A Corrector–Predictor Path-Following Method for Convex Quadratic Symmetric Cone Optimization

Behrouz Kheirfam

Abstract After a brief introduction to Euclidean Jordan algebra, we present a new corrector–predictor path-following interior-point algorithm for convex, quadratic, and symmetric cone optimization. In each iteration, the algorithm involves two kind of steps: a predictor (affine-scaling) step and a full Nesterov and Todd (centring) step. Moreover, we derive the complexity for the algorithm, and we obtain the best-known iteration bound for the small-update method.

1 Introduction

This paper aims at introducing a corrector–predictor path-following interior-point method (IPM) for a convex and quadratic optimization (CQO) problem over symmetric cones underlying Euclidean Jordan algebra.

This problem includes the symmetric cone optimization (SCO) problems as a special case. In particular, it includes CQO and semidefinite optimization (SDO) problems as special cases. Li and Toh [1] presented an inexact primal–dual infeasible path-following algorithm for convex, quadratic, and symmetric cone programming (CQSCO). Nesterov and Todd [2] provided a theoretical foundation for efficient

Communicated by Florian A. Potra.

B. Kheirfam (✉)
Department of Mathematics, Azarbaijan Shahid Madani University, Tabriz, I.R. Iran
e-mail: b.kheirfam@azaruniv.edu

primal–dual IPMs on a special class of convex optimization, where the associated cone was self-scaled. Schmieta and Alizadeh [3,4] studied primal–dual IPMs for SCO extensively under the framework of Euclidean Jordan algebra. Darvay [5] proposed a full-Newton step primal–dual path-following interior-point algorithm for linear optimization (LO). The search direction of his algorithm is introduced by using an algebraic equivalent transformation of the centring equation, which define the central path and then applying Newton's method for the new system of equations. Recently, Wang and Bai [6] and Bai and Zhang [7] extended Darvay's algorithm for LO to SCO and CQSCO by using Euclidean Jordan algebras, respectively. The first predictor–corrector interior-point algorithms based on Darvay's directions are introduced in [8,9].

By using a high order corrector–predictor approach, Liu and Potra [10] introduced an interior-point algorithm for solving sufficient horizontal linear complementarity problems acting in a wide neighborhood of the central path that does not depend on the handicap of the problem, and has the best known iteration complexity for sufficient LCP. The algorithm reduces the duality gap both in the corrector and the predictor steps, and hence it is more efficient. Potra [11] presented two corrector–predictor interior-point algorithms for solving monotone linear complementarity problems. The algorithms produce a sequence of iterations in the wide neighborhood of the central path, and the complexity is reduced by the use of higher order information. The latter approach ensures superlinear convergence even in the absence of strict complementarity. Mizuno et al. [12] 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. This algorithm is sometimes referred to MTY predictor–corrector algorithm. The MTY predictor–corrector algorithm extended for sufficient LCP in 1995 by Miao [13]. This class extensively studied in the 1991 monograph by Kojima et al. [14], where it provided a unified framework for studying IPMs. Recently, Kheirfam [15] presented a predictor–corrector interior-point algorithm for horizontal linear complementarity problems based on Darvay's technique. Gurtuna et al. [16] presented a corrector–predictor method for solving sufficient linear complementarity problems that do not depend on the handicap of the matrix, so that it can be applied for any sufficient LCP. The algorithm is quadratically convergent for problems having a strictly complementary solution and has the same computational complexity as Miao's algorithm for sufficient LCP. Using higher order predictors, the authors extended a class of corrector–predictor methods that are superlinearly convergent even in the absence of strict complementarity, and the algorithms of this class are globally convergent for general positive starting points.

Motivated by their work, we propose a corrector–predictor path-following algorithm for solving CQSCO based on the new search directions. Our algorithm follows the central path by predictor step (which moves along the central path to the optimal solution) and a corrector step (which moves back toward the central path from the predicted point). The algorithm in the predictor step uses the full Newton-step, while it operates one damped Newton step in the corrector step. We analyze the algorithm and obtain the complexity bound.

The paper is organized as follows: In Sect. 2, we briefly provide the theory of the Euclidean Jordan algebra and their associated symmetric cones. In Sect. 3, we propose the new search directions. In Sect. 4, the corrector–predictor algorithm based

on the new directions is presented. In Sect. 5, we analyze the algorithm and derive the iteration bound. Finally, we give some conclusions in Sect. 6.

2 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 [17–19].

A Euclidean Jordan algebra $\mathcal{J}$ is a finite dimensional vector space endowed with a bilinear map $\circ : \mathcal{J} \times \mathcal{J} \to \mathcal{J}$ iff, for all $x, y \in \mathcal{J}$, $x \circ y = y \circ x$, $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$, where $x^2 = x \circ x$, and there exists an inner product such that $\langle x \circ y, z \rangle = \langle x, y \circ z \rangle$. A 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 $x \in \mathcal{J}$. The set $\mathcal{K}(\mathcal{J}) := \{x^2 : x \in \mathcal{J}\}$ is called the cone of squares of Euclidean Jordan algebra $(\mathcal{J}, \circ, \langle \cdot, \cdot \rangle)$. A cone is symmetric if and only if it is the cone of squares of a Euclidean Jordan algebra [17].

An element $c \in \mathcal{J}$ is said to be idempotent iff $c \neq 0$ and $c \circ c = c$. Two elements x and y are orthogonal if $x \circ y = 0$. An idempotent c is primitive if it is nonzero and cannot be expressed by sum of two other nonzero idempotents. A set of primitive idempotents $\{c_1, c_2, \ldots, c_k\}$ is called a Jordan frame iff $c_i \circ c_j = 0$, for any $i \neq j \in \{1, 2, \ldots, k\}$ and $\sum_{i=1}^{k} c_i = e$. For any $x \in \mathcal{J}$, let r be the smallest positive integer such that $\{e, x, x^2, \ldots, x^r\}$ is linearly dependent; r is called the degree of x and denoted by $\deg(x)$. The rank of $\mathcal{J}$, denoted by $\mathrm{rank}(\mathcal{J})$, is defined as the maximum of $\deg(x)$ over all $x \in \mathcal{J}$.

Theorem 2.1 (Theorem III.1.2 in [17]) *Let* $(\mathcal{J}, \circ, \langle \cdot, \cdot \rangle)$ *be a Euclidean Jordan algebra with* $rank(\mathcal{J}) = r$. *Then, for any* $x \in \mathcal{J}$, *there exists a Jordan frame* $\{c_1, c_2, \ldots, c_r\}$ *and real numbers* $\lambda_1(x), \lambda_2(x), \ldots, \lambda_r(x)$ *such that* $x = \sum_{i=1}^{r} \lambda_i(x) c_i$.

Every $\lambda_i(x)$ is called an eigenvalue of x. We denote $\lambda_{\min}(x) (\lambda_{\max}(x))$ as the minimal (maximal) eigenvalue of x. We can define the following: the square root, $x^{\frac{1}{2}} := \sum_{i=1}^{r} \sqrt{\lambda_i(x)} c_i$, wherever all $\lambda_i \geq 0$, the inverse, $x^{-1} := \sum_{i=1}^{r} \lambda_i(x)^{-1} c_i$, wherever all $\lambda_i \neq 0$, the square $x^2 := \sum_{i=1}^{r} \lambda_i(x)^2 c_i$; and the trace $\mathrm{Tr}(x) := \sum_{i=1}^{r} \lambda_i(x)$.

Since "$\circ$" is bilinear map, for every $x \in \mathcal{J}$, there exists a linear operator $L(x)$ such that for every $y \in \mathcal{J}$, $x \circ y := L(x)y$. In particular, $L(x)e = x$ and $L(x)x = x^2$. For each $x \in \mathcal{J}$, we define

$$P(x) := 2L(x)^2 - L(x^2),$$

where $L(x)^2 := L(x)L(x)$. The map $P(x)$ is called the quadratic representation of $\mathcal{J}$. For any $x, y \in \mathcal{J}$, x, and y are said to be operator commutable iff $L(x)$ and $L(y)$ commute, i.e., $L(x)L(y) = L(y)L(x)$. In other words, x and y operator commutable iff, for all $z \in \mathcal{J}$, $x \circ (y \circ z) = y \circ (x \circ z)$ (see [4]). For any $x, y \in \mathcal{J}$, the inner product is defined as $\langle x, y \rangle := Tr(x \circ y)$, and the Frobenius norm of x as follows: $\|x\|_F := \sqrt{\mathrm{Tr}(x^2)} = \sqrt{\sum_{i=1}^{r} \lambda_i^2(x)}$. Observe that $\|e\|_F = \sqrt{r}$, since identity element e has eigenvalue 1.

We say that two elements x and y in $\mathcal{J}$ are similar, and briefly denoted as $x \sim y$, if and only if x and y share the same set of eigenvalues. Let $\text{int}\mathcal{K}$ denotes the interior of the symmetric cone $\mathcal{K}$. We say $x \in \mathcal{K}$ if and only if $\lambda_i \geq 0$, and $x \in \text{int}\mathcal{K}$ if and only if $\lambda_i > 0$, for all $i = 1, 2, \ldots, r$. We also say x is positive semidefinite (positive definite) iff $x \in \mathcal{K}$ ($x \in \text{int}\mathcal{K}$).

Lemma 2.1 (Proposition 21 in [4]) *Let $x, s, u \in \text{int}\mathcal{K}$. Then*

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

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

Lemma 2.2 (Proposition 3.2.4 in [19]) *Let $x, s \in \text{int}\mathcal{K}$, and w be the scaling point of x and s. Then*

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

Lemma 2.3 (Lemma 14 in [4]) *Let $x, s \in \mathcal{J}$. Then*

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

Lemma 2.4 (Lemma 30 in [4]) *Let $x, s \in \text{int}\mathcal{K}$. Then*

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

Lemma 2.5 (Theorem 4 in [20]) *Let $x, s \in \text{int}\mathcal{K}$. Then*

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

Lemma 2.6 (Lemma 2.11 in [6]) *Let $t > 0$ and $v \in \mathcal{K}$. Then*

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

Lemma 2.7 (Lemma 4.1 in [6]) *Let $x(\alpha) := x + \alpha\Delta x$, $s(\alpha) := s + \alpha\Delta s$ for $0 \leq \alpha \leq 1$. Let $x, s \in \text{int}\mathcal{K}$, and $x(\alpha) \circ s(\alpha) \in \text{int}\mathcal{K}$ for $0 \leq \alpha \leq \bar{\alpha}$. Then $x(\bar{\alpha}), s(\bar{\alpha}) \in \text{int}\mathcal{K}$.*

3 The New Search Directions

Consider the CQSCO problem as

$$\min\left\{\langle c, x\rangle + \frac{1}{2}\langle x, Qx\rangle : \; Ax = b, \; x \in \mathcal{K}\right\}, \tag{P}$$

along with its dual problem as

$$\max \left\{ b^T y - \frac{1}{2} \langle x, Qx \rangle : \ A^T y + s - Qx = c, \ s \in \mathcal{K} \right\}, \tag{D}$$

where c and the rows a_i of A lie in $\mathcal{J}$, and $b, y \in R^m$, $\langle x, s \rangle = \text{Tr}(x \circ s)$ stands for the trace inner product in $\mathcal{J}$. $Q : \mathcal{J} \longrightarrow \mathcal{J}$ is a symmetric linear operator with the property $\langle x, Qx \rangle \geq 0$, $x \in \mathcal{J}$. Furthermore, $Ax = b$ and $A^T y + s - Qx = c$ mean that

$$\langle a_i, x \rangle = b_i, \ i = 1, 2, \ldots, m, \quad \text{and} \quad \sum_{i=1}^{m} y_i a_i + s - Qx = c,$$

respectively. 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

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

and the matrix A is of rank m. Under the IPC, finding an optimal solution of (P) and (D) is equivalent to solving the following system

$$Ax = b, \ A^T y + s - Qx = c, \ x \circ s = 0, \ x, s \in \mathcal{K}. \tag{1}$$

The basic idea of primal–dual IPMs is to replace the third equation in (1), the so-called complementarity condition for (P) and (D), by the parameterized equation $x \circ s = \mu e$ with $\mu > 0$. The system (1) can be written as

$$Ax = b, \ A^T y + s - Qx = c, \ x \circ s = \mu e. \tag{2}$$

For each $\mu > 0$, the perturbed system (2) has a unique solution $(x(\mu), y(\mu), s(\mu))$, and we call $x(\mu)$ and $(y(\mu), s(\mu))$ the μ-centers of problem (P) and (D), respectively. The set of μ-centers gives a homotopy path, 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, the limit yields an ϵ-approximate solution for (P) and (D) [18].

Lemma 3.1 (Lemma 28 in [3]) *Let x, s, and w be in some Euclidean Jordan algebra $\mathcal{J}$, $x, s \in \text{int}\mathcal{K}$ and w invertible. Then $x \circ s = \mu e$ iff $P(w^{-\frac{1}{2}})x \circ P(w^{\frac{1}{2}})s = \mu e$.*

Suppose that the iterate (x, y, s) is primal–dual feasible solution, and we may write the relaxed complementarity condition $x \circ s = \mu e$ as $P(w^{-\frac{1}{2}})x \circ P(w^{\frac{1}{2}})s = \mu e$. Therefore, applying Newton's method to the resulting system leads us to the linear system

$$A\Delta x = 0, \ A^T \Delta y + \Delta s - Q\Delta x = 0,$$
$$P\left(w^{-\frac{1}{2}}\right)\Delta x \circ P\left(w^{\frac{1}{2}}\right)s + P\left(w^{-\frac{1}{2}}\right)x \circ P\left(w^{\frac{1}{2}}\right)\Delta s \tag{3}$$
$$= \mu e - P\left(w^{-\frac{1}{2}}\right)x \circ P\left(w^{\frac{1}{2}}\right)s.$$

For the choice of

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

we have $P\left(w^{-\frac{1}{2}}\right)x = P\left(w^{\frac{1}{2}}\right)s$, which gives the Nesterov–Todd (NT) direction [2,21]. Define

$$v := \frac{P(w^{-\frac{1}{2}})x}{\sqrt{\mu}} = \frac{P(w^{\frac{1}{2}})s}{\sqrt{\mu}} = \left(\frac{P(w^{-\frac{1}{2}})x \circ P(w^{\frac{1}{2}})s}{\mu}\right)^{\frac{1}{2}}. \tag{4}$$

Note that $x \circ s = \mu e$ if and only if $v = e$ (Proposition 5.7.2 in [19]). Therefore, we have $\mu v = \mu e$. Combining the last equation with the system (3), we have

$$A\Delta x = 0, \quad A^T \Delta y + \Delta s - Q\Delta x = 0,$$
$$P\left(w^{-\frac{1}{2}}\right)\Delta x \circ P\left(w^{\frac{1}{2}}\right)s + P\left(w^{-\frac{1}{2}}\right)x \circ P\left(w^{\frac{1}{2}}\right)\Delta s$$
$$= \mu v - P\left(w^{-\frac{1}{2}}\right)x \circ P\left(w^{\frac{1}{2}}\right)s. \tag{5}$$

Let us denote

$$\bar{A} := \sqrt{\mu}AP\left(w^{\frac{1}{2}}\right), \quad d_x := \frac{P(w^{-\frac{1}{2}})\Delta x}{\sqrt{\mu}}, \quad d_s := \frac{P(w^{\frac{1}{2}})\Delta s}{\sqrt{\mu}}. \tag{6}$$

Consequently, the system (5) can be written in the following form

$$\bar{A}d_x = 0, \quad \bar{A}^T \frac{\Delta y}{\mu} + d_s - \bar{Q}d_x = 0, \quad d_x + d_s = e - v, \tag{7}$$

where $\bar{Q} := P(w^{\frac{1}{2}})QP(w^{-\frac{1}{2}})$. For the analysis of the algorithm, we define a norm-based proximity measure $\sigma(x, s; \mu)$ as follows:

$$\sigma(v) := \sigma(x, s; \mu) := \|e - v\|_F, \tag{8}$$

where v is defined as (4). Due to the first two equations of the system (7), we have

$$\langle d_x, d_s \rangle = \left\langle d_x, -\bar{A}^T \frac{\Delta y}{\mu} + \bar{Q}d_x \right\rangle = \left\langle d_x, -\bar{A}^T \frac{\Delta y}{\mu} \right\rangle + \langle d_x, \bar{Q}d_x \rangle = \langle d_x, \bar{Q}d_x \rangle \geq 0. \tag{9}$$

4 The Corrector–Predictor Algorithm

Here, we propose a corrector–predictor algorithm based on the modified NT-search directions, which uses these directions in the both corrector and predictor steps. Firstly, we define the τ-neighborhood of the central path as follows:

$$\mathcal{N}(\tau, \mu) :=$$
$$\left\{ (x, s) : Ax = b, \ A^T y + s - Qx = c, \ x, s \in \text{int}\mathcal{K}, \ y \in R^m, \ \sigma(x, s; \mu) \leq \tau \right\}.$$

The algorithm starts with (x, y, s) in the τ-neighborhood, which certainly holds at (x^0, y^0, s^0) since $\sigma(x^0, s^0; \mu^0) = 0$. If for the current iterate (x, y, s), $r\mu > \epsilon$, then the algorithm performs corrector and predictor steps. In corrector steps, by solving the system (7), for the scaled-directions $(d_x, \Delta y, d_s)$, that is,

$$\bar{A}d_x = 0, \quad \bar{A}^T \frac{\Delta y}{\mu} + d_s - \bar{Q}d_x = 0, \quad d_x + d_s = e - v, \tag{10}$$

and using (6) for $(\Delta x, \Delta s)$, we obtain

$$x := x + \Delta x, \quad y := y + \Delta y, \quad s := s + \Delta s.$$

In the predictor (affine-scaling) steps, 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

$$\bar{A}d_x^p = 0, \quad \bar{A}^T \frac{\Delta^p y}{\mu} + d_s^p - \bar{Q}d_x^p = 0, \quad d_x^p + d_s^p = -v, \tag{11}$$

where

$$d_x^p := \frac{P(w^{-\frac{1}{2}})\Delta^p x}{\sqrt{\mu}}, \quad d_s^p := \frac{P(w^{\frac{1}{2}})\Delta^p s}{\sqrt{\mu}}. \tag{12}$$

We denote the iterates after a predictor step by

$$x^p := x + \theta\Delta^p x, \quad y^p := y + \theta\Delta^p y, \quad s^p := s + \theta\Delta^p s, \quad \mu^p := (1 - \theta)\mu,$$

where $\theta \in]0, 1[$. The iterate (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. A formal description of the algorithm is given in Fig. 1.

5 Analysis

In this Section, we deal with the analysis of the previous algorithm. At the analysis of the affine-scaling step, we will give sufficient conditions for strict feasibility and the

---

Corrector – Predictor Method

Input :

an accuracy parameter $\epsilon > 0$;

barrier update parameter θ, $0 < \theta < 1$ (default $\theta = \frac{1}{4\sqrt{2r}}$);

proximity parameter τ, $0 < \tau < 1$ (default $\tau = \frac{5}{11}$);

a strictly feasible point $(x^0, y^0, s^0), r\mu^0 = \langle x^0, s^0 \rangle$

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 $\mathrm{Tr}(x \circ s) > \epsilon$ **do**

begin

solve system (10) and via (6) to obtain $(\Delta x, \Delta y, \Delta s)$;

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

solve system (11) and via (12) to obtain $(\Delta^p x, \Delta^p y, \Delta^p s)$;

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

$\mu := (1 - \theta)\mu$;

end

end

Fig. 1 The algorithm

effect on the proximity measure if the proximity measure does not exceed the proximity parameter. At the corrector step, we obtain the relation between the proximity parameters. The following lemmas are essential for analysis of the algorithm.

Lemma 5.1 (Lemma 8 in [22]) *Let $x, s \in \mathcal{J}$ with $\mathrm{Tr}(x \circ s) \geq 0$. Then*

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

Lemma 5.2 (Lemma 9 in [22]) *Let $x, s \in \mathrm{int}\mathcal{K}$ and $\mu > 0$. Assume that $\sigma := \sigma(v)$. Then, $1 - \sigma \leq \lambda_i(v) \leq 1 + \sigma$, for $i = 1, 2, \ldots, r$.*

5.1 The Affine-Scaling Step

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

Lemma 5.3 *Let* $x, s \in \text{int}\mathcal{K}$ *and* $\mu > 0$ *such that* $\sigma := \sigma(x, s; \mu) < 1$. *Furthermore, let* $0 < \theta < 1$. *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* $x^p, s^p \in \text{int}\mathcal{K}$ *if* $K(\sigma, \theta, r) > 0$, *where*

$$K(\sigma, \theta, r) = (1 - \sigma)^2 - \frac{\sqrt{2}r\theta^2(1 + \sigma)^2}{4(1 - \theta)}.$$

Proof Let us introduce

$$x^p(\alpha) := x + \alpha\theta\Delta^p x, \quad s^p(\alpha) := s + \alpha\theta\Delta^p s,$$

for $\alpha \in [0, 1]$. Using (4) and (12), we have

$$x^p(\alpha) = \sqrt{\mu}P\left(w^{\frac{1}{2}}\right)\left(v + \alpha\theta d_x^p\right), \quad s^p(\alpha) = \sqrt{\mu}P\left(w^{-\frac{1}{2}}\right)\left(v + \alpha\theta d_s^p\right). \quad (13)$$

Since $P(w^{\frac{1}{2}})$ and $P(w^{-\frac{1}{2}})$ are automorphisms of $\text{int}\mathcal{K}$ (Theorem III.2.1 in [17]), by (13), x^p and s^p belong to $\text{int}\mathcal{K}$ if and only if $v + \theta d_x^p$ and $v + \theta d_s^p$ belong to $\text{int}\mathcal{K}$. Therefore, using the third equation of (11), we obtain

$$\begin{aligned} v_x^p(\alpha) \circ v_s^p(\alpha) &= (v + \alpha\theta d_x^p) \circ (v + \alpha\theta d_s^p) = v^2 + \alpha\theta v \circ (d_x^p + d_s^p) + \alpha^2\theta^2 d_x^p \circ d_s^p \\ &= v^2 + \alpha\theta v \circ (-v) + \alpha^2\theta^2 d_x^p \circ d_s^p = (1 - \alpha\theta)v^2 + \alpha^2\theta^2 d_x^p \circ d_s^p. \end{aligned} \quad (14)$$

From the above relation and Lemma 2.3, we obtain

$$\begin{aligned} \lambda_{\min}\left(\frac{v_x^p(\alpha) \circ v_s^p(\alpha)}{1 - \alpha\theta}\right) &= \lambda_{\min}\left(v^2 + \frac{\alpha^2\theta^2}{1 - \alpha\theta}d_x^p \circ d_s^p\right) \\ &\geq \lambda_{\min}\left(v^2\right) - \frac{\alpha^2\theta^2}{1 - \alpha\theta}\|d_x^p \circ d_s^p\|_F \geq \lambda_{\min}\left(v^2\right) - \frac{\theta^2}{1 - \theta}\|d_x^p \circ d_s^p\|_F, \quad (15) \end{aligned}$$

the last inequality follows by $f(\alpha) = \frac{\alpha^2\theta^2}{1 - \alpha\theta}$ for $0 \leq \alpha \leq 1$, and each fixed $0 < \theta < 1$ is strictly increasing. From Lemma 5.1, the third equation of (11) and Lemma 5.2 we get

$$\lambda_{\min}\left(\frac{v_x^p(\alpha) \circ v_s^p(\alpha)}{1 - \alpha\theta}\right) \geq (1 - \sigma)^2 - \frac{\sqrt{2}r\theta^2(1 + \sigma)^2}{4(1 - \theta)} = K(\sigma, \theta, r) > 0. \quad (16)$$

This implies that $\det\left(v_x^p(\alpha) \circ v_s^p(\alpha)\right) > 0$ for $0 \leq \alpha \leq 1$. From Lemma 2.7, we have $v + \theta d_x^p \in \text{int}\mathcal{K}$ and $v + \theta d_s^p \in \text{int}\mathcal{K}$ for $\alpha = 1$. This proves the lemma. □

By following (4), we define

$$v^p := \frac{P(w^p)^{\frac{1}{2}} s^p}{\sqrt{\mu^p}}, \tag{17}$$

where w^p is the scaling point of x^p and s^p. Using (13) with $\alpha = 1$, (17) and Lemmas 2.1 and 2.2, we have

$$(v^p)^2 \sim \frac{P(x^p)^{\frac{1}{2}} s^p}{\mu^p} = \frac{\mu\left(P(P(w)^{\frac{1}{2}}(v + \theta d_x^p))^{\frac{1}{2}} P(w)^{-\frac{1}{2}}(v + \theta d_s^p)\right)}{\mu(1 - \theta)}$$
$$\sim \frac{P(v + \theta d_x^p)^{\frac{1}{2}}(v + \theta d_s^p)}{1 - \theta}. \tag{18}$$

Using Lemma 2.5 and (16) with $\alpha = 1$, we obtain

$$\lambda_{\min}((v^p)^2) = \lambda_{\min}\left(P\left(\frac{v + \theta d_x^p}{\sqrt{1 - \theta}}\right)^{\frac{1}{2}}\left(\frac{v + \theta d_s^p}{\sqrt{1 - \theta}}\right)\right)$$
$$\geq \lambda_{\min}\left(\left(\frac{v + \theta d_x^p}{\sqrt{1 - \theta}}\right) \circ \left(\frac{v + \theta d_s^p}{\sqrt{1 - \theta}}\right)\right) \geq K(\sigma, \theta, r). \tag{19}$$

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

Lemma 5.4 *Let* $\sigma := \sigma(x, s; \mu) < 1$, $\mu^p = (1 - \theta)\mu$, *where* $0 < \theta < 1$, $K(\sigma, \theta, r) > 0$ *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*

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

Proof From Lemma 5.3 we deduce that the affine-scaling step is strictly feasible. Using (18), (19), (14) with $\alpha = 1$ and Lemma 2.4, we obtain

$$\sigma^p := \sigma(x^p, s^p; \mu^p) = \left\|\left(e - (v^p)^2\right)(e + v^p)^{-1}\right\|_F \leq \frac{1}{1 + \lambda_{\min}(v^p)}\left\|e - (v^p)^2\right\|_F$$

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

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

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

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

On the other hand, by Lemma 5.2, we have

$$\left\| e - v^2 \right\|_F = \left\| e - v + v - v^2 \right\|_F \leq \sigma + \lambda_{\max}(v) \| e - v \|_F = \sigma^2 + 2\sigma. \quad (20)$$

Finally, we get

$$\sigma^p \leq \frac{\sigma^2 + 2\sigma + \frac{\sqrt{2}\theta^2}{4(1-\theta)} r (1 + \sigma)^2}{1 + \sqrt{K(\sigma, \theta, r)}},$$

which completes the proof.   □

5.2 The Corrector Step

The following lemma gives a condition for strict feasibility of full NT-step. Its proof is similar to the proof of Lemma 11 in [22].

Lemma 5.5 *Let* $\sigma(v)$ *be defined as* (8)*, and* $(x, s) \in \mathrm{int}\mathcal{K} \times \mathrm{int}\mathcal{K}$*. If* $\sigma(v) < \frac{2}{1 + \sqrt{1 + \sqrt{2}}}$*, then the full-NT step for CQSCO is strictly feasible.*

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

Lemma 5.6 *Let* $x^+ = x + \Delta x$ *and* $s^+ = s + \Delta s$ *and* $\sigma(v) < \frac{2}{1 + \sqrt{1 + \sqrt{2}}}$*. Then*

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

Proof Let $v^+ = \frac{P(w^+)^{\frac{1}{2}} s^+}{\sqrt{\mu}}$ then, similar to (18) and (19), we respectively have

$$\left(v^+ \right)^2 \sim P(v + d_x)^{\frac{1}{2}} (v + d_s), \quad \lambda_{\min}\left(\left(v^+ \right)^2 \right) \geq \lambda_{\min}(v + d_x \circ d_s) \geq 1 - \sigma - \frac{1}{2\sqrt{2}} \sigma^2.$$

Using the above relations, Lemma 2.6 with $t = 1$ and Lemma 2.4, we have

$$\sigma\left(x^+, s^+; \mu \right) = \left\| e - \frac{P(w^+)^{\frac{1}{2}} s^+}{\sqrt{\mu}} \right\|_F = \left\| e - v^+ \right\|_F$$

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

$$\leq \frac{1}{1 + \lambda_{\min}(v^+)} \|e - v - d_x \circ d_s\|_F \leq \frac{\sigma(v) + \frac{1}{2\sqrt{2}}\sigma(v)^2}{1 + \sqrt{1 - \sigma(v) - \frac{1}{2\sqrt{2}}\sigma(v)^2}}.$$

This completes the proof. □

The next lemma gives the effect of full NT-step on duality gap. Its proof is similar to the proof of Lemma 12 in [22].

Lemma 5.7 *If $\sigma(v) < \frac{2}{1+\sqrt{1+\sqrt{2}}}$, then*

$$\mathrm{Tr}\left(x^+ \circ s^+\right) < 2\mu r \left(\frac{3 + \sqrt{1 + \sqrt{2}}}{1 + \sqrt{1 + \sqrt{2}}}\right).$$

5.3 The Effect on Duality Gap After a Main Iteration

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

Lemma 5.8 *Let $x, s \in \mathrm{int}\mathcal{K}$, $\mu > 0$ such that $\sigma := \sigma(x, s; \mu) < 1$, and $0 < \theta \leq 1$. If x^p and s^p are the iterates obtained after the affine-scaling step of the algorithm, then*

$$\mathrm{Tr}\left(x^p \circ s^p\right) \leq \frac{4r\mu^p}{1 - \theta}.$$

Proof Letting $\alpha = 1$ in (13) and (14), we obtain

$$\mathrm{Tr}\left(x^p \circ s^p\right) = \mathrm{Tr}\left(\sqrt{\mu}P(w)^{\frac{1}{2}}\left(v + \theta d_x^p\right) \circ \sqrt{\mu}P(w)^{-\frac{1}{2}}\left(v + \theta d_s^p\right)\right)$$

$$= \mu\mathrm{Tr}\left(\left(v + \theta d_x^p\right) \circ \left(v + \theta d_s^p\right)\right) = \mu\mathrm{Tr}\left((1 - \theta)v^2 + \theta^2 d_x^p \circ d_s^p\right)$$

$$\leq \mu\left(1 - \theta + \frac{1}{2}\theta^2\right)\mathrm{Tr}\left(v^2\right) \leq \left(1 - \frac{1}{2}\theta\right)\mathrm{Tr}(x \circ s). \tag{21}$$

The first inequality follows from the third equation of (11). Since x and s are obtained by a full NT-step of the algorithm, by Lemma 5.7, we have

$$\mathrm{Tr}\left(x^p \circ s^p\right) \leq 2\left(1 - \frac{1}{2}\theta\right)r\left(\frac{3 + \sqrt{1 + \sqrt{2}}}{1 + \sqrt{1 + \sqrt{2}}}\right)\frac{\mu^p}{1 - \theta} \leq \frac{4r\mu^p}{1 - \theta}.$$

The proof is complete. □

5.4 Fixing the Parameter

In the following, 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 before.

Let (x, y, s) be the iterate at the start of a main iteration with $x \in \text{int}\mathcal{K}$ and $s \in \text{int}\mathcal{K}$ such that $\sigma(x, s; \mu) \leq \tau$. After a corrector step, by Lemma 5.6, one has

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

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

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

Following the affine-scaling step and a μ-update, by Lemma 5.4, one has

$$\sigma^p = \sigma\left(x^p, s^p; \mu^p\right) \leq \frac{1 + 2\sigma^2 - K(\sigma, \theta, r)}{1 + \sqrt{K(\sigma, \theta, r)}}, \qquad (22)$$

where $K(\sigma, \theta, r)$ is defined as Lemma 5.3. It can be easily verified that the right-hand side of (22) is monotonically increasing with respect to σ, which means that

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

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

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

At this stage, for a fixed threshold $\tau = \frac{5}{11}$ and update parameter $\theta = \frac{1}{4\sqrt{2r}}$, after some simple calculation, we obtain $\sigma^p \leq 0.4468 < \tau = 0.4545$. Thus, the algorithm is well defined. Moreover, $K(\sigma, \theta, r) \geq 0.4493$. This implies, by Lemma 5.3, that the predictor step is strictly feasible.

5.5 Complexity Bound

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

Lemma 5.9 *Let x^0 and s^0 be strictly feasible, $\mu^0 = \frac{\text{Tr}(x^0 \circ s^0)}{r}$ and $\sigma(x^0, s^0; \mu^0) < \tau$. Moreover, let x^k and s^k be the iterates obtained after k iterations. Then $\text{Tr}(x^k \circ s^k) \leq \epsilon$, for $k \geq 1 + \left\lceil \frac{1}{\theta} \log \frac{4\text{Tr}(x^0 \circ s^0)}{\epsilon} \right\rceil$.*

Proof It follows from Lemma 5.8 that

$$\mathrm{Tr}\left(x^k \circ s^k\right) < \frac{4r\mu^k}{1-\theta} = 4r(1-\theta)^{k-1}\mu^0 = 4(1-\theta)^{k-1}\mathrm{Tr}\left(x^0 \circ s^0\right).$$

Then the inequality $\mathrm{Tr}(x^k \circ s^k) \leq \epsilon$ holds if $4(1-\theta)^{k-1}\mathrm{Tr}(x^0 \circ s^0) \leq \epsilon$. Taking logarithms and using $\log(1+\theta) \leq \theta, \theta \geq -1$, we observe that the above inequality holds if $-\theta(k-1) + \log 4\mathrm{Tr}(x^0 \circ s^0) \leq \log \epsilon$. This implies the result. □

Theorem 5.1 *Let* $\tau = \frac{5}{11}$ *and* $\theta = \frac{1}{4\sqrt{2r}}$. *Then the algorithm is well defined and the algorithm requires at most*

$$O\left(\sqrt{r}\log \frac{\mathrm{Tr}\left(x^0 \circ s^0\right)}{\epsilon}\right)$$

iterations. The output is a primal–dual pair (x, s) *satisfying* $\mathrm{Tr}(x \circ s) \leq \epsilon$.

Proof Let $\tau = \frac{5}{11}$ and $\theta = \frac{1}{4\sqrt{2r}}$, the desired result follows immediately from Lemma 5.9. □

6 Conclusions

We presented a corrector–predictor algorithm and its derived complexity for solving CQO over symmetric cones. The complexity results show that the proposed algorithm enjoys the best known complexity bound for interior point methods. An interesting topic for further research may be the development of the analysis to the Cartesian $P_*(\kappa)$ linear complementarity problems over symmetric cones.

References

1. Li, L., Toh, K.C.: A polynomial-time inexact interior-point method for convex quadratic symmetric cone programming. J. Math. Ind. **2**, 199–212 (2010)
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. Schmieta, S.H., Alizadeh, F.: Extension of primal-dual interior-point algorithms to symmetric cones. Math. Program. **96**(3), 409–438 (2003)
4. Schmieta, S.H., Alizadeh, F.: Associative and Jordan algebras and polynomial time interior-point algorithms for symmetric cones. Math. Oper. Res. **26**, 543–564 (2001)
5. Darvay, Z.: New interior-point algorithms in linear programming. Adv. Model. Optim. **5**(1), 51–92 (2003)
6. Wang, G.Q., Bai, Y.Q.: A new full NesterovTodd step primal-dual path-following interior-point algorithm for symmetric optimization. J. Optim. Theory Appl. **154**(3), 966–985 (2012)
7. Bai, Y., Zhang, L.: A full-Newton step interior-point algorithm for symmetric cone convex quadratic optimization. J. Ind. Manag. Optim. **7**(4), 891–906 (2011)
8. Darvay, Z.: A new predictor-corrector algorithm for linear programming. Alkalmazott matematikai lapok **22**, 135–161 (2005). (In hungarian)

9. Darvay, Z.: A predictor-corrector algorithm for linearly constrained convex optimization. Studia Univ. Babes-Bolyai, Informatica LIV(2), (2009).

10. Liu, X., Potra, F.A.: Corrector-predictor methods sufficient linear complementarity problems in a wide neighborhood of the central path. SIAM J. Optim. **17**(3), 871–890 (2006)

11. Potra, F.A.: Corrector-predictor methods for monotone linear complementarity problems in a wide neighborhood of the central path. Math. Program. **111**, 243–272 (2008)

12. Mizuno, S., Todd, M.J., Ye, Y.: On adaptive-step primal-dual interior point algorithms for linear programming. Math. Oper. Res. **18**, 964–981 (1993)

13. Miao, J.: A quadratically convergent $\mathcal{O}((1+\kappa)\sqrt{n}L)$-iteration algorithm for the $P_*(\kappa)$-matrix linear complementarity problem. Math. Program. **69**, 355–368 (1995)

14. Kojima, M., Megiddo, N., Noma, T., Yoshise, A.: A unified approach to interior point algorithms for linear complementarity problems. Lecture notes in Comput. Sci. 538, Springer-Verlag, Berlin (1991).

15. Kheirfam, B.: A predictor-corrector interior-point algorithm for $P_*(\kappa)$ horizontal linear complementarity problem. Numer. Algorithms (2013). doi:10.1007/s11075-013-9738-3

16. Gurtuna, F., Petra, C., Potra, F., Shevchenko, O., Vancea, A.: Corrector-predictor methods for sufficient linear complementarity problems. Comput. Optim. Appl. **48**, 453–485 (2011)

17. Faraut, J., Korányi, A.: Analysis on symmetric cones. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York (1994)

18. Faybusovich, L.: Linear systems in Jordan algebras and primal-dual interior-point algorithms. J. Comput. Appl. Math. **86**, 149–175 (1997)

19. 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).

20. Sturm, J.F.: Similarity and other spectral relations for symmetric cones. Algebra Appl. **312**(1–3), 135–154 (2000)

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

22. Kheirfam, B., Mahdavi-Amiri, N.: A new interior-point algorithm based on modified Nesterov–Todd direction for symmetric cone linear complementarity problem. Optim. Lett. **8**(3), 1017–1029 (2014)