A New Full Nesterov–Todd Step Primal–Dual Path-Following Interior-Point Algorithm for Symmetric Optimization

G.Q. Wang · Y.Q. Bai

Abstract In this paper, we generalize a primal–dual path-following interior-point algorithm for linear optimization to symmetric optimization by using Euclidean Jordan algebras. The proposed algorithm is based on a new technique for finding the search directions and the strategy of the central path. At each iteration, we use only full Nesterov–Todd steps. Moreover, we derive the currently best known iteration bound for the small-update method. This unifies the analysis for linear, second-order cone, and semidefinite optimizations.

1 Introduction

Symmetric optimization (SO) problem is a convex optimization problem that minimizes a linear function over the intersection of an affine subspace and a symmetric cone. This class of optimization problem includes linear optimization (LO), second-order cone optimization (SOCO), and semidefinite optimization (SDO) as special cases. In the past two decades, the SDO problems have been one of the most active research areas in mathematical programming. Nesterov and Todd [1, 2] provided a theoretical foundation for efficient primal–dual interior-point methods (IPMs) for convex

Communicated by Florian A. Potra.

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

Y.Q. Bai
Department of Mathematics, Shanghai University, Shanghai 200444, China
e-mail: yqbai@shu.edu.cn

programming problems, which includes LO, SOCO, and SDO. Potra and Sheng [3] and Kojima, Shida and Shindoh [4] independently investigated the superlinear convergence of primal–dual infeasible-interior-point algorithms for SDO. Ji, Potra, and Sheng [5] studied the local behavior of the predictor–corrector algorithm considered by Monteiro [6] for SDO using the Monteiro–Zhang family of search directions. Potra and Sheng [7] first analyzed the homogeneous embedding idea for SDO. Luo, Sturm, and Zhang [8] proposed a self-dual embedding model for conically constrained convex optimization, including SDO. For an overview of these related results, we refer to the subject monographs [9, 10] and references within.

In the research field of Lie algebras, it is well-known that a symmetric cone has a deep connection with the Euclidean Jordan algebras. The first work connecting Jordan algebras and optimization is due to Güler [11]. He observed that the family of the self-scaled cones is identical to the set of symmetric cones for which there exists a complete classification theory. Faybusovich [12] first extended primal–dual IPMs for SDO to SO by using Euclidean Jordan algebras. Muramatsu [13] presented a commutative class of search directions for LO over symmetric cones, and analyzed the complexities of primal–dual IPMs for SO. Rangarajan [14] proved the polynomial-time convergence of infeasible IPMs for conic programming over symmetric cones using a wide neighborhood of the central path for a commutative family of search directions. Schmieta and Alizadeh [15] presented a way to transfer the Jordan algebra associated with the second-order cone into the so-called Clifford algebra in the cone of matrices and then carried out a unified method of the analysis for many IPMs in symmetric cones. Vieira [16] proposed primal–dual IPMs for SO based on the so-called eligible kernel functions and obtained the currently best known iteration bounds for the large- and small-update methods.

Recently, Darvay [17] proposed a full-Newton step primal–dual path-following interior-point algorithm for LO. The search direction of his algorithm is introduced by using an algebraic equivalent transformation (form) of the nonlinear equations which define the central path and then applying Newton's method for the new system of equations. Later on, Achache [18] and Wang et al. [19, 20] respectively extended Darvay's algorithm for LO to convex quadratic optimization (CQO), SOCO, and SDO. By constructing strictly feasible iterates for a sequence of the perturbed problems of the given problem and its dual problem, Roos [21] presented a full-Newton step primal–dual infeasible interior-point algorithm for LO based on the classical search direction. Using Euclidean Jordan algebras, Gu et al. [22] extended Roos's algorithm for LO to SO. The order of the iteration bound coincides with the bound derived for LO which is the currently best known iteration bound for SO.

The purpose of the paper is to generalize Darvay's full-Newton step primal–dual path-following interior-point algorithm for LO to SO by using Euclidean Jordan algebras. We develop some new analytic tools and adopt the basic analysis used in Darvay [17] and Wang et al. [19, 20] to the SO case. At each iteration, we use only full Nesterov–Todd steps which have the advantage that no line searches are needed. The currently best known iteration bound for the small-update method is derived. Moreover, our analysis is relatively simple and straightforward to the LO analogue in Darvay [17].

The paper is organized as follows. In Sect. 2, we provide the theory of the Euclidean Jordan algebras and their associated symmetric cones, and develop some new

results that are needed in the analysis of the algorithm. In Sect. 3, after briefly reviewing the concept of the central path for SO, we extend Darvay's technique for LO to SO and obtain the new search directions for SO. The generic primal–dual interior-point algorithm for SO is also presented. In Sect. 4, we analyze the algorithm and derive the currently best known iteration bound for the small-update method. Finally, some conclusions and remarks follow in Sect. 5.

2 Euclidean Jordan Algebras and Their Associated Symmetric Cones

In this section, we provide the theory of the Euclidean Jordan algebras and their associated symmetric cones. This will serve as the basic analytic tool for the analysis of our algorithm presented in Fig. 1. For a comprehensive treatment of Jordan algebras, the reader is referred to the monograph by Faraut and Koranyi [23].

Definition 2.1 Let V be an n-dimensional vector space over $\mathbf{R}$ along with the bilinear map $\circ : (x, y) \mapsto x \circ y \in V$. Then $(V, \circ)$ is a Jordan algebra iff for all $x, y \in V$

(i) $x \circ y = y \circ x$;
(ii) $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$, where $x^2 = x \circ x$.

We define $x^n := x \circ x^{n-1}, n \geq 2$. Jordan algebras are not necessarily associative, but they are power associative, $x^m \circ x^n := x^{m+n}$, i.e., the algebra generated by a single element $x \in V$ is associative (see, e.g., [15]).

An element $e \in V$ is said to be an identity element iff $e \circ x = x \circ e = x$ for all $x \in V$. Note that the identity element e is unique, i.e., if e_1 and e_2 are identity elements of V, then $e_1 = e_1 \circ e_2 = e_2 \circ e_1 = e_2$.

Definition 2.2 A Jordan algebra $(V, \circ)$ over $\mathbf{R}$ with an identity element e is said to be a Euclidean Jordan algebra iff there exists a symmetric, positive definite quadratic form Q on V which is also associative, that is,

$$Q(x \circ y, z) = Q(x, y \circ z).$$

Remark 2.1 In the sequel, we always assume that $(V, \circ)$ is a Euclidean Jordan algebra, and simply denoted as V.

Since "$\circ$" is bilinear for every $x \in V$, there exists a matrix $L(x)$ such that for every y, $x \circ y := L(x)y$. In particular, $L(x)e = x$ and $L(x)x = x^2$. Furthermore, the part (ii) of Definition 2.1 implies that the operators $L(x)$ and $L(x^2)$ commute, i.e., $L(x^2)L(x)y = x^2 \circ (x \circ y) = x \circ (x^2 \circ y) = L(x)L(x^2)y$.

For each $x \in V$, we define

$$P(x) := 2L(x)^2 - L(x^2), \tag{1}$$

where $L(x)^2 := L(x)L(x)$. The map $P(x)$ is called the quadratic representation of V, which is an essential concept in the theory of Jordan algebras.

An element $x \in V$ is called invertible iff there exists a $y = \sum_{i=0}^{k} \alpha_i x^i$ for some finite $k < \infty$ and real numbers α_i such that $x \circ y = y \circ x = e$, and denoted as x^{-1}.

It is well-known that the cone K is said to be a symmetric cone iff it is self-dual and homogeneous. A symmetric cone is also precisely the class of self-scaled cone introduced by Nesterov and Todd [1, 2]. For any Euclidean Jordan algebra V, the corresponding cone of squares

$$K(V) := \{x^2 : x \in V\} \tag{2}$$

is indeed a symmetric cone.

The following theorem gives some major properties of the cone of squares in a Euclidean Jordan algebra.

Theorem 2.1 (Theorem III.2.1, Proposition III.2.2 in [23]) *Let V be a Euclidean Jordan algebra. Then $K(V)$ is a symmetric cone, and is the set of elements x in V for which $L(x)$ is positive semidefinite. Furthermore, if x is invertible, then*

$$P(x) \operatorname{int} K = \operatorname{int} K,$$

where $\operatorname{int} K$ *is the interior of* K.

A symmetric cone K in a Euclidean space V is said to be irreducible iff there do not exist non-trivial subspaces V_1, V_2 and symmetric cones $K_1 \subset V_1$, $K_2 \subset V_2$ such that V is the direct sum of V_1 and V_2, and K the direct sum of K_1 and K_2.

Lemma 2.1 (Proposition II.4.5 in [23]) *Any symmetric cone K is, in a unique way, the direct sum of irreducible symmetric cones.*

A Euclidean Jordan algebra is said to be simple iff it cannot be represented as the orthogonal direct sum of two Euclidean Jordan algebras. Simple Jordan algebras have been classified into five kinds of irreducible symmetric cones. The details can be found in [23].

Since a Jordan algebra V is power associative, we can define the concepts of rank, the minimum and the characteristic polynomials, eigenvalues, trace, and determinant for it in the following way (see, e.g., [15]).

For any $x \in V$, 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 we denote as $\deg(x)$. The rank of V, $\operatorname{rank}(V)$, is the largest $\deg(x)$ of any number $x \in V$. An element $x \in V$ is called regular iff its degree equals the rank of V.

Remark 2.2 In the sequel, unless stated otherwise, we always assume that V is a Euclidean Jordan algebra with $\operatorname{rank}(V) = r$.

For a regular element $x \in V$, since $\{e, x, \ldots, x^r\}$ 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

$$f(\lambda; x) = \lambda^r - a_1(x)\lambda^{r-1} + \cdots + (-1)^r a_r(x) \tag{3}$$

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

An element $c \in V$ is said to be idempotent iff $c \neq 0$ and $c^2 = c$. Two idempotents c_1 and c_2 are said to be orthogonal if $c_1 \circ c_2 = 0$. We say that $\{c_1, \ldots, c_r\}$ is a complete system of orthogonal primitive idempotents, or a Jordan frame, iff each c_i is a primitive idempotent, $c_i \circ c_j = 0, i \neq j$, and $\sum_{i=1}^n c_i = e$. Note that Jordan frames always contain r primitive idempotents, where r is the rank of V.

Theorem 2.2 (Spectral decomposition, Theorem III.1.2 in [23]) *Let $x \in V$. Then there exists a Jordan frame $\{c_1, \ldots, c_r\}$ and real numbers $\lambda_1(x), \ldots, \lambda_r(x)$ such that*

$$x = \sum_{i=1}^n \lambda_i(x) c_i. \tag{4}$$

The numbers $\lambda_i(x)$ (with their multiplicities) are the eigenvalues of x. Furthermore,

$$\mathrm{tr}(x) = \sum_{i=1}^r \lambda_i(x) \quad and \quad \det(x) = \prod_{i=1}^r \lambda_i(x). \tag{5}$$

In fact, the above $\lambda_1(x), \ldots, \lambda_r(x)$ are exactly the roots of the characteristic polynomial $f(\lambda; x)$. Note that, since $e = c_1 + \cdots + c_r$ has eigenvalue 1, with multiplicity r, it follows that $\mathrm{tr}(e) = r$ and $\det(e) = 1$. For any $x \in V$, one can easily verify that

$$x \in K \quad \Longleftrightarrow \quad \lambda_i(x) \geq 0, \quad i = 1, \ldots, r, \tag{6}$$

and

$$x \in \mathrm{int}\, K \quad \Longleftrightarrow \quad \lambda_i(x) > 0, \quad i = 1, \ldots, r. \tag{7}$$

The importance of the spectral decomposition (4) is that it enables us to extend the definition of any real-valued, continuous univariate function $\psi(t)$ to elements of a Euclidean Jordan algebra using eigenvalues.

Throughout the paper, we assume that $\psi(t)$ is a real valued univariate function on $[0, +\infty[$ and differentiable on $]0, +\infty[$ such that $\psi'(t) > 0$ for all $t > 0$. Now we are ready to show how a vector-valued function can be obtained from the function $\psi(t)$.

Definition 2.3 Let $x \in V$ with the spectral decomposition defined by (4). The vector-valued function $\psi(x)$ is defined by

$$\psi(x) := \psi\big(\lambda_1(x)\big) c_1 + \cdots + \psi\big(\lambda_r(x)\big) c_r. \tag{8}$$

Furthermore, replacing $\psi(\lambda_i(x))$ in (8) by $\psi'(\lambda_i(x))$ with $i = 1, \ldots, r$, respectively, we can define the vector-valued function as follows

$$\psi'(x) := \psi'\big(\lambda_1(x)\big) c_1 + \cdots + \psi'\big(\lambda_r(x)\big) c_r. \tag{9}$$

It should be pointed out that $x \circ s = s \circ x$, but, in general, $L(x)L(s) \neq L(s)L(x)$. We say that two elements x and y of V operator commute iff $L(x)L(y) = L(y)L(x)$. In other words, x and y operator commute iff for all $z \in V$, $x \circ (y \circ z) = y \circ (x \circ z)$ (see, e.g., [15]).

Lemma 2.2 (Theorem 27 in [15]) *Let x and s be two elements of a Euclidean Jordan algebra V. Then x and s operator commute if and only if there is a Jordan frame $\{c_1, \ldots, c_r\}$ such that*

$$x = \lambda_1(x)c_1 + \cdots + \lambda_r(x)c_r \quad and \quad s = \lambda_1(s)c_1 + \cdots + \lambda_r(s)c_r.$$

Corollary 2.1 *Let $x, s \in V$, and suppose they share a Jordan frame $\{c_1, \ldots, c_r\}$. Then*

$$x \circ s = \lambda_1(x)\lambda_1(s)c_1 + \cdots + \lambda_r(x)\lambda_r(s)c_r.$$

Proof The desired result follows directly from the fact that $\{c_1, \ldots, c_r\}$ is a complete system of orthogonal primitive idempotents. □

Corollary 2.2 *Let $x, s \in V$ and $x + s = e$. Then x and s operator commute.*

Proof Theorem 2.2 implies that there exists a Jordan frame $\{c_1, \ldots, c_r\}$ such that

$$x = \lambda_1(x)c_1 + \cdots + \lambda_r(x)c_r \quad and \quad e = c_1 + \cdots + c_r.$$

Thus, we get

$$s = e - x = \big(1 - \lambda_1(x)\big)c_1 + \cdots + \big(1 - \lambda_r(x)\big)c_r.$$

This implies the desired result. □

Definition 2.4 Let $x, s \in V$. The elements x and s are approximately equivalent, and briefly denoted as $x \approx s$, iff x and s operator commute, and the absolute value of the difference of their respective corresponding eigenvalues is small enough.

Theorem 2.3 *Let $x, s \in V$, and suppose x and s operator commute. Then*

$$\lambda_i(x + s) = \lambda_i(x) + \lambda_i(s), \quad i = 1, \ldots, r.$$

Furthermore, if each $|\lambda_i(s)|$ with $i = 1, \ldots, r$ is small enough, one has

$$\psi(x + s) \approx \psi(x) + \psi'(x) \circ s.$$

Proof Since $x, s \in V$, Theorem 2.2 implies that there exists a Jordan frame $\{c_1, \ldots, c_r\}$ such that

$$x = \lambda_1(x)c_1 + \cdots + \lambda_r(x)c_r \quad and \quad s = \lambda_1(s)c_1 + \cdots + \lambda_r(s)c_r.$$

Hence, we have

$$x + s = \big(\lambda_1(x) + \lambda_1(s)\big)c_1 + \cdots + \big(\lambda_r(x) + \lambda_r(s)\big)c_r.$$

This implies that

$$\lambda_i(x + s) = \lambda_i(x) + \lambda_i(s), \quad i = 1, \ldots, r.$$

On the other hand, since $\psi(t)$ is a differentiable, we know that, if each $|\lambda_i(s)|$ with $i = 1, \ldots, r$ is small enough, then

$$\psi\big(\lambda_i(x) + \lambda_i(s)\big) \approx \psi\big(\lambda_i(x)\big) + \psi'\big(\lambda_i(x)\big)\lambda_i(s), \quad i = 1, \ldots, r.$$

Hence, we have

$$
\begin{aligned}
\psi(x+s) \\
&= \psi\big(\lambda_1(x+s)\big)c_1 + \cdots + \psi\big(\lambda_r(x+s)\big)c_r \\
&= \psi\big(\lambda_1(x)+\lambda_1(s)\big)c_1 + \cdots + \psi\big(\lambda_r(x)+\lambda_r(s)\big)c_r \\
&\approx \big(\psi\big(\lambda_1(x)\big)+\psi'\big(\lambda_1(x)\big)\lambda_1(s)\big)c_1 + \cdots + \big(\psi\big(\lambda_r(x)\big)+\psi'\big(\lambda_r(x)\big)\lambda_r(s)\big)c_r \\
&= \psi\big(\lambda_1(x)\big)c_1 + \cdots + \psi\big(\lambda_r(x)\big)c_r + \psi'\big(\lambda_1(x)\big)\lambda_1(s)c_1 + \cdots + \psi'\big(\lambda_r(x)\big)\lambda_r(s)c_r \\
&= \psi(x) + \psi'(x)\circ s.
\end{aligned}
$$

The last equality holds due to Corollary 2.1. This completes the proof. □

Corollary 2.3 *Let $x, s \in V$ and $x+s = e$. If $|\lambda_i(s)|$ with $i = 1, \ldots, r$ is small enough, one has*

$$
\psi(x+s) \approx \psi(x) + \psi'(x)\circ s.
$$

Proof The corollary follows immediately from Corollary 2.2 and Theorem 2.3. □

For any $x, s \in V$, we define

$$
\langle x, s\rangle := \mathrm{tr}(x \circ s). \tag{10}
$$

and refer to it as the trace inner product. The Frobenius norm induced by this trace inner product, namely $\|\cdot\|_F$, is defined by

$$
\|x\|_F := \sqrt{\langle x, x\rangle}. \tag{11}
$$

Then

$$
\|x\|_F = \sqrt{\mathrm{tr}(x^2)} = \sqrt{\sum_{i=1}^{r} \lambda_i^2(x)}. \tag{12}
$$

Furthermore, we have

$$
\big|\lambda_{\max}(x)\big| \le \|x\|_F \quad \text{and} \quad \big|\lambda_{\min}(x)\big| \le \|x\|_F, \tag{13}
$$

where $\lambda_{\max}(x)$ and $\lambda_{\min}(x)$ denote the largest and the smallest eigenvalue of x, respectively.

Lemma 2.3 *Let $x \in V$. Then*

$$
\lambda_{\max}(x)e - x \in K \quad \text{and} \quad x - \lambda_{\min}(x)e \in K.
$$

Proof Theorem 2.2 implies that there exists a Jordan frame $\{c_1, \ldots, c_r\}$ such that

$$
x = \sum_{i=1}^{r} \lambda_i(x)c_i \quad \text{and} \quad e = \sum_{i=1}^{r} c_i.
$$

Hence, we have

$$
\lambda_{\max}(x)e - x = \lambda_{\max}(x)\sum_{i=1}^{r} c_i - \sum_{i=1}^{r} \lambda_i(x)c_i = \sum_{i=1}^{r}\big(\lambda_{\max}(x)-\lambda_i(x)\big)c_i.
$$

It follows from (6) that

$$\lambda_{\max}(x)e - x \in K.$$

The proof of the second part of the lemma follows in a similar way. □

Corollary 2.4 *Let $x \in V$ and $\|x\|_F \leq 1$. Then*

$$e - x \in K.$$

Proof Since $\lambda_{\max}(x) \leq \|x\|_F$, the corollary follows immediately from Lemma 2.3. □

Lemma 2.4 *Let $x \in V$ and $e - x \circ x \in K$. Then*

$$0 \leq \|e - x \circ x\|_F^2 \leq r.$$

Proof Theorem 2.1 implies that $\lambda_i(x \circ x) \geq 0$ for $i = 1, \ldots, r$. Thus, we have

$$0 \leq \|e - x \circ x\|_F^2 = \sum_{i=1}^{r} \lambda_i^2(e - x \circ x) = \sum_{i=1}^{r} \bigl(1 - \lambda_i(x \circ x)\bigr)^2 \leq r.$$

The last inequality holds due to the fact that $0 \leq 1 - \lambda_i(x \circ x) \leq 1$. This completes the proof. □

The following well-known inequalities are needed in the analysis of the interior-point algorithm presented in Fig. 1.

Lemma 2.5 (Lemma 2.12 in [22]) *Let $x \in V$. Then*

$$\bigl\|x^2\bigr\|_F \leq \|x\|_F^2.$$

Lemma 2.6 (Lemma 14 in [15]) *Let $x, s \in V$. Then*

$$\lambda_{\min}(x + s) \geq \lambda_{\min}(x) + \lambda_{\min}(s) \geq \lambda_{\min}(x) - \|s\|_F$$

and

$$\lambda_{\max}(x + s) \leq \lambda_{\max}(x) + \lambda_{\max}(s) \leq \lambda_{\max}(x) + \|s\|_F.$$

Lemma 2.7 (Lemma 30 in [15]) *Let $x, s \in K$. Then*

$$\bigl\| P(x)^{1/2}s - e \bigr\|_F \leq \|x \circ s - e\|_F.$$

Lemma 2.8 (Theorem 4 in [24]) *Let $x, s \in K$. Then*

$$\lambda_{\min}\bigl(P(x)^{1/2}s\bigr) \geq \lambda_{\min}(x \circ s).$$

Recall that two matrices X and S are said to be similar iff they share the same eigenvalues, including their multiplicities. In this case, we write $X \sim S$. Analogously, we say that two elements x and s in V are similar, and briefly denoted as $x \sim s$, iff x and s share the same eigenvalues, including their multiplicities (see, e.g., [16, 22]).

Lemma 2.9 *Let $x, s \in V$ and $x \sim s$. Then*

$$\det(x) = \det(s) \quad and \quad \mathrm{tr}(x) = \mathrm{tr}(s).$$

Furthermore, one has

$$x \in K \ (\mathrm{int}\, K) \quad \Longleftrightarrow \quad s \in K \ (\mathrm{int}\, K).$$

Proof The desired results follow directly from Theorem 2.2. □

The following lemma gives the so-called NT-scaling of V, which was first established by Faybusovich [25] in the framework of Euclidean Jordan algebras.

Lemma 2.10 (NT-scaling, Lemma 3.2 in [25]) *Let $x, s \in \mathrm{int}\, K$. Then there exists a unique $w \in \mathrm{int}\, K$ such that*

$$x = P(w)s.$$

Moreover,

$$w = P(x)^{\frac{1}{2}} \big(P\big(x^{\frac{1}{2}}\big)s \big)^{-\frac{1}{2}} \big[= P\big(s^{-\frac{1}{2}}\big) \big(P\big(s^{\frac{1}{2}}\big)x \big)^{\frac{1}{2}} \big].$$

The point w is called the scaling point of x and s (in this order).

As a consequence, there exists $v \in \mathrm{int}\, K$ such that

$$v = P(w)^{-\frac{1}{2}} x = P(w)^{\frac{1}{2}} s. \tag{14}$$

Note that $P(w)^{\frac{1}{2}}$ and its inverse $P(w)^{-\frac{1}{2}}$ are automorphisms of $\mathrm{int}\, K$ (see, e.g., [22]).

Lemma 2.11 *Let $t > 0$ and $v \in K$. Then*

$$\|te - v\|_F \le \frac{1}{t + \lambda_{\min}(v)} \|t^2 e - v \circ v\|_F.$$

Proof It follows from Corollary 2.1 and Theorem 2.3 that

$$
\begin{aligned}
\|te - v\|_F &= \big\| \big((te + v)^{-1} \circ (te + v) \big) \circ (te - v) \big\|_F \\
&= \big\| (te + v)^{-1} \circ \big((te + v) \circ (te - v) \big) \big\|_F \\
&= \big\| (te + v)^{-1} \circ \big(t^2 e - v \circ v \big) \big\|_F \\
&= \sqrt{ \sum_{i=1}^{n} \frac{\lambda_i^2 (t^2 e - v \circ v)}{\lambda_i^2 (te + v)} } \\
&\le \frac{1}{\lambda_{\min}(te + v)} \sqrt{ \sum_{i=1}^{n} \lambda_i^2 \big(t^2 e - v \circ v \big) } \\
&= \frac{1}{t + \lambda_{\min}(v)} \big\| \big(t^2 e - v \circ v \big) \big\|_F.
\end{aligned}
$$

This completes the proof. □

3 Primal–Dual Path-Following Interior-Point Algorithm for SO

In this section, after reviewing the concept of the central path for SO, we extend Darvay's technique for LO to SO and obtain the new search directions for SO. The generic primal–dual path-following interior-point algorithm for SO is also presented.

3.1 The Central Path for SO

We consider the SO problem given in the standard form

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

and its dual problem

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

where c and the rows of A lie in V, and $b \in \mathbf{R}^m$. We call (SOP) feasible iff there exists $x \in K$ such that $Ax = b$, and strictly feasible, if in addition, $x \in \text{int } K$. Similarly, we call (SOD) feasible iff there exists $(y, s) \in \mathbf{R}^m \times K$ such that $A^T y + s = c$, and strictly feasible, if in addition, $s \in \text{int } K$. Throughout the paper, we assume that the matrix A has full rank, i.e., $\text{rank}(A) = m$.

Throughout the paper, we assume that both (SOP) and (SOD) satisfy the interior-point condition (IPC), i.e., both (SOP) and (SOD) are strictly feasible. This can be achieved via the so-called homogeneous self-dual embedding (see, e.g., [7, 24]). Under the IPC, finding an optimal solution of (SOP) and (SOD) is equivalent to solving the following system

$$
\begin{aligned}
Ax = b, & \quad x \in K, \\
A^T y + s = c, & \quad s \in K, \\
x \circ s = 0. &
\end{aligned}
\tag{15}
$$

The basic idea of primal–dual IPMs is to replace the third equation in (15), the so-called complementarity condition for (SOP) and (SOD), by the parameterized equation $x \circ s = \mu e$ with $\mu > 0$. The system (15) can be written as

$$
\begin{aligned}
Ax = b, & \quad x \in K, \\
A^T y + s = c, & \quad s \in K, \\
x \circ s = \mu e. &
\end{aligned}
\tag{16}
$$

Since the IPC holds and A has full rank, the parameterized system (16) has a unique solution $(x(\mu), y(\mu), s(\mu))$ for each $\mu > 0$ [12, 16], and we call $x(\mu)$ the μ-center of (SOP) and $(y(\mu), s(\mu))$ the μ-center of (SOD). The set of μ-centers gives a homotopy path (with μ running through all the 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 $x \circ s = 0$, it naturally yields an optimal solution for (SOP) and (SOD) (see, e.g., [12, 16]).

3.2 The New Search Directions for SO

Similarly to the LO case [17], we replace the standard centering equation $x \circ s = \mu e$ by $\psi(\frac{x \circ s}{\mu}) = \psi(e)$, where $\psi(\cdot)$ is the vector-valued function induced by the univariate function $\psi(t)$, Then, we consider the following system

$$
\begin{aligned}
Ax &= b, \quad x \in K, \\
A^T y + s &= c, \quad s \in K, \\
\psi\left(\frac{x \circ s}{\mu}\right) &= \psi(e).
\end{aligned}
\tag{17}
$$

A direct application of Newton's method is to solve the nonlinear system (17) with μ fixed. For any strictly feasible $x \in \text{int } K$ and $s \in \text{int } K$, we want to find displacements $(\Delta x, \Delta y, \Delta s)$ such that

$$
\begin{aligned}
A\Delta x &= 0, \\
A^T \Delta y + \Delta s &= 0, \\
\psi\left(\frac{(x + \Delta x) \circ (s + \Delta s)}{\mu}\right) &= \psi(e).
\end{aligned}
\tag{18}
$$

The third equation of the system (18) is equivalent to

$$
\psi\left(\frac{x \circ s}{\mu} + \frac{x \circ \Delta s + \Delta x \circ s + \Delta x \circ \Delta s}{\mu}\right) = \psi(e).
\tag{19}
$$

Note that

$$
\frac{x \circ s}{\mu} + \frac{x \circ \Delta s + \Delta x \circ s + \Delta x \circ \Delta s}{\mu} = e.
\tag{20}
$$

Neglecting the term $\Delta x \circ \Delta s$, from Corollary 2.3, we can replace (19) by

$$
\psi\left(\frac{x \circ s}{\mu}\right) + \psi'\left(\frac{x \circ s}{\mu}\right) \circ \left(\frac{x \circ \Delta s + \Delta x \circ s}{\mu}\right) = \psi(e).
\tag{21}
$$

This enables us to rewrite the system (18) as follows

$$
\begin{aligned}
A\Delta x &= 0, \\
A^T \Delta y + \Delta s &= 0, \\
x \circ \Delta s + s \circ \Delta x &= \mu\left(\psi'\left(\frac{x \circ s}{\mu}\right)\right)^{-1} \circ \left(\psi(e) - \psi\left(\frac{x \circ s}{\mu}\right)\right).
\end{aligned}
\tag{22}
$$

Due to the fact that $L(x)L(s) \neq L(s)L(x)$, in general, this system does not always have a unique solution in int K. It is well known that this difficulty can be solved by applying a scaling scheme (see, e.g., [12, 15]). It goes as follows.

Lemma 3.1 (Lemma 28 in [15]) *Let $u \in \text{int } K$. Then*

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

Now we replace the third equation of the system (18) by

$$\psi\left(\frac{P(u)(x + \Delta x) \circ P(u)^{-1}(s + \Delta s)}{\mu}\right) = \psi(e). \tag{23}$$

Applying Newton's method again and neglecting the term $P(u)\Delta x \circ P(u)^{-1}\Delta s$, from Corollary 2.3, we get

$$P(u)x \circ P(u)^{-1}\Delta s + P(u)^{-1}s \circ P(u)\Delta x$$
$$= \mu\left(\psi'\left(\frac{P(u)x \circ P(u)^{-1}s}{\mu}\right)\right)^{-1} \circ \left(\psi(e) - \psi\left(\frac{P(u)x \circ P(u)^{-1}s}{\mu}\right)\right). \tag{24}$$

In this paper, we consider the NT-scaling scheme [1, 2]. Let $u = w^{-\frac{1}{2}}$, where w is the NT-scaling point of x and s. We define

$$v := \frac{P(w)^{-\frac{1}{2}}x}{\sqrt{\mu}}\left[= \frac{P(w)^{\frac{1}{2}}s}{\sqrt{\mu}}\right] \tag{25}$$

and

$$\bar{A} := \frac{1}{\sqrt{\mu}}AP(w)^{\frac{1}{2}}, \qquad d_x := \frac{P(w)^{-\frac{1}{2}}\Delta x}{\sqrt{\mu}}, \qquad d_s := \frac{P(w)^{\frac{1}{2}}\Delta s}{\sqrt{\mu}}. \tag{26}$$

Replacing the third equation of the system (22) by (24), and then using (25) and (26), after some elementary reductions, we obtain

$$\bar{A}d_x = 0,$$
$$\bar{A}^T\Delta y + d_s = 0, \tag{27}$$
$$d_x + d_s = p_v,$$

where

$$p_v = v^{-1} \circ \left(\psi'(v \circ v)\right)^{-1} \circ \left(\psi(e) - \psi(v \circ v)\right).$$

The system (27) has a unique solution in int K (see, e.g., [16]). By choosing the function $\psi(t)$ appropriately, this system can be used to define the new search directions. For example:

- $\psi(t) = t$ yields $p_v = v^{-1} - v$, which gives the classical NT-search direction (see, e.g., [22]);
- $\psi(t) = t^2$ yields $p_v = \frac{1}{2}(v^{-3} - v)$, which gives the search direction differs from the defined in [16] only by a constant multiplier;
- $\psi(t) = t^{\frac{q+1}{2}}$, $q \geq 0$ yields $p_v = \frac{2}{q+1}(v^{-q} - v)$, which also includes the above two cases.

In order to facilitate the analysis of the interior-point algorithm, we restrict our analysis to the case where $\psi(t) = \sqrt{t}$, which yields

$$p_v = 2(e - v). \tag{28}$$

The new search directions d_x and d_s are obtained by solving (27) with (28) so that Δx and Δs are computed via (26). If $(x, y, s) \neq (x(\mu), y(\mu), s(\mu))$, then $(\Delta x, \Delta y, \Delta s)$ is nonzero. The new iterate is obtained by taking a full Nesterov–Todd step as follows

$$x_+ := x + \Delta x, \quad y_+ := y + \Delta y, \quad s_+ := s + \Delta s. \tag{29}$$

For the analysis of the algorithm, we define a norm-based proximity measure $\delta(x, s; \mu)$ as follows

$$\delta(v) := \delta(x, s; \mu) := \frac{\|p_v\|_F}{2} = \|e - v\|_F. \tag{30}$$

From the first two equations of the system (27), d_x and d_s are orthogonal, i.e.,

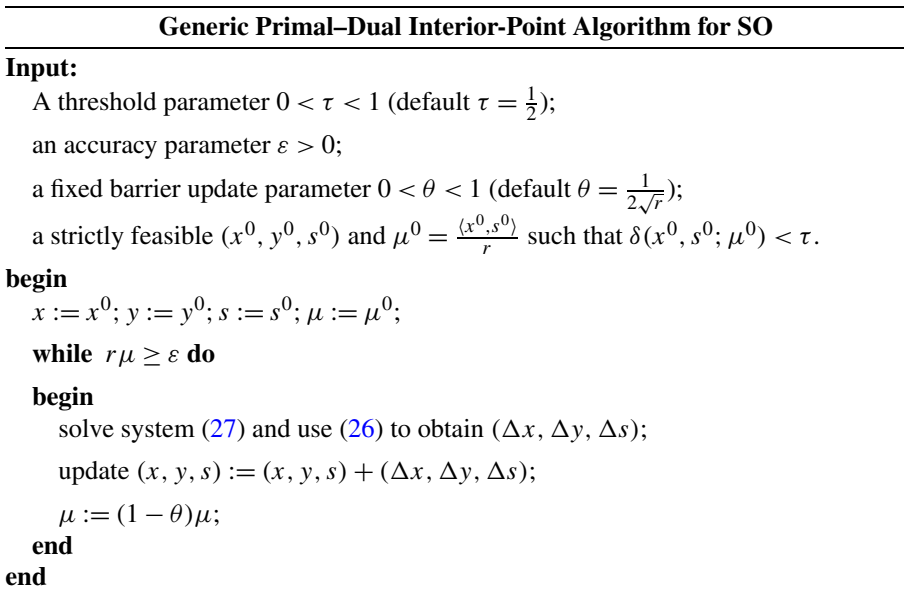$$\langle d_x, d_s \rangle = \langle d_s, d_x \rangle = 0. \tag{31}$$

Furthermore, we can conclude that

$$\delta(v) = 0 \quad \Longleftrightarrow \quad v = e \quad \Longleftrightarrow \quad d_x = d_s = 0 \quad \Longleftrightarrow \quad x \circ s = \mu e. \tag{32}$$

Hence, the value of $\delta(v)$ can be considered as a measure for the distance between the given pair (x, y, s) and the μ-center $(x(\mu), y(\mu), s(\mu))$.

3.3 The Generic Primal–Dual Path-Following Interior-Point Algorithm for SO

The generic primal–dual path-following interior-point algorithm is now presented in Fig. 1.

Generic Primal–Dual Interior-Point Algorithm for SO

Input:

    A threshold parameter $0 < \tau < 1$ (default $\tau = \frac{1}{2}$);

    an accuracy parameter $\varepsilon > 0$;

    a fixed barrier update parameter $0 < \theta < 1$ (default $\theta = \frac{1}{2\sqrt{r}}$);

    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 system (27) and use (26) 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

Fig. 1 Algorithm

4 Analysis of the Algorithm

In this section, we will show that our algorithm can solve the SO problems in polynomial time and prove the local quadratic convergence of the algorithm.

Let us further define

$$q_v = d_x - d_s. \tag{33}$$

Then, we have

$$d_x = \frac{p_v + q_v}{2}, \qquad d_s = \frac{p_v - q_v}{2}, \qquad d_x \circ d_s = \frac{p_v \circ p_v - q_v \circ q_v}{4}. \tag{34}$$

It follows from (31) that $\|p_v\|_F = \|q_v\|_F = 2\delta(v)$.

Let $0 \le \alpha \le 1$, we define

$$x(\alpha) := x + \alpha \Delta x \quad \text{and} \quad s(\alpha) := s + \alpha \Delta s. \tag{35}$$

The following lemma gives a sufficient condition for a feasible step-length $\bar{\alpha} > 0$ such that $x(\bar{\alpha}) \in \text{int } K$ and $s(\bar{\alpha}) \in \text{int } K$.

Lemma 4.1 *Let $x, s \in \text{int } K$ and $x(\alpha) \circ s(\alpha) \in \text{int } K$ for $\alpha \in [0, \bar{\alpha}]$. Then*

$$x(\bar{\alpha}) \in \text{int } K \quad \text{and} \quad s(\bar{\alpha}) \in \text{int } K.$$

Proof Since $x(\alpha) \circ s(\alpha) \in \text{int } K$ for all $\alpha \in [0, \bar{\alpha}]$, Lemma 2.15 in [22], i.e., if $x \circ s \in \text{int } K$, then $\det(x) \ne 0$, implies that $\det(x(\alpha))$ and $\det(s(\alpha))$ do not vanish for all $\alpha \in [0, \bar{\alpha}]$. Since $\det(x(0)) = \det(x) > 0$ and $\det(s(0)) = \det(s) > 0$, by continuity, $\det(x(\alpha)]$ and $\det(s(\alpha))$ stay positive for all $\alpha \in [0, \bar{\alpha}]$. Moreover, by Theorem 2.2, this implies that all the eigenvalues of $x(\alpha)$ and $s(\alpha)$ stay positive for all $\alpha \in [0, \bar{\alpha}]$. Hence, we can conclude that all the eigenvalues of $x(\bar{\alpha})$ and $s(\bar{\alpha})$ are non-negative. This implies the desired result. □

From (29), (25), and (26), we get

$$x_+ = x + \Delta x = \sqrt{\mu} P(w)^{\frac{1}{2}} (v + d_x) \quad \text{and} \quad s_+ = s + \Delta s = \sqrt{\mu} P(w)^{-\frac{1}{2}} (v + d_s).$$

Since $P(w)^{\frac{1}{2}}$ and its inverse $P(w)^{-\frac{1}{2}}$ are automorphisms of $\text{int } K$, Theorem 2.1 implies that x_+ and s_+ belong to $\text{int } K$ if and only if $v + d_x$ and $v + d_s$ belong to $\text{int } K$.

The following lemma shows the strict feasibility of the full Nesterov–Todd step under the condition $\delta(x, s; \mu) < 1$.

Lemma 4.2 *Let $\delta := \delta(x, s; \mu) < 1$. Then the full Nesterov–Todd step is strictly feasible.*

Proof Let $0 \le \alpha \le 1$, we define

$$v_x(\alpha) := v + \alpha d_x \quad \text{and} \quad v_s(\alpha) := v + \alpha d_s. \tag{36}$$

Then

$$
\begin{aligned}
v_x(\alpha) \circ v_s(\alpha) &= (v + \alpha d_x) \circ (v + \alpha d_s) \\
&= v \circ v + \alpha v \circ (d_x + d_s) + \alpha^2 (d_x \circ d_s) \\
&= v \circ v + \alpha v \circ p_v + \alpha^2 \frac{p_v \circ p_v - q_v \circ q_v}{4} \\
&= (1 - \alpha)(v \circ v) + \alpha \left(e - \frac{p_v \circ p_v}{4} \right) + \alpha^2 \left(\frac{p_v \circ p_v - q_v \circ q_v}{4} \right) \\
&= (1 - \alpha)(v \circ v) + \alpha \left(e - (1 - \alpha) \frac{p_v \circ p_v}{4} - \alpha \frac{q_v \circ q_v}{4} \right). \qquad (37)
\end{aligned}
$$

The second to last equality holds due to

$$
v \circ v + v \circ p_v = v \circ v + 2v \circ (e - v) = e - (e - v) \circ (e - v) = e - \frac{p_v \circ p_v}{4}.
$$

Furthermore, since $0 \le \alpha \le 1$, we have

$$
\begin{aligned}
\left\| (1 - \alpha) \frac{p_v \circ p_v}{4} - \alpha \frac{q_v \circ q_v}{4} \right\|_F &\le (1 - \alpha) \left\| \frac{p_v \circ p_v}{4} \right\|_F + \alpha \left\| \frac{q_v \circ q_v}{4} \right\|_F \\
&\le (1 - \alpha) \frac{\|p_v\|_F^2}{4} + \alpha \frac{\|q_v\|_F^2}{4} = \delta^2 < 1.
\end{aligned}
$$

Corollary 2.4 implies that

$$
e - (1 - \alpha) \frac{p_v \circ p_v}{4} - \alpha \frac{q_v \circ q_v}{4} \in \operatorname{int} K.
$$

Thus

$$
(1 - \alpha)(v \circ v) + \alpha \left(e - (1 - \alpha) \frac{p_v \circ p_v}{4} - \alpha \frac{q_v \circ q_v}{4} \right) \in \operatorname{int} K.
$$

From Lemma 4.1, we have

$$
v_x(1) = v + d_x \in \operatorname{int} K \quad \text{and} \quad v_s(1) = v + d_s \in \operatorname{int} K
$$

for $\alpha = 1$. Hence, the result of the lemma holds. $\qquad\square$

According to (25), the v-vector after the step is given by

$$
v_+ := \frac{P(w_+)^{-\frac{1}{2}} x_+}{\sqrt{\mu}} \left[= \frac{P(w_+)^{\frac{1}{2}} s_+}{\sqrt{\mu}} \right], \qquad (38)
$$

where w_+ is the scaling point of x_+ and s_+.

Lemma 4.3 (Proposition 5.9.3 in [16]) *One has*

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

In the next lemma, we proceed to prove the local quadratic convergence of the full Nesterov–Todd step to the target point $(x(\mu), y(\mu), s(\mu))$.

Lemma 4.4 *Let $\delta := \delta(x, s; \mu) < 1$. Then*

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

Thus $\delta(x_+, s_+; \mu) \leq \delta^2$, which shows the quadratical convergence of the algorithm.

Proof From (37), with $\alpha = 1$, we get

$$(v + d_x) \circ (v + d_s) = e - \frac{q_v \circ q_v}{4}. \tag{39}$$

By Lemma 2.6, (13), and Lemma 2.5, we obtain

$$\lambda_{\min}\big((v + d_x) \circ (v + d_s)\big) = \lambda_{\min}\left(e - \frac{q_v \circ q_v}{4}\right)$$

$$\geq 1 - \left\|\frac{q_v \circ q_v}{4}\right\|_F \geq 1 - \frac{\|q_v\|_F^2}{4} = 1 - \delta^2.$$

It follows from Lemma 4.3 and Lemma 2.8 that

$$\lambda_{\min}(v_+) = \lambda_{\min}\big(\big(P(v + d_x)^{\frac{1}{2}}(v + d_s)\big)^{\frac{1}{2}}\big)$$

$$\geq \big(\lambda_{\min}\big((v + d_x) \circ (v + d_s)\big)\big)^{\frac{1}{2}} \geq \sqrt{1 - \delta^2}. \tag{40}$$

Using Lemma 2.11 with $t = 1$, Lemma 2.7, (39), (40), and Lemma 2.5, we have

$$\delta(x_+, s_+; \mu) = \|e - v_+\|_F \leq \frac{1}{1 + \lambda_{\min}(v_+)} \|e - v_+ \circ v_+\|_F$$

$$\leq \frac{1}{1 + \lambda_{\min}(v_+)} \big\|e - (v + d_x) \circ (v + d_s)\big\|_F$$

$$= \frac{1}{1 + \sqrt{1 - \delta^2}} \left\|\frac{q_v \circ q_v}{4}\right\|_F$$

$$\leq \frac{1}{1 + \sqrt{1 - \delta^2}} \frac{\|q_v\|_F^2}{4}$$

$$= \frac{\delta^2}{1 + \sqrt{1 - \delta^2}}.$$

This completes the proof. □

The following lemma gives an upper bound of the duality gap after a full Nesterov–Todd step.

Lemma 4.5 *After a full Nesterov–Todd step. Then*

$$\langle x_+, s_+ \rangle \leq r\mu.$$

Proof Since

$$e - \frac{q_v \circ q_v}{4} = (v + d_x) \circ (v + d_s) \in \text{int } K,$$

and using Lemma 2.4, we obtain

$$\langle x_+, s_+ \rangle = \mu \big\| (v + d_x) \circ (v + d_s) \big\|_F^2 = \mu \left\| e - \frac{q_v \circ q_v}{4} \right\|_F^2 \leq r\mu.$$

This completes the proof. $\qquad\qquad\square$

In the following lemma, we investigate the effect on the proximity measure of a full Nesterov–Todd step followed by an update of the parameter μ.

Lemma 4.6 *Let* $\delta := \delta(x, s; \mu) < 1$ *and* $\mu_+ = (1 - \theta)\mu$, *where* $0 < \theta < 1$. *Then*

$$\delta(x_+, s_+; \mu_+) \leq \frac{\theta\sqrt{r} + \delta^2}{1 - \theta + \sqrt{(1 - \theta)(1 - \delta^2)}}.$$

Proof After updating $\mu_+ = (1 - \theta)\mu$, the vector v_+ is divided by the factor $\sqrt{1 - \theta}$. From Lemma 2.11 with $t = \sqrt{1 - \theta}$, Lemma 2.7, (39), and (40), we have

$$
\begin{aligned}
\delta(x_+, s_+; \mu_+) &= \left\| e - \frac{v_+}{\sqrt{1 - \theta}} \right\|_F = \frac{1}{\sqrt{1 - \theta}} \left\| \sqrt{1 - \theta}\, e - v_+ \right\|_F \\
&\leq \frac{1}{\sqrt{1 - \theta}(\sqrt{1 - \theta} + \lambda_{\min}(v_+))} \left\| (1 - \theta)e - v_+ \circ v_+ \right\|_F \\
&\leq \frac{1}{\sqrt{1 - \theta}(\sqrt{1 - \theta} + \lambda_{\min}(v_+))} \left\| (1 - \theta)e - (v + d_x) \circ (v + d_s) \right\|_F \\
&= \frac{1}{\sqrt{1 - \theta}(\sqrt{1 - \theta} + \lambda_{\min}(v_+))} \left\| -\theta e + \frac{q_v \circ q_v}{4} \right\|_F \\
&\leq \frac{1}{1 - \theta + \sqrt{(1 - \theta)(1 - \delta^2)}} \left(\theta\sqrt{r} + \frac{\|q_v\|^2}{4} \right) \\
&= \frac{1}{1 - \theta + \sqrt{(1 - \theta)(1 - \delta^2)}} \left(\theta\sqrt{r} + \delta^2 \right).
\end{aligned}
$$

This completes the proof. $\qquad\qquad\square$

Corollary 4.1 *Let* $\delta := \delta(x, s; \mu) \leq \frac{1}{2}$ *and* $\theta = \frac{1}{2\sqrt{r}}$ *with* $r \geq 4$. *Then*

$$\delta(x_+, s_+; \mu_+) \leq \frac{1}{2}.$$

Proof Note that

$$1 - \theta = 1 - \frac{1}{2\sqrt{r}} \geq 1 - \frac{1}{2\sqrt{4}} = \frac{3}{4}.$$

From Lemma 4.6 and $\delta \leq \frac{1}{2}$, after some elementary reductions, we get

$$\delta(x_+, s_+; \mu_+) \leq \frac{1}{2}.$$

This completes the proof. $\qquad\qquad\square$

The following lemma gives an upper bound for the total number of iterations produced by our algorithm.

Lemma 4.7 *Suppose that x^0 and s^0 are strictly feasible, $\mu^0 = \frac{\langle x^0, s^0 \rangle}{r}$ and $\delta(x^0, s^0; \mu^0)$ is less than or equal to $\frac{1}{2}$. Moreover, let x^k and s^k be the vectors obtained after k iterations. Then the inequality $\langle x^k, s^k \rangle \leq \varepsilon$ is satisfied for*

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

Proof Lemma 4.5 implies that

$$\langle x^k, s^k \rangle \leq r\mu^k = r(1-\theta)^k \mu^0 = (1-\theta)^k \langle x^0, s^0 \rangle.$$

Then the inequality $\langle x^k, s^k \rangle \leq \varepsilon$ holds if

$$(1-\theta)^k \langle x^0, s^0 \rangle \leq \varepsilon.$$

Taking logarithms, we obtain

$$k \log(1-\theta) + \log \langle x^0, s^0 \rangle \leq \log \varepsilon,$$

and using $-\log(1-\theta) \geq \theta$, we observe that the above inequality holds if

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

Hence, the result of this lemma holds.   □

Theorem 4.1 *Let $\theta = \frac{1}{2\sqrt{r}}$. Then the algorithm requires at most*

$$O\left(\sqrt{r} \log \frac{\langle x^0, s^0 \rangle}{\varepsilon}\right)$$

iterations. The output is a primal–dual pair (x, s) satisfying $\langle x, s \rangle \leq \varepsilon$.

Proof Let $\theta = \frac{1}{2\sqrt{r}}$, Theorem 4.1 follows immediately from Lemma 4.7.   □

Corollary 4.2 *If one takes $x^0 = s^0 = e$, the iteration bound becomes*

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

which is the currently best known iteration bound for the small-update method.

5 Conclusions and Remarks

In this paper, we have shown that a full-Newton step primal–dual path-following interior-point algorithm for LO presented in [17] can be extended to the context of SO. Though the proposed algorithm is exactly an extension from LO to SO, it is not

straightforward due to the non-polyhedral property of a symmetric cone. By employing Euclidean Jordan algebras, we derived the currently best known iteration bound for the small-update method. This unifies the analysis for LO, SOCO, and SDO.

Some interesting topics for further research remain. The search direction used in this paper is based on the NT-scaling scheme. It may be possible to design similar algorithms using other scaling schemes and to obtain polynomial-time iteration bounds. Another topic for further research may be the development of full Nesterov–Todd step primal–dual infeasible interior-point algorithm for SO. Finally, an interesting topic is the generalization of the analysis of the (infeasible) interior-point algorithms to the case where $\psi(t) = t^{\frac{q+1}{2}}$, $q \geq 0$.

References

1. Nesterov, Y.E., Todd, M.J.: Self-scaled barriers and interior-point methods for convex programming. Math. Oper. Res. **22**(1), 1–42 (1997)
2. Nesterov, Y.E., Todd, M.J.: Primal–dual interior-point methods for self-scaled cones. SIAM J. Optim. **8**(2), 324–364 (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. Kojima, M., Shida, M., Shindoh, S.: Local convergence of predictor–corrector infeasible-interior-point algorithms for SDPs and SDLCPs. Math. Program. **96**(80), 129–161 (1998)
5. Ji, J., Potra, F.A., Sheng, R.: On the local convergence of a predictor–corrector method for semidefinite programming. SIAM J. Optim. **10**(1), 195–210 (1999)
6. Monteiro, R.D.C.: Primal–dual path-following algorithms for semidefinite programming. SIAM J. Optim. **7**(3), 663–678 (1997)
7. Potra, F.A., Sheng, R.: On homogeneous interior-point algorithms for semidefinite programming. Optim. Methods Softw. **9**(1–3), 161–184 (1998)
8. Luo, Z.Q., Sturm, J.F., Zhang, S.: Conic convex programming and self-dual embedding. Optim. Methods Softw. **14**(3), 169–218 (2000)
9. Wolkowicz, H., Saigal, R., Vandenberghe, L.: Handbook of Semi-definite Programming, Theory, Algorithms, and Applications. Kluwer Academic, Dordrecht (2000)
10. De Klerk, E.: Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications. Kluwer Academic Publishers, Dordrecht (2002)
11. Güler, O.: Barrier functions in interior point methods. Math. Oper. Res. **21**(4), 860–885 (1996)
12. Faybusovich, L.: Linear system in Jordan algebras and primal-dual interior-point algorithms. J. Comput. Appl. Math. **86**(1), 149–175 (1997)
13. Muramatsu, M.: On a commutative class of search directions for linear programming over symmetric cones. J. Optim. Theory Appl. **112**(3), 595–625 (2002)
14. Rangarajan, B.K.: Polynomial convergence of infeasible-interior-point methods over symmetric cones. SIAM J. Optim. **16**(4), 1211–1229 (2006)
15. Schmieta, S.H., Alizadeh, F.: Extension of primal–dual interior-point algorithms to symmetric cones. Math. Program., Ser. A **96**(3), 409–438 (2003)
16. Vieira, M.V.C.: Jordan algebraic approach to symmetric optimization. Ph.D. thesis, Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, The Netherlands (2007)
17. Darvay, Z.: New interior-point algorithms in linear programming. Adv. Model. Optim. **5**(1), 51–92 (2003)
18. Achache, M.: A new primal–dual path-following method for convex quadratic programming. Comput. Appl. Math. **25**(1), 97–110 (2006)

19. Wang, G.Q., Bai, Y.Q.: A primal–dual path-following interior-point algorithm for second-order cone optimization with full Nesterov–Todd step. Appl. Math. Comput. **215**(3), 1047–1061 (2009)
20. Wang, G.Q., Bai, Y.Q.: A new primal–dual path-following interior-point algorithm for semidefinite optimization. J. Math. Anal. Appl. **353**(1), 339–349 (2009)
21. Roos, C.: A full-Newton step $O(n)$ infeasible interior-point algorithm for linear optimization. SIAM J. Optim. **16**(4), 1110–1136 (2006)
22. Gu, G., Zangiabadi, M., Roos, C.: Full Nesterov–Todd step interior-point methods for symmetric optimization. Eur. J. Oper. Res. **214**(3), 473–484 (2011)
23. Faraut, J., Koranyi, A.: Analysis on Symmetric Cones. Oxford University Press, New York (1994)
24. Sturm, J.F.: Similarity and other spectral relations for symmetric cones. Linear Algebra Appl. **312**(1–3), 135–154 (2000)
25. Faybusovich, L.: A Jordan-algebraic approach to potential-reduction algorithms. Math. Z. **239**(1), 117–129 (2002)