Yugoslav Journal of Operations Research
24 (2014) Number 1, 35-51
DOI: 10.2298/YJOR120904016K

A PREDICTOR-CORRECTOR PATH-FOLLOWING ALGORITHM FOR SYMMETRIC OPTIMIZATION BASED ON DARVAY'S TECHNIQUE

BEHROUZ KHEIRFAM
Department of Mathematics, Azarbaijan Shahid Madani University,
Tabriz, I.R. Iran,
b.kheirfam@azaruniv.edu

Received: September 2012 / Accepted: March 2013

Abstract: In this paper, we present a predictor-corrector path-following interior-point algorithm for symmetric cone optimization based on Darvay's technique. Each iteration of the algorithm contains a predictor step and a corrector step based on a modification of the Nesterov and Todd directions. Moreover, we show that the algorithm is well defined and that the obtained iteration bound is $O(\sqrt{r}\log\frac{r\mu^0}{\varepsilon})$, where r is the rank of Euclidean Jordan algebra.

Keywords: Symmetric cone optimization, interior-point method, predictor-corrector method, polynomial complexity.

MSC: 90C51.

1. INTRODUCTION

In recent years, there have been extensive investigations concerning the analysis of interior-point methods (IPMs) for symmetric cone optimization (SCO). A few optimization problems are special cases of symmetric cones, such as nonnegative orthants, linear optimization (LO), semidefinite optimization (SDO) and second-order cone optimization (SOCO). Basic idea for solving SCO is using feasible interior-point method, as used by Nesterov and Nemirovskii [9]. Their method was primarily either primal or dual based. Later on, Nesterov and Todd [10] proposed symmetric interior-point algorithms on a special class of convex optimization problems, where the associated cone is self-scaled. Later on, it was observed that these cones were precisely symmetric cones [3]. Thus, Nesterov and Todd algorithm was the first primal-dual interior-point algorithm for optimization over symmetric cones. Monteiro and Zhang [8] designed a interior-point path-following algorithm for SDO based on commutative class of search directions. Subsequently, Schmieta and Alizadeh [12] introduced primal-dual IPMs for SCO extensively under the framework of Euclidean Jordan algebra. Roos [11] introduced a full-Newton primal-dual infeasible interior-point (IIPM) algorithm for LO. Gu et al. [5] extended this algorithm to SCO based on Euclidean Jordan algebra. Darvay [1] proposed a new technique for finding a class of search directions. Based on this technique, the author designed a new primal-dual path-following interior-point algorithm for LO with iteration bound $O(\sqrt{n}\log\frac{n}{\varepsilon})$. Recently, Wang and Bai [2] extended the Darvay's technique to SCO.

Sonnevend et al. [13] were the first to introduce the predictor-corrector interior-point algorithm for LO. This algorithm needs more corrector steps after each predictor step in order to return to the appropriate neighborhood of the central path. Mizuno et al. [7] presented a predictor-corrector interior-point algorithm for LO in which each predictor step is followed by a single corrector step, and whose iteration complexity is the best known in LO literature. Ye and Anstreicher [16] extended this result to the linear complementarity (LC) problems with a positive semidefinite matrix with the same iteration complexity. Recently, Illés and Nagy [6] presented a new version of the Mizuno-Todd-Ye predictor-corrector algorithm for $P_*(\varkappa)$-LCP that uses self-regular proximity measure.

Motivated by their work, we propose a predictor-corrector path-following algorithm for solving SCO based on Darvay's technique. Our algorithm uses two kinds of steps: predictor and corrector. The aim of corrector step is to restore the appropriate neighborhood of the central path. After each corrector step, new iterates will be within the region where Newton process is quadratically convergent, which is an advantage of the algorithm. Then the algorithm operates one damped Newton step used to reduce the duality gap. The algorithm is repeated until an ε-approximate solution is followed. We analyze the algorithm and obtain the complexity bound, which coincides with the best known result for SCO.

The paper is organized as follows: In Section 2, firstly we provide the theory of the Euclidean Jordan algebra and their associated symmetric cones; then, after briefly reviewing the central path for SCO, we obtain the search directions based on Darvay's technique for SCO. In Section 3, the predictor-corrector algorithm for SCO is presented. In Section 4, we analyze the algorithm and derive the iteration bound. Finally, we conclude the paper in Section 5.

2. PRELIMINARIES

2.1 Euclidean Jordan algebra

Here, we outline some needed main results on Euclidean Jordan algebra and symmetric cones. For a comprehensive study, the reader is referred to [3, 15].

Jordan algebra $\mathcal{J}$ is a finite dimensional vector space endowed with a bilinear map $\circ: \mathcal{J} \times \mathcal{J} \to \mathcal{J}$ satisfying the following properties for all $x, y \in \mathcal{J}$:

$$1 - x \circ y = y \circ x,$$

$2 - x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$ where $x^2 = x \circ x$.

Moreover, Jordan algebra $(\mathcal{J}, \circ)$ is called Euclidean if there exists an inner product, denoted by "$\langle ., . \rangle$", such that

$\langle x \circ y, z\rangle = \langle x, y \circ z\rangle$,

for all $x, y, z \in \mathcal{J}$.

Jordan algebra has an identity element, if there exists a unique element $e \in \mathcal{J}$ such that $x \circ e = e \circ x = x$, for all $x \in \mathcal{J}$. Throughout the paper, we assume that $\mathcal{J}$ is a Euclidean Jordan algebra with an identity element e. The set $\mathcal{K} = \{x^2 | x \in \mathcal{J}\}$ is called the cone of squares of Euclidean Jordan algebra $(\mathcal{J}, \circ, \langle .,. \rangle)$. Cone is symmetric if and only if it is the cone of squares of some Euclidean Jordan algebra. An element $c \in \mathcal{J}$ is idempotent if $c \circ c = c$. Idempotents x and y are orthogonal if $x \circ y = 0$. An idempotent c is primitive if it is nonzero and can not be expressed by sum of two other nonzero idempotents. A set of primitive idempotents $\{c_1, c_2, \cdots, c_k\}$ is called a Jordan frame if $c_i \circ c_j = 0$, for any $i \neq j \in \{1,2,\cdots,k\}$ and $\sum_{i=1}^{k} c_i = e$. For any $x \in \mathcal{J}$, let r be the smallest positive integer such that $\{e, x, x^2, \cdots, x^r\}$ is linearly dependent; r is called the degree of x and is denoted by $deg\,(x)$. The rank of $\mathcal{J}$, denoted by $rank(\mathcal{J})$ is defined as the maximum of $deg\,(x)$ over all $x \in \mathcal{J}$. The importance of Jordan frame comes from the fact that any element of Euclidean Jordan algebra can be represented by using some Jordan frame, as explained more precisely in the following spectral decomposition theorem.

Theorem 1 *(Theorem III.1.2 in [3]) Let $(\mathcal{J}, \circ, \langle .,. \rangle)$ be an Euclidean Jordan algebra with $rank(\mathcal{J}) = r$. Then, for any $x \in \mathcal{J}$, there exists a Jordan frame $\{c_1, c_2, \cdots, c_r\}$ and real numbers $\lambda_1(x), \lambda_2(x), \cdots, \lambda_r(x)$ such that $x = \sum_{i=1}^{r} \lambda_i(x) c_i$, where the λ_i's are the eigenvalues of x. The numbers $\lambda_i(x)$ (with their multiplicities) are uniquely determined by x. Furthermore, $tr(x) = \sum_{i=1}^{r} \lambda_i(x)$ and $det(x) = \prod_{i=1}^{r} \lambda_i(x)$ where tr and det stand for the trace and determinant, respectively.*

Since ``$\circ$'' is a bilinear map, for each $x \in \mathcal{J}$, there exists a matrix $L(x)$ such that for every $y \in \mathcal{J}, x \circ y = L(x)y$. Moreover, we define $P(x) \coloneqq 2L(x)^2 - L(x^2)$, where $L(x)^2 = L(x)L(x)$. The map $P(x)$ is called the quadratic representation of $\mathcal{J}$, which is an essential concept in the theory of Jordan algebra and plays an important role in the analysis of interior-point algorithms. An element $x \in \mathcal{J}$ is called invertible if there exists a $y = \sum_{i=0}^{m} \alpha_i x^i$ for some finite $m < \infty$ and real numbers α_i such that $x \circ y = y \circ x = e$, and it is denoted as x^{-1}. An element $x \in \mathcal{J}$ is invertible if and only if $P(x)$ is invertible. In this case, $P(x)x^{-1} = x$ and $P(x)^{-1} = P(x^{-1})$.

Let $x = \sum_{i=1}^{r} \lambda_i(x) c_i$ be the spectral decomposition of x. It is possible to extend the definition of any real valued continuous function $f(.)$ to elements of Jordan algebra via their eigenvalues, i.e., $F: \mathcal{J} \to \mathcal{J}$ is given by

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

In particular, we have the square root, $x^{\frac{1}{2}} = \sum_{i=1}^{r} \sqrt{\lambda_i(x)}\, c_i$ when $x \in \mathcal{K}$, and undefined otherwise, the inverse, $x^{-1} = \sum_{i=1}^{r} \lambda_i^{-1}(x)\, c_i$, wherever $\lambda_i \neq 0$, for all $i = 1, 2, \cdots, r$, and undefined otherwise.

The next lemma contains a result of crucial importance in the design of IPMs within the framework of Jordan algebra.

Lemma 2 *(Lemma 2.2 in [4]) Let $x, s \in \mathcal{K}$. Then, $tr(x \circ s) \geq 0$, and we have $tr(x \circ s) = 0$ if and only if $x \circ s = 0$.*

For any $x, y \in \mathcal{J}$, x and y are said to be operator commutable if $L(x)$ and $L(y)$ commute, i.e., $L(x)L(y) = L(y)L(x)$. In other words, x and y operator are commutable if for all $z \in \mathcal{J}, x \circ (y \circ z) = y \circ (x \circ z)$ (see [12]).

Theorem 3 *(Lemma X.2.2 in [3]) Let $x, y \in \mathcal{J}$. The elements x and y operator are commutable if and only if they share a Jordan frame, that is, $x = \sum_{i=1}^{r} \lambda_i(x)c_i$ and $y = \sum_{i=1}^{r} \lambda_i(y)c_i$ for Jordan frame $\{c_1, c_2, \cdots, c_r\}$.*

For any $x, y \in \mathcal{J}$, we define the canonical inner product of $x, y \in \mathcal{J}$ as follows:

$$\langle x, y \rangle = tr(x \circ y),$$

and the Frobenius norm of x as follows:

$$\| x \|_F = \sqrt{\langle x, x \rangle} = \sqrt{tr(x^2)}.$$

It follows that

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

Note that $tr(.)$ is associative, and we have

$$\langle L(x)y, z \rangle = tr\big((x \circ y) \circ z\big) = tr\big((y \circ x) \circ z\big) = tr(y \circ (x \circ z)) = \langle y, L(x)z \rangle,$$

showing that $L(x)$ is a self-adjoint operator. As the definition of $P(x)$ depends only on $L(x)$ and $L(x^2)$, both of which are self-adjoint, $P(x)$ is also self-adjoint. Let $\lambda_{min}(x)$ and $\lambda_{max}(x)$ denote the smallest and the largest eigenvalue of x, respectively. Then

$$|\lambda_{min}(x)| \leq \| x \|_F, \; |\lambda_{max}(x)| \leq \| x \|_F, \; |\langle x, y \rangle| \leq \| x \|_F \| y \|_F.$$

The following lemma shows the existence and the uniqueness of a scaling point w corresponding to any points $x, s \in int\, \mathcal{K}$ such that $P(w)$ takes s into x. This was done by Nesterov and Todd [10] for self-scaled cones. This lemma plays a fundamental role in the design of interior-point algorithms for SCO.

Lemma 4 *(Lemma 3.2 in [4]) Let x, $s \in int\, \mathcal{K}$. Then, there exists a unique $w \in int\, \mathcal{K}$ such that $x = P(w)\, s$. Moreover,*

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

The point w is called the scaling point of x and s. Hence, there exists $\tilde{v} \in int\, \mathcal{K}$ such that

$$\tilde{v} = P\left(w^{-\frac{1}{2}}\right)x = P\left(w^{\frac{1}{2}}\right)s,$$

which is the so-called Nesterov and Todd (NT)-scaling of R^n. We say that two elements $x \in \mathcal{J}$ and $y \in \mathcal{J}$ are similar, as denoted by $x \sim y$, if and only if x and y share the same set of eigenvalues. We say $x \in \mathcal{K}$ if and only if $\lambda_i \geq 0$ and $x \in int\,\mathcal{K}$ if and only if $\lambda_i > 0$, for all $i = 1, 2, \cdots, r$. We also say x is positive semidefinite (positive definite), denoted as $x \succcurlyeq 0$ $(x \succ 0)$, if $x \in \mathcal{K}$ $(x \in int\,\mathcal{K})$.

In what follows, we list some lemmas, which will be used in the analysis later.

Lemma 5 *(Lemma 2.15 in [5]) If $x \circ s \in int\,\mathcal{K}$, then $det\,(x) \neq 0$.*

Lemma 6 (Lemma 2.13 in [5]) Let $x, s \in \mathcal{J}$ with $tr(x \circ s) = 0$. Then

$$-\frac{1}{4} \| x + s \|_F^2\, e \preccurlyeq x \circ s \preccurlyeq \frac{1}{4} \| x + s \|_F^2\, e,$$

$$\| x \circ s \|_F \leq \frac{1}{2\sqrt{2}} \| x + s \|_F^2.$$

Lemma 7 *(Proposition 21 in [12]) Let $x, s, u \in int\,\mathcal{K}$. Then*

$$\text{(i)} - P\left(x^{\frac{1}{2}}\right) s \sim P(s^{\frac{1}{2}})\, x\,.$$

$$\text{(ii)} - P\left((P(u)x)^{\frac{1}{2}}\right) P(u^{-1})s \sim P(x^{\frac{1}{2}})\, s\,.$$

Lemma 8 *(Proposition 3.2.4 in [15]) Let $x, s \in int\,\mathcal{K}$, and w be the scaling point of x and s. Then*

$$(P\left(x^{\frac{1}{2}}\right) s)^{\frac{1}{2}} \sim P(w^{\frac{1}{2}})\, s.$$

Lemma 9 *(Lemma 30 in [12]) Let $x, s \in int\,\mathcal{K}$. Then*

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

Lemma 10 *(Theorem 4 in [14]) Let $x, s \in int\,\mathcal{K}$. Then*

$$\lambda_{\min}\left(P\left(x^{\frac{1}{2}}\right) s\right) \geq \lambda_{\min}(x \circ s).$$

2.2 The problem background

We consider the following symmetric cone optimization (SCO) problem

$$\begin{aligned} &\min\ \langle c, x\rangle \\ &s.t.\ Ax = b, \qquad\qquad (P) \\ &\qquad x \in \mathcal{K}, \end{aligned}$$

where c and the rows of A lie in $\mathcal{J}$, and $b \in R^m, \langle x, s\rangle = tr(x \circ s)$ stands for the trace inner product in $\mathcal{J}$. Moreover, assume that a_i is the i −th row of A, then $Ax = b$ means that

$$\langle a_i, x\rangle = b_i,\ i = 1, 2, \cdots, m. \qquad (1)$$

The dual problem of (P) is as follows

$$\max \ b^T y$$
$$s.t. \ A^T y + s = c, \qquad (D)$$
$$s \in \mathcal{K},$$

where $y \in R^m$ and $A^T y + s = c$ means that

$$\sum_{i=1}^{m} y_i \, a_i + s = c. \tag{2}$$

Throughout the paper, we assume that (P) and (D) satisfy the interior point condition (IPC), i.e., there exists (x^0, y^0, s^0) such that

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

and the matrix A is of rank m. The optimality conditions for (P) and (D) are given by the following system

$$Ax = b, \qquad x \in \mathcal{K}$$
$$A^T y + s = c, \ \ s \in \mathcal{K} \tag{3}$$
$$x \circ s = 0.$$

In path-following IPMs one follows the central path that is given as the set of solutions (μ-centers) of the perturbed optimality conditions

$$Ax = b,$$
$$A^T y + s = c, \tag{4}$$
$$x \circ s = \mu \, e.$$

For each $\mu > 0$, the perturbed system (4) has a unique solution $(x(\mu), y(\mu), s(\mu))$, and we call $x(\mu)$ and $(y(\mu), s(\mu))$ the μ-centers of problems (P) and (D) respectively. The set of μ-centers gives a curve called the central path of (P) and (D). If $\mu \to 0$, then the limit of the central path exists and since the limit points satisfy the complementarity condition, the limit yields an ε-approximate solution for (P) and (D) [4].

Similarly to the LO case [1], Wang and Bai [2] replace the standard centering equation $x \circ s = \mu e$ by $\varphi(\frac{x \circ s}{\mu}) = \varphi(e)$ where $\varphi(.)$ is the vector-valued function induced by the univariate function $\varphi(t)$. Thus, the system (4) becomes

$$Ax = b,$$
$$A^T y + s = c, \tag{5}$$
$$\varphi\left(\frac{x \circ s}{\mu}\right) = \varphi(e).$$

Applying Newton's method to system (5), then using Taylor's theorem to the third equation, lead to

$$A \, \Delta \, x = 0,$$
$$A^T \Delta \, y + \Delta \, s = 0, \tag{6}$$

$$x \circ \Delta s + s \circ \Delta x = \mu(\varphi'(\frac{x \circ s}{\mu}))^{-1} \circ \left(\varphi(e) - \varphi\left(\frac{x \circ s}{\mu}\right)\right).$$

Due to the fact that $L(x)L(s) \neq L(s)L(x)$, system (6) does not always have a unique solution in $int\,\mathcal{K}$. It is well known that this difficulty can be resolved by applying a scaling scheme. This is given in the following lemma.

Lemma 11 *(Lemma 28 in [12]) Let* $u \in int\,\mathcal{K}$. *Then* $x \circ s = \mu e \Leftrightarrow P(u)x \circ P(u^{-1})s = \mu e$.

Replacing the third equation of the system (5) by

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

and applying Newton's method to the result system lead us to the following system

$$\begin{aligned} &A\,\Delta x = 0,\\ &A^T \Delta y + \Delta s = 0,\\ &P(u)x \circ P(u^{-1})\Delta s + P(u^{-1})s \circ P(u)\Delta x = \\ &\quad \mu(\varphi'\left(\frac{P(u)x \circ P(u^{-1})s}{\mu}\right))^{-1} \circ \left(\varphi(e) - \varphi\left(\frac{P(u)x \circ P(u^{-1})s}{\mu}\right)\right). \end{aligned} \tag{7}$$

Let $u = w^{-\frac{1}{2}}$, where w is the NT-scaling point of x and s as defined in Lemma 4. We define

$$v = \frac{P(w^{-\frac{1}{2}})x}{\sqrt{\mu}} \left[= \frac{P(w^{\frac{1}{2}})s}{\sqrt{\mu}}\right] \tag{8}$$

and

$$\bar{A} := \sqrt{\mu} A P\left(w^{\frac{1}{2}}\right), d_x := \frac{P\left(w^{-\frac{1}{2}}\right)\Delta x}{\sqrt{\mu}}, d_s := \frac{P\left(w^{\frac{1}{2}}\right)\Delta s}{\sqrt{\mu}}. \tag{9}$$

This enables us to rewrite the system (7), considering $\varphi(t) = \sqrt{t}$, as follows:

$$\begin{aligned} &\bar{A}\, d_x = 0,\\ &\bar{A}^T \frac{\Delta y}{\mu} + d_s = 0,\\ &d_x + d_s = 2(e - v) := p_v. \end{aligned} \tag{10}$$

The search directions d_x and d_s are obtained by solving (10) so that Δx and Δs are computed via (9). The new iterate is obtained by taking a full NT-step as follows

$$x^+ := x + \Delta x, \qquad y^+ := y + \Delta y, \qquad s^+ := s + \Delta s.$$

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

$$\sigma(v) := \sigma(x, s;\, \mu) := \frac{\|p_v\|_F}{2} = \| e - v \|_F. \tag{11}$$

We can conclude that

$$\sigma(v) = 0 \Leftrightarrow v = e \Leftrightarrow d_x = d_s = 0 \Leftrightarrow x \circ s = \mu e. \quad (12)$$

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

3. THE PREDICTOR-CORRECTOR ALGORITHM

In this section, we propose a predictor-corrector algorithm based on NT-directions obtained by Darvay's technique, which uses these directions in the both predictor- and corrector steps. Firstly, we define the τ-neighborhood of the central path as follows

$$\mathcal{N}(\tau, \mu) := \{(x, s) | Ax = b, A^T y + s = c, x, s \in int\,\mathcal{K}, \sigma(x, s;\ \mu) \leq \tau\},$$

where τ is a threshold parameter. The framework of the algorithm is described as follows

---

Algorithm: A predictor-corrector algorithm for SCO

Input:

An accuracy parameter $\varepsilon > 0$;

barrier update parameter $\theta, 0 < \theta < \frac{1}{2}$ ($default\ \theta = \frac{5}{16\sqrt{r}}$);

proximity parameter τ, $0<\tau<1$ (default $\tau = \frac{1}{2}$);

an initial point (x^0, y^0, s^0) such that $\sigma(x^0, s^0; \mu^0) \leq \tau$.

begin:

$x := x^0, s := s^0, y := y^0;\ \mu := \mu^0$;

While $tr(x \circ s) > \varepsilon$ **do**

begin:

solve system (10) and via (9) to obtain $(\Delta x,\ \Delta y,\ \Delta s)$;

and let $(x, y, s) := (x, y, s) + (\Delta x,\ \Delta y,\ \Delta s),\ \mu = \mu$;

solve system (13) and via (9) to obtain $(\Delta^p x, \Delta^p y, \Delta^p s)$;

$(x^p, y^p, s^p) := (x, y, s) + \theta(\Delta^p x, \Delta^p y, \Delta^p s)$;

$\mu^p := (1 - 2\theta)\mu$;

$x = x^p, y = y^p, s = s^p, \mu = \mu^p$;

end

end

If for the current point (x, y, s) in the neighborhood of the central path $\mathcal{N}(\tau, \mu)$, $tr(x \circ s) > \varepsilon$ then the algorithm performs centering and predictor steps. In centering step, by solving the system (10) for the scaled-directions (d_x, d_s), that is

$$\bar{A}d_x = 0,$$
$$\bar{A}^T \frac{\Delta y}{\mu} + d_s = 0, \tag{13}$$
$$d_x + d_s = 2(e - v),$$

and using (9) for $(\Delta x,\ \Delta y,\ \Delta s)$, we obtain

$$x = x + \Delta x, \qquad y = y + \Delta y, \qquad s = s + \Delta s.$$

In the predictor (affine-scaling) step, starting at the iterate $(x,\ y,\ s)$ and targeting at the μ-centers, the search directions $(\Delta^p x,\ \ \Delta^p y, \Delta^p s)$ are the damped Newton directions, defined by

$$\bar{A}d_x^p = 0,$$
$$\bar{A}^T \frac{\Delta^p y}{\mu} + d_s^p = 0, \tag{14}$$
$$d_x^p + d_s^p = -2v.$$

We denote the iterates after a predictor step by $x^p = x + \theta\Delta^p x,\ y^p = y + \theta\Delta^p\ y,\ s^p = s + \theta\Delta^p s,\ \mu^p = (1 - 2\theta)\mu$, where $\theta \in (0, \frac{1}{2})$. The point $(x^p,\ y^p, s^p)$ will be in the τ-neighborhood again. The algorithm repeats until the duality gap is less than the accuracy parameter ε.

4. ANALYSIS

In this section, we deal with the analysis of the previous algorithm. In the analysis of the affine-scaling step, we will give sufficient conditions for strict feasibility, and the effect on the proximity measure; the proximity measure does not exceed the proximity parameter. At the centering step, we describe the effects of a full Newton-step for the quantity of the proximity measure.

Before dealing with the analysis of the steps, we prove a lemma which yields lower and upper bounds for the eigenvalues of v.

Lemma 12 *Let $x, s\ \in int\ \mathcal{K}$ and $\mu > 0$. Assume that $\sigma := \sigma(v)$. Then*

$$1 - \sigma \leq \lambda_i(v) \leq\ 1 + \sigma, i = 1, 2, \cdots, r.$$

Moreover, the following inequalities hold

$$\lambda_{min} \geq\ (1 - \sigma)^2,\ \ \| v \|_F^2 \leq r(1 + \sigma)^2.$$

Proof *By the definition of σ (cf. 11), we have*

$$\sigma^2 = \| e - v \|_F^2 = \sum_{i=1}^{r} \lambda_i\,(e - v)^2 = \sum_{i=1}^{r} (1 - \lambda_i(v))^2.$$

This implies that

$$|1 - \lambda_i(v)| \leq \sigma, \qquad i = 1, 2, \cdots, r$$

or equivalently

$$1 - \sigma \leq \lambda_i(v) \leq\ 1 + \sigma,\ i = 1, 2, \cdots, r. \tag{15}$$

This proves the first part of the lemma. For the proof of the second part, by definition of Frobenius norm and (15), we have

$$\| v \|_F^2 = \sum_{i=1}^{r} \lambda_i(v)^2 \leq \sum_{i=1}^{r} (1+\sigma)^2 = r(1+\sigma)^2,$$

and

$$\lambda_{min}(v^2) = \lambda_{min}(v)^2 \geq (1-\sigma)^2.$$

This completes the proof. ■

4.1 The affine-scaling step

Next lemma gives a sufficient condition for yielding strict feasibility after an affine-scaling step.

Lemma 13 *Let $x, s \in int\,\mathcal{K}$ and $\mu > 0$ such that $\sigma < 1$. Furthermore, let $0 < \theta < \frac{1}{2}$. Let $x^p = x + \theta \Delta^p x$ and $s^p = s + \theta \Delta^p s$ denote the iterates after an affine-scaling step. Then $x^p, s^p \in int\,\mathcal{K}$ if $K(\sigma, \theta, r) > 0$, where*

$$K(\sigma,\theta,r) = (1-\sigma)^2 - \frac{r\theta^2(1+\sigma)^2}{1-2\theta}.$$

Proof *Introduce a step length α with $\alpha \in [0,1]$ and define*

$$x^p(\alpha) = x + \alpha\theta\Delta^p x, \qquad s^p(\alpha) = x + \alpha\theta\Delta^p s.$$

Using (8) and (9), we have

$$x^p(\alpha) = x + \alpha\theta\Delta^p x = \sqrt{\mu}\, P\left(w^{\frac{1}{2}}\right)\left(v + \alpha\theta\, d_x^p\right),$$

$$s^p(\alpha) = s + \alpha\theta\Delta^p s = \sqrt{\mu}\, P\left(w^{-\frac{1}{2}}\right)\left(v + \alpha\theta\, d_s^p\right). \tag{16}$$

Since $P\left(w^{\frac{1}{2}}\right)$ and $P\left(w^{-\frac{1}{2}}\right)$ are automorphisms of $int\,\mathcal{K}$ (Theorem III.2.1 in [3]), by (16), x^p and s^p belong to $int\,\mathcal{K}$ if and only if $v + \theta\, d_x^p$ and $v + \theta\, d_s^p$ belong to $int\,\mathcal{K}$. Therefore, using the third equation of (14), we obtain

$$v_x^p(\alpha) \circ v_s^p(\alpha) := (v + \alpha\theta d_x^p) \circ (v + \alpha\theta\, d_s^p)$$
$$= v^2 + \alpha\theta v \circ (d_x^p + d_s^p) + \alpha^2\, \theta^2 d_x^p \circ d_s^p$$
$$= v^2 + \alpha\theta v \circ (-2v) + \alpha^2\theta^2 d_x^p \circ d_s^p$$
$$= (1-2\alpha\theta)v^2 + \alpha^2\theta^2 d_x^p \circ d_s^p.$$

From the above relation, we have

$$\frac{v_x^p(\alpha) \circ v_s^p(\alpha)}{1-2\alpha\theta} = v^2 + \frac{\alpha^2\theta^2}{1-2\alpha\theta} d_x^p \circ d_s^p. \tag{17}$$

It follows that

$$\lambda_{min}\left(\frac{v_x^p(\alpha)\circ v_s^p(\alpha)}{1-2\alpha\theta}\right)=\lambda_{min}\left(v^2+\frac{\alpha^2\theta^2}{1-2\alpha\theta}d_x^p\circ d_s^p\right)$$

$$\geq \lambda_{min}(v^2)-\frac{\alpha^2\theta^2}{1-2\alpha\theta}\parallel \lambda(d_x^p\circ d_s^p)\parallel_\infty.$$

For each fixed $0<\theta<\frac{1}{2}$, the function $f(\alpha)=\frac{\alpha^2\theta^2}{1-2\alpha\theta}$ for $0\leq\alpha\leq 1$ is a strictly increasing. Thus, we obtain

$$\lambda_{min}\left(\frac{v_x^p(\alpha)\circ v_s^p(\alpha)}{1-2\alpha\theta}\right)\geq \lambda_{min}(v^2)-\frac{\theta^2}{1-2\theta}\parallel \lambda\big(d_x^p\circ d_s^p\big)\parallel_\infty. \tag{18}$$

From Lemma 6, the third equation of (14) and Lemma 12, we get

$$\parallel \lambda\big(d_x^p\circ d_s^p\big)\parallel_\infty\leq\frac{1}{4}\parallel d_x^p+d_s^p\parallel_F^2=\parallel v\parallel_F^2\leq r(1+\sigma)^2. \tag{19}$$

Now, using (18), (19) and Lemma 12, we obtain

$$\lambda_{min}\left(\frac{v_x^p(\alpha)\circ v_s^p(\alpha)}{1-2\alpha\theta}\right)\geq (1-\sigma)^2-\frac{\theta^2(1+\sigma)^2}{1-2\theta}=K(\sigma,\theta,r). \tag{20}$$

This implies that $det(v_x^p(\alpha)\circ v_s^p(\alpha))>0$ for $0\leq\alpha\leq 1$. By Lemma 5, it follows that $det\left(v_x^p(\alpha)\right)\neq 0$ and $det\left(v_s^p(\alpha)\right)\neq 0$, for $0\leq\alpha\leq 1$. Since $det(v_x^p(0))=det(v_s^p(0))=det(v)>0$ and $v_x^p(\alpha), v_s^p(\alpha)$ are linear functions of α, they do not change sign on [0, 1]. Thus, $det(v_x^p(\alpha))$ and $det(v_s^p(\alpha))$ stay positive for all $0\leq\alpha\leq 1$. Moreover, by Theorem 3, this implies that all the eigenvalues of $v_x^p(\alpha)$ and $v_s^p(\alpha)$ stay positive for all $0\leq\alpha\leq 1$. Hence, all the eigenvalues of $v_x^p(1)$ and v_s^p (1) are positive. Therefore, $v+\theta\, d_x^p\in int\,\mathcal{K}$ and $v+\theta\, d_s^p\in int\,\mathcal{K}$, completing the proof. ■

Following (8), we denote

$$v^p=\frac{P(w^p)^{\frac{1}{2}}s^p}{\sqrt{\mu^p}}, \tag{21}$$

where, w^p is the scaling point of x^p and s^p. Using (16) with $\alpha=1$, (21) and lemmas 7 and 8, we have

$$(v^p)^2\sim\frac{P(x^p)^{\frac{1}{2}}s^p}{\mu^p}$$

$$=\frac{\mu\left(P\left(P(w)^{\frac{1}{2}}(v+\theta d_x^p)\right)^{\frac{1}{2}}P(w)^{-\frac{1}{2}}(v+\theta d_s^p)\right)}{\mu(1-2\theta)} \tag{22}$$

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

Using Lemma 10, (22) and (20) with $\alpha = 1$, we obtain

$$\lambda_{min}((v^p)^2) = \lambda_{min}\left(P\left(\frac{v+\theta d_x^p}{\sqrt{1-2\theta}}\right)^{\frac{1}{2}}\left(\frac{v+\theta d_s^p}{\sqrt{1-2\theta}}\right)\right)$$

$$\geq \lambda_{min}\left(\left(\frac{v+\theta d_x^p}{\sqrt{1-2\theta}}\right)\circ\left(\frac{v+\theta d_s^p}{\sqrt{1-2\theta}}\right)\right) \quad (23)$$

$$\geq K(\sigma,\theta,r).$$

In the following lemma, we investigate the effect on the proximity measure of an affine-scaling step, and the update of the parameter μ.

Lemma 14 *Let* $\sigma := \sigma(x,s;\ \mu) < 1$, $\mu^p = (1-2\theta)\mu$, *where* $0< \theta < \frac{1}{2}$, $K(\sigma,\ \theta,r) > 0$ *and let* x^p, s^p *denote the iterates after an affine-scaling step, i.e.,* $x^p = x + \theta\Delta^p x$ *and* $s^p = s + \theta\Delta^p s$. *Then*

$$\sigma^p := \sigma(x^p,s^p;\ \mu^p) \leq \frac{\varrho(\sigma)-\sqrt{2}K(\sigma,\theta,r)}{1+\sqrt{K(\sigma,\theta,r)}},$$

where $K(\sigma,\ \theta,r)$ is defined as in Lemma 13, and

$$\varrho(\sigma) = \left(1+\sqrt{2}\right)\sigma^2 + 2\left(1-\sqrt{2}\right)\sigma + \sqrt{2}.$$

Proof *From Lemma 13, we deduce that the affine-scaling step is strictly feasible. Using (22), (23), (17) with* $\alpha = 1$ *and Lemma 9, we have*

$$\sigma^p := \sigma(x^p,s^p;\ \mu^p) = \| e - v^p \|_F$$

$$= \| \left(e+(v^p)\right)^{-1} \circ (e-(v^p)^2) \|_F$$

$$\leq \frac{1}{1+\lambda_{min}(v^p)} \| e-(v^p)^2 \|_F$$

$$\leq \frac{1}{1+\lambda_{min}(v^p)} \| e - P\left(\frac{v+\theta d_x^p}{\sqrt{1-2\theta}}\right)^{\frac{1}{2}}\left(\frac{v+\theta d_s^p}{\sqrt{1-2\theta}}\right) \|_F$$

$$\leq \frac{1}{1+\lambda_{min}(v^p)} \| e - \left(\frac{v+\theta d_x^p}{\sqrt{1-2\theta}}\right)\circ\left(\frac{v+\theta d_s^p}{\sqrt{1-2\theta}}\right) \|_F$$

$$\leq \frac{1}{1+\sqrt{K(\sigma,\theta,r)}} \| e - v^2 - \frac{\theta^2}{1-2\theta} d_x^p \circ d_s^p \|_F$$

$$\leq \frac{1}{1+\sqrt{K(\sigma,\theta,r)}} (\| e-v^2 \|_F - \frac{\theta^2}{1-2\theta} \| d_x^p \circ d_s^p \|_F.$$

On the other hand, by Lemma 12, we have

$$\begin{aligned} \| e - v^2 \|_F &= \| e - v + v - v^2 \|_F \\ &\leq \| e - v \|_F + \| v - v^2 \|_F \\ &\leq \sigma + \lambda_{max}(v) \| e - v \|_F \\ &\leq \sigma + (1+\sigma)\sigma. \end{aligned} \tag{24}$$

Moreover, by Lemma 6, the third equation of (14) and Lemma 12, we get

$$\| d_x^p \circ d_s^p \|_F \leq \frac{1}{2\sqrt{2}} \| d_x^p + d_s^p \|_F^2 = \sqrt{2} \| v \|_F^2 \leq \sqrt{2} r (1+\sigma)^2. \tag{25}$$

Finally, using (24) and (25)

$$\sigma^p \leq \frac{2\sigma + \sigma^2 + \frac{\sqrt{2}\theta^2}{1-2\theta} r(1+\sigma)^2}{1 + \sqrt{K(\sigma,\theta,r)}} = \frac{\varrho(\sigma) - \sqrt{2} K(\sigma,\theta,r)}{1 + \sqrt{K(\sigma,\theta,r)}}.$$

We got the desired. ■

4.2 The corrector step

The next lemma gives a condition for strictly feasibility of full Nesterov and Todd step (NT-step).

Lemma 15 *(Lemma 4.2 in [2]) Let $\sigma := \sigma(v) < 1$. Then the full NT-step is strictly feasible.*

The second lemma is devoted to the proximity measure of the iterates obtained by a full NT-step.

Lemma 16 *(Lemma 4.4 in [2]) Let $\sigma(v) < 1$. Suppose that the iterates x^+ and s^+ are produced by a full NT-step, i.e., $x^+ = x + \Delta x$ and $s^+ = s + \Delta s$. Then*

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

Thus, $\sigma(x^+, s^+; \mu) \leq \sigma^2$, which shows the quadratical convergence of the algorithm.

The following lemma gives an upper bound of the duality gap after a full NT-step.

Lemma 17 *(Lemma 4.5 in [2]) After a full NT-step, then*

$$tr(x^+ \circ s^+) \leq \mu r.$$

4.3 The iteration bound

The following lemma gives an upper bound of the duality gap after the main iteration.

Lemma 18 *Let $x, s \in int\, \mathcal{K}$, $\mu > 0$ such that $\sigma := \sigma(x, s; \mu) < 1$, and $0 < \theta < \frac{1}{2}$. If x^p and s^p are the iterates obtained after the affine-scaling step of the algorithm, then*

$$tr(x^p \circ s^p) \leq (1 - 2\theta + 2\theta^2) tr(x \circ s) \leq (1-\theta) tr(x \circ s) < \frac{r\mu^p}{1 - 2\theta}.$$

Proof *Letting $\alpha = 1$ in (16) and (17), we obtain*

$$tr(x^p \circ s^p) = tr(\sqrt{\mu}P(w)^{\frac{1}{2}}(v + \theta d_x^p) \circ \sqrt{\mu}\, P(w)^{-\frac{1}{2}}(v + \theta d_s^p)$$
$$= \mu\, tr\left(\left(v + \theta\, d_x^p\right) \circ \left(v + \theta d_s^p\right)\right) \tag{26}$$
$$= \mu\, tr((1 - 2\theta)v^2 + \theta^2\, d_x^p \circ d_s^p)$$
$$= \mu\,(1 - 2\theta)tr(v^2) + \mu\theta^2\, tr\left(d_x^p \circ d_s^p\right).$$

From the third equation of (14), we get

$$tr\left(d_x^p \circ d_s^p\right) = 2tr(v^2) - \frac{\|d_x^p\|_F^2 + \|d_s^p\|_F^2}{2} \leq 2\, tr(v^2). \tag{27}$$

Now, using (26) and (27), we obtain

$$tr(x^p \circ s^p) \leq (1 - 2\theta + 2\theta^2)\, tr(x \circ s).$$

This proves the first inequality. If $0 < \theta < \frac{1}{2}$ then $\frac{1}{2} < 1 - \theta < 1$, we get

$$1 - 2\theta + 2\theta^2 = 1 - 2\theta(1 - \theta) < 1 - \theta.$$

This proves the second inequality. Since x and s are obtained by a full NT-step of the algorithm, by Lemma 16, we have $tr(x \circ s) \leq r\mu$. Therefore

$$(1 - \theta)tr(x \circ s) \leq (1 - \theta)r\mu = \frac{(1 - \theta)r\mu^p}{1 - 2\theta} < \frac{r\mu^p}{1 - 2\theta},$$

and the proof is completed. ■

4.4 Fixing the parameter

In the following subsection, we want to fix the parameters τ and θ, which guarantee that after a main iteration, the proximity measure will not exceed the proximity parameter got before.

Let (x, y, s) be the iterate at the start of a main iteration with $x \in int\,\mathcal{K}$ and $s \in int\,\mathcal{K}$ such that $\sigma = \sigma(x,\ s;\ \mu) \leq \tau$. After a corrector step, by Lemma 16, one has

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

It can be easily verified that the right-hand side of the above inequality is a monotonically increasing with respect to σ, which implies that

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

Following the predictor step and a μ-update, by Lemma 14, one has

$$\sigma^p \leq \frac{\varrho(\sigma) - \sqrt{2}K(\sigma,\theta,r)}{1 + \sqrt{K(\sigma,\theta,r)}}, \tag{28}$$

where $K(\sigma,\theta,r)$ defined as in Lemma 13. It is easily verifiable that the right-hand side of (28) is a monotonically increasing one with respect to σ, which means that

$$\frac{\varrho(\sigma)-\sqrt{2}K(\sigma,\theta,r)}{1+\sqrt{K(\sigma,\theta,r)}} \le \frac{\varrho(w(\tau))-\sqrt{2}K(w(\tau),\ \theta,r)}{1+\sqrt{K(w(\tau),\theta,r)}}.$$

To keep $\sigma^p \le \tau$, it suffices that

$$\frac{\varrho(w(\tau))-\sqrt{2}K(w(\tau),\ \theta,r)}{1+\sqrt{K(w(\tau),\theta,r)}} \le \tau.$$

At this stage, if we set $\tau = \frac{1}{2}$ and $\theta = \frac{5}{16\sqrt{r}}$, the inequality above certainly holds. This means, that $x, s \in int\, \mathcal{K}$ and $\sigma(x,s;\ \mu) \le \frac{1}{2}$ are maintained during the algorithm. Thus the algorithm is well-defined. Moreover, one has

$$K(\sigma,\theta,r) = (1-\sigma)^2 - \frac{r\theta^2(1+\sigma)^2}{1-2\theta}$$

$$\ge \left(1-w(\tau)\right)^2 - \frac{r\theta^2\left(1+w(\tau)\right)^2}{1-2\theta} \ge 0.4151,$$

by Lemma 13, one conclude that the predictor step is strictly feasible.

4.5 Complexity bound

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

Lemma 19 *Let x^0 and s^0 be strictly feasible, $\mu^0 = \frac{tr(x^0 \circ s^0)}{r}$ and $\sigma(x^0, s^0;\ \mu^0) < \tau$. Moreover, let x^k and s^k be iterates obtained after k iterations. Then $tr(x^k \circ s^k) \le \varepsilon$ for*

$$k \ge 1 + \left\lceil \frac{1}{2\theta} \log \frac{tr(x^0 \circ s^0)}{\varepsilon} \right\rceil.$$

Proof *It follows from Lemma 18 that*

$$tr(x^k \circ s^k) < \frac{r\mu^k}{1-2\theta} = r(1-2\theta)^{k-1}\mu^0 = (1-2\theta)^{k-1} tr(x^0 \circ s^0).$$

Then the inequality $tr(x^k \circ s^k) \le \varepsilon$ holds if

$$(1-2\theta)^{k-1} tr(x^0 \circ s^0) \le \varepsilon.$$

Taking logarithms, we obtain

$$(k-1)\log(1-2\theta) + \log tr(x^0 \circ s^0) \le \varepsilon.$$

Since $\log(1+\theta) \le \theta,\ \theta \ge -1$, we observe that the above inequality holds if

$$-2\theta(k-1) + \log tr(x^0 \circ s^0) \le \log \varepsilon.$$

This implies the result. ■

Theorem 20 *Let $\tau = \frac{1}{2}$ and $\theta = \frac{5}{16\sqrt{r}}$, then the algorithm is well defined and the algorithm requires at most*

$$\mathcal{O}\left(\sqrt{r}\log\frac{tr(x^0 \circ s^0)}{\varepsilon}\right),$$

iterations. The output is a primal-dual pair (x, s) satisfying $tr(x \circ s) \leq \varepsilon$.

Proof *Since $\tau = \frac{1}{2}$ and $\theta = \frac{5}{16\sqrt{r}}$, the proof follows from Lemma 19.* ■

5. CONCLUSION

We have introduced a predictor-corrector path-following algorithm for SCO. We showed that this algorithm can solve SCO problems in polynomial-time, and that it can derive the iteration bound for the algorithm with small-update method.

Acknowledgement

The author would like to thanks the editor and the referees for their useful suggestions which improve the presentation of this paper.

REFERENCES

[1] Darvay, Z., "New interior point algorithms in linear programming", *Adv. Model. Optim.*, 5(1) (2003) 51-92.
[2] 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.*, DOI 10.1007/s10957-012-0013-x.
[3] Faraut, J., Korányi, A., *Analysis on symmetric cones, Oxford Mathematical Monographs*, The Clarendon Press Oxford University Press, Oxford Science Publications, New York, (1994).
[4] Faybusovich, L., "Linear systems in Jordan algebras and primal-dual interior-point algorithms", *J. Comput. Appl. Math.*, 86 (1997) 149-175.
[5] Gu, G., Zangiabadi, M., Roos, C., "Full Nesterov-Todd step infeasible interior-point method for symmetric optimization", *European J. Oper. Res.*, 214(3) (2011) 473-484.
[6] Illés, T., Nagy, M., "A Mizuno-Todd-Ye type predictor-corrector algorithm for sufficient linear complementarity problems", *European J. Oper. Res.*, 181 (2007) 1097-1111.
[7] Mizuno, S., Todd, M.J., Ye, Y., "On adaptive-step primal-dual interior point algorithms for linear programming", *Math. Oper. Res.*, 18 (1993) 964-981.
[8] Monteiro, R. D. C., Zhang, Y., "A unified analysis for a class of path-following primal-dual interior-point algorithms for semidefinite programming", *Math. Program.*, 8 (1998) 281-299.
[9] Nesterov, Y.E., Nemirovskii, A.S., *Interior Point Polynomial Algorithms in Convex Programming*, in: SIAM Stud. Appl. Math., Philadelphia, SIAM, 13,1994.
[10] Nesterov, Y.E., Todd, M.J., "Self-scaled barriers and interior-point methods for convex programming", *Math. Oper. Res.*, 22(1) (1997) 1-42.
[11] Roos, C., "A full-Newton step $O(n)$ infeasible interior-point algorithm for linear optimization", *SIAM J. Optim.*, 16(4) (2006) 1110-1136.
[12] Schmieta, S.H., Alizadeh, F., "Extension of primal-dual interior-point algorithm to symmetric cones", *Math. Program.*, 96(3) (2003) 409-438.

[13] Sonnevend, G., Stoer, G., Zhao, G., "On the complexity of following the central path of linear programs by linear extrapolation", *Methods Oper. Res.*, 63, (1989) 19-31.
[14] Sturm, J.F., "Similarity and other spectral relations for symmetric cones", *Algebra Appl.*, 312(1-3) (2000) 135-154.
[15] 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.
[16] Ye, Y., Anstreicher, K., "On quadratic and $O(\sqrt{n}L)$ convergence of a predictor-corrector algorithm for LCP", *Math. Program.*, 62 (1993) 537-551.