OPERATIONS RESEARCH CENTER

Working Paper

On an Extension of Condition Number Theory to Non-Conic Convex Optimization

by

Robert M. Freund and
Fernando Ordóñez

OR 365-03 **February 2003**

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

INFORMS © 2005

http://dx.doi.org/10.1287/moor.1040.0120

"On an Extension of Condition Number Theory to Nonconic Convex Optimization." Freund, Robert M. and Fernando Ordóñez. Mathematics of Operations Research Vol. 30, No. 1 (2005): 173-194.

On an Extension of Condition Number Theory to Non-Conic Convex Optimization

Robert M. Freund* and Fernando Ordóñez†

February 14, 2003

Abstract

The purpose of this paper is to extend, as much as possible, the modern theory of condition numbers for conic convex optimization:

$$\begin{array}{rll} z_* := & \min_x & c^t x \\ & \text{s.t.} & Ax - b \in C_Y \\ & & x \in C_X \ , \end{array}$$

to the more general non-conic format:

$$\begin{array}{rlll} & z_* := & \min_x & c^t x \\ (GP_d) & & \text{s.t.} & Ax - b \in C_Y \\ & & & x \in P \ , \end{array}$$

where P is any closed convex set, not necessarily a cone, which we call the ground-set. Although any convex problem can be transformed to conic form, such transformations are neither unique nor natural given the natural description of many problems, thereby diminishing the relevance of data-based condition number theory. Herein we extend the modern theory of condition numbers to the problem format (GP_d). As a byproduct, we are able to state and prove natural extensions of many theorems from the conic-based theory of condition numbers to this broader problem format.

Key words: Condition number, convex optimization, conic optimization, duality, sensitivity analysis, perturbation theory.

*MIT Sloan School of Management, 50 Memorial Drive, Cambridge, MA 02142, USA, email: rfreund@mit.edu

†Industrial and Systems Engineering, University of Southern California, GER-247, Los Angeles, CA 90089-0193, USA, email: fordon@usc.edu

1 Introduction

The modern theory of condition numbers for convex optimization problems was developed by Renegar in [16] and [17] for convex optimization problems in the following conic format:

$$(CP_d) \qquad \begin{array}{rll} z_* := & \min_x & c^t x \\ & \text{s.t.} & Ax - b \in C_Y \\ & & x \in C_X \ , \end{array} \tag{1}$$

where $C_X \subseteq \mathcal{X}$ and $C_Y \subseteq \mathcal{Y}$ are closed convex cones, A is a linear operator from the n-dimensional vector space $\mathcal{X}$ to the m-dimensional vector space $\mathcal{Y}$, $b \in \mathcal{Y}$, and $c \in \mathcal{X}^*$ (the space of linear functionals on $\mathcal{X}$). The data d for (CP_d) is defined as $d := (A, b, c)$.

The theory of condition numbers for (CP_d) focuses on three measures – $\rho_P(d), \rho_D(d)$, and $C(d)$, to bound various behavioral and computational quantities pertaining to (CP_d). The quantity $\rho_P(d)$ is called the "distance to primal infeasibility" and is the smallest data perturbation Δd for which $(CP_{d+\Delta d})$ is infeasible. The quantity $\rho_D(d)$ is called the "distance to dual infeasibility" for the conic dual (CD_d) of (CP_d):

$$(CD_d) \qquad \begin{array}{rll} z^* := & \max_y & b^t y \\ & \text{s.t.} & c - A^t y \in C_X^* \\ & & y \in C_Y^* \ , \end{array} \tag{2}$$

and is defined similarly to $\rho_P(d)$ but using the conic dual problem instead (which conveniently is of the same general conic format as the primal problem). The quantity $C(d)$ is called the "condition measure" or the "condition number" of the problem instance d and is a (positively) scale-invariant reciprocal of the smallest data perturbation Δd that will render the perturbed data instance either primal or dual infeasible:

$$C(d) := \frac{\|d\|}{\min\{\rho_P(d), \rho_D(d)\}} \ , \tag{3}$$

for a suitably defined norm $\|\cdot\|$ on the space of data instances d. A problem is called "ill-posed" if $\min\{\rho_P(d), \rho_D(d)\} = 0$, equivalently $C(d) = \infty$. These three condition measure quantities have been shown in theory to be connected to a wide variety of bounds on behavioral characteristics of (CP_d) and its dual, including bounds on sizes of feasible solutions, bounds on sizes of optimal solutions, bounds on optimal objective values, bounds on the sizes and aspect ratios of inscribed balls in the feasible region, bounds on the rate of deformation of the feasible region under perturbation, bounds on changes in optimal objective values under perturbation, and numerical bounds related to the linear algebra computations of certain algorithms, see [16], [5], [4], [6], [7], [8], [21], [19], [22], [20], [14], [15]. In the context of interior-point methods for linear and semidefinite optimization, these same three condition measures have also been shown to

be connected to various quantities of interest regarding the central trajectory, see [10] and [11]. The connection of these condition measures to the complexity of algorithms has been shown in [6], [7], [17], [2], and [3], and some of the references contained therein.

The conic format (CP_d) covers a very general class of convex problems; indeed any convex optimization problem can be transformed to an equivalent instance of (CP_d). However, such transformations are not necessarily unique and are sometimes rather unnatural given the "natural" description and the natural data for the problem. The condition number theory developed in the aforementioned literature pertains only to convex optimization problems in conic form, and the relevance of this theory is diminished to the extent that many practical convex optimization problems are not conveyed in conic format. Furthermore, the transformation of a problem to conic form can result in dramatically different condition numbers depending on the choice of transformation, see the example in Section 2 of [13].

Motivated to overcome these shortcomings, herein we extend the condition number theory to non-conic convex optimization problems. We consider the more general format for convex optimization:

$$\begin{array}{lrll} & z_*(d) = & \min & c^t x \\ (GP_d) & & \text{s.t.} & Ax - b \in C_Y \\ & & & x \in P \ , \end{array} \tag{4}$$

where P is allowed to be any closed convex set, possibly unbounded, and possibly without interior. For example, P could be the solution set of box constraints of the form $l \leq x \leq u$ where some components of l and/or u might be unbounded, or P might be the solution of network flow constraints of the form $Nx = g, x \geq 0$. And of course, P might also be a closed convex cone. We call P the ground-set and we refer to (GP_d) as the "ground-set model" (GSM) format.

We present the definition of the condition number for problem instances of the more general GSM format in Section 2, where we also demonstrate some basic properties. A number of results from condition number theory are extended to the GSM format in the subsequent sections of the paper. In Section 3 we prove that a problem instance with a finite condition number has primal and dual Slater points, which in turn implies that strong duality holds for the problem instance and its dual. In Section 4 we provide characterizations of the condition number as the solution to associated optimization problems. In Section 5 we show that if the condition number of a problem instance is finite, then there exist primal and dual interior solutions that have good geometric properties. In Section 6 we show that the rate of deformation of primal and dual feasible regions and optimal objective function values due to changes in the data are bounded by functions of the condition number. Section 7 contains concluding remarks.

We now present the notation and general assumptions that we will use throughout the paper.

Notation and General Assumptions. We denote the variable space $\mathcal{X}$ by $\mathbb{R}^n$ and the constraint space $\mathcal{Y}$ by $\mathbb{R}^m$. Therefore, $P \subseteq \mathbb{R}^n$, $C_Y \subseteq \mathbb{R}^m$, A is an m by n real matrix, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$. The spaces $\mathcal{X}^*$ and $\mathcal{Y}^*$ of linear functionals on $\mathbb{R}^n$ and $\mathbb{R}^m$ can be identified with $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. For $v, w \in \mathbb{R}^n$ or $\mathbb{R}^m$, we write $v^t w$ for the standard inner product. We denote by $\mathcal{D}$ the vector space of all data instances $d = (A, b, c)$. A particular data instance is denoted equivalently by d or (A, b, c). We define the norm for a data instance d by $\|d\| := \max\{\|A\|, \|b\|, \|c\|_*\}$, where the norms $\|x\|$ and $\|y\|$ on $\mathbb{R}^n$ and $\mathbb{R}^m$ are given, $\|A\|$ denotes the usual operator norm, and $\|\cdot\|_*$ denotes the dual norm associated with the norm $\|\cdot\|$ on $\mathbb{R}^n$ or $\mathbb{R}^m$, respectively. Let $B(v, r)$ denote the ball centered at v with radius r, using the norm for the space of variables v. For a convex cone S, let S^* denote the (positive) dual cone, namely $S^* := \{s \mid s^t x \geq 0 \text{ for all } x \in S\}$. Given a set $Q \subset \mathbb{R}^n$, we denote the closure and relative interior of Q by cl Q and relint Q, respectively. We use the convention that if Q is the singleton $Q = \{q\}$, then relint $Q = Q$. We adopt the standard conventions $\frac{1}{0} = \infty$ and $\frac{1}{\infty} = 0$.

We also make the following two general assumptions:

Assumption 1 *$P \neq \emptyset$ and $C_Y \neq \emptyset$.*

Assumption 2 *Either $C_Y \neq \mathbb{R}^m$ or P is not bounded (or both).*

Clearly if either $P = \emptyset$ or $C_Y = \emptyset$ problem (GP_d) is infeasible regardless of A, b, and c. Therefore Assumption 1 avoids settings wherein all problem instances are trivially inherently infeasible. Assumption 2 is needed to avoid settings where (GP_d) is feasible for every $d = (A, b, c) \in \mathcal{D}$. This will be explained further in Section 2.

2 Condition Numbers for (GP_d) and its Dual

2.1 Distance to Primal Infeasibility

We denote the feasible region of (GP_d) by:

$$X_d := \{x \in \mathbb{R}^n \mid Ax - b \in C_Y, x \in P\} \ . \tag{5}$$

Let $\mathcal{F}_P := \{d \in \mathcal{D} \mid X_d \neq \emptyset\}$, i.e., $\mathcal{F}_P$ is the set of data instances for which (GP_d) has a feasible solution. Similar to the conic case, the primal distance to infeasibility, denoted by $\rho_P(d)$, is defined as:

$$\rho_P(d) := \inf\{\|\Delta d\| \mid X_{d+\Delta d} = \emptyset\} = \inf\left\{\|\Delta d\| \mid d + \Delta d \in \mathcal{F}_P^C\right\} \ . \tag{6}$$

2.2 The Dual Problem and Distance to Dual Infeasibility

In the case when P is a cone, the conic dual problem (2) is of the same basic format as the primal problem. However, when P is not a cone, we must first develop a suitable dual problem, which we do in this subsection. Before doing so we introduce a dual pair of cones associated with the ground-set P. Define the closed convex cone C by homogenizing P to one higher dimension:

$$C := \text{cl } \{(x,t) \in I\!R^n \times I\!R \mid x \in tP,\ t > 0\} \ , \tag{7}$$

and note that $C = \{(x,t) \in I\!R^n \times I\!R \mid x \in tP,\ t > 0\} \cup (R \times \{0\})$ where R is the recession cone of P, namely

$$R := \{v \in I\!R^n \mid \text{there exists } x \in P \text{ for which } x + \theta v \in P \text{ for all } \theta \geq 0\} \ . \tag{8}$$

It is straightforward to show that the (positive) dual cone C^* of C is

$$\begin{array}{rcl} C^* & := & \{(s,u) \in I\!R^n \times I\!R \mid s^t x + u \cdot t \geq 0 \text{ for all } (x,t) \in C\} \\ & = & \{(s,u) \in I\!R^n \times I\!R \mid s^t x + u \geq 0, \text{ for all } x \in P\} \\ & = & \{(s,u) \in I\!R^n \times I\!R \mid \inf_{x \in P} s^t x + u \geq 0\} \ . \end{array} \tag{9}$$

The standard Lagrangian dual of (GP_d) can be constructed as:

$$\max_{y \in C_Y^*} \inf_{x \in P} \{c^t x + (b - Ax)^t y\}$$

which we re-write as:

$$\max_{y \in C_Y^*} \inf_{x \in P} \{b^t y + (c - A^t y)^t x\} \ . \tag{10}$$

With the help of (9) we re-write (10) as:

$$\begin{array}{rccl} & z^*(d) = & \max_{y,u} & b^t y - u \\ (GD_d) & & \text{s.t.} & (c - A^t y, u) \in C^* \\ & & & y \in C_Y^* \ . \end{array} \tag{11}$$

We consider the formulation (11) to be the dual problem of (4). The feasible region of (GD_d) is:

$$Y_d := \{(y,u) \in I\!R^m \times I\!R \mid (c - A^t y, u) \in C^*, y \in C_Y^*\}. \tag{12}$$

Let $\mathcal{F}_D := \{d \in \mathcal{D} \mid Y_d \neq \emptyset\}$, i.e., $\mathcal{F}_D$ is the set of data instances for which (GD_d) has a feasible solution. The dual distance to infeasibility, denoted by $\rho_D(d)$, is defined as:

$$\rho_D(d) := \inf \{\|\Delta d\| \mid Y_{d+\Delta d} = \emptyset\} = \inf \left\{\|\Delta d\| \mid d + \Delta d \in \mathcal{F}_D^C\right\} \ . \tag{13}$$

We also present an alternate form of (11), which does not use the auxiliary variable u, based on the function $u(\cdot)$ defined by

$$u(s) := -\inf_{x \in P} s^t x \ . \tag{14}$$

It follows from Theorem 5.5 in [18] that $u(\cdot)$, the support function of the set P, is a convex function. The epigraph of $u(\cdot)$ is:

$$\text{epi } u(\cdot) := \{(s, v) \in I\!R^n \times I\!R \mid v \geq u(s)\} \ ,$$

and the projection of the epigraph onto the space of the variables s is the effective domain of $u(\cdot)$:

$$\text{effdom } u(\cdot) := \{s \in I\!R^n \mid u(s) < \infty\} \ .$$

It then follows from (9) that

$$C^* \ = \ \text{epi } u(\cdot) \ ,$$

and so (GD_d) can alternatively be written as:

$$\begin{array}{rll} z^*(d) = & \max_y & b^t y - u(c - A^t y) \\ & \text{s.t.} & c - A^t y \in \text{effdom } u(\cdot) \\ & & y \in C_Y^* \ . \end{array} \tag{15}$$

Evaluating the inclusion $(y, u) \in Y_d$ is not necessarily an easy task, as it involves checking the inclusion $(c - A^t y, u) \in C^*$, and C^* is an implicitly defined cone. A very useful tool for evaluating the inclusion $(y, u) \in Y_d$ is given in the following proposition, where recall from (8) that R is the recession cone of P.

Proposition 1 *If y satisfies $y \in C_Y^*$ and $c - A^t y \in \text{relint} R^*$, then $u(c - A^t y)$ is finite, and for all $u \geq u(c - A^t y)$ it holds that (y, u) is feasible for (GD_d).*

Proof: Note from Proposition 11 of the Appendix that cl effdom $u(\cdot) = R^*$ and from Proposition 12 of the Appendix that $c - A^t y \in \text{relint} R^* = \text{relint cl effdom } u(\cdot) = \text{relint effdom } u(\cdot) \subseteq \text{ effdom } u(\cdot)$. This shows that $u(c - A^t y)$ is finite and $(c - A^t y, u(c - A^t y)) \in C^*$. Therefore (y, u) is feasible for (GD_d) for all $u \geq u(c - A^t y)$. ■

2.3 Condition Number

A data instance $d = (A, b, c)$ is consistent if both the primal and dual problems have feasible solutions. Let $\mathcal{F}$ denote the set of consistent data instances, namely $\mathcal{F} :=$

$\mathcal{F}_P \cap \mathcal{F}_D = \{d \in \mathcal{D} \mid X_d \neq \emptyset \text{ and } Y_d \neq \emptyset\}$. For $d \in \mathcal{F}$, the distance to infeasibility is defined as:

$$\begin{array}{rcl} \rho(d) & := & \min\{\rho_P(d), \rho_D(d)\} \\ & = & \inf\{\|\Delta d\| \mid X_{d+\Delta d} = \emptyset \text{ or } Y_{d+\Delta d} = \emptyset\} \ , \end{array} \tag{16}$$

the interpretation being that $\rho(d)$ is the size of the smallest perturbation of d which will render the perturbed problem instance either primal or dual infeasible. The condition number of the instance d is defined as

$$C(d) := \begin{cases} \dfrac{\|d\|}{\rho(d)} & \rho(d) > 0 \\ \infty & \rho(d) = 0 \ , \end{cases}$$

which is a (positive) scale-invariant reciprocal of the distance to infeasibility. This definition of condition number for convex optimization problems was first introduced by Renegar for problems in conic form in [16] and [17].

2.4 Basic Properties of $\rho_P(d), \rho_D(d)$, and $C(d)$, and Alternative Duality Results

The need for Assumptions 1 and 2 is demonstrated by the following:

Proposition 2 *For any data instance $d \in \mathcal{D}$,*

1. *$\rho_P(d) = \infty$ if and only if $C_Y = \mathbb{R}^m$.*
2. *$\rho_D(d) = \infty$ if and only if P is bounded.*

The proof of this proposition relies on Lemmas 1 and 2, which are versions of "theorems of the alternative" for primal and dual feasibility of (GP_d) and (GD_d). These two lemmas are stated and proved at the end of this section.

Proof of Proposition 2: Clearly $C_Y = \mathbb{R}^m$ implies that $\rho_P(d) = \infty$. Also, if P is bounded, then $R = \{0\}$ and $R^* = \mathbb{R}^n$, whereby from Proposition 1 we have that (GD_d) is feasible for any d, and so $\rho_D(d) = \infty$. Therefore for both items it only remains to prove the converse implication. Recall that we denote $d = (A, b, c)$.

Assume that $\rho_P(d) = \infty$, and suppose that $C_Y \neq \mathbb{R}^m$. Then $C_Y^* \neq \{0\}$, and consider a point $\tilde{y} \in C_Y^*$, $\tilde{y} \neq 0$. Define the perturbation $\Delta d = (\Delta A, \Delta b, \Delta c) = (-A, -b + \tilde{y}, -c)$ and $\bar{d} = d + \Delta d$. Then the point $(y, u) = \left(\tilde{y}, \frac{\tilde{y}^t \tilde{y}}{2}\right)$ satisfies the alternative system $(A2_{\bar{d}})$ of Lemma 1 for the data $\bar{d} = (0, \tilde{y}, 0)$, whereby $X_{\bar{d}} = \emptyset$. Therefore $\|\bar{d} - d\| \geq \rho_P(d) = \infty$, a contradiction, and so $C_Y = \mathbb{R}^m$.

Now assume that $\rho_D(d) = \infty$, and suppose that P is not bounded, and so $R \neq \{0\}$. Consider $\tilde{x} \in R$, $\tilde{x} \neq 0$, and define the perturbation $\Delta d = (-A, -b, -c - \tilde{x})$. Then the point $\tilde{x}$ satisfies the alternative system $(B2_{\bar{d}})$ of Lemma 2 for the data $\bar{d} = d + \Delta d = (0, 0, -\tilde{x})$, whereby $Y_{\bar{d}} = \emptyset$. Therefore $\|\bar{d} - d\| \geq \rho_D(d) = \infty$, a contradiction, and so P is bounded. ■

Remark 1 *If $d \in \mathcal{F}$, then $C(d) \geq 1$.*

Proof: Consider the data instance $d_0 = (0, 0, 0)$. Note that $X_{d_0} = P \neq \emptyset$ and $Y_{d_0} = C_Y^* \times I\!R_+ \neq \emptyset$, therefore $d_0 \in \mathcal{F}$. If $C_Y \neq I\!R^m$, consider $b \in I\!R^m \setminus C_Y$, $b \neq 0$, and for any $\varepsilon > 0$ define the instance $d_\varepsilon = (0, -\varepsilon b, 0)$. This instance is such that for any $\varepsilon > 0$, $X_{d_\varepsilon} = \emptyset$, which means that $d_\varepsilon \in \mathcal{F}_P^C$ and therefore $\rho_P(d) \leq \inf_{\varepsilon > 0} \|d - d_\varepsilon\| \leq \|d\|$. If $C_Y = I\!R^m$, then Assumption 2 implies that P is unbounded. This means that there exists a ray $r \in R$, $r \neq 0$. For any $\varepsilon > 0$ the instance $d_\varepsilon = (0, 0, -\varepsilon r)$ is such that $Y_{d_\varepsilon} = \emptyset$, which means that $d_\varepsilon \in \mathcal{F}_D^C$ and therefore $\rho_D(d) \leq \inf_{\varepsilon > 0} \|d - d_\varepsilon\| \leq \|d\|$.

In each case we have $\rho(d) = \min\{\rho_P(d), \rho_P(d)\} \leq \|d\|$, which implies the result. ■

The following two lemmas present weak and strong alternative results for (GP_d) and (GD_d), and are used in the proofs of Proposition 2 and elsewhere.

Lemma 1 *Consider the following systems with data $d = (A, b, c)$:*

$$(X_d) \quad \begin{array}{l} Ax - b \in C_Y \\ x \in P \end{array} , \qquad (A1_d) \quad \begin{array}{l} (-A^t y, u) \in C^* \\ b^t y \geq u \\ y \neq 0 \\ y \in C_Y^* \end{array} , \qquad (A2_d) \quad \begin{array}{l} (-A^t y, u) \in C^* \\ b^t y > u \\ y \in C_Y^* \end{array}$$

If system (X_d) is infeasible, then system $(A1_d)$ is feasible. Conversely, if system $(A2_d)$ is feasible, then system (X_d) is infeasible.

Proof: Assume that system (X_d) is infeasible. This implies that

$$b \notin S := \{Ax - v \mid x \in P, v \in C_Y\} ,$$

which is a nonempty convex set. Using Proposition 10 we can separate b from S and therefore there exists $y \neq 0$ such that

$$y^t(Ax - v) \leq y^t b \ \text{ for all } x \in P, \ v \in C_Y .$$

Set $u := y^t b$, then the inequality implies that $y \in C_Y^*$ and that $(-A^t y)^t x + u \geq 0$ for any $x \in P$. Therefore $(-A^t y, u) \in C^*$ and (y, u) satisfies system $(A1_d)$.

Conversely, if both $(A2_d)$ and (X_d) are feasible then

$$0 \leq y^t(Ax - b) = (A^t y)^t x - b^t y < -\left((-A^t y)^t x + u\right) \leq 0 . \qquad ■$$

Lemma 2 *Consider the following systems with data $d = (A, b, c)$:*

$$(Y_d) \quad \begin{array}{l} (c - A^t y, u) \in C^* \\ y \in C_Y^* \end{array} , \qquad (B1_d) \quad \begin{array}{l} Ax \in C_Y \\ c^t x \leq 0 \\ x \neq 0 \\ x \in R \end{array} , \qquad (B2_d) \quad \begin{array}{l} Ax \in C_Y \\ c^t x < 0 \\ x \in R \end{array}$$

If system (Y_d) is infeasible, then system $(B1_d)$ is feasible. Conversely, if system $(B2_d)$ is feasible, then system (Y_d) is infeasible.

Proof: Assume that system (Y_d) is infeasible, this implies that

$$(0, 0, 0) \notin S := \left\{ (s, v, q) \mid \exists\ y, u \text{ s.t. } (c - A^t y, u) + (s, v) \in C^*, y + q \in C_Y^* \right\} ,$$

which is a nonempty convex set. Using Proposition 10 we separate the point $(0, 0, 0)$ from S and therefore there exists $(x, \delta, z) \neq 0$ such that $x^t s + \delta v + z^t q \geq 0$ for all $(s, v, q) \in S$. For any (y, u), $(\tilde{s}, \tilde{v}) \in C^*$, and $\tilde{q} \in C_Y^*$, define $s = -(c - A^t y) + \tilde{s}$, $v = -u + \tilde{v}$, and $q = -y + \tilde{q}$. By construction $(s, v, q) \in S$ and therefore for any y, u, $(\tilde{s}, \tilde{v}) \in C^*$, $\tilde{q} \in C_Y^*$ we have

$$-x^t c + (Ax - z)^t y + x^t \tilde{s} - \delta u + \delta \tilde{v} + z^t \tilde{q} \geq 0 .$$

The above inequality implies that $\delta = 0$, $Ax = z \in C_Y$, $x \in R$, and $c^t x \leq 0$. In addition $x \neq 0$, because otherwise $(x, \delta, z) = (x, 0, Ax) = 0$. Therefore $(B1_d)$ is feasible.

Conversely, if both $(B2_d)$ and (Y_d) are feasible then

$$0 \leq x^t (c - A^t y) = c^t x - y^t A x < -y^t A x \leq 0 . \qquad \blacksquare$$

3 Slater Points, Distance to Infeasibility, and Strong Duality

In this section we prove that the existence of a Slater point in either (GP_d) or (GD_d) is sufficient to guarantee that strong duality holds for these problems. We then show that a positive distance to infeasibility implies the existence of Slater points, and use these results to show that strong duality holds whenever $\rho_P(d) > 0$ or $\rho_D(d) > 0$. We first state a weak duality result.

Proposition 3 *Weak duality holds between (GP_d) and (GD_d), that is, $z^*(d) \leq z_*(d)$.*

Proof: Consider x and (y, u) feasible for (GP_d) and (GD_d), respectively. Then

$$0 \leq \left(c - A^t y\right)^t x + u = c^t x - y^t A x + u \leq c^t x - b^t y + u,$$

where the last inequality follows from $y^t(Ax - b) \geq 0$. Therefore $z_*(d) \geq z^*(d)$. ■

A classic constraint qualification in the history of constrained optimization is the existence of a Slater point in the feasible region, see for example Theorem 30.4 of [18] or Chapter 5 of [1]. We now define a Slater point for problems in the GSM format.

Definition 1 *A point x is a Slater point for problem (GP_d) if*

$$x \in \text{relint}P \quad \text{and} \quad Ax - b \in \text{relint}C_Y \ .$$

A point (y, u) is a Slater point for problem (GD_d) if

$$y \in \text{relint}C_Y^* \quad \text{and} \quad (c - A^t y, u) \in \text{relint}C^* \ .$$

We now present the statements of the main results of this section, deferring the proofs to the end of the section. The following two theorems show that the existence of a Slater point in the primal or dual is sufficient to guarantee strong duality as well as attainment in the dual or the primal problem, respectively.

Theorem 1 *If x' is a Slater point for problem (GP_d), then $z_*(d) = z^*(d)$. If in addition $z_*(d) > -\infty$, then $Y_d \neq \emptyset$ and problem (GD_d) attains its optimum.*

Theorem 2 *If (y', u') is a Slater point for problem (GD_d), then $z_*(d) = z^*(d)$. If in addition $z^*(d) < \infty$, then $X_d \neq \emptyset$ and problem (GP_d) attains its optimum.*

The next three results show that a positive distance to infeasibility is sufficient to guarantee the existence of Slater point for the primal and the dual problems, respectively, and hence is sufficient to ensure that strong duality holds. The fact that a positive distance to infeasibility implies the existence of an interior point in the feasible region is shown for the conic case in Theorems 15, 17, and 19 in [8] and Theorem 3.1 in [17].

Theorem 3 *Suppose that $\rho_P(d) > 0$. Then there exists a Slater point for (GP_d).*

Theorem 4 *Suppose that $\rho_D(d) > 0$. Then there exists a Slater point for (GD_d).*

Corollary 1 (Strong Duality) *If $\rho_P(d) > 0$ or $\rho_D(d) > 0$, then $z_*(d) = z^*(d)$. If $\rho(d) > 0$, then both the primal and the dual attain their respective optimal values.*

Proof: The proof of this result is a straightforward consequence of Theorems 1, 2, 3, and 4. ■

Note that the contrapositive of Corollary 1 says that if $d \in \mathcal{F}$ and $z_*(d) > z^*(d)$, then $\rho_P(d) = \rho_D(d) = 0$ and so $\rho(d) = 0$. In other words, if a data instance d is primal and dual feasible but has a positive optimal duality gap, then d must necessarily be arbitrarily close to being both primal infeasible and dual infeasible.

Proof of Theorem 1: For simplicity, let z_* and z^* denote the primal and dual optimal objective values, respectively. The interesting case is when $z_* > -\infty$, otherwise weak duality implies that (GD_d) is infeasible and $z_* = z^* = -\infty$. If $z_* > -\infty$ the point $(0, 0, 0)$ does not belong to the non-empty convex set

$$S := \left\{(p, q, \alpha) \mid \exists x \text{ s.t. } x + p \in P, Ax - b + q \in C_Y, c^t x - \alpha < z_*\right\}.$$

We use Proposition 10 to properly separate $(0, 0, 0)$ from S, which implies that there exists $(\gamma, y, \pi) \neq 0$ such that $\gamma^t p + y^t q + \pi\alpha \geq 0$ for all $(p, q, \alpha) \in S$. Note that $\pi \geq 0$ because α is not upper bounded in the definition of S.

If $\pi > 0$, re-scale (γ, y, π) such that $\pi = 1$. For any $x \in I\!R^n$, $\tilde{p} \in P$, $\tilde{q} \in C_Y$, and $\varepsilon > 0$ define $p = -x + \tilde{p}$, $q = -Ax + b + \tilde{q}$, and $\alpha = c^t x - z_* + \varepsilon$. By construction the point $(p, q, \alpha) \in S$ and the proper separation implies that for all x, $\tilde{p} \in P$, $\tilde{q} \in C_Y$, and $\varepsilon > 0$

$$\begin{aligned} 0 &\leq \gamma^t(-x + \tilde{p}) + y^t(-Ax + b + \tilde{q}) + c^t x - z_* + \varepsilon \\ &= (-A^t y + c - \gamma)^t x + \gamma^t \tilde{p} + y^t \tilde{q} + y^t b - z_* + \varepsilon \ . \end{aligned}$$

This expression implies that $c - A^t y = \gamma$, $y \in C_Y^*$, and $(c - A^t y, u) \in C^*$ for $u := y^t b - z_*$. Therefore (y, u) is feasible for (GD_d) and $z^* \geq b^t y - u = b^t y - y^t b + z_* = z_* \geq z^*$, which implies that $z^* = z_*$ and the dual feasible point (y, u) attains the dual optimum.

If $\pi = 0$, the same construction used above and proper separation gives the following inequality for all x, $\tilde{p} \in P$, and $\tilde{q} \in C_Y$

$$\begin{aligned} 0 &\leq \gamma^t(-x + \tilde{p}) + y^t(-Ax + b + \tilde{q}) \\ &= (-A^t y - \gamma)^t x + \gamma^t \tilde{p} + y^t \tilde{q} + y^t b \ . \end{aligned}$$

This implies that $-A^t y = \gamma$ and $y \in C_Y^*$, which implies that $-y^t A\tilde{p} + y^t \tilde{q} + y^t b \geq 0$ for any $\tilde{p} \in P$, $\tilde{q} \in C_Y$. Proper separation also guarantees that there exists $(\hat{p}, \hat{q}, \hat{\alpha}) \in S$ such that $\gamma^t \hat{p} + y^t \hat{q} + \pi\hat{\alpha} = -y^t A\hat{p} + y^t \hat{q} > 0$.

Let x' be the Slater point of (GP_d) and $\hat{x}$ such that $\hat{x} + \hat{p} \in P$, $A\hat{x} - b + \hat{q} \in C_Y$, and $c^t \hat{x} - \hat{\alpha} < z_*$ For all $|\xi|$ sufficiently small, $x' + \xi(\hat{x} + \hat{p} - x') \in P$ and $Ax' - b + \xi(A\hat{x} - b + \hat{q} - (Ax' - b)) \in C_Y$. Therefore

$$\begin{aligned} 0 &\leq -y^t A\left(x' + \xi(\hat{x} + \hat{p} - x')\right) + y^t\left(Ax' - b + \xi(A\hat{x} - b + \hat{q} - (Ax' - b))\right) + y^t b \\ &= \xi\left(-y^t A\hat{x} - y^t A\hat{p} + y^t Ax' + y^t A\hat{x} - y^t b + y^t \hat{q} - y^t Ax' + y^t b\right) \\ &= \xi\left(-y^t A\hat{p} + y^t \hat{q}\right) \ , \end{aligned}$$

a contradiction, since ξ can be negative and $-y^t A\hat{p} + y^t\hat{q} > 0$. Therefore $\pi \neq 0$, completing the proof. ■

Proof of Theorem 2: For simplicity, let z_* and z^* denote the primal and dual optimal objective values respectively. The interesting case is when $z^* < \infty$, otherwise weak duality implies that (GP_d) is infeasible and $z_* = z^* = \infty$. If $z^* < \infty$ the point $(0,0,0,0)$ does not belong to the non-empty convex set

$$S := \left\{(s,v,q,\alpha) \mid \exists y, u \text{ s.t. } (c - A^t y, u) + (s,v) \in C^*, y + q \in C_Y^*, b^t y - u + \alpha > z^*\right\} .$$

We use Proposition 10 to properly separate $(0,0,0,0)$ from S, which implies that there exists $(x,\beta,\gamma,\delta) \neq 0$ such that $x^t s + \beta v + \gamma^t q + \delta\alpha \geq 0$ for all $(s,v,q,\alpha) \in S$. Note that $\delta \geq 0$ because α is not upper bounded in the definition of S.

If $\delta > 0$, re-scale (x,β,γ,δ) such that $\delta = 1$. For any $y \in I\!R^m$, $u \in I\!R$, $(\tilde{s},\tilde{v}) \in C^*$, $\tilde{q} \in C_Y^*$, and $\varepsilon > 0$, define $s = -c + A^t y + \tilde{s}$, $v = -u + \tilde{v}$, $q = -y + \tilde{q}$, and $\alpha = z^* - b^t y + u + \varepsilon$. By construction the point $(s,v,q,\alpha) \in S$ and proper separation implies that for all y, u, $(\tilde{s},\tilde{v}) \in C^*$, $\tilde{q} \in C_Y^*$, and $\varepsilon > 0$

$$\begin{aligned} 0 &\leq x^t(-c + A^t y + \tilde{s}) + \beta(-u + \tilde{v}) + \gamma^t(-y + \tilde{q}) + z^* - b^t y + u + \varepsilon \\ &= (Ax - b - \gamma)^t y + (x,\beta)^t(\tilde{s},\tilde{v}) + (1-\beta)u + \gamma^t\tilde{q} - c^t x + z^* + \varepsilon . \end{aligned}$$

This implies that $Ax - b = \gamma \in C_Y$, $\beta = 1$, $c^t x \leq z^*$, and $(x,1) \in C$, which means that $x \in P$. Therefore x is feasible for (GP_d) and $z^* \geq c^t x \geq z_* \geq z^*$, which implies that $z^* = z_*$ and the primal feasible point x attains the optimum.

If $\delta = 0$, the same construction used above and proper separation gives the following inequality for all y, u, $(\tilde{s},\tilde{v}) \in C^*$, $\tilde{q} \in C_Y^*$

$$\begin{aligned} 0 &\leq x^t(-c + A^t y + \tilde{s}) + \beta(-u + \tilde{v}) + \gamma^t(-y + \tilde{q}) \\ &= (Ax - \gamma)^t y + (x,\beta)^t(\tilde{s},\tilde{v}) - \beta u + \gamma^t\tilde{q} - c^t x . \end{aligned}$$

This implies that $Ax = \gamma \in C_Y$, $\beta = 0$, which means that $x^t\tilde{s} + x^t A^t\tilde{q} - c^t x \geq 0$ for any $(\tilde{s},\tilde{u}) \in C^*$ and $\tilde{q} \in C_Y^*$. The proper separation also guarantees that there exists $(\hat{s},\hat{v},\hat{q},\hat{\alpha}) \in S$ such that $x^t\hat{s} + \beta\hat{v} + \gamma^t\hat{q} = x^t\hat{s} + x^t A^t\hat{q} > 0$.

Let (y',u') be the Slater point of (GD_d) and $(\hat{y},\hat{u})$ such that $(c - A^t\hat{y} + \hat{s}, \hat{u} + \hat{v}) \in C^*$, $\hat{y} + \hat{q} \in C_Y^*$, and $b^t\hat{y} - \hat{u} + \hat{\alpha} > z^*$. Then for all $|\xi|$ sufficiently small, we have that $y' + \xi(\hat{y} + \hat{q} - y') \in C_Y^*$ and

$$\left(c - A^t y' + \xi\left(c - A^t\hat{y} + \hat{s} - c + A^t y'\right), u' + \xi(\hat{u} + \hat{v} - u')\right) \in C^* .$$

Therefore

$$x^t\left(c - A^t y' + \xi\left(c - A^t\hat{y} + \hat{s} - c + A^t y'\right)\right) + x^t A^t(y' + \xi(\hat{y} + \hat{q} - y')) - c^t x \geq 0 .$$

Simplifying and canceling, we obtain

$$\begin{aligned} 0 &\leq \xi\left(-x^t A^t \hat{y} + x^t \hat{s} + x^t A^t y' + x^t A^t \hat{y} + x^t A^t \hat{q} - x^t A^t y'\right) \\ &= \xi\left(x^t \hat{s} + x^t A^t \hat{q}\right) , \end{aligned}$$

a contradiction, since ξ can be negative and $x^t \hat{s} + x^t A^t \hat{q} > 0$. Therefore $\delta \neq 0$, completing the proof. ■

Proof of Theorem 3: Equation (6) and $\rho_P(d) > 0$ imply that $X_d \neq \emptyset$. Assume that X_d contains no Slater point, then relint$C_Y \cap \{Ax - b \mid x \in \text{relint}P\} = \emptyset$ and these nonempty convex sets can be separated using Proposition 10. Therefore there exists $y \neq 0$ such that for any $s \in C_Y$, $x \in P$ we have

$$y^t s \geq y^t (Ax - b) .$$

Let $u = y^t b$; from the inequality above we have that $y \in C_Y^*$ and $-y^t Ax + u \geq 0$ for any $x \in P$, which implies that $(-A^t y, u) \in C^*$. Define $b_\varepsilon = b + \frac{\varepsilon}{\|y\|_*}\hat{y}$, with $\hat{y}$ given by Proposition 9 such that $\|\hat{y}\| = 1$ and $\hat{y}^t y = \|y\|_*$. Then the point (y, u) is feasible for Problem $(A2_{d_\varepsilon})$ of Lemma 1 with data $d_\varepsilon = (A, b_\varepsilon, c)$ for any $\varepsilon > 0$. This implies that $X_{d_\varepsilon} = \emptyset$ and therefore $\rho_P(d) \leq \inf_{\varepsilon > 0} \|d - d_\varepsilon\| = \inf_{\varepsilon > 0} \frac{\varepsilon}{\|y\|_*} = 0$, a contradiction. ■

Proof of Theorem 4: Equation (13) and $\rho_D(d) > 0$ imply that $Y_d \neq \emptyset$. Assume that Y_d contains no Slater point. Consider the nonempty convex set S defined by:

$$S := \left\{\left(c - A^t y, u\right) \mid y \in \text{relint}C_Y^*, u \in I\!R\right\} .$$

No Slater point in the dual implies that relint$C^* \cap S = \emptyset$. Therefore we can properly separate these two nonempty convex sets using Proposition 10, whereby there exists $(x, t) \neq 0$ such that for any $(s, v) \in C^*$, $y \in C_Y^*$, $u \in I\!R$ we have

$$x^t s + tv \geq x^t \left(c - A^t y\right) + tu .$$

The above inequality implies that $Ax \in C_Y$, $c^t x \leq 0$, $(x, t) \in C$, and $t = 0$. This last fact implies that $x \neq 0$ and $x \in R$. Let $\hat{x}$ be such that $\|\hat{x}\|_* = 1$ and $\hat{x}^t x = \|x\|$ (see Proposition 9). For any $\varepsilon > 0$, define $c_\varepsilon = c - \frac{\varepsilon}{\|x\|}\hat{x}$. Then the point x is feasible for Problem $(B2_{d_\varepsilon})$ of Lemma 2 with data $d_\varepsilon = (A, b, c_\varepsilon)$. This implies then that $Y_{d_\varepsilon} = \emptyset$ and consequently $\rho_D(d) \leq \inf_{\varepsilon > 0} \|d - d_\varepsilon\| = \inf_{\varepsilon > 0} \frac{\varepsilon}{\|x\|} = 0$, a contradiction. ■

The contrapositives of Theorems 3 and 4 are not true. Consider for example the data

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} , \; b = \begin{pmatrix} -1 \\ 0 \end{pmatrix} , \text{ and } c = \begin{pmatrix} 1 \\ 0 \end{pmatrix} ,$$

and the sets $C_Y = I\!R_+ \times \{0\}$ and $P = C_X = I\!R_+ \times I\!R$. Problem (GP_d) for this example has a Slater point at $(1, 0)$, and $\rho_P(d) = 0$ (perturbing by $\Delta b = (0, \varepsilon)$ makes the problem infeasible for any ε.) Problem (GD_d) for the same example has a Slater point at $(1, 0)$ and $\rho_D(d) = 0$ (perturbing by $\Delta c = (0, \varepsilon)$ makes the problem infeasible for any ε).

4 Characterization of $\rho_P(d)$ and $\rho_D(d)$ via Associated Optimization Problems

Equation (16) shows that to characterize $\rho(d)$ for consistent data instances $d \in \mathcal{F}$, it is sufficient to express $\rho_P(d)$ and $\rho_D(d)$ in a convenient form. Below we show that these distances to infeasibility can be obtained as the solutions of certain associated optimization problems. These results can be viewed as an extension to problems not in conic form of Theorem 3.5 of [17], and Theorems 1 and 2 of [8].

Theorem 5 *Suppose that $X_d \neq \emptyset$. Then $\rho_P(d) = j_P(d) = r_P(d)$, where*

$$\begin{array}{rll} j_P(d) = & \min & \max\{\|A^t y + s\|_*, |b^t y - u|\} \\ & \|y\|_* = 1 & \\ & y \in C_Y^* & \\ & (s,u) \in C^* & \end{array} \tag{17}$$

and

$$\begin{array}{rll} r_P(d) = & \min & \max \quad \theta \\ & \|v\| \leq 1 & Ax - bt - v\theta \in C_Y \\ & v \in I\!R^m & \|x\| + |t| \leq 1 \\ & & (x,t) \in C \ . \end{array} \tag{18}$$

Theorem 6 *Suppose that $Y_d \neq \emptyset$. Then $\rho_D(d) = j_D(d) = r_D(d)$, where*

$$\begin{array}{rll} j_D(d) = & \min & \max\{\|Ax - p\|, |c^t x + g|\} \\ & \|x\| = 1 & \\ & x \in R & \\ & p \in C_Y & \\ & g \geq 0 & \end{array} \tag{19}$$

and

$$\begin{array}{rll} r_D(d) = & \min & \max \quad \theta \\ & \|v\|_* \leq 1 & -A^t y + c\delta - \theta v \in R^* \\ & v \in I\!R^n & \|y\|_* + |\delta| \leq 1 \\ & & y \in C_Y^* \\ & & \delta \geq 0 \ . \end{array} \tag{20}$$

Proof of Theorem 5: Assume that $j_P(d) > \rho_P(d)$, then there exists a data instance $\bar{d} = (\bar{A}, \bar{b}, \bar{c})$ that is primal infeasible and $\|A - \bar{A}\| < j_P(d)$, $\|b - \bar{b}\| < j_P(d)$, and

$\|c - \bar{c}\|_* < j_P(d)$. From Lemma 1 there is a point $(\bar{y}, \bar{u})$ that satisfies the following:

$$\begin{array}{l} (-\bar{A}^t\bar{y}, \bar{u}) \in C^* \\ \bar{b}^t\bar{y} \geq \bar{u} \\ \bar{y} \neq 0 \\ \bar{y} \in C_Y^*. \end{array}$$

Scale $\bar{y}$ such that $\|\bar{y}\|_* = 1$, then $(y, s, u) = (\bar{y}, -\bar{A}^t\bar{y}, \bar{b}^t\bar{y})$ is feasible for (17) and

$$\begin{array}{rcl} \|A^t y + s\|_* & = & \|A^t\bar{y} - \bar{A}^t\bar{y}\|_* \leq \|A - \bar{A}\|\|\bar{y}\|_* < j_P(d) \\ |b^t y - u| & = & |b^t\bar{y} - \bar{b}^t\bar{y}| \leq \|b - \bar{b}\|\|\bar{y}\|_* < j_P(d). \end{array}$$

In the first inequality above we used the fact that $\|A^t\|_* = \|A\|$. Therefore $j_P(d) \leq \max\{\|A^t y + s\|_*, |b^t y - u|\} < j_P(d)$, a contradiction.

Let us now assume that $j_P(d) < \gamma < \rho_P(d)$ for some γ. This means that there exists $(\bar{y}, \bar{s}, \bar{u})$ such that $\bar{y} \in C_Y^*$, $\|\bar{y}\|_* = 1$, $(\bar{s}, \bar{u}) \in C^*$, and that

$$\|A^t\bar{y} + \bar{s}\|_* < \gamma, \quad |b^t\bar{y} - \bar{u}| < \gamma.$$

From Proposition 9, consider $\hat{y}$ such that $\|\hat{y}\| = 1$ and $\hat{y}^t\bar{y} = \|\bar{y}\|_* = 1$, and define, for $\varepsilon > 0$,

$$\begin{array}{l} \bar{A} = A - \hat{y}\left((A^t\bar{y})^t + \bar{s}^t\right) \\ \bar{b}_\varepsilon = b - \hat{y}\left(b^t\bar{y} - \bar{u} - \varepsilon\right). \end{array}$$

We have that $\bar{y} \in C_Y^*$, $-\bar{A}^t\bar{y} = \bar{s}$, $\bar{b}_\varepsilon^t\bar{y} = \bar{u} + \varepsilon > \bar{u}$, and $(-\bar{A}^t\bar{y}, \bar{u}) \in C^*$. This implies that for any $\varepsilon > 0$, the problem $(A2_{\bar{d}_\varepsilon})$ in Lemma 1 is feasible with data $\bar{d}_\varepsilon = (\bar{A}, \bar{b}_\varepsilon, c)$. Lemma 1 then implies that $X_{\bar{d}_\varepsilon} = \emptyset$ and therefore $\rho_P(d) \leq \|d - \bar{d}_\varepsilon\|$. To finish the proof we compute the size of the perturbation:

$$\begin{array}{l} \|A - \bar{A}\| = \|\hat{y}\left((A^t\bar{y})^t + \bar{s}^t\right)\| \leq \|A^t\bar{y} + \bar{s}\|_*\|\hat{y}\| < \gamma \\ \|b - \bar{b}_\varepsilon\| = |b^t\bar{y} - \bar{u} - \varepsilon|\|\hat{y}\| \leq |b^t\bar{y} - \bar{u}| + \varepsilon < \gamma + \varepsilon, \end{array}$$

which implies, $\rho_P(d) \leq \|d - \bar{d}_\varepsilon\| = \max\left\{\|A - \bar{A}\|, \|b - \bar{b}_\varepsilon\|\right\} < \gamma + \varepsilon < \rho_P(d)$, for ε small enough. This is a contradiction, whereby $j_P(d) = \rho_P(d)$.

To prove the other characterization, we note that $\theta \geq 0$ in Problem (18) and invoke Lemma 6 to rewrite it as

$$\begin{array}{rlll} r_P(d) = & \min & \min & \max\{\|A^t y + s\|_*, |-b^t y + u|\} \\ & \|v\| \leq 1 & y^t v \geq 1 & \\ & v \in I\!R^m & y \in C_Y^* & \\ & & (s, u) \in C^*. & \end{array}$$

The above problem can be written as the following equivalent optimization problem:

$$\begin{array}{rll} r_P(d) = & \min & \max\{\|A^t y + s\|_*, |-b^t y + u|\} \\ & \|y\|_* \geq 1 & \\ & y \in C_Y^* & \\ & (s,u) \in C^*. & \end{array}$$

The equivalence of these problems is verified by combining the minimization operations in the first problem and using the Cauchy-Schwartz inequality. The converse makes use of Proposition 9. To finish the proof, we note that if (y,s,u) is optimal for this last problem then it also satisfies $\|y\|_* = 1$, whereby making it equivalent to (17). Therefore

$$\begin{array}{rll} r_P(d) = & \min & \max\{\|A^t y + s\|_*, |-b^t y + u|\} \quad = j_P(d)\ . \\ & \|y\|_* = 1 & \\ & y \in C_Y^* & \\ & (s,u) \in C^* & \end{array} \qquad \blacksquare$$

Proof of Theorem 6: Assume that $j_D(d) > \rho_D(d)$, then there exists a data instance $\bar{d} = (\bar{A}, \bar{b}, \bar{c})$ that is dual infeasible and $\|A - \bar{A}\| < j_D(d)$, $\|b - \bar{b}\| < j_D(d)$, and $\|c - \bar{c}\|_* < j_D(d)$. From Lemma 2 there exists $\bar{x} \in R$ such that $\bar{x} \neq 0$, $\bar{A}\bar{x} \in C_Y$, and $\bar{c}^t\bar{x} \leq 0$. We can scale $\bar{x}$ such that $\|\bar{x}\| = 1$. Then $(x,p,g) = (\bar{x}, \bar{A}\bar{x}, -\bar{c}^t\bar{x})$ is feasible for (19), and

$$\begin{aligned} \|Ax - p\| &= \|Ax - \bar{A}x\| \leq \|A - \bar{A}\|\|x\| < j_D(d) \\ |c^t x + g| &= |c^t x - \bar{c}^t x| \leq \|c - \bar{c}\|_*\|x\| < j_D(d)\ . \end{aligned}$$

Therefore, $j_D(d) \leq \max\{\|Ax - p\|, |c^t x + g|\} < j_D(d)$, which is a contradiction.

Assume now that $j_D(d) < \gamma < \rho_D(d)$ for some γ. Then there exists $(\bar{x}, \bar{p}, \bar{g})$ such that $\bar{x} \in R$, $\|\bar{x}\| = 1$, $\bar{p} \in C_Y$, and $\bar{g} \geq 0$, and that $\|A\bar{x} - \bar{p}\| \leq \gamma$ and $|c^t\bar{x} + \bar{g}| \leq \gamma$. From Proposition 9, consider $\hat{x}$ such that $\|\hat{x}\|_* = 1$ and $\hat{x}^t\bar{x} = \|\bar{x}\| = 1$, and define: $\bar{A} = A - (A\bar{x} - \bar{p})\,\hat{x}^t$ and $\bar{c}_\varepsilon = c - \hat{x}(c^t\bar{x} + \bar{g} + \varepsilon)$, for $\varepsilon > 0$. By construction $\bar{A}\bar{x} = \bar{p} \in C_Y$ and $\bar{c}_\varepsilon^t\bar{x} = -\bar{g} - \varepsilon < 0$, for any $\varepsilon > 0$. Therefore Problem $(B2_{\bar{d}_\varepsilon})$ in Lemma 2 is feasible for data $\bar{d}_\varepsilon = (\bar{A}, b, \bar{c}_\varepsilon)$, which implies that $Y_{\bar{d}_\varepsilon} = \emptyset$. We can then bound $\rho_D(d)$ as follows:

$$\begin{aligned} \rho_D(d) &\leq \|d - \bar{d}_\varepsilon\| = \max\left\{\|\,(A\bar{x} - \bar{p})\,\hat{x}^t\|, \|\hat{x}(c^t\bar{x} + \bar{g} + \varepsilon)\|_*\right\} \\ &\leq \max\{\gamma, \gamma + \varepsilon\} = \gamma + \varepsilon < \rho_D(d) \end{aligned}$$

for ε small enough, which is a contradiction. Therefore $\rho_D(d) = j_D(d)$.

To prove the other characterization, we note that $\theta \geq 0$ in Problem (20) and invoke Lemma 6 to rewrite it as

$$\begin{array}{rlll} r_D(d) = & \min & \min & \max\{\|-Ax + p\|, |c^t x + g|\} \\ & \|v\|_* \leq 1 & x^t v \geq 1 & \\ & v \in I\!R^n & x \in R & \\ & & p \in C_Y & \\ & & g \geq 0\ . & \end{array}$$

The above problem can be written as the following equivalent optimization problem:

$$\begin{array}{rll} r_D(d) = & \min & \max\{\| - Ax + p\|, |c^t x + g|\} \\ & \|x\| \geq 1 & \\ & x \in R & \\ & p \in C_Y & \\ & g \geq 0 \ . & \end{array}$$

The equivalence of these problems is verified by combining the minimization operations in the first problem and using the Cauchy-Schwartz inequality. The converse makes use of Proposition 9. To finish the proof, we note that if (x, p, g) is optimal for this last problem then it also satisfies $\|x\| = 1$, whereby making it equivalent to (19). Therefore

$$\begin{array}{rll} r_D(d) = & \min & \max\{\| - Ax + p\|, |c^t x + g|\} \quad = j_D(d) \ . \\ & \|x\| = 1 & \\ & x \in R & \\ & p \in C_Y & \\ & g \geq 0 & \end{array} \qquad \blacksquare$$

5 Geometric Properties of the Primal and Dual Feasible Regions

In Section 3 we showed that a positive primal and/or dual distance to infeasibility implies the existence of a primal and/or dual Slater point, respectively. We now show that a positive distance to infeasibility also implies that the corresponding feasible region has a *reliable* solution. We consider a solution in the relative interior of the feasible region to be a reliable solution if it has good geometric properties: it is not too far from a given reference point, its distance to the relative boundary of the feasible region is not too small, and the ratio of these two quantities is not too large, where these quantities are bounded by appropriate condition numbers.

5.1 Distance to Relative Boundary, Minimum Width of Cone

An affine set T is the translation of a vector subspace L, i.e., $T = a + L$ for some a. The minimal affine set that contains a given set S is known as the affine hull of S. We denote the affine hull of S by L_S; it is characterized as:

$$L_S = \left\{ \sum_{i \in I} \alpha_i x_i \mid \alpha_i \in I\!R, \ x_i \in S, \ \sum_{i \in I} \alpha_i = 1, \ I \text{ a finite set} \right\} \ ,$$

see Section 1 in [18]. We denote by $\widehat{L}_S$ the vector subspace obtained when the affine hull L_S is translated to contain the origin; i.e. for any $x \in S$, $\widehat{L}_S = L_S - x$. Note that if $0 \in S$ then L_S is a subspace.

Many results in this section involve the distance of a point $x \in S$ to the relative boundary of the set S, denoted by $\text{dist}(x, \text{rel}\partial S)$, defined as follows:

Definition 2 *Given a non-empty set S and a point $x \in S$, the distance from x to the relative boundary of S is*

$$\begin{array}{rll} \text{dist}(x, \text{rel}\partial S) := & \inf_{\bar{x}} & \|x - \bar{x}\| \\ & \text{s.t.} & \bar{x} \in L_S \setminus S \ . \end{array} \tag{21}$$

Note that if S is an affine set (and in particular if S is the singleton $S = \{s\}$), then $\text{dist}(x, \text{rel}\partial S) = \infty$ for each $x \in S$.

We use the following definition of the min-width of a convex cone:

Definition 3 *For a convex cone K, the min-width of K is defined by*

$$\tau_K := \sup \left\{ \frac{\text{dist}(y, \text{rel}\partial K)}{\|y\|} \mid y \in K, y \neq 0 \right\} \ ,$$

for $K \neq \{0\}$, and $\tau_K := \infty$ if $K = \{0\}$.

The measure τ_K maximizes the ratio of the radius of a ball contained in the relative interior of K and the norm of its center, and so it intuitively corresponds to half of the vertex angle of the widest cylindrical cone contained in K. The quantity τ_K was called the "inner measure" of K for Euclidean norms in Goffin [9], and has been used more recently for general norms in analyzing condition measures for conic convex optimization, see [6]. Note that if K is not a subspace, then $\tau_K \in (0, 1]$, and τ_K is attained for some $y^0 \in \text{relint} K$ satisfying $\|y^0\| = 1$, as well as along the ray αy^0 for all $\alpha > 0$; and τ_K takes on larger values to the extent that K has larger minimum width. If K is a subspace, then $\tau_K = \infty$.

5.2 Geometric Properties of the Feasible Region of GP_d

In this subsection we present results concerning geometric properties of the feasible region X_d of (GP_d). We defer all proofs to the end of the subsection.

The following proposition is an extension of Lemma 3.2 of [16] to the ground-set model format.

Proposition 4 *Consider any $x = \hat{x} + r$ feasible for (GP_d) such that $\hat{x} \in P$ and $r \in R$. If $\rho_D(d) > 0$ then*

$$\|r\| \leq \frac{1}{\rho_D(d)} \max\left\{\|A\hat{x} - b\|, c^t r\right\} \ .$$

The following result is an extension of Assertion 1 of Theorem 1.1 of [16] to the ground-set model format of (GP_d):

Proposition 5 *Consider any $x^0 \in P$. If $\rho_P(d) > 0$ then there exists $\bar{x} \in X_d$ satisfying*

$$\|\bar{x} - x^0\| \leq \frac{\text{dist}(Ax^0 - b, C_Y)}{\rho_P(d)} \max\left\{1, \|x^0\|\right\} \ .$$

The following is the main result of this subsection, and can be viewed as an extension of Theorems 15, 17, and 19 of [8] to the ground-set model format of (GP_d). In Theorem 7 we assume for expository convenience that P is not an affine set and C_Y is not a subspace. These assumptions are relaxed in Theorem 8.

Theorem 7 *Suppose that P is not an affine set, C_Y is not a subspace, and consider any $x^0 \in P$. If $\rho_P(d) > 0$ then there exists $\bar{x} \in X_d$ satisfying:*

1. (a) $\|\bar{x} - x^0\| \leq \dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)} \max\{1, \|x^0\|\}$

   (b) $\|\bar{x}\| \leq \|x^0\| + \dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)}$

2. (a) $\dfrac{1}{\text{dist}(\bar{x}, \text{rel}\partial P)} \leq \dfrac{1}{\text{dist}(x^0, \text{rel}\partial P)} \left(1 + \dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)}\right)$

   (b) $\dfrac{1}{\text{dist}(\bar{x}, \text{rel}\partial X_d)} \leq \dfrac{1}{\min\left\{\text{dist}(x^0, \text{rel}\partial P), \tau_{C_Y}\right\}} \left(1 + \dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)}\right)$

3. (a) $\dfrac{\|\bar{x} - x^0\|}{\text{dist}(\bar{x}, \text{rel}\partial P)} \leq \dfrac{1}{\text{dist}(x^0, \text{rel}\partial P)} \left(\dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)} \max\{1, \|x^0\|\}\right)$

   (b) $\dfrac{\|\bar{x} - x^0\|}{\text{dist}(\bar{x}, \text{rel}\partial X_d)} \leq \dfrac{1}{\min\left\{\text{dist}(x^0, \text{rel}\partial P), \tau_{C_Y}\right\}} \left(\dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)} \max\{1, \|x^0\|\}\right)$

   (c) $\dfrac{\|\bar{x}\|}{\text{dist}(\bar{x}, \text{rel}\partial P)} \leq \dfrac{1}{\text{dist}(x^0, \text{rel}\partial P)} \left(\|x^0\| + \dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)}\right)$

$$(d)\quad \frac{\|\bar{x}\|}{\text{dist}(\bar{x}, \text{rel}\partial X_d)} \leq \frac{1}{\min\{\text{dist}(x^0, \text{rel}\partial P), \tau_{C_Y}\}} \left(\|x^0\| + \frac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)} \right)$$

The statement of Theorem 8 below relaxes the assumptions on P and C_Y not being affine and/or linear spaces:

Theorem 8 *Consider any $x^0 \in P$. If $\rho_P(d) > 0$ then there exists $\bar{x} \in X_d$ with the following properties:*

- *If P is not an affine set, $\bar{x}$ satisfies all items of Theorem 7.*
- *If P is an affine set and C_Y is not a subspace, $\bar{x}$ satisfies all items of Theorem 7, where items 2.(a), 3.(a), and 3.(c) are vacuously valid as both sides of these inequalities are zero.*
- *If P is an affine set and C_Y is a subspace, $\bar{x}$ satisfies all items of Theorem 7, where items 2.(a), 2.(b), 3.(a), 3.(b), 3.(c), and 3.(d) are vacuously valid as both sides of these inequalities are zero.*

We conclude this subsection by presenting a result which captures the thrust of Theorems 7 and 8, emphasizing how the distance to infeasibility $\rho_P(d)$ and the geometric properties of a given point $x^0 \in P$ bound various geometric properties of the feasible region X_d. For $x^0 \in P$, define the following measure:

$$g_{P,C_Y}(x^0) := \frac{\max\{\|x^0\|, 1\}}{\min\{1, \text{dist}(x^0, \text{rel}\partial P), \tau_{C_Y}\}} .$$

Also define the following geometric measure of the feasible region X_d:

$$g_{X_d} := \min_{x \in X_d} \max \left\{ \|x\|, \frac{\|x\|}{\text{dist}(x, \text{rel}\partial X_d)}, \frac{1}{\text{dist}(x, \text{rel}\partial X_d)} \right\} .$$

The following is an immediate consequence of Theorems 7 and 8.

Corollary 2 *Consider any $x^0 \in P$. If $\rho_P(d) > 0$ then*

$$g_{X_d} \leq g_{P,C_Y}(x^0) \left(1 + \frac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)} \right) . \qquad \blacksquare$$

We now proceed with proofs of these results.

Proof of Proposition 4: If $r = 0$ the result is true. If $r \neq 0$, then Proposition 9 shows that there exists $\hat{r}$ such that $\|\hat{r}\|_* = 1$ and $\hat{r}^t r = \|r\|$. For any $\varepsilon > 0$ define the following perturbed problem instance:

$$\bar{A} = A + \frac{1}{\|r\|}(A\hat{x} - b)\,\hat{r}^t, \qquad \bar{b} = b, \qquad \bar{c} = c + \frac{-(c^t r)^+ - \varepsilon}{\|r\|}\hat{r}\ .$$

Note that, for the data $\bar{d} = (\bar{A}, \bar{b}, \bar{c})$, the point r satisfies $(B2_{\bar{d}})$ in Lemma 2, and therefore $(GD_{\bar{d}})$ is infeasible. We conclude that $\rho_D(d) \leq \|d - \bar{d}\|$, which implies

$$\rho_D(d) \ \leq \ \frac{\max\{\|A\hat{x} - b\|, (c^t r)^+ + \varepsilon\}}{\|r\|}$$

and so

$$\rho_D(d) \ \leq \ \frac{\max\{\|A\hat{x} - b\|, c^t r\}}{\|r\|}\ . \qquad \blacksquare$$

The following technical lemma, which concerns the optimization problem (PP) below, is used in the subsequent proofs. Problem (PP) is parametrized by given points $x^0 \in P$ and $w^0 \in C_Y$, and is defined by

$$\begin{array}{lll} (PP) & \max_{x,t,w,\theta} & \theta \\ & \text{s.t.} & Ax - bt - w = \theta\,(b - Ax^0 + w^0) \\ & & \|x\| + |t| \leq 1 \\ & & (x,t) \in C \\ & & w \in C_Y. \end{array} \tag{22}$$

Lemma 3 *Consider any $x^0 \in P$ and $w^0 \in C_Y$ such that $Ax^0 - w^0 \neq b$. If $\rho_P(d) > 0$, then there exists a point (x, t, w, θ) feasible for problem (PP) that satisfies*

$$\theta \geq \frac{\rho_P(d)}{\|b - Ax^0 + w^0\|} > 0\ . \tag{23}$$

Proof: Note that problem (PP) is feasible for any x^0 and w^0 since $(x, t, w, \theta) = (0, 0, 0, 0)$ is always feasible, therefore it can either be unbounded or have a finite optimal objective value. If (PP) is unbounded, we can find feasible points with an objective function large enough such that (23) holds. If (PP) has a finite optimal value, say θ^*, then it follows from elementary arguments that it attains its optimal value. Since $\rho_P(d) > 0$ implies $X_d \neq \emptyset$, Theorem 5 implies that the optimal solution $(x^*, t^*, w^*, \theta^*)$ for (PP) satisfies (23). $\blacksquare$

Proof of Proposition 5: Assume $Ax^0 - b \notin C_Y$, otherwise $\bar{x} = x^0$ satisfies the proposition. We consider problem (PP), defined by (22), with x^0 and $w^0 \in C_Y$ such

that $\|Ax^0 - b - w^0\| = \text{dist}(Ax^0 - b, C_Y)$. From Lemma 3 we have that there exists a point (x, t, w, θ) feasible for (PP) that satisfies

$$\theta \geq \frac{\rho_P(d)}{\|b - Ax^0 + w^0\|} = \frac{\rho_P(d)}{\text{dist}(Ax^0 - b, C_Y)} .$$

Define

$$\bar{x} = \frac{x + \theta x^0}{t + \theta} \text{ and } \bar{w} = \frac{w + \theta w^0}{t + \theta} .$$

By construction we have $\bar{x} \in P$, $A\bar{x} - b = \bar{w} \in C_Y$, therefore $\bar{x} \in X_d$, and

$$\|\bar{x} - x^0\| = \frac{\|x - tx^0\|}{t + \theta} \leq \frac{(\|x\| + t)\max\{1, \|x^0\|\}}{\theta} \leq \frac{\text{dist}(Ax^0 - b, C_Y)}{\rho_P(d)} \max\left\{1, \|x^0\|\right\} .$$

■

Proof of Theorem 7: Note that $\rho_P(d) > 0$ implies $X_d \neq \emptyset$; note also that $\rho_P(d)$ is finite, otherwise Proposition 2 shows that $C_Y = I\!R^m$ which is a subspace. Set $w^0 \in C_Y$ such that $\|w^0\| = \|A\|$ and $\tau_{C_Y} = \frac{\text{dist}(w^0, \text{rel}\partial C_Y)}{\|w^0\|}$. We also assume that $Ax^0 - b \neq w^0$, otherwise we can show that $\bar{x} = x^0$ satisfies the theorem. Let $r_{w^0} = \text{dist}(w^0, \text{rel}\partial C_Y) = \|A\|\tau_{C_Y}$ and let also $r_{x^0} = \text{dist}(x^0, \text{rel}\partial P)$. We invoke Lemma 3 with x^0 and w^0 above to obtain a point (x, t, w, θ), feasible for (PP) and that from inequality (23) satisfies

$$0 < \frac{1}{\theta} \leq \frac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)} . \tag{24}$$

Define the following:

$$\bar{x} = \frac{x + \theta x^0}{t + \theta}, \quad \bar{w} = \frac{w + \theta w^0}{t + \theta}, \quad r_{\bar{x}} = \frac{\theta r_{x^0}}{t + \theta}, \quad r_{\bar{w}} = \frac{\theta \tau_{C_Y}}{t + \theta}.$$

By construction $\text{dist}(\bar{x}, \text{rel}\partial P) \geq r_{\bar{x}}$, $\text{dist}(\bar{w}, \text{rel}\partial C_Y) \geq r_{\bar{w}}\|A\|$, and $A\bar{x} - b = \bar{w} \in C_Y$. Therefore the point $\bar{x} \in X_d$. We now bound its distance to the relative boundary of the feasible region.

Consider any $v \in \widehat{L}_P \cap \{y | Ay \in L_{C_Y}\}$ such that $\|v\| \leq 1$, then

$$\bar{x} + \alpha v \in P, \text{ for any } |\alpha| \leq r_{\bar{x}} ,$$

and

$$A(\bar{x} + \alpha v) - b = \bar{w} + \alpha(Av) \in C_Y, \text{ for any } |\alpha| \leq r_{\bar{w}} .$$

Therefore $(\bar{x} + \alpha v) \in X_d$ for any $|\alpha| \leq \min\{r_{\bar{x}}, r_{\bar{w}}\}$, and the distance to the relative boundary of X_d is then $\text{dist}(\bar{x}, \text{rel}\partial X_d) \geq |\alpha|\|v\| \geq |\alpha|$, for any $|\alpha| \leq \min\{r_{\bar{x}}, r_{\bar{w}}\}$. Therefore $\text{dist}(\bar{x}, \text{rel}\partial X_d) \geq \min\{r_{\bar{x}}, r_{\bar{w}}\} \geq \frac{\theta \min\{r_{x^0}, \tau_{C_Y}\}}{t + \theta}$.

To finish the proof, we just have to bound the different expressions from the statement of the theorem; here we make use of inequality (24):

1. (a) $\|\bar{x} - x^0\| = \dfrac{\|x - tx^0\|}{t+\theta} \leq \dfrac{1}{\theta}\max\{1, \|x^0\|\} \leq \dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)}\max\{1, \|x^0\|\}$.

   (b) $\|\bar{x}\| \leq \dfrac{1}{\theta}\|x\| + \|x^0\| \leq \dfrac{1}{\theta} + \|x^0\| \leq \|x^0\| + \dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)}$.

2. (a) $\dfrac{1}{\text{dist}(\bar{x}, \text{rel}\partial P)} \leq \dfrac{1}{r_{\bar{x}}} = \dfrac{t+\theta}{\theta r_{x^0}} \leq \dfrac{1}{r_{x^0}}\left(1 + \dfrac{1}{\theta}\right) \leq \dfrac{1}{r_{x^0}}\left(1 + \dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)}\right)$.

   (b) $\begin{aligned}\dfrac{1}{\text{dist}(\bar{x}, \text{rel}\partial X_d)} &\leq \dfrac{1}{\min\{r_{x^0}, \tau_{C_Y}\}}\dfrac{t+\theta}{\theta} \leq \dfrac{1}{\min\{r_{x^0}, \tau_{C_Y}\}}\left(1 + \dfrac{1}{\theta}\right) \\ &\leq \dfrac{1}{\min\{r_{x^0}, \tau_{C_Y}\}}\left(1 + \dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)}\right) .\end{aligned}$

3. (a) $\begin{aligned}\dfrac{\|\bar{x} - x^0\|}{\text{dist}(\bar{x}, \text{rel}\partial P)} &\leq \dfrac{\|x - tx^0\|}{\theta r_{x^0}} \leq \dfrac{1}{r_{x^0}}\dfrac{1}{\theta}\max\{1, \|x^0\|\} \\ &\leq \dfrac{1}{r_{x^0}}\dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)}\max\{1, \|x^0\|\} .\end{aligned}$

   (b) $\begin{aligned}\dfrac{\|\bar{x} - x^0\|}{\text{dist}(\bar{x}, \text{rel}\partial X_d)} &\leq \dfrac{\|x - tx^0\|}{\theta\min\{r_{x^0}, \tau_{C_Y}\}} \leq \dfrac{1}{\min\{r_{x^0}, \tau_{C_Y}\}}\dfrac{1}{\theta}\max\{1, \|x^0\|\} \\ &\leq \dfrac{1}{\min\{r_{x^0}, \tau_{C_Y}\}}\dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)}\max\{1, \|x^0\|\} .\end{aligned}$

   (c) $\begin{aligned}\dfrac{\|\bar{x}\|}{\text{dist}(\bar{x}, \text{rel}\partial P)} &\leq \dfrac{\|x + \theta x^0\|}{\theta r_{x^0}} \leq \dfrac{1}{r_{x^0}}\left(\|x^0\| + \dfrac{1}{\theta}\right) \\ &\leq \dfrac{1}{r_{x^0}}\left(\|x^0\| + \dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)}\right) .\end{aligned}$

   (d) $\begin{aligned}\dfrac{\|\bar{x}\|}{\text{dist}(\bar{x}, \text{rel}\partial X_d)} &\leq \dfrac{\|x + \theta x^0\|}{\theta\min\{r_{x^0}, \tau_{C_Y}\}} \leq \dfrac{1}{\min\{r_{x^0}, \tau_{C_Y}\}}\left(\|x^0\| + \dfrac{1}{\theta}\right) \\ &\leq \dfrac{1}{\min\{r_{x^0}, \tau_{C_Y}\}}\left(\|x^0\| + \dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)}\right) .\end{aligned}$ ■

Finally, we note that Theorem 8 can be proved using almost identical arguments as in the proof of Theorem 7, but with a careful analysis to handle the special cases when P is an affine set or C_Y is a subspace, see [12] for exact details.

5.3 Solutions in the relative interior of Y_d

In this subsection we present results concerning geometric properties of the dual feasible region Y_d of (GD_d). We defer all proofs to the end of the subsection. Before proceeding,

we first discuss norms that arise when studying the dual problem. Motivated quite naturally by (18), we define the norm $\|(x,t)\| := \|x\|+|t|$ for points $(x,t) \in C \subset I\!R^n \times I\!R$. This then leads to the following dual norm for points $(s,u) \in C^* \subset I\!R^n \times I\!R$:

$$\|(s,u)\|_* := \max\{\|s\|_*, |u|\} \ . \tag{25}$$

Consistent with the characterization of $\rho_D(d)$ given by (20) in Theorem 6, we define the following dual norm for points $(y,\delta) \in I\!R^m \times I\!R$:

$$\|(y,\delta)\|_* := \|y\|_* + |\delta| \ . \tag{26}$$

It is clear that the above defines a norm on the vector space $I\!R^m \times I\!R$ which contains Y_d.

The following proposition bounds the norm of the y component of the dual feasible solution (y,u) in terms of the objective function value $b^t y - u$; it corresponds to Lemma 3.1 of [16] for the ground-set model format.

Proposition 6 *Consider any (y,u) feasible for (GD_d). If $\rho_P(d) > 0$ then*

$$\|y\|_* \leq \frac{\max\{\|c\|_*, -(b^t y - u)\}}{\rho_P(d)}.$$

The following result corresponds to Assertion 1 of Theorem 1.1 of [16] for the ground-set model format dual problem (GD_d):

Proposition 7 *Consider any $y^0 \in C_Y^*$. If $\rho_D(d) > 0$ then for any $\varepsilon > 0$, there exists $(\bar{y},\bar{u}) \in Y_d$ satisfying*

$$\|\bar{y} - y^0\| \leq \frac{\mathrm{dist}(c - A^t y^0, R^*) + \varepsilon}{\rho_D(d)} \max\left\{1, \|y^0\|\right\} \ .$$

The following is the main result of this subsection, and can be viewed as an extension of Theorems 15, 17, and 19 of [8] to the dual problem (GD_d). In Theorem 9 we assume for expository convenience that C_Y is not a subspace and that R (the recession cone of P) is not a subspace. These assumptions are relaxed in Theorem 10.

Theorem 9 *Suppose that R and C_Y are not subspaces and consider any $y^0 \in C_Y^*$. If $\rho_D(d) > 0$ then for any $\varepsilon > 0$, there exists $(\bar{y},\bar{u}) \in Y_d$ satisfying:*

1. *(a)* $\|\bar{y} - y^0\|_* \leq \dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)} \max\{1, \|y^0\|_*\}$

(b) $\|\bar{y}\|_* \leq \|y^0\|_* + \dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)}$

2. (a) $\dfrac{1}{\operatorname{dist}(\bar{y}, \operatorname{rel}\partial C_Y^*)} \leq \dfrac{1}{\operatorname{dist}(y^0, \operatorname{rel}\partial C_Y^*)} \left(1 + \dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)}\right)$

(b) $\dfrac{1}{\operatorname{dist}((\bar{y}, \bar{u}), \operatorname{rel}\partial Y_d)} \leq \dfrac{(1+\varepsilon)\max\{1, \|A\|\}}{\min\{\operatorname{dist}(y^0, \operatorname{rel}\partial C_Y^*), \tau_{R^*}\}} \left(1 + \dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)}\right)$

3. (a) $\dfrac{\|\bar{y} - y^0\|_*}{\operatorname{dist}(\bar{y}, \operatorname{rel}\partial C_Y^*)} \leq \dfrac{1}{\operatorname{dist}(y^0, \operatorname{rel}\partial C_Y^*)} \left(\dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)} \max\{1, \|y^0\|_*\}\right)$

(b) $\dfrac{\|\bar{y} - y^0\|_*}{\operatorname{dist}((\bar{y}, \bar{u}), \operatorname{rel}\partial Y_d)} \leq$
$$\dfrac{(1+\varepsilon)\max\{1, \|A\|\}}{\min\{\operatorname{dist}(y^0, \operatorname{rel}\partial C_Y^*), \tau_{R^*}\}} \left(\dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)} \max\{1, \|y^0\|_*\}\right)$$

(c) $\dfrac{\|\bar{y}\|_*}{\operatorname{dist}(\bar{y}, \operatorname{rel}\partial C_Y^*)} \leq \dfrac{1}{\operatorname{dist}(y^0, \operatorname{rel}\partial C_Y^*)} \left(\|y^0\|_* + \dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)}\right)$

(d) $\dfrac{\|\bar{y}\|_*}{\operatorname{dist}((\bar{y}, \bar{u}), \operatorname{rel}\partial Y_d)} \leq$
$$\dfrac{(1+\varepsilon)\max\{1, \|A\|\}}{\min\{\operatorname{dist}(y^0, \operatorname{rel}\partial C_Y^*), \tau_{R^*}\}} \left(\|y^0\|_* + \dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)}\right)$$

The statement of Theorem 10 below relaxes the assumptions on R and C_Y not being linear subspaces:

Theorem 10 *Consider any $y^0 \in C_Y^*$. If $\rho_D(d) > 0$ then for any $\varepsilon > 0$, there exists $(\bar{y}, \bar{u}) \in Y_d$ with the following properties:*

- *If C_Y is not a subspace, $(\bar{y}, \bar{u})$ satisfies all items of Theorem 9.*
- *If C_Y is a subspace and R is not a subspace, $(\bar{y}, \bar{u})$ satisfies all items of Theorem 9, where items 2.(a), 3.(a), and 3.(c) are vacuously valid as both sides of these inequalities are zero.*
- *If C_Y and R are subspaces, $(\bar{y}, \bar{u})$ satisfies items 1.(a), 1.(b), 2.(a), 3.(a), and 3.(c) of Theorem 9, where items 2.(a), 3.(a), and 3.(c) are vacuously valid as both sides of these inequalities are zero. The point $(\bar{y}, \bar{u})$ also satisfies*

2'.(b) $\dfrac{1}{\operatorname{dist}((\bar{y}, \bar{u}), \operatorname{rel}\partial Y_d)} \leq \varepsilon$

3'.(b) $\dfrac{\|\bar{y} - y^0\|_*}{\operatorname{dist}((\bar{y}, \bar{u}), \operatorname{rel}\partial Y_d)} \leq \varepsilon$

3'.(d) $$\frac{\|\bar{y}\|_*}{\mathrm{dist}((\bar{y},\bar{u}),\mathrm{rel}\partial Y_d)} \leq \varepsilon$$

We conclude this subsection by presenting a result which captures the thrust of Theorems 9 and 10, emphasizing how the distance to dual infeasibility $\rho_D(d)$ and the geometric properties of a given point $y^0 \in C_Y^*$ bound various geometric properties of the dual feasible region Y_d. For $y^0 \in \mathrm{relint} C_Y^*$, define:

$$g_{C_Y^*,R^*}(y^0) := \frac{\max\{\|y^0\|_*, 1\}}{\min\{1, \mathrm{dist}(y^0, \mathrm{rel}\partial C_Y^*), \tau_{R^*}\}} .$$

We now define a geometric measure for the dual feasible region. We do not consider the whole set Y_d; instead we consider only the projection onto the variables y. Let ΠY_d denote the projection of Y_d onto the space of the y variables:

$$\Pi Y_d := \{y \in I\!R^m \mid \text{there exists } u \in I\!R \text{ for which } (y,u) \in Y_d\} . \tag{27}$$

Note that the set ΠY_d corresponds exactly to the feasible region in the alternate formulation of the dual problem (15). We define the following geometric measure of the set ΠY_d:

$$g_{Y_d} := \inf_{(y,u)\in Y_d} \max\left\{\|y\|_*, \frac{\|y\|_*}{\mathrm{dist}(y, \mathrm{rel}\partial \Pi Y_d)}, \frac{1}{\mathrm{dist}(y, \mathrm{rel}\partial \Pi Y_d)}\right\} .$$

Corollary 3 *Consider any $y^0 \in C_Y^*$. If $\rho_D(d) > 0$ then*

$$g_{Y_d} \leq \max\{1, \|A\|\} g_{C_Y^*,R^*}(y^0) \left(1 + \frac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)}\right) .$$

Proof: We show in Lemma 4, item 4, that for any $(\bar{y},\bar{u}) \in Y_d$, $\mathrm{dist}(\bar{y}, \mathrm{rel}\partial \Pi Y_d) \geq \mathrm{dist}((\bar{y},\bar{u}), \mathrm{rel}\partial Y_d)$. If either C_Y or R is not a subspace, use items *1.(b)*, *2.(b)*, and *3.(d)* from Theorem 9 and apply the definition of g_{Y_d} to obtain

$$g_{Y_d} \leq (1+\varepsilon) \max\{1, \|A\|\} g_{C_Y^*,R^*}(y^0) \left(1 + \frac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)}\right) .$$

Since now the left side is independent of ε, take the limit as $\varepsilon \to 0$. If both C_Y and R are subspaces we obtain the stronger bound

$$g_{Y_d} \leq g_{C_Y^*,R^*}(y^0) \left(1 + \frac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)}\right)$$

by using item *1.(b)* from Theorem 9, items *2'.(b)* and *3'.(d)* from Theorem 10, and the definition of g_{Y_d}. ■

We now state Lemma 4, we start by defining the following set:

$$\widetilde{Y}_d := \left\{(y,u) \in I\!R^m \times I\!R \mid (c - A^t y, u) \in C^*\right\} . \tag{28}$$

Note that the dual feasible region Y_d is recovered from $\widetilde{Y}_d$ as $Y_d = \widetilde{Y}_d \cap (C_Y^* \times I\!R)$. The following lemma, whose proof is deferred to the Appendix, relates a variety of distances to relative boundaries of sets arising in the dual problem:

Lemma 4 *Given a dual feasible point $(y,u) \in Y_d$, let $s = c - A^t y \in$ effdom $u(\cdot)$. Then:*

1. $\text{dist}\left((y,u), \text{rel}\partial(C_Y^* \times I\!R)\right) = \text{dist}\left(y, \text{rel}\partial C_Y^*\right)$.
2. $\text{dist}\left((y,u), \text{rel}\partial\widetilde{Y}_d\right) \geq \frac{1}{\max\{1,\|A\|\}} \text{dist}((s,u), \text{rel}\partial C^*)$.
3. $\text{dist}\left((y,u), \text{rel}\partial Y_d\right) \geq \frac{1}{\max\{1,\|A\|\}} \min\left\{\text{dist}\left((s,u), \text{rel}\partial C^*\right), \text{dist}\left(y, \text{rel}\partial C_Y^*\right)\right\}$.
4. $\text{dist}(y, \text{rel}\partial\Pi Y_d) \geq \text{dist}((y,u), \text{rel}\partial Y_d)$.

We now proceed with the proofs of the results of this subsection.

Proof of Proposition 6: If $y = 0$ the result is true. If $y \neq 0$, then Proposition 9 shows that there exists $\hat{y}$ such that $\|\hat{y}\| = 1$ and $\hat{y}^t y = \|y\|_*$. For any $\varepsilon > 0$, define the following perturbed problem instance:

$$\bar{A} = A - \frac{1}{\|y\|_*}\hat{y}c^t, \qquad \bar{b} = b + \frac{((-b^t y + u)^+ + \varepsilon)}{\|y\|_*}\hat{y}, \qquad \bar{c} = c .$$

We note that, for the data $\bar{d} = (\bar{A}, \bar{b}, \bar{c})$, the point (y,u) satisfies $(A2_{\bar{d}})$ in Lemma 1, and therefore $(GP_{\bar{d}})$ is infeasible. We conclude that $\rho_P(d) \leq \|d - \bar{d}\|$, which implies

$$\rho_P(d) \leq \frac{\max\left\{\|c\|_*, (-b^t y + u)^+ + \varepsilon\right\}}{\|y\|_*}$$

and so

$$\rho_P(d) \leq \frac{\max\left\{\|c\|_*, -(b^t y - u)\right\}}{\|y\|_*} . \qquad \blacksquare$$

The following technical lemma, which concerns the optimization problem (DP) below, is used in the subsequent proofs. Problem (DP) is parameterized by given points $y^0 \in C_Y^*$ and $s^0 \in R^*$, and is defined by

$$\begin{array}{lll} (DP) & \max_{y,\delta,s,\theta} & \theta \\ & \text{s.t.} & -A^t y + \delta c - s = \theta\left(A^t y^0 - c + s^0\right) \\ & & \|y\|_* + |\delta| \leq 1 \\ & & y \in C_Y^* \\ & & \delta \geq 0 \\ & & s \in R^*. \end{array} \tag{29}$$

Lemma 5 *Consider any $y^0 \in C_Y^*$ and $s^0 \in R^*$ such that $A^t y^0 + s^0 \neq c$. If $\rho_D(d) > 0$, then there exists a point (y, δ, s, θ) feasible for problem (DP) that satisfies*

$$\theta \geq \frac{\rho_D(d)}{\|c - A^t y^0 - s^0\|_*} > 0 \ . \tag{30}$$

Proof: Note that problem (DP) is feasible for any y^0 and s^0 since $(y, \delta, s, \theta) = (0, 0, 0, 0)$ is always feasible. Therefore it can either be unbounded or have a finite optimal objective value. If (DP) is unbounded, we can find feasible points with an objective function large enough such that (30) holds. If (DP) has a finite optimal value, say θ^*, then it follows from elementary arguments that it attains this value. Since $\rho_D(d) > 0$ implies $Y_d \neq \emptyset$, Theorem 6 implies that the optimal solution $(y^*, \delta^*, s^*, \theta^*)$ for (DP) satisfies (30). ■

Proof of Proposition 7: Assume $c - A^t y^0 \notin \text{relint} R^*$, otherwise from Proposition 1, the point $(\bar{y}, \bar{u}) = (y^0, u(c - A^t y^0))$ satisfies the assertion of the proposition. We consider problem (DP), defined by (29), with y^0 and $s^0 \in \text{relint} R^*$ such that $\|c - A^t y^0 - s^0\| \leq \text{dist}(c - A^t y^0, R^*) + \varepsilon$. From Lemma 5 we have that there exists a point (y, δ, s, θ) feasible for (DP) that satisfies

$$\theta \geq \frac{\rho_D(d)}{\|c - A^t y^0 - s^0\|} \geq \frac{\rho_D(d)}{\text{dist}(c - A^t y^0, R^*) + \varepsilon} \ .$$

Define

$$\bar{y} = \frac{y + \theta y^0}{\delta + \theta} \quad \text{and} \quad \bar{s} = \frac{s + \theta s^0}{\delta + \theta} \ .$$

By construction we have $\bar{y} \in C_Y^*$, $c - A^t \bar{y} = \bar{s} \in \text{relint} R^*$. Therefore from Proposition 1 $(\bar{y}, u(c - A^t \bar{y})) \in Y_d$, and letting $\xi = \max\{1, \|y^0\|_*\}$ we have

$$\|\bar{y} - y^0\|_* = \frac{\|y - \delta y^0\|_*}{\delta + \theta} \leq \frac{(\|y\|_* + \delta)\xi}{\theta} \leq \frac{\text{dist}(c - A^t y^0, R^*) + \varepsilon}{\rho_D(d)} \ \xi \ .$$

■

Proof of Theorem 9: Note that $\rho_D(d) > 0$ implies $Y_d \neq \emptyset$; note also that $\rho_D(d)$ is finite, otherwise Proposition 2 shows that $R = \{0\}$ which is a subspace. Set $s^0 \in R^*$ such that $\|s^0\|_* = \|A\|$ and $\tau_{R^*} = \frac{\text{dist}(s^0, \text{rel}\partial R^*)}{\|s^0\|_*}$. We also assume for now that $c - A^t y^0 \neq s^0$. We show later in the proof how to handle the case when $c - A^t y^0 = s^0$. Denote $r_{y^0} = \text{dist}(y^0, \text{rel}\partial C_Y^*)$, and $r_{s^0} = \text{dist}(s^0, \text{rel}\partial R^*) = \tau_{R^*} \|A\| > 0$.

With the points y^0 and s^0, use Lemma 5 to obtain a point (y, δ, s, θ) feasible for (DP) such that from inequality (30) satisfies

$$0 < \frac{1}{\theta} \leq \frac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)} \ . \tag{31}$$

Define the following:

$$\bar{y} = \frac{y + \theta y^0}{\delta + \theta}, \quad \bar{s} = \frac{s + \theta s^0}{\delta + \theta}, \quad r_{\bar{y}} = \frac{\theta r_{y^0}}{\delta + \theta}, \quad r_{\bar{s}} = \frac{\theta r_{s^0}}{\delta + \theta}.$$

By construction $\text{dist}(\bar{y}, \text{rel}\partial C_Y^*) \geq r_{\bar{y}}$, $\text{dist}(\bar{s}, \text{rel}\partial R^*) \geq r_{\bar{s}}$, and $c - A^t\bar{y} = \bar{s}$. Therefore, from Proposition 1 the point $(\bar{y}, u(\bar{s})) \in Y_d$. We now choose $\bar{u}$ so that $(\bar{y}, \bar{u}) \in Y_d$ and bound its distance to the relative boundary. Since relint $R^* \subseteq$ effdom $u(\cdot)$, from Proposition 11 and Proposition 12, we have that for any $\varepsilon > 0$, the ball $B\left(\bar{s}, \frac{r_{\bar{s}}}{1+\varepsilon}\right) \cap L_{R^*} \subset$ relint effdom $u(\cdot)$. Define the function $\tilde{\mu}(\cdot, \cdot)$ by

$$\tilde{\mu}(\bar{s}, \kappa) := \begin{array}{ll} \frac{1}{\|A\|}\kappa & + \quad \sup_s u(s) \\ & \quad \|s - \bar{s}\|_* \leq \kappa \\ & \quad s \in R^* \end{array} .$$

Note that $\tilde{\mu}(\cdot, \cdot)$ is finite for every $\bar{s} \in$ relint effdom $u(\cdot)$ and $\kappa \in [0, \text{dist}(\bar{s}, \text{rel}\partial R^*))$, because it is defined as the supremum of the continuous function $u(\cdot)$ over a closed and bounded subset contained in the relative interior of its effective domain, see Theorem 10.1 of [18]. We define $\bar{u} = \tilde{\mu}\left(\bar{s}, \frac{r_{\bar{s}}}{1+\varepsilon}\right)$, and since $\bar{u} \geq \frac{r_{\bar{s}}}{\|A\|(1+\varepsilon)} + u(\bar{s}) \geq u(\bar{s})$ the point $(\bar{y}, \bar{u}) \in Y_d$. Let us now bound $\text{dist}((\bar{y}, \bar{u}), \text{rel}\partial Y_d)$. Consider any vector $v \in L_{C_Y^*} \cap \{y | -A^t y \in L_{R^*}\}$ such that $\|v\|_* \leq 1$, then

$$\bar{y} + \alpha v \in C_Y^* \text{ for any } |\alpha| \leq r_{\bar{y}} ,$$

and

$$c - A^t(\bar{y} + \alpha v) = \bar{s} + \alpha(-A^t v) \in B\left(\bar{s}, \frac{r_{\bar{s}}}{1+\varepsilon}\right) \cap L_{R^*} \text{ for any } |\alpha| \leq \frac{r_{\bar{s}}}{\|A\|(1+\varepsilon)} .$$

This last inclusion implies that $(c - A^t(\bar{y} + \alpha v), \bar{u} + \beta) = (\bar{s} + \alpha(-A^t v), \bar{u} + \beta) \in C^*$ for any $|\alpha|$, $|\beta| \leq \frac{r_{\bar{s}}}{\|A\|(1+\varepsilon)}$. We have shown that $\text{dist}(\bar{y}, \text{rel}\partial C_Y^*) \geq r_{\bar{y}}$ and $\text{dist}((c - A^t\bar{y}, \bar{u}), \text{rel}\partial C^*) \geq \frac{r_{\bar{s}}}{\|A\|(1+\varepsilon)}$. Therefore item 3 of Lemma 4 implies

$$\begin{aligned} \text{dist}((\bar{y}, \bar{u}), \text{rel}\partial Y_d) &\geq \frac{1}{\max\{1, \|A\|\}} \min\left\{r_{\bar{y}}, \frac{r_{\bar{s}}}{\|A\|(1+\varepsilon)}\right\} \\ &\geq \frac{1}{(1+\varepsilon)\max\{1, \|A\|\}} \frac{\theta}{\delta + \theta} \min\left\{r_{y^0}, \frac{r_{s^0}}{\|A\|}\right\} \\ &= \frac{\theta \min\{r_{y^0}, \tau_{R^*}\}}{(1+\varepsilon)\max\{1, \|A\|\}(\delta + \theta)} . \end{aligned}$$

To finish the proof, we bound the different expressions in the statement of the theorem; let $\xi = \max\{1, \|A\|\}$ to simplify notation. Here we use inequality (31):

1. (a) $\|\bar{y} - y^0\|_* = \dfrac{\|y - \delta y^0\|_*}{\delta + \theta} \leq \dfrac{\max\{1, \|y^0\|_*\}}{\theta} \leq \dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)} \max\{1, \|y^0\|_*\}$.

   (b) $\|\bar{y}\|_* \leq \dfrac{1}{\theta}\|y\|_* + \|y^0\|_* \leq \dfrac{1}{\theta} + \|y^0\|_* \leq \|y^0\|_* + \dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)}$.

2. (a) $\dfrac{1}{\operatorname{dist}(\bar{y}, \operatorname{rel}\partial C_Y^*)} \leq \dfrac{1}{r_{\bar{y}}} = \dfrac{\delta + \theta}{\theta r_{y^0}} \leq \dfrac{1}{r_{y^0}}\left(1 + \dfrac{1}{\theta}\right) \leq \dfrac{1}{r_{y^0}}\left(1 + \dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)}\right)$.

   (b) $\begin{aligned} \dfrac{1}{\operatorname{dist}((\bar{y}, \bar{u}), \operatorname{rel}\partial Y_d)} &\leq \dfrac{(1+\varepsilon)\xi}{\min\{r_{y^0}, \tau_{R^*}\}} \dfrac{\delta + \theta}{\theta} \leq \dfrac{(1+\varepsilon)\xi}{\min\{r_{y^0}, \tau_{R^*}\}}\left(1 + \dfrac{1}{\theta}\right) \\ &\leq \dfrac{(1+\varepsilon)\xi}{\min\{r_{y^0}, \tau_{R^*}\}}\left(1 + \dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)}\right). \end{aligned}$

3. (a) $\begin{aligned} \dfrac{\|\bar{y} - y^0\|_*}{\operatorname{dist}(\bar{y}, \operatorname{rel}\partial C_Y^*)} &\leq \dfrac{\|y - \delta y^0\|_*}{\theta r_{y^0}} \leq \dfrac{1}{r_{y^0}\theta} \max\{1, \|y^0\|_*\} \\ &\leq \dfrac{1}{r_{y^0}} \dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)} \max\{1, \|y^0\|_*\}. \end{aligned}$

   (b) $\begin{aligned} \dfrac{\|\bar{y} - y^0\|_*}{\operatorname{dist}((\bar{y}, \bar{u}), \operatorname{rel}\partial Y_d)} &\leq \dfrac{(1+\varepsilon)\xi\|y - \delta y^0\|_*}{\theta \min\{r_{y^0}, \tau_{R^*}\}} \leq \dfrac{(1+\varepsilon)\xi}{\min\{r_{y^0}, \tau_{R^*}\}} \dfrac{1}{\theta} \max\{1, \|y^0\|_*\} \\ &\leq \dfrac{(1+\varepsilon)\xi}{\min\{r_{y^0}, \tau_{R^*}\}} \dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)} \max\{1, \|y^0\|_*\}. \end{aligned}$

   (c) $\begin{aligned} \dfrac{\|\bar{y}\|_*}{\operatorname{dist}(\bar{y}, \operatorname{rel}\partial C_Y^*)} &\leq \dfrac{\|y + \theta y^0\|_*}{\theta r_{y^0}} \leq \dfrac{1}{r_{y^0}}\left(\|y^0\|_* + \dfrac{1}{\theta}\right) \\ &\leq \dfrac{1}{r_{y^0}}\left(\|y^0\|_* + \dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)}\right). \end{aligned}$

   (d) $\begin{aligned} \dfrac{\|\bar{y}\|_*}{\operatorname{dist}((\bar{y}, \bar{u}), \operatorname{rel}\partial Y_d)} &\leq \dfrac{\|y + \theta y^0\|_*(1+\varepsilon)\xi}{\theta \min\{r_{y^0}, \tau_{R^*}\}} \leq \dfrac{(1+\varepsilon)\xi}{\min\{r_{y^0}, \tau_{R^*}\}}\left(\|y^0\|_* + \dfrac{1}{\theta}\right) \\ &\leq \dfrac{(1+\varepsilon)\xi}{\min\{r_{y^0}, \tau_{R^*}\}}\left(\|y^0\|_* + \dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)}\right). \end{aligned}$

For the case $c - A^t y^0 = s^0$, define $\bar{y} = y^0$ and $\bar{u} = \tilde{\mu}\left(s^0, \frac{\tau_{R^*}\|A\|}{1+\varepsilon}\right)$. The proof then proceeds exactly as above, except that now we show that $\operatorname{dist}((c - A^t\bar{y}, \bar{u}), \operatorname{rel}\partial C^*) \geq \frac{\tau_{R^*}}{1+\varepsilon}$, which implies that $\operatorname{dist}((\bar{y}, \bar{u}), \operatorname{rel}\partial Y_d) \geq \frac{1}{\max\{1, \|A\|\}(1+\varepsilon)} \min\{\tau_{R^*}, r_{y^0}\}$ from item 3 of Lemma 4. This inequality is then used to prove each item in the theorem. ■

Finally, we note that Theorem 10 can be proved using almost identical arguments as in the proof of Theorem 9, but with a careful analysis to handle the special cases when R or C_Y are subspaces, see [12] for the exact details.

6 Sensitivity under Perturbation

In this section we present several results that bound the deformation of primal and dual feasible regions and objective function values under data perturbation. All proofs are deferred to the end of the section.

The following two theorems bound the deformation of the primal and dual feasible regions under data perturbation. These results are essentially extensions of Assertion 2 of Theorem 1.1 of [16] to the primal and dual problems in the GSM format.

Theorem 11 *Suppose that $\rho_P(d) > 0$. Let $\Delta d = (\Delta A, \Delta b, \Delta c)$ be such that $X_{d+\Delta d} \neq \emptyset$ and consider any $x' \in X_{d+\Delta d}$. Then there exists $\bar{x} \in X_d$ satisfying*

$$\|\bar{x} - x'\| \leq (\|\Delta b\| + \|\Delta A\|\|x'\|) \frac{\max\{1, \|x'\|\}}{\rho_P(d)} .$$

Theorem 12 *Suppose that $\rho_D(d) > 0$. Let $\Delta d = (\Delta A, \Delta b, \Delta c)$ be such that $Y_{d+\Delta d} \neq \emptyset$ and consider any $(y', u') \in Y_{d+\Delta d}$. Then for any $\varepsilon > 0$, there exists $(\bar{y}, \bar{u}) \in Y_d$ satisfying*

$$\|\bar{y} - y'\|_* \leq (\|\Delta c\|_* + \|\Delta A\|\|y'\|_* + \varepsilon) \frac{\max\{1, \|y'\|_*\}}{\rho_D(d)} .$$

The next two results bound changes in optimal objective function values under data perturbation. Proposition 8 and Theorem 13 below respectively extend to the ground-set model format Lemma 3.9 and Assertion 5 of Theorem 1.1 in [16].

Proposition 8 *Suppose that $d \in \mathcal{F}$ and $\rho(d) > 0$. Let $\Delta d = (0, \Delta b, 0)$ be such that $X_{d+\Delta d} \neq \emptyset$. Then,*

$$z_*(d + \Delta d) - z_*(d) \geq -\|\Delta b\| \frac{\max\{\|c\|_*, -z_*(d)\}}{\rho_P(d)} .$$

Theorem 13 *Suppose that $d \in \mathcal{F}$ and $\rho(d) > 0$. Let $\Delta d = (\Delta A, \Delta b, \Delta c)$ satisfy $\|\Delta d\| < \rho(d)$. Then, if x^* and $\hat{x}$ are optimal solutions for (GP_d) and $(GP_{d+\Delta d})$ respectively,*

$$|z_*(d + \Delta d) - z_*(d)| \leq \|\Delta b\| \frac{\max\{\|c\|_* + \|\Delta c\|_*, -z_*(d)\}}{\rho_P(d) - \|\Delta d\|} +$$

$$+\left(\|\Delta c\|_* + \|\Delta A\| \frac{\max\{\|c\|_* + \|\Delta c\|_*, -z_*(d)\}}{\rho_P(d) - \|\Delta d\|}\right) \max\{\|x^*\|, \|\hat{x}\|\} .$$

Proof of Theorem 11: We consider problem (PP), defined by (22), with $x^0 = x'$ and w^0 such that $(A + \Delta A)x' - (b + \Delta b) = w^0 \in C_Y$. From Lemma 3 we have that there exists a point (x, t, w, θ) feasible for (PP) that satisfies

$$\theta \geq \frac{\rho_P(d)}{\|b - Ax^0 + w^0\|} = \frac{\rho_P(d)}{\|\Delta A x' - \Delta b\|} \geq \frac{\rho_P(d)}{\|\Delta A\|\|x'\| + \|\Delta b\|} .$$

We define

$$\bar{x} = \frac{x + \theta x'}{t + \theta} , \quad \bar{w} = \frac{w + \theta w^0}{t + \theta} .$$

By construction we have that $\bar{x} \in P$, $A\bar{x} - b = \bar{w} \in C_Y$, therefore $\bar{x} \in X_d$, and

$$\|\bar{x} - x'\| = \frac{\|x - tx'\|}{t + \theta} \leq \frac{(\|x\| + t)\max\{1, \|x'\|\}}{\theta} \leq \frac{\|\Delta A\|\|x'\| + \|\Delta b\|}{\rho_P(d)} \max\{1, \|x'\|\} . \quad \blacksquare$$

Proof of Theorem 12: From Proposition 11 we have that for any $\varepsilon > 0$ there exists $\xi \neq \Delta A^t y' - \Delta c$ such that $\|\xi\|_* \leq \varepsilon$ and $c + \Delta c + \xi - (A + \Delta A)^t y' \in \text{relint} R^*$. We consider problem (DP) defined by (29), with $y^0 = y'$ and $s^0 := c + \Delta c + \xi - (A + \Delta A)^t y' \in \text{relint} R^*$. From Lemma 5 we have that there exists a point (y, δ, s, θ) feasible for (DP) that satisfies

$$\theta \geq \frac{\rho_D(d)}{\|c - A^t y^0 - s^0\|_*} = \frac{\rho_D(d)}{\|\Delta A^t y' - \Delta c - \xi\|_*} \geq \frac{\rho_D(d)}{\|\Delta c\|_* + \|\Delta A\|\|y'\|_* + \varepsilon} .$$

We define

$$\bar{y} = \frac{y + \theta y'}{\delta + \theta} , \quad \bar{s} = \frac{s + \theta s^0}{\delta + \theta} .$$

By construction we have that $\bar{y} \in C_Y^*$ and $c - A^t\bar{y} = \bar{s} \in \text{relint} R^* \subseteq \text{effdom}\, u(\cdot)$, from Proposition 11 and Proposition 12. Therefore from Proposition 1, $(\bar{y}, u(c - A^t\bar{y})) \in Y_d$ and

$$\|\bar{y} - y'\|_* = \frac{\|y - \delta y'\|_*}{\delta + \theta} \leq \frac{(\|y\|_* + \delta)\max\{1, \|y'\|_*\}}{\theta} \leq \frac{\|\Delta c\|_* + \|\Delta A\|\|y'\|_* + \varepsilon}{\rho_D(d)} \max\{1, \|y'\|\} .$$

$\blacksquare$

Proof of Proposition 8: The hypothesis that $\rho(d) > 0$ implies that the GSM format problem with data d has zero duality gap and (GP_d) and (GD_d) attain their optimal values, see Corollary 1. Also, since $Y_{d+\Delta d} = Y_d \neq \emptyset$ has a Slater point (since $\rho_D(d) > 0$), and $X_{d+\Delta d} \neq \emptyset$, then $(GP_{d+\Delta d})$ and $(GD_{d+\Delta d})$ have no duality gap and $(GP_{d+\Delta d})$ attains

its optimal value, see Theorem 2. Let $(y, u) \in Y_d$ be an optimal solution of (GD_d), due to the form of the perturbation, point $(y, u) \in Y_{d+\Delta d}$, and therefore

$$z^*(d + \Delta d) \geq (b + \Delta b)^t y - u = z^*(d) + \Delta b^t y \geq z^*(d) - \|\Delta b\| \|y\|_* .$$

The result now follows using the bound on the norm of dual feasible solutions from Proposition 6 and the strong duality for data instances d and $d + \Delta d$. ■

Proof of Theorem 13: The hypothesis that $\rho(d) > 0$ and $\rho(d + \Delta d) > 0$ imply that the GSM format problems with data d and $d + \Delta d$ both have zero duality gap and all problems attain their optimal values, see Corollary 1.

Let $\hat{x} \in X_{d+\Delta d}$ be an optimal solution for $(GP_{d+\Delta d})$. Define the perturbation $\Delta\tilde{d} = (0, \Delta b - \Delta A\hat{x}, 0)$. Then by construction the point $\hat{x} \in X_{d+\Delta\tilde{d}}$. Therefore

$$z_*(d + \Delta d) = (c + \Delta c)^t \hat{x} \geq -\|\Delta c\|_* \|\hat{x}\| + c^t \hat{x} \geq -\|\Delta c\|_* \|\hat{x}\| + z_*(d + \Delta\tilde{d}) .$$

Invoking Proposition 8, we bound the optimal objective function value for the problem instance $d + \Delta\tilde{d}$:

$$z_*(d + \Delta d) + \|\Delta c\|_* \|\hat{x}\| \geq z_*(d + \Delta\tilde{d}) \geq z_*(d) - \|\Delta b - \Delta A\hat{x}\| \frac{\max\{\|c\|_*, -z^*(d)\}}{\rho_P(d)} .$$

Therefore

$$z_*(d + \Delta d) - z_*(d) \geq -\|\Delta c\|_* \|\hat{x}\| - (\|\Delta b\| + \|A\| \|\hat{x}\|) \frac{\max\{\|c\|_*, -z^*(d)\}}{\rho_P(d)} .$$

Changing the roles of d and $d + \Delta d$ we can construct the following upper bound:

$$z_*(d + \Delta d) - z_*(d) \leq \|\Delta c\|_* \|x^*\| + (\|\Delta b\| + \|A\| \|x^*\|) \frac{\max\{\|c + \Delta c\|_*, -z^*(d + \Delta d)\}}{\rho_P(d + \Delta d)} ,$$

where $x^* \in X_d$ is an optimal solution for (GP_s). The value $-z^*(d + \Delta d)$ can be replaced by $-z^*(d)$ on the right side of the previous bound. To see this consider two cases. If $-z^*(d + \Delta d) \leq -z^*(d)$, then we can do the replacement since it yields a larger bound. If $-z^*(d + \Delta d) > -z^*(d)$, the inequality above has a negative left side and a positive right side after the replacement. Note also that because of the hypothesis $\|\Delta d\| < \rho(d)$, the distance to infeasibility satisfies $\rho_P(d + \Delta d) \geq \rho_P(d) - \|\Delta d\| > 0$. We finish the proof combining the previous two bounds, incorporating the lower bound on $\rho_P(d + \Delta d)$, and using strong duality of the data instances d and $d + \Delta d$. ■

7 Concluding Remarks

We have shown herein that most of the essential results regarding condition numbers for conic convex optimization problems can be extended to the non-conic ground-set model

format (GP_d). We have attempted herein to highlight the most important and/or useful extensions; for other results see [12].

It is interesting to note the absence of results that directly bound $z_*(d)$ or the norms of optimal solutions $\|x^*\|$, $\|y^*\|$ of (GP_d) and (GD_d) as in Assertions 3 and 4 of Theorem 1.1 of [16]. Such bounds are very important in relating the condition number theory to the complexity of algorithms. However, we do not believe that such bounds can be demonstrated for (GP_d) without further assumptions. The reason for this is subtle yet simple. Observe from Theorem 6 that $\rho_D(d)$ depends only on $d = (A, b, c)$, C_Y, and the recession cone R of P. That is, P only affects $\rho_D(d)$ through its recession cone, and so information about the "bounded" portion of P is irrelevant to the value of $\rho_D(d)$. For this reason it is not possible to bound the norm of primal optimal solutions x directly, and hence one cannot bound $z_*(d)$ directly either. Under rather mild additional assumptions, it is possible to analyze the complexity of algorithms for solving (GP_d), see [12] as well as a forthcoming paper on this topic.

Note that the characterization results for $\rho_P(d)$ and $\rho_D(d)$ presented herein in Theorems 5 and 6 pertain only to the case when $d \in \mathcal{F}$. A characterization of $\rho(d)$ for $d \notin \mathcal{F}$ is the subject of future research.

8 Appendix

This appendix contains supporting mathematical results that are used in the proofs of the results of this paper. We point the reader to existing proofs for the most well known results.

Proposition 9 (Proposition 2 of [8]) *Let X be an n-dimensional normed vector space with dual space X^*. For every $x \in X$, there exists $\bar{x} \in X^*$ with the property that $\|\bar{x}\|_* = 1$ and $\|x\| = \bar{x}^t x$.* ■

Proposition 10 (Theorems 11.1 and 11.3 of [18]) *Given two non-empty convex sets S and T in $\mathbb{R}^n$, then* relint $S \cap$ relint $T = \emptyset$ *if and only if S and T can be properly separated, i.e., there exists $y \neq 0$ such that*

$$\begin{aligned} \inf_{x \in S} y^t x &\geq \sup_{z \in T} y^t z \\ \sup_{x \in S} y^t x &> \inf_{z \in T} y^t z \ . \end{aligned}$$ ■

The following is a restatement of Corollary 14.2.1 of [18], which relates the effective domain of $u(\cdot)$ of (14) to the recession cone of P, where recall that R^* denotes the dual of the recession cone R defined in (8).

Proposition 11 (Corollary 14.2.1 of [18]) *Let R denote the recession cone of the nonempty convex set P, and define $u(\cdot)$ by (14). Then* cl effdom $u(\cdot) = R^*$. ■

Proposition 12 (Theorem 6.3 of [18]) *For any convex set $Q \subseteq I\!R^n$,* cl relint $Q =$ cl Q, *and* relint cl $Q =$ relint Q. ■

The following lemma is central in relating the two alternative characterizations of the distance to infeasibility and is used in the proofs in Section 4.

Lemma 6 *Consider two closed convex cones $C \subseteq I\!R^n$ and $C_Y \subseteq I\!R^m$, and data $(M, v) \in I\!R^{m\times n} \times I\!R^m$. Strong duality holds between*

$$
\begin{array}{lrlll}
(P): & z_* = & \min & \|M^t y + q\|_* & \\
 & & s.t. & y^t v \geq 1 & \\
 & & & y \in C_Y^* & \\
 & & & q \in C^* &
\end{array}
\quad \textit{and} \quad
\begin{array}{lrll}
(D): & z^* = & \max & \theta \\
 & & s.t. & Mx - \theta v \in C_Y \\
 & & & \|x\| \leq 1 \\
 & & & \theta \geq 0 \\
 & & & x \in C \ .
\end{array}
$$

Proof: The proof that weak duality holds between (P) and (D) is straightforward, therefore $z^* \leq z_*$. Assume $z^* < z_*$, and set $\varepsilon > 0$ such that $0 \leq z^* < z_* - \varepsilon$. Consider the following nonempty convex set S:

$$
S := \left\{ (u, \delta, \alpha) \mid \exists y, q \text{ s.t. } y + u \in C_Y^*, q + \delta \in C^*, y^t v \geq 1 - \alpha, \|M^t y + q\|_* \leq z_* - \varepsilon \right\} \ .
$$

Then $(0, 0, 0) \notin S$, and from Proposition 10 there exists $(z, x, \theta) \neq 0$ such that $z^t u + x^t \delta + \theta\alpha \geq 0$ for any $(u, \delta, \alpha) \in S$. For any $y \in I\!R^m$, $\tilde{u} \in C_Y^*$, $\tilde{\delta} \in C^*$, $\pi \geq 0$, and $\tilde{q}$ such that $\|\tilde{q}\|_* \leq z_* - \varepsilon$, define $q = -M^t y + \tilde{q}$, $u = -y + \tilde{u}$, $\delta = -q + \tilde{\delta}$, and $\alpha = 1 - y^t v + \pi$. This construction implies that the point $(u, \delta, \alpha) \in S$, and that for all y, $\tilde{u} \in C_Y^*$, $\tilde{\delta} \in C^*, \pi \geq 0$, and $\|\tilde{q}\|_* \leq z_* - \varepsilon$ it holds that:

$$
\begin{aligned}
0 &\leq z^t(-y + \tilde{u}) + x^t(M^t y - \tilde{q} + \tilde{\delta}) + \theta(1 - y^t v + \pi) \\
&= y^t \left(Mx - \theta v - z\right) + z^t \tilde{u} + x^t \tilde{\delta} - x^t \tilde{q} + \theta + \theta\pi \ .
\end{aligned}
$$

This implies that $Mx - \theta v = z \in C_Y$, $x \in C$, $\theta \geq 0$, and $\theta \geq x^t \tilde{q}$ for $\|\tilde{q}\|_* \leq z_* - \varepsilon$. If $x \neq 0$, re-scale (z, x, θ) such that $\|x\| = 1$ and then (x, θ) is feasible for (D). Set $\tilde{q} = (z_* - \varepsilon)\hat{q}$, where $\hat{q}$ is given by Proposition 9 and is such that $\|\hat{q}\|_* = 1$ and $\hat{q}^t x = \|x\| = 1$. It then follows that $z^* \geq \theta \geq x^t \tilde{q} = z_* - \varepsilon > z^*$, which is a contradiction.

If $x = 0$, the above expression implies $-\theta v = z \in C_Y$, and $\theta \geq 0$. If $\theta > 0$ then $-v \in C_Y$, which means that the point $(0, \beta)$ is feasible for (D) for any $\beta \geq 0$, implying that $z^* = \infty$, a contradiction since $z^* < z_*$. If $\theta = 0$, then $z = 0$, which is a contradiction since $(z, x, \theta) \neq 0$. ■

The next two lemmas concern properties of the distance to the relative boundary of a convex set.

Lemma 7 *Given convex sets A and B, and a point $x \in A \cap B$, then*

$$\text{dist}(x, \text{rel}\partial(A \cap B)) \geq \min\{\text{dist}(x, \text{rel}\partial A),\ \text{dist}(x, \text{rel}\partial B)\} \ .$$

Proof: The proof of this lemma is based in showing that $L_{A\cap B} \setminus (A \cap B) \subset (L_A \setminus A) \cup (L_B \setminus B)$. If this inclusion is true then

$$\begin{aligned}
\text{dist}(x, \text{rel}\partial(A \cap B)) &= \inf_{\bar{x} \in L_{A\cap B} \setminus (A \cap B)} \|x - \bar{x}\| \\
&\geq \inf_{\bar{x} \in (L_A \setminus A) \cup (L_B \setminus B)} \|x - \bar{x}\| \\
&= \min\left\{ \inf_{\bar{x} \in L_A \setminus A} \|x - \bar{x}\|,\ \inf_{\bar{x} \in L_B \setminus B} \|x - \bar{x}\| \right\} \\
&= \min\{\text{dist}(x, \text{rel}\partial A),\ \text{dist}(x, \text{rel}\partial B)\} \ ,
\end{aligned}$$

which proves the lemma. Therefore we now prove the inclusion. Consider some $\bar{x} \in L_{A\cap B}$, this means that there exists $\alpha_i \in I\!R$, $x_i \in A \cap B$, $i \in I$ a finite set, and $\sum_{i\in I} \alpha_i = 1$, such that $\bar{x} = \sum_{i\in I} \alpha_i x_i$. Since $x_i \in A$ and $x_i \in B$, we have that $\bar{x} \in L_A$ and $\bar{x} \in L_B$. Therefore $L_{A\cap B} \subset L_A \cap L_B$. The desired inclusion is then obtained with a little algebra:

$$\begin{aligned}
L_{A\cap B} \setminus (A \cap B) &\subset L_A \cap L_B \cap (A \cap B)^C \\
&= L_A \cap L_B \cap \left(A^C \cup B^C\right) \\
&= \left(L_A \cap L_B \cap A^C\right) \cup \left(L_A \cap L_B \cap B^C\right) \\
&\subset \left(L_A \cap A^C\right) \cup \left(L_B \cap B^C\right) \\
&= (L_A \setminus A) \cup (L_B \setminus B) \ . \qquad \blacksquare
\end{aligned}$$

Last of all, we have:

Proof of Lemma 4: Equality (1.) is a consequence of the fact that $(y, u) \in L_{C_Y^* \times I\!R} \setminus (C_Y^* \times I\!R)$ if and only if $y \in L_{C_Y^*} \setminus C_Y^*$, and that $\|(y, u) - (\bar{y}, u)\|_* = \|y - \bar{y}\|_* + |u - u| = \|y - \bar{y}\|_*$.

To prove inequality (2.), we first show that if $(\bar{y}, \bar{u}) \in L_{\widetilde{Y}_d} \setminus \widetilde{Y}_d$, then $(c - A^t\bar{y}, \bar{u}) \in L_{C^*} \setminus C^*$. First note that if $(\bar{y}, \bar{u}) \notin \widetilde{Y}_d$ then, by the definition of $\widetilde{Y}_d$, $(c - A^t\bar{y}, \bar{u}) \notin C^*$;

therefore we only need to show that if $(\bar{y},\bar{u}) \in L_{\widetilde{Y}_d}$ then $(c - A^t\bar{y},\bar{u}) \in L_{C^*}$. Let $(\bar{y},\bar{u}) \in L_{\widetilde{Y}_d}$. Then there exists $\alpha_i \in I\!R$, $(y_i,u_i) \in \widetilde{Y}_d$, $i \in I$ a finite set, and $\sum_{i\in I}\alpha_i = 1$, such that $(\bar{y},\bar{u}) = \sum_{i\in I}\alpha_i(y_i,u_i)$. Consider

$$(c - A^t\bar{y},\bar{u}) = (c - A^t\sum_{i\in I}\alpha_i y_i, \sum_{i\in I}\alpha_i u_i) = \sum_{i\in I}\alpha_i(c - A^t y_i, u_i) \ .$$

Since $(y_i,u_i) \in \widetilde{Y}_d$ then $(c - A^t y_i, u_i) \in C^*$ and therefore $(c - A^t\bar{y},\bar{u}) \in L_{C^*}$.

The inclusion above means that

$$\begin{aligned}
\mathrm{dist}((s,u),\mathrm{rel}\partial C^*) &= \inf_{(\bar{s},\bar{u})\in L_{C^*}\setminus C^*} \|(s,u)-(\bar{s},\bar{u})\|_* \\
&\le \inf_{(\bar{y},\bar{u})\in L_{\widetilde{Y}_d}\setminus\widetilde{Y}_d} \|(s,u)-(c-A^t\bar{y},\bar{u})\|_* \\
&= \inf_{(\bar{y},\bar{u})\in L_{\widetilde{Y}_d}\setminus\widetilde{Y}_d} \max\{\|s-(c-A^t\bar{y})\|_*, |u-\bar{u}|\} \\
&= \inf_{(\bar{y},\bar{u})\in L_{\widetilde{Y}_d}\setminus\widetilde{Y}_d} \max\{\|A^t\bar{y}-A^t y\|_*, |u-\bar{u}|\} \\
&\le \inf_{(\bar{y},\bar{u})\in L_{\widetilde{Y}_d}\setminus\widetilde{Y}_d} \max\{\|A\|\|\bar{y}-y\|_*, |u-\bar{u}|\} \\
&\le \inf_{(\bar{y},\bar{u})\in L_{\widetilde{Y}_d}\setminus\widetilde{Y}_d} \max\{\|A\|,1\}\max\{\|y-\bar{y}\|_*, |u-\bar{u}|\} \\
&\le \max\{\|A\|,1\}\inf_{(\bar{y},\bar{u})\in L_{\widetilde{Y}_d}\setminus\widetilde{Y}_d} (\|y-\bar{y}\|_* + |u-\bar{u}|) \\
&= \max\{\|A\|,1\}\inf_{(\bar{y},\bar{u})\in L_{\widetilde{Y}_d}\setminus\widetilde{Y}_d} \|(y,u)-(\bar{y},\bar{u})\|_* \\
&= \max\{\|A\|,1\}\mathrm{dist}((y,u),\mathrm{rel}\partial\widetilde{Y}_d) \ .
\end{aligned}$$

Inequality (3.) follows from the observation that $Y_d = \widetilde{Y}_d \cap (C_Y^* \times I\!R)$, Lemma 7, and the bounds obtained in (1.) and (2.).

The proof of item (4.) uses the fact (soon to be proved) that if $\bar{y} \in L_{\Pi Y_d} \setminus \Pi Y_d$ then for any $\bar{u}$, $(\bar{y},\bar{u}) \in L_{Y_d} \setminus Y_d$. Then from the definition of the relative distance to the boundary we have

$$\begin{aligned}
\mathrm{dist}\,((y,u),\mathrm{rel}\partial Y_d) &= \inf_{(\bar{y},\bar{u})\in L_{Y_d}\setminus Y_d} \|(y,u)-(\bar{y},\bar{u})\|_* \\
&\le \inf_{\bar{y}\in L_{\Pi Y_d}\setminus\Pi Y_d,\bar{u}} \|(y,u)-(\bar{y},\bar{u})\|_* \\
&= \inf_{\bar{y}\in L_{\Pi Y_d}\setminus\Pi Y_d,\bar{u}} \|y-\bar{y}\|_* + |u-\bar{u}| \\
&= \inf_{\bar{y}\in L_{\Pi Y_d}\setminus\Pi Y_d} \|y-\bar{y}\|_* \\
&= \mathrm{dist}\,(y,\mathrm{rel}\partial\Pi Y_d) \ ,
\end{aligned}$$

which proves inequality (4.). To finish the proof we now show that if $\bar{y} \in L_{\Pi Y_d} \setminus \Pi Y_d$ then for any $\bar{u}$, $(\bar{y}, \bar{u}) \in L_{Y_d} \setminus Y_d$. The fact that $\bar{y} \notin \Pi Y_d$ implies that for any $\bar{u}$, $(\bar{y}, \bar{u}) \notin Y_d$. Now since $\bar{y} \in L_{\Pi Y_d}$, there exists $\alpha_i \in I\!R$, $(y_i, u_i) \in Y_d$, $i \in I$ a finite set such that $\bar{y} = \sum_{i \in I} \alpha_i y_i$ and $\sum_{i \in I} \alpha_i = 1$. Since for any $(y, u) \in Y_d$ and $\beta \geq 0$ the point $(y, u + \beta) \in Y_d$, we can express the point $(\bar{y}, \bar{u})$ by the following sum of points in Y_d

$$(\bar{y}, \bar{u}) = \left(\sum_{i \in I} \alpha_i y_i + y - y, \sum_{i \in I} \alpha_i u_i + u + \beta_1 - u - \beta_2 \right) ,$$

where for any $\bar{u}$, $\beta_1 = (\bar{u} - \sum_{i \in I} \alpha_i u_i)^+$ and $\beta_2 = (\bar{u} - \sum_{i \in I} \alpha_i u_i)^-$. This shows that for any $\bar{u}$, $(\bar{y}, \bar{u}) \in L_{Y_d}$, completing the proof. ■

References

[1] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear Programming, Theory and Algorithms.* John Wiley & Sons, Inc, New York, second edition, 1993.

[2] F. Cucker and J. Peña. A primal-dual algorithm for solving polyhedral conic systems with a finite-precision machine. Technical report, GSIA, Carnegie Mellon University, 2001.

[3] M. Epelman and R. M. Freund. A new condition measure, preconditioners, and relations between different measures of conditioning for conic linear systems. *SIAM Journal on Optimization*, 12(3):627–655, 2002.

[4] S. Filipowski. On the complexity of solving sparse symmetric linear programs specified with approximate data. *Mathematics of Operations Research*, 22(4):769–792, 1997.

[5] S. Filipowski. On the complexity of solving feasible linear programs specified with approximate data. *SIAM Journal on Optimization*, 9(4):1010–1040, 1999.

[6] R. M. Freund and J. R. Vera. Condition-based complexity of convex optimization in conic linear form via the ellipsoid algorithm. *SIAM Journal on Optimization*, 10(1):155–176, 1999.

[7] R. M. Freund and J. R. Vera. On the complexity of computing estimates of condition measures of a conic linear system. Technical Report, Operations Research Center, MIT, August 1999.

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

[9] J. L. Goffin. The relaxation method for solving systems of linear inequalities. *Mathematics of Operations Research*, 5(3):388–414, 1980.

[10] M. A. Nunez and R. M. Freund. Condition measures and properties of the central trajectory of a linear program. *Mathematical Programming*, 83(1):1–28, 1998.

[11] M. A. Nunez and R. M. Freund. Condition-measure bounds on the behavior of the central trajectory of a semi-definite program. *SIAM Journal on Optimization*, 11(3):818–836, 2001.

[12] F. Ordóñez. *On the Explanatory Value of Condition Numbers for Convex Optimization: Theoretical Issues and Computational Experience.* PhD thesis, Massachusetts Institute of Technology, 2002.

[13] F. Ordóñez and R. M. Freund. Computational experience and the explanatory value of condition measures for linear optimization. Working Paper OR361-02, MIT, Operations Research Center, 2002.

[14] J. Peña. Computing the distance to infeasibility: theoretical and practical issues. Technical report, Center for Applied Mathematics, Cornell University, 1998.

[15] J. Peña and J. Renegar. Computing approximate solutions for convex conic systems of constraints. *Mathematical Programming*, 87(3):351–383, 2000.

[16] J. Renegar. Some perturbation theory for linear programming. *Mathematical Programming*, 65(1):73–91, 1994.

[17] J. Renegar. Linear programming, complexity theory, and elementary functional analysis. *Mathematical Programming*, 70(3):279–351, 1995.

[18] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, New Jersey, 1997.

[19] J. R. Vera. Ill-posedness and the computation of solutions to linear programs with approximate data. Technical Report, Cornell University, May 1992.

[20] J. R. Vera. *Ill-Posedness in Mathematical Programming and Problem Solving with Approximate Data.* PhD thesis, Cornell University, 1992.

[21] J. R. Vera. Ill-posedness and the complexity of deciding existence of solutions to linear programs. *SIAM Journal on Optimization*, 6(3):549–569, 1996.

[22] J. R. Vera. On the complexity of linear programming under finite precision arithmetic. *Mathematical Programming*, 80(1):91–123, 1998.