A Polynomial Arc-Search Interior-Point Algorithm for Linear Programming

Yaguang Yang*

Abstract

In this paper, ellipse is used to approximate the central path of the linear programming. An interior-point algorithm is devised to search the optimizers along the ellipse. The algorithm is proved to be polynomial with the complexity bound $O(n^{\frac{1}{2}}\log(1/\epsilon))$. Numerical test is conducted for problems in Netlib. For most tested Netlib problems, the result shows that the new algorithm uses less iteration to converge than the Matlab optimization toolbox `linprog` which implements the state-of-art Mehrotra's predictor-corrector algorithm. For all the tested problems, the number of total iterations using the new algorithm is about 20% fewer than the one using `linprog`.

Keywords: Arc-search, interior-point method, polynomial algorithm, linear programming.

*NRC, Office of Research, 21 Church Street, Rockville, 20850. Email: yaguang.yang@verizon.net

1 Introduction

Polynomiality has been used to measure the effectiveness and efficiency of algorithms for a long time. Though simplex method is effective and efficient in practice, Klee and Minty [8] showed that simplex algorithm with a common pivot rule for linear programming is not a polynomial algorithm that inspired the research of developing polynomial algorithms for linear programming. The first achievement in this direction was due to Khachiyan [7]. Although this first polynomial algorithm for linear programming is not computationally competitive [1], a different method, interior-point method, has been proved very efficient. Many interior-point polynomial algorithms, for example [6, 9, 10, 13, 14, 15], have been developed. Bewilderingly, the computational performance of the polynomial algorithm with the lowest known complexity bound is not as good as other polynomial algorithms with higher complexity bound [18]. The most successful interior-point algorithm in practice is the MPC proposed by Mehrotra [12] and refined by other researchers. However, MPC is not proved to be polynomial although a lot of effort has been made.

Our main purpose in this paper is to close the gap between the algorithms that have good theoretical polynomiality result and the algorithms that demonstrate superior practical performance. We will devise an algorithm that is not only theoretically attractive (having the lowest polynomial complexity bound) but also competitive in computational test on standard problems. We will extend an arc-search algorithm developed in [19] for linear programming. The proposed algorithm searches optimizers along an ellipse that approximates the entire central path. This strategy is different from other higher order methods, such as [12, 15], which search optimizers either on an arc of power series approximation or on a straight line of a linear combination of the first and higher order derivatives of the central path. We will prove that the proposed algorithm has polynomial complexity bound $O(n^{\frac{1}{2}}\log(1/\epsilon))$ (which is the best known complexity bound for linear programming and better than other higher order algorithms such as [12, 15]). We will also provide preliminary test results on standard Netlib problems. These results show that the newly developed code is very promising (using fewer iterations in most tested problems) compared to the MATLAB code `linprog` which implements the state-of-the-art MPC plus some other enhancements [20].

Throughout the paper, we will denote Hadamard (element-wise) product of two vectors x and s by $x \circ s$, the ith component of x by x_i, the element-wise inverse of x by x^{-1} if $\min |s_i| > 0$, the element-wise division of the two vectors by $s^{-1} \circ x$, or $x \circ s^{-1}$, or $\frac{x}{s}$ if $\min |s_i| > 0$, the inner product of two vectors x and s by $\langle x, s \rangle$, the Euclidean norm of x by $\|x\|$, the infinite norm of x by $\|x\|_\infty$, the identity matrix of any dimension by I, the identity matrix of dimension m by I_m, the vector of all ones with appropriate dimension by e, the transpose of matrix A by A^{T}, a basis for the null space of A by $\hat{A}$. To make the notation simple for block column vectors, we will denote, for example, $[x^{\mathrm{T}}, s^{\mathrm{T}}]^{\mathrm{T}}$ by (x, s). For $x \in \mathbf{R}^n$, we will denote a related diagonal matrix by $X \in \mathbf{R}^{n\times n}$ whose diagonal elements are components of the vector x. Finally, we define an initial point of any algorithm by x^0, the point after the kth iteration by x^k.

2 Problem Descriptions

Consider the Linear Programming (LP) in the standard form:

$$(LP) \quad \min \ c^{\mathrm{T}}x, \quad \text{subject to} \quad Ax = b, \quad x \geq 0, \tag{1}$$

where $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$, $c \in \mathbf{R}^n$ are given, and $x \in \mathbf{R}^n$ is the vector to be optimized. Associated with the linear programming is the dual programming (DP) that is also presented in the standard form:

$$(DP) \quad \max \ b^{\mathrm{T}}\lambda, \quad \text{subject to} \quad A^{\mathrm{T}}\lambda + s = c, \quad s \geq 0, \tag{2}$$

where dual variable vector $\lambda \in \mathbf{R}^m$, and dual slack vector $s \in \mathbf{R}^n$. Denote the feasible set $\mathcal{F}$ as a collection of all points that satisfy the constraints of LP and DP,

$$\mathcal{F} = \{(x, \lambda, s) : Ax = b, A^{\mathrm{T}}\lambda + s = c, (x, s) \geq 0\}, \tag{3}$$

and the strictly feasible set $\mathcal{F}^o$ as a collection of all points that satisfy the constraints of LP and DP and are strictly positive

$$\mathcal{F}^o = \{(x, \lambda, s) : Ax = b, A^{\mathrm{T}}\lambda + s = c, (x, s) > 0\}. \tag{4}$$

Throughout the paper, we make the following assumptions.

Assumptions:

1. A is a full rank matrix.

2. $\mathcal{F}^o$ is not empty.

Assumption 1 is the standard Linear Independence Constraint Qualification (LICQ). Assumption 2 implies the existence of a central path. It is well known that $x \in \mathbf{R}^n$ is an optimal solution of (1) if and only if x, λ, and s satisfy the following KKT conditions

$$Ax = b \tag{5a}$$

$$A^{\mathrm{T}}\lambda + s = c \tag{5b}$$

$$(x, s) \geq 0. \tag{5c}$$

$$x_i s_i = 0, \quad i = 1, \ldots, n \tag{5d}$$

The first three conditions imply that x is a feasible solution of the primal problem and (λ, s) is a feasible solution of the dual problem. The last condition implies that the duality gap is zero. The central path $\mathcal{C}$ is parameterized by a scalar $\tau > 0$ as follows. For each interior point $(x, \lambda, s) \in \mathcal{C}$ on the central path, there is a $\tau > 0$ such that

$$Ax = b \tag{6a}$$

$$A^{\mathrm{T}}\lambda + s = c \tag{6b}$$

$$(x, s) > 0 \tag{6c}$$

$$x_i s_i = \tau, \quad i = 1, \ldots, n. \tag{6d}$$

Therefore, the central path is an arc in $\mathbf{R}^{2n+m}$ parameterized as a function of τ and is denoted as

$$\mathcal{C} = \{(x(\tau), \lambda(\tau), s(\tau)) : \tau > 0\}. \tag{7}$$

As $\tau \rightarrow 0$, the central path $(x(\tau), \lambda(\tau), s(\tau))$ represented by (6) approaches to a solution of LP represented by (1). Theoretical analyses and computational experiments demonstrate [16] that searching along the central path is the most efficient way to find optimizers. However, there is no practical way to calculate the entire arc of the central path. All path-following algorithms try (a) to search, from the current (x, s) along certain directions related to the tangent of the central path, to a new point that reduces the value of $x^{\mathrm{T}}s$ (the duality gap) and simultaneously satisfies (6a), (6b), and (6c), thereby moving the current point towards the solution, and (b) to stay close to the central path, thereby being able to make a good progress in the next search. We will consider a central path-following algorithm that searches the optimizers (located at the boundary of $\mathcal{F}$) along an arc that approximates the central path $\mathcal{C} \in \mathcal{F}^o \subset \mathcal{F}$.

3 A Polynomial Arc-Search Algorithm in Linear Programming

3.1 Ellipse approximation of the central path

We will use an ellipse $\mathcal{E}$ [4] in $2n + m$ dimensional space to approximate the central path $\mathcal{C}$ described by (6), where

$$\mathcal{E} = \{(x(\alpha), \lambda(\alpha), s(\alpha)) : (x(\alpha), \lambda(\alpha), s(\alpha)) = \vec{a}\cos(\alpha) + \vec{b}\sin(\alpha) + \vec{c}\}, \tag{8}$$

$\vec{a} \in \mathbf{R}^{2n+m}$ and $\vec{b} \in \mathbf{R}^{2n+m}$ are the axes of the ellipse, and they are perpendicular to each other, $\vec{c} \in \mathbf{R}^{2n+m}$ is the center of the ellipse. Given a point $y = (x, \lambda, s) = (x(\alpha_0), \lambda(\alpha_0), s(\alpha_0)) \in \mathcal{E}$ which is close to or on the central path, we will determine $\vec{a}$, $\vec{b}$, $\vec{c}$ and α_0 such that the first and second derivatives of $(x(\alpha_0), \lambda(\alpha_0), s(\alpha_0))$ have the form as if they were on the central path (though they may not be on the central path). Therefore, we want the first and second derivatives at $(x(\alpha_0), \lambda(\alpha_0), s(\alpha_0)) \in \mathcal{E}$ to satisfy

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^{\mathrm{T}} & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{\lambda} \\ \dot{s} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ x \circ s \end{bmatrix}. \tag{9}$$

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^{\mathrm{T}} & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \ddot{\lambda} \\ \ddot{s} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -2\dot{x} \circ \dot{s} \end{bmatrix}. \tag{10}$$

It is clear that such an ellipse should approximate the central path well when $(x(\alpha_0), \lambda(\alpha_0), s(\alpha_0))$ is close to the central path and $\alpha_0 \pm \epsilon \to \alpha_0$. To simplify the notation, let

$$y(\alpha) = (x(\alpha), \lambda(\alpha), s(\alpha)) = \vec{a}\cos(\alpha) + \vec{b}\sin(\alpha) + \vec{c}. \tag{11}$$

Then

$$\dot{y}(\alpha) = (\dot{x}(\alpha), \dot{\lambda}(\alpha), \dot{s}(\alpha)) = -\vec{a}\sin(\alpha) + \vec{b}\cos(\alpha), \tag{12}$$

$$\ddot{y}(\alpha) = (\ddot{x}(\alpha), \ddot{\lambda}(\alpha), \ddot{s}(\alpha)) = -\vec{a}\cos(\alpha) - \vec{b}\sin(\alpha). \tag{13}$$

It is straightforward to verify from (11), (12), and (13) that

$$\vec{a} = -\dot{y}\sin(\alpha) - \ddot{y}\cos(\alpha), \tag{14}$$

$$\vec{b} = \dot{y}\cos(\alpha) - \ddot{y}\sin(\alpha), \tag{15}$$

$$\vec{c} = y + \ddot{y}. \tag{16}$$

The search along the ellipse will be carried out on the interval $[\alpha_0 - \alpha, \alpha_0]$ and $\alpha \in [0, \frac{\pi}{2}]$. In the next subsection, we will show that the calculation of α_0 can be avoided.

3.2 Search along the approximate central path

Though one can search a better feasible point with reduced duality gap along the ellipse defined by (11) which needs to compute $\vec{a}$, $\vec{b}$, and $\vec{c}$, we will use a simplified formula that reduces the operation counts slightly and is more convenient for convergence analysis. Denote

$$\vec{a} = \begin{bmatrix} a_x \\ a_\lambda \\ a_s \end{bmatrix} = -\dot{y}\sin(\alpha) - \ddot{y}\cos(\alpha) = \begin{bmatrix} -\dot{x}\sin(\alpha) - \ddot{x}\cos(\alpha) \\ -\dot{\lambda}\sin(\alpha) - \ddot{\lambda}\cos(\alpha) \\ -\dot{s}\sin(\alpha) - \ddot{s}\cos(\alpha) \end{bmatrix},$$

$$\vec{b} = \begin{bmatrix} b_x \\ b_\lambda \\ b_s \end{bmatrix} = \dot{y}\cos(\alpha) - \ddot{y}\sin(\alpha) = \begin{bmatrix} \dot{x}\cos(\alpha) - \ddot{x}\sin(\alpha) \\ \dot{\lambda}\cos(\alpha) - \ddot{\lambda}\sin(\alpha) \\ \dot{s}\cos(\alpha) - \ddot{s}\sin(\alpha) \end{bmatrix},$$

and

$$\vec{c} = \begin{bmatrix} c_x \\ c_\lambda \\ c_s \end{bmatrix} = y + \ddot{y} = \begin{bmatrix} x + \ddot{x} \\ \lambda + \ddot{\lambda} \\ s + \ddot{s} \end{bmatrix}.$$

Let $x(\alpha)$ and $s(\alpha)$ be the updated x and s after the search, we have

$$\begin{aligned} & x(\alpha) = a_x\cos(\alpha_0 - \alpha) + b_x\sin(\alpha_0 - \alpha) + c_x \\ = \; & a_x(\cos(\alpha_0)\cos(\alpha) + \sin(\alpha_0)\sin(\alpha)) + b_x(\sin(\alpha_0)\cos(\alpha) - \cos(\alpha_0)\sin(\alpha)) \\ & + c_x - c_x\cos(\alpha) + c_x\cos(\alpha) \\ = \; & x\cos(\alpha) + a_x\sin(\alpha_0)\sin(\alpha) - b_x\cos(\alpha_0)\sin(\alpha) + c_x(1 - \cos(\alpha)) \\ = \; & x\cos(\alpha) - (\dot{x}\sin(\alpha_0) + \ddot{x}\cos(\alpha_0))\sin(\alpha_0)\sin(\alpha) \\ & -(\dot{x}\cos(\alpha_0) - \ddot{x}\sin(\alpha_0))\cos(\alpha_0)\sin(\alpha) + (x + \ddot{x})(1 - \cos(\alpha)) \\ = \; & x - \dot{x}(\sin^2(\alpha_0)\sin(\alpha) + \cos^2(\alpha_0)\sin(\alpha)) \\ & + \ddot{x}(-\sin(\alpha_0)\cos(\alpha_0)\sin(\alpha) + \sin(\alpha_0)\cos(\alpha_0)\sin(\alpha) + (1 - \cos(\alpha))) \\ = \; & x - \dot{x}\sin(\alpha) + \ddot{x}(1 - \cos(\alpha)). \end{aligned} \tag{17}$$

Similarly

$$
\begin{aligned}
& s(\alpha) = a_s\cos(\alpha_0-\alpha)+b_s\sin(\alpha_0-\alpha)+c_s \\
= \; & a_s(\cos(\alpha_0)\cos(\alpha)+\sin(\alpha_0)\sin(\alpha))+b_s(\sin(\alpha_0)\cos(\alpha)-\cos(\alpha_0)\sin(\alpha)) \\
& +c_s-c_s\cos(\alpha)+c_s\cos(\alpha) \\
= \; & s\cos(\alpha)+a_s\sin(\alpha_0)\sin(\alpha)-b_s\cos(\alpha_0)\sin(\alpha)+c_s(1-\cos(\alpha)) \\
= \; & s\cos(\alpha)-(\dot{s}\sin(\alpha_0)+\ddot{s}\cos(\alpha_0))\sin(\alpha_0)\sin(\alpha) \\
& -(\dot{s}\cos(\alpha_0)-\ddot{s}\sin(\alpha_0))\cos(\alpha_0)\sin(\alpha)+(s+\ddot{s})(1-\cos(\alpha)) \\
= \; & s-\dot{s}(\sin^2(\alpha_0)\sin(\alpha)+\cos^2(\alpha_0)\sin(\alpha)) \\
& +\ddot{s}(-\sin(\alpha_0)\cos(\alpha_0)\sin(\alpha)+\sin(\alpha_0)\cos(\alpha_0)\sin(\alpha)+(1-\cos(\alpha))) \\
= \; & s-\dot{s}\sin(\alpha)+\ddot{s}(1-\cos(\alpha)),
\end{aligned}
\tag{18}
$$

and

$$
\lambda(\alpha) \;=\; \lambda-\dot{\lambda}\sin(\alpha)+\ddot{\lambda}(1-\cos(\alpha)). \tag{19}
$$

As pointed above, (17), (18), and (19) do not depend on α_0 explicitly. We summarize the above discussion as the following

Theorem 3.1 *Let $(x(\alpha),\lambda(\alpha),s(\alpha))$ be an arc defined by (8) passing through a point $(x,\lambda,s)\in\mathcal{E}$, and its first and second derivatives at (x,λ,s) be $(\dot{x},\dot{\lambda},\dot{s})$ and $(\ddot{x},\ddot{\lambda},\ddot{s})$ which are defined by (9) and (10). Then an ellipse approximation of the central path is given by*

$$
x(\alpha)=x-\dot{x}\sin(\alpha)+\ddot{x}(1-\cos(\alpha)). \tag{20}
$$

$$
\lambda(\alpha)=\lambda-\dot{\lambda}\sin(\alpha)+\ddot{\lambda}(1-\cos(\alpha)). \tag{21}
$$

$$
s(\alpha)=s-\dot{s}\sin(\alpha)+\ddot{s}(1-\cos(\alpha)). \tag{22}
$$

∎

Assuming $(x,s)>0$, one can easily see that if $\dot{x}$, $\ddot{x}$, $\dot{s}$, and $\ddot{s}$ are bounded above from a constant, and if α is small enough, then $x(\alpha)>0$ and $s(\alpha)>0$.

Lemma 3.1 *Let $\dot{x}$, $\dot{s}$, $\ddot{x}$, and $\ddot{s}$ be the solution of (9) and (10). Then*

$$
\dot{x}^{\mathrm{T}}\dot{s}=0,
$$

$$
\ddot{x}^{\mathrm{T}}\dot{s}=0,
$$

$$
\dot{x}^{\mathrm{T}}\ddot{s}=0,
$$

$$
\ddot{x}^{\mathrm{T}}\ddot{s}=0.
$$

Proof: Pre-multiplying $\dot{x}^{\mathrm{T}}$ or $\ddot{x}^{\mathrm{T}}$ to the second rows of (9) and (10), and using the first rows of (9) and (10) gives the results. ∎

Let

$$
\mu=\frac{x^{\mathrm{T}}s}{n}. \tag{23}
$$

We will show that searching along the ellipse will reduce the duality gap, i.e., $\mu(\alpha)=\frac{x(\alpha)^{\mathrm{T}}s(\alpha)}{n}<\mu$. If $(x(\alpha),s(\alpha))>0$ holds in all iterations, reducing duality gap to zero means approaching to the solution of the linear programming. Notice that

$$
\mu(\alpha)=\mu(1-\sin(\alpha)) \tag{24}
$$

holds for any choice of $\alpha\in[0,\frac{\pi}{2}]$ due to the previous lemma, this means that the larger the α is, the more improvement the $\mu(\alpha)$ will be.

3.3 The polynomial arc-search algorithm

Let $\theta > 0$, and

$$\mathcal{N}_2(\theta) = \{(x, \lambda, s) : Ax = b, A^{\mathrm{T}}\lambda + s = c, (x, s) > 0, \|x \circ s - \mu e\| \leq \theta\mu\}. \tag{25}$$

Similar to [13], we present a predictor-corrector type polynomial algorithm which uses $\mathcal{N}_2(\theta)$ and $\mathcal{N}_2(2\theta)$. The algorithm starts the iterate inside $\mathcal{N}_2(\theta)$ and restricts the arc-search in $\mathcal{N}_2(2\theta)$. After the search finds an iterate with smaller duality gap, a corrector step brings the iterate from $\mathcal{N}_2(2\theta)$ back to $\mathcal{N}_2(\theta)$ without changing the duality gap.

We will repeatedly use some simple lemmas.

Lemma 3.2 *Let (x, λ, s) is a strictly feasible point of (LP) and (DP), $(\dot{x}, \dot{\lambda}, \dot{s})$ and $(\ddot{x}, \ddot{\lambda}, \ddot{s})$ satisfy (9) and (10), $(x(\alpha), \lambda(\alpha), s(\alpha))$ be calculated using (17), (18), and (19), then the following conditions hold.*

$$Ax(\alpha) = b, \quad A^{\mathrm{T}}\lambda(\alpha) + s(\alpha) = c.$$

Proof: Direct calculation verifies the result. ∎

Lemma 3.3 *Let $p > 0$, $q > 0$, and $r > 0$. If $p + q \leq r$, then $pq \leq \frac{r^2}{4}$.* ∎

Lemma 3.4 *For $\alpha \in [0, \frac{\pi}{2}]$,*

$$\sin(\alpha) \geq \sin^2(\alpha) = 1 - \cos^2(\alpha) \geq 1 - \cos(\alpha).$$

∎

The following Lemma gives some useful estimations and notations to be used later.

Lemma 3.5 *Let $(\dot{x}, \dot{\lambda}, \dot{s})$ be calculated by (9) and $(\ddot{x}, \ddot{\lambda}, \ddot{s})$ be calculated by (10). Assuming that $(\bar{x}^k, \bar{\lambda}^k, \bar{s}^k) \in \mathcal{N}_2(\theta)$, $\bar{\mu}^k = \frac{\bar{x}^{k\mathrm{T}}\bar{s}^k}{n}$, then*

$$\left\|\frac{\dot{x} \circ \dot{s}}{\bar{\mu}^k}\right\| \leq \frac{(1+\theta)^2}{(1-\theta)}n, \quad \left\|\frac{\ddot{x} \circ \dot{s}}{\bar{\mu}^k}\right\| \leq \frac{2(1+\theta)^3}{(1-\theta)^2}n^{\frac{3}{2}}, \quad \left\|\frac{\dot{x} \circ \ddot{s}}{\bar{\mu}^k}\right\| \leq \frac{2(1+\theta)^3}{(1-\theta)^2}n^{\frac{3}{2}}, \quad \left\|\frac{\ddot{x} \circ \ddot{s}}{\bar{\mu}^k}\right\| \leq \frac{4(1+\theta)^4}{(1-\theta)^3}n^2. \tag{26}$$

Proof: For the sake of simplicity, we omit the superscript of k and over-bars over x, s, and μ in the proof. It is easy to see from the first row of (9) that there exist a vector v such that

$$\frac{\dot{x}}{x} = X^{-1}\hat{A}v. \tag{27}$$

From the third row of (9), we have

$$\frac{\dot{s}}{s} + \frac{\dot{x}}{x} = e. \tag{28}$$

From the second row of (9), we have

$$S^{-1}A^{\mathrm{T}}\dot{\lambda} + \frac{\dot{s}}{s} = 0. \tag{29}$$

Combining the above three equations, we get

$$\left[X^{-1}\hat{A}, -S^{-1}A^{\mathrm{T}}\right]\begin{bmatrix} v \\ \dot{\lambda} \end{bmatrix} = e. \tag{30}$$

Since A is full rank, we have

$$\begin{bmatrix} (\hat{A}^{\mathrm{T}}SX^{-1}\hat{A})^{-1}\hat{A}^{\mathrm{T}}S \\ -(AXS^{-1}A^{\mathrm{T}})^{-1}AX \end{bmatrix}\left[X^{-1}\hat{A}, -S^{-1}A^{\mathrm{T}}\right] = I.$$

Taking the inverse in (30) gives

$$\begin{bmatrix} v \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} (\hat{A}^{\mathrm{T}} S X^{-1} \hat{A})^{-1} \hat{A}^{\mathrm{T}} S \\ -(A X S^{-1} A^{\mathrm{T}})^{-1} A X \end{bmatrix} e.$$

Substituting this relation to (27) and (29) gives

$$\begin{aligned} \frac{\dot{x}}{x} &= X^{-1} \hat{A} (\hat{A}^{\mathrm{T}} S X^{-1} \hat{A})^{-1} \hat{A}^{\mathrm{T}} S e \\ &= \left[I - S^{-1} A^{\mathrm{T}} (A X S^{-1} A^{\mathrm{T}})^{-1} A X \right] e, \end{aligned} \tag{31}$$

and

$$\begin{aligned} \frac{\dot{s}}{s} &= S^{-1} A^{\mathrm{T}} (A X S^{-1} A^{\mathrm{T}})^{-1} A X e \\ &= \left[I - X^{-1} \hat{A} (\hat{A}^{\mathrm{T}} S X^{-1} \hat{A})^{-1} \hat{A}^{\mathrm{T}} S \right] e. \end{aligned} \tag{32}$$

Since $(x, \lambda, s) \in \mathcal{N}_2(\theta)$, we have

$$(1-\theta)\mu I \le XS \le (1+\theta)\mu I.$$

Repeatedly using this estimation and (31) yields

$$\begin{aligned} \left\| \frac{\dot{x}}{x} \right\|^2 &= e^{\mathrm{T}} S^{\mathrm{T}} \hat{A} (\hat{A}^{\mathrm{T}} S X^{-1} \hat{A})^{-1} \hat{A}^{\mathrm{T}} X^{-1} X^{-1} \hat{A} (\hat{A}^{\mathrm{T}} S X^{-1} \hat{A})^{-1} \hat{A}^{\mathrm{T}} S e \\ &\le ((1-\theta)\mu)^{-1} e^{\mathrm{T}} S^{\mathrm{T}} \hat{A} (\hat{A}^{\mathrm{T}} S X^{-1} \hat{A})^{-1} \hat{A}^{\mathrm{T}} S X^{-1} \hat{A} (\hat{A}^{\mathrm{T}} S X^{-1} \hat{A})^{-1} \hat{A}^{\mathrm{T}} S e \\ &= ((1-\theta)\mu)^{-1} e^{\mathrm{T}} S \hat{A} (\hat{A}^{\mathrm{T}} S X^{-1} \hat{A})^{-1} \hat{A}^{\mathrm{T}} S e \\ &\le \frac{(1+\theta)}{(1-\theta)} e^{\mathrm{T}} S \hat{A} (\hat{A}^{\mathrm{T}} S^2 \hat{A})^{-1} \hat{A}^{\mathrm{T}} S e. \end{aligned}$$

Using QR decomposition

$$S\hat{A} = Q \begin{bmatrix} R_1 \\ 0 \end{bmatrix} = [Q_1, Q_2] \begin{bmatrix} R_1 \\ 0 \end{bmatrix} = Q_1 R_1,$$

where Q_1 and Q_2 are orthonormal matrices and they are orthogonal to each other, we have

$$\left\| \frac{\dot{x}}{x} \right\|^2 \le \frac{(1+\theta)}{(1-\theta)} e^{\mathrm{T}} Q_1 Q_1^{\mathrm{T}} e \le \frac{(1+\theta)}{(1-\theta)} \|e\|^2 = \frac{(1+\theta)}{(1-\theta)} n. \tag{33}$$

Similarly,

$$\left\| \frac{\dot{s}}{s} \right\|^2 \le \frac{(1+\theta)}{(1-\theta)} e^{\mathrm{T}} X A^{\mathrm{T}} (A X^2 A^{\mathrm{T}})^{-1} A X e \le \frac{(1+\theta)}{(1-\theta)} n. \tag{34}$$

Applying Lemma 3.3 to the above two equations yields

$$\left\| \frac{\dot{x}}{x} \circ \frac{\dot{s}}{s} \right\| \le \left\| \frac{\dot{x}}{x} \right\| \left\| \frac{\dot{s}}{s} \right\| \le \frac{(1+\theta)}{(1-\theta)} n. \tag{35}$$

Since $\|x \circ s - \mu\| \le \theta\mu$, for any i, $(1-\theta)\mu \le x_i s_i \le (1+\theta)\mu$, or equivalently,

$$\frac{x_i s_i}{1+\theta} \le \frac{\max_i x_i s_i}{1+\theta} \le \mu \le \frac{\min_i x_i s_i}{1-\theta} \le \frac{x_i s_i}{1-\theta}. \tag{36}$$

Using (35), we have

$$\frac{\left\| \dot{x} \circ \dot{s} \right\|}{\mu} \le (1+\theta) \frac{\left\| \dot{x} \circ \dot{s} \right\|}{\max_i x_i s_i} \le (1+\theta) \left\| \frac{\dot{x}}{x} \circ \frac{\dot{s}}{s} \right\| \le \frac{(1+\theta)^2}{1-\theta} n.$$

Let $\phi = -2\left(\frac{\dot{x}}{x} \circ \frac{\dot{s}}{s}\right)$. Following the similar proof, it is easy to get

$$\left\|\frac{\ddot{x}}{x}\right\|^2 \leq \frac{1+\theta}{1-\theta}\|\phi\|^2 \leq 4\left(\frac{1+\theta}{1-\theta}\right)^3 n^2, \qquad \left\|\frac{\ddot{s}}{s}\right\|^2 \leq \frac{1+\theta}{1-\theta}\|\phi\|^2 \leq 4\left(\frac{1+\theta}{1-\theta}\right)^3 n^2. \tag{37}$$

From (36), (33), (34), and (37), we get

$$\frac{\left\|\dot{x} \circ \ddot{s}\right\|}{\mu} \leq (1+\theta)\left\|\frac{\dot{x}}{x} \circ \frac{\ddot{s}}{s}\right\| \leq (1+\theta)\left\|\frac{\dot{x}}{x}\right\|\left\|\frac{\ddot{s}}{s}\right\| \leq 2\frac{(1+\theta)^3}{(1-\theta)^2} n^{\frac{3}{2}}.$$

$$\frac{\left\|\ddot{x} \circ \dot{s}\right\|}{\mu} \leq (1+\theta)\left\|\frac{\ddot{x}}{x} \circ \frac{\dot{s}}{s}\right\| \leq (1+\theta)\left\|\frac{\ddot{x}}{x}\right\|\left\|\frac{\dot{s}}{s}\right\| \leq 2\frac{(1+\theta)^3}{(1-\theta)^2} n^{\frac{3}{2}}.$$

$$\frac{\left\|\ddot{x} \circ \ddot{s}\right\|}{\mu} \leq (1+\theta)\left\|\frac{\ddot{x}}{x} \circ \frac{\ddot{s}}{s}\right\| \leq (1+\theta)\left\|\frac{\ddot{x}}{x}\right\|\left\|\frac{\ddot{s}}{s}\right\| \leq 4\frac{(1+\theta)^4}{(1-\theta)^3} n^2.$$

This finishes the proof. ∎

For the sake of the simplicity in the analysis, we assume that Algorithm 1 of [2] is used prior to call the polynomial algorithm and an initial point inside $\mathcal{N}_2(\theta)$ is obtained.

Algorithm 3.1 *Data:* A, b, c, $\theta = 0.292$, $\epsilon > 0$, *initial point* $(\bar{x}^0, \bar{\lambda}^0, \bar{s}^0) \in \mathcal{N}_2(\theta)$, *and* $\bar{\mu}^0 = \frac{\bar{x}^{0\mathrm{T}}\bar{s}^0}{n}$.
for *iteration* $k = 1, 2, \ldots$

Step 1: If $\bar{\mu}^k < \epsilon$ *and*

$$\|r_B\| = \|A\bar{x}^k - b\| \leq \epsilon, \tag{38a}$$

$$\|r_C\| = \|A^{\mathrm{T}}\bar{\lambda}^k + \bar{s}^k - c\| \leq \epsilon, \tag{38b}$$

$$\|r_t\| = \|\bar{x}^k\bar{s}^k e - \bar{\mu}^k e\| \leq \epsilon, \tag{38c}$$

$$(\bar{x}^k, \bar{s}^k) > 0. \tag{38d}$$

holds, stop. Otherwise continue.

Step 2: Solve the linear systems of equations (9) and (10) to get $(\dot{x}, \dot{\lambda}, \dot{s})$ *and* $(\ddot{x}, \ddot{\lambda}, \ddot{s})$.

Step 3: Find the smallest positive $\sin(\alpha)$ *that satisfies quartic polynomial in terms of* $\sin(\alpha)$

$$q(\alpha) = \left(\left\|\ddot{x} \circ \ddot{s}\right\| + \left\|\dot{x} \circ \dot{s}\right\|\right)\sin^4(\alpha) + \left(\left\|\dot{x} \circ \ddot{s}\right\| + \left\|\ddot{x} \circ \dot{s}\right\|\right)\sin^3(\alpha) + \theta\bar{\mu}^k\sin(\alpha) - \theta\bar{\mu}^k = 0. \tag{39}$$

Set

$$(x(\alpha), \lambda(\alpha), s(\alpha)) = (\bar{x}^k, \bar{\lambda}^k, \bar{s}^k) - (\dot{x}, \dot{\lambda}, \dot{s})\sin(\alpha) + (\ddot{x}, \ddot{\lambda}, \ddot{s})(1 - \cos(\alpha)), \tag{40}$$

and

$$\mu(\alpha) = \bar{\mu}^k(1 - \sin(\alpha)). \tag{41}$$

Step 4: Calculate $(\Delta x, \Delta\lambda, \Delta s)$ *by solving*

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^{\mathrm{T}} & I \\ S(\alpha) & 0 & X(\alpha) \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \mu(\alpha)e - x(\alpha) \circ s(\alpha) \end{bmatrix}. \tag{42}$$

Set

$$(\bar{x}^{k+1}, \bar{\lambda}^{k+1}, \bar{s}^{k+1}) = (x(\alpha), \lambda(\alpha), s(\alpha)) + (\Delta x, \Delta\lambda, \Delta s) \tag{43}$$

and $\bar{\mu}^{k+1} = \frac{\bar{x}^{k+1\mathrm{T}}\bar{s}^{k+1}}{n}$. *Set* $k+1 \to k$. *Go back to Step 1.*

end (for) ∎

It is easy to see that the quartic polynomial (39) is a monotonic increasing function of $\alpha \in [0, \frac{\pi}{2}]$, for $(\bar{x}^k, \bar{s}^k) > 0$ (which will be shown in Lemma 3.8), $q(0) < 0$, and $q(\frac{\pi}{2}) \geq 0$. Therefore, $q(\alpha)$ has only one real solution in $[0, \frac{\pi}{2}]$, and the computational count for the quartic polynomial is negligible [17]. Furthermore, $\alpha = \sin^{-1}(\alpha)$ is well-defined. From (9) and (40), it is easy to check that (41) satisfies the definition of (23). In the rest of this section, we will show that if $(\bar{x}^k, \bar{\lambda}^k, \bar{s}^k) \in \mathcal{N}_2(0.292)$, then $(x(\alpha), \lambda(\alpha), s(\alpha)) \in \mathcal{N}_2(0.584)$ and $(\bar{x}^{k+1}, \bar{\lambda}^{k+1}, \bar{s}^{k+1}) \in \mathcal{N}_2(0.292)$. We will also estimate the size of α that will be used to prove the polynomiality in the next section.

Lemma 3.6 *Let $(\bar{x}^k, \bar{\lambda}^k, \bar{s}^k) \in \mathcal{N}_2(\theta)$, $(\dot{x}, \dot{\lambda}, \dot{s})$ and $(\ddot{x}, \ddot{\lambda}, \ddot{s})$ be calculated from (9) and (10), let $\sin(\alpha)$ be the positive real solution of (39) and $(x(\alpha), \lambda(\alpha), s(\alpha))$ be updated using (40). Then $(x(\alpha), \lambda(\alpha), s(\alpha)) \in \mathcal{N}_2(2\theta)$.*

Proof: First, it is worthwhile to note that Lemmas 3.1 and 3.2 hold in this case. From (39), we have

$$\left(\left\|\ddot{x} \circ \ddot{s}\right\| + \left\|\dot{x} \circ \dot{s}\right\|\right) \sin^4(\alpha) + \left(\left\|\dot{x} \circ \ddot{s}\right\| + \left\|\ddot{x} \circ \dot{s}\right\|\right) \sin^3(\alpha) = \theta \bar{\mu}^k (1 - \sin(\alpha)) = \theta \mu(\alpha).$$

Using (9), (10), (40), (41), and Lemma 3.4, we have

$$\begin{aligned}
& \|x(\alpha) \circ s(\alpha) - \mu(\alpha) e\| \\
= \; & \|(\bar{x}^k \circ \bar{s}^k - \bar{\mu}^k e)(1 - \sin(\alpha)) + (\bar{x}^k \circ \ddot{s} + \bar{s}^k \circ \ddot{x})(1 - \cos(\alpha)) + \dot{x} \circ \dot{s} \sin^2(\alpha) \\
& -(\dot{x} \circ \ddot{s} + \dot{s} \circ \ddot{x}) \sin(\alpha)(1 - \cos(\alpha)) + \ddot{x} \circ \ddot{s}(1 - \cos(\alpha))^2\| \qquad (44) \\
\leq \; & (1 - \sin(\alpha))\|\bar{x}^k \circ \bar{s}^k - \bar{\mu}^k e\| + \|(\ddot{x} \circ \ddot{s} - \dot{x} \circ \dot{s})(1 - \cos(\alpha))^2 - (\dot{x} \circ \ddot{s} + \dot{s} \circ \ddot{x}) \sin(\alpha)(1 - \cos(\alpha))\| \\
\leq \; & (1 - \sin(\alpha))\theta \bar{\mu}^k + \left(\left\|\ddot{x} \circ \ddot{s}\right\| + \left\|\dot{x} \circ \dot{s}\right\|\right) \sin^4(\alpha) + \left(\left\|\dot{x} \circ \ddot{s}\right\| + \left\|\ddot{x} \circ \dot{s}\right\|\right) \sin^3(\alpha) \\
= \; & \theta \mu(\alpha) + \theta \mu(\alpha) = 2\theta \mu(\alpha). \qquad (45)
\end{aligned}$$

Since

$$x_i(\alpha) s_i(\alpha) \geq (1 - 2\theta)\mu(\alpha) = \bar{\mu}^k (1 - 2\theta)(1 - \sin(\alpha)) > 0,$$

for all $\alpha \in [0, \frac{\pi}{2})$, $x(\alpha) > 0$ and $s(\alpha) > 0$ for all $\alpha \in [0, \frac{\pi}{2})$. This concludes the proof. ∎

The next technical result is from [18, page 88].

Lemma 3.7 *Let u and v be vectors of the same dimension, and $u^{\mathrm{T}} v \geq 0$. Then*

$$\|u \circ v\| \leq 2^{-\frac{3}{2}} \|u + v\|^2.$$

The following lemma is a minor modification of a result in [18], it is included for its usefulness in implementation.

Lemma 3.8 *Let $(x(\alpha), \lambda(\alpha), s(\alpha)) \in \mathcal{N}_2(2\theta)$ and $0 < \theta \leq 0.292$. Then $(\bar{x}^{k+1}, \bar{\lambda}^{k+1}, \bar{s}^{k+1}) \in \mathcal{N}_2(\theta)$ and $\bar{\mu}^{k+1} = \mu(\alpha)$.*

Proof: Denote $(\bar{x}^{k+1}(t), \bar{s}^{k+1}(t)) = (x(\alpha) + t\Delta x, s(\alpha) + t\Delta s))$, and $(\bar{x}^{k+1}, \bar{s}^{k+1}) = (\bar{x}^{k+1}(1), \bar{s}^{k+1}(1))$. Using the third row of (42), we have $\frac{x(\alpha)^{\mathrm{T}} \Delta s + s(\alpha)^{\mathrm{T}} \Delta x}{n} = 0$. Using the fact that $\Delta x^{\mathrm{T}} \Delta s = 0$, we have

$$\bar{\mu}^{k+1}(t) = \frac{(x(\alpha) + t\Delta x)^{\mathrm{T}} (s(\alpha) + t\Delta s)}{n} = \frac{x(\alpha)^{\mathrm{T}} s(\alpha)}{n} = \mu(\alpha).$$

Let $D = X(\alpha)^{\frac{1}{2}} S(\alpha)^{-\frac{1}{2}}$. Pre-multiplying $(X(\alpha) S(\alpha))^{-\frac{1}{2}}$ in the last row of (42) yields

$$D\Delta s + D^{-1}\Delta x = (X(\alpha) S(\alpha))^{-\frac{1}{2}} (\mu(\alpha) e - X(\alpha) S(\alpha) e).$$

Let $u = D\Delta s$, $v = D^{-1}\Delta x$, use the technical Lemma and the assumption that $(x(\alpha), \lambda(\alpha), s(\alpha)) \in \mathcal{N}_2(2\theta)$, we have

$$\begin{aligned}
\|\Delta x \circ \Delta s\| &= \|u \circ v\| \le 2^{-\frac{3}{2}} \|(X(\alpha)S(\alpha))^{-\frac{1}{2}}(\mu(\alpha)e - X(\alpha)S(\alpha)e)\|^2 \\
&= 2^{-\frac{3}{2}} \sum_{i=1}^{n} \frac{(\mu(\alpha) - x_i(\alpha)s_i(\alpha))^2}{x_i(\alpha)s_i(\alpha)} \\
&\le 2^{-\frac{3}{2}} \frac{\|\mu(\alpha)e - x(\alpha) \circ s(\alpha)\|^2}{\min_i x_i(\alpha)s_i(\alpha)} \\
&\le 2^{-\frac{3}{2}} \frac{(2\theta)^2 \mu(\alpha)^2}{(1-2\theta)\mu(\alpha)} = 2^{\frac{1}{2}} \frac{\theta^2 \mu(\alpha)}{(1-2\theta)} = 2^{\frac{1}{2}} \frac{\theta^2}{(1-2\theta)} \bar{\mu}^{k+1}.
\end{aligned} \tag{46}$$

Using this result, (43), the last row of (42), we have

$$\begin{aligned}
& \|\bar{x}^{k+1}(t) \circ \bar{s}^{k+1}(t) - \bar{\mu}^{k+1}(t)e\| \\
= & \left\| \left(x(\alpha) + t\Delta x\right) \circ \left(s(\alpha) + t\Delta s\right) - \mu(\alpha)e \right\| \\
= & \left\| (1-t)\left(x(\alpha) \circ s(\alpha) - \mu(\alpha)e\right) + t^2 \Delta x \circ \Delta s \right\| \\
\le & (1-t)2\theta\mu(\alpha) + t^2 \|\Delta x \circ \Delta s\| \\
\le & \left((1-t)2\theta + 2^{\frac{1}{2}} t^2 \frac{\theta^2}{(1-2\theta)} \right) \bar{\mu}^{k+1} := h(t,\theta)\bar{\mu}^{k+1}.
\end{aligned} \tag{47}$$

Taking $t = 1$ gives $\|\bar{x}^{k+1} \circ \bar{s}^{k+1} - \bar{\mu}^{k+1}e\| \le 2^{\frac{1}{2}} \frac{\theta^2}{(1-2\theta)}$. It is easy to verify that for $0 < \theta \le 0.292$,

$$2^{\frac{1}{2}} \frac{\theta^2}{(1-2\theta)} \le \theta.$$

Since, for $0 < \theta \le 0.292$ and $t \in [0, 1)$, $0 < h(t, \theta) \le h(t, 0.292) < 1$, we have, for $\alpha \in [0, \frac{\pi}{2})$,

$$\bar{x}_i^{k+1}(t)\bar{s}_i^{k+1}(t) \ge (1 - h(t,\theta))\bar{\mu}^{k+1} = (1 - h(t,\theta))\mu(\alpha) = (1 - h(t,\theta))(1 - \sin(\alpha))\bar{\mu}^k > 0.$$

Therefore $(\bar{x}^{k+1}, \bar{s}^{k+1}) > 0$. This finishes the proof. ∎

$\theta = 0.25$ is used in [13]. Taking a larger θ will allow a longer step in arc-search, which may reduce the number of iterations to converge to the optimal solution.

Lemma 3.9 *Let $0 < \theta \le 0.292$, and $\sin(\alpha)$ be the positive real solution of (39). Then,*

$$\sin(\alpha) \ge \theta^2 n^{-\frac{1}{2}}.$$

Proof: Since $q(\sin(\alpha))$ is a monotonic increasing function of $\sin(\alpha) \in [0, 1]$, $q(\sin(0)) < 0$, and $q(\sin(\frac{\pi}{2})) \ge 0$, we need only to show that $q(\theta^2 n^{-\frac{1}{2}}) < 0$. Using Lemma 3.5, for $\sin(\alpha) \le \theta^2 n^{-\frac{1}{2}}$, we have

$$\begin{aligned}
q(\alpha) &= \left(\left\|\ddot{x} \circ \ddot{s}\right\| + \left\|\dot{x} \circ \dot{s}\right\| \right) \sin^4(\alpha) + \left(\left\|\dot{x} \circ \ddot{s}\right\| + \left\|\ddot{x} \circ \dot{s}\right\| \right) \sin^3(\alpha) + \theta\bar{\mu}^k \sin(\alpha) - \theta\bar{\mu}^k \\
&\le \left(\frac{4(1+\theta)^4}{(1-\theta)^3} n^2 + \frac{(1+\theta)^2}{(1-\theta)} n \right) \left(\theta^8 n^{-2}\right) \bar{\mu}^k + \frac{4(1+\theta)^3}{(1-\theta)^2} n^{\frac{3}{2}} \left(\theta^6 n^{-\frac{3}{2}}\right) \bar{\mu}^k + \theta\bar{\mu}^k \left(\theta^2 n^{-\frac{1}{2}}\right) - \theta\bar{\mu}^k \\
&= \theta\bar{\mu}^k \left(\frac{4\theta^7(1+\theta)^4}{(1-\theta)^3} + \frac{\theta^7(1+\theta)^2}{(1-\theta)n} + \frac{4\theta^5(1+\theta)^3}{(1-\theta)^2} + \frac{\theta^2}{\sqrt{n}} - 1 \right) \\
&:= \theta\bar{\mu}^k f(\theta, n).
\end{aligned} \tag{48}$$

Clearly, $f(\theta, n)$ is a monotonic increasing function of $\theta \in (0,1)$, and a monotonic decreasing function of $n \in \{1, 2, \ldots\}$. For $\theta \leq 0.292$, and $n \geq 1$, $f(\theta, n) < 0$, hence, for $\sin(\alpha) \leq \theta^2 n^{-\frac{1}{2}}$, $q(\sin(\alpha)) < 0$. This proves $\sin(\alpha) \geq \theta^2 n^{-\frac{1}{2}}$.

∎

Remark 3.1 *It is clear that if* $\left\|\frac{\dot{x}}{x}\right\|$, $\left\|\frac{\ddot{x}}{x}\right\|$, $\left\|\frac{\dot{s}}{s}\right\|$, $\left\|\frac{\ddot{s}}{s}\right\|$ *are smaller than some constant independent of m and n, then arc-search algorithm proposed above would reduce duality gap in a constant rate independent of m and n, a nice feature (polynomial bound would be independent of n) that appears to hold but cannot be proved. In randomly generated tens of thousands problems, we observed that these values are smaller than 2. But we cannot prove it in general case.*

3.4 The implementation considerations

We will show that Algorithm 3.1 is a polynomial algorithm. But we need a little more careful in the implementation to make the algorithm really competitive in practice.

In Step 2 of Algorithm 3.1, (9) and (10) are used to directly obtain $(\dot{x}, \dot{\lambda}, \dot{s})$ and $(\ddot{x}, \ddot{\lambda}, \ddot{s})$. A more efficient way is to use (28), (29), (31), and (32) as follows.

Case 1: $m > n - m$

$$\begin{aligned} \dot{x} &= \hat{A}(\hat{A}^{\mathrm{T}} S X^{-1} \hat{A})^{-1} \hat{A}^{\mathrm{T}} S e, \\ \dot{s} &= S e - S X^{-1} \dot{x}, \\ \dot{\lambda} &= -(A A^{\mathrm{T}})^{-1} A \dot{s}. \end{aligned}$$

The most expensive operation in this calculation is to calculate $(\hat{A}^{\mathrm{T}} S X^{-1} \hat{A})^{-1}$ that needs $O((n-m)^2 n + (n-m)^3)$. Since $(AA^{\mathrm{T}})^{-1} A$ is independent of iterations and needs to be computed only once, and all other computations have lower order operation counts, $O((n-m)^2 n + (n-m)^3)$ is needed to get $(\dot{x}, \dot{\lambda}, \dot{s})$. Similarly replacing e by $\phi = -2\frac{\dot{x}}{x} \circ \frac{\dot{s}}{s}$, we can compute $(\ddot{x}, \ddot{\lambda}, \ddot{s})$ in exactly the same way as we did for $(\dot{x}, \dot{\lambda}, \dot{s})$. Since $(\hat{A}^{\mathrm{T}} S X^{-1} \hat{A})^{-1}$ has been calculated already, the operation count is negligible.

Case 2: $m \leq n - m$

$$\begin{aligned} \dot{s} &= A^{\mathrm{T}} (A X S^{-1} A^{\mathrm{T}})^{-1} A X e, \\ \dot{x} &= X e - X S^{-1} \dot{s}, \\ \dot{\lambda} &= -(A A^{\mathrm{T}})^{-1} A \dot{s}. \end{aligned}$$

The most expensive operation in this calculation is to calculate $(AXS^{-1}A^{\mathrm{T}})^{-1}$ that needs $O(m^2 n + m^3)$. Since $(AA^{\mathrm{T}})^{-1} A$ is independent of the iteration, and all other computations have lower order operation counts, $O(m^2 n + m^3)$ is needed to get $(\dot{x}, \dot{\lambda}, \dot{s})$. Similarly replacing e by $\phi = -2\frac{\dot{x}}{x} \circ \frac{\dot{s}}{s}$, we can compute $(\ddot{x}, \ddot{\lambda}, \ddot{s})$ in exactly the same way as we did for $(\dot{x}, \dot{\lambda}, \dot{s})$. Since $(AXS^{-1}A^{\mathrm{T}})^{-1}$ has been calculated already, the operation count is negligible.

Remark 3.2 *It is worthwhile to mention that if A is a sparse matrics, a sparse QR decomposition [3] should be used to generate a sparse $\hat{A}$. This will make the rest computation more efficient for large problems.*

Extensive experience shows that it is crucial to have a large step $\sin(\alpha)$ to make Algorithm 3.1 efficient and competitive. Although the calculation of the search step size α by (39) is simple and the calculated α guarantees the polynomiality of the algorithm, the calculated α is very conservative. In view of the previous analysis, what we really need is to find a large step size α such that $(x(\alpha), \lambda(\alpha), s(\alpha)) \in \mathcal{N}_2(2\theta)$.

We can select a significantly larger α using negligible few extra operation counts as discussed as follows. First, the following simple estimations for sinusoidal functions are needed and can easily be verified.

$$\frac{1}{2}(1-\sin^2(\alpha)) \le 1-\sin(\alpha) \le 1-\sin^2(\alpha).$$

$$\frac{1}{2}\sin^2(\alpha) \le 1-\cos(\alpha) \le \sin^2(\alpha).$$

$$\sin^2(\alpha) \le \sin(\alpha) \le \frac{1}{2}(1+\sin^2(\alpha)).$$

From these estimations, we have

$$\frac{1}{8}(\sin^4(\alpha)-\sin^6(\alpha)) \le (1-\sin(\alpha))(1-\cos(\alpha))^2 \le (\sin^4(\alpha)-\sin^6(\alpha)).$$

$$\frac{1}{4}(\sin^4(\alpha)-\sin^6(\alpha)) \le (1-\sin(\alpha))\sin(\alpha)(1-\cos(\alpha)) \le \frac{1}{2}(\sin^2(\alpha)-\sin^6(\alpha)).$$

$$\frac{1}{8}\sin^8(\alpha) \le \sin(\alpha)(1-\cos(\alpha))^3 \le \frac{1}{2}(\sin^6(\alpha)+\sin^8(\alpha)).$$

Denote $p = x^k \circ s^k - \mu^k e$, $q = \ddot{x} \circ \ddot{s} - \dot{x} \circ \dot{s}$, and $r = -(\dot{x} \circ \ddot{s} + \dot{s} \circ \ddot{x})$, and define vectors ℓ_1, ℓ_2, and ℓ_3 as

$$\ell_1 = [\ell_{14}, \ell_{16}] = \begin{cases} \left[\frac{1}{8}, -\frac{1}{8}\right] & \text{if } p^{\mathrm{T}} q < 0 \\ [1, -1] & \text{if } p^{\mathrm{T}} q \ge 0, \end{cases}$$

$$\ell_2 = [\ell_{22}, \ell_{24}, \ell_{26}] = \begin{cases} \left[0, \frac{1}{4}, -\frac{1}{4}\right] & \text{if } p^{\mathrm{T}} r < 0 \\ \left[\frac{1}{2}, 0, -\frac{1}{2}\right] & \text{if } p^{\mathrm{T}} r \ge 0, \end{cases}$$

$$\ell_3 = [\ell_{36}, \ell_{38}] = \begin{cases} \left[0, \frac{1}{8}\right] & \text{if } q^{\mathrm{T}} r < 0 \\ \left[\frac{1}{2}, \frac{1}{2}\right] & \text{if } q^{\mathrm{T}} r \ge 0, \end{cases}$$

then

$$\begin{aligned}
& \|x(\alpha) \circ s(\alpha) - \mu(\alpha) e\|^2 \\
= \; & \|(\bar{x}^k \circ \bar{s}^k - \bar{\mu}^k e)(1-\sin(\alpha)) + (\ddot{x} \circ \ddot{s} - \dot{x} \circ \dot{s})(1-\cos(\alpha))^2 - (\dot{x} \circ \ddot{s} + \dot{s} \circ \ddot{x})\sin(\alpha)(1-\cos(\alpha))\|^2 \\
= \; & (p(1-\sin(\alpha)) + q(1-\cos(\alpha))^2 + r\sin(\alpha)(1-\cos(\alpha)))^{\mathrm{T}} \\
& (p(1-\sin(\alpha)) + q(1-\cos(\alpha))^2 + r\sin(\alpha)(1-\cos(\alpha))) \\
= \; & p^{\mathrm{T}} p(1-\sin(\alpha))^2 + 2p^{\mathrm{T}} q(1-\sin(\alpha))(1-\cos(\alpha))^2 + 2p^{\mathrm{T}} r(1-\sin(\alpha))\sin(\alpha)(1-\cos(\alpha)) \\
& + q^{\mathrm{T}} q(1-\cos(\alpha))^4 + 2q^{\mathrm{T}} r(1-\cos(\alpha))^3 \sin(\alpha) + r^{\mathrm{T}} r \sin^2(\alpha)(1-\cos(\alpha)))^2 \\
\le \; & p^{\mathrm{T}} p(1-2\sin^2(\alpha)+\sin^4(\alpha)) + 2p^{\mathrm{T}} q(1-\sin(\alpha))(1-\cos(\alpha))^2 + 2p^{\mathrm{T}} r(1-\sin(\alpha))\sin(\alpha)(1-\cos(\alpha)) \\
& + q^{\mathrm{T}} q \sin^8(\alpha) + 2q^{\mathrm{T}} r(1-\cos(\alpha))^3 \sin(\alpha) + r^{\mathrm{T}} r \sin^6(\alpha) \\
\le \; & p^{\mathrm{T}} p(1-2\sin^2(\alpha)+\sin^4(\alpha)) + 2p^{\mathrm{T}} q(\ell_{14}\sin^4(\alpha) + \ell_{16}\sin^6(\alpha)) \\
& + 2p^{\mathrm{T}} r(\ell_{22}\sin^2(\alpha) + \ell_{24}\sin^4(\alpha) + \ell_{26}\sin^6(\alpha)) \\
& + q^{\mathrm{T}} q \sin^8(\alpha) + 2q^{\mathrm{T}} r(\ell_{36}\sin^6(\alpha) + \ell_{38}\sin^8(\alpha)) + r^{\mathrm{T}} r \sin^6(\alpha) \\
= \; & \sin^8(\alpha)(q^{\mathrm{T}} q + 2q^{\mathrm{T}} r\ell_{38}) + \sin^6(\alpha)(r^{\mathrm{T}} r + 2q^{\mathrm{T}} r\ell_{36} + 2p^{\mathrm{T}} r\ell_{26} + 2p^{\mathrm{T}} q\ell_{16}) \\
& + \sin^4(\alpha)(2p^{\mathrm{T}} r\ell_{24} + 2p^{\mathrm{T}} q\ell_{14} + p^{\mathrm{T}} p) + \sin^2(\alpha)(2p^{\mathrm{T}} r\ell_{22} - 2p^{\mathrm{T}} p) + p^{\mathrm{T}} p.
\end{aligned}$$

Since

$$\begin{aligned}
& (\theta\bar{\mu}^k)^2(1-2\sin^2(\alpha)+\sin^4(\alpha)) = (\theta\bar{\mu}^k)^2(1-\sin^2(\alpha))^2 \\
\le \; & 4(\theta\bar{\mu}^k)^2(1-\sin(\alpha))^2 = (2\theta\mu(\alpha))^2,
\end{aligned}$$

finding the smallest positive solution $\sin(\alpha)$ for the following equation

$$\begin{aligned}
& \sin^8(\alpha)(q^{\mathrm{T}} q + 2q^{\mathrm{T}} r\ell_{38}) + \sin^6(\alpha)(r^{\mathrm{T}} r + 2q^{\mathrm{T}} r\ell_{36} + 2p^{\mathrm{T}} r\ell_{26} + 2p^{\mathrm{T}} q\ell_{16}) \\
+ \; & \sin^4(\alpha)(2p^{\mathrm{T}} r\ell_{24} + 2p^{\mathrm{T}} q\ell_{14} + p^{\mathrm{T}} p - (\theta\bar{\mu}^k)^2) \\
+ \; & \sin^2(\alpha)(2p^{\mathrm{T}} r\ell_{22} - 2p^{\mathrm{T}} p + 2(\theta\bar{\mu}^k)^2) + p^{\mathrm{T}} p - (\theta\bar{\mu}^k)^2 = 0
\end{aligned} \tag{53}$$

guarantees $\|x(\alpha) \circ s(\alpha) - \mu(\alpha)e\| \leq 2\theta\mu(\alpha)$. (53) is a quartic equation in terms of $\sin^2(\alpha)$ which has analytic solution [17] and the operation count is negligible. Computational experience shows that the $\sin(\alpha)$ obtained in this way is significantly larger than the one obtained from (39). Let $\sin(\bar{\alpha}) \leq 1$ be the larger real number obtained from either the smallest positive solution of (53) or the solution of (39), $\sin(\bar{\alpha})$ guarantees that Algorithm 3.1 is still polynomial with potentially better complexity bound.

We can find an even larger $\sin(\alpha)$ by searching α over the interval $[\sin^{-1}(\bar{\alpha}), 1]$ that satisfies the inequality

$$\begin{aligned} & p^{\mathrm{T}}p(1-\sin(\alpha))^2 + 2p^{\mathrm{T}}q(1-\sin(\alpha))(1-\cos(\alpha))^2 + 2p^{\mathrm{T}}r(1-\sin(\alpha))\sin(\alpha)(1-\cos(\alpha)) \\ & + q^{\mathrm{T}}q(1-\cos(\alpha))^4 + 2q^{\mathrm{T}}r(1-\cos(\alpha))^3\sin(\alpha) + r^{\mathrm{T}}r\sin^2(\alpha)(1-\cos(\alpha)))^2 \leq (2\theta\bar{\mu}^k(1-\sin(\alpha))^2 \end{aligned}$$

using bisection or golden section iterations. As p, q, and r are given, and the inequality is a scalar, for a fixed number, say 10 iterations, of bisection or golden section, the computational count is negligible, but the improvement can be significant. Therefore, in Step 3, the most expensive operations are the computations of p, q, r, $p^{\mathrm{T}}p$, $p^{\mathrm{T}}q$, $p^{\mathrm{T}}r$, $q^{\mathrm{T}}q$, $q^{\mathrm{T}}r$, and $r^{\mathrm{T}}r$ which need $O(n)$ operation counts. Clearly, this implementation guarantees $\sin(\alpha) \geq \theta^2 n^{-\frac{1}{2}}$ and $(x(\alpha), \lambda(\alpha), s(\alpha)) \in \mathcal{N}_2(2\theta)$, i.e., Lemmas 3.6 and 3.9 hold.

Finally, in Step 4, the operational count is mainly in the computation of $(\Delta x, \Delta\lambda, \Delta s)$. Similar to the implementation in Step 2, but replacing e by $\varphi = \mu(\alpha)x^{-1}(\alpha) \circ s^{-1}(\alpha) - e$, we can simplify the computation as

Case 1: $m > n - m$

$$\begin{aligned} \Delta x &= \hat{A}(\hat{A}^{\mathrm{T}}S(\alpha)X^{-1}(\alpha)\hat{A})^{-1}\hat{A}^{\mathrm{T}}S(\alpha)\varphi, \\ \Delta s &= S(\alpha)\varphi - S(\alpha)X^{-1}(\alpha)\Delta x, \\ \Delta\lambda &= -(AA^{\mathrm{T}})^{-1}A\Delta s. \end{aligned}$$

The most expensive operation in this calculation is to calculate $(\hat{A}^{\mathrm{T}}S(\alpha)X^{-1}(\alpha)\hat{A})^{-1}$ that needs $O((n-m)^2n+(n-m)^3)$. Since $(AA^{\mathrm{T}})^{-1}A$ is independent of the iteration, and all other computations have lower order operation counts, $O((n-m)^2n+(n-m)^3)$ is needed to get $(\Delta x, \Delta\lambda, \Delta s)$.

Case 2: $m \leq n - m$

$$\begin{aligned} \Delta s &= A^{\mathrm{T}}(AX(\alpha)S(\alpha)^{-1}A^{\mathrm{T}})^{-1}AX(\alpha)\varphi, \\ \Delta x &= X(\alpha)\varphi - X(\alpha)S(\alpha)^{-1}\Delta s, \\ \Delta\lambda &= -(AA^{\mathrm{T}})^{-1}A\Delta s. \end{aligned}$$

The most expensive operation in this calculation is to calculate $(AXS^{-1}A^{\mathrm{T}})^{-1}$ that needs $O(m^2n+m^3)$. Since $(AA^{\mathrm{T}})^{-1}A$ is independent of the iteration, and all other computations have lower order operation counts, $O(m^2n+m^3)$ is needed to get $(\Delta x, \Delta\lambda, \Delta s)$.

We summarize the total computational count in each iteration as follows.

- Step 2, Calculate $(\dot{x}, \dot{\lambda}, \dot{s})$ and $(\ddot{x}, \ddot{\lambda}, \ddot{s})$, for $m > n-m$, the operation count is $O((n-m)^2n+(n-m)^3)$, for $m \leq n-m$, the operation count is $O(m^2n+m^3)$;
- Step 3, Calculate α and update $(x(\alpha), \lambda(\alpha), s(\alpha))$, the operation count is negligible;
- Step 4, Calculate $(\Delta x, \Delta\lambda, \Delta s)$ and update $(\bar{x}^k, \bar{\lambda}^k, \bar{s}^k)$, for $m > n-m$, the operation count is $O((n-m)^2n+(n-m)^3)$, for $m \leq n-m$, the operation count is $O(m^2n+m^3)$.

4 Convergence Analyses

Let (x^*, λ^*, s^*) be any solution of (5). Let index sets $\mathcal{B}$, $\mathcal{N}$ be defined as

$$\mathcal{B} = \{j \in \{1, \ldots, n\} \mid x_j^* \neq 0\}. \tag{56}$$

$$\mathcal{N} = \{j \in \{1, \ldots, n\} \mid s_j^* \neq 0\}. \tag{57}$$

From Goldman-Tucker theorem [5], it can be shown [18] that there always exist a solution (x^*, λ^*, s^*) of (5), where x^* is a solution of the primary linear programming and (λ^*, s^*) is a solution of the dual linear programming, such that $\mathcal{B} \cap \mathcal{N} = \emptyset$ and $\mathcal{B} \cup \mathcal{N} = \{1, \ldots, n\}$. I.e., $x^* \circ s^* = 0$, and $x^* + s^* > 0$. An optimal solution with this property is called strictly complementary. The following Lemma, independent of any algorithm, is given in [18].

Lemma 4.1 *Let $\mu^0 > 0$, and $\gamma \in (0,1)$, Then for all points (x, λ, s) with $(x, \lambda, s) \in \mathcal{F}^o$, $x_i s_i > \gamma\mu$ (where $\mu = \frac{x^{\mathrm{T}} s}{n}$), and $\mu < \mu^0$, there are constants M, C_1, and C_2 such that*

$$\|(x, s)\| \leq M, \tag{58}$$

$$0 < x_i \leq \mu / C_1 \quad (i \in \mathcal{N}), \qquad 0 < s_i \leq \mu / C_1 \quad (i \in \mathcal{B}). \tag{59}$$

$$s_i \geq C_2 \gamma \quad (i \in \mathcal{N}), \qquad x_i \geq C_2 \gamma \quad (i \in \mathcal{B}). \tag{60}$$

∎

Lemma 4.2 *Let $(\bar{x}^k, \bar{\lambda}^k, \bar{s}^k)$ be generated by Algorithms 3.1. Then, $(\bar{x}^k, \bar{s}^k)$ has at least one limit point. Moreover, every limit point is a strictly complementary primary-dual solution of the linear programming, i.e.,*

$$s_i^* \geq C_2 \gamma \quad (i \in \mathcal{N}), \qquad x_i^* \geq C_2 \gamma \quad (i \in \mathcal{B}). \tag{61}$$

Proof: Similar to the proof in [19]. ∎

To show that Algorithm 3.1 is polynomial, we need a result in [18].

Theorem 4.1 *Let $\epsilon \in (0,1)$ be given. Suppose that an algorithm for solving (5) generates a sequence of iterations that satisfies*

$$\mu^{k+1} \leq \left(1 - \frac{\delta}{n^\omega}\right) \mu^k, \quad k = 0, 1, 2, \ldots, \tag{62}$$

for some positive constants δ and ω. Suppose that the starting point (x^0, λ^0, s^0) satisfies $\mu^0 \leq 1/\epsilon$. Then there exists an index K with

$$K = O(n^\omega \log(1/\epsilon))$$

such that

$$\mu^k \leq \epsilon \quad \text{for} \quad \forall k \geq K.$$

∎

In view of Lemma 3.9 and Theorem 4.1, we have

Theorem 4.2 *The sequence $(\bar{x}^k, \bar{\lambda}^k, \bar{s}^k)$ generated by Algorithms 3.1 globally converges to a set of limit points (x^*, λ^*, s^*). For every limit point (x^*, λ^*, s^*), x^* is the optimal solution of the primal problem, (λ^*, s^*) is the optimal solution of the dual problem, and (x^*, s^*) is strictly complementary. Moreover, Algorithms 3.1 is a polynomial algorithm with polynomial complexity bound of $O(n^{\frac{1}{2}} \log(1/\epsilon))$.* ∎

5 Numerical Tests

The algorithm developed in this paper is implemented in a Matlab function `arc2`. Numerical tests have been performed for linear programming problems in Netlib LP library. For Netlib LP problems, [2] has classified these problems into two categories: problems with strict interior-point and problems without strict interior-point. Though the newly developed Matlab codes and other existing codes can solve problems without strict interior-point, we are most interested in the problems with strict interior-point that is assumed by all interior-point methods. Among these problems, we only choose problems which

Table 1: Iteration counts for test problems in Netlib and Matlab

	iterations used by different algorithms		
Problem	arc2	linprog	source
AFIRO	4	7	netlib
blend	10	8	netlib
SCAGR25	5	16	netlib
SCAGR7	6	12	netlib
SCSD1	13	10	netlib
SCSD6	17	12	netlib
SCSD8	13	11	netlib
SCTAP1	14	17	netlib
SCTAP2	13	18	netlib
SCTAP3	14	18	netlib
SHARE1B	9	22	netlib

are presented in standard form and their A are full rank matrices. The selected problems are solved using our Matlab function `arc2` and function `linprog` in Matlab optimization toolBox. The iterations used to solve these problems are compared and the iteration numbers are listed in table 1. Only two Netlib problems that are classified as problems with strict interior-point and are presented in standard form are not included in the table because the PC computer used in the test does not have enough memory to handle problems of this size.

In our implementation, the parameters in the proposed algorithms are chosen as follows. The stopping criterion used in Algorithm 3.1 is the same as `linprog` [20]

$$\frac{\|r_B\|}{\max\{1, \|b\|\}} + \frac{\|r_C\|}{\max\{1, \|c\|\}} + \frac{\mu}{\max\{1, \|c^{\mathrm{T}}x\|, \|b^{\mathrm{T}}\lambda\|\}} < 10^{-8}.$$

For all problems, the initial point is set to $x = s = e$. The initial point used in `linprog` is essentially the same as used in [11] with some minor modifications (see [20]).

It is clear that Algorithm 3.1 uses less iteration than `linprog` to find the optimal solutions in most tested Netlib problems. In total, `arc2` uses 118 iterations to solve these 11 problems; `linprog` uses 151 iterations to solve these 11 problems; i.e., the total iterations `arc2` used for all these problems is about 20% less than the total iterations `linprog` used.

6 Conclusions

This paper proposes a polynomial interior-point path-following algorithm that searches the optimizers along the ellipses that approximate central path. The algorithm is proved to be polynomial with the best known complexity bound. Numerical test results in all the tested Netlib problems show that the algorithm is promising compared to Matlab code `linprog` which implements the state-of-art Mehrotra's Algorithm.

7 Acknowledgements

The author wishes to thank Professor André L. Tits and Professor T. Terlaky for their valuable comments on a related paper [19] that helped him in preparing this paper.

References

[1] R. Bland, D. Goldfarb, and M. Todd, *The ellipsoid method: A survey*, Operations Research, 29 (1981), pp. 1039–1091.

[2] C. Cartis and N. I. M. Gould, *Finding a point in the relative interior of a polyhedren*, Technical Report NA-07/01, Computing Laboratory, Oxford University, Oxford, UK, 2007.

[3] T. A. Davis, *Multifrontal multithreaded rank-revealing sparse qr factorization*, technical report, Department of Computer and information Science and engineering, University of Florida, Florida, 2008.

[4] M. P. do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice-Hall, New Jersey, 1976.

[5] A. Goldman and A. Tucker, *Theory of linear programming*, in Linear Equalities and Related Systems, H. Kuhn and Tucker, eds., Princeton University Press, Princeton, N.J, 1956, pp. 53–97.

[6] N. Karmarkar, *A new polynomial-time algorithm for linear programming*, Combinatorica, 4 (1984), pp. 375–395.

[7] L. Khachiyan, *A polynomial algorithm in linear programming (in russian)*, Doklady Akademiia Nauk SSSR, 224 (1979), pp. 1093–1096.

[8] V. Klee and G. Minty, *How good is the simplex algorithm?*, in Inequalities, III, O. Shisha, ed., Academic Press, 1972, pp. 159–175.

[9] M. Kojima, S. Mizuno, and A. Yoshise, *A polyomial-time algorithm for a class of linear complementarity problems*, Mathematical Programming, 44 (1989), pp. 1–26.

[10] ———, *A primal-dual interior point algorithm for linear programming*, in Progress in Mathematical Programming: Interior-Point and Related Methods, N. Megiddo, ed., Springer-Verlag, New York, 1989, pp. 29–47.

[11] I. Lustig, R. Marsten, and D. Shanno, *On implementing mehrotra's predictor-corrector interior point method for linear programming*, SIAM Journal on Optimization, 2 (1992), pp. 432–449.

[12] S. Mehrotra, *On the implementation of a primal-dual interior point method*, SIAM Journal on Optimization, 2 (1992), pp. 575–601.

[13] S. Mizuno, M. Todd, and Y. Ye, *On adaptive step primal-dual interior-point algorithms for linear programming*, Mathematics of Operartions Research, 18 (1993), pp. 964–981.

[14] R. Monteiro and I. Adler, *Interior path following primal-dual algorithms. part i: linear programming*, Mathematical Programming, 44 (1989), pp. 27–41.

[15] R. Monteiro, I. Adler, and M. Resende, *A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension*, Mathematics of Operations Research, 15 (1990), pp. 191–214.

[16] Y. Nesterov, *Long-step strategies in interior-point primal-dual methods*, Mathematical Programming, 76 (1996), pp. 47–94.

[17] A. D. Polyanin and A. V. Manzhirov, *Handbook of Mathematics For Engineers and Scientists*, Chapman & Hall/CRC, Noca Raton, FL, 2007.

[18] S. Wright, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, 1997.

[19] Y. Yang, *Arc-search path-following interior-point algorithms for linear programming*, Optimization Online, August (2009).

[20] Y. Zhang, *Solving large-scale linear programs by interior-point methods under the matlab environment*, Technical Report TR96-01, Department of Mathematics and Statistics, University of Maryland, Baltimore County, Marland, 1996.