A Polynomial Arc-Search Interior-Point Algorithm for Convex Quadratic Programming

Yaguang Yang*

Abstract

Arc-search is developed for linear programming in [24, 25]. The algorithms search for optimizers along an ellipse that are approximations of the central path. In this paper, the arc-search method is applied to primal-dual path-following interior-point method for convex quadratic programming. A simple algorithm with iteration complexity $O(\sqrt{n}\log(1/\epsilon))$ is devised. Several improvements on the simple algorithm, which improve computational efficiency, increase step length, and further reduce duality gap in every iteration, are then proposed and implemented. It is intuitively clear that the iteration with these improvements will significantly reduce the duality gap more than the iteration of the simple algorithm without the improvements, though it is hard to show how much these improvements reduce the complexity bound. The proposed algorithm is implemented in MATLAB and tested on seven quadratic programming problems originating from [13]. The result is compared to the one obtained by LOQO in [22]. The proposed algorithm uses fewer iterations in all seven problems and the number of total iterations is 27% fewer than the one obtained by LOQO. This preliminary result shows that the proposed algorithm is promising.

Keywords: Arc-search, convex quadratic programming, polynomial algorithm.

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

1 Introduction

Interior-point algorithms for convex quadratic programming have been studied for decades by many researchers. One of the earliest and well-known algorithms was proposed by Dikin [6]. After Karmarkar's pioneer work on interior-point polynomial algorithm for linear programming [14], interior-point polynomial algorithms for convex quadratic programming have been investigated by many researchers. For example, Ye and Tse [26] extended Karmarkar's algorithm and proved that it has polynomial complexity bound $O(n\log(1/\epsilon))$. Monteiro and Adler proposed a different algorithm [18] and improved the complexity bound to $O(\sqrt{n}\log(1/\epsilon))$. However, these algorithms do not use higher-order information which is believed intuitively (and demonstrated by numerical test) to be useful in practical implementation [15].

Monteiro, Adler, and Resende [19] proposed a higher-order algorithm and showed that it has complexity bound $O(n^{\frac{1}{2}(1+\frac{1}{r})}\log(1/\epsilon))$, where $r \geq 1$ is the order of derivatives. Similar to many other higher-order algorithms which have been proposed for linear programming and convex quadratic programming, this algorithm searches optimizers along a straight line which is the linear combination of the first order and higher order derivatives of the central path. It is worthwhile to indicate that although this algorithm may be more efficient than the one proposed earlier by the authors [18], the complexity bound of the higher-order algorithm is worse than the one in [18]. [9] considered multiple centrality corrections for linear programming and provided numerical results to demonstrate the potential benefit of higher-order method, but [9] did not show the polynomiality of this method.

A great deal of research on the infinitesimal approximation of central path using higher-order derivatives has been done, for example, [2]. Searching optimizers along some infinitesimal approximation of central path, such as power series approximation, has also been investigated [15]. But there is few literature to show computational superiority of arc-search. Also, there is few literature to report computational experience of polynomial algorithms for convex quadratic programming.

On the other hand, interior point algorithms implemented in popular software, such as MPC [23] for linear programming, HOPDM [1] for quadratic programming, LOQO [22] for quadratic and nonlinear programming, are not proved to be polynomial.

Recently, [24] suggested approximating the central path by using ellipse, and developed an algorithm for linear programming which searches optimizers along the ellipse. [25] proposed a polynomial algorithm with complexity bound $O(\sqrt{n}\log(1/\epsilon))$ for linear programming which searches optimizers along ellipse and uses correction step. It also provided some promising computational results compared to MPC.

This paper extends the arc-search techniques developed in [25] to path-following interior-point method for convex quadratic programming. We will show that the algorithm has polynomial complexity bound $O(\sqrt{n}\log(1/\epsilon))$ which is better than the higher-order algorithm proposed in [19]. Since higher-order information is used in the construction of the search path, we expect that the algorithm is more efficient than [6, 18, 26] in practice. A MATLAB code is implemented for the algorithm. A simple example is used to demonstrate how this algorithm works. A small set of quadratic programming test problems originating from [13] is used to test the algorithm developed here. The result is compared to the one obtained by LOQO [10]. This preliminary result shows that the proposed algorithm is promising because it uses fewer iterations in all seven tested problems than the iterations LOQO uses.

Throughout the paper, we will use notations adopted in [24]. We denote n-dimensional vector space by $\mathbf{R}^n$, $n \times m$-dimensional matrix space by $\mathbf{R}^{n\times m}$, Hadamard (element-wise) product of two vectors $x \in \mathbf{R}^n$ and $s \in \mathbf{R}^n$ by $x \circ s$, the ith component of x by x_i, element-wise division of the two vectors by $\frac{x}{s}$ if $\min|s_i| > 0$, the Euclidean norm of x by $\|x\|$, the identity matrix of any dimension by I, the vector of all ones with appropriate dimension by e, the transpose of matrix $A \in \mathbf{R}^{m\times n}$ by $A^{\mathrm{T}} \in \mathbf{R}^{n\times m}$, element-wise absolute value vector by $|x| = [|x_1|, \ldots, |x_n|]^{\mathrm{T}}$, an orthogonal base of the null space of A by $\hat{A} \in \mathbf{R}^{n\times m}$. 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. For a matrix $H \in \mathbf{R}^{n\times n}$, we use $H \geq 0$ if H is positive semidefinite, and $H > 0$ if H is positive definite. Finally, we define an initial point of any algorithm by x^0, the point after the kth iteration by x^k.

2 Problem Description

Consider the Convex Quadratic Programming (QP) in the standard form:

$$(QP) \quad \min \ \frac{1}{2}x^{\mathrm{T}}Hx + c^{\mathrm{T}}x, \quad \text{subject to} \quad Ax = b, \quad x \geq 0, \tag{1}$$

where $0 \leq H \in \mathbf{R}^{n\times n}$ is a positive semidefinite matrix, $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$, $c \in \mathbf{R}^n$ are given, and $x \in \mathbf{R}^n$ is the vector to be optimized. Associated with the quadratic programming is the dual programming (DP) that is also presented in the standard form:

$$(DP) \quad \max \ -\frac{1}{2}x^{\mathrm{T}}Hx + b^{\mathrm{T}}\lambda, \quad \text{subject to} \quad -Hx + A^{\mathrm{T}}\lambda + s = c, \quad s \geq 0, \quad x \geq 0, \tag{2}$$

where $\lambda \in \mathbf{R}^m$ is the dual variable vector, and $s \in \mathbf{R}^n$ is the dual slack vector.

Denote the feasible set $\mathcal{F}$ as a collection of all points that meet the constraints of (QP) and (DP).

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

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

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

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 existence of a central path. Since $1 \leq m$, assumption 2 implies $2 \leq n$. It is well known that $x \in \mathbf{R}^n$ is an optimal solution of (1) if and only if x, λ, and s meet the following KKT conditions

$$Ax = b \tag{5a}$$
$$A^{\mathrm{T}}\lambda + s - Hx = c \tag{5b}$$
$$x_i s_i = 0, \quad i = 1, \ldots, n \tag{5c}$$
$$(x, s) \geq 0. \tag{5d}$$

For convex (QP) problem, KKT condition is also sufficient for x to be a global optimal solution. Similar to the linear programming, the central path following algorithm proposed here tries to search the optimizers (located at the boundary of $\mathcal{F}$) along an arc that is an approximation of the central path $\mathcal{C} \in \mathcal{F}^o \subset \mathcal{F}$, where the central path $\mathcal{C}$ is parameterized by a scalar $\tau > 0$ as follows. For each interior point (x, λ, s) on the central path, there is a $\tau > 0$ such that

$$Ax = b \tag{6a}$$
$$A^{\mathrm{T}}\lambda + s - Hx = c \tag{6b}$$
$$x_i s_i = \tau, \quad i = 1, \ldots, n \tag{6c}$$
$$(x, s) > 0. \tag{6d}$$

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 $\tau \to 0$, the moving point $(x(\tau), \lambda(\tau), s(\tau))$ on the central path represented by (6) approaches to the solution of (QP) represented by (1).

3 An Arc-search Algorithm for Convex Quadratic Programming

Let $1 > \theta > 0$, and

$$\mu = \frac{x^{\mathrm{T}} s}{n}. \tag{8}$$

We denote

$$\mathcal{N}_2(\theta) = \{(x, \lambda, s) : Ax = b, -Hx + A^{\mathrm{T}}\lambda + s = c, (x, s) > 0, \|x \circ s - \mu e\| \le \theta\mu\} \subset \mathcal{F}^o. \tag{9}$$

For $(x, \lambda, s) \in \mathcal{N}_2(\theta)$, since $(1-\theta)\mu \le x_i s_i \le (1+\theta)\mu$, we have

$$\frac{x_i s_i}{1+\theta} \le \frac{\max_i x_i s_i}{1+\theta} \le \mu \le \frac{\min_i x_i s_i}{1-\theta} \le \frac{x_i s_i}{1-\theta}. \tag{10}$$

The idea of arc-search proposed in this paper is very simple. The algorithm starts from a feasible point in $\mathcal{N}_2(\theta)$ close to the central path, constructs an arc that passes through the point and approximates the central path, searches along the arc to a new point in a larger area $\mathcal{N}_2(2\theta)$ that reduces the duality gap $x^{\mathrm{T}} s$ and meets (6a), (6b), and (6d). The process is repeated by finding a better point close to the central path or on the central path in $\mathcal{N}_2(\theta)$ that simultaneously meets (6a), (6b), and (6d).

We will use an ellipse $\mathcal{E}$ [7] in $2n + m$ dimensional space to approximate the central path $\mathcal{C}$ described by (6), where

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

$\vec{a} \in \mathbf{R}^{2n+m}$ 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, $\vec{a}$, $\vec{b}$, $\vec{c}$ are functions of α, y, $\dot{y}$, and $\ddot{y}$, where $\dot{y}$ and $\ddot{y}$ are defined as

$$\begin{bmatrix} A & 0 & 0 \\ -H & A^{\mathrm{T}} & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{\lambda} \\ \dot{s} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ x \circ s \end{bmatrix}, \tag{12}$$

$$\begin{bmatrix} A & 0 & 0 \\ -H & A^{\mathrm{T}} & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \ddot{\lambda} \\ \ddot{s} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -2\dot{x} \circ \dot{s} \end{bmatrix}. \tag{13}$$

It has been shown in [24] that one can avoid the calculation of $\vec{a}$, $\vec{b}$, and $\vec{c}$ in the expression of the ellipse. The following formulas are used instead.

Theorem 3.1 *Let $(x(\alpha), \lambda(\alpha), s(\alpha))$ be an arc defined by (11) 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 (12) and (13). Then an ellipse approximation of the central path is given by*

$$x(\alpha) = x - \dot{x}\sin(\alpha) + \ddot{x}(1 - \cos(\alpha)). \tag{14}$$

$$\lambda(\alpha) = \lambda - \dot{\lambda}\sin(\alpha) + \ddot{\lambda}(1 - \cos(\alpha)). \tag{15}$$

$$s(\alpha) = s - \dot{s}\sin(\alpha) + \ddot{s}(1 - \cos(\alpha)). \tag{16}$$

∎

Assuming $(x, s) > 0$, one can easily see that if $\frac{\dot{x}}{x}$, $\frac{\ddot{x}}{x}$, $\frac{\dot{s}}{s}$, and $\frac{\ddot{s}}{s}$ are bounded (we will show that this is true), and if α is small enough, then $x(\alpha) > 0$ and $s(\alpha) > 0$. We will also show that searching along this arc will reduce the duality gap, i.e., $\mu(\alpha) = \frac{x(\alpha)^{\mathrm{T}} s(\alpha)}{n} < \mu$.

Lemma 3.1 *Let (x, λ, s) be a strictly feasible point of (QP) and (DP), $(\dot{x}, \dot{\lambda}, \dot{s})$ and $(\ddot{x}, \ddot{\lambda}, \ddot{s})$ meet (12) and (13), $(x(\alpha), \lambda(\alpha), s(\alpha))$ be calculated using (14), (15), and (16), then the following conditions hold.*

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

Proof: Since (x, λ, s) is a strict feasible point, the result follows from direct calculation by using (4), (12), (13), and Theorem 3.1. ∎

Lemma 3.2 *Let $\dot{x}$, $\dot{s}$, $\ddot{x}$, and $\ddot{s}$ be defined in (12) and (13), and H be positive semidefinite matrix. Then the following relations hold.*

$$\dot{x}^{\mathrm{T}}\dot{s} = \dot{x}^{\mathrm{T}}H\dot{x} \geq 0, \tag{17}$$

$$\ddot{x}^{\mathrm{T}}\ddot{s} = \ddot{x}^{\mathrm{T}}H\ddot{x} \geq 0. \tag{18}$$

$$\ddot{x}^{\mathrm{T}}\dot{s} = \dot{x}^{\mathrm{T}}\ddot{s} = \dot{x}^{\mathrm{T}}H\ddot{x}. \tag{19}$$

$$\begin{aligned} & -(\dot{x}^{\mathrm{T}}\dot{s})(1-\cos(\alpha))^2 - (\ddot{x}^{\mathrm{T}}\ddot{s})\sin^2(\alpha) \\ \leq\ & (\ddot{x}^{\mathrm{T}}\dot{s} + \dot{x}^{\mathrm{T}}\ddot{s})\sin(\alpha)(1-\cos(\alpha)) \\ \leq\ & (\dot{x}^{\mathrm{T}}\dot{s})(1-\cos(\alpha))^2 + (\ddot{x}^{\mathrm{T}}\ddot{s})\sin^2(\alpha). \end{aligned} \tag{20}$$

$$\begin{aligned} & -(\dot{x}^{\mathrm{T}}\dot{s})\sin^2(\alpha) - (\ddot{x}^{\mathrm{T}}\ddot{s})(1-\cos(\alpha))^2 \\ \leq\ & (\ddot{x}^{\mathrm{T}}\dot{s} + \dot{x}^{\mathrm{T}}\ddot{s})\sin(\alpha)(1-\cos(\alpha)) \\ \leq\ & (\dot{x}^{\mathrm{T}}\dot{s})\sin^2(\alpha) + (\ddot{x}^{\mathrm{T}}\ddot{s})(1-\cos(\alpha))^2. \end{aligned} \tag{21}$$

For $\alpha = \frac{\pi}{2}$, (20) and (21) reduce to

$$-\left(\dot{x}^{\mathrm{T}}\dot{s} + \ddot{x}^{\mathrm{T}}\ddot{s}\right) \leq (\ddot{x}^{\mathrm{T}}\dot{s} + \dot{x}^{\mathrm{T}}\ddot{s}) \leq \dot{x}^{\mathrm{T}}\dot{s} + \ddot{x}^{\mathrm{T}}\ddot{s}. \tag{22}$$

Proof: Pre-multiplying $\dot{x}^{\mathrm{T}}$ to the second rows of (12) and pre-multiplying $\ddot{x}^{\mathrm{T}}$ to the second rows of (13), then using the first rows of (12) and (13) gives $\dot{x}^{\mathrm{T}}\dot{s} = \dot{x}^{\mathrm{T}}H\dot{x}$ and $\ddot{x}^{\mathrm{T}}\ddot{s} = \ddot{x}^{\mathrm{T}}H\ddot{x}$. (17) and (18) follow from the fact that H is positive semidefinite. Pre-multiply $\ddot{x}^{\mathrm{T}}$ in the second row of (12) and using the first row of (13), we have

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

Pre-multiply $\dot{x}^{\mathrm{T}}$ in the second row of (13) and using the first row of (12), we have

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

This gives,

$$\begin{aligned} & (\dot{x}(1-\cos(\alpha)) + \ddot{x}\sin(\alpha))^{\mathrm{T}}H(\dot{x}(1-\cos(\alpha)) + \ddot{x}\sin(\alpha)) \\ =\ & (\dot{x}^{\mathrm{T}}H\dot{x})(1-\cos(\alpha))^2 + 2(\dot{x}^{\mathrm{T}}H\ddot{x})\sin(\alpha)(1-\cos(\alpha)) + (\ddot{x}^{\mathrm{T}}H\ddot{x})\sin^2(\alpha) \\ =\ & (\dot{x}^{\mathrm{T}}\dot{s})(1-\cos(\alpha))^2 + (\ddot{x}^{\mathrm{T}}\ddot{s})\sin^2(\alpha) + (\ddot{x}^{\mathrm{T}}\dot{s} + \dot{x}^{\mathrm{T}}\ddot{s})\sin(\alpha)(1-\cos(\alpha)) \geq 0, \end{aligned}$$

which is the first inequality of (20), and

$$\begin{aligned} & (\dot{x}(1-\cos(\alpha)) - \ddot{x}\sin(\alpha))^{\mathrm{T}}H(\dot{x}(1-\cos(\alpha)) - \ddot{x}\sin(\alpha)) \\ =\ & (\dot{x}^{\mathrm{T}}H\dot{x})(1-\cos(\alpha))^2 - 2(\dot{x}^{\mathrm{T}}H\ddot{x})\sin(\alpha)(1-\cos(\alpha)) + (\ddot{x}^{\mathrm{T}}H\ddot{x})\sin^2(\alpha) \\ =\ & (\dot{x}^{\mathrm{T}}\dot{s})(1-\cos(\alpha))^2 + (\ddot{x}^{\mathrm{T}}\ddot{s})\sin^2(\alpha) - (\ddot{x}^{\mathrm{T}}\dot{s} + \dot{x}^{\mathrm{T}}\ddot{s})\sin(\alpha)(1-\cos(\alpha)) \geq 0, \end{aligned}$$

which is the second inequality of (20). Replacing $\dot{x}(1-\cos(\alpha))$ and $\ddot{x}\sin(\alpha)$ by $\dot{x}\sin(\alpha)$ and $\ddot{x}(1-\cos(\alpha))$, and following the same method, we can obtain (21). ∎

Two simple Lemmas will be used. The first one is given in [24].

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

The second Lemma is very useful and proved in [17].

Lemma 3.4 *Let u, v, and w be real vectors of same size satisfying $u + v = w$ and $u^{\mathrm{T}}v \geq 0$. Then,*

$$2\|u\| \cdot \|v\| \leq \|u\|^2 + \|v\|^2 \leq \|u\|^2 + \|v\|^2 + 2u^{\mathrm{T}}v = \|u+v\|^2 = \|w\|^2. \tag{23}$$

∎

Lemma 3.5 *Let $(x, \lambda, s) \in \mathcal{N}_2(\theta)$ and $(\dot{x}, \dot{\lambda}, \dot{s})$ meet (12). Then,*

$$\left\|\frac{\dot{x}}{x}\right\|^2 + \left\|\frac{\dot{s}}{s}\right\|^2 \leq \frac{n}{1-\theta}, \tag{24}$$

$$\left\|\frac{\dot{x}}{x}\right\|^2 \left\|\frac{\dot{s}}{s}\right\|^2 \leq \left(\frac{n}{2(1-\theta)}\right)^2, \tag{25}$$

$$0 \leq \frac{\dot{x}^{\mathrm{T}}\dot{s}}{\mu} \leq \frac{n(1+\theta)}{2(1-\theta)} := c_1 n. \tag{26}$$

Proof: From the last row of (12), we have

$$\begin{aligned} & S\dot{x} + X\dot{s} = XSe \\ \iff \quad & X^{-\frac{1}{2}}S^{\frac{1}{2}}\dot{x} + X^{\frac{1}{2}}S^{-\frac{1}{2}}\dot{s} = X^{\frac{1}{2}}S^{\frac{1}{2}}e. \end{aligned}$$

Let $u = X^{-\frac{1}{2}}S^{\frac{1}{2}}\dot{x}$, $v = X^{\frac{1}{2}}S^{-\frac{1}{2}}\dot{s}$, and $w = X^{\frac{1}{2}}S^{\frac{1}{2}}e$, from Lemma 3.2, $u^{\mathrm{T}}v \geq 0$. Using Lemma 3.4, we have

$$\|u\|^2 + \|v\|^2 = \sum_{i=1}^{n} \frac{\dot{x}_i^2 s_i}{x_i} + \sum_{i=1}^{n} \frac{\dot{s}_i^2 x_i}{s_i} \leq \|X^{\frac{1}{2}}S^{\frac{1}{2}}e\|^2 = n\mu.$$

Dividing both sides of the inequality by $\min_j s_j x_j$ and using (10) gives

$$\sum_{i=1}^{n} \frac{\dot{x}_i^2}{x_i^2} + \sum_{i=1}^{n} \frac{\dot{s}_i^2}{s_i^2} \leq \frac{n\mu}{\min_j s_j x_j} \leq \frac{n}{1-\theta},$$

or equivalently

$$\left\|\frac{\dot{x}}{x}\right\|^2 + \left\|\frac{\dot{s}}{s}\right\|^2 \leq \frac{n}{1-\theta}.$$

This proves (24). Combining (24) and Lemma 3.3 yields

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

This leads to,

$$\left\|\frac{\dot{x}}{x}\right\| \left\|\frac{\dot{s}}{s}\right\| \leq \frac{n}{2(1-\theta)}. \tag{27}$$

Therefore, using (10) and Cauchy-Schwarz inequality yields

$$\frac{\dot{x}^{\mathrm{T}}\dot{s}}{\mu} \leq \frac{|\dot{x}|^{\mathrm{T}}|\dot{s}|}{\mu} \leq (1+\theta)\frac{|\dot{x}|^{\mathrm{T}}|\dot{s}|}{\max_i x_i s_i} \leq (1+\theta)\left(\frac{|\dot{x}|}{x}\right)^{\mathrm{T}}\left(\frac{|\dot{s}|}{s}\right) \leq (1+\theta)\left\|\frac{\dot{x}}{x}\right\|\left\|\frac{\dot{s}}{s}\right\| \leq \frac{n(1+\theta)}{2(1-\theta)}, \tag{28}$$

which is the second inequality of (26). From Lemma 3.2, $\dot{x}^{\mathrm{T}}\dot{s} = \dot{x}^{\mathrm{T}}H\dot{x} \geq 0$, we have the first inequality of (26). ∎

Lemma 3.6 *Let $(x, \lambda, s) \in \mathcal{N}_2(\theta)$, $(\dot{x}, \dot{\lambda}, \dot{s})$ and $(\ddot{x}, \ddot{\lambda}, \ddot{s})$ meet (12) and (13). Then*

$$\left\|\frac{\ddot{x}}{x}\right\|^2 + \left\|\frac{\ddot{s}}{s}\right\|^2 \leq \frac{(1+\theta)n^2}{(1-\theta)^3}. \tag{29}$$

$$0 \leq \frac{\ddot{x}^{\mathrm{T}}\ddot{s}}{\mu} \leq \frac{n^2(1+\theta)^2}{2(1-\theta)^3} := c_2 n^2. \tag{30}$$

$$\left|\frac{\dot{x}^{\mathrm{T}}\ddot{s}}{\mu}\right| \leq \frac{n^{\frac{3}{2}}(1+\theta)^{\frac{3}{2}}}{(1-\theta)^2} := c_3 n^{\frac{3}{2}}, \quad \left|\frac{\ddot{x}^{\mathrm{T}}\dot{s}}{\mu}\right| \leq \frac{n^{\frac{3}{2}}(1+\theta)^{\frac{3}{2}}}{(1-\theta)^2} := c_3 n^{\frac{3}{2}}. \tag{31}$$

Proof: Similar to the proof of Lemma 3.5, from the last row of (13), we have

$$\begin{aligned} & S\ddot{x} + X\ddot{s} = -2\,(\dot{x}\circ\dot{s}) \\ \Longleftrightarrow\quad & X^{-\frac{1}{2}}S^{\frac{1}{2}}\ddot{x} + X^{\frac{1}{2}}S^{-\frac{1}{2}}\ddot{s} = -2X^{-\frac{1}{2}}S^{-\frac{1}{2}}\,(\dot{x}\circ\dot{s})\,, \end{aligned}$$

Let $u = X^{-\frac{1}{2}}S^{\frac{1}{2}}\ddot{x}$, $v = X^{\frac{1}{2}}S^{-\frac{1}{2}}\ddot{s}$, and $w = -2X^{-\frac{1}{2}}S^{-\frac{1}{2}}\,(\dot{x}\circ\dot{s})$, from Lemma 3.2, $u^{\mathrm{T}}v \geq 0$. Using Lemma 3.4, we have

$$\|u\|^2 + \|v\|^2 = \sum_{i=1}^{n}\frac{\ddot{x}_i^2 s_i}{x_i} + \sum_{i=1}^{n}\frac{\ddot{s}_i^2 x_i}{s_i} \leq \left\|-2X^{-\frac{1}{2}}S^{-\frac{1}{2}}\,(\dot{x}\circ\dot{s})\right\|^2 = 4\sum_{i=1}^{n}\left(\frac{\dot{x}_i^2}{x_i}\frac{\dot{s}_i^2}{s_i}\right)$$

Dividing both sides of the inequality by μ and using (10) gives

$$(1-\theta)\left(\sum_{i=1}^{n}\frac{\ddot{x}_i^2}{x_i^2} + \sum_{i=1}^{n}\frac{\ddot{s}_i^2}{s_i^2}\right) \leq 4(1+\theta)\left(\sum_{i=1}^{n}\left(\frac{\dot{x}_i^2}{x_i^2}\frac{\dot{s}_i^2}{s_i^2}\right)\right),$$

or equivalently

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

Therefore,

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

and

$$\frac{\ddot{x}^{\mathrm{T}}\ddot{s}}{\mu} \leq \frac{|\ddot{x}|^{\mathrm{T}}|\ddot{s}|}{\mu} \leq (1+\theta)\frac{|\ddot{x}|^{\mathrm{T}}|\ddot{s}|}{\max_i x_i s_i} \leq (1+\theta)\left(\frac{|\ddot{x}|}{x}\right)^{\mathrm{T}}\left(\frac{|\ddot{s}|}{s}\right) \leq (1+\theta)\left\|\frac{\ddot{x}}{x}\right\|\left\|\frac{\ddot{s}}{s}\right\| \leq \frac{n^2(1+\theta)^2}{2(1-\theta)^3},$$

which is the second inequality of (30). From Lemma 3.2, $\ddot{x}^{\mathrm{T}}\ddot{s} = \ddot{x}^{\mathrm{T}}H\ddot{x} \geq 0$, we have the first inequality of (30). Similarly, it is easy to show (31). ∎

We will use a lemma in [24].

Lemma 3.7 *For $\alpha \in [0, \frac{\pi}{2}]$,*

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

∎

Lemma 3.8 *Let $(x,\lambda,s) \in \mathcal{N}_2(\theta)$, $(\dot{x},\dot{\lambda},\dot{s})$ and $(\ddot{x},\ddot{\lambda},\ddot{s})$ meet (12) and (13). Let $x(\alpha)$ and $s(\alpha)$ be defined by (14) and (16). Then,*

$$\begin{aligned} & \mu(1-\sin(\alpha)) - \frac{1}{n}\Big((\dot{x}^{\mathrm{T}}\dot{s})\sin^4(\alpha) + (\dot{x}^{\mathrm{T}}\dot{s})\sin^2(\alpha)\Big) \\ \leq & \mu(\alpha) = \mu(1-\sin(\alpha)) + \frac{1}{n}\Big(\left(\ddot{x}^{\mathrm{T}}\ddot{s} - \dot{x}^{\mathrm{T}}\dot{s}\right)(1-\cos(\alpha))^2 - \left(\dot{x}^{\mathrm{T}}\ddot{s} + \dot{s}^{\mathrm{T}}\ddot{x}\right)\sin(\alpha)(1-\cos(\alpha))\Big) \\ \leq & \mu(1-\sin(\alpha)) + \frac{1}{n}\Big((\ddot{x}^{\mathrm{T}}\ddot{s})\sin^4(\alpha) + (\ddot{x}^{\mathrm{T}}\ddot{s})\sin^2(\alpha)\Big). \end{aligned} \quad (32)$$

Proof: Using (14) and (16), we have

$$
\begin{aligned}
n\mu(\alpha) &= x^{\mathrm{T}}(\alpha)s(\alpha) \\
&= \Big(x^{\mathrm{T}} - \dot{x}^{\mathrm{T}}\sin(\alpha) + \ddot{x}^{\mathrm{T}}(1-\cos(\alpha))\Big)\Big(s - \dot{s}\sin(\alpha) + \ddot{s}(1-\cos(\alpha))\Big) \\
&= x^{\mathrm{T}}s - x^{\mathrm{T}}\dot{s}\sin(\alpha) + x^{\mathrm{T}}\ddot{s}(1-\cos(\alpha)) \\
&\quad - \dot{x}^{\mathrm{T}}s\sin(\alpha) + \dot{x}^{\mathrm{T}}\dot{s}\sin^2(\alpha) - \dot{x}^{\mathrm{T}}\ddot{s}\sin(\alpha)(1-\cos(\alpha)) \\
&\quad + \ddot{x}^{\mathrm{T}}s(1-\cos(\alpha)) - \ddot{x}^{\mathrm{T}}\dot{s}\sin(\alpha)(1-\cos(\alpha)) + \ddot{x}^{\mathrm{T}}\ddot{s}(1-\cos(\alpha))^2 \\
&= x^{\mathrm{T}}s - (x^{\mathrm{T}}\dot{s} + s^{\mathrm{T}}\dot{x})\sin(\alpha) + (x^{\mathrm{T}}\ddot{s} + s^{\mathrm{T}}\ddot{x})(1-\cos(\alpha)) \\
&\quad - (\dot{x}^{\mathrm{T}}\ddot{s} + \dot{s}^{\mathrm{T}}\ddot{x})\sin(\alpha)(1-\cos(\alpha)) + \dot{x}^{\mathrm{T}}\dot{s}\sin^2(\alpha) + \ddot{x}^{\mathrm{T}}\ddot{s}(1-\cos(\alpha))^2 \\
&= n\mu(1-\sin(\alpha)) - 2\dot{x}^{\mathrm{T}}\dot{s}(1-\cos(\alpha)) \qquad \text{use last rows of (12) and (13)} \\
&\quad - (\dot{x}^{\mathrm{T}}\ddot{s} + \dot{s}^{\mathrm{T}}\ddot{x})\sin(\alpha)(1-\cos(\alpha)) \\
&\quad + \dot{x}^{\mathrm{T}}\dot{s}(1-\cos^2(\alpha)) + \ddot{x}^{\mathrm{T}}\ddot{s}(1-\cos(\alpha))^2 \\
&= n\mu(1-\sin(\alpha)) + (\ddot{x}^{\mathrm{T}}\ddot{s} - \dot{x}^{\mathrm{T}}\dot{s})(1-\cos(\alpha))^2 \\
&\quad - (\dot{x}^{\mathrm{T}}\ddot{s} + \dot{s}^{\mathrm{T}}\ddot{x})\sin(\alpha)(1-\cos(\alpha)) \qquad (33) \\
&\le n\mu(1-\sin(\alpha)) + (\ddot{x}^{\mathrm{T}}\ddot{s} - \dot{x}^{\mathrm{T}}\dot{s})(1-\cos(\alpha))^2 \\
&\quad + (\dot{x}^{\mathrm{T}}\dot{s})(1-\cos(\alpha))^2 + (\ddot{x}^{\mathrm{T}}\ddot{s})\sin^2(\alpha) \qquad \text{use (20) in Lemma 3.2} \\
&\le n\mu\Big(1-\sin(\alpha)\Big) + (\ddot{x}^{\mathrm{T}}\ddot{s})\sin^4(\alpha) + (\ddot{x}^{\mathrm{T}}\ddot{s})\sin^2(\alpha) \qquad \text{use Lemma 3.7} \\
&\le n\mu\Big(1-\sin(\alpha) + c_2 n\sin^2(\alpha) + c_2 n\sin^4(\alpha)\Big). \qquad \text{use Lemma 3.6}
\end{aligned}
$$

This proves the second inequality of the lemma. Combining (33) and (21) proves the first inequality of the lemma. ∎

To keep all iterates of the algorithm inside the feasible set, we need $(x(\alpha), s(\alpha)) > 0$ in all the iterations. We will prove that this is guaranteed if $\mu(\alpha) > 0$ holds. The following corollary states the condition for $\mu(\alpha) > 0$ to hold.

Corollary 3.1 *If $\mu > 0$, then for any fixed $\theta \in (0,1)$, there is an $\bar{\alpha}$ depending on θ, such that for any $\sin(\alpha) \le \sin(\bar{\alpha})$, $\mu(\alpha) > 0$. In particular, if $\theta = 0.148$, $\sin(\bar{\alpha}) \ge 0.6286$.*

Proof: Since

$$
\begin{aligned}
\mu(\alpha) &\ge \mu\left((1-\sin(\alpha) - \frac{1}{n\mu}\Big((\dot{x}^{\mathrm{T}}\dot{s})\sin^4(\alpha) + (\dot{x}^{\mathrm{T}}\dot{s})\sin^2(\alpha)\Big)\right) \\
&\ge \mu\left((1-\sin(\alpha) - \frac{(1+\theta)}{2(1-\theta)}\Big(sin^4(\alpha) + \sin^2(\alpha)\Big)\right) := \mu r(\alpha),
\end{aligned}
$$

$\mu > 0$, and $r(\alpha)$ is a monotonic decreasing function in $[0, \frac{\pi}{2}]$ with $r(0) > 0$, $r(\frac{\pi}{2}) < 0$, there is a unique real solution $\sin(\bar{\alpha}) \in (0,1)$ of $r(\alpha) = 0$ such that for all $\sin(\alpha) < \sin(\bar{\alpha})$, $r(\alpha) > 0$, or $\mu(\alpha) > 0$. It is easy to see that if $\theta = 0.148$, $\sin(\bar{\alpha}) = 0.6286$ is the solution of $r(\alpha) = 0$. ∎

For linear programming, it is known [24] that $\mu(\alpha) \le \mu$ for $\alpha \in [0, \hat{\alpha}]$ with $\hat{\alpha} = \frac{\pi}{2}$, and the larger the α in the interval is, the smaller the $\mu(\alpha)$ will be. This claim is not true for the convex quadratic programming and it needs to be modified as follows.

Lemma 3.9 *Let $(x, \lambda, s) \in \mathcal{N}_2(\theta)$, $(\dot{x}, \dot{\lambda}, \dot{s})$ and $(\ddot{x}, \ddot{\lambda}, \ddot{s})$ meet (12) and (13). Let $x(\alpha)$ and $s(\alpha)$ be defined by (14) and (16). Then, there exists*

$$
\hat{\alpha} = \begin{cases} \frac{\pi}{2}, & \text{if } \frac{\ddot{x}^{\mathrm{T}}\ddot{s}}{n\mu} \le \frac{1}{2} \\ \sin^{-1}(g), & \text{if } \frac{\ddot{x}^{\mathrm{T}}\ddot{s}}{n\mu} > \frac{1}{2} \end{cases} \qquad (34)
$$

where

$$g = \sqrt[3]{\frac{n\mu}{2\ddot{x}^{\mathrm{T}}\ddot{s}} + \sqrt{\left(\frac{n\mu}{2\ddot{x}^{\mathrm{T}}\ddot{s}}\right)^2 + \left(\frac{1}{3}\right)^3}} + \sqrt[3]{\frac{n\mu}{2\ddot{x}^{\mathrm{T}}\ddot{s}} - \sqrt{\left(\frac{n\mu}{2\ddot{x}^{\mathrm{T}}\ddot{s}}\right)^2 + \left(\frac{1}{3}\right)^3}},$$

such that for every $\alpha \in [0, \hat{\alpha}]$, $\mu(\alpha) \le \mu$.

Proof: From the second inequality of (32), we have

$$\mu(\alpha) - \mu \le \mu \sin(\alpha) \left(-1 + \frac{\ddot{x}^{\mathrm{T}}\ddot{s}}{n\mu} \sin(\alpha) + \frac{\ddot{x}^{\mathrm{T}}\ddot{s}}{n\mu} \sin^3(\alpha) \right).$$

Clearly, if $\frac{\ddot{x}^{\mathrm{T}}\ddot{s}}{n\mu} \le \frac{1}{2}$, for any $\alpha \in [0, \frac{\pi}{2}]$, the function

$$f(\alpha) := \left(-1 + \frac{\ddot{x}^{\mathrm{T}}\ddot{s}}{n\mu} \sin(\alpha) + \frac{\ddot{x}^{\mathrm{T}}\ddot{s}}{n\mu} \sin^3(\alpha) \right) \le 0,$$

and $\mu(\alpha) \le \mu$. If $\frac{\ddot{x}^{\mathrm{T}}\ddot{s}}{n\mu} > \frac{1}{2}$, the function f has one real solution $\sin(\alpha) \in (0,1)$ (see [21]). The solution is given as

$$\sin(\hat{\alpha}) = \sqrt[3]{\frac{n\mu}{2\ddot{x}^{\mathrm{T}}\ddot{s}} + \sqrt{\left(\frac{n\mu}{2\ddot{x}^{\mathrm{T}}\ddot{s}}\right)^2 + \left(\frac{1}{3}\right)^3}} + \sqrt[3]{\frac{n\mu}{2\ddot{x}^{\mathrm{T}}\ddot{s}} - \sqrt{\left(\frac{n\mu}{2\ddot{x}^{\mathrm{T}}\ddot{s}}\right)^2 + \left(\frac{1}{3}\right)^3}}.$$

This proves the Lemma. ∎

According to Theorem 3.1, Lemmas 3.1, 3.5, 3.6, and 3.8, if α is small enough, then $(x(\alpha), s(\alpha)) > 0$, and $\mu(\alpha) < \mu$, i.e., the search along the arc defined by Theorem 3.1 will generate a strict feasible point with smaller duality gap. Since $(x, s) > 0$ holds in all iterations, reducing duality gap to zero means approaching to the solution of the convex quadratic programming. We will apply a similar idea used in [16, 24], i.e., starting with an iterate in $\mathcal{N}_2(\theta)$, searching along the approximated central path to reduce the duality gap and to keep the iterate in $\mathcal{N}_2(2\theta)$, and then making a correction to move the iterate back to $\mathcal{N}_2(\theta)$. First we will introduce the following notations.

$$a_0 = -\theta\mu,$$

$$a_1 = \theta\mu,$$

$$a_2 = 2\theta \frac{\dot{x}^{\mathrm{T}}\dot{s}}{n} = 2\theta \frac{\dot{x}^{\mathrm{T}} H \dot{x}}{n},$$

$$a_3 = \left\| \dot{x} \circ \ddot{s} + \dot{s} \circ \ddot{x} - \frac{1}{n}(\dot{x}^{\mathrm{T}}\ddot{s} + \dot{s}^{\mathrm{T}}\ddot{x})e \right\| = \sqrt{\|\dot{x} \circ \ddot{s} + \dot{s} \circ \ddot{x}\|^2 - n(\dot{x}^{\mathrm{T}}\ddot{s} + \dot{s}^{\mathrm{T}}\ddot{x})},$$

$$a_4 = \left\| \ddot{x} \circ \ddot{s} - \dot{s} \circ \dot{x} - \frac{1}{n}(\ddot{x}^{\mathrm{T}}\ddot{s} - \dot{s}^{\mathrm{T}}\dot{x})e \right\| + 2\theta \frac{\dot{x}^{\mathrm{T}}\dot{s}}{n} = \sqrt{\|\ddot{x} \circ \ddot{s} - \dot{s} \circ \dot{x}\|^2 - n(\ddot{x}^{\mathrm{T}}\ddot{s} - \dot{s}^{\mathrm{T}}\dot{x})} + 2\theta \frac{\dot{x}^{\mathrm{T}} H \dot{x}}{n},$$

and a quartic polynomial in terms of $\sin(\alpha)$

$$q(\alpha) = a_4 \sin^4(\alpha) + a_3 \sin^3(\alpha) + a_2 \sin^2(\alpha) + a_1 \sin(\alpha) + a_0 = 0. \tag{35}$$

Since $q(\alpha)$ is a monotonic increasing function of $\alpha \in [0, \frac{\pi}{2}]$, $q(0) < 0$ and $q(\frac{\pi}{2}) \ge 0$, the polynomial has exactly one positive root in $[0, \frac{\pi}{2}]$. Moreover, since (35) is a quartic equation, all the solutions are analytical and the computational cost is independent of the size of A (n and m) and is negligible [12].

Lemma 3.10 *Let* $\theta \le 0.148$ *and* $(x^k, \lambda^k, s^k) \in \mathcal{N}_2(\theta)$, $(\dot{x}, \dot{\lambda}, \dot{s})$ *and* $(\ddot{x}, \ddot{\lambda}, \ddot{s})$ *be calculated from (12) and (13). Let* $\sin(\tilde{\alpha})$ *be the only positive real solution of (35) in* $[0, 1]$. *Assume* $\sin(\alpha) \le \min\{\sin(\tilde{\alpha}), \sin(\bar{\alpha})\}$, *let* $(x(\alpha), \lambda(\alpha), s(\alpha))$ *and* $\mu(\alpha)$ *be updated as follows*

$$(x(\alpha), \lambda(\alpha), s(\alpha)) = (x^k, \lambda^k, s^k) - (\dot{x}, \dot{\lambda}, \dot{s}) \sin(\alpha) + (\ddot{x}, \ddot{\lambda}, \ddot{s})(1 - \cos(\alpha)), \tag{36}$$

$$\mu(\alpha) = \mu^k(1 - \sin(\alpha)) + \frac{1}{n}\left((\ddot{x}^{\mathrm{T}}\ddot{s} - \dot{x}^{\mathrm{T}}\dot{s})(1 - \cos(\alpha))^2 - (\dot{x}^{\mathrm{T}}\ddot{s} + \dot{s}^{\mathrm{T}}\ddot{x}) \sin(\alpha)(1 - \cos(\alpha)) \right). \tag{37}$$

Then $(x(\alpha), \lambda(\alpha), s(\alpha)) \in \mathcal{N}_2(2\theta)$.

Proof: Since $\sin(\tilde{\alpha})$ is the only positive real solution of (35) in $[0, 1]$ and $q(0) < 0$, substituting a_0, a_1, a_2, a_3 and a_4 into (35), we have, for all $\sin(\alpha) \le \sin(\tilde{\alpha})$,

$$\begin{aligned} &\left(\left\|\ddot{x} \circ \ddot{s} - \dot{s} \circ \dot{x} - \frac{1}{n}(\ddot{x}^{\mathrm{T}}\ddot{s} - \dot{s}^{\mathrm{T}}\dot{x})e\right\|\right)\sin^4(\alpha) + \left(\left\|\dot{x} \circ \ddot{s} + \dot{s} \circ \ddot{x} - \frac{1}{n}(\dot{x}^{\mathrm{T}}\ddot{s} + \dot{s}^{\mathrm{T}}\ddot{x})e\right\|\right)\sin^3(\alpha) \\ \le& -\left(2\theta\frac{\dot{x}^{\mathrm{T}}\dot{s}}{n}\right)\sin^4(\alpha) - \left(2\theta\frac{\dot{x}^{\mathrm{T}}\dot{s}}{n}\right)\sin^2(\alpha) + \theta\mu^k(1 - \sin(\alpha)). \end{aligned} \tag{38}$$

From (36) and (37), using Lemmas 3.7, 3.8 and (38), we have

$$\begin{aligned} &\left\|x(\alpha) \circ s(\alpha) - \mu(\alpha)e\right\| \\ =&\left\|(x^k \circ s^k - \mu^k e)(1 - \sin(\alpha)) + \left(\ddot{x} \circ \ddot{s} - \dot{x} \circ \dot{s} - \frac{1}{n}(\ddot{x}^{\mathrm{T}}\ddot{s} - \dot{x}^{\mathrm{T}}\dot{s})e\right)(1 - \cos(\alpha))^2 \right. \\ &\left. - \left(\dot{x} \circ \ddot{s} + \dot{s} \circ \ddot{x} - \frac{1}{n}(\dot{x}^{\mathrm{T}}\ddot{s} + \ddot{x}^{\mathrm{T}}\dot{s})e\right)\sin(\alpha)(1 - \cos(\alpha))\right\| \\ \le&(1 - \sin(\alpha))\left\|x^k \circ s^k - \mu^k e\right\| + \left\|(\ddot{x} \circ \ddot{s} - \dot{x} \circ \dot{s} - \frac{1}{n}(\ddot{x}^{\mathrm{T}}\ddot{s} - \dot{x}^{\mathrm{T}}\dot{s}))e\right\|(1 - \cos(\alpha))^2 \\ &+ \left\|(\dot{x} \circ \ddot{s} + \dot{s} \circ \ddot{x} - \frac{1}{n}(\dot{x}^{\mathrm{T}}\ddot{s} + \ddot{x}^{\mathrm{T}}\dot{s})e\right\|\sin(\alpha)(1 - \cos(\alpha)) \\ \le&\theta\mu^k(1 - \sin(\alpha)) + \left\|(\ddot{x} \circ \ddot{s} - \dot{x} \circ \dot{s} - \frac{1}{n}(\ddot{x}^{\mathrm{T}}\ddot{s} - \dot{x}^{\mathrm{T}}\dot{s}))e\right\|\sin^4(\alpha) + a_3\sin^3(\alpha) \\ \le&2\theta\mu^k(1 - \sin(\alpha)) - \left(2\theta\frac{\dot{x}^{\mathrm{T}}\dot{s}}{n}\right)(\sin^4(\alpha) + \sin^2(\alpha)) \\ \le&2\theta\mu(\alpha). \end{aligned} \tag{39}$$

Hence, the point $(x(\alpha), \lambda(\alpha), s(\alpha))$ satisfies the proximity condition for $\mathcal{N}_2(2\theta)$. To check the positivity condition $(x(\alpha), s(\alpha)) > 0$, note that the initial condition $(x, s) > 0$. It follows from (39) and Corollary 3.1 that, for $\theta \le 0.148$ and $\sin(\alpha) \le \sin(\bar{\alpha})$,

$$x_i(\alpha)s_i(\alpha) \ge (1 - 2\theta)\mu(\alpha) > 0. \tag{40}$$

Therefore, we cannot have $x_i(\alpha) = 0$ or $s_i(\alpha) = 0$ for any index i when $\alpha \in [0, \sin^{-1}(\bar{\alpha})]$. This proves $(x(\alpha), s(\alpha)) > 0$. ∎

Remark 3.1 *It is worthwhile to note that* $\sin(\tilde{\alpha})$ *is selected for the proximity condition (39) to hold, and* $\sin(\bar{\alpha})$ *is selected for* $\mu(\alpha) > 0$*, thereby assuring the positivity condition (40) to hold.*

$\sin(\bar{\alpha})$ is estimated in Corollary 3.1. To estimate the lower bound of $\sin(\tilde{\alpha})$, we need the following lemma.

Lemma 3.11 *Let* $(x, \lambda, s) \in \mathcal{N}_2(\theta)$, $(\dot{x}, \dot{\lambda}, \dot{s})$ *and* $(\ddot{x}, \ddot{\lambda}, \ddot{s})$ *meet (12) and (13). Then*

$$\left\|\dot{x} \circ \dot{s}\right\| \le \frac{\mu(1 + \theta)}{2(1 - \theta)}n, \tag{41}$$

$$\left\|\ddot{x} \circ \ddot{s}\right\| \le \frac{(1 + \theta)^2\mu}{2(1 - \theta)^3}n^2, \tag{42}$$

$$\left\|\ddot{x} \circ \dot{s}\right\| \le \frac{(1 + \theta)^{\frac{3}{2}}\mu}{(1 - \theta)^2}n^{\frac{3}{2}}, \tag{43}$$

$$\left\|\dot{x} \circ \ddot{s}\right\| \le \frac{(1 + \theta)^{\frac{3}{2}}\mu}{(1 - \theta)^2}n^{\frac{3}{2}}. \tag{44}$$

Proof: Since

$$\left\|\frac{\dot{x}}{x}\right\|^2 = \sum_{i=1}^{n}\left(\frac{\dot{x}_i}{x_i}\right)^2, \quad \left\|\frac{\dot{s}}{s}\right\|^2 = \sum_{i=1}^{n}\left(\frac{\dot{s}_i}{s_i}\right)^2,$$

From Lemma 3.5, we have

$$\begin{aligned}
& \left(\frac{n}{2(1-\theta)}\right)^2 \\
\geq \; & \left\|\frac{\dot{x}}{x}\right\|^2 \left\|\frac{\dot{s}}{s}\right\|^2 = \left(\sum_{i=1}^{n}\left(\frac{\dot{x}_i}{x_i}\right)^2\right)\left(\sum_{i=1}^{n}\left(\frac{\dot{s}_i}{s_i}\right)^2\right) \\
\geq \; & \sum_{i=1}^{n}\left(\frac{\dot{x}_i}{x_i}\frac{\dot{s}_i}{s_i}\right)^2 = \left\|\frac{\dot{x}}{x}\circ\frac{\dot{s}}{s}\right\|^2 \\
\geq \; & \sum_{i=1}^{n}\left(\frac{\dot{x}_i\dot{s}_i}{(1+\theta)\mu}\right)^2 = \frac{1}{(1+\theta)^2\mu^2}\left\|\dot{x}\circ\dot{s}\right\|^2.
\end{aligned}$$

This proves (41). Using

$$\left\|\frac{\ddot{x}}{x}\right\|^2 = \sum_{i=1}^{n}\left(\frac{\ddot{x}_i}{x_i}\right)^2, \quad \left\|\frac{\ddot{s}}{s}\right\|^2 = \sum_{i=1}^{n}\left(\frac{\ddot{s}_i}{s_i}\right)^2,$$

and Lemma 3.6, then following the same procedure, it is easy to verify (42). From (24) and (29), we have

$$\begin{aligned}
& \left(\frac{n}{(1-\theta)}\right)\left(\frac{(1+\theta)n^2}{(1-\theta)^3}\right) \\
\geq \; & \left\|\frac{\ddot{x}}{x}\right\|^2\left\|\frac{\dot{s}}{s}\right\|^2 + \left\|\frac{\dot{x}}{x}\right\|^2\left\|\frac{\ddot{s}}{s}\right\|^2 \\
= \; & \left(\sum_{i=1}^{n}\left(\frac{\ddot{x}_i}{x_i}\right)^2\right)\left(\sum_{i=1}^{n}\left(\frac{\dot{s}_i}{s_i}\right)^2\right) + \left(\sum_{i=1}^{n}\left(\frac{\dot{x}_i}{x_i}\right)^2\right)\left(\sum_{i=1}^{n}\left(\frac{\ddot{s}_i}{s_i}\right)^2\right) \\
\geq \; & \sum_{i=1}^{n}\left(\frac{\ddot{x}_i\dot{s}_i}{x_is_i}\right)^2 + \sum_{i=1}^{n}\left(\frac{\dot{x}_i\ddot{s}_i}{x_is_i}\right)^2 \\
\geq \; & \sum_{i=1}^{n}\left(\frac{\ddot{x}_i\dot{s}_i}{(1+\theta)\mu}\right)^2 + \sum_{i=1}^{n}\left(\frac{\dot{x}_i\ddot{s}_i}{(1+\theta)\mu}\right)^2 \\
= \; & \frac{1}{(1+\theta)^2\mu^2}\left(\left\|\ddot{x}\circ\dot{s}\right\|^2 + \left\|\dot{x}\circ\ddot{s}\right\|^2\right).
\end{aligned} \tag{45}$$

This proves the lemma. ∎

Lemma 3.12 *Let $\theta \leq 0.148$. Then $\sin(\tilde{\alpha}) \geq \frac{2\theta}{\sqrt{n}}$ for $n \geq 2$.*

Proof: First notice that $q(\sin(\alpha))$ is a monotonic increasing function of $\sin(\alpha)$ for $\alpha \in [0, \frac{\pi}{2}]$ and $q(\sin(0)) < 0$, therefore, we need only to show that $q(\frac{2\theta}{\sqrt{n}}) < 0$ for $\theta \leq 0.148$ and $n \geq 2$. Using the fact that

$$\begin{aligned}
\left\|\dot{x}\circ\ddot{s} + \dot{s}\circ\ddot{x} - \frac{1}{n}(\dot{x}^{\mathrm{T}}\ddot{s} + \dot{s}^{\mathrm{T}}\ddot{x})\right\| &\leq \left\|\dot{x}\circ\ddot{s}\right\| + \left\|\dot{s}\circ\ddot{x}\right\|, \\
\left\|\ddot{x}\circ\ddot{s} - \dot{s}\circ\dot{x} - \frac{1}{n}(\ddot{x}^{\mathrm{T}}\ddot{s} - \dot{s}^{\mathrm{T}}\dot{x})\right\| &\leq \left\|\ddot{x}\circ\ddot{s}\right\| + \left\|\dot{s}\circ\dot{x}\right\|,
\end{aligned}$$

and Lemmas 3.11, 3.5, and 3.6, we have, for $\alpha \in [0, \frac{\pi}{2}]$,

$$
\begin{aligned}
q(\sin(\alpha)) \leq & \left(\left\|\ddot{x} \circ \ddot{s}\right\| + \left\|\dot{s} \circ \dot{x}\right\| + 2\theta \frac{\dot{x}^{\mathrm{T}}\dot{s}}{n}\right) \sin^4(\alpha) + \left(\left\|\dot{x} \circ \ddot{s}\right\| + \left\|\dot{s} \circ \ddot{x}\right\|\right) \sin^3(\alpha) \\
& + 2\theta \frac{\dot{x}^{\mathrm{T}}\dot{s}}{n} \sin^2(\alpha) + \theta\mu \sin(\alpha) - \theta\mu \\
\leq & \mu\Big(\left(\frac{(1+\theta)^2}{2(1-\theta)^3} n^2 + \frac{n(1+\theta)}{2(1-\theta)} + \frac{\theta(1+\theta)}{(1-\theta)} \right) \sin^4(\alpha) + 2\frac{(1+\theta)^{\frac{3}{2}}}{(1-\theta)^2} n^{\frac{3}{2}} \sin^3(\alpha) \\
& + \frac{\theta(1+\theta)}{(1-\theta)} \sin^2(\alpha) + \theta \sin(\alpha) - \theta \Big).
\end{aligned}
$$

Substituting $\sin(\alpha) = \frac{2\theta}{\sqrt{n}}$ gives

$$
\begin{aligned}
q\Big(\frac{2\theta}{\sqrt{n}}\Big) \leq & \mu\Big(\left(\frac{(1+\theta)^2}{2(1-\theta)^3} n^2 + \frac{n(1+\theta)}{2(1-\theta)} + \frac{\theta(1+\theta)}{(1-\theta)} \right) \frac{16\theta^4}{n^2} + 2\frac{(1+\theta)^{\frac{3}{2}} n^{\frac{3}{2}}}{(1-\theta)^2} \frac{8\theta^3}{n^{\frac{3}{2}}} \\
& + \frac{\theta(1+\theta)}{(1-\theta)} \frac{4\theta^2}{n} + \theta \frac{2\theta}{\sqrt{n}} - \theta \Big) \\
= & \theta\mu\Big(\frac{8\theta^3(1+\theta)^2}{(1-\theta)^3} + \frac{8\theta^3(1+\theta)}{n(1-\theta)} + \frac{16\theta^4(1+\theta)}{(1-\theta)n^2} \\
& + \frac{16\theta^2(1+\theta)^{\frac{3}{2}}}{(1-\theta)^2} + \frac{4\theta^2(1+\theta)}{n(1-\theta)} + \frac{2\theta}{\sqrt{n}} - 1 \Big) := \theta\mu p(\theta).
\end{aligned} \tag{46}
$$

Since $p(\theta)$ is monotonic increasing function of θ, $p(0) < 0$, $n \geq 2$, and it is easy to verify that $p(0.148) < 0$ for $n = 2$, this proves the lemma. ∎

To move the iterate back to $\mathcal{N}_2(\theta)$, we use the direction defined by

$$
\begin{bmatrix} A & 0 & 0 \\ -H & A^{\mathrm{T}} & I \\ S(\alpha) & 0 & X(\alpha) \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \mu(\alpha) e - x(\alpha) \circ s(\alpha) \end{bmatrix}, \tag{47}
$$

and we update $(x^{k+1}, \lambda^{k+1}, s^{k+1})$ and μ^{k+1} by

$$
(x^{k+1}, \lambda^{k+1}, s^{k+1}) = (x(\alpha), \lambda(\alpha), s(\alpha)) + (\Delta x, \Delta \lambda, \Delta s), \tag{48}
$$

$$
\mu^{k+1} = \frac{{x^{k+1}}^{\mathrm{T}} s^{k+1}}{n}. \tag{49}
$$

Now, we show that the correction step brings the iterate from $\mathcal{N}_2(2\theta)$ back to $\mathcal{N}_2(\theta)$. The next technical lemma is from [23, page 88].

Lemma 3.13 *Let u and v are vectors of the same dimension, and $u^{\mathrm{T}} v \geq 0$. Then*

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

Lemma 3.14 *Let $(x(\alpha), \lambda(\alpha), s(\alpha)) \in \mathcal{N}_2(2\theta)$ and $(\Delta x, \Delta \lambda, \Delta s)$ be defined as in (47). Let $(x^{k+1}, \lambda^{k+1}, s^{k+1})$ be updated by using (48). Then, for $\theta \leq 0.148$ and $\sin(\alpha) \leq \sin(\bar{\alpha})$, $(x^{k+1}, \lambda^{k+1}, s^{k+1}) \in \mathcal{N}_2(\theta)$.*

Proof: First, it is easy to see

$$
0 \leq \left\| \Delta x \circ \Delta s - \frac{1}{n} (\Delta x^{\mathrm{T}} \Delta s) e \right\|^2 = \|\Delta x \circ \Delta s\|^2 - n \left(\frac{1}{n} \Delta x^{\mathrm{T}} \Delta s \right)^2 \leq \|\Delta x \circ \Delta s\|^2. \tag{50}
$$

Let $D = X^{\frac{1}{2}}(\alpha) S^{-\frac{1}{2}}(\alpha)$. Pre-multiplying $\Big(X(\alpha)S(\alpha)\Big)^{-\frac{1}{2}}$ in the last row of (47) yields

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

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

$$\begin{aligned}
\left\|\Delta x \circ \Delta s\right\| &= \left\|u \circ v\right\| \le 2^{-\frac{3}{2}} \left\| \left(X(\alpha)S(\alpha)\right)^{-\frac{1}{2}} \left(\mu(\alpha)e - X(\alpha)S(\alpha)e\right) \right\|^2 \\
&= 2^{-\frac{3}{2}} \sum_{i=1}^{n} \frac{(\mu(\alpha) - x_i(\alpha)s_i(\alpha))^2}{x_i(\alpha)s_i(\alpha)} \\
&\le 2^{-\frac{3}{2}} \frac{\|\mu(\alpha)e - x(\alpha)\circ s(\alpha)\|^2}{\min_i x_i(\alpha)s_i(\alpha)} \\
&\le 2^{-\frac{3}{2}} \frac{(2\theta)^2\mu(\alpha)^2}{(1-2\theta)\mu(\alpha)} = 2^{\frac{1}{2}} \frac{\theta^2\mu(\alpha)}{(1-2\theta)}. \qquad (51)
\end{aligned}$$

Define $(x^{k+1}(t), s^{k+1}(t)) = (x(\alpha), s(\alpha)) + t(\Delta x, \Delta s)$. Using the last row of (47), we have

$$\mu^{k+1}(t) = \frac{\left(x(\alpha) + t\Delta x\right)^{\mathrm{T}} \left(s(\alpha) + t\Delta s\right)}{n} = \frac{x(\alpha)^{\mathrm{T}} s(\alpha) + t^2 \Delta x^{\mathrm{T}} \Delta s}{n} = \mu(\alpha) + t^2 \frac{\Delta x^{\mathrm{T}} \Delta s}{n}.$$

Using this relation and (50), (51), and Lemma 3.16, we have

$$\begin{aligned}
& \left\|x^{k+1}(t) \circ s^{k+1}(t) - \mu^{k+1}(t)e\right\| \\
= & \left\|(x(\alpha) + t\Delta x) \circ (s(\alpha) + t\Delta s) - \mu(\alpha)e - \frac{t^2}{n}\Delta x^{\mathrm{T}}\Delta s e\right\| \\
= & \left\|x(\alpha)\circ s(\alpha) + t(\mu(\alpha)e - x(\alpha)\circ s(\alpha)) + t^2 \Delta x \circ \Delta s - \mu(\alpha)e - \frac{t^2}{n}\Delta x^{\mathrm{T}}\Delta s e\right\| \\
= & \left\|(1-t)\left(x(\alpha)\circ s(\alpha) - \mu(\alpha)e\right) + t^2 \left(\Delta x \circ \Delta s - \frac{1}{n}\Delta x^{\mathrm{T}}\Delta s e\right)\right\| \\
\le & (1-t)(2\theta)\mu(\alpha) + t^2 \frac{2^{\frac{1}{2}}\theta^2}{(1-2\theta)}\mu(\alpha) \\
\le & \left((1-t)(2\theta) + t^2 \frac{2^{\frac{1}{2}}\theta^2}{(1-2\theta)}\right)\mu^{k+1} := f(t,\theta)\mu^{k+1}. \qquad (52)
\end{aligned}$$

Therefore, taking $t = 1$ gives $\left\|x^{k+1} \circ s^{k+1} - \mu^{k+1}e\right\| \le \frac{2^{\frac{1}{2}}\theta^2}{(1-2\theta)}\mu^{k+1}$. It is easy to see that, for $\theta \le 0.29$,

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

For $\theta \le 0.148$ and $t \in [0, 1]$, noticing $0 \le f(t,\theta) \le f(t, 0.148) \le 0.296(1-t) + 0.044t^2 < 1$, and using Corollary 3.1, we have, for an additional condition $\sin(\alpha) \le \sin^{-1}(\bar{\alpha})$,

$$\begin{aligned}
x_i^{k+1}(t)s_i^{k+1}(t) &\ge (1 - f(t,\theta))\,\mu^{k+1}(t) \\
&= (1 - f(t,\theta))\left(\mu(\alpha) + \frac{t^2}{n}\Delta x^{\mathrm{T}}\Delta s\right) \\
&\ge (1 - f(t,\theta))\,\mu(\alpha) \\
&> 0, \qquad (53)
\end{aligned}$$

Therefore, $(x^{k+1}(t), s^{k+1}(t)) > 0$ for $t \in [0, 1]$, i.e., $(x^{k+1}, s^{k+1}) > 0$. This finishes the proof. ∎

Next, we show that the combined step (searching along the arc in $\mathcal{N}_2(2\theta)$ and moving back to $\mathcal{N}_2(\theta)$) will reduce the duality gap of the iterate, i.e., $\mu^{k+1} < \mu^k$, if we select some appropriate θ and α.

Lemma 3.15 *Let* $(x(\alpha), \lambda(\alpha), s(\alpha)) \in \mathcal{N}_2(2\theta)$ *and* $(\Delta x, \Delta\lambda, \Delta s)$ *be defined as in (47). Then*

$$0 \le \frac{\Delta x^{\mathrm{T}} \Delta s}{n} \le \frac{2\theta^2(1+2\theta)}{n(1-2\theta)^2}\mu(\alpha) := \frac{c_0}{n}\mu(\alpha). \tag{54}$$

Proof: First, pre-multiply Δx^{T} in the second row of (47) and applying the first row of (47) gives $0 \le \Delta x^{\mathrm{T}} H \Delta x = \Delta x^{\mathrm{T}} \Delta s$, which is the first inequality of (54). From the third row of (47), we have

$$S(\alpha)\Delta x + X(\alpha)\Delta s = \mu(\alpha)e - X(\alpha)S(\alpha)e.$$

Multiplying both sides by $X^{-\frac{1}{2}}(\alpha)S^{-\frac{1}{2}}(\alpha)$ gives

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

Using Lemma 3.4 and the assumption of $(x(\alpha), \lambda(\alpha), s(\alpha)) \in \mathcal{N}_2(2\theta)$, we have

$$\begin{aligned} &\sum_{i=1}^{n}\left(\frac{(\Delta x_i)^2 s_i(\alpha)}{x_i(\alpha)} + \frac{(\Delta s_i)^2 x_i(\alpha)}{s_i(\alpha)}\right) \\ \le &\sum_{i=1}^{n}\frac{(\mu(\alpha) - x_i(\alpha)s_i(\alpha))^2}{x_i(\alpha)s_i(\alpha)} \\ \le &\frac{\sum_{i=1}^{n}(\mu(\alpha) - x_i(\alpha)s_i(\alpha))^2}{\min_i x_i(\alpha)s_i(\alpha)} \\ \le &\frac{(2\theta)^2\mu^2(\alpha)}{(1-2\theta)\mu(\alpha)} = \frac{(2\theta)^2\mu(\alpha)}{(1-2\theta)}. \end{aligned} \tag{55}$$

Dividing both sides by $\mu(\alpha)$ and using $x_i(\alpha)s_i(\alpha) \ge \mu(\alpha)(1-2\theta)$ yields

$$\begin{aligned} &\sum_{i=1}^{n}(1-2\theta)\left(\frac{(\Delta x_i)^2}{x_i^2(\alpha)} + \frac{(\Delta s_i)^2}{s_i^2(\alpha)}\right) \\ = &(1-2\theta)\left(\left\|\frac{\Delta x}{x(\alpha)}\right\|^2 + \left\|\frac{\Delta s}{s(\alpha)}\right\|^2\right) \\ \le &\frac{(2\theta)^2}{(1-2\theta)}, \end{aligned} \tag{56}$$

i.e.,

$$\left\|\frac{\Delta x}{x(\alpha)}\right\|^2 + \left\|\frac{\Delta s}{s(\alpha)}\right\|^2 \le \left(\frac{2\theta}{1-2\theta}\right)^2. \tag{57}$$

Invoking Lemma 3.3, we have

$$\left\|\frac{\Delta x}{x(\alpha)}\right\|^2 \cdot \left\|\frac{\Delta s}{s(\alpha)}\right\|^2 \le \frac{1}{4}\left(\frac{2\theta}{1-2\theta}\right)^4. \tag{58}$$

This gives

$$\left\|\frac{\Delta x}{x(\alpha)}\right\| \cdot \left\|\frac{\Delta s}{s(\alpha)}\right\| \le \frac{2\theta^2}{(1-2\theta)^2}. \tag{59}$$

Using Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&\frac{(\Delta x)^{\mathrm{T}}(\Delta s)}{\mu(\alpha)} \\
\le & \sum_{i=1}^{n} \frac{|\Delta x_i||\Delta s_i|}{\mu(\alpha)} \\
\le & (1+2\theta) \sum_{i=1}^{n} \frac{|\Delta x_i|}{x_i(\alpha)} \frac{|\Delta s_i|}{s_i(\alpha)} \\
= & (1+2\theta) \left|\frac{\Delta x}{x(\alpha)}\right|^{\mathrm{T}} \left|\frac{\Delta s}{s(\alpha)}\right| \\
\le & (1+2\theta) \left\|\frac{\Delta x}{x(\alpha)}\right\| \cdot \left\|\frac{\Delta s}{s(\alpha)}\right\| \\
\le & \frac{2\theta^2(1+2\theta)}{(1-2\theta)^2}.
\end{aligned} \tag{60}$$

Therefore,

$$\frac{(\Delta x)^{\mathrm{T}}(\Delta s)}{n} \le \frac{2\theta^2(1+2\theta)}{n(1-2\theta)^2}\mu(\alpha). \tag{61}$$

This proves the lemma. ∎

Lemma 3.16 *Let $(x(\alpha), \lambda(\alpha), s(\alpha)) \in \mathcal{N}_2(2\theta)$ and $(\Delta x, \Delta\lambda, \Delta s)$ be defined as in (47). Let $(x^{k+1}, \lambda^{k+1}, s^{k+1})$ be defined as in (48). Then*

$$\mu(\alpha) \le \mu^{k+1} := \frac{{x^{k+1}}^{\mathrm{T}} s^{k+1}}{n} \le \mu(\alpha)\left(1 + \frac{2\theta^2(1+2\theta)}{n(1-2\theta)^2}\right) = \mu(\alpha)\left(1+\frac{c_0}{n}\right)$$

Proof: Using the third row of (47), we have $\frac{x(\alpha)^{\mathrm{T}}\Delta s + s(\alpha)^{\mathrm{T}}\Delta x}{n} = 0$. From Lemma 3.15, it is therefore straightforward to obtain

$$\mu(\alpha) \le \frac{x(\alpha)^{\mathrm{T}} s(\alpha)}{n} + \frac{1}{n}\Delta x^{\mathrm{T}}\Delta s = \frac{(x(\alpha)+\Delta x)^{\mathrm{T}}(s(\alpha)+\Delta s)}{n} = \mu^{k+1} \le \mu(\alpha) + \frac{2\theta^2(1+2\theta)}{n(1-2\theta)^2}\mu(\alpha).$$

This proves the lemma. ∎

For linear programming, it is known [24] that $\mu^{k+1} = \mu(\alpha)$. This claim is not always true for the convex quadratic programming as is pointed out in Lemma 3.16. Therefore, some extra work is needed to make sure that the duality gap will be reduced in every iteration.

Lemma 3.17 *For $\theta \le 0.148$, if*

$$\sin(\alpha) = \frac{\theta}{\sqrt{n}}, \tag{62}$$

then $\mu^{k+1} < \mu^k$. Moreover, for $\sin(\alpha) = \frac{\theta}{\sqrt{n}}$,

$$\mu^{k+1} \le \mu^k \left(1 - \frac{0.148\theta}{\sqrt{n}}\right). \tag{63}$$

Proof: From Lemmas 3.16, 3.8, 3.7, 3.2, 3.5, and 3.6, we have

$$\mu^{k+1} \leq \mu(\alpha)\left(1+\frac{2\theta^2(1+2\theta)}{n(1-2\theta)^2}\right) \tag{64a}$$

$$\leq \mu^k\left(1-\sin(\alpha)+\left(\frac{\ddot{x}^{\mathrm{T}}\ddot{s}}{n\mu}-\frac{\dot{x}^{\mathrm{T}}\dot{s}}{n\mu}\right)\sin^4(\alpha)-\left(\frac{\dot{x}^{\mathrm{T}}\ddot{s}}{n\mu}+\frac{\dot{s}^{\mathrm{T}}\ddot{x}}{n\mu}\right)\sin^3(\alpha)\right)\left(1+\frac{c_0}{n}\right)$$

$$\leq \mu^k\left(1-\sin(\alpha)+\frac{n(1+\theta)^2}{2(1-\theta)^3}\sin^4(\alpha)+\frac{2n^{\frac{1}{2}}(1+\theta)^{\frac{3}{2}}}{(1-\theta)^2}\sin^3(\alpha)\right)\left(1+\frac{c_0}{n}\right) \tag{64b}$$

$$= \mu^k\left(1+\frac{c_0}{n}-\left(1+\frac{c_0}{n}\right)\sin(\alpha)+c_2 n\left(1+\frac{c_0}{n}\right)\sin^4(\alpha)+2c_3 n^{\frac{1}{2}}\left(1+\frac{c_0}{n}\right)\sin^3(\alpha)\right). \tag{64c}$$

Substituting $\sin(\alpha)=\frac{\theta}{\sqrt{n}}$ into (64c) gives

$$\begin{aligned}
\mu^{k+1} \leq & \mu^k\Big\{1+\frac{2\theta^2(1+2\theta)}{n(1-2\theta)^2}-\frac{\theta}{\sqrt{n}}-\frac{2\theta^3(1+2\theta)}{n^{\frac{3}{2}}(1-2\theta)^2}+\frac{\theta^4(1+\theta)^2}{2n(1-\theta)^3} \\
& +\frac{\theta^6(1+\theta)^2(1+2\theta)}{n^2(1-\theta)^3(1-2\theta)^2}+\frac{2\theta^3(1+\theta)^{\frac{3}{2}}}{n(1-\theta)^2}+\frac{4\theta^5(1+\theta)^{\frac{3}{2}}(1+2\theta)}{n^2(1-2\theta)^2(1-\theta)^2}\Big\} \\
= & \mu^k\Big\{1-\theta\Big[\frac{1}{\sqrt{n}}-\frac{1}{n}\left(\frac{2\theta(1+2\theta)}{(1-2\theta)^2}+\frac{\theta^3(1+\theta)^2}{2(1-\theta)^3}+\frac{2\theta^2(1+\theta)^{\frac{3}{2}}}{(1-\theta)^2}\right)\Big] \\
& -\theta\Big[\frac{2\theta^2(1+2\theta)}{n^{\frac{3}{2}}(1-2\theta)^2}-\frac{1}{n^2}\left(\frac{\theta^5(1+\theta)^2(1+2\theta)}{(1-\theta)^3(1-2\theta)^2}+\frac{4\theta^4(1+\theta)^{\frac{3}{2}}(1+2\theta)}{(1-2\theta)^2(1-\theta)^2}\right)\Big]\Big\}
\end{aligned} \tag{65}$$

For $\theta \leq 0.148$,

$$\frac{2\theta^2(1+2\theta)}{(1-2\theta)^2} > \frac{\theta^5(1+\theta)^2(1+2\theta)}{(1-\theta)^3(1-2\theta)^2}+\frac{4\theta^4(1+\theta)^{\frac{3}{2}}(1+2\theta)}{(1-2\theta)^2(1-\theta)^2},$$

and

$$\frac{2\theta(1+2\theta)}{(1-2\theta)^2}+\frac{\theta^3(1+\theta)^2}{2(1-\theta)^3}+\frac{2\theta^2(1+\theta)^{\frac{3}{2}}}{(1-\theta)^2} < 0.852,$$

therefore,

$$\mu^{k+1} < \mu^k\Big\{1-\theta\left(\frac{1}{\sqrt{n}}-\frac{0.852}{\sqrt{n}}\right)\Big\} = \mu^k\left(1-\frac{0.148\theta}{\sqrt{n}}\right).$$

This proves (63). ∎

We summarize all the results in this section as the following theorem.

Theorem 3.2 *Let $\theta = 0.148$, $n \geq 2$, and $(x^k,\lambda^k,s^k) \in \mathcal{N}_2(\theta)$. Then, for $\sin(\alpha)=\frac{\theta}{\sqrt{n}}$, $(x(\alpha),\lambda(\alpha),s(\alpha)) \in \mathcal{N}_2(2\theta)$; $(x^{k+1},\lambda^{k+1},s^{k+1}) \in \mathcal{N}_2(\theta)$; and $\mu^{k+1} \leq \mu^k\left(1-\frac{0.148\theta}{\sqrt{n}}\right)$.*

Proof: From Corollary 3.1 and lemma 3.12, we have $\sin(\alpha) \leq \min\{\sin(\tilde{\alpha}), \sin(\bar{\alpha})\}$ Therefore, Lemma 3.10 holds, i.e., $(x(\alpha),\lambda(\alpha),s(\alpha)) \in \mathcal{N}_2(2\theta)$. Since $\sin(\alpha) \leq \sin(\bar{\alpha})$ and $(x(\alpha),\lambda(\alpha),s(\alpha)) \in \mathcal{N}_2(2\theta)$, Lemma 3.14 states $(x^{k+1},\lambda^{k+1},s^{k+1}) \in \mathcal{N}_2(\theta)$. Since $\sin(\alpha)=\frac{\theta}{\sqrt{n}}$, Lemma 3.17 states $\mu^{k+1} \leq \mu^k\left(1-\frac{0.148\theta}{\sqrt{n}}\right)$. This finishes the proof. ∎

We present the proposed method as the following

Algorithm 3.1 (Arc-search path-following)
Data: A, $H \geq 0$, b, c, $\theta = 0.148$, $\epsilon > 0$, initial point $(x^0,\lambda^0,s^0) \in \mathcal{N}_2(\theta)$, and $\mu^0 = \frac{{x^0}^{\mathrm{T}} s^0}{n}$.
for *iteration $k = 1, 2, \ldots$*

Step 1: Solve the linear systems of equations (12) and (13) to get $(\dot{x}, \dot{\lambda}, \dot{s})$ *and* $(\ddot{x}, \ddot{\lambda}, \ddot{s})$.

Step 2: Let $\sin(\alpha) = \frac{\theta}{\sqrt{n}}$. *Update* $(x(\alpha), \lambda(\alpha), s(\alpha))$ *and* $\mu(\alpha)$ *by (36) and (37).*

Step 3: Calculate $(\Delta x, \Delta\lambda, \Delta s)$ *by solving (47), update* $(x^{k+1}, \lambda^{k+1}, s^{k+1})$ *and* μ^{k+1} *by using (48) and (49). Set* $k+1 \to k$. *Go back to Step 1.*

end (for) ∎

4 Convergence Analysis

The first result in this section extends a result of linear programming (c.f. [23]) to convex quadratic programming.

Lemma 4.1 *Suppose Assumption 2 holds, i.e.,* $\mathcal{F}^o \neq \emptyset$. *Then for each* $K \geq 0$, *the set*

$$\{(x, s) \mid (x, \lambda, s) \in \mathcal{F}, \quad x^{\mathrm{T}} s \leq K\}$$

is bounded.

Proof: The proof is almost identical to the proof in [23]. It is given here for completeness. Let $(\bar{x}, \bar{\lambda}, \bar{s})$ be any fixed vector in $\mathcal{F}^o$, and (x, λ, s) be any vector in $\mathcal{F}$ with $x^{\mathrm{T}} s \leq K$. Then

$$A^{\mathrm{T}}(\bar{\lambda} - \lambda) + (\bar{s} - s) - H(\bar{x} - x) = 0,$$

Therefore

$$(\bar{x} - x)^{\mathrm{T}} \left(A^{\mathrm{T}}(\bar{\lambda} - \lambda) + (\bar{s} - s) - H(\bar{x} - x)\right) = 0.$$

Since

$$A(\bar{x} - x) = 0,$$

this means

$$(\bar{x} - x)^{\mathrm{T}}(\bar{s} - s) = (\bar{x} - x)^{\mathrm{T}} H(\bar{x} - x) \geq 0.$$

This leads to

$$\bar{x}^{\mathrm{T}}\bar{s} + K \geq \bar{x}^{\mathrm{T}}\bar{s} + x^{\mathrm{T}} s \geq \bar{x}^{\mathrm{T}} s + x^{\mathrm{T}}\bar{s}.$$

Sine $(\bar{x}, \bar{s}) > 0$ is fixed, let

$$\xi = \min_{i=1,\cdots,n} \quad \min\{\bar{x}_i, \bar{s}_i\}.$$

Then

$$\bar{x}^{\mathrm{T}}\bar{s} + K \geq \xi e^{\mathrm{T}}(x + s) \geq \max_{i=1,\cdots,n} \max\{\xi x_i, \xi s_i\},$$

i.e., for $i \in \{1, \cdots, n\}$,

$$0 \leq x_i \leq \frac{1}{\xi}(K + \bar{x}^{\mathrm{T}}\bar{s}), \qquad 0 \leq s_i \leq \frac{1}{\xi}(K + \bar{x}^{\mathrm{T}}\bar{s}).$$

This proves the lemma. ∎

The following theorem is a direct result of Lemmas 4.1, 3.1, Theorem 3.2, KKT conditions, Theorem A.2 in [23].

Theorem 4.1 *Suppose Assumptions 1 and 2 hold, then the sequence generated by Algorithm 3.1 converges to a set of accumulation points, and all these accumulation points are global optimal solutions of the convex quadratic programming.*

Let (x^*, λ^*, s^*) be any solution of (5), where x^* is a solution of the primary quadratic programming and (λ^*, s^*) is a solution of the dual quadratic programming, following the notation of [3], we denote index sets $\mathcal{B}$, $\mathcal{N}$, and $\mathcal{T}$ as

$$\mathcal{B} = \{j \in \{1, \dots, n\} \mid x_j^* \neq 0\}. \tag{66}$$

$$\mathcal{N} = \{j \in \{1, \dots, n\} \mid s_j^* \neq 0\}. \tag{67}$$

$$\mathcal{T} = \{j \in \{1, \dots, n\} \mid s_j^* = x_j^* = 0\}. \tag{68}$$

From Goldman-Tucker theorem [8], it can be shown (see [23]) that if $H = 0$, then $\mathcal{B} \cap \mathcal{N} = \emptyset = \mathcal{T}$ and $\mathcal{B} \cup \mathcal{N} = \{1, \dots, n\}$. A solution with this property is called strictly complementary. This property has been used in many papers to prove the locally super-linear convergence of interior-point algorithms in linear programming. However, it is pointed out in [11] that this partition does not hold for general quadratic programming problems. We will show that as long as a convex quadratic programming has strictly complementary solution(s), an interior-point algorithm will generate a sequence to approach strict complementary solution(s). As a matter of fact, from Lemma 4.1, we can extend the result of [23, Lemma 5.13] to the case of convex quadratic programming, and obtain the following lemma which is independent of any algorithm.

Lemma 4.2 *Let $\mu^0 > 0$, and $\gamma \in (0,1)$. Assume that the convex QP has strictly complementary solution(s). Then for all points (x, λ, s) with $(x, \lambda, s) \in \mathcal{F}^o$, $x_i s_i > \gamma\mu$, and $\mu < \mu^0$, there are constants M, C_1, and C_2 such that*

$$\|(x, s)\| \le M, \tag{69}$$

$$0 < x_i \le \mu/C_1 \quad (i \in \mathcal{N}), \qquad 0 < s_i \le \mu/C_1 \quad (i \in \mathcal{B}). \tag{70}$$

$$s_i \ge C_2\gamma \quad (i \in \mathcal{N}), \qquad x_i \ge C_2\gamma \quad (i \in \mathcal{B}). \tag{71}$$

Proof: The proof mimics the one in [23, Lemma 5.13]. We present it here for completeness. The first result follows immediately from Lemma 4.1 by setting $K = n\mu^0$. Let (x^*, λ^*, s^*) be any primal-dual strictly complementary solution. Since (x^*, λ^*, s^*) and (x, λ, s) are both feasible,

$$A(x - x^*) = 0, \qquad A^{\mathrm{T}}(\lambda - \lambda^*) + (s - s^*) - H(x - x^*) = 0.$$

$$(x - x^*)^{\mathrm{T}}(s - s^*) = (x - x^*)^{\mathrm{T}} H (x - x^*) \ge 0. \tag{72}$$

Since (x^*, λ^*, s^*) is strictly complementary solution, $\mathcal{T} = \emptyset$, $x_i^* = 0$ for $i \in \mathcal{N}$, and $s_i^* = 0$ for $i \in \mathcal{B}$. Since $x^{\mathrm{T}} s = n\mu$, $(x^*)^{\mathrm{T}} s^* = 0$, from (72), we have

$$n\mu \ge x^{\mathrm{T}} s^* + s^{\mathrm{T}} x^* = \sum_{i \in \mathcal{N}} x_i s_i^* + \sum_{i \in \mathcal{B}} x_i^* s_i.$$

Since each term in the summations is positive and bounded above by $n\mu$, we have for any $i \in \mathcal{N}$, $s_i^* > 0$, therefore,

$$0 < x_i \le \frac{n\mu}{s_i^*}.$$

Denote $\Omega_D = \{(\lambda^*, s^*) | s_i^* > 0\}$ and $\Omega_P = \{(x^*) | x_i^* > 0\}$, we have

$$0 < x_i \le \frac{n\mu}{\sup_{(\lambda^*, s^*) \in \Omega_D} s_i^*}.$$

This leads to

$$\max_{i \in \mathcal{N}} x_i \le \frac{n\mu}{\min_{i \in \mathcal{N}} \sup_{(\lambda^*, s^*) \in \Omega_D} s_i^*}.$$

Similarly,

$$\max_{i \in \mathcal{B}} s_i \le \frac{n\mu}{\min_{i \in \mathcal{B}} \sup_{(x^*) \in \Omega_P} x_i^*}.$$

Combining these 2 inequalities gives

$$\max\{\max_{i\in\mathcal{N}} x_i, \max_{i\in\mathcal{B}} s_i\} \leq \frac{n\mu}{\min\{\min_{i\in\mathcal{N}}\sup_{(\lambda^*,s^*)\in\Omega_D} s_i^*, \min_{i\in\mathcal{B}}\sup_{(x^*)\in\Omega_P} x_i^*\}} = \frac{\mu}{C_1}.$$

This proves (70). Finally, $x_i s_i \geq \gamma\mu$, hence for any $i \in \mathcal{N}$,

$$s_i \geq \frac{\gamma\mu}{x_i} \geq \frac{\gamma\mu}{\mu/C_1} = C_2\gamma.$$

Similarly, for any $i \in \mathcal{B}$,

$$x_i \geq \frac{\gamma\mu}{s_i} \geq \frac{\gamma\mu}{\mu/C_1} = C_2\gamma.$$

∎

Lemma 4.2 leads to the following

Theorem 4.2 *Let $(x^k, \lambda^k, s^k) \in \mathcal{N}_2(\theta)$ be generated by Algorithms 3.1. Assume that the convex QP has strictly complementary solution(s). Then every limit point of the sequence is a strictly complementary primary-dual solution of the convex quadratic programming, i.e.,*

$$s_i^* \geq C_2\gamma \quad (i \in \mathcal{N}), \qquad x_i^* \geq C_2\gamma \quad (i \in \mathcal{B}). \tag{73}$$

Proof: From Lemma 4.2, (x^k, s^k) is bounded, therefore there is at least one limit point (x^*, s^*). Since (x_i^k, s_i^k) is in the neighborhood of the central path, i.e., $x_i^k s_i^k > \gamma\mu^k = (1-3\theta)\mu^k$,

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

every limit point will meet (73) due to the fact that $C_2\gamma$ is a constant. ∎

We now show that the complexity bound of Algorithm 3.1 is $O(\sqrt{n}\log(1/\epsilon))$. We need the following theorem from [23] for this purpose.

Theorem 4.3 *Let $\epsilon \in (0,1)$ be given. Suppose that an algorithm for solving (5) generates a sequence of iterations that satisfies*

$$\mu^{k+1} \leq \left(1 - \frac{\delta}{n^\omega}\right)\mu^k, \quad k = 0, 1, 2, \ldots, \tag{74}$$

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

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

Combining Lemma 3.17 and Theorems 4.3 gives

Theorem 4.4 *The complexity of Algorithm 3.1 is bounded by $O(\sqrt{n}\log(1/\epsilon))$.*

5 Implementation Issues

Algorithm 3.1 is presented in a form that is convenient for convergence analysis. Some implementation details that make the algorithm effective and efficient are discussed in this section.

5.1 Termination Criterion

Algorithm 3.1 needs a termination criterion in real implementation. One can use

$$\mu^k \leq \epsilon, \tag{75a}$$

$$\|r_B\| = \|Ax^k - b\| \leq \epsilon, \tag{75b}$$

$$\|r_C\| = \|A^{\mathrm{T}}\lambda^k + s^k - Hx^k - c\| \leq \epsilon, \tag{75c}$$

$$\|r_t\| = \|X^k S^k e - \mu e\| \leq \epsilon, \tag{75d}$$

$$(x^k, s^k) > 0. \tag{75e}$$

An alternate Criterion is given in `linprog` [27]

$$\frac{\|r_B\|}{\max\{1, \|b\|\}} + \frac{\|r_C\|}{\max\{1, \|c\|\}} + \frac{\mu}{\max\{1, \|c^{\mathrm{T}}x\|, \|b^{\mathrm{T}}\lambda\|\}} \leq \epsilon. \tag{76}$$

5.2 Find Initial $(x^0, \lambda^0, s^0) \in \mathcal{N}_2(\theta)$

Algorithm 3.1 requires an initial point $(x^0, \lambda^0, s^0) \in \mathcal{N}_2(\theta)$. We use a modified algorithm of [4] to provide such an initial point in our simulation. Denote $X = \mathrm{diag}(x_1, \ldots, x_n)$, $S = \mathrm{diag}(s_1, \ldots, s_n)$, and

$$\mu = \frac{x^{\mathrm{T}} s}{n}. \tag{77}$$

Starting from any point (x, λ, s) with $(x, s) > 0$ that may or may not be in $\mathcal{F}^o$, moving the point to a point close to or on the central path amounts to approximately solving

$$F(x(t), \lambda(t), s(t)) = \begin{pmatrix} Ax - b \\ A^{\mathrm{T}}\lambda + s - Hx - c \\ XSe - t\mu e \end{pmatrix} = 0, \qquad (x, s) > 0. \tag{78}$$

(78) can be solved by repeatedly searching along Newton directions while keeping $(x, s) > 0$. In each step, the Newton direction $(dx, d\lambda, ds)$ can be calculated by

$$\begin{bmatrix} A & 0 & 0 \\ -H & A^{\mathrm{T}} & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} dx \\ d\lambda \\ ds \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \mu e - x \circ s \end{bmatrix}. \tag{79}$$

This process is described in the following

Algorithm 5.1 (Find Initial $(x^0, \lambda^0, s^0) \in \mathcal{N}_2(\theta)$)
Data: A, $H \geq 0$, b, c, $\epsilon > 0$, and initial point (x^0, λ^0, s^0) with $(x^0, s^0) > 0$.
for *iteration* $k = 1, 2, \ldots$

Check conditions

$$\|r_B\| = \|Ax^k - b\| \leq \epsilon, \tag{80a}$$

$$\|r_C\| = \|A^{\mathrm{T}}\lambda^k + s^k - Hx^k - c\| \leq \epsilon, \tag{80b}$$

$$\|r_t\| = \|X^k S^k e - \mu e\| \leq \theta\mu, \tag{80c}$$

$$(x^k, s^k) > 0. \tag{80d}$$

If (80) holds, (x^k, λ^k, s^k) is a point in $\mathcal{N}_2(\theta)$. Set the solution $(x^0, \lambda^0, s^0) = (x^k, \lambda^k, s^k)$ and stop.

If (80) does not hold, calculate the Newton direction $(dx^k, d\lambda^k, ds^k)$ from (79). Carry out line search along the Newton direction

$$(x^{k+1}, \lambda^{k+1}, s^{k+1}) = (x^k + \alpha dx^k, \lambda^k + \alpha d\lambda^k, s^k + \alpha ds^k) \tag{81}$$

such that the α satisfies $(x^k + \alpha dx^k, s^k + \alpha ds^k) > 0$ and

$$\begin{aligned} & \left(\|Ax^{k+1} - b\|^2 + \|A^{\mathrm{T}}\lambda^{k+1} + s^{k+1} - Hx^{k+1} - c\|^2 + \|X^{k+1}S^{k+1}e - \mu e\|^2\right) \\ < \; & \left(\|Ax^k - b\|^2 + \|A^{\mathrm{T}}\lambda^k + s^k - Hx^k - c\|^2 + \|X^k S^k e - \mu e\|^2\right) \end{aligned} \tag{82}$$

end (for) ∎

Remark 5.1 *Since $(x, s) > 0$, and the search direction is toward the central path, it is not surprising that we observe in all of our test examples that the condition $(x^k + \alpha dx^k, s^k + \alpha ds^k) > 0$ always holds. However, a rigorous analysis is needed or an alternative algorithm with rigorous analysis should be used if one wants to find initial points for general convex quadratic programming problems.*

5.3 Solving Linear Systems of Equations

In Algorithm 3.1, the majority computational operation in each iteration is to solve linear systems of equations (12), (13), and (47). Directly solving each of these linear systems of equations requires $O(2n+m)^3$ operation count. The following Theorem and its corollaries provide a more efficient way for solving these linear systems of equations.

Theorem 5.1 *Let $\hat{A} \in \mathbf{R}^{n\times(n-m)}$ be a base of the null space of A. Let $(x, \lambda, s) \in \mathcal{F}^o$ and $(\dot{x}, \dot{\lambda}, \dot{s})$ meet (12). Then,*

$$\frac{\dot{x}}{x} = X^{-1}\hat{A}(\hat{A}^{\mathrm{T}}\left(SX^{-1} + H\right)\hat{A})^{-1}\hat{A}^{\mathrm{T}}Se,$$

$$\frac{\dot{s}}{s} = e - \frac{\dot{x}}{x},$$

and

$$\dot{\lambda} = \left(AA^{\mathrm{T}}\right)^{-1} A\left(H\dot{x} - \dot{s}\right).$$

Proof: Pre-multiplying $\left(AA^{\mathrm{T}}\right)^{-1}A$ in $A^{\mathrm{T}}\dot{\lambda} + \dot{s} - H\dot{x} = 0$ gives the last equation. Since $A\dot{x} = 0$, we have $AX\frac{\dot{x}}{x} = 0$, this means that there exists a vector v such that $X\frac{\dot{x}}{x} = \hat{A}v$, i.e.,

$$\frac{\dot{x}}{x} = X^{-1}\hat{A}v. \tag{83}$$

From the last row of (12)

$$\frac{\dot{s}}{s} = e - \frac{\dot{x}}{x} = e - X^{-1}\hat{A}v. \tag{84}$$

Similarly, $A^{\mathrm{T}}\dot{\lambda} + \dot{s} - H\dot{x} = 0$ is equivalent to

$$S^{-1}A^{\mathrm{T}}\dot{\lambda} + \frac{\dot{s}}{s} - S^{-1}H\hat{A}v = 0. \tag{85}$$

Substituting (84) into (85) gives $S^{-1}A^{\mathrm{T}}\dot{\lambda} + e - X^{-1}\hat{A}v - S^{-1}H\hat{A}v = 0$, or in matrix form

$$\left[(X^{-1} + S^{-1}H)\hat{A}, -S^{-1}A^{\mathrm{T}}\right]\begin{bmatrix} v \\ \dot{\lambda} \end{bmatrix} = e. \tag{86}$$

Since H is positive semidefinite, $(SX^{-1} + H)$ is positive definite and invertible, hence $(X^{-1} + S^{-1}H) = S^{-1}(SX^{-1} + H)$ is invertible. Since $(X^{-1} + S^{-1}H)^{-1}S^{-1} = (SX^{-1} + H)^{-1}$ is positive definite, A and $\hat{A}$

are full rank matrices, we have that $\left(A(X^{-1}+S^{-1}H)^{-1}S^{-1}A^{\mathrm{T}}\right)$ and $\hat{A}^{\mathrm{T}}S(X^{-1}+S^{-1}H)\hat{A}$ are positive definite and invertible. It is easy to verify that

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

Taking the inverse in (86) gives

$$\begin{bmatrix} v \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} \left(\hat{A}^{\mathrm{T}}S(X^{-1}+S^{-1}H)\hat{A}\right)^{-1}\hat{A}^{\mathrm{T}}S \\ -\left(A(X^{-1}+S^{-1}H)^{-1}S^{-1}A^{\mathrm{T}}\right)^{-1}A(X^{-1}+S^{-1}H)^{-1} \end{bmatrix} e. \tag{87}$$

Substituting (87) into (83) proves the result. ∎

Remark 5.2 *$\hat{A}$ will be a sparse matrix if A is a sparse matrix and if sparse QR decomposition [5] is used. This feature is important for large size problems. $\left(AA^{\mathrm{T}}\right)^{-1}A$ and H are constants and independent on iterations, therefore, $\left(AA^{\mathrm{T}}\right)^{-1}A$ can be stored, and the computation of $\dot{\lambda}$ can be very efficient. The exactly same idea can be extended to solve (13) and (47).*

Corollary 5.1 *Let $\hat{A}\in\mathbf{R}^{n\times(n-m)}$ be a base of the null space of A. Let $(x,\lambda,s)\in\mathcal{F}^o$, $(\dot{x},\dot{\lambda},\dot{s})$ and $(\ddot{x},\ddot{\lambda},\ddot{s})$ meet (12) and (13). Define $f=-2\frac{\dot{x}}{x}\circ\frac{\dot{s}}{s}$. Then,*

$$\frac{\ddot{x}}{x}=X^{-1}\hat{A}\left(\hat{A}^{\mathrm{T}}\left(SX^{-1}+H\right)\hat{A}\right)^{-1}\hat{A}^{\mathrm{T}}Sf,$$

$$\frac{\ddot{s}}{s}=f-\frac{\ddot{x}}{x},$$

and

$$\ddot{\lambda}=\left(AA^{\mathrm{T}}\right)^{-1}A\left(H\ddot{x}-\ddot{s}\right).$$

∎

Corollary 5.2 *Let $\hat{A}\in\mathbf{R}^{n\times(n-m)}$ be a base of the null space of A. Let $(x,\lambda,s)\in\mathcal{F}^o$, $(\dot{x},\dot{\lambda},\dot{s})$ and $(\ddot{x},\ddot{\lambda},\ddot{s})$ meet (12) and (13). Define $g=\frac{\mu e}{x(\alpha)\circ s(\alpha)}-e$. Then,*

$$\frac{\Delta x}{x(\alpha)}=X^{-1}(\alpha)\hat{A}\left(\hat{A}^{\mathrm{T}}\left(S(\alpha)X^{-1}(\alpha)+H\right)\hat{A}\right)^{-1}\hat{A}^{\mathrm{T}}S(\alpha)g,$$

$$\frac{\Delta s}{s(\alpha)}=g-\frac{\Delta x}{x(\alpha)},$$

and,

$$\Delta\lambda=\left(AA^{\mathrm{T}}\right)^{-1}A\left(H\Delta x-\Delta s\right).$$

∎

Theorem 5.1, Corollaries 5.1 and 5.2 provide efficient formulas to calculate $(\dot{x},\dot{\lambda},\dot{s})$, $(\ddot{x},\ddot{\lambda},\ddot{s})$, and $(\Delta x,\Delta\lambda,\Delta s)$. Let $\hat{A}$ be obtained by QR decomposition $A^{\mathrm{T}}=[Q_1,\hat{A}]\begin{bmatrix} R \\ 0 \end{bmatrix}$. The computational procedure for $\dot{x}$, $\ddot{x}$, $\dot{s}$, $\ddot{s}$, $\dot{\lambda}$, and $\ddot{\lambda}$ is summarized as follows.

Algorithm 5.2 (compute $\dot{x}$, $\ddot{x}$, $\dot{s}$, $\ddot{s}$, $\dot{\lambda}$, and $\ddot{\lambda}$)

Data: Matrices A, $\hat{A}$, $H\geq 0$, X, S, and e.

Compute $\hat{A}^{\mathrm{T}}Se$.

Compute $P=\hat{A}^{\mathrm{T}}\left(SX^{-1}+H\right)\hat{A}$.

Compute $R=P^{-1}$.

Compute $\dot{x}=\hat{A}R\hat{A}^{\mathrm{T}}Se$, $\frac{\dot{x}}{x}=X^{-1}\dot{x}$, $\frac{\dot{s}}{s}=\left(e-\frac{\dot{x}}{x}\right)$, and $\dot{s}=S\frac{\dot{s}}{s}$.

Compute $f=-2\frac{\dot{x}}{x}\circ\frac{\dot{s}}{s}$.

Compute $\ddot{x}=\hat{A}R\hat{A}^{\mathrm{T}}Sf$ and $\ddot{s}=Sf-SX^{-1}\ddot{x}$.

Compute $\dot{\lambda}=\left(AA^{\mathrm{T}}\right)^{-1}A\left(H\dot{x}-\dot{s}\right)$ and $\ddot{\lambda}=\left(AA^{\mathrm{T}}\right)^{-1}A\left(H\ddot{x}-\ddot{s}\right)$. ∎

The computation counts for $(\dot{x}, \dot{\lambda}, \dot{s})$ and $(\ddot{x}, \ddot{\lambda}, \ddot{s})$ are estimated as follows.

- Computing $\hat{A}^{\mathrm{T}} Se$, $O((n-m)n)$;
- Computing $R = \left(\hat{A}^{\mathrm{T}}\left(SX^{-1}+H\right)\hat{A}\right)^{-1}$, $O((n-m)n^2+(n-m)^2 n+(n-m)^3)$;
- Computing $\dot{x} = \hat{A}R\hat{A}^{\mathrm{T}}Se$, $\frac{\dot{x}}{x} = X^{-1}\dot{x}$, $\frac{\dot{s}}{s} = \left(e - \frac{\dot{x}}{x}\right)$, and $\dot{s} = S\frac{\dot{s}}{s}$, $O(n(n-m)+n)$;
- Computing $f = -2\frac{\dot{x}}{x} \circ \frac{\dot{s}}{s}$, $O(n)$;
- Computing $\ddot{x} = \hat{A}R\hat{A}^{\mathrm{T}}Sf$ and $\ddot{s} = Sf - SX^{-1}\ddot{x}$, $O(n(n-m)+n)$.
- Computing $\dot{\lambda} = \left(AA^{\mathrm{T}}\right)^{-1} A\left(H\dot{x} - \dot{s}\right)$ and $\ddot{\lambda} = \left(AA^{\mathrm{T}}\right)^{-1} A\left(H\ddot{x} - \ddot{s}\right)$, $O(n(n+m)+n)$.

Clearly, this implementation is much cheaper than directly solving the systems of equations. A similar algorithm is used to calculate $(\Delta x, \Delta\lambda, \Delta s)$.

Algorithm 5.3 (compute $\Delta x, \Delta\lambda, \Delta s$)
Data: Matrices A, $\hat{A}$, $H \geq 0$, $X(\alpha)$, $S(\alpha)$, and e.
Compute $g = X^{-1}(\alpha)S^{-1}(\alpha)\mu e - e$.
Compute $\hat{A}^{\mathrm{T}}S(\alpha)g$.
Compute $P = \hat{A}^{\mathrm{T}}\left(S(\alpha)X^{-1}(\alpha) + H\right)\hat{A}$.
Compute $R = P^{-1}$.
Compute $\Delta x = \hat{A}R\hat{A}^{\mathrm{T}}S(\alpha)g$ and $\Delta s = S(\alpha)g - S(\alpha)X^{-1}(\alpha)\Delta x$.
Compute $\Delta\lambda = \left(AA^{\mathrm{T}}\right)^{-1} A\left(H\Delta x - \Delta s\right)$. ∎

Remark 5.3 *It may rotate the direction of $x \circ s$ in (12) so that the ellipse will better approximate the central path. Let*

$$w = x \circ s - \frac{\|(\Delta x, \Delta\lambda, \Delta s)\|\mu^2 e}{\|(x^{k+1}, \lambda^{k+1}, s^{k+1}) - (x^k, \lambda^k, s^k)\|}. \tag{88}$$

This can be done by replacing $x \circ s$ with $h = w\frac{\|x \circ s\|}{\|w\|}$, and replacing e by $S^{-1}X^{-1}h$ in Algorithm 5.2.

5.4 Duality Gap Reduction

Directly using $\sin(\alpha) = \frac{\theta}{\sqrt{n}}$ in Algorithm 3.1 provides an effective formula to prove the polynomiality. However, this choice of $\sin(\alpha)$ is too conservative in practice because the search step in $\mathcal{N}_2(2\theta)$ may be small and the duality gap may not be reduced fast enough. A better choice of $\sin(\alpha)$ should have a larger step in every iteration so that the polynomiality is reserved and fast convergence is achieved.

From analysis in Section 3, conditions that restrict step size are proximity conditions, positivity conditions, and duality reduction condition. Assuming $\theta \leq 0.148$, the proximity condition (52) holds for (x^{k+1}, s^{k+1}) without other restriction; but three more factors restrict the search step length. First, proximity condition (39) is met for $\sin(\alpha) \in [0, \sin(\tilde{\alpha})]$, where $\sin(\tilde{\alpha})$ is the smallest positive solution of (35) and it is estimated very conservatively in Lemma 3.12. However, an efficient implementation should use $\sin(\tilde{\alpha})$, the smallest positive solution of (35). Since (35) is a quartic function of $\sin(\alpha)$, the cost of finding the smallest positive solution is negligible [12]. Second, from (40) and (53), $\mu(\alpha) > 0$ is required for positivity conditions $(x(\alpha), s(\alpha)) > 0$ and $(x^{k+1}, s^{k+1}) > 0$ to hold. Since $\sin(\bar{\alpha})$ estimated in Corollary 3.1 may be a little conservative, we directly calculate the smallest positive solution of

$$\mu(\alpha) \geq \mu(1 - \sin(\alpha)) - \frac{1}{n}\left((\dot{x}^{\mathrm{T}}\dot{s})\sin^4(\alpha) + (\dot{x}^{\mathrm{T}}\dot{s})\sin^2(\alpha)\right) = \sigma, \tag{89}$$

where $\sigma > 0$ is a small number. The positivity conditions are guaranteed for all $\sin(\alpha) \in [0, \sin(\bar{\alpha})]$. Since (89) is a quartic function of $\sin(\alpha)$, the cost of finding the smallest positive solution is negligible. Third, from (64a) and Lemma 3.8, for $\mu^{k+1} \leq \mu^k$ to hold, we need

$$\frac{2\theta^2(1+2\theta)}{n(1-2\theta)^2} - \left(1 + \frac{2\theta^2(1+2\theta)}{n(1-2\theta)^2}\right)\sin(\alpha) + \left(1 + \frac{2\theta^2(1+2\theta)}{n(1-2\theta)^2}\right)\frac{\ddot{x}^{\mathrm{T}}\ddot{s}}{n\mu}\left(\sin^2(\alpha) + \sin^4(\alpha)\right) \leq 0.$$

For the sake of convergence analysis, Lemma 3.17 is used. For efficient implementation, the following solution should be adopted. Denote $p = \frac{2\theta^2(1+2\theta)}{n(1-2\theta)^2} > 0$, $q = \frac{\ddot{x}^{\mathrm{T}}\ddot{s}}{n\mu} > 0$, $z = \sin(\alpha) \in [0,1]$, and

$$F(z) = (1+p)qz^4 + (1+p)qz^2 - (1+p)z + p.$$

For $z \in [0,1]$ and $q \leq \frac{1}{6}$, $F'(z) = (1+p)(4qz^3 + 2qz - 1) \leq 0$, therefore, the upper bound of the duality gap is a monotonic decreasing function of $\sin(\alpha)$ for $\alpha \in [0, \frac{\pi}{2}]$. The larger the α is, the smaller the upper bound of the duality gap will be. For $q > \frac{1}{6}$, to minimize the upper bound of the duality gap, we can find the solution of $F'(z) = 0$. It is easy to check from discriminator [21] that the cubic polynomial $F'(z)$ has only one real solution which is given by

$$\sin(\breve{\alpha}) = \sqrt[3]{\frac{n\mu}{8\ddot{x}^{\mathrm{T}}\ddot{s}} + \sqrt{\left(\frac{n\mu}{8\ddot{x}^{\mathrm{T}}\ddot{s}}\right)^2 + \left(\frac{1}{6}\right)^3}} + \sqrt[3]{\frac{n\mu}{8\ddot{x}^{\mathrm{T}}\ddot{s}} - \sqrt{\left(\frac{n\mu}{8\ddot{x}^{\mathrm{T}}\ddot{s}}\right)^2 + \left(\frac{1}{6}\right)^3}}.$$

Since $F''(\sin(\breve{\alpha}) = (1+p)(12q\sin^2(\breve{\alpha}) + 2q) > 0$, at $\sin(\breve{\alpha}) \in [0,1)$, the upper bound of the duality gap is minimized. Therefore, we can define

$$\breve{\alpha} = \begin{cases} \frac{\pi}{2}, & \text{if } \frac{\ddot{x}^{\mathrm{T}}\ddot{s}}{n\mu} \leq \frac{1}{6} \\ \sin^{-1}\left(\sqrt[3]{\frac{n\mu}{8\ddot{x}^{\mathrm{T}}\ddot{s}} + \sqrt{\left(\frac{n\mu}{8\ddot{x}^{\mathrm{T}}\ddot{s}}\right)^2 + \left(\frac{1}{6}\right)^3}} + \sqrt[3]{\frac{n\mu}{8\ddot{x}^{\mathrm{T}}\ddot{s}} - \sqrt{\left(\frac{n\mu}{8\ddot{x}^{\mathrm{T}}\ddot{s}}\right)^2 + \left(\frac{1}{6}\right)^3}} \right), & \text{if } \frac{\ddot{x}^{\mathrm{T}}\ddot{s}}{n\mu} > \frac{1}{6}. \end{cases} \tag{90}$$

It is worthwhile to note that for $\alpha < \breve{\alpha}$, $F'(\sin(\alpha)) < 0$, i.e., $F(\sin(\alpha))$ is a monotonic decreasing function of $\alpha \in [0, \breve{\alpha}]$. As we can see from the above discussion, $\bar{\alpha}$ and $\breve{\alpha}$ are used for satisfying positivity conditions and minimizing the upper bound of the duality gap, and there is little room to improve these values. To further minimize the duality gap in each iteration, we may select the final step size $\sin(\alpha)$ as follows.

Algorithm 5.4 *(***Select step size***)*
Data: Fixed iteration number ℓ, $\sin(\breve{\alpha})$, $\sin(\bar{\alpha})$, *and* $\sin(\tilde{\alpha})$.

Step 1: If $\sin(\tilde{\alpha}) = \min\{\sin(\breve{\alpha}), \sin(\bar{\alpha}), \sin(\tilde{\alpha})\} < \min\{\sin(\breve{\alpha}), \sin(\bar{\alpha})\} = \sin(\breve{\alpha})$, *using golden section search (ℓ iterations) to get an α in the interval* $[\sin(\tilde{\alpha}), \sin(\breve{\alpha})]$ *such that*

$$\left\| x(\alpha) \circ s(\alpha) - \mu(\alpha)e \right\| \leq 2\theta\mu(\alpha).$$

Step 2: Otherwise, select $\alpha = \breve{\alpha}$.

∎

Remark 5.4 *It has been noted that* $\sin(\tilde{\alpha})$ *is significantly smaller than* $\sin(\breve{\alpha})$ *in many iterations. Hence, the proposed method of selecting the step size improves the performance of the algorithm described below. The computation for the step size is negligible compared to the computation of the linear systems of equations.*

Therefore, Algorithm 3.1 can be implemented as follows.

Algorithm 5.5 (Arc-search path-following)
Data: A, $H \geq 0$, b, c, $\theta = 0.148$, $\epsilon > \sigma > 0$.
Step 0: Find initial point $(x^0, \lambda^0, s^0) \in \mathcal{N}_2(\theta)$ using Algorithm 5.1, and $\mu^0 = \frac{{x^0}^{\mathrm{T}} s^0}{n}$.
for *iteration $k = 1, 2, \ldots$*

Step 1: If (76) holds or $\mu \leq \sigma$, stop. Otherwise continue.

Step 2: Compute $(\dot{x}, \dot{\lambda}, \dot{s})$ and $(\ddot{x}, \ddot{\lambda}, \ddot{s})$ using Algorithm 5.2.

Step 3: Find $\sin(\tilde{\alpha})$*, the smallest positive solution of the quartic polynomial of (35),* $\sin(\bar{\alpha})$*, the smallest positive solution of the quartic polynomial of (89), and* $\sin(\breve{\alpha})$ *from (90). Select* α *by Algorithm 5.4. Update* $(x(\alpha), \lambda(\alpha), s(\alpha))$ *and* $\mu(\alpha)$ *by (36) and (37).*

Step 4: Compute $(\Delta x, \Delta\lambda, \Delta s)$ *using Algorithm 5.3, update* $(x^{k+1}, \lambda^{k+1}, s^{k+1})$ *and* μ^{k+1} *by using (48) and (49). Set* $k+1 \to k$*. Go back to Step 1.*

end (for) ∎

Remark 5.5 *The condition* $\mu > \sigma$ *guarantees that the equation (89) has a positive solution before terminate criterion is met.*

6 Preliminary Numerical Test

In this section, we will first use a simple example to demonstrate how the algorithm works. Then we will test QP examples originating from [13] and compare the result with the one reported in [10].

6.1 A Simple Example

In this subsection, we use a problem in [20, page 464] to illustrate the difference between the active set method and the arc-search interior-point method developed in this paper.

Example 6.1

The problem is given as follows.

$$\min_x \quad f(x) = (x_1 - 1)^2 + (x_2 - 2.5)^2 \tag{91a}$$
$$\text{subject to} \quad x_1 - 2x_2 + 2 \geq 0, \tag{91b}$$
$$-x_1 - 2x_2 + 6 \geq 0, \tag{91c}$$
$$-x_1 + 2x_2 + 2 \geq 0, \tag{91d}$$
$$x_1 \geq 0, \quad x_2 \geq 0. \tag{91e}$$

The optimal solution is $x^* = (1.4, 1.7)$. By using the active set method, assuming that every search provides an accurate solution and the initial point is $x^0 = (2, 0)$, the active set algorithm finds the initial active set, constraints 3 and 5, then searches to the point $x^1 = (1, 0)$, then searches to the point $x^2 = (1, 1.5)$, then finds the optimal solution. The detail of the search procedure is provided in [20, pages 464-465]. The search path is depicted in the red line in Figure 1.

The problem is converted to the standard form suitable for interior point algorithms as follows.

$$\min_x \quad f(x) = \frac{1}{2}x^{\mathrm{T}}Hx + c^{\mathrm{T}}x + 7.25 \tag{92a}$$
$$\text{subject to} \quad x_1 - 2x_2 - x_3 + 2 = 0, \tag{92b}$$
$$-x_1 - 2x_2 - x_4 + 6 = 0, \tag{92c}$$
$$-x_1 + 2x_2 - x_5 + 2 = 0, \tag{92d}$$
$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0, \quad x_4 \geq 0, \quad x_5 \geq 0, \tag{92e}$$

where $H = \mathtt{diag}(2, 2, 0, 0, 0)$, $c^{\mathrm{T}} = (-2, -5, 0, 0, 0)$, $b^{\mathrm{T}} = (-2, -6, -2)$. In standard form, the optimal solution is $x^* = (1.4,\ 1.7,\ 0.0,\ 1.2,\ 4.0)$. The initial point $x^0 = (2,\ 0.01,\ 0.01,\ 0.01,\ 0.01)$ and $s^0 = (0.5,\ 100,\ 100,\ 100,\ 100)$ is used so that the initial x is an interior point and close to the initial point of the active set method. This initial point satisfies $x^0 \circ s^0 = e$. $\epsilon = 0.000001$ is used in the termination criterion. $\sigma = 0.000001\epsilon^3$ is used in (89). The central path projected in the (x_1, x_2) plane in Figures 1, 2, and 3 is a dot line in black. The approximations of the central path are ellipses that are projected to the (x_1, x_2) plane in Figures 1, 2, and 3. They are dot lines in blue. Unlike its counterpart in linear programming, the central path is not close to a straight line, but the ellipse approximation of the

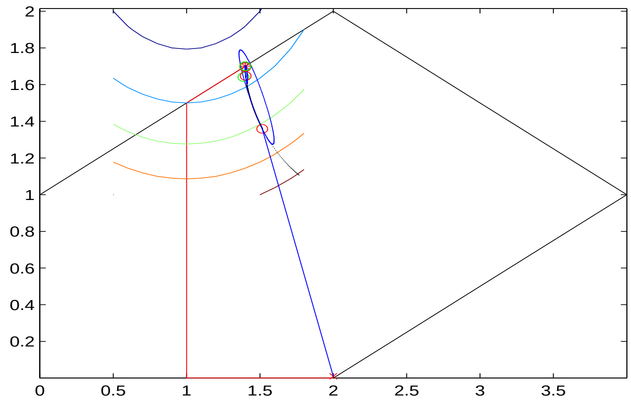

Figure 1: Arc-search for the example in [20, page 464].

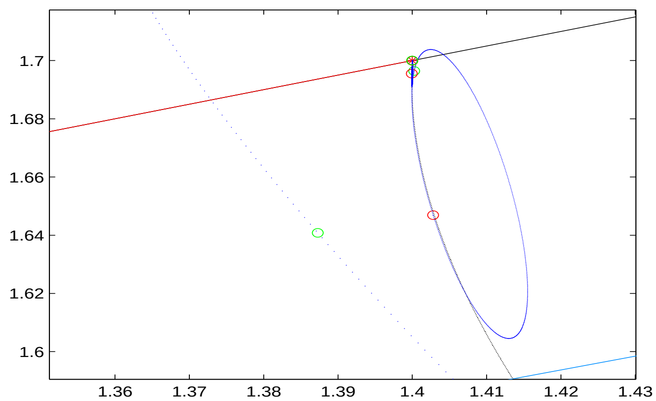

Figure 2: Arc-search approaches optimal solution for the example.

central path in quadratic programming is still very good. The corrector step is also very efficient in bringing the iterate back to the central path. Figures 2 and 3 are magnified parts of Figure 1. They provide more detailed information when the arc-search approaches the optimal solution. In the plots, the red 'x' is the initial point of the test problem; the red circle 'o' are the points obtained by either Algorithm 5.1 (move to the central path initially) or the correct steps (Step 4); the green circle 'o' are the points obtained by searching along the ellipses (Step 3); the red '*' is the optimal solution of the problem. After five (5) iterations, the algorithm finds the optimal solution of this problem. At convergence point, the slack variable is $s^* = (0,\ 0,\ 0.8,\ 0,\ 0)$, therefore, the algorithm converges to a strict complementary solution. The duality gap values in the iterations are $(\mu^0,\ \mu^1,\ \mu^2,\ \mu^3,\ \mu^4,\ \mu^5) = (1,\ 0.45454,\ 0.03578,\ 6.38692e-004,\ 3.08053e-007,\ 7.63864e-014)$. From these figures, we can see intuitively that searching along an ellipse that approximates the central path is attractive. The convergence rate appears to be super-linear.

6.2 Test on Problems in [13]

In [10], a quadratic programming software LOQO is described and numerical test were conducted for many quadratic programming problems. Among these test problems, seven originate from a widely used test set for nonlinear programming published in [13]. The proposed algorithm is implemented in MATLAB and tested for these 7 problems. The result is compared to the one obtained by LOQO [10].

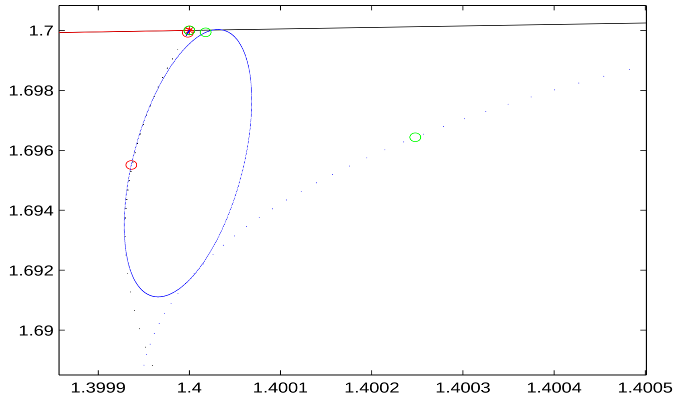

Figure 3: Arc-search approaches optimal solution for the example.

Table 1: Iteration counts for test problems in [13]

Problem	iteration numbers using arc-search	objective value obtained by arc-search	iteration numbers using LOQO	objective value obtained by LOGO
hs021	12	4.0000001e-2	13	4.0000001e-2
hs035	6	-8.8888889e+0	8	-8.8888889e+0
hs035mod	5	-8.7500000e+0	5	-8.7499999e+0
hs051	4	-6.0000000e+0	8	-6.0000000e+0
hs052	8	-6.7335244e-1	9	-6.7335244e-1
hs051	8	-1.9069767e+0	11	-1.9069767e+0
hs076	6	-4.6818182e+0	13	-4.6818182e+0

The polynomial algorithm proposed in this paper uses few iterations and converges to some equally good or better points in all problems. The comparison is listed in Table 1. For these 7 problems, LOQO uses 67 iterations while the proposed algorithm uses 49 iterations, 27% fewer in total iterations than LOQO; and the latter converges to points which have equal or higher accuracy in objective functions.

7 Conclusions

This paper proposed an arc-search interior-point path-following algorithm for convex quadratic programming that searches the optimizers along ellipses that approximate central path. The algorithm is proved to be polynomial with the complexity bound $O(\sqrt{n}\log(1/\epsilon))$. A simple example is provided to demonstrate how the algorithm works. Preliminary test on quadratic programming problems originating from [13] shows that the proposed algorithm is promising. A MATLAB M-file implementation of Algorithm 5.5 is available from the author.

References

[1] A. Altman and J. Gondzio, *Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization*, Optimization Methods and Software, 11 (1999), pp. 275–302.

[2] D. Bayer and J. Lagaris, *The nonlinear geometry of linear programming, i. affine and projective scaling trajectories*, Transactions of the American Mathematical Society, 314 (1989), pp. 499–526.

[3] A. B. Berkelaar, K. Roos, and T. Terlaky, *The optimal set and optimal partition approach to linear and quadratic programming*, in Recent Advances in Sensitivity Analysis and Parametric Programming, H. Greenberg and T. Gal, eds., Hluwer Publishers, Berlin, 1997.

[4] 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.

[5] 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.

[6] I. Dikin, *Iterative solution of problems of linear and quadratic programming*, Doklady Akademii Nauk SSSR, 174 (1967), pp. 747–748.

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

[8] 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.

[9] J. Gondzio, *Multiple centrality corrections in a primal-dual method for linear programming*, Computational Optimization and Applications, 6 (1996), pp. 137–156.

[10] I. Griva, D. F. Shanno, R. J. Vanderbei, and H. Y. Benson, *Global convergence of a primal-dual interior-point method for nonlinear programming*, Algorithmic Operations Research, 3 (2008), pp. 27–52.

[11] O. Guler and Y. Ye, *Convergence behavior of interior-point algorithms*, Mathematical Programming, 60 (1993), pp. 215–228.

[12] D. Herbison-Evans, *Solving quartics and cubics for graphics*, Technical Report TR94-487, Basser Department of Computer Science, University of Sydney, Sydney, Australia, 1994.

[13] W. Hock and K. Schittkowski, *Test Examples for Nonlinear Programming Codes, Lecture Notes in Economics and Mathematical Systems, Vol. 187*, Springer, Philadelphia, PA, 1981.

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

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

[16] 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.

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

[18] ———, *Interior path following primal-dual algorithms. part ii: convex quadratic programming*, Mathematical Programming, 44 (1989), pp. 43–66.

[19] 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.

[20] J. Nocedal and S. Wright, *Numerical Optimization*, Springer-Verlag, New York, 1999.

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

[22] R. J. VANDERBEI, *Loqo: An interior point code for quadratic programming*, Technical Report SOR-94-15, Statistics and Operations Research, Princeton University, Princeton, NJ, 1994.

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

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

[25] ——, *A polynomial arc-search interior-point algorithm for linear programming*, Optimization Online, November (2010).

[26] Y. Ye and E. Tse, *In extension of karmarkar's projective algorithm for convex quadratic programming*, Mathematical Programming, 44 (1989), pp. 157–179.

[27] 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.