilde{d} + A_2 \bar{d} = 0$. Let

$$\bar{\lambda} := \min_i \left\{ \frac{d_i^*}{d_i^* - y_i} : i \in \{1, \dots, n - \bar{n}\}, y_i < 0 \right\}.$$

Note that, since $d_i^* > 0$ for $i \in \{1, \dots, n - \bar{n}\}$, we get $\bar{\lambda} \in (0, 1)$. Now, pick any $\lambda \in (0, \bar{\lambda})$; then the vector

$$d^{**} := (1 - \lambda)d^* + \lambda \begin{bmatrix} \tilde{d} \\ \bar{d} \end{bmatrix} = \begin{bmatrix} (1 - \lambda)d^* + \lambda \tilde{d} \\ \lambda \bar{d} \end{bmatrix} \geq 0$$

and, by the choice of λ , satisfies $d_i^{**} > 0$ for $i = 1, \dots, n - \bar{n}$. And, $d_i^{**} > 0$, for some $i \in \{n - \bar{n} + 1, \dots, n\}$. Also,

$$\begin{aligned} Ad^{**} &= (1 - \lambda)Ad^* + \lambda A \begin{bmatrix} \tilde{d} \\ \bar{d} \end{bmatrix} \\ &= \lambda (A_1 \tilde{d} + A_2 \bar{d}) = 0, \\ \underline{c}^T d^{**} &= (c_Q - A^T y_Q)^T d^{**} = \begin{bmatrix} 0 \\ \bar{c} \end{bmatrix}^T \begin{bmatrix} \tilde{d} \\ \bar{d} \end{bmatrix} = \bar{c}^T \bar{d} = 0. \end{aligned}$$

Therefore we have

$$A \left(\frac{1}{\bar{e}^T d^{**}} d^{**} \right) = 0, \quad c^T \left(\frac{1}{\bar{e}^T d^{**}} d^{**} \right) = 0, \quad e^T \left(\frac{1}{\bar{e}^T d^{**}} d^{**} \right) = 1, \quad \frac{1}{\bar{e}^T d^{**}} d^{**} \geq 0,$$

324 and the number of nonzeros in d^{**} is greater than the number of nonzeros in d^* , contradicting the
325 choice of d^* . \square

326 3 Facial Reduction

327 We now outline our procedure to find the minimal face f_P . Each iteration of our procedure involves
328 two steps. We first identify $0 \neq d \in (f_P)^c$ using Lemma 2.3. This means that $f_P \trianglelefteq K \leftarrow (K \cap \{d\}^\perp)$
329 and the interior of this new K is empty. We then project (P) into $\text{span}(K)$; thus we reduce the
330 dimension of the variables and size of the constraints of our problem, and maintain $\text{int } K \neq \emptyset$ for
331 the new K . (Note that for numerical stability and well-posedness, it is essential that there exists
332 Lagrange multipliers and that $\text{int } K \neq \emptyset$. Regularization involves both finding a minimal face as
333 well as a minimal subspace, see [40].) We then repeat the process.

334 Therefore, in the case that $\text{int } K = \emptyset$, we need to to obtain an equivalent problem to (P) in
335 the subspace $\text{span}(K) = K - K$. One essential step is finding a subspace intersection. We can
336 apply the algorithm in e.g., [14, Thm 12.4.2]. In particular, by abuse of notation, let H_1, H_2 be
337 matrices with orthonormal columns representing the orthonormal bases of the subspaces $\mathcal{H}_1, \mathcal{H}_2$,
338 respectively. Then we need only find a singular value decomposition $H_1^T H_2 = U \Sigma V^T$ and find

339 which singular vectors correspond to singular values $\Sigma_{ii}, i = 1, \dots, r$, (close to) 1. Then both
 340 $H_1U(:, 1 : r)$ and $H_2V(:, 1 : r)$ provide matrices whose ranges yield the intersection. For cones such
 341 as \mathbb{S}_+^n , which possess a “self-replicating” structure, there is an advantage to choosing an isometry
 342 \mathcal{I} so that $\mathcal{I}(\mathbb{S}_+^n \cap (K - K))$ is a smaller dimensional PSD cone \mathbb{S}_+^r .

343 Algorithm 3.1 outlines one iteration of facial reduction. The output returns an equivalent
 344 problem $(\bar{A}, \bar{b}, \bar{c})$ on a smaller face of \mathbb{S}^n that contains the feasible set \mathcal{F}_P ; and we also obtain the
 345 linear transformation \mathcal{P} , and point y_Q that are needed for recovering an optimal solution of the
 346 original problem (P). (See Corollary 2.15.)

Algorithm 3.1: One iteration of facial reduction

- 1 Input($\mathcal{A} : \mathbb{S}^n \rightarrow \mathbb{R}^m, b \in \mathbb{R}^m, c \in \mathbb{S}^n$);
- 2 Obtain an optimal solution (δ^*, d^*) of (AP)

$$\min_{\delta, d} \delta \text{ s.t. } \|\mathcal{A}_c d\| \leq \delta, \langle e, d \rangle = 1, d \succeq 0.$$

3 **if** $\delta^* > 0$, **then**
 4 | STOP; Slater CQ holds for (A, b, c) .
 5 **else**
 6 | **if** $d^* \succ 0$, **then**
 7 | | STOP; generalized Slater CQ holds for (\mathcal{A}, b, c) (see Theorem 2.10);
 8 | | **else**
 9 | | Obtain eigenvalue decomposition
 10 | |
$$d^* = [P \quad Q] \begin{bmatrix} d_+ & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P^T \\ Q^T \end{bmatrix}$$

 11 | | as described in 2.15, with $Q \in \mathbb{R}^{n \times \bar{n}}$;
 12 | | **if** $\mathcal{R}(Q \cdot Q^T) \cap \mathcal{R}(\mathcal{A}^*) = \{0\}$, **then**
 13 | | | STOP; all feasible solutions of $\sup_y \{b^T y : c - \mathcal{A}^* y \succeq 0\}$ are optimal.
 14 | | | **else**
 15 | | | find $\mathcal{P} : \mathbb{R}^{\bar{m}} \rightarrow \mathbb{R}^m$ satisfying the conditions in Corollary 2.15;
 16 | | | solve (2.12) for (y_Q, w_Q) ;
 17 | | | $\bar{c} \leftarrow w_Q$;
 18 | | | $\bar{b} \leftarrow \mathcal{P}^* b$;
 19 | | | $\bar{A}^* \leftarrow Q^T (\mathcal{A}^* \mathcal{P}(\cdot)) Q$;
 20 | | | Output($\bar{\mathcal{A}} : \mathbb{S}^{\bar{n}} \rightarrow \mathbb{R}^{\bar{m}}, \bar{b} \in \mathbb{R}^{\bar{m}}, \bar{c} \in \mathbb{S}^{\bar{n}}; y_Q \in \mathbb{R}^m, \mathcal{P} : \mathbb{R}^{\bar{m}} \rightarrow \mathbb{R}^m$);
 21 | | | **end if**
 22 | | **end if**
 23 | **end if**

We can take advantage of the fact that eigenvalue-eigenvector calculations are efficient and accurate to obtain a more accurate optimal solution (δ^*, d^*) of (AP), i.e., to decide whether the linear system

$$\langle A_i, d \rangle = 0 \quad \forall i = 1, \dots, m+1 \quad (\text{where } A_{m+1} := c), \quad 0 \neq d \succeq 0 \quad (3.1)$$

has a solution, we can use Algorithm 3.2 as a preprocessor for Algorithm 3.1. More precisely,

Algorithm 3.2: Solving (AP)

```

1 Input(  $A_1, \dots, A_m, A_{m+1} := c \in \mathbb{S}^n$ );
2 Output(  $\delta^*, P \in \mathbb{R}^{n \times (n-\bar{n})}, d_+ \in \mathbb{S}^{n-\bar{n}}$  satisfying  $d_+ \succ 0$ ; (so  $d^* = Pd_+P$ );
3 ) if some of the  $A_i$  ( $i \in \{1, \dots, m+1\}$ ) is definite then
4 | STOP; (3.1) does not have a solution.
5 else
6 | if some of the  $A = [U \ \tilde{U}] \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U^T \\ \tilde{U}^T \end{bmatrix} \in \{A_i : i = 1, \dots, m+1\}$  is semidefinite, with
   |  $D \succ 0$ , then
7 | |  $A_i \leftarrow \tilde{U}^T A_i \tilde{U}$ ;
8 | else
9 | | while  $\exists 0 \neq V \in \mathbb{R}^{n \times r}$  such that  $A_i V = 0$  for all  $i = 1, \dots, m+1$ , do
10 | | | % We have that  $\langle A_i, VV^T \rangle = 0 \ \forall i = 1, \dots, m+1$ ;
11 | | | STOP;  $\delta^* = 0, d^* = VV^T$  solves (AP);
12 | | end while
13 | | Use SDP solver to solve (AP).
14 | end if
15 end if

```

347

348

349

350

351

Algorithm 3.2 tries to find a solution d^* satisfying (3.1) without using an SDP solver. It attempts to find a vector v in the nullspace of all the A_i , and then sets $d^* = vv^T$. In addition, any semidefinite A_i allows a reduction to a smaller dimensional space.

351

4 Sensitivity Analysis

352

353

After solving the auxiliary problem (AP) and obtaining $\delta^* = 0$, we can reduce the problem by restricting to the face $K \cap \{d\}^\perp$. However, one cannot (numerically) expect that $\delta^* = 0$.

354

4.1 Rank-revealing rotation and the minimal face $Q\mathbb{S}_+^{n-1}Q^T$

355

The next result shows that δ^* from (AP) is a measure of how close Slater's CQ is to failing.

Theorem 4.1. *Let δ^*, d^* denote the optimum of the auxiliary problem (2.4). Then δ^* bounds how far the feasible primal slacks $\{c - \mathcal{A}^*y : y \in \mathcal{F}_P\}$ are from orthogonality to d^* :*

$$\sup_{y \in \mathcal{F}_P, c - \mathcal{A}^*y \neq 0} \frac{\langle d^*, c - \mathcal{A}^*y \rangle}{\|d^*\| \|c - \mathcal{A}^*y\|} \leq \alpha(\mathcal{A}, c), \quad (4.1)$$

where

$$\alpha(\mathcal{A}, c) = \begin{cases} \delta^*/\sigma_{\min}(\mathcal{A}) & \text{if } c \in \mathcal{R}(\mathcal{A}^*), \\ \delta^*/\min\{\sigma_{\min}(\mathcal{A}^*), \|c_{\mathcal{N}(\mathcal{A})}\|\} & \text{if } c \notin \mathcal{R}(\mathcal{A}^*). \end{cases} \quad (4.2)$$

Proof. Since $\langle e, d^* \rangle = 1$ and recalling $e := \frac{1}{\sqrt{n}}I$, we get

$$\sqrt{n} = \sqrt{n\langle e, d^* \rangle^2} = \text{trace } d^* \geq \|d^*\| \geq \frac{\langle e, d^* \rangle}{\|e\|} = \frac{1}{\frac{1}{\sqrt{n}}\|I\|} = 1.$$

If $c = \mathcal{A}^*y_c$ for some $y_c \in \mathbb{R}^m$, then

$$\begin{aligned} \cos \theta_{d^*z} &= \frac{\langle d^*, c - \mathcal{A}^*y \rangle}{\|d^*\| \|c - \mathcal{A}^*y\|} \leq \frac{\left\langle \mathcal{A}d^*, \begin{pmatrix} y_c - y \\ 1 \end{pmatrix} \right\rangle}{\left\| \mathcal{A}^* \begin{pmatrix} y_c - y \\ 1 \end{pmatrix} \right\|} \\ &\leq \frac{\|\mathcal{A}d^*\| \left\| \begin{pmatrix} y_c - y \\ 1 \end{pmatrix} \right\|}{\sigma_{\min}(\mathcal{A}^*) \left\| \begin{pmatrix} y_c - y \\ 1 \end{pmatrix} \right\|} \\ &\leq \frac{\delta^*}{\sigma_{\min}(\mathcal{A})}. \end{aligned}$$

If $c \notin \mathcal{R}(\mathcal{A}^*)$, then $c - \mathcal{A}^*y = \left\langle \mathcal{A}_{c_{\mathcal{N}(\mathcal{A})}}d^*, \begin{pmatrix} -y \\ 1 \end{pmatrix} \right\rangle$, and by Assumption 1.2, we conclude that $0 \preceq c - \mathcal{A}^*y \neq 0, \forall y \in \mathcal{F}_P$, and $\sigma_{\min}(\mathcal{A}_c^*) = \min \{ \sigma_{\min}(\mathcal{A}^*), \|c_{\mathcal{N}(\mathcal{A})}\| \} > 0$. Therefore the cosine of the angle θ_{d^*z} between d^* and $z = c - \mathcal{A}^*y$ is bounded by

$$\begin{aligned} \cos \theta_{d^*z} &= \frac{\langle d^*, c - \mathcal{A}^*y \rangle}{\|d^*\| \|c - \mathcal{A}^*y\|} \leq \frac{\left\langle \mathcal{A}_c d^*, \begin{pmatrix} -y \\ 1 \end{pmatrix} \right\rangle}{\left\| \mathcal{A}_c^* \begin{pmatrix} -y \\ 1 \end{pmatrix} \right\|} \\ &\leq \frac{\|\mathcal{A}_c d^*\| \left\| \begin{pmatrix} -y \\ 1 \end{pmatrix} \right\|}{\sigma_{\min}(\mathcal{A}_c^*) \left\| \begin{pmatrix} -y \\ 1 \end{pmatrix} \right\|} \\ &= \frac{\delta^*}{\min \{ \sigma_{\min}(\mathcal{A}^*), \|c_{\mathcal{N}(\mathcal{A})}\| \}} \end{aligned}$$

356

□

357 Theorem 4.1 provides a lower bound for the angle and distance between feasible slack vectors
 358 and the vector d^* on the boundary of \mathbb{S}_+^n . For our purposes, the theorem is only useful when
 359 $\alpha(\mathcal{A}, c)$ is small. Given that $\delta^* = \|\mathcal{A}_c d^*\|$, while the lower bound is independent on the scaling
 360 (conditioning) of \mathcal{A}_c , δ^* provides qualitative information about both the scaling of \mathcal{A}_c and the
 361 distance to infeasibility.

Corollary 4.2. *Let δ^*, d^* denote the optimum of the auxiliary problem (2.4), as in Theorem 4.1; and let*

$$d^* = [P \quad Q] \begin{bmatrix} d_+ & 0 \\ 0 & d_\epsilon \end{bmatrix} [P \quad Q]^T, \quad (4.3)$$

with $U = [P \ Q]$ orthogonal, and $d_+ \succ 0$. Then for any non-zero feasible slack $z = c - \mathcal{A}^*y \succeq 0$,

$$z_Q := QQ^T z QQ^T \in \operatorname{argmin}_q \{\|z - q\| : \langle d, q \rangle = 0, q \succeq 0\}, \quad (4.4)$$

and the cosine of the angle between z and z_Q satisfies

$$\cos \theta_{z, z_Q} := \frac{\langle z, z_Q \rangle}{\|z\| \|z_Q\|} = \frac{\|Q^T z Q\|}{\|z\|} \geq 1 - \alpha(\mathcal{A}, c) \frac{\|d^*\|}{\lambda_{\min}(d_+)}, \quad (4.5)$$

and

$$\|z - z_Q\|^2 \leq 2\|z\|^2 \left[\alpha(\mathcal{A}, c) \frac{\|d^*\|}{\lambda_{\min}(d_+)} \right], \quad (4.6)$$

that is, the angle between any feasible slack and the face $\mathcal{R}(Q \cdot Q^T)$ cannot be too large in the sense that

$$\inf_{z=c-\mathcal{A}^*y \succeq 0} \cos \theta_{z, z_Q} \geq 1 - \alpha(\mathcal{A}, c) \frac{\|d^*\|}{\lambda_{\min}(d_+)}.$$

362 The proof of Corollary 4.2 can be found in Appendix A. These results are related to the extreme
 363 angles between vectors in a cone studied in [18, 15]. Moreover, it is related to the distances to
 364 infeasibility in e.g., [24], in which the distance to infeasibility is shown to provide backward and
 365 forward error bounds.

366 We now see that we can use the rotation $U = [P \ Q]$ obtained from the diagonalization of the
 367 optimal d^* in the auxiliary problem (2.4) to reveal *nearness to infeasibility*, as discussed in e.g., [24].
 368 Or, in our approach, this reveals nearness to a facial decomposition. We use the following results
 369 to bound the size of certain blocks of a feasible slack z .

Corollary 4.3. Let δ^*, d^* denote the optimum of the auxiliary problem (2.4), as in Theorem 4.1; and let

$$d^* = [P \ Q] \begin{bmatrix} d_+ & 0 \\ 0 & d_\epsilon \end{bmatrix} [P \ Q]^T, \quad (4.7)$$

with $U = [P \ Q]$ orthogonal, and $d_\epsilon \succeq 0, d_+ \succ 0$. Then for any feasible slack $z = c - \mathcal{A}^*y \succeq 0$,

$$\operatorname{trace} P^T z P \leq \alpha(\mathcal{A}, c) (\|d_+^{-1}\|_2 \|d^*\|) \|z\| \leq \alpha(\mathcal{A}, c) \frac{\|d_+^{-1}\|_2 \|d^*\|}{\sqrt{n}} \operatorname{trace} z. \quad (4.8)$$

Proof. Since

$$\begin{aligned} \langle d^*, z \rangle &= \operatorname{trace} \begin{bmatrix} d_+ & 0 \\ 0 & d_\epsilon \end{bmatrix} \begin{bmatrix} P^T z P & P^T z Q \\ Q^T z P & Q^T z Q \end{bmatrix} \\ &= \operatorname{trace} d_+ P^T z P + d_\epsilon Q^T z Q \\ &\geq \operatorname{trace} d_+ P^T z P \\ &\geq \lambda_{\min}(d_+) \operatorname{trace} P^T z P, \end{aligned} \quad (4.9)$$

370 the claim follows from Theorem 4.1. □

371 4.2 Equivalent problems

372 With a shift of c in (1.1), we may use the result from Theorem 4.1 to get three optimization problems
 373 equivalent to (P). The equivalent problems indicates that, in the case when δ^* is sufficiently small,
 374 it is possible to reduce the dimension of the problem and get a “nearby” problem, under the
 375 assumption that the feasible region is bounded.

376 First we need to find a suitable shift of c , whose use will become apparent in Theorem 4.5:

Lemma 4.4. *Let $\delta^*, d^*, U = [P \ Q], d_+, d_\epsilon$ be defined as in the hypothesis of Corollary 4.3. If $(y_Q, w_Q) \in \mathbb{R}^m \times \mathbb{S}^{\bar{n}}$ is a best least square solution to the equation $Qw_QQ^T + \mathcal{A}^*y_Q = c$, that is, (y_Q, w_Q) is an optimal solution to the linear least square problem*

$$\min_{y,w} \frac{1}{2} \|c - (Qw_QQ^T + \mathcal{A}^*y)\|^2, \quad (4.10)$$

then, letting $c_{\text{res}} := c - (Qw_QQ^T + \mathcal{A}^*y_Q)$,

$$Q^T c_{\text{res}} Q = 0, \quad \text{and} \quad \mathcal{A}(c_{\text{res}}) = 0. \quad (4.11)$$

Moreover, if $\delta^* = 0$, then for any feasible solution y of (1.1),

$$c - \mathcal{A}^*y \in \mathcal{R}(Q \cdot Q^T),$$

377 so $(y, Q^T(c - \mathcal{A}^*y)Q)$ is an optimal solution of (2.12), whose optimal value is zero.

Proof. Let $\Omega(y, w) := \frac{1}{2} \|c - (Qw_QQ^T + \mathcal{A}^*y)\|^2$. Since

$$\Omega(y, w) = \frac{1}{2} \|c\|^2 + \frac{1}{2} \|\mathcal{A}^*y\|^2 + \frac{1}{2} \|w\|^2 + \langle Qw_QQ^T, \mathcal{A}^*y \rangle - \langle Q^T c Q, w \rangle - \langle \mathcal{A}c, y \rangle,$$

we have

$$\nabla_y \Omega = \mathcal{A} [Qw_QQ^T - (c - \mathcal{A}^*y)], \quad (4.12)$$

$$\text{and} \quad \nabla_w \Omega = w - [Q^T (c - \mathcal{A}^*y) Q]. \quad (4.13)$$

378 Hence (y_Q, w_Q) solves (2.12) if, and only if, $\nabla_y \Omega(y_Q, w_Q) = 0$ and $\nabla_w \Omega(y_Q, w_Q) = 0$. Then (4.11)
 379 follows immediately by substitution.

If $\delta^* = 0$, then $\langle d^*, A_i \rangle = 0$ for $i = 1, \dots, m$ and $\langle d^*, c \rangle = 0$. Hence, for any $y \in \mathbb{R}^m$,

$$\langle d_+, P^T(c - \mathcal{A}^*y)P \rangle + \langle d_\epsilon, Q^T(c - \mathcal{A}^*y)Q \rangle = \langle d^*, c - \mathcal{A}^*y \rangle = 0.$$

380 If $c - \mathcal{A}^*y \succeq 0$, we must have $P^T(c - \mathcal{A}^*y)P = 0$ (as $d_+ \succ 0$), and so $P^T(c - \mathcal{A}^*y)Q = 0$. Hence
 381 $QQ^T(c - \mathcal{A}^*y)QQ^T = c - \mathcal{A}^*y$. \square

382 We can now use the rotation from Theorem 4.1 with a shift of c (to $c_{\text{res}} + c_Q$) to get three
 383 equivalent problems to (P).

Theorem 4.5. *Let $\delta^*, d^*, U = [P \ Q], d_+, d_\epsilon$ be defined as in the hypothesis of Corollary 4.3. Let \hat{x} satisfy $\mathcal{A}\hat{x} = b$ and let (w_Q, y_Q) be a best least squares solution of the equation $Qw_QQ^T + \mathcal{A}^*y = c$. Define the matrices*

$$c_Q := Qw_QQ^T, \quad c_{\text{res}} := c - (c_Q + \mathcal{A}^*y_Q)$$

the scalar

$$\beta := \alpha(\mathcal{A}, c) \frac{\|d_+^{-1}\|_2 \|d^*\|}{\sqrt{n}},$$

and the following cone $T_\beta \subseteq \mathbb{S}_+^n$ partitioned appropriately with the matrix in (4.15).

$$T_\beta := \left\{ z = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \in \mathbb{S}_+^n : \text{trace } A \leq \beta \text{trace } z \right\}. \quad (4.14)$$

384 Then we get the following equivalent programs to (P) in (1.1):

1. using the rotation U ,

$$\begin{aligned} v_P &= \sup \{ b^T y : [P \ Q]^T z [P \ Q] \succeq 0, z = c - \mathcal{A}^* y \} \\ (P)_{PQ1} &= \sup \left\{ b^T y : \begin{bmatrix} P^T z P & P^T z Q \\ Q^T z P & Q^T z Q \end{bmatrix} \succeq 0, z = c - \mathcal{A}^* y \right\}; \end{aligned} \quad (4.15)$$

2. using the rotation U and the cone T_β ,

$$(P)_{PQ,\alpha 2} \quad v_P = \sup \left\{ b^T y : \begin{bmatrix} P^T z P & P^T z Q \\ Q^T z P & Q^T z Q \end{bmatrix} \succeq_{T_\beta} 0, z = c - \mathcal{A}^* y \right\}; \quad (4.16)$$

3. with the assumption that (w_Q, y_Q) is a best least square solution of the equation $Qw_Q Q^T + \mathcal{A}^* y_Q = c$ (as in Lemma 4.4), $c_{\text{res}} := c - (c_Q + \mathcal{A}^* y_Q)$ and $c_Q := Qw_Q Q^T$,

$$(P)_{PQ,\hat{x}3} \quad v_P = b^T y_Q + \sup \left\{ b^T y : \begin{bmatrix} P^T z P & P^T z Q \\ Q^T z P & Q^T z Q \end{bmatrix} \succeq_{T_\beta} 0, z = c_{\text{res}} + c_Q - \mathcal{A}^* y \right\} \quad (4.17)$$

Proof. The equivalence of (1.1) with (4.15) follows from the congruence $z \succeq 0$ if, and only if, $[P \ Q]^T z [P \ Q] \succeq 0$, and that with (4.16) follows from Corollary 4.3. For (4.17), first note that for any $y \in \mathbb{R}^m$,

$$z := c_{\text{res}} + c_Q - \mathcal{A}^* y \succeq 0 = c - \mathcal{A}^*(y + y_Q),$$

so $z \succeq 0$ implies $y + y_Q \in \mathcal{F}_Q$ $z \in T_\beta$. Hence

$$\begin{aligned} v_p &= \sup_{y,z} \left\{ b^T y : \begin{bmatrix} P^T z P & P^T z Q \\ Q^T z P & Q^T z Q \end{bmatrix} \succeq 0, z = c - \mathcal{A}^* y = c_{\text{res}} + c_Q - \mathcal{A}^*(y - y_Q) \right\} \\ &= \sup_{\tilde{y},z} \left\{ b^T (\tilde{y} + y_Q) : \begin{bmatrix} P^T z P & P^T z Q \\ Q^T z P & Q^T z Q \end{bmatrix} \succeq 0, z = c_{\text{res}} + c_Q - \mathcal{A}^* \tilde{y} \right\} \\ &= b^T y_Q + \sup_{\tilde{y},z} \left\{ b^T \tilde{y} : \begin{bmatrix} P^T z P & P^T z Q \\ Q^T z P & Q^T z Q \end{bmatrix} \succeq_{T_\beta} 0, z = c_{\text{res}} + c_Q - \mathcal{A}^* \tilde{y} \right\}. \end{aligned}$$

385

□

386 **Remark 4.6.** Results from Theorem 4.1, Corollaries 4.3 and 4.2 still hold for the optimization
387 problem in (4.17), if we replace the feasibility condition $z = c - \mathcal{A}^* y \succeq 0$ by $z = c_{\text{res}} + c_Q - \mathcal{A}^* y \succeq 0$.

388 **4.3 Reduction to a smaller problem**

389 **4.3.1 The $\delta^* = 0$ case**

390 In the case that $\delta^* = 0$, we know that the primal feasible region lies on the face $QS_+^{\bar{n}}Q^T \triangleleft S_+^n$. The
 391 case that $\delta^* = 0$ allows for a reduction to a face $QS_+^{\bar{n}}Q^T \triangleleft S_+^n$. Here we repeat Corollary 2.15.

Corollary 4.7. *Suppose $\delta^* = 0$ and*

$$d^* = [P \quad Q] \begin{bmatrix} d_+ & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P^T \\ Q^T \end{bmatrix},$$

solve (AP), where $[P \quad Q]$ is orthogonal, $Q \in \mathbb{R}^{n \times \bar{n}}$, $d_+ \succ 0$. Suppose $\mathcal{R}(Q \cdot Q^T) \cap \mathcal{R}(\mathcal{A}^) \neq \{0\}$. Let (y_Q, w_Q) be an optimal solution to the least squares problem (2.12), $c_Q := c - \mathcal{A}^*y_Q$, and linear transformation \mathcal{P} be given as in Proposition 2.16. Let*

$$\bar{c} := Q^T c_Q Q = w_Q \in \mathbb{S}^{\bar{n}}; \quad \bar{b} := \mathcal{P}^*b \in \mathbb{R}^{\bar{m}}; \quad \bar{\mathcal{A}}^*(\cdot) := Q^T ((\mathcal{A}^*\mathcal{P})(\cdot)) Q : \mathbb{S}^{\bar{n}} \rightarrow \mathbb{R}^{\bar{m}}.$$

Then

$$\bar{\mathcal{A}} \text{ is onto; } \quad \mathcal{R}(\mathcal{P}) \subseteq \mathcal{N}(P^T(\mathcal{A}^*\cdot)P) \cap \mathcal{N}(P^T(\mathcal{A}^*\cdot)Q), \quad (4.18)$$

and we get the following equivalent reduced program to (P) in (1.1).

$$(P)_{PQ4} \quad v_P = b^T y_Q + \sup \left\{ \langle \bar{b}, v \rangle : \bar{\mathcal{A}}^*v \preceq_{\mathbb{S}_+^{\bar{n}}} \bar{c}, v \in \mathbb{R}^{\bar{m}} \right\}. \quad (4.19)$$

392 Therefore, we have an equivalent smaller problem. The key is finding the linear transformation
 393 \mathcal{P} to satisfy the range condition in (2.15). We can now apply the auxiliary problem to this smaller
 394 problem to see if we can further reduce the problem or if the Slater constraint qualification holds.

395 **4.3.2 The $\delta^* > 0$ case**

396 In general, we cannot determine whether $\delta^* = 0$ exactly. Therefore, we now consider the important
 397 case when the optimal δ^* is *very small* and positive, see Figure 1.

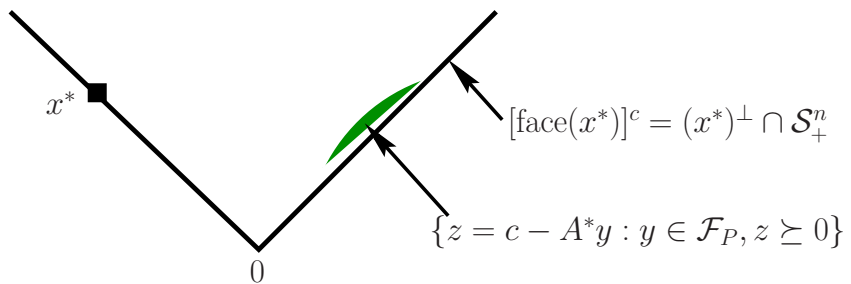


Figure 1: Minimal Face; $0 < \delta^* \ll 1$

398 A natural question is, when we find a very small and positive δ^* , if we decide to view it as equal
 399 to zero and project the primal problem to a smaller PSD cone as in the previous section, does the
 400 solution to the smaller problem correspond to some sensible solution to the original problem? The
 401 following theoretical result indicates that the smaller problem is “nearby” to the original one.

Theorem 4.8. *Following the notations as in Corollary 4.7 and assuming that $\mathcal{R}(Q \cdot Q^T) \cap \mathcal{R}(\mathcal{A}^*) \neq \{0\}$ (but δ^* is not assumed to be zero), for any $v \in \mathbb{R}^{\tilde{m}}$ satisfying $c_Q - \mathcal{A}^* \mathcal{L}v \succeq 0$, there exists $y (= \mathcal{L}\tilde{v} + \tilde{y}) \in \mathbb{R}^m$ such that $z := c_{\text{res}} + c_Q - \mathcal{A}^*y \succeq 0$, and*

$$\|y - \mathcal{L}v\| \leq 3\sqrt{2} \frac{\|z\|}{\sigma_{\min}(\mathcal{A})} \left[\alpha(\mathcal{A}, c) \frac{\|d^*\|}{\lambda_{\min}(d_+)} \right]^{1/2}.$$

402 The proof of Theorem 4.8 follows after the following two lemmas.

Lemma 4.9. *The norm of c_{res} is small in the sense that*

$$\|c_{\text{res}}\| \leq \sqrt{2} \left(\min_{z=c-\mathcal{A}^*y \succeq 0} \|z\| \right) \left[\alpha(\mathcal{A}, c) \frac{\|d^*\|}{\lambda_{\min}(d_+)} \right]^{1/2},$$

and for any $y \in \mathbb{R}^m$ such that $z := c_Q + c_{\text{res}} - \mathcal{A}^*y = c - \mathcal{A}^*(y + \tilde{y}) \succeq 0$,

$$\begin{aligned} \|\mathcal{A}^*y - QQ^T \mathcal{A}^*y QQ^T\| &\leq 2\|z - z_Q\| \\ &\leq 2\sqrt{2}\|z\| \left[\alpha(\mathcal{A}, c) \frac{\|d^*\|}{\lambda_{\min}(d_+)} \right]^{1/2}. \end{aligned}$$

Proof. First we prove the bound on the norm of c_{res} . Recall that y_Q, w_Q solve the least square problem

$$\min_{y, w} \frac{1}{2} \|c - (QwQ^T + \mathcal{A}^*y)\|^2,$$

and c_{res} is defined by

$$c_{\text{res}} := c - Qw_QQ^T - \mathcal{A}^*y_Q.$$

By optimality, for any $\tilde{y} \in \mathcal{F}_p$,

$$\|c_{\text{res}}\| \leq \|c - Q[Q^T(c - \mathcal{A}^*\tilde{y})Q]Q^T - \mathcal{A}^*\tilde{y}\| = \|\tilde{z} - QQ^T \tilde{z} QQ^T\|,$$

where $\tilde{z} := c - \mathcal{A}^*\tilde{y}$. Therefore

$$\|c_{\text{res}}\| \leq \min_{z=c-\mathcal{A}^*y \succeq 0} \sqrt{2}\|z\| \left[\alpha(\mathcal{A}, c) \frac{\|d^*\|}{\lambda_{\min}(d_+)} \right]^{1/2}$$

by Corollary 4.2. This proves the first inequality. The second inequality follows immediately from the first and Corollary 4.2:

$$\begin{aligned} \|\mathcal{A}^*y - QQ^T \mathcal{A}^*y QQ^T\| &\leq \|z - z_Q\| + \|c_{\text{res}}\| \\ &\leq 2\|z - z_Q\| \\ &\leq 2\sqrt{2}\|z\| \left[\alpha(\mathcal{A}, c) \frac{\|d^*\|}{\lambda_{\min}(d_+)} \right]^{1/2}. \end{aligned}$$

403

□

Lemma 4.10. *For any $y \in \mathbb{R}^m$, there exists $\tilde{y} \in \mathbb{R}^m$ and $\tilde{v} \in \mathbb{R}^{\tilde{m}}$ such that*

$$y = \mathcal{P}\tilde{v} + \tilde{y}, \text{ and } \mathcal{A} [QQ^T (\mathcal{A}^*\tilde{y}) QQ^T] = 0.$$

Proof. Consider the linear least square problem

$$\min_{v \in \mathbb{R}^m} \frac{1}{2} \|Q^T \mathcal{A}^*(y - \mathcal{P}v)Q\|^2. \quad (4.20)$$

Recall from Proposition 2.16 that the map $v \mapsto Q^T(\mathcal{A}^*\mathcal{P}v)Q$ is onto. Hence (4.20) has an optimal solution \tilde{v} . Letting $\tilde{y} := y - \mathcal{P}\tilde{v}$, the first order necessary optimality condition is given by

$$\mathcal{P}^* \mathcal{A} [QQ^T(\mathcal{A}^*\tilde{y})QQ^T] = 0.$$

From (2.16), we have that

$$\begin{aligned} \mathcal{N}(\mathcal{P}^*) &= \mathcal{N}(P^T(\mathcal{A}^*\cdot)P)^\perp + \mathcal{N}(P^T(\mathcal{A}^*\cdot)Q)^\perp \\ &= \mathcal{R}(\mathcal{A}(P \cdot P^T)) + \mathcal{R}(\mathcal{A}(P \cdot Q^T + Q \cdot^T P^T)), \end{aligned}$$

404 implying that $\mathcal{A} [QQ^T(\mathcal{A}^*\tilde{y})QQ^T] = 0$. □

Proof of Theorem 4.8. Consider the following convex program:

$$\begin{aligned} \min_{\tilde{v}, \tilde{y}} \quad & \frac{1}{2} \|\tilde{y} + \mathcal{P}(\tilde{v} - v)\|^2 \\ \text{s.t.} \quad & c_{\text{res}} + c_Q - \mathcal{A}^*(\tilde{y} + \mathcal{P}\tilde{v}) \succeq 0 \\ & \mathcal{P}^* \mathcal{A} [QQ^T(\mathcal{A}^*\tilde{y})QQ^T] = 0. \end{aligned} \quad (4.21)$$

By Lemma (4.10) and taking any y such that $c_{\text{res}} + c_Q - \mathcal{A}^*y \succeq 0$, (4.21) is feasible. Hence (4.21) must have an optimal solution. Moreover, all the constraints are linear, so Karush-Kuhn-Tucker condition is necessary and sufficient for optimality. The Lagrangian of (4.21) is given by

$$L(\tilde{y}, \tilde{v}, \Lambda, u) = \frac{1}{2} \|\tilde{y} + \mathcal{P}(\tilde{v} - v)\|^2 - \langle \Lambda, c_{\text{res}} + c_Q - \mathcal{A}^*(\tilde{y} + \mathcal{P}\tilde{v}) \rangle - \langle u, \mathcal{P}^* \mathcal{A} [QQ^T(\mathcal{A}^*\tilde{y})QQ^T] \rangle, \quad (4.22)$$

and

$$\begin{aligned} \nabla_{\tilde{y}} L &= \tilde{y} + \mathcal{P}(\tilde{v} - v) + \mathcal{A}\Lambda - \mathcal{A}[QQ^T(\mathcal{A}^*\mathcal{P}u)QQ^T], \\ \nabla_{\tilde{v}} L &= \mathcal{P}^*[\tilde{y} + \mathcal{P}(\tilde{v} - v)] + \mathcal{P}^* \mathcal{A}\Lambda. \end{aligned}$$

Recalling that $\mathcal{R}(\mathcal{A}^*\mathcal{P}) \subseteq \mathcal{R}(Q \cdot Q^T)$, the KKT conditions are given by

$$\tilde{y} + \mathcal{P}(\tilde{v} - v) + \mathcal{A}\Lambda = \mathcal{A}\mathcal{A}^*\mathcal{P}u, \quad (4.23)$$

$$\mathcal{P}^*[\tilde{y} + \mathcal{P}(\tilde{v} - v)] = -\mathcal{P}^* \mathcal{A}\Lambda, \quad (4.24)$$

$$\mathcal{P}^* \mathcal{A} [QQ^T(\mathcal{A}^*\tilde{y})QQ^T] = 0, \quad (4.25)$$

$$\langle \Lambda, c_{\text{res}} + c_Q - \mathcal{A}^*(\tilde{y} + \mathcal{P}\tilde{v}) \rangle = 0, \quad (4.26)$$

$$c_{\text{res}} + c_Q - \mathcal{A}^*(\tilde{y} + \mathcal{P}\tilde{v}) \succeq 0 \quad , \quad \Lambda \succeq 0. \quad (4.27)$$

From (4.23) and (4.24),

$$\mathcal{P}^* \mathcal{A}\mathcal{A}^*\mathcal{P}u = \mathcal{P}^*[\tilde{y} + \mathcal{P}(\tilde{v} - v) + \mathcal{A}\Lambda] = 0,$$

hence $\mathcal{A}\mathcal{A}^*\mathcal{P}u = 0$:

$$\|\mathcal{A}\mathcal{A}^*\mathcal{P}u\|^2 \leq \|\mathcal{A}\|^2 \|\mathcal{A}^*\mathcal{P}u\|^2 = \|\mathcal{A}\|^2 \langle u, \mathcal{P}^* \mathcal{A}\mathcal{A}^*\mathcal{P}u \rangle = 0.$$

Consequently, by complementary slackness,

$$\begin{aligned}
\langle \Lambda, c_{\text{res}} + c_Q - \mathcal{A}^* \mathcal{P}v \rangle &= \langle \Lambda, \mathcal{A}^*(\tilde{y} + \mathcal{P}(\tilde{v} - v)) \rangle \\
&= \langle \mathcal{A}\Lambda, \tilde{y} + \mathcal{P}(\tilde{v} - v) \rangle \\
&= \langle \mathcal{A}\mathcal{A}^* \mathcal{P}u - [\tilde{y} + \mathcal{P}(\tilde{v} - v)], \tilde{y} + \mathcal{P}(\tilde{v} - v) \rangle \\
&= -\|\tilde{y} + \mathcal{P}(\tilde{v} - v)\|^2.
\end{aligned}$$

If $\Lambda = 0$, we have $\tilde{y} + \mathcal{P}(\tilde{v} - v) = 0$. For proving Theorem 4.8, we may assume without loss of generality that $\Lambda \neq 0$. Observe that, since $c_Q - \mathcal{A}^* \mathcal{P}v \succeq 0$,

$$\begin{aligned}
\|\tilde{y} + \mathcal{P}(\tilde{v} - v)\|^2 &= -\langle \Lambda, c_{\text{res}} + c_Q - \mathcal{A}^* \mathcal{P}v \rangle \\
&\leq -\langle \Lambda, c_{\text{res}} \rangle \\
&= -\langle \Lambda, c_{\text{res}} + c_Q - \mathcal{A}^*(\tilde{y} + \mathcal{P}\tilde{v}) - [c_Q - QQ^T(\mathcal{A}^*(\tilde{y} + \mathcal{P}\tilde{v}))QQ^T] \\
&\quad + [\mathcal{A}^*(\tilde{y} + \mathcal{P}\tilde{v}) - QQ^T[\mathcal{A}^*(\tilde{y} + \mathcal{P}\tilde{v})]QQ^T] \rangle \\
&\leq \|\Lambda\| [\|z - QQ^T z QQ^T\| + \|\mathcal{A}^*(\tilde{y} + \mathcal{P}\tilde{v}) - QQ^T[\mathcal{A}^*(\tilde{y} + \mathcal{P}\tilde{v})]QQ^T\|],
\end{aligned}$$

where $z := c_{\text{res}} + c_Q - \mathcal{A}^*(\tilde{y} + \mathcal{P}\tilde{v})$. Note that the bound in Corollary 3.15 applies on z too. We also know how to bound the very last norm from Lemma 4.9. Finally, since $\tilde{y} + \mathcal{P}(\tilde{v} - v) = -\mathcal{A}\Lambda$, we have

$$\|\tilde{y} + \mathcal{P}(\tilde{v} - v)\|^2 \geq \sigma_{\min}(\mathcal{A})\|\Lambda\|\|\tilde{y} + \mathcal{P}(\tilde{v} - v)\|.$$

In conclusion,

$$\|\tilde{y} + \mathcal{P}(\tilde{v} - v)\| \leq \frac{3\sqrt{2}\|z\|}{\sigma_{\min}(\mathcal{A})} \left[\alpha(\mathcal{A}, c) \frac{\|d^*\|}{\lambda_{\min}(d_+)} \right]^{1/2}$$

405

□

406 Theorem 4.8 also provides a bound on the difference between the optimal values of the original
407 problem and the smaller problem:

Corollary 4.11. *Let*

$$\bar{v}_P := \sup\{\bar{b}^T v : \bar{c} - \bar{\mathcal{A}}^* v \succeq 0\},$$

where \bar{b} and $\bar{\mathcal{A}}$, \bar{b} , \bar{c} are defined as in Corollary 4.7. Then

$$\bar{v}_P - v_P \leq \frac{\|b\|}{\sigma_{\min}(\mathcal{A})} \left\{ \sqrt{2}\|z\| \left[2 \frac{\delta^*}{\sigma_{\min}(\mathcal{A}_c)} \frac{\|d^*\|}{\lambda_{\min}(d_+)} \right]^{1/2} + \frac{\delta^*}{\lambda_{\min}(d_+)} + \frac{\lambda_{\max}(d_\epsilon)}{\lambda_{\min}(d_+)} \|c\| \right\}. \quad (4.28)$$

Proof. For any v such that $\bar{c} - \bar{\mathcal{A}}^* v \succeq 0$, there exists $y \in \mathbb{R}^m$ such that $c_{\text{res}} + c_Q - \mathcal{A}^* y \succeq 0$. By Theorem 4.8,

$$\begin{aligned}
\langle \bar{b}, v \rangle - \langle b, y \rangle &\leq |\langle b, \mathcal{P}v - y \rangle| \\
&\leq 3\sqrt{2}\|b\| \frac{\|z\|}{\sigma_{\min}(\mathcal{A})} \left[\alpha(\mathcal{A}, c) \frac{\|d^*\|}{\lambda_{\min}(d_+)} \right]^{1/2}.
\end{aligned}$$

408

Taking supremum over all v such that $\bar{c} - \bar{\mathcal{A}}^* v \succeq 0$ gives (4.28). □

409

We can use Theorem 4.1 to provide a lower bound on $\bar{v}_P - v_P$.

410 5 Numerical Results

411 5.1 Worst case instance

From Tunçel [41], we consider the following problem instance with $n \geq 3$: $b = e_2 \in \mathbb{R}^n$, $c = 0$, $\mathcal{A} : \mathbb{S}_+^n \rightarrow \mathbb{R}^n$ defined by

$$A_1 = e_1 e_1^T, A_2 = e_1 e_2^T + e_2 e_1^T, A_i = e_{i-1} e_{i-1}^T + e_1 e_i^T + e_i e_1^T \text{ for } i = 3, \dots, n.$$

It is easy to see that

$$\mathcal{F}_p = \{c - \mathcal{A}^*y \in \mathbb{S}_+^n : y \in \mathbb{R}^n\} = \{\mu e_1 e_1^T : \mu \geq 0\},$$

(so \mathcal{F}_p does not contain interior points) and

$$\begin{aligned} \sup\{\langle b, y \rangle : c - \mathcal{A}^*y \succeq 0\} &= \sup\{y_2 : -\mathcal{A}^*y = \mu e_1 e_1^T, \mu \geq 0\} \\ &= \sup\{y_2 : -\mathcal{A}^*y = -\mu e_1, \mu \geq 0\} = 0, \end{aligned}$$

412 which is attained by any feasible solution.

Now consider the auxiliary problem

$$\min \|\mathcal{A}_c d\| = \left[d_{11}^2 + 4d_{12}^2 + \sum_{i=3}^n (d_{i-1,i-1} + 2d_{1i}) \right]^{1/2} \quad \text{s.t.} \quad \langle d, e \rangle = 1, d \succeq 0.$$

413 An optimal solution is $d^* = \sqrt{n} e_n e_n^T$, which attains objective value zero. It is easy to see this is
 414 the only solution⁶. Therefore, Q from Corollary 4.2 must have $n - 1$ columns, implying that the
 415 reduced problem is in \mathbb{S}^{n-1} . Theoretically, each facial reduction step via the auxiliary problem can
 416 only reduce the dimension by one. Moreover, after each reduction step, we get the same SDP with
 417 n reduced by one. Hence it would take $n - 1$ facial reduction steps before a reduced problem with
 418 strictly feasible solutions is found. This realizes the result in [7] on the upper bound of the number
 419 of facial reduction steps needed.

420 5.2 Generating instances with finite nonzero duality gaps

421 In this section we give the theory and procedure for generating SDP instances with finite nonzero
 422 duality gaps. Finite nonzero duality gaps and strict complementarity are closely tied together for
 423 cone optimization problems; using the concept of a *complementarity partition*, we can generate
 424 instances that fail to have strict complementarity; these in turn can be used to generate instances
 425 with finite nonzero duality gaps.

We present these results in a symmetric form. First, set $\bar{s} = c$, fix $\bar{x} \in \mathcal{F}_p$, and let \mathcal{L} denote the nullspace of \mathcal{A} , $\mathcal{L} = \mathcal{N}(\mathcal{A})$. Note that

$$\begin{aligned} -\langle y, b \rangle &= -\langle y, \mathcal{A}\bar{x} \rangle + \langle c, \bar{x} \rangle - \langle c, \bar{x} \rangle \\ &= \langle s, \bar{x} \rangle - \langle \bar{s}, \bar{x} \rangle, \quad \text{with } s = c - \mathcal{A}^*y. \end{aligned}$$

⁶Any solution d attaining objective value 0 must satisfy $d_{11} = 0$, and by the positive semidefiniteness constraint $d_{1i} = 0$ for $i = 2, \dots, n$ and so $d_{ii} = 0$ for $i = 2, \dots, n - 1$. d_{nn} is the only nonzero entry.

Therefore, with given data $(\mathcal{L}, \bar{x}, \bar{s}, K)$, $\bar{s} = c$, $\mathcal{A}\bar{x} = b$, we can now write the primal (1.1) as

$$(P) \quad v_P = \langle \bar{s}, \bar{x} \rangle - \inf_s \left\{ \langle s, \bar{x} \rangle : s \in (\bar{s} + \mathcal{L}^\perp) \cap K \right\}, \quad (5.1)$$

and the dual program in (1.2) becomes

$$(D) \quad v_D = \inf_x \left\{ \langle \bar{s}, x \rangle : x \in (\bar{x} + \mathcal{L}) \cap K^* \right\}. \quad (5.2)$$

Denote the minimal faces for the homogeneous problems (recession directions) by

$$f_P^0 := \text{face} \left(\mathcal{L}^\perp \cap K \right), \quad (5.3)$$

$$f_D^0 := \text{face} \left(\mathcal{L} \cap K^* \right). \quad (5.4)$$

426 Then $f_P^0 \subseteq (f_D^0)^c$ (equivalently, $f_D^0 \subseteq (f_P^0)^c$). We now introduce the notion of a *complementarity*
427 *partition*, see e.g., [40].

428 **Definition 5.1.** *The pair of faces $F_1 \trianglelefteq K, F_2 \trianglelefteq K^*$ form a complementarity partition of K, K^* if*
429 *$F_1 \subseteq F_2^c$. (Equivalently, $F_2 \subseteq F_1^c$.) The partition is proper if both F_1 and F_2 are proper faces. The*
430 *partition is strict if $(F_1)^c = F_2$ or $(F_2)^c = F_1$.*

At the first step, we would like to generate homogeneous problems with data $\mathcal{A} : \mathbb{S}^n \rightarrow \mathbb{R}^m$ whose primal-dual minimal faces do not form a strict complementary partition, that is, minimal faces of the form

$$f_D^0 = \begin{pmatrix} \mathbb{S}_+^{r_1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \trianglelefteq \mathbb{S}_+^n, \quad f_P^0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mathbb{S}_+^{r_3} \end{pmatrix} \trianglelefteq \mathbb{S}_+^n. \quad (5.5)$$

We assume that the matrices $A_i \in \mathbb{S}^n$ that define \mathcal{A} (and more generally all $X \in \mathbb{S}^n$) are partitioned into blocks with diagonal blocks made up of symmetric matrices of size r_1, r_2, r_3 , respectively. E.g.,

$$X = \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{12}^T & X_{22} & X_{23} \\ X_{13}^T & X_{23}^T & X_{33} \end{pmatrix} \in \mathbb{S}^n, \quad X_{ii} \in \mathbb{S}^{r_i}, \quad i \in \{1, 2, 3\}.$$

431 We choose linearly independent matrices $A_i, i = 1, \dots, m$ as follows:

1. As a first step, we choose $A_i \succeq 0, i = 1, \dots, p \leq r_3$ so that the generated face $f_P^0 := \text{face}(\{A_i\}_{i=1}^p)$ satisfies $\dim(f_P^0) = r_3$. After performing a rotation if necessary, we have the face $f_P^0 \trianglelefteq \mathbb{S}_+^n$ of the form (5.5). Define the subspace \mathcal{L}_3 by $\mathcal{L}_3^\perp := \text{span}\{A_i\}_{i=1}^p$. This choice implies that

$$\mathcal{L}_3^\perp = \text{span}\{A_i\}_{i=1}^p \subset \text{span}(f_P^0) = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & X_{33} \end{pmatrix} \in \mathbb{S}^n : X_{33} \in \mathbb{S}^{r_3} \right\}. \quad (5.6)$$

In addition, this choice yields

$$\mathcal{L}_3 \cap \mathbb{S}_+^n = (f_P^0)^c = \left\{ \begin{pmatrix} X_{11} & X_{12} & 0 \\ X_{21} & X_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{S}_+^n : X_{11} \in \mathbb{S}_+^{r_1}, X_{22} \in \mathbb{S}_+^{r_2} \right\}.$$

2. If we choose $\mathcal{L} = \mathcal{L}_3$, with \mathcal{L}_3 defined using (5.6), then $f_D^0 = (f_P^0)^c$, i.e., we have a strict complementarity partition. However, we want to allow f_D^0 to be smaller while maintaining f_P^0 . We now allow additional linearly independent $A_i \in \mathbb{S}^n, i = p + 1, \dots, m$, indefinite, so that

$$\text{face}(\text{span}\{A_i\}_{i=1}^m \cap \mathbb{S}_+^n) = f_P^0, \quad (5.7)$$

but

$$\{0\} \neq \left(\bigcap_{i=p+1}^m \{A_i\}^\perp\right) \cap (f_P^0)^c \subsetneq (f_D^0)^c, \quad (5.8)$$

432 i.e., the face f_P^0 is *not* increased by the addition of these indefinite A_i , but the dual face f_D^0
433 is decreased.

Note that it is simple to guarantee that both (5.7) and (5.8) hold. Recall that the first p semidefinite matrices $A_i \succeq 0$. We could then choose $m = r_1 + r_2 + 1$ with the single block indefinite matrix

$$A_m = \begin{pmatrix} 0 & 0 & A_{13} \\ 0 & A_{22} & A_{23} \\ A_{13}^T & A_{23}^T & A_{33} \end{pmatrix}, \text{ with } A_{22} \succ 0, \text{ and } A_{13} \text{ nonzero,}$$

434 i.e., within the first $r_1 + r_2$ diagonal elements of A_m : we have at least one diagonal element positive;
435 and, we have at least one diagonal element 0 with a corresponding nonzero element in the same
436 row. This choice guarantees that (5.5) holds. These two faces form a *complementarity partition*,
437 $\langle f_P^0, f_D^0 \rangle = 0$, but not a *strict complementarity partition*, i.e. $(f_P^0)^c \cap (f_D^0)^c \neq \{0\}$.

438 The following lemma provides a general condition for choosing the operator \mathcal{A} to ensure that the
439 feasible sets of both the primal and dual recession fail to form a strict complementarity partition.

Lemma 5.2. *Given problem dimensions m and n as above, let r_1, r_2 , and r_3 be positive integers that sum to n . Let $\{A_i\}_{i=1}^m$ be a set of linearly independent matrices such that $A_i \succeq 0, i = 1, \dots, p$, and the cone generated by the matrices $\{A_i\}_{i=1}^p$ takes the form*

$$\text{face}\{A_i\}_{i=1}^p = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mathbb{S}_+^{r_3} \end{pmatrix} \leq \mathbb{S}_+^n,$$

where the diagonal blocks have size r_1, r_2 , and r_3 , respectively. Now let \mathcal{P}_k , for $k = 22$ and $k = 13$, denote the projections onto the $k = 22$ and $k = 13$ blocks, respectively,

$$\mathcal{P}_{22}(x) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & X_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{P}_{13}(x) = \begin{pmatrix} 0 & 0 & X_{13} \\ 0 & 0 & 0 \\ X_{13}^T & 0 & 0 \end{pmatrix},$$

and define the linear transformations $B_k : \mathbb{S}^n \rightarrow \mathbb{R}^m$, for $k = 22$ and $k = 13$, by $(B_k(x))_i = \langle A_i, \mathcal{P}_k(x) \rangle$, for $i = 1, \dots, m$. If

$$A_i = \begin{pmatrix} 0 & 0 & (A_i)_{13} \\ 0 & (A_i)_{22} & * \\ (A_i)_{13}^T & * & * \end{pmatrix} \in \mathbb{S}^n, \quad i = p + 1, \dots, m, \quad (5.9)$$

where the diagonal blocks have the same sizes r_1 , r_2 , and r_3 , respectively, and a star denotes a block having arbitrary elements, and if the two conditions on the peridiagonal blocks

$$\mathcal{N}(B_{22}) \cap \mathbb{S}_+^n = \{0\}, \quad (5.10)$$

$$B_{22}(\mathbb{S}_+^n) + \mathcal{R}(B_{13}) = \begin{pmatrix} 0 \\ \mathbb{R}^{m-p} \end{pmatrix} \subset \mathbb{R}^m, \quad (5.11)$$

440 hold, then the primal and dual faces satisfy (5.5), i.e., these two faces form a complementarity
441 partition but not a strict complementarity partition.

Proof. We first find $\mathcal{L}^\perp \cap \mathbb{S}_+^n = \mathcal{R}(\mathcal{A}^*) \cap \mathbb{S}_+^n$. Taking the dual of each side of (5.11), we obtain the implication

$$B_{22}^* z \succeq 0, \quad B_{13}^* z = 0 \Rightarrow z = 0. \quad (5.12)$$

442 That is, for any nonzero z , if $B_{13}^* z = 0$, then $B_{22}^* z \not\succeq 0$. However the $(1, 1)$ block of matrices lying
443 in $\mathcal{R}(\mathcal{A}^*) \cap \mathbb{S}_+^n$ is zero, so the $(1, 3)$ block must also be zero. Moreover, the $(2, 2)$ block of matrices
444 lying in $\mathcal{R}(\mathcal{A}^*) \cap \mathbb{S}_+^n$ must be positive semidefinite. Thus $\mathcal{A}^* y \in \mathcal{R}(\mathcal{A}^*) \cap \mathbb{S}_+^n$ implies that $y_i = 0$ for
445 $i > p$. Hence $\text{face}(\mathcal{L}^\perp \cap \mathbb{S}_+^n) = \mathcal{L}^\perp \cap \mathbb{S}_+^n$ takes the form given by f_P^0 in (5.5).

Now let us find $\mathcal{L} \cap \mathbb{S}_+^n = \{X \in \mathbb{S}_+^n : \langle A_i, X \rangle = 0, i = 1, \dots, m\}$. First observe that

$$\{X \in \mathbb{S}_+^n : \langle A_i, X \rangle = 0, i = 1, \dots, p\} = \begin{pmatrix} \mathbb{S}_+^{r_1+r_2} & 0 \\ 0 & 0 \end{pmatrix}.$$

Now it follows from (5.10) that the set $\{X \in \mathbb{S}_+^n : \langle A_i, X \rangle = 0, i = p+1, \dots, m\}$ has $(2, 2)$ block zero, which forces the second block row and column to be zero:

$$\{X \in \mathbb{S}_+^n : \langle A_i, X \rangle = 0, i = 1, \dots, m\} = \begin{pmatrix} \mathbb{S}_+^{r_1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

446 Hence $\text{face}(\mathcal{L} \cap \mathbb{S}_+^n) = \mathcal{L} \cap \mathbb{S}_+^n$ takes the form given by f_D^0 in (5.5).

Clearly $\langle f_P^0, f_D^0 \rangle = 0$, so $\text{face}(\mathcal{L} \cap \mathbb{S}_+^n)$ and $\text{face}(\mathcal{L}^\perp \cap \mathbb{S}_+^n)$ form a complementarity partition.
However

$$(f_P^0)^c \cap (f_D^0)^c = \begin{pmatrix} 0 & & \\ & \mathbb{S}_+^{r_2} & \\ & & 0 \end{pmatrix} \neq \{0\},$$

447 so the partition does not satisfy strict complementarity. \square

448 **Remark 5.3.** If the matrices $\{A_i\}_{i=1}^m$ in Lemma 5.2 are such that (a) the $(2, 2)$ block of A_j is
449 definite for some index $j > p$, and (b) the $(1, 3)$ blocks of $\{A_i\}_{i=p+1}^m$ form a linearly independent
450 set, then (5.10) and (5.11) are satisfied.

We may now choose the matrices \bar{x} and \bar{s} from (5.1) and (5.2) in the set $(f_P^0)^c \cap (f_D^0)^c$ to ensure that (1.1) and (1.2) possess a nonzero duality gap. Specifically, we choose \bar{x} and \bar{s} to satisfy

$$\bar{x} := \begin{pmatrix} 0 & 0 & 0 \\ 0 & \bar{X}_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} \succeq 0, \quad \bar{s} := \begin{pmatrix} 0 & 0 & 0 \\ 0 & \bar{S}_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} \succeq 0, \quad \text{with } \text{trace}(\bar{X}_{22}\bar{S}_{22}) > 0. \quad (5.13)$$

451 In Lemma 5.2 we gave conditions on the matrices $\{A_1, A_2, \dots, A_m\}$ such that a strict comple-
452 mentarity partition fails to exist. In the following theorem, we show constructively how (\mathcal{A}, b, c)
453 may be chosen so that the SDP instance has a finite positive duality gap.

454 **Theorem 5.4.** Suppose that the matrices A_i are such that $A_i, i = 1, \dots, p$ are as in Item 1 on
455 page 29, so that f_P^0 satisfies (5.5), and the remaining $A_i, i = p + 1, \dots, m$ are found to satisfy the
456 structure in (5.9). Suppose also that $\{(A_i)_{13}\}_{i=p+1}^m$ is a linearly independent set, and \bar{x}, \bar{s} are chosen
457 to satisfy (5.13), and $b = \mathcal{A}(\bar{x})$ and $c = \bar{s}$ in (1.1) and (1.2).

Consider the following conditions on \bar{x} , and \bar{s} :

$$\langle (A_i)_{22}, \ell \rangle = 0, \quad i = p + 1, \dots, m \implies \langle \bar{S}_{22}, \ell \rangle = 0, \quad (5.14)$$

$$\langle (A_i)_{22}, \bar{X}_{22} \rangle = 0, \quad i = p + 1, \dots, m. \quad (5.15)$$

458 The following statements holds.

- 459 1. All primal feasible points are optimal and $v_P = 0$.
- 460 2. If (5.14) holds, then all dual feasible points are optimal and $v_D = \langle \bar{X}_{22}, \bar{S}_{22} \rangle > 0$. It follows
461 that (1.1) and (1.2) possess a positive duality gap.
- 462 3. If (5.14) fails to hold, but \bar{S}_{22} is positive definite, and (5.15) fails to hold, then not all dual
463 feasible points are optimal, but $v_D > 0$. It follows that (1.1) and (1.2) possess a positive
464 duality gap.

Proof. 1. Since \bar{s} and all matrices in \mathcal{L}^\perp have a zero (1,1) block, all matrices in $(\bar{s} + \mathcal{L}^\perp) \cap K$ have a zero (1,1) block, and hence a zero (1,3) block. In light of the linear independence assumption on the (1,3) blocks of the A_i for $i = p + 1, \dots, m$, we see that

$$(\bar{s} + \mathcal{L}^\perp) \cap K = \bar{s} + (\mathcal{L}^\perp \cap K),$$

i.e.

$$\mathcal{F}_P^s \subset \mathcal{F}_P = \bar{s} + f_P^0.$$

465 Since $\langle \bar{x}, f_P^0 \rangle = 0$, it follows that all feasible slacks satisfy $\langle s, \bar{x} \rangle = \langle \bar{S}_{22}, \bar{X}_{22} \rangle$, and the primal
466 optimal value is $v_P = 0$ for all such points.

2. Using the special structure of the A_i , \bar{s} , and \bar{x} , the dual can be written as

$$v_D = \inf_{\ell \in \mathbb{R}^{r_2 \times r_2}} \{ \langle \bar{S}_{22}, \bar{X}_{22} + \ell \rangle : \bar{X}_{22} + \ell \succeq 0, \langle (A_i)_{22}, \ell \rangle = 0, \quad i = p + 1 \dots, m \}. \quad (5.16)$$

467 If (5.14) holds, then all feasible ℓ have an objective function value of $\langle \bar{S}_{22}, \bar{X}_{22} \rangle$. That is, all
468 feasible points are optimal, and $v_D = \langle \bar{S}_{22}, \bar{X}_{22} \rangle$.

- 469 3. If (5.14) fails to hold, then $\ell = 0$ is feasible but no longer optimal for (5.16). Equivalently,
470 $x = \bar{x}$ is feasible but no longer optimal for (1.2). However, since $\bar{S}_{22} \succ 0$ and $\bar{X}_{22} + \ell \succeq 0$
471 for any feasible ℓ for (5.16), either the inner product between these two matrices must be
472 bounded away from zero on the feasible set of (5.14), or $\ell = -\bar{X}_{22}$ is feasible, in which case
473 $v_D = 0$. The later case is ruled out since by assumption, (5.15) fails to hold. Hence $v_D > 0$,
474 so the duality gap is positive.

475 □

Name	n	m	Optimal values using SeDuMi <u>with</u> facial reduction	Optimal values using SeDuMi <u>without</u> facial reduction
Example 1	3	2	0	-6.30238e-016
Example 2	3	2	0	+0.570395
Example 3	3	4	0	+6.91452e-005
Example 4	3	3	0	+Inf
Example 5	10	5	+5.02950e+02	+5.02950e+02
Example 6	6	8	+1	+1
Example 7	5	3	0	-2.76307e-012
Example 9a	20	20	0	Inf
Example 9b	100	100	0	Inf
RandGen1	10	5	+1.5914e-015	+1.16729e-012
RandGen2	100	67	+1.1056e-010	NaN
RandGen3	200	140	+5.0557e-010	NaN
RandGen4	200	140	+1.02803e-009	NaN
RandGen5	120	45	-5.47393e-015	-1.63758e-015
RandGen6	320	140	+5.9077e-025	NaN
RandGen7	40	27	-5.2203e-029	+5.64118e-011
RandGen8	60	40	-2.03227e-029	NaN
RandGen9	60	40	+5.61602e-015	-3.52291e-012
RandGen10	180	100	+2.47204e-010	NaN
RandGen11	255	150	+7.71685e-010	NaN

Table 1: Comparisons with/without facial reduction

476 5.3 Numerical results from SeDuMi

477 Table 1 shows a comparison of solving SDP instances *with* versus *without* facial reduction. Examples
478 1 through 9 are specially generated problems available online at the URL for this paper given in
479 the footnote on page 1. In particular: Example 3 has a positive duality gap, $v_P = 0 < v_D = 1$;
480 for Example 4, the dual is infeasible; in Example 5, the Slater CQ holds; Examples 9a,9b are
481 instances of the worst case problems presented in Section 5.1. The remaining instances RandGen1-
482 RandGen11 are generated randomly with most of them having a finite positive duality gap, as
483 described in Section 5.2. These instances generically require only one iteration of facial reduction.
484 SeDuMi is used.

485 One general observation is that, if the instance has primal-dual optimal solutions and has zero
486 duality gap, SeDuMi is able to find the optimal solutions. However, if the instance has finite nonzero
487 duality gaps, and if the instance is not too small, SeDuMi is unable to compute any solution, and
488 returns NaN, because the stopping rule (that is, the duality gap being approximately zero) cannot
489 be met.

490 6 Conclusions and future work

491 In this paper we have presented a preprocessing technique for SDP problems where the Slater CQ
 492 (nearly) fails. This is based on solving a stable auxiliary problem that approximately identifies the
 493 minimal face for (P). We have included a backward error analysis and some preliminary tests that
 494 successfully solve problems where the CQ fails and also problems that have a duality gap. The
 495 optimal value of our (AP) has significance as a measure of *nearness to infeasibility*.

Though our stable (AP) satisfied both the primal and dual generalized Slater condition, high accuracy solutions were difficult to obtain for unstructured general problems. (AP) is equivalent to the underdetermined linear least square problem

$$\min \|\mathcal{A}_c d\|_2^2 \quad \text{s.t.} \quad \langle e, d \rangle = 1, \quad d \succeq 0, \quad (6.1)$$

496 which is known to be difficult to solve. High accuracy solutions are essential in performing a proper
 497 facial reduction.

498 Extensions of some of our results can be made to general conic convex programming, in which
 499 case the partial orderings in (1.1) and (1.2) are induced by a proper closed convex cone K and the
 500 dual cone K^* , respectively.

501 A Proof of Corollary 4.2

We would like to prove the following: let δ^*, d^* denote the optimum of the auxiliary problem (2.4), as in Theorem 4.1; and let

$$d^* = [P \quad Q] \begin{bmatrix} d_+ & 0 \\ 0 & d_\epsilon \end{bmatrix} [P \quad Q]^T, \quad (A.1)$$

with $U = [P \quad Q]$ orthogonal, and $d_+ \succ 0$. Then for any non-zero feasible slack $z = c - \mathcal{A}^* y \succeq 0$,

$$z_Q := QQ^T z QQ^T \in \operatorname{argmin} \{ \|z - q\| : \langle P d_+ P^T, q \rangle = 0, q \succeq 0 \}, \quad (A.2)$$

and cosine of the angle between z and z_Q satisfies

$$\cos \theta_{z, z_Q} := \frac{\langle z, z_Q \rangle}{\|z\| \|z_Q\|} = \frac{\|Q^T z Q\|}{\|z\|} \geq 1 - \alpha(\mathcal{A}, c) \frac{\|d^*\|}{\lambda_{\min}(d_+)}.$$

Also,

$$\|z - z_Q\|^2 \leq 2\|z\|^2 \alpha(\mathcal{A}, c) \frac{\|d^*\|}{\lambda_{\min}(d_+)}.$$

Proof. The orthogonality constraint $\langle P d_+ P^T, q \rangle = 0, q \succeq 0$ can be replaced using the substitution $q = Q w Q^T, w \succeq 0$. Then $Q^T z Q \in \operatorname{argmin}_{w \succeq 0} \|z - Q w Q^T\|$ yields the result (A.2). The expression for the angle follows using

$$\frac{\langle z, z_Q \rangle}{\|z\| \|z_Q\|} = \frac{\|Q^T z Q\|^2}{\|z\| \|Q^T z Q\|} = \frac{\|Q^T z Q\|}{\|z\|}. \quad (A.3)$$

From Theorem 4.1, the optimal value of the following optimization problem provides a lower bound on the quantity in (A.3).

$$\begin{aligned} \gamma_0 &:= \min_z && \|Q^T z Q\| \\ &\text{s.t.} && \langle z, d^* \rangle \leq \alpha(\mathcal{A}, c) \|d^*\| \\ &&& \|z\|^2 = 1, \quad z \succeq 0. \end{aligned} \quad (\text{A.4})$$

Since $\langle z, d^* \rangle = \langle P^T z P, d_+ \rangle + \langle Q^T z Q, d_\epsilon \rangle \geq \langle P^T z P, d_+ \rangle$ whenever $z \succeq 0$, we have

$$\begin{aligned} \gamma_0 \geq \gamma &:= \min_z && \|Q^T z Q\| \\ &\text{s.t.} && \langle P^T z P, d_+ \rangle \leq \alpha(\mathcal{A}, c) \|d^*\| \\ &&& \|z\|^2 = 1, \quad z \succeq 0. \end{aligned} \quad (\text{A.5})$$

It is possible to find the optimal value γ of (A.5). After the orthogonal rotation

$$z = \begin{bmatrix} P & Q \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} P & Q \end{bmatrix}^T = PAP^T + PBQ^T + QB^T P^T + QCQ^T,$$

where $A \in \mathbb{S}_+^{n-\bar{n}}$, $C \in \mathbb{S}_+^{\bar{n}}$ and $B \in \mathbb{R}^{(n-\bar{n}) \times \bar{n}}$, (A.5) can be rewritten as

$$\begin{aligned} \gamma &= \min_{A,B,C} && \|C\| \\ &\text{s.t.} && \langle A, d_+ \rangle \leq \alpha(\mathcal{A}, c) \|d^*\| \\ &&& \|A\|^2 + 2\|B\|^2 + \|C\|^2 = 1 \\ &&& \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \in \mathbb{S}_+^n. \end{aligned} \quad (\text{A.6})$$

Since

$$\|B\|^2 \leq \|A\| \|C\| \quad (\text{A.7})$$

holds whenever (A, B, C) is feasible for (A.6), the optimal value of (A.6) occurs only if equality holds in (A.7), i.e., $\|A\|^2 + 2\|B\|^2 + \|C\|^2 = (\|A\| + \|C\|)^2$. Therefore, we get the equivalent problem

$$\gamma = \min_{A,C} \{ \|C\| : \langle A, d_+ \rangle \leq \alpha(\mathcal{A}, c) \|d^*\|, \|A\| + \|C\| = 1, A \in \mathbb{S}_+^{n-\bar{n}}, C \in \mathbb{S}_+^{\bar{n}} \}. \quad (\text{A.8})$$

The optimal value of (A.8) can be obtained by solving the optimization problem

$$\max_A \{ \|A\| : \langle A, d_+ \rangle \leq \alpha(\mathcal{A}, c) \|d^*\|, A \in \mathbb{S}_+^{n-\bar{n}} \}. \quad (\text{A.9})$$

Let u be a normalized eigenvector for the smallest eigenvalue $\lambda_{\min}(d_+)$ and $\begin{bmatrix} V & u \end{bmatrix}$ be an orthogonal matrix. Denote the vector $v = \begin{pmatrix} \sqrt{\beta} u \\ \sqrt{1-\beta} e_1 \end{pmatrix}$ and set

$$\begin{aligned} \beta &:= \min \left\{ \frac{\alpha(\mathcal{A}, c) \|d^*\|}{\lambda_{\min}(d_+)}, 1 \right\}, && A := \beta u u^T, \\ C &= (1 - \|A\|) e_1 e_1^T, && B = \sqrt{\|A\| \|C\|} u e_1^T. \end{aligned}$$

Then

$$\left(\begin{bmatrix} \begin{bmatrix} V & u \\ 0 & I \end{bmatrix} & v \end{bmatrix} \right) \left(\begin{bmatrix} \begin{bmatrix} V & u \\ 0 & I \end{bmatrix} & v \end{bmatrix} \right)^T = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \in \mathbb{S}_+^n,$$

502 and (A, B, C) is feasible for (A.8) (and (A.6)), and the corresponding objective value yields the
 503 desired bound.

The last claim in the corollary follows immediately:

$$\begin{aligned} \|z - z_Q\|^2 &= \|z\|^2 \left(1 - \frac{\|Q^T z_Q\|^2}{\|Q\|^2}\right) \\ &\leq \|z\|^2 \left[1 - \left(1 - \alpha(\mathcal{A}, c) \frac{\|d^*\|}{\lambda_{\min}(d_+)}\right)^2\right] \\ &\leq 2\|z\|^2 \alpha(\mathcal{A}, c) \frac{\|d^*\|}{\lambda_{\min}(d_+)}. \end{aligned}$$

504

□

505 References

- 506 [1] A. Alfakih, A. Khandani, and H. Wolkowicz. Solving Euclidean distance matrix completion
 507 problems via semidefinite programming. *Comput. Optim. Appl.*, 12(1-3):13–30, 1999. A tribute
 508 to Olvi Mangasarian. 6
- 509 [2] M.F. Anjos and H. Wolkowicz. Strengthened semidefinite relaxations via a second lifting for
 510 the Max-Cut problem. *Discrete Appl. Math.*, 119(1-2):79–106, 2002. Foundations of heuristics
 511 in combinatorial optimization. 6
- 512 [3] A. Ben-Israel, A. Charnes, and K. Kortanek. Duality and asymptotic solvability over cones.
 513 *Bulletin of American Mathematical Society*, 75(2):318–324, 1969. 6
- 514 [4] B. Borchers. CSDP, a C library for semidefinite programming. *Optim. Methods Softw.*, 11/12(1-
 515 4):613–623, 1999. projects.coin-or.org/Csdp. 3
- 516 [5] J.M. Borwein and H. Wolkowicz. Characterization of optimality for the abstract convex pro-
 517 gram with finite-dimensional range. *J. Austral. Math. Soc. Ser. A*, 30(4):390–411, 1980/81. 3,
 518 6
- 519 [6] J.M. Borwein and H. Wolkowicz. Facial reduction for a cone-convex programming problem.
 520 *J. Austral. Math. Soc. Ser. A*, 30(3):369–380, 1980/81. 3, 6
- 521 [7] J.M. Borwein and H. Wolkowicz. Regularizing the abstract convex program. *J. Math. Anal.*
 522 *Appl.*, 83(2):495–530, 1981. 3, 6, 7, 28
- 523 [8] S. Boyd, V. Balakrishnan, E. Feron, and L. El Ghaoui. Control system analysis and synthesis
 524 via linear matrix inequalities. *Proc. ACC*, pages 2147–2154, 1993. 3
- 525 [9] X.V. Doan, S. Kruk, and H. Wolkowicz. A robust algorithm for semidefinite programming.
 526 Technical Report CORR 2010-09, University of Waterloo, Waterloo, Ontario, 2010. submitted
 527 in November, 2010. 4

- 528 [10] R.M. Freund. Complexity of an algorithm for finding an approximate solution of a semi-definite
529 program with no regularity assumption. Technical Report OR 302-94, MIT, Cambridge, MA,
530 1994. 4
- 531 [11] R.M. Freund. Complexity of convex optimization using geometry-based measures and a refer-
532 ence point. *Math. Program.*, 99(2, Ser. A):197–221, 2004. 4
- 533 [12] R.M. Freund, F. Ordóñez, and K.C. Toh. Behavioral measures and their correlation with IPM
534 iteration counts on semi-definite programming problems. USC-ISE working paper #2005-02,
535 MIT, 2005. url: www-rcf.usc.edu/~fordon/. 4
- 536 [13] R.M. Freund and J.R. Vera. Some characterizations and properties of the “distance to ill-
537 posedness” and the condition measure of a conic linear system. Technical report, MIT, Cam-
538 bridge, MA, 1997. 4
- 539 [14] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins University Press,
540 Baltimore, Maryland, 3rd edition, 1996. 17
- 541 [15] D. Gourion and A. Seeger. Critical angles in polyhedral convex cones: numerical and statistical
542 considerations. *Math. Programming*, 123(1):173–198, 2010. 21
- 543 [16] G. Gruber and F. Rendl. Computational experience with ill-posed problems in semidefinite
544 programming. *Comput. Optim. Appl.*, 21(2):201–212, 2002. 4
- 545 [17] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1990.
546 Corrected reprint of the 1985 original. 5
- 547 [18] A. Iusem and A. Seeger. Searching for critical angles in a convex cone. *Math. Program.*, 120(1,
548 Ser. B):3–25, 2009. 21
- 549 [19] N. Krislock and H. Wolkowicz. Explicit sensor network localization using semidefinite rep-
550 resentations and facial reductions. *SIAM Journal on Optimization*, 20(5):2679–2708, 2010.
551 6
- 552 [20] Z-Q. Luo, J.F. Sturm, and S. Zhang. Conic convex programming and self-dual embedding.
553 *Optim. Methods Softw.*, 14(3):169–218, 2000. 3, 6
- 554 [21] J. Malick, J. Povh, F. Rendl, and A. Wiegele. Regularization methods for semidefinite pro-
555 gramming. *SIAM Journal on Optimization*, 20(1):336–356, 2009. 4
- 556 [22] G. Pataki. Bad semidefinite programs: they all look the same. Technical report, Department
557 of Operations Research, University of North Carolina, Chapel Hill, 2011. 3, 5
- 558 [23] Gábor Pataki. On the closedness of the linear image of a closed convex cone. *Math. Oper.*
559 *Res.*, 32(2):395–412, 2007. 3
- 560 [24] J. Peña and J. Renegar. Computing approximate solutions for convex conic systems of con-
561 straints. *Math. Program.*, 87(3, Ser. A):351–383, 2000. 10, 21
- 562 [25] I. Pólik and T. Terlaky. New stopping criteria for detecting infeasibility in conic optimization.
563 *Optim. Lett.*, 3(2):187–198, 2009. 3

- 564 [26] I. Pólik and T. Terlaky. On the computational complexity of gap-free duals for semidefinite
565 programming. Technical report, Lehigh University, Bethlehem, PA, 2009. 6
- 566 [27] M.V. Ramana. *An Algorithmic Analysis of Multiquadratic and Semidefinite Programming*
567 *Problems*. PhD thesis, Johns Hopkins University, Baltimore, Md, 1993. 3, 6
- 568 [28] M.V. Ramana. An exact duality theory for semidefinite programming and its complexity
569 implications. *Math. Programming*, 77(2):129–162, 1997. 3, 5, 6
- 570 [29] M.V. Ramana, L. Tunçel, and H. Wolkowicz. Strong duality for semidefinite programming.
571 *SIAM J. Optim.*, 7(3):641–662, 1997. 3, 6, 8
- 572 [30] J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization*.
573 MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics, Philadel-
574 phia, PA, 2001. 6
- 575 [31] R.T. Rockafellar. *Convex analysis*. Princeton Landmarks in Mathematics. Princeton University
576 Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks. 6, 13
- 577 [32] A. Shapiro and A. Nemirovskii. Duality of linear conic problems. Technical report, School of
578 Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia, 2003.
579 16
- 580 [33] Alexander Shapiro. On duality theory of conic linear problems. In *Semi-infinite programming*
581 *(Alicante, 1999)*, volume 57 of *Nonconvex Optim. Appl.*, pages 135–165. Kluwer Acad. Publ.,
582 Dordrecht, 2001. 16
- 583 [34] G.W. Stewart. Rank degeneracy. *SIAM J. Sci. Statist. Comput.*, 5(2):403–413, 1984. 3
- 584 [35] G.W. Stewart. Determining rank in the presence of error. In *Linear algebra for large scale and*
585 *real-time applications (Leuven, 1992)*, volume 232 of *NATO Adv. Sci. Inst. Ser. E Appl. Sci.*,
586 pages 275–291. Kluwer Acad. Publ., Dordrecht, 1993. 3
- 587 [36] J.F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones.
588 *Optim. Methods Softw.*, 11/12(1-4):625–653, 1999. sedumi.ie.lehigh.edu. 3
- 589 [37] M. Todd. Semidefinite optimization. *Acta Numerica*, 10:515–560, 2001. 6
- 590 [38] M. J. Todd and Y. Ye. Approximate Farkas lemmas and stopping rules for iterative infeasible-
591 point algorithms for linear programming. *Math. Programming*, 81(1, Ser. A):1–21, 1998. 3
- 592 [39] Y.E. Nesterov M.J. TODD and Y. YE. Infeasible-start primal-dual methods and infeasibility
593 detectors for nonlinear programming problems. *Math. Program.*, 84(2, Ser. A):227–267, 1999.
594 3
- 595 [40] L. Tunçel and H. Wolkowicz. Strong duality and minimal representations for cone optimization.
596 Technical Report CORR 2008-07, University of Waterloo, Waterloo, Ontario, 2008. submitted.
597 3, 6, 9, 17, 29
- 598 [41] Levent Tunçel. *Polyhedral and semidefinite programming methods in combinatorial optimiza-*
599 *tion*, volume 27 of *Fields Institute Monographs*. American Mathematical Society, Providence,
600 RI, 2010. 28

- 601 [42] R.H. Tütüncü, K.C. Toh, and M.J. Todd. Solving semidefinite-quadratic-
602 linear programs using SDPT3. *Math. Program.*, 95(2, Ser. B):189–217, 2003.
603 www.math.nus.edu.sg/~matttohc/sdpt3.html. 3
- 604 [43] H. Wolkowicz. Some applications of optimization in matrix theory. *Linear Algebra Appl.*,
605 40:101–118, 1981. 3
- 606 [44] H. Wolkowicz. Solving semidefinite programs using preconditioned conjugate gradients. *Optim.*
607 *Methods Softw.*, 19(6):653–672, 2004. 4
- 608 [45] H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors. *Handbook of semidefinite programming*.
609 International Series in Operations Research & Management Science, 27. Kluwer Academic
610 Publishers, Boston, MA, 2000. Theory, algorithms, and applications. 5, 9
- 611 [46] H. Wolkowicz and Q. Zhao. Semidefinite programming relaxations for the graph partitioning
612 problem. *Discrete Appl. Math.*, 96/97:461–479, 1999. Selected for the special Editors’ Choice,
613 Edition 1999. 6
- 614 [47] M. Yamashita, K. Fujisawa, and M. Kojima. Implementation and evaluation of SDPA 6.0
615 (semidefinite programming algorithm 6.0). *Optim. Methods Softw.*, 18(4):491–505, 2003.
616 sdpa.indsys.chuo-u.ac.jp/sdpa/. 3
- 617 [48] M. Yamashita, K. Fujisawa, K. Nakata, M. Nakata, M. Fukuda, K. Kobayashi, and K. Goto. A
618 high-performance software package for semidefinite programs: Sdpa7. Technical report, Dept.
619 of Information Sciences, Tokyo Institute of Technology, Tokyo, Japan, 2010. 4
- 620 [49] Constantin Zălinescu. On zero duality gap and the Farkas lemma for conic programming.
621 *Math. Oper. Res.*, 33(4):991–1001, 2008. 16
- 622 [50] Constantin Zălinescu. On duality gap in linear conic problems. Technical report, University
623 of Al. I. Cusa, Iasi, Romania, 2010. 16
- 624 [51] Q. Zhao, S.E. Karisch, F. Rendl, and H. Wolkowicz. Semidefinite programming relaxations
625 for the quadratic assignment problem. *J. Comb. Optim.*, 2(1):71–109, 1998. Semidefinite
626 programming and interior-point approaches for combinatorial optimization problems (Fields
627 Institute, Toronto, ON, 1996). 6

Index

- 628 E_{ij} , unit matrices, 5
- 629 K^* , polar (dual) cone, 4
- 630 \mathcal{A}_c , homogenized constraint, 9
- 631 \mathcal{Q} , second order cone, 10
- 632 \mathcal{T}^\dagger , Moore-Penrose generalized inverse of \mathcal{T} , 7
- 633 $\text{cond}(\mathcal{A}) = \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)}$, condition number, 5
- 634 $\langle c, x \rangle := \sum c_{ij}x_{ij}$, trace inner product, 3
- 635 s2Mat, 5
- 636 $\sigma_i(A)$, singular values of A , 5
- 637 s2vec, isometry $\mathbb{S}^n \rightarrow \mathbb{R}^{t(n)}$, 5
- 638 $e = \frac{1}{\sqrt{n}}I$, 9
- 639 e_i , unit vector, 5
- 640 $t(n)$, triangular number, 5
- 641 v_D , dual optimal value, 3
- 642 v_P , (finite) primal optimal value, 3
- 643 (AP), auxiliary problem (2.4), 10
- 644 (DAP), dual of auxiliary problem (2.5), 10
- 645 (p-d i-p), primal-dual interior-point, 3
- 646 auxiliary problem, (AP) (2.4), 10
- 647 complementarity partition, 29
- 648 complementarity partition of K, K^* , 29
- 649 condition number, $\text{cond}(\mathcal{A}) = \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)}$, 5
- 650 cone partial order, 4
- 651 conjugate face, 4
- 652 constraint qualification, CQ, 3, 6
- 653 convex cone generated by S , cone(S), 4
- 654 convex cone, K , 4
- 655 CQ, constraint qualification, 3
- 656 dual cone, K^* , 4
- 657 dual of auxiliary problem, (DAP) (2.5), 10
- 658 dual SDP, 3
- 659 exposed face, 4
- 660 extremal ray, 4
- 661 face, $F \trianglelefteq K$, 4
- 662 facial reduction, 3
- 663 facially exposed cone, 4
- 664 feasible sets, $\mathcal{F}_P, \mathcal{F}_P^y, \mathcal{F}_P^z, \mathcal{F}_D$, 4
- 665 Löwner partial order, 3, 4
- 666 minimal face of (1.1), f_P , 4, 17
- 667 Moore-Penrose generalized inverse of \mathcal{T} , \mathcal{T}^\dagger , 7
- 668 optimal value shift, 5
- 669 partial order induced by K , 4
- 670 pointed cone, 4
- 671 polar cone, K^* , 4
- 672 primal SDP, 3
- 673 primal-dual interior-point, p-d i-p, 3
- 674 problem assumptions, 5
- 675 proper complementarity partition of K, K^* , 29
- 676 proper cone, 4
- 677 proper face, 4
- 678 proper subset, 4
- 679 rank-revealing, 3
- 680 rank-revealing rotation, 19
- 681 recession directions, \mathcal{R}_D , 8
- 682 second order cone, \mathcal{Q} , 10
- 683 singular values of A , $\sigma_i(A)$, 5
- 684 Slater's constraint qualification, 3, 6
- 685 strict complementarity partition of K, K^* , 29
- 686 strong duality, 5
- 687 strong infeasibility, 3
- 688 strongly dualized primal problem, 6
- 689 theorems of alternative, 7
- 690 triangular number, 5
- 691 triangular number, $t(n)$, 5
- 692 weak infeasibility, 3