

RESEARCH

Open Access



A Mizuno-Todd-Ye predictor-corrector infeasible-interior-point method for symmetric optimization with the arc-search strategy

Ximei Yang* and Yinkui Zhang

*Correspondence:

yangximeiluoyang@126.com
Henan Engineering Laboratory for
Big Data Statistical Analysis and
Optimal Control, School of
Mathematics and Information
Science, Henan Normal University,
East of Construction Road, Xinxiang,
453007, P.R. China

Abstract

In this paper, we propose a Mizuno-Todd-Ye predictor-corrector infeasible-interior-point method for symmetric optimization using the arc-search strategy. The proposed algorithm searches for optimizers along the ellipses that approximate the central path and ensures that the duality gap and the infeasibility have the same rate of decline. By analyzing, we obtain the iteration complexity $\mathcal{O}(r \log \varepsilon^{-1})$ for the Nesterov-Todd direction, where r is the rank of the associated Euclidean Jordan algebra and ε is the required precision. To our knowledge, the obtained complexity bounds coincide with the currently best known theoretical complexity bounds for infeasible symmetric optimization.

MSC: 90C05; 90C25; 90C51

Keywords: symmetric optimization; infeasible-interior-point method; arc-search; Mizuno-Todd-Ye predictor-corrector; iteration complexity

1 Introduction

The purpose of this paper is to propose a Mizuno-Todd-Ye predictor-corrector (MTY-PC) infeasible-interior-point method (infeasible-IPM) for symmetric optimization (SO) by using Euclidean Jordan algebra (EJA). Recently, SO has caused widespread concern, because it provides a unified framework for various convex optimizations including linear optimization (LO), second-order cone optimization (SOCO), and semi-definite optimization (SDO) as special cases. Meanwhile, there are many methods for solving SO. Particularly, the interior-point method (IPM), which was first proposed by Karmarkar [1], is an important kind of classification algorithm. There is extensive literature on the analysis of IPMs for SO [2–11].

Nowadays, it is broadly accepted that the primal-dual IPM is the most efficient IPM and includes the Mehrotra predictor-corrector (M-PC) algorithm [12] and the MTY-PC algorithm [13] as two typical representatives. For literature on the research of the M-PC algorithm, please see [7, 8, 14]. In the nineties of the last century, researchers began to focus on the MTY-PC algorithm [15–22], because it had the property of the best iteration complexity obtained so far for all the IPMs. Later, some researchers further studied other

aspects of the MTY-PC algorithm [23–25]. Recently, Kitahara [26] proposed a simple variant of the MTY-PC algorithm for LO, and Yang [27] extended the MTY-PC algorithm to SO. Inspired by their works, we present an $\mathcal{O}(r \log \varepsilon^{-1})$ -iteration complexity MTY-PC algorithm for SO. Moreover, the proposed algorithm will use the infeasible starting, which is found to be easy in practice. This kind of IPM is called infeasible-IPM and is studied in the literature [6, 7, 9, 22, 28–31].

Moreover, the proposed algorithm in this paper has another invention, i.e., the arc-search strategy. Yang [32–34] first developed the arc-search algorithm that searches for optimizers along an ellipse that is an approximation of the central path and gave some of the advantages of the arc-search algorithm. In order to further study the advantages of the arc-search algorithm, Yang [35, 36] proposed two infeasible-IPMs for LO and SO, and respectively obtained the $\mathcal{O}(n^{5/4} \log \varepsilon^{-1})$ -iteration complexity for LO and the $\mathcal{O}(r^{5/4} \log \varepsilon^{-1})$ and $\mathcal{O}(r^{7/4} \log \varepsilon^{-1})$ -iteration complexity, where n is the larger dimension of a standard LO, r is the rank of the associated EJA and ε is the required precision. In order to improve the iteration complexity of infeasible-IPM, we will add the arc-search strategy to the MTY-PC algorithm.

In this paper, we propose an MTY-PC infeasible-IPM for SO. The proposed algorithm uses the arc-search strategy and ensures that the duality gap and the infeasibility have the same rate of decline. By analyzing, we achieve the $\mathcal{O}(r \log \varepsilon^{-1})$ iteration complexity for the Nesterov-Todd (NT) direction. To our knowledge, this is the best iteration complexity obtained so far for an infeasible SO problem.

The outline of this paper is organized as follows. In Section 2, we briefly introduce some key results on EJA. In Section 3, we give some preliminary discussions for an algorithm and propose the algorithm. In Section 4, we establish the iteration complexity for the proposed algorithm. Finally, we close the paper by some conclusions.

2 Euclidean Jordan algebra

In order to ensure the integrity of this paper, we give some results for EJA. Most of these can be found in [5, 37].

EJA is a triple $(\mathcal{J}, \circ, \langle \cdot, \cdot \rangle)$, where $(\mathcal{J}, \langle \cdot, \cdot \rangle)$ is an n -dimensional inner product space over \mathbb{R} and $(x, y) \mapsto x \circ y : \mathcal{J} \times \mathcal{J} \mapsto \mathcal{J}$ is a bilinear mapping satisfying the following conditions:

- (a) $x \circ y = y \circ x$ for all $x, y \in \mathcal{J}$.
- (b) $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$ for all $x, y \in \mathcal{J}$, where $x^2 := x \circ x$.
- (c) $\langle x \circ y, z \rangle = \langle y, x \circ z \rangle$ for all $x, y, z \in \mathcal{J}$.

We call $x \circ y$ the Jordan product of x and y and define the inner product as $\langle x, y \rangle := \text{tr}(x \circ y)$. If there exists an element e such that $x \circ e = e \circ x = x$ for all $x \in \mathcal{J}$, then e is called the multiplicative identity element of EJA. For any $x \in \mathcal{J}$, the degree of x is denoted by $\text{deg}(x)$, which is defined as the smallest integer k such that the set $\{e, x, x^2, \dots, x^k\}$ is linearly dependent. The rank of \mathcal{J} , simply denoted by r , is the maximum of $\text{deg}(x)$ for all $x \in \mathcal{J}$. For EJA \mathcal{J} , the corresponding cone of squares $\mathcal{K} := \{x^2 : x \in \mathcal{J}\}$ is indeed a symmetric cone. A cone is symmetric if and only if it is the cone of squares of some EJA. Moreover, $\text{int } \mathcal{K}$ denotes the interior of the symmetric cone \mathcal{K} .

An idempotent c is a nonzero element of \mathcal{J} such that $c^2 = c$. An idempotent is primitive if it cannot be written as the sum of two idempotents. Two idempotents c_1 and c_2 are orthogonal if $c_1 \circ c_2 = 0$. A complete system of orthogonal idempotents is a set $\{c_1, \dots, c_k\}$ of idempotents, where $c_i \circ c_j = 0$ for all $i \neq j$, and $c_1 + \dots + c_k = e$. A complete system of orthogonal primitive idempotents is called a Jordan frame.

Theorem 2.1 (Spectral decomposition [37, Theorem III.1.2]) *Let \mathcal{J} be EJA with rank r . Then, for every $x \in \mathcal{J}$, there exist a Jordan frame $\{c_1, \dots, c_r\}$ and real numbers $\lambda_1, \dots, \lambda_r$ such that*

$$x = \lambda_1 c_1 + \lambda_2 c_2 + \dots + \lambda_r c_r = \sum_{i=1}^r \lambda_i c_i,$$

where the λ_i 's are called the eigenvalues of x .

Let $x = \sum_{i=1}^r \lambda_i c_i$ be the spectral decomposition of x . We say $x \in \mathcal{K}$ if and only if $\lambda_i \geq 0$ and $x \in \text{int } \mathcal{K}$ if and only if $\lambda_i > 0$ for all $i = 1, \dots, r$. Define the square root $x^{1/2} := \sum \lambda_i^{1/2} c_i$ for $x \in \mathcal{K}$, the inverse $x^{-1} := \sum \lambda_i^{-1} c_i, \forall \lambda_i \neq 0$ and the trace $\text{tr}(x) := \sum \lambda_i$ for $x \in \mathcal{J}$ and the determinant $\det(x) := \prod \lambda_i$ for $x \in \mathcal{J}$. We also define the spectral norm $\|x\|_2 := \max_i |\lambda_i|$ and the Frobenius norm $\|x\|_F := \sqrt{\langle x, x \rangle} = \sqrt{\sum \lambda_i^2}$.

Since 'o' is bilinear for every $x \in \mathcal{J}$, there exists a linear operator L_x such that, for every $y \in \mathcal{J}, x \circ y = L_x y$. In particular, $L_x e = x$ and $L_x x = x^2$. We say that two elements $x, y \in \mathcal{J}$ operator commute if $L_x L_y = L_y L_x$. It can be proven that x and s operator commute if and only if they share a common Jordan frame [5, Theorem 27]. For each $x, y \in \mathcal{J}$, define $Q_{x,y} := L_x L_y + L_y L_x - L_{x \circ y}, Q_x := Q_{x,x} = 2L_x^2 - L_{x^2}$ and Q_x is called the quadratic representation of x . The following is a useful proposition of quadratic representation.

Proposition 2.2 ([5, Proposition 21]) *Let $x, y, p \in \text{int } \mathcal{K}$ and define $\tilde{x} := Q_p x$ and $\tilde{y} := Q_{p^{-1}} y$, then $Q_{\tilde{x}^{1/2}} y, Q_{y^{1/2}} \tilde{x}$ and $Q_{\tilde{x}^{1/2}} \tilde{y}$ have the same spectrum.*

3 Preliminary discussions and algorithm

3.1 SO problem and ellipse approximate center

First, we give the standard form of SO and its dual form, as follows:

$$(P) \quad \min \langle c, x \rangle, \quad \text{s.t.} \quad Ax = b, \quad x \in \mathcal{K}, \tag{1}$$

$$(D) \quad \max \langle b, y \rangle, \quad \text{s.t.} \quad A^* y + s = c, \quad s \in \mathcal{K}, y \in \mathbb{R}^m, \tag{2}$$

where $c \in \mathcal{J}, b \in \mathbb{R}^m, A$ is a linear operator that maps \mathcal{J} into \mathbb{R}^m and A^* is its adjoint operator such that $\langle x, A^* y \rangle = \langle Ax, y \rangle$ for all $x \in \mathcal{J}, y \in \mathbb{R}^m$.

Moreover, we denote the sets of optimal solutions of (P) and (D) by \mathcal{P}^* and \mathcal{D}^* , and assume that A is surjective and $\mathcal{F}^0 \neq \emptyset$, where \mathcal{F}^0 indicates a primal-dual strict feasibility set that is defined by

$$\mathcal{F}^0 := \{(x, y, s) \in \text{int } \mathcal{K} \times \mathbb{R}^m \times \text{int } \mathcal{K} : Ax = b, A^* y + s = c\}.$$

The Karush-Kuhn-Tucker (KKT) conditions for (P) and (D) are given by

$$Ax = b, \quad x \in \mathcal{K}, \quad A^* y + s = c, \quad s \in \mathcal{K}, y \in \mathbb{R}^m, \quad x \circ s = 0, \tag{3}$$

where $x \circ s = 0$ is called the complementarity slackness condition.

By relaxing $x \circ s = 0$ with $x \circ s = \mu e$, we obtain

$$Ax = b, \quad x \in \mathcal{K}, \quad A^* y + s = c, \quad s \in \mathcal{K}, y \in \mathbb{R}^m, \quad x \circ s = \mu e, \tag{4}$$

where $\mu = \langle x, s \rangle / r > 0$ is called the duality gap.

System (4) has unique solutions $(x(\mu), y(\mu), s(\mu))$, the set of which is called the central path, which is denoted by

$$C = \{(x(\mu), y(\mu), s(\mu)) : \mu > 0\}. \tag{5}$$

In this paper, we will use the idea of Yang [32–34], which is that the central path C is replaced by an ellipse Ω , where Ω is defined as follows:

$$\Omega = \{(x(\theta), y(\theta), s(\theta)) : (x(\theta), y(\theta), s(\theta)) = \mathbf{a} \cos(\theta) + \mathbf{b} \sin(\theta) + \mathbf{c}\}, \tag{6}$$

where $\mathbf{a} \in \mathbb{R}^{2n+m}$ and $\mathbf{b} \in \mathbb{R}^{2n+m}$ are the axes of the ellipse perpendicular to each other, and $\mathbf{c} \in \mathbb{R}^{2n+m}$ is the center of the ellipse.

For the point $z = (x, y, s) = (x(\theta_0), y(\theta_0), s(\theta_0)) \in \Omega$, we require its first and second derivatives such that

$$A\dot{x} = r_p, \quad A^*\dot{y} + \dot{s} = r_d, \quad s \circ \dot{x} + x \circ \dot{s} = x \circ s, \tag{7}$$

$$A\ddot{x} = 0, \quad A^*\ddot{y} + \ddot{s} = 0, \quad s \circ \ddot{x} + x \circ \ddot{s} = -2\dot{x}\dot{s}, \tag{8}$$

where $r_p = Ax - b$ and $r_d = A^*y + s - c$.

Systems (7) and (8) do not always have a unique solution due to the fact that x and s do not operator commute in general. To overcome this difficulty, we apply a scaling scheme that follows from [5, Lemma 28]. For the scaling point $p \in \text{int } \mathcal{K}$, there are several appropriate choices (see [38]). In this paper, we select the classical NT-scaling point that is

$$p = [Q_{x^{1/2}}(Q_{x^{1/2}}s)^{-1/2}]^{-1/2} = [Q_{s^{-1/2}}(Q_{s^{-1/2}}x)^{1/2}]^{-1/2}, \tag{9}$$

which was first proposed by Nesterov and Todd for self-scaled cones [4] and then adapted by Faybusovich [3] for symmetric cones.

3.2 Foundation of the MTY-PC algorithm

Since the MTY-PC algorithm requires two matrix factorizations and at most three back-solves for each iteration, it is generally divided into two steps, which are the predictor step and the corrector step.

In the predictor step, using p in (9), systems (7) and (8) are rewritten as

$$\tilde{A}\dot{\tilde{x}} = \tilde{r}_p, \quad \tilde{A}^*\dot{\tilde{y}} + \dot{\tilde{s}} = \tilde{r}_d, \quad \tilde{s} \circ \dot{\tilde{x}} + \tilde{x} \circ \dot{\tilde{s}} = \tilde{x} \circ \tilde{s}, \tag{10}$$

$$\tilde{A}\ddot{\tilde{x}} = 0, \quad \tilde{A}^*\ddot{\tilde{y}} + \ddot{\tilde{s}} = 0, \quad \tilde{s} \circ \ddot{\tilde{x}} + \tilde{x} \circ \ddot{\tilde{s}} = -2\dot{\tilde{x}}\dot{\tilde{s}}, \tag{11}$$

where $\tilde{A} = AQ_{p^{-1}}$, $\tilde{c} = Q_{p^{-1}}c$, $\tilde{x} = Q_p x$, $Q_{p^{-1}}s = \tilde{s}$ and $\dot{\tilde{x}} = Q_p \dot{x}$, $\dot{\tilde{s}} = Q_{p^{-1}}\dot{s}$, $\ddot{\tilde{x}} = Q_p \ddot{x}$, $\ddot{\tilde{s}} = Q_{p^{-1}}\ddot{s}$, $\tilde{r}_p = \tilde{A}\tilde{x} - b$, $\tilde{r}_d = \tilde{A}^*\tilde{y} + \tilde{s} - \tilde{c}$.

By solving systems (10) and (11), we obtain the predictor directions $(\dot{\tilde{x}}, \dot{\tilde{y}}, \dot{\tilde{s}})$ and $(\ddot{\tilde{x}}, \ddot{\tilde{y}}, \ddot{\tilde{s}})$ and have the following lemma.

Lemma 3.1 ([34, Theorem 3.1]) *Let $(\tilde{x}(\theta), y(\theta), \tilde{s}(\theta))$ be an arc defined by (6) passing through a point $(\tilde{x}, y, \tilde{s})$, and its first and second derivatives at $(\tilde{x}, y, \tilde{s})$ be $(\dot{\tilde{x}}, \dot{\tilde{y}}, \dot{\tilde{s}})$ and $(\ddot{\tilde{x}}, \ddot{\tilde{y}}, \ddot{\tilde{s}})$,*

which are defined by (10) and (11). Then an ellipsoidal approximation of the central path is given by

$$\tilde{x}(\theta) := \tilde{x} - \sin(\theta)\dot{\tilde{x}} + (1 - \cos(\theta))\ddot{\tilde{x}}, \tag{12a}$$

$$\tilde{s}(\theta) := \tilde{s} - \sin(\theta)\dot{\tilde{s}} + (1 - \cos(\theta))\ddot{\tilde{s}}, \tag{12b}$$

$$y(\theta) := y - \sin(\theta)\dot{y} + (1 - \cos(\theta))\ddot{y}. \tag{12c}$$

Using (12a), (12b), (12c), the third equations in (10) and (11), we have

$$\tilde{x}(\theta) \circ \tilde{s}(\theta) = (1 - \sin(\theta))\tilde{x} \circ \tilde{s} - g^2(\theta)\dot{\tilde{x}} \circ \dot{\tilde{s}} + g(\theta)\sin(\theta)\xi + g^2(\theta)\ddot{\tilde{x}} \circ \ddot{\tilde{s}}, \tag{13}$$

where $g(\theta) = (1 - \cos(\theta))$, $\xi = \dot{\tilde{x}} \circ \ddot{\tilde{s}} + \ddot{\tilde{x}} \circ \dot{\tilde{s}}$.

Furthermore, using (13), we have

$$\langle \tilde{x}(\theta), \tilde{s}(\theta) \rangle = (1 - \sin(\theta))\mu r - g^2(\theta)\text{tr}(\dot{\tilde{x}} \circ \dot{\tilde{s}}) + g(\theta)\sin(\theta)\text{tr}(\xi).$$

In what follows, we discuss a method for selecting the predictor step. Firstly, we give the neighborhood that is used in this paper as follows:

$$\mathcal{N}_F(\gamma) = \{(x, y, s) \in \text{int } \mathcal{K} \times \mathbb{R}^m \times \text{int } \mathcal{K} : \|w - \mu e\|_F \leq \gamma \mu\}, \tag{14}$$

where $w = Q_{x^{1/2}}s$, $0 < \gamma < 1$.

The neighborhood $\mathcal{N}_F(\gamma)$ has some important properties, which are given in the following proposition. For more details, readers are referred to [5].

Proposition 3.2 *Let $\mathcal{N}_F(\gamma)$ be defined in (14) and $w = Q_{x^{1/2}}s$, then*

- (a) *The neighborhood $\mathcal{N}_F(\gamma)$ is scaling invariant.*
- (b) *$\|w - \mu e\|_F \leq \gamma \mu$ implies $\lambda_{\min}(w) \geq \beta \mu$, where $\beta = 1 - \gamma$.*

Now, we give the method of selecting the predictor step, which is to find the largest positive $\bar{\theta} \in (0, \pi/2]$ and to satisfy for all $\theta \in (0, \bar{\theta}]$ that

$$\|\tilde{x}(\theta) \circ \tilde{s}(\theta) - (1 - \sin(\theta))\mu e\|_F \leq 2\gamma(1 - \sin(\theta))\mu, \tag{15}$$

$$\tilde{x}(\theta) \in \text{int } \mathcal{K}, \quad \tilde{s}(\theta) \in \text{int } \mathcal{K}. \tag{16}$$

In the corrector step, we define $(\bar{x}, \bar{y}, \bar{s}) = (Q_{p^{-1}}\tilde{x}(\bar{\theta}), y(\bar{\theta}), Q_p\tilde{s}(\bar{\theta}))$ and calculate the corrector direction $(\Delta x, \Delta y, \Delta s)$ by

$$A\Delta x = 0, \quad A^*\Delta y + \Delta s = 0, \quad \bar{s} \circ \Delta x + \bar{x} \circ \Delta s = r_c, \tag{17}$$

where $r_c = (1 - \sin(\bar{\theta}))\mu e - \bar{x} \circ \bar{s}$.

Similarly, system (17) does not always have a unique solution. Thus, we need to choose an NT-scaling point p_1 such that

$$p_1 = [Q_{\bar{x}^{1/2}}(Q_{\bar{x}^{1/2}}\bar{s})^{-1/2}]^{-1/2} = [Q_{\bar{s}^{-1/2}}(Q_{\bar{s}^{1/2}}\bar{x})^{1/2}]^{-1/2}.$$

The scaling corrector direction is given by solving the following system:

$$\hat{A}\Delta\hat{x} = 0, \quad \hat{A}^*\Delta y + \Delta\hat{s} = 0, \quad \hat{s} \circ \Delta\hat{x} + \hat{x} \circ \Delta\hat{s} = \hat{r}_c, \tag{18}$$

where $\hat{x} = Q_{p_1}\bar{x}$, $\hat{s} = Q_{p_1}\bar{s}$, $\hat{A} = Q_{p_1}A$, $\hat{r}_c = (1 - \sin(\bar{\theta}))\mu e - \hat{x} \circ \hat{s}$.

Eventually, the next iteration point is updated by

$$(\hat{x}(\bar{\theta}), y(\bar{\theta}), \hat{s}(\bar{\theta})) := (\hat{x}, \hat{y}, \hat{s}) + (\Delta\hat{x}, \Delta y, \Delta\hat{s}). \tag{19}$$

In what follows, we give two useful expressions

$$\begin{aligned} \hat{x}(\bar{\theta}) \circ \hat{s}(\bar{\theta}) &= \hat{x} \circ \hat{s} + (1 - \sin(\bar{\theta}))\mu e - \hat{x} \circ \hat{s} + \Delta\hat{x} \circ \Delta\hat{s} \\ &= (1 - \sin(\bar{\theta}))\mu e + \Delta\hat{x} \circ \Delta\hat{s}, \end{aligned} \tag{20}$$

$$\mu(\bar{\theta}) = \frac{1}{r} \langle \hat{x}(\bar{\theta}), \hat{s}(\bar{\theta}) \rangle = (1 - \sin(\bar{\theta}))\mu. \tag{21}$$

3.3 Framework of the MTY-PC algorithm

Based on the previous analysis, we state the generic framework of the proposed MTY-PC algorithm in this paper.

Algorithm 1 Let $\varepsilon > 0$, $\gamma \leq 1/4$, $(x^0, y^0, s^0) \in \mathcal{N}_F(\gamma)$ and $\mu^0 = \langle x^0, s^0 \rangle / r$, $\phi^0 = 1$, $k := 0$.

- Step 1 If $x^k \in \text{int } \mathcal{K}$, $s^k \in \text{int } \mathcal{K}$ and $\phi^k \leq \varepsilon$, then stop.
- Step 2 (Predictor step) The predictor directions $(\tilde{x}^k, \tilde{y}^k, \tilde{s}^k)$, $(\ddot{x}^k, \ddot{y}^k, \ddot{s}^k)$ are obtained by solving the linear system (10), (11) and the largest positive $\bar{\theta}^k \in (0, \pi/2]$ is computed by solving (16) and (15). Set $(\bar{x}^k, \bar{y}^k, \bar{s}^k) = (Q_{p_1}\tilde{x}^{k+1}(\bar{\theta}^k), y^{k+1}(\bar{\theta}^k), Q_p\tilde{s}^{k+1}(\bar{\theta}^k))$.
- Step 3 (Corrector step) Solve corrector direction $(\Delta\hat{x}^k, \Delta y^k, \Delta\hat{s}^k)$ from (18). Let $(\hat{x}^{k+1}, y^{k+1}, \hat{s}^{k+1}) = (\hat{x}^k(\bar{\theta}^k), y^k(\bar{\theta}^k), \hat{s}^k(\bar{\theta}^k))$ and $(x^{k+1}, y^{k+1}, s^{k+1}) = (Q_{p_1}\hat{x}^{k+1}, y^{k+1}, Q_{p_1}\hat{s}^{k+1})$. Go to Step 4.
- Step 4 Compute $\mu^{k+1} = \frac{\langle \hat{x}^k(\bar{\theta}^k), \hat{s}^k(\bar{\theta}^k) \rangle}{r} = \frac{\langle x^{k+1}, s^{k+1} \rangle}{r}$ and $\phi^{k+1} = (1 - \sin(\bar{\theta}^k))\phi^k$. Set $k := k + 1$ and go to Step 1.

To analyze complexity, we give two remarks for Algorithm 1.

Remark 1 Let $\{(x^k, y^k, s^k)\}$ be generated by Algorithm 1, and $\phi^{k+1} \in [0, 1]$ such that $\phi^{k+1} = \prod_{i=0}^k (1 - \sin(\bar{\theta}^i))$. Then $r_p^{k+1} = Ax^{k+1} - b = \phi^{k+1}r_p^0$, $r_d^{k+1} = A^*y^{k+1} + s^{k+1} - c = \phi^{k+1}r_d^0$, $\mu^{k+1} = \frac{\langle x^{k+1}, s^{k+1} \rangle}{r} = \phi^{k+1}\mu^0$ for $k \geq 0$.

Proof Using (7), (8), (12a), (12b), (12c), (17), (18), by calculating directly, we have

$$\begin{aligned} r_p^{k+1} &= Ax^{k+1} - b = AQ_{p_1}\hat{x}^{k+1} - b = AQ_{p_1}\hat{x}^k(\bar{\theta}^k) - b \\ &= AQ_{p_1}(\hat{x}^k + \Delta\hat{x}^k) - b = A\hat{x}^k + A\Delta x^k - b \\ &= AQ_{p_1}\tilde{x}^{k+1}(\bar{\theta}^k) - b \\ &= AQ_{p_1}[\tilde{x}^k - \sin(\bar{\theta}^k)\tilde{x}^k + (1 - \cos(\bar{\theta}^k))\tilde{x}^k] - b \\ &= A[x^k - \sin(\bar{\theta}^k)x^k + (1 - \cos(\bar{\theta}^k))x^k] - b \end{aligned}$$

$$\begin{aligned}
 &= Ax^k - \sin(\bar{\theta}^k)A\hat{x}^k + (1 - \cos(\bar{\theta}^k))A\tilde{x}^k - b \\
 &= [1 - \sin(\bar{\theta}^k)]r_p^k = \phi^{k+1}r_p^0.
 \end{aligned}$$

In the same way, we have $r_d^{k+1} = [1 - \sin(\bar{\theta}^k)]r_p^k = \phi^{k+1}r_d^0$.

In what follows, we focus on proving the last inequality and have

$$\begin{aligned}
 \mu^{k+1} &= \frac{\langle \hat{x}^{k+1}, s^{k+1} \rangle}{r} = \frac{\langle Q_{p_1}^{-1}\hat{x}^{k+1}, Q_{p_1}\hat{s}^{k+1} \rangle}{r} = \frac{\langle \hat{x}^{k+1}, \hat{s}^{k+1} \rangle}{r} \\
 &= \frac{\langle \hat{x}^k(\bar{\theta}^k), \hat{s}^k(\bar{\theta}^k) \rangle}{r} = \frac{\langle \hat{x}^k + \Delta\hat{x}^k, \hat{s}^k + \Delta\hat{s}^k \rangle}{r} \\
 &= \frac{1}{r} [\langle \hat{x}^k \circ \hat{s}^k, e \rangle + \langle \hat{x}^k \circ \Delta\hat{s}^k + \hat{s}^k \circ \Delta\hat{x}^k, e \rangle + \langle \Delta\hat{x}^k \circ \Delta\hat{s}^k, e \rangle] \\
 &= \frac{1}{r} [\langle Q_{p_1}^{-1}\tilde{x}^k(\bar{\theta}^k) \circ Q_{p_1}\tilde{s}^k(\bar{\theta}^k), e \rangle + \langle \hat{r}_c^k, e \rangle] \\
 &= \frac{1}{r} [\langle \tilde{x}^k(\bar{\theta}^k) \circ \tilde{s}^k(\bar{\theta}^k), e \rangle + \langle (1 - \sin(\bar{\theta}^k))\mu^k e - \tilde{x}^k(\bar{\theta}^k) \circ \tilde{s}^k(\bar{\theta}^k), e \rangle] \\
 &= [1 - \sin(\bar{\theta}^k)]\mu^k = \phi^{k+1}\mu^0.
 \end{aligned}$$

This completes the proof. □

From Remark 1, we have $\phi^k = \frac{\|r_p^k\|_F}{\|r_p^0\|_F} = \frac{\|r_d^k\|_F}{\|r_d^0\|_F}$, which implies ϕ^k represents the relative infeasibility at (x^k, y^k, s^k) . Meanwhile, we also have $\phi^k = \frac{\mu^k}{\mu^0}$, which is also the rate of decline of the duality gap μ . Thus, if $\phi^k \leq \varepsilon$, then Algorithm 1 will stop and we obtain an approximate optimal solution of SO.

Remark 2 For Algorithm 1, we choose a particular starting point, which is studied by Zhang [14, 31] and Rangarajan [6]. In what follows, we give the particular starting point.

Let \check{u}^0 and $(\check{r}^0, \check{v}^0)$ be the minimum-norm solutions to the linear systems $Ax = b$ and $A^*y + s = c$, that is,

$$\check{u}^0 = \arg \min \{ \|\check{u}\|_F : A\check{u} = b \}, \quad (\check{r}^0, \check{v}^0) = \arg \min \{ \|\check{v}\|_F : A^*\check{r} + \check{v} = c \}. \tag{22}$$

Let $\rho^0 \geq \max \{ \|\check{u}^0\|_2, \|\check{v}^0\|_2 \}$ and choose (x^0, y^0, s^0) such that

$$x^0 = s^0 = \rho^0 e, \tag{23}$$

which implies that $x^0 \in \text{int } \mathcal{K}, s^0 \in \text{int } \mathcal{K}, x^0 - \check{u}^0 \in \mathcal{K}, s^0 - \check{v}^0 \in \mathcal{K}$.

Let $\rho^* = \min \{ \max \{ \|x^*\|_2, \|s^*\|_2 \} : x^* \in \mathcal{P}^*, (y^*, s^*) \in \mathcal{D}^* \}$. In addition, we assume that for some constant $\Psi > 0$, it has $\rho^0 \geq \rho^*/\Psi$ (note that we can always increase ρ^0).

4 Complexity analysis

For simplicity, we will often write $\tilde{x}, y, \tilde{s}, \bar{x}, \bar{s}, \hat{x}, \hat{s}, \bar{\theta}$ and ϕ for $\tilde{x}^k, y^k, \tilde{s}^k, \bar{x}^k, \bar{s}^k, \hat{x}^k, \hat{s}^k, \bar{\theta}^k$ and ϕ^k , respectively. Moreover, since the NT-scaling point is used in this paper, we can obtain the following special results:

$$v := Q_p x = Q_{p-1} s \iff v = \tilde{x} = \tilde{s} \implies v^2 = \tilde{x} \circ \tilde{s} = \tilde{w} = Q_{\tilde{x}^{1/2}} \tilde{s}, \tag{24}$$

$$\hat{v} := Q_{p_1} \bar{x} = Q_{p_1^{-1}} \bar{s} \Leftrightarrow \hat{v} = \hat{x} = \hat{s} \Rightarrow \hat{v}^2 = \hat{x} \circ \hat{s} = \hat{w} = Q_{\hat{x}^{1/2}} \hat{s}. \tag{25}$$

In what follows, we give some fundamental lemmas. Firstly, by the proof procedure of Lemma 4.1 and Lemma 2 in [6, 7], we have the following lemma.

Lemma 4.1 *Let $\tilde{x} \in \text{int } \mathcal{K}$, $\tilde{s} \in \text{int } \mathcal{K}$, \tilde{x}, \tilde{s} operator commute, then*

- (i) *For $\tilde{q} \in \mathcal{K}$, then $\|Q_{\tilde{x}^{1/2}} \tilde{q}\|_F \leq \langle \tilde{x}, \tilde{q} \rangle$;*
- (ii) $\lambda_{\max}((L_{\tilde{x}} L_{\tilde{s}})^{-1}) \leq 1/\lambda_{\min}(\tilde{w})$,

where $\tilde{w} = Q_{\tilde{x}^{1/2}} \tilde{s}$.

Lemma 4.2 ([6, Lemma 2.9]) *For $x, y \in \mathcal{J}$, then $\|x \circ y\|_F \leq \|x\|_F \|y\|_F$.*

Lemma 4.3 ([39, Lemma 2.15]) *If $x \circ s \in \text{int } \mathcal{K}$, then $\det(x) \neq 0$.*

Lemma 4.4 ([5, Lemma 30]) *Let $(x, s) \in \text{int } \mathcal{K} \times \text{int } \mathcal{K}$, $w = Q_{x^{1/2}} s$, then we have $\|w - \mu e\|_F \leq \|x \circ s - \mu e\|_F$, and with equality holding if x and s operator commute.*

4.1 Technical results

In order to achieve the iteration complexity bounds for the proposed Algorithm 1, we need some technical results.

Lemma 4.5 *Let $(\tilde{x}, \tilde{s}) \in \mathcal{N}_F(\gamma)$ and $(\dot{\tilde{x}}, \dot{\tilde{s}})$ be the solution of (10). Then*

$$\|(\dot{\tilde{x}}, \dot{\tilde{s}})\|_2 \leq \sqrt{\mu r} + (1 + \sqrt{2})\zeta,$$

where $\zeta := \min\{\|(\check{u}, \check{v})\|_2 : \tilde{A}\check{u} = \tilde{r}_p, \tilde{A}^*\check{r} + \check{v} = \tilde{r}_d\}$, $\|(\check{u}, \check{v})\|_2 := \sqrt{\|\check{u}\|_F^2 + \|\check{v}\|_F^2}$.

Proof Let $(\check{u}, \check{r}, \check{v}) \in \mathcal{J} \times \mathbb{R}^m \times \mathcal{J}$ satisfy equations $\tilde{A}\check{u} = \tilde{r}_p$ and $\tilde{A}^*\check{r} + \check{v} = \tilde{r}_d$. Using system (7) and $v = \tilde{x} = \tilde{s}$, we have

$$\begin{aligned} \tilde{A}(\dot{\tilde{x}} - \check{u}) &= 0, \\ \tilde{A}^*(\dot{\tilde{s}} - \check{r}) + (\dot{\tilde{s}} - \check{v}) &= 0, \\ L_v(\dot{\tilde{x}} - \check{u}) + L_v(\dot{\tilde{s}} - \check{v}) &= v^2 - (L_v \check{u} + L_v \check{v}). \end{aligned}$$

Multiplying the last equation by L_v^{-1} , we obtain

$$(\dot{\tilde{x}} - \check{u}) + (\dot{\tilde{s}} - \check{v}) = v - (\check{u} + \check{v}). \tag{26}$$

Using the definition of $\|(\check{u}, \check{v})\|_2 := \sqrt{\|\check{u}\|_F^2 + \|\check{v}\|_F^2}$, we have

$$\|\check{u}\|_F + \|\check{v}\|_F \leq \sqrt{2} \|(\check{u}, \check{v})\|_2. \tag{27}$$

Using (26), (27) and the fact $\langle \dot{\tilde{x}} - \check{u}, \dot{\tilde{s}} - \check{v} \rangle = 0$, we have

$$\begin{aligned} \|(\dot{\tilde{x}}, \dot{\tilde{s}})\|_2 &\leq \|(\dot{\tilde{x}} - \check{u}, \dot{\tilde{s}} - \check{v})\|_2 + \|(\check{u}, \check{v})\|_2 \\ &\leq \|v - (\check{u} + \check{v})\|_F + \|(\check{u}, \check{v})\|_2 \end{aligned}$$

$$\begin{aligned} &\leq \|v\|_F + \|\check{u}\|_F + \|\check{v}\|_F + \|(\check{u}, \check{v})\|_2 \\ &\leq \sqrt{\mu r} + (1 + \sqrt{2})\|(\check{u}, \check{v})\|_2 \\ &= \sqrt{\mu r} + (1 + \sqrt{2})\zeta, \end{aligned}$$

where the last inequality uses the result

$$\|v\|_F = \|\sqrt{\check{x} \circ \check{s}}\|_F \leq \sqrt{\|\check{x} \circ \check{s}\|_F} \leq \sqrt{\langle \check{x}, \check{s} \rangle} \leq \sqrt{\mu r}.$$

The proof is completed. □

Using Remark 2 and the proof techniques of Lemma A.1 in [40], we have the following lemma, which gives the upper bound on ζ .

Lemma 4.6 *Let $(\check{u}^0, \check{r}^0, \check{v}^0)$, (x^0, y^0, s^0) satisfy (22), (23) and $(\check{u}, \check{r}, \check{v})$ satisfy the conditions in Lemma 4.5, then $\zeta \leq (5 + 4\Psi)r\sqrt{\mu}/\sqrt{\beta}$.*

Lemma 4.7 *Let $(\check{x}, \check{s}) \in \mathcal{N}_F(\gamma)$, $\beta = 1 - \gamma$, then*

$$\|\dot{\check{x}}\|_F \|\dot{\check{s}}\|_F \leq \frac{1}{2}\omega^2 \mu r^2,$$

where $\omega = 1 + (1 + \sqrt{2})(5 + 4\Psi)/\sqrt{\beta} \geq 11$.

Proof Using Lemmas 4.5 and 4.6, we have

$$\begin{aligned} \|\dot{\check{x}}\|_F \|\dot{\check{s}}\|_F &\leq \frac{1}{2} [\|\dot{\check{x}}\|_F^2 + \|\dot{\check{s}}\|_F^2] = \frac{1}{2} \|(\dot{\check{x}}, \dot{\check{s}})\|_2^2 \\ &\leq \frac{1}{2} [\sqrt{\mu r} + (1 + \sqrt{2})\zeta]^2 \\ &\leq \frac{1}{2} [\sqrt{\mu r} + (1 + \sqrt{2})(5 + 4\Psi)r\sqrt{\mu}/\sqrt{\beta}]^2 \\ &\leq \frac{1}{2} [\sqrt{1/r} + (1 + \sqrt{2})(5 + 4\Psi)/\sqrt{\beta}]^2 \mu r^2 \leq \frac{1}{2}\omega^2 \mu r^2, \end{aligned}$$

which completes the proof. □

Lemma 4.8 *Let $(\check{x}, \check{s}) \in \mathcal{N}_F(\gamma)$, $\tilde{\omega} = Q_{\check{x}^{1/2}\check{s}}$, $\beta = 1 - \gamma$, then*

- (i) $\|(\ddot{\check{x}}, \ddot{\check{s}})\|_2^2 \leq \omega^4 \mu r^4 / \beta$,
- (ii) $\|\ddot{\check{x}}\|_F \|\ddot{\check{s}}\|_F \leq \omega^4 \mu r^4 / (2\beta)$.

Proof Multiplying the equation of (11) by L_v^{-1} and taking norm-squared on both sides, we have

$$\begin{aligned} \|\ddot{\check{x}} + \ddot{\check{s}}\|_F^2 &= \|L_v^{-1}(-2\ddot{\check{x}} \circ \ddot{\check{s}})\|_F^2 = \|(L_{\check{x}}L_{\check{s}})^{-1/2}(-2\ddot{\check{x}} \circ \ddot{\check{s}})\|_F^2 \\ &\leq \lambda_{\max}((L_{\check{x}}L_{\check{s}})^{-1}) \|2\ddot{\check{x}} \circ \ddot{\check{s}}\|_F^2 \leq \frac{4}{\lambda_{\min}(\tilde{\omega})} \|\ddot{\check{x}} \circ \ddot{\check{s}}\|_F^2 \\ &\leq \frac{4}{\beta\mu} [\|\ddot{\check{x}}\|_F \|\ddot{\check{s}}\|_F]^2 \leq \frac{4}{\beta\mu} \left[\frac{1}{2}\omega^2 \mu r^2 \right]^2 = \frac{1}{\beta}\omega^4 \mu r^4, \end{aligned} \tag{28}$$

where the second equality uses (24), the first two inequalities follow from Lemma 4.1, the last two inequalities are due to Lemmas 4.2 and 4.7.

Using the fact $\langle \ddot{\tilde{x}}, \ddot{\tilde{s}} \rangle = 0$ and (28), we have

$$\|\ddot{\tilde{x}}\|_F \|\ddot{\tilde{s}}\|_F \leq \frac{1}{2} [\|\ddot{\tilde{x}}\|^2 + \|\ddot{\tilde{s}}\|_F^2] = \frac{1}{2} \|\ddot{\tilde{x}} + \ddot{\tilde{s}}\|_F^2 \leq \frac{1}{2\beta} \omega^4 \mu r^4.$$

Therefore, the proof of the lemma is completed. □

The next result follows from Lemmas 4.7 and 4.8.

Lemma 4.9 *Let $(\tilde{x}, \tilde{s}) \in \mathcal{N}_F(\gamma)$, $\beta = 1 - \gamma$, then*

- (a) $\|\dot{\tilde{x}}\|_F \|\dot{\tilde{s}}\|_F \leq \omega^3 \mu r^3 / \sqrt{\beta}$,
- (b) $\|\dot{\tilde{s}}\|_F \|\dot{\tilde{x}}\|_F \leq \omega^3 \mu r^3 / \sqrt{\beta}$.

4.2 The lower bounds on $\bar{\theta}$

In this subsection, we will find a lower bounds of $\bar{\theta}$ to satisfy (15) and (16). They will play a key role in complexity analysis. Let $\bar{\theta}^0 = \arg \sin(\frac{\beta\gamma}{2\omega r})$. If we can prove that (15) and (16) hold for all $\theta \in (0, \bar{\theta}^0]$, then $\bar{\theta}^0$ is one of the lower bounds on $\bar{\theta}$. For this purpose, we first give an important lemma.

Lemma 4.10 *Let $(\tilde{x}, \tilde{s}) \in \mathcal{N}_F(\gamma)$, $\beta = 1 - \gamma$ and $\mu(\theta) = (1 - \sin(\theta))\mu$ be defined in (21), then, for all $\theta \in (0, \bar{\theta}^0]$,*

$$\|\tilde{x}(\theta) \circ \tilde{s}(\theta) - \mu(\theta)e\|_F \leq 2\gamma\mu(\theta).$$

Proof In order to express convenience, we give some notations as follows:

$$\begin{aligned} f_1(\theta) &= (1 - \sin(\theta))(\tilde{x} \circ \tilde{s} - \mu e), & f_2(\theta) &= g^2(\theta)\dot{\tilde{x}} \circ \dot{\tilde{s}}, \\ f_3(\theta) &= g(\theta)\sin(\theta)\xi, & f_4(\theta) &= g^2(\theta)\ddot{\tilde{x}} \circ \ddot{\tilde{s}}. \end{aligned}$$

Using (14), Lemma 4.2, Lemmas 4.7, 4.8, 4.9, we have

$$\begin{aligned} \|f_1(\theta)\|_F &\leq (1 - \sin(\theta))\|\tilde{x} \circ \tilde{s} - \mu e\|_F \leq (1 - \sin(\theta))\gamma\mu, \\ \|f_2(\theta)\|_F &= \|g^2(\theta)\dot{\tilde{x}} \circ \dot{\tilde{s}}\| \leq g^2(\theta)\|\dot{\tilde{x}}\|_F \|\dot{\tilde{s}}\|_F \leq \frac{1}{2} \sin^4(\theta)\omega^2 \mu r^2 := a_1, \\ \|f_3(\theta)\|_F &\leq g(\theta)\sin(\theta)\|\xi\|_F \leq \frac{2}{\sqrt{\beta}} \sin^3(\theta)\omega^3 \mu r^3 := a_2, \\ \|f_4(\theta)\|_F &\leq g^2(\theta)\|\ddot{\tilde{x}} \circ \ddot{\tilde{s}}\|_F \leq \frac{1}{2\beta} \sin^4(\theta)\omega^4 \mu r^4 := a_3, \end{aligned}$$

where we use $g(\theta) = 1 - \cos(\theta) \leq \sin^2(\theta)$ and the fact

$$\|\xi\|_F = \|\dot{\tilde{x}} \circ \dot{\tilde{s}} + \dot{\tilde{s}} \circ \dot{\tilde{x}}\|_F \leq \|\dot{\tilde{x}}\|_F \|\dot{\tilde{s}}\|_F + \|\dot{\tilde{s}}\|_F \|\dot{\tilde{x}}\|_F.$$

In what follows, we will estimate upper bounds on a_1, a_2, a_3 in the interval $(0, \bar{\theta}^0]$. Using $\sin(\bar{\theta}^0) = \frac{\beta\gamma}{2\omega r}$, we have

$$\begin{aligned} a_1 &= \frac{1}{2} \sin^4(\theta) \omega^2 \mu r^2 \leq \frac{1}{2} \sin^4(\bar{\theta}^0) \omega^2 \mu r^2 \leq \frac{\gamma^4 \beta^4}{2^5 \omega^2 r^2} \mu := b_1, \\ a_2 &= \frac{2}{\sqrt{\beta}} \sin^3(\theta) \omega^3 \mu r^3 \leq \frac{2}{\sqrt{\beta}} \sin^3(\bar{\theta}^0) \omega^3 \mu r^3 \leq \frac{\gamma^3 \beta^{5/2}}{2^2} \mu := b_2, \\ a_3 &= \frac{1}{2\beta} \sin^4(\theta) \omega^4 \mu r^4 \leq \frac{1}{2\beta} \sin^4(\bar{\theta}^0) \omega^4 \mu r^4 \leq \frac{\gamma^4 \beta^3}{2^5} \mu := b_3. \end{aligned}$$

Based on the analysis, we give the desired result, which is

$$\begin{aligned} &2\gamma\mu(\theta) - \|\tilde{x}(\theta) \circ \tilde{s}(\theta) - \mu(\theta)e\|_F \\ &\geq 2\gamma(1 - \sin(\theta))\mu - [\|f_1(\theta)\|_F + \|f_2(\theta)\|_F + \|f_3(\theta)\|_F + \|f_4(\theta)\|_F] \\ &\geq \gamma(1 - \sin(\theta))\mu - [a_1 + a_2 + a_3] \\ &\geq \gamma(1 - \sin(\theta))\mu - [b_1 + b_2 + b_3] \\ &\geq \gamma\mu \left[1 - \frac{\gamma\beta}{2\omega r} - \frac{\gamma^3\beta^4}{2^5\omega^2r^2} - \frac{\gamma^2\beta^{5/2}}{2^2} - \frac{\gamma^3\beta^3}{2^5} \right] \\ &\geq \gamma\mu \left[1 - \frac{1}{2^3} - \frac{1}{2^{11}} - \frac{1}{2^6} - \frac{1}{2^{11}} \right] \geq 0, \end{aligned}$$

where we use the result of $\tilde{x}(\theta) \circ \tilde{s}(\theta)$ in (13). □

Lemma 4.11 *Let $(\tilde{x}, \tilde{s}) \in \mathcal{N}_F(\gamma)$, then we have $\tilde{x}(\theta) \in \text{int } \mathcal{K}$, $\tilde{s}(\theta) \in \text{int } \mathcal{K}$ for all $\theta \in (0, \bar{\theta}^0]$.*

Proof Using Lemma 4.10 and $\gamma \leq 1/4$, we have

$$\lambda_{\min}(\tilde{x}(\theta) \circ \tilde{s}(\theta)) - \mu(\theta) \geq -2\gamma\mu(\theta),$$

which is equivalent to

$$\lambda_{\min}(\tilde{x}(\theta) \circ \tilde{s}(\theta)) \geq (1 - 2\gamma)\mu(\theta) \geq 0, \tag{29}$$

which furthermore implies $\tilde{x}(\theta) \circ \tilde{s}(\theta) \in \text{int } \mathcal{K}$.

From Lemma 4.3, we have $\det(\tilde{x}(\theta)) \neq 0$ and $\det(\tilde{s}(\theta)) \neq 0$. Furthermore, since $\tilde{x} \in \text{int } \mathcal{K}$, $\tilde{s} \in \text{int } \mathcal{K}$, by the continuity, it follows that both $\tilde{x}(\theta) \in \text{int } \mathcal{K}$ and $\tilde{s}(\theta) \in \text{int } \mathcal{K}$ in $[0, \bar{\theta}^0]$. The proof is completed. □

From the above analysis, we obtain the result that $\bar{\theta}^0$ is one of the lower bounds on $\bar{\theta}$.

4.3 Corrector step and iteration complexity

It is well known that an important requirement for the MTY-PC algorithm is that the new iteration point must stay in the given neighborhood, which is equivalent to proving $(\hat{x}(\theta), \hat{s}(\theta)) \in \mathcal{N}_F(\gamma)$. In what follows, we will complete this task.

Using $(\bar{x}, \bar{s}) = (Q_{p-1}\tilde{x}(\bar{\theta}), Q_p\tilde{s}(\bar{\theta}))$ and $(\hat{x}, \hat{s}) = (Q_{p_1}\bar{x}, Q_{p_1^{-1}}\bar{s})$, we have

$$(\hat{x}, \hat{s}) = (Q_{p_1}\bar{x}, Q_{p_1^{-1}}\bar{s}) = (Q_{p_1}Q_{p-1}\tilde{x}(\bar{\theta}), Q_{p_1^{-1}}Q_p\tilde{s}(\bar{\theta})),$$

which implies (\hat{x}, \hat{s}) in interval $(0, \bar{\theta}^0]$ to satisfy the condition in Lemma 4.10, Lemma 4.11. Thus, by Lemma 4.10, Lemma 4.11, we have

$$\|\hat{x} \circ \hat{s} - \mu(\theta)e\|_F \leq 2\gamma\mu(\theta), \quad \hat{x} \in \text{int } \mathcal{K}, \hat{s} \in \text{int } \mathcal{K}. \tag{30}$$

Lemma 4.12 *Let $(\Delta\hat{x}, \Delta\hat{s})$ be the solution of (18), then we have*

$$\|\Delta\hat{x}\|_F \|\Delta\hat{s}\|_F \leq \frac{2\gamma^2}{(1-2\gamma)}\mu(\theta)$$

for all $\theta \in (0, \bar{\theta}^0]$.

Proof Multiplying the last equation in (18) by $L_{\hat{v}}^{-1}$, we obtain

$$\Delta\hat{x} + \Delta\hat{s} = L_{\hat{v}}^{-1}\hat{r}_c.$$

Taking norm-squared on both sides of the above equation, we have

$$\begin{aligned} \|\Delta\hat{x} + \Delta\hat{s}\|_F^2 &= \|L_{\hat{v}}^{-1}\hat{r}_c\|_F^2 = \|(L_{\hat{x}}L_{\hat{s}})^{-1/2}\hat{r}_c\|_F^2 \\ &\leq \frac{1}{\lambda_{\min}((L_{\hat{x}}L_{\hat{s}})^{-1})} \|\hat{r}_c\|_F^2 \leq \frac{1}{\lambda_{\min}(\hat{w})} \|\mu(\theta)e - \hat{x} \circ \hat{s}\|_F^2 \\ &\leq \frac{1}{(1-2\gamma)\mu(\theta)} (2\gamma\mu(\theta))^2 = \frac{4\gamma^2}{(1-2\gamma)}\mu(\theta), \end{aligned}$$

where $\hat{r}_c = (1 - \sin(\bar{\theta}))\mu e - \hat{x} \circ \hat{s}$, the second equality uses (25), the second inequality follows from Lemma 4.1, the third inequality is due to (29), (30).

Using the conclusion as above and the fact $\langle \Delta\hat{x}, \Delta\hat{s} \rangle = 0$, we have

$$\|\Delta\hat{x}\|_F \|\Delta\hat{s}\|_F \leq \frac{1}{2} [\|\Delta\hat{x}\|_F^2 + \|\Delta\hat{s}\|_F^2] = \frac{1}{2} \|\Delta\hat{x} + \Delta\hat{s}\|_F^2 \leq \frac{2\gamma^2}{(1-2\gamma)}\mu(\theta),$$

which completes the proof. □

Lemma 4.13 *Let $(\hat{x}(\theta), \hat{s}(\theta))$ be defined in (19), $\theta \in (0, \bar{\theta}^0]$, then we have*

$$(\hat{x}(\theta), \hat{s}(\theta)) \in \mathcal{N}_F(\gamma).$$

Proof Using Lemma 4.4, (20) and Lemma 4.12, we have

$$\begin{aligned} \|\hat{w}(\theta) - \mu(\theta)e\|_F &\leq \|\hat{x}(\theta) \circ \hat{s}(\theta) - \mu(\theta)e\|_F \\ &= \|\Delta\hat{x} \circ \Delta\hat{s}\|_F \leq \|\Delta\hat{x}\|_F \|\Delta\hat{s}\|_F \\ &\leq \frac{2\gamma^2}{(1-2\gamma)}\mu(\theta) \leq \frac{2\gamma}{1-2\gamma}\gamma\mu(\theta) \leq \gamma\mu(\theta), \end{aligned}$$

where $\hat{w}(\theta) = Q_{\hat{x}(\theta)^{1/2}\hat{s}(\theta)}$, the last inequality follows from $\gamma \leq 1/4$.

Using the proof technique that is similar to Lemma 4.11, we have

$$\hat{x}(\theta) \in \text{int } \mathcal{K}, \quad \hat{s}(\theta) \in \text{int } \mathcal{K}.$$

Taking into account the above factors, we have $(\hat{x}(\theta), \hat{s}(\theta)) \in \mathcal{N}_F(\gamma)$. □

The following theorem gives an upper bound for the number of iterations in which Algorithm 1 stops with an ε -approximate solution.

Theorem 4.14 *Let r_p^k, r_d^k, μ^k be defined in Remark 1, then Algorithm 1 will terminate in $\mathcal{O}(r \log \varepsilon^{-1})$ iterations such that*

$$\|r_p^k\| \leq \varepsilon \|r_p^0\|, \quad \|r_d^k\| \leq \varepsilon \|r_d^0\|, \quad \mu^k \leq \varepsilon \mu^0.$$

Proof By using $\bar{\theta}^0 = \arg \sin(\frac{\beta\gamma}{2\omega r}) \leq \bar{\theta}$, we have

$$\phi^k = \prod_{i=0}^{k-1} (1 - \sin(\bar{\theta}^i)) \leq \prod_{i=0}^{k-1} (1 - \sin(\bar{\theta}^0)) \leq (1 - \sin(\bar{\theta}^0))^k \leq \varepsilon,$$

which implies

$$k \geq \frac{1}{\sin(\bar{\theta}^0)} \log \varepsilon^{-1} \geq \frac{2\omega}{\beta\gamma} r \log \varepsilon^{-1},$$

where we use the identity $\log(1 + t) \leq t$ for all $t > -1$.

Therefore, Algorithm 1 terminates after at most $\mathcal{O}(r \log \varepsilon^{-1})$ steps. Meanwhile, by using Remark 1, we have

$$\phi^k = \frac{\|r_p^k\|_F}{\|r_p^0\|_F} = \frac{\|r_d^k\|_F}{\|r_d^0\|_F} = \frac{\mu^k}{\mu^0} \leq \varepsilon.$$

This completes the proof. □

5 Conclusion

For the SO problem, we have proposed an MTY-PC infeasible-IPM, which requires two matrix factorizations and at most three back-solves for each iteration. In order to improve the iteration complexity, we adopt the arc-search strategy that was proposed by Yang [32–34]. Moreover, the proposed algorithm can ensure that the duality gap and the infeasibility have the same rate of decline. Finally, by applying the EJA tool to our analysis, we established the iteration complexity $\mathcal{O}(r \log \varepsilon^{-1})$ for the NT-direction.

Acknowledgements

We would like to thank the support of National Natural Science Foundation of China (NNSFC) under Grant No. 11501180, 11601134 and 11671122, Chinese Postdoctoral Science Foundation No. 2016M590346, Henan Normal University Doctoral Startup Issues No. qd14150 and Young Scientists Foundation No. 2014QK03, and Innovative Research Team (in Science and Technology) in University of Henan Province No. 14IRTSTHN023.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors contributed equally to this work. All authors read and approved the final manuscript.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 23 March 2017 Accepted: 13 November 2017 Published online: 21 November 2017

References

- Karmarkar, N: A new polynomial-time algorithm for linear programming. *Combinatorica* **4**, 373-393 (1984)
- Güler, O: Barrier functions in interior-point methods. *Math. Oper. Res.* **21**, 860-885 (1996)
- Faybusovich, L: Linear systems in Jordan algebras and primal-dual interior-point algorithms. *J. Comput. Appl. Math.* **86**, 149-175 (1997)
- Nesterov, Y, Todd, M: Primal-dual interior-point methods for self-scaled cones. *SIAM J. Optim.* **8**, 324-364 (1998)
- Schmieta, SH, Alizadeh, F: Extension of primal-dual interior point algorithm to symmetric cones. *Math. Program.* **96**, 409-438 (2003)
- Rangarajan, BK: Polynomial convergence of infeasible-interior-point methods over symmetric cones. *SIAM J. Optim.* **16**, 1211-1229 (2006)
- Zhang, J, Zhang, K: Polynomial complexity of an interior point algorithm with a second order corrector step for symmetric cone programming. *Math. Methods Oper. Res.* **73**, 75-90 (2011)
- Liu, C, Liu, H, Liu, X: Polynomial convergence of second-order Mehrotra-type predictor-corrector algorithms over symmetric cones. *J. Optim. Theory Appl.* **154**, 949-965 (2012)
- Liu, H, Yang, X, Liu, C: A new wide neighborhood primal-dual infeasible-interior-point method for symmetric cone programming. *J. Optim. Theory Appl.* **158**, 796-815 (2013)
- Wang, G, Yu, C, Teo, K: A new full Nesterov-Todd step feasible interior-point method for convex quadratic symmetric cone optimization. *Appl. Math. Comput.* **221**, 329-343 (2013)
- Wang, G, Kong, L, Tao, J, Lesaja, G: Improved complexity analysis of full Nesterov-Todd step feasible interior-point method for symmetric optimization. *J. Optim. Theory Appl.* **166**, 588-604 (2015)
- Mehrotra, S: On the implementation of a primal-dual interior point method. *SIAM J. Optim.* **2**, 575-601 (1992)
- Mizuno, S, Todd, MJ, Ye, Y: On adaptive-step primal-dual interior-point algorithms for linear programming. *Math. Oper. Res.* **18**, 964-981 (1993)
- Zhang, D, Zhang, Y: On polynomiality of the Mehrotra-type predictor-corrector interior-point algorithms. *Math. Program.* **68**, 303-318 (1995)
- Ye, Y, Güler, O, Tapia, RA, Zhang, Y: A quadratically convergent $O(\sqrt{nL})$ -iteration algorithm for linear programming. *Math. Program.* **59**, 151-162 (1993)
- Potra, FA: A quadratically convergent predictor-corrector method for solving linear programs from infeasible starting points. *Math. Program.* **67**, 383-406 (1994)
- Zhang, Y, El-Bakry, A: Modified predictor-corrector algorithm for locating weighted centers in linear programming. *J. Optim. Theory Appl.* **80**, 319-334 (1994)
- Ji, J, Potra, FA: Predictor-corrector method for linear complementarity problems with polynomial complexity and superlinear convergence. *J. Optim. Theory Appl.* **85**, 187-199 (1995)
- Miao, J: A quadratically convergent $O((1 + \kappa)\sqrt{nL})$ -iteration algorithm for the $P_*(\kappa)$ -matrix linear complementarity problem. *Math. Program.* **69**, 355-368 (1995)
- Potra, FA: An $O(nl)$ infeasible-interior-point algorithm for LCP with quadratic convergence. *Ann. Oper. Res.* **62**, 81-102 (1996)
- Gonzaga, CC, Tapia, RA: On the convergence of the Mizuno-Todd-Ye algorithm to the analytic center of the solution set. *SIAM J. Optim.* **7**, 49-65 (1997)
- Kojima, M, Shida, M, Shindoh, S: Local convergence of predictor-corrector infeasible-interior-point algorithm for SDPs and SDLCPs. *Math. Program.* **80**, 129-160 (1998)
- Renato, DC, Tsuchiya, T: A new iteration-complexity bound for the MTY predictor-corrector algorithm. *SIAM J. Optim.* **15**, 319-347 (2005)
- Illés, T, Nagya, M: A Mizuno-Todd-Ye type predictor-corrector algorithm for sufficient linear complementarity problems. *Eur. J. Oper. Res.* **181**, 1097-1111 (2007)
- Monteiro, RD, Tsuchiya, T: A strong bound on the integral of the central path curvature and its relationship with the iteration-complexity of primal-dual path-following LP algorithms. *Math. Program.* **115**, 105-149 (2008)
- Tsuchiya, T: A simple variant of the Mizuno-Todd-Ye predictor-corrector algorithm and its objective-function-free complexity. *SIAM J. Optim.* **23**, 1890-1903 (2013)
- Yang, X, Zhang, Y, Liu, H, Pei, Y: A Mizuno-Todd-Ye predictor-corrector infeasible-interior-point method for linear programming over symmetric cones. *Numer. Algorithms* **72**, 915-936 (2016)
- Potra, FA: A quadratically convergent infeasible interior-point algorithm for linear programming. Reports on computational mathematics 28, Department of Mathematics, University of Iowa, Iowa City, IA 52242, USA (1992)
- Kojima, M, Megiddo, N, Mizuno, S: A primal-dual infeasible-interior-point algorithm for linear programming. *Math. Program.* **61**, 263-280 (1993)
- Ye, Y, Todd, MJ, Mizuno, S: An $O(\sqrt{nL})$ -iteration homogeneous and self-dual linear programming algorithm. *Math. Oper. Res.* **19**, 53-67 (1994)
- Zhang, Y: On the convergence of a class of infeasible interior-point methods for the horizontal linear complementarity problem. *SIAM J. Optim.* **4**, 208-227 (1994)
- Yang, Y: Arc-search path-following interior-point algorithms for linear programming. *Optim. Online* (2009)
- Yang, Y: A polynomial arc-search interior-point algorithm for convex quadratic programming. *Eur. J. Oper. Res.* **215**, 25-38 (2011)

34. Yang, Y: A polynomial arc-search interior-point algorithm for linear programming. *J. Optim. Theory Appl.* **158**, 859-873 (2013)
35. Yang, X, Liu, H, Zhang, Y: An arc-search infeasible-interior-point method for symmetric optimization in a wide neighborhood of the central path. *Optim. Lett.* **11**, 135-152 (2017)
36. Yang, X, Liu, H, Zhang, Y: A wide neighborhood infeasible-interior-point method with arc-search for linear programming. *J. Appl. Math. Comput.* **51**, 209-225 (2016)
37. Faraut, J, Korányi, A: *Analysis on Symmetric Cone*. Oxford University Press, New York (1994)
38. Muramatsu, M: On a commutative class of search directions for linear programming over symmetric cones. *J. Optim. Theory Appl.* **112**, 595-625 (2002)
39. Gu, G, Zangiabadi, M, Roos, C: Full Nesterov-Todd step infeasible interior-point method for symmetric optimization. *Eur. J. Oper. Res.* **214**, 473-484 (2011)
40. Liu, C: Study on complexity of some interior-point algorithms in conic programming. Ph.D. thesis, Xidian University (2012) (in Chinese)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
