Sensitivity analysis of semidefinite programs without strong duality

Yuen-Lam Cheung Henry Wolkowicz

June 30, 2014

Abstract

Suppose that we are given a feasible conic program with a finite optimal value and with strong duality failing. It is known that there are *small perturbations* of the problem data that lead to relatively *big changes* in the optimal value. We quantify the notion of big change in the case of a semidefinite program (SDP). We first show that for any SDP with a finite optimal value where strong duality fails, and where there is a nonzero duality gap, then for a sufficiently small step along any feasible perturbation direction, the optimal value changes by at least a fixed constant. And next, if there is a zero duality gap, with or without dual attainment, then any sufficiently small $\epsilon > 0$ feasible perturbation changes the optimal value by at most $O(\epsilon^\gamma)$ for some, to be specified, constant $\gamma \in (0,1)$. Our main tool involves the *facial reduction* of SDP.

Contents

List of Algorithms

1 Introduction

We investigate the sensitivity of semidefinite programs (SDP) for which strong duality fails, with respect to feasible right-hand side perturbations. It is well-known that if strong duality holds for the SDP, then any small right-hand side perturbation changes the optimal value in linear order. Here we provide results without assuming that strong duality holds. We show that if strong duality (zero duality gap and dual attainment) does not hold for an SDP with finite optimal value, then one of the following two cases holds.

- *Case 1: nonzero duality gap.* In this case, any sufficiently small feasible perturbation changes the optimal value of the SDP by no less than (half) the duality gap (which may be infinite).

- *Case 2: zero duality gap, but no dual attainment.* In this case, there exist a constant $\kappa > 0$ and a positive integer d such that any feasible perturbation of sufficiently small norm ϵ changes the optimal value by at most $\kappa\epsilon^{1/2^d}$. The positive integer d is the number of *facial reduction* iterations needed to find the *minimal face* containing the feasible region; and it is bounded above by $n-1$ where n is the size of the matrix variable.

There is a large volume of results on sensitivity analysis for nonlinear programs. To mention a few, Goldfarb and Scheinberg [12] performed sensitivity analysis on SDP assuming that the Slater condition holds for both the primal and the dual. In [18], Rockafellar used the optimal value function $\phi(u) := \inf_x F(x,u)$ and its directional derivative to study the sensitivity of the general convex program $\inf_x f(x) = F(x,0)$. Bonnans and Shapiro [4] (and also [19, 3]) studied the perturbation theory of nonlinear programs, and in [3, Section 7.3] they focused on the case where the dual is not solvable. Ben-Israel, Ben-Tal and Zlobec used a *feasible direction* approach to identify the regions of stability for convex optimization problems (i.e., sets of perturbations on which the changes in optimal solutions and values depend continuously) [2, Section 8]. Classical results for nonlinear programs appeared in [11].

The degree of singularity, which is essential in the second half of this article, is coined by Wang and Pang [22] for convex quadratic inequalities and by Sturm [20] for linear matrix inequalities.

1.1 Preliminaries

We consider the primal-dual pair of (linear) semidefinite programs, SDPs,

$$v_{\mathrm{P}} := \sup_y \left\{ b^T y : C - \mathcal{A}^* y \succeq 0 \right\}, \tag{P}$$

$$v_{\mathrm{D}} := \inf_X \left\{ \langle C, X \rangle : \mathcal{A}(X) = b,\ X \succeq 0 \right\}, \tag{D}$$

where $b \in \mathbb{R}^m$, $C \in \mathbb{S}^n$, $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ is a linear map, $\mathbb{S}^n$ is the Euclidean space (i.e., finite dimensional inner product space) of $n \times n$ real symmetric matrices equipped with the trace inner product $\langle X, Y \rangle := \operatorname{tr}(X^T Y)$ for all $X, Y \in \mathbb{S}^n$, and $X \succeq 0$ (resp. $X \succ 0$) denotes $X \in \mathbb{S}^n_+$, is positive semidefinite, (resp. $X \in \mathbb{S}^n_{++}$, is positive definite). We define the *spectral norm* $\|B\|_2 := \max_x \{\|Bx\|_2 : \|x\|_2 \leq 1\}$ for all $B \in \mathbb{R}^{m \times n}$, and the *Frobenius norm* $\|X\| := \sqrt{\langle X, X \rangle} = \sqrt{\operatorname{tr}(X^2)}$ for all $X \in \mathbb{S}^n$.

The primal SDP (P) is said to be *feasible* if there exists $\hat{y} \in \mathbb{R}^m$ such that $C - \mathcal{A}^* \hat{y} \succeq 0$. (Similarly, (D) is said to be feasible if there exists $\hat{X} \in \mathbb{S}^n$ such that $\mathcal{A}(\hat{X}) = b$ and $\hat{X} \succeq 0$.) If (P) is infeasible, then we take $v_{\mathrm{P}} = -\infty$; if (D) is infeasible, then we take $v_{\mathrm{D}} = +\infty$. *Strong duality* is said to hold for (P) if $v_{\mathrm{P}} = v_{\mathrm{D}}$ and v_{D} is attained. A sufficient condition for strong duality to hold for (P) is that v_{P} is finite and (P) satisfies the *Slater condition*,

$$C - \mathcal{A}^* \hat{y} \succ 0, \text{ for some } \hat{y} \in \mathbb{R}^m. \tag{Slater}$$

We concentrate on SDP of the form (P) that satisfy the following assumption.

Assumption 1.1 *The SDP (P) is feasible and has finite optimal value,* v_{P}*.*

We consider the (linear) perturbed problem

$$\mathrm{val}_{\mathrm{P}}(S) := \sup_{y} \left\{ b^T y : C - \mathcal{A}^* y \succeq S \right\}, \qquad (\mathrm{P}_{\mathrm{pert}}(S))$$

and study the relationship between the size of the perturbation S, and the change $\mathrm{val}_{\mathrm{P}}(S) - \mathrm{val}_{\mathrm{P}}(0)$ in the optimal value. We show that if (P) is feasible with finite optimal value, then whenever the perturbed problem ($\mathrm{P}_{\mathrm{pert}}(S)$) is feasible and $S \in \mathbb{S}^n$ is small, we have

$$\mathrm{val}_{\mathrm{P}}(S) - \mathrm{val}_{\mathrm{P}}(0) \begin{cases} = O(\|S\|) & \text{if strong duality holds for (P);} \\ \geq \frac{1}{2}(v_{\mathrm{D}} - v_{\mathrm{P}}) & \text{if } v_{\mathrm{P}} < v_{\mathrm{D}}\ ; \\ = O(\|S\|^{1/2^d}) \text{ for some integer } d > 0 & \text{if strong duality fails for (P)} \\ & \text{and } v_{\mathrm{P}} = v_{\mathrm{D}}. \end{cases} \tag{1.1}$$

(See Theorems 4.6 and 4.10.) Here the integer $d > 0$ is the *degree of singularity* of the linear matrix inequality, LMI, $C - \mathcal{A}^* y \succeq 0$, i.e., the number of steps needed to *facially reduce* the LMI.

Remark 1.2 *While our results are stated for an SDP of the form (P), they are easily applicable to any feasible SDP in standard equality form, (D), or more generally, subspace formulation [8]. In fact, let $\hat{X}$ satisfy the equation $\mathcal{A}(\hat{X}) = b$ and let $\mathcal{V} : \mathbb{S}^n \to \mathbb{R}^{\dim(\ker(\mathcal{A}))}$ be a linear map satisfying $\mathcal{R}(\mathcal{V}^*) = \ker(\mathcal{A})$. Then for any $X \in \mathbb{S}^n$,*

$$\mathcal{A}(X) = b,\ X \succeq 0 \quad \iff \quad X = \hat{X} + \mathcal{V}^* v \succeq 0 \text{ for some } v.$$

The corresponding perturbed problem for (D) would then be

$$\inf \left\{ \langle C, X \rangle : \mathcal{A}(X) = b - s,\ X \succeq 0 \right\},$$

where $s \in \mathcal{R}(\mathcal{A})$ is the right-hand side perturbation. Let $S \in \mathbb{S}^n$ satisfy $\mathcal{A}(S) = s$. Then

$$\mathcal{A}(X) = b - s,\ X \succeq 0 \quad \iff \quad X = \hat{X} + \mathcal{V}^* v \succeq S \text{ for some } v.$$

1.2 Main Contributions

While there is a vast amount of literature on the sensitivity analysis of linear and nonlinear optimization problems, most results rely on some regularity assumption. On the other hand, there are a few results (e.g. [20, 22]) on the Hölderian error bounds for convex quadratic inequalities and for linear matrix inequalities that depend on the degree of singularity. The results in this article only assume feasibility and a finite optimal value, and at the same time highlight the importance of the degree of singularity in estimating the change in the optimal value with respect to right-hand side perturbations. The bound on the changes in the optimal value provided in this article are also tight.

1.3 Outline

In Section 2, we recall some relevant well-known results, including: asymptotic properties of SDP (Section 2.1); the facial structure of the cone of positive semidefinite matrices (Section 2.2); and facial reduction for SDP (Section 2.3). Illustrative examples for the perturbation results are provided in Section 2.4. In Section 3, we take a closer look at the set of *feasible perturbations*. Then we formalize the main results presented in (1.1) and provide the proofs in Section 4. The last case in (1.1) requires some results concerning the degree of singularity; we provide those in the Appendix.

2 Asymptotics and Facial Reduction

2.1 Asymptotic feasibility and optimal value of SDP

A sequence $\{y^{(k)}\}_k$ is said to be *asymptotically feasible* for (P) if there exists a sequence $\{Z^{(k)}\}_k \subset \mathbb{S}^n_+$ such that $Z^{(k)} + \mathcal{A}^* y^{(k)} \to C$ as $k \to \infty$. We say that (P) is *weakly infeasible* if (P) is not feasible but possesses an asymptotically feasible sequence; and (P) is *strongly infeasible* if (P) does not have an asymptotically feasible sequence. Similarly, a sequence $\{X^{(k)}\}_k$ is said to be asymptotically feasible for (D) if $X^{(k)} \succeq 0$ for all k and $\lim_k \mathcal{A}(X^{(k)}) = b$. Strong infeasibility and weak infeasibility of (D) are defined similarly as for (P).

Define the *asymptotic optimal value* of (P) as

$$v_{\mathrm{P}}^a := \sup \left\{ \limsup_k b^T y^{(k)} : \left\{ y^{(k)} \right\}_k \text{ is asymptotically feasible for (P)} \right\}, \tag{2.1}$$

and the asymptotic optimal value of (D) as

$$v_{\mathrm{D}}^a := \inf \left\{ \liminf_k \langle C, X^{(k)} \rangle : \{X^{(k)}\}_k \text{ is asymptotically feasible for (D)} \right\}.$$

We take the convention that $v_{\mathrm{P}}^a = -\infty$ (resp., $v_{\mathrm{D}}^a = +\infty$) if (P) (resp., (D)) is strongly infeasible. Note that if (P) is feasible, then $v_{\mathrm{P}}^a \geq v_{\mathrm{P}}$. As we can see in Example 2.9 below, strict inequality may hold.

We say that $\hat{y} \in \mathbb{R}^m$ is an *improving direction* for (P) if $-\mathcal{A}^*\hat{y} \succeq 0$ and $b^T\hat{y} \geq 1$; and $\{y^{(k)}\}_k \subset \mathbb{R}^m$ is an *improving direction sequence* if there exists a sequence $\{Z^{(k)}\}_k \subset \mathbb{S}^n_+$ such that $Z^{(k)} + \mathcal{A}^*y^{(k)} \to 0$ and $b^Ty^{(k)} \geq 1$ for all k. Improving direction sequences and improving directions for (P), respectively, serve as certificates of infeasibility and strong infeasibility of the dual (D).

Lemma 2.1 ([14, Lemmas 5 and 6]) *The SDP (D) is infeasible if, and only if, (P) possesses an improving direction sequence. (D) is strongly infeasible if, and only if, (P) possesses an improving direction.*

The dual of an SDP satisfying Assumption 1.1 cannot be strongly infeasible.

Theorem 2.2 ([9]) *If (P) is feasible and $v_{\rm P} < +\infty$, then (D) is either feasible or weakly infeasible, and $v^a_{\rm D} = v_{\rm P}$.*

If both (P) and (D) are feasible, then *weak duality*, i.e., $v_{\rm P} \leq v_{\rm D}$, implies that both (P) and (D) have finite optimal values, and Theorem 2.2 implies that

$$v^a_{\rm P} = v_{\rm D} \geq v_{\rm P} = v^a_{\rm D}. \tag{2.2}$$

2.2 Facial properties of $\mathbb{S}^n_+$

We first introduce the notion of faces for general convex sets and convex cones. Let $(\mathbb{V}, \langle\cdot,\cdot\rangle_{\mathbb{V}})$ be a Euclidean space. A *face* of a nonempty convex set $\mathcal{S} \subseteq \mathbb{V}$, denoted by $\mathcal{F} \trianglelefteq \mathcal{S}$, is a nonempty convex set $\mathcal{F} \subseteq \mathcal{S}$ that satisfies the condition

$$\left\{ x, y \in \mathcal{S},\ \alpha \in (0,1) \text{ and } \alpha x + (1-\alpha)y \in \mathcal{F} \right\} \implies x, y \in \mathcal{F}.$$

A set $\mathcal{F}$ is called a *proper face* of $\mathcal{S}$, denoted by $\mathcal{F} \triangleleft \mathcal{S}$, if $\mathcal{F} \trianglelefteq \mathcal{S}$ and $\mathcal{F} \neq \mathcal{S}$. A face $\mathcal{F}$ of $\mathcal{S}$ is *exposed* if $\mathcal{F} = \{x \in \mathcal{S} : \langle s, x\rangle_{\mathbb{V}} = 0\}$, for some $s \in \mathbb{V}$ satisfying $\langle s, x\rangle_{\mathbb{V}} \geq 0$ for all $x \in \mathcal{S}$. Given a face $\mathcal{F}$ of a nonempty convex cone $\mathcal{K} \subseteq \mathbb{V}$, the *conjugate face* of $\mathcal{F}$ is defined as the set $\mathcal{F}^c := \mathcal{F}^\perp \cap \mathcal{K}^*$, where $\mathcal{K}^* := \{x \in \mathbb{V} : \langle x, y\rangle_{\mathbb{V}} \geq 0,\ \forall\, y \in \mathcal{K}\}$ is the *dual cone* of $\mathcal{K}$. It is immediate that

$$\{\, \mathcal{F}_1, \mathcal{F}_2 \trianglelefteq \mathcal{K} \ \text{ and } \ \mathcal{F}_1 \subseteq \mathcal{F}_2 \,\} \implies \mathcal{F}_2^c \subseteq \mathcal{F}_1^c,$$

and that $\mathcal{F}^c = \{x\}^\perp \cap \mathcal{K}^*$, where x is any element in the relative interior of $\mathcal{F}$, denoted by $\mathrm{ri}(\mathcal{F})$. Moreover, for any $\mathcal{F} \triangleleft \mathcal{K}$, $\mathcal{F}^{cc} := (\mathcal{F}^c)^c = \mathcal{F}$ if, and only if, $\mathcal{F}$ is an exposed face, e.g., [8, Proposition 3.1].

Given a nonempty convex set $\mathcal{S}$ and $\emptyset \neq \mathcal{T} \subseteq \mathcal{S}$, the *minimal face* of $\mathcal{S}$ containing $\mathcal{T}$ is defined as the intersection of all faces of $\mathcal{S}$ that contain $\mathcal{T}$:

$$\mathrm{face}(\mathcal{T}, \mathcal{S}) := \bigcap \{\mathcal{F} : \mathcal{F} \unlhd \mathcal{S},\ \mathcal{T} \subseteq \mathcal{F}\}.$$

Since a nonempty intersection of two faces is a face, we have that $\mathrm{face}(\mathcal{T}, \mathcal{S}) \unlhd \mathcal{S}$. By definition, $\mathrm{face}(\mathcal{T}, \mathcal{S})$ is inclusion-wise the smallest face of $\mathcal{S}$ containing the set $\mathcal{T}$.

The facial structure of $\mathbb{S}^n_+$ is well-known, see e.g., [23]:

Proposition 2.3 *The faces of $\mathbb{S}^n_+$ satisfy the following properties.*

1. *Any face of $\mathbb{S}^n_+$ is either $\{0\}$, $\mathbb{S}^n_+$ or $Q\, \mathbb{S}^{\bar{n}}_+ Q^T$, where $Q \in \mathbb{R}^{n\times\bar{n}}$ (with $1 \le \bar{n} < n$) is of full column rank.*
2. *$\mathbb{S}^n_+$ is facially exposed, i.e., all the faces of $\mathbb{S}^n_+$ are exposed faces.*
3. *Let $0 \neq X \in \mathbb{S}^n_+$. If $Q \in \mathbb{R}^{n\times\bar{n}}$ (with $1 \le \bar{n} < n$) is a full column rank such that $\mathcal{R}(Q) = \mathcal{R}(X)$, then $\mathrm{face}(\{X\}, \mathbb{S}^n_+) = Q\, \mathbb{S}^{\bar{n}}_+ Q^T$.*

2.3 Facial reduction and degree of singularity

In this section we review some concepts regarding the failure of the Slater condition for (P) and the *facial reduction algorithm*, a regularization technique for SDP instances that fail the Slater condition. We also recall the notion of the *degree of singularity* for LMI.

We first extend the notion of minimal faces to LMI and SDP. Given $\hat{C} \in \mathbb{S}^n$ and a linear map $\hat{\mathcal{A}} : \mathbb{S}^n \to \mathbb{R}^m$ such that $\hat{C} - \hat{\mathcal{A}}^*\hat{y} \succeq 0$ for some $\hat{y}$, the *minimal face* of the LMI $\hat{C} - \hat{\mathcal{A}}^* y \succeq 0$ is the minimal face of $\mathbb{S}^n_+$ containing the feasible slacks $\{Z \in \mathbb{S}^n : Z = \hat{C} - \hat{\mathcal{A}}^* y \succeq 0\}$. The *minimal face of a feasible SDP* of the form (P) is defined as the minimal face of the LMI $C - \mathcal{A}^* y \succeq 0$ that defines the feasible region of (P). It is immediate that the Slater condition holds for (P) if, and only if, the minimal face of (P) is $\mathbb{S}^n_+$. Similarly, since the constraints of (D) can be written as an LMI provided (D) is feasible, the minimal face of (D) is defined as the minimal face of $\mathbb{S}^n_+$ containing $\{X \in \mathbb{S}^n : \mathcal{A}(X) = b,\ X \succeq 0\}$. Under the assumption that (P) is feasible, it is immediate that

$$\mathrm{face}(\mathcal{F}^Z_{\mathrm{P}}, \mathbb{S}^n_+) = \mathrm{face}(\mathbb{S}^n_+ \cap \mathcal{L}, \mathbb{S}^n_+) \unlhd \mathbb{S}^n_+,$$

where

$$\mathcal{F}^Z_{\mathrm{P}} := \{Z \in \mathbb{S}^n : Z = C - \mathcal{A}^* y \succeq 0\}, \quad \text{and} \quad \mathcal{L} := \mathrm{span}\,\{\{C\} \cup \mathcal{R}(\mathcal{A}^*)\} \subset \mathbb{S}^n. \tag{2.3}$$

The notion of minimal face is important because we can *regularize* an SDP (or a conic program in general) by restricting it to the minimal face, resulting in an equivalent SDP for which strong duality holds.

Theorem 2.4 ([6]) *Suppose that (P) is feasible. Let f_{P} denote the minimal face of (P). Then*

$$\begin{aligned} v_{\mathrm{P}} &= \sup_y \left\{ b^T y : C - \mathcal{A}^* y \in f_{\mathrm{P}} \right\} && (2.4a) \\ &= \inf_X \left\{ \langle C, X \rangle : \mathcal{A}(X) = b,\ X \in (f_{\mathrm{P}})^* \right\}, && (2.4b) \end{aligned}$$

and (2.4b) is solvable.

One way to find the minimal face of an LMI is to use *facial reduction*. The facial reduction relies on a theorem of the alternative for the Slater condition. We recall a general version of the theorem of the alternative.

Proposition 2.5 ([14]) *Let $(\mathbb{V}, \langle \cdot, \cdot \rangle_{\mathbb{V}})$ be a Euclidean space, $\mathcal{K} \subseteq \mathbb{V}$ be a nonempty closed convex cone and $\mathcal{L} \subseteq \mathbb{V}$ be a linear subspace. Then*

$$\mathrm{ri}(\mathcal{K}) \cap \mathcal{L} \neq \emptyset \iff \mathcal{K}^* \cap \mathcal{L}^\perp \subseteq -\mathcal{K}^*. \tag{2.5}$$

In fact,

$$\mathcal{K} \cap \mathcal{L} \subseteq \mathrm{face}(\mathcal{K}^* \cap \mathcal{L}^\perp, \mathcal{K}^*)^c \trianglelefteq \mathcal{K}.$$

In particular, the Slater condition holds for (P) if, and only if, $\mathbb{S}^n_+ \cap \mathcal{L}^\perp = \{0\}$ by Proposition 2.5 (where $\mathcal{L}$ is defined in (2.3)). If the Slater condition fails for (P), then Proposition 2.5 provides a proper face $\mathrm{face}(\mathbb{S}^n_+ \cap \mathcal{L}^\perp, \mathbb{S}^n_+)^c$ that contains the minimal face of (P) (since $(\mathbb{S}^n_+)^* = \mathbb{S}^n_+$).

In Algorithm A.1, Page 26, we display one version of the facial reduction algorithm for finding the minimal face of the LMI $C - \mathcal{A}^* y \succeq 0$. (More details are available in [7, 8, 16, 20].) The facial reduction takes the linear subspace $\mathcal{L}$ and the cone $\mathcal{K} := \mathbb{S}^n_+$ as input. In each iteration of the facial reduction algorithm, we find a matrix $D \in \mathcal{K}^* \cap \mathcal{L}^\perp \backslash (-\mathcal{K}^*)$, which certifies that the Slater condition fails, i.e., $\mathrm{ri}(\mathcal{K}) \cap \mathcal{L} = \emptyset$; then we replace $\mathcal{K}$ by $\mathcal{K} \cap \{D\}^\perp$.

The number of iterations of the facial reduction algorithm to find the minimal face of an LMI is called the *degree of singularity* of the LMI. The importance of the degree of singularity is highlighted in the error bound result from [20] on the distance to the set of feasible slacks of an LMI; see Theorem 4.7. We will show in Theorem 4.10 that if $v_{\mathrm{P}} = v_{\mathrm{D}}$ but v_{D} is unattained, then for any small feasible perturbation S an upper bound of the difference $\mathrm{val}_{\mathrm{P}}(S) - \mathrm{val}_{\mathrm{P}}(0)$ can be expressed in terms of the degree of singularity.

Given the importance of the degree of singularity in this article, we now formalize its definition. We first define the set of all *chains of certificates for the facial reduction.* (A similar notion was introduced by Pataki in [16].)

Definition 2.6 *For any linear subspace $\mathcal{L} \subseteq \mathbb{S}^n$, $\mathcal{C}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+)$ is defined as the set consisting of all finite sequences of matrices $(D^{(1)}, \dots, D^{(k)})$ that satisfy the following conditions:*

- $D^{(j)} \in \mathrm{ri}(\mathcal{L}^\perp \cap (\mathcal{K}_{j-1})^*) \backslash (-\mathcal{K}_{j-1})^*$, *where* $\mathcal{K}_0 := \mathbb{S}^n_+$ *and* $\mathcal{K}_j = \mathcal{K}_{j-1} \cap \{D^{(j)}\}^\perp$, $\forall\, j = 1, \dots, k$,
- $\mathcal{L} \cap \mathrm{ri}(\mathcal{K}_k) \neq \emptyset$.

In particular, if $\mathcal{L} \cap \mathbb{S}^n_{++} \neq \emptyset$ or $\mathcal{L} \cap \mathbb{S}^n_+ = \{0\}$, then $\mathcal{C}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+) = \emptyset$. That all sequences satisfying the conditions described in Definition 2.6 are finite in length follows from the fact that the facial reduction algorithm has finitely many iterations. In fact, we will show in Proposition B.1, Page 28, that any two finite sequences in $\mathcal{C}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+)$ must be of the same length. Now we can define the degree of singularity.

Definition 2.7 *Let $\mathcal{L} \subseteq \mathbb{S}^n$ be a linear subspace. The degree of singularity of the set $\mathcal{L} \cap \mathbb{S}^n_+$ is the length of any finite sequence of matrices in $\mathcal{C}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+)$, and is denoted by $\mathrm{d}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+)$. The degree of singularity of a feasible LMI $\hat{C} - \hat{\mathcal{A}}^* y \succeq 0$ (where $\hat{C} \in \mathbb{S}^n$ and $\hat{\mathcal{A}} : \mathbb{S}^n \to \mathbb{R}^m$) is defined as $\mathrm{d}(\hat{\mathcal{A}}, \hat{C}) := \mathrm{d}(\mathrm{span}(\{\hat{C}\} \cup \mathcal{R}(\hat{\mathcal{A}})) \cap \mathbb{S}^n_+, \mathbb{S}^n_+)$. The degree of singularity of an SDP of the form (P) is defined to be the degree of singularity of the LMI defining the feasible region of (P).*

2.4 Examples

In this section we consider some examples of SDP instances from [15, 17, 21]. In each of the examples, strong duality fails for the SDP instance and we consider one single feasible perturbation direction and the resultant change in optimal value along that direction. By abuse of notation, in this section we restrict the function $\mathrm{val}_{\mathrm{P}}(\cdot)$ defined on ($\mathrm{P}_{\mathrm{pert}}(S)$) to a fixed direction $\epsilon \mapsto \epsilon \hat{S}$ for some particular $\hat{S}$; so $\mathrm{val}_{\mathrm{P}}(\cdot)$ is a function on $\mathbb{R}_+$.

Example 2.8 ((D) is infeasible.) *For $\epsilon \geq 0$, consider*

$$\mathrm{val}_{\mathrm{P}}(\epsilon) := \sup_y \left\{ y_2 : \begin{bmatrix} y_1 & y_2 & y_3 \\ y_2 & y_3 & 0 \\ y_3 & 0 & 0 \end{bmatrix} \preceq \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \epsilon \end{bmatrix},\ y \in \mathbb{R}^3 \right\}, \tag{2.6}$$

$$\mathrm{val}_{\mathrm{D}}(\epsilon) := \inf_X \left\{ \epsilon X_{33} : X_{11} = 0,\ X_{12} = 1,\ 2X_{13} + X_{22} = 0,\ X \succeq 0 \right\}. \tag{2.7}$$

The dual (2.7) is infeasible. When $\epsilon = 0$, the degree of singularity of the primal (2.6) is $\mathrm{d}(\mathcal{A}, C) = 2$; (2.6) has optimal value $\mathrm{val}_{\mathrm{P}}(0) = 0$ while the asymptotic optimal value is $+\infty$. Indeed, consider

$$Z^{(k)} = \begin{bmatrix} k^2 & -k & 1 \\ -k & 1 & 0 \\ 1 & 0 & \frac{1}{k} \end{bmatrix}, \ y^{(k)} = \begin{bmatrix} -k^2 \\ k \\ 1 \end{bmatrix}, \quad \forall\, k \in \mathbb{N}.$$

Then

$$Z^{(k)} + \mathcal{A}^* y^{(k)} = \begin{bmatrix} k^2 & -k & 1 \\ -k & 1 & 0 \\ 1 & 0 & \frac{1}{k} \end{bmatrix} + \begin{bmatrix} -k^2 & k & -1 \\ k & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{k} \end{bmatrix} \to 0 \quad as \quad k \to \infty$$

(meaning that $\{y^{(k)}\}_k$ is asymptotically feasible) and $b^T y^{(k)} = k$ for all k. Hence $v_{\mathrm{P}}^a = +\infty$.

Now fix any $\epsilon > 0$. For all sufficiently large k, $\begin{bmatrix} k^2 & -k & 1 \\ -k & 1 & 0 \\ 1 & 0 & \epsilon \end{bmatrix} \succeq 0$, so $\begin{pmatrix} -k^2 \\ k \\ -1 \end{pmatrix}$ is feasible for the perturbed problem (2.6). Hence $\mathrm{val}_{\mathrm{P}}(\epsilon) = +\infty$. This confirms the middle case in (1.1), i.e., the perturbation is bounded below by (half) the duality gap, ∞ here.

Example 2.9 (nonzero finite duality gap.) *Fix any $\alpha > 0$. For $\epsilon \geq 0$, consider*

$$\mathrm{val}_{\mathrm{P}}(\epsilon) := \sup_y \left\{ y_2 : \begin{bmatrix} y_2 & 0 & 0 \\ 0 & y_1 & y_2 \\ 0 & y_2 & 0 \end{bmatrix} \preceq \begin{bmatrix} \alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \epsilon \end{bmatrix}, \ y \in \mathbb{R}^2 \right\}, \tag{2.8}$$

$$\mathrm{val}_{\mathrm{D}}(\epsilon) := \inf_X \left\{ \alpha X_{11} + \epsilon X_{33} : X_{22} = 0, \ X_{11} + X_{23} = 1, \ X \succeq 0 \right\}. \tag{2.9}$$

Observe that the primal requires only one iteration of facial reduction to identify the minimal face, i.e., $\mathrm{d}(\mathcal{A}, C) = 1$. But the dual of the reduced primal still fails the Slater condition.

Let $\mathcal{F}_{\mathrm{P}}^y(\epsilon)$ be the set of feasible solutions y of (2.8), $\mathcal{F}_{\mathrm{P}}^Z(\epsilon)$ be the set of feasible slacks of (2.8) and $\mathcal{F}_{\mathrm{D}}^X(\epsilon)$ be the set of feasible solutions X for (2.9). For $\epsilon = 0$, we get

$$\mathcal{F}_{\mathrm{P}}^y(0) = \mathbb{R}_- \times 0, \quad \mathcal{F}_{\mathrm{P}}^Z(0) = \left\{ \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \gamma & 0 \\ 0 & 0 & 0 \end{bmatrix} : \gamma \geq 0 \right\}, \quad \mathcal{F}_{\mathrm{D}}(0) = \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \beta \end{bmatrix} : \beta \geq 0 \right\}.$$

So $\mathrm{val}_{\mathrm{P}}(0) = 0 < \alpha = \mathrm{val}_{\mathrm{D}}(0) = v_{\mathrm{P}}^{a}(0)$.[1] *Now consider* $\epsilon > 0$*. Then*

$$\begin{bmatrix} \alpha - y_2 & 0 & 0 \\ 0 & -y_1 & -y_2 \\ 0 & -y_2 & \epsilon \end{bmatrix} \succeq 0 \iff y_2 \le \alpha \quad \text{and} \quad y_1 \le -y_2^2/\epsilon.$$

Hence $\mathrm{val}_{\mathrm{P}}(\epsilon) = \alpha = v_{\mathrm{P}}^{a}(0)$*. On the other hand, the constraints of the dual* (2.9) *are unchanged. Hence* $\mathrm{val}_{\mathrm{D}}(\epsilon) = \alpha = \mathrm{val}_{\mathrm{D}}(\epsilon)$*, and* $\mathrm{val}_{\mathrm{P}}(\epsilon) - \mathrm{val}_{\mathrm{P}}(0) = \alpha = \mathrm{val}_{\mathrm{D}}(0) - \mathrm{val}_{\mathrm{P}}(0)$*, i.e., this again confirms the middle case in* (1.1).

Example 2.10 (zero duality gap but v_{D} is unattained.) *For* $\epsilon \ge 0$*, consider*

$$\mathrm{val}_{\mathrm{P}}(\epsilon) := \sup_{y} \left\{ 2y : \begin{bmatrix} 0 & y \\ y & 0 \end{bmatrix} \preceq \begin{bmatrix} 1 & 0 \\ 0 & \epsilon \end{bmatrix},\ y \in \mathbb{R} \right\}, \tag{2.10}$$

$$\mathrm{val}_{\mathrm{D}}(\epsilon) := \inf_{X} \left\{ X_{11} + \epsilon X_{22} : \begin{bmatrix} X_{11} & 1 \\ 1 & X_{22} \end{bmatrix} \succeq 0 \right\}. \tag{2.11}$$

Note that y is feasible for (2.10) *if, and only if,* $|y| \le \sqrt{\epsilon}$*. The primal* (2.10) *can be seen as a positive semidefinite completion problem:* (2.10) *aims at finding the range of y so that* $\begin{bmatrix} 1 & -y \\ -y & \epsilon \end{bmatrix} \succeq 0$*, where the fixed data has a parameter* ϵ*.*

When $\epsilon = 0$*,* $y = 0$ *is the only feasible solution for* (2.10) *so* $\mathrm{val}_{\mathrm{P}}(0) = 0$*. On the other hand, the dual optimal value* $\mathrm{val}_{\mathrm{D}}(0) = \mathrm{val}_{\mathrm{P}}(0) = 0$ *but is not attained. Hence strong duality fails for* (2.10) *when* $\epsilon = 0$*. In fact, the degree of singularity of* (2.10) *is 1 when* $\epsilon = 0$*.*

When $\epsilon > 0$*,* (2.10) *satisfies the Slater condition, so* $\mathrm{val}_{\mathrm{P}}(\epsilon) = \mathrm{val}_{\mathrm{D}}(\epsilon) = 2\sqrt{\epsilon}$ *(and* $\mathrm{val}_{\mathrm{D}}(\epsilon)$ *is attained). Yet* $\mathrm{val}_{\mathrm{P}}(\epsilon) - \mathrm{val}_{\mathrm{P}}(0) = 2\sqrt{\epsilon}$*, which is not of linear order. This illustrate the last case in* (1.1).

Example 2.11 (Zero duality gap but v_{D} is unattained.) *This example generalizes Example 2.10 and illustrates the last case in* (1.1)*. We consider an SDP on* $\mathbb{S}^n_+$ *that requires* $n-1$ *iterations of facial reduction to identify the minimal face of* $\mathbb{S}^n_+$ *containing its feasible region, i.e.* $\mathrm{d}(\mathcal{A}, C) = n-1$*. We show that there exists a feasible perturbation S such that* $\mathrm{val}_{\mathrm{P}}(S) - \mathrm{val}_{\mathrm{P}}(0) = 2\|S\|^{1/2^{n-1}}$*.*

Let $n \ge 3$ *and* $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^{n-1}$ *be a linear map defined by the matrices*

$$A_1 = e_1 e_2^T + e_2 e_1^T, \quad \text{and} \quad A_k = e_k e_k^T + e_1 e_{k+1}^T + e_{k+1} e_1^T,\ \forall\, k \in 2 : n-1.$$

---

[1] To see that $v_{\mathrm{P}}^{a} = \alpha$, consider $y^{(k)} = \begin{pmatrix} -\alpha k^2 \\ \alpha \end{pmatrix} \in \mathbb{R}^2$ and $Z^{(k)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \alpha^2 k & -\alpha \\ 0 & -\alpha & \frac{1}{k} \end{bmatrix} \in \mathbb{S}^3_+$. Then $Z^{(k)} + \mathcal{A}^* y^{(k)} = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{k} \end{bmatrix}$ and $b^T y^{(k)} = \alpha$ for all k.

For $\epsilon \geq 0$ consider

$$\mathrm{val}_{\mathrm{P}}(\epsilon) := \sup_y \left\{2y_1 : \mathcal{A}^* y \preceq e_1 e_1^T + \epsilon e_n e_n^T\right\}, \tag{2.12}$$

$$\mathrm{val}_{\mathrm{D}}(\epsilon) := \inf_X \left\{X_{11} + \epsilon X_{nn} : \mathcal{A}(X) = b,\ X \succeq 0\right\}. \tag{2.13}$$

Again, the primal (2.12) can be considered as a positive semidefinite completion problem, where we look at the growth in the free $(1,2)$ element of the feasible slack $Z = e_1 e_1^T + \epsilon e_n e_n^T - \mathcal{A}^* y \succeq 0$ with respect to perturbation in the given (n,n) data element.

Suppose that $\epsilon = 0$. For any $y \in \mathbb{R}^{n-1}$, the matrix $Z = e_1 e_1^T - \mathcal{A}^* y$ always satisfies $Z_{nn} = 0$. In particular, $Z \succeq 0$ if, and only if, $Z = e_1 e_1^T$, i.e., $y = 0$. Hence $\mathrm{val}_{\mathrm{P}}(0) = 0$. On the other hand, (2.13) has an optimal value $\mathrm{val}_{\mathrm{D}}(0) = 0$ that is not attained. Hence strong duality does not hold for (2.12) when $\epsilon = 0$. In fact, as noted above, $\mathrm{d}(\mathcal{A}, e_1 e_1^T) = n - 1$.

Now suppose that $\epsilon > 0$. It is not difficult to see that (2.12) satisfies the Slater condition, i.e., suppose $D \succeq 0$ satisfies $\langle C, D\rangle = 0$ and $\mathcal{A}(D) = 0$. It suffices to show that $D = 0$ (see Proposition 2.5). Indeed, $D \succeq 0$ and $\langle C, D\rangle = 0$ imply $D_{11} = D_{nn} = 0$. But this in turn implies that D has zero diagonal. Hence $D = 0$.

Now note that y is feasible for (2.12) if, and only if,

$$0 \geq y_{n-1} \geq -\epsilon^{1/2}, \quad 0 \geq y_{n-2} \geq -\epsilon^{1/4}, \quad \ldots, \quad 0 \geq y_2 \geq -\epsilon^{1/2^{n-2}}, \quad |y_1| \leq \epsilon^{1/2^{n-1}}.$$

Hence $\mathrm{val}_{\mathrm{P}}(\epsilon) = \mathrm{val}_{\mathrm{D}}(\epsilon) = 2\epsilon^{1/2^{n-1}}$, and $\mathrm{val}_{\mathrm{P}}(\epsilon) - \mathrm{val}_{\mathrm{P}}(0) = 2\epsilon^{1/2^{n-1}}$.

Example 2.12 (Zero duality gap but v_{D} is unattained.) *Let*

$$A_1 = -E_{11}, \quad A_2 = -E_{22}, \quad A_3 = e_3 e_4^T + e_4 e_3^T, \quad A_4 = e_1 e_3^T + e_3 e_1^T, \quad C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

and $b = \begin{bmatrix} 0 & -1 & 2 & 0 \end{bmatrix}^T$. Then (P) reads

$$\sup_y \left\{ -y_2 + 2y_3 : \begin{bmatrix} y_1 & 1 & -y_4 & 0 \\ 1 & y_2 & 0 & 0 \\ -y_4 & 0 & 0 & -y_3 \\ 0 & 0 & -y_3 & 1 \end{bmatrix} \succeq 0,\ y \in \mathbb{R}^4 \right\} = 0, \tag{2.14}$$

which is unattained. (Note that y feasible must satisfy $y_3 = y_4 = 0$.) Meanwhile, (D) *yields*

$$\inf_X \{2X_{12} + X_{44} : X_{11} = 0,\ X_{22} = 1,\ X_{34} = 1,\ X_{13} = 0,\ X \succeq 0\}$$
$$= \inf_X \left\{ X_{44} : X = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & * & * \\ 0 & * & * & 1 \\ 0 & * & 1 & * \end{bmatrix} \succeq 0 \right\} = 0,$$

which is unattained too. Hence strong duality does not hold for (2.14)*. In fact, the degree of singularity of* (2.14) *is 1. Now for any* $\epsilon > 0$*,*

$$\mathrm{val}_{\mathrm{P}}(\epsilon) := \sup_y \left\{ -y_2 + 2y_3 : \begin{bmatrix} y_1 & 1 & -y_4 & 0 \\ 1 & y_2 & 0 & 0 \\ -y_4 & 0 & \epsilon & -y_3 \\ 0 & 0 & -y_3 & 1 \end{bmatrix} \succeq 0 \right\} = 2\sqrt{\epsilon}.$$

(Note that $\mathrm{val}_{\mathrm{P}}(\epsilon)$ *is unattained.) In other words,* $\mathrm{val}_{\mathrm{P}}(\epsilon) - \mathrm{val}_{\mathrm{P}}(0) = 2\sqrt{\epsilon}$*.*

In summary, we considered five SDP instances where strong duality does not hold and a feasible perturbation that leads to big change in optimal value exists. Together with the results in Section 4, we see that the bounds in the second and third cases (1.1) are indeed tight.

3 Feasible perturbations

Before we present our main results (Theorems 4.6 and 4.10), we review in this section the classical notion of *feasible perturbations* and *feasible perturbation directions*, which will be used in the following section. Specifically, we derive a characterization of feasible perturbation directions for (P); see Theorem 3.3.

Define the set of *feasible perturbations* for (P):

$$\mathcal{P} := \mathcal{P}(\mathcal{A}, C) := \{S \in \mathbb{S}^n : C - \mathcal{A}^* y \succeq S, \text{ for some } y \in \mathbb{R}^m\}. \tag{3.1}$$

If the LMI $C - \mathcal{A}^* y \succeq 0$ is feasible, then $0 \in \mathcal{P}$. Moreover, $\mathcal{P}$ is a convex set,[2] and $\mathcal{P}$ has nonempty interior if, and only if, $C - \mathcal{A}^* y \succeq 0$ satisfies the Slater condition. However, $\mathcal{P}$ is not necessarily closed.[3] Related to the set $\mathcal{P}$ is the set of *feasible perturbation directions*

[2] Let $S, T \in \mathcal{P}$ and $y_S, y_T \in \mathbb{R}^m$ satisfy $C - \mathcal{A}^* y_S \succeq S$ and $C - \mathcal{A}^* y_T \succeq T$ respectively. Then for any $\alpha \in [0,1]$, we have $C - \mathcal{A}^*(\alpha y_S + (1-\alpha) y_T) \succeq \alpha S + (1-\alpha) T$. Hence $\alpha S + (1-\alpha) T \in \mathcal{P}$.

[3] Consider $C = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ and $A = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$. Then $C - yA = \begin{bmatrix} 1 & 1 \\ 1 & y \end{bmatrix} \succeq 0$ if, and only if, $y \geq 1$. Note

$$\begin{aligned}\hat{\mathcal{P}} := \hat{\mathcal{P}}(\mathcal{A}, C) \quad &:= \quad \{S \in \mathbb{S}^n : \; C - \mathcal{A}^* y \succeq \epsilon S, \quad \text{for some } y \in \mathbb{R}^m \text{ and } \epsilon > 0\} \qquad (3.2)\\ &= \quad \{S \in \mathbb{S}^n : \; \epsilon S \in \mathcal{P}, \quad \text{for some } \epsilon > 0\},\end{aligned}$$

which is a convex cone. As we see in Example 3.2 below, there exist instances of (P) such that one can find $S \in \hat{\mathcal{P}}$ of arbitrarily small norm such that $\epsilon S \in \mathcal{P}$ for $\epsilon > 0$ only if ϵ is "small", in the sense that $\epsilon \leq O(\|S\|^2)$. And, we show in Section 4.1 that,

$$S \in \hat{\mathcal{P}} \implies \left\{ \text{face}\left(\mathcal{F}_{\mathrm{P}}^{Z}, \mathbb{S}_+^n\right) \subseteq \text{face}\left(\mathcal{F}_{\mathrm{P}}^{Z}(\epsilon S), \mathbb{S}_+^n\right), \; \forall \epsilon > 0, \text{ suff. small, with } \epsilon S \in \mathcal{P} \right\}.$$

where

$$\mathcal{F}_{\mathrm{P}}^{Z}(S) \quad := \quad \{Z \in \mathbb{S}^n : Z = C - \mathcal{A}^* y - S \succeq 0, \text{ for some } y \in \mathbb{R}^m\}.$$

It is possible to characterize the set $\hat{\mathcal{P}}$ (and the proof will be given in Section 3.2, below):

let $\text{face}\left(\mathcal{F}_{\mathrm{P}}^{Z}, \mathbb{S}_+^n\right) = Q \, \mathbb{S}_+^{\bar{n}} Q^T$ with $Q \in \mathbb{R}^{n \times \bar{n}}$ full column rank, and let $U = \begin{bmatrix} P & Q \end{bmatrix} \in \mathbb{R}^{n \times n}$ be nonsingular; then we have

$$\hat{\mathcal{P}} = \left\{ USU^T \in \mathbb{S}^n : S = \begin{array}{c c} & \begin{array}{cc} n-\bar{n} & \bar{n} \end{array} \\ \begin{array}{c} n-\bar{n} \\ \bar{n} \end{array} & \begin{bmatrix} S_1 & S_2 \\ S_2^T & S_3 \end{bmatrix} \end{array}, \; S_1 \preceq 0, \; \mathcal{R}(S_2) \subseteq \mathcal{R}(S_1) \right\}.$$

Throughout the remainder of this section, by using an orthogonal rotation if needed, we assume without loss of generality that the minimal face of (P) is of the form $\begin{bmatrix} 0 & 0 \\ 0 & \mathbb{S}_+^{\bar{n}} \end{bmatrix}$.

3.1 A technical lemma

We first prove a technical result, that given any $Z \in \mathbb{S}_+^n$ and any feasible perturbation direction $S \in \mathbb{S}^n$, the range of Z does not *shrink* when taking sufficiently small step along the direction S.

---

that $\begin{bmatrix} 1 & 1 \\ 1 & y \end{bmatrix} \succeq \begin{bmatrix} s & 0 \\ 0 & 0 \end{bmatrix}$ is feasible for all $s \in (0, 1)$ but infeasible when $s = 1$. Hence the set of feasible perturbations is not closed for this instance.

Lemma 3.1 *Let* $Z = \begin{bmatrix} 0 & 0 \\ 0 & \hat{Z} \end{bmatrix} \in \mathbb{S}^n$ *and* $\hat{Z} \succ 0$*. Then for any* $S = \begin{bmatrix} S_1 & S_2 \\ S_2^T & S_3 \end{bmatrix} \in \mathbb{S}^n$ *with*

$$S_1 \succeq 0 \quad \textit{and} \quad \mathcal{R}(S_2) \subseteq \mathcal{R}(S_1), \tag{3.3}$$

we have $Z + \epsilon S \succeq 0$ *and* $\mathcal{R}(Z) \subseteq \mathcal{R}(Z + \epsilon S)$ *for any* $\epsilon \in (0, \delta)$*, where*

$$\delta := \delta(\hat{Z}, S) := \frac{\lambda_{\min}(\hat{Z})}{2(\|S_3\| + \|S_2\|_2^2\, \lambda_{\max}(S_1^\dagger) + 1)}.[4] \tag{3.4}$$

Proof. For any $\epsilon \in (0, 2\delta)$, we have $\lambda_{\min}(\hat{Z}) > \epsilon\|S_3\|$ which yields

$$0 \prec (\lambda_{\min}(\hat{Z}) - \epsilon\|S_3\|)I \preceq \lambda_{\min}(\hat{Z})I + \epsilon S_3 \preceq \hat{Z} + \epsilon S_3. \tag{3.5}$$

Let B be a full column rank matrix such that $S_1 = BB^T$. Then $\mathcal{R}(S_2) \subseteq \mathcal{R}(S_1)$ implies that $S_2 = BR$ for some matrix R. It is immediate that

$$R = (B^TB)^{-1}B^TBR = B^\dagger S_2,$$

so $\|R\|_2 \le \|B^\dagger\|_2\|S_2\|_2$, i.e.,

$$\lambda_{\max}(RR^T) = \|R\|_2^2 \le \frac{\|S_2\|_2^2}{(\sigma_{\min}(B))^2} = \|S_2\|_2^2\lambda_{\max}(S_1^\dagger). \tag{3.6}$$

Hence $\epsilon < \frac{\lambda_{\min}(\hat{Z})}{\|S_1\| + \lambda_{\max}(RR^T) + 1}$ holds for any $\epsilon \in (0, 2\delta)$ by (3.6). Then by (3.5),

$$\epsilon R(\hat{Z} + \epsilon S_3)^{-1}R^T \preceq \epsilon(\lambda_{\min}(\hat{Z}) - \epsilon\|S_3\|)^{-1}RR^T \preceq \frac{\epsilon\lambda_{\max}(RR^T)}{\lambda_{\min}(\hat{Z}) - \epsilon\|S_3\|}I \preceq I.$$

Using the conjugation $B \cdot B^T$, we get

$$BB^T \succeq \epsilon BR(\hat{Z} + \epsilon S_3)^{-1}R^TB^T \implies \epsilon S_1 - \epsilon^2 S_2(\hat{Z} + \epsilon S_3)^{-1}S_2^T \succeq 0.$$

Therefore

$$Z + \epsilon S = \begin{bmatrix} \epsilon S_1 & \epsilon S_2 \\ \epsilon S_2^T & \hat{Z} + \epsilon S_3 \end{bmatrix} \succeq 0.$$

Next we show that $\mathcal{R}(Z) \subseteq \mathcal{R}(Z + \epsilon S)$ for all $\epsilon \in (0, \delta)$. Suppose that $x \in \ker(Z + \epsilon S)$. Then $0 = \frac{1}{2}x^TZx + \frac{1}{2}x^T(Z + 2\epsilon S)x$; $Z \succeq 0$ and $Z + 2\epsilon S \succeq 0$ imply that $x^TZx = 0$, i.e., $x \in \ker(Z)$. Therefore $\ker(Z + \epsilon S) \subseteq \ker(Z)$, i.e., $\mathcal{R}(Z) \subseteq \mathcal{R}(Z + \epsilon S)$. $\square$

[4] Note that the quantity δ in (3.4) is small when: $\|S_3\|$ is large; when $\lambda_{\min}(\hat{Z})$ is small; or when S is "ill-conditioned" in the sense that the quantity $\|S_2\|_2^2\, \lambda_{\max}(S_1^\dagger) = \frac{\|S_2\|_2^2}{\text{smallest positive eigenvalue of } S_1}$, is large.

It is immediate that the bound (3.6) used in the proof is tight, as we can see in Example 3.2, below. In particular, we can show that there exists a feasible instance of (P) such that for any arbitrarily small $\gamma > 0$, there exists $S \in \hat{\mathcal{P}}$ such that $\|S\| \le \gamma$ and $\epsilon S \in \mathcal{P}$ only if $\epsilon = O(\gamma^2)$.

Example 3.2 *Consider*

$$C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Then $\mathcal{F}_{\mathrm{P}}^{Z} = \{C\}$. *Define the matrix*

$$S = \begin{bmatrix} \delta^2 & 0 & \sqrt{\delta} \\ 0 & \delta & 0 \\ \sqrt{\delta} & 0 & 0 \end{bmatrix}, \quad i.e., \quad S_1 = \begin{bmatrix} \delta^2 & 0 \\ 0 & \delta \end{bmatrix}, \quad S_2 = \begin{bmatrix} \sqrt{\delta} \\ 0 \end{bmatrix}, \quad S_3 = 0,$$

where $\delta \in (0,1)$.[5] *Then* $\|S\| = \sqrt{2\delta + \delta^2 + \delta^4} \le 2\sqrt{\delta}$. *Note that for any* $\epsilon \ge 0$, $C + \epsilon S - \mathcal{A}^* y \succeq 0$ *is feasible if, and only if, there exists* $y \in \mathbb{R}^2$ *such that*

$$\begin{bmatrix} \epsilon\delta^2 & y_1 & \epsilon\sqrt{\delta} \\ y_1 & \epsilon\delta & y_2 \\ \epsilon\sqrt{\delta} & y_2 & 1 \end{bmatrix} \succeq 0,$$

which holds only if $\epsilon \le \delta$ *(by considering the (1,3) principal minor). Note also that for any* $\epsilon \in [0,\delta]$, $C + \epsilon S \succeq 0$. *Hence* $S \in \hat{\mathcal{P}}$ *and* $\epsilon S \in \hat{\mathcal{P}}$ *only if* $\epsilon \in [0,\delta]$.

This example shows that there exists a feasible instance of (P) *such that for any arbitrarily small* $\gamma > 0$, *there exists* $S \in \hat{\mathcal{P}}$ *such that* $\|S\| \le \gamma$ *and* $\epsilon S \in \mathcal{P}$ *only if* $\epsilon \le O(\gamma^2)$.

3.2 Characterizing the set of feasible perturbation directions

We now characterize the set $\hat{\mathcal{P}}$.

Theorem 3.3 *Suppose that* $\mathrm{face}(\mathcal{F}_{\mathrm{P}}^{Z}, \mathbb{S}^n_+) = \begin{bmatrix} 0 & 0 \\ 0 & \mathbb{S}^{\bar{n}}_+ \end{bmatrix}$, *where* $0 < \bar{n} < n$. *Then*

$$\hat{\mathcal{P}} = \left\{ S = \begin{bmatrix} S_1 & S_2 \\ S_2^T & S_3 \end{bmatrix} : \; S_1 \preceq 0, \; \mathcal{R}(S_2) \subseteq \mathcal{R}(S_1) \right\}. \tag{3.7}$$

[5] Relative to the proof of Lemma 3.1, We then have

$$S_1 = B^2, \quad \text{where } B = \begin{bmatrix} \delta & 0 \\ 0 & \sqrt{\delta} \end{bmatrix}, \quad \text{and} \quad S_2^T = BR, \quad \text{where } R = \begin{bmatrix} \frac{1}{\sqrt{\delta}} \\ 0 \end{bmatrix}.$$

Proof. Let $\hat{y} \in \mathbb{R}^m$ satisfy

$$C - \mathcal{A}^* \hat{y} = \begin{array}{c c} & \begin{array}{cc} n-\bar{n} & \bar{n} \end{array} \\ \begin{array}{c} n-\bar{n} \\ \bar{n} \end{array} & \left[\begin{array}{cc} 0 & 0 \\ 0 & \hat{Z} \end{array} \right] \end{array} \succeq 0, \ \hat{Z} \succ 0.$$

Suppose that $S \in \hat{\mathcal{P}}$, i.e., $C - \mathcal{A}^* \tilde{y} \succeq \epsilon S$ for some $\tilde{y} \in \mathbb{R}^m$ and $\epsilon > 0$. Then

$$\begin{bmatrix} -\alpha\epsilon S_1 & -\alpha\epsilon S_2 \\ -\alpha\epsilon S_2^T & (1-\alpha)\hat{Z} - \alpha\epsilon S_3 \end{bmatrix} = C - \mathcal{A}^* \left((1-\alpha)\hat{y} + \alpha\tilde{y} \right) - \alpha\epsilon S \succeq 0,$$

for all $\alpha \in [0,1]$; in particular, $S_1 \preceq 0$. Pick a sufficiently small $\alpha \in (0,1)$ such that $(1-\alpha)\hat{Z} - \alpha\epsilon S_1 \succ 0$. Then using the Schur complement, we have

$$-\alpha\epsilon S_1 - \alpha^2\epsilon^2 S_2 \big((1-\alpha)\hat{Z} - \alpha\epsilon S_3 \big)^{-1} S_2^T \succeq 0,$$

or equivalently,

$$S_1 + \alpha\epsilon S_2 \big((1-\alpha)\hat{Z} - \alpha\epsilon S_3 \big)^{-1} S_2^T \preceq 0.$$

Let $x \in \ker(S_1)$. Then $x^T \big(S_2 ((1-\alpha)\hat{Z} - \alpha\epsilon S_3)^{-1} S_2^T \big) x \le 0$. But $((1-\alpha)\hat{Z} - \alpha\epsilon S_1)^{-1}$ is positive definite, so we have $S_2^T x = 0$. This implies that $\mathcal{R}(S_2) \subseteq \mathcal{R}(S_1)$.

Conversely, suppose that $S = \begin{bmatrix} S_1 & S_2 \\ S_2^T & S_3 \end{bmatrix}$ satisfies $S_1 \preceq 0$ and $\mathcal{R}(S_2) \subseteq \mathcal{R}(S_1)$. By Lemma 3.1, $Z - \epsilon S \succeq 0$ for sufficiently small $\epsilon > 0$. Therefore $S \in \hat{\mathcal{P}}$. $\square$

4 Main results

We first recall that if strong duality holds for (P), then a small feasible perturbation leads to little change in the optimal value. Here f^* denotes the *convex conjugation* of an extended function $f : \mathbb{V} \to \mathbb{R} \cup \{+\infty\}$, where $(\mathbb{V}, \langle \cdot, \cdot \rangle_{\mathbb{V}})$ is a Euclidean space, i.e., $f^*(\phi) := \sup_x \{ \langle \phi, x \rangle_{\mathbb{V}} - f(x) : x \in \mathbb{V} \}$.

Theorem 4.1 ([5, 18]) *The value function* $\mathrm{val}_{\mathrm{P}} : \mathbb{S}^n \to [-\infty, \infty]$ *is concave. Moreover,* $v_{\mathrm{D}} = -(-\mathrm{val}_{\mathrm{P}})^{**}(0)$*, and the equality* $v_{\mathrm{P}} = v_{\mathrm{D}}$ *holds if, and only if,* $\mathrm{val}_{\mathrm{P}}$ *is upper semicontinuous at* 0*. In this case,* X^* *is an optimal solution of* (D) *if, and only if,* $\mathrm{val}_{\mathrm{P}}(S) - \mathrm{val}_{\mathrm{P}}(0) \le \langle X^*, S \rangle$ *for all* $S \in \mathbb{S}^n$*.*

In particular, if strong duality holds for (P), then for any $S \in \mathbb{S}^n$, either ($\mathrm{P}_{\mathrm{pert}}(S)$) is infeasible or ($\mathrm{P}_{\mathrm{pert}}(S)$) has a finite optimal value. As a corollary of Theorem 4.1 we get the following linear order type result.

Corollary 4.2 *Suppose that strong duality holds for* (P). *Then there exists a constant* $\kappa > 0$ *such that for any* $S \in \mathbb{S}^n$ *with* ($\mathrm{P}_{\text{pert}(S)}$) *feasible,* $\mathrm{val}_{\mathrm{P}}(S) - \mathrm{val}_{\mathrm{P}}(0) \leq \kappa\|S\|$.

What if strong duality does not hold? We consider the two different cases: (1) (P)-(D) has a nonzero duality gap, and (2) the duality gap is zero, i.e., $v_{\mathrm{P}} = v_{\mathrm{D}}$, but v_{D} is unattained.

4.1 Case 1: nonzero duality gap

In this case, we show in Theorem 4.6 below that:

> for *any* $S \in \hat{\mathcal{P}}$ and sufficiently small $\epsilon > 0$, we have $\epsilon S \in \mathcal{P}$ and $\mathrm{val}_{\mathrm{P}}(\epsilon S) - \mathrm{val}_{\mathrm{P}}(0) \geq \frac{1}{2}(v_{\mathrm{D}} - v_{\mathrm{P}}) \in (0, +\infty]$.

First we show in Lemma 4.3, as a simple extension of Theorem 2.2, that when Assumption (1.1) holds, we get $v_{\mathrm{P}}^a = v_{\mathrm{D}}$ even when (D) is infeasible. (We already saw from Theorem 2.2 that if both (P) and (D) are feasible, then $v_{\mathrm{P}}^a = v_{\mathrm{D}}$.) In particular, under Assumption 1.1, (2.2) holds and the duality gap between (P) and (D) is given by $v_{\mathrm{D}} - v_{\mathrm{P}} = v_{\mathrm{P}}^a - v_{\mathrm{P}} \in [0, +\infty]$.

Lemma 4.3 *Suppose that* (P) *is feasible but its dual* (D) *is infeasible. Then* $v_{\mathrm{P}}^a = +\infty$.

Proof. If (P) does not have a finite optimal value, then $v_{\mathrm{P}}^a \geq v_{\mathrm{P}} = +\infty$ implies $v_{\mathrm{P}}^a = +\infty$.

Suppose that (P) has a finite optimal value, i.e., (P) satisfies Assumption 1.1. Since (D) is infeasible, by Lemma 2.1 there exists a sequence $\left\{(y^{(k)}, Z^{(k)})\right\}_k$ satisfying

$$b^T y^{(k)} \geq 1 \text{ and } Z^{(k)} \succeq 0 \text{ for all } k, \quad \text{and} \quad \lim_k \left(Z^{(k)} + \mathcal{A}^* y^{(k)}\right) = 0.$$

By Theorem 2.2, (D) cannot be strongly infeasible. Thus $Z^{(k)} + \mathcal{A}^* y^{(k)} \neq 0$ for all k. (Otherwise $y^{(k)}$ would be an improving direction for (P), implying that (D) is strongly infeasible from Lemma 2.1.)

For each k, define

$$\hat{y}^{(k)} := \frac{1}{\|Z^{(k)} + \mathcal{A}^* y^{(k)}\|^{1/2}} y^{(k)}, \quad \hat{Z}^{(k)} := \frac{1}{\|Z^{(k)} + \mathcal{A}^* y^{(k)}\|^{1/2}} Z^{(k)}.$$

Then $\hat{Z}^{(k)} \succeq 0$ for all k,

$$\|\hat{Z}^{(k)} + \mathcal{A}^* \hat{y}^{(k)}\| = \|Z^{(k)} + \mathcal{A}^* y^{(k)}\|^{1/2} \implies \lim_k \left(\hat{Z}^{(k)} + \mathcal{A}^* \hat{y}^{(k)}\right) = 0$$

and

$$b^T\hat{y}^{(k)} \geq \frac{1}{\|Z^{(k)} + \mathcal{A}^* y^{(k)}\|^{1/2}} \to +\infty \implies \lim_k b^T\hat{y}^{(k)} = +\infty.$$

On the other hand, since (P) is feasible, let $\hat{y} \in \mathbb{R}^m$ satisfy $C - \mathcal{A}^*\hat{y} \succeq 0$. Then $\left\{\hat{y} + \hat{y}^{(k)}\right\}_k$ is asymptotically feasible for (P), and $\lim_k b^T(\hat{y} + \hat{y}^{(k)}) = +\infty$. Hence $v_{\mathrm{P}}^a = +\infty$. $\square$

Next we show that under Assumption 1.1, any perturbation $S \in \mathbb{S}^n$ that gives a strictly feasible ($\mathrm{P}_{\mathrm{pert}}(S)$) leads to a *big jump* in the optimal value.

Proposition 4.4 *Suppose that* (P) *satisfies Assumption 1.1 and* $v_{\mathrm{P}} < v_{\mathrm{D}} \in \mathbb{R} \cup \{+\infty\}$. *Suppose that* $S \in \mathbb{S}^n$ *is given, such that the perturbed SDP* ($\mathrm{P}_{\mathrm{pert}}(S)$) *satisfies the Slater condition. Then for all sufficiently small* $\epsilon \in (0, 1)$,

$$\mathrm{val}_{\mathrm{P}}(\epsilon S) - \mathrm{val}_{\mathrm{P}}(0) \geq \frac{1}{2}(v_{\mathrm{D}} - v_{\mathrm{P}}). \tag{4.1}$$

Proof. We consider two different cases: (D) is feasible and (D) is infeasible.

Case 1: (D) *is feasible.* Since ($\mathrm{P}_{\mathrm{pert}}(S)$) satisfies the Slater condition, we have that $S \in \mathrm{int}(\mathrm{dom}(-\mathrm{val}_{\mathrm{P}}))$. Note that $-\mathrm{val}_{\mathrm{P}}(0) \in \mathbb{R}$, so $-\mathrm{val}_{\mathrm{P}}$ is a proper convex function and we have [13, Prop. 1.2.5]

$$v_{\mathrm{D}} = -(-\mathrm{val}_{\mathrm{P}}(0))^{**} = \lim_{\epsilon \searrow 0} \mathrm{val}_{\mathrm{P}}(\epsilon S).$$

In particular, for all sufficiently small $\epsilon > 0$, we have $\mathrm{val}_{\mathrm{P}}(\epsilon S) \geq v_{\mathrm{D}} - \frac{1}{2}(v_{\mathrm{D}} - v_{\mathrm{P}})$. This proves (4.1).

Case 2: (D) *is infeasible.* In this case, ($\mathrm{P}_{\mathrm{pert}}(S)$) is feasible and the dual of ($\mathrm{P}_{\mathrm{pert}}(S)$) is infeasible. Hence Lemma 4.3 applies: there exists a sequence $\left\{(y^{(k)}, Z^{(k)})\right\}_k$ satisfying

$$Z^{(k)} \succeq 0 \;\forall\, k, \quad \lim_k \left(Z^{(k)} + \mathcal{A}^* y^{(k)}\right) = C - S \quad \text{and} \quad \lim_k b^T y^{(k)} = +\infty.$$

Let $\hat{y} \in \mathbb{R}^m$ satisfy $\hat{Z} := C - \mathcal{A}^*\hat{y} - S \succ 0$. Fix any $\epsilon \in (0, \lambda_{\min}(\hat{Z}))$; then there exists K such that for all $k \geq K$, $\|Z^{(k)} + \mathcal{A}^* y^{(k)} - (C - S)\| \leq \epsilon$, implying that

$$C - S - \mathcal{A}^* y^{(k)} - Z^{(k)} \succeq -\epsilon I \succeq -\hat{Z} \implies C - \frac{1}{2}\mathcal{A}^*(y^{(k)} + \hat{y}) \succeq S + \frac{1}{2}Z^{(k)} \succeq S.$$

Hence for all $k \geq K$, $\frac{1}{2}(y^{(k)} + \hat{y})$ is feasible for ($\mathrm{P}_{\mathrm{pert}}(S)$). The fact that $\lim_k b^T y^{(k)} = +\infty$ implies that $\mathrm{val}_{\mathrm{P}}(S) = +\infty = v_{\mathrm{D}}$, so (4.1) holds. $\square$

Remark 4.5 *Since the Slater condition is generic for feasible SDP instances of fixed size, e.g., [1, 10], the set of feasible perturbations* S *such that the perturbed SDP* (P) *fails the*

Slater condition is also of measure zero. In other words, Proposition 4.4 indicates that the set of feasible perturbations of (P) that does not lead to a "big" jump in optimal value is indeed of measure zero.

Theorem 4.6 *Suppose that (P) satisfies Assumption 1.1 and $v_{\mathrm{P}} < v_{\mathrm{D}} \in \mathbb{R} \cup \{+\infty\}$. Suppose that* $\mathrm{face}(\mathcal{F}_{\mathrm{P}}^Z, \mathbb{S}_+^n) = \begin{bmatrix} 0 & 0 \\ 0 & \mathbb{S}_+^{\bar{n}} \end{bmatrix}$ *and* $Z = \begin{bmatrix} 0 & 0 \\ 0 & \hat{Z} \end{bmatrix} \in \mathcal{F}_{\mathrm{P}}^Z$ *with* $\hat{Z} \succ 0$. *Let* $S \in \hat{\mathcal{P}}$ *and* $\epsilon \in (0, \delta)$, *where* δ *is defined in (3.4) as:*

$$\delta := \frac{\lambda_{\min}(\hat{Z})}{2\left(\|S_3\| + \|S_2\|_2^2 \, \lambda_{\max}(S_1^\dagger) + 1\right)}.$$

Then we have $\epsilon S \in \mathcal{P}$ *and* $\mathrm{face}(\mathcal{F}_{\mathrm{P}}^Z, \mathbb{S}_+^n) \subseteq \mathrm{face}\left(\mathcal{F}_{\mathrm{P}}^Z(\epsilon S), \mathbb{S}_+^n\right)$. *In particular, for all sufficiently small* $\epsilon > 0$,

$$\mathrm{val}_{\mathrm{P}}(\epsilon S) - \mathrm{val}_{\mathrm{P}}(0) \geq \frac{1}{2}(v_{\mathrm{D}} - v_{\mathrm{P}}). \tag{4.2}$$

Proof. From Lemma 3.1 we have $Z - \epsilon S \succeq 0$, for all $\epsilon \in (0, \delta)$; therefore $\epsilon S \in \mathcal{P}$ and $Z - \epsilon S \in \mathcal{F}_{\mathrm{P}}^Z(S)$. Also from Lemma 3.1, we have $\mathcal{R}(Z) \subseteq \mathcal{R}(Z - \epsilon S)$, so

$$\mathrm{face}(\mathcal{F}_{\mathrm{P}}^Z, \mathbb{S}_+^n) = \mathrm{face}\left(\{Z\}, \mathbb{S}_+^n\right) \subseteq \mathrm{face}(\{Z - \epsilon S\}, \ \mathbb{S}_+^n) \subseteq \mathrm{face}\left(\mathcal{F}_{\mathrm{P}}^Z(\epsilon S), \mathbb{S}_+^n\right).$$

It remains to prove (4.2). Let $Q \in \mathbb{R}^{n \times \bar{n}}$ be such that Q has orthonormal columns and $\mathrm{face}(\mathcal{F}_{\mathrm{P}}^Z(\epsilon S), \mathbb{S}_+^n) = Q\,\mathbb{S}_+^{\bar{n}} Q^T$. Then

$$\mathrm{val}_{\mathrm{P}}(\epsilon S) = \sup_y \left\{ b^T y : Q^T C Q - Q^T(\mathcal{A}^* y) Q \succeq \epsilon Q^T S Q \right\}, \tag{4.3}$$

which satisfies the Slater condition. On the other hand, consider the primal-dual pair:

$$\widehat{v_{\mathrm{P}}} = \sup_y \left\{ b^T y : Q^T C Q - Q^T(\mathcal{A}^* y) Q \succeq 0 \right\}, \tag{4.4}$$

$$\widehat{v_{\mathrm{D}}} = \inf_{\bar{X}} \left\{ \langle C, Q\bar{X}Q^T \rangle : \mathcal{A}(Q\bar{X}Q^T) = b, \ \bar{X} \succeq 0 \right\}. \tag{4.5}$$

Since $\mathrm{face}(\mathcal{F}_{\mathrm{P}}^Z, \mathbb{S}_+^n) \subseteq Q\,\mathbb{S}_+^{\bar{n}} Q^T$ and $\left\{ Q\bar{X}Q^T : \mathcal{A}(Q\bar{X}Q^T) = b, \ \bar{X} \succeq 0 \right\} \subseteq \{X : \mathcal{A}(X) = b, X \succeq 0\}$, we have

$$\widehat{v_{\mathrm{P}}} = v_{\mathrm{P}} < v_{\mathrm{D}} \leq \widehat{v_{\mathrm{D}}}, \tag{4.6}$$

i.e., the primal-dual pair (4.4) and (4.5) has a nonzero duality gap. Since (4.3) is a perturbation of (4.4), and satisfies the Slater condition by Proposition 4.4, we conclude that (4.6) implies

$$\mathrm{val_P}(\epsilon S) - \mathrm{val_P}(0) \;=\; \mathrm{val_P}(\epsilon S) - \widehat{v_P} \;\geq\; \frac{1}{2}(\widehat{v_D} - \widehat{v_P}) \;\geq\; \frac{1}{2}(v_D - v_P) \in (0, +\infty].$$

This proves (4.2). □

4.2 Case 2: strong duality fails but duality gap is zero

If the duality gap between (P) and (D) is zero and yet strong duality fails, then by Theorem 4.1, the function $\mathrm{val_P}(\cdot)$ is upper semicontinuous at 0 but the subdifferential $\partial(-\mathrm{val_P}(\cdot))(0)$ is empty. Despite the non-existence of subgradients, we can show that, for some fixed positive integer d, $\mathrm{val_P}(S) - \mathrm{val_P}(0) = O(\|S\|^{1/2^d})$ if $S \in \mathcal{P}$ is sufficiently small in norm. This result relies on the following error bound result for LMI.

Theorem 4.7 ([20], Theorem 3.3) *Suppose that the set $\mathcal{F}_\mathrm{P}^Z$ of feasible slacks of (P) is nonempty. Then there exist constants $\kappa > 0$ and $\bar{\epsilon} \in (0, 1)$ such that for any $\epsilon \in (0, \bar{\epsilon})$ and any $\tilde{Z} \in \mathbb{S}^n$ satisfying*

$$\mathrm{dist}(\tilde{Z}, C + \mathcal{R}(\mathcal{A}^*)) \leq \epsilon, \;\; \lambda_{\min}(\tilde{Z}) \geq -\epsilon,$$

we have

$$\mathrm{dist}(\tilde{Z}, \mathcal{F}_\mathrm{P}^Z) \leq \kappa(1 + \|\tilde{Z}\|)\epsilon^{1/2^{\mathrm{d}(\mathcal{A},C)}},$$

where $\mathrm{d}(\mathcal{A}, C)$ is the degree of singularity of the linear subspace $\mathrm{span}(\{C\} \cup \mathcal{R}(\mathcal{A}^*))$, defined in Definition 2.7.

We first deal with the case where (D) satisfies the Slater condition, in Section 4.3; then we use this to prove the general result in Section 4.4.

4.3 Case 2(a): (D) satisfies the Slater condition

We first prove a weaker result assuming that the dual (D) satisfies the Slater condition (which together with the feasibility of (P) implies that $v_P = v_D$).

Proposition 4.8 *Suppose that (P) satisfies Assumption 1.1 and that (D) satisfies the Slater condition. Then there exist constants $\kappa > 0$ and $\bar{\epsilon} \in (0, 1)$ such that for any $S \in \mathbb{S}^n$ with $0 < \|S\| \leq \bar{\epsilon}$ and ($\mathrm{P_{pert}}(S)$) feasible,*

$$\mathrm{val_P}(S) - \mathrm{val_P}(0) \leq \kappa\|S\|^{1/2^{\mathrm{d}(\mathcal{A},C)}}. \tag{4.7}$$

Proof. Let $\kappa_0 > 0$ and $\bar{\epsilon} \in (0,1)$ be such that for any $Y \in \mathbb{S}^n$ and $\epsilon \in (0, \bar{\epsilon})$ with

$$\operatorname{dist}(Y, C + \mathcal{R}(\mathcal{A}^*)) \le \epsilon, \;\; \lambda_{\min}(Y) \ge -\epsilon,$$

the following error bound holds:

$$\operatorname{dist}(Y, \mathcal{F}_{\mathrm{P}}^{Z}) \le \kappa_0 (1 + \|Y\|) \epsilon^{1/2^{\mathrm{d}(\mathcal{A},C)}}. \tag{4.8}$$

Let $\tilde{X}$ be a strictly feasible solution of (D). Fix any $S \in \mathbb{S}^n$ with ($\mathrm{P}_{\mathrm{pert}}(S)$) feasible and $\|S\| \le \bar{\epsilon}$. Since the dual of ($\mathrm{P}_{\mathrm{pert}}(S)$) has a (strictly) feasible solution $\tilde{X}$, we get that $\mathrm{val}_{\mathrm{P}}(S) < +\infty$ and that $\mathrm{val}_{\mathrm{P}}(S)$ is attained, i.e., there exist $\tilde{y}, \tilde{Z}$ satisfying

$$b^T \tilde{y} = \mathrm{val}_{\mathrm{P}}(S), \quad \tilde{Z} = C - S - \mathcal{A}^* \tilde{y} \succeq 0.$$

For any $y \in \mathbb{R}^m$ satisfying $Z := C - \mathcal{A}^* y \succeq 0$,

$$\begin{aligned} \mathrm{val}_{\mathrm{P}}(S) - \mathrm{val}_{\mathrm{P}}(0) &\le b^T \tilde{y} - b^T y \\ &= \langle \tilde{X}, C - S - \tilde{Z} \rangle - \langle \tilde{X}, C - Z \rangle \\ &\le \|\tilde{X}\| \|Z - (\tilde{Z} + S)\|. \end{aligned}$$

Minimizing over all $Z \in \mathcal{F}_{\mathrm{P}}^{Z}$,

$$\mathrm{val}_{\mathrm{P}}(S) - \mathrm{val}_{\mathrm{P}}(0) \le \|\tilde{X}\| \operatorname{dist}(\tilde{Z} + S, \, \mathcal{F}_{\mathrm{P}}^{Z}). \tag{4.9}$$

But $\tilde{Z} + S \in C + \mathcal{R}(\mathcal{A}^*)$ and $\lambda_{\min}(\tilde{Z} + S) \ge -\|S\|$ (because $\tilde{Z} + S \succeq S \succeq -\|S\| I$). Hence by (4.8), $\operatorname{dist}(\tilde{Z} + S, \, \mathcal{F}_{\mathrm{P}}^{Z}) \le \kappa_0 (1 + \|\tilde{Z} + S\|) \|S\|^{1/2^{\mathrm{d}(\mathcal{A},C)}}$. Hence by (4.9),

$$\mathrm{val}_{\mathrm{P}}(S) - \mathrm{val}_{\mathrm{P}}(0) \le \kappa_0 \|\tilde{X}\| (1 + \|\tilde{Z} + S\|) \|S\|^{1/2^{\mathrm{d}(\mathcal{A},C)}}. \tag{4.10}$$

Now observe that

$$\begin{aligned} & \lambda_{\min}(\tilde{X}) \|\tilde{Z}\| \;\le\; \langle C - S - \mathcal{A}^* \tilde{y}, \tilde{X} \rangle \;\le\; \langle C - S, \tilde{X} \rangle - \mathrm{val}_{\mathrm{P}}(S) \\ \Longrightarrow \quad & \|\tilde{Z} + S\| \;\le\; \|S\| + \|\tilde{Z}\| \;\le\; \|S\| + \frac{1}{\lambda_{\min}(\tilde{X})} \Big(\langle C - S, \tilde{X} \rangle - \mathrm{val}_{\mathrm{P}}(S) \Big). \end{aligned} \tag{4.11}$$

Using (4.11) and the assumption that $\|S\| < 1$, the right-hand side of (4.10) no greater than the expression

$$\begin{aligned}&\kappa_0\|\tilde X\|\left(1+\|S\|+\frac{1}{\lambda_{\min}(\tilde X)}\left(\langle C-S,\tilde X\rangle-\mathrm{val}_{\mathrm{P}}(S)\right)\right)\|S\|^{1/2^{\mathrm{d}(\mathcal{A},C)}}\\ \le\ &\kappa_0\frac{\|\tilde X\|}{\lambda_{\min}(\tilde X)}\left(2\lambda_{\min}(\tilde X)+\langle C-S,\tilde X\rangle-\mathrm{val}_{\mathrm{P}}(S)\right)\|S\|^{1/2^{\mathrm{d}(\mathcal{A},C)}}.\end{aligned}$$

Putting back into (4.10), we get

$$\begin{aligned}&\left(1+\kappa_0\frac{\|\tilde X\|}{\lambda_{\min}(\|\tilde X\|)}\|S\|^{1/2^{d(\mathcal{A},C)}}\right)(\mathrm{val}_{\mathrm{P}}(S)-\mathrm{val}_{\mathrm{P}}(0))\\ \le\quad &\kappa_0\frac{\|\tilde X\|}{\lambda_{\min}(\tilde X)}\left(2\lambda_{\min}(\tilde X)+\langle C-S,\tilde X\rangle-\mathrm{val}_{\mathrm{P}}(0)\right)\|S\|^{1/2^{d(\mathcal{A},C)}}.\end{aligned}\tag{4.12}$$

Note that $2\lambda_{\min}(\tilde X)+\langle C-S,\tilde X\rangle-\mathrm{val}_{\mathrm{P}}(0)\le 2\lambda_{\min}(\tilde X)+\|\tilde X\|+\langle C,\tilde X\rangle-\mathrm{val}_{\mathrm{P}}(0)$, which is positive by weak duality. Defining

$$\kappa=\kappa_0\frac{\|\tilde X\|}{\lambda_{\min}(\tilde X)}\left(2\lambda_{\min}(\tilde X)+\|\tilde X\|+\langle C,\tilde X\rangle-\mathrm{val}_{\mathrm{P}}(0)\right),$$

we get from (4.12) that $\mathrm{val}_{\mathrm{P}}(S)-\mathrm{val}_{\mathrm{P}}(0)\le\kappa\|S\|^{1/2^{\mathrm{d}(\mathcal{A},C)}}$. $\square$

4.4 Case 2(b): (D) does not satisfy the Slater condition

Now we consider the case where $v_{\mathrm{P}}=v_{\mathrm{D}}\in\mathbb{R}$ but v_{D} is unattained and (D) fails the Slater condition. Such a scenario can occur, as we can see in Example 2.12. We show that a bound of the form $\mathrm{val}_{\mathrm{P}}(S)-\mathrm{val}_{\mathrm{P}}(0)\le\kappa\|S\|^{1/2^{\mathrm{d}(\mathcal{A},C)}}$ holds even in this case. The proof idea is to restrict (D) on its minimal face, and then to use the fact that such a restriction does not change the degree of singularity of (P):

Lemma 4.9 *Suppose that (P) and (D) are feasible, and the minimal face of (D) is given by $\tilde P\,\mathbb{S}^r_+\tilde P^T$ for some full column rank matrix $\tilde P\in\mathbb{R}^{n\times r}$ (with $r>0$). Then*

$$\sup_y\left\{b^Ty:\tilde P^T(C-\mathcal{A}^*y)\tilde P\succeq 0\right\}\tag{4.13}$$

is also feasible, and $\mathrm{d}(\mathcal{A}(\tilde P\cdot\tilde P^T),\tilde P^TC\tilde P)\le\mathrm{d}(\mathcal{A},C)$.

The proof of Lemma 4.9 is given on Page 33 in Appendix C. Now we prove the main results of this section.

Theorem 4.10 *Assume that (P) satisfies Assumption 1.1, and that $v_{\mathrm{P}}=v_{\mathrm{D}}\in\mathbb{R}$ but v_{D} is unattained. Then there exist $\bar\epsilon\in(0,1)$ and $\kappa>0$ such that for any $S\in\mathbb{S}^n$ with*

(P$_{\text{pert}}(S)$) *feasible and* $\|S\| \le \bar{\epsilon}$,

$$\mathrm{val}_{\mathrm{P}}(S) - \mathrm{val}_{\mathrm{P}}(0) \le \kappa \|S\|^{1/2^{\mathrm{d}(\mathcal{A},C)}}.$$

Proof. If (D) satisfies the Slater condition, then the statement in the theorem holds by Proposition 4.8. In the remainder of the proof we assume that (D) fails the Slater condition.

Since v_{D} is assumed to be unattained, the minimal face of (D) does not equal $\{0\}$.[6] Let $\tilde{P}\,\mathbb{S}^r_+\tilde{P}^T$ be the minimal face of $\mathbb{S}^n_+$ containing the feasible region of (D) with $\tilde{P}^T\tilde{P} = I$. Therefore we have $v_{\mathrm{D}} = \bar{v}_{\mathrm{D}}$, where

$$\bar{v}_{\mathrm{D}} := \inf_W \left\{ \langle C, \tilde{P}W\tilde{P}^T \rangle : \mathcal{A}(\tilde{P}W\tilde{P}^T) = b,\ W \succeq 0 \right\}. \tag{4.14}$$

By definition of minimal face, (4.14) satisfies the Slater condition. Note that since (D) has no optimal solution, (4.14) has no optimal solution either. The dual of (4.14) is given by

$$\bar{v}_{\mathrm{P}} := \sup_y \left\{ b^T y : \tilde{P}^T(C - \mathcal{A}^* y)\tilde{P} \succeq 0 \right\}, \tag{4.15}$$

Any y feasible for (P) is also feasible for (4.15). Hence we have

$$v_{\mathrm{P}} = v_{\mathrm{D}} = \bar{v}_{\mathrm{D}} \ge \bar{v}_{\mathrm{P}} \ge v_{\mathrm{P}}, \quad \text{i.e., } v_{\mathrm{P}} = \bar{v}_{\mathrm{P}}. \tag{4.16}$$

Moreover, the primal-dual pair (4.15)-(4.14) satisfies the assumptions in Proposition 4.8, which together with Lemma 4.9 implies that there exist constants $\kappa > 0$ and $\bar{\epsilon} \in (0,1)$ such that for any $\bar{S} \in \mathbb{S}^r$ with $0 < \|\bar{S}\| \le \bar{\epsilon}$,

$$\tilde{P}^T(C - \mathcal{A}^* y)\tilde{P} \succeq \bar{S} \text{ feasible} \implies \sup_y \left\{ b^T y : \tilde{P}^T(C - \mathcal{A}^* y)\tilde{P} \succeq \bar{S} \right\} - \bar{v}_{\mathrm{P}} \le \kappa \|\bar{S}\|^{1/2^{\mathrm{d}(\mathcal{A},C)}}. \tag{4.17}$$

Fix any $S \in \mathbb{S}^n$ with (P$_{\text{pert}}(S)$) feasible and $\|S\| \le \bar{\epsilon}$. Then by weak duality and the fact that the feasible region of (D) is contained in $\tilde{P}\,\mathbb{S}^r_+\tilde{P}^T$,

$$\begin{aligned} \mathrm{val}_{\mathrm{P}}(S) \le\ & \inf_X \left\{ \langle C - S, X \rangle : \mathcal{A}(X) = b,\ X \succeq 0 \right\} \\ =\ & \inf_W \left\{ \langle C - S, \tilde{P}W\tilde{P}^T \rangle : \mathcal{A}(\tilde{P}W\tilde{P}^T) = b,\ W \succeq 0 \right\}, \end{aligned} \tag{4.18}$$

which satisfies the Slater condition. Since $\tilde{P}^T(C - \mathcal{A}^* y)\tilde{P} \succeq \tilde{P}^T S \tilde{P}$ is feasible, strong

[6] If the minimal face of (D) is $\{0\}$, then $X = 0$ is the only feasible solution. This implies that $b = 0$ and $v_{\mathrm{D}} = 0$, so $v_{\mathrm{P}} = 0 = v_{\mathrm{D}}$ and any primal/dual feasible solution is optimal.

duality holds and

$$\inf_W \left\{ \langle C - S, \tilde{P}W\tilde{P}^T \rangle : \mathcal{A}(\tilde{P}W\tilde{P}^T) = b,\ W \succeq 0 \right\} = \sup_y \left\{ b^T y : \tilde{P}^T(C - \mathcal{A}^* y)\tilde{P} \succeq \tilde{P}^T S \tilde{P} \right\}. \tag{4.19}$$

Since $\|\tilde{P}^T S \tilde{P}\| \leq \|S\| \leq \bar{\epsilon}$, we can use (4.16), (4.17), (4.19) and (4.18) to get

$$\mathrm{val_P}(S) - \mathrm{val_P}(0) = \mathrm{val_P}(S) - \bar{v}_\mathrm{P} \leq \kappa \|\tilde{P}^T S \tilde{P}\|^{1/2^{\mathrm{d}(\mathcal{A},C)}} \leq \kappa \|S\|^{1/2^{\mathrm{d}(\mathcal{A},C)}}.$$

□

Remark 4.11 *The assumption that $v_\mathrm{P} = v_\mathrm{D}$ is important in this proof because this ensures that $v_\mathrm{P} = \bar{v}_\mathrm{P}$ in* (4.16)*, which generally does not hold because the feasible region of* (4.15) *is only a subset of the feasible region of* (P)*.*

5 Conclusion

In this paper we have studied the sensitivity analysis for feasible SDP in the case of where strong duality fails. We have used the notions of *asymptotics*, *facial reduction* and *degree of singularity* to find bounds on the growth of the objective value based on the size of feasible perturbations. In particular, the relative growth is *infinite* along *all* feasible perturbation directions in the case of a nonzero duality gap, while it is bounded by a constant dependent on the degree of singularity if there is a zero duality gap. Examples, including semidefinite completion problems, are included to illustrate the close relationship between the actual growth and the derived error bounds.

One of our main tools is the degree of singularity that is based on the number of facial reduction steps. This illustrates the usefulness of carefully analyzing and taking advantage of the geometry of SDP.

Appendix A Facial reduction

We first present a basic version of the facial reduction algorithm, from [16], in Algorithm A.1. The algorithm takes a linear subspace $\mathcal{L} \subseteq \mathbb{S}^n$ as input, and outputs the degree of singularity $\mathrm{d}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+)$ as well as the minimal face $\mathrm{face}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+)$. We start with $\mathcal{K}_0 := \mathbb{S}^n_+$; in each iteration j of the facial reduction algorithm, we either determine that $\mathcal{L} \cap \mathrm{ri}(\mathcal{K}_j) \neq \emptyset$, or find $D^{(j+1)} \in \mathrm{ri}(\mathcal{L}^\perp \cap (\mathcal{K}_j)^*) \backslash (-\mathcal{K}_j)^*$ and update $\mathcal{K}_{j+1} = \mathcal{K}_j \cap \{D^{(j+1)}\}^\perp$. In particular, Algorithm A.1 finds a finite sequence $(D^{(1)}, \ldots, D^{(\mathrm{d}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+))})$ that is an element of $\mathcal{C}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+)$ and a corresponding sequence of cones $\mathcal{K}_1 \supseteq \mathcal{K}_2 \supseteq \cdots \supseteq \mathcal{K}_{\mathrm{d}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+)} \supseteq \mathcal{L} \cap \mathbb{S}^n_+$.

Algorithm A.1: Facial reduction algorithm

Input: linear subspace $\mathcal{L} \subseteq \mathbb{S}^n$
Output: degree of singularity $\mathrm{d}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+)$, minimal face of $\mathcal{F} = \mathrm{face}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+)$

find $D^{(1)} \in \mathrm{ri}(\mathcal{L}^\perp \cap \mathbb{S}^n_+)$;
if $D^{(1)} = 0$ *or* $D^{(1)} \succ 0$ **then**
    `% minimal face found`;
    if $D^{(1)} = 0$ **then**
        $\mathcal{F} \leftarrow \mathbb{S}^n_+$;
    else
        $\mathcal{F} \leftarrow \{0\}$;
    endif
    $\mathrm{d}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+) \leftarrow 0$;
    STOP;
else
    $\mathcal{K}_1 \leftarrow \mathbb{S}^n_+ \cap \left\{D^{(1)}\right\}^\perp$;
endif

for *j=2,...* **do**
    find $D^{(j)} \in \mathrm{ri}(\mathcal{L}^\perp \cap (\mathcal{K}_{j-1})^*)$;
    if $D^{(j)} \in (-\mathcal{K}_{j-1})^*$ **then**
        $\mathrm{d}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+) \leftarrow j - 1$;
        $\mathcal{F} \leftarrow \mathcal{K}_{j-1}$;
        STOP;
    else
        $\mathcal{K}_j \leftarrow \mathcal{K}_{j-1} \cap \left\{D^{(j)}\right\}^\perp$;
    endif
endfor

Observe that the dimension of $\mathcal{K}_{j+1} = \mathcal{K}_j \cap \left\{D^{(j+1)}\right\}^{\perp}$ must be strictly less than the dimension of $\mathcal{K}_j$, by the choice of $D^{(j+1)}$. Hence the for-loop in Algorithm A.1 must terminate in finitely many iterations.

We can describe more precisely the relationship between any finite sequence $(D^{(1)}, \ldots, D^{(\mathrm{d}(\mathcal{L}\cap\mathbb{S}^n_+, \mathbb{S}^n_+))})$ in $\mathcal{C}(\mathcal{L}\cap\mathbb{S}^n_+, \mathbb{S}^n_+)$ (such as the one found in Algorithm A.1) and the corresponding sequence of cones $\mathcal{K}_1 \supseteq \mathcal{K}_2 \supseteq \cdots \supseteq \mathcal{K}_{\mathrm{d}(\mathcal{L}\cap\mathbb{S}^n_+, \mathbb{S}^n_+)}$.

Proposition A.1 *Let $\mathcal{L} \subseteq \mathbb{S}^n$ be a linear subspace with $k := \mathrm{d}(\mathcal{L}\cap\mathbb{S}^n_+, \mathbb{S}^n_+) > 0$. Pick any $(D^{(1)}, \ldots, D^{(k)}) \in \mathcal{C}(\mathcal{L}\cap\mathbb{S}^n_+, \mathbb{S}^n_+)$, and define*

$$\mathcal{K}_0 := \mathbb{S}^n_+, \quad \mathcal{K}_j := \mathcal{K}_{j-1} \cap \{D^{(j)}\}^{\perp}, \ \forall\, j = 1, \ldots, k.$$

Then there exist full column rank matrices $Q^{(1)} \in \mathbb{R}^{n\times n_1}, Q^{(2)} \in \mathbb{R}^{n_1\times n_2}, \ldots, Q^{(k)} \in \mathbb{R}^{n_{k-1}\times n_k}$ with $\bar{Q}^{(j)} := Q^{(1)}Q^{(2)}\cdots Q^{(j)} \in \mathbb{R}^{n\times n_j}$ (for all $j = 1, \ldots, k$) satisfying

(1) $0 \neq D^{(1)} \succeq 0$ and $0 \neq (\bar{Q}^{(j)})^T D^{(j+1)} \bar{Q}^{(j)} \succeq 0$, for all $j = 1, \ldots, k-1$;

(2) $\ker(D^{(1)}) = \mathcal{R}(Q^{(1)})$ and $\ker((\bar{Q}^{(j)})^T D^{(j+1)} \bar{Q}^{(j)}) = \mathcal{R}(Q^{(j+1)})$ for all $j = 1, \ldots, k-1$; and

(3) $\mathcal{K}_j = \bar{Q}^{(j)}\, \mathbb{S}^{n_j}_+ (\bar{Q}^{(j)})^T$ for all $j = 1, \ldots, k$.

Proof. Since $0 \neq D^{(1)} \in \mathrm{ri}(\mathcal{L}^{\perp}\cap\mathbb{S}^n_+)$, the conditions (1)-(3) immediately hold for $j = 1$.

Suppose that the conditions hold for all $1 \leq j \leq \bar{k}$, where $\bar{k} < k$; we show that they hold for $j = \bar{k}+1$ too. Recall that $D^{(\bar{k}+1)} \in (\mathcal{K}_{\bar{k}})^* \backslash (-\mathcal{K}_{\bar{k}})^*$. By the induction assumption that $\mathcal{K}_{\bar{k}} = \bar{Q}^{(\bar{k})}\, \mathbb{S}^{n_{\bar{k}}}_+ (\bar{Q}^{(\bar{k})})^T$, we get

$$0 \neq (\bar{Q}^{(\bar{k})})^T D^{(\bar{k}+1)} \bar{Q}^{(\bar{k})} \succeq 0, \tag{A.1}$$

so condition (1) holds and there exists a full column rank matrix $Q^{(\bar{k}+1)} \in \mathbb{R}^{n_{\bar{k}}\times n_{\bar{k}+1}}$ such that $\ker((\bar{Q}^{(\bar{k})})^T D^{(\bar{k}+1)} \bar{Q}^{(\bar{k})}) = \mathcal{R}(Q^{(\bar{k}+1)})$, i.e., condition (2) holds. Finally, by (A.1) and the definition of $Q^{(\bar{k}+1)}$,

$$\mathcal{K}_{\bar{k}+1} = \bar{Q}^{(\bar{k})}\, \mathbb{S}^{n_{\bar{k}}}_+ (\bar{Q}^{(\bar{k})})^T \cap \{D^{(\bar{k}+1)}\}^{\perp} = \bar{Q}^{(\bar{k})} Q^{(\bar{k}+1)}\, \mathbb{S}^{n_{\bar{k}}}_+ (Q^{(\bar{k}+1)})^T (\bar{Q}^{(\bar{k})})^T,$$

so condition (3) holds too. $\square$

An immediate consequence of Proposition A.1 is that, by looking at appropriate nullspaces associated with the matrices $D^{(1)}, \ldots, D^{(k)}$, we can get an explicit description of the minimal face $\mathrm{face}(\mathcal{L}\cap\mathbb{S}^n_+, \mathbb{S}^n_+)$, in the form of $Q\,\mathbb{S}^{\bar{n}}_+ Q^T$.

It is also immediate from Proposition A.1 that the facial reduction algorithm presented in [20] is equivalent to Algorithm A.1; in [20], a *rank-revealing rotation* is applied in each step, i.e., $\mathcal{L}$ is updated to be $(\bar{Q}^{(j)})^T \mathcal{L} \bar{Q}^{(j)}$ and $\mathcal{K}_j \leftarrow (\bar{Q}^{(j)})^T \mathcal{K}_j \bar{Q}^{(j)} = \begin{bmatrix} 0 & 0 \\ 0 & \mathbb{S}_+^{\bar{n}_j} \end{bmatrix}$, in the j-th iteration.

Appendix B Well-definedness of the degree of singularity.

The sequence of matrices $(D^{(1)}, \dots, D^{(\mathrm{d}(\mathcal{L} \cap \mathbb{S}_+^n, \mathbb{S}_+^n))})$ found in Algorithm A.1 is an element of the set $\mathcal{C}(\mathcal{L} \cap \mathbb{S}_+^n, \mathbb{S}_+^n)$ (defined on Page 9). Naturally, Algorithm A.1 may use any arbitrary sequence from $\mathcal{C}(\mathcal{L} \cap \mathbb{S}_+^n, \mathbb{S}_+^n)$. Such flexibility in the choice, however, is not an issue: in Proposition B.1 below we show that any two distinct sequences in $\mathcal{C}(\mathcal{L} \cap \mathbb{S}_+^n, \mathbb{S}_+^n)$ provide the same information. In particular, the degree of singularity in Definition 2.7 is well-defined.

Proposition B.1 *Let $\mathcal{L} \subseteq \mathbb{S}^n$ be a linear subspace, and let $(D^{(1)}, \dots, D^{(k)}), (\hat{D}^{(1)}, \dots, \hat{D}^{(\hat{k})}) \in \mathcal{C}(\mathcal{L} \cap \mathbb{S}_+^n, \mathbb{S}_+^n)$. Then $k = \hat{k}$. Moreover, for each $j = 1, \dots, k$, letting $\mathcal{K}_j := \mathbb{S}_+^n \cap \{D^{(1)}\}^\perp \cap \dots \cap \{D^{(j)}\}^\perp$ and $\hat{\mathcal{K}}_j := \mathbb{S}_+^n \cap \{\hat{D}^{(1)}\}^\perp \cap \dots \cap \{\hat{D}^{(j)}\}^\perp$, we have that $\mathcal{K}_j = \hat{\mathcal{K}}_j = \bar{Q}^{(j)} \, \mathbb{S}_+^{n_j} (\bar{Q}^{(j)})^T$ for some full column rank matrix $\bar{Q}^{(j)}$.*

Proof. It is immediate that $\mathcal{R}(D^{(1)}) = \mathcal{R}(\hat{D}^{(1)})$, so $\mathcal{K}_1 = \hat{\mathcal{K}}_1 = \bar{Q}^{(1)} \, \mathbb{S}_+^{n_1} (\bar{Q}^{(1)})^T$, where $\bar{Q}^{(1)}$ is any full column rank matrix such that $\mathcal{R}(\bar{Q}^{(1)}) = \ker(D^{(1)}) = \ker(\hat{D}^{(1)})$. Suppose that $\mathcal{K}_{j-1} = \hat{\mathcal{K}}_{j-1}$ for all $1 \le j \le \bar{k}$, for some $1 \le \bar{k} \le \min\{k, \hat{k}\} - 1$. Since $D^{(j)}, \hat{D}^{(j)} \in \mathrm{ri}(\mathcal{L}^\perp \cap (\mathcal{K}_{j-1})^*)$, by Proposition A.1 we have that $\mathcal{K}_j = \mathcal{K}_{j-1} \cap \{D^{(j)}\}^\perp = \mathcal{K}_{j-1} \cap \{\hat{D}^{(j)}\}^\perp = \hat{\mathcal{K}}_j = \bar{Q}^{(j)} \, \mathbb{S}_+^{n_j} (\bar{Q}^{(j)})^T$ for some full column rank matrix $\bar{Q}^{(j)}$, by Item (3) of Proposition A.1. Hence $\mathcal{K}_j = \hat{\mathcal{K}}_j = \bar{Q}^{(j)} \, \mathbb{S}_+^{n_j} (\bar{Q}^{(j)})^T$ for all $j = 1, \dots, \min\{k, \hat{k}\}$.

Next we show that $k = \hat{k}$. Suppose without loss of generality that $k \le \hat{k}$. Then by Proposition 2.5,

$$\mathcal{L}^\perp \cap (\hat{\mathcal{K}}_k)^* = \mathcal{L}^\perp \cap (\mathcal{K}_k)^* \subseteq (-\mathcal{K}_k)^* = (-\hat{\mathcal{K}}_k)^*,$$

which implies that $\mathcal{L} \cap \mathrm{ri}(\hat{\mathcal{K}}_k) \neq \emptyset$. Hence $\hat{k} = k$. $\square$

From Proposition B.1, we see that for fixed linear subspace $\mathcal{L} \subseteq \mathbb{S}^n$, any finite sequence $(D^{(1)}, \dots, D^{(\mathrm{d}(\mathcal{L} \cap \mathbb{S}_+^n, \mathbb{S}_+^n))})$ satisfying the conditions in Definition 2.6 gives the same chain of cones $\mathcal{K}_1 \supseteq \mathcal{K}_2 \supseteq \dots \supseteq \mathcal{K}_{\mathrm{d}(\mathcal{L} \cap \mathbb{S}_+^n, \mathbb{S}_+^n)}$.

Appendix C Proof of Lemma 4.9.

In the following, we prove Lemma 4.9. We first need a few elementary results. Fix any $\mathcal{K} \trianglelefteq \mathbb{S}_+^n$ and any linear subspace $\mathcal{L} \subseteq \mathrm{span}(\mathcal{K})$. Then $\mathcal{K} = Q \, \mathbb{S}_+^{\bar{n}} Q^T$ for some matrix

$Q \in \mathbb{R}^{n\times\bar{n}}$ with orthonormal columns and $1 \le \bar{n} < n$, and the equality $\mathcal{L} \cap Q\,\mathbb{S}^{\bar{n}}_+ Q^T = Q(Q^T\mathcal{L}Q \cap \mathbb{S}^{\bar{n}}_+)Q^T$ holds. Define

$$\mathrm{d}(\mathcal{L}\cap\mathcal{K},\mathcal{K}) = \mathrm{d}(\mathcal{L}\cap Q\,\mathbb{S}^{\bar{n}}_+Q^T, Q\,\mathbb{S}^{\bar{n}}_+Q^T) := \mathrm{d}(Q^T\mathcal{L}Q\cap\mathbb{S}^{\bar{n}}_+, \mathbb{S}^{\bar{n}}_+).$$

We first prove that one iteration of facial reduction does decrease the degree of singularity by 1.

Lemma C.1 *Let $\mathcal{L} \subseteq \mathbb{S}^n$ be a linear subspace such that $\mathrm{d}(\mathcal{L}\cap\mathbb{S}^n_+, \mathbb{S}^n_+) > 0$. Fix any $D \in \mathrm{ri}(\mathcal{L}^\perp \cap \mathbb{S}^n_+)$. Then*

$$\mathrm{d}(\tilde{\mathcal{L}}\cap\tilde{\mathcal{K}},\tilde{\mathcal{K}}) = \mathrm{d}(\mathcal{L}\cap\mathbb{S}^n_+, \mathbb{S}^n_+) - 1, \quad \textit{where } \tilde{\mathcal{L}} := \mathcal{L}\cap\mathrm{span}(\tilde{\mathcal{K}}) \textit{ and } \tilde{\mathcal{K}} = \mathbb{S}^n_+\cap\{D\}^\perp.$$

Proof. Let $k := \mathrm{d}(\mathcal{L}\cap\mathbb{S}^n_+, \mathbb{S}^n_+) - 1 \ge 0$ and $Q \in \mathbb{R}^{n\times\bar{n}}$ be a matrix with orthonormal columns such that $\ker(D) = \mathcal{R}(Q)$. Then $\tilde{\mathcal{K}} = Q\,\mathbb{S}^{\bar{n}}_+Q^T$.

If $k = 0$, then there exists some $X \in \mathcal{L}\cap\mathrm{ri}(\mathbb{S}^n_+\cap\{D\}^\perp)$, i.e., $X = Q\bar{X}Q^T \in \mathcal{L}$ for some $\bar{X} \succ 0$. Then $\bar{X} \in Q^T\tilde{\mathcal{L}}Q \cap \mathbb{S}^{\bar{n}}_{++}$, i.e., $\mathrm{d}(\tilde{\mathcal{L}}\cap\tilde{\mathcal{K}},\tilde{\mathcal{K}}) = 0 = k$.

Suppose that $k > 0$. Pick any $D^{(1)},\dots,D^{(k)} \in \mathbb{S}^n$ such that $(D, D^{(1)},\dots,D^{(k)}) \in \mathcal{C}(\mathcal{L}\cap\mathbb{S}^n_+, \mathbb{S}^n_+)$. By Definition 2.6, we have that

$$D^{(j)} \in \mathrm{ri}(\mathcal{L}^\perp\cap(\mathcal{K}_{j-1})^*)\backslash(-\mathcal{K}_{j-1})^* \quad \text{and} \quad \mathcal{L}\cap\mathrm{ri}(\mathcal{K}_k) \neq \emptyset, \tag{C.1}$$

where $\mathcal{K}_0 = \mathbb{S}^n_+\cap\{D\}^\perp = Q\,\mathbb{S}^{\bar{n}}_+Q^T$ and $\mathcal{K}_j = \mathcal{K}_{j-1}\cap\{D^{(j)}\}^\perp$ for all $j = 1,\dots,k$. Define

$$\begin{aligned}
\bar{\mathcal{L}} &:= Q^T\tilde{\mathcal{L}}Q \\
\bar{D}^{(j)} &:= Q^TD^{(j)}Q, \quad \forall\, j = 1,\dots,k, \\
\bar{\mathcal{K}}_0 &:= \mathbb{S}^{\bar{n}}_+ \quad \text{and} \quad \bar{\mathcal{K}}_j := \bar{\mathcal{K}}_{j-1}\cap\{\bar{D}^{(j)}\}^\perp, \quad \forall\, j = 1,\dots,k.
\end{aligned}$$

We show that

(1) $\mathcal{K}_j = Q\bar{\mathcal{K}}_jQ^T$, for all $j = 1,\dots,k$;

(2) $\bar{\mathcal{L}}^\perp\cap(\bar{\mathcal{K}}_j)^* = Q^T(\mathcal{L}^\perp\cap(\mathcal{K}_j)^*)Q$, for all $j = 1,\dots,k$;

(3) $\bar{D}^{(j)} \in \mathrm{ri}(\bar{\mathcal{L}}^\perp\cap(\bar{\mathcal{K}}_{j-1})^*)\backslash(-\bar{\mathcal{K}}_{j-1})^*$, for all $j = 1,\dots,k$; and

(4) $\bar{\mathcal{L}}\cap\mathrm{ri}(\bar{\mathcal{K}}_k) \neq \emptyset$.

Fix any $j \in \{1,\dots,k\}$.

For item (1), pick any $X \in \mathcal{K}_j \subseteq \mathcal{K}_0$. Then $X = Q\bar{X}Q^T$ for some $\bar{X} \succeq 0$, and for all $j = 1,\dots,k$, $\langle\bar{X},\bar{D}^{(j)}\rangle = \langle\bar{X}, QD^{(j)}Q^T\rangle = \langle X, D^{(j)}\rangle = 0$. Hence $\bar{X} \in \bar{\mathcal{K}}_j$, implying that

$\mathcal{K}_j \subseteq Q\bar{\mathcal{K}}_j Q^T$. Conversely, pick any $\bar{X} \in \bar{\mathcal{K}}_j$ and define $X := Q\bar{X}Q^T$. Then $X \in \mathcal{K}_0$ and for all $j = 1, \ldots, k$, $\langle X, D^{(j)} \rangle = \langle \bar{X}, \bar{D}^{(j)} \rangle = 0$. Hence $Q\bar{\mathcal{K}}_j Q^T \subseteq \mathcal{K}_j$.

For item (2), pick any $X \in \mathcal{L}^\perp \cap (\mathcal{K}_j)^*$ and let $\bar{X} := Q^T X Q$. For all $\bar{Y} \in \bar{\mathcal{L}}$, $\bar{Y} = Q^T Y Q$ for some $Y \in \tilde{\mathcal{L}} = \mathcal{L} \cap Q\,\mathbb{S}^{\bar{n}} Q^T$, so $0 = \langle X, Y \rangle = \langle X, QQ^T Y QQ \rangle = \langle \bar{X}, \bar{Y} \rangle$. Hence $\bar{X} \in \bar{\mathcal{L}}^\perp$. By (1), we have

$$(\mathcal{K}_j)^* = (Q\bar{\mathcal{K}}_j Q^T)^* = \left\{ W \in \mathbb{S}^n : Q^T W Q \in (\bar{\mathcal{K}}_j)^* \right\},$$

so $\bar{X} \in (\bar{\mathcal{K}}_j)^*$. Therefore we have $Q^T(\mathcal{L}^\perp \cap (\mathcal{K}_j)^*)Q \subseteq \bar{\mathcal{L}}^\perp \cap (\bar{\mathcal{K}}_j)^*$. Conversely, pick any $\bar{X} \in \bar{\mathcal{L}}^\perp \cap (\bar{\mathcal{K}}_j)^*$ and define $X := Q\bar{X}Q^T$. Then $X \in \tilde{\mathcal{L}}^\perp = \mathcal{L}^\perp + (\mathcal{K}_0)^\perp$. Let $X^{(1)} \in \mathcal{L}^\perp$ and $X^{(2)} \in (\mathcal{K}_0)^\perp$ satisfy $X = X^{(1)} + X^{(2)}$. Observe that $X^{(1)} \in (\mathcal{K}_j)^*$: for any $Y \in \mathcal{K}_j$, there exists $\bar{Y} \in \bar{\mathcal{K}}_j$ such that $Y = Q\bar{Y}_j Q^T$ by item (1), and $\langle X^{(1)}, Y \rangle = \langle X - X^{(2)}, Q\bar{Y}Q^T \rangle = \langle \bar{X}, \bar{Y} \rangle \geq 0$ since $\bar{X} \in (\bar{\mathcal{K}}_j)^*$. Therefore $X^{(1)} \in \mathcal{L}^\perp \cap (\mathcal{K}_j)^*$, and $\bar{X} = Q^T X Q = Q^T X^{(1)} Q \in Q^T(\mathcal{L}^\perp \cap (\mathcal{K}_j)^*)Q$, showing that $\bar{X} \in \bar{\mathcal{L}}^\perp \cap (\bar{\mathcal{K}}_j)^* \subseteq Q^T(\mathcal{L}^\perp \cap (\mathcal{K}_j)^*)Q$.

Item (3) follows from (C.1) and item (2):

$$\begin{aligned}
\bar{D}^{(j)} &\in Q^T\big(\mathrm{ri}(\mathcal{L}^\perp \cap (\mathcal{K}_j)^*)\big)Q \backslash (-\bar{\mathcal{K}}_j)^* \\
&= \mathrm{ri}\big(Q^T\big(\mathcal{L}^\perp \cap (\mathcal{K}_j)^*\big)Q\big) \backslash (-\bar{\mathcal{K}}_j)^* = \mathrm{ri}(\bar{\mathcal{L}}^\perp \cap (\bar{\mathcal{K}}_j)^*) \backslash (-\bar{\mathcal{K}}_j)^*.
\end{aligned}$$

Finally, for item (4), by (C.1) there exists

$$\begin{aligned}
X &\in \mathcal{L} \cap \mathrm{ri}(\mathcal{K}_k) = \mathcal{L} \cap \mathrm{ri}(Q\bar{\mathcal{K}}_k Q^T) = \mathcal{L} \cap Q\,\mathrm{ri}(\mathcal{K}_k)Q^T = \tilde{\mathcal{L}} \cap Q\big(\mathrm{ri}(\mathcal{K}_k)\big)Q^T \\
&\subseteq Q\big(\bar{\mathcal{L}} \cap \mathrm{ri}(\mathcal{K}_k)\big)Q^T.
\end{aligned}$$

Hence $\bar{\mathcal{L}} \cap \mathrm{ri}(\mathcal{K}_k) \neq \emptyset$.

Consequently, from items (1)-(4) we get that $(\bar{D}^{(1)}, \ldots, \bar{D}^{(k)}) \in \mathcal{C}(\bar{\mathcal{L}} \cap \mathbb{S}^{\bar{n}}_+, \mathbb{S}^{\bar{n}}_+)$, and

$$\mathrm{d}(\tilde{\mathcal{L}} \cap \tilde{\mathcal{K}}, \tilde{\mathcal{K}}) = \mathrm{d}(\bar{\mathcal{L}} \cap \mathbb{S}^{\bar{n}}_+, \mathbb{S}^{\bar{n}}_+) = k = \mathrm{d}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+) - 1.$$

□

Next we show that the degree of singularity is *monotonic* under a mild assumption.

Lemma C.2 *Let $\mathcal{L}_1, \mathcal{L}_2 \subseteq \mathbb{S}^n$ be linear subspaces. If $\mathcal{L}_1 \subseteq \mathcal{L}_2$ and $\mathcal{L}_1 \cap \mathbb{S}^n_+ \neq \{0\}$, then $\mathrm{d}(\mathcal{L}_2 \cap \mathbb{S}^n_+, \mathbb{S}^n_+) \leq \mathrm{d}(\mathcal{L}_1 \cap \mathbb{S}^n_+, \mathbb{S}^n_+)$.*

Proof. We prove by induction on $\mathrm{d}(\mathcal{L}_1 \cap \mathbb{S}^n_+, \mathbb{S}^n_+) \geq 0$.

If $\mathrm{d}(\mathcal{L}_1 \cap \mathbb{S}^n_+, \mathbb{S}^n_+) = 0$, then $\emptyset \neq \mathcal{L}_1 \cap \mathbb{S}^n_{++} \subseteq \mathcal{L}_2 \cap \mathbb{S}^n_{++}$. Hence $\mathrm{d}(\mathcal{L}_2 \cap \mathbb{S}^n_+, \mathbb{S}^n_+) \leq \mathrm{d}(\mathcal{L}_1 \cap \mathbb{S}^n_+, \mathbb{S}^n_+)$.

Now fix any linear subspace $\mathcal{L}_1 \subseteq \mathcal{L}_2 \subseteq \mathbb{S}^n$ with $\mathcal{L}_1 \cap \mathbb{S}^n_+ \neq \{0\}$ and $\mathrm{d}(\mathcal{L}_1 \cap \mathbb{S}^n_+, \mathbb{S}^n_+) = k > 0$. We prove that $\mathrm{d}(\mathcal{L}_2 \cap \mathbb{S}^n_+, \mathbb{S}^n_+) \leq \mathrm{d}(\mathcal{L}_1 \cap \mathbb{S}^n_+, \mathbb{S}^n_+)$. If $\mathrm{d}(\mathcal{L}_2 \cap \mathbb{S}^n_+, \mathbb{S}^n_+) = 0$, then the inequality trivially holds. Suppose that $\mathrm{d}(\mathcal{L}_2 \cap \mathbb{S}^n_+, \mathbb{S}^n_+) > 0$. Since $k > 0$, there exist

$$0 \neq D_j \in \mathrm{ri}(\mathcal{L}_j^\perp \cap \mathbb{S}^n_+) \quad \text{with} \quad D_j \not\succ 0, \quad \text{for } j = 1, 2.$$

Since $\mathcal{L}_2^\perp \cap \mathbb{S}^n_+ \subseteq \mathcal{L}_1^\perp \cap \mathbb{S}^n_+$, we have that $\mathcal{R}(D_2) \subseteq \mathcal{R}(D_1)$, implying that $\{D_1\}^\perp \subseteq \{D_2\}^\perp$. Therefore $\tilde{\mathcal{L}}_1 \subseteq \tilde{\mathcal{L}}_2$ and $\tilde{\mathcal{K}}_1 \subseteq \tilde{\mathcal{K}}_2$, where

$$\tilde{\mathcal{L}}_j := \mathcal{L}_j \cap \mathrm{span}(\mathbb{S}^n_+ \cap \{D_j\}^\perp) \;\text{ and }\; \tilde{\mathcal{K}}_j := \mathbb{S}^n_+ \cap \{D_j\}^\perp, \quad \text{for } j = 1, 2.$$

Moreover, by Lemma C.1, we have $\mathrm{d}(\tilde{\mathcal{L}}_j \cap \tilde{\mathcal{K}}_j, \tilde{\mathcal{K}}_j) = \mathrm{d}(\mathcal{L}_j \cap \mathbb{S}^n_+, \mathbb{S}^n_+) - 1$ for $j = 1, 2$. Also, since $\mathcal{L}_1 \cap \mathbb{S}^n_+ \neq \{0\}$, we must have $\tilde{\mathcal{L}}_1 \cap \tilde{\mathcal{K}}_1 = \mathcal{L}_1 \cap \mathbb{S}^n_+ \neq \{0\}$. Therefore, by the induction hypothesis, we have $\mathrm{d}(\tilde{\mathcal{L}}_2 \cap \tilde{\mathcal{K}}_2, \tilde{\mathcal{K}}_2) \leq \mathrm{d}(\tilde{\mathcal{L}}_1 \cap \tilde{\mathcal{K}}_1, \tilde{\mathcal{K}}_1)$, implying that $\mathrm{d}(\mathcal{L}_2 \cap \mathbb{S}^n_+, \mathbb{S}^n_+) \leq \mathrm{d}(\mathcal{L}_1 \cap \mathbb{S}^n_+, \mathbb{S}^n_+)$. $\square$

We return our focus on the primal-dual pair (P)-(D). In the context of Lemma 4.9, we are interested in the situation where (D) is feasible but fails the Slater condition. In that case, by Proposition 2.5,

$$\exists v \in \mathbb{R}^m \;\text{ s.t. }\; b^T v = 0, \;\; 0 \neq V := \mathcal{A}^* v \succeq 0;$$

the feasible region of (D) is contained in $\mathbb{S}^n_+ \cap \{V\}^\perp$ and V can be used for one iteration of the facial reduction algorithm on (D). To prove Lemma 4.9, we show that each iteration of facial reduction on (D) does not increase the degree of singularity of the corresponding (new) primal, using the following lemma.

Lemma C.3 *Let $\mathcal{L} \subseteq \mathbb{S}^n$ be a linear subspace and suppose that*

$$\exists\, V = \begin{bmatrix} 0 & 0 \\ 0 & \bar{V} \end{bmatrix} \in \mathcal{L} \;\; s.t. \;\; \bar{V} \in \mathbb{S}^{n-r}_{++} \quad (0 < r < n). \tag{C.2}$$

Define the linear subspace $\hat{\mathcal{L}} := \begin{bmatrix} I_r & 0 \end{bmatrix} \mathcal{L} \begin{bmatrix} I_r \\ 0 \end{bmatrix} \subseteq \mathbb{S}^r$. Then

$$\mathrm{d}(\hat{\mathcal{L}} \cap \mathbb{S}^r_+, \mathbb{S}^r_+) \leq \mathrm{d}(\mathcal{L} \cap \mathbb{S}^n_+, \mathbb{S}^n_+). \tag{C.3}$$

Proof. We prove by induction on $\mathrm{d}(\hat{\mathcal{L}} \cap \mathbb{S}^r_+, \mathbb{S}^r_+)$.

If $\mathrm{d}(\hat{\mathcal{L}} \cap \mathbb{S}^r_+, \mathbb{S}^r_+) = 0$, then (C.3) immediately holds.

Now assume that $\mathrm{d}(\hat{\mathcal{L}} \cap \mathbb{S}^r_+, \mathbb{S}^r_+) > 0$. Then there exists $0 \neq \hat{D} \in \mathrm{ri}(\hat{\mathcal{L}}^\perp \cap \mathbb{S}^r_+)$ with

$\hat{D} \not\succ 0$. Let $\hat{Q} \in \mathbb{R}^{r\times r_1}$ satisfy $\mathcal{R}(\hat{Q}) = \ker(\hat{D})$ and $Q^T Q = I$, and let

$$D := \begin{bmatrix} \hat{D} & 0 \\ 0 & 0 \end{bmatrix} \in \begin{bmatrix} \hat{\mathcal{L}}^\perp \cap \mathbb{S}^r_+ & 0 \\ 0 & 0 \end{bmatrix} = \mathcal{L}^\perp \cap \mathbb{S}^n_+;$$

then we have $D \in \operatorname{ri}(\mathcal{L}^\perp \cap \mathbb{S}^n_+)$ and $\mathcal{R}(Q) = \ker(D)$, where $Q := \begin{bmatrix} \hat{Q} & 0 \\ 0 & I_{n-r} \end{bmatrix}$. Define

$$\begin{aligned} \mathcal{L}_1 &:= \mathcal{L} \cap \operatorname{span}(\mathbb{S}^n_+ \cap \{D\}^\perp) = \mathcal{L} \cap Q\,\mathbb{S}^{n-r+r_1} Q^T \\ \text{and} \qquad \hat{\mathcal{L}}_1 &:= \hat{\mathcal{L}} \cap \operatorname{span}(\mathbb{S}^r_+ \cap \{\hat{D}\}^\perp) = \hat{\mathcal{L}} \cap \hat{Q}\,\mathbb{S}^{r_1} Q^T. \end{aligned}$$

We show that $\begin{bmatrix} I_r & 0 \end{bmatrix} \mathcal{L}_1 \begin{bmatrix} I_r \\ 0 \end{bmatrix} \subseteq \hat{\mathcal{L}}_1$. For any $X \in \mathcal{L}_1$, we have $\bar{X} := \begin{bmatrix} I_r & 0 \end{bmatrix} X \begin{bmatrix} I_r \\ 0 \end{bmatrix} \in \hat{\mathcal{L}}$. It remains to show that $\bar{X} \in \hat{Q}\,\mathbb{S}^{r_1} \hat{Q}^T$. Since $X \in \mathcal{L}_1$, there exists $W \in \mathbb{S}^{n-r+r_1}$ such that $X = QWQ^T$. By definition of Q, we get

$$\bar{X} = \begin{bmatrix} I_r & 0 \end{bmatrix} QWQ^T \begin{bmatrix} I_r \\ 0 \end{bmatrix} = \begin{bmatrix} \hat{Q} & 0 \end{bmatrix} \begin{array}{c} \\ {\scriptstyle r_1} \\ {\scriptstyle n-r} \end{array} \!\!\begin{array}{c} \begin{array}{cc} {\scriptstyle r_1} & {\scriptstyle n-r} \end{array} \\ \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \end{array} \begin{bmatrix} \hat{Q}^T \\ 0 \end{bmatrix} = \hat{Q} W_{11} \hat{Q}^T.$$

Hence $\bar{X} \in \hat{Q}\,\mathbb{S}^{r_1} \hat{Q}^T$, implying that $\bar{X} \in \hat{\mathcal{L}}$.

Next we show that $\begin{bmatrix} \hat{Q}^T & 0 \end{bmatrix} \mathcal{L}_1 \begin{bmatrix} \hat{Q} \\ 0 \end{bmatrix} \cap \mathbb{S}^{r_1}_+ \neq \{0\}$. On the contrary, suppose that

$$\begin{bmatrix} \hat{Q}^T & 0 \end{bmatrix} \mathcal{L}_1 \begin{bmatrix} \hat{Q} \\ 0 \end{bmatrix} \cap \mathbb{S}^{r_1}_+ = \{0\}, \quad \text{or equivalently,} \;\; \exists \hat{F} \in \left(\begin{bmatrix} \hat{Q}^T & 0 \end{bmatrix} \mathcal{L}_1 \begin{bmatrix} \hat{Q} \\ 0 \end{bmatrix} \right)^\perp \cap \mathbb{S}^{r_1}_{++}.$$

Then $\begin{bmatrix} \hat{Q}\hat{F}\hat{Q}^T & 0 \\ 0 & 0 \end{bmatrix} \in \mathcal{L}_1^\perp = \mathcal{L}^\perp + \operatorname{span}(D) = \mathcal{L}^\perp$. Since $\hat{F} \succ 0$ and $\mathcal{R}(\hat{Q}) = \ker(\hat{D})$, we have that $\begin{bmatrix} \hat{F} + \hat{D} & 0 \\ 0 & 0 \end{bmatrix} \in \mathcal{L}^\perp \cap \mathbb{S}^n_{++}$, which would imply that $\mathcal{L} \cap \mathbb{S}^n_+ = \{0\}$, contradicting the assumption (C.2).

Using $\begin{bmatrix} I_r & 0 \end{bmatrix} \mathcal{L}_1 \begin{bmatrix} I_r \\ 0 \end{bmatrix} \subseteq \hat{\mathcal{L}}_1 \subseteq \hat{Q}\,\mathbb{S}^{r_1}_+ \hat{Q}^T$ and $\begin{bmatrix} \hat{Q}^T & 0 \end{bmatrix} \mathcal{L}_1 \begin{bmatrix} \hat{Q} \\ 0 \end{bmatrix} \cap \mathbb{S}^{r_1}_+ \neq \{0\}$, we can apply Lemma C.2 to get

$$\mathrm{d}(\hat{\mathcal{L}}_1 \cap \hat{Q}\,\mathbb{S}^{r_1}_+ \hat{Q}^T, \hat{Q}\,\mathbb{S}^{r_1}_+ \hat{Q}^T) \le \mathrm{d}\left(\begin{bmatrix} I_r & 0 \end{bmatrix} \mathcal{L}_1 \begin{bmatrix} I_r \\ 0 \end{bmatrix} \cap \hat{Q}\,\mathbb{S}^{r_1}_+ \hat{Q}^T, \hat{Q}\,\mathbb{S}^{r_1}_+ \hat{Q}^T \right).$$

Now note that

$$\hat{Q}^T \begin{bmatrix} I_r & 0 \end{bmatrix} \mathcal{L}_1 \begin{bmatrix} I_r \\ 0 \end{bmatrix} \hat{Q} = \begin{bmatrix} \hat{Q}^T & 0 \end{bmatrix} \mathcal{L}_1 \begin{bmatrix} \hat{Q} \\ 0 \end{bmatrix} = \begin{bmatrix} I_{r_1} & 0 \end{bmatrix} Q^T \mathcal{L}_1 Q \begin{bmatrix} I_{r_1} \\ 0 \end{bmatrix},$$

so

$$\mathrm{d}\left(\begin{bmatrix} I_r & 0 \end{bmatrix} \mathcal{L}_1 \begin{bmatrix} I_r \\ 0 \end{bmatrix} \cap \hat{Q}\, \mathbb{S}^{r_1}_+ \hat{Q}^T,\, \hat{Q}\, \mathbb{S}^{r_1}_+ \hat{Q}^T\right) = \mathrm{d}\left(\begin{bmatrix} I_{r_1} & 0 \end{bmatrix} Q^T \mathcal{L}_1 Q \begin{bmatrix} I_{r_1} \\ 0 \end{bmatrix} \cap \mathbb{S}^{r_1}_+,\, \mathbb{S}^{r_1}_+\right)$$
$$\leq \mathrm{d}\left(Q^T \mathcal{L}_1 Q \cap \mathbb{S}^{n-r+r_1}_+,\, \mathbb{S}^{n-r+r_1}_+\right),$$

where the inequality follows from the induction hypothesis, because $V_1 \in \mathbb{S}^{n-r+r_1}$ defined by $V_1 := \begin{bmatrix} 0 & 0 \\ 0 & \bar{V} \end{bmatrix}$ lies in $Q^T \mathcal{L}_1 Q$. Therefore we get

$$\mathrm{d}(\hat{\mathcal{L}}_1 \cap \hat{Q}\, \mathbb{S}^{r_1}_+ \hat{Q}^T,\, \hat{Q}\, \mathbb{S}^{r_1}_+ \hat{Q}^T) \leq \mathrm{d}\left(\mathcal{L}_1 \cap Q\, \mathbb{S}^{n-r+r_1}_+ Q^T, Q\, \mathbb{S}^{n-r+r_1}_+ Q^T\right)$$

By Lemma C.1 and the definitions of $\mathcal{L}_1$ and $\hat{\mathcal{L}}_1$, we get $\mathrm{d}(\hat{\mathcal{L}} \cap \mathbb{S}^r_+,\, \mathbb{S}^r_+) \leq \mathrm{d}(\mathcal{L} \cap \mathbb{S}^n_+,\, \mathbb{S}^n_+)$.
$\square$

It is straightforward to show that a rotation does not change the degree of singularity i.e., for any linear subspace $\mathcal{L} \subseteq \mathbb{S}^n$ and any orthogonal matrix $U \in \mathbb{R}^{n \times n}$, $\mathrm{d}(U^T \mathcal{L} U \cap \mathbb{S}^n_+,\, \mathbb{S}^n_+) = \mathrm{d}(\mathcal{L} \cap \mathbb{S}^n_+,\, \mathbb{S}^n_+)$. Therefore, we can drop the assumption that $V = \begin{bmatrix} 0 & 0 \\ 0 & \bar{V} \end{bmatrix}$ with $\bar{V} \succ 0$ and allow for general V that is singular and nonzero:

Corollary C.4 *Suppose that (P) is feasible and there exist $V \in \mathbb{S}^n$ and $v \in \mathbb{R}^m$ such that*

$$0 \neq V = \mathcal{A}^* v \succeq 0, \tag{C.4}$$

and $\ker(V) = \mathcal{R}(P)$, where $P \in \mathbb{R}^{n \times r}$ has orthonormal columns. Let $\hat{C} = P^T C P \in \mathbb{S}^r$, $\hat{A}_i = P^T A_i P \in \mathbb{S}^r$ for $i \in 1 : m$, and define $\hat{\mathcal{A}} : \mathbb{S}^r \to \mathbb{R}^m$ using $\hat{A}_1, \ldots, \hat{A}_m$. Define $\hat{\mathcal{L}} := \mathrm{span}(\hat{C}, \hat{A}_1, \ldots, \hat{A}_m)$. Then $\mathrm{d}(\hat{\mathcal{L}} \cap \mathbb{S}^r_+,\, \mathbb{S}^r_+) \leq \mathrm{d}(\mathcal{L} \cap \mathbb{S}^n_+,\, \mathbb{S}^n_+)$.

Now we can easily prove Lemma 4.9.

Proof of Lemma 4.9. The feasibility of (4.13) is immediate. The minimal face $\tilde{P}\, \mathbb{S}^r_+ \tilde{P}^T$ can be obtained via facial reduction on (D). At each step of the facial reduction, the corresponding new primal (P) remains feasible and the degree of singularity of LMI defining the new primal feasible region does not increase, by Corollary C.4. In particular, the projection $\tilde{P}^T \cdot \tilde{P}$ on the primal feasible region using the minimal face of (D) does not increase the degree of singularity. $\square$

Index

References

[1] F. Alizadeh, J-P.A. Haeberly, and M.L. Overton. Complementarity and nondegeneracy in semidefinite programming. *Math. Programming*, 77:111–128, 1997. 19

[2] A. Ben-Israel, A. Ben-Tal, and S. Zlobec. Optimality in convex programming: a feasible directions approach. *Math. Programming Stud.*, (19):16–38, 1982. Optimality and stability in mathematical programming. 3

[3] J.F. Bonnans, R. Cominetti, and A. Shapiro. Sensitivity analysis of optimization problems under second order regular constraints. *Math. Oper. Res.*, 23(4):806–831, 1998. 3

[4] J.F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems*. Springer Series in Operations Research. Springer-Verlag, New York, 2000. 3

[5] J.M. Borwein and A.S. Lewis. *Convex analysis and nonlinear optimization*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 3. Springer, New York, second edition, 2006. Theory and examples. 17

[6] J.M. Borwein and H. Wolkowicz. Regularizing the abstract convex program. *J. Math. Anal. Appl.*, 83(2):495–530, 1981. 8

[7] Y.-L. Cheung. *Preprocessing and Reduction for Semidefinite Programming via Facial Reduction: Theory and Practice*. PhD thesis, University of Waterloo, 2013. 8

[8] Y.-L. Cheung, S. Schurr, and H. Wolkowicz. Preprocessing and regularization for degenerate semidefinite programs. In D.H. Bailey, H.H. Bauschke, P. Borwein, F. Garvan, M. Théra, J.D. Vanderwerff, and H. Wolkowicz, editors, *Computational and Analytical Mathematics*, volume 50 of *Springer Proceedings in Mathematics & Statistics*, pages 251–303. Springer New York, 2013. 4, 6, 8

[9] R.J. Duffin. Infinite programs. In A.W. Tucker, editor, *Linear Equalities and Related Systems*, pages 157–170. Princeton University Press, Princeton, NJ, 1956. 6

[10] M. Dür, B. Jargalsaikhan, and G. Still. The Slater condition is generic in linear conic programming, 2012. 19

[11] A.V. Fiacco. *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, volume 165 of *Mathematics in Science and Engineering*. Academic Press, 1983. 3

[12] D. Goldfarb and K. Scheinberg. On parametric semidefinite programming. In *Proceedings of the Stieltjes Workshop on High Performance Optimization Techniques*, 1996. 3

[13] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms. I*, volume 305 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1993. Fundamentals. 19

[14] Z-Q. Luo, J.F. Sturm, and S. Zhang. Duality results for conic convex programming. Technical Report Report 9719/A, April, Erasmus University Rotterdam, Econometric Institute EUR, P.O. Box 1738, 3000 DR, The Netherlands, 1997. 6, 8

[15] G. Pataki. Bad semidefinite programs: they all look the same. Technical report, Department of Operations Research, University of North Carolina, Chapel Hill, 2011. 9

[16] G. Pataki. Strong duality in conic linear programming: Facial reduction and extended duals. In D.H. Bailey, H.H. Bauschke, P. Borwein, F. Garvan, M. Théra, J.D. Vanderwerff, and H. Wolkowicz, editors, *Computational and Analytical Mathematics*, volume 50 of *Springer Proceedings in Mathematics & Statistics*, pages 613–634. Springer New York, 2013. 8, 9, 25

[17] M.V. Ramana. An exact duality theory for semidefinite programming and its complexity implications. *Math. Programming*, 77(2):129–162, 1997. 9

[18] R.T. Rockafellar. *Conjugate Duality and Optimization*. SIAM, Philadelphia, PA, 1974. Regional Conference Series in Applied Mathematics. 3, 17

[19] A. Shapiro. On uniqueness of Lagrange multipliers in optimization problems subject to cone constraints. *SIAM J. Optim.*, 7(2):508–518, 1997. 3

[20] J.F. Sturm. Error bounds for linear matrix inequalities. *SIAM J. Optim.*, 10(4):1228–1248 (electronic), 2000. 3, 5, 8, 21, 28

[21] L. Tunçel. *Polyhedral and semidefinite programming methods in combinatorial optimization*, volume 27 of *Fields Institute Monographs*. American Mathematical Society, Providence, RI, 2010. 9

[22] T. Wang and J.-S. Pang. Global error bounds for convex quadratic inequality systems. *Optimization*, 31(1):1–12, 1994. 3, 5

[23] H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors. *Handbook of semidefinite programming*. International Series in Operations Research & Management Science, 27. Kluwer Academic Publishers, Boston, MA, 2000. Theory, algorithms, and applications. 7