Condition Measures and Properties of the Central Trajectory of a Linear Program

Robert M. Freund and Manuel Nunez

WP#3889 -96-MSA March, 1996

Condition Measures and Properties of the Central Trajectory of a Linear Program

Manuel A. Nunez and Robert M. Freund

March 25, 1996

Abstract

The central trajectory of a linear program consists of the set of optimal solutions $x(\mu)$ and $(y(\mu), s(\mu))$ to the logarithmic barrier problems:

$$\begin{array}{ll} (P_\mu(d)): & \min\{c^T x + \mu p(x) : Ax = b, x > 0\}, \\ (D_\mu(d)): & \max\{b^T y - \mu p(s) : A^T y + s = c, s > 0\}, \end{array}$$

where $p(u) = -\sum_{i=1}^{n} \ln(u_i)$, is the logarithmic barrier function, $d = (A, b, c)$ is a data instance in the space of all data $\mathcal{D} = \{(A, b, c) : A \in \Re^{mn}, b \in \Re^m, c \in \Re^n\}$, and the parameter μ is a positive scalar considered independent of the data instance $d \in \mathcal{D}$.

This study shows that certain properties of solutions along the central trajectory of a linear program are inherently related to the condition number $\mathcal{C}(d)$ of the data instance $d = (A, b, c)$, where the condition number $\mathcal{C}(d)$ and a closely-related measure $\rho(d)$ called the "distance to ill-posedness" were introduced by Renegar in a recent series of papers [17, 15, 16]. In the context of the central trajectory problem, $\rho(d)$ essentially measures how close the data instance $d = (A, b, c)$ is to be being infeasible for $(P_\mu(d))$, and $\mathcal{C}(d) \triangleq \|d\|/\rho(d)$ is a scale-invariant reciprocal of the distance to ill-posedness $\rho(d)$, and so $\mathcal{C}(d)$ goes to ∞ as the data instance $d = (A, b, c)$ approaches infeasibility. We present lower and upper bounds on sizes of optimal solutions along the central trajectory, and on rates of change of solutions along the central trajectory as either μ changes or the data d changes, where these bounds are all polynomial functions of μ and are linear or polynomial functions of the condition number $\mathcal{C}(d)$ and the related distance to ill-posedness $\rho(d)$ of the data instance $d = (A, b, c)$.

1 Introduction, notation, and definitions

The central trajectory of a linear program consists of the set of optimal solutions $x = x(\mu)$ and $(y, s) = (y(\mu), s(\mu))$ to the logarithmic barrier problems:

$$\begin{aligned} (P_\mu(d)): &\quad \min\{c^T x + \mu p(x) : Ax = b, x > 0\}, \\ (D_\mu(d)): &\quad \max\{b^T y - \mu p(s) : A^T y + s = c, s > 0\}, \end{aligned}$$

where $p(u) = -\sum_{i=1}^{n} \ln(u_i)$, is the logarithmic barrier function, $d = (A, b, c)$ is a data instance in the space of all data $\mathcal{D} = \{(A, b, c) : A \in \Re^{mn}, b \in \Re^m, c \in \Re^n\}$, and the parameter μ is a positive scalar considered independent of the data instance $d \in \mathcal{D}$. The central trajectory is fundamental to the study of interior-point algorithms for linear programming, and has been the subject of an enormous volume of research, see among many others, the references cited in the surveys by Gonzaga [8] and Jansen et al [10]. It is well known that programs $(P_\mu(d))$ and $(D_\mu(d))$ are related through Lagrangian duality; if either program is feasible, then both programs attain their optima, and optimal solutions $x = x(\mu)$ and $(y, s) = (y(\mu), s(\mu))$ satisfy $c^T x - b^T y = n\mu$, and hence exhibit a linear programming duality gap of $n\mu$ for the dual linear programming problems associated with $(P_\mu(d))$ and $(D_\mu(d))$.

The purpose of this paper is to explore and demonstrate properties of solutions to $(P_\mu(d))$ and $(D_\mu(d))$ that are inherently related to the condition number $\mathcal{C}(d)$ of the data instance $d = (A, b, c)$, where the condition number $\mathcal{C}(d)$ and a closely-related measure $\rho(d)$ called the "distance to ill-posedness" were introduced by Renegar in a recent series of papers [17, 15, 16]. In the context of the central trajectory problem, $\rho(d)$ essentially measures how close the data instance $d = (A, b, c)$ is to being infeasible for $(P_\mu(d))$, and $\mathcal{C}(d) \triangleq \|d\|/\rho(d)$ is a scale-invariant reciprocal of the distance to ill-posedness $\rho(d)$, and so $\mathcal{C}(d)$ goes to ∞ as the data instance $d = (A, b, c)$ approaches infeasibility. We now present these concepts in more detail.

The data for the programs $(P_\mu(d))$ and $(D_\mu(d))$ is the array $d = (A, b, c)$, where $d = (A, b, c) \in \mathcal{D} = \{(A, b, c) : A \in \Re^{mn}, b \in \Re^m, c \in \Re^n\}$ and the positive scalar μ is treated as a parameter independent of the data $d = (A, b, c)$. Consider the following subset of the data set $\mathcal{D}$:

$$\mathcal{F} = \{(A, b, c) \in \mathcal{D} : \text{ there exists } (x, y) \text{ such that } Ax = b, x > 0, A^T y < c\},$$

that is, the elements in $\mathcal{F}$ correspond to those instances in $\mathcal{D}$ for which $(P_\mu(d))$ and $(D_\mu(d))$ are feasible. The complement of $\mathcal{F}$, denoted by $\mathcal{F}^C$, is the set of data instances $d = (A, b, c)$

for which $(P_\mu(d))$ and $(D_\mu(d))$ are infeasible. The boundary of $\mathcal{F}$ and $\mathcal{F}^C$ is the set

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

where ∂S denotes the boundary of a set S, and $\text{cl}(S)$ is the closure of a set S. Note that $\mathcal{B} \neq \emptyset$ since $(0,0,0) \in \mathcal{B}$. The data instances $d = (A, b, c)$ in $\mathcal{B}$ are called the ill-posed data instances, in that arbitrarily small changes in the data $d = (A, b, c)$ yield data instances in $\mathcal{F}$ as well as data instances in $\mathcal{F}^C$.

In order to measure the "distance to ill-posedness" of a given data instance, we need to define a norm over the data set $\mathcal{D}$; and to do so we first define norms for $\Re^n$ and $\Re^m$. We assume that $\Re^n$ is a normed vector space and that for any $x \in \Re^n$, $\|x\|$ denotes the norm of the vector x. We also assume that $\Re^m$ is a normed vector space and that for any $y \in \Re^m$, $\|y\|$ denotes the norm of the vector y. Observe that even though we are using the same notation for the norm in $\Re^m$ and the norm in $\Re^n$, they are not necessarily the same norms. We do not explicitly make the distinction because when computing the norm of a given vector it is clear from the context or from the dimension of the vector what norm we are employing. We associate with $\Re^n$ and $\Re^m$ the dual spaces $(\Re^n)^*$ and $(\Re^m)^*$ of linear functionals defined on $\Re^n$ and $\Re^m$, respectively, and whose (dual) norms are denoted by $\|c\|_*$ for $c \in (\Re^n)^*$ and $\|v\|_*$ for $v \in (\Re^m)^*$, and where the dual norm $\|c\|_*$ induced on the space $(\Re^n)^*$ is defined as:

$$\|c\|_* = \max\{c^T x : \|x\| \leq 1, x \in \Re^n\},$$

and similarly for $\|v\|_*$ for $v \in (\Re^m)^*$. Observe that there exists a natural isomorphism $* : \Re^n \to (\Re^n)^*$ that assigns to each vector $v \in \Re^n$ a linear functional $v^* \in (\Re^n)^*$ defined as $v^* x = v^T x$ for all $x \in \Re^n$. Hence, we can define a new norm on $\Re^n$, namely $\|v\|_* = \|v^*\|$ for all $v \in \Re^n$, where the norm on the right hand side is the dual defined above. The operator $*$ is an isometry between the spaces $(\Re^n, \|\cdot\|_*)$ and $((\Re^n)^*, \|\cdot\|)$. Similar remarks hold concerning norms arising from $\Re^m$.

We next define norms for linear operators. Let $\mathcal{X}$ and $\mathcal{Y}$ be finite-dimensional normed vector spaces with norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{Y}}$, respectively, and let $L(\mathcal{X}, \mathcal{Y})$ be the set of all linear operators from $\mathcal{X}$ to $\mathcal{Y}$. Then for a given linear operator T in $L(\mathcal{X}, \mathcal{Y})$ we define $\|T\|$ to be the operator norm, namely,

$$\|T\| = \max\left\{\|Tx\|_{\mathcal{Y}} : x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1\right\}.$$

Given a data instance $d = (A, b, c)$, A is both a matrix of mn real numbers as well as a linear operator mapping the vector space $(\Re^n, \|\cdot\|)$ into the vector space $(\Re^m, \|\cdot\|)$. Similarly, A^T is both a matrix of mn real numbers as well as a linear operator mapping the vector space $((\Re^m)^*, \|\cdot\|_*)$ into the vector space $((\Re^n)^*, \|\cdot\|_*)$. It is elementary to show that by using these characterizations that $\|A\| = \|A^T\|$.

Finally, if u and v are vectors in $\Re^k$ and $\Re^l$, respectively, we can define the norm of the product vector (u, v) as $\|(u,v)\| = \max\{\|u\|, \|v\|\}$, whose corresponding dual norm is $\|(u,v)\|_* = \|u\|_* + \|v\|_*$.

For $d = (A, b, c) \in \mathcal{D}$, we define the product norm on the Cartesian product $\Re^{mn} \times \Re^m \times \Re^n$ as

$$\|d\| = \max\{\|A\|, \|b\|, \|c\|_*\},$$

for all $d \in \mathcal{D}$, where $\|A\|$ is the operator norm associated with the linear operator A, $\|b\|$ is the norm specified in $\Re^m$, and $\|c\|_*$ is the isometric dual norm on $\Re^n$.

For $d \in \mathcal{D}$, we define the ball centered at d with radius δ as:

$$B(d, \delta) = \{\bar{d} \in \mathcal{D} : \|\bar{d} - d\| \leq \delta\}.$$

For a data instance $d \in \mathcal{D}$, the "distance to ill-posedness" is defined as follows:

$$\rho(d) = \inf\{\|d - \bar{d}\| : \bar{d} \in \mathcal{B}\},$$

see [17, 15, 16], and so $\rho(d)$ is the distance of the data instance $d = (A, b, c)$ to the set of ill-posed instances $\mathcal{B}$. It is straightforward to show that

$$\rho(d) = \begin{cases} \sup\{\delta : B(d,\delta) \subset \mathcal{F}\} & \text{if } d \in \mathcal{F}, \\ \sup\{\delta : B(d,\delta) \subset \mathcal{F}^C\} & \text{if } d \in \mathcal{F}^C, \end{cases} \tag{1}$$

so that we could also define $\rho(d)$ by employing (1). The "condition number" $\mathcal{C}(d)$ of the data instance d is defined as

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

when $\rho(d) > 0$, and $\mathcal{C}(d) = \infty$ when $\rho(d) = 0$. The condition number $\mathcal{C}(d)$ can be viewed as a scale-invariant reciprocal of $\rho(d)$, as it is elementary to demonstrate that $\mathcal{C}(d) = \mathcal{C}(\alpha d)$ for any positive scalar α. Observe that since $\bar{d} = (\bar{A}, \bar{b}, \bar{c}) = (0, 0, 0) \in \mathcal{B}$ and $\mathcal{B}$ is a closed set, then for any $d \notin \mathcal{B}$ we have $\|d\| = \|d - \bar{d}\| \geq \rho(d) > 0$, so that $\mathcal{C}(d) \geq 1$. The value of

$\mathcal{C}(d)$ is a measure of the relative conditioning of the data instance d.

The study of perturbation theory and information complexity for convex programs in terms of the distance to ill-posedness $\rho(d)$ and the condition number $\mathcal{C}(d)$ of a given data instance d has been the subject of many recent papers. In particular, Renegar in [15] studies perturbations in the very general setting:

$$(RLP): \quad z = \sup\{c^*x : Ax \leq b, x \geq 0, x \in \mathcal{X}\},$$

where $\mathcal{X}$ and $\mathcal{Y}$ denote real normed vector spaces, $A : \mathcal{X} \to \mathcal{Y}$ is a continuous linear operator, $c^* : \mathcal{X} \to \Re$ is a continuous linear functional, and the inequalities $Ax \leq b$ and $x \geq 0$ are induced by any closed convex cones (linear or nonlinear) containing the origin in $\mathcal{X}$ and $\mathcal{Y}$, respectively. Previous to this paper of Renegar, others studied perturbations of linear programs and systems of linear inequalities, but not in terms of the distance to ill-posedness (see [12, 18, 19]). In [16] and [17] Renegar introduces the concept of a *fully efficient* algorithm and provides a fully-efficient algorithm that given any data instance d answers whether the program (RLP) associated with d is consistent or not.

Vera in [23] develops a fully-efficient algorithm for a certain form of linear programming that is a special case of (RLP) in which the spaces are finite-dimensional, the linear inequalities are induced by the nonnegative orthant, and nonnegativity constraints $x \geq 0$ do not appear, that is, when the problem (RLP) is $\min\{c^Tx : Ax \leq b, x \in \Re^n\}$. In [22], Vera establishes similar bounds as Renegar in [15] for norms of optimal primal and dual solutions and optimal objective function values. He then uses these bounds to develop an algorithm for finding approximate optimal solutions of the original instance. In [24] he provides a measure of the precision of a logarithmic barrier algorithm based upon the distance to ill-posedness of the instance. To do this, he follows the same arguments as Den Hertog, Roos, and Terlaky [4], making the appropriate changes when necessary to express their results in terms of the distance to ill-posedness.

Filipowski [5] expands upon Vera's results under the assumption that it is known beforehand that the primal data instance is feasible. In addition, she develops several fully-efficient algorithms that approximate optimal solutions to the original instance under this assumption.

Freund and Vera in [6] address the issue of deciding feasibility of (RLP). The problem that they study is defined as finding x that solves $b - Ax \in C_{\mathcal{Y}}$ and $x \in C_{\mathcal{X}}$, where $C_{\mathcal{X}}$ and

$C_{\mathcal{Y}}$ are closed convex cones in the linear vector spaces $\mathcal{X}$ and $\mathcal{Y}$, respectively. They develop optimization programs that allow one to compute exactly or at least estimate the distance to ill-posedness. They also show additional results relating the distance to ill-posedness to the existence of certain inscribed and circumscribed balls for the feasible region, with implications for Hačijan's ellipsoid algorithm [9].

This paper is organized as follows. In Section 2 we present several properties related to the distance to ill-posedness of the data $d = (A, b, c)$. Lemma 2.1 and Corollary 2.1 state characterizations of sets of ill-posed data instances. Lemma 2.2 and Lemma 2.3 present some elementary properties of the set of ill-posed instances $\mathcal{B}$ and the distance to ill-posedness $\rho(d)$, respectively.

In Section 3 we present results on lower and upper bounds on sizes of optimal solutions along the central trajectory of the dual logarithmic barrier problems $(P_\mu(d))$ and $(D_\mu(d))$. The upper bound results are stated in Theorem 3.1, and the lower bound results are stated in Theorem 3.2 and Theorem 3.3.

In Section 4 we study the sensitivity of optimal solutions along the central trajectory to changes (perturbations) in the data $d = (A, b, c)$. Theorem 4.1 presents upper bounds on changes in optimal solutions along the central trajectory as the data instance $d = (A, b, c)$ is changed to a "nearby" data instance $\bar{d} = (\bar{A}, \bar{b}, \bar{c})$. Theorem 4.2 presents upper bounds on changes in optimal solutions along the central trajectory as the barrier parameter μ is changed and the data instance $d = (A, b, c)$ remains fixed. Corollary 4.3 states upper bounds on the first derivatives $\dot{x} = \dot{x}(\mu)$ and $(\dot{y}, \dot{s}) = (\dot{y}(\mu), \dot{s}(\mu))$ of optimal solutions along the central trajectory with respect to the barrier parameter μ. Finally, Theorem 4.3 presents upper bounds on changes in optimal objective function values along the central trajectory as the data instance $d = (A, b, c)$ is changed to a "nearby" data instance $\bar{d} = (\bar{A}, \bar{b}, \bar{c})$.

Section 5 contains a brief examination of properties of analytic center problems related to condition measures. These properties are used to demonstrate one of the lower bound results in Section 4.

2 Properties related to the distance to ill-posedness

In this section we present several properties related to the distance to ill-posedness of the data (A, b, c) for the logarithmic barrier problem $(P_\mu(d))$ and its dual $(P_\mu(d))$. In Lemma 2.1

and Corollary 2.1, we characterize sets of ill-posed data instances. In Lemma 2.2 and Lemma 2.3, we present some elementary properties of the set of ill-posed instances $\mathcal{B}$ and the distance to ill-posedness $\rho(d)$, respectively.

We first state three elementary propositions. The first two propositions are each a different version of Farkas' Lemma, that are stated for the context of the central trajectory problems studied here.

Proposition 2.1 *Exactly one of the following two systems has a solution:*

- $Ax = b$ *and* $x > 0$.
- $A^T y \geq 0$, $b^T y \leq 0$, *and* $(Ae_n - b)^T y > 0$,

where e_n *denotes the vector* $(1, \ldots, 1)^T$ *in* $\Re^n$.

Proposition 2.2 *Exactly one of the following two systems has a solution:*

- $A^T y < c$.
- $Ax = 0$, $x \geq 0$, $c^T x \leq 0$, *and* $(e_n - c)^T x > 0$,

where e_n *denotes the vector* $(1, \ldots, 1)^T$ *in* $\Re^n$.

The third proposition is a special case of the extension form of the Hahn-Banach Theorem (see Corollary 2 in Luenberger [11], p. 112). A simple and short proof for finite-dimensional spaces is presented in [6].

Proposition 2.3 *Given* $u \in \Re^k$, *there exists* $\bar{u} \in (\Re^k)^*$ *such that* $\bar{u}^T u = \|u\|$ *and* $\|\bar{u}\|_* = 1$.

Now consider the following subsets of the data space $\mathcal{D}$:

$$\mathcal{F}_P = \{(A, b, c) \in \mathcal{D} : \text{ there exists } x \in \Re^n \text{ such that } Ax = b, x > 0\},$$

$$\mathcal{F}_D = \{(A, b, c) \in \mathcal{D} : \text{ there exists } y \in \Re^m \text{ such that } A^T y < c\},$$

that is, $\mathcal{F}_P$ is the set of primal feasible data instances and $\mathcal{F}_D$ is the set of dual feasible data instances. Observe that $\mathcal{F}$, which is the set of instances for which the logarithmic barrier problem $(P_\mu(d))$ (and its dual $(P_\mu(d))$) have optimal solutions, is characterized by $\mathcal{F} = \mathcal{F}_P \cap \mathcal{F}_D$. It is also convenient to introduce the corresponding sets of ill-posed data instances: $\mathcal{B}_P = \text{cl}(\mathcal{F}_P) \cap \text{cl}(\mathcal{F}_P^C) = \partial\mathcal{F}_P = \partial\mathcal{F}_P^C$ and $\mathcal{B}_D = \text{cl}(\mathcal{F}_D) \cap \text{cl}(\mathcal{F}_D^C) = \partial\mathcal{F}_D = \partial\mathcal{F}_D^C$.

Similarly, we define the following distances to ill-posedness of a data instance $d = (A, b, c)$. Let $\rho_P(d) = \inf\{\|d - \bar{d}\| : \bar{d} \in \mathcal{B}_P\}$ and $\rho_D(d) = \inf\{\|d - \bar{d}\| : \bar{d} \in \mathcal{B}_D\}$. Then $\rho_P(d)$ and $\rho_D(d)$ denote the distance to primal ill-posedness and the distance to dual ill-posedness of the data instance d.

We also have alternative definitions of $\rho_P(d)$ and $\rho_D(d)$ analogous to the one given in definition (1):

$$\rho_P(d) = \begin{cases} \sup\{\delta : B(d,\delta) \subset \mathcal{F}_P\} & \text{if } d \in \mathcal{F}_P, \\ \sup\{\delta : B(d,\delta) \subset \mathcal{F}_P^C\} & \text{if } d \in \mathcal{F}_P^C. \end{cases} \tag{2}$$

$$\rho_D(d) = \begin{cases} \sup\{\delta : B(d,\delta) \subset \mathcal{F}_D\} & \text{if } d \in \mathcal{F}_D, \\ \sup\{\delta : B(d,\delta) \subset \mathcal{F}_D^C\} & \text{if } d \in \mathcal{F}_D^C. \end{cases} \tag{3}$$

Likewise, the corresponding condition measures for the primal problem and for the dual problem are $\mathcal{C}_P(d) = \|d\|/\rho_P(d)$ if $\rho_P(d) > 0$ and $\mathcal{C}_P(d) = \infty$, otherwise; $\mathcal{C}_D(d) = \|d\|/\rho_D(d)$ if $\rho_D(d) > 0$ and $\mathcal{C}_D(d) = \infty$, otherwise.

The following lemma describes the closure of various data instance sets.

Lemma 2.1 *The data instance sets cl($\mathcal{F}_P$), cl($\mathcal{F}_P^C$), cl($\mathcal{F}_D$), and cl($\mathcal{F}_D^C$) are characterized as follows:*

$$\begin{aligned} cl(\mathcal{F}_P) &= \{(A,b,c) : \textit{ there exist } x \in \Re^n \textit{ and } r \in \Re \textit{ such that} \\ &\quad\; Ax - br = 0, x \geq 0, r \geq 0, (x,r) \neq 0\}, \\ cl(\mathcal{F}_P^C) &= \{(A,b,c) : \textit{ there exists } u \in \Re^m \textit{ such that} \\ &\quad\; A^T u \leq 0, b^T u \geq 0, u \neq 0\}, \\ cl(\mathcal{F}_D) &= \{(A,b,c) : \textit{ there exist } y \in \Re^m \textit{ and } t \in \Re \textit{ such that} \\ &\quad\; ct - A^T y \geq 0, t \geq 0, (y,t) \neq 0\}, \\ cl(\mathcal{F}_D^C) &= \{(A,b,c) : \textit{ there exists } v \in \Re^n \textit{ such that} \\ &\quad\; Av = 0, c^T v \leq 0, v \geq 0, v \neq 0\}. \end{aligned}$$

Proof: Let $d = (A, b, c) \in \text{cl}(\mathcal{F}_P)$, then there exists a sequence $\{d_h = (A_h, b_h, c_h) : h \in \mathcal{N}\}$, where $\mathcal{N}$ denotes the set of natural numbers, such that $d_h \in \mathcal{F}_P$ for all h, and $\lim_{h\to\infty} d_h =$

d. For each h, we have that there exists x_h such that $A_h x_h = b_h$, and $x_h > 0$. Consider the sequence $\{(\hat{x}_h, \hat{r}_h) : h \in \mathcal{N}\}$, where $\hat{x}_h = \frac{x_h}{\|x_h\|+1}$ and $\hat{r}_h = \frac{1}{\|x_h\|+1}$ for all h. Observe that for each h, $A_h\hat{x}_h - b_h\hat{r}_h = 0$, $\|\hat{x}_h\| + |\hat{r}_h| = 1$, and $(\hat{x}_h, \hat{r}_h) \geq 0$. Hence, there exist a vector $(\hat{x}, \hat{r}) \in \Re^{n+1}$ and a sequence $\{(\hat{x}_{h_k}, \hat{r}_{h_k}) : k \in \mathcal{N}\}$ such that $\lim_{k\to\infty} h_k = \infty$, $\lim_{k\to\infty}(\hat{x}_{h_k}, \hat{r}_{h_k}) = (\hat{x}, \hat{r})$, and $\|\hat{x}\| + |\hat{r}| = 1$. Since, $\lim_{k\to\infty} d_{h_k} = d$, it follows that $A\hat{x} - b\hat{r} = 0, \hat{x} \geq 0, \hat{r} \geq 0, (\hat{x}, \hat{r}) \neq 0$.

On the other hand, for a given data instance $d = (A, b, c)$ assume that there exist x and r such that $Ax - br = 0, x \geq 0, r \geq 0, (x, r) \neq 0$. Then $\|x\| + |r| > 0$. Let $(x_\epsilon, r_\epsilon) = (x + \epsilon e_n, r + \epsilon)$ for any $\epsilon > 0$, where e_n denotes the vector $(1, \ldots, 1)^T$ in $\Re^n$. Then $(x_\epsilon, r_\epsilon) > 0$. From Proposition 2.3, there exists $(\bar{x}_\epsilon, \bar{r}_\epsilon)$ such that $\bar{x}_\epsilon^T x_\epsilon + \bar{r}_\epsilon r_\epsilon = \|x_\epsilon\| + |r_\epsilon| > 0$, and $\max\{\|\bar{x}_\epsilon\|_*, |\bar{r}_\epsilon|\} = 1$. Define $A_\epsilon = A - \frac{\epsilon}{\|x_\epsilon\|+|r_\epsilon|}(Ae_n - b)\bar{x}_\epsilon^T$, $b_\epsilon = b + \frac{\epsilon}{\|x_\epsilon\|+|r_\epsilon|}(Ae_n - b)\bar{r}_\epsilon$. Then, $A_\epsilon x_\epsilon - b_\epsilon r_\epsilon = 0$, $x_\epsilon > 0$, and $r_\epsilon > 0$, whereby $d_\epsilon = (A_\epsilon, b_\epsilon, c) \in \mathcal{F}_P$. Nevertheless, since $\|x_\epsilon\| + |r_\epsilon| \to \|x\| + |r| > 0$ as $\epsilon \to 0$, we have that $\|d_\epsilon - d\| \to 0$ as $\epsilon \to 0$, so that $d \in \text{cl}(\mathcal{F}_P)$. This concludes the proof of the characterization of $\text{cl}(\mathcal{F}_P)$.

Similarly, for a given data instance $d = (A, b, c) \in \text{cl}(\mathcal{F}_P^C)$, there exists a sequence $\{d_h = (A_h, b_h, c_h) : h \in \mathcal{N}\}$, such that $d_h \in \mathcal{F}_P^C$ for all h, and $\lim_{h\to\infty} d_h = d$. For each h, we have from Proposition 2.1 that there exists u_h such that $A_h^T u_h \leq 0$, $b_h^T u_h \geq 0$, and $(b_h - A_h e_n)^T u_h > 0$. Consider the sequence $\{\hat{u}_h : h \in \mathcal{N}\}$, where $\hat{u}_h = \frac{u_h}{\|u_h\|}$ for all h. Observe that for each h, $A_h^T\hat{u}_h \leq 0$, $b_h^T\hat{u}_h \geq 0$, $(b_h - A_h e_n)^T\hat{u}_h > 0$, and $\|\hat{u}_h\| = 1$. Hence, there exists a vector $\hat{u} \in \Re^m$ and a sequence $\{\hat{u}_{h_k} : k \in \mathcal{N}\}$ such that $\lim_{k\to\infty} h_k = \infty$, $\lim_{k\to\infty}\hat{u}_{h_k} = \hat{u}$, and $\|\hat{u}\| = 1$. Since $\lim_{k\to\infty} d_{h_k} = d$, it follows that $A^T\hat{u} \leq 0, \hat{u}^T b \geq 0, \hat{u} \neq 0$.

Now, suppose that for a given data instance $d = (A, b, c)$ there exists u such that $A^T u \leq 0, b^T u \geq 0, u \neq 0$. Without loss of generality we assume that $\|u\|_* = 1$. Using Proposition 2.3, there exists $\bar{u}$ such that $\bar{u}^T u = \|u\|_* = 1$ and $\|\bar{u}\| = 1$. For a given $\epsilon > 0$, let $\Delta b_\epsilon = \epsilon\bar{u}$, then since $A^T u \leq 0$, $u \neq 0$, and $(b + \Delta b_\epsilon)^T u > 0$, it follows from Proposition 2.1 that $d_\epsilon = (A, b + \Delta b_\epsilon, c) \in \mathcal{F}_P^C$. Since $\lim_{\epsilon\to 0} d_\epsilon = d$, it follows that $d \in \text{cl}(\mathcal{F}_P^C)$. This concludes the proof of the characterization of $\text{cl}(\mathcal{F}_P^C)$.

Similar arguments show the remaining characterizations of $\text{cl}(\mathcal{F}_D)$ and $\text{cl}(\mathcal{F}_D^C)$.
q.e.d.

As an immediate consequence of Lemma 2.1 we obtain:

Corollary 2.1

$$\begin{aligned} \mathcal{B}_P = \{&(A, b, c) : \textit{there exist } x \in \Re^n, r \in \Re, \textit{ and } u \in \Re^m \textit{ such that} \\ &Ax - br = 0, x \geq 0, r \geq 0, (x, r) \neq 0, A^T u \leq 0, b^T u \geq 0, u \neq 0\}, \end{aligned}$$

and

$$\begin{aligned}\mathcal{B}_D &= \{(A,b,c): \text{ there exist } y \in \Re^m, t \in \Re, \text{ and } v \in \Re^n \text{ such that} \\ &\quad ct - A^T y \geq 0, t \geq 0, (y,t) \neq 0, Av = 0, c^T v \leq 0, v \geq 0, v \neq 0\}.\end{aligned}$$

The next lemma relates the three sets of ill-posed data instances.

Lemma 2.2 $((\mathcal{B}_P \cup \mathcal{B}_D) \cap \mathcal{F}) \subset \mathcal{B} \subset (\mathcal{B}_P \cup \mathcal{B}_D)$.

Proof: Suppose that $d \in \mathcal{B}$. Then, given any $\epsilon > 0$, we have that there exist $\bar{d}$ and $\hat{d}$ such that $\bar{d} \in B(d,\epsilon) \cap \mathcal{F}$ and $\hat{d} \in B(d,\epsilon) \cap \mathcal{F}^C$. Since $\bar{d} \in \mathcal{F}$, it follows that $B(d,\epsilon) \cap \mathcal{F}_P \neq \emptyset$ and $B(d,\epsilon) \cap \mathcal{F}_D \neq \emptyset$. Therefore, it follows by letting $\epsilon \to 0$ that $d \in \mathrm{cl}(\mathcal{F}_P)$ and $d \in \mathrm{cl}(\mathcal{F}_D)$. On the other hand, $\hat{d} \in \mathcal{F}^C$ implies that $\hat{d} \in \mathcal{F}_P^C$ or $\hat{d} \in \mathcal{F}_D^C$. Therefore, it also follows by letting $\epsilon \to 0$ that $d \in \mathrm{cl}(\mathcal{F}_P^C)$ or $d \in \mathrm{cl}(\mathcal{F}_D^C)$. In conclusion, $d \in \mathcal{B}_P \cup \mathcal{B}_D$.

Now, assume that $d \in (\mathcal{B}_P \cup \mathcal{B}_D) \cap \mathcal{F}$. Since $d \in \mathcal{F}$, then $d \in \mathrm{cl}(\mathcal{F})$ and we only need to show that $d \in \mathrm{cl}(\mathcal{F}^C)$. Given $\epsilon > 0$ and assuming that $d \in \mathcal{B}_P$, it follows that $B(d,\epsilon) \cap \mathcal{F}_P^C \neq \emptyset$, so that $B(d,\epsilon) \cap \mathcal{F}^C \neq \emptyset$, and $d \in \mathrm{cl}(\mathcal{F}^C)$. If $d \in \mathcal{B}_D$, it follows that $B(d,\epsilon) \cap \mathcal{F}_D^C \neq \emptyset$, so that again $B(d,\epsilon) \cap \mathcal{F}^C \neq \emptyset$, and $d \in \mathrm{cl}(\mathcal{F}^C)$, and the result follows.
q.e.d.

Observe that $\bar{d} = (\bar{A},\bar{b},\bar{c}) = (0,0,0) \in \mathcal{B}$ but $\bar{d} \notin \mathcal{F}$, hence the first inclusion of Lemma 2.2 is proper. Moreover, if d is the following data instance:

$$d = \left(\begin{bmatrix} 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right),$$

then $d \in \mathcal{F}^C$, $d \in \mathrm{cl}(\mathcal{F}_P)$, and $d \in \mathrm{cl}(\mathcal{F}_P^C)$, so that $d \in \mathcal{B}_P \cup \mathcal{B}_D$. Nevertheless, $d \notin \mathrm{cl}(\mathcal{F})$, hence $d \notin \mathcal{B}$, and the second inclusion of Lemma 2.2 is also proper.

The next result relates the three distances to ill-posed sets.

Lemma 2.3 $\rho(d) = \min\{\rho_P(d), \rho_D(d)\}$ *for each data instance* $d = (A,b,c) \in \mathcal{F}$.

Proof: In this lemma we use the alternative definitions (1), (2), and (3) of $\rho(d)$, $\rho_P(d)$, and $\rho_D(d)$, respectively. Suppose that $d \in \mathcal{F}$. Given any $\delta \geq 0$ such that $\delta < \rho(d)$, then $B(d,\delta) \subset \mathcal{F}$, and it follows that $B(d,\delta) \subset \mathcal{F}_P$ and $B(d,\delta) \subset \mathcal{F}_D$. Hence, $\delta \leq \rho_P(d)$ and $\delta \leq \rho_D(d)$, that is $\delta \leq \min\{\rho_P(d), \rho_D(d)\}$. Therefore, $\rho(d) \leq \min\{\rho_P(d), \rho_D(d)\}$. On the other hand, let ϵ be an arbitrary positive scalar, and without loss of generality assume that $\min\{\rho_P(d), \rho_D(d)\} = \rho_P(d)$. Since $\rho_P(d) = \sup\{\delta : B(d,\delta) \subset \mathcal{F}_P\}$, it follows that there exists δ such that $B(d,\delta) \subset \mathcal{F}_P$ and $\rho_P(d) - \epsilon \leq \delta < \rho_P(d) \leq \rho_D(d)$. Moreover, $\delta < \rho_D(d)$

implies $B(d,\delta) \subset \mathcal{F}_D$. Hence, $B(d,\delta) \subset \mathcal{F}_P \cap \mathcal{F}_D = \mathcal{F}$, so that $\rho_P(d) - \epsilon \leq \delta \leq \rho(d)$. Therefore, because ϵ is arbitrary, we have that $\rho(d) \geq \rho_P(d) = \min\{\rho_P(d), \rho_D(d)\}$, concluding the proof.
q.e.d.

3 Upper and lower bounds of solutions along the central trajectory

This section presents results on lower and upper bounds on sizes of optimal solutions along the central trajectory, for the pair of dual logarithmic barrier problems $(P_\mu(d))$ and $(D_\mu(d))$, as well as upper bounds on the sizes of changes in optimal solutions as the data is changed. As in the previous section, we assume that $d = (A, b, c)$ represents a data instance. Before presenting the first bound, we define the following constant, denoted $\mathcal{K}_\mu(d)$, which arises in many of the results to come.

$$\mathcal{K}_\mu(d) = \mathcal{C}(d)^2 + \frac{\mu n}{\rho(d)}.$$

The first result concerns upper bounds on sizes of optimal solutions.

Theorem 3.1 *If $d = (A, b, c) \in \mathcal{F}$ and $\rho(d) > 0$, then*

$$\|\hat{x}\| \leq \mathcal{C}(d)^2 + \frac{\mu n}{\rho(d)} = \mathcal{K}_\mu(d),$$

$$\|\hat{y}\|_* \leq \mathcal{C}(d)^2 + \frac{\mu n}{\rho(d)} = \mathcal{K}_\mu(d),$$

$$\|\hat{s}\|_* \leq 2\|d\| \left(\mathcal{C}(d)^2 + \frac{\mu n}{\rho(d)} \right) = 2\|d\|\mathcal{K}_\mu(d),$$

for any optimal solution $\hat{x}$ to $(P_\mu(d))$ and any optimal solution $(\hat{y}, \hat{s})$ to the dual problem $(D_\mu(d))$.

This theorem states that the norms of optimal solutions along the central trajectory are bounded above by quantities only involving the condition number $\mathcal{C}(d)$ and the distance to ill-posedness $\rho(d)$ of the data d, as well as the dimension n and the barrier parameter μ. Furthermore, for example, the theorem shows that the norm of the optimal primal solution along the central trajectory grows at most linearly in the barrier parameter μ, and at a rate no larger than $n/\rho(d)$.

Proof of Theorem 3.1: Let $\hat{x}$ be an optimal solution to $(P_\mu(d))$ and $(\hat{y}, \hat{s})$ an optimal solution to the corresponding dual problem $(D_\mu(d))$. Note that the optimality conditions of $(P_\mu(d))$ and $(D_\mu(d))$ imply that $c^T\hat{x} = b^T\hat{y} + \mu n$. Note also that by Proposition 2.3, there exists a vector $\bar{x}$ such that $\bar{x}^T\hat{x} = \|\hat{x}\|$ and $\|\bar{x}\|_* = 1$. Similarly, by Proposition 2.3, there exists a vector $\bar{y}$ such that $\bar{y}^T\hat{y} = \|\hat{y}\|_*$ and $\|\bar{y}\| = 1$.

Observe that since $\hat{s} = c - A^T\hat{y}$, then $\|\hat{s}\|_* \leq \|c\|_* + \|A^T\|_*\|y\|_*$, where $\|A^T\|_* = \max\{\|A^Ty\|_* : \|y\|_* \leq 1\} = \|A\|$. Thus, $\|\hat{s}\|_* \leq \|d\|(1 + \|y\|_*)$, and using the fact that $\mathcal{C}(d) \geq 1$ the bound on $\|\hat{s}\|_*$ is a consequence of the bound on $\|\hat{y}\|_*$. It therefore is sufficient to prove the bounds on $\|\hat{x}\|$ and on $\|\hat{y}\|_*$.

The rest of the proof proceeds by examining three cases:

1. $c^T\hat{x} \leq 0$,

2. $0 < c^T\hat{x} \leq \mu n$, and

3. $\mu n < c^T\hat{x}$.

In case (1), let $\Delta A = -\frac{1}{\|\hat{x}\|}b\bar{x}^T$. Then $(A + \Delta A)\hat{x} = 0$, $\hat{x} > 0$, and $c^T\hat{x} \leq 0$. If $\bar{d} = (A + \Delta A, b, c)$ is primal infeasible we have $\|\bar{d} - d\| \geq \rho(d) > 0$. If $\bar{d}$ is primal feasible, then $(P_\mu(\bar{d}))$ is unbounded ($\hat{x}$ is a ray of $(P_\mu(\bar{d}))$) and so its dual $(D_\mu(\bar{d}))$ is infeasible, so that again $\|\bar{d} - d\| \geq \rho(d) > 0$. In either instance, $\rho(d) \leq \|\bar{d} - d\| = \|\Delta A\| = \frac{\|\bar{x}\|_*\|b\|}{\|\hat{x}\|} = \frac{\|b\|}{\|\hat{x}\|} \leq \frac{\|d\|}{\|\hat{x}\|}$. Therefore, $\|\hat{x}\| \leq \mathcal{C}(d) \leq \mathcal{C}(d)^2 + \frac{\mu n}{\rho(d)}$, since $\mathcal{C}(d) \geq 1$ for any d. This proves the bound on $\|\hat{x}\|$ for this case.

The bound on $\|\hat{y}\|_*$ is trivial if $\hat{y} = 0$, so we assume that $\hat{y} \neq 0$. Let $\theta = b^T\hat{y}$, $\Delta b = -\theta\frac{\bar{y}}{\|\hat{y}\|_*}$, $\Delta A = -\frac{1}{\|\hat{y}\|_*}\bar{y}c^T$, and $\bar{d} = (A + \Delta A, b + \Delta b, c)$. Observe that $(b + \Delta b)^T\hat{y} = 0$ and $(A + \Delta A)^T\hat{y} < 0$, so that $\rho(d) \leq \|\bar{d} - d\| = \frac{\max\{\|c\|_*, |\theta|\}}{\|\hat{y}\|_*}$. Hence, $\|\hat{y}\|_* \leq \max\{\mathcal{C}(d), \frac{|\theta|}{\rho(d)}\}$. Furthermore, $|\theta| = |c^T\hat{x} - \mu n| \leq \|\hat{x}\|\|c\|_* + \mu n \leq \mathcal{C}(d)\|d\| + \mu n$. Therefore, again using the fact that $\mathcal{C}(d) \geq 1$ for any d, we have $\|\hat{y}\|_* \leq \mathcal{C}(d)^2 + \frac{\mu n}{\rho(d)}$.

In case (2), let $\bar{d} = (A + \Delta A, b, c + \Delta c)$, where $\Delta A = -\frac{1}{\|\hat{x}\|}b\bar{x}^T$ and $\Delta c = -\mu n\frac{\bar{x}}{\|\hat{x}\|}$. Observe that $(A + \Delta A)\hat{x} = 0$ and $(c + \Delta c)^T\hat{x} \leq 0$. Using similar logic to that in the first part of case (1), we conclude that $\rho(d) \leq \|\bar{d} - d\| = \max\{\|\Delta A\|, \|\Delta c\|_*\} = \frac{\max\{\|b\|, \mu n\}}{\|\hat{x}\|} \leq \frac{\|d\| + \mu n}{\|\hat{x}\|}$.

Therefore, $\|\hat{x}\| \leq \mathcal{C}(d) + \frac{\mu n}{\rho(d)} \leq \mathcal{C}(d)^2 + \frac{\mu n}{\rho(d)}$, since $\mathcal{C}(d) \geq 1$ for any d. This proves the bound on $\|\hat{x}\|$ for this case.

The bound on $\|\hat{y}\|_*$ is trivial if $\hat{y} = 0$, so we assume that $\hat{y} \neq 0$. Let $\bar{d} = (A + \Delta A, b + \Delta b, c)$, where $\Delta A = -\frac{1}{\|\hat{y}\|_*}\bar{y}c^T$ and $\Delta b = \mu n \frac{\bar{y}}{\|\hat{y}\|_*}$. Observe that $(b + \Delta b)^T \hat{y} = b^T \hat{y} + \mu n = c^T \hat{x} > 0$ and $(A + \Delta A)^T \hat{y} < 0$. Using similar logic to that in the first part of case (1), we conclude that $\rho(d) \leq \|\bar{d} - d\| = \max\{\|\Delta A\|, \|\Delta b\|\} = \frac{\max\{\|c\|_*, \mu n\}}{\|\hat{y}\|_*} \leq \frac{\|d\| + \mu n}{\|\hat{y}\|_*}$. Therefore, $\|\hat{y}\|_* \leq \mathcal{C}(d) + \frac{\mu n}{\rho(d)} \leq \mathcal{C}(d)^2 + \frac{\mu n}{\rho(d)}$.

In case (3), we first consider the bound on $\|\hat{y}\|_*$. Again noting that this bound on $\|\hat{y}\|_*$ is trivial if $\hat{y} = 0$, we assume that $\hat{y} \neq 0$. Then let $\bar{d} = (A + \Delta A, b, c)$, where $\Delta A = -\frac{1}{\|\hat{y}\|_*}\bar{y}c^T$. Since $(A + \Delta A)^T \hat{y} < 0$ and $b^T \hat{y} = c^T \hat{x} - \mu n > 0$, it follows from the same logic as in the first part of case (1) that $\rho(d) \leq \|\bar{d} - d\| = \frac{\|c\|_*}{\|\hat{y}\|_*}$. Therefore, $\|\hat{y}\|_* \leq \frac{\|c\|_*}{\rho(d)} \leq \mathcal{C}(d) \leq \mathcal{C}(d)^2 + \frac{\mu n}{\rho(d)}$.

Finally, let $\Delta A = -\frac{1}{\|\hat{x}\|} b\bar{x}^T$ and $\Delta c = -\theta \frac{\bar{x}}{\|\hat{x}\|}$, where $\theta = c^T \hat{x}$. Observe that $(A + \Delta A)\hat{x} = 0$ and $(c + \Delta c)^T \hat{x} = 0$. Using the same argument as in the previous cases, we conclude that $\rho(d) \leq \|\bar{d} - d\| = \max\{\|\Delta A\|, \|\Delta c\|_*\} = \frac{\max\{\|b\|, \theta\}}{\|\hat{x}\|}$, so that $\|\hat{x}\| \leq \max\{\mathcal{C}(d), \frac{\theta}{\rho(d)}\}$. Furthermore, $\theta = b^T \hat{y} + \mu n \leq \|b\| \|\hat{y}\|_* + \mu n \leq \|d\| \mathcal{C}(d) + \mu n$. Therefore, $\|\hat{x}\| \leq \mathcal{C}(d)^2 + \frac{\mu n}{\rho(d)}$, because $\mathcal{C}(d) \geq 1$.

q.e.d.

Remark 1 Note that $\mathcal{K}_\mu(d)$ is scale invariant in the sense that $\mathcal{K}_{\lambda\mu}(\lambda d) = \mathcal{K}_\mu(d)$ for any $\lambda > 0$. From this it follows that the bounds from Theorem 3.1 on $\|\hat{x}\|$ and $\|\hat{y}\|_*$ are also scale invariant. However, as one would expect, the bound on $\|\hat{s}\|_*$ is not scale invariant, since $\|\hat{s}\|_*$ is sensitive to scalings of the form λd. Moreover, observe that as $\mu \to 0$ these bounds converge to the bounds presented by Vera in [22] for optimal solutions to linear programs of the form $\min\{c^T x : Ax = b, x \geq 0\}$.

We next consider upper bounds on solutions of $(P_\mu(\bar{d}))$ and $(D_\mu(\bar{d}))$, where $\bar{d}$ is a data instance that is a small perturbation of the data instance d. Let

$$P_\mu^*(d, \delta) = \{x : x \text{ is an optimal solution to } (P_\mu(\bar{d})) \text{ for some } \bar{d} \in B(d, \delta)\},$$

$$D_\mu^*(d, \delta) = \{(y, s) : (y, s) \text{ is an optimal solution to } (D_\mu(\bar{d})) \text{ for some } \bar{d} \in B(d, \delta)\}.$$

Then $P^*_\mu(d,\delta)$ and $D^*_\mu(d,\delta)$ consist of all optimal solutions to perturbed problems of the form $(P_\mu(\bar{d}))$ and $(D_\mu(\bar{d}))$, respectively, for all $\bar{d}$ satisfying $\|d-\bar{d}\| \leq \delta$. Then from Theorem 3.1 we obtain the following corollary, which presents upper bounds on the sizes of solutions to these perturbed problems:

Corollary 3.1 *Let $\alpha \in (0,1)$ be given and fixed, and let δ be such that $\delta \leq \alpha\rho(d)$, where $d \in \mathcal{F}$ and $\rho(d) > 0$. Then*

$$\|x\| \leq \left(\frac{1+\alpha}{1-\alpha}\right)^2 \left(\mathcal{C}(d)^2 + \frac{\mu n}{\rho(d)}\right) = \left(\frac{1+\alpha}{1-\alpha}\right)^2 \mathcal{K}_\mu(d),$$

$$\|y\|_* \leq \left(\frac{1+\alpha}{1-\alpha}\right)^2 \left(\mathcal{C}(d)^2 + \frac{\mu n}{\rho(d)}\right) = \left(\frac{1+\alpha}{1-\alpha}\right)^2 \mathcal{K}_\mu(d),$$

$$\|s\|_* \leq 2(\|d\|+\delta)\left(\frac{1+\alpha}{1-\alpha}\right)^2 \left(\mathcal{C}(d)^2 + \frac{\mu n}{\rho(d)}\right) = 2(\|d\|+\delta)\left(\frac{1+\alpha}{1-\alpha}\right)^2 \mathcal{K}_\mu(d),$$

for all $x \in P^*_\mu(d,\delta)$ and $(y,s) \in D^*_\mu(d,\delta)$.

Proof: The proof follows by observing that for $\bar{d} \in B(d,\delta)$ we have $\|\bar{d}\| \leq \|d\| + \delta$, and $\rho(\bar{d}) \geq (1-\alpha)\rho(d)$, so that

$$\mathcal{C}(\bar{d}) \leq \frac{\|d\|+\delta}{(1-\alpha)\rho(d)} = \left(\frac{1}{1-\alpha}\right)(\mathcal{C}(d) + \delta/\rho(d)) \leq \left(\frac{1}{1-\alpha}\right)(\mathcal{C}(d)+\alpha) \leq \mathcal{C}(d)\left(\frac{1+\alpha}{1-\alpha}\right),$$

since $\mathcal{C}(d) \geq 1$.
q.e.d.

Note that for a fixed value α that Corollary 3.1 shows that the norms of solutions to any suitably perturbed problem are uniformly upper-bounded by a fixed constant times the upper bounds on the solutions to the original problem.

The next result presents a lower bound on the norm of any primal optimal solution to the central trajectory problem $(P_\mu(d))$.

Theorem 3.2 *If the program $(P_\mu(d))$ has an optimal solution $\hat{x}$ and $\rho(d) > 0$, then*

$$\|\hat{x}\| \geq \frac{1}{2\|d\|}\left(\frac{\mu n}{\mathcal{C}(d)^2 + \frac{\mu n}{\rho(d)}}\right) = \frac{\mu n}{2\|d\|\mathcal{K}_\mu(d)}$$

and,

$$\hat{x}_j \geq \frac{1}{2\|d\|}\left(\frac{\mu m_0}{\mathcal{C}(d)^2 + \frac{\mu n}{\rho(d)}}\right) = \frac{\mu m_0}{2\|d\|\mathcal{K}_\mu(d)}$$

for all $j = 1, \ldots, n$*, where* $m_0 = \min\{\|v\|_* : v \in \Re^n, \|v\|_\infty = 1\}$*, and* $\|v\|_\infty = \max\{|v_j| : 1 \leq j \leq n\}$.

This theorem shows that $\|\hat{x}\|$ and $\hat{x}_j$ are bounded from below by functions only involving the quantities $\|d\|$, $\mathcal{C}(d)$, $\rho(d)$, n, and μ (plus the constant m_0, which only depends on the norm used). Furthermore, the theorem shows that for μ close to zero, that $\hat{x}_j$ grows at least linearly in μ, and at a rate that is at least $m_0/(2\|d\|\mathcal{C}(d)^2)$.

The theorem offers less insight when $\mu \to \infty$, since the lower bound on $\|\hat{x}\|$ presented in the theorem converges to $(2\mathcal{C}(d))^{-1}$ as $\mu \to \infty$. When the feasible region is unbounded, it is well known (see also the results at the end of this section) that $\|\hat{x}(\mu)\| \to \infty$ as $\mu \to \infty$, so that as $\mu \to \infty$ the lower bound of Theorem 3.2 does not adequately capture the behavior of the sizes of optimal solutions to $(P_\mu(d))$ when the feasible region is unbounded. We will present a more relevant bound shortly, in Theorem 3.3.

Note also that the constant m_0 is completely independent of the data (A, b, c), and in fact m_0 only depends on the properties of the norm $\|v\|_*$ relative to the infinity norm $\|v\|_\infty$.

Proof of Theorem 3.2: By the Karush-Kuhn-Tucker optimality conditions of the dual pair of problems $(P_\mu(d))$ and $(D_\mu(d))$, we have that $\hat{s}^T\hat{x} = \mu n$, where $\hat{s}$ is the corresponding dual variable. Since $\hat{s}^T\hat{x} \leq \|\hat{s}\|_*\|\hat{x}\|$, it follows that $\|\hat{x}\| \geq \frac{\mu n}{\|\hat{s}\|_*}$ and the first inequality follows from Theorem 3.1.

For the second inequality, observe that $\mu = \hat{s}_j\hat{x}_j$, thus

$$\hat{x}_j = \frac{\mu}{\hat{s}_j} \geq \frac{\mu}{\|\hat{s}\|_\infty} \geq \frac{\mu m_0}{\|\hat{s}\|_*}.$$

Observe that m_0 is such that $\|\hat{s}\|_* \geq m_0\|\hat{s}\|_\infty$. Therefore, the result follows again from Theorem 3.1.

q.e.d.

The following corollary uses Theorem 3.2 to provide lower bounds for solutions to perturbed problems.

Corollary 3.2 *Let $\alpha \in (0,1)$ be given and fixed, and let δ be such that $\delta \leq \alpha\rho(d)$, where $d \in \mathcal{F}$ and $\rho(d) > 0$. If $x \in P^*_\mu(d,\delta)$, then*

$$\|x\| \geq \left(\frac{1-\alpha}{1+\alpha}\right)^2 \frac{\mu n}{2(\|d\|+\delta)\mathcal{K}_\mu(d)},$$

and

$$x_j \geq \left(\frac{1-\alpha}{1+\alpha}\right)^2 \frac{\mu m_0}{2(\|d\|+\delta)\mathcal{K}_\mu(d)},$$

for all $j = 1, \ldots, n$, where m_0 is the constant defined in Theorem 3.2.

Proof: The proof follows the same logic as that of Corollary 3.1.
q.e.d.

Note that for a fixed value α that Corollary 3.2 shows that the norms of solutions to any suitably perturbed problem are uniformly lower-bounded by a fixed constant times the lower bounds on the solutions to the original problem.

The last result of this section, Theorem 3.3, presents different lower bounds on components of $\hat{x}$ along the central trajectory, that are relevant when $\mu \to \infty$ and when the primal feasible region is unbounded. We will prove this theorem in Section 5. In this theorem, $\bar{\mathcal{C}}_D(d_B)$ denotes a certain condition number that is independent of μ and only depends on part of the data instance d associated with a certain partition of the indices of the components of x. We will formally define this other condition number in Section 5.

Theorem 3.3 *If the central trajectory problem $(P_\mu(d))$ has an optimal solution $x(\mu)$, then there exists a unique partition of the indices $\{1, \ldots, n\}$ into two subsets B and N such that*

$$x_j(\mu) \geq \frac{\mu m_0}{2\|d\|\bar{\mathcal{C}}_D(d_B)}$$

for all $j \in B$, and $x_j(\mu)$ is uniformly bounded for all $\mu \geq 0$ for all $j \in N$, where $d_B = (A_B, b, c_B)$ is a data instance in $\Re^{m|B|+m+|B|}$ composed of those elements of d indexed by the set B, and $m_0 = \min\{\|v\|_* : v \in \Re^n, \|v\|_\infty = 1\}$.

Note that the set B is the index set of components of x that are unbounded over the feasible region of $(P_\mu(d))$, and N is the index set of components of x that are bounded over the feasible region of $(P_\mu(d))$. Theorem 3.3 states that as $\mu \to \infty$, that $x_j(\mu)$ for $j \in B$

will go to ∞ at least linearly in μ as $\mu \to \infty$, and at a rate that is at least $m_0/(2\|d\|\bar{\mathcal{C}}_D(d_B))$.

Of course, from Theorem 3.3, it also follows that when the feasible region of $(P_\mu(d))$ is unbounded, that is, $B \neq \emptyset$, that $\lim_{\mu\to\infty} \|x(\mu)\| = \infty$.

Finally, note that Theorem 3.1 combined with Theorem 3.3 state that as $\mu \to \infty$, that $x_j(\mu)$ for $j \in B$ will go to ∞ exactly linearly in μ.

4 Bounds on changes in optimal solutions as the data is changed

In this section, we present upper bounds on changes in optimal solutions to $(P_\mu(d))$ and $(D_\mu(d))$ as the data $d = (A, b, c)$ is changed or as the barrier parameter μ is changed. The major results of this section are contained in Theorem 4.1, Theorem 4.2, Theorem 4.3, and Theorem 4.4. Theorem 4.1 presents upper bounds on the sizes of changes in optimal solutions to $(P_\mu(d))$ and $(D_\mu(d))$ as the data $d = (A, b, c)$ is changed to data $\bar{d} = (\bar{A}, \bar{b}, \bar{c})$ in a specific neighborhood of the original data $d = (A, b, c)$. Theorem 4.2 presents upper bounds on the sizes of changes in optimal solutions to $(P_\mu(d))$ and $(D_\mu(d))$ as the barrier parameter μ is changed. Theorem 4.3 presents an upper bound on the size of the change in the optimal objective function value of $(P_\mu(d))$ as the data $d = (A, b, c)$ is changed to data $\bar{d} = (\bar{A}, \bar{b}, \bar{c})$ in a specific neighborhood of the original data $d = (A, b, c)$. Finally, Theorem 4.4 presents an upper bound on the size of the change in the optimal objective function value of $(P_\mu(d))$ as the barrier parameter μ is changed. Along the way, we also present upper and lower bounds on the norm of the matrix $(A\hat{X}^2A^T)^{-1}$ in Corollary 4.2 as well as upper bounds of the first derivatives of the optimal solutions $x(\mu)$ and $(y(\mu), s(\mu))$ of $(P_\mu(d))$ and $(D_\mu(d))$ with respect to the barrier parameter μ, in Corollary 4.3. Before presenting the main results, we first define some constants that are used in the analysis, and we prove some intermediary results that will be used in the proofs of the main results of this section.

We start by defining the following constants, which relate various norms to various other norms:

$$m_0 = \min\{\|v\|_* : v \in \Re^n, \|v\|_\infty = 1\}, \tag{4}$$

$$M_0 = \max\{\|v\|_* : v \in \Re^n, \|v\|_\infty = 1\}, \tag{5}$$

$$m_2 = \min\{\|v\| : v \in \Re^m, \|v\|_2 = 1\}, \tag{6}$$

$$M_2 = \max\{\|v\| : v \in \Re^m, \|v\|_2 = 1\}, \tag{7}$$

$$m_3 = \min\{\|v\| : v \in \Re^n, \|v\|_2 = 1\}, \tag{8}$$

$$M_3 = \max\{\|v\| : v \in \Re^n, \|v\|_2 = 1\}, \tag{9}$$

$$m_4 = \min\{\|v\| : v \in \Re^n, \|v\|_* = 1\}, \tag{10}$$

$$M_4 = \max\{\|v\| : v \in \Re^n, \|v\|_* = 1\}, \tag{11}$$

where $\|v\|_\infty = \max\{|v_j| : 1 \le j \le n\}$ and $\|v\|_2$ is the Euclidean norm of v. Observe that m_0 is the same constant defined in Theorem 3.2. Note that all of these constants are finite and positive, and are independent of the data $d = (A, b, c)$, and are only dependent on the choice of the norms used.

For the matrix A, recall that $\|A\|$ denotes the usual operator norm for A. Let $\|A\|_2$ denote the norm defined by:

$$\|A\|_2 = \max\{\|Ax\|_2 : \|x\|_2 \le 1\}.$$

The following three propositions establish some elementary properties based on the constants and the above definition.

Proposition 4.1 *The following inequalities hold for the constants (4)-(11).*

(i) $m_0\|v\|_\infty \le \|v\|_* \le M_0\|v\|_\infty$ *for any* $v \in \Re^n$.

(ii) $m_2\|v\|_2 \le \|v\| \le M_2\|v\|_2$ *for any* $v \in \Re^m$.

(iii) $m_3\|v\|_2 \le \|v\| \le M_3\|v\|_2$ *for any* $v \in \Re^n$.

(iv) $m_4\|v\|_* \le \|v\| \le M_4\|v\|_*$ *for any* $v \in \Re^n$.

(v) $(1/M_2)\|v\|_2 \le \|v\|_* \le (1/m_2)\|v\|_2$ *for any* $v \in \Re^m$.

(vi) $(1/M_3)\|v\|_2 \le \|v\|_* \le (1/m_3)\|v\|_2$ *for any* $v \in \Re^n$.

(vii) $(m_2/M_3)\|A\|_2 \le \|A\| \le (M_2/m_3)\|A\|_2$.

Proposition 4.2 *Consider the matrix* AA^T *as a linear operator from* $(\Re^m, \|\cdot\|_*)$ *to* $(\Re^m, \|\cdot\|)$. *Then* *(i)* $(1/M_2^2)\|(AA^T)^{-1}\|_2 \le \|(AA^T)^{-1}\| \le (1/m_2^2)\|(AA^T)^{-1}\|_2$,

(ii) $\rho(d) \le (M_2/m_3)\sqrt{\lambda_1(AA^T)}$,

where $\lambda_1(AA^T)$ *denotes the smallest eigenvalue of* AA^T.

Proof: The proof of (i) follows directly from Proposition 4.1, inequalities (ii) and (v). For the proof of (ii), let $\lambda_1 = \lambda_1(AA^T)$. There exists $\bar{v} \in \Re^m$ with $\|\bar{v}\|_2 = 1$ and $AA^T\bar{v} = \lambda_1\bar{v}$, so that $\|A^T\bar{v}\|_2^2 = \bar{v}^TAA^T\bar{v} = \lambda_1$. Let $\bar{A} = A - \bar{v}\bar{v}^TA$, $\bar{b} = b + \epsilon\bar{v}$ for any $\epsilon > 0$ and small. Then, $\bar{A}^T\bar{v} = 0$ and $\bar{b}^T\bar{v} = b^T\bar{v} + \epsilon \ne 0$, for all $\epsilon > 0$ small. Hence, by Farkas' Lemma, $\bar{A}x = \bar{b}$ and $x > 0$ is an inconsistent system of inequalities. Therefore, $\rho(d) \le \max\{\|\bar{A} - A\|, \|\bar{b} - b\|\} = \|\bar{A} - A\| \le (M_2/m_3)\|\bar{A} - A\|_2 = (M_2/m_3)\|\bar{A}^T\bar{v}\|_2 = (M_2/m_3)\sqrt{\lambda_1}$, thus proving (ii).
q.e.d.

Proposition 4.3 *If* $D \in \Re^{n\times n}$ *is a diagonal matrix with positive diagonal entries, then*

$$\|Dv\|_* \le (M_0/m_0)\max_{1\le j\le n}\{D_{jj}\}\|v\|_*,$$

for any vector $v \in \Re^n$.

Proof: Given any $v \in \Re^n$, we have that $\|Dv\|_* \le M_0\|Dv\|_\infty \le M_0 \max_{1\le j\le n}\{D_{jj}\}\|v\|_\infty \le (M_0/m_0)\max_{1\le j\le n}\{D_{jj}\}\|v\|_*$.
q.e.d.

We now introduce the following notational convention which is standard in the field of interior point methods: if $x \in \Re^n$ and $x > 0$, then $X = \mathrm{diag}(x_1, ..., x_n)$. For any vector $v \in \Re^n$, we regard Xv as a vector in $\Re^n$ as well. We do not regard X as an operator, but rather as a scaling matrix in $\Re^{n\times n}$.

The next result establishes upper and lower bounds on certain quantities as the data $d = (A, b, c)$ is changed to data $\bar{d} = (\bar{A}, \bar{b}, \bar{c})$ in a specific neighborhood of the original data $d = (A, b, c)$. This result will prove useful in proving the theorems in this section. Recall the definition of $P_\mu^*(d, \delta)$ is:

$$P_\mu^*(d,\delta) = \{x : x \text{ is an optimal solution to } (P_\mu(\bar{d})) \text{ for some } \bar{d} \in B(d,\delta)\}.$$

Lemma 4.1 *Suppose that $d = (A, b, c) \in \mathcal{F}$ and $\rho(d) > 0$. Let $\alpha \in (0, 1)$ be given and fixed, and let δ be such that $\delta \leq \alpha\rho(d)$. If $\hat{x}$ is the optimal solution to $(P_\mu(d))$, and $\bar{x} \in P^*_\mu(d, \delta)$, then for $j = 1, ..., n$,*

$$f_1 \left(\frac{\mu(1-\alpha)}{\|d\|\mathcal{K}_\mu(d)}\right)^2 \leq \hat{x}_j\bar{x}_j \leq h_1 \left(\frac{\mathcal{K}_\mu(d)}{1-\alpha}\right)^2, \tag{12}$$

and for any $v \in \Re^n$,

$$\|\hat{X}\bar{X}v\|_* \leq g_1 \left(\frac{\mathcal{K}_\mu(d)}{1-\alpha}\right)^2 \|v\|_*, \tag{13}$$

where $f_1 = \frac{m_0^2}{32}$, $h_1 = \frac{4}{m_3^2}$, $g_1 = \frac{4M_0}{m_0 m_3^2}$, and m_0, M_0, and m_3 are the constants defined in (4), (5), and (8), respectively.

Proof: From Theorem 3.1 we have that $\|\hat{x}\| \leq \mathcal{K}_\mu(d)$, and from Corollary 3.1 we also have that $\|\bar{x}\| \leq (4/(1-\alpha)^2)\mathcal{K}_\mu(d)$. Therefore, using Proposition 4.1, we obtain $\hat{x}_j\bar{x}_j \leq \hat{x}^T\bar{x} \leq \|\hat{x}\|_2\|\bar{x}\|_2 \leq \|\hat{x}\|\|\bar{x}\|/m_3^2 \leq (4/m_3^2)(\mathcal{K}_\mu(d)^2/(1-\alpha)^2) = h_1\mathcal{K}_\mu(d)^2/(1-\alpha)^2$ for all $j = 1, \ldots, n$.

On the other hand, from Theorem 3.2 and Corollary 3.2, it follows that

$$\hat{x}_j \geq \frac{\mu m_0}{2\|d\|\mathcal{K}_\mu(d)},$$

$$\bar{x}_j \geq \frac{(1-\alpha)^2\mu m_0}{8(\|d\| + \delta)\mathcal{K}_\mu(d)} \geq \frac{(1-\alpha)^2\mu m_0}{16\|d\|\mathcal{K}_\mu(d)},$$

for all $j = 1, \ldots, n$. Therefore,

$$\hat{x}_j\bar{x}_j \geq \frac{m_0^2}{32}\frac{\mu^2(1-\alpha)^2}{\|d\|^2\mathcal{K}_\mu(d)^2} = f_1 \left(\frac{\mu(1-\alpha)}{\|d\|\mathcal{K}_\mu(d)}\right)^2,$$

for all $j = 1, \ldots, n$.

Finally, for any $v \in \Re^n$ we have that $\|\hat{X}\bar{X}v\|_* \leq M_0\|\hat{X}\bar{X}v\|_\infty \leq (4M_0/m_3^2)(\mathcal{K}_\mu(d)^2/(1-\alpha)^2)\|v\|_\infty \leq \frac{4M_0}{m_0 m_3^2}(\mathcal{K}_\mu(d)^2/(1-\alpha)^2)\|v\|_* = g_1(\mathcal{K}_\mu(d)^2/(1-\alpha)^2)\|v\|_*$.

q.e.d.

Corollary 4.1 *Let $d = (A, b, c)$ be a data instance in $\mathcal{F}$ such that $\rho(d) > 0$. Let $\hat{x}$ and $\bar{x}$ be the optimal solutions of $(P_\mu(d))$ and $(P_{\bar{\mu}}(d))$, respectively, where $\mu, \bar{\mu} > 0$. Then for $j = 1, ..., n$,*

$$8f_1 \frac{\mu\bar{\mu}}{\|d\|^2 \mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)} \leq \hat{x}_j \bar{x}_j \leq \frac{h_1}{4} \mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d),$$

and for any $v \in \Re^n$,

$$\|\hat{X}\bar{X}v\|_* \leq \frac{g_1}{4} \mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\|v\|_*,$$

$$\|\hat{X}^{-1}\bar{X}^{-1}v\|_* \leq h_2 \frac{\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\|d\|^2}{\mu\bar{\mu}}\|v\|_*,$$

where f_1 , h_1, and g_1 are the constants defined in Lemma 4.1, and $h_2 = \frac{4M_0}{m_0^3}$. In particular, we have that for $j = 1, ..., n$,

$$8f_1 \left(\frac{\mu}{\|d\|\mathcal{K}_\mu(d)}\right)^2 \leq (\hat{x}_j)^2 \leq \frac{h_1}{4} \mathcal{K}_\mu(d)^2$$

and for any $v \in \Re^n$,

$$\|\hat{X}^2 v\|_* \leq \frac{g_1}{4} \mathcal{K}_\mu(d)^2\|v\|_*,$$

$$\|\hat{X}^{-2}v\|_* \leq h_2 \frac{\mathcal{K}_\mu(d)^2\|d\|^2}{\mu^2}\|v\|_*.$$

Proof: From Theorem 3.1 we have that $\|\hat{x}\| \leq \mathcal{K}_\mu(d)$ and $\|\bar{x}\| \leq \mathcal{K}_{\bar{\mu}}(d)$. Hence, by Proposition 4.1, $\hat{x}_j\bar{x}_j \leq \hat{x}^T\bar{x} \leq \|\hat{x}\|_2\|\bar{x}\|_2 \leq (1/m_3^2)\|\hat{x}\|\|\bar{x}\| \leq (h_1/4)\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)$, for $j = 1, \ldots, n$.

On the other hand, from Theorem 3.2 we have that $\hat{x}_j\bar{x}_j \geq \mu\bar{\mu}m_0^2/(4\|d\|^2\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)) = 8f_1\mu\bar{\mu}/(\|d\|^2\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d))$, for $j = 1, \ldots, n$.

Next, for any $v \in \Re^n$ we have that

$$\|\hat{X}\bar{X}v\|_* \leq M_0\|\hat{X}\bar{X}v\|_\infty \leq M_0\frac{h_1}{4}\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\|v\|_\infty \leq$$

$$\frac{M_0}{m_0}\frac{h_1}{4}\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\|v\|_* = \frac{g_1}{4}\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\|v\|_*.$$

Furthermore,

$$\|\hat{X}^{-1}\bar{X}^{-1}v\|_* \leq M_0\|\hat{X}^{-1}\bar{X}^{-1}v\|_\infty \leq M_0 \left(\min_{1\leq j\leq n}\{\hat{x}_j\bar{x}_j\}\right)^{-1} \|v\|_\infty \leq$$

$$\frac{M_0}{m_0}\frac{1}{8f_1}\frac{\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\|d\|^2}{\mu\bar{\mu}}\|v\|_* = h_2\frac{\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\|d\|^2}{\mu\bar{\mu}}\|v\|_*.$$

q.e.d.

Let $\hat{x} > 0$ and $\bar{x} > 0$ be two positive vectors in $\Re^n$. These two vectors can be used to create the matrix $A\hat{X}\bar{X}A^T$ defined by using the diagonal scaling matrices $\hat{X}$ and $\bar{X}$. Then $A\hat{X}\bar{X}A^T$ can also be considered to be a linear operator from $((\Re^m)^*, \|\cdot\|_*)$ to $(\Re^m, \|\cdot\|)$. The next lemma presents lower and upper bounds on the operator norm of the inverse of this linear operator.

Lemma 4.2 *Let $\alpha \in (0,1)$ be given and fixed, and let δ be such that $\delta \leq \alpha\rho(d)$, where $d \in \mathcal{F}$ and $\rho(d) > 0$. If $\hat{x}$ is the optimal solution to $(P_\mu(d))$, and $\bar{x} \in P_\mu^*(d,\delta)$, then*

$$f_2\left(\frac{1-\alpha}{\mathcal{K}_\mu(d)\|d\|}\right)^2 \leq \|(A\hat{X}\bar{X}A^T)^{-1}\| \leq g_2\left(\frac{\mathcal{C}(d)\mathcal{K}_\mu(d)}{\mu(1-\alpha)}\right)^2, \tag{14}$$

where $f_2 = \frac{m_3^2m_2^2}{4M_2^2M_3^2}$, $g_2 = \frac{32M_2^2}{m_0^2m_2^2m_3^2}$ and m_0, m_2, M_2, m_3, and M_3, are the constants defined in (4), (6), (7), (8), and (9), respectively.

Proof: Using Proposition 4.2 part (i), we have that $\|(A\hat{X}\bar{X}A^T)^{-1}\| \leq (1/m_2^2)\|(A\hat{X}\bar{X}A^T)^{-1}\|_2 \leq (1/m_2^2)(\min_{1\leq j\leq n}\{\hat{x}_j\bar{x}_j\})^{-1}\|(AA^T)^{-1}\|_2$. Now, by applying Proposition 4.2, part (ii), and Lemma 4.1, we obtain that

$$\|(A\hat{X}\bar{X}A^T)^{-1}\| \leq \frac{1}{m_2^2}\frac{\|d\|^2\mathcal{K}_\mu(d)^2}{f_1\mu^2(1-\alpha)^2}\frac{1}{\lambda_1(AA^T)} \leq \frac{1}{m_2^2}\frac{\|d\|^2\mathcal{K}_\mu(d)^2}{f_1\mu^2(1-\alpha)^2}\frac{M_2^2}{m_3^2\rho(d)^2} =$$

$$g_2\left(\frac{\mathcal{C}(d)\mathcal{K}_\mu(d)}{\mu(1-\alpha)}\right)^2.$$

On the other hand, by Proposition 4.2 part (i), $\|(A\hat{X}\bar{X}A^T)^{-1}\| \geq (1/M_2^2)\|(A\hat{X}\bar{X}A^T)^{-1}\|_2 \geq (1/M_2^2)(\max_{1\leq j\leq n}\{\hat{x}_j\bar{x}_j\})^{-1}\|(AA^T)^{-1}\|_2$. Now, by applying Proposition 4.2, part (ii), and Lemma 4.1, we obtain that

$$\|(A\hat{X}\bar{X}A^T)^{-1}\| \geq \frac{1}{M_2^2}\frac{(1-\alpha)^2}{h_1\mathcal{K}_\mu(d)^2}\frac{1}{\lambda_1(AA^T)} \geq \frac{1}{M_2^2}\frac{(1-\alpha)^2}{h_1\mathcal{K}_\mu(d)^2}\frac{1}{\lambda_m(AA^T)} =$$

$$\frac{1}{M_2^2}\frac{(1-\alpha)^2}{h_1\mathcal{K}_\mu(d)^2}\frac{1}{\|A\|_2^2} \geq \frac{1}{M_2^2}\frac{(1-\alpha)^2}{h_1\mathcal{K}_\mu(d)^2}\frac{m_2^2}{M_3^2}\frac{1}{\|A\|^2} \geq f_2\left(\frac{1-\alpha}{\mathcal{K}_\mu(d)\|d\|}\right)^2,$$

where $\lambda_m(AA^T)$ is the largest eigenvalue of AA^T.
q.e.d.

The next corollary is important in that it establishes lower and upper bounds on the operator norm of the matrix $(A\hat{X}^2A^T)^{-1}$, which is of central importance in interior point algorithms for linear programming that use Newton's method. Notice that the bounds in the corollary only depend on the condition number $\mathcal{C}(d)$, the distance to ill-posedness $\rho(d)$, the size of the data instance $d = (A, b, c)$, the barrier parameter μ, and certain constants. Also note that as $\mu \to 0$, the upper bound on $\|(A\hat{X}^2A^T)^{-1}\|$ in the corollary goes to ∞ quadratically in $1/\mu$ in the limit. Incidentally, the matrix $(A\hat{X}^2A^T)^{-1}$ differs from the inverse of the Hessian of the dual objective function at its optimum by the scalar $-\mu^2$.

Corollary 4.2 *Let $d = (A, b, c)$ be a data instance in $\mathcal{F}$ such that $\rho(d) > 0$. Let $\hat{x}$ and $\bar{x}$ be the optimal solutions of $(P_\mu(d))$ and $(P_{\bar{\mu}}(d))$, respectively, where $\mu, \bar{\mu} > 0$. Then*

$$4f_2 \frac{1}{\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\|d\|^2} \leq \|(A\hat{X}\bar{X}A^T)^{-1}\| \leq \frac{g_2}{8} \frac{\mathcal{C}(d)^2\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)}{\mu\bar{\mu}},$$

where f_2 and g_2 are the constants defined in Lemma 4.2. In particular, when $\mu = \bar{\mu}$ we have:

$$4f_2 \left(\frac{1}{\mathcal{K}_\mu(d)\|d\|}\right)^2 \leq \|(A\hat{X}^2A^T)^{-1}\| \leq \frac{g_2}{8} \left(\frac{\mathcal{C}(d)\mathcal{K}_\mu(d)}{\mu}\right)^2.$$

Proof: Following the proof of Lemma 4.2, we have from Proposition 4.2 and Corollary 4.1 that

$$\|(A\hat{X}\bar{X}A^T)^{-1}\| \leq \frac{1}{m_2^2} \frac{1}{(\min_{1\leq j\leq n}\{\hat{x}_j\bar{x}_j\})} \frac{M_2^2}{m_3^2\rho(d)^2} \leq$$

$$\frac{1}{m_2^2} \frac{\|d\|^2\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)}{8f_1\mu\bar{\mu}} \frac{M_2^2}{m_3^2\rho(d)^2} = \frac{g_2}{8} \frac{\mathcal{C}(d)^2\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)}{\mu\bar{\mu}}.$$

On the other hand, we have again from Proposition 4.2 and Corollary 4.1 that

$$\|(A\hat{X}\bar{X}A^T)^{-1}\| \geq \frac{\|(AA^T)^{-1}\|_2}{M_2^2\,(\max_{1\leq j\leq n}\{\hat{x}_j\bar{x}_j\})}$$

$$\geq \frac{4}{M_2^2 h_1 \mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\lambda_1(AA^T)}$$

$$\geq \frac{4}{M_2^2 h_1 \mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\lambda_m(AA^T)}$$

$$= \frac{4}{M_2^2 h_1 \mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\|A\|_2^2}$$

$$\geq \frac{4m_2^2}{M_2^2 M_3^2 h_1 \mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\|A\|^2}$$

$$= \frac{4f_2}{\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\|A\|^2}$$

$$\geq \frac{4f_2}{\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\|d\|^2},$$

where $\lambda_1(AA^T)$ and $\lambda_m(AA^T)$ are the smallest and largest eigenvalues of the matrix AA^T, respectively.
q.e.d.

We are now ready to state and prove the first theorem of this section, which presents upper bounds on changes in optimal solutions as the data is changed.

Theorem 4.1 *Let $\alpha \in (0,1)$ be given and fixed, and let δ be such that $\delta \leq \alpha\rho(d)$, where $d \in \mathcal{F}$ and $\rho(d) > 0$. If $\hat{x}$ is the optimal solution to $(P_\mu(d))$, and $(\hat{y}, \hat{s})$ is the optimal solution to $(D_\mu(d))$, then*

$$\|\bar{x} - \hat{x}\| \leq g_3\ \delta\ \frac{\mathcal{C}(d)^2\mathcal{K}_\mu(d)^5(\mu + \|d\|)}{\mu^2(1-\alpha)^6},$$

$$\|\bar{y} - \hat{y}\|_* \leq g_4\ \delta\ \frac{\mathcal{C}(d)^2\mathcal{K}_\mu(d)^5(\mu + \|d\|)}{\mu^2(1-\alpha)^6},$$

and

$$\|\bar{s} - \hat{s}\|_* \leq g_5\ \delta\ \frac{\mathcal{C}(d)^2\mathcal{K}_\mu(d)^5(\mu + \|d\|)^2}{\mu^2(1-\alpha)^6},$$

for all $\bar{x} \in P_\mu^*(d,\delta)$ and for all $(\bar{y}, \bar{s}) \in D_\mu^*(d,\delta)$, where $g_3 = 5M_3h_1\max\{M_3, \frac{g_2}{m_2}\}$, $g_4 = 5g_2\max\{1, M_4g_1\}$, $g_5 = 6\max\{1, g_4\}$, and h_1, g_1, and g_2 are the constants defined in Lemmas 4.1 and 4.2, respectively.

Before proving the theorem, we offer the following comments. Notice that the bounds are linear in δ which indicates that the central trajectory associated with d changes at most linearly and in direct proportion to perturbations in d as long as the perturbations are smaller than $\alpha\rho(d)$. Also, the bounds are polynomial in the condition number $\mathcal{C}(d)$ and the barrier parameter μ. Furthermore, notice that as $\mu \to 0$ these bounds diverge to ∞. This is because small perturbations in d can produce extreme changes in the limit of the central trajectory associated with d as $\mu \to 0$.

Proof of Theorem 4.1: Let $\hat{x}$ be the primal optimal solution to $(P_\mu(d))$ and let $\bar{x} \in P_\mu^*(d,\delta)$. Then from the Karush-Kuhn-Tucker optimality conditions we have that for some $\bar{d} = (\bar{A}, \bar{b}, \bar{c}) \in B(d, \delta)$:

$$\mu\hat{X}^{-1}e_n = \hat{s}, \mu\bar{X}^{-1}e_n = \bar{s},$$
$$\hat{s} = c - A^T\hat{y}, \bar{s} = \bar{c} - \bar{A}^T\bar{y},$$
$$A\hat{x} = b, \bar{A}\bar{x} = \bar{b},$$
$$\hat{x}, \bar{x} > 0,$$

where $\hat{y}, \bar{y} \in \Re^m$. Therefore,

$$\bar{x} - \hat{x} = \frac{1}{\mu}\hat{X}\bar{X}(\hat{s} - \bar{s}) = \frac{1}{\mu}\left(\hat{X}\bar{X}\left((c - A^T\hat{y}) - (\bar{c} - \bar{A}^T\bar{y})\right)\right) =$$
$$\frac{1}{\mu}\left(\hat{X}\bar{X}\left(c - \bar{c} + (\bar{A} - A)^T\bar{y}\right) + \hat{X}\bar{X}A^T(\bar{y} - \hat{y})\right). \tag{15}$$

On the other hand, $A(\bar{x} - \hat{x}) = \bar{b} - b - (\bar{A} - A)\bar{x}$. Since A has rank m (otherwise $\rho(d) = 0$), then $P = A\hat{X}\bar{X}A^T$ is a positive definite matrix. By combining these statements together with (15), we obtain

$$\bar{b} - b - (\bar{A} - A)\bar{x} = \frac{1}{\mu}A\hat{X}\bar{X}\left(c - \bar{c} + (\bar{A} - A)^T\bar{y}\right) + \frac{1}{\mu}P(\bar{y} - \hat{y}),$$

$$\mu P^{-1}\left(\bar{b} - b - (\bar{A} - A)\bar{x}\right) = P^{-1}A\hat{X}\bar{X}\left(c - \bar{c} + (\bar{A} - A)^T\bar{y}\right) + \bar{y} - \hat{y},$$

and so

$$\bar{y} - \hat{y} = \mu P^{-1}\left(\bar{b} - b - (\bar{A} - A)\bar{x}\right) - P^{-1}A\hat{X}\bar{X}\left(c - \bar{c} + (\bar{A} - A)^T\bar{y}\right). \tag{16}$$

Therefore, we obtain

$$\|\bar{y}-\hat{y}\|_* \leq \|P^{-1}\| \left(\mu\|\bar{b}-b-(\bar{A}-A)\bar{x}\| + \|A\|\|\hat{X}\bar{X}\left(c-\bar{c}+(\bar{A}-A)^T\bar{y}\right)\|\right)$$

$$\leq \|P^{-1}\| \left(\mu\|\bar{b}-b-(\bar{A}-A)\bar{x}\| + M_4\|A\|\|\hat{X}\bar{X}\left(c-\bar{c}+(\bar{A}-A)^T\bar{y}\right)\|_*\right)$$

using Proposition 4.1. From Corollary 3.1, we have that

$$\|\bar{b}-b-(\bar{A}-A)\bar{x}\| \leq \delta(1+\|\bar{x}\|) \leq \delta\left(1+\frac{4}{(1-\alpha)^2}\mathcal{K}_\mu(d)\right) \leq \frac{5\delta}{(1-\alpha)^2}\mathcal{K}_\mu(d), \tag{17}$$

$$\|c-\bar{c}+(\bar{A}-A)^T\bar{y}\|_* \leq \delta(1+\|\bar{y}\|_*) \leq \delta\left(1+\frac{4}{(1-\alpha)^2}\mathcal{K}_\mu(d)\right) \leq \frac{5\delta}{(1-\alpha)^2}\mathcal{K}_\mu(d). \tag{18}$$

Therefore, by combining (13), (14), (17), and (18), we obtain the following bound on $\|\bar{y}-\hat{y}\|_*$:

$$\|\bar{y}-\hat{y}\|_* \leq g_2\left(\frac{\mathcal{C}(d)\mathcal{K}_\mu(d)}{\mu(1-\alpha)}\right)^2\left(\frac{5\delta}{(1-\alpha)^2}\mathcal{K}_\mu(d)\right)\left(\mu+M_4g_1\|d\|\left(\frac{\mathcal{K}_\mu(d)}{1-\alpha}\right)^2\right)$$

$$\leq 5g_2\max\{1, M_4g_1\}\ \delta\ \frac{\mathcal{C}(d)^2\mathcal{K}_\mu(d)^5(\mu+\|d\|)}{\mu^2(1-\alpha)^6},$$

thereby demonstrating the bound for $\|\bar{y}-\hat{y}\|_*$.

Now, by substituting equation (16) into equation (15), we obtain that

$$\bar{x}-\hat{x} = \frac{1}{\mu}\hat{X}\bar{X}\left(I-A^TP^{-1}A\hat{X}\bar{X}\right)\left(c-\bar{c}+(\bar{A}-A)^T\bar{y}\right)+\hat{X}\bar{X}A^TP^{-1}\left(\bar{b}-b-(\bar{A}-A)\bar{x}\right)$$

$$= \frac{1}{\mu}D^{\frac{1}{2}}\left(I-D^{\frac{1}{2}}A^TP^{-1}AD^{\frac{1}{2}}\right)D^{\frac{1}{2}}\left(c-\bar{c}+(\bar{A}-A)^T\bar{y}\right)+DA^TP^{-1}\left(\bar{b}-b-(\bar{A}-A)\bar{x}\right),$$

where $D=\hat{X}\bar{X}$. Observe that the matrix $Q=I-D^{\frac{1}{2}}A^TP^{-1}AD^{\frac{1}{2}}$ is a projection matrix, and so $\|Qx\|_2 \leq \|x\|_2$ for all $x \in \Re^n$. Hence, from Proposition 4.1 part (iii), we obtain that

$$\|\bar{x}-\hat{x}\| \leq M_3\|\bar{x}-\hat{x}\|_2 \leq \frac{M_3}{\mu}\|D^{\frac{1}{2}}\|_2\|D^{\frac{1}{2}}\|_2\|c-\bar{c}+(\bar{A}-A)^T\bar{y}\|_2$$

$$+M_3\|D\|_2\|A^TP^{-1}\left(\bar{b}-b-(\bar{A}-A)\bar{x}\right)\|_2.$$

It follows from Proposition 4.1 parts (ii), (v) and (vi), Lemma 4.1, Lemma 4.2, and inequalities (17) and (18) that

$$\|\bar{x}-\hat{x}\| \le \frac{M_3^2 h_1}{\mu}\left(\frac{\mathcal{K}_\mu(d)}{1-\alpha}\right)^2 \frac{5\delta}{(1-\alpha)^2}\mathcal{K}_\mu(d)$$

$$+\frac{M_3 h_1 g_2}{m_2}\left(\frac{\mathcal{K}_\mu(d)}{1-\alpha}\right)^2 \|d\| \left(\frac{\mathcal{C}(d)\mathcal{K}_\mu(d)}{\mu(1-\alpha)}\right)^2 \frac{5\delta}{(1-\alpha)^2}\mathcal{K}_\mu(d),$$

from which we obtain the following bound:

$$\|\bar{x}-\hat{x}\| \le 5M_3h_1\max\{M_3, \frac{g_2}{m_2}\}\ \delta\ \frac{\mathcal{C}(d)^2\mathcal{K}_\mu(d)^5(\mu+\|d\|)}{\mu^2(1-\alpha)^6},$$

which thereby demonstrates the bound on $\|\bar{x}-\hat{x}\|$.

Finally, observe that $\bar{s}-\hat{s}=\bar{c}-c+(A-\bar{A})^T\bar{y}+A^T(\hat{y}-\bar{y})$, so that $\|\bar{s}-\hat{s}\|_* \le \|\bar{c}-c+(A-\bar{A})^T\bar{y}\|_* + \|A\|\|\hat{y}-\bar{y}\|_*$. Using our previous results, we obtain

$$\|\bar{s}-\hat{s}\|_* \le \frac{5\delta}{(1-\alpha)^2}\mathcal{K}_\mu(d) + \|d\|\left(g_4\ \delta\ \frac{\mathcal{C}(d)^2\mathcal{K}_\mu(d)^5(\mu+\|d\|)}{\mu^2(1-\alpha)^6}\right) \le$$

$$6\max\{1,g_4\}\ \delta\ \frac{\mathcal{C}(d)^2\mathcal{K}_\mu(d)^5(\mu+\|d\|)^2}{\mu^2(1-\alpha)^6},$$

and this concludes the proof of this theorem.

q.e.d.

The next theorem presents upper bounds on changes in optimal solutions as the barrier parameter μ is changed.

Theorem 4.2 *Let $d=(A,b,c)$ be a data instance in $\mathcal{F}$ such that $\rho(d)>0$. Let $\hat{x}$ and $\bar{x}$ be the optimal solutions of $(P_\mu(d))$ and $(P_{\bar{\mu}}(d))$, respectively, where $\mu,\bar{\mu}>0$. Let $(\hat{y},\hat{s})$ and $(\bar{y},\bar{s})$ be the optimal solutions of $(D_\mu(d))$ and $(D_{\bar{\mu}}(d))$, respectively. Then*

$$\|\bar{x}-\hat{x}\| \le g_6\ \frac{|\bar{\mu}-\mu|}{\mu\bar{\mu}}\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\|d\|,$$

$$\|\bar{y}-\hat{y}\|_* \le g_7\ \frac{|\bar{\mu}-\mu|}{\mu\bar{\mu}}\mathcal{C}(d)^2\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\|d\|,$$

and

$$\|\bar{s} - \hat{s}\|_* \leq g_7 \, \frac{|\bar{\mu} - \mu|}{\mu\bar{\mu}} \mathcal{C}(d)^2 \mathcal{K}_\mu(d) \mathcal{K}_{\bar{\mu}}(d) \|d\|^2,$$

where $g_6 = M_3^2 h_1/4$, $g_7 = g_2/8$, *and* h_1 *and* g_2 *are the constants defined in Lemma 4.1 and Corollary 4.2, respectively.*

Before proving the theorem, we offer the following comments. Notice that the bounds are linear in $|\bar{\mu} - \mu|$ which indicates that solutions along the central trajectory associated with d change at most linearly and in direct proportion to changes in μ. Also, the bounds are polynomial in the condition number $\mathcal{C}(d)$ and the barrier parameter μ.

Proof of Theorem 4.2: From the Karush-Kuhn-Tucker optimality conditions we have that

$$\begin{aligned} \mu \hat{X}^{-1} e_n &= \hat{s}, \bar{\mu} \bar{X}^{-1} e_n = \bar{s}, \\ \hat{s} &= c - A^T \hat{y}, \bar{s} = c - A^T \bar{y}, \\ A\hat{x} &= b, A\bar{x} = b, \\ \hat{x}, \bar{x} &> 0, \end{aligned}$$

where $\hat{y}, \bar{y} \in \Re^m$. Therefore,

$$\bar{x} - \hat{x} = \frac{1}{\mu\bar{\mu}} \hat{X}\bar{X} \left(\bar{\mu}\hat{s} - \mu\bar{s}\right) = \frac{1}{\mu\bar{\mu}} \hat{X}\bar{X} \left(\bar{\mu}(c - A^T\hat{y}) - \mu(c - A^T\bar{y})\right) =$$

$$\frac{1}{\mu\bar{\mu}} \hat{X}\bar{X} \left((\bar{\mu} - \mu)c - A^T(\bar{\mu}\hat{y} - \mu\bar{y})\right). \tag{19}$$

On the other hand, $A(\bar{x} - \hat{x}) = b - b = 0$. Since A has rank m (otherwise $\rho(d) = 0$), then $P = A\hat{X}\bar{X}A^T$ is a positive definite matrix. By combining these statements together with (19), we obtain

$$0 = \frac{1}{\mu\bar{\mu}} A\hat{X}\bar{X} \left((\bar{\mu} - \mu)c - A^T(\bar{\mu}\hat{y} - \mu\bar{y})\right),$$

and so

$$P\left(\bar{\mu}\hat{y} - \mu\bar{y}\right) = (\bar{\mu} - \mu) A\hat{X}\bar{X}c,$$

equivalently

$$\bar{\mu}\hat{y} - \mu\bar{y} = (\bar{\mu} - \mu) P^{-1} A\hat{X}\bar{X}c. \tag{20}$$

By substituting equation (20) into equation (19), we obtain:

$$\bar{x} - \hat{x} = \frac{\bar{\mu} - \mu}{\mu\bar{\mu}} \hat{X}\bar{X}\left(c - A^T P^{-1} A\hat{X}\bar{X}c\right) =$$

$$\frac{\bar{\mu} - \mu}{\mu\bar{\mu}} D\left(c - A^T P^{-1} ADc\right) =$$

$$\frac{\bar{\mu} - \mu}{\mu\bar{\mu}} D^{\frac{1}{2}}\left(I - D^{\frac{1}{2}} A^T P^{-1} AD^{\frac{1}{2}}\right) D^{\frac{1}{2}} c,$$

where $D = \hat{X}\bar{X}$. Observe that the matrix $Q = I - D^{\frac{1}{2}} A^T P^{-1} AD^{\frac{1}{2}}$ is a projection matrix, and so $\|Qx\|_2 \leq \|x\|_2$ for all $x \in \Re^n$. Hence, from Proposition 4.1 parts (iii) and (v) and Corollary 4.1, we obtain:

$$\|\bar{x} - \hat{x}\| \leq M_3 \|\bar{x} - \hat{x}\|_2 \leq M_3 \frac{|\bar{\mu} - \mu|}{\mu\bar{\mu}} \|D^{\frac{1}{2}}\|_2 \|D^{\frac{1}{2}}\|_2 \|c\|_2 \leq$$

$$M_3^2 \frac{|\bar{\mu} - \mu|}{\mu\bar{\mu}} \frac{h_1}{4} \mathcal{K}_\mu(d) \mathcal{K}_{\bar{\mu}}(d) \|d\|,$$

which demonstrates the bound for $\|\bar{x} - \hat{x}\|$.

Now, since $c = A^T\hat{y} + \hat{s}$ and $c = A^T\bar{y} + \bar{s}$, it follows that

$$A^T(\bar{y} - \hat{y}) + \bar{s} - \hat{s} = 0,$$

which yields the following equalities in logical sequence:

$$A^T(\bar{y} - \hat{y}) + \hat{X}^{-1}\bar{X}^{-1}(\bar{\mu}\hat{x} - \mu\bar{x}) = 0,$$

$$A^T(\hat{y} - \bar{y}) = \hat{X}^{-1}\bar{X}^{-1}(\bar{\mu}\hat{x} - \mu\bar{x}),$$

$$\hat{X}\bar{X}A^T(\hat{y} - \bar{y}) = \bar{\mu}\hat{x} - \mu\bar{x},$$

so that by premultiplying by A, we obtain

$$A\hat{X}\bar{X}A^T(\hat{y} - \bar{y}) = (\bar{\mu} - \mu)b,$$

$$P(\hat{y} - \bar{y}) = (\bar{\mu} - \mu)b,$$

$$\hat{y} - \bar{y} = (\bar{\mu} - \mu)P^{-1}b.$$

Therefore, from Corollary 4.2,

$$\|\hat{y} - \bar{y}\|_* \leq |\bar{\mu} - \mu| \|P^{-1}\| \|b\| \leq |\bar{\mu} - \mu| \frac{g_2}{8} \frac{\mathcal{C}(d)^2 \mathcal{K}_\mu(d) \mathcal{K}_{\bar{\mu}}(d) \|d\|}{\mu \bar{\mu}},$$

which establishes the bound for $\|\hat{y} - \bar{y}\|_*$.

Finally, using that $\hat{s} - \bar{s} = A^T(\bar{y} - \hat{y})$, we obtain $\|\hat{s} - \bar{s}\|_* = \|A\| \|\hat{y} - \bar{y}\|_*$, and so this concludes the proof of this theorem.

q.e.d.

Using the same arguments as in Theorem 4.2 and the smoothness of the function $x(\mu) = \arg\min\{c^T x + \mu p(x) : Ax = b, x > 0\}$ for $\mu > 0$, it is straightforward to demonstrate that

$$\dot{x}(\mu) = \lim_{\bar{\mu} \to \mu} \frac{x(\bar{\mu}) - x(\mu)}{\bar{\mu} - \mu} = \frac{1}{\mu^2} X(\mu)^2 \left(c - A^T P^{-1} A X(\mu)^2 c \right),$$

where $P = A(X(\mu))^2 A^T$, and similarly

$$\dot{y}(\mu) = \lim_{\bar{\mu} \to \mu} \frac{y(\bar{\mu}) - y(\mu)}{\bar{\mu} - \mu} = \frac{1}{\mu} \left(y(\mu) - P^{-1} A X(\mu)^2 c \right),$$

$$\dot{s}(\mu) = \lim_{\bar{\mu} \to \mu} \frac{s(\bar{\mu}) - s(\mu)}{\bar{\mu} - \mu} = -A^T \dot{y}(\mu) = \frac{1}{\mu} \left(s(\mu) + A^T P^{-1} A X(\mu)^2 c - c \right).$$

These same results were previously derived by Adler and Monteiro in [1]. However, with the help of Theorem 4.2, we have the following upper bounds on these derivatives, whose proofs are immediate from the theorem:

Corollary 4.3 *Let $d = (A, b, c)$ be a data instance in $\mathcal{F}$ such that $\rho(d) > 0$. Let $x(\mu)$ and $(y(\mu), s(\mu))$ be the optimal solutions of $(P_\mu(d))$ and $(D_\mu(d))$, respectively, where $\mu > 0$. Then*

$$\|\dot{x}(\mu)\| \leq g_6 \frac{1}{\mu^2} \mathcal{K}_\mu(d)^2 \|d\|,$$

$$\|\dot{y}(\mu)\| \leq g_7 \frac{1}{\mu^2} \mathcal{C}(d)^2 \mathcal{K}_\mu(d)^2 \|d\|,$$

$$\|\dot{s}(\mu)\| \leq g_7 \frac{1}{\mu^2} \mathcal{C}(d)^2 \mathcal{K}_\mu(d)^2 \|d\|^2.$$

The next theorem establishes a relation between the condition number and changes in the optimal objective value of $(P_\mu(d))$ as the data (A, b, c) changes.

Theorem 4.3 *Suppose that $d = (A, b, c) \in \mathcal{F}$ and $\rho(d) > 0$. Let $\alpha \in (0, 1)$ be given and fixed, and let δ be such that $\delta \leq \alpha\rho(d)$, and let $\bar{d} = (\bar{A}, \bar{b}, \bar{c}) \in B(d, \delta)$. Define $z = \min\{c^T x + \mu p(x) : Ax = b, x > 0\}$ and define $\bar{z} = \min\{\bar{c}^T x + \mu p(x) : \bar{A}x = \bar{b}, x > 0\}$. Then*

$$| \bar{z} - z | \leq 3\delta \left(\frac{1+\alpha}{1-\alpha}\right)^4 K_\mu(d)^2.$$

Notice that the upper bound in this theorem is linear in δ so long as δ is no larger than $\alpha\rho(d)$, which indicates that optimal objective values along the central trajectory will change at most linearly and in direct proportion to changes in d for small changes in d. Note also that the bound is polynomial in the condition number $\mathcal{C}(d)$ and in the barrier parameter μ.

Proof of Theorem 4.3: Consider the Lagrangian functions associated with these problems,

$$\begin{aligned} L(x, y) &= c^T x + \mu p(x) + y^T(b - Ax), \\ \bar{L}(x, y) &= \bar{c}^T x + \mu p(x) + y^T(\bar{b} - \bar{A}x), \end{aligned}$$

and define $\Phi(x, y) = L(x, y) - \bar{L}(x, y)$. Observe that,

$$\begin{aligned} z &= \max_y \min_{x>0} L(x, y) &= \min_{x>0} \max_y L(x, y), \\ \bar{z} &= \max_y \min_{x>0} \bar{L}(x, y) &= \min_{x>0} \max_y \bar{L}(x, y). \end{aligned}$$

Hence, if $(\hat{x}, \hat{y})$ is a pair of optimal solutions to the primal and dual programs corresponding to (A, b, c), and $(\bar{x}, \bar{y})$ is a pair of optimal solutions to the primal and dual programs corresponding to $(\bar{A}, \bar{b}, \bar{c})$, then

$$\begin{aligned} z = L(\hat{x}, \hat{y}) &= \max_y \{L(\hat{x}, y)\} \\ &= \max_y \{\bar{L}(\hat{x}, y) + \Phi(\hat{x}, y)\} \\ &\geq \bar{L}(\hat{x}, \bar{y}) + \Phi(\hat{x}, \bar{y}) \\ &\geq \bar{z} + \Phi(\hat{x}, \bar{y}). \end{aligned}$$

Thus, $z - \bar{z} \geq \Phi(\hat{x}, \bar{y})$. Similarly, we can prove that $z - \bar{z} \leq \Phi(\bar{x}, \hat{y})$.

Therefore, we obtain the following bounds

$$\begin{aligned} | z - \bar{z} | &\leq | \Phi(\hat{x}, \bar{y}) |, \text{ or} \\ | z - \bar{z} | &\leq | \Phi(\bar{x}, \hat{y}) |. \end{aligned}$$

On the other hand, using Hölder's inequality and the bounds from Corollary 3.1 we have

$$\begin{aligned}
\mid \Phi(\hat{x},\bar{y}) \mid &= \mid (c-\bar{c})^T\hat{x} + \bar{y}^T(b-\bar{b}) - \bar{y}^T(A-\bar{A})\hat{x} \mid \\
&\leq \|c-\bar{c}\|_*\|\hat{x}\| + \|\bar{y}\|_*\|b-\bar{b}\| + \|\bar{y}\|_*\|(A-\bar{A})\hat{x}\| \\
&\leq \delta\|\hat{x}\| + \delta\|\bar{y}\|_* + \delta\|\bar{y}\|_*\|\hat{x}\| \\
&\leq 3\delta \left(\tfrac{1+\alpha}{1-\alpha}\right)^4 K_\mu(d)^2
\end{aligned}$$

Similarly, we can show that

$$\mid \Phi(\bar{x},\hat{y}) \mid \leq 3\delta \left(\frac{1+\alpha}{1-\alpha}\right)^4 K_\mu(d)^2,$$

and the result follows.
q.e.d.

The last theorem of this section establishes an upper bound on changes in the optimal objective function value of $(P_\mu(d))$ as μ changes.

Theorem 4.4 *Let* $z(\mu) = \min\{c^Tx + \mu p(x) : Ax = b, x > 0\}$ *where* $d = (A,b,c) \in \mathcal{F}$ *and* $\mu > 0$. *Then*

$$|z(\mu) - z(\bar{\mu})| \leq n\ |\mu - \bar{\mu}|\left(g_9 + \ln\left(\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)\right) + |\ln(\|d\|)| + \max\{|\ln(\mu)|, |\ln(\bar{\mu})|\}\right),$$

for all $\mu, \bar{\mu} > 0$, *where* $g_9 = \max\{|\ln(m_0/2)|, |\ln(m_3)|\}$.

Before proving the theorem, we offer the following comments. Notice that this upper bound is linear in $|\bar{\mu} - \mu|$ which indicates that optimal objective function values along the central trajectory associated with d change at most linearly and in direct proportion to changes in μ. Also, the bounds are logarithmic in the condition number $\mathcal{C}(d)$ and in the barrier parameter μ.

Proof of Theorem 4.4: For any $\mu > 0$ define $x(\mu) = \arg\min\{c^Tx + \mu p(x) : Ax = b, x > 0\}$ and $(y(\mu), s(\mu)) = \arg\max\{b^Ty - \mu p(s) : A^Ty + s = c, s > 0\}$. As in Theorem 4.3, for given $\mu, \bar{\mu} > 0$, consider the following Lagrangian functions: $L(x,y) = c^Tx + \mu p(x) + y^T(b - Ax)$ and $\bar{L}(x,y) = c^Tx + \bar{\mu}p(x) + y^T(b - Ax)$. Define $\Phi(x,y) = L(x,y) - \bar{L}(x,y) = (\mu - \bar{\mu})p(x)$.

By a similar argument as in the proof of Theorem 4.3, we have that $z(\mu) - z(\bar{\mu}) \geq \Phi(x(\mu), y(\bar{\mu}))$ and $z(\mu) - z(\bar{\mu}) \leq \Phi(x(\bar{\mu}), y(\mu))$. Therefore, we obtain the following bounds:

either $|z(\mu)-z(\bar{\mu})| \leq |\Phi(x(\mu),y(\bar{\mu}))| = |\mu-\bar{\mu}||p(x(\mu))|$, or $|z(\mu)-z(\bar{\mu})| \leq |\Phi(x(\bar{\mu}),y(\mu))| = |\mu-\bar{\mu}||p(x(\bar{\mu}))|$. In other words,

$$|z(\mu) - z(\bar{\mu})| \leq |\mu - \bar{\mu}| \max\{|p(x(\mu))|, |p(x(\bar{\mu}))|\}.$$

On the other hand, from Theorem 3.2 and Corollary 4.1, we have that

$$\frac{m_0}{2}\frac{\mu}{\|d\|\mathcal{K}_\mu(d)} \leq x_j(\mu) \leq \frac{1}{m_3}\mathcal{K}_\mu(d),$$

for all $j = 1, \ldots, n$. Hence,

$$n\left(\ln\left(\frac{m_0}{2}\right) + \ln\left(\frac{\mu}{\|d\|\mathcal{K}_\mu(d)}\right)\right) \leq -p(x(\mu)) \leq n\left(\ln\left(\frac{1}{m_3}\right) + \ln(\mathcal{K}_\mu(d))\right),$$

so that

$$|p(x(\mu))| \leq n \max\left\{|\ln\left(\frac{m_0}{2}\right)| + |\ln\left(\frac{\mu}{\|d\|\mathcal{K}_\mu(d)}\right)|, |\ln\left(\frac{1}{m_3}\right)| + |\ln(\mathcal{K}_\mu(d))|\right\} \leq$$

$$n\left(g_9 + \ln(\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)) + |\ln(\|d\|)| + \max\{|\ln(\mu)|, |\ln(\bar{\mu})|\}\right).$$

Similarly, using $\bar{\mu}$ instead of μ we also obtain

$$|p(x(\bar{\mu}))| \leq n\left(g_9 + \ln(\mathcal{K}_\mu(d)\mathcal{K}_{\bar{\mu}}(d)) + |\ln(\|d\|)| + \max\{|\ln(\mu)|, |\ln(\bar{\mu})|\}\right),$$

and the result follows.
q.e.d.

Remark 2 *Since $z(\mu) = c^T x(\mu) + p(x(\mu))$, it follows from the smoothness of $x(\mu)$ that $z(\mu)$ is also a smooth function. Furthermore, from Theorem 4.4 we have that*

$$|\dot{z}(\mu)| \leq 2n\ \left(g_9 + \ln(\mathcal{K}_\mu(d)) + |\ln(\|d\|)| + |\ln(\mu)|\right).$$

5 Bounds for analytic center problems

In this section, we study some elementary properties of primal and dual analytic center problems, that are used in the proof of Theorem 3.3, which is presented at the end of this section.

Given a data instance $d = (A, b, c)$ for a linear program, the primal analytic center problem, denoted $AP(d)$, is defined as:

$$(AP(d)): \ \min\{p(x) : Ax = b, x > 0\}.$$

Structurally, the program $(AP(d))$ is closely related to the central trajectory problem $(P_\mu(d))$, and was first extensively studied by Sonnevend, see [20] and [21]. In terms of data dependence, note that the program $(AP(d))$ does not depend on the data c. It is well known that $(AP(d))$ has a unique solution when its feasible region is bounded and non empty. We call this unique solution the (primal) analytic center.

Similarly, we define the dual analytic center problem, denoted $AD(d)$, as:

$$(AD(d)): \ \max\{-p(s) : s = c - A^T y, s > 0\}.$$

In terms of data dependence, the program $(AD(d))$ does not depend on the data b. The program $(AD(d))$ has a unique solution when its feasible region is bounded and non empty, and we call this unique solution the (dual) analytic center. Note in particular that the two programs $(AP(d))$ and $(AD(d))$ are *not* duals of each other. As we will show soon, the study of these problems is relevant to obtain certain results on the central trajectory problem.

We will now present some particular upper bounds on the norms of feasible solutions of the analytic center problems $(AP(d))$ and $(AD(d))$, that are similar in spirit to certain results of the previous sections on the central trajectory problems $(P_\mu(d))$ and $(D_\mu(d))$. In order to do so, we first introduce a bit more notation. Define the following data sets: $\mathcal{D}_P = \{(A, b) : A \in \Re^{mn}, b \in \Re^m\}$ and $\mathcal{D}_D = \{(A, c) : A \in \Re^{mn}, c \in \Re^n\}$. In a manner similar to the central trajectory problem, we define the following feasibility sets for analytic center problems:

$$\bar{\mathcal{F}}_P = \{(A, b) \in \mathcal{D}_P : \text{ there exists } (x, y) \text{ such that } Ax = b, , x > 0, \text{ and } A^T y < 0\},$$

$$\bar{\mathcal{F}}_D = \{(A, c) \in \mathcal{D}_D : \text{ there exists } (x, y) \text{ such that } A^T y < c, \text{ and } Ax = 0, x > 0\},$$

that is, $\bar{\mathcal{F}}_P$ consists of data instances d for which $(AP(d))$ is feasible and $\bar{\mathcal{F}}_D$ consists of data instances d for which $(AD(d))$ is feasible. It is also appropriate to introduce the corresponding sets of ill-posed data instances: $\bar{\mathcal{B}}_P = \ \text{cl}(\bar{\mathcal{F}}_P) \cap \ \text{cl}(\bar{\mathcal{F}}_P^C) = \partial \bar{\mathcal{F}}_P = \partial \bar{\mathcal{F}}_P^C$, and

$\bar{\mathcal{B}}_D = \text{cl}(\bar{\mathcal{F}}_D) \cap \text{cl}(\bar{\mathcal{F}}_D^C) = \partial\bar{\mathcal{F}}_D = \partial\bar{\mathcal{F}}_D^C$.

For the primal analytic center problem $AP(d)$, the distance to ill-posedness of a data instance $d = (A, b, c)$ is defined as $\bar{\rho}_P(d) = \inf\{\|(A, b) - (\bar{A}, \bar{b})\|_P : (\bar{A}, \bar{b}) \in \bar{\mathcal{B}}_P\}$. For the dual analytic center problem $AD(d)$, the distance to ill-posedness of a data instance $d = (A, b, c)$ is defined as $\bar{\rho}_D(d) = \inf\{\|(A, c) - (\bar{A}, \bar{c})\|_D : (\bar{A}, \bar{c}) \in \bar{\mathcal{B}}_D\}$, where $\|(A, b)\|_P = \max\{\|A\|, \|b\|\}$ and $\|(A, c)\|_D = \max\{\|A\|, \|c\|_*\}$. Likewise, the corresponding condition measures are $\bar{\mathcal{C}}_P(d) = \|(A, b)\|_P/\bar{\rho}_P(d)$ if $\bar{\rho}_P(d) > 0$ and $\bar{\mathcal{C}}_P(d) = \infty$ otherwise; $\bar{\mathcal{C}}_D(d) = \|(A, c)\|_D/\bar{\rho}_D(d)$ if $\bar{\rho}_D(d) > 0$ and $\bar{\mathcal{C}}_D(d) = \infty$ otherwise.

Proposition 5.1 *If $d = (A, b, c)$ is such that $(A, b) \in \bar{\mathcal{F}}_P$, then $\bar{\rho}_P(d) \leq \rho(d)$.*

Proof: Given any $\epsilon > 0$, consider $\delta = \bar{\rho}_P(d) - \epsilon$. If $\bar{d} = (\bar{A}, \bar{b}, \bar{c})$ is a data instance such that $\|\bar{d} - d\| \leq \delta$, then $\|(\bar{A}, \bar{b}) - (A, b)\|_P \leq \delta$. Hence, $(\bar{A}, \bar{b}) \in \bar{\mathcal{F}}_P$, so that the system $\bar{A}x = \bar{b}$, $x > 0$, $\bar{A}^T y < 0$ has a solution, and so the system $\bar{A}x = \bar{b}$, $x > 0$, $\bar{A}^T y < c$ also has a solution, that is, $\bar{d} \in \mathcal{F}$. Therefore, $\rho(d) \geq \delta = \bar{\rho}_P(d) - \epsilon$, and the result follows by letting $\epsilon \to 0$.
q.e.d.

The following two lemmas present upper bounds on the norms of all feasible solutions for primal and dual analytic center problems, respectively.

Lemma 5.1 *Let $d = (A, b, c)$ be such that $(A, b) \in \bar{\mathcal{F}}_P$ and $\bar{\rho}_P(d) > 0$. Then*

$$\|x\| \leq \bar{\mathcal{C}}_P(d)$$

for any feasible x of $(AP(d))$.

Proof: Let x be a feasible solution of $(AP(d))$. By Proposition 2.3, there is a vector $\bar{x}$ such that $\bar{x}^T x = \|x\|$ and $\|\bar{x}\|_* = 1$. Define $\Delta A = \frac{-b\bar{x}^T}{\|x\|}$ and $\bar{d} = (A + \Delta A, b, c)$. Then, $(A + \Delta A)x = 0$ and $x > 0$. Now, consider the program $(AP(\bar{d}))$ defined as $\min\{p(x) : (A + \Delta A)x = b, x > 0\}$. Because $(A + \Delta A)x = 0$, $x > 0$, has a solution, there cannot exist y for which $(A + \Delta A)^T y < 0$, and so $(A + \Delta A, b) \in \bar{\mathcal{F}}_P^C$, whereby $\bar{\rho}_P(d) \leq \|(A + \Delta A, b) - (A, b)\|_P$. On the other hand, $\|(A + \Delta A, b) - (A, b)\|_P \leq \frac{\|\bar{x}\|_* \|b\|}{\|x\|} \leq \frac{\|(A,b)\|_P}{\|x\|}$, so that $\|x\| \leq \|(A, b)\|_P/\bar{\rho}_P(d) = \bar{\mathcal{C}}_P(d)$.
q.e.d.

Lemma 5.2 *Let $d = (A, b, c)$ be such that $(A, c) \in \bar{\mathcal{F}}_D$ and $\bar{\rho}_D(d) > 0$. Then*

$$\|y\|_* \leq \bar{\mathcal{C}}_D(d),$$

$$\|s\|_* \leq 2\|(A, c)\|_D \bar{\mathcal{C}}_D(d),$$

for any feasible (y, s) of $(AD(d))$.

Proof: Let (y, s) be a feasible solution of $(AD(d))$. If $y = 0$, then $s = c$ and the bounds are trivially true, so that we assume $y \neq 0$. By Proposition 2.3, there is a vector $\bar{y}$ such that $\|y\|_* = \bar{y}^T y$ and $\|\bar{y}\| = 1$. Let $\Delta A = -\frac{\bar{y}c^T}{\|y\|_*}$ and $\bar{d} = (A + \Delta A, b, c)$. Hence, $(A + \Delta A)^T y = A^T y - c < 0$. Because $(A + \Delta A)^T y$ has a solution, there cannot exist x for which $(A + \Delta A)x = 0$ and $x > 0$, and so $(A + \Delta A, c) \in \bar{\mathcal{F}}_D^C$, whereby $\bar{\rho_D}(d) \leq \|(A + \Delta A, c) - (A, c)\|_D$. On the other hand, $\|(A + \Delta A, c) - (A, c)\|_D = \frac{\|c\|_*}{\|y\|_*} \leq \frac{\|(A,c)\|_D}{\|y\|_*}$, so that $\|y\|_* \leq \|(A, c)\|_D / \bar{\rho_D}(d) = \bar{\mathcal{C}}_D(d)$. The bound for $\|s\|_*$ can be easily derived using the fact that $\|s\|_* \leq \|c\|_* + \|A\|\|y\|_*$ and $\bar{\mathcal{C}}_D(d) \geq 1$.
q.e.d.

With the aid of Lemma 5.2, we are now in position to present the proof of Theorem 3.3.

Proof of Theorem 3.3: From Tucker's strict complementarity theorem (see Dantzig [3], p. 139), there exists a unique partition of the set $\{1, \ldots, n\}$ into subsets B and N, $B \cap N = \emptyset$ and $B \cup N = \{1, \ldots, n\}$, such that

1. $Au = 0$, $u \geq 0$ implies $u_N = 0$ and there exists $\hat{u}$ for which $A\hat{u} = 0$, $\hat{u}_B > 0$, and $\hat{u}_N = 0$,

2. $A^T y = v$, $v \geq 0$ implies $v_B = 0$ and there exists $(\hat{y}, \hat{v})$ for which $A^T \hat{y} = \hat{v}$, $\hat{v}_B = 0$, and $\hat{v}_N > 0$.

Consider the set $S = \{s_B \in \Re^{|B|} : s_B = c_B - A_B^T y \text{ for some } y \in \Re^m, s_B > 0\}$. Because $(P_\mu(d))$ has an optimal solution, S is non empty. Also, S is bounded. To see this, suppose instead that S is unbounded, in which case there exists $\tilde{y}$ such that $A_B^T \tilde{y} \geq 0$ and $A_B^T \tilde{y} \neq 0$. Then $A^T(\tilde{y} + \lambda\hat{y}) \geq 0$ for λ sufficiently large, whereby $A_B^T(\tilde{y} + \lambda\hat{y}) = 0$. This in turn implies that $A_B^T \tilde{y} = 0$, a contradiction.

Because S is non empty and bounded, $d_B = (A_B, b, c_B) \in \bar{\mathcal{F}}_D$. Therefore, by Lemma 5.2, for any $s_B \in S$, $\|s_B\|_* \leq 2\|(A_B, c_B)\|_D \bar{\mathcal{C}}_D(d_B)$, and in particular

$$\|s_B(\mu)\|_* \leq 2\|(A_B, c_B)\|_D \bar{\mathcal{C}}_D(d_B) \leq 2\|d\|\bar{\mathcal{C}}_D(d_B),$$

where $s(\mu)$ is the optimal solution of $(D_\mu(d))$. Hence, for any $j \in B$, $s_j(\mu) \leq \|s_B(\mu)\|_\infty \leq \|s_B(\mu)\|_*/m_0 \leq 2\|d\|\bar{\mathcal{C}}_D(d_B)/m_0$. Nevertheless, since $x_j(\mu)s_j(\mu) = \mu$, then $x_j(\mu) \geq \frac{\mu m_0}{2\|d\|\bar{\mathcal{C}}_D(d_B)}$ for $j \in B$.

Finally, by definition of the partition of $\{1, \ldots, n\}$ into B and N, $x_j(\mu)$ is bounded for all $j \in N$ and for all $\mu > 0$. This also ensures that B is unique.
q.e.d.

References

[1] Ilan Adler and Renato D. C. Monteiro. Limiting behavior of the affine scaling continuous trajectories for linear programming problems. *Mathematical Programming*, 50(1):29–51, 1991.

[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] George B. Dantzig. *Linear Programming and Extensions.* Princeton University Press, Princeton, New Jersey, 1963.

[4] D. Den Hertog, C. Roos, and T. Terlaky. On the classical logarithmic barrier function method for a class of smooth convex programming problems. *Journal of Optimization Theory and Applications*, 73(1):1–25, 1992.

[5] Sharon Filipowski. On the complexity of solving linear programs specified with approximate data and known to be feasible. Technical Report May, Cornell University, 1994.

[6] Robert M. Freund and Jorge R. Vera. Some characterizations and properties of the "distance to ill-posedness" in conic linear systems. Technical Report W.P. #3862-95-MSA, Sloan School of Management, Massachusetts Institute of Technology, 1995.

[7] Gene H. Golub and Charles F. Van Loan. *Matrix Computations.* The Johns Hopkins University Press, New York, second edition, 1989.

[8] C. C. Gonzaga. Path following methods for linear programming. *SIAM Review*, 34(2):167–227, 1992.

[9] L.G. Hačijan. A polynomial algorithm in linear programming. *Soviet Math Dokl.*, 20(1):191–194, 1979.

[10] B. Jansen, C. Roos, and Terlaky T. A short survey on ten years interior point methods. Technical Report 95-45, Delft University of Technology, 1995.

[11] David G. Luenberger. *Optimization by Vector Space Methods.* John Wiley & Sons, Inc, New York, 1968.

[12] O. L. Mangasarian. A stable theorem of the alternative: An extension of the Gordan theorem. *Linear Algebra and Its Applications*, 41:209–223, 1981.

[13] G.L. Nemhauser and L.A. Wolsey. *Integer and Combinatorial Optimization.* John Wiley & Sons, Inc, New York, 1988.

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

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

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

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

[18] Stephen M. Robinson. Stability theory for systems of inequalities. Part I: Linear systems. *SIAM Journal on Numerical Analysis*, 12(5):754–769, 1975.

[19] Stephen M. Robinson. Stability theory for systems of inequalities. Part II: Nonlinear systems. *SIAM Journal on Numerical Analysis*, 13(4):497–513, 1976.

[20] G. Sonnevend. An 'analytic' center for polyhedrons and new classes of global algorithms for linear (smooth, convex) optimization. Technical report, Dept. of Numerical Analysis, Institute of Mathematics, Eötvös University, 1088, Budapest, Muzeum Körut 6-8, 1985. Preprint.

[21] G. Sonnevend. A new method for solving a set of linear (convex) inequalities and its applications for identification and optimization. Technical report, Dept. of Numerical Analysis, Institute of Mathematics, Eötvös University, 1088, Budapest, Muzeum Körut 6-8, 1985. Preprint.

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

[23] Jorge R. Vera. Ill-posedness and the complexity of deciding existence of solutions to linear programs. Technical Report June, Cornell University, 1994.

[24] Jorge R. Vera. On the complexity of linear programming under finite precision arithmetic. Technical Report August, Cornell University, 1994.