OPERATIONS RESEARCH CENTER

working paper

**MASSACHUSETTS INSTITUTE
OF TECHNOLOGY**

**A METHOD FOR THE PARAMETRIC CENTER
PROBLEM, WITH A STRICTLY MONOTONE
POLYNOMIAL-TIME ALGORITHM FOR
LINEAR PROGRAMMING**

by Robert M. Freund
and
Kok-Choon Tan

OR 192-89

March, 1989

Abstract

Given a system of linear inequalities and equalities $Ax \leq b + dt$ and $Mx = g + ht$ where the right-hand-sides (RHS) are parametrically deformed over the scalar t , the *parametric center problem* is to trace the parametric family of approximate solutions $\bar{x}(t)$ to the center problems $P(t)$, where $P(t)$ is the problem: maximize $\sum_{i=1}^m \ln(b_i + d_i t - A_i x)$ subject to $Ax < b + dt$ and $Mx = g + ht$. We present an algorithm for tracing the parametric family of solutions $\bar{x}(t)$ over the given range $t \in [\underline{t}, \bar{t}]$. At each iterate of the algorithm, the value of the parameter t is strictly increased and a Newton step is taken. The sequence of values of t exhibit the following geometric rate of change: If t^k and t^{k+1} are two successive values of the parameter t generated by the algorithm, then either $(t^{k+1} - T_{\min}) \geq (t^k - T_{\min})\left(1 + \frac{1}{128m}\right)$ or $(T_{\max} - t^{k+1}) \leq (T_{\max} - t^k)\left(1 - \frac{1}{128m}\right)$, where $T_{\min}$ ($T_{\max}$) is a lower (upper) bound on the smallest (largest) value of t for which $Ax \leq b + dt$, $Mx = g + ht$ has a solution. Thus the iterates exhibit either linear growth away from $T_{\min}$ or linear convergence toward $T_{\max}$, with a rate of change of $\frac{1}{128m}$, where m is the number of inequality constraints.

When applied to the linear programming problem, the algorithm is an $O(mL)$ iteration algorithm for linear programming, that strictly improves the primal objective value at each iteration, and requires no dual feasible solution (or even dual feasibility) to start. After $O(mL)$ iterations, the algorithm either detects primal unboundedness or produces an interior solution that can be rounded to an optimal solution to the linear program.

Key Words: Newton step, center, linear program, interior-point algorithm.

Introduction

Given a system of m linear inequalities in $\mathbb{R}^n$ of the form $Ax \leq b$ and k equations in $\mathbb{R}^n$ of the form $Mx = g$, the (analytic) center of the system (A, b, M, g) is the optimal solution to the convex program:

$$\begin{aligned} \text{P:} \quad & \text{maximize} \quad \sum_{i=1}^m \ln(b_i - A_i x) \\ & \text{s.t.} \quad Ax < b, \\ & \quad \quad \quad Mx = g \end{aligned}$$

where (A_i, b_i) denotes i^{th} row of A and i^{th} component of b , respectively. (See Sonnevend [12, 13]. Assuming $\{x \in \mathbb{R}^n \mid Ax \leq b, Mx = g\}$ is nonempty and bounded, the center of the system (A, b, M, g) , denoted $\hat{x}$, is uniquely defined. The computation of points near the center and their properties are important for interior point algorithms for linear programming and extensions, see Karmarkar [6], Renegar [11], Megiddo [8], Kojima et. al. [7], Vaidya [16], Monteiro and Adler [10], Mehrotra and Sun [9], Jarre [5], Barnes et. al. [1], and Todd and Ye [14], among others. Algorithms for finding the center are presented in Censor and Lent [2], Vaidya [15], and [4].

This study is concerned with the parametric analysis of the family of centers as the right-hand-side (RHS) (b, g) of the system (A, b, M, g) varies parametrically. We define the parametric center problem (PCP) to be the problem of tracing the parametric family of optimal solutions $\hat{x}_t$ to the problems:

$$\begin{aligned}
P(t): \quad & \text{maximize} && \sum_{i=1}^m \ln (b_i + d_i t - A_i x) \\
& \text{s.t.} && Ax < b + dt, \\
& && Mx = g + ht.
\end{aligned}$$

(where $d \in \mathbb{R}^m$ and $h \in \mathbb{R}^k$ are given), as t varies over a given interval $t \in [\underline{t}, \bar{t}]$, and $\underline{t}, \bar{t}$ are finite or infinite. We present an algorithm for generating a piecewise-linear function $\bar{x}(t) : [\underline{t}, \bar{t}] \rightarrow \mathbb{R}^n$ with the property that $\bar{x}(t)$ is an approximation to the center $\hat{x}(t)$, i.e., $\bar{x}(t)$ is an approximate solution to $P(t)$, as t is varied. We refer to $\bar{x}(t)$ as the approximate path of solutions. (The sense of the approximation and its properties are defined in Section 3.)

The algorithm starts with an approximate solution $\bar{x}(\underline{t})$ to the center problem $P(t)$ at $t = \underline{t}$. At iteration k , the value of t is $t = t^k$ and the approximate center for $P(t^k)$ is $\bar{x}(t^k)$. The next value of t is chosen as $t = t^{k+1}$ where $t^{k+1} > t^k$ and a Newton step is performed to solve for $\bar{x}(t^{k+1})$. The path $\bar{x}(t)$ is then extended for $t \in [t^k, t^{k+1}]$ by linear interpolation. The important feature of the algorithm is the guaranteed increase in t at each iteration. In particular, t^{k+1} can be chosen so that either $(t^{k+1} - T_{\min}) \geq (t^k - T_{\min}) \left(1 + \frac{1}{128m}\right)$ or $(T_{\max} - t^{k+1}) \leq (T_{\max} - t^k) \left(1 - \frac{1}{128m}\right)$, where $T_{\min}$ ($T_{\max}$) is a lower (upper) bound on the smallest (largest) value of t for which $Ax \leq b + dt$, $Mx = g + ht$, has a solution. Thus the iterate values of t demonstrate either geometric growth away from $T_{\min}$ or geometric

contraction toward T_{MAX} , with a rate of change of $\frac{1}{128m}$ (where m is the number of inequality constraints).

This algorithm can be applied to solve the linear programming problem in a new way. Suppose we wish to solve the linear problem

$$\begin{array}{ll} \text{LP:} & \text{maximize} & c^T x \\ & \text{s.t.} & Ax \leq b \end{array}$$

Then we can use the algorithm for PCP to solve for the path of centers of the system

$$\begin{array}{ll} \text{LP}(t): & \text{maximize} & \sum_{i=1}^m \ln(b_i - A_i x) \\ & \text{s.t.} & Ax < b \\ & & cx = t \end{array},$$

as t is increased. This yields a new "central-trajectory-following" algorithm for linear programming that differs from other central-trajectory methods in two ways. First, it is a strictly monotone algorithm for linear programming, i.e., an algorithm that strictly increases the objective function value of the primal at each iteration, unlike other central-trajectory-following algorithms. Second, it requires no prior information or bound on the optimal objective value, and will process a linear program that is unbounded in the primal objective value (i.e., dual infeasible), unlike other central-trajectory methods. However, the complexity of the algorithm is $O(mL)$ iterations, as opposed to $O(\sqrt{m}L)$ for most other central-trajectory methods (where L is the bit-size of the problem instance) and so has an inferior

complexity bound (by $\sqrt{m}$) . Perhaps it is the strict monotonicity of the primal objective value in the algorithm that is responsible for the inferior complexity bound.

This paper is organized as follows. In Section 2, we present the main results regarding the parametric center problem and we present the algorithm for tracing the approximate parametric path of centers. The remaining three sections are devoted to proofs of the results of Section 2. Section 3 presents notation and preliminary results. Section 4 contains an analysis of one step of the algorithm and presents the results on the use of Newton's method. Section 5 contains the results regarding bounds on feasible values of t that are generated by the algorithm.

2. The Parametric Center Problem

Given a system of m linear inequalities in $\mathbb{R}^n$ of the form $Ax \leq b$ and equations $Mx = g$, the (analytic) center of the system referred to as (A, b, M, g) is the optimal solution to the program

$$\begin{aligned} P: \quad & \text{maximize} \quad \sum_{i=1}^m \ln(b_i - A_i x) \\ & \text{s.t.} \quad Ax < b, \\ & \quad \quad \quad Mx = g \end{aligned}$$

(see Sonnevend [12, 13].)

Suppose $\hat{x}$ is the unique solution to P . Our interest lies in tracing the center $\hat{x}$ as the right-hand-side (RHS) of the system (A, b, M, g) varies parametrically. In particular, we are interested in generating the parametric family of optimal solutions $\hat{x}_t$ to the problems

$$\begin{aligned} P'(t): \quad & \text{maximize} \quad \sum_{i=1}^m \ln(b_i + d_i t - A_i x) \\ & \text{s.t.} \quad Ax < b + dt, \\ & \quad \quad \quad Mx + g + ht. \end{aligned} \tag{2.1}$$

In this section, we present an algorithm for generating a piecewise-linear path of solutions $\bar{x}(t)$ such that $\bar{x}(t)$ is close to $\hat{x}_t$ in a suitable measure, and as t is varied strictly monotonically over a prespecified range.

First note that there is no loss of generality in assuming that the equations $Mx = g + ht$ are not present. To see this, we can assume without loss of generality that $M = [B, N]$ is a $k \times n$ matrix where B is $k \times k$ and is nonsingular. By suitably partitioning $A = [C, D]$ and

$x = (y, z)$, we can eliminate the y variables to obtain the equivalent problem

$$\begin{aligned} P''(t): \quad & \text{maximize} \quad \sum_{i=1}^m \ln (\bar{b}_i + \bar{d}_i t - \bar{A}_i z) \\ & \text{s.t.} \quad \bar{A} z < \bar{b} \quad , \end{aligned}$$

where $\bar{A} = D - CB^{-1}N$, $\bar{b} = b - CB^{-1}g$, and $\bar{d} = d - CB^{-1}h$. It is straightforward to show that $\hat{z}_t$ solves $P''(t)$ if and only if $\hat{x}_t = (\hat{y}_t, \hat{z}_t)$ solves $P'(t)$, where $\hat{y}_t = B^{-1}(g + ht - N\hat{z}_t)$. We thus can concentrate on the more convenient problem

$$\begin{aligned} P(t): \quad & \text{maximize} \quad \sum_{i=1}^m \ln (b_i + d_i t - A_i x) \\ & \text{s.t.} \quad Ax < b + dt \quad . \end{aligned} \tag{2.2}$$

Let $X = \{x \in R^n \mid Ax \leq b\}$ and $X_t = \{x \in R^n \mid Ax \leq b + dt\}$. Suppose $\hat{x}_t$ is the center of X_t , i.e., $\hat{x}_t$ solves $P(t)$. Let $\hat{s}_t = b + dt - A\hat{x}_t$, and let $\hat{S}_t = \text{diag}(\hat{s}_t)$.

Because $P(t)$ is a convex program, $\hat{x}_t$ solves $P(t)$ if and only if the Karush-Kuhn-Tucker (K-K-T) conditions are satisfied at $\hat{x}_t$, namely

$$\hat{s}_t = b + dt - A\hat{x}_t > 0 \tag{2.3a}$$

$$e^T \hat{S}_t^{-1} A = 0 \quad , \tag{2.3b}$$

where e is the vector of ones of appropriate dimension.

We assume that our initial value of t is $t^0 = 0$, and the interior of $X_{t^0} = \{x \in R^n \mid Ax \leq b\}$ is nonempty and bounded. In this case, it is

straightforward to show that X_t will be bounded for all values of t . Let $T = \{t \mid \text{int } X_t \neq \emptyset\}$. Then T will be an open interval, and it is straightforward to extend the analysis in Megiddo [8] to show that the path of parametric centers $\Gamma: t \rightarrow \hat{x}_t, t \in T$, is continuous and differentiable.

Let $T_{\text{MAX}} = \sup_t \{t \in T\}$ and $T_{\text{MIN}} = \inf_t \{t \in T\}$. It is also straightforward to show that $T_{\text{MAX}} = +\infty$ and $T_{\text{MIN}} = -\infty$ if and only if $d = Ar$ for some $r \in \mathbb{R}^n$, in which case the sets X_t are just a family of translations of X_0 by the translation vector tr . We therefore assume throughout this paper that d does not lie in the column range of A . In this case, either $T_{\text{MIN}} > -\infty$, or $T_{\text{MAX}} < +\infty$, or both.

The algorithm presented in this section will trace a piecewise-linear path $\bar{x}(t)$ that is "close" to the parametric center path $\hat{x}_t$ in a suitable measure. At each iteration the value of t is strictly increased; thus the algorithm is strictly monotone in the parameter t . Furthermore, at each iteration, the magnitude of the increase in t is bounded (from below) in at least one of two ways, as follows. Suppose t^k is the current value of t . Then at the next iteration the algorithm will produce either a finite lower bound $LB \leq T_{\text{MIN}}$ or a finite upper bound $UB \geq T_{\text{MAX}}$, or both. If UB is produced, then t^{k+1} , the new value of t , will satisfy

$$t^{k+1} - t^k \geq \frac{1}{128m} (UB - t^k), \text{ so that } (T_{\text{MAX}} - t) \text{ decreases geometrically,}$$

$$\text{i.e., } (T_{\text{MAX}} - t^{k+1}) \leq \left(1 - \frac{1}{128m}\right) (T_{\text{MAX}} - t^k). \text{ If } LB \text{ is produced, then } t^{k+1}$$

$$\text{will satisfy } (t^{k+1} - t^k) \geq \frac{1}{128m} (t^k - LB), \text{ so that } (t - T_{\text{MIN}}) \text{ grows}$$

$$\text{geometrically, i.e., } (t^{k+1} - T_{\text{MIN}}) \geq \left(1 + \frac{1}{128m}\right) (t^k - T_{\text{MIN}}).$$

At least one of the above two bounds must be satisfied. Note that in either of the two cases the geometric rate of change of t relative to a bound is at least $\frac{1}{128m}$, where m is the number of inequality constraints.

Given a linear program in the form:

$$\begin{array}{ll} \overline{\text{LP}}: & \text{maximize} \quad c^T x \\ & x \\ & \text{s.t.} \quad Ax \leq b \end{array}$$

we can reformulate $\overline{\text{LP}}$ as the following equivalent problem

$$\begin{array}{ll} \text{LP}: & \text{maximize} \quad t \\ & x, t \\ & \text{s.t.} \quad Ax \leq b + 0t \\ & \quad \quad c^T x = 0 + t \end{array}$$

Thus an algorithm that traces the parametric path of centers to the above system for strictly increasing values of t will be a strictly monotone algorithm for solving $\overline{\text{LP}}$. The application of the parametric center problem algorithm to linear programming is presented at the end of this section, and is an $O(mL)$ iteration algorithm for linear programming.

2.1 Properties of the Parametric Center

For $t \in (T_{\text{MIN}}, T_{\text{MAX}})$, let $\hat{x}_t$ be the center of X_t , and let $\hat{S}_t = b + dt - A\hat{x}_t$, and let u_t be the projection of $\hat{S}_t^{-1}d$ onto the range space of $\hat{S}_t^{-1}A$, i.e., $u_t = [I - \hat{S}_t^{-1}A(A^T\hat{S}_t^{-2}A)^{-1}A^T\hat{S}_t^{-1}]\hat{S}_t^{-1}d$. Then note that $d = \hat{S}_t u_t + Ar_t$ for some $r_t \in \mathbb{R}^n$. We now define the path indicator function $f(t)$ as $f(t) = e^T u_t$, and note immediately that

$f(t) = e^T u_t = e^T \hat{S}_t^{-1} (d - Ar_t) = e^T \hat{S}_t^{-1} d$, since $e^T \hat{S}_t^{-1} A = 0$, by (2.3).

Therefore an alternative equivalent definition of $f(t)$ is $f(t) = e^T \hat{S}_t^{-1} d$.

The motivation for considering the path indicator function $f(t)$ is as follows. It is obvious that T_{MAX} is the optimal objective value of the following linear programming problem:

$$\begin{aligned} \text{LP}^\circ: \quad T_{MAX} &= \underset{x, t}{\text{maximize}} \quad t \\ \text{s.t.} \quad &Ax - dt \leq b \end{aligned}$$

Suppose for a given value of t that $f(t) < 0$. Let $\hat{x}_t$ be the solution to $P(t)$, let $\hat{s}_t = b + dt - A\hat{x}_t$, and $\hat{S}_t = \text{diag}(\hat{s}_t)$. Then for any feasible solution $(\tilde{x}, \tilde{t})$ to LP° ,

$$\begin{aligned} \tilde{t} &= \frac{e^T \hat{S}_t^{-1} d \tilde{t}}{f(t)} \leq \frac{1}{f(t)} [e^T \hat{S}_t^{-1} (A\tilde{x} - b)] = \frac{1}{f(t)} [e^T \hat{S}_t^{-1} (A\tilde{x} - A\hat{x}_t - \hat{s}_t + dt)] \\ &= \frac{1}{f(t)} [-m + t f(t)] \end{aligned}$$

whereby $T_{MAX} \leq t - \frac{m}{f(t)}$. Similarly, if $f(t) > 0$, $T_{MIN} \geq t - \frac{m}{f(t)}$.

Now suppose that for a given value of t , say t^* , that $f(t^*) = 0$. Let $\hat{x}_{t^*}$ be the center of the system $Ax \leq b + dt^*$, and let $\hat{S}_{t^*} = \text{diag}(b + dt^* - A\hat{x}_{t^*})$. Then $e^T \hat{S}_{t^*}^{-1} A = 0$ and $e^T \hat{S}_{t^*}^{-1} d = 0$, whereby $(\hat{x}_{t^*}, t^*)$ is the center of the system $(A, -d)(x, t) \leq b$, from (2.3). Thus the set $X^\circ = \{(x, t) \in \mathbb{R}^{n+1} \mid Ax - dt \leq b\}$ is bounded, and in particular T_{MIN} and T_{MAX} are finite. We say X° is bounded in t if both T_{MIN} and T_{MAX} are finite. If T_{MIN} (T_{MAX}) is finite, we say that X° is bounded from below (above) in t . Note that if X° is

bounded in t , then $\dot{t}^*$, the value of t which yields $f(t^*) = e^T \hat{S}_t^{-1} d = 0$, is guaranteed to exist.

Returning to the path indicator $f(t)$ defined earlier, we define $\hat{x}_t$ to be on the upper path if $f(t) \leq 0$, and to be on the lower path if $f(t) \geq 0$. The intuition behind this definition is provided in the next two propositions.

Proposition 2.1 The path indicator function $f(t) = e^T u_t$ is strictly decreasing for $t \in (T_{\min}, T_{\max})$.

Proof: The K-K-T conditions of $P(t)$ require that

$$e^T \hat{S}_t^{-1} A = 0 \quad \text{and} \quad A \hat{x}_t + \hat{s}_t = b + dt.$$

Let $\dot{s}_t$ and $\dot{x}_t$ be the vector of derivatives of $\hat{s}_t$ and $\hat{x}_t$ at t , respectively. Then differentiating the above expressions yields $\dot{s}_t \hat{S}_t^{-2} A = 0$ and $A \dot{x}_t + \dot{s}_t = d$.

Furthermore, since $f(t) = e^T u_t = e^T \hat{S}_t^{-1} d$,

$$\text{then } f'(t) = -\dot{s}_t \hat{S}_t^{-2} d = -\dot{s}_t \hat{S}_t^{-2} (A \dot{x}_t + \dot{s}_t) = -\dot{s}_t \hat{S}_t^{-2} \dot{s}_t < 0,$$

unless $\dot{s}_t = 0$, in which case $d = A \dot{x}_t$. But this last possibility is ruled out by the assumption that d does not lie in the column range of A . ■

Proposition 2.2 (Upper and Lower Paths)

- (i) T_{MAX} and T_{MIN} are both finite if and only if there exists $t^* \in \mathbb{R}$ such that $f(t^*) = 0$, and

$$\begin{aligned} f(t) &> 0 \quad \text{for all } t \in (T_{MIN}, t^*), \\ f(t) &< 0 \quad \text{for all } t \in (t^*, T_{MAX}); \end{aligned}$$

- (ii) T_{MAX} is finite and $T_{MIN} = -\infty$ if and only if $f(t) < 0$ for all $t < T_{MAX}$;

- (iii) T_{MIN} is finite and $T_{MAX} = \infty$ if and only if $f(t) > 0$ for all $t > T_{MIN}$.

These three cases are illustrated in Figure 1, (a), (b) and (c).

Proof: (i) We have seen earlier that if $f(t^*) = 0$, then $(\hat{x}_t^*, t^*)$ is the center of $X^\circ = \{(x, t) \mid Ax - dt \leq b\}$. That being the case, X° is bounded and so T_{MAX} and T_{MIN} are both finite. Conversely, if T_{MAX} and T_{MIN} are both finite, then the center (x^*, t^*) of X° , which is the solution to the problem

$$\begin{aligned} P^\circ: \quad & \underset{x, t}{\text{maximize}} \quad \sum_{i=1}^m \ln(b_i + d_i t - A_i x) \\ & \text{s.t.} \quad Ax - dt < b \end{aligned}$$

exists uniquely. The K-K-T conditions for P° require that $e^T \hat{S}_t^{-1} A = 0$ and $e^T \hat{S}_t^{-1} d = 0$, i.e., $f(t^*) = 0$, where $\hat{s}_t = b + dt - Ax^*$ and $\hat{S}_t = \text{diag}(\hat{s}_t)$. The rest follows from Proposition 2.1.

Figure 1 (a)

Figure 1 (b)

Figure 1 (c)

(ii) Suppose T_{MAX} is finite and $T_{\text{MIN}} = -\infty$. Let $\bar{t} < T_{\text{MAX}}$. We have seen earlier that $f(\bar{t}) > 0$ implies T_{MIN} is finite. Thus, $f(\bar{t}) \leq 0$. But, from (i), if $f(\bar{t}) = 0$ then T_{MIN} is finite, which is a contradiction. Thus, $f(\bar{t}) < 0$. Conversely, if $f(\bar{t}) < 0$ for some $\bar{t}$, then T_{MAX} is finite. Consequently, if $f(t) < 0$ for all $t < T_{\text{MAX}}$ then $T_{\text{MIN}} = -\infty$ follows from (i).

The proof of (iii) is similar to (ii). ■

2.2 Algorithm PCP

Before presenting the algorithm for the parametric center problem, we introduce some more notation and the main improvement theorem.

Recall that $X = \{x \in \mathbb{R}^n \mid Ax \leq b\}$

and $X_t = \{x \in \mathbb{R}^n \mid Ax \leq b + dt\}$.

We assume that X is bounded, and so A has full column rank and X_t is bounded for all t . Let $\|v\|$ denote the Euclidean norm. If M is a symmetric positive-definite matrix,

we denote $\|y - z\|_M = \sqrt{(y - z)^T M (y - z)}$. Let $\hat{x}$ be the center of X ,

let $\bar{x} \in \text{int } X$ be given, let $\bar{s} = b - A\bar{x}$, and let $Q(\bar{x}) = A^T \bar{S}^{-2} A$,

where $\bar{S} = \text{diag}(\bar{s})$. We say that $\bar{x}$ is close to the center of X if

$\|\bar{x} - \hat{x}\|_{Q(\bar{x})} \leq 1/21$. The motivation for this criterion of closeness to the

center will be apparent from the following theorem, which also serves as a basis for the PCP (Parametric Center Problem) algorithm.

Theorem 2.1 (Improvement Theorem). Suppose $\bar{x} \in \text{int } X = \{x \mid Ax < b\}$

Let $\bar{s} = b - A\bar{x}$, $\bar{S} = \text{diag}(\bar{s})$, and $Q = A^T \bar{S}^{-2} A$. Suppose $\|\bar{x} - \hat{x}\|_Q < \frac{1}{21}$, where $\hat{x}$ is the center of the system (A, b) . Define

$$\bar{u} = \left[I - \bar{S}^{-1} A (A^T \bar{S}^{-2} A)^{-1} A^T \bar{S}^{-1} \right] \bar{S}^{-1} d \quad (\text{projection of } \bar{S}^{-1} d)$$

$$\bar{r} = (A^T \bar{S}^{-2} A)^{-1} A^T \bar{S}^{-2} d \quad (\text{translation})$$

$$\alpha = \frac{1}{80 \|\bar{u}\|} \quad (\text{step length})$$

Furthermore, define $\bar{s}_\alpha = b + \bar{S} \bar{u} \alpha - A\bar{x}$,

$$\bar{S}_\alpha = \text{diag}(\bar{s}_\alpha)$$

$$\bar{q} = -(A^T \bar{S}_\alpha^{-2} A)^{-1} A^T \bar{S}_\alpha^{-1} e \quad (\text{Newton step})$$

$$\bar{x}_{\text{NEW}} = \bar{x} + \bar{q} + \alpha \bar{r} \quad (\text{New approximate center})$$

Then $\|\bar{x}_{\text{NEW}} - \hat{x}_\alpha\|_{\bar{Q}} \leq \frac{1}{21}$, where $\hat{x}_\alpha$ is the center of the system $(A, b + \alpha d)$, and $\bar{Q} = A^T \bar{S}_{\text{NEW}}^{-2} A$ where $\bar{S}_{\text{NEW}} = \text{diag}(b + \alpha d - A \bar{x}_{\text{NEW}})$. ■

In this theorem, $\bar{x}$ is given and is assumed to be close to the center $\hat{x}$ of X . The vectors $\bar{u}$ and $\bar{r}$ are defined and satisfy $\bar{S}^{-1} d = \bar{S}^{-1} A \bar{r} + \bar{u}$, so that $d = A \bar{r} + \bar{S} \bar{u}$. The increase in t from $t = 0$ to $t = \alpha$ is defined next, and is a function of $\|\bar{u}\|$. Because $d \neq A \bar{r}$ for any $\bar{r} \in \mathbb{R}^n$, $\|\bar{u}\| > 0$, and so α is well-defined. The Newton step $\bar{q}$ is then

defined using the modified slacks $\bar{s}_\alpha$, and $\bar{x}_{\text{NEW}}$ is composed, using $\bar{x}$, $\bar{q}$ (the Newton step) and $\bar{r}$ (a translation vector). According to the theorem, $\bar{x}_{\text{NEW}}$ will be suitably close to $\hat{x}_\alpha$, the center of X_α . This theorem will be proved in Section 4.

The increase in t from $t = 0$ is given as $\alpha = \frac{1}{80\|\bar{u}\|}$, which is a function of $\|\bar{u}\|$. Ideally, we would like this increase to be as large as possible, to speed algorithmic convergence. However, the fact that α is proportional to $1/\|\bar{u}\|$ makes good intuitive sense. The quantity $\|\bar{u}\|$ is the minimum (least-square) distance of d from the column range of A , in the weighted norm with weights $1/\bar{s}_i$, $i = 1, \dots, m$. Suppose d lies very close to the range space of A , i.e., $d = \bar{S}\bar{u} + Ar$, and so $\|\bar{u}\|$ is small. Then changing the RHS by αd is the same as translating X by αr and then changing the RHS by only $\alpha \bar{S}\bar{u}$. Then because $\|\bar{u}\|$ is small, we can take a big step. If, however, $\|\bar{u}\|$ is large, then "most of d " lies outside of the range space of A , so that a change in the RHS of αd is shifting the shape of X substantially, not just translating the polyhedron. Thus, the step we can take will be smaller.

However, as the next theorem indicates, the value of $\|\bar{u}\|$ also gives important bounds on the values of T_{MAX} and T_{MIN} , and thus will show that if α is small, so is T_{MAX} or T_{MIN} .

Recall that $f(t)$, the path indicator function, gives important information about the boundedness/unboundedness of the path of centers $\hat{x}_t$ in both directions, according to Proposition 2.2.

Theorem 2.2 (Bounds on T_{MAX} and T_{MIN}). Under the conditions and definitions of Theorem 2.1,

$$(i) \text{ if } e^{T\bar{u}} / ||\bar{u}|| > 1/20, \text{ then } T_{MIN} \geq \overline{LB} \equiv -\frac{22\sqrt{m(m+1)} + 1}{21||\bar{u}||}$$

$$(ii) \text{ if } e^{T\bar{u}} / ||\bar{u}|| < -1/20, \text{ then } T_{MAX} \leq \overline{UB} \equiv \frac{22\sqrt{m(m+1)} + 1}{21||\bar{u}||}$$

$$(iii) \text{ if } |e^{T\bar{u}} / ||\bar{u}||| \leq 1/20, \text{ then } T_{MIN} \geq \widetilde{LB} \equiv -\frac{1.6\sqrt{m(m-1)} + .6}{||\bar{u}||}$$

$$\text{and } T_{MAX} \leq \widetilde{UB} \equiv \frac{1.6\sqrt{m(m-1)} + .6}{||\bar{u}||} \quad \blacksquare$$

This theorem will be proved in Section 5.

Note that either (i), (ii), or (iii) must be satisfied, so that either a finite lower bound on T_{MIN} or a finite upper bound on T_{MAX} is produced in the course of increasing t from $t = 0$ to $t = \alpha$, or both are produced. Case (i) corresponds to being (approximately) on the lower path at $t = 0$, i.e., $f(0) \gtrsim 0$. Case (ii) corresponds to being (approximately) on the upper path, i.e., $f(0) \lesssim 0$. Case (iii) corresponds to $f(0) \approx 0$, so that $(\bar{x}, 0)$ is close to the center of X° .

Note also that these bounds are $O(m/||\bar{u}||)$. Thus, even though t is increased from $t = 0$ to $t = \alpha = 1/80||\bar{u}||$, the ratio of α to one or both bounds is at least $\frac{1}{128m}$. Therefore repeated increases in t using the methodology of Theorem 2.1 will result in either geometric growth

in the quantity $t - T_{\min}$, with a growth rate of at least $\left(1 + \frac{1}{128m}\right)$, or geometric contraction in the quantity $T_{\max} - t$, with a contraction rate of at least $\left(1 - \frac{1}{128m}\right)$, as the next proposition indicates.

Proposition 2.3 (Geometric Change in t Relative to $T_{\min}$ or $T_{\max}$).

Suppose $m \geq 2$. With the notation $\bar{x}_{\text{OLD}} \equiv \bar{x}$, $t_{\text{OLD}} = 0$, $t_{\text{NEW}} = \alpha$, where α and $\bar{x}_{\text{NEW}}$ are defined as in Theorem 2.1, then either

$$(i) \quad (t_{\text{NEW}} - T_{\min}) \geq \left(1 + \frac{1}{128m}\right)(t_{\text{OLD}} - T_{\min}) \quad , \quad \text{or}$$

$$(ii) \quad (T_{\max} - t_{\text{NEW}}) \leq \left(1 - \frac{1}{128m}\right)(T_{\max} - t_{\text{OLD}}) \quad .$$

Proof: Suppose a lower bound, either $\bar{LB}$ or $\widetilde{LB}$, is generated through Theorem 2.2. Notice that $\widetilde{LB} \leq \bar{LB}$ if $m \geq 2$, so in either case,

$$T_{\min} \geq \widetilde{LB} = -(1.6\sqrt{m(m-1)} + .6)/\|\bar{u}\| \geq -1.6m/\|\bar{u}\| \quad . \quad \text{Thus}$$

$$\frac{(t_{\text{NEW}} - T_{\min})}{t_{\text{OLD}} - T_{\min}} = \frac{t_{\text{NEW}}}{-T_{\min}} + 1 \geq -\left(\frac{1}{80\|\bar{u}\|}\right) \left/\left(\frac{-1.6m}{\|\bar{u}\|}\right)\right. + 1 = 1 + \frac{1}{128m} \quad .$$

A parallel analysis demonstrates (ii) if an upper bound is generated. ■

Now let us return to our initial interest – to trace the parametric center path $\hat{x}_t$ for the program $P(t)$. Suppose we want to trace the path as t ranges in an interval $t \in [\underline{t}, \bar{t}]$, where $\bar{t} \geq \underline{t}$. Suppose we are given a

point x° that lies close to the center of X_t . (Such a point can be found by using the algorithm in Vaidya [15] or in [4].)

The following algorithm, denoted PCP for Parametric Center Problem, is an iterative algorithm that invokes Theorems 2.1 and 2.2. At Step 0, initial lower and upper bounds are set to their extreme values. The initial value of t is chosen as $t^\circ = \underline{t}$, and the counter k (for the number of iterations) is set equal to zero. The current value of t is t^k . The current value $\bar{x}$ is set, as is the RHS. In Step 1, the values $\bar{u}$ and $\bar{r}$ are computed as defined in Theorem 2.1. In Step 2, the constant α , which is the increase in t at the iteration, is defined as in Theorem 2.1. The next "close-to-center" point, $\bar{x}_{\text{NEW}}$, is then computed; $\bar{x}_{\text{NEW}}$ is close to the center of $X_{t^k + \alpha}$, i.e., $\bar{x}_{\text{NEW}}$ approximately solves $P(t^k + \alpha)$, according to Theorem 2.1. In Step 3, a piecewise-linear path $\bar{x}(t)$ is defined in the range $t \in [t^k, t^k + \alpha] = [t^k, t^{k+1}]$, using $\bar{x}(t^k) = \bar{x}$ and $\bar{x}(t^{k+1}) = \bar{x}_{\text{NEW}}$ as endpoints and interpolating. In Step 4, the current bounds on T_{MIN} and T_{MAX} are updated, in accordance with Theorem 2.2. In Step 5, the algorithm checks if $t^{k+1} \geq \bar{t}$. If so, it stops. If not, it returns to Step 1. The output of the algorithm is the piecewise-linear path $\bar{x}(t)$ and the incremental values of t , namely $t^\circ, t^1, \dots, t^k, \dots$.

Algorithm PCP (Parametric Center Problem)

Input: $A \in \mathbb{R}^{m \times n}$, $b, d \in \mathbb{R}^m$, $\underline{t}, \bar{t}, x^\circ$.

Step 0 (Initialization)

Set $UB = +\infty$, $LB = -\infty$, $k = 0$, $t^\circ = \underline{t}$, $\bar{x} = x^\circ$, $\text{RHS} = b + d\underline{t}$.

Step 1 (Projection of d)

Set $\bar{s} = \text{RHS} - A\bar{x}$, $\bar{S} = \text{diag}(\bar{s})$.

Compute $\bar{u} = [I - \bar{S}^{-1} A (A^T \bar{S}^{-2} A)^{-1} A^T \bar{S}^{-1}] \bar{S}^{-1} d$;

$$\bar{r} = (A^T \bar{S}^{-2} A)^{-1} A^T \bar{S}^{-2} d .$$

Step 2 (Compute Step Length and Compute New Approximate Center)

Set $\alpha = \frac{1}{80 \|\bar{u}\|}$;

$$\bar{s}_\alpha = \text{RHS} + \alpha \bar{S} \bar{u} - A\bar{x} , \quad \bar{S}_\alpha = \text{diag}(\bar{s}_\alpha) ;$$

$$\bar{q} = -(A^T \bar{S}_\alpha^{-2} A)^{-1} A^T \bar{S}_\alpha^{-1} e ;$$

$$\bar{x}_{\text{NEW}} = \bar{x} + \alpha \bar{r} + \bar{q} .$$

Step 3 (Extend Piecewise-Linear Path)

Set $t^{k+1} = t^k + \alpha$.

For $t^k \leq t \leq t^{k+1}$, define $\bar{x}(t) = \bar{x} + \left[\frac{(t - t^k)}{\alpha} \right] (\bar{x}_{\text{NEW}} - \bar{x})$.

Step 4 (Update Lower and Upper Bounds)

(i) if $e^T \bar{u} / \|\bar{u}\| > 1/20$, then $\text{LB} = \max \{ \text{LB}, t^k - (22 \sqrt{m(m+1)} + 1) / (21 \|\bar{u}\|) \}$;

(ii) if $e^T \bar{u} / \|\bar{u}\| < -1/20$, then $UB = \min \{UB, t^k + (22 \sqrt{m(m+1)} + 1) / (21 \|\bar{u}\|)\}$;

(iii) if $|e^T \bar{u} / \|\bar{u}\|| \leq 1/20$, then $LB = \max \{LB, t^k - (1.6 \sqrt{m(m-1)} + .6) / \|\bar{u}\|\}$

and $UB = \min \{UB, t^k + (1.6 \sqrt{m(m-1)} + .6) / \|\bar{u}\|\}$.

Step 5 If $t^{k+1} < \bar{t}$, set $RHS = RHS + d\alpha$, $k = k + 1$, $\bar{x} = \bar{x}_{NEW}$, and go to Step 1.

If $t^{k+1} \geq \bar{t}$, STOP.

2.3 Algorithmic Performance

According to Theorem 2.1, if x° is close to the center of X_t , then for each $k = 1, 2, \dots$, $\bar{x}(t^k)$ will be close to the center of X_{t^k} . Thus the break points of the piecewise-linear path $\bar{x}(t)$ will be near the parameterized center. Furthermore, we will prove in Section 4:

Lemma 2.4 For all values of t generated by algorithm PCP, $\bar{x}(t)$ is near the center X_t , in the sense that $\|\bar{x}(t) - \hat{x}_t\|_Q \leq .585$, where $Q = A^T \bar{S}_t^{-2} A$, and $\bar{s}_t = b + dt - A\bar{x}(t)$, $\bar{S}_t = \text{diag}(\bar{s}_t)$. ■

We are now ready to discuss the performance of the algorithm.

According to Proposition 2.3, we obtain at each iteration either a geometric decrease in the gap $T_{MAX} - t$ at each iteration, or a geometric increase in the gap $t - T_{MIN}$. We thus can measure algorithmic performance according to the change in $T_{MAX} - t$, or $t - T_{MIN}$, or both, depending on whether we are approximating the upper path ($f(t) \leq 0$), the lower path ($f(t) \geq 0$), or both. Suppose that in the course of running algorithm

PCP , that a lower bound on $T_{\min}$ is never generated. Then all iterates will satisfy criterion (ii) or (iii) at Step 4, so that all iterates will generate an upper bound, and all iterates will lie approximately on the upper path.

Lemma 2.5 (Algorithm Performance Based Only on $T_{\max}$) If

$T_{\min} = -\infty$ or if no iterates of the algorithm PCP generate a lower bound on $T_{\min}$, then the sequence of t values will satisfy

$$T_{\max} - t^k \leq \left(1 - \frac{1}{128m}\right)^k (T_{\max} - \underline{t}) .$$

In particular, if $\bar{t} < T_{\max}$, the algorithm will stop after at most $K = \left\lceil 128m \ln \left((T_{\max} - \underline{t}) / (T_{\max} - \bar{t}) \right) \right\rceil$ iterations.

Proof: Under the hypothesis of the lemma, the algorithm must satisfy either criterion (ii) or (iii) at Step 4. Thus, by Proposition 2.3, $T_{\max} - t^{k+1} \leq \left(1 - \frac{1}{128m}\right) (T_{\max} - t^k)$. Thus we obtain the geometric decrease of the Lemma. If $\bar{t} < T_{\max}$, let $K = \left\lceil 128m \ln \left((T_{\max} - \underline{t}) / (T_{\max} - \bar{t}) \right) \right\rceil$.

$$\text{Then } \ln (T_{\max} - t^K) \leq K \ln \left(1 - \frac{1}{128m}\right) + \ln (T_{\max} - \underline{t})$$

$$\leq -K \left(\frac{1}{128m}\right) + \ln (T_{\max} - \underline{t})$$

$$\leq -(\ln ((T_{\max} - \underline{t}) / (T_{\max} - \bar{t}))) + \ln (T_{\max} - \underline{t}) .$$

Thus, $T_{\max} - t^K \leq T_{\max} - \bar{t}$, whereby $t^K \geq \bar{t}$. Thus the algorithm will stop. ■

Suppose instead that none of the iterates of the algorithm generate an upper bound at Step 4. Analogous to Lemma 2.5 we have:

Lemma 2.6 (Algorithm Performance Based Only on T_{MIN}) If $T_{\text{MAX}} = +\infty$ or if none of the iterates of algorithm PCP generate an upper bound on T_{MAX} at Step 4, then the sequence of t values will satisfy

$$t^k - T_{\text{MIN}} \geq \left(1 + \frac{1}{128m}\right)^k (t - T_{\text{MIN}}).$$

In particular, the algorithm will stop after $K = \left\lceil 128m \ln ((\bar{t} - T_{\text{MIN}}) / (t - T_{\text{MIN}})) \right\rceil$ iterations. ■

We next examine the case when the algorithm generates both upper and lower bounds. We first need the following result, which will be proved in Section 5.

Lemma 2.7 If criterion (ii) of Step 4 of the algorithm PCP is satisfied at iteration k , then in all subsequent iterations, criteria (ii) or (iii) of Step 4 will be satisfied. ■

The significance of Lemma 2.7 is as follows: if at iteration k an upper bound is generated, then an upper bound is generated at every subsequent iteration.

Lemma 2.8 (Algorithm Performance Based on Lower and Upper Bounds) If the algorithm PCP generates both lower and upper bounds, then there is some

iterate j such that for $k \leq j$, the algorithm generates lower bounds only and for all $k > j$ the algorithm generates upper bounds, and

$$(i) \quad \text{for all } k \leq j, \quad t^k - T_{\min} \geq \left(1 + \frac{1}{128m}\right)^k (\underline{t} - T_{\min})$$

$$(ii) \quad \text{for all } k \geq j, \quad T_{\max} - t^k \leq \left(1 - \frac{1}{128m}\right)^{k-j} (T_{\max} - t^j).$$

Furthermore, if $\bar{t} < T_{\max}$, then the algorithm will stop after at most

$$K = \left\lceil 256m \ln \left(\frac{T_{\max} - T_{\min}}{2} \right) - 128m \ln (\underline{t} - T_{\min}) - 128m \ln (T_{\max} - \bar{t}) \right\rceil + 2$$

iterations.

Proof: The existence of j is guaranteed by Lemma 2.7. The geometric convergence rates are then a consequence of Proposition 2.3. Finally, suppose $\bar{t} < T_{\max}$, and let K be as defined above. Let $\hat{t} = t^j$, and note that

$$.5 (\ln (\hat{t} - T_{\min}) + \ln (T_{\max} - \hat{t})) \leq \ln \left(\frac{T_{\max} - T_{\min}}{2} \right),$$

from the arithmetic-geometric mean inequality. Thus,

$$\begin{aligned} K &\geq 128m \ln (\hat{t} - T_{\min}) + 128m \ln (T_{\max} - \hat{t}) \\ &\quad - 128m \ln (T_{\max} - \bar{t}) - 128m \ln (\underline{t} - T_{\min}) \end{aligned}$$

$$= 128m \ln \left(\frac{\hat{t} - T_{\text{MIN}}}{\underline{t} - T_{\text{MIN}}} \right) + 128m \ln \left(\frac{T_{\text{MAX}} - \hat{t}}{T_{\text{MAX}} - \underline{t}} \right).$$

According to Lemma 2.5, with $\bar{t}$ replaced by $\hat{t}$, $t^k \geq \hat{t}$ after at most $128m \ln \left(\frac{\hat{t} - T_{\text{MIN}}}{\underline{t} - T_{\text{MIN}}} \right)$ iterations.

Furthermore, according to Lemma 2.6, with $\underline{t}$ replaced by $\underline{t}^{j+1}$, $t^k \geq \bar{t}$ after at most K iterations. ■

2.4 A Strictly Monotone Algorithm for Linear Programming that requires $O(mL)$ iterations.

Suppose we wish to solve the problem

$$\begin{aligned} \widetilde{\text{LP}}: \quad & \text{maximize} \quad \tilde{c}^T \tilde{x} \\ & \text{s. t.} \quad \tilde{A}\tilde{x} \leq \tilde{b} \end{aligned}$$

where $\tilde{x}, \tilde{c} \in \mathbb{R}^{n+1}$, $\tilde{A} \in \mathbb{R}^{m \times (n+1)}$, and $\tilde{b} \in \mathbb{R}^m$. Upon setting $t = \tilde{c}^T \tilde{x}$ and eliminating one of the $(n+1)$ variables of $\tilde{x}$, $\widetilde{\text{LP}}$ is easily transformed to the form

$$\begin{aligned} \text{LP}: \quad & \text{maximize} \quad t \\ & \text{s. t.} \quad Ax \leq b + dt \end{aligned}$$

where $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and the data (A, b) are a linear transformation of the data $(\tilde{A}, \tilde{b}, \tilde{c})$. We can solve LP by using the algorithm PCP to trace the path $\bar{x}(t)$ of center to the parametric problem $Ax \leq b + dt$.

Suppose $x^\circ \in \mathbb{R}^n$ is a given starting point for which $(x^\circ, t^\circ) = (x^\circ, c^T x^\circ)$ satisfies the starting criterion of the algorithm PCP, namely $\|x^\circ - \hat{x}_t\|_{Q^\circ} \leq \frac{1}{21}$, where $s^\circ = b + dt^\circ - A x^\circ > 0$, $S^\circ = \text{diag}(s^\circ)$

and $Q^\circ = A^T (S^\circ)^{-2} A$. Then we can use algorithm PCP to generate the path $\bar{x}(t)$ for $t \in [t^\circ, T_{\text{MAX}}] = [c^T x^\circ, z^*]$ where z^* is the optimal objective value of the linear program $\widetilde{\text{LP}}$. The sequence of values of t^k , $k = 0, \dots$, will be strictly increasing, according to Theorem 2.1, i.e., the objective value will be strictly increasing at each iteration. Let L be the total number of bits in a binary encoding of LP . In order to evaluate the algorithm's complexity, we consider three cases.

Case (i): The linear program is unbounded. In this case, the algorithm will never generate a finite upper bound on T_{MAX} , which equals infinity.

After $k = O(mL)$ iterations, $t^k \geq (t^\circ - T_{\text{MIN}}) \left(1 + \frac{1}{128m}\right)^k$ will exceed

2^L , and we can conclude that LP is unbounded.

Case (ii): The linear program is bounded and $f(t^\circ)$, the value of the path indicator function at $t = t^\circ$, is negative. This being the case, the algorithm will always generate upper bounds, and after $O(mL)$ iterations,

$(z^* - t^k) \leq (z^* - t^\circ) \left(1 - \frac{1}{128m}\right)^k$ is less than 2^{-L} , whereby $\bar{x}(t^k)$ can be

rounded to an optimal solution, see Karmarkar [6].

Case (iii): The linear program is bounded and $f(t^\circ) > 0$. In this case, one can show as in case (i) that after $\bar{k} = O(mL)$ iterations, that $f(t^k) \leq 0$, for otherwise the LP would be unbounded. Furthermore, after an additional

$\bar{k} = O(mL)$ iterations, we will obtain via case (ii) that we can round to an optimal solution. Thus, after $O(mL)$ iterations, we can round to an optimal solution.

Note that in either of the three cases, that algorithm PCP will process LP (by detecting unboundedness or producing an optimal solution) after $O(mL)$ iterations.

This algorithm falls into the class of central-trajectory based algorithms, but is inferior in that the bound of $O(mL)$ iterations is worse than the bound of $O(\sqrt{m}L)$ iterations for algorithms such as Renegar [11] or Vaidya [16] that trace the (weighted) center of the system

$$Ax \leq b$$

$$cx \geq \delta$$

as δ is increased, or to the bound of $O(\sqrt{m}L)$ iterations for algorithms based on barrier penalty methods that trace the solution to

$$\begin{aligned} &\text{maximize} && c^T x + \varepsilon \sum_{i=1}^m \ln s_i \\ &\text{s.t.} && Ax + s = b \\ &&& s > 0, \end{aligned}$$

see Monteiro and Adler [10], among others.

All three methods follow the same path in their idealized version. Yet the latter two obtain convergence in $O(\sqrt{m}L)$ iterations, which is superior to our algorithm. However, these other algorithms do not guarantee strict improvement in the objective value, (but do guarantee strict improvement in the duality gap). In contrast, our algorithm will guarantee strict improvement in the objective function of $(z^* - c^T \bar{x})/128m$ at each

iteration. Perhaps it is the implicit imposition of the strict improvement in objective value that increases the iteration bound by a factor of $O(\sqrt{m})$.

Furthermore, our algorithm does not assume that LP is bounded.

Instead, our algorithm will detect unboundedness of LP directly.

As a final note, note that our algorithm can be used to mimic Renegar's algorithm [11], tracing the center of

$$\begin{aligned} Ax &\leq b \\ -cx &\leq -t \end{aligned}$$

as t is increased. Thus, Renegar's set-up is a special case of the problem PCP we are considering. However, we see no way to cast problem (2.2) as a special case of the set-up used by Renegar [11]; we allow all RHS values to vary simultaneously, which is apparently more general than in his work [11].

The remainder of the paper is devoted to proofs of the results presented in this section. In Section 3, we present notation and preliminaries. Section 4 contains an analysis of a single step of the algorithm PCP , and contains proofs of Theorem 2.1 and Lemma 2.4. Section 5 contains an analysis of bounds generated by the algorithm, and contains proofs of Theorems 2.2 and Lemma 2.7

3. Notation and Preliminary Results

In this section we present notation and some preliminary results that will be used in the proofs of Theorems 2.1 and 2.2 and in subsequent analysis.

3.1 Notation and Translations

For a vector v , $\|v\|$ denotes the Euclidean norm, and for a matrix M , $\|M\|$ denotes the usual matrix norm, i. e.,

$$\|M\| = \sup_{v \neq 0} \|Mv\|/\|v\|.$$

Note that if M is a diagonal matrix, $\|M\| = \max_i |m_{ii}|$.

If M is a positive definite matrix, the M -norm of v is $\|v\|_M = \sqrt{v^T M v}$.

The matrix $P_M := I - M(M^T M)^{-1} M^T$ denotes the orthogonal projection matrix which projects onto the null space of M^T . Let $Q_t(x; u)$ denote the negative of the Hessian of the function

$f_t(x; u) := \sum_{i=1}^m \ln(b_i + tu_i - A_i x)$, where $u \in \mathbb{R}^m$ is a given vector parameter.

Let $\Delta(x) := \text{diag}(b - Ax)$ and $\Delta_t(x; u) := \text{diag}(b + tu - Ax)$.

Then $Q_t(x; u) = A^T \Delta_t^{-2}(x; u) A$. When $t = 0$, we denote $Q(x) = A^T \Delta^{-2}(x) A$.

Let $\hat{x}_t(u)$ denote the center of the system $Ax \leq b + tu$, and let $\hat{x}$ denote the center of the system $Ax \leq b$.

Suppose $d = u + Ar$ for some $u \in \mathbb{R}^m$ and $r \in \mathbb{R}^n$. We observe that $\Delta_t(x; d) = \Delta_t(x - tr; u)$ and hence $Q_t(x; d) = Q_t(x - tr; u)$. Thus, the difference between modifying the RHS by d and by u simply corresponds to a translation of the inequality system by $-tr$, in that

$$\{x \in \mathbb{R}^n \mid Ax \leq b + td\} = \{x \in \mathbb{R}^n \mid Ax \leq b + tu\} + tr.$$

The following Lemma is therefore obvious.

Lemma 3.1: Suppose $d = u + Ar$ for some $u \in \mathbb{R}^m$ and $r \in \mathbb{R}^n$. Then

$$(i) \quad \hat{x}_t(d) = \hat{x}_t(u) + tr,$$

$$(ii) \quad \|x - \hat{x}_t(d)\|_{Q_t(\hat{x}_t(d); d)} = \|(x - tr) - \hat{x}_t(u)\|_{Q_t(\hat{x}_t(u); u)} \text{ for any } x \in \mathbb{R}^n, \text{ and}$$

$$(iii) \quad \max \{t \mid Ax \leq b + td \text{ for some } x\} = \max \{t \mid Ax \leq b + tu \text{ for some } x\}. \blacksquare$$

In the sequel, we shall be working with appropriate choices of $u \in \mathbb{R}^m$ and $r \in \mathbb{R}^n$ instead of d . As we shall see in Section 5, this in fact is central to the construction of the proof of Theorem 2.2. For the appropriate $u \in \mathbb{R}^m$, where convenient and the context is clear,

$$\text{we let } \Delta_t(x) := \Delta_t(x; u), \quad \Delta(x) = \Delta_0(x), \quad \hat{x}_t = \hat{x}_t(u), \quad \bar{s} = b - A\bar{x},$$

$$\bar{S} = \text{diag}(\bar{s}) , \quad \hat{s} = b - A\hat{x} , \quad \hat{S} = \text{diag}(\hat{s}) , \quad \bar{s}_t = b + tu - A\bar{x} ,$$

and $\bar{S}_t = \text{diag}(\bar{s}_t)$. We also will abbreviate $Q_t(x; u)$ by $Q_t(x)$.

The next Lemma presents some basic inequalities. It is essentially Proposition 7.2 of [4] with some simple extensions.

Lemma 3.2: Suppose $\bar{x} \in \mathbb{R}^n$ satisfies $\bar{s} = b - A\bar{x} > 0$ and let $Q(\bar{x}) = A^T \bar{S}^{-2} A$. Then for any $\tilde{x} \in \mathbb{R}^n$ such that $\|\tilde{x} - \bar{x}\|_{Q(\bar{x})} \leq \delta < 1$, we have

$$(i) \quad \tilde{s} = b - A\tilde{x} > 0 ,$$

$$(ii) \quad \|\Delta^{-1}(\tilde{x}) \Delta(\bar{x})\| = \|\tilde{S}^{-1} \bar{S}\| \leq \frac{1}{1 - \delta} ,$$

$$(iii) \quad \|\Delta^{-1}(\bar{x}) \Delta(\tilde{x})\| = \|\bar{S}^{-1} \tilde{S}\| \leq 1 + \delta ,$$

and for any $v \in \mathbb{R}^n$,

$$(iv) \quad \|v\|_{Q(\tilde{x})} \leq \frac{1}{1 - \delta} \|v\|_{Q(\bar{x})} , \quad \text{where } Q(\tilde{x}) = A^T \tilde{S}^{-2} A , \text{ and}$$

$$(v) \quad \|v\|_{Q(\bar{x})} \leq (1 + \delta) \|v\|_{Q(\tilde{x})} . \quad \blacksquare$$

3.2 Equivalent Measures of Closeness

In this study, we measure how close a point $\bar{x}$ is to the center $\hat{x}$ of the system $Ax \leq b$ with the norm $\|\bar{x} - \hat{x}\|_{Q(\bar{x})}$. We shall also make

use of a different measure of closeness to $\hat{x}$ that was introduced in [4] , and we will show a certain equivalence of the two measures.

Define

$$Q = \frac{1}{m} A^T \bar{S}^{-2} A, \quad y = y(\bar{x}) = \frac{1}{m} A^T \bar{S}^{-1} e \quad (3.1a)$$

$$\text{and } \gamma = \gamma(\bar{x}) = \sqrt{\frac{(m-1)y^T Q^{-1} y}{1 - y^T Q^{-1} y}}. \quad (3.1b)$$

Note that $Q = \frac{1}{m} Q(\bar{x})$. Let $\hat{x}$ denote the center of the system $Ax \leq b$. Then $y(\hat{x}) = 0$ and so $\gamma(\hat{x}) = 0$. In [4], the scalar $\gamma = \gamma(\bar{x})$ is used to measure the closeness of $\bar{x}$ to the center $\hat{x}$.

Lemma 3.3 ([4]). Let $\hat{x}$ be the center of the system $Ax \leq b$.

Let $h > 0$ be a given parameter.

Suppose $\gamma = \gamma(\bar{x}) \leq g\left(\frac{\ln(1+h)}{h}\right)$, where $g(\alpha) = 1 - (\alpha + \sqrt{(\alpha-1)^2 + 1})/2$.

Then $\|\hat{x} - \bar{x}\|_{Q(\bar{x})}^2 \leq \left(\frac{m}{m-1}\right) \frac{h^2(1+\gamma^2)}{(1-h\gamma)^2}$.

Proof: Follows from the proof Lemma 7.2 of [4] with weights $w = (1/m)e$.

■

We shall say that $\bar{x}$ is approximately centered if $\gamma(\bar{x}) \leq .0072$.

Corollary 3.1: If $\gamma(\bar{x}) \leq .0072$, then $\|\bar{x} - \hat{x}\|_{Q(\bar{x})} < 1/21$.

Proof: Let $h = .03$ and $\alpha = \frac{\ln(1+h)}{h}$. Then from Lemma 3.3,

$$\text{we have } \|\bar{x} - \hat{x}\|_{Q(\bar{x})}^2 \leq \frac{m}{m-1} \left(\frac{h^2(1+\gamma^2)}{(1-h\gamma)^2} \right).$$

Substituting for h and γ , and noting $m \geq 2$ gives the desired result. ■

Therefore, for appropriate values of h , (e.g., $h = .03$) $\bar{x}$ is approximately centered implies $\bar{x}$ is close to the center by our criterion, that is $\|\hat{x} - \bar{x}\|_{Q(\bar{x})} < 1/21$. Next, we show the converse implication.

Lemma 3.4: Suppose $\|\hat{x} - \bar{x}\|_{Q(\bar{x})} \leq \delta < 1/2$, where $\hat{x}$ is the center of the system $Ax \leq b$.

$$\text{Then } \bar{\gamma} = \gamma(\bar{x}) \leq a + \sqrt{2a}, \text{ where } a = \frac{\delta^2}{2(1-\delta)(1-2\delta)}.$$

Proof: Let $\bar{s} = b - A\bar{x}$ and $\hat{s} = b - A\hat{x}$. From Lemma 3.2,

$$\|\hat{S}^{-1}(\hat{s} - \bar{s})\| = \|\hat{x} - \bar{x}\|_{Q(\hat{x})} \leq \frac{1}{1-\delta} \|\hat{x} - \bar{x}\|_{Q(\bar{x})} \leq \frac{\delta}{1-\delta} < 1.$$

From Lemma 2.1 of [4], with weights $w = (1/m)e$,

$$\sum_{i=1}^m \ln \hat{s}_i - \sum_{i=1}^m \ln \bar{s}_i \leq \left(\frac{m}{m-1} \right) \frac{\alpha^2}{2(1-\alpha)}, \text{ where } \alpha = \frac{\delta}{1-\delta} < 1.$$

On the other hand, from (2.4) of [4],

$$\sum_{i=1}^m \ln \hat{s}_i - \sum_{i=1}^m \ln \bar{s}_i \geq \left(\frac{m}{m-1} \right) (1 + \bar{\gamma} - \sqrt{1 + 2\bar{\gamma}}) , \quad \text{where } \bar{\gamma} = \gamma(\bar{x}) .$$

Therefore, $a = \frac{\alpha^2}{2(1-\alpha)} \geq (1 + \bar{\gamma} - \sqrt{1 + 2\bar{\gamma}}) .$

Thus, $\bar{\gamma} \leq a + \sqrt{2a} , \quad \text{where } a = \frac{\alpha^2}{2(1-\alpha)} = \frac{\delta^2}{2(1-\delta)(1-2\delta)} .$ ■

Finally, we present some elementary inequalities.

Lemma 3.5: Assume $m \geq 2$. Let Q, y and γ be defined as in (3.1) .

- (i) For all $\varepsilon \geq 0$, $\gamma^2 \leq \varepsilon$ implies $(m-1)y^T Q^{-1} y \leq \varepsilon$
- (ii) For all $\varepsilon < 1$, $(m-1)y^T Q^{-1} y \leq \varepsilon$ implies $\gamma^2 \leq \varepsilon/(1-\varepsilon)$.

Proof: Note that $\gamma^2 = \frac{(m-1)y^T Q^{-1} y}{1-y^T Q^{-1} y}$ whereby $y^T Q^{-1} y = \frac{\gamma^2}{m-1+\gamma^2}$.

Since $\frac{\gamma^2}{m-1+\gamma^2} \leq \frac{\gamma^2}{m-1}$, (i) follows immediately.

To prove (ii), note that $(m-1)y^T Q^{-1} y \leq \varepsilon$ implies

$$\gamma^2 \leq \frac{\varepsilon}{1-\varepsilon/(m-1)} \leq \frac{\varepsilon}{1-\varepsilon} .$$
 ■

4. Analysis of One Iteration of the Algorithm

In this section, we analyze one iteration of the algorithm and prove the Improvement Theorem (Theorem 2.1) and Lemma 2.4. First, we show that if $|t|$ is small then the two centers $\hat{x}$ and $\hat{x}_t$ are sufficiently close to each other with respect to some appropriate norm (so that Newton's method, when applied, will converge).

Theorem 4.1: Let $\hat{x}$ denote the center of the inequality system $Ax \leq b$ and let $\hat{x}_t$ denote the center of the inequality system $Ax \leq b + tu$. Let $\hat{u} = \Delta^{-1}(\hat{x})u$. Suppose $|t| \leq 1/(76\|\hat{u}\|)$. Then $\|\hat{x} - \hat{x}_t\|_{Q(\hat{x})} < 1/12$.

Proof: The proof makes use of Lemma 3.3 of the previous section. We want to show that the quantity $\gamma = \gamma(\hat{x})$ for the system $Ax \leq (b + tu)$ is sufficiently small. We begin by giving expressions for y and Q . First note that $A^T \Delta^{-1}(\hat{x})e = 0$ (see (2.3)), and so for the system $Ax \leq (b + tu)$,

$$\begin{aligned} y = y(\hat{x}) &:= \frac{1}{m} A^T \Delta_t^{-1}(\hat{x})e = \frac{1}{m} A^T [\Delta_t^{-1}(\hat{x})e - \Delta^{-1}(\hat{x})e] \\ &= \frac{-1}{m} A^T \Delta_t^{-1}(\hat{x}) \Delta^{-1}(\hat{x}) [\Delta_t(\hat{x})e - \Delta(\hat{x})e] \\ &= -\frac{1}{m} A^T \Delta_t^{-1}(\hat{x}) \Delta^{-1}(\hat{x}) (tu) \\ &= -\frac{1}{m} A^T \Delta_t^{-1}(\hat{x}) (t\hat{u}), \text{ by definition of } \hat{u} \text{ above.} \end{aligned}$$

$$\text{Also, } Q = \frac{1}{m} Q_t(\hat{x}) = \frac{1}{m} A^T \Delta_t^{-2}(\hat{x}) A,$$

and so ,

$$y^T Q^{-1} y = \frac{1}{m} (t \hat{u})^T \Delta_t^{-1}(\hat{x}) A \left(A^T \Delta_t^{-2}(\hat{x}) A \right)^{-1} A^T \Delta_t^{-1}(\hat{x}) (t \hat{u}) \leq \frac{1}{m} \|t \hat{u}\|^2$$

because the eliminated matrix is a projection matrix. Hence

$$\gamma^2 = \gamma^2(\hat{x}) = \frac{y^T Q^{-1} y (m-1)}{1 - y^T Q^{-1} y} \leq \frac{\|t \hat{u}\|^2}{1 - \|t \hat{u}\|^2} \leq \frac{1}{5,775} \text{ and } \gamma \leq .0132 .$$

Thus, from Lemma 3.3 with $h = 1/18$,

$$\|\hat{x} - \hat{x}_t\|_{Q_t(\hat{x})}^2 \leq \frac{2h^2(1 + \gamma^2)}{(1 - h\gamma)^2} < \left(\frac{1}{12}\right)^2 .$$

Corollary 4.1: Under the conditions of Theorem 4.1,

$$\|\hat{x} - \hat{x}_t\|_{Q_t(\hat{x})} < 1/11$$

Proof: Follows from Lemma 3.2(iv). ■

Furthermore, using Lemma 3.2(ii), the following corollary is immediate.

Corollary 4.2: Suppose $\|\bar{x} - \hat{x}\|_{Q(\bar{x})} \leq 1/21$.

Let $d \in \mathbb{R}^m$ be given and define $\bar{u} := P_{\bar{S}^1_A} \bar{S}^{-1} d$.

Let $\hat{x}_t$ denote the center of the inequality system $Ax \leq b + t \bar{S} \bar{u}$.

Suppose $|t| \leq \frac{1}{80 \|\bar{u}\|}$. Then $t \|\Delta^{-1}(\hat{x}) \bar{S} \bar{u}\| \leq 1/76$ and $\|\hat{x} - \hat{x}_t\|_{Q_t(\hat{x}_t)} < 1/11$.

Proof: Let $u = \bar{S} \bar{u}$ and $\hat{u} = \Delta^{-1}(\hat{x}) u$. Note then that $\hat{u} = \Delta^{-1}(\hat{x}) \bar{S} \bar{u}$, so that

$$\|\hat{u}\| \leq \|\Delta^{-1}(\hat{x}) \bar{S}\| \|\bar{u}\| \leq \frac{21}{20} \|\bar{u}\|,$$

from Lemma 3.2(ii). Thus,

$$|t| \leq \frac{1}{80 \|\bar{u}\|} \leq \frac{21}{20} \cdot \frac{1}{80 \|\hat{u}\|} \leq \frac{1}{76 \|\hat{u}\|}.$$

Hence, from Theorem 4.1, we have $\|\hat{x} - \hat{x}_t\|_{Q_t(\hat{x})} < 1/12$, and by Corollary 4.1, $\|\hat{x} - \hat{x}_t\|_{Q_t(\hat{x}_t)} < 1/11$. ■

We next use a theorem of Renegar [11] which gives the region and rate of convergence of Newton's method for our problem.

Theorem 4.2 (Renegar [11]): Assume

$$\bar{x} \in \text{int } X_t = \{x \in \mathbb{R}^n \mid Ax < b + tu\} \quad \text{and}$$

$$\varepsilon := \|\bar{x} - \hat{x}_t\|_{Q_t(\hat{x}_t)} < 1, \text{ where } \hat{x}_t \text{ is the center of the system } Ax \leq (b + tu).$$

Let $\bar{s}_t = b + ut - A\bar{x}$, $\bar{S}_t = \text{diag}(\bar{s}_t)$ and $\bar{x}_t = \bar{x} - Q_t^{-1}(\bar{x}) A^T \bar{S}_t^{-1} e$, where $Q_t = A^T \bar{S}_t^{-2} A$.

Then $\|\bar{x}_t - \hat{x}_t\|_{Q_t(\hat{x}_t)} \leq \frac{(1 + \varepsilon)^2 \varepsilon^2}{1 - \varepsilon}$. ■

We are now ready to prove the Improvement Theorem. For the reader's convenience, we restate the theorem before proving it.

Theorem 2.1 (Improvement Theorem):

Suppose $\bar{x} \in \text{int } X = \{x \mid Ax < b\}$ satisfies $\|\bar{x} - \hat{x}\|_{Q(\bar{x})} \leq 1/21$.

Define $\bar{u} := P_{\bar{S}^{-1}A} \bar{S}^{-1}d$ (projection of $\bar{S}^{-1}d$) ;

$\bar{r} := (A^T \bar{S}^{-2} A)^{-1} A^T \bar{S}^{-2} d$ (translation vector of system) ;

$\alpha := \frac{1}{80 \|\bar{u}\|}$ (step length) .

Further, define $\bar{s}_\alpha := b + \alpha \bar{S} \bar{u} - A\bar{x}$, $\bar{S}_\alpha := \text{diag}(\bar{s}_\alpha)$;

$\bar{q} := -(A^T \bar{S}_\alpha^{-2} A)^{-1} A^T \bar{S}_\alpha^{-1} e$ (Newton step) ;

$\bar{x}_{\text{NEW}} := \bar{x} + \bar{q} + \alpha \bar{r}$ (New approximate center) .

Then $\|\bar{x}_{\text{NEW}} - \hat{x}_\alpha\|_{Q_\alpha(\bar{x}_{\text{NEW}})} \leq 1/21$.

Proof: First note that $d = u + A\bar{r}$, where $u = \bar{S}\bar{u}$. Thus by Lemma 3.1,

$$\|\bar{x}_{\text{NEW}} - \hat{x}_\alpha(d)\|_{Q_\alpha(\hat{x}_\alpha(d); d)} = \|\bar{x}_\alpha - \hat{x}_\alpha(u)\|_{Q_\alpha(\hat{x}_\alpha(u); u)} ,$$

where $\bar{x}_\alpha = \bar{x} + \bar{q}$. Hence, by Lemma 3.2 (iv) (with $\bar{x} = \hat{x}_\alpha(d)$, $\tilde{x} = \bar{x}_{\text{NEW}}$ and $\delta = 1/22$) , it suffices to show that $\|\bar{x}_\alpha - \hat{x}_\alpha\|_{Q_\alpha(\hat{x}_\alpha)} \leq 1/22$, where $\hat{x}_\alpha = \hat{x}_\alpha(u)$ and $Q_\alpha(\hat{x}_\alpha) = Q_\alpha(\hat{x}_\alpha(u); u)$. In Lemma 4.1 (iii) below we show that, $\varepsilon = \|\bar{x} - \hat{x}_\alpha\|_{Q_\alpha(\hat{x}_\alpha)} \leq .1462$, which, by Theorem 4.2, implies that $\|\bar{x}_\alpha - \hat{x}_\alpha\|_{Q_\alpha(\hat{x}_\alpha)} \leq .033 < 1/22$. ■

Lemma 4.1: Under the conditions of Theorem 2.1, define $\bar{u}, \bar{r}$, and α as in Theorem 2.1. Let $u = \bar{S}\bar{u}$, let $\hat{x}_t$ denote the center of the inequality system $Ax \leq b + ut$ and let $Q_t(x) = Q_t(x; u)$. For all $t \in [0, \alpha]$,

$$(i) \quad \|v\|_{Q(\hat{x})} \leq \left(\frac{11}{10}\right)\left(\frac{77}{76}\right) \|v\|_{Q_t(\hat{x}_t)} \quad \text{for all } v \in \mathbb{R}^n ;$$

$$(ii) \quad \|v\|_{Q_t(\hat{x}_t)} \leq \left(\frac{12}{11}\right)\left(\frac{76}{75}\right) \|v\|_{Q(\hat{x})} \quad \text{for all } v \in \mathbb{R}^n ; \text{ and}$$

$$(iii) \quad \|\bar{x} - \hat{x}_t\|_{Q_t(\hat{x}_t)} \leq .1462 .$$

Proof: From Corollary 4.2, we have

$$\|t \Delta^{-1}(\bar{x}) u\| \leq \frac{1}{76} \quad \text{and} \quad \|\hat{x} - \hat{x}_t\|_{Q(\hat{x}_t)} < 1/11 \quad (4.1)$$

Thus, by Lemma 3.2(ii) and (iii)

$$\|\Delta_t^{-1}(\hat{x}_t) \Delta_t(\hat{x})\| \leq \frac{12}{11} \quad \text{and} \quad \|\Delta_t^{-1}(\hat{x}) \Delta_t(\hat{x}_t)\| \leq \frac{11}{10} . \quad (4.2)$$

$$\begin{aligned}
\text{Also, } \left\| \Delta_t^{-1}(\hat{x}) \Delta(\hat{x}) \right\| &= \max_i \left| \frac{b_i - A_i \hat{x}}{b_i + t u_i - A_i \hat{x}} \right| \\
&= \max_i \left| \frac{1}{1 + t \left(\frac{u_i}{b_i - A_i \hat{x}} \right)} \right| \\
&\leq \left| \frac{1}{1 - \left\| t \Delta^{-1}(\hat{x}) u \right\|} \right| \leq \frac{76}{75} \quad , \tag{4.3}
\end{aligned}$$

because $\left\| t \Delta^{-1}(\hat{x}) u \right\| \leq \frac{1}{76}$, from (4.1). Similarly,

$$\begin{aligned}
\left\| \Delta^{-1}(\hat{x}) \Delta_t(\hat{x}) \right\| &= \max_i \left| \frac{b_i + t u_i - A_i \hat{x}}{b_i - A_i \hat{x}} \right| \\
&= \max_i \left| 1 + t \frac{u_i}{b_i - A_i \hat{x}} \right| \\
&\leq 1 + \left\| t \Delta^{-1}(\hat{x}) u \right\| \leq \frac{77}{76} \quad . \tag{4.4}
\end{aligned}$$

$$\begin{aligned}
\text{Hence, } \left\| v \right\|_{Q(\hat{x})} &= \left\| \Delta^{-1}(\hat{x}) A v \right\| \\
&\leq \left\| \Delta^{-1}(\hat{x}) \Delta_t(\hat{x}) \right\| \left\| \Delta_t^{-1}(\hat{x}) \Delta_t(\hat{x}_t) \right\| \left\| \Delta_t^{-1}(\hat{x}_t) A v \right\| \\
&\leq \left(\frac{77}{76} \right) \left(\frac{11}{10} \right) \left\| v \right\|_{Q_t(\hat{x}_t)} \quad , \quad \text{from (4.2) and (4.4)} \quad .
\end{aligned}$$

This proves (i). The proof of (ii) is similar, using (4.2) and (4.3). Next, to show (iii), we have by the triangle inequality,

$$\begin{aligned}
\|\bar{x} - \hat{x}_t\|_{Q_t(\hat{x}_t)} &\leq \|\bar{x} - \hat{x}\|_{Q_t(\hat{x}_t)} + \|\hat{x} - \hat{x}_t\|_{Q_t(\hat{x}_t)} \\
&\leq \left(\frac{12}{11}\right)\left(\frac{76}{75}\right)\|\bar{x} - \hat{x}\|_{Q(\hat{x})} + \frac{1}{11}, \quad \text{from (ii) and (4.1),} \\
&\leq \left(\frac{12}{11}\right)\left(\frac{76}{75}\right)\left(\frac{1}{20}\right) + \frac{1}{11} < .1462,
\end{aligned}$$

since $\|\bar{x} - \hat{x}\|_{Q(\hat{x})} \leq \frac{21}{20}\|\bar{x} - \hat{x}\|_{Q(\bar{x})} \leq \frac{1}{20}$, from Lemma 3.2(iv). ■

We are now ready to prove:

Lemma 2.4: For all values of t generated by algorithm PCP, $\bar{x}(t)$ is near the center of X_t , in the sense that

$$\|\bar{x}(t) - \hat{x}_t\|_{Q_t(\bar{x}(t))} \leq .585.$$

Proof: Let $\alpha = 1/(80\|\bar{u}\|)$ as in Theorem 2.1. Let $t \in [0, \alpha]$. By Lemma 3.2(ii), it suffices to show that $\|\bar{x}(t) - \hat{x}_t\|_{Q_t(\hat{x}_t)} \leq 0.369$.

We have, from the proof of Theorem 2.1, and Lemma 4.1(iii),

$$\|\bar{x}_\alpha - \bar{x}\|_{Q_\alpha(\hat{x}_\alpha)} \leq \|\bar{x}_\alpha - \hat{x}_\alpha\|_{Q_\alpha(\hat{x}_\alpha)} + \|\bar{x} - \hat{x}_\alpha\|_{Q_\alpha(\hat{x}_\alpha)} < .033 + .1462 < 0.18.$$

$$\begin{aligned}
\text{Hence, } \|\bar{x}_\alpha - \bar{x}\|_{Q(\hat{x}_t)} &\leq \left(\frac{12}{11}\right)\left(\frac{76}{75}\right)\|\bar{x}_\alpha - \bar{x}\|_{Q(\hat{x})} && \text{(by Lemma 4.1(ii))} \\
&\leq \left(\frac{12}{11}\right)\left(\frac{76}{75}\right)\left(\frac{11}{10}\right)\left(\frac{77}{76}\right)\|\bar{x}_\alpha - \bar{x}\|_{Q_\alpha(\hat{x}_\alpha)} && \text{(by Lemma 4.1(i))} \\
&< 0.222.
\end{aligned}$$

Next, observe that $\bar{x}(t) - \bar{x} = \frac{t}{\alpha} (\bar{x}_\alpha - \bar{x})$. Therefore,

$$\|\bar{x}(t) - \bar{x}\|_{Q_t(\hat{x}_t)} = \frac{t}{\alpha} \|(\bar{x}_\alpha - \bar{x})\|_{Q_t(\hat{x}_t)} < 0.222 (t/\alpha) .$$

Thus, by the triangle inequality and Lemma 4.1(iii),

$$\|\bar{x}(t) - \hat{x}_t\|_{Q_t(\hat{x}_t)} \leq \|\bar{x}(t) - \bar{x}\|_{Q_t(\hat{x}_t)} + \|\bar{x} - \hat{x}_t\|_{Q_t(\hat{x}_t)} < .222 + .1462 < .369 .$$

■

5. Lower and Upper Bounds

In this section, we analyze the upper and/or lower bounds generated by algorithm PCP , and we will prove Theorem 2.2 and Lemma 2.7.

Recall that $T = \{t \mid \text{int } X_t \neq \emptyset\}$, $T_{\text{MAX}} := \sup T$ and $T_{\text{MIN}} := \inf T$. We shall derive upper bounds on T_{MAX} and lower bounds on T_{MIN} . We shall, for convenience, adopt the following notation and assumptions throughout this section. The current value of t is $t = 0$, and the current iterate $\bar{x}$ satisfies $\bar{s} = b - A\bar{x} > 0$ and $\|\hat{x} - \bar{x}\|_{Q(\bar{x})} = \|\bar{S}^{-1}(\hat{s} - \bar{s})\| \leq 1/21$, where $\hat{x}$ is the center of the system $Ax \leq b$, $\hat{s} = b - A\hat{x}$, $\bar{S} = \text{diag}(\bar{s})$ and $\hat{S} = \text{diag}(\hat{s})$. We decompose d into

$$d = u + A\bar{r} , \text{ where } \bar{S}^{-1}u = \bar{u} = P_{\bar{S}^{-1}A} \bar{S}^{-1}d \text{ and}$$

$$\bar{r} = (A^T \bar{S}^{-2} A)^{-1} A^T \bar{S}^{-2} d . \quad (5.1)$$

By assumption, d does not lie in the column space of A and so $\bar{u} \neq 0$. We shall prove Theorem 2.2 and Lemma 2.7 in this section. We begin with three fundamental Lemmas.

Lemma 5.1: $\left| e^T \hat{S}^{-1} u - e^T \bar{u} \right| \leq \|\bar{u}\|/20$

Proof:

$$\begin{aligned} \left| e^T \hat{S}^{-1} u - e^T \bar{u} \right| &= \left| e^T (\hat{S}^{-1} - \bar{S}^{-1}) u \right| \\ &= \left| (\bar{s} - \hat{s}) \hat{S}^{-1} \bar{S}^{-1} u \right| \\ &\leq \left\| \hat{S}^{-1} (\bar{s} - \hat{s}) \right\| \left\| \bar{S}^{-1} u \right\| \\ &\leq (1/20) \|\bar{u}\| . \end{aligned}$$

The first inequality is the Cauchy-Schwartz inequality and the last follows from Lemma 3.2 and the assumption that $\|\bar{x} - \hat{x}\|_{Q(\bar{x})} \leq 1/21$. ■

Recall that $f(0) = e^T \hat{S}^{-1} u$ is the path indicator function at $t = 0$. If $e^T \bar{u} > \|\bar{u}\|/20$ then from Lemma 5.1, $e^T \hat{S}^{-1} u \geq e^T \bar{u} - \|\bar{u}\|/20 > 0$.

Therefore, the current center $\hat{x}$ is on the lower path and $T_{\min}$ is bounded below, as we have seen in Section 2, by the bound $-m/(e^T \hat{S}^{-1} u)$.

Similarly, if $e^T \bar{u} < -\|\bar{u}\|/20$ then $T_{\max}$ is bounded above by $-m/(e^T \hat{S}^{-1} u)$. However, we typically cannot deduce the exact value of $e^T \hat{S}^{-1} u$ from the current iterate. Therefore, we shall later in this section derive an alternate bound using the next Lemma which is a variant of Lemma 7.1 of [4].

Lemma 5.2: Let $\hat{x}$ be the center of the system $Ax \leq b$. Suppose $\bar{x} \in \text{int } X = \{x \mid Ax < b\}$ is given such that $\|\bar{x} - \hat{x}\|_{Q(\bar{x})} \leq \delta < 1$. Define the ellipsoid F_{OUT} by:

$$F_{\text{OUT}} := \left\{ x \in \mathbb{R}^n \mid \|x - \bar{x}\|_{Q(\bar{x})} \leq (1 + \delta) \sqrt{m(m-1)} + \delta \right\} \quad (5.2)$$

Then $\text{int } X \subseteq F_{\text{OUT}}$.

Proof: Let $x \in \text{int } X$ be given and let $s = b - Ax$. Then from properties of the center, (see for example [4], Theorem 2.1), $\|x - \hat{x}\|_{Q(\hat{x})} \leq \sqrt{m(m-1)}$. From

Lemma 3.2(v), we also have $\|x - \hat{x}\|_{Q(\bar{x})} \leq (1 + \delta) \sqrt{m(m-1)}$. Thus,
 $\|x - \bar{x}\|_{Q(\bar{x})} \leq \|x - \hat{x}\|_{Q(\bar{x})} + \|\hat{x} - \bar{x}\|_{Q(\bar{x})} \leq (1 + \delta) \sqrt{m(m-1)} + \delta$. ■

The next Lemma concerns the well-known classical least-square problem or minimum-norm problem.

Lemma 5.3: Given an $m \times n$ matrix M and an m -vector d , for all scalar t ,

$$\|Mx - td\| \geq |t| \|P_M d\| \quad \text{for all } x \in \mathbb{R}^n,$$

where $P_M = I - M(M^T M)^{-1} M^T$. ■

We are now ready to prove Theorem 2.2 and Lemma 2.7. We shall prove Theorem 2.2 in two parts.

Proposition 5.1: Under the definitions and conditions of Theorem 2.2,

- (i) if $e^T \bar{u} / \|\bar{u}\| < -1/20$ then $T_{\text{MAX}} \leq UB \equiv (22 \sqrt{m(m+1)} + 1) / 21 \|\bar{u}\|$,
- (ii) if $e^T \bar{u} / \|\bar{u}\| > 1/20$ then $T_{\text{MIN}} \geq LB \equiv -(22 \sqrt{m(m+1)} + 1) / 21 \|\bar{u}\|$.

Proof: (i) Suppose $e^T \bar{u} / \|\bar{u}\| < -1/20$. We first show that $\begin{bmatrix} \bar{x} \\ 0 \end{bmatrix}$ is close to the center of the following extended system with one additional variable and one additional constraint:

$$\overline{A} \begin{bmatrix} x \\ t \end{bmatrix} \leq \begin{bmatrix} b \\ 1 \end{bmatrix} \quad (5.3)$$

where $\bar{A} = \begin{bmatrix} A & -u \\ 0 & \theta \end{bmatrix}$ and $\theta = e^T \hat{S}^{-1} u$.

Note that from the remark following Lemma 5.1, $\theta < 0$. Therefore,

$$\begin{aligned} T_{\text{MAX}} &= \sup_t \left\{ t \mid Ax \leq b + dt \text{ for some } x \right\} \\ &= \sup_t \left\{ t \mid Ax \leq b + ut \text{ for some } x \right\} \\ &= \sup_t \left\{ t \mid \bar{A} \begin{bmatrix} x \\ t \end{bmatrix} \leq \begin{bmatrix} b \\ 1 \end{bmatrix} \text{ for some } x \right\}. \end{aligned}$$

For system (5.3), define

$$\bar{\Delta}(x, t) := \text{diag} \left(\begin{bmatrix} b \\ 1 \end{bmatrix} - \bar{A} \begin{bmatrix} x \\ t \end{bmatrix} \right) \text{ and } \bar{Q}(x, t) := \bar{A}^T \bar{\Delta}^{-2}(x, t) \bar{A}.$$

Then, $\bar{\Delta}(\hat{x}, 0) = \begin{bmatrix} \hat{S} & 0 \\ 0 & 1 \end{bmatrix}$, $\bar{\Delta}(\bar{x}, 0) = \begin{bmatrix} \bar{S} & 0 \\ 0 & 1 \end{bmatrix}$ and

$$[e^T, 1] \bar{\Delta}^{-1}(\hat{x}, 0) \bar{A} = [e^T \hat{S}^{-1} A, -e^T \hat{S}^{-1} u + \theta] = 0.$$

Therefore, $\begin{bmatrix} \hat{x} \\ 0 \end{bmatrix}$ is the center of system (5.3) and

$$\left\| \begin{bmatrix} \hat{x} \\ 0 \end{bmatrix} - \begin{bmatrix} \bar{x} \\ 0 \end{bmatrix} \right\|_{\bar{Q}(\bar{x}, 0)} = \|\bar{x} - \hat{x}\|_{Q(\bar{x})} \leq 1/21.$$

Let $\bar{X} = \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \in \mathbb{R}^{n+1} \mid \bar{A} \begin{bmatrix} x \\ t \end{bmatrix} \leq \begin{bmatrix} b \\ 1 \end{bmatrix} \right\}$, and

$$\bar{F}_{\text{OUT}} = \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \in \mathbb{R}^{n+1} \mid \left\| \begin{bmatrix} x \\ t \end{bmatrix} - \begin{bmatrix} \bar{x} \\ 0 \end{bmatrix} \right\|_{\bar{Q}(\bar{x},0)} \leq \frac{22}{21} \sqrt{m(m+1)} + \frac{1}{21} \right\} .$$

Then, from Lemma 5.2, $\text{int } \bar{X} \subseteq \bar{F}_{\text{OUT}}$.

Furthermore, because $0 = A^T \bar{S}^{-1} \bar{u} = A^T \bar{S}^{-2} u$, we have

$$\bar{Q}(\bar{x}, 0) = \begin{bmatrix} A^T \bar{S}^{-2} A & -A^T \bar{S}^{-2} u \\ -u^T \bar{S}^{-2} A & \theta^2 + u^T \bar{S}^{-2} u \end{bmatrix} = \begin{bmatrix} A^T \bar{S}^{-2} A & 0 \\ 0 & \theta^2 + \|\bar{u}\|^2 \end{bmatrix} .$$

Thus

$$\begin{aligned} \left\| \begin{bmatrix} x \\ t \end{bmatrix} - \begin{bmatrix} \bar{x} \\ 0 \end{bmatrix} \right\|_{\bar{Q}(\bar{x},0)}^2 &= \left\| \bar{S}^{-1} A (x - \bar{x}) \right\|^2 + t^2 (\theta^2 + \|\bar{u}\|^2) \\ &\geq t^2 \|\bar{u}\|^2 . \end{aligned}$$

$$\text{Also, for any } \begin{bmatrix} x \\ t \end{bmatrix} \in \bar{F}_{\text{OUT}} , \quad \left\| \begin{bmatrix} x \\ t \end{bmatrix} - \begin{bmatrix} \bar{x} \\ 0 \end{bmatrix} \right\|_{\bar{Q}(\bar{x},0)} \leq \frac{22 \sqrt{m(m+1)} + 1}{21} ,$$

$$\text{whereby } t \leq \frac{22 \sqrt{m(m+1)} + 1}{21 \|\bar{u}\|} .$$

It then follows from Lemma 3.1 that

$$\begin{aligned} T_{\text{MAX}} &= \sup \left\{ t \mid Ax - tu < b \text{ for some } x \right\} \\ &= \sup \left\{ t \mid Ax - tu < b \text{ for some } x, \theta t < 1 \right\} \quad (\text{since } \theta < 0) \\ &= \sup \left\{ t \mid \begin{bmatrix} x \\ t \end{bmatrix} \in \text{int } \bar{X} \text{ for some } x \right\} \end{aligned}$$

$$\leq \max \left\{ t \mid \begin{bmatrix} x \\ t \end{bmatrix} \in \bar{F}_{\text{OUT}} \text{ for some } x \right\} \quad (\text{since } \text{int } \bar{X} \subseteq \bar{F}_{\text{OUT}})$$

$$\leq (22 \sqrt{m(m+1)} + 1) / 21 \|\bar{u}\| .$$

The proof of (ii) exactly parallels that of (i). ■

Proposition 5.2: Suppose $|e^T \bar{u}| / \|\bar{u}\| \leq 1/20$. Then

$$(i) \quad T_{\text{MAX}} \leq \bar{UB} \equiv (1.6 \sqrt{m(m-1)} + .6) / \|\bar{u}\|$$

$$(ii) \quad T_{\text{MIN}} \geq \bar{LB} \equiv -(1.6 \sqrt{m(m-1)} + .6) / \|\bar{u}\|$$

Proof: We shall show that $\begin{bmatrix} \bar{x} \\ 0 \end{bmatrix}$ is approximately centered (in the sense of Section 3.2) for the extended system

$$[A, -u] \begin{bmatrix} x \\ t \end{bmatrix} \leq b \tag{5.4}$$

Let Q, y and $\gamma = \gamma(\bar{x})$ for system $Ax \leq b$ be as defined in Section 3.2. Let $\tilde{Q}, \tilde{y}$ and $\tilde{\gamma} = \tilde{\gamma}(\bar{x}, 0)$ be the corresponding parameters for system (5.4). Observe that

$$b - [A, -u] \begin{bmatrix} \bar{x} \\ 0 \end{bmatrix} = b - A\bar{x} = \bar{s} > 0$$

Therefore, since $\bar{u} = \bar{S}^{-1} u$ and $A^T \bar{S}^{-1} \bar{u} = 0$,

$$\tilde{Q} = \frac{1}{m} \begin{bmatrix} A^T \\ -u^T \end{bmatrix} \bar{S}^{-2} [A, -u] = \begin{bmatrix} Q & 0 \\ 0 & \|\bar{u}\|^2 / m \end{bmatrix}$$

$$\text{and } \tilde{y} = \frac{1}{m} \begin{bmatrix} A^T \\ -u^T \end{bmatrix} \bar{S}^{-1} e = \begin{bmatrix} y \\ -e^T \bar{u}/m \end{bmatrix}.$$

$$\text{Therefore, } \tilde{y}^T \tilde{Q}^{-1} \tilde{y} = y^T Q^{-1} y + (e^T \bar{u})^2 / m \|\bar{u}\|^2.$$

Recalling that $\|\bar{x} - \hat{x}\|_{Q(\bar{x})} \leq \delta = 1/21$, we compute $a = \frac{1}{760}$ and conclude from Lemma 3.4 that, $\gamma \leq a + \sqrt{2a}$. Next from Lemma 3.5(i), $(m-1)y^T Q^{-1} y \leq \gamma^2$. Thus,

$$\begin{aligned} (m-1)\tilde{y}^T \tilde{Q}^{-1} \tilde{y} &\leq (a + \sqrt{2a})^2 + \left(\frac{m-1}{m}\right) \left(\frac{e^T \bar{u}}{\|\bar{u}\|}\right)^2 \\ &< \left(\frac{1}{760} + \frac{1}{\sqrt{380}}\right)^2 + \left(\frac{1}{20}\right)^2 < .0053. \end{aligned}$$

Therefore, by Lemma 3.5(ii), $\tilde{\gamma}^2 < .0054$ and so $\tilde{\gamma} < .0735$. Taking $h = .41$ in Lemma 3.3, we have

$$\left\| \begin{bmatrix} x^* \\ t^* \end{bmatrix} - \begin{bmatrix} \bar{x} \\ 0 \end{bmatrix} \right\|_{\tilde{Q}(\bar{x}, 0)}^2 \leq \left(\frac{m}{m-1}\right) \frac{h^2(1 + \tilde{\gamma}^2)}{(1 - \tilde{\gamma}h)^2} < .36 < (.6)^2, \text{ (since } m/(m-1) \leq 2)$$

where $\begin{bmatrix} x^* \\ t^* \end{bmatrix}$ is the center of system (5.4) and $\tilde{Q}(\bar{x}, 0) = m\tilde{Q}$.

$$\text{Note that } \left\| \begin{bmatrix} x \\ t \end{bmatrix} - \begin{bmatrix} \bar{x} \\ 0 \end{bmatrix} \right\|_{\tilde{Q}(\bar{x}, 0)} = \left\| \bar{S}^{-1} A(x - \bar{x}) - t\bar{u} \right\|.$$

$$\text{Let } \tilde{X} = \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \mid [A, -u] \begin{bmatrix} x \\ t \end{bmatrix} \leq b \right\} \text{ and}$$

$$\tilde{F}_{\text{OUT}} = \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \mid \left\| \bar{S}^{-1} A (x - \bar{x}) - t \bar{u} \right\| \leq 1.6 \sqrt{m(m-1)} + .6 \right\}.$$

Then by Lemma 5.2, $\text{int } \tilde{X} \subseteq \tilde{F}_{\text{OUT}}$. By Lemma 5.3, for all

$$\begin{bmatrix} x \\ t \end{bmatrix} \in \tilde{F}_{\text{OUT}}, \quad 1.6 \sqrt{m(m-1)} + .6 \geq \left\| \bar{S}^{-1} A (x - \bar{x}) - t \bar{u} \right\| \geq |t| \|\bar{u}\| \quad \text{since}$$

$$A^T \bar{S}^{-1} \bar{u} = 0. \quad \text{Thus}$$

$$\begin{aligned} \sup \{ |t| \mid t \in T \} &= \sup \left\{ |t| \mid \begin{bmatrix} x \\ t \end{bmatrix} \in \text{int } \tilde{X} \quad \text{for some } x \right\} \\ &\leq \max \left\{ |t| \mid \begin{bmatrix} x \\ t \end{bmatrix} \in \tilde{F}_{\text{OUT}} \quad \text{for some } x \right\} \\ &\leq (1.6 \sqrt{m(m-1)} + .6) / \|\bar{u}\| \end{aligned}$$

This completes the proof. ■

Theorem 2.2 follows immediately from Propositions 5.1 and 5.2. Finally, we prove

Lemma 2.7: If criterion (ii) of Step 4 of algorithm PCP is satisfied at iteration k , then in all subsequent iterations, criteria (ii) or (iii) of Step 4 will be satisfied.

Proof: Suppose $e^T \bar{u} / \|\bar{u}\| < -1/20$ in iteration k . It suffices to show that criterion (i) will not be satisfied in all subsequent iterations. By Lemma 5.1, we have $f(t^k) = e^T \hat{S}^{-1} u < 0$. By Proposition 2.2,

$f(t) \leq f(t^k) < 0$ for all $t \geq t^k$. Therefore, in all subsequent iterations, criterion (i) will not be satisfied, as it would imply that $f(t) \geq 0$ for some $t \geq t^k$. ■

References

- [1] Barnes, E. R., S. Chopra, and D. L. Jensen (1988). "A polynomial time version of the affine scaling algorithm." Working paper, Graduate School of Business Administration, New York University, New York, NY.
- [2] Censor, Y. and A. Lent (1987). "Optimization of 'log x' entropy over linear equality constraints." **SIAM Journal of Control and Optimization** 25, 921-933.
- [3] Freund, R. M. (1988). "Projective transformations for interior point methods, Part I: Basic theory and linear programming." Working paper OR 179-88, Operations Research Center, M.I.T., Cambridge, MA.
- [4] Freund, R. M. (1988). "Projective transformations for interior point methods, Part II: Analysis of an algorithm for finding the weighted center of a polyhedral system" Working paper OR 180-88, Operations Research Center, M.I.T., Cambridge, MA.
- [5] Jarre, F. (1988). "On the convergence of the method of centers when applied to convex quadratic programs." Manuscript, Institut für Angewandte Mathematik und Statistik, Universität Würzburg, Am Hubland, 8700 Würzburg, West Germany.
- [6] Karmarkar, N. (1984). "A new polynomial time algorithm for linear programming." **Combinatorica** 4, 373-395.
- [7] Kojima, M., S. Mizuno, and A. Yoshise (1987). "A primal-dual interior point method for linear programming." Research report No. B-193, Dept. of Information Sciences, Tokyo Institute of Technology, Japan.
- [8] Megiddo, N. (1986). "Pathways to the optimal set in linear programming." Research report RJ 5295, IBM Almaden Research Center, San Jose, CA.
- [9] Mehotra, S. and J. Sun (1988). "A method of analytic centers for quadratically constrained convex quadratic programs." Technical report 88-01, Dept. of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL.
- [10] Monteiro, R. C. and I. Adler (1987). "An $O(n^3L)$ primal-dual interior point algorithm for linear programming." Report ORC 87-4, Dept. of Industrial Engineering and Operations Research, University of California, Berkeley, CA.
- [11] Renegar, J. (1988). "A polynomial-time algorithm, based on Newton's method, for linear programming." **Mathematical Programming** 40, 59-94.

- [12] Sonnevend, G. (1985). "An 'analytic' center for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming." Preprint, Dept. of Numerical Analysis, Institute of Mathematics, Eötvös University, 1088, Budapest, Muzeum Körút 6-8.
- [13] Sonnevend, G. (1985). "A new method for solving a set of linear (convex) inequalities and its applications for identification and optimization." Preprint, Dept. of Numerical Analysis, Institute of Mathematics, Eötvös University, 1088, Budapest, Muzeum Körút 6-8.
- [14] Todd, M. J. and Y. Ye (1987). "A centered projective algorithm for linear programming." Technical report No. 763, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY.
- [15] Vaidya, P. (1987). "A locally well-behaved potential function and a simple Newton-type method for finding the center of a polytope." AT&T Bell Laboratories, Murray Hill, NJ.
- [16] Vaidya, P. (1987). "An algorithm for linear programming which requires $O(((m+n)n^2 + (m+n)^{1.5} n) L)$ arithmetic operations." **Proceeding of the Nineteenth ACM Symposium on the Theory of Computing**, 29-38.