Improved Complexity Analysis of Full Nesterov–Todd Step Feasible Interior-Point Method for Symmetric Optimization

G. Q. Wang · L. C. Kong · J. Y. Tao · G. Lesaja

Abstract In this paper, an improved complexity analysis of full Nesterov–Todd step feasible interior-point method for symmetric optimization is considered. Specifically, we establish a sharper quadratic convergence result using several new results from Euclidean Jordan algebras, which leads to a wider quadratic convergence neighbourhood of the central path for the iterates in the algorithm. Furthermore, we derive the currently best known iteration bound for full Nesterov–Todd step feasible interior-point method.

G. Q. Wang (✉)
College of Fundamental Studies, Shanghai University of Engineering Science,
Shanghai 201620, People's Republic of China
e-mail: guoq_wang@hotmail.com

L. C. Kong
Department of Applied Mathematics, Beijing Jiaotong University, Beijing 100044,
People's Republic of China
e-mail: konglchen@126.com

J. Y. Tao
Department of Mathematics and Statistics, Loyola University Maryland, Baltimore,
MD 21210, USA
e-mail: jtao@loyola.edu

G. Lesaja
Department of Industrial Engineering, Yasar University, Izmir, Turkey
e-mail: goran.lesaja@yasar.edu.tr

1 Introduction

Symmetric optimization (SO) problem is a convex optimization problem, which minimizes a linear function over the intersection of an affine subspace and a symmetric cone. The SO problem is an important class of convex optimization problems, because it includes linear optimization (LO), second-order cone optimization (SOCO), and semidefinite optimization (SDO) as special cases. Particularly, many interior-point methods (IPMs) for LO are successfully extended to SOCO and SDO (see, e.g., [1–5]), due to their polynomial complexity and practical efficiency.

Güler [6] was the first to realize that symmetric cones serve as a unifying framework to which the important types of cones used in optimization such as nonnegative orthant, second-order cone, and cone of positive semidefinite matrices belong. Faybusovich [7] made the first attempt to generalize IPMs from LO to SO using EJAs and associated symmetric cones. Consequently, several efficient IPMs for SO have been developed (see, e.g., [8–15]). For an overview of the relevant results, we refer to the recent book on this subject [16] and the references therein.

The primal-dual full Newton-step feasible IPM for LO was first analysed by Roos et al. in [17] and was later extended to infeasible version by Roos [18]. Subsequently, Gu et al. [19] proposed an improved version of full Newton-step infeasible IPM for LO where the convergence analysis of the algorithm was simplified. However, it still maintains the currently best known iteration bound. Both versions of the method were extended by Gu et al. [8] to SO using EJAs. The search direction used in [8] is based on the Nesterov–Todd (NT) scheme. The obtained iteration bounds coincide with the ones derived for LO, where n is replaced by r, the rank of EJAs, which are the currently best known iteration bounds for SO. Some other references, related to full step IPMs, without any attempt to be complete, are [12,20–24].

In this paper, we present an improved convergence and complexity analysis of the full NT-step feasible IPM for SO discussed in [8]. More specifically, a sharper quadratic convergence result is established, which leads to a wider neighbourhood of the central path for the iterates in the algorithm. The purpose of the paper is mainly theoretical, which is to show that despite full NT-steps in the wider neighbourhood of the central path, the best complexity known for these types of methods is still maintained. Moreover, these features give a certain numerical appeal to the methods although they are still small-update methods and cannot quite compare in efficiency with large-update methods whose theoretical complexity is far worse (the "irony" of IPMs).

The outline of the paper is as follows. In Sect. 2, we recall some known results that are needed in this paper. More importantly, some new results are developed that are used to perform the new complexity analysis of feasible IPM. In Sect. 3, we present the full NT-step feasible IPM for SO with its complexity analysis. Finally, some concluding remarks are made in Sect. 4.

2 Preliminaries

In what follows we assume that the reader is familiar with the basic concepts of EJAs and symmetric cones. A comprehensive treatment of EJAs can be found in the monograph [25] and in [7,8,11,12,26] as it relates to optimization.

Let $(\mathcal{V}, \circ)$ be an n-dimensional EJA with rank r equipped with the standard inner product $\langle x, s \rangle = \text{tr}(x \circ s)$ (see, (8)) and $\mathcal{K}$ be the corresponding symmetric cone. Moreover, we always assume that there exists an identity element $e \in \mathcal{V}$ such that $e \circ x = x$ for all $x \in \mathcal{V}$.

For any element $x \in \mathcal{V}$, the Lyapunov transformation $L(x) : \mathcal{V} \to \mathcal{V}$ is given by

$$L(x)y := x \circ y, \quad \forall y \in \mathcal{V}. \tag{1}$$

Furthermore, we define the quadratic representation of x in $\mathcal{V}$ as follows:

$$P(x) := 2L(x)^2 - L\left(x^2\right), \tag{2}$$

where $L(x)^2 = L(x)L(x)$.

For any EJA $\mathcal{V}$, the corresponding cone of squares

$$\mathcal{K}(\mathcal{V}) := \left\{ x^2 : x \in \mathcal{V} \right\} \tag{3}$$

is indeed a symmetric cone, i.e. self-dual closed convex and homogeneous cone [25].

For any $x \in \mathcal{V}$, let r be the smallest integer such that the set $\{e, x, \ldots, x^r\}$ is linearly dependent. Then r is the degree of x which is denoted as $\deg(x)$. Clearly, this degree of x is bounded by the dimension of the vector space $\mathcal{V}$. The rank of $\mathcal{V}$, denoted by $\text{rank}(\mathcal{V})$, is the largest $\deg(x)$ of any element $x \in \mathcal{V}$. An element $x \in \mathcal{V}$ is called regular if its degree equals the rank of $\mathcal{V}$. In the sequel, otherwise stated, $\mathcal{V}$ is used to be denoted an EJA with $\text{rank}(\mathcal{V}) = r$.

For a regular element $x \in \mathcal{V}$, since $\{e, x, x^2, \ldots, x^r\}$ is linearly dependent, there are real numbers $a_1(x), \ldots, a_r(x)$ such that the minimal polynomial of every regular element x is given by

$$f(\lambda; x) = \lambda^r - a_1(x)\lambda^{r-1} + \cdots + (-1)^r a_r(x), \tag{4}$$

which is the characteristic polynomial of the regular element x. The coefficient $a_1(x)$ is called the trace of x, denoted as $\text{tr}(x)$. And the coefficient $a_r(x)$ is called the determinant of x, denoted as $\det(x)$.

An element $c \in \mathcal{V}$ is said to be an idempotent if $c^2 = c$. Two idempotents c_1 and c_2 are said to be orthogonal if $c_1 \circ c_2 = 0$. Moreover, an idempotent is primitive if it is non-zero and cannot be written as the sum of two (necessarily orthogonal) non-zero idempotents. We say that $\{c_1, \ldots, c_r\}$ is a complete system of orthogonal primitive idempotents, or Jordan frame, if each c_i is a primitive idempotent, $c_i \circ c_j = 0, i \neq j$, and $\sum_{i=1}^{r} c_i = e$. The Löwner partial ordering "$\succeq_{\mathcal{K}}$" of $\mathcal{V}$ defined by a cone $\mathcal{K}$ is defined by $x \succeq_{\mathcal{K}} s$ if $x - s \in \mathcal{K}$. The interior of $\mathcal{K}$ is denoted as $\text{int}\mathcal{K}$ and we write $x \succ_{\mathcal{K}} s$ if $x - s \in \text{int}\mathcal{K}$.

The following theorem describes a spectral decomposition of elements in $\mathcal{V}$, which plays an important role in the analysis of the IPMs for SO.

Theorem 2.1 (Theorem III.1.2 in [25]) *Let* $x \in \mathcal{V}$. *Then there exist a Jordan frame* $\{c_1, \ldots, c_r\}$ *and real numbers* $\lambda_1(x), \ldots, \lambda_r(x)$ *such that*

$$x = \sum_{i=1}^{r} \lambda_i(x)c_i. \tag{5}$$

The numbers $\lambda_i(x)$ *(with their multiplicities) are the eigenvalues of* x. *Furthermore, the trace and the determinant of* x *are given by*

$$\mathrm{tr}(x) = \sum_{i=1}^{r} \lambda_i(x) \text{ and } \det(x) = \prod_{i=1}^{r} \lambda_i(x). \tag{6}$$

Fix a Jordan frame $\{c_1, c_2, \ldots, c_r\}$ in an EJA $\mathcal{V}$. For $i, j \in \{1, 2, \ldots, r\}$, define the eigenspaces

$$\mathcal{V}_{ii} := \{x \in \mathcal{V} : x \circ c_i = x\} = \mathbb{R}\, c_i, \mathcal{V}_{ij} := \left\{x \in \mathcal{V} : x \circ c_i = \frac{1}{2}x = x \circ c_j\right\}, i \neq j.$$

The following theorem provides another important decomposition, the Peirce decomposition, of the space $\mathcal{V}$.

Theorem 2.2 (Theorem IV.2.1 in [25]) *The space* $\mathcal{V}$ *is the orthogonal direct sum of the spaces* $\mathcal{V}_{ij}$ $(i \leq j)$, *i.e.*

$$\mathcal{V} = \bigoplus_{i \leq j} \mathcal{V}_{ij}.$$

Furthermore, we have $\mathcal{V}_{ij} \circ \mathcal{V}_{ij} \subset \mathcal{V}_{ii} + \mathcal{V}_{jj}$, $\mathcal{V}_{ij} \circ \mathcal{V}_{jk} \subset \mathcal{V}_{ik}$, *if* $i \neq k$, *and*

$$\mathcal{V}_{ij} \circ \mathcal{V}_{kl} = \{0\}, \ if\{i, j\} \cap \{k, l\} = \emptyset.$$

Thus, the Peirce decomposition of $x \in \mathcal{V}$ with respect to the Jordan frame $\{c_1, \ldots, c_r\}$ is given by

$$x = \sum_{i=1}^{r} x_i c_i + \sum_{i < j} x_{ij} \tag{7}$$

with $x_i \in \mathbb{R}, i = 1, \ldots, r$ and $x_{ij} \in \mathcal{V}_{ij}, 1 \leq i < j \leq r$.

Corollary 2.1 (Lemma 12 in [11]) *Lex* $x \in \mathcal{V}$ *and its spectral decomposition with respect to the Jordan frame* $\{c_1, \ldots, c_r\}$ *is given by* (5). *Then the following statements hold.*

(i) *The matrices,* $L(x)$ *and* $P(x)$ *commute and thus share a common system of eigenvectors; in fact the* c_i *are among their common eigenvectors.*

(ii) *The eigenvalues of* $L(x)$ *have the form* $\frac{\lambda_i + \lambda_j}{2}$, $1 \leq i \leq j \leq r$.

(iii) *The eigenvalues of* $P(x)$ *have the form* $\lambda_i \lambda_j$, $1 \leq i \leq j \leq r$.

As already indicated, for any $x, s \in \mathcal{V}$, the trace inner product is defined by

$$\langle x, s \rangle := \text{tr}(x \circ s). \tag{8}$$

Thus, $\text{tr}(x) = \langle x, e \rangle$. Hence, it is easy to verify that

$$\text{tr}(x + s) = \text{tr}(x) + \text{tr}(s) \quad \text{and} \quad x \preceq_{\mathcal{K}} s \Rightarrow \text{tr}(x) \le \text{tr}(s). \tag{9}$$

The Frobenius norm induced by this trace inner product is given by

$$\|x\|_F := \sqrt{\langle x, x \rangle}. \tag{10}$$

It follows from Theorem 2.1 that

$$\|x\|_F = \sqrt{\text{tr}(x^2)} = \sqrt{\sum_{i=1}^{r} \lambda_i^2(x)}. \tag{11}$$

One can easily verify that

$$|\lambda_{\min}(x)| \le \|x\|_F \quad \text{and} \quad |\lambda_{\max}(x)| \le \|x\|_F. \tag{12}$$

Furthermore, we have

$$\left\| x^2 \right\|_F \le \|x\|_F^2. \tag{13}$$

In the following two lemmas, we recall several important inequalities used in the sequel.

Lemma 2.1 (Lemma 2.13 in [8]) *Let $x, s \in \mathcal{V}$ and $\langle x, s \rangle = 0$. Then*

$$-\frac{1}{4}\|x + s\|_F^2 \, e \preceq_{\mathcal{K}} x \circ s \preceq_{\mathcal{K}} \frac{1}{4}\|x + s\|_F^2 \, e.$$

Lemma 2.2 (Lemma 2.16 in [8]) *Let $x, s \in \mathcal{V}$. Then*

$$\|x \circ s\|_F \le \frac{1}{2} \left\| x^2 + s^2 \right\|_F.$$

Next lemma provides an important inequality connecting eigenvalues of $x \circ s$ with the sum of Frobenius norms of x and s.

Lemma 2.3 *Let $x, s \in \mathcal{V}$. Then*

$$\sum_{i=1}^{r} |\lambda_i(x \circ s)| \le \frac{1}{2} \left(\|x\|_F^2 + \|s\|_F^2 \right).$$

Proof For any $x, s \in \mathcal{V}$, we have

$$x^2 + s^2 - 2x \circ s = (x - s)^2 \in \mathcal{K} \text{ and } x^2 + s^2 + 2x \circ s = (x + s)^2 \in \mathcal{K}.$$

This implies that

$$-\frac{1}{2}(x^2 + s^2) \preceq_{\mathcal{K}} x \circ s \preceq_{\mathcal{K}} \frac{1}{2}(x^2 + s^2).$$

From Theorem 8 in [27], we have

$$\lambda_i^{\downarrow}(x \circ s) \leq \frac{1}{2}\lambda_i^{\downarrow}(x^2 + s^2) \text{ and } -\lambda_i^{\downarrow}(x \circ s) \leq \frac{1}{2}\lambda_i^{\downarrow}(x^2 + s^2)$$

for all i, where for $z \in \mathcal{V}$, $\lambda_i^{\downarrow}(z)$ $(i = 1, 2, \ldots, r)$ denote the eigenvalues of z written in the decreasing order. Thus,

$$|\lambda_i^{\downarrow}(x \circ s)| \leq \frac{1}{2}\lambda_i^{\downarrow}(x^2 + s^2)$$

for all i. This implies that

$$\sum_{i=1}^{r} |\lambda_i(x \circ s)| \leq \frac{1}{2}\sum_{i=1}^{r} \lambda_i(x^2 + s^2) = \frac{1}{2}\mathrm{tr}(x^2 + s^2) = \frac{1}{2}(\mathrm{tr}(x^2) + \mathrm{tr}(s^2)).$$

Since $\mathrm{tr}(x^2) = \|x\|_F^2$ and $\mathrm{tr}(s^2) = \|s\|_F^2$, we have

$$\sum_{i=1}^{r} |\lambda_i(x \circ s)| \leq \frac{1}{2}\left(\|x\|_F^2 + \|s\|_F^2\right).$$

This completes the proof.   □

If $\langle x, s \rangle = 0$, then $\|x+s\|_F^2 = \|x\|_F^2 + \|s\|_F^2$. Thus, the following corollary follows immediately from Lemma 2.3.

Corollary 2.2 *Let $x, s \in \mathcal{V}$ and $\langle x, s \rangle = 0$. Then*

$$\sum_{i=1}^{r} |\lambda_i(x \circ s)| \leq \frac{1}{2}\|x + s\|_F^2.$$

Lemma 2.4 (Lemma C.6 in [17]) *Let γ be a vector in $\mathbb{R}^p$ such that $\gamma > -1$ and $\mathbf{1}^T \gamma = \sigma$, where $\mathbf{1}$ denotes the all-one vector in $\mathbb{R}^p$. Then if either $\gamma \geq 0$ or $\gamma \leq 0$,*

$$\sum_{i=1}^{p} \frac{-\gamma_i}{1 + \gamma_i} \leq \frac{-\sigma}{1 + \sigma}.$$

Equality holds if and only if at most one of the coordinates of γ is non-zero.

Lemma 2.5 *Let $u, v \in \mathcal{V}$ and $\langle u, v \rangle = 0$, and suppose $\|u + v\|_F = 2a$ with $a < 1$.*
Then

$$\left\langle e, (e + u \circ v)^{-1} - e \right\rangle \leq \frac{2a^4}{1 - a^4}.$$

Proof It follows from Lemma 2.1 that

$$|\lambda_i(u \circ v)| \leq \frac{1}{4}\|u + v\|_F^2 = \frac{1}{4} \cdot 4a^2 = a^2 < 1, \quad i = 1, \ldots, r.$$

This implies that $-e \prec_\mathcal{K} u \circ v \prec_\mathcal{K} e$. Let $\beta := u \circ v$, then $\mathrm{tr}(\beta) = 0$ by assumption.
It easily follows that $\langle e, \beta \rangle = 0$. Next, let

$$I_+(\beta) := \{i : \lambda_i(\beta) > 0\} \text{ and } I_-(\beta) := \{i : \lambda_i(\beta) < 0\}.$$

Then

$$\sigma := \sum_{i \in I_+(\beta)} \lambda_i(\beta) = - \sum_{i \in I_-(\beta)} \lambda_i(\beta).$$

Let $\gamma_i = \lambda_i(\beta)$ for $i \in I_+(\beta)$ and $\gamma_i = \lambda_i(\beta)$ for $i \in I_-(\beta)$. From Lemma 2.4, we
have

$$
\begin{aligned}
\left\langle e, (e + u \circ v)^{-1} - e \right\rangle &= \left\langle e, (e + \beta)^{-1} - e \right\rangle = \left\langle e, -\beta \circ (e + \beta)^{-1} \right\rangle \\
&= \sum_{i=1}^{r} \frac{-\lambda_i(\beta)}{1 + \lambda_i(\beta)} = \sum_{i \in I_+(\beta)} \frac{-\lambda_i(\beta)}{1 + \lambda_i(\beta)} + \sum_{i \in I_-(\beta)} \frac{-\lambda_i(\beta)}{1 + \lambda_i(\beta)} \\
&\leq \frac{-\sigma}{1 + \sigma} + \frac{\sigma}{1 - \sigma} = \frac{2\sigma^2}{1 - \sigma^2}.
\end{aligned}
\tag{14}
$$

The last expression is monotonically increasing in σ. Hence we may replace it by
an upper bound, which can be obtained from Corollary 2.2 as follows:

$$\sigma = \frac{1}{2} \sum_{i=1}^{r} |\lambda_i(\beta)| = \frac{1}{2} \sum_{i=1}^{r} |\lambda_i(x \circ s)| \leq \frac{1}{4}\|u + v\|_F^2 = a^2.$$

Substitution of this bound for σ into (14) yields the desired inequality

$$\left\langle e, (e + u \circ v)^{-1} - e \right\rangle \leq \frac{2a^4}{1 - a^4}.$$

This completes the proof.   □

3 Full Nesterov–Todd Step Feasible Interior-Point Method

In this section, we present a new convergence and complexity analysis for the full NT-step feasible IPM for SO given in [8]. It is shown that the iterates can be taken in the wider quadratic convergence neighbourhood of the central path than previously obtained in [8]. Furthermore, we establish the currently best known iteration bound for full NT-step feasible small-update methods.

3.1 The SO Problems

In this paper, we consider the primal SO problem in the standard form

$$\text{(SOP)} \qquad \min \ \{\langle c, x \rangle : \ Ax = b, \ x \in \mathcal{K}\},$$

and its dual problem

$$\text{(SOD)} \qquad \max \ \{b^{\mathrm{T}} y : \ A^{\mathrm{T}} y + s = c, \ s \in \mathcal{K}\},$$

where A is a matrix representation of the linear operator from $\mathcal{V}$ to $\mathbb{R}^m$. Vector c and the rows of A lie in $\mathcal{V}$, $b \in \mathbb{R}^m$, and A^{T} is the adjoint of A. Throughout the paper, we assume that the rows of A are linearly independent. Many researchers have studied SO and achieved numerous important results [8–12].

3.2 A Brief Outline of the Full NT-Step Feasible IPM

Without loss of generality, as established in [28], we can assume that both (SOP) and (SOD) satisfy the interior-point condition (IPC), i.e. there exists (x^0, y^0, s^0) such that

$$Ax^0 = b, \ x^0 \in \text{int}\mathcal{K}, \ A^{\mathrm{T}} y^0 + s^0 = c, \ s^0 \in \text{int}\mathcal{K}.$$

The Karush–Kuhn–Tucker (KKT) conditions for (SOP) and (SOD) are given by

$$Ax = b, \ x \in \mathcal{K}, \quad A^{\mathrm{T}} y + s = c, \ s \in \mathcal{K}, \quad x \circ s = 0. \qquad (15)$$

The standard approach of primal-dual IPMs is to replace the third equation in (15), that is called complementarity condition for (SOP) and (SOD), by the parameterized equation $x \circ s = \mu e$ with $\mu > 0$, which yields the following system

$$Ax = b, \ x \in \mathcal{K}, \quad A^{\mathrm{T}} y + s = c, \ s \in \mathcal{K}, \quad x \circ s = \mu e. \qquad (16)$$

The parameterized system (16) has a unique solution $(x(\mu), y(\mu), s(\mu))$ for each $\mu > 0$. The $x(\mu)$ component is called the μ-centre of (SOP) and the pair $(y(\mu), s(\mu))$ is called the μ-centre of (SOD). The set of μ-centres forms a homotopy path with μ running through all positive real numbers, which is called the central path. If $\mu \to 0$,

then the limit of the central path exists and since the limit points satisfy the complementarity condition, i.e., $x \circ s = 0$, it naturally yields an optimal solution for (SOP) and (SOD) (see, e.g., [7, 11]).

The core idea of primal-dual IPMs is to follow the central path approximately, i.e. within a certain neighbourhood, and to approach the optimal set of SO by gradually reducing μ to zero. The standard 'tool' for tracing the central path approximately is a Newton method. Given an iterate (x, y, s) 'close' to the central path (the measure of the closeness will be defined in the sequel), the new iterate is obtained by taking full NT-step as described below. Applying Newton's method to the system (16), we have

$$A\Delta x = 0, \quad A^{\mathrm{T}}\Delta y + \Delta s = 0, \quad x \circ \Delta s + s \circ \Delta x = \mu e - x \circ s. \qquad (17)$$

The system does not always have a unique solution due to the fact that x and s do not operator commute in general, i.e. $L(x)L(s) \neq L(s)L(x)$. To overcome this difficulty, the third equation of the system (16) is replaced by the following equivalent scaled equation (cf. Lemma 28 in [11])

$$P(u)x \circ P(u^{-1})s = \mu e,$$

where u is a scaling point from the interior of the cone $\mathcal{K}$.

Applying Newton's method to this modified system, we have

$$A\Delta x = 0,$$
$$A^{\mathrm{T}}\Delta y + \Delta s = 0, \qquad (18)$$
$$P(u)x \circ P(u^{-1})\Delta s + P(u^{-1})s \circ P(u)\Delta x = \mu e - P(u)x \circ P(u^{-1})s.$$

There are several appropriate choices for the scaling point u that lead to obtaining unique search directions from system (18) (see, e.g. [9]). In this paper, we use the classical NT-scaling point to find unique NT-direction. Let $u := w^{-\frac{1}{2}}$, where

$$w := P(x)^{\frac{1}{2}} \left(P\left(x^{\frac{1}{2}}\right)s \right)^{-\frac{1}{2}} \left[= P(s^{-\frac{1}{2}}) \left(P\left(s^{\frac{1}{2}}\right)x \right)^{\frac{1}{2}} \right] \qquad (19)$$

is the NT-scaling point of x and s. This scaling point was first proposed by Nesterov and Todd for self-scaled cones [29, 30] and then adapted by Faybusovich [7] for symmetric cones.

Introducing the variance vector

$$v := \frac{P(w)^{-\frac{1}{2}}x}{\sqrt{\mu}} \left[= \frac{P(w)^{\frac{1}{2}}s}{\sqrt{\mu}} \right], \qquad (20)$$

and the scaled search directions

$$d_x := \frac{P(w)^{-\frac{1}{2}}\Delta x}{\sqrt{\mu}}, \ d_s := \frac{P(w)^{\frac{1}{2}}\Delta s}{\sqrt{\mu}}, \qquad (21)$$

the system (18) is further simplified

$$\overline{A}d_x = 0, \quad \overline{A}^{\mathrm{T}}\Delta y + d_s = 0, \quad d_x + d_s = v^{-1} - v, \tag{22}$$

where $\overline{A} := \frac{1}{\sqrt{\mu}}AP(w)^{\frac{1}{2}}$. This system has a unique solution $(d_x, \Delta y, d_s)$.

From the first two equations of the system (22), one can easily verify that the scaled search directions d_x and d_s are orthogonal with respect to the trace inner product, i.e. $\langle d_x, d_s \rangle = 0$. Hence, we have

$$d_x = d_s = 0 \Leftrightarrow v^{-1} - v = 0 \Leftrightarrow v = e,$$

which implies that the iterate (x, y, s) coincides with the corresponding μ-centre if and only if $v = e$. The original search directions can then be obtained from (21)

$$\Delta x = \sqrt{\mu}P(w)^{\frac{1}{2}}d_x \text{ and } \Delta s = \sqrt{\mu}P(w)^{-\frac{1}{2}}d_s. \tag{23}$$

If $(x, y, s) \neq (x(\mu), y(\mu), s(\mu))$, then $(\Delta x, \Delta y, \Delta s)$ is non-zero. The new iterate is obtained by taking full NT-steps

$$x^+ := x + \Delta x, \quad y^+ := y + \Delta y, \text{ and } s^+ := s + \Delta s. \tag{24}$$

To measure the distance of an iterate to the corresponding μ-centre, a norm-based proximity measure $\delta(x, s; \mu)$ is introduced

$$\delta(v) := \delta(x, s; \mu) := \frac{1}{2}\left\| v^{-1} - v \right\|_F. \tag{25}$$

One can easily verify that

$$\delta(v) = 0 \Leftrightarrow v = e \Leftrightarrow d_x = d_s = 0 \Leftrightarrow x \circ s = \mu e, \tag{26}$$

which implies that the value of $\delta(v)$ can indeed be considered as a measure for the distance between the given iterate and the corresponding μ-centre.

The τ-neighbourhood of the central path can now be defined as

$$\mathcal{N}(\tau) = \left\{ (x, y, s) \in \mathrm{int}\mathcal{K} \times I\!\!R^m \times \mathrm{int}\mathcal{K} : Ax = b, A^{\mathrm{T}}y + s = c, \delta(v) \leq \tau \right\}. \tag{27}$$

The brief outline of the feasible algorithm given above can now be summarized as follows. At the start of the algorithm, we choose a strictly feasible point (x^0, y^0, s^0) and $\mu^0 = \frac{\langle x^0, s^0 \rangle}{r}$ such that $\delta(x^0, s^0; \mu^0) \leq \tau$ with $0 < \tau < 1$. Using Newton's method with NT scaling, the NT-search directions are calculated. A new iterate (x, y, s) is constructed by taking a full NT-step. Then, μ is reduced by the factor $1 - \theta$ with $0 < \theta < 1$. The appropriate choice of parameters τ and θ is crucial to assure that the new iterate stays in the τ-neighbourhood (27) of the central path. This process is

repeated until μ is small enough, say until $r\mu < \varepsilon$. At this stage an ε-approximate optimal solution of SO has been found.

The generic full NT-step feasible IPM is as follows:

> **Input:**
>
>> A threshold parameter $0 < \tau < 1$;
>>
>> an accuracy parameter $\varepsilon > 0$;
>>
>> a fixed barrier update parameter $0 < \theta < 1$;
>>
>> a strictly feasible (x^0, y^0, s^0) and $\mu^0 = \frac{\langle x^0, s^0 \rangle}{r}$ such that $\delta(x^0, s^0; \mu^0) < \tau$.
>
> **begin**
>> $x := x^0$; $y := y^0$; $s := s^0$; $\mu := \mu^0$.
>>
>> **while** $r\mu \geq \varepsilon$ **do**
>>
>> **begin**
>>> solve the system (22) and use (23) to obtain $(\Delta x, \Delta y, \Delta s)$;
>>>
>>> update $(x, y, s) := (x, y, s) + (\Delta x, \Delta y, \Delta s)$;
>>>
>>> $\mu := (1 - \theta)\mu$.
>> **end**
> **end**

3.3 Analysis of the Full NT-Step Feasible IPM

The condition for strict feasibility of the full NT-step is established in the lemma below.

Lemma 3.1 (Lemma 3.3 in [8]) *Let* $\delta := \delta(x, s; \mu) < 1$. *Then the full NT-step is strictly feasible.*

The following lemma shows that, after a full NT-step, the duality gap assumes the same value as at the μ-centres, namely $r\mu$.

Lemma 3.2 (Lemma 3.4 in [8]) *After a full NT-step, the duality gap is given by*

$$\langle x^+, s^+ \rangle = r\mu.$$

Lemma 3.3 (Proposition 5.6 in [31]) *One has*

$$v^+ \sim \left(P(v + d_x)^{\frac{1}{2}} (v + d_s) \right)^{\frac{1}{2}}.$$

Symbol $\sim$ denotes the similarity of elements in EJA. Recall, two elements x and s in $\mathcal{V}$ are similar, denoted by $x \sim s$, if they share the same set of eigenvalues, including their multiplicities. Furthermore, v^+ represents the variance vector after a full NT-step, which, according to (20), is given by

$$v^+ := \frac{P(w^+)^{-\frac{1}{2}} x^+}{\sqrt{\mu}} \left[= \frac{P(w^+)^{\frac{1}{2}} s^+}{\sqrt{\mu}} \right], \tag{28}$$

where w^+ is the NT-scaling point of x^+ and s^+.

The following technical lemma is a slight modification of Lemma 30 in [11].

Lemma 3.4 *Let $x, s \in \text{int}\,\mathcal{K}$. Then*

$$\|x \circ s\|_F \geq \|P(x)^{\frac{1}{2}} s\|_F.$$

Proof Corollary 2.1 implies that $L(x)$ and $L(x^2)$ commute and the matrix

$$L\left(x^2\right) - L(x)^2 = L(x)^2 - \left(2L(x)^2 - L\left(x^2\right)\right) = L(x)^2 - P(x)$$

has the eigenvalues

$$\left(\frac{\lambda_i + \lambda_j}{2}\right)^2 - \lambda_i \lambda_j \geq 0, \quad 1 \leq i \leq j \leq r.$$

Thus, $L\left(x^2\right) \succeq L(x)^2$ and we have

$$\begin{aligned}
\|x \circ s\|_F^2 &= \langle x \circ s, x \circ s \rangle = \langle s, L(x)^2 s \rangle = \langle s, 2L(x)^2 s \rangle - \langle s, L(x)^2 s \rangle \\
&\geq \langle s, 2L(x)^2 s \rangle - \langle s, L(x^2)s \rangle = \langle s, (2L(x)^2 - L(x^2))s \rangle \\
&= \langle s, P(x)s \rangle = \langle P(x)^{\frac{1}{2}} s, P(x)^{\frac{1}{2}} s \rangle = \|P(x)^{\frac{1}{2}} s\|_F^2.
\end{aligned}$$

This completes the proof. $\square$

Corollary 3.1 *Let $x, s \in \text{int}\,\mathcal{K}$. Then*

$$\|(x \circ s)^{-\frac{1}{2}}\|_F \leq \|(P(x)^{\frac{1}{2}} s)^{-\frac{1}{2}}\|_F.$$

The key part in proving the quadratic convergence is to investigate the effect on proximity measure $\delta(x, s; \mu)$ of a full NT-step to the target point $(x(\mu), y(\mu), s(\mu))$. For this purpose, Gu et al. [8] extended Theorem II.50 in [17] for LO to the SO case. This result is stated in the theorem below.

Theorem 3.1 (Theorem 3.6 in [8]) *Let $\delta := \delta(x, s; \mu) < 1$. Then*

$$\delta(x^+, s^+; \mu) \leq \frac{\delta^2}{\sqrt{2(1 - \delta^2)}}.$$

Furthermore, if $\delta \leq \frac{1}{\sqrt{2}}$, then $\delta(x^+, s^+; \mu) \leq \delta^2$.

In what follows, we derive a sharper quadratic convergence result than the one mentioned above. Our derivation is based on the generalization of Theorem II.52 in [17] for LO. This leads to a wider quadratic convergence neighbourhood of the central path for the algorithm presented in this paper.

Theorem 3.2 *Let $\delta := \delta(x, s; \mu) < 1$. Then, the full NT-step is strictly feasible and*

$$\delta(x^+, s^+; \mu) \leq \frac{\delta^2}{\sqrt{2(1 - \delta^4)}}.$$

Proof Let

$$u := P(v + d_x)^{\frac{1}{2}}(v + d_s) \text{ and } \bar{u} := (v + d_x) \circ (v + d_s).$$

Then, it follows from Lemma 3.3 that $v^+ \sim u^{\frac{1}{2}}$. Furthermore, we have

$$\bar{u} = v^2 + v \circ (d_x + d_s) + d_x \circ d_s = v^2 + v \circ (v^{-1} - v) + d_x \circ d_s = e + d_x \circ d_s.$$

From (28) and Lemma 3.2, we have

$$\|u^{\frac{1}{2}}\|_F^2 = \|v^+\|_F^2 = \langle v^+, v^+\rangle = \frac{1}{\mu}\langle x^+, s^+\rangle = r.$$

Corollary 3.1 implies that

$$\|u^{-\frac{1}{2}}\|_F^2 \leq \|\bar{u}^{-\frac{1}{2}}\|_F^2.$$

Thus, we have

$$4\delta(x^+, s^+; \mu)^2 = \|(v^+)^{-1} - v^+\|_F^2 = \|u^{-\frac{1}{2}} - u^{\frac{1}{2}}\|_F^2 = \|u^{-\frac{1}{2}}\|_F^2 + \|u^{\frac{1}{2}}\|_F^2 - 2r$$
$$= \|u^{-\frac{1}{2}}\|_F^2 - r \leq \|\bar{u}^{-\frac{1}{2}}\|_F^2 - r = \langle e, (e + d_x \circ d_s)^{-1} - e\rangle.$$

Application of the Lemma 2.5 to the last expression above with $u = d_x$ and $v = d_s$ yields the result of the theorem, since $\|d_x + d_s\| = 2\delta$, with $\delta < 1$. This completes the proof. □

Corollary 3.2 *Let $\delta := \delta(x, s; \mu) \leq \frac{1}{\sqrt[4]{2}}$. Then the full NT-step is strictly feasible and*

$$\delta(x^+, s^+; \mu) \leq \delta^2.$$

The corollary above shows the quadratic convergence of the full NT-step to the target μ-centre $(x(\mu), y(\mu), s(\mu))$ in the wider neighbourhood determined by $1/\sqrt[4]{2}$ as opposed to $1/\sqrt{2}$ in [8].

In the previous theorem, the μ was kept fixed and (x, y, s) was updated to (x^+, y^+, s^+). In the following theorem, the effect on the proximity measure is investigated when (x, y, s) is kept fixed and μ is updated to $\mu^+ = (1 - \theta)\mu$. The theorem is a generalization of Lemma II.54 in [17].

Theorem 3.3 *Let* $\delta := \delta(x, s; \mu) < 1$ *and* $\mu^+ = (1 - \theta)\mu$ *with* $0 < \theta < 1$. *Then*

$$\delta(x, s; \mu^+)^2 = (1 - \theta)\delta^2 + \frac{r\theta^2}{4(1 - \theta)}.$$

Proof From (20), we have $\mu \mathrm{tr}(v^2) = \mathrm{tr}(x \circ s)$. This shows that $\mathrm{tr}(v^2) = \mathrm{tr}(e) = r$. Therefore, v and $v^{-1} - v$ are orthogonal with respect to the trace inner product. Thus, we have

$$4\delta(x, s; \mu^+)^2 = \left\| \sqrt{1 - \theta} v^{-1} - \frac{v}{\sqrt{1 - \theta}} \right\|_F^2 = \left\| \sqrt{1 - \theta}(v^{-1} - v) + \frac{\theta v}{\sqrt{1 - \theta}} \right\|_F^2$$

$$= (1 - \theta)\|(v^{-1} - v)\|_F^2 + \frac{\theta^2 \|v\|_F^2}{1 - \theta} = 4(1 - \theta)\delta^2 + \frac{r\theta^2}{1 - \theta}.$$

This completes the proof.   □

Corollary 3.3 *Let* $\delta := \delta(x, s; \mu) \leq \frac{1}{\sqrt[4]{2}}$ *and* $\theta = \frac{1}{\sqrt{2r}}$ *with* $r \geq 2$. *Then*

$$\delta(x^+, s^+; \mu^+) \leq \frac{1}{\sqrt[4]{2}}.$$

Proof From Corollary 3.2 and the fact that $\delta \leq \frac{1}{\sqrt[4]{2}}$, we have

$$\delta(x^+, s^+; \mu)^2 \leq (\delta^2)^2 = \delta^4 \leq \frac{1}{\sqrt{2}}.$$

Then, after the barrier parameter is updated to $\mu^+ = (1 - \theta)\mu$, with $\theta = \frac{1}{\sqrt{2r}}$, Theorem 3.3 yields the following upper bound for $\delta(x^+, s^+; \mu^+)^2$:

$$\delta(x^+, s^+; \mu^+)^2 \leq \frac{1 - \theta}{2} + \frac{1}{8(1 - \theta)} \leq \max \left\{ \frac{5}{8}, \frac{1}{2} \right\} = \frac{5}{8}.$$

The last inequality holds due to the fact that the left-hand side is a convex function of θ, whose values are 5/8 and 1/2 at $\theta = 0$ and $\theta = 1/2$, respectively. Since $\theta \in [0, 1/2]$, the left-hand side does not exceed 5/8. Note that $5/8 < 1/\sqrt{2}$. Hence, the desired inequality follows

$$\delta \left(x^+, s^+; \mu^+ \right) \leq \frac{1}{\sqrt[4]{2}}.$$

This completes the proof.   □

The consequence of the above results is that if we set the values of the threshold parameter and the barrier update parameter to

$$\tau = \frac{1}{\sqrt[4]{2}} \text{ and } \theta = \frac{1}{\sqrt{2r}}, \tag{29}$$

the algorithm will generate a sequence of iterates that are not only feasible; they, in addition, always lay in the τ-neighbourhood (27) of the central path. Moreover, the duality gap will be reduced according to $\langle x^+, s^+ \rangle = r\mu$. Hence, the algorithm is well defined and globally convergent.

Similar to the proof of Lemma 4.7 in [12], we can easily verify the validity of the following lemma.

Lemma 3.5 *Let x^0 and s^0 are strictly feasible, $\mu^0 = \frac{\langle x^0, s^0 \rangle}{r} \leq \frac{1}{\sqrt[4]{2}}$. Moreover, let (x^k, y^k, s^k) be the k-th iterate of the algorithm. Then the inequality $\langle x^k, s^k \rangle \leq \varepsilon$ is satisfied if*

$$k \geq \frac{1}{\theta} \log \frac{\langle x^0, s^0 \rangle}{\varepsilon}.$$

From Lemma 3.5, we have the following theorem which provides an upper bound for the total number of the iterations produced by the algorithm.

Theorem 3.4 *Let $\tau = \frac{1}{\sqrt[4]{2}}$ and $\theta = \frac{1}{\sqrt{2r}}$ with $r \geq 2$. Then the algorithm requires*

$$\sqrt{2r} \log \frac{\langle x^0, s^0 \rangle}{\varepsilon}$$

iterations to obtain an iterate (x, y, s) satisfying $\langle x, s \rangle \leq \varepsilon$ which is an ε-approximate optimal solution of (SOP) *and* (SOD).

Thus, the feasible algorithm is well defined, globally convergent, and achieves quadratic convergence of full NT-steps in the wider neighbourhood while still maintaining the best known iteration bound known for these types of methods, namely

$$O\left(\sqrt{r} \log \frac{r\mu_0}{\varepsilon}\right).$$

Although the neighbourhood is wider, it is not substantially wider, hence, the algorithm is still a small-update method.

4 Conclusions

In this paper, we have presented an improved convergence and complexity analysis of the full NT-step feasible IPM for SO discussed in [8]. It has been shown that the iterates in this modified version of the algorithm can be taken in the wider quadratic convergence neighbourhood of the central path characterized by $1/\sqrt[4]{2}$ as opposed to $1/\sqrt{2}$ in [8]. However, the currently best iteration bound known for feasible algorithm of the type described in the paper is still achieved. For the analysis of the algorithm, some new results on EJAs has to be developed including Lemmas 2.3 and 2.5 as two important ones. Moreover, these results are interesting in their own right.

Some interesting topics for further research include the design and analysis of a full NT-step infeasible IPM with polynomial iteration bound and the generalization of these algorithms to the Cartesian $P_*(\kappa)$-linear complementarity problem over symmetric cones [32,33].

References

1. Zhang, Y.: On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming. SIAM J. Optim. **8**(2), 365–386 (1998)
2. Potra, F.A., Sheng, R.Q.: On homogeneous interior-point algorithms for semidefinite programming. Optim. Methods Softw. **9**(1–3), 161–184 (1998)
3. Potra, F.A., Sheng, R.: A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming. SIAM J. Optim. **8**(4), 1007–1028 (1998)
4. De Klerk, E.: Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications. Kluwer Academic Publishers, Dordrecht, The Netherlands (2002)
5. Alizadeh, F., Goldfard, D.: Second order cone optimization. Math. Program. **95**(1), 3–51 (2003)
6. Güler, O.: Barrier functions in interior-point methods. Math. Oper. Res. **21**(4), 860–885 (1996)
7. Faybusovich, L.: Euclidean Jordan algebras and interior-point algorithms. Positivity **1**, 331–35 (1997)
8. Gu, G., Zangiabadi, M., Roos, C.: Full Nesterov–Todd step infeasible interior-point method for symmetric optimization. Eur. J. Oper. Res. **214**(3), 473–484 (2011)
9. Muramatsu, M.: On a commutative class of search directions for linear programming over symmetric cones. J. Optim. Theory Appl. **112**(3), 595–625 (2002)
10. Rangarajan, B.K.: Polynomial convergence of infeasible-interior-point methods over symmetric cones. SIAM J. Optim. **16**(4), 1211–1229 (2006)
11. Schmieta, S.H., Alizadeh, F.: Extension of primal-dual interior-point algorithms to symmetric cones. Math. Program. **96**(3), 409–438 (2003)
12. Wang, G.Q., Bai, Y.Q.: A new full Nesterov–Todd step primal-dual path-following interior-point algorithm for symmetric optimization. J. Optim. Theory Appl. **154**(3), 966–985 (2012)
13. Liu, C.H., Liu, H.W., Liu, X.Z.: Polynomial convergence of second-order mehrotra type predictor–corrector algorithms over symmetric cones. J. Optim. Theory Appl. **154**(3), 949–965 (2012)
14. Liu, H.W., Yang, X.M., Liu, C.H.: A new wide neighbourhood primal-dual infeasible-interior-point method for symmetric cone programming. J. Optim. Theory Appl. **158**(3), 796–815 (2013)
15. Kheirfam, B.: A corrector–predictor path-following method for convex quadratic symmetric cone optimization. J. Optim. Theory Appl. (2014). doi:10.1007/s10957-014-0554-2
16. Anjos, M.F., Lasserre, J.B.: Handbook on semidefinite, conic and polynomial optimization: theory, algorithms, software and applications. In: International Series in Operational Research and Management Science, Vol. 166. Springer, New York, USA (2012)
17. Roos, C., Terlaky, T., Vial, J.-Ph.: Theory and Algorithms for Linear Optimization (1st Edition, Theory and Algorithms for Linear Optimization. An Interior-Point Approach. John Wiley & Sons, Chichester, UK, 1997). Springer, New York, USA (2005)
18. Roos, C.: A full-Newton step $O(n)$ infeasible interior-point algorithm for linear optimization. SIAM J. Optim. **16**(4), 1110–1136 (2006)
19. Gu, G., Mansouri, H., Zangiabadi, M., Bai, Y.Q., Roos, C.: Improved full-Newton step $O(nL)$ infeasible interior-point method for linear optimization. J. Optim. Theory Appl. **145**(2), 271–288 (2010)
20. Darvay, Z.: New interior-point algorithms in linear programming. Adv. Model. Optim. **5**(1), 51–92 (2003)
21. Mansouri, H., Roos, C.: A new full-Newton step $O(n)$ infeasible interior-point algorithm for semidefinite optimization. Numer. Algorithms **52**(2), 225–255 (2009)

22. Zhang, L.P., Xu, Y.H.: A full-Newton step interior-point algorithm based on modified Newton direction. Oper. Res. Lett. **39**(5), 318–322 (2011)

23. Zangiabadi, M., Gu, G., Roos, C.: Full Nesterov–Todd step primal-dual interior-point methods for second-order cone optimization. J. Optim. Theory Appl. **158**(3), 816–858 (2013)

24. Wang, G.Q., Bai, Y.Q., Gao, X.Y., Wang, D.Z.: Improved complexity analysis of full Nesterov–Todd step interior-point methods for semidefinite optimization. J. Optim. Theory Appl. (2014). doi:10.1007/s10957-014-0619-2

25. Faraut, J., Korányi, A.: Analysis on Symmetric Cones. Oxford University Press, New York (1994)

26. Faybusovich, L.: A Jordan-algebraic approach to potential-reduction algorithms. Math. Z. **239**(1), 117–129 (2002)

27. Gowda, M.S., Tao, J., Moldovan, M.: Some inertia theorems in Euclidean Jordan algebras. Linear Alg. Appl. **430**(8–9), 1992–2011 (2009)

28. Luo, Z.Q., Sturm, J.F., Zhang, Y.: Conic convex programming and self-dual embedding. Optim. Methods Softw. **14**(3), 169–218 (2000)

29. Nesterov, Y.E., Todd, M.J.: Self-scaled barriers and interior-point methods for convex programming. Math. Oper. Res. **22**(1), 1–42 (1997)

30. Nesterov, Y.E., Todd, M.J.: Primal-dual interior-point methods for self-scaled cones. SIAM J. Optim. **8**(2), 324–364 (1998)

31. Vieira, M.V.C.: Interior-point methods based on kernel functions for symmetric optimization. Optim. Methods Softw. **27**(3), 513–537 (2012)

32. Wang, G.Q., Bai, Y.Q.: A class of polynomial interior-point algorithms for the Cartesian P-matrix linear complementarity problem over symmetric cones. J. Optim. Theory Appl. **152**(3), 739–772 (2012)

33. Wang, G.Q., Lesaja, G.: Full Nesterov–Todd step feasible interior-point method for the Cartesian $P_*(\kappa)$-SCLCP. Optim. Methods Softw. **28**(3), 600–618 (2013)