SOME CHARACTERIZATIONS AND PROPERTIES OF THE "DISTANCE TO ILL-POSEDNESS" AND THE CONDITION MEASURE OF A CONIC LINEAR SYSTEM [1]

Robert M. Freund[2]
M.I.T.

Jorge R. Vera[3]
Catholic University of Chile

October, 1995, Revised June, 1997, Revised June, 1998

Abstract

A conic linear system is a system of the form

$$P(d): \quad \text{find } x \text{ that solves } b - Ax \in C_Y,\ x \in C_X,$$

where C_X and C_Y are closed convex cones, and the data for the system is $d = (A, b)$. This system is "well-posed" to the extent that (small) changes in the data (A, b) do not alter the status of the system (the system remains solvable or not). Renegar defined the "distance to ill-posedness," $\rho(d)$, to be the smallest change in the data $\Delta d = (\Delta A, \Delta b)$ for which the system $P(d + \Delta d)$ is "ill-posed," i.e., $d + \Delta d$ is in the intersection of the closure of feasible and infeasible instances $d' = (A', b')$ of $P(\cdot)$. Renegar also defined the "condition measure" of the data instance d as $\mathcal{C}(d) := \|d\|/\rho(d)$, and showed that this measure is a natural extension of the familiar condition measure associated with systems of linear equations. This study presents two categories of results related to $\rho(d)$, the distance to ill-posedness, and $\mathcal{C}(d)$, the condition measure of d. The first category of results involves the approximation of $\rho(d)$ as the optimal value of certain mathematical programs. We present ten different mathematical programs each of whose optimal values provides an approximation of $\rho(d)$ to within certain constants, depending on whether $P(d)$ is feasible or not, and where the constants depend on properties of the cones and the norms used. The second category of results involves the existence of certain inscribed and intersecting balls involving the feasible region of $P(d)$ or the feasible region of its alternative system, in the spirit of the ellipsoid algorithm. These results roughly state that the feasible region of $P(d)$ (or its alternative system when $P(d)$ is not feasible) will contain a ball of radius r that is itself no more than a distance R from the origin, where the ratio R/r satisfies $R/r = c_1 O(\mathcal{C}(d))$, and such that $r = c_2 \Omega\left(\frac{1}{\mathcal{C}(d)}\right)$ and $R = c_3 O(\mathcal{C}(d))$, where c_1, c_2, c_3 are constants that depend only on properties of the cones and the norms used. Therefore the condition measure $\mathcal{C}(d)$ is a relevant tool in proving the existence of an inscribed ball in the feasible region of $P(d)$ that is not too far from the origin and whose radius is not too small.

AMS Subject Classification: 90C, 90C05, 90C60
Keywords: Complexity of Linear Programming, Infinite Programming, Interior Point Methods, Conditioning, Error Analysis

[1] This research has been partially supported through a grant from FONDECYT (project number 1930948), by NSF Grant INT-9703803, and by a research fellowship from CORE, Catholic University of Louvain.

[2] MIT O.R. Center, 77 Massachusetts Ave., Cambridge, MA 02139, USA. email: rfreund@mit.edu

[3] Department of Industrial and System Engineering, Catholic University of Chile, Campus San Joaquín, Vicuña Mackenna 4860, Santiago, CHILE. email: jvera@ing.puc.cl

"Some Characterizations and Properties of the 'Distance to Ill-Posedness' and the Condition Measure of a Conic Linear System." Freund, Robert M., and Jorge R. Vera. Mathematical Programming Vol. 86, No. 2 (1999): 225-260.

© Springer-Verlag Berlin Heidelberg 1999 http://doi.org/10.1007/s10107990063a

1 Introduction

This paper is concerned with characterizations and properties of the "distance to ill-posedness" and of the condition measure of a conic linear system, i.e., a system of the form:

$$P(d): \quad \text{find } x \text{ that solves } b - Ax \in C_Y,\ x \in C_X, \tag{1}$$

where $C_X \subset X$ and $C_Y \subset Y$ are each a closed convex cone in the (finite) n-dimensional normed linear vector space X (with norm $\|x\|$ for $x \in X$) and in the (finite) m-dimensional linear vector space Y (with norm $\|y\|$ for $y \in Y$), respectively. Here $b \in Y$, and $A \in L(X,Y)$ where $L(X,Y)$ denotes the set of all linear operators $A : X \longrightarrow Y$. At the moment, we make no assumptions on C_X and C_Y except that each is a closed convex cone. The reader will recognize immediately that when $X = R^n$ and $Y = R^m$, and either (i) $C_X = \{x \in R^n \mid x \geq 0\}$ and $C_Y = \{y \in R^m \mid y \geq 0\}$, (ii) $C_X = \{x \in R^n \mid x \geq 0\}$ and $C_Y = \{0\} \subset R^m$, or (iii) $C_X = R^n$ and $C_Y = \{y \in R^m \mid y \geq 0\}$, then $P(d)$ is a linear inequality system of the format (i) $Ax \leq b, x \geq 0$, (ii) $Ax = b, x \geq 0$, or (iii) $Ax \leq b$, respectively.

The problem $P(d)$ is a very general format for studying the feasible region of a mathematical program, and even lends itself to analysis by interior-point methods, see Nesterov and Nemirovskii[8] and Renegar [12] and [13].

The concept of the "distance to ill-posedness" and a closely related condition measure for problems such as $P(d)$ was introduced by Renegar in [10] in a more specific setting, but then generalized more fully in [11] and in [12]. We now describe these two concepts in detail.

We denote by $d = (A, b)$ the "data" for the problem $P(d)$. That is, we regard the cones C_X and C_Y as fixed and given, and the data for the problem is the linear operator A together with the vector b. We denote the set of solutions of $P(d)$ as X_d to emphasize the dependence on the data d, i.e.,

$$X_d = \{x \in X \mid b - Ax \in C_Y, x \in C_X\}.$$

We define

$$\mathcal{F} = \{(A,b) \in L(X,Y) \times Y \mid \text{ there exists x satisying } b - Ax \in C_Y, x \in C_X\} \ . \tag{2}$$

Then $\mathcal{F}$ corresponds to those data instances (A, b) for which $P(d)$ is consistent, i.e., $P(d)$ has a solution.

For $d = (A, b) \in L(X,Y) \times Y$ we define the product norm on the cartesian product $L(X,Y) \times Y$ as

$$\|d\| = \|(A,b)\| = \max\{\|A\|, \|b\|\} \tag{3}$$

where $\|b\|$ is the norm specified for Y and $\|A\|$ is the operator norm, namely

$$\|A\| = \max\{\|Ax\| \mid \|x\| \leq 1\}. \tag{4}$$

We denote the complement of $\mathcal{F}$ by $\mathcal{F}^C$. Then $\mathcal{F}^C$ consists precisely of those data instances $d = (A, b)$ for which $P(d)$ is inconsistent.

The boundary of $\mathcal{F}$ and of $\mathcal{F}^C$ is precisely the set

$$\mathcal{B} = \partial\mathcal{F} = \partial\mathcal{F}^C = cl\,(\mathcal{F}) \cap cl(\mathcal{F}^C) \tag{5}$$

where ∂S denotes the boundary of a set S and $cl(S)$ is the closure of a set S. Note that if $d = (A, b) \in \mathcal{B}$, then $P(d)$ is ill-posed in the sense that arbitrary small changes in the data $d = (A, b)$ will yield consistent instances of $P(d)$ as well as inconsistent instances of $P(d)$.

For any $d = (A, b) \in L(X, Y) \times Y$, we define

$$\begin{array}{rlllll} \rho(d) = & \inf & \|\Delta d\| & = & \inf & \|(\Delta A, \Delta b)\| \\ & \Delta d & & & \Delta A, \Delta b & \\ & s.t. & d + \Delta d \in \mathcal{B} & s.t. & (A + \Delta A, b + \Delta b) \in cl(\mathcal{F}) \cap cl(\mathcal{F}^{\mathcal{C}}) & . \end{array} \tag{6}$$

Then $\rho(d)$ is the "distance to ill-posedness" of the data d, i.e., $\rho(d)$ is the distance of d to the set $\mathcal{B}$ of ill-posedness instances. In addition to the work of Renegar cited earlier, further analysis of the distance to ill-posedness has been studied by Vera [17], [18], [16], Filipowski [4], [5], and Nunez and Freund [9].

In addition to the general case $P(d)$, we will also be interested in two special cases when one of the cones is either the entire space or only the zero-vector. When $C_Y = \{0\}$, then $P(d)$ specializes to

$$Ax = b,\ x \in C_X\ .$$

When $C_X = X$, then $P(d)$ specializes to

$$b - Ax \in C_Y,\ x \in X\ .$$

One of the purposes of this paper is to explore approximate characterizations of the distance to ill-posedness $\rho(d)$ as the optimal value of a mathematical program whose solution is relatively easy to obtain. By "relatively easy," we roughly mean that such a program is either a convex program or is solvable through $O(m)$ or $O(n)$ convex programs. Vera [17] and [16] explored such characterizations for linear programming problems, and the results herein expand the scope of this line of research in two ways: first by expanding the problem context from linear equations and linear inequalities to conic linear systems, and second by developing more efficient mathematical programs that characterize $\rho(d)$. Renegar [12] presents a characterization of the distance to ill-posedness as the solution of a certain mathematical program, but this characterization is not in general easy to solve.

There are a number of reasons for exploring various characterizations of $\rho(d)$, not the least of which is to better understand the underlying nature of $\rho(d)$. First, we anticipate that such characterization results for $\rho(d)$ will be useful in the complexity analysis of a variety of algorithms for convex optimization of problems in conic linear form. There is also the intellectual issue of the complexity of computing $\rho(d)$ or an approximation thereof, and there is the prospect of using such characterizations to further understand the behavior of the underlying problem $P(d)$. Furthermore, when an approximation of $\rho(d)$ can be computed efficiently, then there is promise that the problem

of deciding the feasibility of $P(d)$ or the infeasibility of $P(d)$ can be processed "efficiently", say in polynomial time, as shown in [17]. In Section 3 of this paper, we present ten different mathematical programs each of whose optimal values provides an approximation of $\rho(d)$ to within certain constant factors, depending on whether $P(d)$ is feasible or not, and where the constants depend only on the "structure" of the cones C_X and C_Y and not on the dimension or on the data $d = (A, b)$.

The second purpose of this paper is to prove the existence of certain inscribed and intersecting balls involving the feasible region of $P(d)$ (or the feasible region of the alternative system of $P(d)$ if $P(d)$ is infeasible), in the spirit of the ellipsoid algorithm and in order to set the stage for an analysis of the ellipsoid algorithm, hopefully in a subsequent paper. Recall that when $P(d)$ is specialized to the case of non-degenerate linear inequalities and the data $d = (A, b)$ is an array of rational numbers of bitlength L, then the feasible region of $P(d)$ will intersect a ball of radius R centered at the origin, and will contain a ball of radius r where $r = (1/n)2^{-L}$ and $R = n2^L$. Furthermore, the ratio R/r is of critical importance in the analysis of the complexity of using the ellipsoid algorithm to solve the system $P(d)$ in this particular case. (For the general case of $P(d)$, the Turing machine model of computation is not very appropriate for analyzing issues of complexity, and indeed other models of computation have been proposed (see Blum et al. [3], also Smale [15].))

By analogy to the properties of rational non-degenerate linear inequalities mentioned above, Renegar [12] has shown that the feasible region X_d, if nonempty, must intersect a ball of radius R centered at the origin where $R \leq \|d\|/\rho(d)$. Renegar [11] defines the *condition measure* of the data $d = (A, b)$ to be $\mathcal{C}(d)$:

$$\mathcal{C}(d) = \frac{\|d\|}{\rho(d)} \ ,$$

and so $R \leq \mathcal{C}(d)$. Here we see the value $n2^L$ has been replaced by the condition measure $\mathcal{C}(d)$.

For the problem $P(d)$ considered herein in (1), the feasible region is the set X_d. In Sections 4 and 5 of this paper, we utilize the characterization results of Section 3 to prove that the feasible region X_d (or the feasible region of the alternative system when $P(d)$ is infeasible) must contain an inscribed ball of radius r that is no more than a distance R from the origin, and where the ratio R/r satisfies $R/r = c_1 O(\mathcal{C}(d))$. Furthermore, we prove that $r = c_2 \Omega\left(\frac{1}{\mathcal{C}(d)}\right)$ and $R = c_3 O(\mathcal{C}(d))$, where the constants c_1, c_2, c_3 depend on properties of the cones and the norms used (and $c_1 = c_2 = c_3 = 1$ if the norms of the spaces are chosen in a particular way). Note that by analogy to rational non-degenerate linear inequalities, the quantity $n2^L$ is replaced by $\mathcal{C}(d)$. Therefore the condition measure $\mathcal{C}(d)$ is a very relevant tool in proving the existence of an inscribed ball in the feasible region of $P(d)$ that is not too far from the origin and whose radius is not too small. This should prove effective in the analysis of the ellipsoid algorithm as applied to solving $P(d)$.

The paper is organized as follows. Section 2 contains preliminary results, definitions, and analysis. Section 3 contains the ten different mathematical programs each of whose optimal values provides approximations of $\rho(d)$ to within certain constant factors, as discussed earlier. Section 4 contains four lemmas that give partial or full characterizations of certain inscribed and intersecting balls related to the feasible region of $P(d)$ (or its alternative region in the case when $P(d)$ is infeasible). Section 5 presents a synthesis of all of the results in the previous two sections into theorems that give a complete treatment both of the characterization results and of the inscribed and intersecting ball results.

Acknowledgment. We gratefully acknowledge comments from James Renegar on the first draft of this paper, which have contributed to a restatement and simplification of the proofs of Lemma 4.2 and Lemma 4.4. We also gratefully acknowledge the comments of the associate editor, which have contributed to improved exposition in the paper. We are also grateful for the fellowship and research environment at CORE, Catholic University of Louvain, where this work was initiated.

2 Preliminaries and Some More Notation

We will work in the setup of finite dimensional normed linear vector spaces. Both X and Y are normed linear spaces of finite dimension n and m, respectively, endowed with norms $\|x\|$ for $x \in X$ and $\|y\|$ for $y \in Y$. For $\bar{x} \in X$, let $B(\bar{x}, r)$ denote the ball centered at $\bar{x}$ with radius r, i.e.,

$$B(\bar{x}, r) = \{x \in X \mid \|x - \bar{x}\| \leq r\},$$

and define $B(\bar{y}, r)$ analogously for $\bar{y} \in Y$.

For $\bar{d} = (\bar{A}, \bar{b}) \in L(X, Y) \times Y$, we define the ball

$$B(\bar{d}, r) = \{d = (A, b) \in L(X, Y) \times Y | \|d - \bar{d}\| \leq r\}.$$

With this additional notation, it is easy to see that the definition of $\rho(d)$ given in (6) is equivalent to:

$$\rho(d) = \begin{cases} \sup\{\delta \mid B(d, \delta) \subset \mathcal{F}\} & \text{if } d \in \mathcal{F} \\ \\ \sup\left\{\delta \mid B(d, \delta) \subset \mathcal{F}^C\right\} & \text{if } d \in \mathcal{F}^C. \end{cases} \tag{7}$$

We associate with X and Y the dual spaces X^* and Y^* of linear functionals defined on X and Y, respectively, and whose induced (dual) norms are denoted by $\|u\|_*$ for $u \in X^*$ and $\|w\|_*$ for $w \in Y^*$. Let $c \in X^*$. In order to maintain consistency with standard linear algebra notation in mathematical programming, we will consider c to be a column vector in the space X^* and will denote the linear function $c(x)$ by $c^T x$. Similarly, for $A \in L(X, Y)$ and $f \in Y^*$, we denote $A(x)$ by Ax and $f(y)$ by $f^T y$. We denote the adjoint of A by A^T.

If C is a convex cone in X, C^* will denote the dual convex cone defined by

$$C^* = \{z \in X^* \mid z^T x \geq 0 \text{ for any } x \in C\} .$$

Remark 2.1 *If we identify $(X^*)^*$ with X, then $(C^*)^* = C$ whenever C is a closed convex cone.*

Remark 2.2 *If $C_X = X$, then $C_X^* = \{0\}$. If $C_X = \{0\}$, then $C_X^* = X$.*

We denote the set of real numbers by R and the set of nonnegative real numbers by R_+.

Regarding the consistency of $P(d)$, we have the following partial "theorem of the alternative," the proof of which is a straightforward exercise using a separating hyperplane argument.

Proposition 2.1 *If $P(d)$ has no solution, then the system (8) has a solution:*

$$\begin{aligned} A^T y &\in C_X^* \\ y &\in C_Y^* \\ y^T b &\leq 0 \\ y &\neq 0. \end{aligned} \tag{8}$$

If the system (9) has a solution:

$$\begin{aligned} A^T y &\in C_X^* \\ y &\in C_Y^* \\ y^T b &< 0, \end{aligned} \tag{9}$$

then $P(d)$ has no solution. ∎

Using Proposition 2.1, it is elementary to prove the following:

Lemma 2.1 *Consider the set of ill-posed instances $\mathcal{B}$. Then $\mathcal{B}$ can be characterized as:*

$$\begin{aligned} \mathcal{B} = \quad & \{d = (A, b) \in L(X, Y) \times Y \mid \text{ there exists } (x, r) \in X \times R \text{ with} \\ & (x, r) \neq 0 \text{ and } y \in Y^* \text{ with } y \neq 0 \text{ satisfying } br - Ax \in C_Y, x \in C_X, r \geq 0, \\ & y \in C_Y^*, A^T y \in C_X^*, \text{ and } y^T b \leq 0\} \ . \end{aligned}$$

∎

We now recall some facts about norms. Given a finite dimensional linear vector space X endowed with a norm $\|x\|$ for $x \in X$, the dual norm induced on the space X^* is denoted by $\|z\|_*$ for $z \in X^*$, and is defined as:

$$\|z\|_* = \max\{z^T x \mid \|x\| \leq 1\}. \tag{10}$$

If we denote the unit balls in X and X^* by B and B^*, then it is straightforward to verify that

$$B = \{x \in X \mid \|x\| \leq 1\} = \{x \in X \mid z^T x \leq 1 \text{ for all } z \text{ with } \|z\|_* \leq 1\},$$

and

$$B^* = \{z \in X^* \mid \|z\|_* \leq 1\} = \{z \in X^* \mid z^T x \leq 1 \text{ for all } x \text{ with } \|x\| \leq 1\}.$$

Furthermore,

$$z^T x \leq \|z\|_* \|x\| \text{ for any } x \in X \text{ and } z \in X^*, \tag{11}$$

which is the Hölder inequality. Finally, note that if $A = uv^T$, then it is easy to derive that $\|A\| = \|v\|_* \|u\|$ using (10) and (4).

If X and V are finite-dimensional normed linear vector spaces with norm $\|x\|$ for $x \in X$ and norm $\|v\|$ for $v \in V$, then for $(x, v) \in X \times V$, the function $f(x, v)$ defined by

$$f(x, v) = \|(x, v)\| \ := \ \|x\| + \|v\|$$

defines a norm on $X \times V$, whose dual norm is given by

$$\|(w, u)\|_* \ := \ \max \{\|w\|_*, \|u\|_*\} \text{ for } (w, u) \in (X \times V)^* = X^* \times V^*.$$

The following result, which is a special case of the Hahn-Banach Theorem (see, e.g., [19]), will be used extensively in our analysis. We include a short proof based on the subdifferential operator of a convex function.

Proposition 2.2 *For every $x \in X$, there exists $z \in X^*$ with the property that $\|z\|_* = 1$ and $\|x\| = z^T x$.*

Proof: If $x = 0$, then any $z \in X^*$ with $\|z\|_* = 1$ will satisfy the statement of the proposition. Therefore, we suppose that $x \neq 0$. Consider $\|x\|$ as a function of x, i.e., $f(x) = \|x\|$. Then $f(\cdot)$ is a real-valued convex function, and so the subdifferential operator $\partial f(x)$ is non-empty for all $x \in X$, see [2]. Consider any $x \in X$, and let $z \in \partial f(x)$. Then

$$f(w) \geq f(x) + z^T (w - x) \text{ for any } w \in X. \tag{12}$$

Substituting $w = 0$ we obtain $\|x\| = f(x) \leq z^T x$. Substituting $w = 2x$ we obtain $2f(x) = f(2x) \geq f(x) + z^T(2x - x)$, and so $f(x) \geq z^T x$, whereby $f(x) = z^T x$. From (11) it then follows that $\|z\|_* \geq 1$. Now if we let $u \in X$ and set $w = x + u$, we obtain from (12) that $f(u) + f(x) \geq f(u + x) = f(w) \geq f(x) + z^T(w - x) = f(x) + z^T(u + x - x) = f(x) + z^T u$. Therefore, $z^T u \leq f(u) = \|u\|$, and so from (10) we obtain $\|z\|_* \leq 1$. Therefore, $\|z\|_* = 1$. ∎

Because X and Y are normed linear vector spaces of finite dimension, all norms on each space are equivalent, and one can specify a particular norm for X and a particular norm for Y if so desired. If $X = R^n$, the L_p norm is given by

$$\|x\|_p = \left(\sum_{j=1}^{n} |x_j|^p \right)^{1/p},$$

for $p \geq 1$. The norm dual to $\|x\|_p$ is $\|z\|_* = \|z\|_q$ where q satisfies $1/p + 1/q = 1$, with appropriate limits as $p \to 1$ and $p \to +\infty$.

We will say that a cone C is *regular* if C is a closed convex cone, has a nonempty interior and is pointed (i.e., contains no line).

Remark 2.3 *If C is a closed convex cone, then C is regular if and only if C^* is regular.*

Let C be a regular cone in the normed linear vector space X. A critical component of our analysis concerns the extent to which the norm function $\|x\|$ can be approximated by some linear function $u^T x$ over the cone C for some particularly good choice of $u \in X^*$. Let $u \in \text{int}C^*$ be given, and suppose that u has been normalized so that $\|u\|_* = 1$. Let $f(u) = \text{minimum}\{u^T x \mid x \in C, \|x\| = 1\}$. Then it is elementary to see that $0 < f(u) \leq 1$, and also that $f(u)\|x\| \leq u^T x \leq \|x\|$ for any $x \in C$. Therefore the linear function $u^T x$ approximates $\|x\|$ over all $x \in C$ to within the factor $f(u)$. Put another way, the larger $f(u)$ is, the closer $u^T x$ approximates $\|x\|$ over all $x \in C$. Maximizing the value of $f(u)$ over all $u \in X^*$ satisfying $\|u\|_* = 1$, we are led to the following definition:

Definition 2.1 *If C is a regular cone in the normed linear vector space X, the coefficient of linearity for the cone C is given by:*

$$\begin{array}{lll} \beta = & \sup & \inf\ u^T x \\ & u \in X^* & x \in C \\ & \|u\|_* = 1 & \|x\| = 1 \quad . \end{array} \tag{13}$$

Let $\bar{u}$ denote that value of $u \in X^*$ that achieves the supremum in (13). We refer to $\bar{u}$ generically as the "norm approximation vector" for the cone C. Then for all $x \in C$, $\beta\|x\| \leq \bar{u}^T x \leq \|x\|$, and so $\|x\|$ is approximated by the linear function $\bar{u}^T x$ to within the factor β over the cone C. Therefore, β measures the extent to which $\|x\|$ can be approximated by a linear function $\bar{u}^T x$ on the cone C. Also, $\bar{u}^T x$ is the "best" such linear approximation of $\|x\|$ over this cone. It is easy to see that $\beta \leq 1$, since $u^T x \leq \|u\|_* \|x\| = 1$ for u and x as in (13). The larger the value of β, the more closely $\|x\|$ is approximated by a linear function $u^T x$ over $x \in C$. For this reason, we refer to β as the "coefficient of linearity" for the cone C.

We have the following properties of the coefficient of linearity β:

Proposition 2.3 *Suppose that C is a regular cone in the normed linear vector space X, and let β denote the coefficient of linearity for C. Then $0 < \beta \leq 1$. Furthermore, the norm approximation vector $\bar{u}$ exists and is unique, and satisfies the following properties:(i) $\bar{u} \in \text{int}C^*$,*

(ii) $\|\bar{u}\|_* = 1$,

(iii) $\beta = \min\{\bar{u}^T x \mid x \in C,\ \|x\| = 1\}$, *and*

(iv) $\beta\|x\| \leq \bar{u}^T x \leq \|x\|$ *for any* $x \in C$.

The proof of Proposition 2.3 follows easily from the following observation:

Remark 2.4 *Suppose C is a closed convex cone. Then $u \in \text{int}C^*$ if and only if $u^T x > 0$ for all $x \in C/\{0\}$. Also, if $u \in \text{int}C^*$, the set $\{x \in C \mid u^T x = 1\}$ is a closed and bounded convex set.*

We illustrate the construction of the coefficient of linearity on two families of cones, the nonnegative orthant R^n_+ and the positive semi-definite cone $S^{n\times n}_+$. We first consider the nonnegative orthant. Let $X = R^n$ and $C = R^n_+ = \{x \in R^n \mid x \geq 0\}$. Then we can identify X^* with X and in so doing, $C^* = R^n_+$ as well. If $\|x\| = \|x\|_p$, then for $x \in R^n_+$, it is straightforward to show that $\bar{u} = \left(n^{\left(\frac{1}{p}-1\right)}\right) e$, where $e = (1, \ldots, 1)^T$, i.e., the linear function given by $\bar{u}^T x$ is the "best" linear approximation of the function $\|x\|$ on the set R^n_+. Furthermore, straightforward calculation yields that $\beta = n^{\left(\frac{1}{p}-1\right)}$. Then if $p = 1$, $\beta = 1$, but if $p > 1$ then $\beta < 1$.

Now consider the positive semi-definite cone, which has been shown to be of enormous importance in mathematical programming (see Alizadeh [1] and Nesterov and Nemirovskii [8]). Let $X = S^{n\times n}$ denote the set of real $n \times n$ symmetric matrices, and let $C = S^{n\times n}_+ = \{x \in S^{n\times n} \mid x \succeq 0\}$, where "$\succeq$" is the Löwner partial ordering, i.e., $x \succeq w$ if $x - w$ is a positive semi-definite symmetric matrix. Then C is a closed convex cone. We can identify X^* with X, and in so doing it is elementary to derive that $C^* = S^{n\times n}_+$, i.e., $C = S^{n\times n}_+$ is self-dual. For $x \in X$, let $\lambda(x)$ denote the n-vector of ordered eigenvalues of x. That is, $\lambda(x) = (\lambda_1(x), \ldots, \lambda_n(x))^T$ where $\lambda_i(x)$ is the i$^{\text{th}}$ largest eigenvalue of X. For any $p \in [1, \infty)$, let the norm of x be defined by

$$\|x\| = \|x\|_p \ = \ \left(\sum_{j=1}^{n} |\lambda_j(x)|^p\right)^{\frac{1}{p}} \quad ,$$

i.e., $\|x\|_p$ is the L_p-norm of the vector of eigenvalues of x. (see [7], e.g., for a proof that $\|x\|_p$ is a norm.)

When $p = 2$, $\|x\|_2$ corresponds precisely to the Frobenius norm of x. When $p = 1$, $\|x\|_1$ is the sum of the absolute values of the eigenvalues of x. Therefore, when $x \in S^{n\times n}_+$, $\|x\|_1 = tr(x) = \sum_{i=1}^{n} x_{ii}$ where x_{ii} is the i$^{\text{th}}$ diagonal entry of the real matrix x, and so is linear on $C = S^{n\times n}_+$. It is easy to show for the norm $\|x\|_p$ over $S^{n\times n}_+$ that $\bar{u} = \left(n^{\left(\frac{1}{p}-1\right)}\right) I$ has $\|\bar{u}\|_* = \|\bar{u}\|_q = 1$ and that $\beta = n^{\left(\frac{1}{p}-1\right)}$. Thus, for the Frobenius norm we have $\beta = \frac{1}{\sqrt{n}}$ and for the L_1-norm, we have $\beta = 1$.

The coefficient of linearity β for the regular cone C is essentially the same as the scalar α defined in Renegar [12] on page 328. In [12], α is referred to as a measure of "pointedness" of the

cone C. In fact, one can define pointedness in a geometrically intuitive way and it can be shown that β corresponds precisely to the pointedness of the cone C. However, this result is beyond the scope of this paper.

The coefficients of linearity for the cones C_X and/or C_Y play a role in virtually all of the results in this paper. Generally, the results in Section 3 and Section 5 will be stronger to the extent that these coefficients of linearity are large. The following remark shows that by a judicious choice of the norm on the vector space X, one can ensure that the coefficient of linearity for a cone C or the coefficient of linearity for the dual cone C^* are equal to 1 (but not both).

Remark 2.5 *If C is a regular cone, then it is possible to choose the norm on X in such a way that the coefficient of linearity for C is $\beta = 1$. Alternatively, it is possible to choose the norm on X in such a way that the coefficient of linearity for C^* is $\beta^* = 1$.*

To see why this remark is true, recall that for finite dimensional linear vector spaces, that all norms are equivalent. Now suppose that C is a regular cone. Pick any $\bar{u} \in \text{int}C^*$. Let the unit ball for X, denoted as B, be defined as:

$$B = \text{conv}\left(\{x \in C \mid \bar{u}^T x \leq 1\} \cup \{x \in -C \mid -\bar{u}^T x \leq 1\}\right),$$

where "conv(S,T)" denotes the convex hull of the sets S and T. It can then easily be verified that this ball induces a norm $\|\cdot\|$ on X. Furthermore, it is easy to see that for all $x \in C$, that $\|x\| = \bar{u}^T x$, whereby $\beta = 1$. Alternatively, a similar type of construction can be applied to the dual cone C^* to ensure that the coefficient of linearity β^* for C^* satisfies $\beta^* = 1$. However, because the norm on X (or on X^*) induces the dual norm on the dual space, it is not generally possible to construct the dually paired norms $\|\cdot\|$ and $\|\cdot\|_*$ in such a way that both $\beta = 1$ and $\beta^* = 1$.

3 Characterization Results for $\rho(d)$

Given a data instance $d = (A, b) \in L(X, Y) \times Y$, we now present characterizations of the distance to ill-posedness $\rho(d)$ for the feasibility problem $P(d)$ given in (1) .

The characterizations of $\rho(d)$ will depend on whether $d \in \mathcal{F}$ or $d \in \mathcal{F}^C$ (recall (2)), i.e., whether $P(d)$ is consistent or not. We first study the case when $d \in \mathcal{F}$ ($P(d)$ is consistent), followed by the case when $d \in \mathcal{F}^C$ ($P(d)$ is not consistent). Before proceeding, we adopt the following notational conventions.

For the remainder of this study, we make the following modification of our notation.

Definition 3.1 *Whenever the cone C_X is regular, the coefficient of linearity for C_X is denoted by β, and the coefficient of linearity for C_X^* is denoted by β^*. Whenever the cone C_Y is regular, the*

coefficient of linearity for C_Y is denoted by $\bar{\beta}$, and the coefficient of linearity for C_Y^* is denoted by $\bar{\beta}^*$.

Furthermore, when the cone C_X is regular, we denote the norm approximation vector for the cone C_X by $\bar{u}$. Also, when the cone C_Y is regular, we denote the norm approximation vector for the cone C_Y^* by $\bar{z}$.

In particular, then, we have the following partial restatement of Proposition 2.3.

Corollary 3.1 *If C_X is regular, then $\bar{u} \in \text{int}C_X^*$ and $\|\bar{u}\|_* = 1$, and*

$$\beta\|x\| \leq \bar{u}^T x \leq \|x\| \text{ for any } x \in C_X.$$

If C_Y is regular, then $\bar{z} \in \text{int}C_Y$ and $\|\bar{z}\| = 1$, and

$$\bar{\beta}^*\|y\|_* \leq \bar{z}^T y \leq \|y\|_* \text{ for any } y \in C_Y^*.$$

We emphasize that the four coefficient of linearity constants, β, β^*, $\bar{\beta}$, and $\bar{\beta}^*$, depend only on the norms $\|x\|$ and $\|y\|$ and the cones C_X and C_Y, and are independent of the data (A, b) defining the problem $P(d)$.

3.1 Characterization Results when $P(d)$ is consistent

The starting point of our analysis is the following result of Renegar [12], which we motivate as follows. Consider the following homogenization and normalization of $P(d)$:

H :

$$\begin{array}{rcrcl} br & - & Ax & \in & C_Y \\ & & x & \in & C_X \\ r & & & \geq & 0 \\ |r| & + & \|x\| & \leq & 1. \end{array}$$

Recall that $\rho(d)$ measures the extent to which the data $d = (A, b)$ can be altered and yet (1) will still be feasible for the new system. A modification of this view is that $\rho(d)$ measures the extent to which the system (1) can be modified while ensuring its feasibility. Consider the following program:

$P_r(d)$:

$$\begin{array}{rlll} r(d) = & \underset{v \in Y}{\text{minimum}} & \underset{r,x,\theta}{\text{maximum}} & \theta \\ & \|v\| \leq 1 & s.t. & \begin{array}{rcrcrl} br & - & Ax & - & v\theta & \in C_Y \\ & & x & & & \in C_X \\ r & & & & & \geq 0 \\ |r| & + & \|x\| & & & \leq 1. \end{array} \end{array} \tag{14}$$

Then $r(d)$ is the largest scaling factor θ such that for any v with $\|v\| \leq 1$, $v\theta$ can be added to the first inclusion of H without affecting the feasibility of the system. The following is a slightly altered restatement of a result due to Renegar:

Theorem (Theorem 3.5 of [12]) *Suppose that $d \in \mathcal{F}$. Then*

$$r(d) = \rho(d). \tag{15}$$

Now note that the inner maximization program of (14) is a convex program. If we replace this inner maximization program by an appropriately constructed Lagrange dual, we obtain the following modification of program (14):

$$\begin{array}{lll} \underset{\substack{v \in Y \\ \|v\| \leq 1}}{\text{minimum}} & \underset{y,q,g}{\text{minimum}} & \max\left\{ \|A^T y - q\|_*, |b^T y + g| \right\} \\ & s.t. & \begin{array}{rl} y^T v & \geq 1 \\ y & \in C_Y^* \\ q & \in C_X^* \\ g & \geq 0 \quad . \end{array} \end{array} \tag{16}$$

By combining the inner and outer minimizations in (16) and using the duality properties of norms (see (10)), we obtain the following program:

$P_j(d)$:

$$\begin{array}{rll} j(d) = & \underset{y,q,g}{\text{minimum}} & \max\left\{ \|A^T y - q\|_*, \ |b^T y + g| \right\} \\ & s.t. & \begin{array}{rl} y & \in C_Y^* \\ q & \in C_X^* \\ g & \geq 0 \\ \|y\|_* & = 1 \ . \end{array} \end{array} \tag{17}$$

Now note that program $P_j(d)$ is a measure of how close the system (1) is to being infeasible. To see this, note that if $d = (A, b)$ were in $\mathcal{F}^C$, then from Proposition 2.1 it would be true that

$j(d) = 0$. The nonnegative quantity $j(d)$ measures the extent to which the alternative system (9) is not feasible. The smaller the value of $j(d)$ is, the closer the conditions (9) are to being satisfied, and so the smaller the value of $\rho(d)$ should be. These arguments are imprecise, but the next theorem validates the intuition of this line of thinking:

Theorem *Suppose that $d \in \mathcal{F}$. Then*

$$j(d) = \rho(d). \tag{18}$$

One way to prove (18) would be to prove that the duality constructs employed in modifying (14) to (16) to (17) are indeed valid, thereby showing that $j(d) = r(d) = \rho(d)$ and establishing that program $P_j(d)$ is just a partial dualization of $P_r(d)$. Instead we offer the following proof which is more direct and does not rely explicitly on (15).

Proof: Suppose that $j(d) > \rho(d)$. Then there exists $\bar{d} = (\bar{A}, \bar{b})$ such that $\|\bar{A} - A\| < j(d)$ and $\|\bar{b} - b\| < j(d)$ and $\bar{d} \in \mathcal{F}^C$. From Proposition 2.1, there exists $\bar{y} \in C_Y^*$ with $\|\bar{y}\|_* = 1$ that satisfies $\bar{A}^T\bar{y} \in C_Y^*$, $\bar{b}^T\bar{y} \leq 0$. Let $\bar{q} = \bar{A}^T\bar{y}$ and $\bar{g} = -\bar{b}^T\bar{y} \geq 0$. Then

$$\|A^T\bar{y} - \bar{q}\|_* = \|\bar{A}^T\bar{y} - \bar{q} + (A - \bar{A})^T\bar{y}\|_* \leq \|A - \bar{A}\|\|\bar{y}\|_* < j(d)$$

and

$$|b^T\bar{y} + \bar{g}| = |\bar{b}^T\bar{y} + (b - \bar{b})^T\bar{y} - \bar{b}^T\bar{y}| = |(b - \bar{b})^T\bar{y}| \leq \|b - \bar{b}\|\|\bar{y}\|_* < j(d),$$

and so $(\bar{y}, \bar{q}, \bar{g})$ is feasible for $P_j(d)$ with objective value less than $j(d)$, which is a contradiction. Therefore $j(d) \leq \rho(d)$.

Now suppose that $j(d) < \delta < \rho(d)$ for some δ. Then there exists $(\bar{y}, \bar{q}, \bar{g})$ such that $\bar{y} \in C_Y^*$, $\bar{q} \in C_X^*$, $\bar{g} \geq 0$, and $\|A^T\bar{y} - \bar{q}\|_* < \delta$, $|b^T\bar{y} + \bar{g}| < \delta$, and $\|\bar{y}\|_* = 1$. Let $\hat{y}$ satisfy $\|\hat{y}\| = 1$, $\bar{y}^T\hat{y} = \|\bar{y}\|_* = 1$, see Proposition 2.2, and let $\bar{A} = A - \hat{y}\left(\bar{y}^TA - \bar{q}^T\right)$, $\bar{b}_\epsilon = b - \hat{y}\left(b^T\bar{y} + \bar{g} + \epsilon\right)$ for all $\epsilon > 0$. Then $\bar{A}^T\bar{y} = A^T\bar{y} - A^T\bar{y} + \bar{q} = \bar{q} \in C_X^*$, and $\bar{b}_\epsilon^T\bar{y} = -\bar{g} - \epsilon < 0$ for all $\epsilon > 0$. Therefore $\bar{d}_\epsilon = (\bar{A}, \bar{b}_\epsilon) \in \mathcal{F}^C$ for all $\epsilon > 0$. However, $\|\bar{A} - A\| = \|A^T\bar{y} - \bar{q}\|_* < \delta$ and $\|\bar{b}_\epsilon - b\| = |b^T\bar{y} + \bar{g} + \epsilon| \leq |b^T\bar{y} + \bar{g}| + \epsilon < \delta$ for $\epsilon > 0$ and sufficiently small, and so $\rho(d) \leq \|\bar{d}_\epsilon - d\| = \max\{\|\bar{A} - A\|, \|\bar{b}_\epsilon - b\|\} < \delta$, for all $\epsilon > 0$ and sufficiently small. This too is a contradiction, and so $j(d) \geq \rho(d)$, whereby $j(d) = \rho(d)$. ∎

Remark 3.1 *$P_j(d)$ is not in general a convex program due to the non-convex constraint "$\|y\|_* = 1$". However, in the case when $Y = R^m$, (then Y^* can also be identified with R^m), if we choose the norm on Y^* to be the L_∞ norm (so the norm on Y is the L_1 norm), then $P_j(d)$ can be solved by solving $2m$ convex programs. To see this, observe that when $1 = \|y\|_* = \|y\|_\infty$ in $P_j(d)$, then the constraint "$\|y\|_* = 1$" can be replaced by the constraint "$y_i = \pm 1$ for some $i \in \{1, \ldots, m\}$" without changing the optimal objective value of $P_j(d)$. This implies that $j(d) =$*

$\min\{j_{+1}(d),\ldots,j_{+m}(d),j_{-1}(d),\ldots,j_{-m}(d)\}$, *where :*

$$\begin{array}{rlll} j_{\pm i}(d) = & \text{minimum} & \max\left\{\|A^Ty-q\|_*,\ |b^Ty+g|\right\} & \\ & y,q,g & & \\ & s.t. & y\in C_Y^* & \\ & & q\in C_X^* & \\ & & g\geq 0 & \\ & & y_i=\pm 1\ , & \end{array}$$

and $j_{\pm i}(d)$ is the optimal objective value of a convex program, $i=1,\ldots,m$.

We now proceed to present five different mathematical programs each of whose optimal values provides an approximation of the value of the distance to ill-posedness $\rho(d)$, in the case when $P(d)$ is consistent. For each of these five mathematical programs, the nature of the approximation of $\rho(d)$ is specified in a theorem stating the result. For the first program, suppose that C_X is a regular cone, and consider:

$P_\alpha(d)$:

$$\begin{array}{rlll} \alpha(d) = & \text{minimum} & \gamma & \\ & y,\gamma & & \\ & s.t. & A^Ty+\gamma\bar{u} & \in C_X^* \\ & & -b^Ty+\gamma & \geq 0 \\ & & \|y\|_* & =1 \\ & & y & \in C_Y^*. \end{array} \tag{19}$$

Theorem 3.1 *If $d\in\mathcal{F}$ and C_X is regular, then*

$$\beta\cdot\alpha(d)\leq\rho(d)\leq\alpha(d)\ .$$

Proof: Recall from Corollary 3.1 that $\bar{u}\in \text{int}C_X^*$. Therefore $\alpha(d)\geq 0$, since otherwise $d\in\mathcal{F}^{\mathcal{C}}$ via Proposition 2.1, which would violate the supposition of the theorem. Suppose that (y,γ) is feasible for $P_\alpha(d)$. Let $q=A^Ty+\gamma\bar{u}$ and notice that $\|A^Ty-q\|_*=\gamma\|\bar{u}\|_*=\gamma$. Also, if we let $g=-b^Ty+\gamma$, then $g\geq 0$ and $|b^Ty+g|=|\gamma|=\gamma$. Therefore, $\max\left\{\|A^Ty-q\|_*,\ |b^Ty+g|\right\}=\gamma$, and (y,q,g) is feasible for $P_j(d)$ with objective value γ. It then follows that $\alpha(d)\geq j(d)=\rho(d)$ from (18).

On the other hand, suppose that (y,q,g) is feasible for $P_j(d)$, and let

$$\delta=\max\left\{\|A^Ty-q\|_*,\ |b^Ty+g|\right\}.$$

Then it must be true that

$$A^Ty+(\delta/\beta)\bar{u}\in C_X^*. \tag{20}$$

To demonstrate the validity of (20), suppose the contrary. Then there exists $x \in C_X$ with $\|x\| = 1$ for which $x^T \left(A^T y + (\delta/\beta)\bar{u}\right) < 0$. But then

$$\delta = \max\left\{ \|A^T y - q\|_*,\ |b^T y + g| \right\} \geq \|A^T y - q\|_* = \|q - A^T y\|_* \|x\| \geq q^T x - x^T A^T y$$

$$\geq -x^T A^T y > (\delta/\beta)\bar{u}^T x \geq (\delta\beta\|x\|)/\beta = \delta$$

(where the last inequality is from Corollary 3.1), a contradiction. Therefore (20) is true. Then also $\frac{\delta}{\beta} \geq \delta \geq |b^T y + g| \geq b^T y$, and so $(y, \gamma) = \left(y, \frac{\delta}{\beta}\right)$ is feasible for $P_\alpha(d)$. It then follows that $\alpha(d) \leq j(d)/\beta = \rho(d)/\beta$ (from (18)), completing the proof. ∎

Similar to $P_j(d)$, $P_\alpha(d)$ is generally a nonconvex program due to the constraint "$\|y\|_* = 1$." When C_Y is also regular, then from Corollary 3.1 the linear function $\bar{z}^T y$ is a "best" linear approximation of $\|y\|_*$ on C_Y^*, and if we replace "$\|y\|_* = 1$" by "$\bar{z}^T y = 1$" in $P_\alpha(d)$ we obtain the following convex program:

$P_{\tilde{\alpha}}(d)$:

$$\begin{array}{rlrcll} \tilde{\alpha}(d) = & \underset{y,\gamma}{\text{minimum}} & \gamma & & & \\ & s.t. & A^T y & + & \gamma\bar{u} & \in C_X^* \\ & & -b^T y & + & \gamma & \geq 0 \\ & & \bar{z}^T y & & & = 1 \\ & & y & & & \in C_Y^*. \end{array} \tag{21}$$

Replacing the norm constraint by its linear approximation will reduce (by a constant) the extent to which the program computes an approximation of $\rho(d)$, and the analog of Theorem 3.1 becomes:

Theorem 3.2 *If $d \in \mathcal{F}$ and both C_X and C_Y are regular, then*

$$\bar{\beta}^* \beta \cdot \tilde{\alpha}(d) \leq \rho(d) \leq \tilde{\alpha}(d).$$

Proof: Suppose that (y, γ) is a feasible solution of $P_\alpha(d)$. Then $\left(y/\bar{z}^T y, \gamma/\bar{z}^T y\right)$ is a feasible solution of $P_{\tilde{\alpha}}(d)$ with objective function value $\gamma/\bar{z}^T y \leq \gamma/\left(\bar{\beta}^* \|y\|_*\right) = \gamma/\bar{\beta}^*$ (from Corollary 3.1). It then follows that $\tilde{\alpha}(d) \leq \alpha(d)/\bar{\beta}^*$. Applying Theorem 3.1, we obtain $\bar{\beta}^* \beta \tilde{\alpha}(d) \leq \rho(d)$.

Next suppose that (y, γ) is a feasible solution of $P_{\tilde{\alpha}}(d)$. Then $(y/\|y\|_*, \gamma/\|y\|_*)$ is a feasible solution of $P_\alpha(d)$ with objective function value $\gamma/\|y\|_* \leq \gamma/(\bar{z}^T y) = \gamma$ (from Corollary 3.1). It then follows that $\alpha(d) \leq \tilde{\alpha}(d)$. Applying Theorem 3.1, we obtain $\rho(d) \leq \tilde{\alpha}(d)$. ∎

The next mathematical program supposes that the cone C_X is regular. Consider the following program:

$P_w(d)$:

$$
\begin{array}{rlll}
w(d) = & \underset{v \in Y}{\text{minimum}} & \underset{r,x,\theta}{\text{maximum}} & \theta \\
 & \|v\| \leq 1 & s.t. & \begin{array}{rcrcrl} br & - & Ax & - & v\theta & \in C_Y \\ & & x & & & \in C_X \\ r & + & \bar{u}^T x & & & \leq 1 \\ r & & & & & \geq 0. \end{array}
\end{array}
\tag{22}
$$

Notice that $P_w(d)$ is identical to $P_r(d)$ except that the norm constraint "$|r| + \|x\| \leq 1$" in $P_r(d)$ is replaced by the linearized version "$r + \bar{u}^T x \leq 1$". We have:

Theorem 3.3 *If $d \in \mathcal{F}$ and C_X is regular, then*

$$\beta \cdot w(d) \leq \rho(d) \leq w(d) \ .$$

Proof: The proof follows from (15) using the inequalities $\beta\|x\| \leq \bar{u}^T x \leq \|x\|$ of Corollary 3.1, using the same logic as in the proof of Theorem 3.2. ∎

The fourth mathematical program supposes that the cone C_Y is regular. Consider the following convex program:

$P_u(d)$:

$$
\begin{array}{rlll}
u(d) = & \underset{y,q,g}{\text{minimum}} & \max\left\{\|A^T y - q\|_*,\ |b^T y + g|\right\} & \\
 & s.t. & \begin{array}{rcrcl} y & & & & \in C_Y^* \\ & & q & & \in C_X^* \\ & & & g & \geq 0 \\ \bar{z}^T y & & & & = 1 \ . \end{array} &
\end{array}
\tag{23}
$$

Notice that $P_u(d)$ is identical to $P_j(d)$ except that the norm constraint "$\|y\|_* = 1$" in $P_j(d)$ is replaced by the linearized version "$\bar{z}^T y = 1$". We have:

Theorem 3.4 *If $d \in \mathcal{F}$ and C_Y is regular, then*

$$\bar{\beta}^* u(d) \leq \rho(d) \leq u(d) \ .$$

Proof: The proof follows from (18) using the inequalities $\bar{\beta}^*\|y\|_* \leq \bar{z}^T y \leq \|y\|_*$ of Corollary 3.1, using the same logic as in the proof of Theorem 3.2. ∎

Notice that the feasible region of $P_u(d)$ is a convex set, and that the objective function is a gauge function, i.e., a nonnegative convex function that is positively homogeneous of degree 1, see

[14]. A mathematical program that minimizes a gauge function over a convex set is called a gauge program, and corresponding to every gauge program is a dual gauge program that also minimizes a (dual) gauge function over a (dual) convex set, see [6]. For the program $P_u(d)$, its dual gauge program is given by the following convex program:

$P_v(d)$:

$$\begin{array}{rll} v(d) = & \underset{x,\,r}{\text{minimum}} & \|x\| + |r| \\ & br - & Ax - \bar{z} \in C_Y \\ & & x \in C_X \\ & & r \geq 0 \ . \end{array} \tag{24}$$

One can interpret $P_v(d)$ as measuring the extent to which $P(d)$ has a solution x for which $b - Ax$ is in the interior of the cone C_Y. To see this, note from Corollary 3.1 that $\bar{z} \in \text{int}C_Y$, and so $P_v(d)$ will only be feasible if $P(d)$ has a solution x for which $b - Ax$ is in the interior to C_Y. The more interior a solution there is, the smaller (r, x) can be scaled and still satisfy $br - Ax - \bar{z} \in C_Y$. One would then expect $v(d)$ to be inversely proportional to $\rho(d)$ (and to $u(d)$), as the next theorem indicates. Indeed, the theorem states that $u(d) \cdot v(d) = 1$, where we employ the convention that $0 \cdot \infty = 1$ when $\{u(d), v(d)\} = \{0, \infty\}$.

Theorem 3.5 *Suppose that $d \in \mathcal{F}$ and C_Y is regular. Then $u(d) \cdot v(d) = 1$, and*

$$\frac{\bar{\beta}^*}{v(d)} \leq \rho(d) \leq \frac{1}{v(d)} \ .$$

Proof: Suppose that $\rho(d) = 0$. Then $u(d) = 0$ from Theorem 3.4 and from (17) and (18), there exists $\hat{y} \in C_Y^*$ satisfying $A^T\hat{y} \in C_X^*$, $b^T\hat{y} \leq 0$, and $\|\hat{y}\|_* = 1$, which in turn implies that $P_v(d)$ cannot have a feasible solution (for if (x, r) is feasible for $P_v(d)$, then $0 = \hat{y}^T (br - Ax - \bar{z}) < 0$, a contradiction). Thus $v(d) = \infty$, and so $u(d) \cdot v(d) = 1$ by convention, and also $\frac{\bar{\beta}^*}{v(d)} = 0 = \rho(d) = \frac{1}{v(d)}$.

Therefore suppose that $\rho(d) > 0$. Then $u(d) > 0$ from Theorem 3.4 and also it is straightforward to show that both $P_u(d)$ and $P_v(d)$ are feasible and attain their optima. Note that for any (y, q, g) and (x, r) feasible for $P_u(d)$ and $P_v(d)$, respectively, we have

$$1 = \bar{z}^T y \leq y^T br - y^T Ax \leq y^T br + gr - y^T Ax + q^T x \leq$$

$$(\|x\| + |r|) \max\{\|A^T - q\|_*, |b^T y + g|\}$$

whereby $u(d) \cdot v(d) \geq 1$, and so in particular $v(d) > 0$. We now will show that $u(d) \cdot v(d) = 1$, which will complete the proof.

Define the following set:

$$S = \{(\gamma, w, v) \in \Re \times Y \times X \mid \text{there exists } r \geq 0, x \in X, s \in C_Y, \text{ and } p \in C_X \text{ which}$$
$$\text{satisfy } \|x\| + |r| \leq \gamma, br - Ax - \bar{z} - w = s, x - v = p\}.$$

Then S is a nonempty convex set, and basic limit arguments easily establish that S is also a closed set. For any given and fixed $\epsilon \in (0, v(d))$, the point $(v(d) - \epsilon, 0, 0) \notin S$ (for otherwise the optimal value of $P_v(d)$ would be less than or equal to $v(d) - \epsilon$, a contradiction). Since S is a closed nonempty convex set, $(v(d) - \epsilon, 0, 0)$ can be strictly separated from S by a hyperplane, i.e., there exists $(\theta, y, q) \neq 0$ and $\alpha \in \Re$ such that

$$(i) \quad \theta(v(d) - \epsilon) < \alpha$$

$$(ii) \quad \theta\gamma - y^T w - q^T w - q^T v > \alpha \quad \text{for any } (\gamma, w, v) \in S.$$

In particular, (ii) implies that

$$(\|x\| + |r| + \delta)\theta - y^T(br - Ax - \bar{z} - s) - q^T(x - p) > \alpha$$
$$\text{for any } x \in X, r \geq 0, \delta \geq 0, s \in C_Y, \text{ and } p \in C_X \ . \tag{25}$$

This implies that $\theta \geq 0$, $y \in C_Y^*$, $q \in C_X^*$, and $\alpha > 0$.

Suppose first that $\theta > 0$. Then we can rescale (θ, y, q) and α so that $\theta = 1$. Then notice that (25) implies that $1 - b^T y \geq 0$. Also, we claim that (25) implies that $\|A^T y - q\|_* \leq 1$. (To see this, suppose instead that $\|A^T y - q\|_* > 1$. Then there exists $\hat{x} \in X$ such that $\|\hat{x}\| = 1$ and $\hat{x}^T(q - A^T y) > 1$, and then setting $x = \lambda\hat{x}$ for $\lambda > 0$ and sufficiently large, we can drive the left-hand-side of (25) to a negative number, which would yield a contradiction.) Also from (25) and (i), note that $y^T\bar{z} > \alpha > v(d) - \epsilon > 0$. Define $(y', q') = \left(\frac{y}{y^T\bar{z}}, \frac{q}{y^T\bar{z}}\right)$, and $g' = \frac{(b^T y)^-}{y^T\bar{z}}$. Then $y' \in C_Y^*$, $q' \in C_X^*, g' \geq 0, (y')^T\bar{z} = 1$, and $\max\{\|A^T y' - q'\|_*, |b^T y' + g'|\} \leq \frac{1}{y^T\bar{z}} < \frac{1}{v(d)-\epsilon}$, and so (y', q', g') is feasible for $P_u(d)$ with objective value at most $\frac{1}{v(d)-\epsilon}$. Therefore $u(d) \leq \frac{1}{v(d)-\epsilon}$. Since this is true for any $\epsilon \in (0, v(d))$ then $u(d) = \frac{1}{v(d)}$, and then the second assertion of the theorem follows from Theorem 3.4.

It remains to consider the case where $\theta = 0$. Then $\alpha > 0$ and (25) implies that $y^T b \leq 0$, $A^T y = q \in C_X^*$, and $y^T\bar{z} > \alpha > 0$. Then we can rescale y so that $y^T\bar{z} = 1$, and if we define $g = (b^T y)^-$, then (y, q, g) is feasible for $P_u(d)$ with an objective value of zero. Therefore $u(d) = 0$ which implies via Theorem 3.4 that $\rho(d) = 0$, which contradicts the supposition. Therefore $\theta = 0$ is an impossibility, and the theorem is proved. ∎

A simplifying perspective on the results in this subsection is that all five characterization theorems of this subsection are either directly or indirectly derived from Theorem 3.5 of [12]. To see this, first recall that Theorem 3.5 of [12] shows that $\rho(d)$ is obtained as the optimal value of the program $P_r(d)$. Theorem 3.3 was obtained by linearizing the norm constraint "$|r| + \|x\| \leq 1$" in $P_r(d)$. Theorem 3.4 was obtained by linearizing the norm constraint "$\|y\|_* = 1$" of $P_j(d)$, but $P_j(d)$ was itself constructed from $P_r(d)$ via two partial duality derivations. Also, Theorem 3.1 and Theorem 3.2 were obtained by taking particular advantage of properties of the coefficients of linearity β and $\bar{\beta}^*$ as they pertain to modifications of $P_j(d)$ as well. Finally, Theorem 3.5 was

obtained by applying gauge duality to $P_u(d)$, which itself was obtained from $P_j(d)$ by linearization of the norm constraint "$\|y\|_* = 1$" of $P_j(d)$.

We conclude this subsection with the following comment. The five characterization theorems in this subsection provide approximations of $\rho(d)$, but are exact characterizations when $\beta = 1$ and/or $\bar{\beta}^* = 1$. However, from Remark 2.5, we can choose the norms on X and on Y in such a way as to guarantee that $\beta = 1$ and $\bar{\beta}^* = 1$. If the norms are so chosen, then all five theorems provide exact characterizations of $\rho(d)$.

3.2 Characterization Results when $P(d)$ is not consistent

In this subsection, we parallel the results of the previous subsection for the case when $P(d)$ is not consistent. That is, we present five different mathematical programs and we prove that the optimal value of each of these mathematical programs provides an approximation of the value of $\rho(d)$, in the case when $P(d)$ is not consistent. For each of these five mathematical programs, the nature of the approximation of $\rho(d)$ is specified in a theorem stating the result.

As in the previous subsection, the starting point of our analysis is an application of Theorem 3.5 of Renegar [12], which we motivate as follows. Consider the following normalization of the alternative system (8):

HD :

$$\begin{array}{c} A^T y \in C_X^* \\ -b^T y \geq 0 \\ y \in C_Y^* \\ \|y\|_* \leq 1 \ . \end{array} \tag{26}$$

Consider the following program based on HD:

$P_\pi(d)$:

$$\begin{array}{lccll} \pi(d) = & \text{minimum} & \text{maximum} & \theta & \\ & v \in X^* & y, \theta & & \\ & & & & \\ & \|v\|_* \leq 1 & s.t. & A^T y - v\theta \in C_X^* & \\ & & & -b^T y - \theta \geq 0 & \\ & & & y \in C_Y^* & \\ & & & \|y\|_* \leq 1 \ . & \end{array} \tag{27}$$

Then $\pi(d)$ is the largest scaling factor θ such that for any v with $\|v\|_* \leq 1$, $-v\theta$ can be added to the first inclusion of HD and $-\theta$ can be added to the second inclusion of HD without affecting the feasibility of the system HD. The following is also a slightly altered restatement of a result due to Renegar:

Theorem (Theorem 3.5 of [12]) *Suppose that $d \in \mathcal{F}^C$. Then*

$$\pi(d) = \rho(d). \tag{28}$$

Exactly as in the previous subsection, we can use partial duality constructs to create the following program from $P_\pi(d)$:

$P_k(d)$:

$$\begin{array}{rcl} k(d) = & \underset{x,\, r,\, w}{\text{minimum}} & \|br - Ax - w\| \\ & & \\ & s.t. & x \in C_X \\ & & r \geq 0 \\ & & w \in C_Y \\ & & \|x\| + r = 1 \ . \end{array} \tag{29}$$

Note that program $P_k(d)$ is a measure of how close the system $P(d)$ is to being feasible. To see this, note that if $d = (A, b)$ were in $\mathcal{F}$, then it would be true that $k(d) = 0$. The nonnegative quantity $k(d)$ measures the extent to which (1) is not feasible. The smaller the value of $k(d)$ is, the closer the conditions (1) are to being satisfied, and so the smaller the value of $\rho(d)$ should be. These arguments are validated in the following theorem:

Theorem *Suppose that $d \in \mathcal{F}^C$. Then*

$$k(d) = \rho(d). \tag{30}$$

Proof: Suppose that $k(d) > \rho(d)$. Then there exists $\bar{d} = (\bar{A}, \bar{b})$ such that $\|\bar{A} - A\| < k(d)$ and $\|\bar{b} - b\| < k(d)$ and $\bar{d} \in \mathcal{F}$. Therefore there exists $(\bar{x}, \bar{r})$ with $\bar{r} > 0$, $\bar{x} \in C_X$, $\bar{b}\bar{r} - \bar{A}\bar{x} \in C_Y$, and $|\bar{r}| + \|\bar{x}\| = 1$. Let $\bar{w} = \bar{b}\bar{r} - \bar{A}\bar{x}$. Then

$$\begin{array}{rcl} \|b\bar{r} - A\bar{x} - \bar{w}\| & = & \|\bar{b}\bar{r} - \bar{A}\bar{x} - \bar{w} + (b - \bar{b})\,\bar{r} - (A - \bar{A})\,\bar{x}\| \\ & & \\ & \leq & \|b - \bar{b}\|\|\bar{r}| + \|A - \bar{A}\|\|\bar{x}\| < k(d). \end{array}$$

But then $(\bar{x}, \bar{r}, \bar{w})$ is feasible for $P_k(d)$ with objective value less than $k(d)$, which is a contradiction. Therefore $k(d) \leq \rho(d)$.

Now suppose that $k(d) < \delta < \rho(d)$ for some δ. Then there exists $(\bar{x}, \bar{r}, \bar{w})$ such that $\bar{x} \in C_X$, $\bar{r} \geq 0$, $\bar{w} \in C_Y$, and $\|b\bar{r} - A\bar{x} - \bar{w}\| \leq \delta$, and $|\bar{r}| + \|\bar{x}\| = 1$. Let $\tilde{x}$ satisfy $\|\tilde{x}\|_* = 1$ and $\tilde{x}^T\bar{x} = \|\bar{x}\|$, see Proposition 2.2. For $\epsilon > 0$, let

$$\bar{A}_\epsilon = A + (b(\bar{r} + \epsilon) - A\bar{x} - \bar{w})\tilde{x}^T.$$

Then $b(\bar{r} + \epsilon) - \bar{A}_\epsilon\bar{x} = \bar{w} \in C_Y$, and $\bar{r} + \epsilon > 0$ and $\bar{x} \in C_X$. Therefore $\bar{d}_\epsilon := (\bar{A}_\epsilon, b) \in \mathcal{F}$. However,

$$\|\bar{A}_\epsilon - A\| \leq \|b\bar{r} - A\bar{x} - \bar{w}\| + \|b\|\epsilon \leq \delta + \|b\|\epsilon.$$

For $\epsilon < \frac{\rho(d) - \delta}{\|b\|}$, we have $\|\bar{A}_\epsilon - A\| < \rho(d)$, whereby $\bar{d}_\epsilon = (\bar{A}_\epsilon, b) \in \mathcal{F}^C$, a contradiction. Therefore $k(d) \geq \rho(d)$, and so $k(d) = \rho(d)$. ∎

(We point out that because $P_k(d)$ was constructed by using partial duality constructs applied to $P_\pi(d)$ as illustrated above, one can also view (30) as an application of Theorem 3.5 of [12].)

Remark 3.2 *$P_k(d)$ is not in general a convex program due to the non-convex constraint "$r + \|x\| = 1$". However, in the case when $X = R^n$, if we choose the norm on X to be the L_∞ norm, then $P_k(d)$ can be solved by solving $2n$ convex programs, where the construction exactly parallels that given for $P_j(d)$ earlier in this section. One can easily show that $k(d) = min\{k_{+1}, \ldots, k_{+m}, k_{-1}, \ldots, k_{-m}\}$, where :*

$$k_{\pm j}(d) = \begin{array}{ll} \text{minimum} & \|br - Ax - w\| \\ x, r, w & \end{array}$$

$$\begin{array}{ll} s.t. & x \in C_X \\ & r \geq 0 \\ & w \in C_Y \\ & x_j = \pm(1 - r) \ . \end{array}$$

We now proceed to present five different mathematical programs each of whose optimal values provides an approximation of the value of the distance to ill-posedness $\rho(d)$ when $P(d)$ is not consistent. For the first program, suppose that C_Y is a regular cone, and consider:

$P_\sigma(d)$:

$$\sigma(d) = \begin{array}{ll} \text{minimum} & \gamma \\ r, x, \gamma & \end{array}$$

$$\begin{array}{ll} s.t. & br - Ax + \bar{z}\gamma \in C_Y \\ & r + \|x\| = 1 \\ & r \geq 0 \\ & x \in C_X \ . \end{array} \tag{31}$$

Theorem 3.6 *If $d \in \mathcal{F}^C$ and C_Y is regular, then*

$$\bar{\beta}^* \cdot \sigma(d) \leq \rho(d) \leq \sigma(d) \ .$$

Proof: Recall from Corollary 3.1 that $\bar{z} \in \text{int}C_Y$. Therefore $\sigma(d) \geq 0$, since otherwise there would exist (x, r) satisfying $br - Ax \in \text{int}C_Y$, $r > 0$, $x \in C_X$, contradicting the hypothesis that $d \in \mathcal{F}^C$. Suppose that (r, x, γ) is feasible for $P_\sigma(d)$, and let $w = br - Ax + \bar{z}\gamma$. Then (r, x, w) is feasible for $P_k(d)$ with objective value $\|br - Ax - w\| = \|\gamma\bar{z}\| = \gamma\|\bar{z}\| = \gamma$. It then follows that $k(d) \leq \sigma(d)$, and so $\rho(d) \leq \sigma(d)$ from (30).

On the other hand, suppose that (x, r, w) is feasible for $P_k(d)$, and let $\delta = \|br - Ax - w\|$. Then

$$br - Ax + \left(\frac{\delta}{\bar{\beta}^*}\right)\bar{z} \in C_Y. \tag{32}$$

To demonstrate the validity of (32), suppose the contrary. Then there exists $y \in C_Y^*$ with $\|y\|_* = 1$ and $y^T\left(br - Ax + \left(\frac{\delta}{\bar{\beta}^*}\right)\bar{z}\right) < 0$. But then

$$\begin{aligned} \delta = \|br - Ax - w\| &\geq y^T(w + Ax - br) \geq y^T(Ax - br) > \tfrac{\delta}{\bar{\beta}^*} y^T\bar{z} \\ &\geq \tfrac{\delta}{\bar{\beta}^*}(\bar{\beta}^*\|y\|_*) = \delta, \end{aligned}$$

where the last inequality is from Corollary 3.1. As this is a contradiction, (32) is true. Therefore $\gamma = \frac{\delta}{\bar{\beta}^*}$ is a feasible objective value of $P_\sigma(d)$, and so $\sigma(d) \leq \frac{k(d)}{\bar{\beta}^*}$, whereby $\rho(d) = k(d) \geq \bar{\beta}^*\sigma(d)$ from (30). ∎

Similar to $P_k(d)$, $P_\sigma(d)$ is generally a non-convex program due to the constraint "$r + \|x\| = 1$". When C_X is also regular, if we replace "$r + \|x\| = 1$" by "$r + \bar{u}^T x = 1$" in $P_\sigma(d)$ we obtain the following convex program:

$P_{\tilde{\sigma}}(d)$:

$$\begin{aligned} \tilde{\sigma}(d) = \quad & \underset{r,\,x,\,\gamma}{\text{minimum}} \quad \gamma \\ & s.t. \qquad br - Ax + \bar{z}\gamma \in C_Y \\ & \qquad\quad r + \bar{u}^T x = 1 \\ & \qquad\quad r \geq 0 \\ & \qquad\quad x \in C_X\ . \end{aligned} \tag{33}$$

The analog of Theorem 3.6 becomes:

Theorem 3.7 *If $d \in \mathcal{F}^C$ and both C_X and C_Y are regular, then*

$$\bar{\beta}^*\beta \cdot \tilde{\sigma}(d) \leq \rho(d) \leq \tilde{\sigma}(d).$$

Proof: Suppose that (r, x, γ) is a feasible solution of $P_\sigma(d)$. Then $(r/(r + \bar{u}^T x), x/(r + \bar{u}^T x), \gamma/(r + \bar{u}^T x))$ is a feasible solution of $P_{\tilde{\sigma}}(d)$ with objective function value $\gamma/(r + \bar{u}^T x) \leq \gamma/(r + \beta\|x\|) \leq \gamma/\beta$ (from Corollary 3.1). It then follows that $\tilde{\sigma}(d) \leq \sigma(d)/\beta$. Applying Theorem 3.6, we obtain $\bar{\beta}^*\beta\tilde{\sigma}(d) \leq \rho(d)$.

Next suppose that (r, x, γ) is a feasible solution of $P_{\tilde{\sigma}}(d)$. Then $(r/(r + \|x\|), x/(r + \|x\|), \gamma/(r + \|x\|))$ is a feasible solution of $P_\sigma(d)$ with objective function value $\gamma/(r + \|x\|) \leq \gamma/(r + \bar{u}^T x) = \gamma$ (from Corollary 3.1). It then follows that $\sigma(d) \leq \tilde{\sigma}(d)$. Applying Theorem 3.6, we obtain $\rho(d) \leq \tilde{\sigma}(d)$. ∎

For the next mathematical program, suppose that the cone C_Y is regular, and consider:

$P_\delta(d)$:

$$\begin{array}{rlll} \delta(d) = & \underset{v \in X^*}{\text{minimum}} & \underset{y,\theta}{\text{maximum}} & \theta \\ & \|v\|_* \leq 1 & s.t. & A^T y - v\theta \in C_X^* \\ & & & -b^T y - \theta \geq 0 \\ & & & y \in C_Y^* \\ & & & \bar{z}^T y \leq 1 \ . \end{array} \tag{34}$$

Notice that $P_\delta(d)$ is identical to $P_\pi(d)$ except that the norm constraint "$\|y\|_* \leq 1$" in $P_\pi(d)$ is replaced by the linearized version "$\bar{z}^T y \leq 1$". We have:

Theorem 3.8 *If $d \in \mathcal{F}^C$ and C_Y is regular, then*

$$\bar{\beta}^* \cdot \delta(d) \leq \rho(d) \leq \delta(d) \ .$$

Proof: The proof follows from (28) using the inequalities $\bar{\beta}^* \|y\|_* \leq \bar{z}^T y \leq \|y\|_*$ of Corollary 3.1, using the same logic as in the proof of Theorem 3.7. ∎

The fourth mathematical program supposes that the cone C_X is regular. Consider the following convex program:

$P_g(d)$:

$$\begin{array}{rlll} g(d) = & \underset{x,r,w}{\text{minimum}} & \|br - Ax - w\| & \\ & s.t. & x \in C_X & \\ & & r \geq 0 & \\ & & w \in C_Y & \\ & & \bar{u}^T x + r = 1 \ . & \end{array} \tag{35}$$

Notice that $P_g(d)$ is identical to $P_k(d)$ except that the norm constraint "$r + \|x\| = 1$" in $P_k(d)$ is replaced by the linearized version "$r + \bar{u}^T x = 1$". We have:

Theorem 3.9 *If $d \in \mathcal{F}^C$ and C_X is regular, then*

$$\beta g(d) \leq \rho(d) \leq g(d) \ .$$

Proof: The proof follows from (30) using the inequalities $\beta\|x\| \leq \bar{u}^T x \leq \|x\|$ of Corollary 3.1, using the same logic as in the proof of Theorem 3.7. ∎

Notice that $P_g(d)$ is a gauge program; its dual gauge program is given by:

$P_h(d)$:

$$\begin{array}{rll} h(d) = & \underset{y}{\text{minimum}} & \|y\|_* \\ & s.t. & A^T y - \bar{u} \in C_X^* \\ & & -b^T y - 1 \geq 0 \\ & & y \in C_Y^* \ . \end{array} \tag{36}$$

Note that $P_h(d)$ is also a convex program. One can interpret $P_h(d)$ as measuring the extent to which (8) has a solution y for which $A^T y \in \text{int}C_X^*$ and that satisfies $b^T y < 0$. To see this, note from Corollary 3.1 that $\bar{u} \in \text{int}C_X^*$, and so $P_h(d)$ will only be feasible if the first and the third conditions of (8) are satisfied in their interior. The more interior a solution there is, the smaller y can be scaled and still satisfy $A^T y - \bar{u} \in C_Y$ and $-b^T y - 1 \geq 0$. One would then expect $h(d)$ to be inversely proportional to $\rho(d)$ (and to $g(d)$), as Theorem 3.10 indicates. Just as in the case of Theorem 3.5, we employ the convention that $0 \cdot \infty = 1$ when $\{g(d), h(d)\} = \{0, \infty\}$.

Theorem 3.10 *Suppose that $d \in \mathcal{F}^C$ and C_X is regular. Then $g(d) \cdot h(d) = 1$, and*

$$\frac{\beta}{h(d)} \leq \rho(d) \leq \frac{1}{h(d)} \ .$$

Proof: This proof parallels that of Theorem 3.5. Suppose first that $\rho(d) = 0$. Then $g(d) = 0$ from Theorem 3.9. And from (29) and (30), there exists $(\hat{x}, \hat{r}, \hat{w})$ satisfying $b\hat{r} - A\hat{x} - \hat{w} = 0$, $\hat{r} \geq 0$, $\hat{x} \in C_X$, $\hat{w} \in C_Y$, $\|\hat{x}\| + \hat{r} = 1$, which in turn implies that $P_h(d)$ cannot have a feasible solution (for if y is feasible for $P_h(d)$, then $\hat{x}^T(A^T y - \bar{u}) \geq 0$, $\hat{r}(-b^T y - 1) \geq 0$, $\hat{w}^T y \geq 0$, and so $0 = y^T(b\hat{r} - A\hat{x} - \hat{w}) \leq -\bar{u}^T\hat{x} - \hat{r} < 0$, a contradiction). Thus $h(d) = \infty$, and so $g(d) \cdot h(d) = 1$ by convention, and also $\frac{\beta}{h(d)} = 0 = \rho(d) = \frac{1}{h(d)}$.

Therefore suppose that $\rho(d) > 0$. Then $g(d) > 0$ from Theorem 3.9, and also it is straightforward to show that both $P_g(d)$ and $P_h(d)$ are feasible and attain their optima. Note that for any (x, r, w) and y feasible for $P_g(d)$ and $P_h(d)$, respectively, that

$$1 = \bar{u}^t x + r \leq y^T Ax - y^T br \leq y^T Ax - y^T br + w^T y \leq \|y\|_* \|br - Ax - w\| \,,$$

whereby $g(d) \cdot h(d) \geq 1$, and so in particular $h(d) > 0$. We now will show $g(d) \cdot h(d) = 1$, which will complete the proof.

Define the following set:

$$\begin{aligned} S = \{(\gamma, q, s, p) \in \ & \Re \times \Re \times X^* \times Y^* \mid \text{there exists } y \in Y^*, v \in C_X^*, \pi \geq 0, u \in C_Y^* \\ & \text{which satisfy } \|y\|_* \leq \gamma, A^T y - \bar{u} - s = v, -b^T y - 1 - q = \pi, y - p = u\}. \end{aligned}$$

Then S is a nonempty convex set, and basic limit arguments easily establish that S is also a closed set. For any $\epsilon \in (0, h(d))$, the point $(h(d)) - \epsilon, 0, 0, 0) \notin S$ (for otherwise the optimal value of $P_h(d)$ would be no greater than $h(d) - \epsilon$, a contradiction). Since S is a closed nonempty convex set, $(h(d)) - \epsilon, 0, 0, 0)$ can be strictly separated from S by a hyperplane, i.e., there exists $(\theta, r, x, w) \neq 0$

and $\alpha \in \Re$ such that

(i) $\theta(h(d) - \epsilon) < \alpha$,

(ii) $\theta\gamma - rq - x^T s - w^T p > \alpha$ for any $(\gamma, q, s, p) \in S$.

In particular, (ii) implies that

$$(\|y\|_* + \delta)\theta + -r(-b^T y - 1 - \pi) - x^T(A^T y - \bar{u} - v) - w^T(y - u) > \alpha \quad \text{for any } \delta \geq 0, y \in Y^*, v \in C_X^*, \pi \geq 0, \text{ and } u \in C_Y^*. \tag{37}$$

This implies that $\theta \geq 0, r \geq 0, x \in C_X, w \in C_Y$, and $r + \bar{u}^T x > \alpha > 0$.

Suppose first that $\theta > 0$, and so by rescaling (θ, r, x, w) and α we can presume that $\theta = 1$. Then (37) implies that $\|br - Ax - w\| \leq 1$. (To see this, note that if $\|br - Ax - w\| > 1$, then there exists $\bar{y} \in Y^*$ for which $\|\bar{y}\|_* = 1$ and $\bar{y}^T(w + Ax - br) > 1$, and then setting $y = \gamma\bar{y}$ for $\gamma > 0$ and sufficiently large, we can drive the left-hand-side of (37) to a negative number, which is a contradiction.) Also not that (37) implies that $r + \bar{u}^T x > \alpha > h(d) - \epsilon > 0$ from (i). Define $(x', r', w') = (\frac{1}{r + \bar{u}^T x})(x, r, w)$. Then (x, r, w) is feasible for $P_g(d)$, and $g(d) \leq \|br' - Ax' - w'\| = \frac{\|br - Ax - w\|}{r + \bar{u}^T x} \leq \frac{1}{r + \bar{u}^T x} < \frac{1}{h(d) - \epsilon}$. Since this is true for any $\epsilon \in (0, h(d))$, then $g(d) = \frac{1}{h(d)}$, and then the second assertion of the theorem follows from Theorem 3.9.

It only remains to consider the case when $\theta = 0$. Then $\alpha > 0$ and (37) implies that $r \geq 0$, $x \in C_X$, $br - Ax - w = 0$, $w \in C_Y$, and $r + \bar{u}^T x > \alpha > 0$. We can rescale (r, x, w) and α so that $r + \bar{u}^T x = 1$, and then (r, x, w) is feasible for $P_g(d)$ with an objective value of zero. Therefore $g(d) = 0$, which implies via Theorem 3.9 that $\rho(d) = 0$, which contradicts the supposition that $\rho(d) > 0$. Therefore $\theta = 0$ is an impossibility, and the theorem is proved. ∎

The comments at the end of the previous subsection apply to this subsection as well: all five characterization theorems of this subsections are either directly or indirectly derived from Theorem 3.5 of [12]. Also, by appropriate choice of norms on X and/or Y, all five characterization theorems provide exact characterizations of $\rho(d)$.

4 Bounds on Radii of Contained and Intersecting Balls

In this section, we develop four results concerning the radii of certain inscribed balls in the feasible region of the system (1) or, in the case when $P(d)$ is not consistent, of the alternative system (8). These results are stated as Lemmas 4.1, 4.2, 4.3, and 4.4 of this section. While these results are of an intermediate nature, it is nevertheless useful to motivate them, which we do now, by thinking in terms of the ellipsoid algorithm for finding a point in a convex set.

Consider the ellipsoid algorithm for finding a feasible point in a convex set S. Roughly speaking, the main ingredients that are needed to apply the ellipsoid algorithm and to produce a

complexity bound on the number of iterations of the ellipsoid algorithm are the existence of:

(i) a ball $B(\hat{x}, r)$ with the property that $B(\hat{x}, r) \subset S$,

(ii) a ball $B(0, R)$ with the property that $B(\hat{x}, r) \subset B(0, R)$, and

(iii) an upper bound on the ratio R/r .

With these three ingredients, it is then possible to produce a complexity bound on the number of iterations of the ellipsoid algorithm, which will be $O(n^2 \ \ln(R/r))$. In addition, it is also convenient to have the following:

(iv) a lower bound on the radius r of the contained ball $B(\hat{x}, r)$, and

(v) an upper bound on the radius R of the initial ball $B(0, R)$.

In the bit model of complexity as applied to linear inequality systems, one is usually able to set $r = (1/n)2^{-L}$ and $R = n2^L$, where L is the number of bits needed to represent the system. (Of course, these values of r and R break down when the system is degenerate (in our parlance, "ill-posed") in which case the system must be perturbed first.)

By analogy for the problem $P(d)$ considered herein in (1), the convex set in mind is the set X_d, which is the feasible region of the problem $P(d)$, and $n2^L$ is generally replaced by the *condition measure* of $d = (A, b)$, denoted $\mathcal{C}(d)$, which is defined to be

$$\mathcal{C}(d) = \frac{\|d\|}{\rho(d)} \ , \tag{38}$$

see Renegar [11]. The value of $\mathcal{C}(d)$ is a measure of the relative conditioning of the data instance d. (The condition measure $\mathcal{C}(d)$ can be viewed as a scale-invariant reciprocal of the distance to ill-posedness $\rho(d)$, as it is elementary to demonstrate that $\mathcal{C}(\alpha d) = \alpha \mathcal{C}(d)$ for any positive scalar α.)

The results in this section will be used in Section 5 to demonstrate in general that we can find a point $\hat{x} \in X_d$ and radii r and R with the five properties below, that are analogs of the five properties listed above:

(i) $B(\hat{x}, r) \subset X_d$

(ii) $B(\hat{x}, r) \subset B(0, R)$

(iii) $R/r = c_1 O(\mathcal{C}(d))$

(iv) $r = c_2 \Omega \left(\frac{1}{\mathcal{C}(d)} \right)$, and

(v) $R = c_3 O(\mathcal{C}(d))$,

where the constants c_1, c_2, and c_3 depend only on the coefficients of linearity for the cones C_X, C_X^*, C_Y, and C_Y^*, and are independent of the data $d = (A, b)$ of the problem $P(d)$. Here the quantity $n2^L$ is roughly replaced by $\mathcal{C}(d)$.

The above remarks pertain to the the case when $P(d)$ is consistent, i.e., when $P(d)$ has a solution. When $P(d)$ is not consistent, then the convex set in mind is the feasible region for the alternative system (8), denoted by Y_d. The results in this section will also be used in Section 5 to demonstrate in general that we can find a point $\hat{y}$ in Y_d and radii r and R with the three properties below, that are analogs of the first three properties listed above:

$$(i) \quad B(\hat{y}, r) \subset Y_d$$

$$(ii) \quad B(\hat{y}, r) \subset B(0, R)$$

$$(iii) \quad R/r = c_4 O(\mathcal{C}(d)) \quad ,$$

where again the constant c_4 depends on the coefficients of linearity for the cones C_X, C_X^*, C_Y, and C_Y^*, and is independent of the problem data d. Because the system (8) is homogeneous, it makes little sense to bound r from below or R from above, as all constructions can be scaled by any positive quantity. Therefore properties (iv) and (v) above are not relevant.

The results in this section are rather technical, and the reader may first want to read Section 5 before pondering the results in this section in detail.

We first examine the case when $P(d)$ is consistent, in which case the feasible region $X_d = \{x \in X | b - Ax \in C_Y, x \in C_X\}$ is nonempty.

Lemma 4.1 *Suppose that $d \in \mathcal{F}$ and C_Y is regular. If $\rho(d) > 0$, then there exists $\hat{x} \in X_d$ and positive scalars r_1 and R_1 satisfying:*

$$(i) \quad B(\hat{x}, r_1) \subset \{x \in X | b - Ax \in C_Y\},$$

$$(ii) \quad B(\hat{x}, r_1) \subset B(0, R_1),$$

$$(iii) \quad \frac{R_1}{r_1} \leq 1 + \frac{2\|d\|}{\bar{\beta}^* \rho(d)},$$

$$(iv) \quad r_1 \geq \frac{\bar{\beta}^* \rho(d)}{3\|d\|}, \text{ and}$$

$$(v) \quad R_1 \leq 1 + \frac{2\|d\|}{\bar{\beta}^* \rho(d)} \quad .$$

In order to prove Lemma 4.1, we first prove:

Proposition 4.1 $\rho(d) \leq \|d\|$.

Proof: If $d \in \mathcal{F}$ (respectively, $\mathcal{F}^C$), then $\theta d \in \mathcal{F}$ (respectively, $\mathcal{F}^C$) for all $\theta > 0$. Therefore, $\bar{d} = (\bar{A}, \bar{b}) = (0,0) \in \mathcal{B} = cl(\mathcal{F}) \cap cl\left(\mathcal{F}^C\right)$, and so $\rho(d) \leq \|d - \bar{d}\| = \|d - 0\| = \|d\|$. ∎

Proof of Lemma 4.1: For any $w \in C_Y^*$ with $\|w\|_* = 1$, we have

$$\bar{z}^T w + \frac{\bar{\beta}^* b^T w}{\|d\|} \geq \bar{\beta}^* - \frac{\bar{\beta}^* \|d\| \ \|w\|_*}{\|d\|} = 0 \ ,$$

so that $\frac{1}{2}\left(\bar{z} + \frac{\bar{\beta}^*}{\|d\|} b\right) \in C_Y$. Now let $(\tilde{x}, \tilde{r})$ solve $P_v(d)$ (see (24)), and let

$$\hat{x} = \frac{\tilde{x}}{\tilde{r} + \frac{\bar{\beta}^*}{2\|d\|}} = \frac{\tilde{x}}{\delta} \quad ,$$

where $\delta = \tilde{r} + \frac{\bar{\beta}^*}{2\|d\|}$. Let $q = b\tilde{r} - A\tilde{x} - \bar{z}$. Then $q \in C_Y$ and we have $b\tilde{r} - A\tilde{x} + \frac{\bar{\beta}^*}{2\|d\|} b - \frac{1}{2}\bar{z} = \bar{z} + q + \frac{\bar{\beta}^*}{2\|d\|} b - \frac{1}{2}\bar{z} = \frac{1}{2}\left(\bar{z} + \frac{\bar{\beta}^*}{\|d\|} b\right) + q \in C_Y$, so that $\delta b - A\tilde{x} - \frac{1}{2}\bar{z} \in C_Y$, whereby $b - A\hat{x} - \frac{1}{2\delta}\bar{z} \in C_Y$. Thus $\hat{x} \in C_X$ and $b - A\hat{x} \in C_Y$, so $\hat{x} \in X_d$. Let $r_1 = \frac{\bar{\beta}^*}{2\delta\|d\|}$. Then if $\|x - \hat{x}\| \leq r_1$, we have

$$\begin{aligned} b - Ax = b - A\hat{x} + A(\hat{x} - x) \quad &= \tfrac{1}{\delta}\left(\delta b - A\tilde{x} - \tfrac{1}{2}\bar{z}\right) + \tfrac{1}{2\delta}\bar{z} + A(\hat{x} - x) \\ &= y + \tfrac{1}{2\delta}\bar{z} + A(\hat{x} - x) \end{aligned}$$

where $y \in C_Y$. Thus for any $w \in C_Y^*$ with $\|w\|_* = 1$,

$$\begin{aligned} w^T(b - Ax) \quad &\geq \tfrac{1}{2\delta}\bar{z}^T w + w^T A(\hat{x} - x) \quad \geq \tfrac{\bar{\beta}^*}{2\delta} - \|w\|_* \|A\| \ \|\hat{x} - x\| \\ &\geq \tfrac{\bar{\beta}^*}{2\delta} - \|d\| r_1 = 0 \ . \end{aligned}$$

Therefore $b - Ax \in C_Y$, proving (i).

Next, let $R_1 = \|\hat{x}\| + r_1$, and so (ii) is satisfied.

To prove (iii), observe that

$$\begin{aligned} \tfrac{R_1}{r_1} = \tfrac{\|\hat{x}\|}{r_1} + 1 = \tfrac{\|\tilde{x}\|}{\delta r_1} + 1 \quad &\leq \tfrac{v(d)}{\delta r_1} + 1 \qquad\qquad \text{(from (24))} \\ &\leq \tfrac{1}{\delta r_1 \rho(d)} + 1 \qquad \text{(from Theorem 3.5)} \\ &= \tfrac{2\|d\|}{\bar{\beta}^* \rho(d)} + 1 \quad , \end{aligned}$$

proving (iii). To compute the bounds in (iv) and (v), notice first that $\delta = \tilde{r} + \frac{\bar{\beta}^*}{2\|d\|} \geq \frac{\bar{\beta}^*}{2\|d\|}$, so that $r_1 \leq 1$. Therefore, from (iii) we have

$$R_1 \leq \frac{R_1}{r_1} \leq \frac{2\|d\|}{\bar{\beta}^* \rho(d)} + 1,$$

proving (v). We also have:

$$\begin{aligned} \delta = \tilde{r} + \tfrac{\bar{\beta}^*}{2\|d\|} &\leq v(d) + \tfrac{\bar{\beta}^*}{2\|d\|} && \text{(from (24))} \\ &\leq v(d) + \tfrac{1}{2\rho(d)} && \text{(from Proposition 4.1)} \\ &\leq \tfrac{3}{2\rho(d)} \ . && \text{(from Theorem 3.5)} \end{aligned}$$

Therefore, $r_1 = \frac{\bar{\beta}^*}{2\delta\|d\|} \geq \frac{\bar{\beta}^*\rho(d)}{3\|d\|}$, proving (iv). ∎

Remark 4.1 *The proof of Lemma 4.1 can be modified to yield different constants on the bounds on* $\frac{R_1}{r_1}, r_1$, *and* R_1. *By changing the constant "*$\frac{1}{2}$*" in the third line of the proof to "*$\frac{1}{1+\epsilon}$*" for* $\epsilon > 0$, *and making suitable changes in the proof, one can obtain the following bounds:* $\frac{R_1}{r_1} \leq 1 + \frac{(1+\epsilon)\|d\|}{\epsilon\bar{\beta}^*\rho(d)}, r_1 \geq \frac{\epsilon\bar{\beta}^*\rho(d)}{(2+\epsilon)\|d\|}$, *and* $R_1 \leq \epsilon + \frac{(1+\epsilon)\|d\|}{\bar{\beta}^*\rho(d)}$. *One can then choose* ϵ *to optimize one of these bounds, for example.*

We next have:

Lemma 4.2 *Suppose that* $d \in \mathcal{F}$ *and* C_X *is regular. If* $\rho(d) > 0$, *then there exists* $\hat{x} \in X_d$ *and positive scalars* r_2 *and* R_2 *satisfying:*

$$(i) \quad B(\hat{x}, r_2) \subset C_X$$

$$(ii) \quad B(\hat{x}, r_2) \subset B(0, R_2)$$

$$(iii) \quad \frac{R_2}{r_2} \leq 1 + \frac{3\|d\|}{\bar{\beta}^*\rho(d)}$$

$$(iv) \quad r_2 \geq \frac{\bar{\beta}^*\rho(d)}{3\|d\|}$$

$$\text{and } (v) \quad R_2 \leq 2 + \frac{2\|d\|}{\bar{\beta}^*\rho(d)} \ .$$

Proof: Let $\bar{x}$ denote the norm approximation vector for the cone C_X^*. Consider the following optimization problem:

$$
\begin{array}{lll}
Q: & \underset{r,\,x,\,\theta}{\text{maximize}} & \theta \\
& s.t. & \begin{array}{rcrcrl}
br & - & Ax & + & \theta(b - A\bar{x}) & \in C_Y \\
 & & x & & & \in C_X \\
r & & & & & \geq 0 \\
|r| & + & \|x\| & & & \leq 1.
\end{array}
\end{array}
$$

From (14) and (15), the optimal value of Q is at least $\frac{\rho(d)}{\|b-A\bar{x}\|}$, and so there exits $(\tilde{r}, \tilde{x}, \tilde{\theta})$ feasible for Q with $\tilde{\theta} \geq \frac{\rho(d)}{\|b-A\bar{x}\|}$, and so

$$
\tilde{\theta} \;\geq\; \frac{\rho(d)}{\|b-A\bar{x}\|} \;\geq\; \frac{\rho(d)}{\|b\|+\|A\|} \;\geq\; \frac{\rho(d)}{2\|d\|}. \tag{39}
$$

Define

$$
\hat{x} = \frac{\tilde{x} + \bar{x}\tilde{\theta}}{\tilde{\theta} + \tilde{r}} \quad , \quad r_2 = \frac{\tilde{\theta}\beta^*}{\tilde{\theta} + \tilde{r}}, \quad \text{and} \quad R_2 = \|\hat{x}\| + r_2 \; .
$$

Then the feasibility $(\tilde{r}, \tilde{x}, \tilde{\theta})$ in Q ensures that $b - A\hat{x} \in C_Y, \hat{x} \in C_X$, so that $\hat{x}$ is feasible for $P(d)$. For any $v \in X$ satisfying $\|v\| \leq r_2$, and for any $u \in C_X^*$, we have

$$
\begin{array}{rcl}
u^T(\hat{x} + v) & = & \frac{u^T\tilde{x} + u^T\bar{x}\tilde{\theta}}{\tilde{\theta}+\tilde{r}} + u^T v \\
& \geq & \frac{0 + \tilde{\theta}\beta^*\|u\|_*}{\tilde{\theta}+\tilde{r}} - \|v\|\|u\|_* \quad \text{(from Proposition 2.3)} \\
& \geq & \|u\|_* \left(\frac{\tilde{\theta}\beta^*}{\tilde{\theta}+\tilde{r}} - r_2\right) = 0 \quad ,
\end{array}
$$

and so $B(\hat{x}, r_2) \subset C_X$, which shows (i); and (ii) follows from the definition of R_2.

To prove (iii), observe that

$$\begin{aligned}
\tfrac{R_2}{r_2} = \tfrac{\|\hat{x}\|}{r_2} + 1 &\leq \tfrac{\|\tilde{x}\|+\tilde{\theta}}{\tilde{\theta}\beta^*} + 1 \\
&\leq \left(\tfrac{1}{\beta^*}\right)\left(\tfrac{1}{\tilde{\theta}} + 1\right) + 1 \\
&\leq \tfrac{1}{\beta^*}\left(\tfrac{2\|d\|}{\rho(d)} + 1\right) + 1 \qquad \text{(from (39))} \\
&\leq 1 + \tfrac{3\|d\|}{\beta^*\rho(d)} \; . \qquad \text{(from Propsition 4.1)}
\end{aligned}$$

To prove (iv) and (v), note first that $r_2 = \frac{\tilde{\theta}\beta^*}{\tilde{r}+\tilde{\theta}} \leq 1$ since $\beta^* \leq 1$ and $\tilde{r} \geq 0$. Then from (iii) we have that

$$R_2 \leq \frac{R_2}{r_2} \leq 1 + \frac{3\|d\|}{\beta^*\rho(d)},$$

which proves (v). To prove (iv), observe that

$$\begin{aligned}
\tfrac{1}{r_2} = \tfrac{\tilde{\theta}+\tilde{r}}{\tilde{\theta}\beta^*} = \tfrac{1}{\beta^*} + \tfrac{\tilde{r}}{\tilde{\theta}\beta^*} \leq \tfrac{1}{\beta^*} + \tfrac{1}{\tilde{\theta}\beta^*} &\leq \tfrac{1}{\beta^*} + \tfrac{2\|d\|}{\rho(d)\beta^*} \qquad \text{(from (39))} \\
&\leq \tfrac{3\|d\|}{\beta^*\rho(d)}, \qquad \text{(from Proposition 4.1)}
\end{aligned}$$

and so $r_2 \geq \frac{\beta^*\rho(d)}{3\|d\|}$, completing the proof. ∎

Remark 4.2 *Similarly to Remark 4.1, the proof of Lemma 4.2 can be modified to yield different constants on the bounds on* $\frac{R_2}{r_2}, r_2$, *and* R_2. *By changing the first inclusion in the feasibility conditions of program Q above to "*$br - Ax + \theta(\epsilon b - A\bar{x}) \in C_Y$*" for* $\epsilon > 0$ *and making suitable changes in the proof, one can obtain the following bounds:* $\frac{R_2}{r_2} \leq 1 + \frac{1}{\beta^*} + \frac{(1+\epsilon)\|d\|}{\beta^*\rho(d)}, r_2 \geq \frac{\beta^*\rho(d)}{(1+2\epsilon)\|d\|}$, *and* $R_2 \leq \frac{2}{\epsilon} + \frac{(1+\epsilon)\|d\|}{\rho(d)}$. *One can then choose* ϵ *to optimize one of these bounds as well.*

We now turn to the case when $P(d)$ is inconsistent, i.e., $P(d)$ has no solution. In this case, from Proposition 2.1, the system (8) has a solution, and let us then examine the set of all solutions to (8), which we denote by Y_d to emphasize the dependence on the data $d = (A, b)$:

$$Y_d = \{y \in Y^* | A^T y \in C_X^*, \; y \in C_Y^*, \; y^T b \leq 0\} \; . \tag{40}$$

Lemma 4.3 *Suppose that* $d \in \mathcal{F}^C$ *and* C_X *is regular. If* $\rho(d) > 0$, *then there exists* $\hat{y} \in Y_d$ *and positive scalars* r_3 *and* R_3 *satisfying* $\frac{R_3}{r_3} \leq \frac{\|d\|}{\beta\rho(d)}$, *and that satisfy:*

$$B(\hat{y}, r_3) \subset \{y \in Y^* | A^T y \in C_X^*, \; b^T y \leq 0\} \tag{41}$$

and

$$\|\hat{y}\|_* \leq R_3 \ . \tag{42}$$

Proof of Lemma 4.3: Let $\tilde{y}$ solve $P_h(d)$. Then $A^T\tilde{y} - \bar{u} \in C_X^*$, $-b^T\tilde{y} \geq 1$, and $\tilde{y} \in C_Y^*$. Then since $\tilde{y} \neq 0$ (otherwise $-\bar{u} \in C_X^*$ and so C_X is not regular, via Proposition 2.3), let $\hat{y} = \frac{\tilde{y}}{\|\tilde{y}\|_*}$. Let $r_3 = \frac{\beta g(d)}{\|d\|}$, and let $R_3 = 1$.

To prove (41) it suffices to show that if $\|y - \hat{y}\|_* \leq r_3$, then $A^T y \in C_X^*$ and $y^T b \leq 0$. We have that $q = A^T\hat{y} - \frac{\bar{u}}{\|\tilde{y}\|_*} \in C_X^*$. For any $x \in C_X$ with $\|x\| = 1$,

$$\begin{aligned}
x^T A^T y &= x^T A^T (y - \hat{y}) + x^T A^T \hat{y} - \tfrac{\bar{u}^T x}{\|\tilde{y}\|_*} + \tfrac{\bar{u}^T x}{\|\tilde{y}\|_*} \\
&\geq -\|x\| \ \|A\| \ \|y - \hat{y}\|_* + x^T q + \tfrac{\bar{u}^T x}{\|\tilde{y}\|_*} \\
&\geq -\|A\| r_3 + \tfrac{\bar{u}^T x}{\|\tilde{y}\|_*} && \text{(since } x^T q \geq 0\text{)} \\
&\geq -\beta g(d) + \tfrac{\bar{u}^T x}{\|\tilde{y}\|_*} && \text{(since } \|d\| \geq \|A\|\text{)} \\
&\geq -\beta g(d) + \tfrac{\beta}{\|\tilde{y}\|_*} && \text{(from Corollary 3.1)} \\
&\geq -\beta \left(g(d) - \tfrac{1}{h(d)} \right) \\
&= 0 \ . && \text{(from Theorem 3.10)}
\end{aligned}$$

Therefore, $A^T y \in C_X^*$. Similarly,

$$\begin{aligned}
-b^T y &= -b^T (y - \hat{y}) - b^T \hat{y} \\
&\geq -\|d\| \ \|y - \hat{y}\|_* - \tfrac{b^T \tilde{y}}{\|\tilde{y}\|_*} \\
&\geq -\beta g(d) + \tfrac{1}{\|\tilde{y}\|_*} \\
&\geq -\beta g(d) + \tfrac{1}{h(d)} = g(d)(1 - \beta) && \text{(from Theorem 3.10)} \\
&\geq 0. && \text{(from Proposition 2.3)}
\end{aligned}$$

Therefore, $A^T y \in C_X^*$ and $b^T y \leq 0$, which proves (41).

To prove (42), note that $\|\hat{y}\|_* = 1 = R_3$, which demonstrates (42). Finally note that $\frac{R_3}{r_3} = \frac{1}{r_3} = \frac{\|d\|}{\beta g(d)} \leq \frac{\|d\|}{\beta \rho(d)}$ from Theorem 3.9. ∎

We next prove:

Lemma 4.4 *Suppose that $d \in \mathcal{F}^C$ and C_Y is regular. If $\rho(d) > 0$, then there exists $\hat{y} \in Y_d$ and positive scalars r_4 and R_4 satisfying $\frac{R_4}{r_4} \leq \frac{2\|d\|}{\bar{\beta}\rho(d)}$, and that satisfy:*

$$B(\hat{y}, r_4) \subset C_Y^* \tag{43}$$

and

$$\|\hat{y}\|_* \leq R_4 \ . \tag{44}$$

Proof: Let $\bar{y}$ denote the norm approximation vector for the cone C_Y. Then $\|\bar{y}\|_* = 1$ and $y^T\bar{y} \geq \bar{\beta}\|y\|$ for any $y \in C_Y$, see Proposition 2.3. Suppose that d satisfies $\|d\|_* \leq \bar{\beta}$. Then for any $y \in C_Y$, we have $(\bar{y}+d)^T y \geq \bar{\beta}\|y\| - \|d\|_*\|y\| \geq \bar{\beta}\|y\| - \bar{\beta}\|y\| = 0$, and so $\bar{y} + d \in C_Y^*$. Therefore $B(\bar{y}, \bar{\beta}) \subset C_Y^*$, and recall that $\|\bar{y}\|_* = 1$.

Consider now the following system in the variable y:

$$\begin{array}{c} A^T y + \left(\frac{1}{\mathcal{C}(d)}\right) A^T\bar{y} \in C_X^* \\ -b^T y + \frac{1}{\mathcal{C}(d)}\left(-b^T\bar{y}\right) \geq 0 \\ y \in C_Y^* \\ \|y\|_* \leq 1. \end{array} \tag{45}$$

Then $\|\frac{1}{\mathcal{C}(d)}(A^T\bar{y})\|_* \leq \frac{\|A\|\|\bar{y}\|_*\rho(d)}{\|d\|} \leq \rho(d)$, and $|\frac{1}{\mathcal{C}(d)}(-b^T\bar{y})| \leq \frac{\|b\|\|\bar{y}\|_*\rho(d)}{\|d\|} \leq \rho(d)$. Then from (27) and (28) it follows that (45) has a solution $\tilde{y}$.

Define

$$\hat{y} = \left(\frac{\mathcal{C}(d)}{1+\mathcal{C}(d)}\right)\tilde{y} + \left(\frac{1}{1+\mathcal{C}(d)}\right)\bar{y}.$$

Then $\|\hat{y}\|_* \leq 1$, since $\hat{y}$ is a convex combination of $\tilde{y}$ and $\bar{y}$. Also $\hat{y} \in C_Y^*$ and $A^T\hat{y} \in C_X^*$ and $-b^T\hat{y} \geq 0$ from convexity and from (45). Therefore $\hat{y} \in Y_d$. Let $r_4 = \frac{\bar{\beta}}{\mathcal{C}(d)+1}$ and $R_4 = 1$. Then $B(\bar{y}, \bar{\beta}) \subset C_Y^*$ and $\tilde{y} \subset C_Y^*$ imply that $B(\hat{y}, r_4) = B\left(\hat{y}, \frac{\bar{\beta}}{\mathcal{C}(d)+1}\right) \subset C_Y^*$. Also $\|\hat{y}\|_* \leq 1 = R_4$. Finally, note that $\frac{R_4}{r_4} = \frac{1}{r_4} = \frac{\mathcal{C}(d)+1}{\bar{\beta}} \leq \frac{2\mathcal{C}(d)}{\bar{\beta}} = \frac{2\|d\|}{\bar{\beta}\rho(d)}$. ∎

5 Synthesis of Results

In this section, we synthesize the results of the previous two sections into theorems that characterize aspects of the distance to ill-posedness for the three particular cases of problem $P(d)$ of (1), namely

(i) Case 1 : C_X and C_Y are both regular,

(ii) Case 2: C_X is regular and $C_Y = \{0\}$,

(iii) Case 3: $C_X = X$ and C_Y is regular,

and for the status of solvability of $P(d)$ of (1), namely

(a) $P(d)$ is consistent, i.e., (1) has a solution, and

(b) $P(d)$ is inconsistent, i.e., (8) has a solution.

Each of the six theorems of this section synthesizes our results of the previous two sections, as applied to the one of the three cases above and one of the two status' of the solvability of $P(d)$. Each theorem summarizes the applicable approximation characterizations of $\rho(d)$ of Section 3, and also synthesizes the appropriate bounds on radii of contained and intersecting balls developed in Section 4. For a motivation of the importance of these bounds on radii of contained and intersecting balls contained herein, the reader is referred to the opening discussion at the beginning of Section 4.

Each case is treated as a separate subsection, and all proofs are deferred to the end of the section.

5.1 Case 1: C_X and C_Y are both regular.

Theorem 5.1 *Suppose that C_X and C_Y are both regular. If $P(d)$ is consistent, i.e., $d \in \mathcal{F}$, then*

(i) $\beta \cdot \alpha(d) \leq \rho(d) \leq \alpha(d)$

(ii) $\bar{\beta}^* \beta \cdot \tilde{\alpha}(d) \leq \rho(d) \leq \tilde{\alpha}(d)$

(iii) $\beta \cdot w(d) \leq \rho(d) \leq w(d)$

(iv) $\bar{\beta}^* \cdot u(d) \leq \rho(d) \leq u(d)$

(v) $\frac{\bar{\beta}^*}{v(d)} \leq \rho(d) \leq \frac{1}{v(d)}$.

(vi) If $\rho(d) > 0$, then there exists $\hat{x} \in X_d$ and positive scalars r and R satisfying:

$$(a) \quad B(\hat{x}, r) \subset X_d$$

$$(b) \quad B(\hat{x}, r) \subset B(0, R)$$

$$(c) \quad \frac{R}{r} \leq 2 + \frac{5\|d\|}{\min\{\beta^*, \bar{\beta}^*\}\rho(d)}$$

$$(d) \quad r \geq \frac{\min\{\beta^*, \bar{\beta}^*\}\rho(d)}{6\|d\|}$$

$$(e) \quad R \leq 2 + \frac{2\|d\|}{\min\{\beta^*, \bar{\beta}^*\}\rho(d)} \quad .$$

Theorem 5.2 *Suppose that C_X and C_Y are both regular. If $P(d)$ is not consistent, i.e., $d \in \mathcal{F}^C$, then*

(i) $\bar{\beta}^* \cdot \sigma(d) \leq \rho(d) \leq \sigma(d)$

(ii) $\bar{\beta}^* \beta \cdot \tilde{\sigma}(d) \leq \rho(d) \leq \tilde{\sigma}(d)$

(iii) $\bar{\beta}^* \cdot \delta(d) \leq \rho(d) \leq \delta(d)$

(iv) $\beta \cdot g(d) \leq \rho(d) \leq g(d)$

(v) $\frac{\beta}{h(d)} \leq \rho(d) \leq \frac{1}{h(d)}$.

(vi) If $\rho(d) > 0$, then there exists $\hat{y} \in Y_d$ and positive scalars r and R satisfying:

$$(a) \quad B(\hat{y}, r) \subset Y_d$$

$$(b) \quad B(\hat{y}, r) \subset B(0, R)$$

$$(c) \quad \frac{R}{r} \leq 1 + \frac{3\|d\|}{\min\{\beta, \bar{\beta}\} \cdot \rho(d)} \quad .$$

5.2 Case 2: C_X is regular and $C_Y = \{0\}$.

Theorem 5.3 *Suppose that C_X is regular and $C_Y = \{0\}$. If $P(d)$ is consistent, i.e., $d \in \mathcal{F}$, then*

(i) $\beta \cdot \alpha(d) \leq \rho(d) \leq \alpha(d)$

(ii) $\beta \cdot w(d) \leq \rho(d) \leq w(d)$

(iii) If $\rho(d) > 0$, then there exists $\hat{x} \in X_d$ and positive scalars r and R satisfying:

(a) $\{x \in X |\ |\|x - \hat{x}\| \leq r,\ Ax = b\} \subset X_d$

(b) $B(\hat{x}, r) \subset B(0, R)$

(c) $\frac{R}{r} \leq 1 + \frac{3\|d\|}{\beta^* \rho(d)}$

(d) $r \geq \frac{\beta^* \rho(d)}{3\|d\|}$

(e) $R \leq 2 + \frac{2\|d\|}{\beta^* \rho(d)}$.

Theorem 5.4 *Suppose that C_X is regular and $C_Y = \{0\}$. If $P(d)$ is not consistent, i.e., $d \in \mathcal{F}^C$, then*

(i) $\beta \cdot g(d) \leq \rho(d) \leq g(d)$

(ii) $\frac{\beta}{h(d)} \leq \rho(d) \leq \frac{1}{h(d)}$

(iii) If $\rho(d) > 0$, then there exists $\hat{y} \in Y_d$ and positive scalars r and R satisfying:

(a) $B(\hat{y}, r) \subset Y_d$

(b) $B(\hat{y}, r) \subset B(0, R)$

(c) $\frac{R}{r} \leq 1 + \frac{\|d\|}{\beta \cdot \rho(d)}$.

5.3 Case 3: $C_X = X$ and C_Y is regular.

Theorem 5.5 *Suppose that $C_X = X$ and C_Y is regular. If $P(d)$ is consistent, i.e., $d \in \mathcal{F}$, then*

(i) $\bar{\beta}^* \cdot u(d) \leq \rho(d) \leq u(d)$

(ii) $\frac{\bar{\beta}^*}{v(d)} \leq \rho(d) \leq \frac{1}{v(d)}$

(iii) If $\rho(d) > 0$, then there exists $\hat{x} \in X_d$ and positive scalars r and R satisfying:

(a) $B(\hat{x}, r) \subset X_d$

(b) $B(\hat{x}, r) \subset B(0, R)$

(c) $\frac{R}{r} \leq 1 + \frac{2||d||}{\bar{\beta}^* \rho(d)}$

(d) $r \geq \frac{\bar{\beta}^* \rho(d)}{3||d||}$

(e) $R \leq 1 + \frac{2||d||}{\bar{\beta}^* \rho(d)}$.

Theorem 5.6 *Suppose that $C_X = X$ and C_Y is regular. If $P(d)$ is not consistent, i.e., $d \in \mathcal{F}^C$, then*

(i) $\bar{\beta}^* \cdot \sigma(d) \leq \rho(d) \leq \sigma(d)$

(ii) $\bar{\beta}^* \cdot \delta(d) \leq \rho(d) \leq \delta(d)$

(iii) If $\rho(d) > 0$, then there exists $\hat{y} \in Y_d$ and positive scalars r and R satisfying:

(a) $\{y \in Y^* | \|y - \hat{y}\|_* \leq r,\ A^T y = 0\} \subset Y_d$

(b) $B(\hat{y}, r) \subset B(0, R)$

(c) $\frac{R}{r} \leq 1 + \frac{3||d||}{\min\{\bar{\beta}, \bar{\beta}^*\} \cdot \rho(d)}$.

Proof of Theorem 5.1: Parts (i), (ii), (iii), (iv) and (v) follow directly from Theorems 3.1, 3.2, 3.3, 3.4 and 3.5, respectively. It remains to prove part (vi).

Let $S = \{x \in X | b - Ax \in C_Y\}$ and $T = C_X$. Then $S \cap T = X_d$. From Lemma 4.1, there exists $\hat{x}_1 \in X_d$ and $r_1,\ R_1$ satisfying conditions $(i) - (v)$ of Lemma 4.1. From Lemma 4.2, there exists $\hat{x}_2 \in X_d$ and $r_2,\ R_2$ satisfying conditions $(i) - (v)$ of Lemma 4.2. Then the conditions of Proposition A.2 of the Appendix are satisfied, and so there exists $\hat{x}$ and $r, \hat{R}$ satisfying the five conditions of Proposition A.2. Therefore, $(i)\ B(\hat{x}, r) \subset S \cap T = X_d$, which is (a). Also from (ii), $B(\hat{x}, r) \subset B(0, \hat{R})$, which is (b). From (iii), we have

$$\frac{\hat{R}}{r} \leq \frac{R_1}{r_1} + \frac{R_2}{r_2} \leq 2 + \frac{5\|d\|}{\min\{\beta^*, \bar{\beta}^*\}\rho(d)}$$

(invoking Lemma 4.1 (iii) and Lemma 4.2 (iii)), which is (c). Similarly applying Lemma 4.1 and 4.2 and Proposition A.2 in parts (iv) and (v) yields

$$r \geq \frac{1}{2}\min\{r_1, r_2\} \geq \frac{\min\{\beta^*, \bar{\beta}^*\}\rho(d)}{6\|d\|}$$

and

$$\hat{R} \leq \max\{R_1, R_2\} \leq 2 + \frac{2\|d\|}{\min\{\beta^*, \bar{\beta}^*\}\rho(d)} \ . \blacksquare$$

Proof of Theorem 5.2: Parts (i), (ii), (iii), (iv) and (v) follow directly from Theorems 3.6, 3.7, 3.8, 3.9 and 3.10, respectively. It remains to prove part (vi).

Let $S = \{y \in Y^* | A^T y \in C_X^*,\ b^T y \leq 0\}$ and $T = C_Y^*$. Then $S \cap T = Y_d$. From Lemma 4.3, there exists $\hat{y}_3,\ r_3,\ R_3$ satisfying $\hat{y}_3 \in S \cap T,\ B(\hat{y}_3,\ r_3) \subset S$ and $\|\hat{y}_3\|_* \leq R_3$, and $\frac{R_3}{r_3} \leq \frac{\|d\|}{\beta\rho(d)}$. From Lemma 4.4 there exists $\hat{y}_4,\ r_4,\ R_4$ satisfying $\hat{y}_4 \in S \cap T,\ B(\hat{y}_4,\ r_4) \subset T$, and $\|\hat{y}_4\|_* \leq R_4$, and $\frac{R_4}{r_4} \leq \frac{2\|d\|}{\bar{\beta}\rho(d)}$. Then from Proposition A.1 of the Appendix, there exists $\hat{y}$ and $r, \hat{R}$ satisfying $B(\hat{y},\ r) \subset S \cap T = Y_d$, and $\|\hat{y}\| \leq \hat{R}$, and

$$\frac{\hat{R}}{r} \ \leq \frac{R_3}{r_3} + \frac{R_4}{r_4} \ \leq \frac{3\|d\|}{\min\{\beta, \bar{\beta}\}\rho(d)}.$$

Now let $R = \hat{R} + r$. Then for any $y \in B(\hat{y},\ r),\ \|y\|_* \leq \|\hat{y}\|_* + r \leq \hat{R} + r = R$, and

$$\frac{R}{r} = \frac{\hat{R}}{r} + 1 \ \leq 1 + \frac{3\|d\|}{\min\{\beta, \bar{\beta}\}\rho(d)} \ . \blacksquare$$

Proof of Theorem 5.3: Parts (i) and (ii) follow directly from Theorems 3.1 and 3.3, respectively. To prove (iii) we apply Lemma 4.2; there exists $\hat{x} \in X_d$ and $r_2,\ R_2$ satisfying the five conditions of Lemma 4.2. Let $r = r_2$ and $R = R_2$. Then $(b),\ (c),\ (d)$, and (e) follow directly. To prove (a), observe that from Lemma 4.2 (i) that

$$\{x \in X |\ \|x - \hat{x}\| \leq r\} \subset C_X \quad ,$$

and intersecting both sides with the affine set $\{x \in X | Ax = b\}$ gives

$$\{x \in X |\ \|x - \hat{x}\| \leq r,\ Ax = b\} \subset C_X \cap \{x \in X | Ax = b\} = X_d \ . \blacksquare$$

Proof of Theorem 5.4: Parts (i) and (ii) follow directly from Theorems 3.9 and 3.10, respectively. To prove (iii) we apply Lemma 4.3; there exists $\hat{y} \in Y_d$ and $r_3,\ R_3$ satisfying $\frac{R_3}{r_3} \leq \frac{\|d\|}{\beta\rho(d)}$ and (41) and (42). Let $r = r_3$ and $R = R_3 + r_3$. Then from (41) we obtain

$$\{y \in Y^* |\ \|y - \hat{y}\|_* \leq r\} \subset \{y \in Y^* | A^T y \in C_X^*,\ b^T y \leq 0\} = Y_d \quad .$$

Also, for any y satisfying $\|y - \hat{y}\| \leq r$, $\|y\| \leq \|\hat{y}\| + r \leq R_3 + r_3 = R$. Finally, note that

$$\frac{R}{r} = \frac{R_3}{r_3} + 1 \quad \leq \frac{\|d\|}{\beta\rho(d)} + 1 \ . \blacksquare$$

Proof of Theorem 5.5: Parts (i) and (ii) follow directly from Theorems 3.4 and 3.5, respectively. To prove (iii) we apply Lemma 4.1; there exists $\hat{x} \in X_d$ and $r_1,\ R_1$ satisfying the five conditions of Lemma 4.1. Let $r = r_1$ and $R = R_1$. Then (b), (c), (d), and (e) following directly. To prove (a), observe from Lemma 4.1(i) that

$$\{x \in X |\ \|x - \hat{x}\| \leq r\} \subset \{x \in X | b - Ax \in C_Y\} = X_d \ . \blacksquare$$

Proof of Theorem 5.6: Parts (i) and (ii) follow from Theorems 3.6 and 3.8, respectively. It remains to prove (iii).

Let $S = \{y \in Y^* | b^T y \leq 0\}$. If we let $v = 0$ in (34), we see that there exists $\hat{y}_1 \in C_Y^*$ satisfying $A^T \hat{y}_1 = 0, \bar{z}^T \hat{y}_1 \leq 1$ and $-b^T \hat{y}_1 \geq \delta(d) \geq \rho(d)$, from Theorem 3.8. Therefore, if we set $r_1 = \frac{\rho(d)}{\|d\|}$ and $R_1 = \frac{1}{\bar{\beta}^*}$, we have $\|\hat{y}_1\|_* \leq \frac{\bar{z}^T \hat{y}_1}{\bar{\beta}^*} \leq \frac{1}{\bar{\beta}^*} = R_1$ (from Corollary 3.1), and for any y satisfying $\|y - \hat{y}_1\|_* \leq r_1$, we have $b^T y = b^T(y - \hat{y}_1) + b^T \hat{y}_1 \leq \|b\|\ \|\hat{y}_1 - y\|_* - \rho(d) \leq \|d\| r_1 - \rho(d) = 0$, and so $B(\hat{y}_1,\ r_1) \subset S$. If we let $T = C_Y^*$, we have $S \cap T \cap \{y \in Y^* | A^T y = 0\} = Y_d$, and $\hat{y}_1 \in Y_d$. From Lemma 4.4, there exists $\hat{y}_2$ and $r_4,\ R_4$ satisfying $\frac{R_4}{r_4} \leq \frac{2\|d\|}{\bar{\beta}\rho(d)}$, and (43) and (44). Then from Proposition A.1 of the Appendix, there exists $\hat{y}$ and $r,\ \hat{R}$ satisfying $B(\hat{y}, r) \subset S \cap T$ and $\|\hat{y}\| \leq \hat{R}$, and

$$\frac{\hat{R}}{r} \leq \frac{R_1}{r_1} + \frac{R_4}{r_4} \leq \frac{3\|d\|}{\min\{\bar{\beta}, \bar{\beta}^*\}\rho(d)} \quad .$$

Note also that

$$\{y \in Y^* |\ \|y - \hat{y}\|_* \leq r,\ A^T y = 0\} \subset S \cap T \cap \{y \in Y^* | A^T y = 0\} = Y_d \ .$$

Let $R = \hat{R} + r$. Then for any $y \in B(\hat{y}, r),\ \|y\|_* \leq \|\hat{y}\|_* + r = \hat{R} + r = R$, and

$$\frac{R}{r} = \frac{\hat{R}}{r} + 1 \quad \leq 1 + \frac{3\|d\|}{\min\{\bar{\beta}, \bar{\beta}^*\}\rho(d)} \ . \blacksquare$$

APPENDIX

This appendix contains two simple constructions with balls on the intersection of two convex sets.

Proposition A.1 *Let X be a finite-dimensional normed linear vector space with norm $\|\cdot\|$ and let S and T be convex subsets of X. Suppose that*

(i) $\hat{x}_1 \in S \cap T$, $B(\hat{x}_1, r_1) \subset S$, where $r_1 > 0$, and $\|\hat{x}_1\| \leq R_1$, and

(ii) $\hat{x}_2 \in S \cap T$, $B(\hat{x}_2, r_2) \subset T$, where $r_2 > 0$, and $\|\hat{x}_2\| \leq R_2$.

Let $\alpha = \frac{r_2}{r_1+r_2}$, and $r = \frac{r_1 r_2}{r_1+r_2}$, and $\hat{R} = \alpha R_1 + (1-\alpha)R_2$.
Then the point $\hat{x} = \alpha\hat{x}_1 + (1-\alpha)\hat{x}_2$ will satisfy:

$$(i) \quad B(\hat{x}, r) \subset S \cap T \ ,$$

$$(ii) \quad \|\hat{x}\| \leq \hat{R} \ ,$$

$$\text{and} \quad (iii) \quad \frac{\hat{R}}{r} \leq \frac{R_1}{r_1} + \frac{R_2}{r_2} \quad .$$

Proof: First note that $0 \leq \alpha \leq 1$. Because $B(\hat{x}_1, r_1) \subset S$ and $\hat{x}_2 \in S$, $B(\alpha\hat{x}_1 + (1-\alpha)\hat{x}_2, \alpha r_1) \subset S$. Similarly, because $B(\hat{x}_2, r_2) \subset T$ and $\hat{x}_1 \in T$, $B(\alpha\hat{x}_1 + (1-\alpha)\hat{x}_2,\ (1-\alpha)r_2) \subset T$. Noticing that $\alpha r_1 = (1-\alpha)r_2 = r$, we have $B(\hat{x}, r) = B(\alpha\hat{x}_1 + (1-\alpha)\hat{x}_2,\ r) \subset S \cap T$. Also $\|\hat{x}\| \leq \alpha\|\hat{x}_1\| + (1-\alpha)\|\hat{x}_2\| \leq \alpha R_1 + (1-\alpha)R_2 = \hat{R}$. Finally, to show (iii), we have

$$\frac{\hat{R}}{r} = \frac{\alpha R_1 + (1-\alpha)R_2}{r} = \frac{R_1}{r_1} + \frac{R_2}{r_2} \ . \ \blacksquare$$

Proposition A.2 *Let X be a finite-dimensional normed linear vector space with norm $\|\cdot\|$ and let S and T be convex subsets of X. Suppose that*

(i) $\hat{x}_1 \in S \cap T$, $B(\hat{x}_1, r_1) \subset S$, where $r_1 > 0$, and $B(\hat{x}_1, r_1) \subset B(0, R_1)$ and

(ii) $\hat{x}_2 \in S \cap T$, $B(\hat{x}_2, r_2) \subset T$, where $r_2 > 0$, and $B(\hat{x}_2, r_2) \subset B(0, R_2)$.

Let $\alpha = \frac{r_2}{r_1+r_2}$, and $r = \frac{r_1 r_2}{r_1+r_2}$, and $\hat{R} = \alpha R_1 + (1-\alpha)R_2$. Then the point $\hat{x} = \alpha\hat{x}_1 + (1-\alpha)\hat{x}_2$

will satisfy:

$$(i) \quad B(\hat{x}, r) \subset S \cap T \ ,$$

$$(ii) \quad B(\hat{x}, r) \subset B(0, \hat{R}) \ ,$$

$$(iii) \quad \frac{\hat{R}}{r} \leq \frac{R_1}{r_1} + \frac{R_2}{r_2} \ \ ,$$

$$(iv) \quad r \geq \tfrac{1}{2} \min\{r_1, r_2\} \ \ ,$$

$$\textit{and} \quad (v) \quad \hat{R} \leq \max\{R_1, R_2\} \ \ .$$

Proof: Parts (i) and (iii) follow identically the proof of Proposition A.1. To see (iv), note that by definition of r, $r \geq \frac{\min\{r_1,r_2\}\max\{r_1,r_2\}}{2\max\{r_1,r_2\}} = \frac{1}{2}\min\{r_1, r_2\}$. Part (v) follows from the fact that $\hat{R}$ is a convex combination of R_1 and R_2. To prove (ii), note that for any $x \in B(\hat{x}, r)$, we have $\|x\| \leq \|\hat{x}\| + r \leq \alpha\|\hat{x}_1\| + (1-\alpha)\|\hat{x}_2\| + r$. However, $\|\hat{x}_i\| + r_i \leq R_i$, $i = 1, 2$, so that $\|x\| \leq \alpha(R_1 - r_1) + (1-\alpha)(R_2 - r_2) + r = \hat{R} - r \leq \hat{R}$, which completes the proof. ∎

References

[1] F. Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal on Optimization*, 5(1):13–51, 1995.

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

[3] L. Blum, M. Shub, and S. Smale. On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines. *Bull. Amer. Math. Soc.*, 21(1):1–46, 1989.

[4] Sharon Filipowski. On the complexity of solving linear programs specified with approximate data and known to be feasible. Technical Report, Dept. of Industrial and Manufacturing Systems Engineering, Iowa State University, May 1994.

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

[6] Robert M. Freund. Dual gauge programs, with applications to quadratic programming and the minimum-norm problem. *Mathematical Programming*, 38(1):47–67, 1987.

[7] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, New York, 1985.

[8] Yurii Nesterov and Arkadii Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1994.

[9] Manuel A. Nunez and Robert M. Freund. Condition measures and properties of the central trajectory of a linear program. Technical Report W.P. #3889-96-MSA, Sloan School of Management, Massachusetts Institute of Technology, 1996.

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

[11] James Renegar. Incorporating condition measures into the complexity theory of linear programming. *SIAM Journal on Optimization*, 5(3):506–524, 1995.

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

[13] James Renegar. Condition numbers, the barrier method, and the conjugate gradient method. *SIAM Journal on Optimization*, 64(4):879–912, 1996.

[14] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, N.J., 1970.

[15] S. Smale. Some remarks on the foundations of numerical analysis. *SIAM Review*, 32(4):211–220, 1990.

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

[17] Jorge R. Vera. Ill-posedness and the complexity of deciding existence of solutions to linear programs. Technical Report, Cornell University, June 1994. To appear in *SIAM Journal on Optimization.*

[18] Jorge R. Vera. On the complexity of linear programming under finite precision arithmetic. Technical Report, Dept. of Industrial Engineering, University of Chile, August 1994.

[19] K. Yosida. *Functional Analysis.* Springer-Verlag, Berlin, 1987.