Math. Program., Ser. A 88: 61–83 (2000)

Digital Object Identifier (DOI) 10.1007/s101070000137

Renato D.C. Monteiro · Takashi Tsuchiya

Polynomial convergence of primal-dual algorithms for the second-order cone program based on the MZ-family of directions

Received: June 5, 1998 / Accepted: September 8, 1999
Published online April 20, 2000 – © Springer-Verlag 2000

Abstract. In this paper we study primal-dual path-following algorithms for the second-order cone programming (SOCP) based on a family of directions that is a natural extension of the Monteiro-Zhang (MZ) family for semidefinite programming. We show that the polynomial iteration-complexity bounds of two well-known algorithms for linear programming, namely the short-step path-following algorithm of Kojima et al. and Monteiro and Adler, and the predictor-corrector algorithm of Mizuno et al., carry over to the context of SOCP, that is they have an $O(\sqrt{n}\,\log \varepsilon^{-1})$ iteration-complexity to reduce the duality gap by a factor of ε, where n is the number of second-order cones. Since the MZ-type family studied in this paper includes an analogue of the Alizadeh, Haeberly and Overton pure Newton direction, we establish for the first time the polynomial convergence of primal-dual algorithms for SOCP based on this search direction.

Key words. second-order cone programming – ice-cream cone – interior-point methods – polynomial complexity – path-following methods – primal-dual methods – Newton method

1. Introduction

The second-order cone programming (SOCP) problem is to minimize or maximize a linear function over the intersection of an affine space with the Cartesian product of a finite number of second-order cones. Recently, this problem has received considerable attention for its wide range of applications (see [10, 14, 31]) and for being "easily" solvable via interior-point algorithms (see [23–26]). In this paper, we study primal-dual path-following algorithms for the SOCP based on a family of search directions which is a natural extension of the Monteiro-Zhang family of directions introduced in the context of the semidefinite programming (SDP) (see [17], [32] and [22]). We establish polynomial convergence of two path-following algorithms that are natural extensions of standard linear programming (LP) algorithms, namely the short-step method of Kojima et al. [12] and Monteiro and Adler [19, 20] and the predictor-corrector method of Mizuno et al. [16].

R.D.C. Monteiro: School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, USA, e-mail: monteiro@isye.gatech.edu

The work of this author was based on research supported by the National Science Foundation under grants INT-9600343, CCR-9700448 and CCR-9902010.

T. Tsuchiya: The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-Ku, Tokyo, 106-8569, Japan, e-mail: tsuchiya@sun312.ism.ac.jp

This research is supported in part by the Grant-in-Aid for Encouragement of Young Scientists (A) 09780417 (1997) of the Ministry of Education, Science and Culture of Japan.

Mathematics Subject Classification (1991): 65K05, 90C25, 90C30

While primal-only interior-point algorithms for solving SOCP, and the important special case of quadratically constrained convex quadratic programming, were developed about eight years ago (see [9,15,23,24]), it was only recently that primal-dual algorithms for SOCP have been developed. The first polynomial primal-dual path-following algorithm for SOCP was proposed by Nesterov and Todd [25,26]. They develop a general approach for solving the homogeneous and self-dual cone programming which includes SOCP, LP and SDP as special cases. In their work, a direction called the Nesterov-Todd (NT) direction is proposed and the short-step path-following algorithm based on this direction is shown to have $O(\sqrt{n}\ \log \varepsilon^{-1})$ iteration-complexity, where n is the number of second-order cones.

Adler and Alizadeh [1] study a unified primal-dual approach for SDP and SOCP, and propose a direction for SOCP analogous to the Alizadeh-Haeberly-Overton (AHO) direction introduced in [2] for SDP. Recently, Alizadeh and Schmieta studied several theoretical and practical aspects of SOCP mainly from the viewpoint of nondegeneracy [3].

Faybusovich studies the homogeneous self-dual cone programming from the viewpoint of Euclidean Jordan algebra [7], which provides another framework for handling homogeneous self-dual cones [4–6]. He extends the AHO direction to homogeneous self-dual cone programming, gives conditions for it to be well-defined, and studies some nondegeneracy issues in the context of this problem.

In the recent papers [29] and [30], Tsuchiya extends standard path-following algorithms for LP [11,12,19] and SDP [13,17,22,21] to SOCP. He introduces a primal-dual product in the context of SOCP which plays exactly the same role as Xs and $S^{1/2}XS^{1/2}$ do in the context of LP and SDP, respectively, and defines neighborhoods of the central path in terms of the eigenvalue decomposition of this product. Path-following algorithms using two major scaling invariant directions, namely the Helmberg-Rendl-Vanderbei-Wolkowicz/Kojima-Shindoh-Hara/Monteiro (HRVW/KSH/M) [8,13,17] direction and the NT direction, are analyzed in his works [29] and [30], where it is shown that: (i) the short-step, semilong-step and long-step path-following algorithms using the HRVW/KSH/M direction have $O(\sqrt{n}\ \log \varepsilon^{-1})$, $O(n\ \log \varepsilon^{-1})$ and $O(n^{3/2}\ \log \varepsilon^{-1})$ iteration-complexities, respectively; (ii) the short-step path-following algorithm using the NT direction has an $O(\sqrt{n}\ \log \varepsilon^{-1})$ iteration-complexity, and the semilong-step and long-step algorithms have both $O(n\ \log \varepsilon^{-1})$ iteration-complexity.

In this paper, we show that the short-step path-following and the Mizuno-Todd-Ye predictor-corrector algorithms based on an extension of the Monteiro and Zhang family of search directions have both $O(\sqrt{n}\ \log \varepsilon^{-1})$ iteration-complexity bound. Briefly speaking, a direction of this family is equivalent to the AHO direction in a certain scaled space determined by a linear transformation which leaves the cone invariant. Since the central-path neighborhoods, and the distances used to define them, remain invariant under these linear transformations, an iteration in the original space can be analyzed from the viewpoint of the scaled space. A main advantage of this viewpoint is that the analysis of an iteration along the AHO direction suffices to describe the behavior of an iteration of the algorithm along any direction of the family. The relevance of the MZ-family is that it contains the three major standard directions, namely the AHO direction, the HRVW/KSH/M direction and the NT direction, and hence it provides a unifying framework for the polynomial convergence analysis of path-following algorithms. Polynomiality of the short-step path-following algorithm for SDP based on the MZ-family was first established by Monteiro [18].

A major difficulty in analyzing path-following algorithms using the AHO direction is that it is necessary to estimate a first-order error term (i.e., depending linearly on the step-size) that appears in the centrality measure for the new iterate. Careful analysis is needed to show that this term can be bounded by a quantity that depends quadratically on the opening of the cone and the deviation of the centrality parameter from one. By choosing the opening of the cone small enough and the centrality parameter not too far from one, it is possible to show that the error coming from this term is small enough to ensure that the algorithm have an $O(\sqrt{n}\,\log \varepsilon^{-1})$ iteration-complexity. The same kind of difficulty occurs in the context of SDP (see [18]). However, the technique used to bound the effect of the first order term in the context of SOCP is quite different from the one in SDP, and is a major new development of this paper. As in the case of SDP, the AHO direction is shown to be well-defined only for points close to the central path, more specifically in the ∞-norm neighborhood with opening less than $1/3$.

Finally, after the release of the first version of this paper, Schmieta and Alizadeh [27] have independently shown that the results of this paper can be extended to most symmetric cones by using a completely different approach than ours.

This paper is organized as follows. In Sect. 2, we describe the SOCP problem, define the central path and its neighborhoods, and introduce the Newton system that determines the AHO search direction. In Sect. 3, we develop several technical results that are used in the analysis of Sect. 4, where the short-step path-following algorithm and the predictor-corrector algorithm using the AHO direction are described and their polynomial convergence are established. In Sect. 5, we introduce the MZ family of directions for SOCP and show that the convergence results obtained for the AHO direction in Sect. 4 also holds for the MZ-family.

1.1. Notation and terminology

The following notation is used throughout the paper. The superscript ^{T} denotes transpose. $\Re^p$ denotes the p-dimensional Euclidean space. The set of all $p \times q$ matrices with real entries is denoted by $\Re^{p\times q}$. If P and Q are square symmetric matrices, we write $P \succeq Q$, or $Q \preceq P$, to indicate that $P - Q$ is positive semidefinite. For a square matrix Q with all real eigenvalues, we denote its smallest and largest eigenvalues by $\lambda_{\min}[Q]$ and $\lambda_{\max}[Q]$, respectively. Given a finite number of square matrices $Q_1, \ldots, Q_n$, we denote the block diagonal matrix with these matrices as block diagonals by diag $(Q_1, \ldots, Q_n)$, or by diag $(Q_i : i = 1, \ldots, n)$. The Euclidean norm and its associated operator norm are both denoted by $\|\cdot\|$; hence, $\|Q\| \equiv \max_{\|u\|=1} \|Qu\|$ for any $Q \in \Re^{p\times p}$. We denote the interior of a set $\Omega \subset \Re^p$ by $\operatorname{int}\Omega$.

2. The second-order cone program and preliminary discussion

This section describes the SOCP problem, the central path and the neighborhoods of the central path that will be used in our presentation. It also introduces the Newton direction, referred to as the AHO direction, for the system which characterizes the central path in terms of a certain Jordan algebra product between the primal variable and the dual slack variable.

2.1. The second-order cone program

In this paper we consider the following second-order program

$$\text{(P)} \qquad \min \left\{ \sum_{i=1}^{n} c_i^T x_i \ : \ \sum_{i=1}^{n} A_i x_i = b, \ x_i \in \mathcal{K}_i, \ \ i = 1, \ldots, n \right\}, \tag{1}$$

where $x_i \in \Re^{k_i}$, $i = 1, \ldots, n$, are the variables, $b \in \Re^m$, $A_i \in \Re^{m \times k_i}$ and $c_i \in \Re^{k_i}$, $i = 1, \ldots, n$, are the data, and the set $\mathcal{K}_i$, $i = 1, \ldots, n$, is the second-order cone of dimension k_i defined as

$$\mathcal{K}_i = \left\{ x_i = (x_{i0}, x_{i1}) \in \Re \times \Re^{k_i - 1} : \ x_{i0} - \|x_{i1}\| \geq 0 \right\}.$$

It is well-known that the cone $\mathcal{K}_i$ is self-dual, that is $\mathcal{K}_i = \mathcal{K}_i^* \equiv \{s_i \in \Re^{k_i} : s_i^T x_i \geq 0, \forall x_i \in K_i\}$. The dual of problem (1) is

$$\text{(D)} \qquad \max \left\{ b^T y \ : \ A_i^T y + s_i = c_i, \ s_i \in \mathcal{K}_i, \ i = 1, \ldots, n \right\}. \tag{2}$$

Defining

$$\begin{aligned}
&K \equiv k_1 + \ldots + k_n, \quad \mathcal{K} = \mathcal{K}_1 \times \ldots \times \mathcal{K}_n, \\
&A = (A_1 \ A_2 \ \ldots \ A_n) \in \Re^{m \times K}, \ \ c = (c_1, \cdots, c_n) \in \Re^K, \\
&x = (x_1, \ldots, x_n) \in \Re^K, \quad s = (s_1, \ldots, s_n) \in \Re^K,
\end{aligned}$$

problems (P) and (D) can be simply written as

$$\begin{aligned}
&\text{(P)} \quad \min \left\{ c^T x : Ax = b, \ x \in \mathcal{K} \right\}, \\
&\text{(D)} \quad \max \left\{ b^T y : A^T y + s = c, \ s \in \mathcal{K} \right\}.
\end{aligned}$$

The set of *interior feasible solutions* of (1) and (2) are:

$$\begin{aligned}
F^0(\text{P}) &\equiv \{x : Ax = b, \ x \in \mathcal{K}^0\}, \\
F^0(\text{D}) &\equiv \{(s, y) : A^T y + s = c, \ s \in \mathcal{K}^0\},
\end{aligned}$$

respectively, where $\mathcal{K}^0$ denotes the interior of the cone $\mathcal{K}$.

Throughout this paper, we make the following assumptions:

A1) $F^0(P) \times F^0(D) \neq \emptyset$;
A2) the rows of the matrix $A = (A_1 \ \ldots \ A_n)$ are linearly independent.

Under assumption A1, it is well-known that (P) and (D) have optimal solutions and their optimal values coincide. Moreover, solving (P) and (D) is equivalent to finding $(x, s, y) \in \mathcal{K} \times \mathcal{K} \times \Re^m$ such that

$$\text{(PD)} \quad x^T s = 0, \ Ax = b, \ A^T y + s = c.$$

2.2. *Euclidean Jordan algebra, central path and Newton direction*

The primal-dual algorithms studied in this paper are based on the Euclidean Jordan algebra associated with the second-order cone (see for example [4–7]). The Euclidean Jordan algebra for the second-order cone $\mathcal{K}_i$ is the algebra defined by the following bilinear form from $\Re^{k_i} \times \Re^{k_i}$ to $\Re^{k_i}$:

$$x_i \circ s_i = \left(x_i^T s_i\ ,\ x_{i0}s_{i1} + s_{i0}x_{i1}\right), \quad \forall x_i, s_i \in \Re^{k_i}.$$

The element $e_i = (1, 0, \ldots, 0)$ is the unit element of this algebra. The Jordan algebra associated with the cone $\mathcal{K} = \mathcal{K}_1 \times \ldots \times \mathcal{K}_n$ is given by

$$x \circ s = (x_1 \circ s_1, \ldots, x_n \circ s_n), \quad \forall x, s \in \Re^K,$$

with $e \equiv (e_1, \ldots, e_n)$ being its unit element. From now on, the space $\Re^K$ will always be assumed to be endowed with the above Jordan algebra. Given an element $x \in \Re^K$, we denote by $\mathbf{mat}(x)$ the matrix diag $(X_1, \ldots, X_n)$ with

$$X_i = \begin{pmatrix} x_{i0} & x_{i1}^T \\ x_{i1} & x_{i0} I \end{pmatrix}, \quad i = 1, \ldots, n. \tag{3}$$

It is easy to verify that

$$x \circ s = \mathbf{mat}(x)s = \mathbf{mat}(s)x.$$

It is known that $\mathbf{mat}(x)$ is a symmetric matrix with smallest and largest eigenvalues given by

$$\begin{aligned} \lambda_{\min}(\mathbf{mat}(x)) &= \min\{x_{i0} - \|x_{i1}\| : i = 1, \ldots, n\}, \\ \lambda_{\max}(\mathbf{mat}(x)) &= \max\{x_{i0} + \|x_{i1}\| : i = 1, \ldots, n\}. \end{aligned} \tag{4}$$

Verifying this fact is an easy exercise of linear algebra, and a proof can be found in Lemma 2.13 of [30]. Note that $\mathbf{mat}(x)$ is symmetric positive semidefinite (positive definite) if and only if $x \in \mathcal{K}$ ($x \in \mathcal{K}^0$).

The central path for (P) and (D) is defined as the set of solutions $(x, s, y) \in \mathcal{K} \times \mathcal{K} \times \Re^m$ to the system (see for example [1,5]):

$$x \circ s = \nu e, \quad Ax = b, \quad A^T y + s = c, \tag{5}$$

for all $\nu > 0$. Under assumptions A1 and A2, it can be shown that: i) system (5) has exactly one solution $(x, s, y) = (x_\nu, s_\nu, y_\nu)$ in $\mathcal{K} \times \mathcal{K} \times \Re^m$, which in fact lies in $\mathcal{K}^0 \times \mathcal{K}^0 \times \Re^m$; ii) (x_ν, s_ν, y_ν) depends continuously (and, analytically) on the parameter $\nu > 0$; iii) any accumulation point of the path (x_ν, s_ν, y_ν), as ν tends to zero, is a solution of the optimality condition (PD).

The algorithms studied in Sects. 3 and 4 use as search direction the Newton direction for system (5) for some appropriate value of $\nu > 0$. More specifically, the search direction, which depends on a given centrality parameter $\sigma \in \Re$, is a solution $(\Delta x, \Delta s, \Delta y) \in \Re^K \times \Re^K \times \Re^m$ of the linear system of equations

$$S\Delta x + X\Delta s = \sigma\mu e - Xs, \tag{6a}$$
$$A^T\Delta y + \Delta s = c - s - A^T y, \tag{6b}$$
$$A\Delta x = b - Ax, \tag{6c}$$

where $X \equiv \mathbf{mat}(x)$, $S \equiv \mathbf{mat}(s)$, $\mu = \mu(x, s) \equiv x^T s/n$. Results about the well-definedness of this direction, which we refer to as the AHO direction, will be given in Sect. 3 where it is shown that $(\Delta x, \Delta s, \Delta y)$ exists and is unique for points (x, s, y) lying close to the central path. Note that (6) is exactly the Newton system at the point (x, s, y) with respect to (5) when $\nu = \sigma\mu$.

2.3. *Scaling, eigenvalues and the neighborhoods of the central path*

We next introduce a group of scaling automorphisms that maps the cone $\mathcal{K}$ onto itself and define the scaling-invariant neighborhoods of the central path. Consider the following group of matrices

$$\mathcal{G}_i \equiv \left\{ \lambda\widetilde{T}_i \,:\, \lambda > 0,\; \widetilde{T}_i \in \Re^{k_i \times k_i},\; \widetilde{T}_i^T J_{k_i} \widetilde{T}_i = J_{k_i},\; (\widetilde{T}_i)_{00} > 0 \right\}, \tag{7}$$

where

$$J_{k_i} \equiv \begin{pmatrix} 1 & 0 \\ 0 & -I \end{pmatrix} \in \Re^{k_i \times k_i}.$$

It is well-known that $\mathcal{G}_i$ is exactly the auto-morphism group of the cone $\mathcal{K}_i$, namely the set of all nonsingular matrices T_i such that $\mathcal{K}_i = T_i(\mathcal{K}_i)$, where $T_i(\mathcal{K}_i) \equiv \{T_i x_i : x_i \in \mathcal{K}_i\}$. (Since we have not been able to find a reference for this fact, for the sake of completeness, we include its proof in the appendix.) Let

$$\mathcal{G} \equiv \{ T = \text{diag}\,(T_1, \ldots, T_n) \,:\, T_i \in \mathcal{G}_i,\; i = 1, \ldots, n \}.$$

It is easy to see that $\mathcal{G}$ is a subgroup of the auto-morphism group of the cone $\mathcal{K}$.

The following proposition gives an explicit formula for the unique symmetric positive definite matrix in $\mathcal{G}$ which carries e to x.

Proposition 1. *(Proposition 2.1 of [30]) For any $x \in \mathcal{K}^0$, there exists a unique symmetric matrix in $\mathcal{G}$ which maps e to x given by $T_x \equiv diag\,(T_{x_1}, \ldots, T_{x_n})$ where, for all $i = 1, \ldots, n$,*

$$T_{x_i} = \begin{pmatrix} x_{i0} & x_{i1}^T \\ x_{i1} & \beta_{x_i} I + \dfrac{x_{i1} x_{i1}^T}{\beta_{x_i} + x_{i0}} \end{pmatrix}, \tag{8}$$

and

$$\beta_{x_i} = \sqrt{x_{i0}^2 - \|x_{i1}\|^2}. \tag{9}$$

Moreover, T_x is positive definite.

We define the set of $2n$ eigenvalues $\{\lambda_i^j : i = 1, \ldots, n,\ \ j = 0, 1\}$ associated with an element $v = (v_1, \ldots, v_n) \in \Re^{k_1} \times \ldots \times \Re^{k_n}$ of the Cartesian algebra as

$$\lambda_i^0 = \lambda_i^0(v) \equiv v_{i0} - \|v_{i1}\|, \quad \lambda_i^1 = \lambda_i^1(v) \equiv v_{i0} + \|v_{i1}\|,$$

for $i = 1, \ldots, n$. Clearly, $v \in \mathcal{K}$ if and only if $\lambda_i^0 \geq 0$ for all i, and $v \in \mathcal{K}^0$ if and only if $\lambda_i^0 > 0$. See Sect. 2.4 of [30] for a motivation and a more detailed explanation of the notion of eigenvalues in the context of SOCP.

We are now ready to introduce the neighborhoods of the central path. For a pair $(x, s) \in \mathcal{K}^0 \times \mathcal{K}^0$, define the distances

$$d_2(x, s) \equiv \sqrt{\sum_{\substack{i=1,\ldots,n \\ j=0,1}} \left(\lambda_i^j(w_{xs}) - \mu\right)^2} = \sqrt{2}\,\|w_{xs} - \mu e\|.$$

$$d_\infty(x, s) \equiv \max_{\substack{i=1,\ldots,n \\ j=0,1}} \left|\lambda_i^j(w_{xs}) - \mu\right| = \max_{i=1,\ldots,n} \left[|w_{i0} + \|w_{i1}\| - \mu|\,, |w_{i0} - \|w_{i1}\| - \mu|\right],$$

where $\mu \equiv \mu(x, s)$ and

$$w_{xs} = (w_1, \ldots, w_n) \equiv T_x s. \tag{10}$$

The neighborhoods of the central path with opening $\gamma \in (0, 1)$ determined by the above distances are:

$$\mathcal{N}_2(\gamma) = \left\{ (x, s, y) \in F^0(\mathrm{P}) \times F^0(\mathrm{D}) \ :\ d_2(x, s) \leq \gamma\mu(x, s) \right\},$$
$$\mathcal{N}_\infty(\gamma) = \left\{ (x, s, y) \in F^0(\mathrm{P}) \times F^0(\mathrm{D}) \ :\ d_\infty(x, s) \leq \gamma\mu(x, s) \right\}.$$

It is easy to show that $d_\infty(x, s) \leq d_2(x, s)$ for every $(x, s) \in \mathcal{K}^0 \times \mathcal{K}^0$, and hence that $\mathcal{N}_2(\gamma) \subset \mathcal{N}_\infty(\gamma)$.

In the above definitions, the product $w_{xs} = T_x s$ arises in exactly the same way as the quantity $X^{1/2}SX^{1/2}$ does in the context of SDP, namely $X^{1/2}SX^{1/2}$ is the scaled dual variable when the primal variable X is scaled to the identity matrix I, which plays the role of the identity in the associated Euclidean Jordan Algebra. Likewise, w_{xs} is the scaled dual variable when x is scaled to e. Then, it is intuitive that the quantity $\|X^{1/2}SX^{1/2} - \mu I\|_F$ should be replaced by $\sqrt{2}\,\|w_{xs} - \mu e\|$ when the distance of a point to the central path is being defined in in the context of SOCP. In fact, in terms of eigenvalues, the definition of the above distances is completely identical in both contexts.

The following invariance property of the eigenvalues of w_{xs} has been established in Proposition 2.4 of [30].

Proposition 2. *Suppose that* $(x, s) \in \mathcal{K}^0 \times \mathcal{K}^0$ *and* $G \in \mathcal{G}$. *Let* $(\widetilde{x}, \widetilde{s}) \equiv (G^T x, G^{-1} s)$, $w \equiv w_{xs}$ *and* $\widetilde{w} \equiv w_{\widetilde{x}\widetilde{s}}$. *Then:*

a) $\widetilde{w}_{i0} = w_{i0}$ *and* $\|\widetilde{w}_{i1}\| = \|w_{i1}\|$ *for every* $i = 1, \ldots, n$;

b) $\lambda_i^j(\widetilde{w}) = \lambda_i^j(w)$ *for every* $i = 1, \ldots, n$ *and* $j = 0, 1$;

c) $d_2(\widetilde{x}, \widetilde{s}) = d_2(x, s)$ *and* $d_\infty(\widetilde{x}, \widetilde{s}) = d_\infty(x, s)$.

Remark. As it is known from Faraut and Koráyi [7], every homogeneous and self-dual cone is in one-to-one correspondence with an Euclidean Jordan algebra. Since this Jordan algebra plays an important role in the generalization of primal-dual algorithms from LP and/or SDP to SOCP, it may be worthwhile to mention the correspondence of the concepts used in this paper with the ones used in a more general treatment of Euclidean Jordan algebra such as the one of [7]. Indeed, the quantities $\mathbf{mat}(x)$ and T_x correspond to the terms $L(x)$ and $P(x^{1/2})$ used in [7], respectively, where $x^{1/2}$ is the unique element whose square Jordan product is x. Here, $L(x)$ is the linear operator defined as $L(x)s \equiv x \circ s$, where $x \circ s$ denotes the product of the Jordan algebra, and $P(x)$ is the quadratic representation operator associated with x. Using this correspondence, it is possible to extend some of the concepts and results introduced here to any Euclidean Jordan algebra. We refer the reader to [6] and [28] for some preliminary results along this direction.

3. Technical results

In this section, we develop the technical results needed to establish the polynomial convergence of the algorithms presented in Sect. 4. Lemmas 2 and 3 are key results towards establishing the well-definedness of the AHO direction and obtaining a bound on the centrality measure of the next iterate in terms of that for the current iterate. In Lemma 4 and Theorem 1, we show that the AHO direction is well-defined in any neighborhood $\mathcal{N}_\infty(\gamma)$ with $\gamma \in (0, 1/3)$. Lemmas 8 and 9 are the main results used in the analysis of the algorithms presented in Sect. 4 and allow us to show that all iterates remain inside some 2-norm neighborhood and eventually approach the primal-dual optimal set.

Given $(x, s) \in \mathcal{K}^0 \times \mathcal{K}^0$, let

$$X \equiv \mathbf{mat}(x),$$
$$S \equiv \mathbf{mat}(s),$$
$$R_{xs} \equiv T_x X^{-1} S T_x, \tag{11}$$
$$W_{xs} \equiv \mathbf{mat}(w_{xs}). \tag{12}$$

Lemma 1. *For any $x \in \mathcal{K}^0$, the matrices X and T_x satisfy:*

a) $X - T_x = U_x = \mathrm{diag}\ (U_{x_i} : i = 1, \ldots, n)$, *where*

$$U_{x_i} = \begin{bmatrix} 0 & 0 \\ 0 & (x_{i0} - \beta_{x_i}) P_{x_i} \end{bmatrix}$$

and P_{x_i} is the orthogonal projection matrix onto the subspace orthogonal to x_{i1}, namely

$$P_{x_i} \equiv I - \frac{x_{i1} x_{i1}^T}{\|x_{i1}\|^2}; \tag{13}$$

b) $T_x X^{-1} = X^{-1} T_x = \mathrm{diag}\ (I - x_{i0}^{-1} U_{x_i} : i = 1, \ldots, n)$*; as a consequence,* $T_x X^{-1} e = e$*;*

c) *X and T_x commute and* $X \succeq T_x$.

Proof. The i-th diagonal block of $X - T_x$ is the matrix $X_i - T_{x_i}$, where X_i and T_{x_i} are given by (3) and (8), respectively. A simple algebraic manipulation involving (3), (8) and (9) reveals that $X_i - T_{x_i} = U_{x_i}$, and hence that a) holds. It is easy to see that $X_i U_{x_i} = U_{x_i} X_i = x_{i0} U_{x_i}$, and hence $X_i^{-1} U_{x_i} = x_{i0}^{-1} U_{x_i}$. Using this and a), we obtain

$$X^{-1} T_x = X^{-1}(X - U_x) = I - X^{-1} U_x = \text{diag}\left(I - X_i^{-1} U_{x_i} : i = 1, \ldots, n\right)$$
$$= \text{diag}\left(I - x_{i0}^{-1} U_{x_i} : i = 1, \ldots, n\right),$$

that is b) holds. This implies that $X^{-1} T_x$ is a symmetric matrix, or equivalently that X^{-1} and T_x commute. Hence, X and T_x commute. The other claim in c) that $X \succeq T_x$ follows immediately from a) since $U_x \succeq 0$.

□

The next two lemmas play a fundamental role in our analysis.

Lemma 2. *We have* $R_{xs} = diag\ (R_i : i = 1, \ldots, n)$, *where*

$$R_i = \begin{bmatrix} w_{i0} & w_{i1}^T \\ w_{i1} & \tilde{R}_i \end{bmatrix} \tag{14}$$

with $(w_{i0}, w_{i1}) = T_{x_i} s_i \in \Re \times \Re^{k_i - 1}$ *and*

$$\tilde{R}_i \equiv \frac{1}{x_{i0}} \left[w_{i1} x_{i1}^T + \beta_{x_i}^2 s_{i0} I \right] = \frac{w_{i1} x_{i1}^T}{x_{i0}} + \left(w_{i0} - \frac{w_{i1}^T x_{i1}}{x_{i0}} \right) I. \tag{15}$$

Proof. By (11) and Lemma 1(b), we have

$$R_i = T_{x_i} X_i^{-1} S_i T_{x_i} = \left(T_{x_i} X_i^{-1}\right) S_i X_i \left(X_i^{-1} T_{x_i}\right)$$
$$= \left(I - x_{i0}^{-1} U_{x_i}\right) S_i X_i \left(I - x_{i0}^{-1} U_{x_i}\right) \tag{16}$$

where X_i and S_i are the i-th diagonal blocks of X and S, respectively. Now, using (13) and the definition of X_i and S_i, we easily see that

$$I - \frac{1}{x_{i0}} U_{x_i} = \begin{bmatrix} 1 & 0 \\ 0 & I - \tau_i P_{x_i} \end{bmatrix}, \qquad S_i X_i = \begin{bmatrix} p_{i0} & p_{i1}^T \\ p_{i1} & Z_i \end{bmatrix}, \tag{17}$$

where

$$p_i \equiv x_i \circ s_i = X_i s_i, \qquad \tau_i \equiv 1 - \frac{\beta_{x_i}}{x_{i0}}, \qquad Z_i \equiv x_{i0} s_{i0} I + s_{i1} x_{i1}^T. \tag{18}$$

Substituting the identities in (17) into (16), we obtain

$$R_i = \begin{bmatrix} p_{i0} & (p_{i1} - \tau_i P_{x_i} p_{i1})^T \\ p_{i1} - \tau_i P_{x_i} p_{i1} & V_i \end{bmatrix},$$

where

$$V_i \equiv \left(I - \tau_i P_{x_i}\right) Z_i \left(I - \tau_i P_{x_i}\right). \tag{19}$$

Hence relation (14) follows once we show that $w_{i0} = p_{i0}$, $w_{i1} = p_{i1} - \tau_i P_{x_i} p_{i1}$ and $\tilde{R}_i = V_i$. Indeed, using (17), (18) and Lemma 1(b), we obtain

$$w_i = T_{x_i} s_i = \left(T_{x_i} X_i^{-1}\right) X_i s_i = \left(I - x_{i0}^{-1} U_{x_i}\right) p_i = \begin{pmatrix} p_{i0} \\ (I - \tau_i P_{x_i}) p_{i1} \end{pmatrix},$$

from which it follows that $w_{i0} = p_{i0}$ and $w_{i1} = p_{i1} - \tau_i P_{x_i} p_{i1}$. We will now show that $\tilde{R}_i = V_i$. By (9) and (18), we have

$$\begin{aligned} Z_i = x_{i0} s_{i0} I + s_{i1} x_{i1}^T \; &= \; x_{i0} s_{i0} I + \frac{p_{i1} - s_{i0} x_{i1}}{x_{i0}} x_{i1}^T \\ &= \frac{s_{i0}}{x_{i0}} \left(x_{i0}^2 - \|x_{i1}\|^2\right) I + \frac{s_{i0} \|x_{i1}\|^2}{x_{i0}} P_{x_i} + \frac{p_{i1} x_{i1}^T}{x_{i0}} \\ &= \frac{s_{i0}}{x_{i0}} \beta_{x_i}^2 I + \frac{s_{i0} \|x_{i1}\|^2}{x_{i0}} P_{x_i} + \frac{p_{i1} x_{i1}^T}{x_{i0}}. \end{aligned}$$

Substituting this relation into (19) and using the definition of $\tilde{R}_i$ given in (15) and the identities $w_{i1} = (I - \tau_i P_{x_i}) p_{i1}$, $P_{x_i} x_{i1} = 0$ and $P_{x_i}^2 = P_{x_i}$, we obtain

$$\begin{aligned} V_i &= \frac{s_{i0}}{x_{i0}} \beta_{x_i}^2 \left(I - \tau_i P_{x_i}\right)^2 + \frac{s_{i0} \|x_{i1}\|^2}{x_{i0}} \left(I - \tau_i P_{x_i}\right)^2 P_{x_i} + \frac{w_{i1} x_{i1}^T}{x_{i0}} \\ &= \frac{s_{i0}}{x_{i0}} \beta_{x_i}^2 \left(I - 2\tau_i P_{x_i} + \tau_i^2 P_{x_i}\right) + \frac{s_{i0} \|x_{i1}\|^2}{x_{i0}} \left(P_{x_i} - 2\tau_i P_{x_i} + \tau_i^2 P_{x_i}\right) + \frac{w_{i1} x_{i1}^T}{x_{i0}} \\ &= \tilde{R}_i + \frac{s_{i0}}{x_{i0}} \left[\left(\beta_{x_i}^2 + \|x_{i1}\|^2\right)\left(\tau_i^2 - 2\tau_i\right) + \|x_{i1}\|^2\right] P_{x_i} \; = \; \tilde{R}_i, \end{aligned}$$

where the last equality follows from the fact that the coefficient of P_{x_i} on its left hand side is zero, a fact that can be easily verified by using the definitions of β_{x_i} and τ_i in (9) and (18). We have thus shown that (14) holds.

It remains to show the identity in (15). A simple calculation using (9) and the fact that $w_{i0} = p_{i0}$ reveal that $\beta_{x_i}^2 s_{i0} = w_{i0} x_{i0} - p_{i1}^T x_{i1}$. Using the identities $w_{i1} = p_{i1} - \tau_i P_{x_i} p_{i1}$ and $P_{x_i} x_{i1} = 0$, we easily see that $w_{i1}^T x_{i1} = p_{i1}^T x_{i1}$. Hence, $\beta_{x_i}^2 s_{i0} = w_{i0} x_{i0} - w_{i1}^T x_{i1}$, from which (15) immediately follows.

□

Lemma 3. *Let $(x, s, y) \in \mathcal{K}^0 \times \mathcal{K}^0 \times \Re^m$ be a triple such that*

$$\max_{i,j} \left| \lambda_i^j (w_{xs}) \; - \; \nu \right| \; \leq \gamma \, \nu$$

for some scalars $\gamma > 0$ and $\nu > 0$. Then,

$$\| R_{xs} - W_{xs} \| \leq 2 \gamma \nu, \tag{20}$$

$$\| W_{xs} - \nu I \| \leq \gamma \nu. \tag{21}$$

As a consequence,

$$\| R_{xs} - \nu I \| \; \leq \; 3 \gamma \nu.$$

Proof. We first show that (20) holds. Let $(w_{i0}, w_{i1}) = T_{x_i} s_i$, and let W_i be the i-th block of W_{xs}. By Lemma 2 and the definition of W_{xs}, we have

$$\|R_{xs} - W_{xs}\| \; = \; \max_{i=1,\dots,n} \|R_i - W_i\| \; = \; \max_{i=1,\dots,n} \|\tilde{R}_i - w_{i0} I\|.$$

On the other hand, using (15) we obtain for all i that

$$\|\tilde{R}_i - w_{i0} I\| = \left\| \frac{1}{x_{i0}} \left[w_{i1} x_{i1}^T - \left(w_{i1}^T x_{i1} \right) I \right] \right\| \; \leq \; \frac{2\,\|w_{i1}\|\,\|x_{i1}\|}{x_{i0}} \; \leq \; 2\,\|w_{i1}\|$$
$$= \left(\lambda_i^1(w_{xs}) - \lambda_i^0(w_{xs}) \right) \; \leq \; \left| \lambda_i^1(w_{xs}) - \nu \right| + \left| \lambda_i^0(w_{xs}) - \nu \right| \; \leq \; 2\gamma\nu.$$

Inequality (20) now follows from the above two relations. Inequality (21) follows from (4).

□

In the following two results, we establish the well-definedness of the AHO direction for points lying in any neighborhood $\mathcal{N}_\infty(\gamma)$ with $\gamma \in (0, 1/3)$.

Lemma 4. *Let $(x, s, y) \in \mathcal{K}^0 \times \mathcal{K}^0 \times \Re^m$ be a triple such that*

$$\|R_{xs} - \nu I\| \leq \tau\nu, \tag{22}$$

for some scalars $\tau \in (0, 1)$ and $\nu > 0$. Assume that $(u, v) \in \Re^K \times \Re^K$ and $h \in \Re^K$ satisfy

$$Su + Xv = h, \quad u^T v \geq 0, \tag{23}$$

and define $\delta_u \equiv \|T_x^{-1} u\|$ and $\delta_v \equiv \|T_x v\|$. Then,

$$\delta_u \leq \frac{\left\|T_x X^{-1} h\right\|}{(1-\tau)\nu}, \quad \delta_v \leq \frac{2\left\|T_x X^{-1} h\right\|}{1-\tau}. \tag{24}$$

Proof. It is easy to see that (22) implies that

$$d^T R_{xs}\, d \; \geq \; (1-\tau)\, \nu\, \|d\|^2. \tag{25}$$

Multiplying the first relation in (23) on the left by $u^T X^{-1}$ and using the second relation in (23), we obtain

$$u^T X^{-1} S u \; \leq \; u^T X^{-1} S u + u^T v \; = \; u^T X^{-1} h.$$

This relation, (11) and (25) with $d = T_x^{-1} u$ imply that

$$(1-\tau)\, \nu\, \delta_u^2 = (1-\tau)\nu \left\| T_x^{-1} u \right\|^2 \; \leq \; \left(T_x^{-1} u \right)^T R_{xs} \left(T_x^{-1} u \right)$$
$$= u^T X^{-1} S u \; \leq \; u^T X^{-1} h \; = \; \left(T_x^{-1} u \right)^T \left(T_x X^{-1} h \right)$$
$$\leq \left\| T_x^{-1} u \right\| \left\| T_x X^{-1} h \right\| \; = \; \delta_u \left\| T_x X^{-1} h \right\|,$$

from which the first inequality in (24) follows. We now prove the second inequality of (24). Multiplying the first relation in (23) on the left by $T_x X^{-1}$ and using the definition of R_{xs}, we obtain

$$R_{xs} T_x^{-1} u + T_x v = T_x X^{-1} h.$$

This relation together with (22) and the first inequality of (24) implies that

$$\begin{aligned} \delta_v = \|T_x v\| \;&=\; \left\|T_x X^{-1} h - R_{xs} T_x^{-1} u\right\| \;\le\; \left\|T_x X^{-1} h\right\| + \left\|R_{xs} T_x^{-1} u\right\| \\ &\le \left\|T_x X^{-1} h\right\| + \|R_{xs}\|\, \delta_u \;\le\; \left\|T_x X^{-1} h\right\| + (1+\tau)\,\nu\, \frac{\left\|T_x X^{-1} h\right\|}{(1-\tau)\nu}, \end{aligned}$$

from which the second inequality of (24) follows.

□

As a consequence of the above lemma, we obtain the following result about the well-definedness of the AHO direction.

Theorem 1. *Let $(x, s, y) \in \mathcal{K}^0 \times \mathcal{K}^0 \times \Re^m$ be a point such that*

$$\max_{i,j} \left|\lambda_i^j(w_{xs}) - \nu\right| \le \gamma\, \nu$$

for some scalars $\gamma \in (0, 1/3)$ and $\nu > 0$. Then, system (6) has exactly one solution. In particular, the AHO direction is well-defined at every point $(x, s, y) \in \mathcal{K}^0 \times \mathcal{K}^0 \times \Re^m$ such that $d_\infty(x, s) < \mu(x, s)/3$.

Proof. To show that (6) has a unique solution, let $(u, v, q) \in \Re^K \times \Re^K \times \Re^m$ be a solution of the homogeneous system associated with (6). Then, $Su + Xv = 0$, $Au = 0$ and $A^T q + v = 0$. The last two relations imply that $u^T v = 0$. By Lemma 3, (22) holds with $\tau = 3\gamma < 1$. Using Lemma 4 with $h = 0$, we conclude that $u = v = 0$, and hence that $A^T q = 0$. Since the rows of A are linearly independent, we have $q = 0$. We have thus shown that $(u, v, q) = (0, 0, 0)$. This implies that system (6) has a unique solution.

□

Let $x(\alpha) \equiv x + \alpha\Delta x$, $s(\alpha) \equiv s + \alpha\Delta s$, $y(\alpha) \equiv y + \alpha\Delta y$, and let $\mu(\alpha) \equiv x(\alpha)^T s(\alpha)/n$. In the next four lemmas we develop a bound on the quantity $\sqrt{2}\,\|T_x^{-1} x(\alpha) \circ T_x s(\alpha) - \mu(\alpha) e\|$ which, as we will see in Lemma 9, majorizes the centrality measure $d_2(x(\alpha), s(\alpha))$.

Lemma 5. *Let $(x, s, y) \in F^0(\mathrm{P}) \times F^0(\mathrm{D})$ and let $(\Delta x, \Delta s, \Delta y)$ be a solution of (6) for some $\sigma \in \Re$. Then, for every $\alpha \in \Re$, we have:*

$$\mu(\alpha) = (1 - \alpha + \sigma\alpha)\,\mu, \tag{26}$$

$$\begin{aligned} T_x^{-1} x(\alpha) \circ T_x s(\alpha) - \mu(\alpha) e = (1-\alpha)(w_{xs} - \mu e) + \alpha(W_{xs} - R_{xs})\widetilde{\Delta x} \\ + \alpha^2\, \widetilde{\Delta x} \circ \widetilde{\Delta s}, \end{aligned} \tag{27}$$

where $\mu \equiv \mu(x, s)$ and

$$\widetilde{\Delta x} \;\equiv\; T_x^{-1} \Delta x, \qquad \widetilde{\Delta s} \;\equiv\; T_x \Delta s. \tag{28}$$

Proof. Using the assumption that $(x, s, y) \in F^0(\mathrm{P}) \times F^0(\mathrm{D})$ and relations (6b) and (6c), we easily see that $\Delta x^T \Delta s = 0$. Multiplying (6a) on the left by e^T, we obtain

$$s^T \Delta x + x^T \Delta s = \sigma n\mu - x^T s = -(1-\sigma)n\mu.$$

Using these two last relations, we obtain

$$x(\alpha)^T s(\alpha) = (x + \alpha\Delta x)^T (s + \alpha\Delta s) = x^T s + \alpha(s^T \Delta x + x^T \Delta s) = n\mu[1 - \alpha(1-\sigma)].$$

Dividing both sides of this relation by n, we obtain (26). Multiplying (6a) on the left by $T_x X^{-1}$ and using (10), (11), (28) and Lemma 1(b), we obtain

$$\widetilde{\Delta s} \;=\; T_x \Delta s \;=\; T_x X^{-1} (\sigma\mu e - Xs - S\Delta x) \;=\; \sigma\mu e - w_{xs} - R_{xs}\widetilde{\Delta x}.$$

The last identity, relations (10), (12) and (28) and the fact that $u \circ v = \mathbf{mat}(u)\, v$ for all $u, v \in \Re^K$ imply that

$$\begin{aligned} T_x^{-1} x(\alpha) \circ T_x\, s(\alpha) &= T_x^{-1}(x + \alpha\Delta x) \circ T_x\,(s + \alpha\Delta s) \\ &= \left(e + \alpha\widetilde{\Delta x}\right) \circ \left(w_{xs} + \alpha\widetilde{\Delta s}\right) \\ &= w_{xs} + \alpha\left(w_{xs} \circ \widetilde{\Delta x} + \widetilde{\Delta s}\right) + \alpha^2\, \widetilde{\Delta x} \circ \widetilde{\Delta s} \\ &= w_{xs} + \alpha\left[\sigma\mu e - w_{xs} + (W_{xs} - R_{xs})\,\widetilde{\Delta x}\right] + \alpha^2\, \widetilde{\Delta x} \circ \widetilde{\Delta s}. \end{aligned}$$

Combining this identity with (26), we obtain (27).

□

Lemma 6. *Assume that* $(x, s, y) \in \mathcal{N}_2(\gamma)$ *for some scalar* $\gamma \in (0, 1/3)$ *and let* $(\Delta x, \Delta s, \Delta y)$ *be the unique solution of (6) for some* $\sigma \in \Re$. *Then, the directions* $\widetilde{\Delta x}$ *and* $\widetilde{\Delta s}$ *defined in (28) satisfy:*

$$\|\widetilde{\Delta x}\| \;\le\; \frac{\Theta}{2}, \qquad \|\widetilde{\Delta s}\| \;\le\; \Theta\mu, \tag{29}$$

where $\mu \equiv \mu(x, s)$ *and*

$$\Theta \equiv \frac{2\left[\,\gamma^2/2 + (1-\sigma)^2 n\,\right]^{1/2}}{1 - 3\gamma}.$$

Proof. Using the fact that $w_{xs}^T e = s^T T_x e = s^T x = n\mu$ and $\|w_{xs} - \mu e\| \le \gamma\mu/\sqrt{2}$, we obtain

$$\begin{aligned} \|w_{xs} - \sigma\mu e\|^2 &= \|w_{xs} - \mu e\|^2 + \|\mu e - \sigma\mu e\|^2 + 2(1-\sigma)\mu(w_{xs} - \mu e)^T e \\ &\le \left[\frac{\gamma^2}{2} + (1-\sigma)^2 n\right]\mu^2. \end{aligned} \tag{30}$$

Since $d_\infty(x, s) \le d_2(x, s) \le \gamma\mu$, it follows from Lemma 3 with $\nu = \mu$ that (22) holds with $\tau = 3\gamma < 1$ and $\nu = \mu$. Hence, it follows from (10), (30), Lemma 1(b) and Lemma 4 with $\nu = \mu$, $(u, v) = (\Delta x, \Delta s)$, $h = \sigma\mu e - Xs$ and $\tau = 3\gamma$ that

$$\begin{aligned} \|\widetilde{\Delta x}\| &\le \frac{\left\|T_x X^{-1}(\sigma\mu e - Xs)\right\|\, \mu^{-1}}{1 - 3\gamma} \;=\; \frac{\|w_{xs} - \sigma\mu e\|\, \mu^{-1}}{1 - 3\gamma} \;\le\; \frac{\Theta}{2}, \\ \|\widetilde{\Delta s}\| &\le \frac{2\left\|T_x X^{-1}(\sigma\mu e - Xs)\right\|}{1 - 3\gamma} \;=\; \frac{2\,\|w_{xs} - \sigma\mu e\|}{1 - 3\gamma} \;\le\; \Theta\mu. \end{aligned}$$

□

Lemma 7. *Let $u_i, v_i \in \Re^{k_i}$ for $i = 1, ..., n$ and define $u \equiv (u_1, ..., u_n)$ and $v \equiv (v_1, ..., v_n)$. Then,*

$$\|u \circ v\| \le \sqrt{2}\,\|u\|\,\|v\|.$$

Proof. See Lemma 2.12 of [30].

□

Lemma 8. *Assume that $(x, s, y) \in \mathcal{N}_2(\gamma)$ for some scalar $\gamma \in (0, 1/3)$ and let $(\Delta x, \Delta s, \Delta y)$ be the unique solution of (6) for some $\sigma \in \Re$. Then, for any $\alpha \in [0, 1]$, there holds*

$$\sqrt{2}\,\big\|\, T_x^{-1} x(\alpha) \circ T_x\, s(\alpha) - \mu(\alpha) e \,\big\| \;\le\; \big[\, (1-\alpha)\gamma + \sqrt{2}\,\alpha\,\gamma\,\Theta + \alpha^2\,\Theta^2 \,\big]\,\mu, \tag{31}$$

where $\mu \equiv \mu(x, s)$.

Proof. Since $d_\infty(x, s) \le d_2(x, s)$, the assumption of Lemma 3 is satisfied with $\nu = \mu$. Using (20) with $\nu = \mu$, (27), (29) and Lemma 7, we obtain for $\alpha \in [0, 1]$ that

$$\begin{aligned}
&\sqrt{2}\,\big\|\, T_x^{-1} x(\alpha) \circ T_x\, s(\alpha) - \mu(\alpha) e \,\big\| \\
&\le (1-\alpha)\sqrt{2}\,\|w_{xs} - \mu e\| + \alpha\sqrt{2}\,\|R_{xs} - W_{xs}\|\,\|\widetilde{\Delta x}\| + \alpha^2\sqrt{2}\,\|\widetilde{\Delta x} \circ \widetilde{\Delta s}\| \\
&\le (1-\alpha)\gamma\,\mu + \sqrt{2}\,\alpha\,\gamma\,\Theta\,\mu + 2\,\alpha^2\|\widetilde{\Delta x}\|\,\|\widetilde{\Delta s}\| \\
&\le \big[\,(1-\alpha)\,\gamma + \sqrt{2}\,\alpha\,\gamma\,\Theta + \alpha^2\,\Theta^2\,\big]\,\mu.
\end{aligned}$$

□

The next lemma allows us to show that the left hand side of (31) majorizes $d_2(x(\alpha), s(\alpha))$ (see the proof of Theorem 2).

Lemma 9. *Suppose that $(x, s) \in \mathcal{K}^0 \times \mathcal{K}^0$ and let $\mu \equiv \mu(x, s)$. Then,*

$$d_2(x, s) \equiv \sqrt{2}\,\|w_{xs} - \mu e\| = \min_{G \in \mathcal{G}} \sqrt{2}\,\|x_G \circ s_G - \mu e\|, \tag{32}$$

where $x_G \equiv G^T x$ and $s_G \equiv G^{-1} s$ for every $G \in \mathcal{G}$.

Proof. The lemma immediately follows from Lemma 2.10 of [30]. Here we give an alternative proof by using Lemma 1(c). Indeed, first note that if $G = T_x^{-1}$ then $d_2(x, s) = \sqrt{2}\,\|x_G \circ s_G - \mu e\|$, from which it follows that the above minimum is less than or equal to $d_2(x, s)$. Now let $G \in \mathcal{G}$ be given and let $(\widetilde{x}, \widetilde{s}) \equiv (x_G, s_G)$. Lemma 1(c) implies that $\widetilde{X}^2 \succeq (T_{\widetilde{x}})^2$, where $\widetilde{X} \equiv \mathbf{mat}(\widetilde{x})$. Hence,

$$\|\widetilde{X} v\|^2 = v^T \widetilde{X}^2 v \ge v^T (T_{\widetilde{x}})^2 v = \|T_{\widetilde{x}}\, v\|^2, \quad \forall v \in \Re^K.$$

This inequality with $v = \widetilde{s} - \mu \widetilde{X}^{-1} e$ together with Proposition 2(c) and Lemma 1(b) implies that

$$\|x_G \circ s_G - \mu e\| = \|\widetilde{X}\widetilde{s} - \mu(\widetilde{x}, \widetilde{s}) e\| \ge \|T_{\widetilde{x}}\, \widetilde{s} - \mu(\widetilde{x}, \widetilde{s}) e\| = d_2(\widetilde{x}, \widetilde{s}) = d_2(x, s).$$

Since this inequality holds for every $G \in \mathcal{G}$, (32) follows.

□

The next lemma, which is the last result of this section, is used in Sect. 4 to establish feasibility of the sequence of iterates.

Lemma 10. *Let* $(x, s) \in \mathcal{K} \times \mathcal{K}$ *be given. If* $x \circ s \in \mathcal{K}^0$*, then* $(x, s) \in \mathcal{K}^0 \times \mathcal{K}^0$*. In particular, if* $\sqrt{2}\,\|x \circ s - \nu e\| \le \gamma\nu$ *for some* $\gamma \in (0, 1)$ *and* $\nu > 0$*, then* $(x, s) \in \mathcal{K}^0 \times \mathcal{K}^0$.

Proof. Let $p \equiv x \circ s$. Then, for every $i = 1, \ldots, n$,

$$p_{i0} = x_{i0}s_{i0} + x_{i1}^T s_{i1}, \tag{33}$$

$$p_{i1} = x_{i0}s_{i1} + s_{i0}x_{i1}. \tag{34}$$

To show the first statement, assume that $p \in \mathcal{K}^0$. Then, $\|p_{i1}\| < p_{i0}$ for all i. This implies that $x_{i0} > 0$ since otherwise we would have $x_i = 0$ due to the fact that $x_i \in \mathcal{K}_i$, and hence $p_{i0} = 0$ due to (33), obtaining a contradiction. By (34), we have $s_{i1} = (p_{i1} - s_{i0}x_{i1})/x_{i0}$, which combined with (33) yields

$$\begin{aligned} p_{i0} = x_{i0}s_{i0} + \frac{1}{x_{i0}} x_{i1}^T(p_{i1} - s_{i0}x_{i1}) &= \frac{s_{i0}}{x_{i0}}\left(x_{i0}^2 - \|x_{i1}\|^2\right) + \frac{x_{i1}^T p_{i1}}{x_{i0}} \\ \le \frac{s_{i0}}{x_{i0}}\left(x_{i0}^2 - \|x_{i1}\|^2\right) + \frac{\|x_{i1}\|\,\|p_{i1}\|}{x_{i0}} &\le \frac{s_{i0}}{x_{i0}}\left(x_{i0}^2 - \|x_{i1}\|^2\right) + \|p_{i1}\|, \end{aligned}$$

where the last inequality follows from the fact that $x_i \in \mathcal{K}_i$. Since $\|p_{i1}\| < p_{i0}$, we conclude from the last relation that $x_{i0} > \|x_{i1}\|$ for all i, that is, $x \in \mathcal{K}^0$. In a similar way, one can show that $s \in \mathcal{K}^0$.

To show the second statement, assume that $\sqrt{2}\,\|p - \nu e\| \le \gamma\nu$ for some $\gamma \in (0, 1)$ and $\nu > 0$. Then, for every i, we have

$$\begin{aligned} \gamma\nu \ge \sqrt{2}\|p_i - \nu e\| = \sqrt{2}\left((p_{i0} - \nu)^2 + \|p_{i1}\|^2\right)^{1/2} &\ge |p_{i0} - \nu| + \|p_{i1}\| \\ &\ge \nu - p_{i0} + \|p_{i1}\|, \end{aligned}$$

from which it follows that $p_{i0} - \|p_{i1}\| \ge (1 - \gamma)\nu > 0$. Hence, $p \in \mathcal{K}^0$.

□

4. Algorithms and polynomial convergence

In this section, we establish polynomial iteration-complexity bounds for two primal-dual feasible interior-point algorithms based on the Newton direction determined by (6). Both algorithms are extensions of well-known algorithms for linear programming: the first one is a short-step path-following method which generalizes the algorithms presented in Kojima, Mizuno and Yoshise [12] and Monteiro and Adler [19,20]; the second one is a predictor-corrector algorithm similar to the predictor-corrector LP method of Mizuno, Todd and Ye [16].

4.1. Short-step path following algorithm

In this subsection, we analyze the polynomial convergence of a short-step path following algorithm based on the search direction (6).

We start by stating the algorithm that will be considered in this subsection.

Algorithm-I:

Choose constants $\gamma \in (0, 1/3)$ and $\delta \in (0, 1)$ satisfying condition (36) below and let $\sigma \equiv 1 - \delta/\sqrt{2n}$. Let $\varepsilon \in (0, 1)$ and $(x^0, s^0, y^0) \in \mathcal{N}_2(\gamma)$, and set $\mu_0 \equiv \mu(x^0, s^0)$.

Repeat until $\mu_k \le \varepsilon\mu_0$, **do**

(1) Compute the solution $(\Delta x^k, \Delta s^k, \Delta y^k)$ of system (6) with $\mu = \mu_k$ and $(x, s, y) = (x^k, s^k, y^k)$;

(2) Set $(x^{k+1}, s^{k+1}, y^{k+1}) \equiv (x^k, s^k, y^k) + (\Delta x^k, \Delta s^k, \Delta y^k)$;

(3) Set $\mu_{k+1} \equiv \mu(x^{k+1}, s^{k+1})$ and increment k by 1.

End

Setting $\Gamma = \gamma$ in the following result, we obtain the analysis of one iteration of Algorithm-I for suitable choices of the constants γ and δ.

Theorem 2. *Let* $\gamma \in (0, 1/3)$ *and* $\delta \in (0, 1)$ *be constants satisfying*

$$\Gamma \equiv \frac{4(\gamma^2 + \delta^2)}{(1 - 3\gamma)^2} \left(1 - \frac{\delta}{\sqrt{2n}}\right)^{-1} < 1. \tag{35}$$

Suppose that $(x, s, y) \in \mathcal{N}_2(\gamma)$ *and let* $(\Delta x, \Delta s, \Delta y)$ *denote the solution of system (6) with* $\sigma \equiv 1 - \delta/\sqrt{2n}$*. Then,*

(a) $(\widehat{x}, \widehat{s}, \widehat{y}) \equiv (x + \Delta x, s + \Delta s, y + \Delta y) \in \mathcal{N}_2(\Gamma)$*;*
(b) $\mu(\widehat{x}, \widehat{s}) = \left(1 - \delta/\sqrt{2n}\right)\mu(x, s)$.

Proof. It follows from Lemma 8, the definition of σ, the fact that $\Theta \ge \sqrt{2}\gamma$ and (35) that for every $\alpha \in [0, 1]$,

$$\begin{aligned}
\sqrt{2}\, \| T_x^{-1} x(\alpha) \circ T_x\, s(\alpha) - \mu(\alpha) e \| &\le \left[(1-\alpha)\gamma + 2\alpha\Theta^2 \right] \mu \\
&\le \left\{ (1-\alpha)\gamma + \alpha \frac{8\,[\gamma^2/2 + (1-\sigma)^2 n\,]}{(1-3\gamma)^2} \right\} \mu \\
&= \left\{ (1-\alpha)\gamma + \alpha \frac{4\,(\gamma^2 + \delta^2\,)}{(1-3\gamma)^2} \right\} \mu \\
&= \left\{ (1-\alpha)\gamma + \alpha \left(1 - \frac{\delta}{\sqrt{2n}}\right) \Gamma \right\} \mu \\
&= \{(1-\alpha)\,\gamma + \alpha\,\Gamma\,\sigma\}\,\mu,
\end{aligned}$$

and hence, in view of (26),

$$\begin{aligned}\sqrt{2}\left\|T_x^{-1}x(\alpha)\circ T_x s(\alpha)-\mu(\alpha)e\right\| &\le \max\{\gamma,\Gamma\}(1-\alpha+\alpha\sigma)\mu\\ &= \max\{\gamma,\Gamma\}\mu(\alpha)<\mu(\alpha).\end{aligned}$$

Using the last relation together with Lemma 10, we easily see that $T_x^{-1}x(\alpha)\in\mathcal{K}^0$ and $T_x s(\alpha)\in\mathcal{K}^0$ for every $\alpha\in[0,1]$. In view of Proposition 1, this implies that $(x(\alpha),s(\alpha))\in\mathcal{K}^0\times\mathcal{K}^0$ for every $\alpha\in[0,1]$. Moreover, using the fact that $Ax=b$, $A\Delta x=0$, $A^Ty+s=c$ and $A^T\Delta y+\Delta s=0$, we easily see $Ax(\alpha)=b$ and $A^Ty(\alpha)+s(\alpha)=c$. We have thus shown that $(x(\alpha),s(\alpha),y(\alpha))\in F^0(\mathrm{P})\times F^0(\mathrm{D})$, for every $\alpha\in[0,1]$. Using Lemma 9 with $\widetilde{x}=T_x^{-1}x(\alpha)$ and $\widetilde{s}=T_x s(\alpha)$, we conclude that

$$d_2(x(\alpha),s(\alpha))\le\sqrt{2}\left\|T_x^{-1}x(\alpha)\circ T_x\, s(\alpha)-\mu(\alpha)e\right\|\le\Gamma\mu(\alpha),$$

for any $\alpha\in[0,1]$. Hence, $(\widehat{x},\widehat{s},\widehat{y})=(x(1),s(1),y(1))\in\mathcal{N}_2(\Gamma)$. Statement (b) is due to (26).

□

As an immediate consequence of Theorem 2, we have the following convergence result for Algorithm-I.

Corollary 1. *Assume that* $\gamma\in(0,1/3)$ *and* $\delta\in(0,1)$ *are constants satisfying*

$$\frac{4(\gamma^2+\delta^2)}{(1-3\gamma)^2}\le\left(1-\frac{\delta}{\sqrt{2n}}\right)\gamma. \tag{36}$$

Then, every iterate (x^k,s^k,y^k) *generated by Algorithm-I is in the neighborhood* $\mathcal{N}_2(\gamma)$ *and satisfies* ${x^k}^T s^k=\left(1-\delta/\sqrt{2n}\right)^k {x^0}^T s^0$. *Moreover, Algorithm-I terminates in at most* $\mathcal{O}(\sqrt{n}\,\log\varepsilon^{-1})$ *iterations.*

Examples of constants γ and δ satisfying the conditions of Corollary 1 are $\gamma=\delta=1/50$.

4.2. *Predictor-corrector algorithm*

In this subsection, we give the polynomial convergence analysis of a predictor-corrector algorithm which is a direct extension of the LP predictor-corrector algorithm studied by Mizuno, Todd and Ye [16].

The algorithm considered in this subsection is as follows.

Algorithm-II:

Choose a constant $0 < \tau \leq 1/30$.
Let $\varepsilon \in (0,1)$ and $(x^0, s^0, y^0) \in \mathcal{N}_2(\tau)$, and set $\mu_0 \equiv \mu(x^0, s^0)$.
Repeat until $\mu_k \leq \varepsilon\mu_0$, **do**

(1) Compute the solution $(\Delta x^k, \Delta s^k, \Delta y^k)$ of system (6) with $\sigma = 0$ and $(x, s, y) = (x^k, s^k, y^k)$;

(2) Let $\alpha_k \equiv \max\{\alpha \in [0,1] : (x^k(\alpha'), x^k(\alpha'), y^k(\alpha')) \in \mathcal{N}_2(2\tau), \forall \alpha' \in [0, \alpha]\}$, where $(x^k(\alpha), s^k(\alpha), y^k(\alpha)) \equiv (x^k + \alpha\Delta x^k, s^k + \alpha\Delta s^k, y^k + \alpha\Delta y^k)$;

(3) Let $(\widehat{x}^k, \widehat{s}^k, \widehat{y}^k) \equiv (x^k, s^k, y^k) + \alpha_k(\Delta x^k, \Delta s^k, \Delta y^k)$;

(4) Compute the solution $(\widehat{\Delta x}^k, \widehat{\Delta s}^k, \widehat{\Delta y}^k)$ of system (6) with $\sigma = 1$ and $(x, s, y) = (\widehat{x}^k, \widehat{s}^k, \widehat{y}^k)$;

(5) Set $(x^{k+1}, s^{k+1}, y^{k+1}) \equiv (\widehat{x}^k, \widehat{s}^k, \widehat{y}^k) + (\widehat{\Delta x}^k, \widehat{\Delta s}^k, \widehat{\Delta y}^k)$;

(6) Set $\mu_{k+1} \equiv \mu(x^{k+1}, s^{k+1})$ and increment k by 1.

End

The following result provides the polynomial convergence analysis of the above algorithm.

Theorem 3. *Assume that* $\tau \in (0, 1/30]$. *Then, Algorithm-II satisfies the following statements:*

a) for every $k \geq 0$, $(x^k, s^k, y^k) \in \mathcal{N}_2(\tau)$ *and* $(\widehat{x}^k, \widehat{s}^k, \widehat{y}^k) \in \mathcal{N}_2(2\tau)$;
b) for every $k \geq 0$, $\mu(x^{k+1}, s^{k+1}) = \mu(\widehat{x}^k, \widehat{s}^k) = (1 - \alpha_k)\mu(x^k, s^k)$ *and*

$$\alpha_k = \frac{1}{\mathcal{O}(\sqrt{n})}.$$

c) the algorithm terminates in at most $\mathcal{O}(\sqrt{n}\,\log \varepsilon^{-1})$ *iterations.*

Proof. Statement (c) and the well-definedness of Algorithm-II follow directly from (a) and (b). In turn, these two statements follow by a simple induction argument and the two lemmas below.

□

The following lemma analyzes the predictor step of Algorithm-II, namely the step described in items (1)–(4) of Algorithm-II.

Lemma 11. *Suppose that* $(x, s, y) \in \mathcal{N}_2(\tau)$ *for some* $\tau \in (0, 1/6)$. *Let* $(\Delta x, \Delta s, \Delta y)$ *denote the solution of (6) with* $\sigma = 0$. *Let* $\bar{\alpha}$ *denote the unique positive root of the second-order polynomial* $p(\alpha) \equiv \Theta^2\alpha^2 + (\sqrt{2}\,\tau\,\Theta + \tau)\alpha - \tau$ *where*

$$\Theta \equiv \frac{2(\tau^2/2 + n)^{1/2}}{1 - 3\tau}.$$

Then, for any $\alpha \in [0, \bar{\alpha}]$, *we have:*

(a) $(x(\alpha), s(\alpha), y(\alpha)) \in \mathcal{N}_2(2\tau)$*;*
(b) $\mu(\alpha) = (1-\alpha)\mu$.

Moreover, $\bar{\alpha} = 1/\mathcal{O}(n^{1/2})$.

Proof. Using Lemma 8 with $\gamma = \tau$ and $\sigma = 0$, the fact that $p(\alpha) \le 0$ for $\alpha \in [0, \bar{\alpha}]$, relation (26) with $\sigma = 0$ and the definition of $p(\alpha)$, we obtain

$$\sqrt{2}\, \| T_x^{-1} x(\alpha) \circ T_x\, s(\alpha) - \mu(\alpha) e \| \le \left\{(1-\alpha)\tau + \sqrt{2}\alpha\, \tau\, \Theta + \alpha^2 \Theta^2\right\}\mu$$
$$= 2\tau\mu(\alpha) + p(\alpha)\mu \ \le\ 2\,\tau\mu(\alpha).$$

An argument similar to the one used in Theorem 2 together with the fact that $2\tau < 1/3$ can be used to show that (a) holds. Statement (b) follows from (26) with $\sigma = 0$. The last assertion of the lemma is due to (26).

□

The following lemma analyzes the corrector step of Algorithm-II, namely the step described in items (5)–(7) of Algorithm-II.

Lemma 12. *Suppose* $(\widehat{x}, \widehat{s}, \widehat{y})$ *is in* $\mathcal{N}_2(2\tau)$ *for some* $\tau \in (0, 1/30]$. *Let* $(\widehat{\Delta x}, \widehat{\Delta s}, \widehat{\Delta y})$ *denote the solution of (6) with* $(x, s, y) = (\widehat{x}, \widehat{s}, \widehat{y})$ *and* $\sigma = 1$. *Then,*

$$(\widehat{x}, \widehat{s}, \widehat{y}) + (\widehat{\Delta x}, \widehat{\Delta s}, \widehat{\Delta y}) \in \mathcal{N}_2(\tau),$$
$$\mu(\widehat{x} + \widehat{\Delta x}\,, \widehat{s} + \widehat{\Delta s}) = \mu(\widehat{x}\,, \widehat{s}).$$

Proof. The result follows immediately from Theorem 2 with $\delta = 0$, $(x, s, y) = (\widehat{x}, \widehat{s}, \widehat{y})$, $\gamma = 2\tau$ and $\Gamma = \tau$, and the fact that (35) holds when $0 < \tau \le 1/30$.

□

5. The MZ family of directions

In this section we introduce the MZ family of directions which is a natural extension of the Monteiro and Zhang family of directions for SDP to the context of SOCP. As in the context of SDP, this family arises by computing the AHO direction (6) with respect to a scaled problem and mapping the direction back to the original space. Each direction of the family is then associated with the scaling matrix chosen to construct the scaled problem.

Given a matrix $G \in \mathcal{G}$, consider the following change of variables

$$\widetilde{x} \equiv G^T x, \quad (\widetilde{s}, \widetilde{y}) \equiv (G^{-1}s, y). \tag{37}$$

Letting

$$\tilde{c} = G^{-1}c, \quad \widetilde{A} \equiv AG^{-T}, \quad \widetilde{b} \equiv b,$$

we easily see that problem (P) and (D) of Sect. 2 can be written in terms of these new variables as

$$(\tilde{\mathrm{P}}) \quad \min\{\tilde{c}^T\widetilde{x} : \widetilde{A}\widetilde{x} = \widetilde{b},\ \widetilde{x} \in \mathcal{K}\},$$
$$(\tilde{\mathrm{D}}) \quad \max\{\widetilde{b}^T\widetilde{y} : \widetilde{A}^T\widetilde{y} + \widetilde{s} = \tilde{c},\ \widetilde{s} \in \mathcal{K}\}.$$

Due to Proposition 2 and the fact that $\mu(x, s) = \mu(\widetilde{x}, \widetilde{s})$, we have

$$(\widetilde{x}, \widetilde{s}, \widetilde{y}) \in \widetilde{\mathcal{N}}_2(\gamma) \Longleftrightarrow (x, s, y) \in \mathcal{N}_2(\gamma),$$
$$(\widetilde{x}, \widetilde{s}, \widetilde{y}) \in \widetilde{\mathcal{N}}_\infty(\gamma) \Longleftrightarrow (x, s, y) \in \mathcal{N}_\infty(\gamma),$$

where $\widetilde{\mathcal{N}}_2(\gamma)$ and $\widetilde{\mathcal{N}}_\infty(\gamma)$ denote the 2-norm and ∞-norm neighborhoods associated with the pair of problems $(\widetilde{\mathrm{P}}, \widetilde{\mathrm{D}})$. Moreover, if $(\widetilde{x}_\nu, \widetilde{s}_\nu, \widetilde{y}_\nu)$ denote the point on the central path with parameter $\nu > 0$ for the pair $(\widetilde{\mathrm{P}}, \widetilde{\mathrm{D}})$, then $(\widetilde{x}_\nu, \widetilde{s}_\nu, \widetilde{y}_\nu) = (G^T x_\nu, G^{-1}s_\nu, y_\nu)$.

The matrix $G \in \mathcal{G}$ also determines a scaled Newton direction (with parameter $\sigma > 0$) as follows. An interior feasible point (x, s, y) for (P,D) determines an interior feasible point $(\widetilde{x}, \widetilde{s}, \widetilde{y})$ for $(\widetilde{\mathrm{P}}, \widetilde{\mathrm{D}})$ as in (37). At the scaled point $(\widetilde{x}, \widetilde{s}, \widetilde{y})$, the AHO direction is computed and the resulting direction $(\widetilde{\Delta x},\ \widetilde{\Delta s},\ \widetilde{\Delta y})$ is mapped back into the original space to yield the scaled AHO direction, or MZ direction with scaling G, $(\Delta x, \Delta s, \Delta y) = (\Delta x_G, \Delta s_G, \Delta y_G)$ given by

$$\Delta x = G^{-T}\,\widetilde{\Delta x}, \quad (\Delta s, \Delta y) = (G\widetilde{\Delta s}, \widetilde{\Delta y}).$$

Hence, $(\Delta x, \Delta s, \Delta y) = (\Delta x_G, \Delta s_G, \Delta y_G)$ is a solution of

$$\widetilde{S}G^T\Delta x + \widetilde{X}G^{-1}\Delta s = \sigma\mu e - \widetilde{X}\widetilde{s}, \tag{39a}$$
$$A^T\Delta y + \Delta s = c - s - A^T y, \tag{39b}$$
$$A\Delta x = b - Ax, \tag{39c}$$

where $\widetilde{X} \equiv \mathbf{mat}(\widetilde{x})$ and $\widetilde{S} \equiv \mathbf{mat}(\widetilde{s})$.

The scaled AHO direction $(\Delta x_G, \Delta s_G, \Delta y_G)$ at the point (x, s, y) depends on G, and as G varies over the set of nonsingular matrices, we obtain a family of search directions, which we refer to as the MZ-family. When $G = I$, $G = T_s$ and $G = T_x^{-1}$, one obtains the AHO direction, the HRVW/KSH/M direction and the dual counterpart of the HRVW/KSH/M direction. The NT direction is also a member of the MZ family and is equal to the direction $(\Delta x_G, \Delta s_G, \Delta y_G)$ with $G = G_{xs}$, where G_{xs} is the unique positive definite symmetric matrix $G \in \mathcal{G}$ satisfying $Gx = G^{-1}s$, or equivalently $G^2x = s$.

The results obtained in Sects. 3 and 4 for the AHO direction can be extended to the whole MZ-family due to the fact that any member of this family reduces to the AHO direction in the scaled space and the fact that the duality gap and the centrality measures remain invariant. In what follows, we summarize these results.

Corollary 2. *Let $G \in \mathcal{G}$ and $(x, s, y) \in \mathcal{K}^0 \times \mathcal{K}^0 \times \Re^m$ be a point such that*

$$\max_{i,j} \left|\lambda_i^j(w_{xs}) - \nu\right| \le \gamma\,\nu$$

for some scalars $0 < \gamma < 1/3$ and $\nu > 0$. Then, system (39) has exactly one solution.

Consider the short-step path following algorithm and the Mizuno-Todd-Ye predictor-corrector algorithm in which at each iteration an arbitrary direction from the MZ family is computed, that is, at the k-th iteration, a matrix $G_k \in \mathcal{G}$ is chosen and the direction $(\Delta x_{G_k}, \Delta s_{G_k}, \Delta y_{G_k})$ is computed. Then, the results obtained in Corollary 1 and Theorem 3 hold exactly as stated.

Appendix

Proposition 3. *The auto-morphism group of the cone $\mathcal{K}_i$ is equal to the set $\mathcal{G}_i$ defined in (7).*

Proof. We first claim that the auto-morphism group of the set $\mathcal{K}_i \cup (-\mathcal{K}_i)$ is equal to

$$\left\{ \lambda \widetilde{T}_i \,:\, \lambda > 0,\ \widetilde{T}_i \in \Re^{k_i \times k_i},\ \widetilde{T}_i^T J_{k_i} \widetilde{T}_i = J_{k_i} \right\}, \tag{40}$$

Indeed, using the equivalence $x \in \mathcal{K}_i \cup (-\mathcal{K}_i) \Leftrightarrow x^T J_{k_i} x \geq 0$, it is easy to see that that the set (40) is contained in the auto-morphism group of $\mathcal{K}_i \cup (-\mathcal{K}_i)$. Conversely, assume that T_i is in the auto-morphism group of $\mathcal{K}_i \cup (-\mathcal{K}_i)$. Then, we have $x \in \mathcal{K}_i \cup (-\mathcal{K}_i) \Leftrightarrow T_i x \in \mathcal{K}_i \cup (-\mathcal{K}_i)$, which in turn is equivalent to $x^T J_{k_i} x \geq 0 \Leftrightarrow x^T T_i^T J_{k_i} T_i x \geq 0$, in view of the above equivalence. Consider now the minimization problem $\min\{x^T J_{k_i} x : x^T T_i^T J_{k_i} T_i x \geq 0\}$ and observe that its optimal value is zero. Moreover, any point in $\partial\mathcal{K}_i$ is an optimal solution of this problem. Hence, for any $0 \neq x \in \partial\mathcal{K}_i$, there exists $\lambda(x) \in \Re^{k_i}_+$ such that

$$(J_{k_i} - \lambda(x) B_i)x = 0. \tag{41}$$

where $B_i \equiv T_i^T J_{k_i} T_i$. From this equation, we easily see that $\lambda(x)$ is a continuous function of $x \in (\partial\mathcal{K}_i)\backslash\{0\}$ and that $\lambda(x) > 0$ is an eigenvalue of $J_{k_i} B_i^{-1}$ for all $x \in (\partial\mathcal{K}_i)\backslash\{0\}$. Since $J_{k_i} B_i^{-1}$ has a finite number of eigenvalues and $\lambda(x)$ is a continuous function over the connected set $(\partial\mathcal{K}_i)\backslash\{0\}$, there exists $\bar{\lambda} > 0$ such that $\lambda(x) = \bar{\lambda}$ for all $x \in (\partial\mathcal{K}_i)\backslash\{0\}$. Hence, by (41), it follows that $(J_{k_i} - \bar{\lambda} B_i)x = 0$ for all $x \in (\partial\mathcal{K}_i)\backslash\{0\}$. Since the set $(\partial\mathcal{K}_i)\backslash\{0\}$ is not contained in any manifold of dimension less than k_i, a simple argument reveals that $J_{k_i} = \bar{\lambda} B_i = \bar{\lambda}\, T_i^T J_{k_i} T_i$, and hence that J_{k_i} is in the set (40).

Now, assume that T_i is in the auto-morphism group of $\mathcal{K}_i$. Then, it is easy to see that $T_i(\operatorname{int}\mathcal{K}_i) = \operatorname{int}\mathcal{K}_i$. Hence, $(T_i)_{00} = (T_i e_i)_0 > 0$ since $e_i \in \operatorname{int}\mathcal{K}_i$. This fact together with the above claim shows that the auto-morphism group of $\mathcal{K}_i$ is contained in $\mathcal{G}_i$. Assume now that $T_i \in \mathcal{G}_i$. Then, by the above claim, T_i is in the auto-morphism group of $\mathcal{K}_i \cup (-\mathcal{K}_i)$. Moreover, we have $0 < (T_i)_{00} = (T_i e_i)_0$, showing that $T_i e_i \in \mathcal{K}_i$. Using the fact that T_i is invertible and a simple continuity argument, it is now easy to see that $T_i x \in \mathcal{K}_i$ for all $x \in \mathcal{K}_i$, or equivalently $T_i(\mathcal{K}_i) \subset \mathcal{K}_i$. This fact and the above claim imply that T_i is in the auto-morphism group of $\mathcal{K}_i$.

□

Acknowledgements. This work was partly done while the first author was visiting the Institute of Statistical Mathematics in Japan on March of 1998 and the second author was visiting the School of Industrial and Systems Engineering of Georgia Institute of Technology on April/May of 1998. The first author is grateful to this institute for providing a congenial scientific atmosphere during his stay. He would like to thank the second author for hosting his visit to the institute. The second author would also like to thank the first author for his support and hospitality during his visit to Georgia Tech on April/May of 1998, when final details about this work were developed.

References

1. Adler, I., Alizadeh, F. (1994): Primal-dual interior-point algorithms for convex quadratically constrained and semidefinite optimization problems. Manuscript, May, 1994 (Revised: August, 1995)
2. Alizadeh, F., Haeberly, J.-P., Overton, M. (1998): Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results. SIAM J. Optim., **8**, 746–768
3. Alizadeh, F., Schmieta, S. (1997): Optimization with semidefinite, quadratic and linear constraints. Technical Report RRR 23-97, RUTCOR Research Report, Rutgers University, NJ, USA, November 1997
4. Faybusovich, L. (1995): Jordan algebras, symmetric cones and interior-point methods. Manuscript, Department of Mathematics, Notre Dame, IN, USA
5. Faybusovich, L. (1997): Linear systems in Jordan algebras and primal-dual interior-point algorithms. J. Comput. Appl. Math. **86**, 149–175
6. Faybusovich, L. (1997): Euclidean Jordan algebras and interior-point algorithms. Positivity **1**, 331–357
7. Faraut, J., Koráyi, A. (1994): Analysis on Symmetric Cones. Oxford University Press, New York
8. Helmberg, C., Rendl, F., Vanderbei, R.J., Wolkowicz, H. (1996): An interior-point method for semidefinite programming. SIAM J. Optim. **6**, 342–361
9. Jarre, F. (1990/91): On the convergence of the method of analytic centers when applied to convex quadratic programs. Math. Program. **49**, 341–359
10. Jarre, F., Kocvara, M., Zowe, J. (1995): Optimal truss topology design by interior point methods. Cooperative Research Report of the Institute of Statistical Mathematics **77**, 236–252
11. Kojima, M., Mizuno, S., Yoshise, A. (1989): A primal-dual interior point algorithm for linear programming. In: Megiddo, N., ed., Progress in Mathematical Programming, Interior-point and Related Methods, Springer-Verlag, New York, pp. 29–47
12. Kojima, M., Mizuno, S., Yoshise, A. (1989): A polynomial-time algorithm for a class of linear complementarity problems. Math. Program. **44**, 1–26
13. Kojima, M., Shindoh, S., Hara, S. (1997): Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices. SIAM J. Optim. **7**, 86–125
14. Lobo, M., Vandenberghe, L., Boyd, S., Lebret, H. (1998): Applications of second-order cone programming. Linear Algebra Appl. **284**, 193–228
15. Mehrotra, S., Sun, J. (1991): A method of analytic centers for quadratically constrained convex quadratic programs. SIAM J. Numer. Anal. **28**, 529–544
16. Mizuno, S., Todd, M.J.,Ye, Y. (1993): On adaptive-step primal-dual interior-point algorithms for linear programming. Math. Oper. Res. **18**, 964–981
17. Monteiro, R.D.C. (1997): Primal-dual path-following algorithms for semidefinite programming. SIAM J. Optim. **7**, 663–678
18. Monteiro, R.D.C. (1998): Polynomial convergence of primal-dual algorithms for semidefinite programming based on Monteiro and Zhang family of directions. SIAM J. Optim. **8**, 797–812
19. Monteiro, R.D.C., Adler, I. (1989): Interior path-following primal-dual algorithms, Part I: Linear programming. Math. Program. **44**, 27–41
20. Monteiro, R.D.C., Adler, I. (1989): Interior path-following primal-dual algorithms, Part II: Convex quadratic programming. Math. Program. **44**, 43–66
21. Monteiro, R.D.C., Tsuchiya, T. (1999): Polynomial convergence of a new family of primal-dual algorithms for semidefinite programming. SIAM J. Optim. **9**, 551–577
22. Monteiro, R.D.C., 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
23. Nesterov, Yu., Nemirovskii, A. (1988): Polynomial barrier methods in convex programming. Ekonomika i Matem. Metody **24**, 1084–1091
24. Nesterov, Yu., Nemirovskii, A (1994): Interior point polynomial methods in convex programming. SIAM Publications, Philadelphia, Pennsylvania, USA
25. Nesterov, Y.E., Todd, M. (1997): Self-scaled barriers and interior-point methods for convex programming. Math. Oper. Res. **22**, 1–42
26. Nesterov, Y.E., Todd, M. (1998): Primal-dual interior-point methods for self-scaled cones. SIAM J. Optim. **8**, 324–364

27. Schmieta, S., Alizadeh, F. (1998): Associative algebras, symmetric cones and polynomial time interior point algorithms. Report RRR 17-98, RUTCOR, Rutgers University, 640 Bartholomew Road, Piscataway, NJ 08854-8003
28. Sturm, J.F. (1998): Similarity and other spectral relations for symmetric cones. Manuscript. Communications Research Laboratory, McMaster University, Hamilton, Canada
29. Tsuchiya, T. (1997): A polynomial primal-dual path-following algorithm for second-order cone programming. Research Memorandum No. 649, The Institute of Statistical Mathematics, Tokyo, Japan, October (Revised: December 1997)
30. Tsuchiya, T. (1999): A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming. Optim. Methods Software **11**&**12**, 141–182
31. Xue, G., Ye, Y. (1997): Efficient algorithms for minimizing a sum of Euclidean norms with applications. SIAM J. Optim. **7**, 1017–1036
32. Zhang, Y. (1998): On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming. SIAM J. Optim. **8**, 365–386