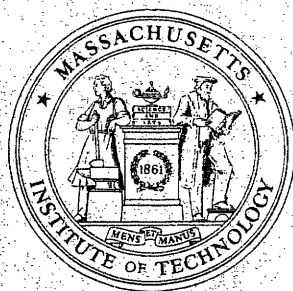
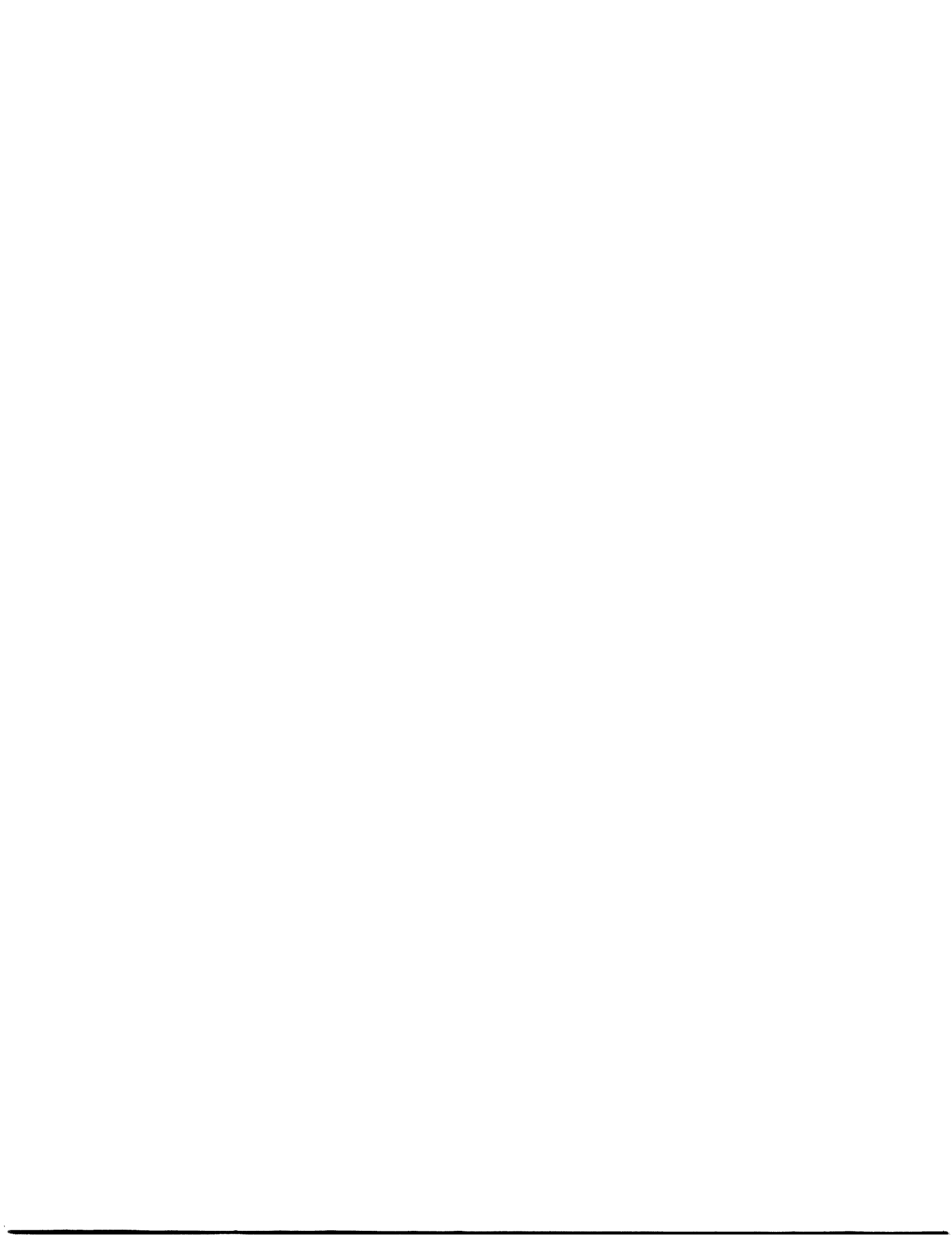


OPERATIONS RESEARCH CENTER

working paper



**MASSACHUSETTS INSTITUTE
OF TECHNOLOGY**



**THEORETICAL EFFICIENCY OF A
SHIFTED BARRIER FUNCTION
ALGORITHM FOR LINEAR PROGRAMMING**

Robert M. Freund

OR 194-89

April, 1989



Abstract

This paper examines the theoretical efficiency of solving a standard-form linear program by solving a sequence of shifted-barrier problems of the form

$$\begin{aligned} & \underset{x}{\text{minimize}} && c^T x - \varepsilon \sum_{j=1}^n \ln(x_j + \varepsilon h_j) \\ & \text{s.t.} && Ax = b, \quad x + \varepsilon h > 0, \end{aligned}$$

for a given and fixed shift vector $h > 0$, and for a sequence of values of $\varepsilon > 0$ that converges to zero. The resulting sequence of solutions to the shifted barrier problems will converge to a solution to the standard form linear program. The advantage of using the shifted-barrier approach is that a starting feasible solution is unnecessary, and there is no need for a Phase I-Phase II approach to solving the linear program, either directly or through the addition of an artificial variable. Furthermore, the algorithm can be initiated with a "warm start," i.e., an initial guess of a primal solution \tilde{x} that need not be feasible. The number of iterations needed to solve the linear program to a desired level of accuracy will depend on a measure of how close the initial solution \tilde{x} is to being feasible. The number of iterations will also depend on the judicious choice of the shift vector h . If an approximate center of the dual feasible region is known, then h can be chosen so that the guaranteed fractional decrease in ε at each iteration is $(1 - 1/(6\sqrt{n}))$, which contributes a factor of $6\sqrt{n}$ to the number of iterations needed to solve the problem. The paper also analyzes the complexity of computing an approximate center of the dual feasible region from a "warm start," i.e., an initial (possibly infeasible) guess $\tilde{\pi}$ of a solution to the center problem of the dual.

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



1. Introduction

This paper examines the theoretical efficiency of an algorithm for solving a standard-form linear program

$$\begin{aligned} & \text{minimize} && c^T x \\ & && x \\ & \text{s.t.} && Ax = b, \quad x \geq 0, \end{aligned}$$

by solving a sequence of shifted-barrier problems $Sh(\epsilon)$ of the form:

$$\begin{aligned} Sh(\epsilon): & \quad \text{minimize} && c^T x - \epsilon \sum_{j=1}^n \ln(x_j + \epsilon h_j) \\ & && x \\ & \text{s.t.} && Ax = b, \quad x + \epsilon h > 0, \end{aligned}$$

for a given and fixed shift vector $h > 0$, and for a sequence of values of $\epsilon > 0$ that converges to zero. At the beginning of each iteration, the algorithm has an approximate solution \bar{x} to the problem $Sh(\epsilon)$, for the current value of ϵ . A fractional quantity $\alpha < 1$ is then computed, and the new value of ϵ is chosen as $\epsilon' = \alpha \epsilon$. The algorithm then computes a Newton step, and the resulting new value of x is an approximate solution to problem $Sh(\epsilon')$. The resulting sequence of solutions to the shifted barrier problems will converge to a solution to the standard form linear program.

Problem $Sh(\epsilon)$ given above is a specific instance of a more general problem introduced in Gill et al. [7], namely

$$\begin{aligned} SBP(w, f): & \quad \text{minimize} && c^T x - \sum_{j=1}^n w_j \ln(x_j + f_j) \\ & && x \\ & \text{s.t.} && Ax = b, \quad x + f > 0, \end{aligned}$$

where in addition to the shift vector $f > 0$ there is a positive vector $w = (w_1, \dots, w_n)^T$ used to weight the contribution of each logarithm term. The results of Gill et al. [7] treat the problem of determining simultaneous sequences of values of $w = w^k$ and $f = f^k$ for $k = 1, \dots$, so that the resulting optimal solutions x^k to $SBP(w^k, f^k)$ converge to an optimal solution x^* to the linear program. There are also results in [7] regarding generic properties of $SBP(w, f)$ and the use of Newton's method for solving $SBP(w, f)$. The problem $Sh(\epsilon)$ considered in this study is recovered from $SBP(w, f)$ by setting $w^k = (\epsilon^k, \dots, \epsilon^k)^T$ and

by only considering shift vectors $f = f^k$ of the form $f^k = \epsilon^k h$, where $h > 0$ is given and fixed, and the scalar ϵ^k is varied.

There are a number of advantages in using the shifted-barrier problem $Sh(\epsilon)$ to solve a linear program. Perhaps the most important advantage is that the algorithm presented here can be initiated from a "warm start," i.e., a guess of a solution \tilde{x} to the linear program that is perhaps not feasible for the current linear program, but perhaps is very close to the optimal solution. This situation arises often in practice when solving a sequence of slightly-modified versions of a given linear programming problem. In this case, the optimal solution to a previous version of the linear program is infeasible for the current linear program, but is very close to the optimal solution of the current linear program. Thus, valuable information about slightly different versions of the current linear program can be used to great advantage in solving the current linear program, as opposed to other interior-point algorithms that must be initiated from a "cold start."

A second advantage of the shifted-barrier algorithm presented here is that a starting feasible solution is unnecessary, and hence there is no need for a Phase I-Phase II approach to solving the linear program, either directly or through the addition of an artificial variable. Most interior-point algorithms handle the Phase I-Phase II problem by introducing an artificial row or column with large coefficients either in the objective function or in the right-hand-side, see Anstreicher [1], Gay [6], Gonzaga [8], Steger [16], Todd and Burrell [17], Ye and Kojima [21], Renegar [12], Vaidya [19], and Monteiro and Adler [9], among others. In those algorithms, which use the "big M" method of initializing the algorithm, coefficients whose size is $O(L)$ must be chosen (where L is the length of the binary encoding of the linear program data), which is not usually implementable in practice. Anstreicher [2] was the first to present a polynomial-time interior point algorithm for linear programming that mitigates the need to modify the given linear program with an artificial row or column with large coefficients. The shifted-barrier algorithm presented here also shares this property.

The efficiency of the shifted-barrier algorithm depends critically on three factors. The first factor is the choice of the shift vector h . A naive approach is to choose h as the vector of ones, i.e., $h = (1, 1, 1, \dots, 1)$. Not too surprisingly, a much better choice of h can be determined by using knowledge of the center of the dual feasible region. In particular, suppose $(\bar{\pi}, \bar{s})$ is a dual feasible solution, i.e. $A^T \bar{\pi} + \bar{s} = c$, $\bar{s} \geq 0$, and $(\bar{\pi}, \bar{s})$ is close to the center of the dual feasible. Then a judicious choice of h is $h_j = 1/(n\bar{s}_j)$, $j = 1, \dots, n$. With h chosen in this manner, the guaranteed decrease in ϵ at each

iteration, which is measured by the fraction α , is $\alpha \leq (1 - 1/(6\sqrt{n}))$. This leads to a factor of $6\sqrt{n}$ in the analysis of the number of iterations of the shifted-barrier algorithm.

If $\epsilon = \epsilon^*$ is a desired level of accuracy for solving the shifted-barrier problem, and the algorithm is initiated with a value of $\epsilon = \epsilon^*$, then the number of iterations required to achieve $\epsilon \leq \epsilon^*$ will be $K = \lceil 6\sqrt{n} (\ln(\epsilon^*) - \ln(\epsilon^*)) \rceil$, if the shift vector h is chosen as above. Thus the second major factor affecting the efficiency of the shifted-barrier algorithm for linear programming is the initial value of $\epsilon = \epsilon^*$, which we would like to choose to be as small as possible. In Section 4 of the paper, we show how to choose ϵ^* as a function of the initial guess \tilde{x} of the optimal solution. Given the initial guess of \tilde{x} , and also given the choice of the shift vector h above, we present a way to efficiently choose the initial value of $\epsilon = \epsilon^*$. Furthermore, the value of ϵ^* will be roughly proportional to the degree of infeasibility of the initial (possibly infeasible) guess \tilde{x} . Part of the value of ϵ^* will be proportional to the degree of infeasibility of \tilde{x} in the equations $Ax = b$, and will be a function of the size of the vector $v = b - A\tilde{x}$, and another part of the value of ϵ^* will be proportional to the extent to which \tilde{x} is not nonnegative. Thus, if \tilde{x} is almost feasible, the initial value of $\epsilon = \epsilon^*$ can be chosen to be quite small. Hence, the algorithm can be initiated with a good "warm start."

Because knowing an approximate solution to the center of the dual is so important in using a shifted-barrier algorithm for linear programming, the third critical factor affecting the efficiency of the shifted-barrier approach is the complexity of computing an approximate center of the dual feasible solution. Algorithms for computing an approximate center from a known interior feasible solution are given in Vaidya [19] and in [4]. An algorithm for computing an approximate center from a possibly infeasible dual solution is presented in this paper in Section 5, and is a direct application of the algorithm of [4] and the parametric center-finding algorithm of [5]. The general complexity of computing an approximate center of the dual is analyzed in Section 5, and is based on an analysis using the two algorithms in [4] and in [5]. Suppose that $(\tilde{\pi}, \tilde{s})$ is an initial (possibly dual infeasible) guess of the center of dual feasible region. The main result of Section 5 gives a bound on the number of iterations needed to compute an approximate center of the dual feasible region. This bound is roughly proportional to how far $(\tilde{\pi}, \tilde{s})$ is from the center of the dual feasible region, in an appropriate measure.

This paper is organized as follows. In Section 2, we analyze the use of Newton's method for obtaining solutions to the shifted-barrier problem $Sh(\epsilon)$ for a decreasing sequence of values of ϵ . The main results, Theorems 2.1 and Proposition 2.2, show how the value of ϵ

can be decreased in conjunction with the computation of a Newton step. Section 3 applies the results of Section 2 and contains a path-following algorithm for a shifted-barrier approach to solving a linear program. Section 4 presents results regarding initializing the algorithm from knowledge of an approximate center of the dual feasible region. If an approximate center of the dual feasible region is known, then it is shown that the shift vector h can be chosen so that $\alpha \leq (1 - 1/(6\sqrt{n}))$ at each iteration of the algorithm, yielding the desired $6\sqrt{n}$ factor in the algorithm's iteration count. In Theorem 4.1, it is shown that initial value of $\epsilon = \epsilon^\circ$ can be chosen so that ϵ° is roughly proportional to degree of infeasibility of the initial guess \tilde{x} of a solution. In Section 5, we present an algorithm for computing an approximate center of the dual feasible region from an initial (possibly infeasible) dual solution $(\tilde{\pi}, \tilde{s})$. The complexity of computing an approximate center $(\bar{\pi}, \bar{s})$ of the dual feasible solution from the given possibly infeasible guess $(\tilde{\pi}, \tilde{s})$ is analyzed as well.

Notation. This paper will utilize the following notation. Regarding norms, $\|v\|$ will denote the Euclidean norm of a vector v , and $\|v\|_1$ will denote the L_1 -norm. The matrix norm $\|M\|$ is defined as $\|M\| = \sup\{\|Mv\| \mid \|v\| = 1\}$. We assume throughout the paper that the matrix A is $m \times n$ and has rank m , and that $n \geq 2$. The vector of ones is denoted by e , namely $e = (1, 1, 1, \dots, 1)^T$. If s, z, d, y, x, h , and w are vectors, then S, Z, D, Y, X, H , and W denote the diagonal matrixes whose diagonal entries correspond to the vector components. Then note, for example, that $\|S\| = \max_j \{s_j\}$ if $s \geq 0$.

2. An Improvement Theorem for Shifted Barrier Functions for Linear Programming

For the given linear program (LP) and its dual (DP)

$$\begin{array}{ll}
 \text{LP:} & \text{minimize} & c^T x \\
 & & x \\
 & \text{s.t.} & Ax = b \\
 & & x \geq 0 \\
 \\
 \text{DP:} & \text{maximize} & b^T \pi \\
 & & \pi, s \\
 & \text{s.t.} & A^T \pi + s = c \\
 & & s \geq 0,
 \end{array}$$

we propose to solve LP by introducing a shifted barrier function as follows. Let $h \in \mathbb{R}^n$ be a given strictly positive vector. Then for a given value of $\epsilon > 0$, we relax the nonnegativity conditions on x to the conditions $x + \epsilon h \geq 0$. As $\epsilon > 0$ is shrunk to zero, this condition will in the limit be the usual nonnegativity condition $x \geq 0$. With this in mind, we propose to solve LP by considering the following shifted barrier problem (see Polyak [11], also see Gill et al. [7]):

$$\begin{aligned} \text{Sh}(\epsilon): \text{ minimize} \quad & c^T x - \epsilon \sum_{j=1}^n \ln(x_j + \epsilon h_j) \\ \text{ s.t.} \quad & Ax = b \\ & x + \epsilon h > 0 . \end{aligned}$$

The Karush-Kuhn-Tucker (K-K-T) conditions ensure that for a given $\epsilon > 0$, that x solves $\text{Sh}(\epsilon)$ if and only if there exists $\pi \in \mathbb{R}^m$ for which

$$Ax = b, \quad x + \epsilon h > 0, \quad (2.1a)$$

$$c_j - \epsilon / (x_j + \epsilon h_j) = (A^T \pi)_j, \quad j = 1, \dots, n. \quad (2.1b)$$

Conditions (2.1) can be rewritten in the following different format:

$$Ax = b \quad (2.2a)$$

$$y = x + \epsilon h > 0 \quad (2.2b)$$

$$A^T \pi + s = c, \quad s > 0 \quad (2.2c)$$

$$e - \left(\frac{1}{\epsilon}\right) Ys = 0. \quad (2.2d)$$

For a given value of $\epsilon > 0$, we will say that x and (π, s) are β -approximate solutions to $\text{Sh}(\epsilon)$ if x and (π, s) satisfy:

$$Ax = b \quad (2.3a)$$

$$y = x + \epsilon h > 0 \quad (2.3b)$$

$$A^T \pi + s = c, \quad s > 0 \quad (2.3c)$$

$$\|r\| \leq \beta, \quad \text{where } r = e - \left(\frac{1}{\epsilon}\right) Ys \quad (2.3d)$$

The next proposition presents properties of a β -approximate solution.

Proposition 2.1. (Properties of a β -approximate Solution to $Sh(\epsilon)$.) Suppose x and (π, s) are a β -approximate solution to $Sh(\epsilon)$. Then

- (i) $Ax = b$, $x_j > -\epsilon h_j$, $j=1, \dots, n$, i.e., x is almost primal feasible.
- (ii) $A^T \pi + s = c$, $s \geq 0$, i.e., (π, s) is dual feasible.
- (iii) $-\epsilon(\|e - Hs\| + \beta) \leq x_j s_j \leq \epsilon(1 + \beta)$, $j=1, \dots, n$.
- (iv) $-\epsilon\|e - Hs\| + \beta \leq x^T s \leq \epsilon(1 + \beta)$, i.e., x and s are almost complementary.

Proof: (i) and (ii) follow from 2.3a, 2.3b, and 2.3c directly. Let

$r = e - \left(\frac{1}{\epsilon}\right) \gamma s = e - Hs - \left(\frac{1}{\epsilon}\right) \chi s$. Then $\|r\| \leq \beta$. Furthermore, expanding r yields for $j=1, \dots, n$,

$$r_j = 1 - h_j s_j - \left(\frac{1}{\epsilon}\right) x_j s_j.$$

Thus, $x_j s_j = \epsilon(1 - r_j - h_j s_j) \leq \epsilon(1 - r_j) \leq \epsilon(1 + \beta)$.

This shows the right part of (iii). To see the left part, we have

$$x_j s_j = \epsilon(1 - h_j s_j) - \epsilon r_j \geq -\epsilon\|e - Hs\| - \epsilon\|r\| \geq -\epsilon\|e - Hs\| - \epsilon\beta.$$

We have now shown (iii), and (iv) is an immediate consequence of (iii). ■

Note that the upper bound on the almost-complementarity condition in (iv) depends only on n , ϵ , and β . However the lower bound also depends on $\|e - Hs\|$, which could possibly be arbitrarily large. However, we will show in Section 3 that if h is chosen judiciously, then $\|e - Hs\|$ can be bounded by $1.5\sqrt{n} - 50$, and we have:

Corollary 2.1. Suppose all dual feasible solutions (π, s) satisfy $\|e - Hs\| \leq 1.5\sqrt{n} - 50$. Then if x and (π, s) are a β -approximate solution to $Sh(\epsilon)$, then $|x^T s| \leq \epsilon n(1.5\sqrt{n} - 50 + \beta)$. ■

Corollary 2.2. Suppose $h_j \leq O(2^L)$, $j = 1, \dots, n$, where L is the length of a binary encoding of the data for LP. Then if x and (π, s) are a β -approximate solution to $Sh(\epsilon)$, where $\epsilon \leq O(2^{-2L})$, then x can be rounded to an optimal solution to LP in $O(n^3)$ operations.

Proof: From (i) of Proposition 2.1, $x_j \geq -\epsilon h_j \geq -O(2^{-L})$, whereby x can be rounded to feasible solution \bar{x} of LP, see Papadimitriou and Steiglitz [10]. Furthermore, since $x^T s \leq n\epsilon(1+\beta) \leq O(2^{-L})$, it is also straightforward to show that $\bar{x}^T s \leq O(2^{-L})$, whereby \bar{x} and s can be rounded to optimal primal and dual solutions in $O(n^3)$ operations, also see [10]. ■

In Section 4, we will show that the same choice of h that yields $\|e - Hs\| \leq 1.5\sqrt{n} - .50$ also yields $h_j \leq O(2^L)$, $j = 1, \dots, n$.

We are now interested in generating β -approximate solutions to $Sh(\epsilon)$ for a sequence of values of $\epsilon > 0$ that converges to zero. The following Improvement Theorem shows that if \bar{x} and $(\bar{\pi}, \bar{s})$ are a β -approximate solution to $Sh(\epsilon)$, then a Newton step will generate new values of x and (π, s) such that x and (π, s) are a β -approximate solution to $Sh(\epsilon')$ where $\epsilon' = \alpha\epsilon < \epsilon$ (i.e., $\alpha < 1$).

Theorem 2.1 (Improvement Theorem). Suppose \bar{x} and $(\bar{\pi}, \bar{s})$ are a β -approximate solution to $Sh(\epsilon)$ for some $\epsilon > 0$ and $0 < \beta < 1$. Let

$$\alpha = \frac{\beta + \|e - H\bar{s}\|}{\sqrt{\beta + \|e - H\bar{s}\|}}$$

and let $\epsilon' = \alpha\epsilon$.

Then x and (π, s) are a β -approximate solution to $Sh(\epsilon')$ where x and (π, s) are defined as follows:

$$\text{Let } \bar{d} = \epsilon'h + \bar{x}$$

$$z = \bar{D} \left[I - \bar{D} A^T (A \bar{D}^2 A^T)^{-1} A \bar{D} \right] \left(e - \left(\frac{1}{\epsilon'} \right) \bar{D} c \right)$$

$$\text{Then } x = \bar{x} + z$$

$$\pi = (A \bar{D}^2 A^T)^{-1} A \bar{D} (\bar{D} c - \epsilon' e)$$

$$s = c - A^T \pi. \quad \blacksquare$$

Note above that if $\|e - H\bar{s}\|$ is small, then α will be small, and so ϵ' will be small relative to ϵ . Therefore, just as in Proposition 2.1, if h is chosen wisely so that $\|e - Hs\|$ can be bounded for all dual feasible values of s , then so can α . In fact, we have

Proposition 2.2. If $\|e - H\bar{s}\| \leq 1.5\sqrt{n} - 50$ and $\beta = .25$, then $\alpha \leq \left(1 - \frac{1}{6\sqrt{n}}\right)$.

Proof: From the definition of α , we have

$$\alpha \leq \frac{.25 + 1.5\sqrt{n} - 50}{.50 + 1.5\sqrt{n} - 50} = 1 - \frac{.25}{1.5\sqrt{n}} = 1 - \frac{1}{6\sqrt{n}}. \quad \blacksquare$$

Next note in the theorem that $x = \bar{x} + z$, where z is a Newton direction for the quadratic approximation to $Sh(\epsilon')$. In order to prove Theorem 2.1, we will need to prove that $\bar{d} > 0$, that $s > 0$, that

$$y = x + \epsilon' h > 0,$$

and that $r = e - \left(\frac{1}{\epsilon}\right) Ys$ satisfies $\|r\| < \beta$.

The method of proof draws on many of the constructions presented in Tseng [18], see also Roos and Vial [13]. The proof of Theorem 2.1 will follow as a consequence of the following sequence of lemmas.

Lemma 2.1. Under the hypothesis and notation of Theorem 2.1, $\bar{d} > 0$.

Proof: Let $\bar{y} = \bar{x} + \epsilon h$, and let $\bar{r} = e - \left(\frac{1}{\epsilon}\right) \bar{Y}\bar{s}$. Then because \bar{x} and $(\bar{\pi}, \bar{s})$ are a β -approximate solution to $Sh(\epsilon)$, $\|\bar{r}\| \leq \beta$. We can write

$$\begin{aligned} \bar{r} &= e - H\bar{s} - \left(\frac{1}{\epsilon}\right) \bar{X}\bar{s} \\ &= \alpha \left(e - H\bar{s} - \left(\frac{1}{\alpha\epsilon}\right) \bar{X}\bar{s} \right) + (1-\alpha)(e - H\bar{s}) \\ &= \alpha \left(e - H\bar{s} - \left(\frac{1}{\epsilon}\right) \bar{X}\bar{s} \right) + (1-\alpha)(e - H\bar{s}) \\ &= \alpha \left(e - \left(\frac{1}{\epsilon}\right) \bar{D}\bar{s} \right) + (1-\alpha)(e - H\bar{s}) \end{aligned}$$

$$\begin{aligned} \text{Thus, } \left\| e - \left(\frac{1}{\epsilon}\right) \bar{D}\bar{s} \right\| &\leq \frac{\|\bar{r}\|}{\alpha} + \frac{(1-\alpha)}{\alpha} \|e - H\bar{s}\| \\ &\leq \frac{\beta}{\alpha} + \frac{(1-\alpha)}{\alpha} \|e - H\bar{s}\| \\ &= \frac{\beta + \|e - H\bar{s}\|}{\alpha} - \|e - H\bar{s}\| \\ &= \sqrt{\beta} + \|e - H\bar{s}\| - \|e - H\bar{s}\| = \sqrt{\beta} < 1. \end{aligned}$$

Thus $\left|1 - \left(\frac{1}{\epsilon}\right) \bar{d}_j \bar{s}_j\right| < 1$, whereby $\bar{d}_j > 0$, since $\bar{s}_j > 0$, (see 2.3c), $j=1, \dots, n$. ■

Lemma 2.2. Under the hypothesis and notation of Theorem 2.1, $\|\bar{D}^{-1} z\| \leq \sqrt{\beta}$.

$$\begin{aligned} \text{Proof: } \bar{D}^{-1} z &= \left[I - \bar{D} A^T (A \bar{D}^2 A^T)^{-1} A \bar{D} \right] \left(e - \left(\frac{1}{\epsilon}\right) \bar{D} c \right) \\ &= \left[I - \bar{D} A^T (A \bar{D}^2 A^T)^{-1} A \bar{D} \right] \left(e - \left(\frac{1}{\epsilon}\right) \bar{D} (A^T \bar{\pi} + \bar{s}) \right) \\ &= \left[I - \bar{D} A^T (A \bar{D}^2 A^T)^{-1} A \bar{D} \right] \left(e - \left(\frac{1}{\epsilon}\right) \bar{D} \bar{s} \right), \end{aligned}$$

the latter equality following from the fact that the matrix in brackets projects onto the null space of $\bar{D} A^T$. Because the bracketed matrix is a projection matrix,

$$\begin{aligned} \|\bar{D}^{-1} z\| &\leq \left\| e - \left(\frac{1}{\epsilon}\right) \bar{D} \bar{s} \right\| = \left\| e - \left(\frac{1}{\alpha \epsilon}\right) (\alpha \epsilon H + \bar{X}) \bar{s} \right\| \\ &= \left\| e - H \bar{s} - \left(\frac{1}{\alpha \epsilon}\right) \bar{X} \bar{s} \right\| \\ &= \left\| \frac{1}{\alpha} \left(e - H \bar{s} - \left(\frac{1}{\epsilon}\right) \bar{X} \bar{s} \right) + \left(1 - \frac{1}{\alpha}\right) (e - H \bar{s}) \right\| \\ &\leq \frac{1}{\alpha} \left\| e - H \bar{s} - \left(\frac{1}{\epsilon}\right) \bar{X} \bar{s} \right\| + \frac{(1-\alpha)}{\alpha} \|e - H \bar{s}\| \\ &\leq \frac{\beta}{\alpha} + \frac{(1-\alpha)}{\alpha} \|e - H \bar{s}\|, \end{aligned}$$

where the last equality follows from the fact that \bar{x} and $(\bar{\pi}, \bar{s})$ are a β -approximate solution to $Sh(\epsilon)$, see (2.3b) and (2.3d). We now obtain

$$\|\bar{D}^{-1} z\| \leq \frac{\beta + \|e - H \bar{s}\|}{\alpha} - \|e - H \bar{s}\| = \sqrt{\beta}. \quad \blacksquare$$

Lemma 2.3. Under the hypotheses and notation of Theorem 2.1, define

$$y = x + \epsilon h \tag{2.4}$$

$$r = e - \left(\frac{1}{\epsilon}\right) Ys \tag{2.5}$$

Then $r = (\bar{D}^{-1} Z)^2 e$.

Proof: First note from the definition of z , x , π , and s in Theorem 2.1 that

$$s = c - A^T \pi = \epsilon' \bar{D}^{-1} (e - \bar{D}^{-1} z) . \quad (2.6)$$

From this expression we can write

$$\bar{D}^{-1} z = I - \left(\frac{1}{\epsilon'} \right) \bar{D} s ,$$

$$\text{and so } (\bar{D}^{-1} z)^2 = I - \left(\frac{2}{\epsilon'} \right) \bar{D} s + \left(\frac{1}{\epsilon'} \right)^2 \bar{D}^2 s^2 . \quad (2.7)$$

However, from (2.6) we also have

$$z = \bar{d} - \left(\frac{1}{\epsilon'} \right) \bar{D}^2 s ,$$

$$\text{and so } z s = \bar{D} s - \left(\frac{1}{\epsilon'} \right) \bar{D}^2 s^2 ,$$

$$\text{and } \bar{D}^2 s^2 = \epsilon' (\bar{D} s - z s) . \quad (2.8)$$

Substituting (2.8) into (2.7) gives

$$(\bar{D}^{-1} z)^2 = I - \left(\frac{1}{\epsilon'} \right) \bar{D} s - \left(\frac{1}{\epsilon'} \right) z s .$$

Finally,

$$\begin{aligned} (\bar{D}^{-1} z)^2 e &= e - \left(\frac{1}{\epsilon'} \right) \bar{D} s - \left(\frac{1}{\epsilon'} \right) z s \\ &= e - \left(\frac{1}{\epsilon'} \right) (\bar{X} + \epsilon' H) s - \left(\frac{1}{\epsilon'} \right) z s = e - H s - \left(\frac{1}{\epsilon'} \right) X s = r , \end{aligned}$$

because $x = \bar{x} + z$ and $r = e - H s - \left(\frac{1}{\epsilon'} \right) X s$. ■

Proof of Theorem 2.1. In order to prove the theorem, we need to show that x and (π, s) satisfy (2.3a) – (2.3d) .

$$(2.3a): \quad \text{Because } A z = 0 , \quad A x = A (\bar{x} + z) = A \bar{x} = b .$$

$$(2.3b): \quad \bar{D}^{-1} (x + \epsilon' h) = \bar{D}^{-1} (\bar{x} + \epsilon' h + z) = \bar{D}^{-1} (\bar{d} + z) = e + \bar{D}^{-1} z .$$

But $\| \bar{D}^{-1} z \| \leq \sqrt{\beta} < 1$ from Lemma 2.2 , whereby $\bar{D}^{-1} (x + \epsilon' h) = e + \bar{D}^{-1} z > 0$.

Thus $x + \epsilon' h > 0$, since \bar{D}^{-1} has positive components on the diagonal from Lemma 2.1.

(2.3d): Let y and r be defined as in (2.4) and (2.5). We must show that $\|r\| \leq \beta$.

From Lemma 2.3,

$$\|r\| = \|(\bar{D}^{-1} z)^T e\| \leq \|(\bar{D}^{-1} z)^T e\|_1 = \|(\bar{D}^{-1} z) e\|^2 = \|(\bar{D}^{-1} z)\|^2 \leq \beta,$$

from Lemma 2.2.

(2.3c): From (2.4) and (2.5) we have for $j=1, \dots, n$,

$$r_j = 1 - \left(\frac{1}{\epsilon'}\right) y_j s_j = 1 - \left(h_j + \left(\frac{1}{\epsilon'}\right) x_j\right) s_j, \text{ so that}$$

$$s_j \left(h_j + \left(\frac{1}{\epsilon'}\right) x_j\right) = 1 - r_j \geq 1 - \beta > 0, \text{ since } \|r\| \leq \beta.$$

However, $h_j + \left(\frac{1}{\epsilon'}\right) x_j > 0$ from (2.3b), whereby $s_j > 0$, $j=1, \dots, n$. ■

3. A Path-Following Algorithm for Shifted Barrier Functions

In this section, we utilize the Improvement Theorem (Theorem 2.1) as a basis for a path-following algorithm for linear programming using a shifted barrier function. The problem we are interested in solving is

$$\begin{aligned} \text{LP:} \quad & \text{minimize} && c^T x \\ & && x \\ & \text{s.t.} && Ax = b \\ & && x \geq 0. \end{aligned}$$

The dual of LP is

$$\begin{aligned} \text{LD:} \quad & \text{maximize} && b^T \pi \\ & && \pi, s \\ & && A^T \pi + s = c \\ & && s \geq 0. \end{aligned}$$

We presume in this section that we are given a positive shift vector $h \in \mathbb{R}^n$ and that we are interested in solving the problem $Sh(\epsilon)$ presented in the beginning to Section 2, for a decreasing sequence of values of $\epsilon > 0$ that converges to zero. We suppose for the moment that we are given an initial value of $x = x^0 \in \mathbb{R}^n$ and initial values of $(\pi, s) = (\pi^0, s^0)$ such that x^0 and (π^0, s^0) are a β -approximate solution to $Sh(\epsilon^0)$ for some given values of $\epsilon = \epsilon^0$ and of β . (This assumption will be relaxed in the next two sections.) Thus the data for the problem consists of the array $(c, A, b, h, x^0, \pi^0, s^0, \epsilon^0, \beta)$. The following algorithm is a sequential implementation of Theorem 2.1:

Algorithm Shifted Barrier $(c, A, b, h, x^0, \pi^0, s^0, \epsilon^0, \beta)$

0. $k = 0$
1. $\bar{x} = x^k, (\bar{\pi}, \bar{s}) = (\pi^k, s^k), \epsilon = \epsilon^k$
2. $\alpha = \frac{\beta + \|e - H\bar{s}\|}{\sqrt{\beta + \|e - H\bar{s}\|}}, \epsilon' = \alpha\epsilon$
3. $\bar{d} = \epsilon' h + \bar{x}, z = \bar{D} [I - \bar{D} A^T (A \bar{D}^2 A^T)^{-1} A \bar{D}] (e - (1/\epsilon') \bar{D} c)$
4. $x^{k+1} = \bar{x} + z$
 $\pi^{k+1} = (A \bar{D}^2 A^T)^{-1} A \bar{D} (\bar{D} c - \epsilon' e)$
 $s^{k+1} = c - A^T \pi^{k+1}$
 $\epsilon^{k+1} = \epsilon'$
5. $k = k + 1$. Go to 1.

Notice that the work per iteration of this algorithm is $O(n^3)$, which is the complexity of solving the least squares problem in Step 3. Also notice that performance of the algorithm hinges on being able to obtain the initial β -approximate solution x^0 and (π^0, s^0) . We defer discussion of this initialization issue until the next section. One measure of performance of this algorithm is given below:

Proposition 3.1. Suppose that all dual solutions (π, s) satisfy $\|e - Hs\| \leq 1.5\sqrt{n} - .50$, and that $\beta = .25$. Let $\epsilon^* > 0$ be a desired level of accuracy. Then algorithm Shifted Barrier will yield a $.25$ -approximate solution to $Sh(\epsilon)$ for some $\epsilon > 0$, $\epsilon \leq \epsilon^*$, after at most $K = \lceil 6\sqrt{n}(\ln(\epsilon^0) - \ln(\epsilon^*)) \rceil$ iterations.

Proof: Theorem 2.1 guarantees that x^k and (π^k, s^k) are a .25-approximate solution to $Sh(\epsilon^k)$ for $k=1, \dots$. From Proposition 2.2, $\alpha \leq \left(1 - \frac{1}{6\sqrt{n}}\right)$ at each iteration.

Let $K = \lceil 6\sqrt{n}(\ln(\epsilon^0) - \ln(\epsilon^*)) \rceil$. Then $\epsilon^K \leq \left(1 - \frac{1}{6\sqrt{n}}\right)^K \epsilon^0$, whereby

$$\ln(\epsilon^K) - \ln(\epsilon^0) \leq K \ln\left(1 - \frac{1}{6\sqrt{n}}\right) \leq -K/(6\sqrt{n}) \leq \ln(\epsilon^*) - \ln(\epsilon^0),$$

from which we obtain $\ln(\epsilon^K) \leq \ln(\epsilon^*)$, i.e., $\epsilon^K \leq \epsilon^*$. ■

Corollary 3.1. Suppose that in addition to the conditions of Proposition 3.1, that $h_j \leq O(2^L)$, $j = 1, \dots, n$. Then algorithm Shifted Barrier will generate a solution that can be rounded to an optimal solution to LP after at most

$$K = \lceil 6\sqrt{n}(\ln(\epsilon^0) + \ln O(L)) \rceil \text{ iterations.}$$

Proof: The proof is an immediate consequence of Proposition 3.1 and Corollary 2.2, with $\epsilon^* = O(2^{-2L})$. ■

In light of Proposition 3.1, the efficiency of the algorithm will depend on the choice of the shift vector h , and we seek a value of h that will ensure that $\|e - Hs\|$ is small for any dual feasible solution (π, s) . The efficiency of the algorithm will also depend on the initial value ϵ^0 , and we seek to keep ϵ^0 as small as possible. Thus we seek values of $x = x^0$ and $(\pi, s) = (\pi^0, s^0)$ so that x^0 and (π^0, s^0) are a .25-approximate solution to $Sh(\epsilon^0)$, where $\epsilon^0 > 0$ is preferably a small number. In Sections 4 and 5, we will examine ways to choose h and $x^0, (\pi^0, s^0)$, and ϵ^0 in an efficient manner.

4. Efficient Choice of the Shift Vector h and the Initial Values $x^0, (\pi^0, s^0)$, and ϵ^0 .

In this section, we present a method for choosing the shift vector h and the initial values $x^0, (\pi^0, s^0)$, and ϵ^0 . We will show that if this method is used for choosing h , then $\|e - Hs\| \leq 1.5\sqrt{n} - .5$ for all dual feasible solutions (π, s) , thus establishing the efficiency of the Shifted Barrier algorithm in terms of the geometric reduction constant α (See Proposition 2.2 and Proposition 3.1). We will also show that if \tilde{x} is a guess of a feasible or optimal solution to LP, then the initial values $x^0, (\pi^0, s^0)$, and ϵ^0 can be chosen so that ϵ^0 roughly measures the degree to which \tilde{x} is infeasible, and thus ϵ^0 will be a small number if \tilde{x} is almost feasible.

Before we present the results, we need to examine some concepts related to the center of the dual feasible region. Suppose that the feasible region of the dual LD is bounded. (This supposition will be relaxed in Section 5). The center of the dual feasible region is that value of $(\pi, s) = (\hat{\pi}, \hat{s})$ that solves the logarithmic barrier problem:

$$\begin{aligned} & \text{maximize} && \sum_{j=1}^n \ln(s_j) \\ & \pi, s \\ & \text{s.t.} && A^T \pi + s = c \\ & && s > 0, \end{aligned}$$

see Sonnevend [14, 15], also Vaidya [19], and [4]. If the feasible region of the dual LD is bounded and has an interior, the center $(\hat{\pi}, \hat{s})$ will exist uniquely. We are interested in working with a dual feasible solution $(\tilde{\pi}, \tilde{s})$ that is close to the center $(\hat{\pi}, \hat{s})$ of the dual feasible region in an appropriate measure of closeness. One measure of closeness is the length of the gradient of the negative logarithm barrier function

$$f(\pi) = - \sum_{j=1}^n \ln(c_j - (A^T \pi)_j).$$

We note that $\nabla f(\tilde{\pi}) = A \tilde{S}^{-1} e$ (where $\tilde{s} = c - A^T \tilde{\pi}$) and that the Hessian of $f(\tilde{\pi})$ is $\nabla^2 f(\tilde{\pi}) = A \tilde{S}^{-2} A^T$. We will say that $(\tilde{\pi}, \tilde{s})$ is a τ -approximate center of the dual feasible region if $(\tilde{\pi}, \tilde{s})$ is dual feasible and $\tilde{s} > 0$ and

$$\sqrt{\nabla f(\tilde{\pi})^T (\nabla^2 f(\tilde{\pi}))^{-1} \nabla f(\tilde{\pi})} = \sqrt{e^T \tilde{S}^{-1} A^T (A \tilde{S}^{-2} A^T)^{-1} A \tilde{S}^{-1} e} \leq \tau.$$

Thus $(\tilde{\pi}, \tilde{s})$ is a τ -approximate center if the norm of the gradient of $f(\tilde{\pi})$ is less than or equal to τ , where the norm is measured using the inverse of the Hessian of $f(\pi)$ at $\pi = \tilde{\pi}$. The next Lemma relates the notion of a τ -approximate center to a more standard measure of the distance of $(\tilde{\pi}, \tilde{s})$ to the center $(\hat{\pi}, \hat{s})$. We say that a dual feasible solution $(\tilde{\pi}, \tilde{s})$ is δ -close to the center $(\hat{\pi}, \hat{s})$ of the dual feasible region if $\tilde{s} > 0$ and $\|\tilde{S}^{-1}(\hat{s} - \tilde{s})\| \leq \delta$.

Lemma 4.1 (see [4], Lemma 7.2). Suppose $(\hat{\pi}, \hat{s})$ is the center of the dual feasible region and that $(\tilde{\pi}, \tilde{s})$ is a τ -approximate center of the dual feasible region, for $\tau \leq .08$. Then $(\tilde{\pi}, \tilde{s})$ is δ -close to the center $(\hat{\pi}, \hat{s})$ for $\delta = \sqrt{.55}$, i.e., $\|\tilde{S}^{-1}(\hat{s} - \tilde{s})\| \leq \sqrt{.55}$.

The proof of this Lemma is given in the Appendix. The complexity of computing a τ -approximate center for $\tau = .08$ is analyzed in Section 5. The following result is a partial converse of Lemma 4.1, and is also proved in the Appendix.

Lemma 4.2. Suppose $(\hat{\pi}, \hat{s})$ is the center of the dual feasible region and that $(\tilde{\pi}, \tilde{s})$ is δ -close to the center for $\delta \leq 1/21$. Then $(\tilde{\pi}, \tilde{s})$ is a τ -approximate center for $\tau \geq .075$.

We now turn our attention to the problem of choosing the shift vector h . An efficient choice of h is given in the following Lemma.

Lemma 4.3 (Choice of a shift vector h). Let $(\tilde{\pi}, \tilde{s})$ be δ -close to the center of the dual feasible region, where $\delta = 1/21$. Let $h = \left(\frac{1}{n}\right) \tilde{S}^{-1} e$. Then for all dual feasible values of (π, s) , $\|e - Hs\| \leq 1.5\sqrt{n} - 5$.

Proof: Let $(\hat{\pi}, \hat{s})$ be the center of the dual feasible region and let (π, s) be any dual feasible solution. Then from properties of the center (see Sonnevend [14, 15], also [3], Theorem 2.1), $\|\hat{S}^{-1}(s - \hat{s})\| \leq \sqrt{n(n-1)}$. Also from the hypothesis of the Lemma, $\|\tilde{S}^{-1}(\tilde{s} - \hat{s})\| \leq \delta$, where $\delta = 1/21$. Thus

$$\|\tilde{S}^{-1}(s - \hat{s})\| \leq \|\tilde{S}^{-1}\hat{S}\| \|\hat{S}^{-1}(s - \hat{s})\| \leq (1 + \delta) \|\hat{S}^{-1}(s - \hat{s})\| \leq (1 + \delta) \sqrt{n(n-1)}.$$

Therefore $\|\tilde{S}^{-1}(s - \tilde{s})\| \leq \|\tilde{S}^{-1}(s - \hat{s})\| + \|\tilde{S}^{-1}(\hat{s} - \tilde{s})\| \leq (1 + \delta) \sqrt{n(n-1)} + \delta$.

$$\begin{aligned} \text{Let } h &= \left(\frac{1}{n}\right) \tilde{S}^{-1} e. \text{ Then } \|e - Hs\| = \left\| e - \left(\frac{1}{n}\right) \tilde{S}^{-1} s \right\| \\ &= \left\| \left(\frac{1}{n}\right) \tilde{S}^{-1}(\tilde{s} - s) + \left(1 - \frac{1}{n}\right) e \right\| \leq \frac{1}{n} \|\tilde{S}^{-1}(\tilde{s} - s)\| + \|e\| \left(\frac{n-1}{n}\right) \\ &\leq \left(\frac{1}{n}\right) (1 + \delta) \sqrt{n(n-1)} + \frac{\delta}{n} + \sqrt{n} \left(\frac{n-1}{n}\right) = \left(\frac{22}{21}\right) \sqrt{\frac{n-1}{n}} + \frac{1}{21n} + \frac{n-1}{\sqrt{n}} \\ &\leq 1.5\sqrt{n} - 5 \text{ for } n \geq 2. \quad \blacksquare \end{aligned}$$

Corollary 4.1. With h chosen as in Lemma 4.3, $h_j \leq O(2^L)$, $j = 1, \dots, n$.

Proof: For each $j = 1, \dots, n$, let $M_j = \text{maximum } s_j$

$$\text{s.t. } A^T \pi + s = c$$

$$s \geq 0.$$

Then because the dual feasible region is bounded and has an interior (otherwise the center $(\hat{\pi}, \hat{s})$ would not exist), M_j is finite, and also $2^{-L} \leq M_j \leq 2^L$, see Papadimitrou and Steiglitz [10]. Furthermore, from properties of the center, we must have $\hat{s}_j \geq \left(\frac{1}{n-1}\right)M_j$, see Sonnevend [14, 15], also [3]. Now let $h = \left(\frac{1}{n}\right)\tilde{S}^{-1}e$, where $(\tilde{\pi}, \tilde{s})$ is δ -close to the center $(\hat{\pi}, \hat{s})$ for $\delta = 1/21$. Then $\|\tilde{S}^{-1}(\tilde{s}-\hat{s})\| \leq 1/21$ implies

$$\tilde{s}_j \geq \hat{s}_j \left(\frac{21}{22}\right) \geq \left(\frac{21}{22}\right)\left(\frac{M_j}{n-1}\right). \text{ Thus}$$

$$h_j = \frac{1}{n\tilde{s}_j} \leq \left(\frac{22}{21}\right)\left(\frac{n-1}{n}\right)\left(\frac{1}{M_j}\right) < \left(\frac{22}{21}\right)(2^L) = O(2^L), \quad j = 1, \dots, n. \quad \blacksquare$$

From Lemma 4.3, we obtain the iteration bound of Corollary 3.1 for processing a linear program with algorithm Shifted Barrier.

In Section 5, we will analyze the complexity of computing $h = \left(\frac{1}{n}\right)\tilde{S}^{-1}e$ efficiently (which is the complexity of computing a solution that is δ -close to the center of the dual feasible region).

We now turn our attention to choosing initial values of $x = x^\circ$, $(\pi, s) = (\pi^\circ, s^\circ)$, and $\epsilon = \epsilon^\circ$. We assume that we have a guess of a good value of x , which we denote by $x = \tilde{x} \in R^n$. The choice of \tilde{x} can be arbitrary, and in fact we need neither $A\tilde{x} = b$ nor $\tilde{x} \geq 0$. A good choice of \tilde{x} may be a feasible or optimal solution to a previous version of the linear program, that is (possibly) infeasible for the current version of the linear program. Once again, we assume that we have at hand a dual feasible solution $(\tilde{\pi}, \tilde{s})$ such that $(\tilde{\pi}, \tilde{s})$ is a δ -close center for the dual feasible region for $\delta = 1/21$. We now use $(\tilde{\pi}, \tilde{s})$ and \tilde{x} to define the initial values of algorithm Shifted Barrier as follows:

$$s^\circ = \tilde{s} \tag{4.1a}$$

$$\pi^\circ = \tilde{\pi} \tag{4.1b}$$

$$\epsilon^\circ = 8 \left\| \tilde{S}\tilde{x} - \tilde{S}^{-1}A^T(A\tilde{S}^{-2}A^T)^{-1}(A\tilde{x}-b) \right\| \tag{4.1c}$$

$$x^\circ = \tilde{S}^{-2}A^T(A\tilde{S}^{-2}A^T)^{-1}b + \tilde{S}^{-1}\left[I - \tilde{S}^{-1}A^T(A\tilde{S}^{-2}A^T)^{-1}A\tilde{S}^{-1}\right]\left(\epsilon^\circ\left(1 - \frac{1}{n}\right)e + \tilde{S}\tilde{x}\right) \tag{4.1d}$$

We will prove below that these initial values are a .25-approximate solution to $Sh(\epsilon^\circ)$. Note that in terms of efficiency of the Shifted Barrier algorithm, that the value of ϵ° is very important, and it should ideally be a small number. From (4.1c), we note that

$$\begin{aligned}\epsilon^\circ &\leq 8\|\tilde{S}\tilde{x}\| + 8\|\tilde{S}^{-1}A^T(A\tilde{S}^{-2}A^T)^{-1}(A\tilde{x}-b)\| \\ &= 8\|\tilde{S}\tilde{x}\| + 8\sqrt{(A\tilde{x}-b)^T(A\tilde{S}^{-2}A^T)^{-1}(A\tilde{x}-b)}\end{aligned}$$

In the above expression, ϵ° is bounded above by two quantities which indicate how far away x° is from being feasible. The second quantity is a measure of the distance from $A\tilde{x}$ to b and measures the infeasibilities of \tilde{x} in terms of the equations $Ax = b$ in the matrix norm $(A\tilde{S}^{-2}A^T)^{-1}$. The first quantity measures the length of \tilde{x} scaled by \tilde{S} . Note that the more negative a component \tilde{x}_j is, the larger the value of $\|\tilde{S}\tilde{x}\|$ is. Roughly speaking, the bound on ϵ° decreases the closer \tilde{x} lies to the region $\{x \mid Ax = b\}$ and to the region $\{x \mid x \geq 0\}$.

Theorem 4.1 (Initial Value Theorem). Suppose \tilde{x} is a guess of the value of a feasible solution to the primal LP. If $(\tilde{\pi}, \tilde{s})$ is a δ -close center of the dual feasible region for $\delta = 1/21$, and (π°, s°) , x° , and ϵ° are defined as in (4.1a) - (4.1d), then x° and (π°, s°) are a β -approximate solution to $Sh(\epsilon^\circ)$, where $\beta = .25$ and $h = \left(\frac{1}{n}\right)\tilde{S}^{-1}e$.

Proof: We must verify conditions (2.3a) - (2.3d) for the quantities ϵ° , x° , and (π°, s°) .

(2.3a): Direct multiplication of (4.1d) yields $Ax^\circ = b$.

(2.3c): Because $(\tilde{\pi}, \tilde{s})$ is δ -close to the center of the dual feasible region, $A^T\tilde{\pi} + \tilde{s} = c$ and $\tilde{s} > 0$. Thus $A^T\pi^\circ + s^\circ = c$ and $s^\circ > 0$.

(2.3d): Let $y^\circ = x^\circ + \epsilon^\circ h$ and $r^\circ = e - \left(\frac{1}{\epsilon^\circ}\right)(Y^\circ)s^\circ = e - Hs^\circ - \left(\frac{1}{\epsilon^\circ}\right)X^\circ s^\circ$. Substituting $h = \left(\frac{1}{n}\right)\tilde{S}^{-1}e$ and the values of s° and x° from (4.1) yields

$$\begin{aligned}r^\circ &= e - \left(\frac{1}{n}\right)\tilde{S}^{-1}\tilde{s} - \left(\frac{1}{\epsilon^\circ}\right)\tilde{S}^{-1}A^T(A\tilde{S}^{-2}A^T)^{-1}b - \left(\frac{1}{\epsilon^\circ}\right)\left(\epsilon^\circ\left(1 - \left(\frac{1}{n}\right)\right)e + \tilde{S}\tilde{x}\right) \\ &\quad + \left(\frac{1}{\epsilon^\circ}\right)\tilde{S}^{-1}A^T(A\tilde{S}^{-2}A^T)A\tilde{S}^{-1}\left(\epsilon^\circ\left(1 - \frac{1}{n}\right)e + \tilde{S}\tilde{x}\right) \\ &= -\left(\frac{1}{\epsilon^\circ}\right)\tilde{S}^{-1}A^T(A\tilde{S}^{-2}A^T)^{-1}(b - Ax^\circ) - \left(\frac{1}{\epsilon^\circ}\right)\tilde{S}\tilde{x} + \left(1 - \frac{1}{n}\right)\tilde{S}^{-1}A^T(A\tilde{S}^{-2}A^T)^{-1}A\tilde{S}^{-1}e.\end{aligned}$$

Thus

$$\begin{aligned} \|r^\circ\| &\leq \left(\frac{1}{\varepsilon^\circ}\right) \left\| \tilde{S}^{-1} A^T (A \tilde{S}^{-2} A^T)^{-1} (A \tilde{x} - b) - \tilde{S} \tilde{x} \right\| + \left\| \tilde{S}^{-1} A^T (A \tilde{S}^{-2} A^T)^{-1} A \tilde{S}^{-1} e \right\| \\ &= \left(\frac{\varepsilon^\circ}{8\varepsilon^\circ}\right) + \sqrt{e^T \tilde{S}^{-1} A^T (A \tilde{S}^{-2} A^T)^{-1} A \tilde{S}^{-1} e} . \end{aligned}$$

However, from Lemma 4.2, $(\tilde{\pi}, \tilde{s})$ is a τ -approximate center of the dual feasible region, for $\tau = .075$. Thus this last expression becomes

$$\|r^\circ\| \leq \frac{1}{8} + .075 < .25 .$$

(2.3b): Let $y^\circ = x^\circ + \varepsilon^\circ h$. Because $r^\circ = e - \left(\frac{1}{\varepsilon^\circ}\right)(Y^\circ)s^\circ$ and $\|r^\circ\| \leq .25$, we have

$$\begin{aligned} r_j^\circ &= 1 - y_j^\circ s_j^\circ / \varepsilon^\circ , \text{ or } y_j^\circ = (1 - r_j^\circ) \varepsilon^\circ / s_j^\circ , \quad j=1, \dots, n . \text{ Now } s_j^\circ > 0 , \varepsilon^\circ > 0 \text{ and} \\ r_j^\circ &< \|r^\circ\| \leq .25 , \text{ whereby } y_j^\circ > 0 , \quad j=1, \dots, n . \quad \blacksquare \end{aligned}$$

In conclusion, Lemmas 4.2, 4.3, and Theorem 4.1 point out the fact that having a δ -close center $(\tilde{\pi}, \tilde{s})$ of the dual feasible region provides us with an efficient choice of the shift vector $h = \left(\frac{1}{n}\right)\tilde{S}^{-1} e$ as well as efficient initial values of x° , (π°, s°) , and ε° for initiating the Shifted Barrier algorithm of Section 3. Lemma 4.2 relates the value of $\delta \leq \frac{1}{21}$ to the value of $\tau \leq .075$. Lemma 4.3 states that if $h = \left(\frac{1}{n}\right)\tilde{S}^{-1} e$, then all dual feasible points (π, s) satisfy $\|e - Hs\| \leq 1.5\sqrt{n} - .50$, whereby from Proposition 2.2, $\alpha \leq 1 - \frac{1}{6\sqrt{n}}$ at each iteration of the algorithm, yielding the complexity measure of the algorithm that is presented in Proposition 3.1. Finally, Theorem 4.1 shows that if $(\tilde{\pi}, \tilde{s})$ is a δ -close center, then the algorithm can be initiated with values of x° , (π°, s°) , and ε° given in (4.1) and that the value of ε° roughly reflects the degree of infeasibility of the given vector \tilde{x} in terms of the satisfiability of the equations $Ax = b$ as well as the nonnegativity conditions $x \geq 0$. In this next section, we present a method for computing a δ -close center $(\tilde{\pi}, \tilde{s})$ of the dual feasible region for $\delta \leq \frac{1}{21}$ and the complexity of this method is analyzed as well.

5. Analysis of a Method for Finding a δ -close Center $(\bar{\pi}, \bar{s})$ of the Dual Feasible Region for $\delta \leq 1/21$.

This section analyzes the use of two algorithms that can be used together to find a δ -close center of the dual feasible region, namely

$$Y = \left\{ (\pi, s) \in \mathbb{R}^m \times \mathbb{R}^n \mid A^T \pi + s = c, s \geq 0 \right\} .$$

In the previous section, we assumed Y was bounded. However, there is no loss of generality in assuming the boundedness of Y so long as we are given a bound B on the optimal objective value of LP . (In practice, such a bound is usually easy to obtain based on a good understanding of the problem at hand.) Then we can replace LP and LD by the problems:

$$LP': \quad \text{minimize} \quad c^T x - B x_{n+1}$$

$$x, x_{n+1}$$

$$Ax - b x_{n+1} = b$$

$$x \geq 0, x_{n+1} \geq 0$$

$$LD': \quad \text{maximize} \quad b^T \pi$$

$$\pi, s, s_{n+1}$$

$$A^T \pi + s = c$$

$$-b^T \pi + s_{n+1} = -B$$

$$s \geq 0, s_{n+1} \geq 0 .$$

The following Lemma shows the equivalence of LP to LP' as well as the boundedness of LD' .

Lemma 5.1 (Equivalency of LP to LP' and Boundedness of LD'). Suppose that the set of optimal solutions to LP is nonempty and bounded, and that B is a strict lower bound on the optimal objective value z^* of LP . Let $\bar{s}_{n+1} = z^* - B$.

- (i) $(\bar{x}, \bar{x}_{n+1}) = (\bar{x}, 0)$ and $(\bar{\pi}, \bar{s}, \bar{s}_{n+1})$ are a pair of optimal primal-dual solutions to LP' and LD' if and only if \bar{x} and $(\bar{\pi}, \bar{s})$ are a pair of optimal primal-dual solutions to LP and LD .
- (ii) The feasible region of LD' is bounded. Furthermore, there exists a feasible solution (π, s, s_{n+1}) of LD' with $s > 0$ and $s_{n+1} > 0$.

- Proof:** (i) Because B is strict lower bound on the optimal objective value z^* of LP, the last constraint of LD' will never be binding in an optimal solution to LD', i.e., $b^T \bar{\pi} > B$ in any optimal solution $(\bar{\pi}, \bar{s}, \bar{s}_{n+1})$ to LD'. Thus $\bar{s}_{n+1} > 0$ and $\bar{x}_{n+1} = 0$ (from complementarity) in any optimal solution to LP. The rest of the assertion follows in a straightforward manner.
- (ii) By hypothesis, the set of optimal solutions to LP is nonempty and bounded. Hence, by a theorem of the alternative, it is straightforward to show that there exists π^1 for which $A^T \pi^1 < c$ and that the feasible region of LD' is bounded. Furthermore, since B is a strict lower bound on the optimal objective value of LD (and hence LD', from (i)), there exists π^2 for which $A^T \pi^2 \leq c$ and $b^T \pi^2 > B$. By taking the appropriate convex combination of π^1 and π^2 , we obtain a vector π for which $A^T \pi < c$ and $b^T \pi > B$. ■

With Lemma 5.1 in mind, we now assume throughout this section that the feasible region of LD is bounded and has an interior, i.e., there exists a point (π, s) for which $A^T \pi + s = c$ and $s > 0$. Therefore, the center $(\hat{\pi}, \hat{s})$ of the dual feasible region exists uniquely, and we can now concentrate on finding a δ -close center of the dual feasible region. As in the case of analysis of the primal, we suppose that we have a guess $\bar{\pi}$ of the value of $\hat{\pi}$, and that $\bar{\pi}$ will be the starting point of a method for finding a δ -close center. We do not assume that $A^T \bar{\pi} \leq c$, i.e., that $\bar{\pi}$ is feasible for the dual. We will make use of two different center-finding algorithms. Each is described below.

Algorithm PT

The first algorithm we will utilize is a projective transformation-based algorithm for computing an approximation to the center of a given system of linear inequalities $A^T \pi \leq g$ starting from a given initial interior solution to that system. The algorithm is described and analyzed in [4]. We will call this algorithm PT for "projective transformation" algorithm. At each iteration, algorithm PT performs a projective transformation, and then computes a direction d from the current feasible solution $\bar{\pi}$ by solving a system of equations corresponding to a least-squares problem. (Thus the work per iteration is $O(n^3)$ operations.) A steplength α is then computed, either by an analytic formula (much as in Karmarkar's algorithm), or by performing a line-search. The new iterate is $\bar{\pi} \leftarrow \bar{\pi} + \alpha d$. Performance of the algorithm can be measured by considering the difference in the logarithmic barrier function

$$f(\pi) = \sum_{j=1}^n \ln(g_j - (A^T \pi)_j)$$

at the starting point $\pi = \bar{\pi}$ and at the center $\hat{\pi}$, as follows:

Lemma 5.2 (Complexity of Algorithm PT). Suppose $\bar{\pi}$ satisfies $\bar{s} = g - A^T \bar{\pi} > 0$, and that algorithm PT is initiated at $\pi = \bar{\pi}$. Suppose $(\hat{\pi}, \hat{s})$ is the center of the system $A^T \pi \leq g$. Then algorithm PT will compute a δ -close center $(\tilde{\pi}, \tilde{s})$ of the system

$A^T \pi \leq g$ after at most $K = 7 + \left\lceil \left(\frac{n-1}{n} \right) \left(\frac{f(\hat{\pi}) - f(\bar{\pi})}{.0033} \right) \right\rceil$ iterations, where $\delta = \frac{1}{21}$, and

the work per iteration is at most $O(n^3)$ operations (i.e., the complexity of computing the direction). ■

Note that the number of iterations is bounded by $7 + (f(\hat{\pi}) - f(\bar{\pi})) / .0033$, which is independent of n , the number of inequalities. This bound indicates that if $f(\bar{\pi})$ is large, i.e., $f(\bar{\pi})$ is close to $f(\hat{\pi})$, then the number of iterations will be small. Thus the bound on the number of iterations is roughly proportional to how close $\bar{\pi}$ is to $\hat{\pi}$ in the measure of the logarithmic barrier objective function $f(\pi)$. It should be noted that the bound above is probably not very tight in practice so long as the projective transformation algorithm is implemented with a line search. In that case, the author's own experience on small problems indicates that the algorithm converges to the center at least as efficiently as Karmarkar's algorithm converges to solutions of a linear program. (This is not surprising, because both algorithms are based on the same projective transformation methodology and strategy, see [4].) The proof of Lemma 5.2 is given in the Appendix.

Algorithm PCP

The second algorithm we will utilize is a path-following algorithm for tracing the path of centers $(\hat{\pi}_t, \hat{s}_t)$ to the system $A^T \pi \leq (g + dt)$ as the scalar parameter t is varied over a given range. The algorithm is described in [5]. We will call this algorithm PCP for "parametric center problem." At a given iteration k , the value of t is $t = t^k$. The current point $(\bar{\pi}^k, \bar{s}^k)$ is a δ -close center of the system $A^T \pi \leq g + dt^k$ for

$\delta = \frac{1}{21}$. A constant α is computed, which is the increase in the value of t , and the new value of t is computed as $t^{k+1} = t^k + \alpha$. A Newton step is then computed and a new value of π is chosen, namely $\bar{\pi}^{k+1}$, whose slacks are $\bar{s}^{k+1} = g + dt^{k+1} - A^T \bar{\pi}^{k+1}$. The new

value $(\bar{\pi}^{k+1}, \bar{s}^{k+1})$ is a δ -close center of the system $A^T \pi \leq g + dt^{k+1}$, for $\delta = \frac{1}{21}$.

The work per iteration of the algorithm is $O(n^3)$. Performance of algorithm PCP is measured as follows:

Lemma 5.3 (Complexity of Algorithm PCP). Suppose $(\bar{\pi}, \bar{s})$ is a δ -close center of the system $A^T \pi \leq g + dt$ for $\delta = \frac{1}{21}$, at $t = 0$. Suppose that $d < 0$, and define

$$T_{MAX} = \text{maximum } t \\ \pi, t \\ \text{s.t. } A^T \pi \leq g + dt .$$

Suppose $T_{MAX} > 1$.

Then after at most $K = \left\lceil 128 n \ln \left(\frac{T_{MAX}}{T_{MAX} - 1} \right) \right\rceil$ iterations of algorithm PCP, the

algorithm will compute of δ -close center of the system $A^T \pi \leq g + dt$ for $\delta = \frac{1}{21}$, at $t = 1$. ■

Note that the value of K increases linearly with n . The following discussion is an interpretation of K . Because $(\bar{\pi}, \bar{s})$ is a δ -close center of the system $A^T \pi \leq g$, then the set $Y_0 = \{\pi \in \mathbb{R}^m \mid A^T \pi \leq g\}$ is bounded. As t is increased,

$Y_t = \{\pi \in \mathbb{R}^m \mid A^T \pi \leq g + dt\}$ shrinks, because $d < 0$ and the RHS is strictly decreasing, i.e., $Y_t \subset Y_{t'}$ for $t' > t$. Furthermore, T_{MAX} is guaranteed to be finite. The

quantity $\ln \left(\frac{T_{MAX}}{T_{MAX} - 1} \right)$ measures how close the set Y_1 (i.e., Y_t at $t = 1$) is to the set

Y_0 . If T_{MAX} is large, increasing t from $t = 0$ to $t = 1$ will not contract the boundary of Y_t very much on a relative basis. Thus Y_0 and Y_1 are shaped similarly, and so their

centers should be near to one another. Because T_{MAX} is large, $\ln \left(\frac{T_{MAX}}{T_{MAX} - 1} \right)$ will be small.

Conversely, if T_{MAX} is small, (e.g., if $T_{MAX} = 1 + \epsilon$ for some small ϵ), then Y_1 will be a substantial contraction of Y_0 , and the centers of Y_0 and Y_1 may be very far from one

another. Because T_{MAX} is small, $\ln\left(\frac{T_{MAX}}{T_{MAX} - 1}\right) = \ln\left(\frac{1 + \epsilon}{\epsilon}\right) = \ln\left(\frac{1}{\epsilon} + 1\right)$ will be large.

The proof of Lemma 5.3 is given in the Appendix.

We now are ready to present the method for finding a δ -close center $(\tilde{\pi}, \tilde{s})$ of the dual feasible region from a given starting point. Suppose $(\hat{\pi}, \hat{s})$ is the center of the dual feasible region. As in the case of Section 4 for the primal, we assume that we have a given initial value of $\pi = \bar{\pi}$ that is a guess of the value of $\hat{\pi}$. If $\bar{\pi}$ is dual feasible and $A^T \bar{\pi} < c$ we can use algorithm PT to find a dual feasible point $(\tilde{\pi}, \tilde{s})$ that is a δ -close

center of the dual feasible region in $K = 7 + \left\lceil \frac{f(\hat{\pi}) - f(\bar{\pi})}{.0033} \right\rceil$ iterations, where

$$f(\pi) = \sum_{j=1}^n \ln(c_j - (A^T \pi)_j), \text{ according to Lemma 5.2, where } \delta = \frac{1}{21}.$$

Suppose, however, that $\bar{\pi}$ does not satisfy $A^T \bar{\pi} < c$. Then the strategy we propose is to replace the RHS c by a vector g for which $A^T \bar{\pi} < g$. We can then use algorithm PT to find a point $(\bar{\pi}, \bar{s})$ near the center of the system $A^T \pi \leq g$, and then use algorithm PCP to trace a sequence of points near the center of the system $A^T \pi \leq g + t(c - g)$ as t is increased from $t = 0$ to $t = 1$. At the final iterate, we will have a point $(\tilde{\pi}, \tilde{s})$ that is near the center of the system $A^T \pi \leq g + 1(c - g) = c$. The method is as follows:

Step 1. Let $g \in R^n$ be any vector that satisfies $g > c$ and $g > A^T \bar{\pi}$, for example $g_j = \max\{c_j + 1, (A^T \bar{\pi})_j + 1\}$, $j=1, \dots, n$.

Step 2. Use algorithm PT to find a δ -close center $(\bar{\pi}, \bar{s})$ of the system $A^T \pi \leq g$, for $\delta = \frac{1}{21}$.

Step 3. Define $d = c - g$. Use algorithm PCP to generate a sequence of δ -close center points of the system $A^T \pi \leq g + dt$ for $t \in [0, 1]$, for $\delta = \frac{1}{21}$.

At the final iterate, we will have a point $(\tilde{\pi}, \tilde{s})$ that is a δ -close center of the system $A^T \pi \leq g + dt$ at $t = 1$, i.e. $A^T \pi \leq g + (c - g) = c$, and so $(\tilde{\pi}, \tilde{s})$ will be a

δ -close center of the dual feasible region for $\delta = \frac{1}{21}$. The complexity of the above method is as follows, whose proof is an immediate consequence of Lemmas 5.2 and 5.3.

Lemma 5.4 (Complexity of computing a δ -close center). Using the above method for computing a δ -close center of the dual feasible region for $\delta = \frac{1}{21}$, the total number of iterations is bounded above by $K = K^1 + K^2$, where

$$(i) \quad K^1 = 7 + \left\lceil \frac{f(\hat{\pi}) - f(\bar{\pi})}{.0033} \right\rceil$$

where $(\hat{\pi}, \hat{s})$ is the center of system $A^T \pi \leq g$, and

$$f(\pi) = \sum_{j=1}^n \ln(g_j - (A^T \pi)_j), \text{ and}$$

$$(ii) \quad K^2 = \left\lceil 128 n \ln \left(\frac{T_{MAX}}{T_{MAX} - 1} \right) \right\rceil, \text{ where}$$

$$T_{MAX} = \text{maximum } t$$

$$\pi, t$$

$$A^T \pi \leq g + dt. \quad \blacksquare$$

Note that in the above method, the choice of g is fairly arbitrary.

As was discussed in this section, the value of K^1 roughly measures how close $\bar{\pi}$ is to the center of the system $A^T \pi \leq g$, and the value of K^2 roughly measures how close the center of the system $A^T \pi \leq g$ is to the center of the system $A^T \pi \leq c$. Thus $K = K^1 + K^2$ roughly measures how close $\bar{\pi}$ is to the center of the system $A^T \pi \leq c$.

Through Lemma 5.4, we have a method that will compute of δ -close center (for $\delta = \frac{1}{21}$) of the dual feasible region, from any starting point, and whose complexity roughly corresponds to how close the starting point is to the actual center of dual feasible region.

Appendix

The purpose of this Appendix is to prove Lemmas 4.1 and 4.2, and Lemmas 5.2 and 5.3. The results in these four Lemmas are slight modifications of results contained in the papers [4] and [5], but with different notation. Thus, none of these results in the Appendix are truly new to this paper. There is a problem, however, in proving these Lemmas in a brief yet cogent manner, because the notation in the papers [4] and [5] vary substantially from that of this paper. Therefore, we begin with a discussion of notational issues.

We first start with the algorithm PT presented in paper [4]. In Section 2 of that paper, an algorithm is presented for solving the following center problem

$$\begin{aligned} & \underset{x}{\text{maximize}} && F(x) = \sum_{i=1}^m w_i \ln(b_i - A_i x) \\ & \text{subject to} && Ax < b \\ & && Mx = g, \end{aligned}$$

where A is an $m \times n$ matrix, and where $w = (w_1, \dots, w_m)^T$ is a vector of positive weights that satisfy $e^T w = 1$. Adapting this problem to the problem of finding the center of dual feasible region presented in Sections 4 and 5 of this study, we replace A by A^T , b by c , and interchange the roles of m and n , replace x by π , delete M and g , set $w = (1/n)e$ and

note that $F(x) = \left(\frac{1}{m}\right) \sum_{i=1}^m \ln(b_i - A_i x)$, which in the notation of this paper is

$F(x) = \left(\frac{1}{n}\right) f(\pi) = \left(\frac{1}{n}\right) \sum_{j=1}^n \ln(c_j - (A^T \pi)_j)$. Therefore, when citing results from [4] about

$F(x)$, we can replace $F(x)$ by $\left(\frac{1}{n}\right) f(\pi)$. Two other key notational points in the algorithm

are the definition in [4] of $\bar{w} = \min_i \{w_i\} = \left(\frac{1}{n}\right)$ in the notation of this paper, and

$k = \bar{w}/(1 - \bar{w}) = 1/(n - 1)$ in the notation of this paper. Performance and analysis of the algorithm PT in [4] frequently makes use of the quantity γ which is a constant defined in

Step 3 and Step 4 of the algorithm PT in [4]. As it turns out, the constant γ is intimately related to the value of τ defined in Section 4 of this paper, as follows:

Proposition A.1 (Values of γ and τ). Let $(\bar{\pi}, \bar{s})$ be a dual feasible solution that satisfies

$\bar{s} > 0$, and let $f(\pi) = \sum_{j=1}^n \ln(c_j - (A^T \pi)_j)$. Let τ be the quantity

$$\tau = \sqrt{(\nabla f(\bar{\pi}))^T (\nabla^2 f(\bar{\pi}))^{-1} \nabla f(\bar{\pi})} = \sqrt{e^T \bar{S}^{-1} A^T (A \bar{S}^{-2} A^T)^{-1} A \bar{S}^{-1} e}.$$

If $\bar{\pi}$ is an iterate of the algorithm PT presented in [4], let γ be the value defined in Step 3 and 4 of the algorithm for $\pi = \bar{\pi}$. Then

$$\gamma = \left(\sqrt{\frac{n-1}{n-\tau^2}} \right) \tau.$$

Proof: In the notation of this study, at Step 3 of the algorithm PT (see [4]), we must solve

$$\begin{aligned} & \text{maximize} && -y^T d \\ & && d \end{aligned}$$

$$\text{s.t. } d^T (A - y \bar{s}^T) (\bar{S}^{-1} W \bar{S}^{-1}) (A^T - \bar{s} y^T) d \leq k,$$

where $\bar{s} = c - A^T \bar{\pi}$ and $y = \left(\frac{1}{n}\right) A \bar{S}^{-1} e$. Let $Q = A \bar{S}^{-1} W \bar{S}^{-1} A^T$ and

$\tilde{Q} = Q - y y^T$, and note that $k = 1/(n-1)$, $w = (1/n)e$, and $W = (1/n)I$.

Then direct substitution shows the above problem is

$$\begin{aligned} & \text{maximize} && -y^T d \\ & && d \end{aligned}$$

$$\text{s.t. } d^T \tilde{Q} d \leq k,$$

and the optimal value of d is $\bar{d} = \frac{-\tilde{Q}^{-1} y \sqrt{k}}{\sqrt{y^T \tilde{Q}^{-1} y}}$. Hence

$\gamma = (-y^T \bar{d})/k = \sqrt{y^T \tilde{Q}^{-1} y/k}$. Next note that because \tilde{Q} is a rank-one modification of Q , then from the Sherman-Morrison formula,

$$\tilde{Q}^{-1} = Q^{-1} + \frac{Q^{-1} y y^T Q^{-1}}{1 - y^T Q^{-1} y},$$

$$\text{and so } \gamma = \sqrt{\frac{y^T Q^{-1} y/k}{1 - y^T Q^{-1} y}}.$$

However $Q = A \bar{S}^{-1} W \bar{S}^{-1} A^T = \left(\frac{1}{n}\right) A \bar{S}^{-2} A^T = \left(\frac{1}{n}\right) \nabla^2 f(\bar{\pi})$,

and $y = \left(\frac{1}{n}\right) A \bar{S}^{-1} e = \left(\frac{1}{n}\right) \nabla f(\bar{\pi})$, and so

$y^T Q^{-1} y = \left(\frac{1}{n}\right) \nabla f(\bar{\pi}) \left(\nabla^2 f(\bar{\pi})\right)^{-1} \nabla f(\bar{\pi}) = \frac{1}{n} \tau^2$. Substituting this last expression and

remembering that $k = 1/(n-1)$ yields

$$\gamma = \sqrt{\frac{\left(\frac{(n-1)/n}{1 - \frac{1}{n} \tau^2}\right) \tau^2}{1 - \frac{1}{n} \tau^2}} \tau,$$

which is $\gamma = \left(\sqrt{\frac{n-1}{n - \tau^2}}\right) \tau$. ■

We also need to translate some notation from the paper [5]. In that paper, if $A \bar{x} + \bar{s} = b$, $\bar{s} > 0$, the quantity $\|v\|_{Q(\bar{x})} = \left\|v^T A^T \bar{S}^{-2} A v\right\|$ is defined. In our

notation, this is $\|v\|_{Q(\bar{\pi})} = \left\|v^T A \bar{S}^{-2} A^T v\right\|$ where $\bar{s} = c - A^T \bar{\pi}$.

Proof of Lemma 4.1: The proof is based on Lemma 3.3 of [5] . Suppose $(\tilde{\pi}, \tilde{s})$ is a τ -approximate center where $\tau = .08$. Then $\gamma \leq \tau \leq .08$ from Proposition A.1. Let $h = .5$. Then in the terminology of Lemma 3.3 of [5] ,

$$\|\hat{\pi} - \tilde{\pi}\|_{Q(\tilde{w})}^2 \leq \left(\frac{n}{n-1}\right) \frac{h^2(1+\gamma^2)}{(1-h\gamma)^2} \leq .55 ,$$

where $\|\hat{\pi} - \tilde{\pi}\|_{Q(\tilde{w})}$ is defined to be $\sqrt{(\hat{\pi} - \tilde{\pi})^T A \tilde{S}^{-2} A^T (\hat{\pi} - \tilde{\pi})} = \|\tilde{S}^{-1}(\hat{s} - \tilde{s})\|$. Thus

$$\sqrt{(\hat{\pi} - \tilde{\pi})^T A \tilde{S}^{-2} A^T (\hat{\pi} - \tilde{\pi})} = \|\tilde{S}^{-1}(\hat{s} - \tilde{s})\| \leq \sqrt{.55} . \quad \blacksquare$$

Proof of Lemma 4.2: The proof is based on Lemma 3.4 of [5] . Suppose $(\tilde{\pi}, \tilde{s})$ is a δ -close center and $\delta = \frac{1}{21}$. Then $\|\hat{\pi} - \tilde{\pi}\|_{Q(\tilde{w})} \leq \frac{1}{21}$. Thus from Lemma 3.4 of [5] , with

$a = \frac{\delta^2}{2(1-\delta)(1-2\delta)}$, we obtain $\gamma \leq .0527$. But from Proposition A.1, $\tau \leq \sqrt{2} \gamma$, so $\gamma \leq .075$. \blacksquare

Towards proof of Lemma 5.1, we have the following two Propositions.

Proposition A.2: If $\bar{\pi}$ is the current iterate of the algorithm PT and the value of γ is computed and $\gamma = \bar{\gamma} \leq .08567$, then in all subsequent iterations, we will have the value of $\gamma \leq .08567$.

Proof: This proof is based on Lemmas 4.1 and 4.2 of [4] . Let $\bar{\pi}$ and $\tilde{\pi}$ be two successive iterates of the algorithm PT , and let $\bar{\gamma}$ and $\tilde{\gamma}$ be the corresponding values of γ produced at Steps 3 and 4 in those iterations. Suppose $\bar{\gamma} \leq .08567$ but $\tilde{\gamma} > .08567$. Then from Lemma 4.1(ii) of [4] , $f(\tilde{\pi}) \leq f(\bar{\pi}) + .669(kn)\bar{\gamma}^2$.

On the other hand, from Lemma 4.2(ii) of [4] ,

$$f(\tilde{\pi}) \geq f(\bar{\pi}) + .4612(kn)\bar{\gamma}^2 ,$$

and from Lemma 4.2(i) of [4] ,

$$f(\hat{\pi}) \geq f(\tilde{\pi}) + .0033(kn) .$$

Combining these yields

$$\begin{aligned} (kn) .669 \bar{\gamma}^2 &\geq f(\hat{\pi}) - f(\bar{\pi}) = f(\hat{\pi}) - f(\tilde{\pi}) + f(\tilde{\pi}) - f(\bar{\pi}) \\ &\geq .0033(kn) + .4612(kn) \bar{\gamma}^2 . \end{aligned}$$

Thus $\bar{\gamma}^2 \geq \frac{.0033}{.669 - .4612}$, or $\bar{\gamma} > .125$, a contradiction. Thus if $\bar{\gamma} \leq .08567$, then

$$\tilde{\gamma} \leq .08567 . \quad \blacksquare$$

Proposition A.3: Under the hypothesis of Proposition A.2 ,

$$\tilde{\gamma} = .68 \bar{\gamma} .$$

Proof: We have from Lemma 4.1(ii) and Lemma 4.2(ii) of [4] ,

$$\begin{aligned} (kn) .4612 \tilde{\gamma}^2 &\leq f(\tilde{\pi}) - f(\bar{\pi}) = f(\hat{\pi}) - f(\bar{\pi}) - (f(\hat{\pi}) - f(\tilde{\pi})) \\ &\leq .669(kn) \bar{\gamma}^2 - .4612(kn) \tilde{\gamma}^2 . \end{aligned}$$

Thus $\tilde{\gamma}^2 \leq \frac{.669 - .4612}{.4612} \bar{\gamma}^2$, so that $\tilde{\gamma} \leq .68 \bar{\gamma}$. \blacksquare

Proposition A.2 states that once γ drops below $\gamma = .08567$ in algorithm PT , then it decreases at least by a factor of .68 at all subsequent iterations.

Proof of Lemma 5.2: From Remark 7.2 of [4] , algorithm PT must have $\gamma < .08567$ after

at most $\left\lceil \left(\frac{1}{nk} \right) \left(\frac{f(\hat{\pi}) - f(\bar{\pi})}{.0033} \right) \right\rceil$ iterations, i.e., after at most $\left\lceil \left(\frac{n-1}{n} \right) \left(\frac{f(\hat{\pi}) - f(\bar{\pi})}{.0033} \right) \right\rceil$

iterations. Then after at most seven additional iterations, Proposition A.3 ensures that at the current iterate, $\gamma \leq (.68)^7 (.08567) < .0058$. Then from Lemma 3.3 of [5] with $h = .03$, we obtain

$$\|\tilde{S}^{-1}(\hat{s} - \tilde{s})\| = \sqrt{(\tilde{\pi} - \hat{\pi})^T A \tilde{S}^{-2} A^T (\tilde{\pi} - \hat{\pi})} = \|\tilde{\pi} - \hat{\pi}\|_{Q(\tilde{\pi})} \leq \frac{1}{21},$$

i.e., $(\tilde{\pi}, \tilde{s})$ is a δ -close center for $\delta = \frac{1}{21}$. ■

Proof of Lemma 5.3: The proof of Lemma 5.3 is an application of Lemma 2.5 of [5]. At each iterate of algorithm PCP, the algorithm computes either a finite upper bound on T_{MAX} at Step 4, or a finite lower bound on T_{MIN} at Step 4, or both, where

$$\begin{aligned} T_{MIN} = & \text{minimum } t \\ & \pi, s, t \\ \text{s.t.} & \quad A^T \pi + s = g + dt \\ & \quad s \geq 0. \end{aligned}$$

However, if $d < 0$, $T_{MIN} = -\infty$, and so we can apply Lemma 2.5 of [5] with $t \in [\underline{t}, \bar{t}] = [0, 1]$, i.e., $\underline{t} = 0$ and $\bar{t} = 1$. Thus the algorithm will stop after at most

$$K = \left\lceil 128 n \ln \left(\frac{T_{MAX}}{T_{MAX} - 1} \right) \right\rceil \text{ iterations.} \quad \blacksquare$$

References

- [1] Anstreicher, K. M. (1986). "A monotonic projective algorithm for fractional linear programming," *Algorithmica* I, pp. 483-498.
- [2] Anstreicher, K. M. (1986). "A combined Phase I – Phase II projective algorithm for linear programming," Yale School of Management, New Haven, CT.
- [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] Freund, R. M., and K. C. Tan (1989) . "A method for the parametric center problem, with a strictly Polynomial-time algorithm for linear programming" Working paper OR 182-89, Operations Research Center, M.I.T., Cambridge, MA.
- [6] Gay, D. M. (1987). "A variant of Karmarkar's linear programming algorithm for problems in standard form," *Mathematical Programming* 37, pp. 81-90.
- [7] Gill, P., W. Murray, M. Saunders, and M. Wright (1989). "Shifted Barrier Methods for Linear Programming," forthcoming.
- [8] Gonzaga, C. (1989). "A conical projection algorithm for linear programming," *Mathematical Programming* 43, pp. 151-174.
- [9] Montiero, R. C. and I. Adler (1987). "An $O(n^3L)$ interior point algorithm for convex quadratic programming," Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA.
- [10] Papadimitriou, P., and K. Steiglitz (1982). Combinatorial Optimization: Algorithms and Complexity, Prentice Hall.
- [11] Polyak, R. (1988). Informal presentation at the Joint Summer Research Conference: Mathematical Developments from Linear Programming, Bowdoin College, Brunswick, ME.
- [12] Renegar, J. (1988). "A polynomial algorithm, based on Newton's method, for linear programming," *Mathematical Programming* 40 (1), 59-94.
- [13] Roos, C., and J. P. Vial (1988). "A simple polynomial method of approximate centres for linear programming," Department of Mathematics and Computer Science, Delft University of Technology, Delft, Netherlands.
- [14] 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 Körut 6-8.
- [15] 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 Körut 6-8.

- [16] Steger, A. E. (1985). "An extension of Karmarkar's algorithm for bounded linear programming problems," M. S. Thesis, State University of New York, Stony Brook, NY.
- [17] Todd, M. J. and B. P. Burrell (1985). "An extension of Karmarkar's algorithm for linear programming using dual variables," Technical Report No. 648, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY.
- [18] Tseng, P. (1988). "A simple polynomial-time algorithm for convex quadratic programming," Working Paper LIDS-P-1819, Laboratory for Information and Decision Systems, M.I.T. Cambridge, MA.
- [19] 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.
- [20] 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.
- [21] Ye, Y. and M. Kojima (1986). "Recovering optimal dual solutions in Karmarkar's Polynomial algorithm for linear programming," manuscript, Department of Engineering-Economic Systems, Stanford University, Stanford, CA.