http://doi.org/10.1007/BF02680548

"Condition Measures and Properties of the Central Trajectory of a Linear Program." Nunez, Manuel A., and Robert M. Freund. *Mathematical Programming* Vol. 83, No. 1 (1998): 1-28.

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

Manuel A. Nunez and Robert M. Freund

OR 316-96 (Revised) May 1997

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

Manuel A. Nunez and Robert M. Freund

March 25, 1996
Revised May 27, 1997

Abstract

Given a data instance $d = (A, b, c)$ of a linear program, we show that certain properties of solutions along the central trajectory of the linear program are inherently related to the condition number $\mathcal{C}(d)$ of the data instance $d = (A, b, c)$, where $\mathcal{C}(d)$ is a scale-invariant reciprocal of a closely-related measure $\rho(d)$ called the "distance to ill-posedness." (The distance to ill-posedness essentially measures how close the data instance $d = (A, b, c)$ is to being primal or dual infeasible.) 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 the barrier parameter μ or the data $d = (A, b, c)$ of the linear program is changed. These bounds are all linear or polynomial functions of certain natural parameters associated with the linear program, namely the condition number $\mathcal{C}(d)$, the distance to ill-posedness $\rho(d)$, the norm of the data $\|d\|$, and the dimensions m and n.

1 Introduction

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:

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

where for $u > 0$ in $\Re^n$, $p(u) = -\sum_{j=1}^{n} \ln(u_j)$ is the logarithmic barrier function, $d = (A, b, c)$ is a data instance in the space of all data $\mathcal{D} = \{(A, b, c) : A \in \Re^{m\times n}, b \in \Re^m, c \in \Re^n\}$, and the parameter μ is a positive scalar considered independent of the data instance $d = (A, b, c) \in \mathcal{D}$. 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 [12] and Jansen et al. [13], and the book by Wright [32]. It is well known that programs $P_\mu(d)$ and $D_\mu(d)$ are related through Lagrangian duality; if each 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 [19, 20, 21, 22]. In the context of the central trajectory problem, $\rho(d)$ essentially is the minimum change $\Delta d = (\Delta A, \Delta b, \Delta c)$ in the data $d = (A, b, c)$ necessary to create a data instance $d + \Delta d$ that is an infeasible instance of $P_\mu(\cdot)$ or $D_\mu(\cdot)$. The condition number of the data instance $d = (A, b, c)$, denoted $\mathcal{C}(d)$, is defined to be $\mathcal{C}(d) := \|d\|/\rho(d)$ and is a scale-invariant reciprocal of the distance to ill-posedness $\rho(d)$, so that $\mathcal{C}(d)$ goes to ∞ as the data instance d approaches infeasibility.

The main results in the paper are stated in Sections 3 and 4. In Section 3 we present upper and lower bounds on sizes of optimal solutions to the barrier problems $P_\mu(d)$ and $D_\mu(d)$ in terms of the conditioning of the data instance d. Theorems 3.1 and 3.2 state bounds on such solutions that are linear in μ, where the constants in the bounds are polynomial functions of the condition number $\mathcal{C}(d)$, the distance to ill-posedness $\rho(d)$, the dimension n, the norm of the data $\|d\|$, or their inverses. These theorems show in particular that as μ goes to zero, that $x_j(\mu)$ grows at least linearly in μ; and as μ goes to ∞, $x_j(\mu)$ grows at most linearly in μ. Moreover, in Theorem 3.3, we also show that when the feasible region of $P_\mu(d)$ is unbounded, then certain coordinates of $x(\mu)$ grow exactly linearly in μ as $\mu \rightarrow \infty$, all at rates bounded by polynomial functions of the condition number $\mathcal{C}(d)$, the distance to ill-posedness $\rho(d)$, the dimension n, the norm of the data $\|d\|$, or their inverses.

In Section 4, we study the sensitivity of the optimal solutions to $P_\mu(d)$ and $D_\mu(d)$ as either the data $d = (A, b, c)$ changes or the barrier parameter μ changes. Theorems 4.1 and 4.4 state upper bounds on the sizes of the changes on optimal solutions as well as in the optimal objective values as the data $d = (A, b, c)$ is changed. Theorems 4.3 and 4.5 state upper bounds on the sizes of changes in optimal solutions and optimal objective values as the barrier parameter μ is changed. Along the way, we prove Theorem 4.2, which states bounds on the norm of the matrix $(AX^2(\mu)A^T)^{-1}$. This matrix is the main computational matrix in interior-point central trajectory methods. All of the bounds in this section are polynomial functions of the condition number $\mathcal{C}(d)$, the distance to ill-posedness $\rho(d)$, the dimension n, the norm of the data $\|d\|$, or their inverses.

Literature review. 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 [19] studied 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, $A : \mathcal{X} \rightarrow \mathcal{Y}$ is a continuous linear operator, $c^* : \mathcal{X} \rightarrow \Re$ 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 [16, 23, 24, 25]). In [20] and [21] Renegar introduced the concept of a *fully efficient* algorithm; and provided a fully-efficient algorithm that given any data instance d answers whether the program RLP associated with d is consistent or not.

Vera in [30] developed 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, the problem RLP is $\min\{c^Tx : Ax \leq b, x \in \Re^n\}$. In [28], Vera established bounds similar to those in [19] for norms of optimal primal and dual solutions and optimal objective function values. He then used these bounds to develop an algorithm for finding approximate optimal solutions of the original instance. In [29] he provided a measure of the precision of a logarithmic barrier algorithm based upon the distance to ill-posedness of the instance. To do this, he followed the same arguments as Den Hertog, Roos, and Terlaky [5], making the appropriate changes when necessary to express their results in terms of the distance to ill-posedness.

Filipowski [6, 7, 8] expanded upon Vera's results under the assumption that it is known beforehand that the primal data instance is feasible. In addition, she developed several fully-efficient algorithms that approximate optimal solutions to the original instance under this assumption.

Freund and Vera [9] addressed the issue of deciding feasibility of RLP. The problem that they studied 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 developed optimization problems that allow one to compute exactly or at least estimate the distance to ill-posedness. They also showed additional results relating the distance to ill-posedness to the existence of certain inscribed and circumscribed balls for the feasible region, which has implications for Khachiyan's ellipsoid algorithm [14].

Organization of the paper. This paper is organized as follows. In Section 2 we formally review the concept of *ill-posed* data instances, the distance to ill-posedness $\rho(d)$, and the condition number $\mathcal{C}(d)$. In this section we also discuss the notational conventions and present a few preliminary results that are used throughout the paper.

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 $d = (A, b, c)$ and in the barrier parameter μ. Theorems 4.1 and 4.3 state upper bounds on changes in optimal solutions and objective values along the central trajectory as the data instance d is changed to a "nearby" data instance $d + \Delta d$. Theorems 4.4 and 4.5 state upper bounds on changes in optimal solutions and objective values along the central trajectory as the barrier parameter μ is changed. Theorem 4.2 states upper and lower bounds on the norm of the matrix $(AX^2(\mu)A^T)^{-1}$. Corollary 4.1 states upper bounds on the first derivatives $\dot{x}(\mu)$ and $(\dot{y}(\mu), \dot{s}(\mu))$ of optimal solutions along the central trajectory with respect to the barrier parameter μ.

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 3.

2 Notation, Definitions, and Preliminaries

We denote by $\mathcal{D}$ the space of data instances, that is, $\mathcal{D} = \{(A, b, c) : A \in \Re^{m \times n}, b \in \Re^m, c \in \Re^n\}$, where $m \leq n$. The data for the programs $P_\mu(d)$ and $D_\mu(d)$ is the array $d = (A, b, c) \in \mathcal{D}$. As observed in the Introduction, the positive scalar μ is treated as a parameter independent of the data $d = (A, b, c)$. Given a subset of data instances $S \subset \mathcal{D}$, we denote by $cl(S)$ the closure of S, by $int(S)$ the interior of S, and by ∂S the boundary of S.

Consider the following subset of the data set $\mathcal{D}$:

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

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

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

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, due to the fact that arbitrarily small changes in the data d can 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 $\mathcal{D}$. To do so we define the following norms on the space $\Re^k$:

$$\begin{aligned} \|v\|_\alpha &= \left(\sum_{i=1}^{k} |v_i|^\alpha\right)^{1/\alpha}, \\ \|v\|_\infty &= \max_{1 \leq i \leq k} |v_i|, \end{aligned}$$

for each $v \in \Re^k$, where $1 \leq \alpha < \infty$, and where $k = m$ or $k = n$. When computing the norm of a given vector using one of these norms, we do not explicitly make the distinction between the spaces $\Re^m$ and $\Re^n$ because the dimension will always be clear from the context. Given an $m \times n$-matrix A, we define the norm of A to be the operator norm:

$$\|A\|_{\alpha,\beta} = \max\left\{\|Ax\|_\beta : x \in \Re^n, \|x\|_\alpha \leq 1\right\},$$

where $1 \leq \alpha, \beta \leq \infty$. In particular, we omit the subscripts when $\alpha = \beta = 1$, that is,

$$\|A\| := \|A\|_{1,1}.$$

It follows that $\|A\| = \max_{1 \leq j \leq n} \sum_{i=1}^{m} |A_{ij}|$. Furthermore, it follows that $\|A^T\|_{\infty,\infty} = \|A\|_{1,1} = \|A\|$. Finally, let $\|A\|_2$ denote the norm defined by:

$$\|A\|_2 := \|A\|_{2,2}.$$

Observe that if $A = uv^T$, where $u \in \Re^m$ and $v \in \Re^n$, then $\|A\| = \|u\|_1 \|v\|_\infty$.

The following proposition states well known bounds among the norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$.

Proposition 2.1 *The following inequalities hold:*

(i) $\|v\|_2 \leq \|v\|_1 \leq \sqrt{k}\|v\|_2$ *for any* $v \in \Re^k$.

(ii) $\|v\|_\infty \leq \|v\|_1 \leq k\|v\|_\infty$ *for any* $v \in \Re^k$.

(iii) $(1/\sqrt{k})\|v\|_2 \leq \|v\|_\infty \leq \|v\|_2$ *for any* $v \in \Re^k$.

(iv) $(1/\sqrt{n})\|A\|_2 \leq \|A\| \leq \sqrt{m}\|A\|_2$ *for any* $A \in \Re^{m \times n}$.

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

$$\|d\| = \max\left\{\|A\|, \|b\|_1, \|c\|_\infty\right\}.$$

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

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

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

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

see [19, 20, 21, 22], and so $\rho(d)$ is the distance of the data instance $d = (A, b, c)$ to the set of ill-posed instances $\mathcal{B}$. Observe that under the particular choice of norms used to define the norm on $\mathcal{D}$, the distance to ill-posedness $\rho(d)$ can be computed in polynomial time whenever d is rational (see Remark 3.1 of [9]).

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. Moreover, for $d = (A, b, c)$, let $\Delta d = (-A, -b, -c)$ and observe that since $d + \Delta d = (0, 0, 0) \in \mathcal{B}$ and $\mathcal{B}$ is a closed set, then for any $d \notin \mathcal{B}$ we have $\|d\| = \|\Delta d\| \geq \rho(d) > 0$, 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 interior of $\mathcal{F}$, $int(\mathcal{F})$, is characterized in the following lemma. For a proof of this lemma see Robinson [25] or Ashmanov [2].

Lemma 2.1

$$int(\mathcal{F}) = \{d = (A, b, c) : d \in \mathcal{F} \text{ and } rank(A) = m\}.$$

Observe that given a data instance $d \in \mathcal{F}$ and $\rho(d) > 0$ (so that $d \in int(\mathcal{F})$), it follows from the strict convexity of the logarithmic barrier and the full rank of A that the programs $P_\mu(d)$ and $D_\mu(d)$ will each have a unique optimal solution.

Furthermore, we state two elementary propositions that are well known variants of classical "theorems of the alternative" for linear inequality systems, see Gale [10], and are stated in the context of the central trajectory problems studied here.

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

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

Proposition 2.3 *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* $x \neq 0$.

Finally, we introduce the following notational convention which is standard in the field of interior point methods: if $x \in \Re^n$, then $X = \mathrm{diag}(x_1, ..., x_n)$. Moreover, we denote by e a vector of ones whose dimension depends on the context of the expression where this vector appears, so that no confusion should arise.

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 developed in the previous section, $d = (A, b, c)$ represents a data instance. Before presenting the first bound, we define the following scalar quantity, denoted $\mathcal{K}(d, \mu)$, which appears in many of the results of this section as well as in Section 4:

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

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*

$$\begin{aligned} \|x(\mu)\|_1 &\leq \mathcal{K}(d, \mu), & (3)\\ \|y(\mu)\|_\infty &\leq \mathcal{K}(d, \mu), & (4)\\ \|s(\mu)\|_\infty &\leq 2\|d\|\mathcal{K}(d, \mu), & (5) \end{aligned}$$

for the optimal solution $x(\mu)$ to $P_\mu(d)$ and the optimal solution $(y(\mu), s(\mu))$ to the dual problem $D_\mu(d)$, where $\mathcal{K}(d, \mu)$ is the scalar defined in (2).

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 $n/\rho(d)$, as μ goes to ∞.

Proof of Theorem 3.1: Let $\hat{x} = x(\mu)$ be the optimal solution to $P_\mu(d)$ and $(\hat{y}, \hat{s}) = (y(\mu), s(\mu))$ be the 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 $c^T\hat{x} = b^T\hat{y} + \mu n$.

Observe that since $\hat{s} = c - A^T\hat{y}$, then $\|\hat{s}\|_\infty \leq \|c\|_\infty + \|A^T\|_{\infty,\infty}\|\hat{y}\|_\infty$. Since $\|A^T\|_{\infty,\infty} = \|A\|$, we have that $\|\hat{s}\|_\infty \leq \|d\|(1 + \|\hat{y}\|_\infty)$, and using the fact that $\mathcal{C}(d) \geq 1$ the bound (5) on $\|\hat{s}\|_\infty$ is a consequence of the bound (4) on $\|\hat{y}\|_\infty$. It therefore is sufficient to prove the bounds on $\|\hat{x}\|_1$ and on $\|\hat{y}\|_\infty$. In addition, the bound on $\|\hat{y}\|_\infty$ is trivial if $\hat{y} = 0$, so from now on we assume that $\hat{y} \neq 0$. Also, let $\bar{y}$ be a vector in $\Re^m$ such that $\bar{y}^T\hat{y} = \|\hat{y}\|_\infty$ and $\|\bar{y}\|_1 = 1$.

The rest of the proof proceeds by examining three cases:

(i) $c^T\hat{x} \leq 0$,

(ii) $0 < c^T\hat{x} \leq \mu n$, and

(iii) $\mu n < c^T\hat{x}$.

In case (i), let $\Delta A = -be^T/\|\hat{x}\|_1$. Then $(A + \Delta A)\hat{x} = 0$, $\hat{x} > 0$, and $c^T\hat{x} \leq 0$. From Proposition 2.3, we have that $D_\mu(d + \Delta d)$ is infeasible, and so $\rho(d) \leq \|\Delta d\| = \|\Delta A\| = \|b\|_1/\|\hat{x}\|_1 \leq \|d\|/\|\hat{x}\|_1$. Therefore, $\|\hat{x}\|_1 \leq \|d\|/\rho(d) = \mathcal{C}(d) \leq \mathcal{K}(d, \mu)$, since $\mathcal{C}(d) \geq 1$ for any d. This proves (3) in this case.

Let $\theta = b^T\hat{y}$, $\Delta b = -\theta\bar{y}/\|\hat{y}\|_\infty$, $\Delta A = -\bar{y}c^T/\|\hat{y}\|_\infty$, and $d + \Delta d = (A + \Delta A, b + \Delta b, c)$. Observe that $(b + \Delta b)^T\hat{y} = 0$ and $(A + \Delta A)^T\hat{y} < 0$, so that $P_\mu(d + \Delta d)$ is infeasible from Proposition 2.2. Therefore, $\rho(d) \leq \|\Delta d\| = \max\{\|c\|_\infty, |\theta|\}/\|\hat{y}\|_\infty$. Hence, $\|\hat{y}\|_\infty \leq \max\{\mathcal{C}(d), |\theta|/\rho(d)\}$. Furthermore, $|\theta| = |b^T\hat{y}| = |c^T\hat{x} - \mu n| \leq \|\hat{x}\|_1\|c\|_\infty + \mu n \leq \mathcal{C}(d)\|d\| + \mu n$. Therefore, again using the fact that $\mathcal{C}(d) \geq 1$ for any d, we have (4).

In case (ii), let $d + \Delta d = (A + \Delta A, b, c + \Delta c)$, where $\Delta A = -be^T/\|\hat{x}\|_1$ and $\Delta c = -\mu ne/\|\hat{x}\|_1$. Observe that $(A + \Delta A)\hat{x} = 0$ and $(c + \Delta c)^T\hat{x} \leq 0$. From Proposition 2.3, $D_\mu(d + \Delta d)$ is infeasible, and so we conclude that $\rho(d) \leq \|\Delta d\| = \max\{\|\Delta A\|, \|\Delta c\|_\infty\} = \max\{\|b\|_1, \mu n\}/\|\hat{x}\|_1 \leq (\|d\| + \mu n)/\|\hat{x}\|_1$. Therefore, $\|\hat{x}\|_1 \leq \mathcal{C}(d) + \mu n/\rho(d) \leq \mathcal{K}(d, \mu)$. This proves (3) for this case.

Now, let $d + \Delta d = (A + \Delta A, b + \Delta b, c)$, where $\Delta A = -\bar{y}c^T/\|\hat{y}\|_\infty$ and $\Delta b = \mu n\bar{y}/\|\hat{y}\|_\infty$. Observe that $(b + \Delta b)^T\hat{y} = b^T\hat{y} + \mu n = c^T\hat{x} > 0$ and $(A + \Delta A)^T\hat{y} < 0$. Again, from Proposition 2.2, $P_\mu(d + \Delta d)$ is infeasible, and so we conclude that $\rho(d) \leq \|\Delta d\| = \max\{\|\Delta A\|, \|\Delta b\|_1\} = \max\{\|c\|_\infty, \mu n\}/\|\hat{y}\|_\infty \leq (\|d\| + \mu n)/\|\hat{y}\|_\infty$. Therefore, $\|\hat{y}\|_\infty \leq \mathcal{C}(d) + \mu n/\rho(d) \leq \mathcal{K}(d, \mu)$.

In case (iii), we first consider the bound on $\|\hat{y}\|_\infty$. Let $d + \Delta d = (A + \Delta A, b, c)$, where $\Delta A = -\bar{y}c^T/\|\hat{y}\|_\infty$. Since $(A + \Delta A)^T\hat{y} < 0$ and $b^T\hat{y} = c^T\hat{x} - \mu n > 0$, it follows from Proposition 2.2 that $P_\mu(d + \Delta d)$ is infeasible and so, $\rho(d) \leq \|\Delta d\| = \|c\|_\infty/\|\hat{y}\|_\infty$. Therefore, $\|\hat{y}\|_\infty \leq \mathcal{C}(d) \leq \mathcal{K}(d, \mu)$.

Finally, let $\Delta A = -be^T/\|\hat{x}\|_1$ and $\Delta c = -\theta e/\|\hat{x}\|_1$, where $\theta = c^T\hat{x}$. Observe that $(A + \Delta A)\hat{x} = 0$ and $(c + \Delta c)^T\hat{x} = 0$. Using Proposition 2.3, we conclude that $D_\mu(d + \Delta d)$ is infeasible and so, $\rho(d) \leq \|\Delta d\| = \max\{\|\Delta A\|, \|\Delta c\|_\infty\} = \max\{\|b\|_1, \theta\}/\|\hat{x}\|_1$, so that $\|\hat{x}\|_1 \leq \max\{\mathcal{C}(d), \theta/\rho(d)\}$. Furthermore, $\theta = c^T\hat{x} = b^T\hat{y} + \mu n \leq \|b\|_1\|\hat{y}\|_\infty + \mu n \leq \|d\|\mathcal{C}(d) + \mu n$. Therefore, $\|\hat{x}\|_1 \leq \mathcal{K}(d, \mu)$.
q.e.d.

Note that the scalar quantity $\mathcal{K}(d, \mu)$ appearing in Theorem 3.1 is scale invariant in the sense that $\mathcal{K}(\lambda d, \lambda\mu) = \mathcal{K}(d, \mu)$ for any $\lambda > 0$. From this it follows that the bounds in Theorem 3.1 on $\|x(\mu)\|_1$ and $\|y(\mu)\|_\infty$ are also scale invariant. However, as one would expect, the bound on $\|s(\mu)\|_\infty$ is not scale invariant, since $\|s(\mu)\|_\infty$ is sensitive to positive scalings of the data. Moreover, observe that as $\mu \to 0$ the bounds in Theorem 3.1 converge to the bounds presented by Vera in [28] for optimal solutions to linear programs of the form $\min\{c^Tx : Ax = b, x \geq 0\}$.

Examining the proof of Theorem 3.1, it is clear that the bounds stated in Theorem 3.1 will not generally be achieved. Indeed, implicit in the proof is the fact that bounds tighter than those in the theorem can be proved, and will depend on which of the three cases in the proof are applicable. However, our goal lies mainly in establishing bounds that are polynomial in the condition number $\mathcal{C}(d)$, the parameter μ, the size of the data $\|d\|$, and the dimensions m and n, and not necessarily in establishing the best achievable bounds.

We now present a simple example illustrating that the bounds in Theorem 3.1 are not necessarily tight. Let $m = 1$, $n = 2$, and

$$d = (A, b, c) = \left([1, 1], [1], \begin{bmatrix} -1 \\ -1 \end{bmatrix}\right).$$

For this data instance, we have that $\|d\| = 1$ and $\rho(d) = 1$, so that $\mathcal{C}(d) = 1$ and $\mathcal{K}(d, \mu) = 1 + n\mu$. Now observe that $x(\mu) = (1/2, 1/2)^T$ for all $\mu > 0$, so that $\|x(\mu)\|_1 = 1 < \mathcal{K}(d, \mu) = 1 + n\mu$ for all $\mu > 0$, which demonstrates that (3) is not tight in general. Furthermore, notice that in this example $c^Tx(\mu) < 0$, and so case (i) of the proof implies that $\|x(\mu)\|_1 \leq \mathcal{C}(d)$ (in fact, $\|x(\mu)\|_1 = \mathcal{C}(d) = 1$ in this example), which is a tighter bound than (3).

Corollary 3.1 *Let $\alpha \in (0,1)$ be given and fixed, and let δ be such that $\delta \le \alpha\rho(d)$, where $d \in \mathcal{F}$ and $\rho(d) > 0$. If $d + \Delta d \in \mathcal{D}$ is such that $\|\Delta d\| \le \delta$, then*

$$\begin{aligned}
\|x(\mu)\|_1 &\le \left(\frac{1+\alpha}{1-\alpha}\right)^2 \mathcal{K}(d,\mu), \\
\|y(\mu)\|_\infty &\le \left(\frac{1+\alpha}{1-\alpha}\right)^2 \mathcal{K}(d,\mu), \\
\|s(\mu)\|_\infty &\le 2(\|d\|+\delta)\left(\frac{1+\alpha}{1-\alpha}\right)^2 \mathcal{K}(d,\mu),
\end{aligned}$$

where $x(\mu)$ is the optimal solution to $P_\mu(d+\Delta d)$, $(y(\mu), s(\mu))$ is the optimal solution to $D_\mu(d+\Delta d)$, and $\mathcal{K}(d,\mu)$ is the scalar defined in (2).

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

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

since $\mathcal{C}(d) \ge 1$.
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 the optimal solutions $x(\mu)$ and $s(\mu)$ to the central trajectory problems $P_\mu(d)$ and $D_\mu(d)$, respectively.

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

$$\begin{aligned}
\|x(\mu)\|_1 &\ge \frac{\mu n}{2\|d\|\mathcal{K}(d,\mu)}, \\
\|s(\mu)\|_\infty &\ge \frac{\mu n}{\mathcal{K}(d,\mu)}, \\
x_j(\mu) &\ge \frac{\mu}{2\|d\|\mathcal{K}(d,\mu)}, \\
s_j(\mu) &\ge \frac{\mu}{\mathcal{K}(d,\mu)},
\end{aligned}$$

for all $j = 1, \ldots, n$, where $x(\mu)$ is the optimal solution to $P_\mu(d)$, $(y(\mu), s(\mu))$ is the optimal solution to $D_\mu(d)$, and $\mathcal{K}(d,\mu)$ is the scalar defined in (2).

This theorem shows that $\|x(\mu)\|_1$ and $x_j(\mu)$ are bounded from below by functions only involving the quantities $\|d\|$, $\mathcal{C}(d)$, $\rho(d)$, n, and μ. In addition, the theorem shows that for μ close to zero, that $x_j(\mu)$ grows at least linearly in μ, and at a rate that is at least $1/(2\|d\|\mathcal{C}(d)^2)$ (since $\mathcal{K}(d,\mu) = \mathcal{C}(d)^2 + \mu n/\rho(d) \approx \mathcal{C}(d)^2$ near $\mu = 0$). Furthermore, the theorem also shows that for μ close to zero, that $s_j(\mu)$ grows at least linearly in μ, and at a rate that is at least $1/\mathcal{C}(d)^2$.

The theorem offers less insight when $\mu \to \infty$, since the lower bound on $\|x(\mu)\|_1$ 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 $\|x(\mu)\| \to \infty$ 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. Similar remarks apply to the bound on $\|s(\mu)\|_\infty$ as $\mu \to \infty$.

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 $s(\mu)^T x(\mu) = \mu n$. Since $s(\mu)^T x(\mu) \leq \|s(\mu)\|_\infty \|x(\mu)\|_1$, it follows that $\|x(\mu)\|_1 \geq \mu n / \|s(\mu)\|_\infty$ and $\|s(\mu)\|_\infty \geq \mu n / \|x(\mu)\|_1$. Therefore, the first two inequalities follow from Theorem 3.1.

For the remaining inequalities, observe that for each $j = 1, \ldots, n$, $\mu = s_j(\mu) x_j(\mu)$, $x_j(\mu) \leq \|x(\mu)\|_1$, and $s_j(\mu) \leq \|s(\mu)\|_\infty$. 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 $d + \Delta d \in \mathcal{D}$ is such that $\|\Delta d\| \leq \delta$, then*

$$\begin{aligned} \|x(\mu)\|_1 &\geq \left(\frac{1-\alpha}{1+\alpha}\right)^2 \frac{\mu n}{2(\|d\| + \delta)\mathcal{K}(d,\mu)}, \\ \|s(\mu)\|_\infty &\geq \left(\frac{1-\alpha}{1+\alpha}\right)^2 \frac{\mu n}{\mathcal{K}(d,\mu)}, \end{aligned}$$

$$
\begin{aligned}
x_j(\mu) &\geq \left(\frac{1-\alpha}{1+\alpha}\right)^2 \frac{\mu}{2(\|d\|+\delta)\mathcal{K}(d,\mu)}, \\
s_j(\mu) &\geq \left(\frac{1-\alpha}{1+\alpha}\right)^2 \frac{\mu}{\mathcal{K}(d,\mu)},
\end{aligned}
$$

for all $j = 1, \ldots, n$, where $x(\mu)$ is the optimal solution to $P_\mu(d + \Delta d)$, $(y(\mu), s(\mu))$ is the optimal solution to $D_\mu(d + \Delta d)$, and $\mathcal{K}(d, \mu)$ is the scalar defined in (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 $x(\mu)$ 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, $\mathcal{C}_I(d_B)$ 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 *Let $x(\mu)$ denote the optimal solution to $P_\mu(d)$ and $(y(\mu), s(\mu))$ denote the optimal solution to $D_\mu(d)$. Then there exists a unique partition of the indices $\{1, \ldots, n\}$ into two subsets B and N such that*

$$
\begin{aligned}
x_j(\mu) &\geq \frac{\mu}{2\|d\|\mathcal{C}_I(d_B)}, \\
s_j(\mu) &\leq 2\|d\|\mathcal{C}_I(d_B),
\end{aligned}
$$

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 $\Re^{m \times |B| + m + |B|}$ composed of those elements of d indexed by the set B.

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 $1/(2\|d\|\mathcal{C}_I(d_B))$. 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 $\lim_{\mu\to\infty} \|x(\mu)\|_1 = \infty$. 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.

We end this section with the following remark concerning the scalar quantity $\mathcal{K}(d,\mu)$ defined in (2). Rather than using the quantity $\mathcal{K}(d,\mu)$, the results in this section could alternatively have been expressed in terms of the following scalar quantity:

$$\mathcal{R}(d,\mu) = \left(\frac{\max\{\|d\|, n\mu\}}{\min\{\rho(d), \mu\}}\right)^2. \tag{6}$$

One can think of the quantity $\mathcal{R}(d,\mu)$ as the square of the condition number of the data instance (A,b,c,μ) associated with the problem $P_\mu(d)$, where now $\mu > 0$ is considered as part of the data. The use of $\mathcal{R}(d,\mu)$ makes more sense intuitively relative to other results obtained in similar contexts (see for instance [28]). In this case, the norm on the data space would be defined as $\|(A,b,c,\mu)\| = \max\{\|A\|, \|b\|_1, \|c\|_\infty, n\mu\}$, and the corresponding distance to ill-posedness would be defined by $\rho(A,b,c,\mu) = \min\{\rho(d),\mu\}$. However, we prefer to use the scalar $\mathcal{K}(d,\mu)$ of (2), which arises more naturally in the proofs and conveniently leads to slightly tighter results, and also because it more accurately conveys the behavior of the optimal solutions to $P_\mu(d)$ as μ changes.

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 Theorems 4.1, 4.2, 4.3, 4.4, and 4.5. We first present all five theorems; the proofs of the theorems are deferred to the end of the section. As in the previous section, the bounds stated in these theorems are not necessarily the best achievable. Rather, it has been our goal to establish bounds that are polynomial in terms of the condition number $\mathcal{C}(d)$, the parameter μ, the size of the data $\|d\|$, and the dimensions m and n.

The first theorem, 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 $d + \Delta d = (A + \Delta A, b + \Delta b, c + \Delta c)$ in a suitably small neighborhood of the original data d.

Theorem 4.1 *Let $d = (A, b, c)$ be a data instance in $\mathcal{F}$ such that $\rho(d) > 0$, and let $\mu > 0$ be given and fixed. Given $\alpha \in (0, 1)$ fixed, let $\Delta d = (\Delta A, \Delta b, \Delta c) \in \mathcal{D}$ be such that $\|\Delta d\| \leq \alpha\rho(d)$. Then,*

$$\|\bar{x}(\mu) - x(\mu)\|_1 \leq \|\Delta d\| \frac{640n\, \mathcal{C}(d)^2 \mathcal{K}(d, \mu)^5 (\mu + \|d\|)}{\mu^2 (1-\alpha)^6}, \tag{7}$$

$$\|\bar{y}(\mu) - y(\mu)\|_\infty \leq \|\Delta d\| \frac{640m\, \mathcal{C}(d)^2 \mathcal{K}(d, \mu)^5 (\mu + \|d\|)}{\mu^2 (1-\alpha)^6}, \tag{8}$$

$$\|\bar{s}(\mu) - s(\mu)\|_\infty \leq \|\Delta d\| \frac{640m\, \mathcal{C}(d)^2 \mathcal{K}(d, \mu)^5 (\mu + \|d\|)^2}{\mu^2 (1-\alpha)^6}, \tag{9}$$

where $x(\mu)$ and $\bar{x}(\mu)$ are the optimal solutions to $P_\mu(d)$ and $P_\mu(d + \Delta d)$, respectively; $(y(\mu), s(\mu))$ and $(\bar{y}(\mu), \bar{s}(\mu))$ are the optimal solutions to $D_\mu(d)$ and $D_\mu(d+\Delta d)$, respectively; and $\mathcal{K}(d, \mu)$ is the scalar defined in (2).

Notice that the bounds are linear in $\|\Delta d\|$ 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 $\mu \to 0$.

The next theorem is important in that it establishes lower and upper bounds on the operator norm of the matrix $(AX^2(\mu)A^T)^{-1}$, where $x(\mu)$ is the optimal solution of $P_\mu(d)$. This is of central importance in interior point algorithms for linear programming that use Newton's method.

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

$$\frac{1}{mn}\left(\frac{1}{\mathcal{K}(d, \mu)\|d\|}\right)^2 \leq \|(AX^2(\mu)A^T)^{-1}\|_{1,\infty} \leq 4m\left(\frac{\mathcal{C}(d)\mathcal{K}(d, \mu)}{\mu}\right)^2,$$

where $\mathcal{K}(d, \mu)$ is the scalar defined in (2).

Notice that the bounds in the theorem 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 the dimensions m and n. Also note that as $\mu \rightarrow 0$, the upper bound on $\|(AX^2(\mu)A^T)^{-1}\|_{1,\infty}$ in the theorem goes to ∞ quadratically in $1/\mu$ in the limit. Incidentally, the matrix $(AX^2(\mu)A^T)^{-1}$ differs from the inverse of the Hessian of the dual objective function at its optimum by the scalar $-\mu^2$.

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

$$\begin{aligned} \|x(\bar{\mu}) - x(\mu)\|_1 &\leq \frac{n}{\mu\bar{\mu}} \, |\bar{\mu} - \mu| \, \mathcal{K}(d, \mu)\mathcal{K}(d, \bar{\mu})\|d\|, & (10) \\ \|y(\bar{\mu}) - y(\mu)\|_\infty &\leq \frac{4m}{\mu\bar{\mu}} \, |\bar{\mu} - \mu| \, \mathcal{K}(d, \mu)\mathcal{K}(d, \bar{\mu})\|d\|\mathcal{C}(d)^2, & (11) \\ \|s(\bar{\mu}) - s(\mu)\|_\infty &\leq \frac{4m}{\mu\bar{\mu}} \, |\bar{\mu} - \mu| \, \mathcal{K}(d, \mu)\mathcal{K}(d, \bar{\mu})\|d\|^2\mathcal{C}(d)^2, & (12) \end{aligned}$$

where $\mathcal{K}(d, \cdot)$ is the scalar defined in (2).

Notice that these 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.

The next result, Corollary 4.1, states upper bounds on the first derivatives of the optimal solutions $x(\mu)$ and $(y(\mu), s(\mu))$ of $P_\mu(d)$ and $D_\mu(d)$, respectively, with respect to the barrier parameter μ. We first define the derivatives along the central trajectory as follows:

$$\begin{aligned} \dot{x}(\mu) &= \lim_{\bar{\mu} \rightarrow \mu} \frac{x(\bar{\mu}) - x(\mu)}{\bar{\mu} - \mu}, \\ \dot{y}(\mu) &= \lim_{\bar{\mu} \rightarrow \mu} \frac{y(\bar{\mu}) - y(\mu)}{\bar{\mu} - \mu}, \\ \dot{s}(\mu) &= \lim_{\bar{\mu} \rightarrow \mu} \frac{s(\bar{\mu}) - s(\mu)}{\bar{\mu} - \mu}. \end{aligned}$$

See Adler and Monteiro [1] for the application of these derivatives to the limiting behavior of central trajectories in linear programming.

Corollary 4.1 *Let $d = (A, b, c)$ be a data instance in $\mathcal{F}$ such that $\rho(d) > 0$, and let $\mu > 0$ be given and fixed. Let $x(\mu)$ and $(y(\mu), s(\mu))$ be the optimal solutions of $P_\mu(d)$ and $D_\mu(d)$, respectively. Then*

$$\begin{aligned} \|\dot{x}(\mu)\|_1 &\leq \frac{n}{\mu^2}\,\mathcal{K}(d,\mu)^2\|d\|, \\ \|\dot{y}(\mu)\|_\infty &\leq \frac{4m}{\mu^2}\,\mathcal{K}(d,\mu)^2\|d\|\mathcal{C}(d)^2, \\ \|\dot{s}(\mu)\|_\infty &\leq \frac{4m}{\mu^2}\,\mathcal{K}(d,\mu)^2\|d\|^2\mathcal{C}(d)^2, \end{aligned}$$

where $\mathcal{K}(d, \mu)$ is the scalar defined in (2).

The proof of this corollary follows immediately from Theorem 4.3.

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 data d is changed to data $d + \Delta d$ in a specific neighborhood of the original data d. Before stating this theorem, we introduce the following notation. Let d be a data instance in $\mathcal{F}$, then we denote by $z(d)$ the corresponding optimal objective value associated with $P_\mu(d)$ by keeping the parameter μ fixed, that is,

$$z(d) = \min\{c^T x + \mu p(x) : Ax = b, x > 0\}.$$

Theorem 4.4 *Let $d = (A, b, c)$ be a data instance in $\mathcal{F}$ such that $\rho(d) > 0$, and let $\mu \geq 0$ be given and fixed. Given $\alpha \in (0, 1)$ fixed, let $\Delta d = (\Delta A, \Delta b, \Delta c) \in \mathcal{D}$ be such that $\|\Delta d\| \leq \alpha\rho(d)$. Then,*

$$|z(d + \Delta d) - z(d)| \leq 3\ \|\Delta d\| \left(\frac{1+\alpha}{1-\alpha}\right)^4 K(d,\mu)^2, \tag{13}$$

where $\mathcal{K}(d, \mu)$ is the scalar defined in (2).

Observe that, as in Theorem 4.1, the upper bound in the change in the objective function value is linear in $\|\Delta d\|$ so long as $\|\Delta d\|$ 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.

Last of all, Theorem 4.5 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. As before, it is convenient to introduce the following notation. Let d be a data instance in $\mathcal{F}$, then we denote by $z(\mu)$ the corresponding optimal objective value associated with $P_\mu(d)$ by keeping the data instance d fixed, that is,

$$z(\mu) = \min\{c^T x + \mu p(x) : Ax = b, x > 0\}.$$

Theorem 4.5 *Let $d = (A, b, c)$ be a data instance in $\mathcal{F}$ such that $\rho(d) > 0$. Then,*

$$|z(\bar{\mu}) - z(\mu)| \leq |\bar{\mu} - \mu|\, n \left(\ln(2) + \ln\left(\mathcal{K}(d,\mu)\mathcal{K}(d,\bar{\mu})\right) + |\ln(\|d\|)| + \max\{|\ln(\mu)|, |\ln(\bar{\mu})|\} \right),$$

for given $\mu, \bar{\mu} > 0$, where $\mathcal{K}(d, \cdot)$ is the scalar defined in (2).

As in Theorem 4.3, the upper bound given by this theorem 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.

Remark 1 *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, and from Theorem 4.5 it then follows that*

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

Before proving the five theorems, we first prove a variety of intermediary results that will be used in the proofs of the five theorems. The following proposition is a key proposition that relates the distance to ill-posedness of a data instance $d = (A, b, c)$ to the smallest eigenvalue of the matrix AA^T.

Proposition 4.1 *Let $d = (A, b, c) \in \mathcal{F}$ and $\rho(d) > 0$. Then*

(i) $(1/m)\|(AA^T)^{-1}\|_2 \leq \|(AA^T)^{-1}\|_{1,\infty} \leq \|(AA^T)^{-1}\|_2$,

and

(ii) $\rho(d) \leq \sqrt{m\lambda_1(AA^T)}$,

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

Proof: From Lemma 2.1, A has rank m, so that $(AA^T)^{-1}$ exists. The proof of (i) follows directly from Proposition 2.1, inequalities (i) and (iii). For the proof of (ii), let $\lambda_1 = \lambda_1(AA^T)$. There exists $v \in \Re^m$ with $\|v\|_2 = 1$ and $AA^Tv = \lambda_1 v$, so that $\|A^Tv\|_2^2 = v^TAA^Tv = \lambda_1$. Let $\Delta A = -vv^TA$, $\Delta b = \epsilon v$ for any $\epsilon > 0$ and small. Then, $(A + \Delta A)^T v = 0$ and $(b+\Delta b)^T v = b^Tv + \epsilon \neq 0$, for all $\epsilon > 0$ small. Hence, $(A+\Delta A)x = b + \Delta b$ is an inconsistent system of equations for all $\epsilon > 0$ and small. Therefore, by Proposition 2.1, inequality (iv), $\rho(d) \leq \max\{\|\Delta A\|, \|\Delta b\|_1\} = \|\Delta A\| \leq \sqrt{m}\|\Delta A\|_2 = \sqrt{m}\|A^Tv\|_2 = \sqrt{m\lambda_1}$, thus proving (ii).
q.e.d.

The next three results establish upper and lower bounds on certain quantities as the data $d = (A, b, c)$ is changed to data $d + \Delta d = (A + \Delta A, b + \Delta b, c + \Delta c)$ in a specific neighborhood of the original data d; or as the parameter μ is changed along the central trajectory. These results will also be used in the proofs of the theorems of this section.

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

$$\frac{1}{32}\left(\frac{\mu(1-\alpha)}{\|d\|\mathcal{K}(d,\mu)}\right)^2 \leq x_j(\mu)\bar{x}_j(\mu) \leq 4\left(\frac{\mathcal{K}(d,\mu)}{1-\alpha}\right)^2, \tag{14}$$

where $\mu > 0$ is given and fixed, and $\mathcal{K}(d, \mu)$ is the scalar defined in (2).

Proof: Let $x = x(\mu)$ and $\bar{x} = \bar{x}(\mu)$. From Theorem 3.1 we have that $\|x\|_1 \leq \mathcal{K}(d, \mu)$, and from Corollary 3.1 we also have that $\|\bar{x}\|_1 \leq (4/(1-\alpha)^2)\mathcal{K}(d, \mu)$. Therefore, we obtain $x_j\bar{x}_j \leq \|x\|_1\|\bar{x}\|_1 \leq 4(\mathcal{K}(d, \mu)^2/(1-\alpha)^2)$ for all $j = 1, \ldots, n$.

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

$$\begin{aligned}
x_j &\geq \frac{\mu}{2\|d\|\mathcal{K}(d,\mu)}, \\
\bar{x}_j &\geq \frac{(1-\alpha)^2\mu}{8(\|d\| + \|\Delta d\|)\mathcal{K}(d,\mu)} \\
&\geq \frac{(1-\alpha)^2\mu}{16\|d\|\mathcal{K}(d,\mu)},
\end{aligned}$$

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

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

for all $j = 1, \ldots, n$.
q.e.d.

Lemma 4.2 *Suppose that $d = (A, b, c) \in \mathcal{F}$ and $\rho(d) > 0$. Let $\alpha \in (0,1)$ be given and fixed, and let Δd be such that $\|\Delta d\| \leq \alpha \rho(d)$. If $x = x(\mu)$ is the optimal solution to $P_\mu(d)$, and $\bar{x} = \bar{x}(\mu)$ is the optimal solution to $P_\mu(d + \Delta d)$, then*

$$\frac{1}{4mn} \left(\frac{1-\alpha}{\mathcal{K}(d,\mu)\|d\|} \right)^2 \leq \|(AX\bar{X}A^T)^{-1}\|_{1,\infty} \leq 32m \left(\frac{\mathcal{C}(d)\mathcal{K}(d,\mu)}{\mu(1-\alpha)} \right)^2, \tag{15}$$

where $\mu > 0$ is given and fixed, and $\mathcal{K}(d, \mu)$ is the scalar defined in (2).

Proof: Using identical logic to Proposition 4.1 part (i), we have that

$$\begin{aligned} \|(AX\bar{X}A^T)^{-1}\|_{1,\infty} &\leq \|(AX\bar{X}A^T)^{-1}\|_2 \\ &\leq \frac{\|(AA^T)^{-1}\|_2}{\min_j\{x_j\bar{x}_j\}}. \end{aligned}$$

Now, by applying Proposition 4.1, part (ii), and Lemma 4.1, we obtain that

$$\begin{aligned} \|(AX\bar{X}A^T)^{-1}\|_{1,\infty} &\leq \frac{32\|d\|^2\mathcal{K}(d,\mu)^2}{\mu^2(1-\alpha)^2\lambda_1(AA^T)} \\ &\leq \frac{32m\|d\|^2\mathcal{K}(d,\mu)^2}{\mu^2(1-\alpha)^2\rho(d)^2} \\ &= \frac{32m\mathcal{C}(d)^2\mathcal{K}(d,\mu)^2}{\mu^2(1-\alpha)^2}. \end{aligned}$$

On the other hand, by identical logic to Proposition 4.1 part (i),

$$\begin{aligned} \|(AX\bar{X}A^T)^{-1}\|_{1,\infty} &\geq \frac{1}{m}\|(AX\bar{X}A^T)^{-1}\|_2 \\ &\geq \frac{\|(AA^T)^{-1}\|_2}{m\max_j\{x_j\bar{x}_j\}}. \end{aligned}$$

Now, by applying Proposition 2.1, part (iv), and Lemma 4.1, we obtain that

$$\begin{aligned}
\|(AX\bar{X}A^T)^{-1}\|_{1,\infty} &\geq \frac{(1-\alpha)^2}{4m\mathcal{K}(d,\mu)^2\lambda_1(AA^T)} \\
&\geq \frac{(1-\alpha)^2}{4m\mathcal{K}(d,\mu)^2\lambda_m(AA^T)} \\
&= \frac{(1-\alpha)^2}{4m\mathcal{K}(d,\mu)^2\|A\|_2^2} \\
&\geq \frac{(1-\alpha)^2}{4mn\mathcal{K}(d,\mu)^2\|A\|^2} \\
&\geq \frac{(1-\alpha)^2}{4mn\mathcal{K}(d,\mu)^2\|d\|^2},
\end{aligned}$$

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

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

$$\frac{1}{mn\mathcal{K}(d,\mu)\mathcal{K}(d,\bar{\mu})\|d\|^2} \leq \|(AX\bar{X}A^T)^{-1}\|_{1,\infty} \leq \frac{4m\mathcal{C}(d)^2\mathcal{K}(d,\mu)\mathcal{K}(d,\bar{\mu})}{\mu\bar{\mu}},$$

where $\mathcal{K}(d,\cdot)$ is the scalar defined in (2).

Proof: Following the proof of Lemma 4.2, we have from Proposition 4.1 and Theorem 3.2 that

$$\begin{aligned}
\|(AX\bar{X}A^T)^{-1}\|_{1,\infty} &\leq \frac{m}{\min_j\{x_j\bar{x}_j\}\rho(d)^2} \\
&\leq \frac{4m\|d\|^2\mathcal{K}(d,\mu)\mathcal{K}(d,\bar{\mu})}{\mu\bar{\mu}\rho(d)^2} \\
&= \frac{4m\mathcal{C}(d)^2\mathcal{K}(d,\mu)\mathcal{K}(d,\bar{\mu})}{\mu\bar{\mu}}.
\end{aligned}$$

On the other hand, we have again from Proposition 2.1, Theorem 3.1, and Proposition 4.1 that

$$
\begin{aligned}
\|(AX\bar{X}A^T)^{-1}\|_{1,\infty} &\geq \frac{\|(AA^T)^{-1}\|_2}{m\max_j\{x_j\bar{x}_j\}} \\
&\geq \frac{1}{m\mathcal{K}(d,\mu)\mathcal{K}(d,\bar{\mu})\lambda_1(AA^T)} \\
&\geq \frac{1}{m\mathcal{K}(d,\mu)\mathcal{K}(d,\bar{\mu})\lambda_m(AA^T)} \\
&= \frac{1}{m\mathcal{K}(d,\mu)\mathcal{K}(d,\bar{\mu})\|A\|_2^2} \\
&\geq \frac{1}{mn\mathcal{K}(d,\mu)\mathcal{K}(d,\bar{\mu})\|A\|^2} \\
&\geq \frac{1}{mn\mathcal{K}(d,\mu)\mathcal{K}(d,\bar{\mu})\|d\|^2}.
\end{aligned}
$$

q.e.d.

Note that the proof of Theorem 4.2 follows as an immediate application of Lemma 4.3, by setting $\mu = \bar{\mu}$.

We are now ready to prove Theorem 4.1.

Proof of Theorem 4.1: Let $x = x(\mu)$ and $\bar{x} = \bar{x}(\mu)$ be the optimal solutions to $P_\mu(d)$ and $P_\mu(d + \Delta d)$, respectively; and let $(y, s) = (y(\mu), s(\mu))$ and $(\bar{y}, \bar{s}) = (\bar{y}(\mu), \bar{s}(\mu))$ be the optimal solutions to $D_\mu(d)$ and $D_\mu(d + \Delta d)$, respectively. Then from the Karush-Kuhn-Tucker optimality conditions we have that:

$$
\begin{aligned}
Xs &= \mu e, & \bar{X}\bar{s} &= \mu e, \\
A^Ty + s &= c, & (A + \Delta A)^T\bar{y} + \bar{s} &= c + \Delta c, \\
Ax &= b, & (A + \Delta A)\bar{x} &= b + \Delta b, \\
x &> 0, & \bar{x} &> 0.
\end{aligned}
$$

Therefore,

$$
\bar{x} - x = \frac{1}{\mu}X\bar{X}(s - \bar{s})
$$

$$
\begin{aligned}
&= \frac{1}{\mu}X\bar{X}\left((c - A^T y) - (c + \Delta c - (A + \Delta A)^T \bar{y})\right) \\
&= \frac{1}{\mu}X\bar{X}\left(\Delta A^T \bar{y} - \Delta c\right) + \frac{1}{\mu}X\bar{X}A^T(\bar{y} - y). \qquad (16)
\end{aligned}
$$

On the other hand, $A(\bar{x} - x) = \Delta b - \Delta A\bar{x}$. Since A has rank m (otherwise $\rho(d) = 0$), then $P = AX\bar{X}A^T$ is a positive definite matrix. By combining these statements together with (16), we obtain

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

and so

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

Therefore, we have the following identity:

$$
\bar{y} - y = \mu P^{-1}\left(\Delta b - \Delta A\bar{x}\right) + P^{-1}AX\bar{X}\left(\Delta c - \Delta A^T\bar{y}\right). \qquad (17)
$$

From this identity, it follows that

$$
\|\bar{y} - y\|_\infty \leq \|P^{-1}\|_{1,\infty}\left(\mu\|\Delta b - \Delta A\bar{x}\|_1 + \|A\|\|X\bar{X}(\Delta c - \Delta A^T\bar{y})\|_1\right). \qquad (18)
$$

Note that

$$
\|X\bar{X}(\Delta c - \Delta A^T\bar{y})\|_1 \leq \|X\bar{X}\|_{\infty,1}\|\Delta c - \Delta A^T\bar{y}\|_\infty \leq \|x\|_1\|\bar{x}\|_1\|\Delta c - \Delta A^T\bar{y}\|_\infty. \qquad (19)
$$

From Corollary 3.1, we have that

$$
\begin{aligned}
\|\Delta b - \Delta A\bar{x}\|_1 &\leq \|\Delta d\|(1 + \|\bar{x}\|_1) \\
&\leq \|\Delta d\|\left(1 + \frac{4}{(1-\alpha)^2}\mathcal{K}(d,\mu)\right) \\
&\leq \frac{5\|\Delta d\|}{(1-\alpha)^2}\mathcal{K}(d,\mu), \qquad (20) \\
\|\Delta c - \Delta A^T\bar{y}\|_\infty &\leq \|\Delta d\|(1 + \|\bar{y}\|_\infty) \\
&\leq \|\Delta d\|\left(1 + \frac{4}{(1-\alpha)^2}\mathcal{K}(d,\mu)\right) \\
&\leq \frac{5\|\Delta d\|}{(1-\alpha)^2}\mathcal{K}(d,\mu). \qquad (21)
\end{aligned}
$$

Therefore, by combining (18), (19), (20), and (21), and by using Theorem 3.1, Corollary 3.1, and Lemma 4.2, we obtain the following bound on $\|\bar{y} - y\|_\infty$:

$$\begin{aligned}
\|\bar{y} - y\|_\infty &\leq 32m \left(\frac{\mathcal{C}(d)\mathcal{K}(d,\mu)}{\mu(1-\alpha)}\right)^2 \left(\frac{5\|\Delta d\|}{(1-\alpha)^2}\mathcal{K}(d,\mu)\right) \left(\mu + 4\|d\| \left(\frac{\mathcal{K}(d,\mu)}{1-\alpha}\right)^2\right) \\
&\leq 640m\|\Delta d\| \frac{\mathcal{C}(d)^2\mathcal{K}(d,\mu)^5(\mu + \|d\|)}{\mu^2(1-\alpha)^6},
\end{aligned}$$

thereby demonstrating the bound (8) on $\|\bar{y} - y\|_\infty$.

Now, by substituting identity (17) into equation (16), we obtain

$$\begin{aligned}
\bar{x} - x &= \frac{1}{\mu} X\bar{X}\left(I - A^T P^{-1} A X\bar{X}\right)\left(\Delta A^T \bar{y} - \Delta c\right) + X\bar{X}A^T P^{-1}\left(\Delta b - \Delta A\bar{x}\right) \\
&= \frac{1}{\mu} D^{\frac{1}{2}}\left(I - D^{\frac{1}{2}} A^T P^{-1} A D^{\frac{1}{2}}\right) D^{\frac{1}{2}}\left(\Delta A^T \bar{y} - \Delta c\right) + DA^T P^{-1}\left(\Delta b - \Delta A\bar{x}\right),
\end{aligned}$$

where $D = X\bar{X}$. Observe that the matrix $Q = I - D^{\frac{1}{2}} A^T P^{-1} A D^{\frac{1}{2}}$ is a projection matrix, and so $\|Qx\|_2 \leq \|x\|_2$ for all $x \in \Re^n$. Hence, from Proposition 2.1 parts (i) and (iii), we obtain that

$$\begin{aligned}
\|\bar{x} - x\|_1 &\leq \frac{1}{\mu}\|D^{\frac{1}{2}}\left(I - D^{\frac{1}{2}} A^T P^{-1} A D^{\frac{1}{2}}\right) D^{\frac{1}{2}}\left(\Delta A^T \bar{y} - \Delta c\right)\|_1 + \\
&\qquad\qquad \|DA^T P^{-1}\left(\Delta b - \Delta A\bar{x}\right)\|_1 \\
&\leq \frac{\sqrt{n}}{\mu}\|D^{\frac{1}{2}}\|_2\|D^{\frac{1}{2}}\|_2\|\Delta A^T\bar{y} - \Delta c\|_2 + \|DA^T P^{-1}\left(\Delta b - \Delta A\bar{x}\right)\|_1 \\
&= \frac{\sqrt{n}}{\mu}\|D\|_2\|\Delta A^T\bar{y} - \Delta c\|_2 + \|DA^T P^{-1}\left(\Delta b - \Delta A\bar{x}\right)\|_1 \\
&\leq \frac{n}{\mu}\max_j\{x_j\bar{x}_j\}\|\Delta A^T\bar{y} - \Delta c\|_\infty + \|D\|_{\infty,1}\|A^T\|_{\infty,\infty}\|P^{-1}\|_{1,\infty}\|\Delta b - \Delta A\bar{x}\|_1 \\
&= \frac{n}{\mu}\max_j\{x_j\bar{x}_j\}\|\Delta A^T\bar{y} - \Delta c\|_\infty + \sum_{j=1}^{n}|x_j\bar{x}_j|\|A\|\|P^{-1}\|_{1,\infty}\|\Delta b - \Delta A\bar{x}\|_1 \\
&\leq \frac{n}{\mu}\max_j\{x_j\bar{x}_j\}\|\Delta A^T\bar{y} - \Delta c\|_\infty + \|x\|_1\|\bar{x}\|_1\|d\|\|P^{-1}\|_{1,\infty}\|\Delta b - \Delta A\bar{x}\|_1
\end{aligned}$$

It follows from Lemma 4.1, Theorem 3.1, Corollary 3.1, Lemma 4.2, and inequalities (20) and (21) that

$$\|\bar{x} - x\|_1 \leq \frac{4n}{\mu}\left(\frac{\mathcal{K}(d,\mu)}{1-\alpha}\right)^2 \frac{5\|\Delta d\|}{(1-\alpha)^2}\mathcal{K}(d,\mu) +$$

$$128m\left(\frac{\mathcal{K}(d,\mu)}{1-\alpha}\right)^2\|d\|\left(\frac{\mathcal{C}(d)\mathcal{K}(d,\mu)}{\mu(1-\alpha)}\right)^2\frac{5\|\Delta d\|}{(1-\alpha)^2}\mathcal{K}(d,\mu),$$

from which we obtain the following bound (recall that $n \geq m$):

$$\|\bar{x}-x\|_1 \leq 640n\|\Delta d\|\frac{\mathcal{C}(d)^2\mathcal{K}(d,\mu)^5(\mu+\|d\|)}{\mu^2(1-\alpha)^6},$$

which thereby demonstrates the bound (7) on $\|\bar{x}-x\|_1$.

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

$$\begin{aligned}\|\bar{s}-s\|_\infty &\leq \frac{5\|\Delta d\|}{(1-\alpha)^2}\mathcal{K}(d,\mu) + 640m\|d\|\|\Delta d\|\frac{\mathcal{C}(d)^2\mathcal{K}(d,\mu)^5(\mu+\|d\|)}{\mu^2(1-\alpha)^6} \\ &\leq 640m\|\Delta d\|\frac{\mathcal{C}(d)^2\mathcal{K}(d,\mu)^5(\mu+\|d\|)^2}{\mu^2(1-\alpha)^6},\end{aligned}$$

and this concludes the proof of this theorem.
q.e.d.

We now present the proof of Theorem 4.3.

Proof of Theorem 4.3: Let $x = x(\mu)$ and $\bar{x} = x(\bar{\mu})$ be the primal optimal solutions to $P_\mu(d)$ and $P_{\bar{\mu}}(d)$, respectively; and let $(y,s) = (y(\mu), s(\mu))$ and $(\bar{y},\bar{s}) = (y(\bar{\mu}), s(\bar{\mu}))$ be the dual optimal solutions to $D_\mu(d)$ and $D_{\bar{\mu}}(d)$, respectively. From the Karush-Kuhn-Tucker optimality conditions we have that

$$\begin{array}{rclrcl} Xs &=& \mu e, & \bar{X}\bar{s} &=& \bar{\mu}e, \\ A^Ty+s &=& c, & A^T\bar{y}+\bar{s} &=& c, \\ Ax &=& b, & A\bar{x} &=& b, \\ x &>& 0, & \bar{x} &>& 0. \end{array}$$

Therefore,

$$\begin{aligned}\bar{x}-x &= \frac{1}{\mu\bar{\mu}}X\bar{X}\left(\bar{\mu}s-\mu\bar{s}\right) \\ &= \frac{1}{\mu\bar{\mu}}X\bar{X}\left(\bar{\mu}(c-A^Ty)-\mu(c-A^T\bar{y})\right) \\ &= \frac{1}{\mu\bar{\mu}}X\bar{X}\left((\bar{\mu}-\mu)c-A^T(\bar{\mu}y-\mu\bar{y})\right).\end{aligned} \tag{22}$$

On the other hand, $A(\bar{x}-x) = b - b = 0$. Since A has rank m (otherwise $\rho(d) = 0$), then $P = AX\bar{X}A^T$ is a positive definite matrix. By combining these statements together with (22), we obtain

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

and so

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

equivalently

$$\bar{\mu}y - \mu\bar{y} = (\bar{\mu}-\mu)P^{-1}AX\bar{X}c. \tag{23}$$

By substituting identity (23) into equation (22) and by letting $D = X\bar{X}$, we obtain:

$$\begin{aligned}
\bar{x} - x &= \frac{\bar{\mu}-\mu}{\mu\bar{\mu}} X\bar{X}\left(c - A^T P^{-1} AX\bar{X}c\right) \\
&= \frac{\bar{\mu}-\mu}{\mu\bar{\mu}} D\left(c - A^T P^{-1} ADc\right) \\
&= \frac{\bar{\mu}-\mu}{\mu\bar{\mu}} D^{\frac{1}{2}}\left(I - D^{\frac{1}{2}} A^T P^{-1} AD^{\frac{1}{2}}\right) D^{\frac{1}{2}}c,
\end{aligned}$$

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 2.1, parts (i) and (iii), and Theorem 3.1, we have

$$\begin{aligned}
\|\bar{x} - x\|_1 &\leq \sqrt{n}\|\bar{x} - x\|_2 \\
&\leq \frac{\sqrt{n}}{\mu\bar{\mu}}|\bar{\mu}-\mu|\|D^{\frac{1}{2}}\|_2\|D^{\frac{1}{2}}\|_2\|c\|_2 \\
&= \frac{\sqrt{n}}{\mu\bar{\mu}}|\bar{\mu}-\mu|\|D\|_2\|c\|_2 \\
&\leq \frac{n}{\mu\bar{\mu}}|\bar{\mu}-\mu|\|x\|_1\|\bar{x}\|_1\|c\|_\infty \\
&\leq \frac{n}{\mu\bar{\mu}}|\bar{\mu}-\mu|\mathcal{K}(d,\mu)\mathcal{K}(d,\bar{\mu})\|d\|,
\end{aligned}$$

which demonstrates the bound (10) for $\|\bar{x} - x\|_1$.

Now, since $c = A^T y + s$ and $c = A^T\bar{y} + \bar{s}$, it follows that $A^T(\bar{y} - y) + \bar{s} - s = 0$, which yields the following equalities in logical sequence:

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

so that by premultiplying by A, we obtain

$$\begin{aligned}
AX\bar{X}A^T(y-\bar{y}) &= (\bar{\mu}-\mu)b, \\
P(y-\bar{y}) &= (\bar{\mu}-\mu)b, \\
y-\bar{y} &= (\bar{\mu}-\mu)P^{-1}b.
\end{aligned}$$

Therefore, from Corollary 4.3,

$$\begin{aligned}
\|\bar{y}-y\|_\infty &\leq |\bar{\mu}-\mu|\|P^{-1}\|_{1,\infty}\|b\|_1 \\
&\leq \frac{4m}{\mu\bar{\mu}}|\bar{\mu}-\mu|\mathcal{C}(d)^2\mathcal{K}(d,\mu)\mathcal{K}(d,\bar{\mu})\|d\|,
\end{aligned}$$

which establishes the bound (11) for $\|\bar{y}-y\|_\infty$.

Finally, using the fact that $s-\bar{s}=A^T(\bar{y}-y)$, we obtain $\|s-\bar{s}\|_\infty \leq \|A^T\|_{\infty,\infty}\|y-\bar{y}\|_\infty = \|A\|\|y-\bar{y}\|_\infty \leq \|d\|\|y-\bar{y}\|_\infty$, which establishes (12) from (11), and so this concludes the proof of this theorem.
q.e.d.

We now prove Theorem 4.4.

Proof of Theorem 4.4: Consider the Lagrangian functions associated with these problems,

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

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

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

Hence, if $(x(\mu),y(\mu))$ is a pair of optimal solutions to the primal and dual programs corresponding to d, and $(\bar{x}(\mu),\bar{y}(\mu))$ is a pair of optimal solutions to the primal and dual programs corresponding to $d+\Delta d$, then

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

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

Therefore, we obtain the following bounds: either

$$\begin{array}{rcl} |z(d+\Delta d) - z(d)| & \leq & |\Phi(x(\mu), \bar{y}(\mu))|, \\ & \text{or} & \\ |z(d+\Delta d) - z(d)| & \leq & |\Phi(\bar{x}(\mu), y(\mu))|. \end{array}$$

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

$$\begin{aligned} |\Phi(x(\mu), \bar{y}(\mu))| &= |\Delta c^T x(\mu) + \bar{y}(\mu)^T \Delta b - \bar{y}(\mu)^T \Delta A x(\mu)| \\ &\leq \|\Delta c\|_\infty \|x(\mu)\|_1 + \|\bar{y}(\mu)\|_\infty \|\Delta b\|_1 + \|\bar{y}(\mu)\|_\infty \|\Delta A\| \|x(\mu)\|_1 \\ &\leq \|\Delta d\| \left(\|x(\mu)\|_1 + \|\bar{y}(\mu)\|_\infty + \|\bar{y}(\mu)\|_\infty \|x(\mu)\|_1 \right) \\ &\leq 3\|\Delta d\| \left(\frac{1+\alpha}{1-\alpha} \right)^4 \mathcal{K}(d, \mu)^2. \end{aligned}$$

Similarly, we can show that

$$|\Phi(\bar{x}(\mu), y(\mu))| \leq 3\|\Delta d\| \left(\frac{1+\alpha}{1-\alpha} \right)^4 \mathcal{K}(d, \mu)^2,$$

and the result follows.
q.e.d.

Finally, we prove Theorem 4.5.

Proof of Theorem 4.5: Let $x(\mu)$ and $x(\bar{\mu})$ be the optimal solutions to $P_\mu(d)$ and $P_{\bar{\mu}}(d)$, respectively; and $(y(\mu), s(\mu))$ and $(y(\bar{\mu}), s(\bar{\mu}))$ be the optimal solutions to $D_\mu(d)$ and $D_{\bar{\mu}}(d)$, respectively. As in Theorem 4.4, for given $\mu, \bar{\mu} > 0$, consider the following Lagrangian functions: $L(x, y) = c^T x + \mu p(x) + y^T(b - Ax)$ and $\bar{L}(x, y) = c^T x + \bar{\mu} p(x) + y^T(b - Ax)$. 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 4.4, 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

$$\begin{array}{rcccl} |z(\bar{\mu}) - z(\mu)| & \leq & |-\Phi(x(\mu), y(\bar{\mu}))| & = & |\bar{\mu} - \mu||p(x(\mu))|, \\ & & \text{or} & & \\ |z(\bar{\mu}) - z(\mu)| & \leq & |-\Phi(x(\bar{\mu}), y(\mu))| & = & |\bar{\mu} - \mu||p(x(\bar{\mu}))|. \end{array}$$

Therefore,

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

On the other hand, from Theorem 3.1 and Theorem 3.2, we have

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

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

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

so that

$$\begin{aligned} |p(x(\mu))| &\leq n \max\left\{\left|\ln\left(\frac{\mu}{2\|d\|\mathcal{K}_\mu(d)}\right)\right|, \ln(\mathcal{K}(d,\mu))\right\} \\ &\leq n\Big(\ln(2) + \ln(\mathcal{K}(d,\mu)\mathcal{K}(d,\bar{\mu})) + |\ln(\|d\|)| + \max\{|\ln(\mu)|, |\ln(\bar{\mu})|\}\Big). \end{aligned}$$

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

$$|p(x(\bar{\mu}))| \leq n\Big(\ln(2) + \ln(\mathcal{K}(d,\mu)\mathcal{K}(d,\bar{\mu})) + |\ln(\|d\|)| + \max\{|\ln(\mu)|, |\ln(\bar{\mu})|\}\Big),$$

and the result follows.
q.e.d.

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 analytic center problem in equality form, denoted $AE(d)$, is defined as:

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

Structurally, the program $AE(d)$ is closely related to the central trajectory problem $P_\mu(d)$, and was first extensively studied by Sonnevend, see [26] and [27]. In terms of data dependence, note that the program $AE(d)$ does not depend on the data c. It is well known that

$AE(d)$ 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 analytic center problem in inequality form, denoted $AI(d)$, as:

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

In terms of data dependence, the program $AI(d)$ does not depend on the data b. The program $AI(d)$ 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 $AE(d)$ and $AI(d)$ are *not* duals of each other. (In fact, direct calculation reveals that $AE(d)$ and $AI(d)$ cannot both be solvable, since at least one of $AE(d)$ and $AI(d)$ must be unbounded.) 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 $AE(d)$ and $AI(d)$, 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: $\mathcal{D}_E = \{(A, b) : A \in \Re^{m\times n}, b \in \Re^m\}$ and $\mathcal{D}_I = \{(A, c) : A \in \Re^{m\times n}, c \in \Re^n\}$. In a manner similar to the central trajectory problem, we define the following feasibility sets for analytic center problems:

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

in other words, $\mathcal{F}_E$ consists of data instances d for which $AE(d)$ is feasible and attains its optimal value, that is, $AE(d)$ is solvable; and $\mathcal{F}_I$ consists of data instances d for which $AI(d)$ is feasible and attains its optimal value, that is, $AI(d)$ is solvable. It is also appropriate to introduce the corresponding sets of ill-posed data instances: $\mathcal{B}_E = cl(\mathcal{F}_E) \cap cl(\mathcal{F}_E^C) = \partial\mathcal{F}_E = \partial\mathcal{F}_E^C$, and $\mathcal{B}_I = cl(\mathcal{F}_I) \cap cl(\mathcal{F}_I^C) = \partial\mathcal{F}_I = \partial\mathcal{F}_I^C$.

For the analytic center problem in equality form $AE(d)$, the distance to ill-posedness of a data instance $d = (A, b, c)$ is defined as $\rho_E(d) = \inf\{\|(\Delta A, \Delta b)\|_E : (A+\Delta A, b+\Delta b) \in \mathcal{B}_E\}$. For the analytic center problem in inequality form $AD(d)$, the distance to ill-posedness of a data instance $d = (A, b, c)$ is defined as $\rho_I(d) = \inf\{\|(\Delta A, \Delta c)\|_I : (A+\Delta A, c+\Delta c) \in \mathcal{B}_I\}$, where $\|(A, b)\|_E = \max\{\|A\|, \|b\|_1\}$ and $\|(A, c)\|_I = \max\{\|A\|, \|c\|_\infty\}$. Likewise, the corresponding condition measures are $\mathcal{C}_E(d) = \|(A, b)\|_E/\rho_E(d)$ if $\rho_E(d) > 0$ and $\mathcal{C}_E(d) = \infty$

otherwise; $\mathcal{C}_I(d) = \|(A,c)\|_I/\rho_I(d)$ if $\rho_I(d) > 0$ and $\mathcal{C}_I(d) = \infty$ otherwise.

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

Proof: Given any $\epsilon > 0$, consider $\delta = \rho_E(d) - \epsilon$. If $d + \Delta d = (A + \Delta A, b + \Delta b, c + \Delta c)$ is a data instance such that $\|\Delta d\| \le \delta$, then $\|(\Delta A, \Delta b)\|_E \le \delta$. Hence, $(A + \Delta A, b + \Delta b) \in \mathcal{F}_E$, so that the system $(A + \Delta A)x = b + \Delta b$, $x > 0$, $(A + \Delta A)^T y < 0$ has a solution, and therefore the system $(A + \Delta A)x = b + \Delta b$, $x > 0$, $(A + \Delta A)^T y < c$ also has a solution, that is, $d + \Delta d \in \mathcal{F}$. Therefore, $\rho(d) \ge \delta = \rho_E(d) - \epsilon$, 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 analytic center problems in equality form and in inequality form, respectively.

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

$$\|x\|_1 \le \mathcal{C}_E(d)$$

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

Proof: Let x be a feasible solution of $AE(d)$. Define $\Delta A = -be^T/\|x\|_1$ and $\Delta d = (\Delta A, 0, 0)$. Then, $(A + \Delta A)x = 0$ and $x > 0$. Now, consider the program $AE(d + \Delta d)$. 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 $(A + \Delta A, b) \in \mathcal{F}_E^C$, whereby $\rho_E(d) \le \|(\Delta A, 0)\|_E$. On the other hand, $\|(\Delta A, 0)\|_E = \|b\|_1/\|x\|_1 \le \|(A,b)\|_E/\|x\|_1$, so that $\|x\|_1 \le \|(A,b)\|_E/\rho_E(d) = \mathcal{C}_E(d)$.
q.e.d.

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

$$\begin{aligned} \|y\|_\infty &\le \mathcal{C}_I(d), \\ \|s\|_\infty &\le 2\|(A,c)\|_I \mathcal{C}_I(d), \end{aligned}$$

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

Proof: Let (y,s) be a feasible solution of $AI(d)$. If $y = 0$, then $s = c$ and the bounds are trivially true, so that we assume $y \ne 0$. Let $\bar{y}$ be such that $\|y\|_\infty = \bar{y}^T y$ and $\|\bar{y}\|_1 = 1$. Let $\Delta A = -\bar{y}c^T/\|y\|_\infty$ and $\Delta d = (\Delta A, 0, 0)$. Hence, $(A + \Delta A)^T y = A^T y - c < 0$. Because

$(A + \Delta A)^T y < 0$ has a solution, there cannot exist x for which $(A + \Delta A)x = 0$ and $x > 0$, and so $(A + \Delta A, c) \in \mathcal{F}_I^C$, whereby $\rho_I(d) \leq \|(\Delta A, 0)\|_I$. On the other hand, $\|(\Delta A, 0)\|_I = \|c\|_\infty / \|y\|_\infty \leq \|(A, c)\|_I / \|y\|_\infty$, so that $\|y\|_\infty \leq \|(A, c)\|_I / \rho_I(d) = \mathcal{C}_I(d)$. The bound for $\|s\|_\infty$ is easily derived using the fact that $\|s\|_\infty \leq \|c\|_\infty + \|A^T\|_{\infty,\infty} \|y\|_\infty = \|c\|_\infty + \|A\| \|y\|_\infty$ and $\mathcal{C}_I(d) \geq 1$.
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 [4], p. 139, and [31]), there exists a unique partition $[B, N]$ of the set $\{1, \ldots, n\}$ into subsets B and N, $B \cap N = \emptyset$ and $B \cup N = \{1, \ldots, n\}$ satisfying the following two properties:

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. $A^T y = v$, $v \leq 0$ 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 $A_B^T \tilde{y} \leq 0$ and $A_B^T \tilde{y} \neq 0$. Then, using the vector $\hat{y}$ from property 2 above, we obtain that $A_N^T(\tilde{y} + \lambda \hat{y}) = A_N^T \tilde{y} + \lambda \hat{v}_N \leq 0$ for λ sufficiently large, and since $A_B^T \hat{y} = \hat{v}_B = 0$, it follows that $A^T(\tilde{y} + \lambda \hat{y}) \leq 0$ for λ sufficiently large. By the definition of the partition $[B, N]$, we have that $A_B^T(\tilde{y} + \lambda \hat{y}) = 0$. This in turn implies that $A_B^T \tilde{y} = 0$, a contradiction.

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

$$\|s_B(\mu)\|_\infty \leq 2\|(A_B, c_B)\|_I \mathcal{C}_I(d_B) \leq 2\|d\| \mathcal{C}_I(d_B).$$

Hence, for any $j \in B$, $s_j(\mu) \leq \|s_B(\mu)\|_\infty \leq 2\|d\|\mathcal{C}_I(d_B)$. Moreover, since $x_j(\mu)s_j(\mu) = \mu$, then

$$x_j(\mu) \geq \frac{\mu}{2\|d\|\mathcal{C}_I(d_B)},$$

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] S. A. Ashmanov. Stability conditions for linear programming problems. *U.S.S.R. Comput. Maths. Math. Phys*, 21(6):40–49, 1981.

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

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

[5] 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.

[6] 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.

[7] Sharon Filipowski. On the complexity of solving sparse symmetric linear programs specified with approximate data. Technical report, Dept. of Industrial and Manufacturing Systems Engineering, Iowa State University, December 1994.

[8] Sharon Filipowski. On the complexity of solving feasible systems of linear inequalities specified with approximate data. *Mathematical Programming*, 71(3):259–288, December 1995.

[9] Robert M. Freund and Jorge R. Vera. Some characterizations and properties of the "distance to ill-posedness" in conic linear systems. MIT Sloan Working Paper 3862-95-MSA, Sloan School of Management, Massachusetts Institute of Technology, 1995.

[10] D. Gale. *The Theory of Linear Economic Models.* McGraw-Hill Book Company, Inc, New York, 1960.

[11] Gene H. Golub and Charles F. Van Loan. *Matrix Computations.* The Johns Hopkins University Press, New York, third edition, 1997.

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

[13] 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.

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

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

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

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

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

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

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

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

[22] James Renegar. Condition numbers, the barrier method, and the conjugate gradient method. *SIAM Journal on Optimization*, 6(4):879–912, November 1996.

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

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

[25] Stephen M. Robinson. A characterization of stability in linear programming. *Operations Research*, 25(3):435–447, 1977.

[26] 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.

[27] 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.

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

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

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

[31] A. C. Williams. Complementarity theorems for linear programming. *SIAM Review*, 12(1):135–137, January 1970.

[32] Stephen J. Wright. *Primal-Dual Interior-Point Methods.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1997.