Acta Mathematicae Applicatae Sinica, English Series
Vol. 33, No. 3 (2017) 771–788
DOI: 10.1007/s10255-017-0697-7
http://www.ApplMath.com.cn & www.SpringerLink.com

Acta Mathematicae Applicatae Sinica, English Series

A New Infeasible-Interior-Point Algorithm for Linear Programming over Symmetric Cones

Chang-he LIU[1,2], You-lin SHANG[1,†], Ping HAN[1]

[1]School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang 471023, China
[2]College of Mathematics and System Science, Shandong University of Science and Technology, Qingdao 266590, China ([†]E-mail: changheliu@163.com, mathshang@sina.com)

Abstract In this paper we present an infeasible-interior-point algorithm, based on a new wide neighbourhood $\mathcal{N}(\tau_1, \tau_2, \eta)$, for linear programming over symmetric cones. We treat the classical Newton direction as the sum of two other directions. We prove that if these two directions are equipped with different and appropriate step sizes, then the new algorithm has a polynomial convergence for the commutative class of search directions. In particular, the complexity bound is $\mathcal{O}(r^{1.5} \log \varepsilon^{-1})$ for the Nesterov-Todd (NT) direction, and $\mathcal{O}(r^2 \log \varepsilon^{-1})$ for the xs and sx directions, where r is the rank of the associated Euclidean Jordan algebra and $\varepsilon > 0$ is the required precision. If starting with a feasible point (x^0, y^0, s^0) in $\mathcal{N}(\tau_1, \tau_2, \eta)$, the complexity bound is $\mathcal{O}(\sqrt{r} \log \varepsilon^{-1})$ for the NT direction, and $\mathcal{O}(r \log \varepsilon^{-1})$ for the xs and sx directions. When the NT search direction is used, we get the best complexity bound of wide neighborhood interior-point algorithm for linear programming over symmetric cones.

Keywords symmetric cone; Euclidean Jordan algebra; interior-point methods; linear programming; polynomial complexity

2000 MR Subject Classification 90C51; 90C05; 90C22; 90C25

Introduction

During the last two decades, major developments in convex programming were focusing on conic programming, which optimizes a linear objective function subject to linear constraints and over a pointed, closed, convex cone. [13] first led to a general theory of interior-point-methods (IPMs) in convex programming. Their methods were primarily either primal or dual based. Later, [14,15] provided a theoretical foundation of efficient primal-dual IPMs on a special class of conic programming, where the associated cone is so-called self-scaled cone. [8] observed that the self-scaled cones are precisely symmetric cones, which have been much studied in other areas of mathematical sciences (see, for example, [4]). This special subclass of conic programming includes linear programming (LP), semidefinite programming (SDP) and second order cone programming (SOCP) as special cases. We denote this special subclass of conic programming as $\mathcal{K}$-LP, where $\mathcal{K}$ is the associated symmetric cone. In a more general context, [5,6] analyzed the IPMs over symmetric cones characterized by Jordan algebras, where Jordan algebras played a crucial role. [18] extended the analysis by [12] for SDP to $\mathcal{K}$-LP problem. They proved polynomial iteration complexities for variants of the short, semi-long, and long step path following algorithms based on commutative class of search directions. For short

Manuscript received December 2, 2015. Revised December 13, 2016.
Supported by the National Natural Science Foundation of China (No. 11471102) and the Key Basic Research Foundation of the Higher Education Institutions of Henan Province (No. 16A110012).

step method, the iteration bound is $\mathcal{O}(\kappa\sqrt{r}\log\varepsilon^{-1})$, and for semi-long and long step methods, the iteration bound is $\mathcal{O}(\kappa r\log\varepsilon^{-1})$, where r is the rank of the associated Euclidean Jordan algebra, $\varepsilon > 0$ is the required precision, and $0 < \kappa < \infty$ will be defined in Section 6. Later, based on the commutative class of search directions, [17] analyzed an infeasible-IPM(IIPM) over symmetric cones using the wide neighbourhood N_∞^-. This algorithm does not require the iterates be feasible to the relevant linear systems, but only be in the interior of the cone $\mathcal{K}$. The complexity bound obtained there is $\mathcal{O}(\kappa r^2\log\varepsilon^{-1})$. Other IPMs over symmetric cones can be found in [3,7,10,11,16,19,20], for example.

Based on Ai's original idea[1], an important result was given by [2] for linear complementarity problem (LCP). Their algorithm decomposes the classical Newton direction into two orthogonal ones and proceeds in a new wide neighbourhood $\mathcal{N}(\tau_1,\tau_2,\eta)$. It is proved that their algorithm stops after at most $\mathcal{O}(\sqrt{n}L)$ iterations, where n is the number of variables and L is the input data length. This result yields the first wide neighbourhood path-following algorithm having the same theoretical complexity as a small neighbourhood algorithm for monotone LCPs. Later, [9] generalized the Ai-Zhang's idea to SDP and showed $\mathcal{O}(\sqrt{n}L)$ iteration complexity of their algorithm when using the NT direction. In addition, they proposed a question: whether Ai-Zhang's scheme can be applied to SOCP problems and further to general conic programming. We answer this question in the affirmative. Recently, Liu et. al.[10] extend the neighborhood-following algorithm of LP in [1] to symmetric cones. In the paper, the authors proved a key property about the new wide neighbourhood, which plays a crucial role in the complexity analysis. In this paper, we propose a new IIPM for LP over symmetric cones. This unifies the analysis for linear, second-order cone and semidefinite programming. It is proved that the new algorithm stops after at most $\mathcal{O}(\kappa r^{1.5}\log\varepsilon^{-1})$ iteration. The complexity bound obtained here is better than that obtained by [17]. For the feasible case, the iteration complexity of the algorithm was reduced to $\mathcal{O}(\kappa\sqrt{r}\log\varepsilon^{-1})$, which is the same complexity as small neighbourhood (short-step) IPMs over symmetric cones analyzed by [18].

In Section 2, we review the theory of Jordan algebras and symmetric cones. In Section 3, we introduce the $\mathcal{K}$-LP problems and a new wide neighbourhood. In Section 4, we explain the way to decompose the Newton direction and state the generic framework of our algorithm. In Section 5, we first demonstrate several technical lemmas, and then establish the iteration complexity of the proposed algorithm based on the commutative class of directions. We also give the complexity bound for the case of feasible starting point. Finally, some conclusions are given in Section 6.

2 Euclidean Jordan Algebras and Symmetric Cones

In this section, we introduce Jordan algebras and symmetric cones as well as some of their basic properties. This theory serves as our basic toolbox for the analysis of IPMs. Our presentation mostly follows[4,18].

Let $\mathcal{J}$ be an n-dimensional vector space over $\mathcal{R}$, along with a bilinear map $\circ : \mathcal{J}\times\mathcal{J} \longmapsto \mathcal{J}$. Then $(\mathcal{J},\circ)$ is a *Jordan algebra* if for all $x, y \in \mathcal{J}$, $x\circ y = y\circ x$ and $x\circ(x^2\circ y) = x^2\circ(x\circ y)$ where $x^2 = x\circ x$. A Jordan algebra $\mathcal{J}$ is called Euclidean if there exists a symmetric positive definite quadratic form $\mathcal{Q}$ on $\mathcal{J}$ such that $\mathcal{Q}(x\circ y, z) = \mathcal{Q}(x, y\circ z)$. An element $e \in \mathcal{J}$ is an identity element if $x\circ e = e\circ x = x$ for all $x \in \mathcal{J}$. The cone of squares of a Euclidean Jordan algebra $\mathcal{J}$ is the set $\mathcal{K} := \{x^2 : x \in \mathcal{J}\}$.

Let $G(\mathcal{K})$ denote the group of automorphisms of a cone $\mathcal{K}$. $\mathcal{K}$ is a homogeneous cone if $G(\mathcal{K})$

acts on it transitively. That is, if $x, y \in \operatorname{int} \mathcal{K}$, then there exists $g \in G(\mathcal{K})$ such that $g(x) = y$. Symmetric cones are also precisely the class of self-scaled cones introduced by [14] (see also [6] and [8]). The relevance of the theory of Euclidean Jordan algebras for $\mathcal{K}$-LP problem stems from the following theorem, which can be found in [4, Theorems III.2.1 and III.3.1].

Theorem 2.1. *A cone is symmetric if and only if it is the cone of squares of some Euclidean Jordan algebra.*

Since "$\circ$" is bilinear for every $x \in \mathcal{J}$, there exists a linear operator L_x such that $x \circ y = L_x y$ for all $y \in \mathcal{J}$. For each $x \in \mathcal{J}$ define $Q_x := 2L_x^2 - L_{x^2}$, which is called the *quadratic representation* of x, and it plays an important role in our subsequent analysis.

For $x \in \mathcal{J}$, let r be the smallest integer such that the set $\{e, x, x^2, \cdots, x^r\}$ is linearly dependent. Then r is called the *degree* of x and denoted by $\deg(x)$. The *rank* of $\mathcal{J}$, denoted by $\operatorname{rank}(\mathcal{J})$, is the maximum of $\deg(x)$ over all members $x \in \mathcal{J}$. An idempotent c is a nonzero element of $\mathcal{J}$ such that $c^2 = c$. A complete system of orthogonal idempotents is a set $\{c_1, \cdots, c_k\}$ of idempotents, where $c_i \circ c_j = 0$ for all $i \neq j$, and $c_1 + \ldots + c_k = e$. An idempotent is primitive if it is not the sum of two other idempotents. A complete system of orthogonal primitive idempotents is called a *Jordan frame*. We have the following *spectral decomposition theorem*.

Theorem 2.2. ([4, Theorem III.1.2]). *Let $\mathcal{J}$ be a Euclidean Jordan algebra with rank r. Then for every $x \in \mathcal{J}$, there exist a Jordan frame $\{c_1, \cdots, c_r\}$ and real numbers $\lambda_1, \cdots, \lambda_r$ such that $x = \lambda_1 c_1 + \cdots + \lambda_r c_r$. The numbers λ_i are called the eigenvalues of x.*

We define the following. The inverse $x^{-1} := \lambda_1^{-1} c_1 + \ldots + \lambda_r^{-1} c_r$, whenever all $\lambda_i \neq 0$; The square root $x^{1/2} := \lambda_1^{1/2} c_1 + \ldots + \lambda_r^{1/2} c_r$, whenever all $\lambda_i \geq 0$; The trace $\operatorname{tr}(x) := \lambda_1 + \ldots + \lambda_r$; The determinant $\det(x) := \lambda_1 \cdots \lambda_r$. Denote the minimum (maximum) eigenvalues of $x \in \mathcal{J}$ by $\lambda_{\min}(x)(\lambda_{\max}(x))$. If x^{-1} is well defined, we call x invertible. We call $x \in \mathcal{J}$ positive semidefinite (positive definite), denoted by $x \succeq 0$ $(x \succ 0)$, if all its eigenvalues are nonnegative (positive). It is clear that an element is positive semidefinite (positive definite) if and only if it belongs to (the interior of) the cone of squares.

Lemma 2.1. ([7, Lemma 2.15]). *If $x \circ y \in \operatorname{int} \mathcal{K}$, then $\det(x) \neq 0$.*

Since $\operatorname{tr}(x \circ y)$ is a bilinear function, the inner product can be defined as $\langle x, y \rangle := \operatorname{tr}(x \circ y)$. Since $\operatorname{tr}(\cdot)$ is associative, it follows that the inner product is associative, that is $\langle x \circ y, z \rangle = \langle x, y \circ z \rangle$. The spectrum of $x \in \mathcal{J}$ is the multiset of its eigenvalues. For $x \in \mathcal{J}$ with spectrum $\lambda_1, \lambda_2, \cdots, \lambda_r$, the Frobenius norm and the spectral norm can be defined as $\|x\|_F := \sqrt{\langle x, x \rangle} = \sqrt{\left(\sum \lambda_i^2\right)}$, and $\|x\|_2 := \max_i |\lambda_i|$. Observe that $\|e\|_F = \sqrt{r}$ and $\|e\|_2 = 1$, since identity element e has eigenvalue 1, with multiplicity r. Since the inner product is associative, it follows that L_x and L_x^{-1} are symmetric with respect to $\langle \cdot, \cdot \rangle$, that is $\langle L_x y, z \rangle = \langle y, L_x z \rangle$, and $\langle L_x^{-1} y, z \rangle = \langle y, L_x^{-1} z \rangle$.

Lemma 2.2 ([17, Lemma 2.9]). *For $x, y \in \mathcal{J}$, we have $\|x \circ y\|_F \leq \|x\|_F \|y\|_F$.*

Let $\{c_1, c_2, \ldots, c_r\}$ be a Jordan frame in $\mathcal{J}$. For $i, j \in \{1, 2, \cdots, r\}$, the Peirce spaces are given by $\mathcal{J}_{ii} := \{x \in \mathcal{J} : x \circ c_i = x\}$ and when $i \neq j$, $\mathcal{J}_{ij} := \{x \in \mathcal{J} : x \circ c_i = \frac{1}{2} x = x \circ c_j\}$. Then we have the following.

Theorem 2.3 ([4, Theorem IV.2.1]). *The space $\mathcal{J}$ is the orthogonal direct sum of spaces $\mathcal{J}_{ij}$ $(i \leq j)$.*

Thus, given a Jordan frame $\{c_1, c_2, \cdots, c_r\}$, we can write $x \in \mathcal{J}$ as

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

where $x_i \in \mathcal{R}$ and $x_{ij} \in \mathcal{J}_{ij}$. This expression is the *Peirce decomposition* of x with respect to $\{c_1, c_2, \cdots, c_r\}$.

We state two useful propositions about the quadratic representation.

Proposition 2.1 ([4, Proposition III.2.2]). *If $x, y \in \operatorname{int}\mathcal{K}$, then $Q_x y \in \operatorname{int}\mathcal{K}$.*

Proposition 2.2 ([18, Proposition 21]). *Let $x, y, p \in \operatorname{int}\mathcal{K}$ and define $\widetilde{x} := Q_p x$ and $\widetilde{y} := Q_{p^{-1}} y$. Then*

1. *$Q_{x^{1/2}} y$ and $Q_{y^{1/2}} x$ have the same spectrum.*
2. *$Q_{x^{1/2}} y$ and $Q_{\widetilde{x}^{1/2}} \widetilde{y}$ have the same spectrum.*

We say two elements $x, y \in \mathcal{J}$ *operator commute* if $L_x L_y = L_y L_x$. The tool of operator commutativity is very useful in the analysis of algorithms.

Theorem 2.4 ([18, Theorem 27]). *Let x and y be two elements of Euclidean Jordan algebra $\mathcal{J}$. Then x and y operator commute if and only if there is a Jordan frame $c_1, \cdots, c_r$ such that $x = \sum_{i=1}^{r} \lambda_i c_i$ and $y = \sum_{i=1}^{r} \mu_i c_i$.*

Lemma 2.3 ([18, Lemma 30]). *Let $x, y \in \operatorname{int}\mathcal{K}$ and define $w := Q_{x^{1/2}} y$, then $\operatorname{tr}(x \circ y) = \operatorname{tr}(w)$, $\lambda_{\min}(x \circ y) \le \lambda_{\min}(w)$, and $\lambda_{\max}(x \circ y) \ge \lambda_{\max}(w)$. Moreover, if x and y operator commute then $x \circ y = w$.*

For any $x \in \mathcal{J}$ with spectral decomposition $x = \lambda_1 c_1 + \ldots + \lambda_r c_r$, we define the *positive part* and *negative part* of x by

$$x^+ = \lambda_1^+ c_1 + \cdots + \lambda_r^+ c_r \qquad \text{and} \qquad x^- = \lambda_1^- c_1 + \cdots + \lambda_r^- c_r,$$

where $\lambda_i^+ = \max(\lambda_i, 0)$ and $\lambda_i^- = \min(\lambda_i, 0)$.

Lemma 2.4. $\|(x+y)^+\|_F \le \|x^+ + y^+\|_F \le \|x^+\|_F + \|y^+\|_F, \ \forall\, x, y \in \mathcal{J}.$

Proof. Let $x+y$ have the spectral decomposition $x+y = \lambda_1 c_1 + \cdots + \lambda_r c_r$, where $\{c_1, \cdots, c_r\}$ is a Jordan frame and the eigenvalues satisfy

$$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_k \ge 0 \ge \lambda_{k+1} \ge \cdots \ge \lambda_r.$$

By the definition of the positive part, we have $(x+y)^+ = \lambda_1 c_1 + \cdots + \lambda_k c_k$. Thus

$$\|(x+y)^+\|_F^2 = \lambda_1^2 + \cdots + \lambda_k^2.$$

On the other hand, let the Peirce decomposition of $a = x^+ + y^+$ with respect to $\{c_1, c_2, \cdots, c_r\}$ be $a = \sum_{i=1}^{r} a_i c_i + \sum_{i<j} a_{ij}$, where $a_i \in \mathcal{R}$ and $a_{ij} \in \mathcal{J}_{ij}$. Then, for $1 \le i \le k$, we have $a_i = \langle x^+ + y^+,\ c_i \rangle \ge \langle x + y,\ c_i \rangle = \lambda_i \ge 0$, where the first inequality follows from $x^+ + y^+ = x + y - (x^- + y^-) \succeq x + y$.

By the orthogonality of the spaces $\mathcal{J}_{ij}$, we have

$$\|a\|_F^2 = \|\sum_{i=1}^{r} a_i c_i + \sum_{i<j} a_{ij}\|_F^2 = \|\sum_{i=1}^{r} a_i c_i\|_F^2 + \|\sum_{i<j} a_{ij}\|_F^2,$$

which implies

$$\|x^+ + y^+\|_F^2 \geq \|\sum_{i=1}^{r} a_i c_i\|_F^2 = \sum_{i=1}^{r} a_i^2 \geq \sum_{i=1}^{k} a_i^2 \geq \sum_{i=1}^{k} \lambda_i^2 = \|(x+y)^+\|_F^2,$$

from which the first inequality follows. The second inequality follows the triangle inequality. □

3 $\mathcal{K}$-LP Problems and Wide Neighbourhood

Let $\mathcal{J}$ be a Euclidean Jordan algebra with dimension n, rank r, and cone of squares $\mathcal{K}$. Consider the primal-dual pair of $\mathcal{K}$-LP problems

$$\begin{aligned} \text{(P)} \quad & \min \ \langle c, x \rangle, \\ & \text{s.t.} \ Ax = b, \ x \in \mathcal{K}, \end{aligned} \tag{2}$$

and

$$\begin{aligned} \text{(D)} \quad & \max \ \langle b, y \rangle, \\ & \text{s.t.} \ \ A^* y + s = c, \ s \in \mathcal{K}, \ y \in \mathcal{R}^m, \end{aligned} \tag{3}$$

where $c \in \mathcal{J}$ and $b \in \mathcal{R}^m$. Here A is a linear operator that maps $\mathcal{J}$ into $\mathcal{R}^m$ and A^* is its adjoint operator. We call x and (y, s) primal and dual feasible solutions if they satisfy the primal and dual constraints respectively. We denote the sets of optimal solutions of (P) and (D) by $\mathcal{P}^*$ and $\mathcal{D}^*$ respectively. A problem (P) (resp. (D)) is called solvable if $\mathcal{P}^*$ (resp. $\mathcal{D}^*$) is nonempty. For convenience of reference, we define the following two sets:

$$\begin{aligned} \mathcal{F} &:= \{(x, y, s) \in \mathcal{K} \times \mathcal{R}^m \times \mathcal{K} : Ax = b, A^* y + s = c\}, \\ \mathcal{F}^0 &:= \{(x, y, s) \in \operatorname{int} \mathcal{K} \times \mathcal{R}^m \times \operatorname{int} \mathcal{K} : Ax = b, \ A^* y + s = c\}. \end{aligned}$$

We call $\mathcal{F}$ and $\mathcal{F}^0$, respectively, the (primal-dual) feasibility set and strictly feasibility set. (x, y, s) is said to be feasible if it is in $\mathcal{F}$ and strictly feasible if it is in $\mathcal{F}^0$. In this paper, we assume that A is surjective and $\mathcal{F}^0 \neq \phi$.

It has shown that [6,13], under the assumptions above, the sets of optimal solutions $\mathcal{P}^*$ and $\mathcal{D}^*$ are nonempty and bounded, and moreover $\langle x^*, s^* \rangle = 0$ for $x^* \in \mathcal{P}^*$ and $(y^*, s^*) \in \mathcal{D}^*$. He also proved that, for $x, s \in \mathcal{K}$, $\langle x, s \rangle = 0$ is equivalent to $x \circ s = 0$. Therefore, x^* and (y^*, s^*) are optimal solutions if and only if they satisfy the following system

$$\begin{aligned} & Ax = b, && x \in \mathcal{K}, \\ & A^* y + s = c, && s \in \mathcal{K}, \quad y \in \mathcal{R}^m, \\ & x \circ s = 0, \end{aligned} \tag{4}$$

where the last equality is called the complementarity slackness condition. Replace $x \circ s = 0$ in (4) with the perturbed complementary slackness condition, $x \circ s = \mu e$ for $\mu > 0$, we have the relaxed system

$$\begin{aligned} & Ax = b, && x \in \mathcal{K}, \\ & A^* y + s = c, && s \in \mathcal{K}, \quad y \in \mathcal{R}^m, \\ & x \circ s = \mu e. \end{aligned} \tag{5}$$

Primal-dual path-following interior-point algorithms follow the solutions of the relaxed System (5) as μ goes to zero. The relaxed system have unique solutions for all $\mu > 0$, and these solutions form the so-called central trajectory (central path), denoted by $\mathcal{C}$. Moreover, the limit of the trajectory as μ goes to 0 yields optimal solution for (P) and (D).

In the classical IPMs, the iterates are allowed to move in a neighbourhood of the central path. The so-called negative infinity neighbourhood that is a wide neighbourhood, is defined as

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

where $\gamma \in (0, 1)$ and $\mu = \langle x, s\rangle/r$ is the normalized duality gap. [2] introduced a new neighbourhood of the central path for LCP. Later, [9] extended it to SDP problems. Analogously, we define our neighbourhood as

$$\mathcal{N}(\tau_1, \tau_2, \eta) := \mathcal{N}_\infty^-(1-\tau_2) \cap \big\{(x, y, s) : \|(\tau_1\mu e - Q_{x^{1/2}}s)^+\|_F \leq \eta(\tau_1 - \tau_2)\mu\big\}, \tag{6}$$

where $\eta \geq 1$ and $0 < \tau_2 < \tau_1 < 1$.

This neighbourhood is a wide neighbourhood since one can verify that

$$\mathcal{N}_\infty^-(1-\tau_1) \subseteq \mathcal{N}(\tau_1, \tau_2, \eta) \subseteq \mathcal{N}_\infty^-(1-\tau_2), \qquad \forall\, \eta \geq 1, \quad 0 < \tau_2 < \tau_1 < 1. \tag{7}$$

Specially, if $\eta = 1$, it can be expressed more simply as follows:

$$\mathcal{N}(\tau_1, \tau_2, 1) = \{(x, y, s) \in \operatorname{int}\mathcal{K} \times \mathcal{R}^m \times \operatorname{int}\mathcal{K} : \|(\tau_1\mu e - Q_{x^{1/2}}s)^+\|_F \leq (\tau_1 - \tau_2)\mu\}.$$

For simplicity, we shall choose $\eta = 1$ in this paper, and we introduce a new notation $\mathcal{N}(\tau_1, \beta)$ to indicate this neighbourhood, that is

$$\mathcal{N}(\tau_1, \beta) := \{(x, y, s) \in \operatorname{int}\mathcal{K} \times \mathcal{R}^m \times \operatorname{int}\mathcal{K} : \|(\tau_1\mu e - Q_{x^{1/2}}s)^+\|_F \leq \beta\tau_1\mu\}. \tag{8}$$

where $\beta = (\tau_1 - \tau_2)/\tau_1$. Note that by part (i) of Proposition 2.2, $Q_{x^{1/2}}s$ and $Q_{s^{1/2}}x$ have the same spectrum, and thus $\mathcal{N}_\infty^-(1-\gamma)$ and $\mathcal{N}(\tau_1, \tau_2, \eta)$ are symmetric with respect to x and s.

4 Search Direction and Algorithm Framework

Most classic primal-dual path-following algorithms take Newton steps toward points on the central path $\mathcal{C}$, defined by System (5) for $\mu > 0$, rather than pure Newton steps for the optimality System (4), sometimes known as the affine-scaling direction. Since these steps are biased toward the interior of $\mathcal{K}$, it usually is possible to take longer steps along them than along the pure Newton steps for (4) before violating the positive definite condition. To move from the current point (x, y, s) towards the target on the central path corresponds to $\tau\mu$ leads us to the linear system

$$\begin{aligned} &A\triangle x = b - Ax, \\ &A^*\triangle y + \triangle s = c - s - A^*y, \\ &\triangle x \circ s + x \circ \triangle s = \tau\mu e - x \circ s, \end{aligned} \tag{9}$$

where $(\triangle x, \triangle y, \triangle s) \in \mathcal{J} \times \mathcal{R}^m \times \mathcal{J}$ is the search direction, $\tau \in [0, 1]$ is called centering parameter. [18, Lemma 28] show an equivalent way of writing the complementarity condition

$x \circ s = \mu e$. Let $x, s \in \operatorname{int} \mathcal{K}$ and p invertible. Then $x \circ s = \mu e$ if and only if $Q_p x \circ Q_{p^{-1}} s = \mu e$. Thus, the System (5) can be equivalently written as

$$\begin{aligned} &\widetilde{A}\widetilde{x} = b, && \widetilde{x} \in \mathcal{K}, \\ &\widetilde{A}^* y + \widetilde{s} = \widetilde{c}, && \widetilde{s} \in \mathcal{K}, \quad y \in \mathcal{R}^m, \\ &\widetilde{x} \circ \widetilde{s} = \mu e, && \end{aligned} \tag{10}$$

where $\widetilde{A} = AQ_{p^{-1}}$, $\widetilde{c} = Q_{p^{-1}} c$, $\widetilde{x} = Q_p x$, and $\widetilde{s} = Q_{p^{-1}} s$. Denote by $\mathcal{C}(x, s)$ the set of all elements so that the scaled elements operator commute, i.e.

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

This is a subclass of the Monteiro-Zhang family of search directions called the commutative class. In particular, choosing $p = s^{1/2}$ and $p = x^{-1/2}$ we get the xs and sx search directions respectively. For the choice of

$$p = [Q_{x^{1/2}} (Q_{x^{1/2}} s)^{-1/2}]^{-1/2} = [Q_{s^{-1/2}} (Q_{s^{1/2}} x)^{1/2}]^{-1/2}, \tag{11}$$

we obtain the NT search direction.

In this paper, we restrict the scaling $p \in \mathcal{C}(x, s)$. Corresponding to the scaling System (10), the Newton System (9) becomes

$$\begin{aligned} &\widetilde{A} \triangle \widetilde{x} = b - \widetilde{A}\widetilde{x}, \\ &\widetilde{A}^* \triangle y + \triangle \widetilde{s} = \widetilde{c} - \widetilde{A}^* y - \widetilde{s}, \\ &\triangle \widetilde{x} \circ \widetilde{s} + \widetilde{x} \circ \triangle \widetilde{s} = \tau \mu e - \widetilde{x} \circ \widetilde{s}. \end{aligned} \tag{12}$$

In our new algorithm, we decompose the Newton System (12) into the following two systems:

$$\begin{aligned} &\widetilde{A} \triangle \widetilde{x}_- = b - \widetilde{A}\widetilde{x}, \\ &\widetilde{A}^* \triangle y_- + \triangle \widetilde{s}_- = \widetilde{c} - \widetilde{A}^* y - \widetilde{s}, \\ &\triangle \widetilde{x}_- \circ \widetilde{s} + \widetilde{x} \circ \triangle \widetilde{s}_- = (\tau \mu e - \widetilde{x} \circ \widetilde{s})^-. \end{aligned} \tag{13}$$

and

$$\begin{aligned} &\widetilde{A} \triangle \widetilde{x}_+ = 0, \\ &\widetilde{A}^* \triangle y_+ + \triangle \widetilde{s}_+ = 0, \\ &\triangle \widetilde{x}_+ \circ \widetilde{s} + \widetilde{x} \circ \triangle \widetilde{s}_+ = (\tau \mu e - \widetilde{x} \circ \widetilde{s})^+. \end{aligned} \tag{14}$$

As pointed in [2,9], the negative part $(\tau\mu e - \widetilde{x} \circ \widetilde{s})^-$ is responsible for reducing the duality gap, and the positive part $(\tau\mu e - \widetilde{x} \circ \widetilde{s})^+$ is used to control the centrality. [2] suggested to treat the negative part and the positive part separately to obtain a better iteration complexity bound for wide neighbourhood IPMs.

Let $\alpha := (\alpha_1, \alpha_2) \in [0, 1]^2$ be the step sizes taken along $(\triangle \widetilde{x}_-, \triangle y_-, \triangle \widetilde{s}_-)$ and $(\triangle \widetilde{x}_+, \triangle y_+, \triangle \widetilde{s}_+)$ respectively. The new iterate is

$$(\widetilde{x}(\alpha), y(\alpha), \widetilde{s}(\alpha)) := (\widetilde{x}, y, \widetilde{s}) + \alpha_1 (\triangle \widetilde{x}_-, \triangle y_-, \triangle \widetilde{s}_-) + \alpha_2 (\triangle \widetilde{x}_+, \triangle y_+, \triangle \widetilde{s}_+). \tag{15}$$

The following proposition shows the scale-invariance of the neighbourhood.

Proposition 4.1. *The neighbourhood $\mathcal{N}(\tau_1,\beta)$ is scaling invariant, that is (x,y,s) is in the neighbourhood if and only if $(\widetilde{x},y,\widetilde{s})$ is.*

Proof. Let $\widetilde{w}=Q_{\widetilde{x}^{1/2}}\widetilde{s}$. Note that the neighbourhood $\mathcal{N}(\tau_1,\beta)$ can be defined in terms of eigenvalues of w and by part (ii) of Proposition 2.2 w and $\widetilde{w}$ have the same eigenvalues. The required result follows. □

Having introduced the key elements for the new algorithm, we state the generic framework of our algorithm.

Algorithm 4.1. *Input parameters: an accuracy parameter $\varepsilon>0$, neighbourhood parameters $0<\tau_1,\ \beta<1$, a centering parameter $0\le\tau\le 1$ and an initial point $(x^0,y^0,s^0)\in\mathcal{N}(\tau_1,\beta)$. Set $\mu_0=\langle x^0,s^0\rangle/r, k:=0$.*

Step 1 If $\mu_k\le\varepsilon\mu_0$, then stop.

Step 2 Choose a scaling element $p\in\mathcal{C}(x^k,s^k)$ and compute $(\widetilde{x}^k,\widetilde{s}^k)$.

Step 3 Compute the directions $(\triangle\widetilde{x}^k_-,\triangle y^k_-,\triangle\widetilde{s}^k_-)$ and $(\triangle\widetilde{x}^k_+,\triangle y^k_+,\triangle\widetilde{s}^k_+)$ by solving the scaled Newton systems (13) and (14) respectively.

Step 4 Choose step size vector $\alpha^k=(\alpha^k_1,\alpha^k_2)$, such that the new iterates

$$(\widetilde{x}^{k+1},y^{k+1},\widetilde{s}^{k+1}):=(\widetilde{x}^k,y^k,\widetilde{s}^k)+\alpha^k_1(\triangle\widetilde{x}^k_-,\triangle y^k_-,\triangle\widetilde{s}^k_-)+\alpha^k_2(\triangle\widetilde{x}^k_+,\triangle y^k_+,\triangle\widetilde{s}^k_+),$$

remain in $\mathcal{N}(\tau_1,\beta)$.

Step 5 Let $(x^{k+1},y^{k+1},s^{k+1})=(Q_{p^{-1}}\widetilde{x}^{k+1},y^{k+1},Q_p\widetilde{s}^{k+1})$ and $\mu_{k+1}=\langle x^{k+1},s^{k+1}\rangle/r$. Set $k:=k+1$ and go to Step 1.

We note that in the practical implementations, the step sizes $\alpha^k=(\alpha^k_1,\alpha^k_2)$ are chosen to be a large fraction, for example, 98% of $\hat{\alpha}^k=(\hat{\alpha}^k_1,\hat{\alpha}^k_2)$, the sizes of the step to the boundary of the symmetric cone $\mathcal{K}$. This choice for α^k may be adequate for practical purposes, but we need more elaborated choices for theoretically guaranteed convergence. In next section, we specify our choice for α^k and present the convergency and iteration-complexity of Algorithm 4.1. Our choice is based on several factors, including keeping the centrality, improvement of the infeasibility, and decreasing of the duality gap.

Using (13) and (14) the following proposition is readily verified.

Proposition 4.2. *Let $\{(\widetilde{x}^k,y^k,\widetilde{s}^k)\}$ be generated by Algorithm 4.1. Then for $k\ge 0$, one has*

$$\widetilde{A}\widetilde{x}^{k+1}-b=\nu^{k+1}(\widetilde{A}\widetilde{x}^0-b),\qquad \widetilde{A}^*y^{k+1}+\widetilde{s}^{k+1}-\widetilde{c}=\nu^{k+1}(\widetilde{A}^*y^0+\widetilde{s}^0-\widetilde{c}),$$

where $\nu^0=1$ and

$$\nu^{k+1}=(1-\alpha^k_1)\nu^k=\prod_{i=0}^{k}(1-\alpha^i_1)\in[0,1]. \tag{16}$$

From Proposition 4.2, we have

$$\nu^k=\frac{\|\widetilde{A}\widetilde{x}^k-b\|_F}{\|\widetilde{A}\widetilde{x}^0-b\|_F}=\frac{\|\widetilde{A}^*y^k+\widetilde{s}^k-\widetilde{c}\|_F}{\|\widetilde{A}^*y^0+\widetilde{s}^0-\widetilde{c}\|_F},$$

which implies ν^k represents the relative infeasibility at $(\widetilde{x}^k,y^k,\widetilde{s}^k)$. Hence at every iterate we maintain the condition:

$$\langle\widetilde{x}^k,\widetilde{s}^k\rangle\ge\nu^k\langle\widetilde{x}^0,\widetilde{s}^0\rangle, \tag{17}$$

which ensures that the infeasibility approaches to zero as the complementarity $\langle x, s\rangle$ approaches to zero. Observe that $\langle \widetilde{x}^k, \widetilde{s}^k\rangle = 0$ is possible only if $\nu^k = 0$. In the case of $\langle \widetilde{x}^k, \widetilde{s}^k\rangle = 0$, $(Q_{p^{-1}}\widetilde{x}^k, y^k, Q_p\widetilde{s}^k)$ is a solution to (4) and Algorithm 4.1 terminates. However, it seems extremely unlikely for $\langle \widetilde{x}^k, \widetilde{s}^k\rangle = 0$ to happen in practice. Thus, we will not consider this finite termination case in our analysis of convergence.

We now specify a particular starting point for Algorithm 4.1. Let u^0 and (r^0, v^0) be the minimum-norm solutions to the linear systems $Ax = b$ and $A^*y + s = c$ respectively. That is

$$u^0 = \arg\min\{\|u\|_F : Au = b\}, \qquad (r^0, v^0) = \arg\min\{\|v\|_F : A^*r + v = c\}. \tag{18}$$

We choose (x^0, y^0, s^0) such that

$$x^0 = s^0 = \rho^0 e, \qquad \rho^0 \geq \max\{\|u^0\|_2, \|v^0\|_2\}. \tag{19}$$

This implies that $x^0, s^0 \in \operatorname{int}\mathcal{K}$, $x^0 - u^0 \in \mathcal{K}$ and $s^0 - v^0 \in \mathcal{K}$.

Let

$$\rho^* = \min\{\max(\|x^*\|_2, \|s^*\|_2) : x^* \in \mathcal{P}^*, (y^*, s^*) \in \mathcal{D}^*\}, \tag{20}$$

and in addition, we assume that for some constant $\Psi > 0$, it has $\rho^0 \geq \rho^*/\Psi$. Note that we can always increase ρ^0.

We constructed an auxiliary sequence $\{(u^k, r^k, v^k)\}$ as follows:

$$(u^{k+1}, r^{k+1}, v^{k+1}) = (x^{k+1}, y^{k+1}, s^{k+1}) - (1 - \alpha_1^k)(x^k - u^k, y^k - r^k, s^k - v^k). \tag{21}$$

The auxiliary sequence will be used in our analysis of complexity and need not be actually computed in Algorithm 4.1. The following lemma gives useful properties of the auxiliary sequence $\{(u^k, r^k, v^k)\}$.

Lemma 4.1. *Let $\{(x^k, y^k, s^k)\}$ be generated by Algorithm 4.1, $\{(u^k, r^k, v^k)\}$ be given by (21), and $\{\nu^k\}$ be given by (16). Then for $k \geq 0$*

(1) $Au^k = b$ and $A^*r^k + v^k = c$;

(2) $x^k - u^k = \nu^k(x^0 - u^0) \in \mathcal{K}$ and $s^k - v^k = \nu^k(s^0 - v^0) \in \mathcal{K}$.

Proof. The proof follows from direct substitution. □

5 Analysis of Polynomial Convergence for Algorithm

In this section, we first give the strategy of choice for step size α^k. Then we develop several technical lemmas. At the end of this section, we present our main result of polynomial convergence. For simplicity, from now on we will suppress the superscript k, except for $k = 0$ whenever no confusion arises. However, we will denote α^k by $\hat{\alpha}$ while using α as a free variable.

Our choice of $\hat{\alpha}$ is based on several considerations. We require that the step size vector $\widehat{\alpha} = (\hat{\alpha}_1, \hat{\alpha}_2)$ satisfies the following three conditions:

$$(\widetilde{x}(\alpha), y(\alpha), \widetilde{s}(\alpha)) \in \mathcal{N}(\tau_1, \beta), \tag{22}$$

$$\langle \widetilde{x}(\alpha), \widetilde{s}(\alpha)\rangle \geq (1 - \alpha_1)\nu\langle \widetilde{x}^0, \widetilde{s}^0\rangle, \tag{23}$$

$$\langle \widetilde{x}(\alpha), \widetilde{s}(\alpha)\rangle \leq (1 - (1 - \delta)\alpha_1)\langle \widetilde{x}, \widetilde{s}\rangle, \tag{24}$$

where $\nu = \nu^k$ is defined in (16), and $\delta \in (0, 1)$ is a constant independent of r. Condition (22) is a centrality condition that prevents iterates from prematurely getting too close to the boundary of the symmetric cone $\mathcal{K}$. Condition (23), as we see in (17), ensures that the infeasibility approaches to zero as the complementarity approaches to zero. Condition (24) is needed in order to make a comparable progress in the complementarity. From this point on, by Algorithm 4.1 we mean that the the step size $\hat{\alpha}$ satisfies (22)–(24). We note that Lemma 5.9, 5.10 and 5.11 guarantee the existence of $\widehat{\alpha}$ that satisfies conditions (22)–(24) simultaneously. For example, set $\alpha_2 = \sqrt{\tau_1}/(\sqrt{\mathrm{cond}\,(G)}\omega^2 r)$ and $\alpha_1 = \alpha_2\sqrt{\beta\tau_1/r}$.

5.1 Technical Lemmas

In this part, we let the centering parameter $\tau = \tau_1$. We will use the notation: $(\triangle\widetilde{x}(\alpha), \triangle y(\alpha), \triangle\widetilde{s}(\alpha)) = \alpha_1(\triangle\widetilde{x}_-, \triangle y_-, \triangle\widetilde{s}_-) + \alpha_2(\triangle\widetilde{x}_+, \triangle y_+, \triangle\widetilde{s}_+)$, and $\chi(\alpha) = \widetilde{x}\circ\widetilde{s} + \alpha_1(\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^- + \alpha_2(\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^+$. It can be easily verified that $\langle\triangle\widetilde{x}_+, \triangle\widetilde{s}_+\rangle = 0$, $\widetilde{x}(\alpha)\circ\widetilde{s}(\alpha) = \chi(\alpha) + \triangle\widetilde{x}(\alpha)\circ\triangle\widetilde{s}(\alpha)$,

$$\triangle\widetilde{x}(\alpha)\circ\triangle\widetilde{s}(\alpha) = \alpha_1^2\triangle\widetilde{x}_-\circ\triangle\widetilde{s}_- + \alpha_1\alpha_2(\triangle\widetilde{x}_-\circ\triangle\widetilde{s}_+ + \triangle\widetilde{s}_-\circ\triangle\widetilde{x}_+) + \alpha_2^2\triangle\widetilde{x}_+\circ\triangle\widetilde{s}_+, \tag{25}$$

and

$$\langle\widetilde{x}(\alpha), \widetilde{s}(\alpha)\rangle = \mathrm{tr}\,(\chi(\alpha)) + \alpha_1^2\langle\triangle\widetilde{x}_-, \triangle\widetilde{s}_-\rangle + \alpha_1\alpha_2(\langle\triangle\widetilde{x}_-, \triangle\widetilde{s}_+\rangle + \langle\triangle\widetilde{s}_-, \triangle\widetilde{x}_+\rangle). \tag{26}$$

By using $\mathrm{tr}\,((\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^-) + \mathrm{tr}\,((\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^+) = \mathrm{tr}\,(\tau_1\mu e - \widetilde{x}\circ\widetilde{s}) = -(1-\tau_1)r\mu$ we have

$$\mathrm{tr}\,((\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^-) \le -(1-\tau_1)r\mu. \tag{27}$$

When $(\widetilde{x}, y, \widetilde{s}) \in \mathcal{N}(\tau_1, \beta)$, we have

$$\mathrm{tr}\,((\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^+) \le \sqrt{r}\|(\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^+\|_F \le \sqrt{r}\beta\tau_1\mu. \tag{28}$$

The following lemma will be used frequently during the analysis.

Lemma 5.1 ([18, Lemma 33]). *Let $p, q \in \mathcal{J}$ and G a positive definite matrix which is symmetric with respect to the inner product $\langle\cdot,\cdot\rangle$. Then*

$$\begin{aligned}\|p\|_F\|q\|_F \le& \sqrt{\mathrm{cond}\,(G)}\|G^{-1/2}p\|_F\|G^{1/2}q\|_F \\ \le& \frac{1}{2}\sqrt{\mathrm{cond}\,(G)}\big(\|G^{-1/2}p\|_F^2 + \|G^{1/2}q\|_F^2\big),\end{aligned}$$

where $\mathrm{cond}\,(G) = \lambda_{\max}(G)/\lambda_{\min}(G)$.

We note that $\|\cdot\|_G$, defined by $\|(u,v)\|_G = (\|G^{-1/2}u\|_F^2 + \|G^{1/2}v\|_F^2)^{1/2}$, $u, v \in \mathcal{J}$, is a norm on $\mathcal{J}\times\mathcal{J}$.

Lemma 5.2. *Let $G = L_{\widetilde{s}}^{-1}L_{\widetilde{x}}$. If $\beta \le 1/2$, then $\|(\triangle\widetilde{x}_+, \triangle\widetilde{s}_+)\|_G^2 \le \beta\tau_1\mu$.*

Proof. Since $\widetilde{x}$ and $\widetilde{s}$ operator commute, there is a Jordan frame $c_1, \ldots, c_r$ such that $\widetilde{x} = \sum_{i=1}^r \lambda_i c_i$ and $\widetilde{s} = \sum_{i=1}^r \mu_i c_i$. Then, $\widetilde{x}\circ\widetilde{s} = \sum_{i=1}^r \lambda_i\mu_i c_i$, and for $i = 1, \cdots, r$, $L_{\widetilde{x}}L_{\widetilde{s}}c_i = \widetilde{x}\circ(\widetilde{s}\circ c_i) = \lambda_i\mu_i c_i$, which implies

$$(L_{\widetilde{x}}L_{\widetilde{s}})^{-1}c_i = c_i/(\lambda_i\mu_i). \tag{29}$$

Multiplying the last equation of (14) by $(L_{\widetilde{x}}L_{\widetilde{s}})^{-1/2}$, we obtain

$$G^{-1/2}\triangle\widetilde{x}_+ + G^{1/2}\triangle\widetilde{s}_+ = (L_{\widetilde{x}}L_{\widetilde{s}})^{-1/2}(\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^+.$$

Taking norm-squared on both sides, together with the fact $\langle \Delta\widetilde{x}_+, \Delta\widetilde{s}_+\rangle = 0$, we have

$$\begin{aligned}
&\|G^{-1/2}\triangle\widetilde{x}_+\|_F^2 + \|G^{1/2}\triangle\widetilde{s}_+\|_F^2\\
=&\|(L_{\widetilde{x}}L_{\widetilde{s}})^{-1/2}(\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^+\|_F^2\\
=&\langle (L_{\widetilde{x}}L_{\widetilde{s}})^{-1/2}(\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^+, (L_{\widetilde{x}}L_{\widetilde{s}})^{-1/2}(\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^+\rangle\\
=&\langle (\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^+, (L_{\widetilde{x}}L_{\widetilde{s}})^{-1}(\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^+\rangle\\
=&\Big\langle \sum(\tau_1\mu - \lambda_i\mu_i)^+c_i, (L_{\widetilde{x}}L_{\widetilde{s}})^{-1}\sum(\tau_1\mu - \lambda_i\mu_i)^+c_i\Big\rangle\\
=&\Big\langle \sum(\tau_1\mu - \lambda_i\mu_i)^+c_i, \sum(\tau_1\mu - \lambda_i\mu_i)^+c_i/(\lambda_i\mu_i)\Big\rangle\\
=&\sum\big[(\tau_1\mu - \lambda_i\mu_i)^+\big]^2/(\lambda_i\mu_i)\\
\le&(\beta\tau_1\mu)^2/(\tau_2\mu)\\
\le&\beta\tau_1\mu,
\end{aligned}$$

Here, the fifth equality follows from (29), the first inequality follows from $(\widetilde{x}, y, \widetilde{s}) \in \mathcal{N}(\tau_1, \beta)$, and the last inequality holds due to $\tau_2 = (1-\beta)\tau_1$ and $\beta \le 1/2$. □

Lemma 5.3. *Let* $G = L_{\widetilde{s}}^{-1}L_{\widetilde{x}}$. *Then* $\|(\triangle\widetilde{x}_-, \triangle\widetilde{s}_-)\|_G \le \sqrt{r\mu} + (1+\sqrt{2})\xi$, *where* $\xi := \min\{\|(\bar{u},\bar{v})\|_G : \widetilde{A}\bar{u} = b - \widetilde{A}\widetilde{x},\ \widetilde{A}^*\bar{r} + \bar{v} = \widetilde{c} - \widetilde{A}^*y - \widetilde{s}\}$.

Proof. Let $(\bar{u}, \bar{r}, \bar{v}) \in \mathcal{J}\times\mathcal{R}^m\times\mathcal{J}$ satisfy the equations $\widetilde{A}\bar{u} = b - \widetilde{A}\widetilde{x}$ and $\widetilde{A}^*\bar{r} + \bar{v} = \widetilde{c} - \widetilde{A}^*y - \widetilde{s}$, then by System (13) we have

$$\begin{aligned}
&\widetilde{A}(\triangle\widetilde{x}_- - \bar{u}) = 0,\\
&\widetilde{A}^*(\triangle y_- - \bar{r}) + (\triangle\widetilde{s}_- - \bar{v}) = 0,\\
&L_{\widetilde{s}}(\triangle\widetilde{x}_- - \bar{u}) + L_{\widetilde{x}}(\triangle\widetilde{s}_- - \bar{v}) = (\tau\mu e - \widetilde{x}\circ\widetilde{s})^- - (L_{\widetilde{s}}\bar{u} + L_{\widetilde{x}}\bar{v}).
\end{aligned}$$

Multiplying the last equation by $(L_{\widetilde{x}}L_{\widetilde{s}})^{-1/2}$, we obtain

$$G^{-1/2}(\triangle\widetilde{x}_- - \bar{u}) + G^{1/2}(\triangle\widetilde{s}_- - \bar{v}) = (L_{\widetilde{x}}L_{\widetilde{s}})^{-1/2}(\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^- - (G^{-1/2}\bar{u} + G^{1/2}\bar{v}).$$

Therefore

$$\begin{aligned}
&\|(\triangle\widetilde{x}_-, \triangle\widetilde{s}_-)\|_G\\
\le&\|(\triangle\widetilde{x}_- - \bar{u}, \triangle\widetilde{s}_- - \bar{v})\|_G + \|(\bar{u},\bar{v})\|_G\\
=&\|(L_{\widetilde{x}}L_{\widetilde{s}})^{-1/2}(\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^- - (G^{-1/2}\bar{u} + G^{1/2}\bar{v})\|_F + \|(\bar{u},\bar{v})\|_G\\
\le&\|(L_{\widetilde{x}}L_{\widetilde{s}})^{-1/2}(\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^-\|_F + \|G^{-1/2}\bar{u}\|_F + \|G^{1/2}\bar{v}\|_F + \|(\bar{u},\bar{v})\|_G
\end{aligned}$$

where the equality holds due to $\langle \triangle\widetilde{x}_- - \bar{u}, \triangle\widetilde{s}_- - \bar{v}\rangle = 0$.

Similar to the proof of Lemma 5.2,

$$\|(L_{\widetilde{x}}L_{\widetilde{s}})^{-1/2}(\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^-\|_F^2 = \sum_{i=1}^r[(\tau_1\mu - \lambda_i\mu_i)^-]^2/(\lambda_i\mu_i) \le \sum_{i=1}^r \lambda_i\mu_i = r\mu.$$

By the definition of $\|\cdot\|_G$, we have $\|G^{-1/2}\bar{u}\|_F + \|G^{1/2}\bar{v}\|_F \le \sqrt{2}\|(\bar{u},\bar{v})\|_G$. Hence the required result follows. □

Lemma 5.4. *Let* (u^0, r^0, v^0) *and* (x^0, y^0, s^0) *satisfy (18) and (19) respectively. Then* $\xi \le (5+4\Psi)r\sqrt{\mu}/\sqrt{\tau_2}$.

Proof. Let $\widetilde{u} = Q_p u$ and $\widetilde{v} = Q_{p^{-1}} v$. Then by Lemma 4.1 we have $\widetilde{A}\widetilde{u} = b$ and $\widetilde{A}^* r + \widetilde{v} = \widetilde{c}$. Moreover, $\widetilde{x} - \widetilde{u} \in \mathcal{K}$ and $\widetilde{s} - \widetilde{v} \in \mathcal{K}$. Let $(\bar{u}, \bar{r}, \bar{v}) := (\widetilde{x} - \widetilde{u}, y - r, \widetilde{s} - \widetilde{v})$, then, one has $\widetilde{A}(-\bar{u}) = b - \widetilde{A}\widetilde{x}$, $\widetilde{A}^*(-\bar{r}) + (-\bar{v}) = \widetilde{c} - \widetilde{A}^* y - \widetilde{s}$. Hence,

$$\xi \leq \|(-\bar{u}, -\bar{v})\|_G \leq \|G^{-1/2}\bar{u}\|_F + \|G^{1/2}\bar{v}\|_F.$$

Since $\widetilde{x}$ and $\widetilde{s}$ operator commute, then G and $Q_{\widetilde{x}}$ commute, and we have

$$\|G^{1/2}\bar{v}\|_F^2 = \langle \bar{v}, G\bar{v} \rangle = \langle Q_{\widetilde{x}}^{1/2}\bar{v}, Q_{\widetilde{x}}^{-1} G Q_{\widetilde{x}}^{1/2}\bar{v} \rangle \leq \lambda_{\max}(Q_{\widetilde{x}}^{-1}G)\|Q_{\widetilde{x}}^{1/2}\bar{v}\|_F^2.$$

Then, by using [17, Lemma 4.1] we have

$$\|G^{1/2}\bar{v}\|_F^2 \leq \frac{\langle \widetilde{x}, \bar{v} \rangle^2}{\lambda_{\min}(\widetilde{w})} = \frac{\langle \widetilde{x}, \widetilde{v} - \widetilde{s} \rangle^2}{\lambda_{\min}(\widetilde{w})} \leq \frac{\langle \widetilde{x}, \widetilde{v} - \widetilde{s} \rangle^2}{\tau_2 \mu} = \frac{\langle x, v - s \rangle^2}{\tau_2 \mu}.$$

Similarly it can be shown that $\|G^{-1/2}\bar{u}\|_F^2 \leq \frac{\langle s, u - x \rangle^2}{\tau_2 \mu}$. Therefore,

$$\xi \leq \frac{\langle x, s - v \rangle + \langle x - u, s \rangle}{\sqrt{\tau_2 \mu}}. \tag{30}$$

For $x^* \in \mathcal{P}^*$, and $(y^*, s^*) \in \mathcal{D}^*$, we have $A(x^* - u) = 0$ and $A^*(y^* - r) + (s^* - v) = 0$ by using part one of Lemma 4.1. Hence,

$$\begin{aligned}
0 =& \langle x^* - u, s^* - v \rangle = \langle x^* - x + x - u, s^* - s + s - v \rangle \\
=& \langle x^*, s^* \rangle + \langle x, s \rangle + \langle x^*, s - v \rangle + \langle x - u, s^* \rangle + \langle x - u, s - v \rangle \\
& - \langle x^*, s \rangle - \langle x, s^* \rangle - \langle x, s - v \rangle - \langle x - u, s \rangle.
\end{aligned}$$

It follows that

$$\begin{aligned}
\langle x, s - v \rangle + \langle x - u, s \rangle =& \langle x^*, s^* \rangle + \langle x, s \rangle + \langle x^*, s - v \rangle + \langle x - u, s^* \rangle \\
& + \langle x - u, s - v \rangle - \langle x^*, s \rangle - \langle x, s^* \rangle \\
\leq& \langle x, s \rangle + \langle x^*, s - v \rangle + \langle x - u, s^* \rangle + \langle x - u, s - v \rangle \\
=& \Big(1 + \frac{\langle x^*, s - v \rangle + \langle x - u, s^* \rangle + \langle x - u, s - v \rangle}{\langle x, s \rangle}\Big)\langle x, s \rangle \\
=& \Big(1 + \frac{\nu\langle x^*, s^0 - v^0 \rangle + \nu\langle x^0 - u^0, s^* \rangle + \nu^2\langle x^0 - u^0, s^0 - v^0 \rangle}{\langle x, s \rangle}\Big) r\mu \\
\leq& \Big(1 + \frac{\langle x^*, s^0 - v^0 \rangle + \langle x^0 - u^0, s^* \rangle + \langle x^0 - u^0, s^0 - v^0 \rangle}{\langle x^0, s^0 \rangle}\Big) r\mu,
\end{aligned} \tag{31}$$

where the third equation follows from part two of Lemma 4.1, and the last inequality follows from (17), (19) and $0 < \nu < 1$. For the initial points choice as in Section 4, it hold that

$$\begin{aligned}
& \frac{\langle x^*, s^0 - v^0 \rangle + \langle x^0 - u^0, s^* \rangle + \langle x^0 - u^0, s^0 - v^0 \rangle}{\langle x^0, s^0 \rangle} \\
\leq& \frac{2r\rho^*\rho^0 + 2r\rho^*\rho^0 + 4r(\rho^0)^2}{r(\rho^0)^2} = 4 + 4\rho^*/\rho^0 \leq 4 + 4\Psi,
\end{aligned} \tag{32}$$

where first inequality follows from (18)–(20), and the facts: $\|p\|_F \leq \sqrt{r}\|p\|_2$, $\langle p, q \rangle \leq \|p\|_F\|q\|_F \leq r\|p\|_2\|q\|_2$.

By substituting (31) and (32) into (30), we obtain the required result. □

Lemma 5.5. *Let* $G = L_{\widetilde{s}}^{-1} L_{\widetilde{x}}$. *Then* $\|(\triangle \widetilde{x}_-, \triangle \widetilde{s}_-)\|_G^2 \leq \omega^2 r^2 \mu$, *where*

$$\omega = \big(1 + (1+\sqrt{2})(5+4\Psi)\big)\big/\sqrt{\tau_2} \geq 13. \tag{33}$$

By Lemma 5.1, 5.2 and 5.5, we have the following corollary.

Corollary 5.1. *Let* $G = L_{\widetilde{s}}^{-1} L_{\widetilde{x}}$. *If* $\beta \leq 1/2$, *then*

(1) $\|\triangle \widetilde{x}_+\|_F \|\triangle \widetilde{s}_+\|_F \leq \sqrt{\text{cond}\,(G)}\beta\tau_1\mu/2$;

(2) $\|\triangle \widetilde{x}_-\|_F \|\triangle \widetilde{s}_-\|_F \leq \sqrt{\text{cond}\,(G)}\omega^2 r^2\mu/2$;

(3) $\|\triangle \widetilde{x}_-\|_F \|\triangle \widetilde{s}_+\|_F \leq \sqrt{\text{cond}\,(G)}\sqrt{\beta\tau_1}\omega r\mu$;

(4) $\|\triangle \widetilde{s}_-\|_F \|\triangle \widetilde{x}_+\|_F \leq \sqrt{\text{cond}\,(G)}\sqrt{\beta\tau_1}\omega r\mu$,

where $\text{cond}(G) = \lambda_{\max}(G)/\lambda_{\min}(G)$.

Lemma 5.6. *Let* $\tau_1 \leq 1/8, \beta \leq 1/2$ *and* $(\widetilde{x}, y, \widetilde{s}) \in \mathcal{N}(\tau_1, \beta)$. *If* $\alpha_1 = \alpha_2\sqrt{\beta\tau_1/r}$ *and* $\alpha_2 \leq \sqrt{\tau_1}/(\sqrt{\text{cond}\,(G)}\omega^2 r)$, *then we have* $\mu(\alpha) \leq \big(1 - \frac{\alpha_2\sqrt{\beta\tau_1}}{2\sqrt{r}}\big)\mu$.

Proof. By using (27) and (28), one has $\text{tr}\,(\chi(\alpha)) \leq r\mu - \alpha_1(1-\tau_1)r\mu + \alpha_2\sqrt{r}\beta\tau_1\mu$. Then from (26), we have

$$\begin{aligned}\langle \widetilde{x}(\alpha), \widetilde{s}(\alpha)\rangle =& \text{tr}\,(\chi(\alpha)) + \alpha_1^2\langle \triangle\widetilde{x}_-, \triangle\widetilde{s}_-\rangle + \alpha_1\alpha_2(\langle \triangle\widetilde{x}_-, \triangle\widetilde{s}_+\rangle + \langle \triangle\widetilde{s}_-, \triangle\widetilde{x}_+\rangle)\\ \leq& r\mu - \alpha_1(1-\tau_1)r\mu + \alpha_2\sqrt{r}\beta\tau_1\mu + \alpha_1^2\|\triangle\widetilde{x}_-\|_F\|\triangle\widetilde{s}_-\|_F\\ &+ \alpha_1\alpha_2(\|\triangle\widetilde{x}_-\|_F\|\triangle\widetilde{s}_+\|_F + \|\triangle\widetilde{s}_-\|_F\|\triangle\widetilde{x}_+\|_F).\end{aligned}$$

From Corollary 5.1,

$$\begin{aligned}\langle \widetilde{x}(\alpha), \widetilde{s}(\alpha)\rangle \leq& r\mu - \alpha_1(1-\tau_1)r\mu + \alpha_2\sqrt{r}\beta\tau_1\mu + \alpha_1^2\sqrt{\text{cond}\,(G)}\omega^2 r^2\mu/2\\ &+ 2\alpha_1\alpha_2\sqrt{\text{cond}\,(G)}\sqrt{\beta\tau_1}\omega r\mu\\ =& r\mu - \alpha_2(1-\tau_1)\sqrt{\beta\tau_1 r}\mu + \alpha_2\sqrt{r}\beta\tau_1\mu + \alpha_2^2\sqrt{\text{cond}\,(G)}\omega^2\beta\tau_1 r\mu/2\\ &+ 2\alpha_2^2\sqrt{\text{cond}\,(G)}\omega\beta\tau_1\sqrt{r}\mu\\ =& \big[1 - \alpha_2\sqrt{\beta\tau_1/r}\big(1 - \tau_1 - \sqrt{\beta\tau_1} - \alpha_2\sqrt{\text{cond}\,(G)}\omega^2\sqrt{\beta\tau_1 r}/2\\ &- 2\alpha_2\sqrt{\text{cond}\,(G)}\omega\sqrt{\beta\tau_1}\big)\big]r\mu\\ \leq& \big[1 - \alpha_2\sqrt{\beta\tau_1/r}\big(1 - \tau_1 - \sqrt{\beta\tau_1} - \sqrt{\beta}\tau_1/2 - 2\sqrt{\beta}\tau_1/\omega\big)\big]r\mu\\ \leq& \Big[1 - \frac{\alpha_2\sqrt{\beta\tau_1}}{2\sqrt{r}}\Big]\,r\mu,\end{aligned} \tag{34}$$

where, the last inequality follows from the fact $\tau_1 \leq 1/8, \beta \leq 1/2$ and $\omega \geq 13$. Then, by using $\mu(\alpha) = \langle \widetilde{x}(\alpha), \widetilde{s}(\alpha)\rangle/r$, we obtain the required result. □

Lemma 5.7. *If* $(\widetilde{x}, y, \widetilde{s}) \in \mathcal{N}(\tau_1, \beta)$ *and* $\mu(\alpha) \leq \mu$, *then*

$$\|(\tau_1\mu(\alpha)e - \chi(\alpha))^+\|_F \leq (1-\alpha_2)\beta\tau_1\mu(\alpha).$$

Proof. Let $\widetilde{x}\circ\widetilde{s}=\lambda_1c_1+\cdots+\lambda_rc_r$, where $\{c_1,\cdots,c_r\}$ is a Jordan frame and the spectral eigenvalues satisfy

$$\tau_1\mu-\lambda_1\le\tau_1\mu-\lambda_2\le\cdots\le\tau_1\mu-\lambda_k\le 0\le\tau_1\mu-\lambda_{k+1}\le\cdots\le\tau_1\mu-\lambda_r.$$

Then, one has $(\tau_1\mu e-\widetilde{x}\circ\widetilde{s})^+=(\tau_1\mu-\lambda_{k+1})c_{k+1}+\cdots+(\tau_1\mu-\lambda_r)c_r$, and

$$\begin{aligned}\chi(\alpha)=&\sum_{i=1}^r\lambda_ic_i+\alpha_1\sum_{i=1}^k(\tau_1\mu-\lambda_i)c_i+\alpha_2\sum_{i=k+1}^r(\tau_1\mu-\lambda_i)c_i\\=&\sum_{i=1}^k((1-\alpha_1)\lambda_i+\alpha_1\tau_1\mu)c_i+\sum_{i=k+1}^r((1-\alpha_2)\lambda_i+\alpha_2\tau_1\mu)c_i,\end{aligned}\tag{35}$$

which implies $\lambda_i(\chi(\alpha))\ge 0,\ i=1,\cdots,r$. Using (35), we have

$$\begin{aligned}\|(\tau_1\mu(\alpha)e-\chi(\alpha))^+\|_F^2=&\sum\left([\tau_1\mu(\alpha)-\lambda_i(\chi(\alpha))]^+\right)^2\\\le&\sum\left(\left[\tau_1\mu(\alpha)-\frac{\mu(\alpha)}{\mu}\lambda_i(\chi(\alpha))\right]^+\right)^2\\=&(1-\alpha_2)^2\Big(\frac{\mu(\alpha)}{\mu}\Big)^2\sum_{i=k+1}^r\left([\tau_1\mu-\lambda_i]^+\right)^2,\end{aligned}$$

where the inequality holds due to $\mu(\alpha)\le\mu$ and $\chi(\alpha)\succeq 0$. Hence,

$$\|(\tau_1\mu(\alpha)e-\chi(\alpha))^+\|_F\le(1-\alpha_2)\frac{\mu(\alpha)}{\mu}\|(\tau_1\mu e-\widetilde{x}\circ\widetilde{s})^+\|_F\le(1-\alpha_2)\beta\tau_1\mu(\alpha).$$

□

Lemma 5.8. *Let $\tau_1\le 1/8,\beta\le 1/2$ and $(\widetilde{x},y,\widetilde{s})\in\mathcal{N}(\tau_1,\beta)$. If $\alpha_1=\alpha_2\sqrt{\beta\tau_1/r}$ and $\alpha_2\le\sqrt{\tau_1}/(\sqrt{\mathrm{cond}\,(G)}\omega^2r)$, then we have $\|\triangle\widetilde{x}(\alpha)\circ\triangle\widetilde{s}(\alpha)\|_F\le\alpha_2\beta\tau_1\mu(\alpha)$.*

Proof. Let

$$\begin{aligned}\eta_1&=\|\triangle\widetilde{x}_-\|_F\|\triangle\widetilde{s}_-\|_F,\\\eta_2&=\|\triangle\widetilde{x}_+\|_F\|\triangle\widetilde{s}_+\|_F,\\\eta_3&=\|\triangle\widetilde{x}_-\|_F\|\triangle\widetilde{s}_+\|_F+\|\triangle\widetilde{s}_-\|_F\|\triangle\widetilde{x}_+\|_F.\end{aligned}$$

Then, from Corollary 5.1,

$$\eta_1\le\frac{1}{2}\sqrt{\mathrm{cond}\,(G)}\omega^2r^2\mu,\qquad\eta_2\le\frac{1}{2}\sqrt{\mathrm{cond}\,(G)}\beta\tau_1\mu,\qquad\eta_3\le 2\sqrt{\mathrm{cond}\,(G)}\sqrt{\beta\tau_1}\omega r\mu.\tag{36}$$

By (25) and Lemma 2.2, we have

$$\begin{aligned}\|\triangle\widetilde{x}(\alpha)\circ\triangle\widetilde{s}(\alpha)\|_F\le&\alpha_1^2\|\triangle\widetilde{x}_-\circ\triangle\widetilde{s}_-\|_F+\alpha_1\alpha_2(\|\triangle\widetilde{x}_-\circ\triangle\widetilde{s}_+\|_F\\&+\|\triangle\widetilde{s}_-\circ\triangle\widetilde{x}_+\|_F)+\alpha_2^2\|\triangle\widetilde{x}_+\circ\triangle\widetilde{s}_+\|_F\\\le&\alpha_1^2\eta_1+\alpha_2^2\eta_2+\alpha_1\alpha_2\eta_3.\end{aligned}$$

On the other hand, by using $\alpha_1\le\alpha_2$, we have

$$\begin{aligned}\mathrm{tr}\,(\chi(\alpha))=&\mathrm{tr}\,(\widetilde{x}\circ\widetilde{s})+\alpha_1\mathrm{tr}\,(\tau_1\mu e-\widetilde{x}\circ\widetilde{s})+(\alpha_2-\alpha_1)\mathrm{tr}\,((\tau_1\mu e-\widetilde{x}\circ\widetilde{s})^+)\\\ge&r\mu-\alpha_1(1-\tau_1)r\mu\\\ge&r\mu-\alpha_1r\mu.\end{aligned}\tag{37}$$

Then, using Cauchy-Schwarz inequality, we have

$$\begin{aligned}\langle\widetilde{x}(\alpha),\widetilde{s}(\alpha)\rangle =&\mathrm{tr}\,(\chi(\alpha))+\alpha_1^2\langle\triangle\widetilde{x}_-,\triangle\widetilde{s}_-\rangle+\alpha_1\alpha_2(\langle\triangle\widetilde{x}_-,\triangle\widetilde{s}_+\rangle+\langle\triangle\widetilde{s}_-,\triangle\widetilde{x}_+\rangle)\\ \geq& r\mu-\alpha_1 r\mu-\alpha_1^2\eta_1-\alpha_1\alpha_2\eta_3.\end{aligned}$$

A straightforward calculation shows that

$$\begin{aligned}&\|\triangle\widetilde{x}(\alpha)\circ\triangle\widetilde{s}(\alpha)\|_F-\alpha_2\beta\tau_1\mu(\alpha)\\ \leq&\alpha_1^2\eta_1+\alpha_2^2\eta_2+\alpha_1\alpha_2\eta_3-\alpha_2\beta\tau_1(r\mu-\alpha_1 r\mu-\alpha_1^2\eta_1-\alpha_1\alpha_2\eta_3)/r\\ =&\alpha_2^2\eta_2+(1+\alpha_2\beta\tau_1/r)(\alpha_2^2\beta\tau_1\eta_1/r+\alpha_2^2\sqrt{\beta\tau_1/r}\eta_3)-\alpha_2\beta\tau_1(1-\alpha_2\sqrt{\beta\tau_1/r})\mu\\ \leq&\alpha_2\big[\alpha_2\eta_2+(1+\alpha_2\beta\tau_1)(\alpha_2\beta\tau_1\eta_1/r+\alpha_2\sqrt{\beta\tau_1/r}\eta_3)-\beta\tau_1(1-\alpha_2\sqrt{\beta\tau_1})\mu\big]\\ \leq&\alpha_2\beta\tau_1\mu\Big[\frac{\sqrt{\tau_1}}{2\omega^2 r}+\Big(1+\frac{\beta\tau_1^{3/2}}{\omega^2 r}\Big)\Big(\frac{\sqrt{\tau_1}}{2}+\frac{2\sqrt{\tau_1}}{\omega\sqrt{r}}\Big)-\Big(1-\frac{\sqrt{\beta}\tau_1}{\omega^2 r}\Big)\Big]\\ \leq&0.\end{aligned}$$

Here, the third inequality follows from (36) and $\alpha_2\leq\sqrt{\tau_1}/(\sqrt{\mathrm{cond}\,(G)}\omega^2 r)$, the last inequality holds due to the facts $\tau_1\leq 1/8,\beta\leq 1/2,r\geq 1$, and $\omega\geq 13$ by (33). □

The following lemma gives a sufficient condition which guarantees all the iterates in the neighbourhood $\mathcal{N}(\tau_1,\beta)$.

Lemma 5.9. *Let $\tau_1\leq 1/8,\beta\leq 1/2$ and $(\widetilde{x},y,\widetilde{s})\in\mathcal{N}(\tau_1,\beta)$. If $\alpha_1=\alpha_2\sqrt{\beta\tau_1/r}$ and $\alpha_2\leq\sqrt{\tau_1}/(\sqrt{\mathrm{cond}\,(G)}\omega^2 r)$, then $(\widetilde{x}(\alpha),y(\alpha),\widetilde{s}(\alpha))\in\mathcal{N}(\tau_1,\beta)$.*

Proof. By Lemma 5.6, it holds that $\mu(\alpha)\leq\mu$. Furthermore, by using Lemma 5.7 and Lemma 5.8, we have

$$\begin{aligned}&\|(\tau_1\mu(\alpha)e-\widetilde{x}(\alpha)\circ\widetilde{s}(\alpha))^+\|_F\\ =&\|(\tau_1\mu(\alpha)e-\chi(\alpha)-\triangle\widetilde{x}(\alpha)\circ\triangle\widetilde{s}(\alpha))^+\|_F\\ \leq&\|(\tau_1\mu(\alpha)e-\chi(\alpha))^+\|_F+\|(-\triangle\widetilde{x}(\alpha)\circ\triangle\widetilde{s}(\alpha))^+\|_F\\ \leq&\|(\tau_1\mu(\alpha)e-\chi(\alpha))^+\|_F+\|\triangle\widetilde{x}(\alpha)\circ\triangle\widetilde{s}(\alpha)\|_F\\ \leq&(1-\alpha_2)\beta\tau_1\mu(\alpha)+\alpha_2\beta\tau_1\mu(\alpha)\\ =&\beta\tau_1\mu(\alpha),\end{aligned}$$

where, the first inequality follows from Lemma 2.4. Then, one has $\lambda(\widetilde{x}(\alpha)\circ\widetilde{s}(\alpha))\geq(1-\beta)\tau_1\mu(\alpha)>0$. Thus, by Lemma 2.1 we have $\det(\widetilde{x}(\alpha))\neq 0$ and $\det(\widetilde{s}(\alpha))\neq 0$. Then, since $\widetilde{x},\widetilde{s}\succ 0$, by continuity it follows that $\widetilde{x}(\alpha)\succ 0$ and $\widetilde{s}(\alpha)\succ 0$. Moreover, by [10, Theorem 3.1], we have

$$\|(\tau_1\mu(\alpha)e-\widetilde{w}(\alpha))^+\|_F\leq\|(\tau_1\mu(\alpha)e-\widetilde{x}(\alpha)\circ\widetilde{s}(\alpha))^+\|_F\leq\beta\tau_1\mu(\alpha).$$

where $\widetilde{w}(\alpha)=Q_{\widetilde{x}(\alpha)^{1/2}}\widetilde{s}(\alpha)$.

Consequently, we have $(\widetilde{x}(\alpha),y(\alpha),\widetilde{s}(\alpha))\in\mathcal{N}(\tau_1,\beta)$. □

Lemma 5.10. *Let $(\widetilde{x},y,\widetilde{s})\in\mathcal{N}(\tau_1,\beta)$ and $\beta\leq 1/2$. If*

$$\alpha_1=\alpha_2\sqrt{\beta\tau_1/r},\qquad\alpha_2\leq\sqrt{\tau_1}/(\sqrt{\mathrm{cond}\,(G)}\omega^2 r),$$

then

$$\langle\widetilde{x}(\alpha),\widetilde{s}(\alpha)\rangle\geq(1-\alpha_1)\nu\langle\widetilde{x}^0,\widetilde{s}^0\rangle.$$

Proof. Firstly, by (37), we have $\operatorname{tr}(\chi(\alpha)) \geq r\mu - \alpha_1(1-\tau_1)r\mu$. Thus, using Cauchy-Schwarz inequality,

$$\begin{aligned}\langle \widetilde{x}(\alpha), \widetilde{x}(\alpha)\rangle \geq & r\mu - \alpha_1(1-\tau_1)r\mu - \alpha_1^2\eta_1 - \alpha_1\alpha_2\eta_3 \\ = & (1-\alpha_1)r\mu + \alpha_1\tau_1 r\mu - \alpha_1^2\eta_1 - \alpha_1\alpha_2\eta_3 \\ \geq & (1-\alpha_1)\nu\langle \widetilde{x}^0, \widetilde{s}^0\rangle + \alpha_1(\tau_1 r\mu - \alpha_1\eta_1 - \alpha_2\eta_3),\end{aligned}$$

where in the last inequality we used (17).

Therefore, to complete the proof, it is sufficient to show that

$$\begin{aligned}\tau_1 r\mu - \alpha_1\eta_1 - \alpha_2\eta_3 \geq & \tau_1 r\mu - \alpha_1\sqrt{\operatorname{cond}(G)}\omega^2 r^2\mu/2 - 2\alpha_2\sqrt{\operatorname{cond}(G)}\sqrt{\beta\tau_1}\omega r\mu \\ = & \tau_1 r\mu - \alpha_2\sqrt{\operatorname{cond}(G)}\sqrt{\beta\tau_1}\omega^2 r^2\mu[1/(2\sqrt{r}) + 2/(\omega r)] \\ \geq & \tau_1 r\mu - \sqrt{\beta}\tau_1 r\mu[1/2 + 2/\omega] \geq 0.\end{aligned}$$

Here, the last inequality inequality holds due to $\beta \leq 1/2$ and $\omega \geq 13$. □

For the Condition (24), we let $\tau_1 + \sqrt{\tau_1} \leq \delta < 1$.

Lemma 5.11. *Let $\tau_1 \leq 1/8, \beta \leq 1/2$ and $(\widetilde{x}, y, \widetilde{s}) \in \mathcal{N}(\tau_1, \beta)$. If $\alpha_1 = \alpha_2\sqrt{\beta\tau_1/r}$ and $\alpha_2 \leq \sqrt{\tau_1}/(\sqrt{\operatorname{cond}(G)}\omega^2 r)$, then we have $\langle \widetilde{x}(\alpha), \widetilde{s}(\alpha)\rangle \leq (1-(1-\delta)\alpha_1)\langle \widetilde{x}, \widetilde{s}\rangle$.*

Proof. From (34), it holds that

$$\begin{aligned}\langle \widetilde{x}(\alpha), \widetilde{s}(\alpha)\rangle \leq & \big[1 - \alpha_2\sqrt{\beta\tau_1/r}\big(1 - \tau_1 - \sqrt{\beta\tau_1} - \sqrt{\beta}\tau_1/2 - 2\sqrt{\beta}\tau_1/\omega\big)\big] r\mu \\ = & \big[1 - \alpha_1\big(1 - \tau_1 - \sqrt{\tau_1}(\sqrt{\beta} + \sqrt{\beta\tau_1}/2 + 2\sqrt{\beta\tau_1}/\omega)\big)\big]\langle \widetilde{x}, \widetilde{s}\rangle \\ \leq & \big[1 - \alpha_1(1 - \tau_1 - \sqrt{\tau_1})\big]\langle \widetilde{x}, \widetilde{s}\rangle \\ \leq & \big[1 - \alpha_1(1-\delta)\big]\langle \widetilde{x}, \widetilde{s}\rangle,\end{aligned}$$

where, the second inequality follows from the fact $\tau_1 \leq 1/8, \beta \leq 1/2$. □

In view of Lemma 5.9–5.11, we may find step size in the following way. First, set $\alpha_2 = \sqrt{\tau_1}/(\sqrt{\operatorname{cond}(G)}\omega^2 r)$. Second, find the greatest $\alpha_1 \in [0, 1]$ such that conditions (22)–(24) hold. Lemma 5.9–5.11 guarantee that $\alpha_1 \geq \alpha_2\sqrt{\beta\tau_1/r}$.

5.2 Polynomial Complexity

The following theorem gives an upper bound for the number of iterations in which Algorithm 4.1 stops with an ε-approximate solution.

Theorem 5.1. *Suppose that $\sqrt{\operatorname{cond}(G)} \leq \kappa < \infty$ for all iterations. Then Algorithm 4.1 terminates in at most $\mathcal{O}(\kappa r^{1.5}\log\varepsilon^{-1})$ iterations.*

Proof. At each iteration, if we let $\alpha_2 = \sqrt{\tau_1}/(\sqrt{\operatorname{cond}(G)}\omega^2 r)$ and $\alpha_1 = \alpha_2\sqrt{\beta\tau_1/r}$, then by using (34) we have

$$\mu(\alpha) = \langle x(\alpha), s(\alpha)\rangle/r = \langle \widetilde{x}(\alpha), \widetilde{s}(\alpha)\rangle/r \leq \Big[1 - \frac{\alpha_2\sqrt{\beta\tau_1}}{2\sqrt{r}}\Big]\mu = \Big[1 - \frac{\alpha_1}{2}\Big]\mu.$$

Therefore,

$$\mu(\alpha) \leq \Big[1 - \frac{\alpha_2\sqrt{\beta\tau_1/r}}{2}\Big]\mu = \Big[1 - \frac{\sqrt{\beta}\tau_1}{2\sqrt{\operatorname{cond}(G)}\omega^2 r^{3/2}}\Big]\mu \leq \Big[1 - \frac{\sqrt{\beta}\tau_1}{2\kappa\omega^2 r^{3/2}}\Big]\mu,$$

from which the statement of the theorem follows. □

By (7) and [18, Lemma 36], we have cond $(G) = 1$ for the NT direction, and cond $(G) \le r/\tau_2$ for the xs and sx directions.

Corollary 5.2. *If the NT search direction is used, the iteration complexity of Algorithm 4.1 is $\mathcal{O}(r^{1.5}\log\varepsilon^{-1})$. If the xs and sx search directions are used, the iteration complexities of Algorithm 4.1 are $\mathcal{O}(r^2\log\varepsilon^{-1})$.*

5.3 Complexity for Feasible Staring Points

In this subsection, we demonstrate that if strictly feasible starting points are used, then the complexity bounds for Algorithm 4.1 can be lowered. Since the proof techniques are exactly the same as those used in the infeasible starting point case, we will only give a brief outline of the proof, omitting the details.

We start with the observation that for feasible interior-point method, we always have $\widetilde{A}\widetilde{x} = b$ and $\widetilde{A}^* y + \widetilde{s} = \widetilde{c}$, which imply that $\langle \triangle\widetilde{x}_-, \triangle\widetilde{s}_-\rangle = 0, \langle \triangle\widetilde{x}_+, \triangle\widetilde{s}_+\rangle = 0$. Therefore, one has $\langle \triangle\widetilde{x}(\alpha), \triangle\widetilde{s}(\alpha)\rangle = 0$, and

$$\langle \widetilde{x}(\alpha), \widetilde{s}(\alpha)\rangle = \mathrm{tr}\,(\widetilde{x}\circ\widetilde{s}) + \alpha_1 \mathrm{tr}\,((\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^-) + \alpha_2 \mathrm{tr}\,((\tau_1\mu e - \widetilde{x}\circ\widetilde{s})^+). \tag{38}$$

As a key result, one has $\xi = 0$ and $\|(\triangle\widetilde{x}_-, \triangle\widetilde{s}_-)\|_G \le \sqrt{r\mu}$ in Lemma 5.3. Therefore, in place of Lemma 5.8, we have the following lemma.

Lemma 5.12. *Let $\tau_1 \le 1/8, \beta \le 1/2$ and feasible point $(\widetilde{x}, y, \widetilde{s}) \in \mathcal{N}(\tau_1, \beta)$. If $\alpha_1 = 0.2\alpha_2\sqrt{\beta\tau_1/r}$ and $\alpha_2 \le 1/\sqrt{\mathrm{cond}\,(G)}$, then we have $\|\triangle\widetilde{x}(\alpha)\circ\triangle\widetilde{s}(\alpha)\|_F \le \alpha_2\beta\tau_1\mu(\alpha)$.*

Under the condition of Lemma 5.12, by using (27), (28) and (38) one has

$$\begin{aligned}\langle \widetilde{x}(\alpha), \widetilde{s}(\alpha)\rangle \le & r\mu - \alpha_1(1-\tau_1)r\mu + \alpha_2\sqrt{r}\beta\tau_1\mu \\ = & \big(1 - 0.2\alpha_2(1-\tau_1)\sqrt{\beta\tau_1/r} + \alpha_2\beta\tau_1/\sqrt{r}\big)r\mu \\ \le & \Big(1 - \frac{\alpha_2 3\sqrt{\beta\tau_1}}{40\sqrt{r}}\Big)r\mu.\end{aligned}$$

As the proof of Lemma 5.9, we have $(\widetilde{x}(\alpha), y(\alpha), \widetilde{s}(\alpha)) \in \mathcal{N}(\tau_1, \beta)$. Therefore, we have the following iteration complexity bound for the feasible interior-point algorithm.

Theorem 5.2. *Let a feasible point $(x^0, y^0, s^0) \in \mathcal{N}(\tau_1, \beta)$, and suppose $\sqrt{\mathrm{cond}\,(G)} \le \kappa < \infty$ for all iterations. Then the feasible algorithm terminates in at most $\mathcal{O}(\kappa\sqrt{r}\log\varepsilon^{-1})$ iterations. Hence, for NT direction the algorithm takes $\mathcal{O}(\sqrt{r}\log\varepsilon^{-1})$ iterations, and for xs and sx directions the algorithm takes $\mathcal{O}(r\log\varepsilon^{-1})$ iterations.*

We note that, when the NT search direction is used, the feasible interior-point algorithm achieves its best complexity bound which coincides with the best known complexity of interior-point methods.

6 Conclusions

We have established complexity bound of an infeasible-interior-point algorithm, based on a new wide neighbourhood, for linear programming over symmetric cones. We summarize the obtained complexity results in Table 1, where r is the rank of the associated Euclidean Jordan algebra and $\varepsilon > 0$ is the required precision. For comparison, we also include the complexity bounds for the IIPM in [17].

Table 1. Summary of Complexity Bounds

	Infeasible initial point		Feasible initial point	
	xs/sx	NT	xs/sx	NT
New-IIPM	$\mathcal{O}(r^2 \log \varepsilon^{-1})$	$\mathcal{O}(r^{1.5} \log \varepsilon^{-1})$	$\mathcal{O}(r \log \varepsilon^{-1})$	$\mathcal{O}(\sqrt{r} \log \varepsilon^{-1})$
IIPM	$\mathcal{O}(r^{2.5} \log \varepsilon^{-1})$	$\mathcal{O}(r^2 \log \varepsilon^{-1})$	$\mathcal{O}(r^{1.5} \log \varepsilon^{-1})$	$\mathcal{O}(r \log \varepsilon^{-1})$

Acknowledgements. The authors are grateful to the anonymous referees and editors for their constructive and valuable suggestions improving this paper.

References

[1] Ai, W. Neighborhood-following algorithms for linear programming. *Sci. China Ser. A.*, 47: 812–820 (2004)
[2] Ai, W., Zhang, S. An $O(\sqrt{n}L)$ iteration primal-dual path-following method, based on wide neighborhoods and large updates, for monotone LCP. *SIAM J. Optim.*, 16: 400–417 (2005)
[3] Alizadeh, F., Xia, Y. The Q method for symmetric cone programming. *J. Optim. Theory Appl.*, 149: 102–137 (2011)
[4] Faraut, J., Korányi, A. Analysis on Symmetric Cones. Oxford University Press, New York, 1994
[5] Faybusovich, L. Euclidean Jordan algebras and interior-point algorithms. *Positivity*, 1: 331–357 (1997)
[6] Faybusovich, L. Linear systems in Jordan algebras and primal-dual interior-point algorithms. *J. Compu. Appl. Math.*, 86: 149–175 (1997)
[7] Gu, G., Zangiabadi, M., Roos, C. Full Nesterov-Todd step infeasible interior-point method for symmetric optimization. *Eur. J. Oper. Res.*, 214: 473–484 (2011)
[8] Güler, O. Barrier functions in interior-point methods. *Math. Oper. Res.*, 21: 860–885 (1996)
[9] Li, Y., Terlaky, T. A new class of large neighborhood path-following interior point algorithms for semidefinite optimization with $O(\sqrt{n}\log\frac{\mathrm{Tr}(X^0S^0)}{\varepsilon})$ iteration complexity. *SIAM J. Optim.*, 20: 2853–2875 (2010)
[10] Liu, C., Liu, H., Shang, Y. Neighborhood-following algorithms for symmetric cone programming. *Sci. Sin. Math.*, 43: 691–702 (2013) (in Chinese)
[11] Liu, H., Yang, X., Liu, C. A new wide neighborhood primal-dual infeasible-interior-point method for symmetric cone programming. *J. Optim. Theory Appl.*, 158(3): 796–815 (2013)
[12] Monteiro, R.D.C., Zhang, Y. A unified analysis for a class of long-step primal-dual path-following interior-point algorithms for semidefinie programming. *Math. Program.*, 81: 281–299 (1998)
[13] Nesterov, Yu., Nemirovski, A. Interior Point Polynomial Algorithms in Convex Programming. SIAM, Philadelphia, 1994
[14] Nesterov, Yu., Todd, M.J. Self-scaled barriers and interior-point methods for convex programming. *Math. Oper. Res.*, 22: 1–42 (1997)
[15] Nesterov, Yu., Todd, M.J. Primal-dual interior-point methods for self-scaled cones. *SIAM J. Optim.*, 8: 324–364 (1998)
[16] Potra, F.A. An infeasible interior point method for linear complementarity problems over symmetric cones. *AIP Conf. Proc.*, 1168: 1403–1406 (2009)
[17] Rangarajan, B.K. Polynomial convergence of infeasible-interior-point methods over symmetric cones. *SIAM J. Optim.*, 16: 1211–1229 (2006)
[18] Schmieta, S.H., Alizadeh, F. Extension of primal-dual interior-point algorithm to symmetric cones. *Math. Program.*, 96: 409–438 (2003)
[19] Wang, G.Q., Bai, Y.Q. A new full Nesterov-Todd step primal-dual path-following interior-point algorithm for symmetric optimization. *J. Optim. Theory Appl.*, 154: 966–985 (2012)
[20] Zhang, J., Zhang, K. Polynomial complexity of an interior point algorithm with a second order corrector step for symmetric cone programming. *Math. Methods Oper., Res.*, 73: 75–90 (2011)