ORIGINAL PAPER

A strict complementarity approach to error bound and sensitivity of solution of conic programs

Lijun Ding[1] **· Madeleine Udell**[2]

Abstract

In this paper, we provide an elementary, geometric, and unified framework to analyze conic programs that we call the strict complementarity approach. This framework allows us to establish error bounds and quantify the sensitivity of the solution. The framework uses three classical ideas from convex geometry and linear algebra: linear regularity of convex sets, facial reduction, and orthogonal decomposition. We show how to use this framework to derive error bounds for linear programming, second order cone programming, and semidefinite programming.

Keywords Conic program · Strict complementarity · Error bound

1 Introduction

Given two finite dimensional Euclidean spaces $\mathbf{E}$ and $\mathbf{F}$, each equipped with an inner product denoted as $\langle \cdot, \cdot \rangle$, we consider a conic program in standard form with decision variable $x \in \mathbf{E}$:

$$
\begin{aligned}
\text{minimize} \quad & \langle c, x \rangle \\
\text{subject to} \quad & \mathcal{A}x = b, \qquad\qquad (\mathcal{P}) \\
& x \in \mathcal{K}.
\end{aligned}
$$

Here the problem data comprises a linear map $\mathcal{A} : \mathbf{E} \to \mathbf{F}$, a right hand side vector $b \in \mathbf{F}$, and a cost vector $c \in \mathbf{E}$. The cone $\mathcal{K} \subset \mathbf{E}$ is proper [1, Section 2.4.1]. The solution set and optimal value of $(\mathcal{P})$ are denoted as $\mathcal{X}_\star$ and $p_\star$ respectively. When

✉ Lijun Ding
lijunbrianding@gmail.com

Madeleine Udell
udell@stanford.edu

1 Wisconsin Institute for Discovery, University of Wisconsin - Madison, Madison, WI 53715, USA

2 Management Science and Engineering, Stanford University, Stanford, CA 94305, USA

$\mathcal{K}$ is the nonnegative orthant $\mathbf{R}_+^n$, the second order cone SOC_n, or the set of positive semidefinite matrices $\mathbf{S}_+^n$, we call the corresponding problem ($\mathcal{P}$) an linear programming (LP), second order cone programming (SOCP), or semidefinite programming (SDP), respectively.

In this paper, we provide an elementary framework based on *strict complementarity* (see Sect. 2.1) to establish *error bounds* and quantify the *sensitivity of the solution* of Problem ($\mathcal{P}$). In the following, $\|\cdot\|_2$ denotes the Euclidean norm induced by the inner product, while $\|\cdot\|$ is a generic norm that will be specified when we instantiate these bounds later in the paper.

- **Error bound:** Given $x \in \mathbf{E}$, define three error metrics: suboptimality $\epsilon_{opt}(x) := \langle c, x \rangle - p_\star$, linear infeasibility $\mathcal{A}x - b$, and conic infeasibility as $x_- = x - x_+$, where the conic part $x_+ := \mathcal{P}_\mathcal{K}(x)$.[1] These errors are easy to measure, while the distance of a given point x to the solution is not. This paper shows how to establish an error bound for ($\mathcal{P}$) that bounds the distance to the solution in terms of these measurable error metrics, for some constants $c_i, i = 1, 2, 3$ and exponent $p > 0$ independent of x:

$$\mathbf{dist}(x, \mathcal{X}_\star)^p \le c_1 |\epsilon_{opt}(x)| + c_2 \|\mathcal{A}x - b\| + c_3 \|x_-\|, \tag{ERB}$$

  where $\mathbf{dist}(x, \mathcal{X}_\star) := \inf_{x_\star \in \mathcal{X}_\star} \|x - x_\star\|_2$ is the distance to $x_\star$.
- **Sensitivity of the solution:** We often wish to understand how the solution of the problem changes with perturbations of the problem data. Given new problem data $(c', \mathcal{A}', b') \in \mathbf{E} \times \mathbf{L}(\mathbf{E}, \mathbf{F}) \times \mathbf{F}$ for Problem ($\mathcal{P}$), where $\mathbf{L}(\mathbf{E}, \mathbf{F})$ is the set of linear maps from $\mathbf{E}$ to $\mathbf{F}$, Problem ($\mathcal{P}$) admits a new optimal solution set $\mathcal{X}'_\star$. This paper shows how to quantify the sensitivity of the solution, for some constants $c'_i, i = 1, 2, 3$ and exponent $p' > 0$, via the inequality

$$\mathbf{dist}^{p'}(\mathcal{X}_\star, \mathcal{X}'_\star) \le c'_1 \|c - c'\| + c'_2 \|\mathcal{A} - \mathcal{A}'\| + c'_3 \|b - b'\|, \tag{SSB}$$

  where $\mathbf{dist}(\mathcal{X}_\star, \mathcal{X}'_\star) := \inf_{x_\star \in \mathcal{X}_\star, x' \in \mathcal{X}'_\star} \|x_\star - x'_\star\|_2$ is the distance between solution sets. In fact, once an error bound of the form (ERB) is available, we can prove an inequality of this form by bounding the error metrics of the new solution $x'_\star \in \mathcal{X}'_\star$ with respect to the original problem data $(c, \mathcal{A}, b)$ in terms of the perturbation $(c - c', \mathcal{A} - \mathcal{A}', b - b')$.

Importance of the error bound and sensitivity of solution. The error bound and sensitivity of the solution can be regarded as condition numbers for Problem ($\mathcal{P}$). They guarantee that the output of iterative algorithms to solve ($\mathcal{P}$) is still useful despite optimization error (of the algorithm) and measurement error (of the problem data) [2, 3]. The error bound is also vital in proving faster convergence for first order algorithms [4–7]. Hence a huge body of work has devoted to establish error bounds and sensitivity of solutions [2, 4, 5, 8–10].

---

[1] The orthogonal projector $\mathcal{P}_\mathcal{K}$ is defined as $\mathcal{P}_\mathcal{K}(x) = \mathrm{argmin}_{y \in \mathcal{K}} \|y - x\|_2$.

Our Contribution. In this paper, we use the notion of *strict complementarity* (defined in Sect. 2.1) to provide an elementary, geometric, and unified framework, described in detail in Sect. 2, to establish bounds of the form (ERB) and (SSB) for the conic program ($\mathcal{P}$). Specifically, in Sects. 3 and 4, we show how to construct a bound with exponents $p = p' = 1$ for LP and $p = p' = 2$ for SOCP and SDP, under strict complementarity, and provide a way to obtain explicit estimates of $c_i, i = 1, 2, 3$ in terms of the primal and dual solutions and problem data when the primal solution is unique. Table 1 summarizes our results.

The main contribution of this paper is a new and simple framework for proving bounds of this form. As discussed in Sect. 5, many particular bounds that we present here have been discovered before. On the other hand, we believe that some of the bounds are new: in particular, bounds on the sensitivity of the solution that pertain when the primal or dual solution are not unique.

Paper organization. The rest of the paper is organized as follows. In Sect. 2.1, we discuss two important analytical conditions assumed throughout this paper: strong duality and dual strict complementarity. In Sect. 2.2, we describe the basic framework of the strict complementarity approach: linear regularity of convex sets, facial reduction, and extension via orthogonal decomposition. In Sect. 3, we apply the framework to specific examples, LP, SOCP, and SDP, to establish error bounds. We next demonstrate how to use the error bound established to characterize the sensitivity of solutions in Sect. 4 by bounding the error measures of the new solution $x'_\star \in \mathcal{X}'_\star$ in terms of the perturbation $(c - c', \mathcal{A} - \mathcal{A}', b - b')$. Finally, we discuss previous results regarding (ERB) and (SSB), how this work relates to them, and potential new directions.

Notation. We use $\mathbf{E}, \mathbf{F}, \mathbf{E}', \mathbf{F}'$ to represent generic finite dimensional Euclidean spaces. For a set $\mathcal{C}$ in $\mathbf{E}$, we denote its interior, boundary, affine hull, and relative interior as $\mathbf{int}(\mathcal{C})$, $\partial\mathcal{C}$, $\mathrm{affine}(\mathcal{C})$, and $\mathbf{relint}(\mathcal{C})$ respectively. We equip $\mathbf{R}^n$ with the dot inner product, and $\mathbf{S}^n$ and $\mathbf{R}^{n \times n}$ with the trace inner product. The distance to a set $\mathcal{C}$ is defined as $\mathbf{dist}(x, \mathcal{C}) = \inf_{z \in \mathcal{C}} \|x - z\|_2$. We write $\|\cdot\|$ for an arbitrary norm, and $\|\cdot\|_2$ for the ℓ_2 norm induced by the underlying inner product. For matrices, the operator norm (maximum singular value), Frobenius norm, and nuclear norm (sum of singular values) are denoted as $\|\cdot\|_{op}$, $\|\cdot\|_F$, and $\|\cdot\|_*$ respectively. For a linear map $\mathcal{B} : \mathbf{E} \to \mathbf{F}$ and a linear space $\mathcal{V} \subset \mathbf{E}$, we write the restriction of $\mathcal{B}$ to $\mathcal{V}$ as $\mathcal{B}_\mathcal{V}$. We define the largest and smallest singular value of $\mathcal{B}$ as $\sigma_{\max}(\mathcal{B}) := \max_{\|x\|_2=1} \|\mathcal{B}(x)\|_2$ and $\sigma_{\min}(\mathcal{B}) := \min_{\|x\|_2=1} \|\mathcal{B}(x)\|_2$ respectively.

2 The strict complementary slackness approach

In Sect. 2.1, we introduce two important structural conditions, strong duality and dual strict complementarity, that are essential to our approach. Next in Sect. 2.2, we describe the main ingredients of the strict complementary slackness approach: linear regularity (Sect. 2.2.1), facial reduction (Sect. 2.2.2), and orthogonal decomposition (Sect. 2.2.3). Our main result, Theorem 1, is in Sect. 2.2.3.

2.1 Analytical conditions

Here we define two conditions that are essential to our framework: strong duality and dual strict complementarity. To start, let us recall the dual problem of $(\mathcal{P})$ is

$$\begin{aligned} \text{maximize} \quad & \langle b, y \rangle \\ \text{subject to} \quad & c - \mathcal{A}^* y \in \mathcal{K}^*. \end{aligned} \tag{$\mathcal{D}$}$$

The vector $y \in \mathbf{F}$ is the decision variable, the linear map $\mathcal{A}^*$ is the adjoint of the linear map $\mathcal{A}$, and the cone $\mathcal{K}^*$ is the dual cone of $\mathcal{K}$, i.e., $\mathcal{K}^* := \{s \in \mathbf{E} \mid \langle s, x \rangle \geq 0, \forall x \in \mathcal{K}\}$. Let us introduce strong duality first.

Definition 1 (Strong duality) The primal and dual problems $(\mathcal{P})$ and $(\mathcal{D})$ satisfy strong duality (SD) if the primal and dual solution sets $\mathcal{X}_\star, \mathcal{Y}_\star$ are nonempty, $\mathcal{X}_\star$ is compact, and there exists a primal and dual solution pair $(x_\star, y_\star) \in \mathcal{X}_\star \times \mathcal{Y}_\star$ such that

$$\langle c, x_\star \rangle = \langle b, y_\star \rangle = \langle \mathcal{A} x_\star, y_\star \rangle. \tag{SD}$$

Equivalently, define the slack vector $s_\star = c - \mathcal{A}^* y_\star$ to rewrite the equality (SD) as

$$0 = \langle c - \mathcal{A}^*(y_\star), x_\star \rangle = \langle s_\star, x_\star \rangle.$$

Note that we require the existence of primal and dual optimal solutions instead of just equality of optimal values. Strong duality in the stated form is ensured by *primal and dual Slater's condition*: there is a primal and dual feasible pair (x, y) with $(x, c - \mathcal{A}^* y) \in \mathbf{int}(\mathcal{K}) \times \mathbf{int}(\mathcal{K}^*)$.

Next we state the second condition: *dual strict complementarity*. This condition is the key to established error bounds for a variety of optimization problems [4, 5].

Definition 2 (Dual strict complementarity (DSC)) Given a solution pair $(x_\star, y_\star) \in \mathcal{X}_\star \times \mathcal{Y}_\star$, define the *complementary face* $\mathcal{F}_{s_\star} := \{x \mid \langle x, s_\star \rangle = 0, \ s_\star = c - \mathcal{A}^* y_\star\} \cap \mathcal{K}$. The solution pair $(x_\star, y_\star)$ satisfies dual strict complementarity if

$$x_\star \in \mathrm{rel}(\mathcal{F}_{s_\star}). \tag{DSC}$$

If $(\mathcal{P})$ and $(\mathcal{D})$ has one such pair, we say $(\mathcal{P})$ and $(\mathcal{D})$ (or simply $(\mathcal{P})$) admits dual strict complementarity.

Let us now unpack the definition of $\mathcal{F}_{s_\star}$ and dual strict complementarity. Also see Fig. 1 for a graphical illustration of the condition.

Understanding the complementary face $\mathcal{F}_{s_\star}$. To understand the name complementary face, let us introduce the *complementary hyperplane*, the hyperplane $\mathcal{H}_{s_\star} = \{x \mid \langle x, s_\star \rangle = 0\}$ orthogonal to the slack vector $s_\star = c - \mathcal{A}^* y_\star$. The complementary face $\mathcal{F}_{s_\star}$ is simply the intersection of $\mathcal{H}_{s_\star}$ and the cone $\mathcal{K}$. The intersection is nonempty due to strong duality. We can see that $\mathcal{F}_{s_\star}$ is indeed a face because its intersection with $\mathcal{K}$ is nonempty (it contains $x_\star$), and every $x \in \mathcal{K}$ lies on the same
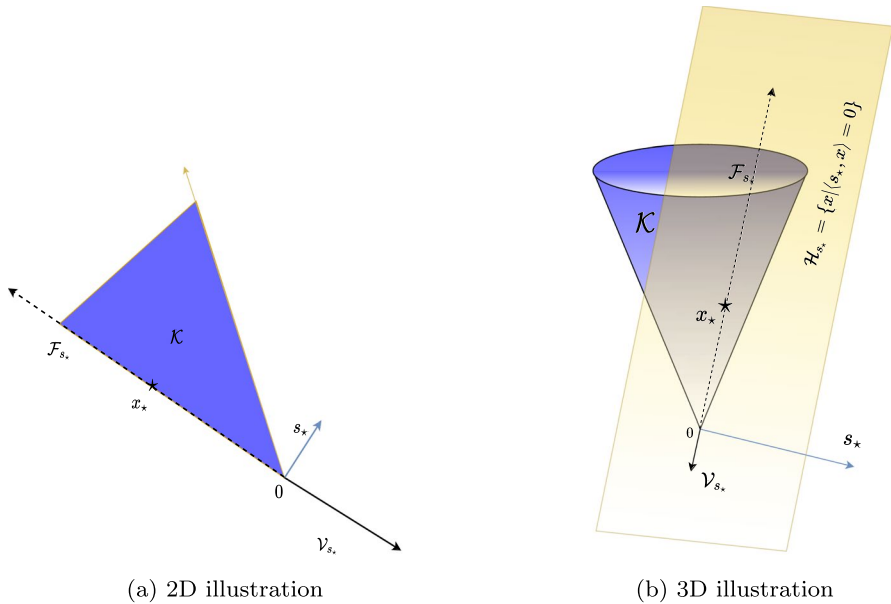
(a) 2D illustration   (b) 3D illustration

Fig. 1 Strict Complementarity: For both plots, the indigo cone is the cone $\mathcal{K}$; the slack vector $s_\star$ is the blue ray; the complementary face $\mathcal{F}_{s_\star}$ is the dashed black ray; and the complementary space $\mathcal{V}_{s_\star}$ is the black line (both solid and dashed parts). In the 2D case, the complementary hyperplane $\mathcal{H}_{s_\star}$ and $\mathcal{V}_{s_\star}$ coincide. In the 3D case, the complementary hyperplane $\mathcal{H}_{s_\star}$ is the yellow plane, which is tangent to the purple cone $\mathcal{K}$ at the point $x_\star$ and is orthogonal to $s_\star$

side of the hyperplane $\mathcal{H}_{s_\star}$, $\langle x, s_\star \rangle \geq 0$, as $s_\star \in \mathcal{K}^*$. In particular, we see the face $\mathcal{F}_{s_\star}$ is exposed.

Dual strict complementarity (DSC) and Slater's condition. To better understand dual strict complementarity, define the *complementary space* as the affine hull of $\mathcal{F}_{s_\star}$, $\mathcal{V}_{s_\star} := \text{affine}(\mathcal{F}_{s_\star})$. The complementary face $\mathcal{F}_{s_\star} = \mathcal{H}_{s_\star} \cap \mathcal{K}$ is a cone, so the complementary space $\mathcal{V}_{s_\star}$ is a linear subspace. Imagine modifying problem ($\mathcal{P}$) by replacing the cone $\mathcal{K}$ by $\mathcal{F}_{s_\star}$ in problem ($\mathcal{P}$) and restricting the decision variable x to the subspace $\mathcal{V}_{s_\star}$. Note that $x_\star$ is still a solution to this problem, so this procedure is related to facial reduction: the modified problem restricts x to a face of the original cone $\mathcal{K}$. DSC means that there is a primal x in the interior of the cone $\mathcal{F}_{s_\star} \subset \mathcal{V}_{s_\star}$, where the interior is taken w.r.t. the subspace $\mathcal{V}_{s_\star}$. Hence DSC is equivalent to the usual Slater's condition for the modified problem.

Primal strict complementarity and strict complementarity. Given dual strict complementarity (DSC), a natural way to define primal strict complementarity (PSC) is to reverse the role of $s_\star$ and $x_\star$ in the definition of DSC. Precisely, PSC means that there exists $(x_\star, y_\star)$ with $c - \mathcal{A}^* y_\star = s_\star \in \text{rel}(\{s \mid \langle x_\star, s \rangle = 0 \text{ and } s \in \mathcal{K}^*\})$. Primal and dual strict complementarity are not always equivalent unless the cone $\mathcal{K}$ is exposed.[2] Happily, all the symmetric cones are exposed, including $\mathcal{K} = \mathbf{R}_+^n$, SOC_n,

---

[2] For a discussion on primal and dual strict complementarity, see [11, Remark 4.10].

and $\mathbf{S}_+^n$. [11] shows that DSC and PSC actually hold "generically"[3] for general conic programs. It is worth noting that the standard notion of strict complementarity (SC) for SDP [12, Definition 4] and LP, both defined algebraically, are equivalent to the geometric notion of DSC here.[4] SC always holds for LP [13], holds "generically" for SDP as shown in [12], and even holds for some structured instances of SDP [3]. Due to the equivalence, DSC also holds under the same conditions for LP and SDP.

2.2 Defining the strict complementary slackness approach

In this section, we explain how to use the two assumptions in Sect. 2.1 to establish a framework to prove error bounds of the form (ERB) and sensitivity bounds of the form (SSB). As we explained in the introduction, an error bound (ERB) can be used to derive a sensitivity bound (SSB). Hence, we focus on proving an error bound first. Our main theorem, Theorem 1 in Sect. 2.2.3, reduces the task to bounding the quantity $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+) \right\|_2$.[5] We explain how to further bound this quantity in Sect. 3.

2.2.1 From optimality to feasibility: linear regularity of convex sets

Our first step is to identify problem ($\mathcal{P}$) with the feasibility problem of finding x such that

$$\langle c, x \rangle = p_\star, \quad \mathcal{A}(x) = b, \quad \text{and} \quad x \in \mathcal{K}. \tag{1}$$

This transformation activates the following geometric result, called linear regulairty regularity of convex sets [14, Theorem 4.6], [15, Theorem 2.1], a classical result on error bounds for feasibility systems. This result states that the distance to the intersection of two sets is bounded by the sum of distances to the two sets.

Lemma 1 (Linear regularity) *Suppose the $C \subset \mathbf{E}$ is an affine space with $C = \{x \mid \mathcal{B}x = d\}$, where $\mathcal{B} : \mathbf{E} \to \mathbf{F}$ is a linear map, and $D \subset \mathbf{E}$ is a closed convex cone. If $C \cap \mathbf{int}(D) \neq \emptyset$ and $C \cap D$ is compact, then there are some $\gamma, \gamma' > 0$ such that for all $x \in \mathbf{E}$,*

$$\mathbf{dist}(x, C \cap D) \le \gamma \| \mathcal{B}x - d \|_2 + \gamma' \mathbf{dist}(x, D). \tag{2}$$

Now it is tempting to set $C = \{x \mid \langle c, x \rangle = p_\star, \mathcal{A}(x) = b\}$ and $D = \mathcal{K}$, and conclude (ERB) holds with exponent $p = 1$, since $C \cap D = \mathcal{X}_\star$. The catch is that for most problems of interest, the optimal solution $x_\star \in \partial\mathcal{K}$ lies on the boundary of

---

[3] Roughly speaking, this condition holds except on a measure 0 set of problems parameterized by $\mathcal{A}, b, c$, conditioning on the existence of a primal dual solution pair. We refer the reader to the references for more details.

[4] The definition of SC for LP and SDP, and the proof of the equivalence can be found in Sect. 1.

[5] Here $\mathcal{V}_{s_\star}^\perp$ is orthogonal complementary space of $\mathcal{V}_{s_\star} = \text{affine}(\mathcal{F}_{s_\star})$, and $\mathcal{P}_{\mathcal{V}_{s_\star}^\perp}$ is the corresponding projection. Recall the conic part x_+ is $x_+ = \mathcal{P}_{\mathcal{K}}(x)$.

the cone $\mathcal{K}$, and the condition $\textbf{int}(D) \cap C = \emptyset$ *does not* hold! Indeed, unless $c = 0$, which makes $(\mathcal{P})$ a feasibility problem, we always have $x_\star \in \partial \mathcal{K}$.

2.2.2 Facial reduction

We may still use Lemma 1 to establish an error bound. The key is to use the facial reduction idea mentioned earlier. Recall the condition required is $C \cap \textbf{int}(D) = \{x \mid \langle c, x \rangle = p_\star, \mathcal{A}(x) = b\} \cap \textbf{int}(\mathcal{K}) \neq \emptyset$, which does not hold for $(\mathcal{P})$ with nonzero c. The problem is that the cone $\mathcal{K}$ lies in the large space $\mathbf{E}$, so its interior (with respect to $\mathbf{E}$) does not contain $x_\star$. Instead, consider restricting the variable x to the complementary space $\mathcal{V}_{s_\star} = \text{affine}(\mathcal{F}_{s_\star})$ and replacing the cone $\mathcal{K}$ by the complementary face $\mathcal{F}_{s_\star}$. The interior of the cone $\mathcal{F}_{s_\star}$ with respect to the space $\mathcal{V}_{s_\star}$ does contain $x_\star$ under strict complementarity, so we may activate Lemma 1.

This modification enables an error bound for $x \in \mathcal{V}_{s_\star}$ as stated in Lemma 2 below. We also provide a more concrete estimate of the constants γ, γ' when $\mathcal{X}_\star$ is a singleton.

Lemma 2 *Suppose strong duality and dual strict complementarity hold. Then there are constants γ, γ' such that for any $x \in \mathcal{V}_{s_\star}$:*

$$\textbf{dist}(x, \mathcal{X}_\star) \leq \gamma \|\mathcal{A}(x) - b\|_2 + \gamma' \textbf{dist}(x, \mathcal{F}_{s_\star}). \tag{3}$$

Moreover, if $\mathcal{X}_\star$ is a singleton, then we may take $\gamma' = 0$ and $\gamma = \dfrac{1}{\sigma_{\min}(\mathcal{A}_{\mathcal{V}_{s_\star}})}$. Here, the linear map $\mathcal{A}_{\mathcal{V}_{s_\star}}$ is $\mathcal{A}$ restricted to $\mathcal{V}_{s_\star}$ and $\sigma_{\min}(\mathcal{A}_{\mathcal{V}_{s_\star}})$ is the smallest singular value of $\mathcal{A}_{\mathcal{V}_{s_\star}}$.

Proof The inequality 3 is immediate by using Lemma 1 with $C = \{x \in \mathcal{V}_{s_\star} \mid \mathcal{A}(x) = b\} \subset \mathcal{V}_{s_\star}$, $D = \mathcal{F}_{s_\star}$, and $\mathbf{E} = \mathcal{V}_{s_\star}$.

Indeed, this choice gives $C \cap D = \mathcal{X}_\star$. The condition $C \cap \textbf{int}(D) \neq \emptyset$, where the interior is taken with respect to the space $\mathcal{V}_{s_\star}$, is exactly (DSC): $\exists x_\star \in \mathcal{X}_\star$ such that $x_\star \in \textbf{relint}(\mathcal{F}_{s_\star})$.

Now we show that $\gamma' = 0$ and $\gamma = \frac{1}{\sigma_{\min}(\mathcal{A}_{\mathcal{V}_{s_\star}})}$ when $\mathcal{X}_\star$ is a singleton. First assume $\mathcal{A}_{\mathcal{V}_{s_\star}}$ has a trivial nullspace, so $x_\star$ is the only solution in $\mathcal{V}_{s_\star}$ to $\mathcal{A}_{\mathcal{V}_{s_\star}}(x) = b$. Hence $\sigma_{\min}(\mathcal{A}_{\mathcal{V}_{s_\star}}) > 0$ and so for any $x \in \mathcal{V}_{s_\star}$, $\|x - x_\star\|_2 \leq \frac{1}{\sigma_{\min}(\mathcal{A}_{\mathcal{V}_{s_\star}})} \|\mathcal{A}_{\mathcal{V}_{s_\star}}(x - x_\star)\|_2 = \frac{1}{\sigma_{\min}(\mathcal{A}_{\mathcal{V}_{s_\star}})} \|\mathcal{A}(x) - b\|_2$. Finally, we show by contradiction that the nullspace of $\mathcal{A}_{\mathcal{V}_{s_\star}}$ is trivial whenever $\mathcal{X}_\star$ is a singleton. If the nullspace is not trivial, then there is some $x' \in \mathcal{V}_{s_\star}$ such that $\mathcal{A}_{\mathcal{V}_{s_\star}}(x') = 0$. Hence $x_\star + \alpha x'$ for some small enough α is still optimal, as $x_\star \in \textbf{relint}(\mathcal{F}_{s_\star})$, which contradicts our hypothesis that the solution set $\mathcal{X}_\star$ is a singleton. $\square$

This choice of the face $\mathcal{F}_{s_\star}$ and the corresponding linear space $\mathcal{V}_{s_\star}$ correspond to the idea of facial reduction [16, 17]. Facial reduction is a conceptual and numerical technique designed to handle conic feasibility problems for which constraint qualifications (such as Slater's condition) fail. Note that such failure is the interesting case

for a feasibility system (1) when the optimal solution $x_\star \in \partial \mathcal{K}$. Indeed, our choice of face can be considered as one step of the facial reduction procedure.

2.2.3 Extension to the whole space: orthogonal decomposition

In this section, we derive our main result, Theorem 1, by extending the previous result to the whole space using the orthogonal decomposition $\mathbf{E} = \mathcal{V}_{s_\star} \oplus \mathcal{V}_{s_\star}^\perp$ with $\mathcal{V}_{s_\star} \perp \mathcal{V}_{s_\star}^\perp$.

Theorem 1 *Suppose strong duality and dual strict complementarity hold. Then for some constants γ, γ' described in Lemma 2 and for all $x \in \mathbf{E}$, we have*

$$
\begin{aligned}
\mathbf{dist}(x, \mathcal{X}_\star) \leq &(1 + \gamma \sigma_{\max}(\mathcal{A})) \left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+) \right\|_2 + \gamma \|\mathcal{A}(x) - b\|_2 \\
&+ (1 + \gamma \sigma_{\max}(\mathcal{A})) \left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_-) \right\|_2 + \gamma' \left\| \mathcal{P}_{\mathcal{V}_{s_\star}}(x_-) \right\|_2 \\
&+ \gamma' \mathbf{dist}(\mathcal{P}_{\mathcal{V}_{s_\star}}(x_+), \mathcal{F}_{s_\star}),
\end{aligned} \tag{4}
$$

where $\mathcal{P}_{\mathcal{V}_{s_\star}}$ and $\mathcal{P}_{\mathcal{V}_{s_\star}^\perp}$ are orthogonal projections to $\mathcal{V}_{s_\star}$ and $\mathcal{V}_{s_\star}^\perp$ respectively. The terms $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}}(x_-) \right\|_2$ and $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_-) \right\|_2$ can themselves be bounded by $\|x_-\|_2$.

Proof Recall Lemma 2 establishes an error bound for only those $x \in \mathcal{V}_{s_\star}$. Using the orthogonal decomposition proposed, for any $x \in \mathbf{E}$,

$$
x = \mathcal{P}_{\mathcal{V}_{s_\star}}(x) + \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x).
$$

This decomposition immediately gives

$$
\mathbf{dist}(x, \mathcal{X}_\star) \leq \left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x) \right\|_2 + \mathbf{dist}(\mathcal{P}_{\mathcal{V}_{s_\star}}(x), \mathcal{X}_\star). \tag{5}
$$

The second term $\mathbf{dist}(\mathcal{P}_{\mathcal{V}_{s_\star}}(x), \mathcal{X}_\star)$ can be bounded using Lemma 2:

$$
\mathbf{dist}(\mathcal{P}_{\mathcal{V}_{s_\star}}(x), \mathcal{X}_\star) \leq \gamma \left\| \mathcal{A}(\mathcal{P}_{\mathcal{V}_{s_\star}}(x)) - b \right\|_2 + \gamma' \mathbf{dist}(\mathcal{P}_{\mathcal{V}_{s_\star}}(x), \mathcal{F}_{s_\star}). \tag{6}
$$

To translate the above bound to linear infeasibility $\mathcal{A}(x) - b$ and conic infeasibility x_-, we note that $x = \mathcal{P}_{\mathcal{V}_{s_\star}}(x) + \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x)$ and $x = x_+ + x_-$[6]. With these two decompositions, and tad more algebra, we arrive at the theorem. $\square$

To go further, we need to bound the following two terms in terms of $\epsilon_{opt}(x)$, $\mathcal{A}(x) - b$, and x_-:

---

[6] Recall $x_+ = \mathcal{P}_{\mathcal{K}}(x)$ and $x = x_+ + x_-$.

Table 1 This table presents the bound $f(\epsilon(x_+), \|x\|)$ for $\left\|\mathcal{P}_{\mathcal{V}_{s_\star}}(x_+)\right\|_2$, the power p in (ERB) and (SSB), and an estimate of c_1 for *feasible x* for different cases based on $s_\star$ and $\mathcal{K}$

Conic program	$s_\star = 0$	$s_\star \in \mathbf{int}(\mathcal{K}^*)$ LP	SOCP	SDP
Violation of complementarity $f(\epsilon(x_+), \|x\|)$	0	$c_\star \epsilon(x_+)$ $\frac{\epsilon(x_+)}{s_{\min>0}}$	$\sqrt{\frac{2\sqrt{2}\|x\|_2 \epsilon(x_+)}{\|s_\star\|_2}}$	$\frac{\epsilon(X_+)}{T} + \sqrt{\frac{2\epsilon(X_+)\|X\|_{op}}{T}}$
Exponents p, p' in (ERB) and (SSB)	1	1	1 2	2
Constant c_1 for $\epsilon_{opt}(x)$	0	$\kappa c_\star$ $\frac{\kappa}{s_{\min>0}}$	$2\sqrt{2}\kappa^2 \frac{B}{\|s_\star\|_2}$	$\kappa^2 \frac{8B}{T}$

For $s_\star \in \mathbf{int}(\mathcal{K}*)$, the quantity $c_\star = \sup_{x \in \mathcal{K}} \frac{1}{\langle s_\star, \frac{x}{\|x\|}\rangle} < \infty$. For LP, the quantity $s_{\min>0}$ is the smallest nonzero element of $s_\star$. For SDP, the quantity T is the smallest nonzero eigenvalue of $S_\star$. The condition number $\kappa = 1 + \gamma \sigma_{\max}(\mathcal{A})$ and is $1 + \frac{\sigma_{\max}(\mathcal{A})}{\sigma_{\min}(\mathcal{A}_{\mathcal{V}_{s_\star}})}$ when $\mathcal{X}_\star$ is a singleton. We assume $\|x\|_2 \leq B$ for SOCP and $\|X\|_{op} \leq B$ for SDP. We also assume $\frac{\epsilon(X_+)}{T} \leq \sqrt{2\frac{\epsilon(X_+)B}{T}}$ for SDP

1. The distance to the space $\mathcal{V}_{s_\star}$: $\mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+)$.
2. The term $\mathbf{dist}(\mathcal{P}_{\mathcal{V}_{s_\star}}(x_+), \mathcal{F}_{s_\star})$.

In Sect. 3, we show how to bound both terms for the special cases of LP, SOCP, SDP, and more general conic programs ($\mathcal{P}$) where $\mathcal{K}$ is a finite product of LP, SOCP, or SDP cones. A quick summary of results can be found in Table 1.

As we shall see in Sect. 3, the term $\mathbf{dist}(\mathcal{P}_{\mathcal{V}_{s_\star}}(x_+), \mathcal{F}_{s_\star})$ is usually zero. Thus the major challenge is bounding $\left\|\mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+)\right\|_2$[7]. Note that under this condition, for feasible x of ($\mathcal{P}$), the bound (4) reduces to

$$\mathbf{dist}(x, \mathcal{X}_\star) \leq (1 + \gamma \sigma_{\max}(\mathcal{A})) \left\|\mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x)\right\|_2. \tag{7}$$

If the solution set $\mathcal{X}_\star$ is a singleton, then from Lemma 2, we know $\gamma = \frac{1}{\sigma_{\min}(\mathcal{A}_{\mathcal{V}_{s_\star}})}$, and we encounter a condition number like quantity $\frac{\sigma_{\max}(\mathcal{A})}{\sigma_{\min}(\mathcal{A}_{\mathcal{V}_{s_\star}})}$ in (7). Depending on applications, the condition number may scale with the problem dimension but the bound is still tight as the following example shows.

Example 1 Consider an SDP with $C = -\mathbf{1}\mathbf{1}^\top$, where $\mathbf{1} \in \mathbf{R}^n$ is the all one vector, $\mathcal{A} = \mathbf{diag}(\cdot)$, and $b = \mathbf{1}$. This is a simplification of the SDP for $\mathbb{Z}_2$ synchronization [3, 18]. For this SDP, it is easily verified that the unique optimal solution is $\mathbf{1}\mathbf{1}^\top$ and dual strict complementarity holds with dual optimal slack $S_\star = -\mathbf{1}\mathbf{1}^\top + nI$. The condition number like quantity $\frac{\sigma_{\max}(\mathcal{A})}{\sigma_{\min}(\mathcal{A}_{\mathcal{V}_{s_\star}})}$ in this case is $\sqrt{n}$ which does scale with the

---

[7] Bounding this term is in some sense necessary in establishing an error bound. See more discussion in Appendix C.

dimension n. However, if in (7) we let $x = I_n$, the identity matrix which is feasible, then the LHS and RHS of (7) are $\sqrt{n^2 - n}$ and $\sqrt{n - 1} + \sqrt{n^2 - n}$ respectively. Thus the bound is actually tight for large n.

3 Application: error bounds

In this section, we show how to use the framework established in Sect. 2 to analyze conic programs ($\mathcal{P}$) over the nonnegative orthant, the second order cone, the set of positive semidefinite matrices, or a finite product of these cones. Our analysis has two main steps:

1. Identify and write out the complementary face $\mathcal{F}_{s_\star}$ and $\mathcal{V}_{s_\star}$.
2. Bound the term $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+) \right\|_2$ via a function $f(\langle s_\star, x_+ \rangle, \|x\|)$, called the violation of complementarity, using the explicit structure of $\mathcal{V}_{s_\star}$.

We summarize the findings of this section as the following lemma and corollary. Refer to Table 1 for a quick summary of the results.

Lemma 3 *Define the complementarity error $\epsilon(x) = \langle s_\star, x \rangle$. Suppose strong duality holds. The quantity $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+) \right\|_2$ can be bounded by several different functions $f(\langle s_\star, x_+ \rangle, \|x\|)$, which we call the violation of complementarity, depending on the slack vector $s_\star$ and the cone $\mathcal{K}$. The first two trivial cases are the following.*

1. *If $s_\star = 0$, then $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+) \right\| = 0 =: f_0(\epsilon(x_+))$.*
2. *If $s_\star \in \mathbf{int}(\mathcal{K}^*)$, then $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+) \right\|_2 = \|x_+\|_2 \leq c_\star \epsilon(x_+) =: f_{\mathbf{int}}(\epsilon(x_+))$, where $c_\star = \sup_{x \in \mathcal{K}} \frac{1}{\langle s_\star, \frac{x}{\|x\|} \rangle} < \infty$.*

Moreover, for the nontrivial case $s_\star \in \partial \mathcal{K}^* / \{0\}$, we have the following bounds:

3. *$\mathcal{K} = \mathbf{R}_+^n$: $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+) \right\|_2 \leq \frac{1}{s_{\min>0}} \epsilon(x_+) =: f_{\mathbf{R}_+^n}(\epsilon(x_+))$, where $s_{\min>0}$ is the smallest nonzero element of $s_\star$.*
4. *$\mathcal{K} = \mathrm{SOC}_n$: $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+) \right\|_2 \leq \sqrt{2\sqrt{2} \frac{\|x\|_2 \epsilon(x_+)}{\|s_\star\|_2}} =: f_{\mathrm{SOC}_n}(\epsilon(x_+), \|x\|_2)$.*
5. *$\mathcal{K} = \mathbf{S}_+^n$: $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(X_+) \right\|_F \leq \frac{\epsilon(X_+)}{T} + \sqrt{2\frac{\epsilon(X_+)}{T} \|X\|_{op}} =: f_{\mathbf{S}_+^n}(\epsilon(X_+), \|X\|_{op})$. Here T is the smallest nonzero eigenvalue of $S_\star$.*

Proof Let us first consider the two trivial cases (1) $s_\star = 0$ and (2) $s_\star \in \mathbf{int}(\mathcal{K}^*)$. These cases are excluded whenever c and b are both nonzero. In the first case, $\mathcal{V}_{s_\star}^\perp = \{0\}$, and we simply have $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+) \right\| = 0$. In the second case, we have $\mathcal{V}_{s_\star}^\perp = \mathbf{E}$,

and $\left\|\mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+)\right\| = \|x_+\| \le c_\star \langle s_\star, x_+ \rangle$, where $c_\star = \sup_{x \in \mathcal{K}} \frac{1}{\langle s_\star, \frac{x}{\|x\|}\rangle} < \infty$. We defer the proof for the other cases to Sect. 3.1, 3.2, and Sect. 3.3 for LP, SOCP, and SDP respectively.

Combining Lemma 3, and Theorem 1, we reach the following corollary. The quantity $\mathbf{dist}(\mathcal{P}_{\mathcal{V}_{s_\star}}(x_+), \mathcal{F}_{s_\star})$ can be verified to be zero for the two trivial cases by noting (i) the complementary space $\mathcal{V}_{s_\star} = \mathbf{E}$ and $\mathcal{F}_{s_\star} = \mathcal{K}$ for the case $s_\star = 0$, and (ii) the complementary space $\mathcal{V}_{s_\star} = \{0\}$ and $\mathcal{F}_{s_\star}$ is a closed cone for the case $s_\star \in \mathbf{int}(\mathcal{K}^*)$. It is also zero for other three cases as shown in Sects. 3.1–3.3.

Corollary 1 *Suppose strong duality and dual strict complementarity hold, and one of the five cases in Lemma 3 pertains. Then there exists constants γ, γ' so that for all $x \in \mathbf{E}$,*

$$\mathbf{dist}(x, \mathcal{X}_\star) \le \kappa f(\epsilon(x_+), \|x\|) + \gamma \|\mathcal{A}(x) - b\|_2$$
$$+ \kappa \left\|\mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_-)\right\|_2 + \gamma' \left\|\mathcal{P}_{\mathcal{V}_{s_\star}}(x_-)\right\|_2, \tag{8}$$

where the condition number $\kappa = 1 + \gamma \sigma_{\max}(\mathcal{A})$. In particular, when $\mathcal{X}_\star$ is a singleton, then $\gamma' = 0$ and $\gamma = \frac{1}{\sigma_{\min}(\mathcal{A}_{\mathcal{V}_{s_\star}})}$. Here the formula for $f(\epsilon(x), \|x\|)$ can be found in Lemma 3 for each of the different cases, and we can further decompose the complementarity error $\epsilon(x_+) = \epsilon_{opt}(x) + \langle y_\star, b - \mathcal{A}(x)\rangle - \langle s_\star, x_-\rangle$ using $x = x_+ + x_-$.

A few remarks regarding the lemma and the corollary are in order.

Remark 1 (Global and local error bound) Note the formula for the violation of complementarity $f(\epsilon(x), \|x\|)$ uses $\|x\|_2$ for SOCP and $\|X\|_{op}$ for SDP. Hence the bound (8) for these two cases does not quite align with the form of the error bound (ERB) we seek. To eliminate the dependence on this norm (by bounding the norm), either of the following conditions suffices:

- $\|x\| \le B$ for some constant B,
- $\max\{\epsilon_{opt}(x), \|\mathcal{A}x - b\|, \|x_-\|\} \le \bar{c}$ for some constant $\bar{c}$.

The second requirement combined with (8) for SOCP (SDP) implies $\|x\|_2$ ($\|X\|_{op}$) is bounded by some $\bar{B}$ depending on $\bar{c}$ but independent of x (X). We may then replace the term $\|x\|_2$ ($\|X\|_{op}$) by B or $\bar{B}$ in f. Requiring either of these two conditions produces a *local* error bound. Interestingly, no such condition on the norm is not necessary for the LP case and the other two trivial cases; hence the bounds (8) in these cases are *global* error bounds.

Remark 2 (Value of p and estimate of c_i in (ERB)) Ignoring the term f, the bound (8) in Lemma 3 is linear in $\epsilon_{opt}(x), \|\mathcal{A}x - b\|, \|x_-\|$. For LP and the two trivial cases, f is linear in $\epsilon(x_+)$, hence the error bound (ERB) holds with exponent $p = 1$. For SOCP and SDP, the square root of $\epsilon(x_+)$ appears in f, hence (8) gives an error bound of the form (ERB) with exponent $p = 2$, under the assumption $\|x\| \le B$.

Now let's consider the constants c_1, c_2, and c_3 in the error bound (ERB). It is cumbersome to estimate these for general x; here, suppose x is feasible. For the SOCP and SDP cases, also suppose $\|x\| \leq B$. Then the bound (8) reduces to

$$\mathbf{dist}(x, \mathcal{X}_\star) \leq (1 + \gamma\sigma_{max}(\mathcal{A}))f(\epsilon_{opt}(x), B). \tag{9}$$

The resulting constant c_1 for $\epsilon_{opt}(x)$ in the error bound (ERB) for each of the five cases appears in Table 1.

Remark 3 (Conditions for LP) Recall that for linear programming, dual strict complementarity is the same as strict complementarity, which always holds under strong duality [13]. Hence we need not explicitly require the dual strict complementarity condition. Also the compactness condition for strong duality in Sect. 2.1 can be dropped if we establish Theorem 1 using Hoffman's lemma [8] instead of Lemma 1.

Remark 4 (Finite product of cones) Error bounds for a conic program whose cone is a finite product of $\mathbf{R}_+^n$, SOC_n, and $\mathbf{S}_+^n$ can be established by bounding the term $\left\|\mathcal{P}_{\mathcal{V}_{s_\star}}(x_+)\right\|_2$ by a sum of te correponding fs in Lemma 3. We omit the details.

Next, we prove the bound f, violation of complementarity, in Lemma 3 for the LP, SOCP, and SDP comes by following the aforementioned procedure: (i) identify and write out $\mathcal{F}_{s_\star}$ and $\mathcal{V}_{s_\star}$, and (ii) bound the term $\left\|\mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+)\right\|_2$.

3.1 Linear programming (LP)

In linear programming, the cone $\mathcal{K} = \mathbf{R}_+^n = \{x \in \mathbf{R}^n \mid x_i \geq 0, \ \forall\, i = 1, \dots, n\}$.

Identify $\mathcal{F}_{s_\star}$ and $\mathcal{V}_{s_\star}$. For a particular dual optimal solution $(y_\star, s_\star)$, satisfying dual strict complementarity, the complementary face $\mathcal{F}_{s_\star} = \{x \in \mathbf{R}_+^n \mid x_i = 0, \text{ for all } (s_\star)_i > 0\}$, and the complementary space $\mathcal{V}_{s_\star} = \{x \in \mathbf{R}^n \mid x_i = 0 \text{ for all } (s_\star)_i > 0\}$. Hence, the term $\mathbf{dist}(\mathcal{P}_{\mathcal{V}_{s_\star}}(x_+), \mathcal{F}_{s_\star})$ is simply zero as $\mathcal{P}_{\mathcal{V}_{s_\star}}(x_+) \in \mathcal{F}_{s_\star}$.

Bound the term $\left\|\mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+)\right\|_2$. For the term $\left\|\mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+)\right\|_2$, denote $I_{s_\star} = \{i \mid (s_\star)_i > 0\}$, $I_{s_\star}^c = \{1, \dots, n\} - I_{s_\star}$, and $s_{\min>0} = \min_{i \in I_s} s_i$, we have

$$\left\|\mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+)\right\|_2 = \left\|(x_+)_{I_{s_\star}}\right\|_2 \leq \|(x_+)_{I_{s_\star}}\|_1 \ \leq \frac{1}{s_{\min>0}}\langle s_\star, x_+\rangle. \tag{10}$$

Hence, Lemma 3 for the LP case is established.

3.2 Second order programming (SOCP)

In second order cone programming, the cone $\mathcal{K}$ is $\mathrm{SOC}_n = \{x = (x_{1:n}, x_{n+1}) \mid \|x_{1:n}\|_2 \leq x_{n+1}, x_{1:n} \in \mathbf{R}^n, x_{n+1} \in \mathbf{R}\}$.

Identify $\mathcal{F}_{s_\star}$ and $\mathcal{V}_{s_\star}$. Given a dual solution $s_\star = (s_{\star,1:n}, s_{\star,n+1}) \in (\mathrm{SOC}_n)^* = \mathrm{SOC}_n$ satisfying dual strict complementarity, the complementary face is defined as $\mathcal{F}_{s_\star} = \{x \mid \langle s_{\star,1:n}, x_{1:n} \rangle + x_{n+1} s_{\star,n+1} = 0, ; x \in \mathrm{SOC}_n\}$. We can further simplify this expression as

$$\mathcal{F}_{s_\star} = \{\lambda(-s_{\star,1:n}, s_{\star,n+1}) \mid \lambda \geq 0\}.$$

The complementary space $\mathcal{V}_{s_\star}$ for the nontrivial case $s_\star \in \partial \, \mathrm{SOC}_n \setminus (0,0)$ is simply

$$\mathcal{V}_{s_\star} = \mathbf{span}\{(-s_{\star,1:n}, s_{\star,n+1})\}.$$

For the dual optimal solution $s_\star$, denote $\check{s}_\star = \frac{1}{\|s_\star\|}(-s_{\star,1:n}, s_{\star,n+1})$. Note that the term $\mathbf{dist}(\mathcal{P}_{\mathcal{V}_{s_\star}}(x_+), \mathcal{F}_{s_\star})$ is again simply zero as $\mathcal{P}_{\mathcal{V}_{s_\star}}(x_+) \in \mathcal{F}_{s_\star}$.

Bound the term $\left\|\mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+)\right\|_2$. We now turn to analyze $\mathcal{P}_{\mathcal{V}^\perp}(x_+)$, which can be written explicitly as

$$\mathcal{P}_{\mathcal{V}^\perp}(x_+) = x_+ - \langle x_+, \check{s}_\star \rangle \check{s}_\star. \tag{11}$$

Now introduce the shorthand $\epsilon(x) = \langle x_+, \frac{s_\star}{\|s_\star\|} \rangle = \langle x_{+,1:n}, \frac{s_{\star,1:n}}{\|s_\star\|} \rangle + x_{+,n+1} \frac{s_{\star,n+1}}{\|s_\star\|}$. The norm square of $\mathcal{P}_{\mathcal{V}^\perp}(\mathbf{x}_+)$ can be written as

$$
\begin{aligned}
\|\mathcal{P}_{\mathcal{V}^\perp}(x_+)\|_2^2 &= \|x_+\|_2^2 - \langle x_+, \check{s}_\star \rangle^2 \\
&= \|x_+\|_2^2 + x_{+,n+1}^2 - \left(-\langle x_{+,1:n}, s_{\star,1:n} \rangle + x_{+,n+1} \frac{s_{\star,n+1}}{\|s_\star\|}\right)^2 \\
&\overset{(a)}{\leq} 2x_{+,n+1}^2 - (2x_{+,n+1} \frac{s_{\star,n+1}}{\|s_\star\|} - \epsilon(x))^2 \overset{(b)}{\leq} -\epsilon^2(x) + 2\sqrt{2} x_{+,n+1} \epsilon(x),
\end{aligned}
\tag{12}
$$

where in step (a), we use the fact that $x_+ \in \mathrm{SOC}_n$ and the definition of ϵ. In step (b), we use the fact $s_\star \in \partial \, \mathrm{SOC}_n$. Lemma 3 for the SOCP case is established by noting $x_{+,n+1} \leq \|x_+\| \leq \|x\|$.

3.3 Semidefinite programming (SDP)

In semidefinite programming, the cone $\mathcal{K}$ is $\mathbf{S}_+^n = \{X \in \mathbf{S}^n \mid X \succeq 0\}$. Note that we use capital letter X and S for matrices.

Identify $\mathcal{F}_{s_\star}$ and $\mathcal{V}_{s_\star}$. For a dual optimal solution $(y_\star, S_\star)$ with $S_\star = C - \mathcal{A}y_\star \succeq 0$ satisfying dual strict complementarity, the complementary face $\mathcal{F}_{S_\star} := \{X \mid \langle X, S_\star \rangle = 0, X \succeq 0\} = \{X \mid X = VRV^\top, R \in \mathbf{S}^r, R \succeq 0\}$, where $\mathbf{rank}(S_\star) = n - r$, and $V \in \mathbf{R}^{n \times r}$ is a matrix with orthonormal columns that span

range$(S_\star)$. In this case, the complementary space $\mathcal{V}_{S_\star} = \{VRV^\top \mid R \in \mathbf{S}^r\}$. We note that the term **dist**$(\mathcal{P}_{\mathcal{V}_{S_\star}}(X_+), \mathcal{F}_{S_\star})$ is again zero as $\mathcal{P}_{\mathcal{V}_{S_\star}}(X_+) \in \mathcal{F}_{S_\star}$.

Bound the term $\left\|\mathcal{P}_{\mathcal{V}_{S_\star}^\perp}(X_+)\right\|_{\mathrm{F}}$. We now turn to bound the term $\left\|\mathcal{P}_{\mathcal{V}_{S_\star}^\perp}(X_+)\right\|_{\mathrm{F}}$. We utilize Lemma 4 [19, Lemma 4.3] to bound the term $\left\|\mathcal{P}_{\mathcal{V}_{S_\star}^\perp}(X_+)\right\|_{\mathrm{F}}$ with $S = S_\star$, and use the observation that $\|X_+\|_{op} \leq \|X\|_{op}$.

Lemma 4 *Suppose* $X, S \in \mathbf{S}^n$ *are both positive semidefinite. Let* $V \in \mathbf{R}^{n \times r}$ *be the matrix formed by the eigenvectors with the r smallest eigenvalues of S and define* $\mathcal{V} = \mathbf{range}(V)$. *Let* $\epsilon = \mathbf{tr}(XS)$. *If* $T = \lambda_{n-r}(S) > 0$, *then*

$$\left\|\mathcal{P}_{\mathcal{V}^\perp}(X)\right\|_{\mathrm{F}} \leq \frac{\epsilon}{T} + \sqrt{2\frac{\epsilon}{T}\|X\|_{op}} \quad \text{and} \quad \left\|\mathcal{P}_{\mathcal{V}^\perp}(X)\right\|_* \leq \frac{\epsilon}{T} + 2\sqrt{r\frac{\epsilon}{T}\|X\|_{op}}.$$

4 Application: sensitivity of solution

As discussed in the introduction, to study the sensitivity of the solution, we consider a solution $x_\star'$ of the perturbed problem

$$\begin{aligned}
\underset{x}{\text{minimize}} \quad & \langle c', x\rangle \\
\text{subject to} \quad & \mathcal{A}'x = b', \\
& x \in \mathcal{K}.
\end{aligned} \qquad (\mathcal{P}')$$

where the problem data $(\mathcal{A}', b', c') = (\mathcal{A}, b, c) + (\Delta\mathcal{A}, \Delta b, \Delta c)$ for some small perturbation $\Delta = (\Delta\mathcal{A}, \Delta b, \Delta c)$, and ask how the distance $\|x_\star' - x_\star\|_2$ changes according to Δ.

Note that once the error bound (ERB) is established, we can understand the sensitivity of the solution by estimating the suboptimality, linear and conic infeasibility of the new solution $x_\star'$ with respect to the original problem $(\mathcal{P})$ via the perturbation Δ. Following this strategy, we prove the following theorem:

Theorem 2 *Suppose the primal and dual Slater's condition holds for some (x_0, y_0), and the map $\mathcal{A}$ is surjective. For any small enough $\epsilon > 0$, there is some constant $\bar{c}$, such that for any optimal $x_\star'$ of $(\mathcal{P}')$, and all $\|\Delta\| := \|\Delta\mathcal{A}\|_{op} + \|\Delta b\|_2 + \|\Delta c\|_2 \leq \epsilon$, we have*

$$\max\{\epsilon_{opt}(x_\star'), \left\|\mathcal{A}(x_\star') - b\right\|_2\} \leq \bar{c}\|\Delta\|.$$

Hence if (ERB) holds, then (SSB) holds with $p = p'$.

To facilitate the proof, we define the smallest nonzero singular value of $\mathcal{A}$ as $\sigma_{\min>0}(\mathcal{A}) = \min_{\|x\|_2 = 1, x \perp \mathbf{nullspace}(\mathcal{A})} \|\mathcal{A}(x)\|_2$, and the pseudoinverse of $\mathcal{A}'$ as $(\mathcal{A}')^\dagger(y) = \operatorname{argmin}_{\mathcal{A}'(x)=y} \|x\|_2$.

Proof Consider any solution $x'_\star$ to the problem $(\mathcal{P}')$. Suppose the following assumptions are satisfied (proved in Appendix A):

1. *Primal and dual Slater's condition for* $(\mathcal{P}')$: there exist a primal x'_0 and dual y'_0 solution feasible for problem $(\mathcal{P}')$ and its dual that satisfy $\min\{\mathbf{dist}(x'_0, \partial\mathcal{K}), \mathbf{dist}(c - \mathcal{A}^* y'_0, \partial(\mathcal{K}^*))\} > \rho$ for some ρ independent of Δ, and $\max\{\|x_0 - x'_0\|_2, \|y_0 - y'_0\|_2\} < \xi$ for some ξ independent of Δ.
2. *Boundedness of primal solutions of* $(\mathcal{P}')$: There is some $B > 0$ independent of Δ such that any solution $x_\star$ to $(\mathcal{P}')$ satisfies $\|x_\star\|_2 \le B$.

Let us start with the linear infeasibility $\mathcal{A}(x'_\star) - b$. Using the linear feasibility of $x'_\star$, $\mathcal{A}' x'_\star = b'$, w.r.t. $(\mathcal{P}')$, we have

$$
\begin{aligned}
\mathcal{A}'(x'_\star) = b' &\implies (\mathcal{A} + \Delta\mathcal{A})x'_\star = b + \Delta b \implies \mathcal{A}x'_\star - b = \Delta b - \Delta\mathcal{A}x'_\star \\
&\implies \|\mathcal{A}x'_\star - b\|_2 \le \|\Delta b\|_2 + \|\Delta\mathcal{A}\|_{op}B.
\end{aligned}
\tag{13}
$$

This shows $\|\mathcal{A}x'_\star - b\|_2 \le c\|\Delta\|$ for any $c > B + 1$.

Next, consider $\epsilon_{opt}(x'_\star)$. We would like to use the optimality of $x'_\star$ to $(\mathcal{P}')$ and compare it against $x_\star$. However, since $x_\star$ is not necessarily feasible for $(\mathcal{P}')$, we need more subtle reasoning. Consider $\hat{x} := (1 - \alpha)(x_\star + (\mathcal{A}')^\dagger(\Delta b - \Delta\mathcal{A}x_\star)) + \alpha x'_0$ with $\alpha = \frac{\|(\mathcal{A}')^\dagger(\Delta b - \Delta\mathcal{A}x_\star)\|_2}{\rho + \|(\mathcal{A}')^\dagger(\Delta b - \Delta\mathcal{A}x_\star)\|_2}$. Here $(\mathcal{A}')^\dagger$ exists and has largest singular value at most $\frac{2}{\sigma_{\min>0}(\mathcal{A})}$ as long as $\sigma_{\max}(\Delta\mathcal{A}) \le \frac{\sigma_{\min>0}(\mathcal{A})}{2}$. We have $\|\hat{x}\|_2 \le B_1$ for some Δ independent B_1 as $\|x_0 - x'_0\|_2 < \xi$. Note that $\hat{x} \in \mathcal{K}$ since

$$
\hat{x} = (1 - \alpha)\underbrace{x_\star}_{\in\mathcal{K}} + \alpha\underbrace{\left(x'_0 + \frac{1 - \alpha}{\alpha}(\mathcal{A}')^\dagger(\Delta b - \Delta\mathcal{A}x_\star)\right)}_{\in\mathcal{K} \text{ since } \left\|\frac{1-\alpha}{\alpha}(\mathcal{A}')^\dagger(\Delta b - \Delta\mathcal{A}x_\star)\right\|_2 \le \rho},
$$

if $\alpha > 0$. The case of $\alpha = 0$ is trivial. We also know $\hat{x}$ is feasible with respect to the linear constraints of $(\mathcal{P}')$:

$$
\begin{aligned}
\mathcal{A}'(\hat{x}) &= (1 - \alpha)\mathcal{A}'(x_\star + (\mathcal{A}')^\dagger(\Delta b - \Delta\mathcal{A}x_\star)) + \alpha\mathcal{A}'x'_0 \\
&= (1 - \alpha)(b + \Delta\mathcal{A}(x_\star) + \Delta b - \Delta\mathcal{A}x_\star) + \alpha b' = b'.
\end{aligned}
\tag{14}
$$

Note by construction $\hat{x} - x_\star = \alpha(x'_0 - x_\star - \alpha(\mathcal{A}')^\dagger(\Delta b - \Delta\mathcal{A}x_\star)) =: \Delta x$ with $\alpha = \frac{\|(\mathcal{A}')^\dagger(\Delta b - \Delta\mathcal{A}x_\star)\|_2}{\rho + \|(\mathcal{A}')^\dagger(\Delta b - \Delta\mathcal{A}x_\star)\|_2}$. Hence $\|\Delta x\|_2 \le c'\left(\|\Delta b\|_2 + \|\Delta\mathcal{A}\|_{op}\right)$ for some constant c'. Now using the optimality of $x'_\star$, we have

$$
c'x'_\star \le c'\hat{x} \implies \epsilon_{opt}(x'_\star) = c^\top x'_\star - c^\top x_\star \le \Delta c^\top(\hat{x} - x'_\star) + c^\top \Delta x.
\tag{15}
$$

Hence $\epsilon_{opt}(x'_\star) \le \bar{c}\|\Delta\|$ for large enough $\bar{c}$ using $\|x'_\star\|_2 \le B$.

5 Discussion

Using the framework established in Sect. 2, we have shown an error bound of the form (ERB) and a bound on the sensitivity of the solution with respect to problem data (SSB) for a broad class of problems: LP, SOCP, and SDP, and conic programs for which the cone is a finite product of the LP, SOCP, and SDP cones. Let us now compare the results we have obtained with the literature on error bounds and sensitivity of solution.

Error bound The celebrated results of [8] show that the error bound is linear for linear programming: in (ERB), the exponent $p = 1$. The work of Sturm [9, Section 4] shows that under strict complementarity and compactness of the solution set, SDPs satisfy a quadratic error bound: $p = 2$ in (ERB). Sturm also discusses the exponent of $\textbf{dist}(x, \mathcal{X}_\star)$ without strict complementarity: it can be bounded by 2^d where d is the *singularity degree*, which is at most $n - 1$ [9, Lemma 3.6]. A recent result shows that $p = 2$ under dual strict complementarity type conditions when the cone $\mathcal{K}$ is a symmetric cone under the framework of *amenable cones* [20]. When the cone $\mathcal{K}$ is defined as a semialgebraic set (LP, SOCP, SDP are special cases), Drusvyatskiy, Ioffe, and Lewis [21, Corrollary 4.8] showed that for generic C, the exponent p is always 2 when the inequality is restricted to feasible x. We note that the proofs for these bounds do not provide estimates for c_i. In our framework, estimates of c_i (expressed in terms of the primal and dual solution) can be obtained supposing the primal solution is unique.

Sensitivity of solution When $p' = 1$, the bound describing the sensitivity of the solution (SSB) is also called stability or metric regularity [22, Definition A.6], discussed in detail in [22, Appendix A], and [23, 24]. In the context of semidefinite programming, when the primal and dual solutions are unique and strict complementarity holds, Nayakkankuppam and Overton [10] shows that (SSB) holds with $p' = 1$. When the cone $\mathcal{K}$ is a semialgebraic set, Drusvyatskiy, Ioffe, and Lewis [21, section 5] showed that for generic perturbations $C - C'$, $b - b'$, the sensitivity bound (SSB) holds with $p' = 1$.

In Sect. 4, we have seen that (SSB) holds whenever (ERB) with $p' = p$. Hence, using our results from Sect. 3, we see $p' = 1$ for LP, and $p' = 2$ for SOCP and SDP. This result improves on the previous bound for SOCPs and SDPs assuming only strict complementarity.

Extension to quadratic programming (QP)? A potential future direction is to use the approach of the paper to establish error bounds for QP. The strategy consists of three steps: (1) reducing the QP to an SOCP, (2) utilizing the error bound for the SOCP, and (3) translating the error bound to the QP setting. We leave the detail to future work.

Other Cones? An interesting future direction is the extension to other cones, e.g., the copositive cone, the completely positive cone, and the doubly positive cone (the intersection of nonnegative matrices and positive semidefinite matrices). Can we still bound the term $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x) \right\|_2$? For the cones $\mathbf{R}_+^n$, SOC_n, and $\mathbf{S}_+^n$, our technique relies on the explicit structure of $\mathbf{R}_+^n$, SOC_n, and $\mathbf{S}_+^n$ to bound $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x) \right\|_2$. Characterizing the facial structure seems to be challenging for other cones.

Appendix A: Proof for Section 4

We first show the Slater's condition for $(\mathcal{P}')$ and its dual. Recall the Slater's condition for the original $(\mathcal{P})$ means that the two points x_0, y_0 satisfying $(x_0, c - \mathcal{A}^* y) \in \textbf{int}(\mathcal{K}) \times \textbf{int}(\mathcal{K}^*)$. This implies that $\min\{\textbf{dist}(x_0, \partial\mathcal{K}), \textbf{dist}(c - \mathcal{A}^* y_0, \partial\mathcal{K}^*)\} \geq \eta$ for some constant $\eta > 0$. We now construct x_0' and y_0' from x_0 and y_0. It can be easily verified if $\|\Delta\mathcal{A}\| \leq \frac{\sigma_{\min>0}(\mathcal{A})}{2}$, $\frac{2}{\sigma_{\min>0}(\mathcal{A})}(\|\Delta b\| + \|\Delta\mathcal{A}\|\|x_0\|) \leq \frac{\eta}{2}$, and $\|\Delta c\| + \|\Delta\mathcal{A}\|\|y_0\| \leq \frac{\eta}{2}$, then the choice

$$x_0' = x_0 + \mathcal{A}\simeq^\dagger(\Delta b - \Delta\mathcal{A}x_0), \quad y_0' = y_0 \tag{16}$$

satisfy $\min\{\textbf{dist}(x_0', \partial\mathcal{K}), \textbf{dist}(c' - (\mathcal{A}^*)' y_0', \partial(\mathcal{K}^*))\} \geq \frac{\eta}{2}$, $\max\{\|x_0 - x_0'\|, \|y_0 - y_0'\|\} \leq \frac{\eta}{2}$, and are feasible for $(\mathcal{P}')$ and its dual.

Now for the boundedness condition of any solution $x_\star'$ to $(\mathcal{P}')$. Using the previous constructed x_0' and y_0', we know

$$
\begin{aligned}
\langle x_\star', c' - (\mathcal{A}^*)' y_0' \rangle &\leq \langle x_0', c' - (\mathcal{A}^*)' y_0' \rangle = \langle c', x_0' \rangle - \langle b', y_0' \rangle \\
&= \langle c, x_0 \rangle - \langle b, y_0 \rangle + \langle \Delta c, x_0' \rangle - \langle \Delta b, y_0' \rangle \\
&\leq \langle c, x_0 \rangle - \langle b, y_0 \rangle + \frac{\eta}{2}(\|x_0\| + \frac{\eta}{2} + \frac{\sigma_{\min>0}(\mathcal{A})}{2}\|y_0\|)
\end{aligned}
\tag{17}
$$

The rest is a simple consequence of the following lemma.

Lemma 5 *Suppose $\mathcal{K}$ is closed and convex. Given $\epsilon > 0$ and $s_0 \in (\mathcal{K}^*)^\circ$ with $d = \textbf{dist}(s_0, \partial\mathcal{K}^*) > 0$, there is a $B > 0$ such that for any x satisfying $x \in \mathcal{K}$, and $\langle x, s \rangle \leq \epsilon$ for some s with $\|s - s_0\| \leq \frac{d}{2}$, its norm satisfies $\|x\| \leq B$.*

Proof Suppose such B does not exist, then there is a sequence $(x_n, s_n) \in \mathcal{K} \times \mathcal{K}^*$ with $\|s_n - s_0\| \leq \frac{d}{2}, \langle x_n, s_n \rangle \leq \epsilon$, and $\lim_{n\to\infty} \|x_n\| = +\infty$.

Now consider $(\frac{x_n}{\|x_n\|}, s_n) \in \mathcal{K} \times \mathcal{K}^*$. Since $\frac{x_n}{\|x_n\|}$ and s_n are bounded, we can choose a appropriate subsequence of $(\frac{x_n}{\|x_n\|}, s_n)$ which converges to certain $(x, s) \in \mathcal{K} \times \textbf{int}(\mathcal{K}^*)$ as $\|s_n - s_0\| \leq \frac{d}{2}$. Call the subsequence $(\frac{x_n}{\|x_n\|}, s_n)$ still. Using $\langle \frac{x_n}{\|x_n\|}, s_n \rangle \leq \frac{\epsilon}{\|x_n\|}$ and $\|x_n\| \to +\infty$, we see

$$\langle x, s \rangle = 0.$$

This is not possible as $s \in \textbf{int}(\mathcal{K}^*)$. Hence such B must exist.

Appendix B: Equivalence between DSC and SC for LP and SDP

We define the strict complementarity of LP and SDP, and show it is equivalent to DSC defined in Sect. 2.1. For a vector $x \in \mathbf{R}^n$, denote $\mathbf{nnz}(x)$ as its number of nonzeros.

Definition 3 For LP, if there exists optimal primal dual pair $(x_\star, y_\star) \in \mathcal{X}_\star \times \mathcal{Y}_\star \subset \mathbf{R}^n_+ \times \mathbf{R}^m$ with $s_\star = c - \mathcal{A}^* y_\star \in \mathbf{R}^n_+$ such that

$$\mathbf{nnz}(x_\star) + \mathbf{nnz}(s_\star) = n,$$

we say $(\mathcal{P})$ satisfies strict complementarity. Similary, for SDP, if there exists optimal primal dual pair $(X_\star, y_\star) \in \mathcal{X}_\star \times \mathcal{Y}_\star \subset \mathbf{S}^n_+ \times \mathbf{R}^m$ with $S_\star = C - \mathcal{A}^* y_\star \in \mathbf{S}^n_+$ such that

$$\mathbf{rank}(X_\star) + \mathbf{rank}(S_\star) = n,$$

we say $(\mathcal{P})$ satisfies strict complementarity.

Lemma 6 *For both LP and SDP, under strong duality, the strict complementarity defined is equivalent to dual strict complementarity.*

Proof For LP, under strong duality (see (SD)), we have for any optimal $x_\star$, $x_\star \in \mathcal{F}_{s_\star} = \{x \in \mathbf{R}^n_+ \mid x_i = 0, \text{ for all } (s_\star)_i > 0\}$. The relative interior of $\mathcal{F}_{s_\star}$ is

$$\mathrm{rel}(\mathcal{F}_{s_\star}) = \{x \in \mathbf{R}^n_+ \mid x_i = 0 \text{ for all } (s_\star)_i > 0, \text{ and } x_i > 0 \text{ for all } (s_\star)_i = 0\}.$$

The equivalence between DSC and SC is then immediate.

For SDP, under strong duality (see (SD)), we know that for any optimal $X_\star$, $X_\star \in \mathcal{F}_{S_\star} = \{X \mid \langle X, S_\star \rangle = 0, X \geq 0\} = \{X \mid \mathbf{range}(X) \subset \mathbf{nullspace}(S_\star), \text{ and } X \geq 0\}$. The relative interior of $\mathcal{F}_{S_\star}$ is

$$\mathrm{rel}(\mathcal{F}_{S_\star}) = \{X \mid \mathbf{range}(X) = \mathbf{nullspace}(S_\star), \text{ and } X \geq 0\}.$$

The equivalence is immediate by using Rank-nullity theorem for $S_\star$.

Appendix C: A lower bound on distance to optimality: $\left\| \mathcal{P}_{\mathcal{V}^\perp_{s_\star}}(x) \right\|_2 \leq \mathbf{dist}(x, \mathcal{X}_\star)$

We have shown how to establish upper bounds on $\mathbf{dist}(x, \mathcal{X}_\star)$ with an upper bound on $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+) \right\|_2$. In this section, we show that the same quantity $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+) \right\|_2$ also yields a lower bound. Hence, it is important to understand the behavior of $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+) \right\|_2$. For simplicity, we suppose in this section that x is feasible for the problem $(\mathcal{P})$: in this case, $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+) \right\|_2 = \left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x) \right\|_2$. The argument for infeasible x is essentially the same.

First recall for feasible x, the only nonzero error metric is the suboptimality $\epsilon_{opt}(x) = \langle c, x \rangle - p_\star$. Also note that $\epsilon_{opt}(x) = 0 \iff \left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x) \right\|_2$ using complementary slackness. Hence, there is some nonnegative function $g : \mathbf{R} \to \mathbf{R}_+$ such that $g(\epsilon_{opt}(x)) \le \left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x) \right\|_2$. Now note that for any feasible x, we always have the lower bound on distance given by $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x) \right\|_2$ as $\mathcal{V}_{s_\star} \supset \mathcal{X}_\star$,

$$g(\epsilon_{opt}(x)) \le \left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x) \right\|_2 = \mathbf{dist}(x, \mathcal{V}_{s_\star}) \le \mathbf{dist}(x, \mathcal{X}_\star). \tag{18}$$

The lower bound (18) hence shows that $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+) \right\|_2$ provides a lower bound on the distance $\mathbf{dist}(x, \mathcal{X}_\star)$, and provides hope that this bound might scale with the suboptimality $\epsilon_{opt}(x)$. We summarize our findings in the following theorem.

Theorem 3 *Suppose there exists a increasing continuous g with $g(0) = 0$ so that for any x feasible for Problem* $(\mathcal{P})$[8] *with $\epsilon(x) \le \bar{c}$,*

$$\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x) \right\|_2 \ge g(\epsilon_{opt}(x)). \tag{19}$$

Then the following inequality holds:

$$\mathbf{dist}(x, \mathcal{X}_\star) \ge g(\epsilon_{opt}(x)). \tag{20}$$

Remark 5 We also have a partial converse for the above theorem that follows from the same proof. Assume $\mathbf{dist}(\mathcal{P}_{\mathcal{V}_{s_\star}}(x), \mathcal{F}_{s_\star})$ is 0. If the relation (20) holds for all feasible x with $\epsilon_{opt}(x) \le \bar{c}$, then $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x) \right\|_2 \ge \frac{g(\epsilon_{opt}(x))}{1 + \gamma \sigma_{\max}(\mathcal{A})}$. Here γ is defined in (3).

Proof We have proved that (19) implies (20) using the motivating logic laid out at the beginning of this section.

Conversely, assume the term $\mathbf{dist}(\mathcal{P}_{\mathcal{V}_{s_\star}}(x_+), \mathcal{F}_{s_\star})$ is zero, x is feasible, $\epsilon_{opt}(x) \le \bar{c}$, and the inequality (20) holds. We see that for all feasible x,

$$\beta g(\epsilon_{opt}(x)) \le \mathbf{dist}(x, \mathcal{X}_\star) \overset{(a)}{\le} (1 + \gamma \sigma_{\max}(\mathcal{A})) \left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x) \right\|_2. \tag{21}$$

where we use (7) for step (a).

Appendix D: Conic decomposition

In the main paper, we decompose a general $x \in \mathbf{E}$ according to the subspace $\mathcal{V}_{s_\star}$. A different decomposition uses the cone $\mathcal{F}_{s_\star}$: every $x \in \mathbf{E}$ admits the conic decomposition $x = \mathcal{P}_{\mathcal{F}_{s_\star}}(x) + \mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x)$ where $\mathcal{F}_{s_\star}^\circ$ is the *polar cone* of $\mathcal{F}_{s_\star}$, i.e., the negative dual cone $-\mathcal{F}_{s_\star}^*$.

---

[8] The assumption x being feasible is just for convenience of presentation. The equivalence still holds for all x with suboptimality, infeasibility, and conic infeasibility bounded above by some constant $\bar{c} > 0$.

Theorem 4 *Suppose strong duality and dual strict complementarity hold. Then for some constants γ, γ' described in Lemma* 2 *and for all $x \in \mathbf{E}$, we have*

$$\text{dist}(x, \mathcal{X}_\star) \leq (1 + \gamma \sigma_{\max}(\mathcal{A})) \left\| \mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x) \right\|_2 + \gamma \|\mathcal{A}(x) - b\|_2. \tag{22}$$

Let us compare the above bound (22) and (4) in Theorem 1. To make the comparison easier, first note that from the proof of Theorem 1, we can bound the distance from x to $\mathcal{X}_\star$ using the decomposition $x = \mathcal{P}_{\mathcal{V}_{s_\star}}(x) + \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x)$:

$$\text{dist}(x, \mathcal{X}_\star) \leq (1 + \gamma \sigma_{\max}(\mathcal{A})) \left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x) \right\|_2 + \gamma \|\mathcal{A}(x) - b\|_2 + \gamma' \text{dist}(\mathcal{P}_{\mathcal{V}_{s_\star}}(x), \mathcal{F}_{s_\star}). \tag{23}$$

The bound (4) in Theorem 1 is further obtained via the decomposition $x = x_+ + x_-$.

Comparing (23) and (22), we find that there is an extra term $\gamma' \text{dist}(\mathcal{P}_{\mathcal{V}_{s_\star}}(x), \mathcal{F}_{s_\star})$ in (23) and the term $\left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x) \right\|_2$ in (23) is replaced by $\left\| \mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x) \right\|_2$. Since $\left\| \mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x) \right\|_2 \geq \left\| \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x) \right\|_2$, it is not immediately clear which bound is tighter.

A more subtle difference between (22) and (4) in Theorem 1 is that we are not able to further bound $\left\| \mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x) \right\|_2$ using the decomposition $x = x_+ + x_-$ with respect to $\mathcal{K}$. We reach this impasse because the projection operator $\mathcal{P}_{\mathcal{F}_{s_\star}^\circ}$ is not linear and so we cannot rely on the triangle inequality $\left\| \mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x) \right\|_2 \leq \left\| \mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x_+) \right\|_2 + \left\| \mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x_-) \right\|_2$. Hence we cannot bound $\left\| \mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x) \right\|_2$ using conic infeasibility.

Thus, to use (23), we must use x (which may be infeasible with respect to the cone $\mathcal{K}$) directly to bound $\left\| \mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x) \right\|_2$. We now consider how to bound this term for the cases considered in the main text.

Case $s_\star = 0$. For $s_\star = 0$, we have $\mathcal{F}_{s_\star} = \mathcal{K}$. Thus $\mathcal{F}_{s_\star}^\circ = \mathcal{K}^\circ$ and $\mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x) = x_-$. For $s_\star \in \mathbf{int}(\mathcal{K}^*)$, we have $\mathcal{F}_{s_\star} = \{0\}$. Thus $\mathcal{F}_{s_\star}^\circ = \mathbf{E}$ and $\left\| \mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x) \right\|_2 = \|x\|_2 \leq \|x_+\|_2 + \|x_-\|_2$. We can bound $\|x_+\|_2$ as in Lemma 3.

Case $s_\star = \mathbf{R}_+$. For $\mathcal{K} = \mathbf{R}_+$, the projection of $\mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x) = x - (x_{I_{s_\star}^c})_+ = (x_+)_{I_{s_\star}} + x_-$ where $x_{I_{s_\star}^c}$ is the vector x zeroing all entries in the support of $s_\star$ and $(x_+)_{I_{s_\star}}$ is the vector x_+ zeroing out all entries not in the support of $s_\star$. Hence, $\left\| \mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x) \right\|_2 \leq \left\| (x_+)_{I_{s_\star}} \right\|_2 + \|x_-\|_2$ and we can further bound $\left\| (x_+)_{I_{s_\star}} \right\|_2$ as in Lemma 3.

Case $s_\star = \text{SOC}_n$. For $\mathcal{K} = \text{SOC}_n$, considering the nontrivial case $s_\star \neq 0$, the projection of $\mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x)$ is

$$x - (\langle x, \check{s}_\star \rangle)_+ \check{s}_\star = x_+ - (\langle x, \check{s}_\star \rangle)_+ \check{s}_\star + x_-.$$

Thus we have

$$\left\|\mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x) - \mathcal{P}_{\mathcal{V}_{s_\star}^\perp}(x_+)\right\|_2 = \left\|(\langle x_+, \check{s}_\star\rangle - (\langle x, \check{s}_\star\rangle)_+)\check{s}_\star + x_-\right\|_2$$
$$\leq |\langle x_+, \check{s}_\star\rangle - (\langle x, \check{s}_\star\rangle)_+| + \|x_-\|_2 \qquad (24)$$
$$\overset{(a)}{\leq} |\langle x_+, \check{s}_\star\rangle| + \|x_-\|_2.$$

In the step (a), we use $x = x_+ + x_-$ and $\langle x_-, \check{s}_\star\rangle \leq 0$. Hence, one can bound $\left\|\mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x)\right\|_2$ by combining the above bound and the bound on $\mathcal{P}_{\mathcal{V}^\perp}(x_+)$ established in the main text.

Case $s_\star = \mathbf{S}_+^n$. For $\mathcal{K} = \mathbf{S}_+^n$, $\mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(X)$ is

$$\mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(X) = X - V(V^\top X V)_+ V^\top = X_+ - V(V^\top X V)_+ V^\top + X_-.$$

Since $\mathcal{P}_{\mathcal{V}_{s_\star}^\top}(X+) = X_+ - VV^\top X_+ VV^\top$, we know

$$\left\|\mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(X) - \mathcal{P}_{\mathcal{V}^\perp}(X_+)\right\|_F = \left\|V(V^\top X V)_+ V^\top - VV^\top X_+ VV^\top + X_-\right\|_F$$
$$\overset{(a)}{\leq} \left\|(V^\top X V)_+ - V^\top X_+ V\right\|_F + \|X_-\|_F \qquad (25)$$
$$\overset{(b)}{\leq} 2\|X_-\|_F.$$

In step (a), we use the fact that V has orthonormal columns, and in step (b), we use the fact that $V^\top X_+ V$ is still positive semidefinite and projection to the convex set $\mathbf{S}_+^r$ is nonexpansive. Thus, we can bound $\left\|\mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(X)\right\|_F$ using the result for $\left\|\mathcal{P}_{\mathcal{V}^\perp}(X_+)\right\|_F$ in the main text.

Case x is feasible. Finally, note that when x is feasible for $(\mathcal{P})$, then in each of the five cases considered in the paper, the bound (22) and the bound (4) in Theorem 1 coincide.

Proof Recall Lemma 2 establishes an error bound only for $x \in \mathcal{V}_{s_\star}$. Using the conic decomposition into the face $\mathcal{F}$ and its polar, for any $x \in \mathbf{E}$,

$$x = \mathcal{P}_{\mathcal{F}_{s_\star}}(x) + \mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x).$$

This decomposition immediately gives

$$\mathbf{dist}(x, \mathcal{X}_\star) \leq \left\|\mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x)\right\|_2 + \mathbf{dist}(\mathcal{P}_{\mathcal{F}_{s_\star}}(x), \mathcal{X}_\star). \qquad (26)$$

The second term $\mathbf{dist}(\mathcal{P}_{\mathcal{F}_{s_\star}}(x), \mathcal{X}_\star)$ can be bounded using Lemma 2 as $\mathcal{F}_{s_\star} \subset \mathcal{V}_{s_\star}$:

$$\mathbf{dist}(\mathcal{P}_{\mathcal{F}_{s_\star}}(x), \mathcal{X}_\star) \leq \gamma \left\|\mathcal{A}(\mathcal{P}_{\mathcal{F}_{s_\star}}(x)) - b\right\|_2. \qquad (27)$$

Using the decomposition $x = \mathcal{P}_{\mathcal{F}_{s_\star}}(x) + \mathcal{P}_{\mathcal{F}_{s_\star}^\circ}(x)$ again for the term $\left\|\mathcal{A}(\mathcal{P}_{\mathcal{F}_{s_\star}}(x)) - b\right\|_2$, we reach the bound (22).
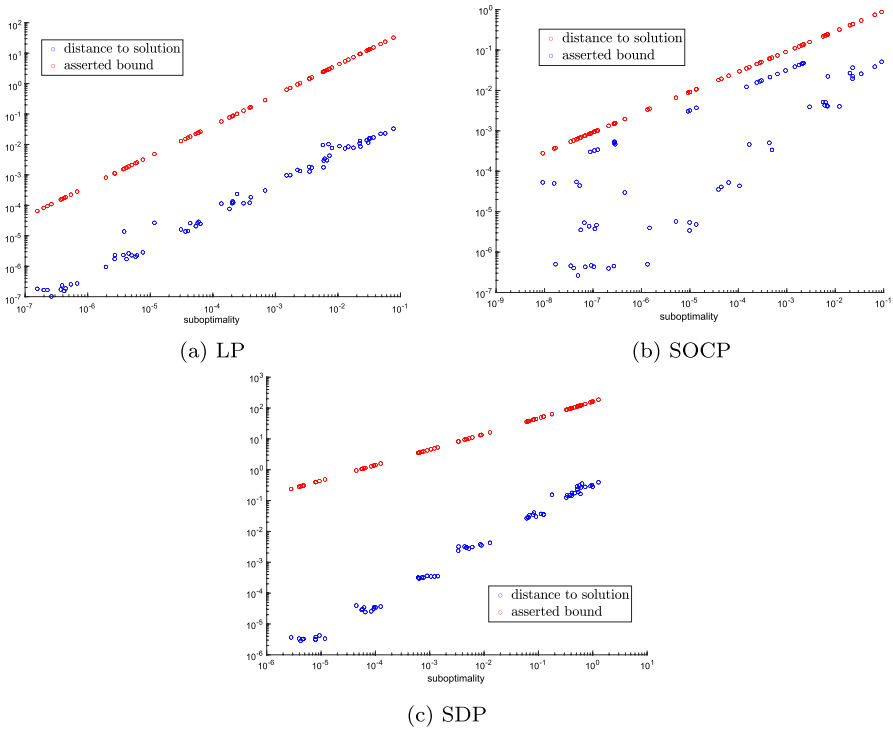
Fig. 2 Verification of the inequality (28). The asserted bound in red is $(1 + \gamma \sigma_{\max}(\mathcal{A})) f(\epsilon_{opt}(x), \|x\|)$. The blue points represent suboptimality versus distance to solution (color figure online)

Appendix E: Numerical simulation for the bound (8)

Here we numerically verify the correctness of Inequality (8) for feasible x:

$$\mathbf{dist}(x, \mathcal{X}_\star) \leq (1 + \gamma \sigma_{\max}(\mathcal{A})) f(\epsilon_{opt}(x), \|x\|). \tag{28}$$

The function f can be found in Table 1.

Experiment setup We generated a random instance $\mathcal{A}, b, c$ for each of LP, SOCP, and SDP. We solved the corresponding conic problem and obtained the optimal solution $x_\star$ and a dual optimal $s_\star$. We numerically verified that the strict complementarity (by checking Definition 3 for LP and SDP and (DSC) for SOCP) and the uniqueness of the primal (by checking whether $\sigma_{\min}(\mathcal{A}_{\mathcal{V}_\star}) > 0$) both hold for the three cases. We compute $\gamma = \frac{1}{\sigma_{\min}(\mathcal{A}_{\mathcal{V}_\star})}$ according to Lemma 2. Next, we randomly perturbed the solution $x_\star$ 70 many times and obtained (possibly infeasible) $x'_i, i = 1, \ldots, 70$. We then projected x'_i to the feasible set to obtain x_i. Finally, we plotted the suboptimality of x_i versus the distance to $x_\star$ (in blue), and the the suboptimality of x_i versus the bound $(1 + \gamma \sigma_{\max}(\mathcal{A})) f(\epsilon_{opt}(x_i), \|x_i\|)$ (in red) in Fig. 2.

From Fig. 2, we observe that the bounds are valid, as they lie uniformly above the true distance, so (28) holds. The bounds for SDP appear to be looser compared to the bounds for LP and SOCP.

References

1. Boyd, S., Boyd, S.P., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
2. Lewis, A.: Nonsmooth optimization: conditioning, convergence and semialgebraic models. In: Proceedings of the International Congress of Mathematicians, Seoul, vol. 4, pp. 872–895 (2014)
3. Ding, L., Udell, M.: On the simplicity and conditioning of low rank semidefinite programs. arXiv preprint arXiv:2002.10673 (2020)
4. Drusvyatskiy, D., Lewis, A.S.: Error bounds, quadratic growth, and linear convergence of proximal methods. Math. Oper. Res. **43**(3), 919 (2018)
5. Zhou, Z., So, A.M.C.: A unified approach to error bounds for structured convex optimization problems. Math. Program. **165**(2), 689 (2017)
6. Necoara, I., Nesterov, Y., Glineur, F.: Linear convergence of first order methods for non-strongly convex optimization. Math. Program. **175**(1–2), 69 (2019)
7. Johnstone, P.R., Moulin, P.: Faster Subgradient Methods for Functions with Hölderian Growth. arXiv preprint arXiv:1704.00196 (2017)
8. Hoffman, A.J.: On approximate solutions of systems of linear inequalities. J. Res. Natl. Bur. Stand. **49**(4), 263–265 (1952)
9. Sturm, J.F.: Error bounds for linear matrix inequalities. SIAM J. Optim. **10**(4), 1228 (2000)
10. Nayakkankuppam, M.V., Overton, M.L.: Conditioning of semidefinite programs. Math. Program. **85**(3), 525 (1999)
11. Dür, M., Jargalsaikhan, B., Still, G.: Genericity results in linear conic programming—a tour d'horizon. Math. Oper. Res. **42**(1), 77 (2017)
12. Alizadeh, F., Haeberly, J.P.A., Overton, M.L.: Complementarity and nondegeneracy in semidefinite programming. Math. Program. **77**(1), 111 (1997)
13. Goldman, A.J., Tucker, A.W.: Theory of linear programming. In: Kuhn, H.W., Tucker, A.W.(eds.) Linear Inequalities and Related Systems, pp. 53–97. Princeton University Press, New Jersey (1956)
14. Bauschke, H.H., Borwein, J.M., Li, W.: Strong conical hull intersection property, bounded linear regularity, Jameson's property (G), and error bounds in convex optimization. Math. Program. **86**(1), 135 (1999)
15. Zhang, S.: Global error bounds for convex conic problems. SIAM J. Optim. **10**(3), 836 (2000)
16. Drusvyatskiy, D., Wolkowicz, H.: The many faces of degeneracy in conic optimization. arXiv preprint arXiv:1706.03705 (2017)
17. Borwein, J.M., Wolkowicz, H.: Facial reduction for a cone-convex programming problem. J. Aust. Math. Soc. **30**(3), 369 (1981)
18. Bandeira, A.S.: Random Laplacian matrices and convex relaxations. Found. Comput. Math. **18**(2), 345 (2018)
19. Ding, L., Yurtsever, A., Cevher, V., Tropp, J.A., Udell, M.: An optimal-storage approach to semidefinite programming using approximate complementarity. arXiv preprint arXiv:1902.03373 (2019)
20. Lourenço, B.F.: Amenable cones: error bounds without constraint qualifications. Math. Program. **186**(1), 1–48 (2021)
21. Drusvyatskiy, D., Ioffe, A.D., Lewis, A.S.: Generic minimizing behavior in semialgebraic optimization. SIAM J. Optim. **26**(1), 513 (2016)
22. Ruszczynski, A.: Nonlinear Optimization. Princeton University Press, Princeton (2011)

23. Klatte, D., Kummer, B.: Nonsmooth Equations in Optimization: Regularity, Calculus, Methods and Applications, vol. 60. Springer, Berlin (2006)
24. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer, Berlin (2013)