A new Primal-Dual Interior-Point Algorithm for Second-Order Cone Optimization*

Y. Q. Bai † G. Q. Wang† C. Roos‡

November 14, 2004

† *Department of Mathematics, College Science,*
Shanghai University, Shanghai, 200436
‡ *Faculty of Electrical Engineering, Mathematics and Computer Science,*
Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands
e-mail: yqbai@satff.shu.edu.cn, guoq_wang@hotmail.com, c.roos@ewi.tudelft.nl

Abstract

We present a primal-dual interior-point algorithm for second-order conic optimization problems based on a specific class of kernel functions. This class has been investigated earlier for the case of linear optimization problems. In this paper we derive the complexity bounds $O(\sqrt{N}\,(\log N)\,\log\frac{N}{\epsilon})$ for large- and $O(\sqrt{N}\,\log\frac{N}{\epsilon})$ for small- update methods, respectively. Here N denotes the number of second order cones in the problem formulation.

Keywords: second-order conic optimization, interior-point methods, primal-dual method, large- and small-update methods, polynomial complexity.
AMS Subject Classification: **90C22, 90C31**

1 Introduction

Second-order conic optimization (SOCO) problems are convex optimization problems because their objective is a linear function and their feasible set is the intersection of an affine space with the Cartesian product of a finite number of second-order (also called Lorentz or ice−cream) cones. Any second order cone in $\mathbf{R}^n$ has the form

$$K = \left\{ (x_1,\, x_2,\, \dots,\, x_n) \in \mathbf{R}^n \;:\; x_1^2 \geq \sum_{i=2}^{n} x_i^2,\, x_1 \geq 0 \right\}, \tag{1}$$

where n is some natural number. Thus a second-order conic optimization problem has a form

$$(P) \qquad \min\left\{ c^T x \;:\; Ax = b,\, x \in \mathcal{K} \right\},$$

where $\mathcal{K} \subseteq \mathbf{R}^n$ is the Cartesian product of several second-order cones, i.e.,

$$\mathcal{K} = \mathcal{K}^1 \times \mathcal{K}^2 ... \times \mathcal{K}^N, \tag{2}$$

*The first author kindly acknowledges the support of Dutch Organization for Scientific Researches (NWO grant 613.000.110).

with $\mathcal{K}^j \in \mathbf{R}^{n_j}$ for each j, and $n = \sum_{j=1}^{N} n_j$. We partition the vector x accordingly:

$$x = (x^1, x^2, \ldots, x^N)$$

with $x^j \in \mathcal{K}^j$. Furthermore, $A \in R^{m\times n}$, $c \in \mathbf{R}^n$ and $b \in \mathbf{R}^m$.
The dual problem of (P) is given by

$$(D) \qquad \max\left\{b^T y \;:\; A^T y + s = c,\, s \in \mathcal{K}\right\},$$

Without loss of generality we assume that A has full rank: $\operatorname{rank} A = m$. As a consequence, if the pair (y, s) is dual feasible then y is uniquely determined by s. Therefore, we will feel free to say that s is dual feasible, without mentioning y.

It is well-known that SOCO problems include linear and convex quadratic programs as special cases. On the other hand, SOCO problems are special cases of semidefinite optimization (SDO) problems, and hence can be solved by using an algorithm for SDO problems. Interior-point methods (IPMs) that exploit the special structure of SOCO problems, however, have much better complexity than when using an IPM for SDO for solving SOCO problems.

In the last few years the SOCO problem has received considerable attention from researchers because of its wide range of applications (see, e.g., [8, 21]) and because of the existence of efficient IPM algorithms (see, e.g., [2, 3, 17, 16, 18, 19]). Many researchers have studied SOCO and achieved plentiful and beautiful results.

Several IPMs designed for LO (see e.g., [14]) have been successfully extended to SOCO. Important work in this direction was done by Nesterov and Todd [10, 11] who showed that the primal-dual algorithm maintains its theoretical efficiency when the nonnegativity constrains in LO are replaced by a convex cone, as long as the cone is homogeneous and self-dual. Adler and Alizadeh [1] studied a unified primal-dual approach for SDO and SOCO, and proposed a direction for SOCO analogous to the AHO-direction for SDO. Later, Schmieta and Alizadeh [15] presented a way to transfer the Jordan algebra associated with the second-order cone into the so-called Clifford algebra in the cone of matrices and then carried out a unified analysis of the analysis for many IPMs in symmetric cones. Faybusovich [6], using Jordan algebraic techniques, analyzed the Nesterov-Todd method for SOCO. Monteiro [9] and Tsuchiya [19] applied Jordan algebra to the analysis of IPMs for SOCO with specialization to various search directions. Other researchers have worked on IPMs for special cases of SOCO, such as convex quadratic programming, minimizing a sum of norms, ...etc. For an overview of these results we refer to [21] and its related references.

Recently, J. Peng et al. [12] designed primal-dual interior-point algorithms for LO, SDO and SOCO based on so-called self-regular (SR) proximity functions. Moreover, they derived a $O(\sqrt{N}\log N)\log\frac{N}{\epsilon}$ complexity bound for SOCO with large-update methods, the currently best bound for such methods. Their work was extended in [4, 5] to other proximity functions based on univariate so-called kernel functions.

Motivated by [12] and [4, 5], in this paper we present a primal-dual IPM for SOCO problems based on kernel functions of the form

$$\psi(t) := \frac{t^{p+1}-1}{p+1} + \frac{t^{-q}-1}{q},\; t > 0, \tag{3}$$

where $p \in [0, 1]$ and $q > 0$ are the parameters; p is called the growth degree and q the barrier degree of the kernel function.

One may easily verify that $\psi(1) = \psi'(1) = 0$, $\lim_{t\to 0}\psi(t) = \lim_{t\to\infty}\psi(t) = +\infty$ and that $\psi(t)$ is strictly convex. It is worth pointing out that only when $p = 1$ the function $\psi(t)$ is SR. Thus the functions considered here do not belong to the class of functions studied in [12], except if $p = 1$. The same class of kernel functions (3) has been studied for the LO case in [4], and for the SDO [20] in case $p = 0$.

As discussed in [4], every kernel function $\psi(t)$ gives rise to an IPM. We will borrow several tools for the analysis of the algorithm in this paper from [4], and some of them from [12]. These analytic tools reveal that the iteration bound highly depends on the choice of $\psi(t)$, especially on the inverse functions of $\psi(t)$ and its derivatives. Our aim will be to investigate the dependence of the iteration bound on the parameters p and q. We will consider both large- and small-update methods.

The outline of the paper is as follows. In Section 2, after briefly recalling some relevant properties of the second-order cone and its associated Euclidean Jordan algebra, we review some basic concepts for IPMs for solving the SOCO problem, such as central path, NT-search direction, etc. We also present the primal-dual IPM for SOCO considered in this paper at the end of Section 2. In Section 3, we study the properties of the kernel function $\psi(t)$ and its related vector-valued barrier function and real-valued barrier function. The step size and the resulting decrease of the barrier function are discussed in Section 4. In Section 5, we analyze the algorithm to derive the complexity bound for large- and small-update methods. Finally, some concluding remarks follow in Section 6.

Some notations used throughout the paper are as follows. $\mathbf{R}^n$, $\mathbf{R}^n_+$ and $\mathbf{R}^n_{++}$ denote the set of all vectors (with n components), the set of nonnegative vectors and the set of positive vectors, respectively. As usual, $\|\,.\,\|$ denotes the Frobenius norm for matrices, and the 2-norm for vectors. The Löwner partial ordering "$\succeq$" of $\mathbf{R}^n$ defined by a second-order cone K is defined by $x \succeq_K s$ if $x - s \in K$. The interior of K is denoted as K_+ and we write $x \succ_K s$ if $x - s \in K_+$. Finally, $\mathbf{E}_n$ denotes the $n \times n$ identity matrix.

2 Preliminaries

2.1 Algebraic properties of second-order cones

In this section we briefly recall some algebraic properties of the second-order cone K as defined by (1) and its associated Euclidean Jordan algebra. Our main sources for this section are [1, 7, 9, 12, 15, 18].

If, for any two vectors $x, s \in \mathbf{R}^n$, the bilinear operator $\circ$ is defined by[1]

$$x \circ s := (x^T s;\ x_1 s_2 + s_1 x_2;\ \ldots;\ x_1 s_n + s_1 x_n),$$

then $(\mathbf{R}^n, \circ)$ is a commutative Jordan algebra. Note that the map $s \mapsto x \circ s$ is linear. The matrix of this linear map is denoted as $L(x)$, and one may easily verify that it is an arrow-shaped matrix:

$$L(x) := \begin{bmatrix} x_1 & x_{2:n}^T \\ x_{2:n} & x_1 \mathbf{E}_{n-1} \end{bmatrix}, \tag{4}$$

[1]We use here and elsewhere Matlab notation: $(u; v)$ is the column vector obtained by concatenating the column vectors u and v.

where $x_{2:n} = (x_2;\ \ldots;\ x_n)$. We have the following five properties.

1. $x \in K$ if and only if $L(x)$ is positive semidefinite;
2. $x \circ s = L(x)s = L(s)x = s \circ x$;
3. $x \in K$ if and only if $x = y \circ y$ for some $y \in K$.
4. $e \circ x = x$ for all $x \in K$, where $e = (1;\ 0;\ \ldots;\ 0)$;
5. if $x \in K_+$ then $L(x)^{-1}e = x^{-1}$, i.e., $x \circ L(x)^{-1}e = e$.

The first property implies that $x \mapsto L(x)$ provides a natural embedding of K into the cone of positive semidefinite matrices, which makes SOCO essentially a specific case of SDO. From the third property we see that K is just the set of squares in the Jordan algebra. Due to the fourth property, the vector e is the (unique) unit element of the Jordan algebra.

Let us point out that the cone K is not closed under the product '$\circ$' and that this product is not associative.

The maximal and minimal eigenvalues of $L(x)$ are denoted as $\lambda_{\max}(x)$ and $\lambda_{\min}(x)$, respectively. These are given by

$$\lambda_{\max}(x) := x_1 + \|x_{2:n}\|, \quad \lambda_{\min}(x) := x_1 - \|x_{2:n}\|. \tag{5}$$

It readily follows that

$$x \in K \Leftrightarrow \lambda_{\min}(x) \geq 0, \qquad x \in K_+ \Leftrightarrow \lambda_{\min}(x) > 0.$$

Lemma 2.1 *If $x \in \mathbf{R}^n$, then $|\lambda_{\max}(x)| \leq \sqrt{2}\,\|x\|$ and $|\lambda_{\min}(x)| \leq \sqrt{2}\,\|x\|$.*

Proof: By (5), we have $\lambda^2_{\max}(x) + \lambda^2_{\min}(x) = 2\|x\|^2$, which implies the lemma. □

The trace and the determinant of $x \in \mathbf{R}^n$ are defined by

$$\mathbf{Tr}(x) := \lambda_{\max}(x) + \lambda_{\min}(x) = 2x_1, \quad \mathbf{det}(x) := \lambda_{\max}(x)\lambda_{\min}(x) = x_1^2 - \|x_{2:n}\|^2. \tag{6}$$

From the above definition, for any $x, s \in K$, it is obvious that

$$\mathbf{Tr}(x \circ s) = 2x^T s, \quad \mathbf{Tr}(x \circ x) = 2\,\|x\|^2.$$

It is also straightforward to verify that[2]

$$\mathbf{Tr}((x \circ s) \circ t) = \mathbf{Tr}(x \circ (s \circ t)). \tag{7}$$

Lemma 2.2 *Let $x,\ s \in \mathbf{R}^n$. Then*

$$\lambda_{\min}(x + s) \geq \lambda_{\min}(x) - \sqrt{2}\,\|s\|.$$

[2]Recall that the Jordan product itself is not associative. But one may easily verify that the first coordinate of $(x \circ s) \circ t$ equals

$$x_1 s_1 t_1 + \sum_{i=2}^{n} (x_1 s_i t_i + s_1 x_i t_i + t_1 x_i s_i) = \left(x^T s + x^T t + s^T t\right) - 2x_1 s_1 t_1,$$

which is invariant under permutations of x, s and t, and hence (7) follows.

Proof: By using the well-known triangle inequality we obtain

$$\lambda_{\min}(x+s) = x_1 + s_1 - \|(x+s)_{2:n}\| \geq x_1 + s_1 - \|x_{2:n}\| - \|s_{2:n}\| = \lambda_{\min}(x) + \lambda_{\min}(s).$$

By Lemma 2.1, we have $|\lambda_{\min}(s)| \leq \sqrt{2}\,\|s\|$. This implies the lemma. □

Lemma 2.3 (Lemma 6.2.3 in [12]) *Let $x,\ s \in \mathbf{R}^n$. Then*

$$\lambda_{\max}(x)\lambda_{\min}(s) + \lambda_{\min}(x)\lambda_{\max}(s) \leq \mathbf{Tr}(x \circ s) \leq \lambda_{\max}(x)\lambda_{\max}(s) + \lambda_{\min}(x)\lambda_{\min}(s),$$

$$\mathbf{det}\,(x \circ s) \leq \mathbf{det}\,(x)\,\mathbf{det}\,(s)\,.$$

The last inequality holds with equality if and only if the vectors $x_{2:n}$ and $s_{2:n}$ are linearly dependent.

Corollary 2.4 *Let $x \in \mathbf{R}^n$ and $s \in K$. Then*

$$\lambda_{\min}(x)\,\mathbf{Tr}(s) \leq \mathbf{Tr}(x \circ s) \leq \lambda_{\max}(x)\,\mathbf{Tr}(s).$$

Proof: Since $s \in K$, we have $\lambda_{\min}(s) \geq 0$, and hence also $\lambda_{\max}(s) \geq 0$. Now Lemma 2.3 implies that

$$\lambda_{\min}(x)\left(\lambda_{\max}(s) + \lambda_{\min}(s)\right) \leq \mathbf{Tr}(x \circ s) \leq \lambda_{\max}(x)\left(\lambda_{\max}(s) + \lambda_{\min}(s)\right).$$

Since $\lambda_{\max}(s) + \lambda_{\min}(s) = \mathbf{Tr}(s)$, this implies the lemma. □

We conclude this section by introducing the so-called spectral decomposition of a vector $x \in \mathbf{R}^n$. This is given by

$$x := \lambda_{\max}(x)\,z^1 + \lambda_{\min}(x)\,z^2, \tag{8}$$

where

$$z^1 := \frac{1}{2}\left(1;\ \frac{x_{2:n}}{\|x_{2:n}\|}\right), \quad z^2 := \frac{1}{2}\left(1;\ \frac{-x_{2:n}}{\|x_{2:n}\|}\right).$$

Here by convention $x_{2:n}/\|x_{2:n}\| = 0$ if $x_{2:n} = 0$. Note that the vectors z_1 and z_2 belong to K (but not to K_+). The importance of this decomposition is that it enables us to extend the definition of any function $\psi : \mathbf{R}_{++} \to \mathbf{R}_+$ to a function that maps the interior of K into K. In particular this holds for our kernel function $\psi(t)$, as given by (3).

Definition 2.5 *Let $\psi : \mathbf{R} \to \mathbf{R}$ and $x \in \mathbf{R}^n$. With z_1 and z_2 as defined in (8), we define*

$$\psi(x) := \psi(\lambda_{\max}(x))\,z^1 + \psi(\lambda_{\min}(x))\,z^2.$$

In other words,

$$\psi(x) = \begin{cases} \left(\dfrac{\psi(\lambda_{\max}(x)) + \psi(\lambda_{\min}(x))}{2};\ \dfrac{\psi(\lambda_{\max}(x)) - \psi(\lambda_{\min}(x))}{2}\dfrac{x_{2:n}}{\|x_{2:n}\|}\right) & \text{if } x_{2:n} \neq 0 \\ (\psi(\lambda_{\max}(x));\ 0;\ \ldots;\ 0) & \text{if } x_{2:n} = 0. \end{cases}$$

Obviously, $\psi(x)$ is a well-defined vector-valued function. In the sequel we use $\psi(\cdot)$ both to denote a vector function (if the argument is a vector) and a univariate function (if the argument is a scalar). The above definition respects the Jordan product in the sense described by the following lemma.

Lemma 2.6 (Lemma 6.2.5 in [12]) *Let ψ_1 and ψ_2 be two functions $\psi(t) = \psi_1(t)\psi_2(t)$ for $t \in \mathbf{R}$. Then one has*

$$\psi(x) = \psi_1(x) \circ \psi_2(x), \quad x \in K.$$

The above lemma has important consequences. For example, for $\psi(t) = t^{-1}$ then $\psi(x) = x^{-1}$, where x^{-1} is the inverse of x in the Jordan algebra (if it exists); if $\psi(t) = t^2$ then $\psi(x) = x \circ x$, etc.

Lemma 2.7 *Let $\psi : \mathbf{R}_{++} \to \mathbf{R}_+$ and $x \in K_+$. Then $\psi(x) \in K$.*

Proof: Since $x \in K_+$, its eigenvalues are positive. Hence $\psi(\lambda_{\max}(x))$ and $\psi(\lambda_{\min}(x))$ are well-defined and nonnegative. Since z_1 and z_2 belong to K, also $\psi(\lambda_{\max}(x))\, z^1 + \psi(\lambda_{\min}(x))\, z^2 \in K$. □

If $\psi(t)$ is twice differentiable, like our kernel function, then the derivatives $\psi'(t)$ and $\psi''(t)$ exist for $t > 0$, and we also have vector-valued functions $\psi'(x)$ and $\psi''(x)$, namely:

$$\psi'(x) = \psi'(\lambda_{\max}(x))\, z^1 + \psi'(\lambda_{\min}(x))\, z^2,$$
$$\psi''(x) = \psi''(\lambda_{\max}(x))\, z^1 + \psi''(\lambda_{\min}(x))\, z^2.$$

If $x_{2:n} \neq 0$ then z_1 and z_2 satisfy $\|z_1\| = \|z_2\| = 1/\sqrt{2}$ and $z_1^T z_2 = 0$. From this one easily deduces that

$$\|\psi(x)\| = \frac{1}{\sqrt{2}} \sqrt{\psi(\lambda_{\max}(x))^2 + \psi(\lambda_{\min}(x))^2}. \tag{9}$$

Note that (9) also holds if $x_{2:n} = 0$, because then $\lambda_{\max}(x) = \lambda_{\min}(x) = x_1$, whence $\|\psi(x)\| = |x_1|$ and $\psi(\lambda_{\max}(x))^2 + \psi(\lambda_{\min}(x))^2 = 2x_1^2$.

From now we assume that $\psi(t)$ is the kernel function (3). We associate to $\psi(t)$ a real-valued barrier function $\Psi(x)$ on K_+ as follows.

$$\Psi(x) := \mathbf{Tr}(\psi(x)) = 2(\psi(x))_1 = \psi(\lambda_{\max}(x)) + \psi(\lambda_{\min}(x)), \quad x \in K_+. \tag{10}$$

Obviously, since $\psi(t) \geq 0$ for all $t > 0$, and $\lambda_{\max}(x) \geq \lambda_{\min}(x) > 0$, we have $\Psi(x) \geq 0$ for all $x \in K_+$. Moreover, since $\psi(t) = 0$ if and only if $t = 1$, we have $\Psi(x) = 0$ if and only if $\lambda_{\max}(x) = \lambda_{\min}(x) = 1$, which occurs if and only if $x = e$. One may also verify that $\psi(e) = 0$ and that e is the only vector with this property.

Similarly, one easily verifies that $\psi'(x) = 0$ holds if and only if $\psi(\lambda_{\max}(x)) + \psi(\lambda_{\min}(x)) = 0$ and $\psi(\lambda_{\max}(x)) - \psi(\lambda_{\min}(x)) = 0$. This is equivalent to $\psi(\lambda_{\max}(x)) = \psi(\lambda_{\min}(x)) = 0$, which holds if and only if $\lambda_{\max}(x) = \lambda_{\min}(x) = 1$, i.e., if and only if $x = e$. Summarizing, we have

$$\Psi(x) = 0 \quad \Leftrightarrow \quad \psi(x) = 0 \quad \Leftrightarrow \quad \psi'(x) = 0 \quad \Leftrightarrow \quad x = e. \tag{11}$$

We recall the following lemma without proof.

Lemma 2.8 (cf. Proposition 6.2.9 in [12]) *Since $\psi(t)$ is strictly convex for $t > 0$, $\Psi(x)$ is strictly convex for $x \in K_+$.*

In the analysis of our algorithm, that will be presented below, we need to consider derivatives with respect to a real parameter t of the functions $\psi'(x(t))$ and $\Psi(x(t))$, where $x(t) =$

$(x_1(t);\ x_2(t);\ \ldots;\ x_n(t))$. The usual concepts of continuity, differentiability and integrability can be naturally extended to vectors of functions, by interpreting them entry-wise. Denoting

$$x'(t) = (x'_1(t);\ x'_2(t);\ \ldots;\ x'_n(t)),$$

we then have

$$\frac{d}{dt}(x(t) \circ s(t)) = x'(t) \circ s(t) + x(t) \circ s'(t), \tag{12}$$

$$\frac{d}{dt}\mathbf{Tr}(\psi(x(t))) = 2\frac{d\,\psi(x(t))}{dt} = 2\psi'(x(t))^T x'(t) = \mathbf{Tr}(\psi'(x(t)) \circ x'(t)). \tag{13}$$

2.2 Re-scaling the cone

When defining the search direction in our algorithm, we need a re-scaling of the space in which the cone lives. Given x^0, $s^0 \in K_+$, we use an automorphism $W(x^0, s^0)$ of the cone K such that

$$W(x^0, s^0)\, x^0 = W(x^0, s^0)^{-1} s^0.$$

The existence and uniqueness of such an automorphism is well-known. To make the paper self-containing, a simple derivation of this result is given in the Appendix. In the sequel we denote $W(x^0, s^0)$ simply as W. Let us point out that the matrix W is symmetric and positive definite. For any x, $s \in \mathbf{R}^n$ we define

$$\tilde{x} := Wx, \quad \tilde{s} := W^{-1}s.$$

We call this Nesterov-Todd (NT)-scaling of $\mathbf{R}^n$, after the inventors. In the following lemma we recall several properties of the NT-scaling scheme.

Lemma 2.9 (cf. Proposition 6.3.3 in [12]) *For any x, $s \in \mathbf{R}^n$ one has*

(i) $\mathbf{Tr}(\tilde{x} \circ \tilde{s}) = \mathbf{Tr}(x \circ s)$;

(ii) $\mathbf{det}\,(\tilde{x}) = \lambda\,\mathbf{det}\,(x)\,,\ \mathbf{det}\,(\tilde{s}) = \lambda^{-1}\mathbf{det}\,(s),$ *where* $\lambda = \sqrt{\frac{\mathbf{det}(s^0)}{\mathbf{det}(x^0)}}$;

(iii) $x \succeq_K 0,\ (x \succ_K 0) \Leftrightarrow \tilde{x} \succeq_K 0,\ (\tilde{x} \succ_K 0)$.

Proof: The proof of (i) is straightforward:

$$\mathbf{Tr}(\tilde{x} \circ \tilde{s}) = \mathbf{Tr}\left(Wx \circ W^{-1}s\right) = 2\,(Wx)^T \left(W^{-1}s\right) = 2x^T W^T W^{-1} s = 2x^T s = \mathbf{Tr}(x \circ s).$$

For the proof of (ii) we need the matrix

$$Q = \mathrm{diag}\,(1,\ -1,\ \ldots,\ -1) \in \mathbf{R}^{n \times n}. \tag{14}$$

Obviously, $Q^2 = \mathbf{E}_n$ where $\mathbf{E}_n$ denotes the identity matrix of size $n \times n$. Moreover, $\mathbf{det}\,(x) = x^T Q x$, for any x. By Lemma A.2 and Lemma A.4 we have $WQW = \lambda W$. Hence we may write

$$\mathbf{det}\,(\tilde{x}) = \tilde{x}^T Q \tilde{x} = (Wx)^T Q\,(Wx) = x^T WQWx = \lambda\, x^T Q x = \lambda\,\mathbf{det}\,(x)\,.$$

In a similar way we can prove $\mathbf{det}\,(\tilde{s}) = \lambda^{-1}\mathbf{det}\,(s)$. Finally, since W is an automorphism of K, we have $x \in K$ if and only $\tilde{x} \in K$; since $\mathbf{det}\,(x) > 0$ if and only $\mathbf{det}\,(\tilde{x}) > 0$, by (ii), also $x \in K_+$ if and only $\tilde{x} \in K_+$. □

2.3 The central path for SOCO

To simplify the presentation in this and the first part of the next section, for the moment we assume that $N = 1$ in (2). So $\mathcal{K}$ is itself a second-order cone, which we denote as K.

We assume that both (P) and (D) satisfy the interior-point condition (IPC), i.e., there exists (x^0, y^0, s^0) such that $Ax^0 = b$, $x^0 \in K_+$, $A^T y^0 + s^0 = c$, $s^0 \in K_+$. Assuming the IPC, it is well-known that the optimality conditions for the pair of problems (P) and (D) are

$$\begin{aligned} Ax &= b, \quad x \in K, \\ A^T y + s &= c, \quad s \in K, \\ L(x)\, s &= 0. \end{aligned} \tag{15}$$

The basic idea of primal-dual IPMs is to replace the third equation in (15) by the parameterized equation $L(x)\, s = \mu e$ with $\mu > 0$. Thus we consider the following system

$$\begin{aligned} Ax &= b, \quad x \in K, \\ A^T y + s &= c, \quad s \in K, \\ L(x)\, s &= \mu e. \end{aligned} \tag{16}$$

For each $\mu > 0$ the parameterized system (16) has a unique solution $(x(\mu), y(\mu), s(\mu))$ and we call $x(\mu)$ the μ-center of (P) and $(y(\mu), s(\mu))$ the μ-center of (D). Note that at the μ-center we have

$$x(\mu)^T s(\mu) = \tfrac{1}{2}\mathbf{Tr}(x \circ s) = \tfrac{1}{2}\mathbf{Tr}(L(x)\, s) = \frac{1}{2}\mathbf{Tr}(\mu e) = \mu. \tag{17}$$

The set of μ-center gives a homotopy path, which is called the central path. If $\mu \to 0$ then the limit of the central path exists and since the limit points satisfy the complementarity condition $L(x)\, s = 0$, it naturally yields optimal solution for both (P) and (D) (see, e.g., [21]).

2.4 New search direction for SOCO

IPMs follow the central path approximately and find an approximate solution of the underlying problems (P) and (D) as μ goes to zero. The search directions is usually derived from a certain Newton-type system. We first want to point out that a straightforward approach to obtain such a system fails to define unique search directions.

For the moment we assume $N = 1$. By linearizing system (16) we obtain the following linear system for the search directions.

$$\begin{aligned} A\Delta x &= 0, \\ A^T \Delta y + \Delta s &= 0, \\ L(x)\Delta s + L(s)\Delta x &= \mu e - L(x)s. \end{aligned} \tag{18}$$

This system has a unique solution if and only if the matrix $A\, L(s)^{-1} L(x) A^T$ is nonsingular as one may easily verify. Unfortunately this might not be the case, even if A has full rank. This is due to the fact that the matrix $Z = L(s)^{-1} L(x)$, which has positive eigenvalues, may not be symmetric. As a consequence, the symmetrized matrix $Z + Z^T$ may have negative eigenvalues. For a discussion of this phenomenon and an example we refer to [12, page 143].

The above discussion makes clear that we need some symmetrizing scheme. There exist many such schemes. See, e.g., [1, 13, 18]. In this paper we follow the approach taken in [12] and we choose for the NT-scaling scheme, as introduced in Section 2.2. This choice can be justified by recalling that until now large-update IMPs based on the NT search direction have the best known theoretical iteration bound.

Thus, letting W denote the (unique) automorphism of K that satisfies $Wx = W^{-1}s$ we define

$$v := \frac{Wx}{\sqrt{\mu}} \quad \left(= \frac{W^{-1}s}{\sqrt{\mu}} \right) \tag{19}$$

and

$$\bar{A} := \frac{1}{\sqrt{\mu}} A W^{-1}, \quad d_x := \frac{W\Delta x}{\sqrt{\mu}}, \quad d_s := \frac{W^{-1}\Delta s}{\sqrt{\mu}}. \tag{20}$$

Lemma 2.10 *One has*

$$\mu^2 \, \mathbf{det}\left(v^2\right) = \mathbf{det}\,(x)\ \mathbf{det}\,(s)\,, \qquad \mu\,\mathbf{Tr}(v^2) = \mathbf{Tr}(x \circ s).$$

Proof: Defining $\tilde{x} = Wx$ and $\tilde{s} = W^{-1}s$, Lemma 2.9 yields $\mathbf{det}\,(\tilde{x}) = \sqrt{\mathbf{det}\,(x)\,\mathbf{det}\,(s)}$, which implies $\mu\,\mathbf{det}\,(v) = \sqrt{\mathbf{det}\,(x)\,\mathbf{det}\,(s)}$. Since $(\mathbf{det}\,(v))^2 = \mathbf{det}\,(v \circ v)$, by Lemma 2.3, the first equality follows. The proof of the second equality goes in the same way. □

The system (18) can be rewritten as follows:

$$\begin{aligned} \bar{A} d_x &= 0, \\ \bar{A}^T \Delta y + d_s &= 0, \\ L(W^{-1}v) W d_s + L(Wv) W^{-1} d_x &= e - L(W^{-1}v) W v. \end{aligned}$$

Since this system is equivalent to (18), it may not have a unique solution. To overcome this difficulty we replace the last equation by

$$L(v)\, d_s + L(v)\, d_x = e - L(v)\, v,$$

which is equivalent to

$$d_s + d_x = L(v)^{-1} e - v = v^{-1} - v.$$

Thus the system defining the scaled search directions becomes

$$\begin{aligned} \bar{A} d_x &= 0, \\ \bar{A}^T \Delta y + d_s &= 0, \\ d_s + d_x &= v^{-1} - v. \end{aligned} \tag{21}$$

Since the matrix $\bar{A}^T \bar{A}$ is positive definite, this system has a unique solution.

We just outlined the approach to obtain the uniquely defined classical search direction for SOCO. The approach in this paper differs only in one detail: we replace the right hand side in the last equation by $-\psi'(v)$, where $\psi(t)$ is the kernel function given by (3). Thus we will use the following system to define our search direction:

$$\begin{aligned} \bar{A} d_x &= 0, \\ \bar{A}^T \Delta y + d_s &= 0, \\ d_s + d_x &= -\psi'(v). \end{aligned} \tag{22}$$

Since (22) has the same matrix of coefficients as (21), also (22) has a unique solution.[3] Recall from (11) that

$$\Psi(v)=0 \quad\Leftrightarrow\quad \psi(v)=0 \quad\Leftrightarrow\quad \psi'(v)=0 \quad\Leftrightarrow\quad v=e.$$

Since d_x and d_s are orthogonal, we will have $d_x = d_s = 0$ in (22) if and only if $\psi'(v)=0$, i.e., if and only if $v=e$. The latter implies $x\circ s=\mu e$, which means that $x=x(\mu)$ and $s=s(\mu)$. This is the content of the next lemma.

Lemma 2.11 *One has $\psi'(v)=0$ if and only if $x\circ s=\mu e$.*

Proof: We just established that $\psi'(v)=0$ if and only if $v=e$. According to (19), $v=e$ holds if and only if $x=\sqrt{\mu}\,W^{-1}e$ and $s=\sqrt{\mu}\,We$. By Lemma A.4 the matrix W has the form

$$W=\sqrt{\lambda}\begin{bmatrix} a_1 & \bar{a}^T \\ \bar{a} & \mathbf{E}_{n-1}+\frac{\bar{a}\bar{a}^T}{1+a_1}\end{bmatrix}$$

where $a=(a_1;\bar{a})$ is a vector such that $\mathbf{det}\,(a)=1$ and $\lambda>0$. Using that $W^{-1}=\frac{1}{\sqrt{\lambda}}W_{Qa}$, with Q as defined in (14), so $Qa == (a_1;-\bar{a})$ (see Lemma A.3), it follows that $v=e$ holds if and only if

$$x=\frac{\sqrt{\mu}}{\sqrt{\lambda}}\begin{bmatrix} a_1 \\ -\bar{a}\end{bmatrix}, \qquad s=\sqrt{\mu}\sqrt{\lambda}\begin{bmatrix} a_1 \\ \bar{a}\end{bmatrix}.$$

If x and s have this form then

$$x\circ s=\frac{\sqrt{\mu}}{\sqrt{\lambda}}\sqrt{\mu}\sqrt{\lambda}\begin{bmatrix} a_1^2-\bar{a}^T\bar{a} \\ a_1\bar{a}+a_1(-\bar{a})\end{bmatrix}=\mu\begin{bmatrix}\mathbf{det}\,(a) \\ 0\end{bmatrix}=\mu\, e,$$

and conversely, if $x\circ s=\mu e$ then there must exist a $\lambda>0$ and a vector a such that x and s are of the above form. This proves the lemma. □

We conclude from Lemma 2.11 that if $(x,\,y,\,s)\neq(x(\mu),\,y(\mu),\,s(\mu))$ then $\psi'(v)\neq 0$ and hence $(\Delta x,\,\Delta y,\,\Delta s)$ is nonzero.

We proceed by adapting the above definitions and those in the previous sections to the case where $N>1$, when the cone underlying the given problems (P) and (D) is the cartesian product of N cones $\mathcal{K}^j$, as given in (2). First we partition any vector $x\in\mathbf{R}^n$ according to the dimensions of the successive cones $\mathcal{K}^j$, so

$$x=\left(x^1;\;\ldots;\;x^N\right),\qquad x^j\in\mathbf{R}^{n_j},$$

and we define the algebra $(\mathbf{R}^n,\diamond)$ as a direct product of Jordan algebras:

$$x\diamond s:=\left(x^1\circ s^1;\;x^2\circ s^2;\;\ldots;\;x^N\circ s^N\right).$$

[3] It may be worth mentioning that if we use the kernel function of the classical logarithmic barrier function, i.e., $\psi(t)=\frac{1}{2}(t^2-1)-\log t$, then $\psi'(t)=t-t^{-1}$, whence $-\psi'(v)=v^{-1}-v$, and hence system (22) then coincides with the classical system (21).

Obviously, if $e^j \in \mathcal{K}^j$ is the unit element in the Jordan algebra for the j-th cone, then the vector

$$e = (e^1;\ e^2;\ \ldots;\ e^N)$$

is the unit element in $(\mathbf{R}^n, \diamond)$.

The NT-scaling scheme in this general case is now obtained as follows. Let $x, s \in \mathcal{K}_+$, and for each j, let W^j denote the (unique) automorphism of $\mathcal{K}^j$ that satisfies $W^j x^j = \left(W^j\right)^{-1} s^j$ and

$$\lambda^j = \sqrt{\frac{\mathbf{det}\,(s^j)}{\mathbf{det}\,(x^j)}}.$$

Let us denote

$$W := \mathrm{diag}\,(W^1,\ \ldots,\ W^N).$$

Obviously, $W = W^T$ and W is positive definite. The matrix W can be used to re-scale x and s to the same vector v, as defined in (19). Note that we then may write

$$v = \left(v^1,\ \ldots,\ v^N\right), \quad v^j := \frac{W^j x^j}{\sqrt{\mu}} \quad \left(= \frac{{W^j}^{-1} s^j}{\sqrt{\mu}} \right),$$

Modifying the definitions of $\psi(x)$ and $\Psi(x)$ as follows,

$$\psi(v) = (\psi(v^1);\ \psi(v^2);\ \ldots;\ \psi(v^N)), \quad \Psi(v) = \sum_{j=1}^{N} \Psi(v^j), \tag{23}$$

and defining the matrix $\bar{A}$ as in (20) the scaled search directions are then defined by the system (22), with

$$\psi'(v) = (\psi'(v^1);\ \psi'(v^2);\ \ldots;\ \psi'(v^N)).$$

By transforming back to the x- and s-space, respectively, using (20), we obtain search directions Δx, Δy and Δs in the original spaces, with

$$\Delta x = \sqrt{\mu}\, W^{-1} d_x, \quad \Delta s = \sqrt{\mu}\, W d_s.$$

By taking a step along the search direction, with a step size α defined by some line search rule, that will be specified later on, we construct a new triple (x, y, s) according to

$$\begin{aligned} x^+ &= x + \alpha \Delta x, \\ y^+ &= y + \alpha \Delta y, \\ s^+ &= s + \alpha \Delta s. \end{aligned} \tag{24}$$

2.5 Primal-dual interior-point algorithm for SOCO

Before presenting our algorithm, we define a barrier function in terms of the original variables x and s, and the barrier parameter μ, according to

$$\Phi(x, s, \mu) := \Psi(v).$$

The algorithm now is as presented in Figure 1. It is clear from this description that closeness

Primal-Dual Algorithm for SOCO

Input:
A threshold parameter $\tau > 1$;
an accuracy parameter $\varepsilon > 0$;
a fixed barrier update parameter $\theta \in (0, 1)$;
a strictly feasible pair (x^0, s^0) and $\mu^0 > 0$ such that $\Psi(x^0, s^0, \mu^0) \leq \tau$.
begin
  $x := \mathbf{x^0}$; $s := \mathbf{s^0}$; $\mu := \mu^0$;
  while $N\mu \geq \varepsilon$ **do**
  begin
    $\mu := (1 - \theta)\mu$;
    while $\Phi(x, s, \mu) > \tau$ **do**
    begin
      $x := x + \alpha\Delta x$;
      $s := s + \alpha\Delta s$;
      $y := y + \alpha\Delta y$;
    end
  end
end

Figure 1: Algorithm

of (x, y, s) to $(x(\mu), y(\mu), s(\mu))$ is measured by the value of $\Psi(v)$, with τ as a threshold value: if $\Psi(v) \leq \tau$ then we start a new *outer iteration* by performing a μ-update, otherwise we enter an *inner iteration* by computing the search directions at the current iterates with respect to the current value of μ and apply (24) to get new iterates. The parameters τ, θ and the step size α should be chosen in such a way that the algorithm is 'optimized' in the sense that the number of iterations required by the algorithm is as small as possible. The choice of the so-called barrier update parameter θ plays an important role both in theory and practice of IPMs. Usually, if θ is a constant independent of the dimension n of the problem, for instance $\theta = \frac{1}{2}$, then we call the algorithm a *large-update* (or *long-step*) method. If θ depends on the dimension of the problem, such as $\theta = \frac{1}{\sqrt{n}}$, then the algorithm is named a *small-update* (or *short-step*) method. The choice of the step size α $(0 \leq \alpha \leq 1)$ is another crucial issue in the analysis of the algorithm. It has to be taken such that the closeness of the iterates to the current μ-center improves by a sufficient amount. In the theoretical analysis the step size α is usually given a value that depends on the closeness of the current iterates to the μ-center. Note that at the μ-center the duality gap equals $N\mu$, according to (17). The algorithm stops if the duality gap at the μ-center is less than ε.

For the analysis of this algorithm we need to derive some properties of our (scaled) barrier function $\Psi(v)$. This is the subject of the next section.

3 Properties of the barrier function $\Psi(v)$

3.1 Properties of the kernel function $\psi(t)$

First we derive some properties of $\psi(t)$ as given by (3). The first three derivatives of $\psi(t)$ are

$$\psi'(t) = t^p - \frac{1}{t^{q+1}}, \tag{25}$$

$$\psi''(t) = pt^{p-1} + \frac{q+1}{t^{q+2}} > 0, \tag{26}$$

$$\psi'''(t) = p(p-1)t^{p-2} - \frac{(q+1)(q+2)}{t^{q+3}} < 0. \tag{27}$$

As mentioned before $\psi(t)$ is strictly convex and $\psi''(t)$ is monotonically decreasing for $t \in (0, \infty)$. The following lemmas list some properties of $\psi(t)$ that are needed for analysis of the algorithm. When a proof is known in the literature we simply give a reference and omit the proof.

Lemma 3.1 *If $t_1 > 0,\ t_2 > 0$, then one has*

$$\psi(\sqrt{t_1 t_2}) \leq \frac{1}{2}(\psi(t_1) + \psi(t_2)).$$

Proof: By Lemma 2.1.2 in [12], we know that the property in the lemma holds if and only if $t\psi''(t) + \psi'(t) \geq 0$, whenever $t > 0$. Using (25), one has

$$t\psi''(t) + \psi'(t) = pt^p + \frac{q+1}{t^{q+1}} + t^p - \frac{1}{t^{q+1}} = (p+1)t^p + \frac{q}{t^{q+1}} > 0.$$

This implies the lemma. □

Lemma 3.2 *If $t \geq 1$, then*

$$\psi(t) \leq \frac{p+q+1}{2}(t-1)^2.$$

Proof: By using the Taylor expansion around $t = 1$ and using $\psi'''(t) < 0$ we obtain the inequality in the lemma, since $\psi(1) = \psi'(1) = 0$ and $\psi''(1) = p + q + 1$. □

Lemma 3.3 *If $q \geq 1 - p$ and $t \geq 1$ then*

$$t \leq 1 + \sqrt{t\psi(t)}.$$

Proof: Defining $f(t) := t\psi(t) - (t-1)^2$, one has $f(1) = 0$ and $f'(t) = \psi(t) + t\psi'(t) - 2(t-1)$. Hence $f'(1) = 0$ and

$$f''(t) = 2\psi'(t) + t\psi''(t) - 2 = (p+2)\ t^p + \frac{q-1}{t^{q+1}} - 2 \geq pt^p + \frac{q-1}{t^{q+1}} \geq p\left(t^p - \frac{1}{t^{q+1}}\right) \geq 0.$$

Thus we obtain

$$t\psi(t) \geq (t-1)^2. \tag{28}$$

Since $t-1 \geq 0$, this implies the lemma. □

In the sequel $\varrho : [0,\infty) \to [1,\infty)$ will denote the inverse function of $\psi(t)$ for $t \geq 1$ and $\rho : [0,\infty) \to (0,1]$ the inverse function of $-\frac{1}{2}\psi'(t)$ for $t \in (0,1]$. Thus we have

$$\begin{aligned} \varrho(s) = t \quad &\Leftrightarrow \quad \psi(t) \quad = s, \qquad s \geq 0,\, t \geq 1, \\ \rho(s) = t \quad &\Leftrightarrow \quad -\psi'(t) = 2s, \qquad s \geq 0,\, 0 < t \leq 1. \end{aligned} \tag{29}$$

Lemma 3.4 *For each $q > 0$ one has*

$$((p+1)s+1)^{\frac{1}{p+1}} \leq \varrho(s) \leq \left((p+1)s + \frac{p+q+1}{q}\right)^{\frac{1}{p+1}},\ s \geq 0, \tag{30}$$

Moreover, if $q \geq 1-p$ then

$$\varrho(s) \leq 1 + s^{\frac{1}{2}} \left((p+1)\,s + \frac{p+q+1}{q}\right)^{\frac{1}{2(p+1)}},\ s \geq 0. \tag{31}$$

Proof: Let $\varrho\,(s) = t$, $t \geq 1$. Then $s = \psi(t)$. Hence, using $t \geq 1$,

$$\frac{t^{p+1}-1}{p+1} - \frac{1}{q} \leq s = \frac{t^{p+1}-1}{p+1} + \frac{t^{-q}-1}{q} \leq \frac{t^{p+1}-1}{p+1}.$$

The left inequality gives

$$\varrho\,(s) = t \leq \left(1 + (1+p)\left(s + \frac{1}{q}\right)\right)^{\frac{1}{p+1}} = \left((p+1)s + \frac{p+q+1}{q}\right)^{\frac{1}{p+1}} \tag{32}$$

and the right inequality

$$\varrho\,(s) = t \geq (1 + (1+p)s)^{\frac{1}{p+1}}\,,$$

proving (28).

Now we turn to the case that $q \geq 1-p$. By Lemma 3.3 we have $t \leq 1 + \sqrt{t\psi(t)} = 1 + \sqrt{ts}$. Substituting the upper bound for t given by (32) we obtain (31). The proof of the lemma is now completed. □

Remark 3.5 *Note that at $s = 0$ the upper bound (31) for $\varrho\,(s)$ is tighter than in (28). This will be important later on when we deal with the complexity of small-update methods.*

Lemma 3.6 (Theorem 4.9 in [4]) *For any positive vector $z \in \mathbf{R}^n$ one has*

$$\sqrt{\sum_{i=1}^{n} \left(\psi'(z_i)\right)^2} \geq \psi'\left(\varrho\left(\sum_{i=1}^{n} \psi(z_i)\right)\right).$$

Equality holds if and only if $z_i \neq 1$ for at most 1 coordinate i and $z_i \geq 1$.

Lemma 3.7 (Theorem 3.2 in [4]) *For any positive vector $z \in \mathbf{R}^n$ and any $\beta \geq 1$ we have*

$$\sum_{i=1}^{n} \psi(\beta z_i) \leq n\psi\left(\beta\varrho\left(\frac{1}{n}\sum_{i=1}^{n}\psi(z_i)\right)\right).$$

Equality holds if and only if there exists a scalar $z \geq 1$ such that $z_i = z$ for all i.

Lemma 3.8 *One has*

$$\rho(s) \geq \frac{1}{(2s+1)^{\frac{1}{q+1}}}, \quad s > 0.$$

Proof: Let $\rho(s) = t$, for some $s \geq 0$. Then $s = -\frac{1}{2}\psi'(t)$, with $t \in (0, 1]$. Using (25) we may write

$$2s = \frac{1}{t^{q+1}} - t^p \geq \frac{1}{t^{q+1}} - 1,$$

which implies

$$\rho(s) = t \geq \frac{1}{(2s+1)^{\frac{1}{q+1}}},$$

proving the lemma. □

3.2 Properties of the barrier function $\Psi(v)$

The following lemma is a consequence of Lemma 3.1.

Lemma 3.9 (Proposition 6.2.9 in [12]) *If $x, s,$ and $v \in \mathcal{K}_+$ satisfy*

$$\mathbf{det}\left(v^2\right) = \mathbf{det}\,(x)\;\mathbf{det}\,(s), \quad \mathbf{Tr}(v^2) = \mathbf{Tr}(x \circ s),$$

then

$$\Psi(v) \leq \frac{1}{2}(\Psi(x) + \Psi(s)).$$

Below we use the following norm-based proximity $\delta(v)$ for v.

$$\delta(v) = \frac{1}{\sqrt{2}}\sqrt{\sum_{i=1}^{N} \|\psi'(v^j)\|^2}. \tag{33}$$

The following lemma gives a lower bound for $\delta(v)$ in terms of $\Psi(v)$.

Lemma 3.10 *If $v \in \mathcal{K}_+$, then*

$$\delta(v) \geq \frac{1}{2}\frac{(1+p)\Psi(v)}{(1+(1+p)\Psi(v))^{\frac{1}{1+p}}}.$$

Proof: Due to (33) and (9) we have

$$\delta(v) = \frac{1}{2}\sqrt{\sum_{i=1}^{N}((\psi'(\lambda_{\max}(v^j)))^2 + (\psi'(\lambda_{\min}(v^j))))^2)}$$

and, due to (23) and (10),

$$\Psi(v) = \sum_{j=1}^{N}\left(\psi(\lambda_{\max}(v^j)) + \psi(\lambda_{\min}(v^j))\right).$$

This makes clear that $\delta(v)$ and $\Psi(v)$ depend only on the eigenvalues $\lambda_{\max}(v^j)$ and $\lambda_{\min}(v^j)$ of the vectors v^j, for $j = 1 : N$. This observation makes it possible to apply Lemma 3.6, with z being the vector in $\mathbf{R}^{2N}$ consisting of all the eigenvalues. This gives

$$\delta(v) \geq \frac{1}{2}\psi'(\varrho\,(\Psi(v))).$$

Since $\psi'(t)$ is monotonically increasing for $t \geq 1$, we may replace $\varrho\,(\Psi(v))$ by a smaller value. Thus, by using the left hand side inequality in 30 in Lemma 3.4 we find

$$\delta(v) \geq \frac{1}{2}\psi'(((p+1)\Psi(v)+1)^{\frac{1}{p+1}}).$$

Finally, using (25) and $q > 0$, we obtain

$$\begin{aligned}\delta(v) &\geq \frac{1}{2}\left(((p+1)\Psi(v)+1)^{\frac{p}{p+1}} - \frac{1}{((p+1)\Psi(v)+1)^{\frac{q+1}{p+1}}}\right)\\ &\geq \frac{1}{2}\left(((p+1)\Psi(v)+1)^{\frac{p}{p+1}} - \frac{1}{((p+1)\Psi(v)+1)^{\frac{1}{p+1}}}\right) = \frac{1}{2}\frac{(p+1)\Psi(v)}{((p+1)\Psi(v)+1)^{\frac{1}{p+1}}}.\end{aligned}$$

This proves the lemma. □

Lemma 3.11 *If $v \in K_+$ and $\beta \geq 1$, then*

$$\Psi(\beta v) \leq 2N\psi\left(\beta\varrho\left(\frac{\Psi(v)}{2N}\right)\right).$$

Proof: Due to (10), and (5), we have

$$\Psi(\beta v) = \sum_{j=1}^{N}\left(\psi(\beta\lambda_{\max}(v^j)) + \psi(\beta\lambda_{\min}(v^j))\right).$$

As in the previous lemma, the variables are essentially only the eigenvalues $\lambda_{\max}(v^j)$ and $\lambda_{\min}(v^j)$ of the vectors v^j, for $j = 1 : N$. Applying Lemma 3.7, with z being the vector in $\mathbf{R}^{2N}$ consisting of all the eigenvalues, the inequality in the lemma immediately follows. □

Corollary 3.12 *If* $\Psi(v) \leq \tau$ *and* $v_+ = \dfrac{v}{\sqrt{1-\theta}}$, *with* $0 \leq \theta \leq 1$, *then one has*

$$\Psi(v_+) \leq 2N\psi\left(\frac{\varrho\left(\frac{\tau}{2N}\right)}{\sqrt{1-\theta}}\right).$$

Proof: By taking $\beta = 1/\sqrt{1-\theta}$ in Lemma 3.11, and using that $\varrho(s)$ is monotonically increasing, this immediately follows. □

As we will show in the next section, each subsequent inner iteration will give rise to a decrease of the value of $\Psi(v)$. Hence during the course of the algorithm the largest values of $\Psi(v)$ occur after the μ-updates. Therefore, Corollary 3.12 provides a uniform upper bound for $\Psi(v)$ during the execution of the algorithm. As a consequence, by using the upper bounds for $\varrho(s)$ in (28) and (31) we find that the numbers L_1 and L_2 introduced below are upper bounds for $\Psi(v)$ during the course of the algorithm.

$$L_1 := 2N\psi\left(\frac{\left((p+1)\frac{\tau}{2N} + \frac{p+q+1}{q}\right)^{\frac{1}{p+1}}}{\sqrt{1-\theta}}\right), \quad q > 0, \tag{34}$$

$$L_2 := 2N\psi\left(\frac{1 + \sqrt{\frac{\tau}{2N}}\left((p+1)\left(\frac{\tau}{2N}\right) + \frac{p+q+1}{q}\right)^{\frac{1}{2(p+1)}}}{\sqrt{1-\theta}}\right), \quad q \geq 1-p. \tag{35}$$

4 Analysis of complexity bound for SOCO

4.1 Estimate of decrease of $\Psi(v)$ during an inner iteration

In each iteration we start with a primal-dual pair (x, s) and then get the new iterates $x^+ = x + \alpha\Delta x$ and $s^+ = s + \alpha\Delta s$ from (24). The step size α is strictly feasible if and only if $x + \alpha\Delta x \in \mathcal{K}_+$ and $s + \alpha\Delta s \in \mathcal{K}_+$, i.e., if and only if $x^j + \alpha\Delta x^j \in \mathcal{K}_+^j$ and $s^j + \alpha\Delta s^j \in \mathcal{K}_+^j$ for each j.

Let W^j denote the automorphism of $\mathcal{K}^j$ that satisfies $W^j x^j = {W^j}^{-1} s^j$ and $v^j = W^j x^j/\sqrt{\mu}$. Then we have

$$W^j(x^j + \alpha\Delta x^j) = \sqrt{\mu}\,\left(v^j + \alpha d_x^j\right)$$
$${W^j}^{-1}(s^j + \alpha\Delta s^j) = \sqrt{\mu}\,\left(v^j + \alpha d_s^j\right).$$

Since W^j is an automorphism of $\mathcal{K}^j$, we conclude that step size α is strictly feasible if and only if $v^j + \alpha d_x^j \in \mathcal{K}^j$ and $v^j + \alpha d_s^j \in \mathcal{K}^j$, for each j. Using Lemma 2.9, with $\tilde{x} = \sqrt{\mu}\,\left(v^j + \alpha d_x^j\right)$ and $\tilde{s} = \sqrt{\mu}\,\left(v^j + \alpha d_s^j\right)$, we obtain

$$\mu^2\,\mathbf{det}\left(v^j + \alpha d_x^j\right)\mathbf{det}\left(v^j + \alpha d_s^j\right) = \mathbf{det}\left(x^j + \alpha\Delta x^j\right)\mathbf{det}\left(s^j + \alpha\Delta s^j\right)$$
$$\mu\,\mathbf{Tr}\left(\left(v^j + \alpha d_x^j\right)\circ\left(v^j + \alpha d_s^j\right)\right) = \mathbf{Tr}\left(\left(x^j + \alpha\Delta x^j\right)\circ\left(s^j + \alpha\Delta s^j\right)\right).$$

On the other hand, if W_+^j is the automorphism that satisfies $W_+^j x^{+j} = {W_+^j}^{-1} s^{+j}$ and $v_+^j = W_+^j x^{+j}/\sqrt{\mu}$, then, by Lemma 2.10 we have

$$\mu^2 \,\mathbf{det}\left((v_+^j)^2\right) = \mathbf{det}\left(x^j + \alpha\Delta x^j\right)\,\mathbf{det}\left(s^j + \alpha\Delta s^j\right)$$

$$\mu\,\mathbf{Tr}\left((v_+^j)^2\right) = \mathbf{Tr}(\left(x^j + \alpha\Delta x^j\right) \circ \left(s^j + \alpha\Delta s^j\right)).$$

We conclude from this that, for each j,

$$\mathbf{det}\left((v_+^j)^2\right) = \mathbf{det}\left(v^j + \alpha d_x^j\right)\mathbf{det}\left(v^j + \alpha d_s^j\right)$$

$$\mathbf{Tr}\left((v_+^j)^2\right) = \mathbf{Tr}\left(\left(v^j + \alpha d_x^j\right) \circ \left(v^j + \alpha d_s^j\right)\right).$$

By Lemma 3.9 this implies that, for each j,

$$\Psi(v_+^j) \leq \frac{1}{2}\left(\Psi(v^j + \alpha d_x^j) + \Psi(v^j + \alpha d_s^j)\right).$$

Taking the sum over all j, $1 \leq j \leq N$, we get

$$\Psi(v_+) \leq \frac{1}{2}\left(\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s)\right).$$

Hence, denoting the decrease in $\Psi(v)$ during an inner iteration as

$$f(\alpha) := \Psi(v_+) - \Psi(v).$$

we have

$$f(\alpha) \leq f_1(\alpha) := \frac{1}{2}\left(\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s)\right) - \Psi(v).$$

One can easily verify that $f(0) = f_1(0) = 0$.

In the sequel we derive an upper bound for the function $f_1(\alpha)$ in stead of $f(\alpha)$, thereby taking advantage of the fact that $f_1(\alpha)$ is convex ($f(\alpha)$ may be not convex!). We start by considering the first and second derivatives of $f_1(\alpha)$. Using (13), (12) and (7), one can easily verify that

$$f_1'(\alpha) = \frac{1}{2}\mathbf{Tr}(\psi'(v + \alpha d_x) \diamond d_x + \psi'(v + \alpha d_s) \diamond d_s) \tag{36}$$

and

$$f_1''(\alpha) = \frac{1}{2}\mathbf{Tr}((d_x \diamond d_x) \diamond \psi''(v + \alpha d_x) + (d_s \diamond d_s) \diamond \psi''(v + \alpha d_s)) \tag{37}$$

Using (36), the third equality in system (22), and (33), we find

$$f_1'(0) = \frac{1}{2}\mathbf{Tr}(\psi'(v) \diamond (d_x + d_s)) = -\frac{1}{2}\mathbf{Tr}(\psi'(v) \diamond \psi'(v)) = -2\delta(v)^2. \tag{38}$$

To facilitate the analysis below, we define

$$\lambda_{\max}(v) = \max\left\{\lambda_{\max}(v^j)\ :\ 1 \leq j \leq N\right\}, \text{ and }\ \lambda_{\min}(v) = \min\left\{\lambda_{\min}(v^j)\ :\ 1 \leq j \leq N\right\}.$$

Furthermore, we denote $\delta(v)$ simply as δ.

Lemma 4.1 *One has*

$$f_1''(\alpha) \leq 2\delta^2 \psi''(\lambda_{\min}(v) - 2\alpha\delta).$$

Proof: Since d_x and d_s are orthogonal, and $d_x + d_s = -\psi'(v)$, we have

$$\|d_x + d_s\|^2 = \|d_x\|^2 + \|d_s\|^2 = 2\delta^2,$$

whence $\|d_x\| \leq \delta\sqrt{2}$ and $\|d_s\| \leq \delta\sqrt{2}$. Hence, by Lemma 2.2, we have

$$\lambda_{\max}(v + \alpha d_x) \geq \lambda_{\min}(v + \alpha d_x) \geq \lambda_{\min}(v) - 2\alpha\delta,$$
$$\lambda_{\max}(v + \alpha d_s) \geq \lambda_{\min}(v + \alpha d_s) \geq \lambda_{\min}(v) - 2\alpha\delta.$$

By (26) and (27), $\psi''(v)$ is positive and monotonically decreasing, respectively. We thus obtain

$$0 < \psi''(\lambda_{\max}(v + \alpha d_x)) \leq \psi''(\lambda_{\min}(v) - 2\alpha\delta),$$
$$0 < \psi''(\lambda_{\max}(v + \alpha d_s)) \leq \psi''(\lambda_{\min}(v) - 2\alpha\delta).$$

Applying Corollary 2.4, and using that $\mathbf{Tr}(z \circ z) = 2z^T z = 2\|z\|^2$, for any z, we obtain

$$\mathbf{Tr}\left((d_x \diamond d_x) \diamond \psi''(v + \alpha d_x)\right) \leq \psi''(\lambda_{\min}(v) - 2\alpha\delta)\, \mathbf{Tr}(d_x \diamond d_x) = 2\psi''(\lambda_{\min}(v) - 2\alpha\delta)\, \|d_x\|^2,$$
$$\mathbf{Tr}\left((d_x \diamond d_x) \diamond \psi''(v + \alpha d_s)\right) \leq \psi''(\lambda_{\min}(v) - 2\alpha\delta)\, \mathbf{Tr}(d_s \diamond d_s) = 2\psi''(\lambda_{\min}(v) - 2\alpha\delta)\, \|d_s\|^2.$$

Substitution into (37) yields that

$$f_1''(\alpha) \leq \psi''(\lambda_{\min}(v) - 2\alpha\delta)\, \left(\|d_x\|^2 + \|d_s\|^2\right) = 2\delta^2 \psi''(\lambda_{\min}(v) - 2\alpha\delta).$$

This prove the lemma. □

Our aim is to find the step size α for which $f_1(\alpha)$ is minimal. The last lemma is exactly the same as Lemma 4.1 in [4] (with $v_1 := \lambda_{\min}(v)$), where we dealt with the case of linear optimization. As a consequence, as we use below the same arguments as in [4], we present several lemmas without repeating their proofs, and simply refer the reader to the corresponding result in [4].

4.2 Selection of step size α

In this section we find a suitable default step size for the algorithm. Recall that α should be chosen such that x_+ and s_+ are strictly feasible and such that $\Psi(v_+) - \Psi(v)$ decreases sufficiently. Let us also recall that $f_1(\alpha)$ is strictly convex, and $f_1(0) < 0$ (due to (38), assuming that $\delta > 0$) whereas $f_1(\alpha)$ approaches infinity if the new iterates approach the boundary of the feasible region. Hence, the new iterates will certainly be strictly feasible if $f_1(\alpha) \leq 0$. Ideally we should find the unique α that minimizes $f_1(\alpha)$. Our default step size will be an approximation for this ideal step size.

Lemma 4.2 (Lemma 4.2 in [4]) $f_1'(\alpha) \leq 0$ *certainly holds if the step size* α *satisfies*

$$-\psi'(\lambda_{\min}(v) - 2\alpha\delta) + \psi'(\lambda_{\min}(v)) \leq 2\delta.$$

Lemma 4.3 (Lemma 4.3 in [4]) *The largest possible the step size α satisfying the inequality in Lemma 4.2 is given by*

$$\bar{\alpha} := \frac{\rho(\delta) - \rho(2\delta)}{2\delta},$$

with ρ as defined in (29).

Lemma 4.4 (Lemma 4.4 in [4]) *With ρ and $\bar{\alpha}$ as in Lemma 4.3, one has*

$$\bar{\alpha} \geq \frac{1}{\psi''(\rho(2\delta))}.$$

Since $\psi''(t)$ is monotonically decreasing, Lemma 4.4 implies, when using also Lemma 3.8,

$$\bar{\alpha} \geq \frac{1}{\psi''(\rho(2\delta))} \geq \frac{1}{\psi''\left((1+4\delta)^{\frac{-1}{q+1}}\right)}.$$

Using (26) we obtain

$$\bar{\alpha} \geq \frac{1}{p(1+4\delta)^{\frac{1-p}{q+1}} + (q+1)(1+4\delta)^{\frac{q+2}{q+1}}} \geq \frac{1}{(p+q+1)(1+4\delta)^{\frac{q+2}{q+1}}}.$$

We define our default step size as $\tilde{\alpha}$ by the last expression:

$$\tilde{\alpha} := \frac{1}{(p+q+1)(1+4\delta)^{\frac{q+2}{q+1}}}. \tag{39}$$

Lemma 4.5 (Lemma 4.5 in [4]) *If the step size α is such that $\alpha \leq \bar{\alpha}$, then*

$$f(\alpha) \leq -\alpha\delta^2.$$

The last lemma implies that

$$f(\tilde{\alpha}) \leq -\frac{\delta^2}{(p+q+1)(1+4\delta)^{\frac{q+2}{q+1}}}. \tag{40}$$

According to the algorithm, at the start of each inner iteration we have $\Psi(v) > \tau$. So, assuming $\tau \geq 1$, and using Lemma 3.10, we conclude that

$$\delta(v) \geq \frac{1}{2}\frac{(1+p)\Psi(v)}{(1+(1+p)\Psi(v))^{\frac{1}{1+p}}} \geq \frac{1}{2}\frac{\Psi(v)}{(1+2\Psi(v))^{\frac{1}{1+p}}} \geq \frac{1}{2}\frac{\Psi(v)}{(3\Psi(v))^{\frac{1}{1+p}}} \geq \frac{1}{6}\Psi(v))^{\frac{p}{1+p}}.$$

Since the right hand side expression in (40) is monotonically decreasing in δ we obtain

$$f(\tilde{\alpha}) \leq -\frac{\Psi^{\frac{2p}{1+p}}}{36(p+q+1)\left(1+\frac{2}{3}\Psi^{\frac{p}{p+1}}\right)^{\frac{q+2}{q+1}}} \leq -\frac{\Psi^{\frac{2p}{1+p}}}{36(p+q+1)\left(\frac{5}{3}\Psi^{\frac{p}{p+1}}\right)^{\frac{q+2}{q+1}}} \leq -\frac{\Psi^{\frac{pq}{(q+1)(p+1)}}}{100(p+q+1)},$$

where we omitted the argument v in $\Psi(v)$.

5 Iteration bounds

5.1 Complexity of large-update methods

We need to count how many inner iterations are required to return to the situation where $\Psi(v) \leq \tau$. We denote the value of $\Psi(v)$ after the μ-update as Ψ_0 and the subsequent values in the same outer iteration as Ψ_i, $i = 1, 2, ..., T$, where T denotes the total number of inner iterations in the outer iteration.

Recall that $\Psi_0 \leq L_1$, with L_1 as given by (34). Since $\psi(t) \leq t^{p+1}/(p+1)$ for $t \geq 1$ and $0 \leq p \leq 1$, we obtain

$$\Psi_0 \leq 2N\psi\left(\frac{\left(\frac{\tau}{N}+\frac{q+2}{q}\right)^{\frac{1}{p+1}}}{\sqrt{1-\theta}}\right) \leq \frac{2\left(\tau+\frac{q+2}{q}N\right)}{(p+1)\,(1-\theta)^{\frac{p+1}{2}}}.$$

According to the expression for $f(\tilde{\alpha})$, we have

$$\Psi_{i+1} \leq \Psi_i - \beta(\Psi_i)^{1-\gamma}, \quad i = 0, 1, ..., T-1, \tag{41}$$

where $\beta = \frac{1}{100(p+q+1)}$ and $\gamma = \frac{p+q+1}{(q+1)(p+1)}$.

Lemma 5.1 (Proposition 2.2 in [12]) *If $t_0, t_1, \cdots, t_T$ is a sequence of positive numbers and*

$$t_{i+1} \leq t_i - \beta t_i^{1-\gamma}, \qquad i = 0, 1, \cdots, T-1,$$

with $\beta > 0$ and $0 < \gamma \leq 1$, then $T \leq \frac{t_0^\gamma}{\beta\gamma}$.

Hence, due to this lemma we obtain

$$T \leq 100(p+1)(q+1)\Psi_0^{\frac{p+q+1}{(q+1)(p+1)}} \leq 100(p+1)(q+1)\left(\frac{2\left(\tau+\frac{q+2}{q}N\right)}{(p+1)\,(1-\theta)^{\frac{p+1}{2}}}\right)^{\frac{p+q+1}{(q+1)(p+1)}}. \tag{42}$$

The number of outer iterations is bounded above by $\frac{1}{\theta}\log\frac{N}{\epsilon}$, as one easily may verify. By multiplying the number of outer iterations and the upper bound for the number T of inner iterations, we obtain an upper bound for the total number of iterations, namely,

$$\frac{100(p+1)(q+1)}{\theta\,(1-\theta)^{\frac{p+q+1}{2(q+1)}}}\left(\frac{2\left(\tau+\frac{q+2}{q}N\right)}{p+1}\right)^{\frac{p+q+1}{(q+1)(p+1)}}\log\frac{N}{\epsilon}. \tag{43}$$

Large-update methods have $\theta = \Theta(1)$, and $\tau = O(N)$. Note that our bound behaves bad for small values of q, i.e. if q approaches 0. Assuming $q \geq 1$, the iteration bound becomes

$$O\left(qN^{\frac{p+q+1}{(q+1)(p+1)}}\right)\log\frac{N}{\epsilon}.$$

It is interesting to consider the special case where $\tau = \frac{q-2}{q}$. Then, when omitting expressions that are $\Theta(1)$ and the factor $\log\frac{N}{\epsilon}$, the bound becomes

$$(p+1)(q+1)\,(4N)^{\frac{p+q+1}{(q+1)(p+1)}}. \tag{44}$$

Given q, the value for p that minimizes the above expression turns out to be

$$p = \frac{q \log(4N)}{q+1} - 1,$$

which in the cases of interest lies outside the interval $[0, 1]$. Thus, let us take $p = 1$. Then the best value of q is given by

$$q = \log \sqrt{4N} - 1,$$

and this value of q reduces (44) to

$$2 \log \sqrt{4N}\, (4N)^{\frac{1+\log \sqrt{4N}}{2 \log \sqrt{4N}}} = \sqrt{4N}(\log(4N))\,(4N)^{\frac{1}{\log(4N)}} = \exp(1)\sqrt{4N}\,(\log(4N)).$$

Thus we may conclude that for a suitable choice of the parameters p and q the iteration bound for large-update methods is

$$O\left(\sqrt{4N}\,(\log(4N))\, \log \frac{N}{\epsilon}\right).$$

Note that this bound closely resembles the bound obtained in [4] for the kernel functions $\psi_3(t)$ and $\psi_7(t)$ in that paper.

5.2 Complexity of small-update methods

Small-update methods are characterized by $\theta = \Theta(1/\sqrt{N})$ and $\tau = O(1)$. The exponent of the expression between brackets in the iteration bound (43) is always larger than $\frac{1}{2}$. Hence, if $\theta = \Theta(1/\sqrt{N})$, the exponent of N in the bound will be at least 1. For small-update methods we can do better, however. For this we need the upper bound L_2 for $\Psi(v)$, as given by (35). Recall that this bound is valid only if $q \geq 1 - p$.

Thus, using also Lemma 3.2, we find

$$\Psi_0 \leq (p+q+1)\, N \left(\frac{1 + \sqrt{\frac{\tau}{2N}} \left(\frac{(p+1)\tau}{2N} + \frac{p+q+1}{q} \right)^{\frac{1}{2(p+1)}}}{\sqrt{1-\theta}} - 1 \right)^2.$$

Since $1 - \sqrt{(1-\theta)} \leq \theta$, this can be simplified to

$$\Psi_0 \leq \frac{p+q+1}{1-\theta} N \left(\theta + \sqrt{\frac{\tau}{2N}} \left(\frac{(p+1)\tau}{2N} + \frac{p+q+1}{q} \right)^{\frac{1}{2(p+1)}} \right)^2.$$

which, using $p \geq 0$, can be further simplified to

$$\Psi_0 \leq \frac{p+q+1}{2(1-\theta)} \left(\theta\sqrt{2N} + \sqrt{\tau \left(\frac{(p+1)\tau}{2N} + \frac{p+q+1}{q} \right)} \right)^2.$$

By using the first inequality in (42) we find that the total number of iterations per outer iteration is bounded above by

$$100(p+1)(q+1) \left(\frac{p+q+1}{2(1-\theta)} \left(\theta\sqrt{2N} + \sqrt{\tau \left(\frac{(p+1)\tau}{2N} + \frac{p+q+1}{q} \right)} \right)^2 \right)^{\frac{p+q+1}{(q+1)(p+1)}}.$$

Hence, using $\frac{p+q+1}{(q+1)(p+1)} \leq 1$, we find that the total number of iterations is bounded above by

$$\frac{50(p+1)(q+1)(p+q+1)}{\theta(1-\theta)}\left(\theta\sqrt{2N}+\sqrt{\tau\left(\frac{(p+1)\tau}{2N}+\frac{p+q+1}{q}\right)}\right)^2\log\frac{N}{\epsilon}.$$

Taking, e.g., $p=0$, $q=1$, $\tau=1$ and $\theta=1/\sqrt{2N}$, this yields the iteration bound

$$\frac{200}{\theta(1-\theta)}\left(1+\sqrt{\frac{1}{2N}+2}\right)^2\log\frac{N}{\epsilon}\leq\frac{1333}{\theta(1-\theta)}\log\frac{N}{\epsilon},$$

which is $O\left(\sqrt{N}\log\frac{N}{\varepsilon}\right)$.

6 Conclusions and remarks

We presented a primal-dual interior-point algorithm for SOCO based on a specific kernel function and derived the complexity analysis for the corresponding algorithm, both with large- and small-update methods. Besides, we developed some new analysis tools for SOCO. Our analysis is a relatively simple and straightforward extension of analogous results for LO.

Some interesting topics for further research remain. First, the search directions used in this paper are based on the NT-symmetrization scheme. It may be possible to design similar algorithms using other symmetrization schemes and still obtain polynomial-time iteration bounds. Finally, numerical tests should be performed to investigate the computational behavior of the algorithms in comparison with other existing approaches.

Acknowledgement

The authors thank Manual Vieirra for some comments on an earlier version of the paper.

A Some technical lemmas

Lemma A.1 *For any vector $a \in \mathbf{R}^n$ one has*

$$\left(\mathbf{E}_n + aa^T\right)^{\frac{1}{2}} = \mathbf{E}_n + \frac{aa^T}{1+\sqrt{1+a^Ta}}.$$

Proof: Let $A = \mathbf{E}_n + aa^T$. The eigenvalues of A are $1+a^Ta$ (multiplicity 1) and 1 (multiplicity $n-1$). Hence, if $e_2, \ldots, e_n$ is any orthonormal basis of the orthocomplement of a, then

$$A = \left(1+a^Ta\right)\frac{aa^T}{a^Ta} + \sum_{i=2}^n e_i.$$

Therefore,

$$A^{\frac{1}{2}} = \sqrt{1+a^Ta}\,\frac{aa^T}{a^Ta} + \sum_{i=2}^n e_i = \sqrt{1+a^Ta}\,\frac{aa^T}{a^Ta} + A - \left(1+a^Ta\right)\frac{aa^T}{a^Ta} = \mathbf{E}_n + \left(\sqrt{1+a^Ta}-1\right)\frac{aa^T}{a^Ta}.$$

Multiplication of the numerator and denominator in the last term with $1+\sqrt{1+a^Ta}$ yields the lemma. $\square$

Obviously, we have for every $x \in K$ that $Qx \in K$, and $Q(K) = K$. This means that Q is an automorphism of K. Without repeating the proof we recall the following lemma.

Lemma A.2 (Proposition 3 in [9]) *The automorphism group of the cone K consists of all positive definite matrices W such that $WQW = \lambda Q$ for some $\lambda > 0$.*

To any nonzero $a = (a_1; \bar{a}) \in K$ we associate the matrix

$$W_a = \begin{bmatrix} a_1 & \bar{a}^T \\ \bar{a} & \mathbf{E}_{n-1} + \frac{\bar{a}\bar{a}^T}{1+a_1} \end{bmatrix} = -Q + \frac{(e+a)(e+a)^T}{1+a_1}.$$

Note that $W_e = \mathbf{E}_n$ and $W_a e = a$. Moreover, the matrix W_a is positive definite. The latter can be made clear as follows. By Schur's lemma W_a is positive definite if and only if

$$\mathbf{E}_{n-1} + \frac{\bar{a}\bar{a}^T}{1+a_1} - \frac{\bar{a}\bar{a}^T}{a_1} \succ 0.$$

This is equivalent to

$$a_1(1+a_1)\mathbf{E}_{n-1} - \bar{a}\bar{a}^T \succ 0,$$

which certainly holds if $a_1^2 \mathbf{E}_{n-1} - \bar{a}\bar{a}^T \succeq 0$. This holds if and only if for every $\bar{x} \in \mathbf{R}^{n-1}$ one has $a_1^2 \|\bar{x}\|^2 \geq (\bar{a}^T\bar{x})^2$. By the Cauchy-Schwartz inequality the latter is true since $a_1 \geq \|\bar{a}\|$.

Lemma A.3 *Let the matrix $W \succ 0$ be such that $WQW = \lambda Q$ for some $\lambda > 0$. Then there exists $a = (a_1; \bar{a}) \in K$ with $\det(a) = 1$ such that $W = \sqrt{\lambda}\, W_a$. Moreover, $W^{-1} = \frac{1}{\sqrt{\lambda}} W_{Qa}$.*

Proof: First assume $\lambda = 1$. Since $Q^2 = \mathbf{E}_n$, $WQW = Q$ implies $QWQ = W^{-1}$. Without loss of generality we may assume that W has the form

$$W = \begin{bmatrix} a_1 & \bar{a}^T \\ \bar{a} & C \end{bmatrix},$$

where C denotes a symmetric matrix. Thus we obtain

$$W^{-1} = QWQ = \begin{bmatrix} a_1 & -\bar{a}^T \\ -\bar{a} & C \end{bmatrix}.$$

This holds if and only if

$$a_1^2 - \bar{a}^T\bar{a} = 1, \quad a_1\bar{a} - C\bar{a} = 0, \quad C^2 - \bar{a}\bar{a}^T = \mathbf{E}_{n-1}. \tag{45}$$

By Lemma A.1, the last equation implies

$$C = \mathbf{E}_{n-1} + \frac{\bar{a}\bar{a}^T}{1+\sqrt{1+\bar{a}^T\bar{a}}} = \mathbf{E}_{n-1} + \frac{\bar{a}\bar{a}^T}{1+a_1},$$

where the last equality follows from the first equation in (45); here we used that $W \succ 0$, which implies $a_1 > 0$. One may easily verify that this C also satisfies the second equation in (45), i.e., $C\bar{a} = a_1\bar{a}$. Note that $\det(a) = a_1^2 - \bar{a}^T\bar{a} = 1$. This completes the proof if $\lambda = 1$. If $\lambda \neq 1$, then multiplication of W_a by $\sqrt{\lambda}$ yields the desired expression. □

Lemma A.4 *Let $x, s \in K_+$. Then there exists a unique automorphism W of K such that $Wx = W^{-1}s$. Using the notation of Lemma A.3, this automorphism is given by $W = \sqrt{\lambda}\, W_a$, where*

$$a = \frac{s + \lambda\, Qx}{\sqrt{2\lambda}\sqrt{x^T s + \sqrt{\mathbf{det}\,(x)\,\mathbf{det}\,(s)}}}, \qquad \lambda = \sqrt{\frac{\mathbf{det}\,(s)}{\mathbf{det}\,(x)}}.$$

Proof: By Lemma A.3 every automorphism of K has the form $W = \sqrt{\lambda}\, W_a$, with $\lambda > 0$ and $a = (a_1; \bar{a}) \in K$ with $\det(a) = 1$. Since $Wx = W^{-1}s$ holds if and only if $W^2 x = s$, we need to find λ and a such that $W^2 x = s$. Some straightforward calculations yield

$$W^2 = \lambda \begin{bmatrix} 2a_1^2 - 1 & 2a_1\bar{a}^T \\ 2a_1\bar{a} & \mathbf{E}_{n-1} + 2\bar{a}\bar{a}^T \end{bmatrix}.$$

So, denoting $x = (x_1;\ \bar{x})$ and $s = (s_1;\ \bar{s})$, we need to find λ and a such that

$$\lambda \begin{bmatrix} 2a_1^2 - 1 & 2a_1\bar{a}^T \\ 2a_1\bar{a} & \mathbf{E}_{n-1} + 2\bar{a}\bar{a}^T \end{bmatrix} \begin{bmatrix} x_1 \\ \bar{x} \end{bmatrix} = \begin{bmatrix} s_1 \\ \bar{s} \end{bmatrix}.$$

This is equivalent to

$$\begin{aligned} (2a_1^2 - 1)\, x_1 + 2a_1\, \bar{a}^T\bar{x} &= \frac{s_1}{\lambda} \\ 2\, a_1 x_1\, \bar{a} + \bar{x} + 2(\bar{a}^T\bar{x})\, \bar{a} &= \frac{\bar{s}}{\lambda}. \end{aligned}$$

Using $a_1 x_1 + \bar{a}^T\bar{x} = a^T x$, this can be written as

$$\begin{aligned} 2\,(a^T x)\, a_1 &= \frac{s_1}{\lambda} + x_1 \\ 2\,(a^T x)\, \bar{a} &= \frac{\bar{s}}{\lambda} - \bar{x}. \end{aligned}$$

Thus we find

$$a = \begin{bmatrix} a_1 \\ \bar{a} \end{bmatrix} = \frac{1}{2\, a^T x} \begin{bmatrix} \frac{s_1}{\lambda} + x_1 \\ \frac{\bar{s}}{\lambda} - \bar{x} \end{bmatrix} = \frac{s + \lambda Qx}{2\,\lambda\, a^T x}. \tag{46}$$

Now observe that what we have shown so far is that if $W = \sqrt{\lambda}\, W_a$ and $W^2 x = s$ then a satisfies (46). Hence, since $W^{-1} = \frac{1}{\sqrt{\lambda}} W_{Qa}$ and $W^{-2}s = x$, we conclude that Qa can be expressed as follows.

$$Qa = \frac{x + \frac{1}{\lambda}Qs}{2\,\frac{1}{\lambda}\,(Qa)^T s} = \frac{\lambda\, x + Qs}{2\, a^T Qs}.$$

Multiplying both sides with Q, while using that $Q^2 = \mathbf{E}_n$, we obtain

$$a = \frac{s + \lambda Q x}{2\, a^T Q s}. \tag{47}$$

Comparing this with the expression for a in (46) we conclude that

$$\lambda\, a^T x = a^T Q s. \tag{48}$$

Taking the inner product of both sides in (47) with x and Qs, respectively, we get

$$a^T x = \frac{x^T s + \lambda x^T Q x}{2\, a^T Q s}, \qquad a^T Q s = \frac{s^T Q s + \lambda x^T s}{2\, a^T Q s}.$$

Since $\lambda\, a^T x = a^T Q s$, this implies

$$\lambda \left(x^T s + \lambda x^T Q x \right) = s^T Q s + \lambda x^T s,$$

which gives

$$\lambda^2 x^T Q x = s^T Q s.$$

Observing that $x^T Q x = \mathbf{det}\,(x)$ (and $s^T Q s = \mathbf{det}\,(s)$), and $\lambda > 0$, we obtain

$$\lambda = \sqrt{\frac{\mathbf{det}\,(s)}{\mathbf{det}\,(x)}}. \tag{49}$$

Now using that $\det(a) = a^T Q a = 1$ we derive from (47) that

$$\frac{\left(s + \lambda Q x\right)^T Q \left(s + \lambda Q x\right)}{4 \left(a^T Q s\right)^2} = 1.$$

This implies

$$\lambda^2 x^T Q x + s^T Q s + 2 \lambda x^T s = 4 \left(a^T Q s\right)^2,$$

or, equivalently,

$$\lambda^2 \mathbf{det}\,(x) + \mathbf{det}\,(s) + 2 \lambda x^T s = 4 \left(a^T Q s\right)^2.$$

Note that the expression at the left is positive, because $x \in K_+$ and $s \in K_+$. Because $a \in K_+$ and $Qs \in K_+$ also $a^T Q s > 0$. Substituting the value of λ we obtain

$$2 \mathbf{det}\,(s) + 2 \sqrt{\frac{\mathbf{det}\,(s)}{\mathbf{det}\,(x)}}\, x^T s = 4 \left(a^T Q s\right)^2,$$

whence

$$2\, a^T Q s = \sqrt{2} \sqrt{\mathbf{det}\,(s) + \sqrt{\frac{\mathbf{det}\,(s)}{\mathbf{det}\,(x)}}\, x^T s} = \sqrt{2} \sqrt[4]{\frac{\mathbf{det}\,(s)}{\mathbf{det}\,(x)}} \sqrt{x^T s + \sqrt{\mathbf{det}\,(x)\, \mathbf{det}\,(s)}}.$$

Hence, with λ as defined in (49), the vector a is given by

$$a = \frac{s + \lambda\, Q x}{\sqrt{2 \lambda} \sqrt{x^T s + \sqrt{\mathbf{det}\,(x)\, \mathbf{det}\,(s)}}}.$$

This proves the lemma. □

Lemma A.5 *Let $x,\, s \in K_+$ and let W be unique automorphism W of K such that $Wx = W^{-1}s$. Then*

$$Wx = W^{-1}s = \frac{\alpha}{2}\left[\begin{array}{c} 1 \\ \frac{\left(u\bar{x}+u^{-1}\bar{s}\right)+2\alpha^{-1}(x_1\bar{s}+s_1\bar{x})}{\alpha+u^{-1}s_1+u\,x_1} \end{array}\right],$$

where

$$u = \sqrt{\lambda}, \quad \alpha = \sqrt{2}\sqrt{x^Ts + \sqrt{\mathbf{det}\,(x)\,\mathbf{det}\,(s)}}.$$

Proof: By Lemma A.4,

$$Wx = \sqrt{\lambda}W_a x = \sqrt{\lambda}\left[\begin{array}{cc} a_1 & \bar{a}^T \\ \bar{a} & \mathbf{E}_{n-1} + \frac{\bar{a}\bar{a}^T}{1+a_1} \end{array}\right]\left[\begin{array}{c} x_1 \\ \bar{x} \end{array}\right] = u\left[\begin{array}{c} a^Tx \\ x_1\bar{a} + \bar{x} + \frac{\alpha(\bar{a}^T\bar{x})\,\bar{a}}{\alpha+ux_1+u^{-1}s_1} \end{array}\right],$$

$$W^{-1}s = \frac{1}{\sqrt{\lambda}}W_{Qa}s = \frac{1}{\sqrt{\lambda}}\left[\begin{array}{cc} a_1 & -\bar{a}^T \\ -\bar{a} & \mathbf{E}_{n-1} + \frac{\bar{a}\bar{a}^T}{1+a_1} \end{array}\right]\left[\begin{array}{c} s_1 \\ \bar{s} \end{array}\right] = u^{-1}\left[\begin{array}{c} a^TQs \\ -s_1\bar{a} + \bar{s} + \frac{\alpha(\bar{a}^T\bar{s})\,\bar{a}}{\alpha+ux_1+u^{-1}s_1} \end{array}\right],$$

with

$$a = \frac{1}{\alpha u}\left(s + \lambda\, Qx\right) = \frac{1}{\alpha}\left(u^{-1}s + u\,Qx\right). \tag{50}$$

and hence

$$1 + a_1 = 1 + \frac{1}{\alpha}\left(u^{-1}s_1 + u\,x_1\right) = \frac{\alpha + ux_1 + u^{-1}s_1}{\alpha}.$$

Since $Wx = W^{-1}s$ we get

$$\begin{aligned} Wx &= \frac{1}{2}\left(u\left[\begin{array}{c} a^Tx \\ x_1\bar{a} + \bar{x} + \frac{\alpha(\bar{a}^T\bar{x})\,\bar{a}}{\alpha+ux_1+u^{-1}s_1} \end{array}\right] + u^{-1}\left[\begin{array}{c} a^TQs \\ -s_1\bar{a} + \bar{s} + \frac{\alpha(\bar{a}^T\bar{s})\,\bar{a}}{\alpha+ux_1+u^{-1}s_1} \end{array}\right]\right) \\ &= \frac{1}{2}\left[\begin{array}{c} u(a^Tx) + u^{-1}(a^TQs) \\ (ux_1 - u^{-1}s_1)\bar{a} + u\bar{x} + u^{-1}\bar{s} + \frac{\alpha\,\bar{a}^T\left(u\bar{x}+u^{-1}\bar{s}\right)\bar{a}}{\alpha+ux_1+u^{-1}s_1} \end{array}\right]. \end{aligned}$$

Due to (48) one has

$$u(a^Tx) + u^{-1}(a^TQs) = u(a^Tx) + u^{-1}(\lambda a^Tx) = 2u\,a^Tx.$$

Moreover, using (50),

$$\begin{aligned} ua^Tx = a^T(ux) &= \frac{1}{\alpha}\left(u^{-1}s + u\,Qx\right)^T(ux) = \frac{1}{\alpha}\left(s^Tx + \lambda\,x^TQx\right) = \frac{1}{\alpha}\left(s^Tx + \lambda\,\mathbf{det}\,(x)\right) \\ &= \frac{1}{\alpha}\left(s^Tx + \sqrt{\mathbf{det}\,(x)\,\mathbf{det}\,(s)}\right) = \frac{1}{\alpha}\frac{\alpha^2}{2} = \frac{\alpha}{2}. \end{aligned}$$

Using (48) once more we obtain

$$\begin{aligned} \bar{a}^T\left(u\bar{x} + u^{-1}\bar{s}\right) &= u\bar{a}^T\bar{x} + u^{-1}\bar{a}^T\bar{s} = u\left(a^Tx - a_1x_1\right) + u^{-1}\left(a_1s_1 - a^TQs\right) \\ &= u(a^Tx) - u^{-1}(a^TQs) + u\left(-a_1x_1\right) + u^{-1}\left(a_1s_1\right) \\ &= u(a^Tx) - u^{-1}(\lambda a^Tx) - a_1\left(ux_1 - u^{-1}s_1\right) \\ &= a_1\left(u^{-1}s_1 - ux_1\right). \end{aligned}$$

From (50) we derive that $a_1 = (u^{-1}s_1 + u\,x_1)/\alpha$. Thus we find that

$$\bar{a}^T\left(u\bar{x} + u^{-1}\bar{s}\right) = \left(u^{-1}s_1 - ux_1\right)\frac{u^{-1}s_1 + u\,x_1}{\alpha} = \frac{u^{-2}s_1^2 - u^2x_1^2}{\alpha}.$$

Substitution of these results in the above expression for W gives,

$$Wx = \frac{1}{2}\begin{bmatrix} \alpha \\ (ux_1 - u^{-1}s_1)\bar{a} + u\bar{x} + u^{-1}\bar{s} + \frac{u^{-2}s_1^2 - u^2x_1^2}{\alpha + u^{-1}s_1 + u\,x_1}\,\bar{a} \end{bmatrix}.$$

The rest of the proof consists of simplifying the lower part of the above vector. We may write

$$\begin{aligned}
(ux_1 - u^{-1}s_1)\bar{a} + u\bar{x} + u^{-1}\bar{s} + \frac{u^{-2}s_1^2 - u^2x_1^2}{\alpha + u^{-1}s_1 + u\,x_1}\,\bar{a}& \\
&= \left(ux_1 - u^{-1}s_1 + \frac{u^{-2}s_1^2 - u^2x_1^2}{\alpha + u^{-1}s_1 + u\,x_1}\right)\bar{a} + u\bar{x} + u^{-1}\bar{s} \\
&= \frac{\alpha\left(ux_1 - u^{-1}s_1\right)}{\alpha + u^{-1}s_1 + u\,x_1}\,\bar{a} + u\bar{x} + u^{-1}\bar{s} \\
&= \frac{ux_1 - u^{-1}s_1}{\alpha + u^{-1}s_1 + u\,x_1}\,\left(u^{-1}\bar{s} - u\bar{x}\right) + u\bar{x} + u^{-1}\bar{s} \\
&= \left(1 + \frac{ux_1 - u^{-1}s_1}{\alpha + u^{-1}s_1 + u\,x_1}\right)u^{-1}\bar{s} + \left(1 - \frac{ux_1 - u^{-1}s_1}{\alpha + u^{-1}s_1 + u\,x_1}\right)u\bar{x} \\
&= \frac{(\alpha + 2ux_1)\,u^{-1}\bar{s} + \left(\alpha + 2u^{-1}s_1\right)u\bar{x}}{\alpha + u^{-1}s_1 + u\,x_1} \\
&= \frac{\alpha\left(u\bar{x} + u^{-1}\bar{s}\right) + 2x_1\bar{s} + 2s_1\bar{x}}{\alpha + u^{-1}s_1 + u\,x_1}.
\end{aligned}$$

Substitution of the last expression gives

$$Wx = \frac{1}{2}\begin{bmatrix} \alpha \\ \frac{\alpha\left(u\bar{x}+u^{-1}\bar{s}\right)+2x_1\bar{s}+2s_1\bar{x}}{\alpha+u^{-1}s_1+u\,x_1} \end{bmatrix} = \frac{\alpha}{2}\begin{bmatrix} 1 \\ \frac{\left(u\bar{x}+u^{-1}\bar{s}\right)+2\alpha^{-1}(x_1\bar{s}+s_1\bar{x})}{\alpha+u^{-1}s_1+u\,x_1} \end{bmatrix},$$

which completes the proof. □

References

[1] I. Adler and F. Alizadeh. Primal-dual interior point algorithms for convex quadratically constrained and semidefinite optimization promblems. Technical Report RRR-111-95, Rutger Center for Operations Research, Brunswick, NJ, 1995.

[2] E.D. Andersen, J. Gondzio, Cs. Mészáros, and X. Xu. Implementation of interior point methods for large scale linear programming. In T. Terlaky, editor, *Interior Point Methods of Mathematical Programming*, pages 189–252. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.

[3] E.D. Andersen, C. Roos, and T. Terlaky. On implementating a primal-dual interior-point method for conic quadratic optimization. *Mathematical Programming (Series B)*, 95:249–277, 2003.

[4] Y.Q. Bai, M. El Ghami, and C. Roos. A comparative study of kernel functions for primal-dual interior-point algorithms in linear optimization. *SIAM Journal on Optimization*, 15(1):101–128 (electronic), 2004.

[5] Y.Q. Bai and C. Roos. A primal-dual interior-point method based on a new kernel function with linear growth rate, 2002. To appear in *Proceedings Perth Industrial Optimization meeting*.

[6] L. Faybusovich. Linear system in jordan algebras and primal-dual interior-point algorithms. *Journal of Computational and Applitied Mathematics*, 86(1):149–175, 1997.

[7] M. Fukushima, Z.Q. Luo, and P. Tseng. A smoothing function for second-order-cone complementary problems. *SIAM Journal on Optimization*, 12:436–460, 2001.

[8] M.S. Lobo, L. Vandenberghe, S.E. Boyd, and H. Lebret. Applications of the second-order cone programming. *Linear Algebra and Its Applications*, 284:193–228, 1998.

[9] R.D.C. Monteiro and T. Tsuchiya. Polynomial convergence of primal-dual algorithms for the second-order program based the MZ-family of directions. *Mathematical Progamming*, 88:61–93, 2000.

[10] Y.E. Nesterov and M.J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research*, 22(1):1–42, 1997.

[11] Y.E. Nesterov and M.J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM Journal on Optimization*, 8(2):324 – 364, 1998.

[12] J. Peng, C. Roos, and T. Terlaky. *Self-Regularity. A New Paradigm for Primal-Dual Interior-Point Algorithms*. Princeton University Press, 2002.

[13] S. M. Robinson. Some continuity properties of polyhedral multifunctions. *Math. Programming Stud.*, 14:206–214, 1981. Mathematical Programming at Oberwolfach (Proc. Conf., Math. Forschungsinstitut, Oberwolfach, 1979).

[14] C. Roos, T. Terlaky, and J.-Ph. Vial. *Theory and Algorithms for Linear Optimization. An Interior-Point Approach*. John Wiley & Sons, Chichester, UK, 1997.

[15] S.H. Schmieta and F. Alizadeh. Associative and jordan algebras, and polynimal time interior-point algorithms for symmetric cones. *Mathematics of Operations Research*, 26(3):543 – 564, 2001.

[16] J. F. Sturm. Implementation of interior point methods for mixed semidefinite and second order cone optimization problems. *Optimization Methods & Software*, 17(6):1105–1154, 2002.

[17] J.F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods & Software*, 11:625–653, 1999.

[18] T. Tsuchiya. A polynomial primal-dual path-following algorithm for second-order cone programming. Research Memorandum No. 649, The Institute of Statistical Mathematics, Tokyo, Japan, 1997.

[19] T. Tsuchiya. A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programing. *Optimization Mehtods & Software*, 11-12:141–182, 1999.

[20] G.Q. Wang, Y.Q. Bai, and C. Roos. Primal-dual interior-point algorithms for semidefinite optimization based on a simple kernel function, 2004. To appear in *International Journal of Mathematical Algorithms*.

[21] H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors. *Handbook of semidefinite programming*. International Series in Operations Research & Management Science, 27. Kluwer Academic Publishers, Boston, MA, 2000. Theory, algorithms, and applications.