MATHEMATICS OF OPERATIONS RESEARCH
Vol. 30, No. 1, February 2005, pp. 173–194
ISSN 0364-765X | EISSN 1526-5471 | 05 | 3001 | 0173

DOI 10.1287/moor.1040.0120

On an Extension of Condition Number Theory to Nonconic Convex Optimization

Robert M. Freund
Sloan School of Management, Massachusetts Institute of Technology,
50 Memorial Drive, Cambridge, Massachusetts 02142, rfreund@mit.edu

Fernando Ordóñez
Industrial and Systems Engineering, University of Southern California, GER-247,
Los Angeles, California 90089-0193, fordon@usc.edu

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

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

to the more general nonconic format:

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

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
MSC2000 subject classification: Primary: 90C25, 90C31, 49K40, 65K99; secondary: 90C22
OR/MS subject classification: Primary: Programming/nonlinear/theory; mathematics/convexity
History: Received February 16, 2003; revised February 11, 2004, and May 11, 2004.

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

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

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

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

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 Renegar [16], Filipowski [4, 5], Freund and Vera [6, 7, 8], Vera [19, 20, 21, 22], Peña [14], Peña and Renegar [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 Nunez and Freund [10, 11]). The connection of these condition measures to the complexity of algorithms has been shown in Freund and Vera [6, 7], Renegar [17], Cucker and Peña [2], and Epelman and Freund [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 Ordóñez and Freund [13, §2]).

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

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

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 $l \le x \le u$ where some components of l and/or u might be unbounded, or P might be the solution of network flow constraints of the form $Nx = g$, $x \ge 0$. 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 §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 §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 §4, we provide characterizations of the condition number as the solution to associated optimization problems. In §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 §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 $\mathscr{X}$ by $\mathbb{R}^n$ and the constraint space $\mathscr{Y}$ 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 $\mathscr{X}^*$ and $\mathscr{Y}^*$ 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 $\mathscr{D}$ 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, relative interior, and complement of Q by $\operatorname{cl} Q$, $\operatorname{relint} Q$, and Q^C, respectively. We use the convention that if Q is the singleton $Q = \{q\}$, then $\operatorname{relint} Q = Q$. We adopt the standard conventions $1/0 = \infty$ and $1/\infty = 0$.

We also make the following two general assumptions:

Assumption 1.1. *$P \neq \varnothing$ and $C_Y \neq \varnothing$.*

Assumption 1.2. *Either $C_Y \neq \mathbb{R}^m$ or P is not bounded (or both).*

Clearly, if either $P = \varnothing$ or $C_Y = \varnothing$, problem (GP_d) is infeasible regardless of A, b, and c. Therefore, Assumption 1.1 avoids settings wherein all problem instances are trivially inherently infeasible. Assumption 1.2 is needed to avoid settings where (GP_d) is feasible for every $d = (A, b, c) \in \mathscr{D}$. This will be explained further in §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 $\mathscr{F}_P := \{d \in \mathscr{D} \mid X_d \neq \varnothing\}$, i.e., $\mathscr{F}_P$ 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

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

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:

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

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

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

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

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

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 rewrite 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 rewrite (10) as

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

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

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

Let $\mathcal{F}_D := \{d \in \mathcal{D} \mid Y_d \neq \varnothing\}$, 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

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

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 Rockafellar [18, Theorem 5.5] that $u(\cdot)$, the support function of the set $-P$, is a convex function. The epigraph of $u(\cdot)$ is

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

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

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

It then follows from (9) that

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

and so (GD_d) can alternatively be written as

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

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 2.1. *If y satisfies $y \in C_Y^*$ and $c - A^t y \in \operatorname{relint} R^*$, 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 A.3 in the appendix that $\operatorname{cl} \operatorname{effdom} u(\cdot) = R^*$ and from Proposition A.4 in the appendix that $c - A^t y \in \operatorname{relint} R^* = \operatorname{relint} \operatorname{cl} \operatorname{effdom} u(\cdot) = \operatorname{relint} \operatorname{effdom} u(\cdot) \subseteq \operatorname{effdom} u(\cdot)$. 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 $u \geq u(c - A^t y)$. $\square$

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

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

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 (see Renegar [16, 17]).

2.4. Basic properties of $\rho_P(d)$, $\rho_D(d)$, and $C(d)$ and alternative duality results. The need for Assumptions 1.1 and 1.2 is demonstrated by the following:

Proposition 2.2. *For any data instance $d \in \mathcal{D}$,*
1. *$\rho_P(d) = \infty$ if and only if $C_Y = \mathbb{R}^m$, and*
2. *$\rho_D(d) = \infty$ if and only if P is bounded.*

The proof of this proposition relies on Lemmas 2.1 and 2.2, 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 2.2. 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 2.1 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 $(y, u) = (\tilde{y}, \tilde{y}^t \tilde{y}/2)$ satisfies the alternative system $(A2_{\bar{d}})$ of

Lemma 2.1 for the data $\bar{d} = (0, \tilde{y}, 0)$, whereby $X_{\bar{d}} = \varnothing$. 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 2.2 for the data $\bar{d} = d + \Delta d = (0, 0, -\tilde{x})$, whereby $Y_{\bar{d}} = \varnothing$. Therefore, $\|\bar{d} - d\| \geq \rho_D(d) = \infty$, a contradiction, and so P is bounded. $\square$

Remark 2.1. The set $\mathcal{F} \neq \varnothing$, and if $d \in \mathcal{F}$, then $C(d) \geq 1$.

Proof. If $C_Y \neq \mathbb{R}^m$, consider $b \in \mathbb{R}^m \backslash C_Y$ (hence $b \neq 0$), and for any $\varepsilon > 0$ define the instance $d_\varepsilon = (0, -\varepsilon b, 0)$. This instance is such that for any $\varepsilon > 0$, $X_{d_\varepsilon} = \varnothing$, 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 $C_Y = \mathbb{R}^m$, then Assumption 1.2 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 $Y_{d_\varepsilon} = \varnothing$, 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. $\square$

The following two lemmas present weak and strong alternative results for (GP_d) and (GD_d), and are used in the proofs of Proposition 2.2 and elsewhere.

Lemma 2.1. *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 A.2 we can separate b from S and therefore there exists $y \neq 0$ such that

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

Setting $u := y^t b$, this 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 there exist $x \in P$, $u \in \mathbb{R}$, and $y \in C_Y^*$ such that

$$0 \leq y^t(Ax - b) = (A^t y)^t x - b^t y < -((-A^t y)^t x + u) \leq 0. \quad \square$$

Lemma 2.2. *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

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

which is a nonempty convex set. Using Proposition A.2 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 $s=-(c-A^ty)+\tilde{s}$, $v=-u+\tilde{v}$, 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

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

The above inequality implies that $\delta=0$, $Ax=z\in C_Y$, $x\in R$, and $c^tx\leq 0$. 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

$$0\leq x^t(c-A^ty)=c^tx-y^tAx<-y^tAx\leq 0. \quad \square$$

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 3.1. *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,

$$0\leq (c-A^ty)^tx+u=c^tx-y^tAx+u\leq c^tx-b^ty+u,$$

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

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

DEFINITION 3.1. A point x is a Slater point for problem (GP_d) if

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

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

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

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

THEOREM 3.2. *If (y',u') is a Slater point for problem (GD_d), then $z_*(d)=z^*(d)$. If in addition $z^*(d)<\infty$, then $X_d\neq\varnothing$ 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 points 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 Freund and Vera [8, Theorems 15, 17, and 19] and Renegar [17, Theorem 3.1].

THEOREM 3.3. *Suppose that $\rho_P(d)>0$. Then, there exists a Slater point for (GP_d).*

THEOREM 3.4. *Suppose that $\rho_D(d)>0$. Then, there exists a Slater point for (GD_d).*

COROLLARY 3.1 (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. This result is a straightforward consequence of Theorems 3.1, 3.2, 3.3, and 3.4. $\square$

Note that the contrapositive of Corollary 3.1 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 3.1. 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 nonempty convex set

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

We use Proposition A.2 in the appendix 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 $\pi > 0$, rescale (γ, y, π) such that $\pi = 1$. For any $x \in \mathbb{R}^n$, $\tilde{p} \in P$, $\tilde{q} \in C_Y$, 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,

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

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

The proof of Theorem 3.2 uses arguments that parallel those used in the proof of Theorem 3.1, and so is omitted. We refer the interested reader to Ordóñez [12, Theorem 4].

Proof of Theorem 3.3. Equation (6) and $\rho_P(d) > 0$ imply that $X_d \neq \varnothing$. Assume that X_d contains no Slater point. Then, $\operatorname{relint} C_Y \cap \{Ax - b \mid x \in \operatorname{relint} P\} = \varnothing$ and these nonempty convex sets can be separated using Proposition A.2. Therefore, there exists $y \neq 0$ such that for any $s \in C_Y$, $x \in P$, we have

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

From the inequality above and setting $u = y^t b$, 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 $b_\varepsilon = b + (\varepsilon / \|y\|_*)\hat{y}$, with $\hat{y}$ given by Proposition A.1 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 2.1 with data $d_\varepsilon = (A, b_\varepsilon, c)$ for any $\varepsilon > 0$. This implies that $X_{d_\varepsilon} = \varnothing$ and therefore $\rho_P(d) \le \inf_{\varepsilon>0} \|d - d_\varepsilon\| = \inf_{\varepsilon>0} \varepsilon/\|y\|_* = 0$, a contradiction. □

The proof of Theorem 3.4 uses arguments that parallel those used in the proof of Theorem 3.3, and so is omitted. We refer the interested reader to Ordóñez [12, Theorem 8].

The contrapositives of Theorems 3.3 and 3.4 are not true. Consider, for example, the data

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

and the sets $C_Y = \mathbb{R}_+ \times \{0\}$ and $P = C_X = \mathbb{R}_+ \times \mathbb{R}$. 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 Renegar [17, Theorem 3.5] and Freund and Vera [8, Theorems 1 and 2].

THEOREM 4.1. *Suppose that $X_d \neq \varnothing$. Then, $\rho_P(d) = j_P(d) = r_P(d)$, where*

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

and

$$r_P(d) = \begin{array}{ll} \min\limits_{v} & \begin{array}{ll} \max\limits_{x,t,\theta} & \theta \end{array} \\ \|v\| \le 1 & Ax - bt - v\theta \in C_Y \\ v \in \mathbb{R}^m & \|x\| + |t| \le 1 \\ & (x, t) \in C. \end{array} \tag{18}$$

THEOREM 4.2. *Suppose that $Y_d \neq \varnothing$. Then, $\rho_D(d) = j_D(d) = r_D(d)$, where*

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

and

$$r_D(d) = \begin{array}{ll} \min\limits_{v} & \begin{array}{ll} \max\limits_{y,\delta,\theta} & \theta \end{array} \\ \|v\|_* \le 1 & -A^t y + c\delta - \theta v \in R^* \\ v \in \mathbb{R}^n & \|y\|_* + |\delta| \le 1 \\ & y \in C_Y^* \\ & \delta \ge 0. \end{array} \tag{20}$$

PROOF OF THEOREM 4.1. 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 2.1 there is a point $(\bar{y}, \bar{u})$ that satisfies the following:

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

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

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

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

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

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

$$\begin{aligned}
\bar{A} &= A - \hat{y}((A^t\bar{y})^t + \bar{s}^t),\\
\bar{b}_\varepsilon &= b - \hat{y}(b^t\bar{y} - \bar{u} - \varepsilon).
\end{aligned}$$

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 $\varepsilon > 0$, Problem $(A2_{\bar{d}_\varepsilon})$ in Lemma 2.1 is feasible with data $\bar{d}_\varepsilon = (\bar{A}, \bar{b}_\varepsilon, c)$. Lemma 2.1 then implies that $X_{\bar{d}_\varepsilon} = \varnothing$ and therefore $\rho_P(d) \leq \|d - \bar{d}_\varepsilon\|$. To finish the proof we compute the size of the perturbation:

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

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

To prove the other characterization, note we can add $\theta \geq 0$ to (18) and then invoke Lemma A.1 to rewrite it as

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

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

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

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 A.1. 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,

$$r_P(d) = \min_{y,s,u} \quad \max\{\|A^t y + s\|_*, |-b^t y + u|\} = j_P(d)$$
$$\|y\|_* = 1$$
$$y \in C_Y^*$$
$$(s, u) \in C^*. \quad \square$$

The proof of Theorem 4.2 uses arguments that parallel those used in the proof of Theorem 4.1, and so is omitted. We refer the interested reader to Ordóñez [12, Theorem 6].

5. Geometric properties of the primal and dual feasible regions. In §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 and 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

$$L_S = \left\{ \sum_{i\in I} \alpha_i x_i \;\middle|\; \alpha_i \in \mathbb{R}, x_i \in S, \sum_{i\in I} \alpha_i = 1, I \text{ a finite set} \right\}$$

(see Rockafellar [18, §1]). We denote by $\hat{L}_S$ the vector subspace obtained when the affine hull L_S is translated to contain the origin; i.e., for any $x \in S$, $\hat{L}_S = L_S - x$. 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 $\operatorname{dist}(x, \operatorname{rel}\partial S)$, defined as follows:

Definition 5.1. Given a nonempty set S and a point $x \in S$, the distance from x to the relative boundary of S is

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

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

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

Definition 5.2. For a convex cone K, the min-width of K is defined by

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

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 Freund and Vera [6]). Note that if K is not a subspace, then $\tau_K \in (0, 1]$, and τ_K is attained for some $y^0 \in \operatorname{relint} K$ 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 Renegar [16, Lemma 3.2] to the ground-set model format.

PROPOSITION 5.1. *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*

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

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

PROPOSITION 5.2. *Consider any $x^0 \in P$. If $\rho_P(d) > 0$, then there exists $\bar{x} \in X_d$ satisfying*

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

The following is the main result of this subsection, and can be viewed as an extension of Freund and Vera [8, Theorems 15, 17, and 19] to the ground-set model format of (GP_d). In Theorem 5.1 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 5.2.

THEOREM 5.1. *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) $\dfrac{1}{\operatorname{dist}(\bar{x}, \operatorname{rel}\partial P)} \leq \dfrac{1}{\operatorname{dist}(x^0, \operatorname{rel}\partial P)} \left(1 + \dfrac{\|Ax^0 - b\| + \|A\|}{\rho_P(d)}\right)$,

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

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

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

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

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

The statement of Theorem 5.2 below relaxes the assumptions on P and C_Y not being affine and/or linear spaces:

THEOREM 5.2. *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 5.1.*
- *If P is an affine set and C_Y is not a subspace, $\bar{x}$ satisfies all items of Theorem 5.1, 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* 5.1, *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 5.1 and 5.2, 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:

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

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

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

The following is an immediate consequence of Theorems 5.1 and 5.2.

Corollary 5.1. *Consider any $x^0 \in P$. If $\rho_P(d) > 0$, then*

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

We now proceed with the proofs of these results.

Proof of Proposition 5.1. If $r = 0$, the result is true. If $r \neq 0$, then Proposition A.1 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 2.2, and therefore $(GD_{\bar{d}})$ is infeasible. We conclude that $\rho_D(d) \le \|d - \bar{d}\|$, which implies

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

and so

$$\rho_D(d) \le \frac{\max\{\|A\hat{x} - b\|, c^t r\}}{\|r\|}. \quad \square$$

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

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

Lemma 5.1. *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*

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

Proof. Note that problem (PP) is feasible for any x^0 and w^0 because $(x, t, w, \theta) = (0, 0, 0, 0)$ 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. Because $\rho_P(d) > 0$ implies $X_d \neq \varnothing$, Theorem 4.1 implies that the optimal solution $(x^*, t^*, w^*, \theta^*)$ for (PP) satisfies (23). $\square$

Proof of Proposition 5.2. Assume that $Ax^0 - b \notin C_Y$, otherwise $\bar{x} = x^0$ satisfies the proposition. We consider problem (PP), defined by (22), with $w^0 \in C_Y$ chosen such that $\|Ax^0 - b - w^0\| = \operatorname{dist}(Ax^0 - b, C_Y)$. From Lemma 5.1 we have that there exists a point (x, t, w, θ) feasible for (PP) that satisfies

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

Define

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

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

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

Proof of Theorem 5.1. Note that $\rho_P(d) > 0$ implies $X_d \neq \varnothing$; note also that $\rho_P(d)$ is finite, otherwise Proposition 2.2 shows that $C_Y = \mathbb{R}^m$, which is a subspace. For convenience we suppose for now that $A \neq 0$. Set $w^0 \in C_Y$ such that $\|w^0\| = \|A\|$ and $\tau_{C_Y} = \operatorname{dist}(w^0, \operatorname{rel}\partial C_Y)/\|w^0\|$. We also assume that $Ax^0 - b \neq w^0$, otherwise we can show that $\bar{x} = x^0$ satisfies the theorem. Let $r_{w^0} = \operatorname{dist}(w^0, \operatorname{rel}\partial C_Y) = \|A\|\tau_{C_Y}$ and let also $r_{x^0} = \operatorname{dist}(x^0, \operatorname{rel}\partial P)$. We invoke Lemma 5.1 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:

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

By construction $\operatorname{dist}(\bar{x}, \operatorname{rel}\partial P) \geq r_{\bar{x}}$, $\operatorname{dist}(\bar{w}, \operatorname{rel}\partial C_Y) \geq r_{\bar{w}}\|A\|$, 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 $v \in \hat{L}_P \cap \{y \mid Ay \in L_{C_Y}\}$ such that $\|v\| \leq 1$. Then,

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

and

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

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

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) $$\frac{1}{\operatorname{dist}(\bar{x}, \operatorname{rel}\partial P)} \le \frac{1}{r_{\bar{x}}} = \frac{t+\theta}{\theta r_{x^0}} \le \frac{1}{r_{x^0}}\left(1+\frac{1}{\theta}\right) \le \frac{1}{r_{x^0}}\left(1+\frac{\|Ax^0-b\|+\|A\|}{\rho_P(d)}\right),$$

(b) $$\frac{1}{\operatorname{dist}(\bar{x}, \operatorname{rel}\partial X_d)} \le \frac{1}{\min\{r_{x^0}, \tau_{C_Y}\}}\frac{t+\theta}{\theta} \le \frac{1}{\min\{r_{x^0}, \tau_{C_Y}\}}\left(1+\frac{1}{\theta}\right)$$
$$\le \frac{1}{\min\{r_{x^0}, \tau_{C_Y}\}}\left(1+\frac{\|Ax^0-b\|+\|A\|}{\rho_P(d)}\right).$$

3. (a) $$\frac{\|\bar{x}-x^0\|}{\operatorname{dist}(\bar{x}, \operatorname{rel}\partial P)} \le \frac{\|x-tx^0\|}{\theta r_{x^0}} \le \frac{1}{r_{x^0}}\frac{1}{\theta}\max\{1, \|x^0\|\}$$
$$\le \frac{1}{r_{x^0}}\frac{\|Ax^0-b\|+\|A\|}{\rho_P(d)}\max\{1, \|x^0\|\}.$$

(b) $$\frac{\|\bar{x}-x^0\|}{\operatorname{dist}(\bar{x}, \operatorname{rel}\partial X_d)} \le \frac{\|x-tx^0\|}{\theta\min\{r_{x^0}, \tau_{C_Y}\}} \le \frac{1}{\min\{r_{x^0}, \tau_{C_Y}\}}\frac{1}{\theta}\max\{1, \|x^0\|\}$$
$$\le \frac{1}{\min\{r_{x^0}, \tau_{C_Y}\}}\frac{\|Ax^0-b\|+\|A\|}{\rho_P(d)}\max\{1, \|x^0\|\}.$$

(c) $$\frac{\|\bar{x}\|}{\operatorname{dist}(\bar{x}, \operatorname{rel}\partial P)} \le \frac{\|x+\theta x^0\|}{\theta r_{x^0}} \le \frac{1}{r_{x^0}}\left(\|x^0\|+\frac{1}{\theta}\right)$$
$$\le \frac{1}{r_{x^0}}\left(\|x^0\|+\frac{\|Ax^0-b\|+\|A\|}{\rho_P(d)}\right).$$

(d) $$\frac{\|\bar{x}\|}{\operatorname{dist}(\bar{x}, \operatorname{rel}\partial X_d)} \le \frac{\|x+\theta x^0\|}{\theta\min\{r_{x^0}, \tau_{C_Y}\}} \le \frac{1}{\min\{r_{x^0}, \tau_{C_Y}\}}\left(\|x^0\|+\frac{1}{\theta}\right)$$
$$\le \frac{1}{\min\{r_{x^0}, \tau_{C_Y}\}}\left(\|x^0\|+\frac{\|Ax^0-b\|+\|A\|}{\rho_P(d)}\right).$$

Finally, note that in the case when $A=0$, the point $\bar{x}=x^0$ is feasible and thus satisfies the theorem. ☐

We note that Theorem 5.2 can be proved using almost identical arguments as in the proof of Theorem 5.1, but with a careful analysis to handle the special cases when P is an affine set or C_Y is a subspace (see Ordóñez [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). Due to the similarity to the primal case, we omit all proofs in this subsection and refer the reader to Ordóñez [12, Chapter 4] for detailed proofs of these results. 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 $(x,t)\in C\subset\mathbb{R}^n\times\mathbb{R}$. This then leads to the following dual norm for points $(s,u)\in C^*\subset\mathbb{R}^n\times\mathbb{R}$:

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

Consistent with the characterization of $\rho_D(d)$ given by (20) in Theorem 4.2, we define the following dual norm for points $(y,\delta)\in\mathbb{R}^m\times\mathbb{R}$:

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

It is clear that the above defines a norm on the vector space $\mathbb{R}^m\times\mathbb{R}$ 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 Renegar [16, Lemma 3.1] for the ground-set model format.

Proposition 5.3. *Consider any (y, u) feasible for (GD_d). If $\rho_P(d) > 0$, then*

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

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

Proposition 5.4. *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*

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

The following is the main result of this subsection, and can be viewed as an extension of Freund and Vera [8, Theorems 15, 17, and 19] to the dual problem (GD_d). In Theorem 5.3 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 5.4.

Theorem 5.3. *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) $\|\bar{y} - y^0\|_* \le \dfrac{\|c - A^t y^0\|_* + \|A\|}{\rho_D(d)} \max\{1, \|y^0\|_*\}$,

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

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

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

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

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

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

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

The statement of Theorem 5.4 below relaxes the assumptions on R and C_Y not being linear subspaces:

Theorem 5.4. *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 5.3.*
- *If C_Y is a subspace and R is not a subspace, $(\bar{y}, \bar{u})$ satisfies all items of Theorem 5.3, 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* 5.3, *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) $\dfrac{1}{\operatorname{dist}((\bar{y}, \bar{u}), \operatorname{rel} \partial Y_d)} \le \varepsilon.$

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

  3′. (d) $\dfrac{\|\bar{y}\|_*}{\operatorname{dist}((\bar{y}, \bar{u}), \operatorname{rel} \partial Y_d)} \le \varepsilon.$

The next result captures the thrust of Theorems 5.3 and 5.4, 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 $y^0 \in \operatorname{relint} C_Y^*$, define

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

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:

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

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:

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

Corollary 5.2. *Consider any $y^0 \in C_Y^*$. If $\rho_D(d) > 0$, then*

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

We conclude this subsection with a technical lemma that concerns the optimization problem (DP) below. Problem (DP) is parameterized by given points $y^0 \in C_Y^*$ and $s^0 \in R^*$, and is defined by

$$\begin{array}{lll}
(DP) & \displaystyle\max_{y, \delta, s, \theta} & \theta \\
& \text{s.t.} & -A^t y + \delta c - s = \theta\left(A^t y^0 - c + s^0\right), \\
& & \|y\|_* + |\delta| \le 1, \\
& & y \in C_Y^*, \\
& & \delta \ge 0, \\
& & s \in R^*.
\end{array} \tag{28}$$

Lemma 5.2. *Consider any $y^0 \in C_Y^*$ and $s^0 \in R^*$ such that $A^t y^0 + s^0 \ne c$. If $\rho_D(d) > 0$, then there exists a point (y, δ, s, θ) feasible for problem (DP) that satisfies*

$$\theta \ge \frac{\rho_D(d)}{\|c - A^t y^0 - s^0\|_*} > 0. \tag{29}$$

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 Renegar [16, Theorem 1.1, Assertion 2] to the primal and dual problems in the GSM format.

THEOREM 6.1. *Suppose that $\rho_P(d) > 0$. Let $\Delta d = (\Delta A, \Delta b, \Delta c)$ be such that $X_{d+\Delta d} \neq \varnothing$ 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 6.2. *Suppose that $\rho_D(d) > 0$. Let $\Delta d = (\Delta A, \Delta b, \Delta c)$ be such that $Y_{d+\Delta d} \neq \varnothing$ 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 6.1 and Theorem 6.3 below extend, respectively, Renegar [16, Lemma 3.9 and Theorem 1.1, Assertion 5] to the ground-set model format.

PROPOSITION 6.1. *Suppose that $d \in \mathcal{F}$ and $\rho(d) > 0$. Let $\Delta d = (0, \Delta b, 0)$ be such that $X_{d+\Delta d} \neq \varnothing$. Then,*

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

THEOREM 6.3. *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 6.1. The result is trivial if $\Delta A = 0$ and $\Delta b = 0$, so we presume that $\Delta A \neq 0$ and/or $\Delta b \neq 0$. We consider problem (PP), defined by (22), with $x^0 = x'$ and w^0 such that $(A + \Delta A)x' - (b + \Delta b) = w^0 \in C_Y$. Let us first suppose that $b - Ax^0 + w^0 \neq 0$. From Lemma 5.1 we have that there exists a point (x, t, w, θ) feasible for (PP) that satisfies

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

On the other hand, if $b - Ax^0 + w^0 = 0$, then it is trivial to show that there exists a point (x, t, w, θ) feasible for (PP) that satisfies

$$\theta \geq \frac{\rho_P(d)}{\|\Delta A\| \|x'\| + \|\Delta b\|}.$$

We define

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

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

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

Proof of Theorem 6.2. From Proposition A.3 we have that for any $\varepsilon > 0$ there exists $\xi \neq \Delta A^t y' - \Delta c$ such that $\|\xi\|_* \leq \varepsilon$ and $c + \Delta c + \xi - (A + \Delta A)^t y' \in \operatorname{relint} R^*$. We consider problem (DP) defined by (28), with $y^0 = y'$ and $s^0 := c + \Delta c + \xi - (A + \Delta A)^t y' \in \operatorname{relint} R^*$. From Lemma 5.2 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

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

By construction we have that $\bar{y} \in C_Y^*$ and $c - A^t \bar{y} = \bar{s} \in \operatorname{relint} R^* \subseteq \operatorname{effdom} u(\cdot)$ from Propositions A.3 and A.4. Therefore, from Proposition 2.1, $(\bar{y}, u(c - A^t \bar{y})) \in Y_d$ and

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

Proof of Proposition 6.1. 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 3.1). Also, because $Y_{d+\Delta d} = Y_d \neq \varnothing$ has a Slater point (because $\rho_D(d) > 0$) and $X_{d+\Delta d} \neq \varnothing$, 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 3.2). 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 5.3 and the strong duality for data instances d and $d + \Delta d$. $\square$

Proof of Theorem 6.3. 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 3.1).

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 6.1, 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,

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

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

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

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 $-z^*(d+\Delta d) \le -z^*(d)$, then we can do the replacement because 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 $\rho_P(d+\Delta d) \ge \rho_P(d) - \|\Delta d\| > 0$. 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 $d+\Delta d$. $\square$

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 nonconic ground-set model format (GP_d). We have attempted herein to highlight the most important and/or useful extensions; for other results see Ordóñez [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 Renegar [16, Theorem 1.1, Assertions 3, 4]. 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 4.2 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. Curiously, this loss of information is not present in the characterization of the primal distance to infeasibility; the characterization of $\rho_P(d)$ uses all of the information about P through its conic extension C, as shown in Theorem 4.1. Under rather mild additional assumptions, it is possible to analyze the complexity of algorithms for solving (GP_d) (see Ordóñez [12]).

Note that the characterization results for $\rho_P(d)$ and $\rho_D(d)$ presented herein in Theorems 4.1 and 4.2 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.

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

Proposition A.1 (Freund and Vera [8, Proposition 2]). *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 A.2 (Rockafellar [18, Theorems 11.1 and 11.3]). *Given two nonempty convex sets S and T in $\mathbb{R}^n$, then $\operatorname{relint} S \cap \operatorname{relint} T = \varnothing$ if and only if S and T can be properly separated, i.e., there exists $y \neq 0$ such that*

$$\inf_{x\in S} y^t x \ge \sup_{z\in T} y^t z,$$
$$\sup_{x\in S} y^t x > \inf_{z\in T} y^t z.$$

The following is a restatement of Rockafellar [18, Corollary 14.2.1] 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 A.3 (Rockafellar [18, Corollary 14.2.1]). *Let R denote the recession cone of the nonempty convex set P and define $u(\cdot)$ by (14). Then,* $\operatorname{cl}\operatorname{effdom} u(\cdot) = R^*$.

Proposition A.4 (Rockafellar [18, Theorem 6.3]). *For any convex set $Q \subseteq \mathbb{R}^n$,* $\operatorname{cl}\operatorname{relint} Q = \operatorname{cl} Q$ *and* $\operatorname{relint}\operatorname{cl} Q = \operatorname{relint} Q$.

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

Lemma A.1. *Consider two nonempty closed convex cones $C \subseteq \mathbb{R}^n$ and $C_Y \subseteq \mathbb{R}^m$, and data $(M, v) \in \mathbb{R}^{m\times n} \times \mathbb{R}^m$. Strong duality holds between*

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

Proof. The proof that weak duality holds between (P) and (D) is straightforward. Therefore, $z^* \leq z_*$. Note that if $z_* = \infty$, then $-v \in C_Y$, and so $z^* = \infty = z_*$. Let us therefore assume $z^* < z_* < \infty$ and set $\varepsilon > 0$ such that $0 \leq z^* < z_* - \varepsilon$. Consider the following nonempty convex set S:

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

Then, $(0, 0, 0) \notin S$, and from Proposition A.2 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 $y \in \mathbb{R}^m$, $\tilde{u} \in C_Y^*$, $\tilde{\delta} \in C^*$, $\pi \geq 0$, 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

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

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 $x \neq 0$, rescale (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 A.1 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 because $z^* < z_*$. If $\theta = 0$, then $z = 0$, which is a contradiction because $(z, x, \theta) \neq 0$. $\square$

References

[1] Bazaraa, M. S., H. D. Sherali, C. M. Shetty. 1993. *Nonlinear Programming, Theory and Algorithms*, 2nd ed. John Wiley & Sons, New York.

[2] Cucker, F., J. Peña. 2002. A primal-dual algorithm for solving polyhedral conic systems with a finite-precision machine. *SIAM J. Optim.* **12**(2) 522–554.

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

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

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

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

[7] Freund, R. M., J. R. Vera. 2003. On the complexity of computing estimates of condition measures of a conic linear system. *Math. Oper. Res.* **28**(4) 625–648.

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

[9] Goffin, J. L. 1980. The relaxation method for solving systems of linear inequalities. *Math. Oper. Res.* **5**(3) 388–414.

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

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

[12] Ordóñez, F. 2002. On the explanatory value of condition numbers for convex optimization: Theoretical issues and computational experience. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

[13] Ordóñez, F., R. M. Freund. 2003. Computational experience and the explanatory value of condition measures for linear optimization. *SIAM J. Optim.* **14**(2) 307–333.

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

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

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

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

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

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

[20] Vera, J. R. 1992. Ill-posedness in mathematical programming and problem solving with approximate data. Ph.D. thesis, Cornell University, Ithaca, NY.

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

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