Math Meth Oper Res (2011) 73:75–90
DOI 10.1007/s00186-010-0334-1

ORIGINAL ARTICLE

Polynomial complexity of an interior point algorithm with a second order corrector step for symmetric cone programming

Jian Zhang · Kecun Zhang

Received: 24 June 2010 / Accepted: 25 October 2010 / Published online: 13 November 2010

Abstract In this paper, we propose a second order interior point algorithm for symmetric cone programming using a wide neighborhood of the central path. The convergence is shown for commutative class of search directions. The complexity bound is $O(r^{3/2}\log\epsilon^{-1})$ for the NT methods, and $O(r^2\log\epsilon^{-1})$ for the XS and SX methods, where r is the rank of the associated Euclidean Jordan algebra and $\epsilon > 0$ is a given tolerance. If the staring point is strictly feasible, then the corresponding bounds can be reduced by a factor of $r^{3/4}$. The theory of Euclidean Jordan algebras is a basic tool in our analysis.

Keywords Linear programming · Symmetric cone · Euclidean Jordan algebra · Interior point method · Polynomial complexity

Mathematics Subject Classification (2000) 90C05 · 90C25 · 90C51

1 Introduction

There is extensive literatures on the analysis of interior point methods (IPMs) for symmetric cone programming (SCP). Nonnegative orthants, second-order cones, and positive semidefinite cones are important special cases of symmetric cones. SCP includes solving problems such as linear programming (LP), semidefinite programming (SDP) and second order cone programming (SOCP) problems. The foundation

J. Zhang (✉) · K. Zhang
Faculty of Science, Xi'an Jiaotong University, P.O.Box 2325, 710049 Xi'an, People's Republic of China
e-mail: jianzhang1985@gmail.com

K. Zhang
e-mail: zhangkc@mail.xjtu.edu.cn

for solving these problems using IPMs was laid by Nesterov and Nemirovskii (1994). These methods were primarily either primal or dual based. Later, Nesterov and Todd (1997) introduced symmetric primal dual interior point algorithms on a special class of cones called self-scaled cones, which allowed a symmetric treatment of the primal and the dual. It turns out that self-scaled cones are precisely the symmetric cones, and thus Nesterov and Todd (NT) algorithm was the first primal-dual method for optimization over symmetric cones. Later, several authors derived IPMs for SDP and SOCP.

Monteiro and Zhang (1998) designed unified primal dual path following algorithms for SDP based on so-called commutative class of search directions. These search directions include the popular directions such as the NT, XS and SX directions. Tsuchiya (1999) used Jordan algebraic techniques to extend the methods of Monteiro and Zhang (1998) to SOCP. He obtained polynomial time iteration complexity for short step, semi-long step, and long step algorithms, but restricted his analysis to the NT and XS methods. Later, Schmieta and Alizadeh (2003) extended this kind of approach to SCP. They proved polynomial iteration complexities for variants of the short, semi-long, and long step path following algorithms based on commutative class directions. For short step method, the iteration bound is $O(\sqrt{\kappa}\sqrt{r}\log\epsilon^{-1})$, and for semi-long and long step methods, the iteration bound is $O(\sqrt{\kappa}r\log\epsilon^{-1})$ (The meaning of κ will be mentioned in Sect. 4). Muramatsu (2002) proposed a subclass of the commutative class for which he proved polynomial complexities of the IPMs using semi-long and long steps. The author also obtained $O(r\log\epsilon^{-1})$ and $O(r^{1+\frac{|q-1|}{2(q+1)}}\log\epsilon^{-1})$ complexities respectively. Monteiro (1998) proposed polynomial convergence short step algorithms for SDP based on the Menteiro and Zhang (MZ) family of search directions. This is a wider class of directions which includes the commutative class of directions. This approach was later extended to SOCP by Monteiro and Tsuchiya (2000). Schmieta and Alizadeh (2001) showed that Monteiro's short step analysis (Monteiro 1998) applies to SCP whose underlying Euclidean Jordan algebra is derived from an associative algebra. The complexity bound is $O(\sqrt{r}\log\epsilon^{-1})$. Long step methods have worse theoretical iteration complexity bounds than short step methods. However, they have practical advantages.

The aforementioned algorithms concern on the so-called feasible IPMs. They require the starting point is strictly feasible. This requirement may be difficult to obtain in practical implementation. Unlike feasible IPMs, infeasible IPMs do not require the iterates be feasible to the relevant linear systems but only be in the interior of the cone constraints. At the same time, the main difficulties in analysis of infeasible IPMs lie in the nonorthogonality of search directions. Zhang (1998) extended an infeasible IPM from LP to SDP using the XS direction. Recently, based on the commutative class of search directions, Rangarajan (2006) established a long step infeasible interior point algorithm for SCP with polynomial complexity $O(\sqrt{\kappa}r^2\log\epsilon^{-1})$.

For LP, predictor corrector methods have attracted much attention due to its high efficiency. Many predictor corrector IPMs adopted a heuristics proposed first by Mehrotra in his paper (Mehrotra 1992). It was an infeasible second order IPM. Zhang and Zhang (1995) established convergence theory and complexity bounds for two Mehrotra type algorithms, both using second corrector steps. Recently Salahi and Mahdavi-Amiri (2006) proposed feasible Mehrotra type second order corrector IPMs for LP and established the corresponding complexity bounds. But for general SCP, there has been no literature concerning IPMs using a second order corrector step. This motivated us to consider a Mehrotra type algorithm for SCP using a second order corrector step. We use the wide neighborhood in our algorithm. The complexity for infeasible starting points is $O(r^{3/2}\sqrt{\kappa}\log\epsilon^{-1})$. If a strictly feasible starting point is available, then the complexity bound can be lowered to $O(r^{3/4}\sqrt{\kappa}\log\epsilon^{-1})$.

The rest of this paper is organized as follows. In the next section, we give a brief introduction to Euclidean Jordan algebras. In Sect. 3, we present our interior point algorithm for SCP. The complexity analysis of the algorithm is presented in Sect. 4. Some final remarks are given in Sect. 5.

2 Euclidean Jordan algebras

In this section we outline a minimal foundation of the theory of Euclidean Jordan algebras. This theory serves as our basic toolbox for the analysis of IPMs. Our presentation mostly follows Faraut and Korányi (1994).

Let V be a finite dimensional vector space over the field of real numbers. Then $\langle V, \circ\rangle$ is called a Jordan algebra if a bilinear map $V \times V \to V$ denoted by "$\circ$" is defined which satisfies $x \circ y = y \circ x$ and $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$ for any $x, y \in V$, where $x^2 = x \circ x$. A Jordan algebra $\langle V, \circ\rangle$ is called Euclidean if an associative inner product "$\langle\cdot,\cdot\rangle$" is defined, i.e., $\langle x \circ y, z\rangle = \langle x, y \circ z\rangle$ holds for any $x, y, z \in V$.

A Jordan algebra has an identity, if there exists a unique element $e \in V$ such that $x \circ e = e \circ x = x$ holds for all $x \in V$. The set $K = \{x^2 : x \in V\}$ is called the cone of squares of Euclidean Jordan algebra $(V, \circ, \langle\cdot,\cdot\rangle)$. A cone is symmetric if and only if it is the cone of squares of some Euclidean Jordan algebra. An element $c \in V$ is called idempotent if $c \circ c = c$. 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. For any $x \in V$, let $m(x)$ be the minimal positive integer such that $\{e, x, \ldots, x^{m(x)}\}$ is linearly dependent. The rank of V, denoted by $rank(V)$, is defined by $\max\{m(x), x \in V\}$. A set of primitive idempotents $\{c_1, \ldots, c_k\}$ is called a Jordan frame if $c_i \circ c_j = 0$ for any $i \neq j \in \{1, \ldots, k\}$ and $\sum_{i=1}^{k} c_i = e$. We have the following spectral decomposition theorem.

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

Every λ_i is called an eigenvalue of x. We denote $\lambda_{\min}(x)$ ($\lambda_{\max}(x)$) be the minimal (maximal) eigenvalue of x.

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

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

In particular, the square root: $x^{1/2} = \sum_{i=1}^{r} \sqrt{\lambda_i} c_i$ whenever $x \in K$ and undefined otherwise; the inverse $x^{-1} = \sum_{i=1}^{r} \lambda_i^{-1} c_i$ whenever $\lambda_i \neq 0$ for all $i = 1, \ldots, r$ and undefined otherwise. If x^{-1} is defined, we call x invertible.

Since "$\circ$" is bilinear for every $x \in V$, there exists a matrix $L(x)$ such that for every $y \in V$, $x \circ y = L(x)y$. In particular, $L(x)e = x$ and $L(x)x = x^2$. For each $x, y \in V$, define $Q_{x,y} = L(x)L(y) + L(y)L(x) - L(x \circ y)$ and $Q_x = 2L^2(x) - L(x^2)$. Q_x is called the quadratic representation of x. For any $x, y \in V$, x and y are said to be operator commute if $L(x)$ and $L(y)$ commute, i.e., $L(x)L(y) = L(y)L(x)$. It is well known that x and y operator commute if and only if x and y share a common Jordan frame. Let $x = \sum_{i=1}^{r} \lambda_i c_i$ be the spectral decomposition of x. Then every eigenvalue of $L(x)$ can be written as $\frac{\lambda_i + \lambda_j}{2}$ for some $i, j \leq r$. Every eigenvalue of Q_x can be written as $\lambda_i \lambda_j$ for some $i, j \leq r$.

Two elements x and s are similar, denoted as $x \sim s$, if and only if x and s share the same set of eigenvalues. We say $x \in K$ if and only if $\lambda_i \geq 0$ and $x \in \mathtt{int}K$ if and only if $\lambda_i > 0$ for all $i = 1, \ldots, r$. We also say x is positive semidefinite (positive definite) if $x \in K$ ($x \in \mathtt{int}K$, respectively). Define $tr(x) = \sum_{i=1}^{r} \lambda_i$ and $det(x) = \prod_{i=1}^{r} \lambda_i$, they are the trace and the determinant of $x \in V$. For the identity element, $tr(e) = r$. The inner product $\langle \cdot, \cdot \rangle$ is defined by $\langle x, y \rangle = tr(x \circ y)$ for any $x, y \in V$. Thus, we can define norm on V by $\|x\|_F = \sqrt{\langle x, x \rangle} = \sqrt{tr(x^2)} = \sqrt{\sum_{i=1}^{r} \lambda_i^2}$, $\forall x \in V$. Since $tr(\cdot)$ is associative, i.e., $tr(x \circ (y \circ z)) = tr((x \circ y) \circ z)$.

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

shows that $L(x)$ is a self-adjoint operator. As the definition of Q_x depends only on $L(x)$ and $L(x^2)$, both of which are self-adjoint, Q_x is also self-adjoint.

3 Algorithm

3.1 Problem background

Let V be a Euclidean Jordan algebra of dimension n with rank r, and K be its associated cone of squares. Consider the following primal and dual problem:

$$(P) \quad \min\{\langle c, x \rangle : \langle a_i, x \rangle = b_i, i = 1, \ldots, m, x \in K\}$$

and

$$(D) \quad \max \left\{ b^T y : \sum_{i=1}^{m} y_i a_i + s = c, s \in K, y \in R^m \right\}.$$

Here, $c, a_i \in V, i = 1, \ldots, m, b \in R^m$. $x \in K$ is called primal feasible if $\langle a_i, x \rangle = b_i$ for $i = 1, \ldots, m$. Similarly, $(s, y) \in K \times R^m$ is called dual feasible if $\sum_{i=1}^m y_i a_i + s = c$. Let $A^{m \times n}$ be the matrix corresponding to the linear transformation that maps x to the m-vector whose i-th component is $\langle a_i, x \rangle$. The sets of primal and dual interior feasible points are

$$\mathcal{F}^0(P) = \{x \in V : Ax = b, x \in \mathtt{int}K\}$$

and

$$\mathcal{F}^0(D) = \left\{(s, y) \in K \times R^m : A^T y + s = c, s \in \mathtt{int}K\right\},$$

respectively. Throughout this paper we assume that $\operatorname{rank}(A) = m$ and $\mathcal{F}^0(P) \times \mathcal{F}^0(D) \neq \varnothing$. To find an optimal solution of (P) and (D) is equivalent to solving the following system (Faybusovich 1997):

$$\begin{cases} Ax = b, x \in K \\ A^T y + s = c, s \in K \\ x \circ s = 0. \end{cases}$$

The third equation is usually referred to as the complementarity slackness condition. By replacing the complementarity slackness condition with $x \circ s = \mu e$, one may consider

$$(PC_\mu) \quad \begin{cases} Ax = b, x \in K \\ A^T y + s = c, s \in K \\ x \circ s = \mu e. \end{cases}$$

Interior point algorithms follow the solutions to (PC_μ) as μ goes to zero. For each $\mu > 0$, (PC_μ) has unique solution, and these solutions form a curve parameterized by μ. This curve is called the central path and most IPMs approximately follow the central path to reach the optimal set. If $\mu \to 0$, then the limit of the path exists and yields optimal solutions for (P) and (D), see Faybusovich (1997).

A natural way to define a search direction is to follow the Newton approach and to linearize the third equation in (PC_μ). This leads to the following system:

$$\begin{cases} A\Delta x = b - Ax \\ A^T \Delta y + \Delta s = c - s - A^T y \\ \Delta x \circ s + x \circ \Delta s = \sigma\mu e - x \circ s \end{cases} \tag{1}$$

where $(\Delta x, \Delta s, \Delta y) \in V \times V \times R^m$. $\sigma \in [0, 1]$ is called centering parameter and $\mu = \langle x, s \rangle / r$ is the normalized duality gap. However, due to the fact that x and s do not operator commute in general, this system does not always have a unique solution. In particular, for SDP, the direction obtained above is known as the AHO direction (Alizadeh et al. 1998), which is not necessarily unique. This difficulty can be solved by applying a scaling scheme. It goes as follows. Let $x, s, p \in \mathtt{int}K$, then $x \circ s = \alpha e$

if and only if $Q_p x \circ Q_{p^{-1}} s = \alpha e$ (Schmieta and Alizadeh 2003, Lemma 28). Now replacing the third equation in (PC_μ) by $Q_p x \circ Q_{p^{-1}} s = \mu e$ and then applying Newton method, we obtain the following linear system:

$$\begin{cases} A\Delta x = b - Ax \\ A^T \Delta y + \Delta s = c - s - A^T y \\ (Q_p \Delta x) \circ (Q_{p^{-1}} s) + (Q_p x) \circ (Q_{p^{-1}} \Delta s) = \sigma \mu e - (Q_p x) \circ (Q_{p^{-1}} s) \end{cases} \tag{2}$$

This class of directions is called Monteiro-Zhang family of search directions. In this paper, we restrict our attention to the following set of scalings named commutative class scalings:

$$\mathcal{C}(x, s) := \{p \in \mathtt{int} K : Q_p x \ \ and \ \ Q_{p^{-1}} s \ \text{ operator commute}\}.$$

Though $\mathcal{C}(x, s)$ seems to be a restrictive class, it does include some of the most interesting choice of scalings such as XS, SX and NT scalings. For $p = s^{1/2}$ we get the XS scaling, for $p = x^{-1/2}$ we get the SX scaling, and for the choice of $p = [Q_{x^{1/2}}(Q_{x^{1/2}} s)^{-1/2}]^{-1/2} = [Q_{s^{-1/2}}(Q_{s^{1/2}} x)^{1/2}]^{-1/2}$ we obtain the NT scaling. For now we always assume that $p \in \mathcal{C}(x, s)$. Let $\widetilde{A} = AQ_{p^{-1}}$, $\widetilde{c} = Q_{p^{-1}} c$, $\widetilde{x} = Q_p x$, $\widetilde{s} = Q_{p^{-1}} s$. The Newton system (2) can be equivalently written as

$$\begin{cases} \widetilde{A}\Delta\widetilde{x} = b - \widetilde{A}\widetilde{x} \\ \widetilde{A}^T \Delta y + \Delta\widetilde{s} = \widetilde{c} - \widetilde{s} - \widetilde{A}^T y \\ \Delta\widetilde{x} \circ \widetilde{s} + \widetilde{x} \circ \Delta\widetilde{s} = \sigma\mu e - \widetilde{x} \circ \widetilde{s} \end{cases} \tag{3}$$

Our algorithm will restrict the iterates to the so-called wide neighborhood:

$$\mathcal{N}_{-\infty}(\gamma) := \{(x, s, y) \in \mathtt{int} K \times \mathtt{int} K \times R^m : \lambda_{\min}(Q_{x^{1/2}} s) \geq (1 - \gamma)\mu\}$$

where $\mu = \langle x, s \rangle / r$ and $\gamma \in (0, 1)$. The neighborhood contains the central path and γ represents the size of the neighborhood as it can be shown that $\mathcal{N}_{-\infty}(0) \bigcap [\mathcal{F}^0(P) \times \mathcal{F}^0(D)]$ is exactly the central path and $\mathcal{N}_{-\infty}(1) \bigcap [\mathcal{F}^0(P) \times \mathcal{F}^0(D)] = \mathcal{F}^0(P) \times \mathcal{F}^0(D)$.

Proposition 1 *The neighborhood $\mathcal{N}_{-\infty}(\gamma)$ is scaling invariant, i.e., (x, s, y) is in the neighborhood if and only if $(\widetilde{x}, \widetilde{s}, y)$ is.*

3.2 Algorithm

In this subsection, we describe our algorithm in detail.

Step 1: Let $0 < \sigma < 1$, $\epsilon > 0$, $\gamma \in (0, 1)$. Initial point $x_0, s_0 \in \mathtt{int} K$, $y_0 \in R^m$ such that $(x_0, s_0, y_0) \in \mathcal{N}_{-\infty}(\gamma)$. Set $k = 0$, $\phi_0 = 1$.

Step 2: If $\mu_k \le \epsilon\mu_0$, then stop; Otherwise, choose $p \in \mathcal{C}(x_k, s_k)$. Solve $(\Delta\widetilde{x}_k^a, \Delta\widetilde{s}_k^a, \Delta y_k^a)$ from the following scaled Newton system at $(\widetilde{x}_k, \widetilde{s}_k, y_k) = (Q_p x_k, Q_{p^{-1}} s_k, y_k)$:

$$\begin{cases} \widetilde{A}\Delta\widetilde{x}_k^a = b - \widetilde{A}\widetilde{x}_k \\ \widetilde{A}^T\Delta y_k^a + \Delta\widetilde{s}_k^a = \widetilde{c} - \widetilde{s}_k - \widetilde{A}^T y_k \\ \Delta\widetilde{x}_k^a \circ \widetilde{s}_k + \widetilde{x}_k \circ \Delta\widetilde{s}_k^a = \sigma\mu_k e - \widetilde{x}_k \circ \widetilde{s}_k \end{cases} \tag{4}$$

where $\mu_k = \langle x_k, s_k\rangle/r$. Let $(\Delta x_k^a, \Delta s_k^a, \Delta y_k^a) = (Q_{p^{-1}}\Delta\widetilde{x}_k^a, Q_p\Delta\widetilde{s}_k^a, \Delta y_k^a)$.

Step 3: Compute a corrector step $(\Delta\widetilde{x}_k^c, \Delta\widetilde{s}_k^c, \Delta y_k^c)$:

$$\begin{cases} \widetilde{A}\Delta\widetilde{x}_k^c = 0 \\ \widetilde{A}^T\Delta y_k^c + \Delta\widetilde{s}_k^c = 0 \\ \Delta\widetilde{x}_k^c \circ \widetilde{s}_k + \widetilde{x}_k \circ \Delta\widetilde{s}_k^c = -\Delta\widetilde{x}_k^a \circ \Delta\widetilde{s}_k^a \end{cases} \tag{5}$$

Let $(\Delta x_k^c, \Delta s_k^c, \Delta y_k^c) = (Q_{p^{-1}}\Delta\widetilde{x}_k^c, Q_p\Delta\widetilde{s}_k^c, \Delta y_k^c)$.

Step 4: Set

$$(x(\alpha), s(\alpha), y(\alpha)) = (x_k, s_k, y_k) + \alpha(\Delta x_k^a, \Delta s_k^a, \Delta y_k^a) + \alpha^2(\Delta x_k^c, \Delta s_k^c, \Delta y_k^c).$$

Find $\alpha_{k1} \in (0, 1)$ such that for all $\alpha \in (0, \alpha_{k1}]$, $(x(\alpha), s(\alpha), y(\alpha)) \in \mathcal{N}_{-\infty}(\gamma)$. Compute the largest step size α_{k2} such that $\langle x(\alpha), s(\alpha)\rangle \ge (1-\alpha)\phi_k\langle x_0, s_0\rangle$.

Step 5: Compute $\alpha_k := \operatorname{argmax}\{\delta(\alpha) : \alpha \in (0, \min(\alpha_{k1}, \alpha_{k2}))\}$, where

$$\delta(\alpha) = \alpha\left[1 - \sigma - \alpha^2\frac{\langle\Delta\widetilde{x}_k^c, \Delta\widetilde{s}_k^a\rangle + \langle\Delta\widetilde{x}_k^a, \Delta\widetilde{s}_k^c\rangle}{r\mu_k}\right]. \tag{6}$$

Step 6: Set $\phi_{k+1} = (1 - \alpha_k)\phi_k$,

$$(x_{k+1}, s_{k+1}, y_{k+1}) = (x(\alpha_k), s(\alpha_k), y(\alpha_k)).$$

Set k: = k + 1, return to Step 2.

Remark 1 From the Newton systems (4) and (5), we have the following statements:

$$\begin{aligned} A\Delta x_k^a &= b - Ax_k \\ A\Delta x_k^c &= 0 \\ A^T\Delta y_k^a + \Delta s_k^a &= c - s_k - A^T y_k \\ A^T\Delta y_k^c + \Delta s_k^c &= 0 \end{aligned}$$

Using these statements we can show the relations

$$Ax_k - b = \phi_k(Ax_0 - b) \tag{7}$$

and

$$A^T y_k + s_k - c = \phi_k(A^T y_0 + s_0 - c), \tag{8}$$

and they represent the relative infeasibilities at (x_k, s_k, y_k). Since we maintain the condition $\langle x_k, s_k\rangle \geq \phi_k\langle x_0, s_0\rangle$ at every iterate, it ensures that the infeasibilities approach zero as the complementarity $\langle x_k, s_k\rangle$ approaches zero.

Remark 2 In order to obtain the polynomial convergence of our algorithm, we specify a particular initial point choice. This choice was proposed first by Zhang (1994) for LP and recently by Rangarajan (2006) for SCP. Let (u_0, r_0, v_0) be the solution to $\min\{\|u\|_F : Au = b\}$ and $\min\{\|v\|_F : A^T r + v = c\}$, and

$$x_0 = s_0 = \rho_0 e \in \mathtt{int}K \tag{9}$$

with $\rho_0 > \max(\|u_0\|_2, \|v_0\|_2)$, where $\|x\|_2 := \max_i |\lambda_i(x)|$. This implies that $x_0 - u_0 \in \mathtt{int}K$ and $s_0 - v_0 \in \mathtt{int}K$. Without loss of generality, we assume that for some constant $\Phi > 0$, $\rho_0 \geq \frac{1}{\Phi}\rho_* := \frac{1}{\Phi}\min\{\max(\|x_*\|_2, \|s_*\|_2), (x_*, s_*)\ solves\ (P)\ and\ (D)\}$.

Remark 3 Let us assume a reference point (u_0, v_0, r_0) satisfies the equality constraints such that $x_0 - u_0, s_0 - v_0 \in \mathtt{int}K$. For a given sequence of iterates $\{(x_k, s_k, y_k)\}$, we construct an auxiliary sequence $\{(u_k, v_k, r_k)\}$:

$$\begin{aligned}
u_{k+1} &= u_k + \alpha_k(\Delta x_k^a + x_k - u_k) + \alpha_k^2 \Delta x_k^c\\
v_{k+1} &= v_k + \alpha_k(\Delta s_k^a + s_k - v_k) + \alpha_k^2 \Delta s_k^c\\
r_{k+1} &= r_k + \alpha_k(\Delta y_k^a + y_k - r_k) + \alpha_k^2 \Delta y_k^c.
\end{aligned}$$

From the above definitions, we can show the following properties by induction:

$$\begin{aligned}
x_{k+1} - u_{k+1} = \phi_{k+1}(x_0 - u_0) \in \mathtt{int}K;\\
s_{k+1} - v_{k+1} = \phi_{k+1}(s_0 - v_0) \in \mathtt{int}K;\\
Au_k = b \quad \text{and} \quad A^T r_k + v_k = c;\\
A(x_k + \Delta x_k^a - u_k) = 0;\\
A^T(y_k + \Delta y_k^a - r_k) + (s_k + \Delta s_k^a - v_k) = 0.
\end{aligned}$$

4 Analysis

In this section, we analyse the complexity of our algorithm. Before we proceed to bounding the step lengths, we establish some technical lemmas. For simplicity, we will often write x, Δx^a, Δx^c for x_k, Δx_k^a, Δx_k^c etc. The indices should be clear from the context. We will use the following notations:

$$\begin{aligned}
\widetilde{x}(\alpha) &= \widetilde{x} + \alpha \Delta\widetilde{x}^a + \alpha^2 \Delta\widetilde{x}^c \\
\widetilde{s}(\alpha) &= \widetilde{s} + \alpha \Delta\widetilde{s}^a + \alpha^2 \Delta\widetilde{s}^c \\
\mu(\alpha) &= \frac{\langle x(\alpha), s(\alpha)\rangle}{r} \\
\widetilde{\mu}(\alpha) &= \frac{\langle \widetilde{x}(\alpha), \widetilde{s}(\alpha)\rangle}{r}
\end{aligned}$$

It is easy to see that $\widetilde{x}(\alpha) = Q_p x(\alpha)$, $\widetilde{s}(\alpha) = Q_{p^{-1}} s(\alpha)$ and $\widetilde{\mu}(\alpha) = \mu(\alpha)$.

Lemma 1 *Let* $(\Delta\widetilde{x}^a, \Delta\widetilde{x}^c)$ *and* $(\Delta\widetilde{s}^a, \Delta\widetilde{s}^c)$ *be generated by the algorithm, then we have*

$$\begin{aligned}
\widetilde{x}(\alpha) \circ \widetilde{s}(\alpha) &= (1-\alpha)\widetilde{x} \circ \widetilde{s} + \alpha\sigma\mu e + \alpha^3 \Delta\widetilde{x}^a \circ \Delta\widetilde{s}^c + \alpha^3 \Delta\widetilde{x}^c \circ \Delta\widetilde{s}^a + \alpha^4 \Delta\widetilde{x}^c \circ \Delta\widetilde{s}^c \\
\langle \widetilde{x}(\alpha), \widetilde{s}(\alpha)\rangle &= (1-\alpha)\langle \widetilde{x}, \widetilde{s}\rangle + \alpha\sigma r\mu + \alpha^3(\langle \Delta\widetilde{x}^c, \Delta\widetilde{s}^a\rangle + \langle \Delta\widetilde{x}^a, \Delta\widetilde{s}^c\rangle) \\
\widetilde{\mu}(\alpha) &= (1-\delta(\alpha))\mu.
\end{aligned}$$

Proof From definition of $\widetilde{x}(\alpha)$, $\widetilde{s}(\alpha)$ and the third equations of (4) and (5), we get

$$\begin{aligned}
\widetilde{x}(\alpha) \circ \widetilde{s}(\alpha) &= (\widetilde{x} + \alpha\Delta\widetilde{x}^a + \alpha^2\Delta\widetilde{x}^c) \circ (\widetilde{s} + \alpha\Delta\widetilde{s}^a + \alpha^2\Delta\widetilde{s}^c) \\
&= \widetilde{x} \circ \widetilde{s} + \alpha(\widetilde{x} \circ \Delta\widetilde{s}^a + \Delta\widetilde{x}^a \circ \widetilde{s}) + \alpha^2(\Delta\widetilde{x}^c \circ \widetilde{s} + \widetilde{x} \circ \Delta\widetilde{s}^c + \Delta\widetilde{x}^a \circ \Delta\widetilde{s}^a) \\
&\quad + \alpha^3 \Delta\widetilde{x}^c \circ \Delta\widetilde{s}^a + \alpha^3 \Delta\widetilde{x}^a \circ \Delta\widetilde{s}^c + \alpha^4 \Delta\widetilde{x}^c \circ \Delta\widetilde{s}^c \\
&= (1-\alpha)\widetilde{x} \circ \widetilde{s} + \alpha\sigma\mu e + \alpha^3 \Delta\widetilde{x}^c \circ \Delta\widetilde{s}^a + \alpha^3 \Delta\widetilde{x}^a \circ \Delta\widetilde{s}^c + \alpha^4 \Delta\widetilde{x}^c \circ \Delta\widetilde{s}^c.
\end{aligned}$$

The second statement follows straightforwardly from $\langle \Delta\widetilde{x}^c, \Delta\widetilde{s}^c\rangle = 0$ and $\langle \widetilde{x}(\alpha), \widetilde{s}(\alpha)\rangle = \langle \widetilde{x}(\alpha) \circ \widetilde{s}(\alpha), e\rangle$. The last statement is obvious. We complete the proof. □

Lemma 2 *Suppose* $(x, s) \in \mathcal{N}_{-\infty}(\gamma)$, *p is chosen as commutative class scaling corresponding to x and s. Then, there exists constants* $\beta_1 > 1$, $\beta_2 = \frac{\beta_1^2}{4(1-\gamma)}$, $\beta_3 = \sqrt{\beta_1\beta_2} = \frac{{\beta_1}^{3/2}}{2\sqrt{1-\gamma}}$ *such that:*

(1): $\|\Delta\widetilde{x}^a \circ \Delta\widetilde{s}^a\|_F \le \frac{\sqrt{\mathrm{Cond}(G)}}{2}\beta_1 r^2\mu.$

(2): $\|\Delta\widetilde{x}^c \circ \Delta\widetilde{s}^c\|_F \le \frac{\mathrm{Cond}(G)^{3/2}}{2}\beta_2 r^4\mu.$

(3): $|\frac{\langle \Delta\widetilde{x}^a, \Delta\widetilde{s}^c\rangle}{r}| \le \|\Delta\widetilde{x}^a \circ \Delta\widetilde{s}^c\|_F \le \mathrm{Cond}(G)\beta_3 r^3\mu.$

(4): $|\frac{\langle \Delta\widetilde{x}^c, \Delta\widetilde{s}^a\rangle}{r}| \le \|\Delta\widetilde{x}^c \circ \Delta\widetilde{s}^a\|_F \le \mathrm{Cond}(G)\beta_3 r^3\mu.$

Proof We introduce $G = L(\widetilde{s})^{-1}L(\widetilde{x})$ and $t^2 := \|G^{-1/2}\Delta\widetilde{x}^a\|_F^2 + \|G^{1/2}\Delta\widetilde{s}^a\|_F^2$. Pay attention to Remarks 2 and 3, by following a similar proof of Proposition 3.7 and (3.24) (Rangarajan 2006, pp. 1124–1226), we can show

$$t^2 \leq \beta_1 r^2 \mu \tag{10}$$

holds for some constant $\beta_1 > 1$. Statement (1) follows from Lemma 33 (Schmieta and Alizadeh 2003), (10) and $\|x \circ y\|_F \leq \|x\|_F \|y\|_F$.

Since $\widetilde{x}$ and $\widetilde{s}$ operator commute, they share a common Jordan frame, see $\{c_1, \ldots, c_r\}$. Suppose the spectral decomposition of $\widetilde{x}$ and $\widetilde{s}$ are $\widetilde{x} = \sum_{i=1}^r \lambda_i c_i$ and $\widetilde{s} = \sum_{i=1}^r \mu_i c_i$. Since $\widetilde{x} = Q_p x \in \mathtt{int} K$, $\widetilde{s} = Q_{p^{-1}} s \in \mathtt{int} K$, we have $\lambda_i, \mu_i > 0, i = 1, 2, \ldots, r$. Following a similar argument of Lemma 4.1 (Rangarajan 2006), the operator norm

$$\begin{aligned} \|L(\widetilde{x})^{-1} L(\widetilde{s})^{-1}\| &= \lambda_{\max}(L(\widetilde{x})^{-1} L(\widetilde{s})^{-1}) \\ &= \max_{i,j} \left[\frac{2}{\lambda_i + \lambda_j} \frac{2}{\mu_i + \mu_j} \right] \\ &\leq \max_{i,j} \left[\left(\frac{1}{\lambda_i} + \frac{1}{\lambda_j} \right) \frac{1}{\mu_i + \mu_j} \right] \\ &= \lambda_{\max}(Q_{\widetilde{x}}^{-1} G) \end{aligned} \tag{11}$$

where the last equality comes from the last three lines in Rangarajan (2006, p. 1226). It follows from (11), Lemma 3.5 and Lemma 4.1 (Rangarajan 2006) that

$$\|L(\widetilde{x})^{-1} L(\widetilde{s})^{-1}\| \leq \lambda_{\max}(Q_{\widetilde{x}}^{-1} G) = \frac{1}{\lambda_{\min}(Q_{\widetilde{x}^{1/2}} \widetilde{s})} = \frac{1}{\lambda_{\min}(\widetilde{x} \circ \widetilde{s})} \tag{12}$$

Now we turn to the third equation of (5). It can be equivalently written as:

$$L(\widetilde{s}) \Delta\widetilde{x}^c + L(\widetilde{x}) \Delta\widetilde{s}^c = -\Delta\widetilde{x}^a \circ \Delta\widetilde{s}^a.$$

i.e.,

$$G^{-1/2} \Delta\widetilde{x}^c + G^{1/2} \Delta\widetilde{s}^c = -L(\widetilde{x})^{-1/2} L(\widetilde{s})^{-1/2} \Delta\widetilde{x}^a \circ \Delta\widetilde{s}^a.$$

Again, taking norm-squared on both sides of this equation and using the fact $\langle \Delta\widetilde{x}^c, \Delta\widetilde{s}^c \rangle = 0$, it follows that

$$\begin{aligned} T^2 &:= \|G^{-1/2} \Delta\widetilde{x}^c\|_F^2 + \|G^{1/2} \Delta\widetilde{s}^c\|_F^2 \\ &\leq \|L(\widetilde{x})^{-1} L(\widetilde{s})^{-1}\| \cdot \|\Delta\widetilde{x}^a \circ \Delta\widetilde{s}^a\|_F^2 \\ &\leq \frac{1}{\lambda_{\min}(\widetilde{x} \circ \widetilde{s})} \left(\frac{\sqrt{\mathrm{Cond}(G)}}{2} \beta_1 r^2 \mu \right)^2 \\ &\leq \frac{\mathrm{Cond}(G)}{4(1-\gamma)} \beta_1^2 r^4 \mu \end{aligned}$$

where the second inequality comes from (12) and statement (1), the last inequality follows from the fact that $\lambda_{\min}(\widetilde{x} \circ \widetilde{s}) = \lambda_{\min}(Q_{\widetilde{x}^{1/2}} \widetilde{s}) = \lambda_{\min}(Q_{x^{1/2}} s) \geq (1-\gamma)\mu$.

We obtain

$$T^2 \leq \text{Cond}(G)\beta_2 r^4 \mu. \tag{13}$$

Again, from Lemma 33 (Schmieta and Alizadeh 2003), the second statement holds.

By replacing $u = \Delta\widetilde{x}^a$ and $v = \Delta\widetilde{s}^c$ in the proof of Lemma 33 (Schmieta and Alizadeh 2003), we can get

$$\begin{aligned}\|\Delta\widetilde{x}^a \circ \Delta\widetilde{s}^c\|_F &\leq \|\Delta\widetilde{x}^a\|_F \|\Delta\widetilde{s}^c\|_F \\ &\leq \sqrt{\text{Cond}(G)}\|G^{-1/2}\Delta\widetilde{x}^a\|_F \|G^{1/2}\Delta\widetilde{s}^c\|_F \\ &\leq \sqrt{\text{Cond}(G)}\sqrt{t^2}\sqrt{T^2} \\ &\leq \text{Cond}(G)\beta_3 r^3 \mu.\end{aligned}$$

Similarly, we can get

$$\|\Delta\widetilde{x}^c \circ \Delta\widetilde{s}^a\|_F \leq \text{Cond}(G)\beta_3 r^3 \mu.$$

The first inequality of statements (3) and (4) follow directly from the fact that $|\frac{\langle x,s\rangle}{r}| = |\frac{\langle x \circ s, e\rangle}{r}| \leq \|x \circ s\|_F \|e\|_F / r \leq \|x \circ s\|_F$ for any $x, s \in V$. The proof is completed. $\square$

The following three lemmas provide lower bounds of α_{k1}, α_{k2} and α_k.

Lemma 3 $\alpha_{k1} \geq \frac{\theta_1}{r^{3/2}\sqrt{\text{Cond}(G)}}$, *where* $\theta_1 = \min\left\{\sqrt{\frac{\gamma\sigma}{4\beta_3(2-\gamma)}}, \sqrt[3]{\frac{\gamma\sigma}{\beta_2}}\right\}$.

Proof According to Proposition 1, we need only to find the largest step length satisfying $\lambda_{\min}(Q_{\widetilde{s}(\alpha)^{1/2}}\widetilde{s}(\alpha)) \geq (1-\gamma)\widetilde{\mu}(\alpha)$. We first bound the left and right hand side of the inequality defining the neighborhood $\mathcal{N}_{-\infty}(\gamma)$. To begin with a bound on the left side, we have

$$\begin{aligned}&\lambda_{\min}(Q_{\widetilde{x}(\alpha)^{1/2}}\widetilde{s}(\alpha)) \\ &\quad \geq \lambda_{\min}(\widetilde{x}(\alpha) \circ \widetilde{s}(\alpha)) \\ &\quad \geq (1-\alpha)\lambda_{\min}(\widetilde{x} \circ \widetilde{s}) + \alpha\sigma\mu + \alpha^3\lambda_{\min}(\Delta\widetilde{x}^a \circ \Delta\widetilde{s}^c) \\ &\qquad + \alpha^3\lambda_{\min}(\Delta\widetilde{x}^c \circ \Delta\widetilde{s}^a) + \alpha^4\lambda_{\min}(\Delta\widetilde{x}^c \circ \Delta\widetilde{s}^c) \\ &\quad \geq (1-\alpha)\lambda_{\min}(\widetilde{x} \circ \widetilde{s}) + \alpha\sigma\mu - \alpha^3\|\Delta\widetilde{x}^a \circ \Delta\widetilde{s}^c\|_F \\ &\qquad - \alpha^3\|\Delta\widetilde{x}^c \circ \Delta\widetilde{s}^a\|_F - \alpha^4\|\Delta\widetilde{x}^c \circ \Delta\widetilde{s}^c\|_F \\ &\quad \geq (1-\alpha)(1-\gamma)\mu + \alpha\sigma\mu - 2\alpha^3\text{Cond}(G)\beta_3 r^3\mu - \alpha^4/2(\text{Cond}(G))^{3/2}\beta_2 r^4\mu \\ &\quad =: g(\alpha).\end{aligned}$$

The first inequality follows from Lemma 3.5 (Rangarajan 2006), the second inequality comes from Lemma 14 (Schmieta and Alizadeh 2003). The third equality follows from the fact $\lambda_{\min}(x) \geq -\|x\|_F$, and the last inequality holds because of Lemma 2.

From Lemma 1 and Lemma 2, we get

$$\begin{aligned}\widetilde{\mu}(\alpha) &= \mu\left[1-\alpha(1-\sigma)+\alpha^3\frac{\langle\Delta\widetilde{x}^c,\Delta\widetilde{s}^a\rangle+\langle\Delta\widetilde{x}^a,\Delta\widetilde{s}^c\rangle}{r\mu}\right]\\ &\le \mu[1-\alpha(1-\sigma)+2\alpha^3\mathrm{Cond}(G)\beta_3 r^3]:=f(\alpha).\end{aligned}$$

In order to the iterate stay in the neighborhood, it is sufficient to have

$$\lambda_{\min}(Q_{\widetilde{x}(\alpha)^{1/2}}\widetilde{s}(\alpha))\ge g(\alpha)\ge(1-\gamma)f(\alpha)\ge(1-\gamma)\widetilde{\mu}(\alpha). \tag{14}$$

We want to find a lower bound for α satisfying $g(\alpha)\ge(1-\gamma)f(\alpha)$, that is

$$2\alpha^2\mathrm{Cond}(G)\beta_3 r^3(2-\gamma)+\alpha^3/2(\mathrm{Cond}(G))^{3/2}\beta_2 r^4\le\gamma\sigma. \tag{15}$$

The first item of Eq. 15

$$2\alpha^2\mathrm{Cond}(G)\beta_3 r^3(2-\gamma)\le\gamma\sigma/2$$

for all $\alpha\in\left(0,\sqrt{\frac{\gamma\sigma}{4\beta_3(2-\gamma)}}\frac{1}{r^{3/2}\sqrt{\mathrm{Cond}(G)}}\right]$. The second item of Eq. 15

$$\alpha^3/2(\mathrm{Cond}(G))^{3/2}\beta_2 r^4\le\alpha^3/2(\mathrm{Cond}(G))^{3/2}\beta_2 r^{9/2}\le\gamma\sigma/2$$

for all $\alpha\in\left(0,\sqrt[3]{\frac{\gamma\sigma}{2\beta_2}}\frac{1}{r^{3/2}\sqrt{\mathrm{Cond}(G)}}\right]$. From the definition of θ_1, (15) holds for all $\alpha\in\left(0,\frac{\theta_1}{r^{3/2}\sqrt{\mathrm{Cond}(G)}}\right]$. So in this interval, $\lambda_{\min}(Q_{\widetilde{x}(\alpha)^{1/2}}\widetilde{s}(\alpha))>0$, which implies $det(Q_{\widetilde{x}(\alpha)^{1/2}}\widetilde{s}(\alpha))>0$. We now need to prove that in this interval both $\widetilde{x}(\alpha)$ and $\widetilde{s}(\alpha)$ are positive definite. By continuity of $\lambda_{\min}$, we can without loss of generality assume that there exists a value α_0 in the interval such that $\widetilde{x}(\alpha_0)=0$ (or $\widetilde{s}(\alpha_0)=0$). From Proposition 5 (Schmieta and Alizadeh 2003), $det(Q_{\widetilde{x}(\alpha_0)^{1/2}}\widetilde{s}(\alpha_0))=det(\widetilde{x}(\alpha_0))det(\widetilde{s}(\alpha_0))=0$. This is a contradiction. Thus, $\widetilde{x}(\alpha)$ and $\widetilde{s}(\alpha)$ are positive definite. Together with (14) we conclude $(\widetilde{x}(\alpha),\widetilde{s}(\alpha))\in\mathcal{N}_{-\infty}(\gamma)$ for all $\alpha\in\left(0,\frac{\theta_1}{r^{3/2}\sqrt{\mathrm{Cond}(G)}}\right]$. We finish the proof. □

Lemma 4 $\alpha_{k2}\ge\frac{\theta_2}{r^{3/2}\sqrt{\mathrm{Cond}(G)}}$, *where* $\theta_2=\sqrt{\frac{\sigma}{2\beta_3}}$.

Proof Using Lemma 1 and Lemma 2 we have

$$\begin{aligned}&\langle x(\alpha),s(\alpha)\rangle-(1-\alpha)\phi\langle x_0,s_0\rangle\\ &=(1-\alpha)(\langle x,s\rangle-\phi\langle x_0,s_0\rangle)+\alpha\sigma\mu r+\alpha^3(\langle\Delta\widetilde{x}^c,\Delta\widetilde{s}^a\rangle+\langle\Delta\widetilde{x}^a,\Delta\widetilde{s}^c\rangle)\\ &\ge\alpha\sigma\mu r-2\mathrm{Cond}(G)\alpha^3\beta_3 r^4\mu>0\end{aligned}$$

for all $\alpha\in\left(0,\frac{\theta_2}{r^{3/2}\sqrt{\mathrm{Cond}(G)}}\right]$. The inequality follows because $\langle x,s\rangle\ge\phi\langle x_0,s_0\rangle$. The proof is completed. □

Lemma 5 $\alpha_k \geq \frac{\theta}{r^{3/2}\sqrt{\text{Cond}(G)}}$, *where* $\theta = \min\left\{\theta_1, \theta_2, \sqrt{\frac{1-\sigma}{6\beta_3}}\right\} < 1$.

Proof From the definition of $\delta(\alpha)$ and Lemma 2, we can show that

$$\begin{aligned}\delta'(\alpha) &= 1-\sigma-3\alpha^2\frac{\langle\Delta\widetilde{x}^c, \Delta\widetilde{s}^a\rangle+\langle\Delta\widetilde{x}^a, \Delta\widetilde{s}^c\rangle}{r\mu}\\ &\geq 1-\sigma-6\alpha^2\text{Cond}(G)\beta_3 r^3 \geq 0\end{aligned}$$

for all $\alpha \in (0, \frac{\theta_3}{r^{3/2}\sqrt{\text{Cond}(G)}}]$ where $\theta_3 = \sqrt{\frac{1-\sigma}{6\beta_3}}$. Let $\bar{\alpha} = \frac{\theta}{r^{3/2}\sqrt{\text{Cond}(G)}}$. $\delta(\alpha)$ is monotonically increasing in the interval $(0, \bar{\alpha}] \subseteq (0, \min(\alpha_{k1}, \alpha_{k2})]$. From the definition of α_k, we conclude $\alpha_k \geq \bar{\alpha}$. The proof is completed. □

Now we are in a position to state our main result.

Theorem 2 *Assume that* Cond(G) *can be bounded from above by* $\kappa < \infty$ *for all iterations of the algorithm. Then the algorithm will terminate in*

$$O(r^{3/2}\sqrt{\kappa}\log\epsilon^{-1})$$

iterations such that $\|Ax_k-b\| \leq \epsilon\|Ax_0-b\|$, $\|A^Ty_k+s_k-c\| \leq \epsilon\|A^Ty_0+s_0-c\|$ *and* $\langle x_k, s_k\rangle \leq \epsilon\langle x_0, s_0\rangle$.

Proof From the definition of $\delta(\alpha)$ and Lemma 1, we have

$$\begin{aligned}\delta(\alpha) &= \alpha\left[1-\sigma-\alpha^2\frac{\langle\Delta\widetilde{x}^c, \Delta\widetilde{s}^a\rangle+\langle\Delta\widetilde{x}^a, \Delta\widetilde{s}^c\rangle}{r\mu}\right]\\ &\geq \alpha[1-\sigma-2\alpha^2\text{Cond}(G)\beta_3 r^3].\end{aligned}$$

From Lemma 5 we can see $\delta(\alpha_k) \geq \delta(\alpha)$ for all $\alpha \in \left(0, \frac{\theta}{r^{3/2}\sqrt{\text{Cond}(G)}}\right]$. Specifically, $\delta(\alpha_k) \geq \delta(\bar{\alpha})$ holds for $\bar{\alpha} = \frac{\theta}{r^{3/2}\sqrt{\text{Cond}(G)}}$. Thus,

$$\begin{aligned}\delta(\alpha_k) &\geq \left[1-\sigma-2\beta_3\theta^2\right]\frac{\theta}{r^{3/2}\sqrt{\text{Cond}(G)}}\\ &\geq \left[1-\sigma-\frac{1-\sigma}{3}\right]\frac{\theta}{r^{3/2}\sqrt{\text{Cond}(G)}}\\ &\geq \frac{c_0}{r^{3/2}\sqrt{\kappa}}\end{aligned}$$

where $c_0 = 2(1-\sigma)\theta/3 < 1$. Using Lemma 1, we conclude

$$\mu_{k+1} = \mu(\alpha_k) = (1-\delta(\alpha_k))\mu_k \leq \left(1-\frac{c_0}{r^{3/2}\sqrt{\kappa}}\right)\mu_k.$$

Then we have

$$\mu_k \leq \left(1 - \frac{c_0}{r^{3/2}\sqrt{\kappa}}\right)^k \mu_0.$$

The relation $\mu_k \leq \epsilon\mu_0$ holds if $\left(1 - \frac{c_0}{r^{3/2}\sqrt{\kappa}}\right)^k \leq \epsilon$. Taking logarithm, we obtain:

$$k \log\left(1 - \frac{c_0}{r^{3/2}\sqrt{\kappa}}\right) \leq \log \epsilon.$$

Using the fact that $-\log(1-t) \geq t$ for $0 < t < 1$, we observe that the above inequality certainly holds if $k\frac{c_0}{r^{3/2}\sqrt{\kappa}} \geq \log \epsilon^{-1}$, i.e.,

$$k \geq \frac{r^{3/2}\sqrt{\kappa}}{c_0} \log \epsilon^{-1}.$$

This implies after at most $O(r^{3/2}\sqrt{\kappa} \log \epsilon^{-1})$ iterations, $\mu_k \leq \epsilon\mu_0$. i.e.,

$$\langle x_k, s_k\rangle \leq \epsilon\langle x_0, s_0\rangle.$$

As mentioned in Remark 1, the condition $\langle x_k, s_k\rangle \geq \phi_k\langle x_0, s_0\rangle$ is maintained at every iterate, so we have $\phi_k \leq \epsilon$. Together with (7) and (8), we get

$$\|Ax_k - b\| \leq \epsilon\|Ax_0 - b\|, \|A^T y_k + s_k - c\| \leq \epsilon\|A^T y_0 + s_0 - c\|.$$

The proof is completed. □

We can specialize the algorithm further by prescribing the scaling element p.

Corollary 1 *If in the algorithm p is chosen as NT scaling, then the algorithm stops in $O(r^{3/2} \log \epsilon^{-1})$ iterations. If p is chosen either XS or SX scaling, the algorithm takes $O(r^2 \log \epsilon^{-1})$ iterations.*

Proof According to Lemma 36 in Schmieta and Alizadeh (2003), for the NT method, $\kappa = 1$ and for the XS and SX methods, $\kappa = r/(1-\gamma)$. The conclusion follows immediately from Theorem 2. □

Remark 4 Compared with the results in Rangarajan (2006), the iteration bound is reduced by a factor of $\sqrt{r}$. For the NT scaling, the complexity coincides with the result for LP by Zhang and Zhang (1995).

Now we consider a feasible case of our algorithm. Since the proof techniques are exactly the same as those used in Sect. 4, we will only give a brief outline of the proof, omitting the details. We observe that for feasible interior point methods, in place of (10), we have $t^2 \leq \beta_1 r\mu$. As a result, corresponding to Lemma 2, we have (1): $\|\Delta\widetilde{x}^a \circ \Delta\widetilde{s}^a\|_F \leq \frac{\sqrt{\mathrm{Cond}(G)}}{2}\beta_1 r\mu$. (2): $\|\Delta\widetilde{x}^c \circ$

$\Delta\widetilde{s}^c\|_F \le \frac{\mathrm{Cond}(G)^{3/2}}{2}\beta_2 r^2\mu$. (3): $|\frac{\langle\Delta\widetilde{x}^a,\Delta\widetilde{s}^c\rangle}{r}| \le \|\Delta\widetilde{x}^a \circ \Delta\widetilde{s}^c\|_F \le \mathrm{Cond}(G)\beta_3 r^{3/2}\mu$. (4): $|\frac{\langle\Delta\widetilde{x}^c,\Delta\widetilde{s}^a\rangle}{r}| \le \|\Delta\widetilde{x}^c \circ \Delta\widetilde{s}^a\|_F \le \mathrm{Cond}(G)\beta_3 r^{3/2}\mu$. Moreover, Lemmas 3–5 become, with possibly different values for the constants involved,

$$\alpha_{k1} \ge \frac{\theta_1}{r^{3/4}\sqrt{\mathrm{Cond}(G)}}, \quad \alpha_{k2} \ge \frac{\theta_2}{r^{3/4}\sqrt{\mathrm{Cond}(G)}}, \quad \alpha_k \ge \frac{\theta}{r^{3/4}\sqrt{\mathrm{Cond}(G)}}.$$

Therefore, we have the following complexity bound.

Theorem 3 *If the algorithm starts with a strictly feasible point and* Cond(G) *can be bounded from above by* $\kappa < \infty$ *for all iterations. Then the algorithm will terminate in*

$$O(r^{3/4}\sqrt{\kappa}\log\epsilon^{-1})$$

iterations. Furthermore, for NT method, the algorithm takes $O(r^{3/4}\log\epsilon^{-1})$ *iterations; for XS and SX method, the algorithm takes* $O(r^{5/4}\log\epsilon^{-1})$ *iterations.*

Remark 5 For long step methods, the complexity bound obtained by Schmieta and Alizadeh (2003) is $O(r\sqrt{\kappa}\log\epsilon^{-1})$, which is worse than $O(r^{3/4}\sqrt{\kappa}\log\epsilon^{-1})$. We also observe that the bound obtained for NT method coincides the corresponding result in Zhang and Zhang (1995).

5 Conclusions

We have established polynomial convergence of infeasible and feasible IPMs for three important methods: the XS, SX and the NT method. The complexity obtained for the NT method coincides with the bound obtained for LP by Zhang and Zhang (1995).

Acknowledgments We are grateful to Editor Sven Leyffer and two anonymous referees for their insightful comments and constructive suggestions.

References

Alizadeh F, Haeberly JP, Overton ML (1998) Primal-dual interior point methods for semidefinite programming: convergence rates, stability and numerical results. SIAM J Optim 8:746–768

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

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

Mehrotra S (1992) On the implementation of a primal dual interior point method. SIAM J Optim 2:575–601

Monteiro RDC (1998) Polynomial convergence of primal-dual algorithms for semidefinite programming based on Monteiro and Zhang family of directions. SIAM J Optim 8:797–812

Monteiro RDC, Tsuchiya T (2000) Polynomial convergence of primal-dual algorithms for the second-order cone program based on the MZ-family of directions. Math Program 88:61–83

Monteiro RDC, Zhang Y (1998) A unified analysis for a class of path-following primal-dual interior-point algorithms for semidefinite programming. Math Program 81:281–299

Muramatsu M (2002) On a commutative class of search directions for linear programming over symmetric cones. J Optim Theory Appl 112:595–625

Nesterov YE, Nemirovskii AS (1994) Interior-point polynomial algorithms in convex programming. SIAM, Philadelphia

Nesterov YE, Todd MJ (1997) Self-scaled barriers and interior point methods for convex programming. Math Oper Res 22:1–42

Rangarajan BK (2006) Polynomial convergence of infeasible-interior-point methods over symmetric cones. SIAM J Optim 16:1211–1229

Salahi M, Mahdavi-Amiri N (2006) Polynomial time second order Mehrotra-type predictor-corrector algorithms. Appl Math Comput 183:646–658

Schmieta SH, Alizadeh F (2001) Associative and Jordan algebras, and polynomial time interior-point algorithms for symmetric cones. Math Oper Res 26:543–564

Schmieta SH, Alizadeh F (2003) Extension of primal-dual interior-point algorithms to symmetric cones. Math Program 96:409–438

Tsuchiya T (1999) A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming. Optim Methods Softw 11:141–182

Zhang Y (1994) On the convergence of a class of infeasible-interior-point methods for the horizontal linear complementarity problem. SIAM J Optim 4:208–227

Zhang Y (1998) On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming. SIAM J Optim 8:365–386

Zhang Y, Zhang D (1995) On polynomial of the Mehrotra-type predictor-corrector interior-point algorithms. Math Program 68:303–318