McMaster University

Advanced Optimization Laboratory

Title:

New Primal-Dual Algorithms for Second-Order Conic Optimization Based on Self-Regular Proximities

Author:

Jiming Peng, Cornelis Roos and Tamás Terlaky

AdvOl-Report No. 2000/5

October 2000, Hamilton, Ontario, Canada

New Primal-Dual Algorithms for Second-Order Conic Optimization Based on Self-Regular Proximities*

Jiming Peng[#†] Cornelis Roos[#] Tamás Terlaky[‡]

October 2000

[#]*Faculty of Information Technology and Systems,*
Delft University of Technology
P.O.Box 5031, 2600 GA Delft, The Netherlands.
Email: C.Roos@its.tudelft.nl.

[‡]*Department of Computing and Software,*
McMaster University, Hamilton, Ontario, Canada, L8S 4L7.
Email: terlaky@mcmaster.ca

Abstract

Recently the authors introduced the notions of *self-regular* functions and *self-regular* proximities and used them in the design and analysis of interior-point methods (IPMs) for linear and semidefinite optimization (LO and SDO). In this paper, we consider an extension of these concepts to second-order conic optimization (SOCO). General analytical functions whose arguments lie in the second-order cone are introduced and versatile properties of these functions are exploited as well. Based on these self-regular proximities, new primal-dual Newton methods for solving SOCO problems are proposed. It will be shown that these new large-update IPMs for SOCO enjoy analogous polynomial $\mathcal{O}\left(N^{\frac{q+1}{2q}}\log\frac{N}{\varepsilon}\right)$ iteration bound as their LO and SDO cousins where N is the number of constraining cones and q is a constant, the so-called barrier degree of the corresponding proximity. Our analysis allows to choose not only a constant q, but also a large q, as large as, $\log n$. In this case, our new algorithm has the best known $\mathcal{O}\left(N^{\frac{1}{2}}\log N\log\frac{N}{\varepsilon}\right)$ iteration bound for large-update IPMs.

Keywords: Second-order Conic Optimization, Interior Point Method, Primal-Dual Newton Method, Self-Regularity, Polynomial Complexity.

AMS Subject Classification: **90C05**

*The research of the first two authors are mainly supported by the project *High Performance Methods for Mathematical Optimization* under the Dutch SWON-grant 613-304-200. Both the first and third authors were partially supported by the National Science and Engineering Research Council of Canada, grant # : 227650-00.

[†]This work was finished when the first author visited the Department of Computing and Software, McMaster University, Canada. Email: pengj@mcmaster.ca, J.Peng@its.tudelft.nl.

1 Introduction

Mathematically, a typical second-order cone can be defined by

$$K = \{(x_1, x_2, \cdots, x_n) \in \Re^n : x_1^2 - \sum_{i=2}^{n} x_i^2 \geq 0, x_1 \geq 0\}.$$

This cone is often referred to as the Lorentz cone in physics and orally we also use its descriptive nickname: the ice-cream cone[1]. Second-order conic optimization (SOCO) is the problem of minimizing a linear objective function subject to the intersection of an affine set and the direct product of several second-order cones. Hence, from a pure mathematical viewpoint, the constraint functions defining the second-order cone are nothing others than some specific quadratic functions.

In light of the above-mentioned relation, SOCO is always recognized as a generalization of LO. Several important types of problems can be modeled as SOCO problems. For example, a general convex quadratic optimization problem with convex quadratic constraints can be cast as a SOCO problem [14]. It also includes as specific cases: robust linear optimization, robust least-squares, matrix-fractional problems and problems involving sums and maxima of norms etc. SOCO has been widely applied in several areas for long. To mention few of them, one can list e.g., antenna array weight design, grasping force optimization, FIR filter design, portfolio optimization with loss risk constraints, truss design etc. For more details about different applications of SOCO we refer to the survey paper [12] and the references therein.

An alternative way to describe the second-order cone is via matrix expression. For any $x = (x_1, \cdots, x_n)^T \in \Re^n$, let us define the matrix

$$\mathrm{mat}\,(x) = \begin{pmatrix} x_1 & x_{2:n} \\ x_{2:n}^T & x_1 E_{n-1} \end{pmatrix}, \tag{1}$$

where $x_{2:n} = (x_2, x_3, \cdots, x_n)$ and E_{n-1} denotes the identity matrix in $\Re^{(n-1)\times(n-1)}$. Using the above definition, one can easily prove that the vector $x \in K$ if and only if the matrix $\mathrm{mat}\,(x)$ is positive semidefinite, i.e., $\mathrm{mat}\,(x) \succeq 0$. This observation means that SOCO is essentially a specific case of semidefinite optimization (SDO). Thus one has the 'sandwich' relation $LO \subset SOCO \subset SDO$. This delicate circumstance addresses partially why SOCO did not attract much attention as its counterparts LO and SDO.

We consider in this paper the standard SOCO problem which takes the following form:

$$\begin{array}{ll} \text{(SOCO)} & \min c^T x \\ & Ax = b, \quad x \succeq_K 0, \end{array}$$

and its dual

$$\begin{array}{ll} \text{(SOCD)} & \max b^T y \\ & A^T y + s = c, \quad s \succeq_K 0, \end{array}$$

where K is the product of several second-order cones, i.e., $K = K^1 \times K^2 \times \cdots \times K^N$ with

$$K^j = \{(x_1^j, \cdots, x_{n_j}^j)^T \in \Re^{n_j} : \left(x_1^j\right)^2 \geq \sum_{i=2}^{n_j} \left(x_i^j\right)^2, x_1^j \geq 0\},$$

[1] This interesting name comes from the similarity between the shape of a general second-order cone in the space $\Re^3$ and the enjoyable summer-food.

$A \in \Re^{m\times n}$ with $n = \sum_{j=1}^{N} n_j$, and

$$x = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^N \end{pmatrix}, \quad x^j \in \Re^{n_i},\ j = 1, 2, \cdots, N, \quad x \in \Re^n.$$

Further, K_+ denotes the interior region of K. As standard, the notation $x \succeq_K s$ (or $x \succ_K s$) means that $x - s \in K$ (or $x - s \in K_+$). In this paper the matrix A is further assumed to be of full row rank, i.e., rank $A = m$.

An efficient approach to SOCO problems is to solve them via using IPMs. Since Karmarkar's epoch-making paper [10], the study of IPMs has dominated the continuous optimization literature for more than fifteen years and thousands of papers have been published about diverse aspects of IPMs for classes of problems ranging from LO to SDO. Due to its subtle relations with LO and SDO, SOCO had been neglected for a long time in the early stage of research on IPMs.

Recently this situation started to change gradually since it was observed that (as pointed out in the book [15]) albeit an SOCO problem can be solved via using an SDO approach, IPMs solving the SOCO problem directly have much better complexity than an IPM applied to the semidefinite formulation of the SOCO problem. The reason behind this observation is clear: IPMs with small-update (or large-update) working directly on SOCO have polynomial $\mathcal{O}\left(\sqrt{N}\log\frac{(x^0)^T s^0}{n\varepsilon}\right)$ (or $\mathcal{O}\left(N\log\frac{(x^0)^T s^0}{n\varepsilon}\right)$) iteration bound where N is the number of cones involved in the underlying problem and n is number of variables in the matrix, while for the SDO algorithm, the iteration number is bounded above by $\mathcal{O}\left(\sqrt{n}\log\frac{(x^0)^T s^0}{n\varepsilon}\right)$ (or $\mathcal{O}\left(n\log\frac{(x^0)^T s^0}{n\varepsilon}\right)$) and in general n is a much bigger number than N. Another important observation is that both in theory and in practice, at each iteration much less work is needed for the IPMs based on SOCO. Moreover, even in the case when the same type of search direction, say the A.H.O. (see our discussion in Section 6.3 later) direction is employed, the theoretical properties of the corresponding algorithm for SOCO can be enhanced slightly compared with its analogue for SDO [13]. These theoretical conclusions are further strengthened by recent extensive numerical experiments on SOCO [4] where several SOCO problems with hundreds of thousands of variables are reported to be solved efficiently.

To be more specific, let us go into more details. Throughout this paper, we assume that both (SOCO) and (SOCD) satisfy the interior point condition (IPC), i.e., there exists (x^0, y^0, s^0) such that

$$Ax^0 = b,\ x^0 \succ_K 0; \qquad A^T y^0 + s^0 = c,\ s^0 \succ_K 0.$$

It is known that the IPC is a rather mild assumption in the study of SOCO, since by using the homogeneous self-dual model described in [4], we could cast the original problem as a slightly larger SOCO problem such that a strictly feasible point for this new problem can be easily obtained. For this and other properties of SOCO, we refer to [4], the recent book [24] and the references therein. Under the IPC, finding an optimal solution of SOCO is equivalent to solving the following system

$$\begin{aligned} Ax &= b, \quad x \succeq_K 0, \\ A^T y + s &= c, \quad s \succeq_K 0, \\ \mathrm{mat}\,(x)\, s &= 0. \end{aligned} \tag{2}$$

The basic idea of primal-dual IPMs is to replace the third equation in (2), the so-called *complementarity condition* for (SOCO) and (SOCD) by the parameterized equation $\mathrm{mat}\,(x)\,s = \mu\tilde{e}$ with $\mu > 0$, where

$$\tilde{e} = \begin{pmatrix} \tilde{e}^1 \\ \tilde{e}^2 \\ \vdots \\ \tilde{e}^N \end{pmatrix}, \quad \tilde{e}^j = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \Re^{n_j},\ j = 1, 2, \cdots, N.$$

Thus we consider the system

$$\begin{aligned} Ax &= b, \quad x \succeq_K 0, \\ A^T y + s &= c, \quad s \succeq_K 0, \\ \mathrm{mat}\,(x)\,s &= \mu\tilde{e}. \end{aligned} \tag{3}$$

If the IPC holds, then for each $\mu > 0$, the parameterized system (3) has a unique solution. This solution is denoted by $(x(\mu), y(\mu), s(\mu))$ and we call $x(\mu)$ the μ-center of (P) and $(y(\mu), s(\mu))$ the μ-center of (D). The set of μ-centers (with μ running through all positive real numbers) gives a homotopy path, which is called *the central path*. The central path converges to the solution set of SOCO as μ reduces to zero [5, 15, 21]. IPMs trace the central path appropriately and find an approximate solution to the underlying SOCO problem as μ goes to zero.

To trace the central path efficiently, various strategies have been introduced to keep the iterative sequence in a certain neighborhood of the central path as well as reducing the parameter μ. These strategies have played an important role both in the analysis and practice of IPMs. It is worth to point out that two general strategies are widely used in IPMs with respect to the update of the parameter μ. These are the so-called small-update and large-update IPMs. It has been proven and generally accepted that the worst-case iteration bound of small-update IPMs is better than that for large-update IPMs while the latter ones are much more efficient in practice (see the discussion in the introduction of [19]). This is a big gap between the theory and practice of IPMs.

Recently, the authors of this paper introduced the concept of an univariate *self-regular* function, and showed that any such function can be naturally extended to a proximity function on the positive orthant and the cone of positive definite matrices [19]. The *self-regular* proximities obtained in this way, can be used in IPMs to keep control on the distance of an iterative sequence to the central path as well as to define the corresponding search directions. By using some new analysis tools developed in [18, 19] and employing the new search directions, they were able to show that the resulting new large-update IPMs for LO and SDO have polynomial $\mathcal{O}\left(n^{\frac{q+1}{2q}} \log \frac{n}{\epsilon}\right)$ iteration bounds where q is a constant, the so-called barrier degree of the proximity. This is a significant improvement over the before known $\mathcal{O}\left(n \log \frac{n}{\epsilon}\right)$ iteration bound of large-update IPMs.

The present work aims at extending the results of [19] to SOCO. We mention that there exist already many works about IPMs for SOCO [3, 4, 13, 15, 22, 23, 24]. As claimed by Alizadeh and Schmieta in [3], the polynomial convergence proof of some IPMs for LO can be extended to SOCO and SDO by following their 'recipe' word by word. It is worth to note that the class of IPMs addressed in [3] is the class of so-called potential reduction methods based on classic logarithmic barrier functions in different cones whose behavior has been well studied by many authors (see, i.e., [16, 17]). However, as we observed in [19], the proof of the global convergence of IPMs based on the *self-regular* approach involves the analysis of general analytic functions

which are defined on the associated cone, and to analyze the properties of a general analytical function in the cone is much harder than that of the logarithmic barrier function. As we will see in our later analysis, in order to establish the complexity of our algorithm, we will first study versatile properties of general functions in the second-order cone K. We remark that there exist already some works [6, 7] dealing with functions in the second-order cone, however most of the results presented in [6, 7] are in very deep abstract mathematical form and not directly applicable for our specific purpose.

The paper is organized as follows. In Section 2, we will first introduce the definition of general analytic functions defined on K and discuss versatile properties of these functions. Section 3 is devoted to introduce the new search direction and the scaling technique for SOCO. The notions of *self-regular* functions and *self-regular* proximities in the second-order cone K are discussed as well. In Section 4 we first describe our new algorithm based on a *self-regular* proximity. Then we will establish the polynomial complexity of the algorithm and finally close this paper by some concluding remarks.

We mention that in the rest of this paper, we denote by $\Re_+$ the nonnegative axis, i.e., $\Re_+ = [0, \infty)$, and by $\Re_{++}$ the positive axis, i.e., $\Re_{++} = (0, \infty)$.

2 Preliminary Results on Functions Associated with Second-Order Cone

In the design and analysis of IPMs, we always resort to some functions defined in a suitable space. In this section we will present some fundamental results about general functions defined in the second-order cone. We note that in [6], via using Jordan algebra, the authors developed some deep and abstract theory for various functions in the so-called v-space. However, these results presented in [6] are not easily understandable and can not be applied to IPMs directly. Later on, Faybusovich [7, 8] recognized the significance of Euclidean Jordan algebra and used it to analyze IPMs. With regard to the algorithmic aspect, Faybusovich's results rely on, more or less the logarithmic barrier approach which originated from the book [15] or the papers [16, 17]. Tsuchiya in [22] and Monteiro and Tsuchiya [13] then applied Jordan algebra to analyze IPMs for SOCO with specification to various search directions.

In this section, following the aforementioned approach, we study general functions on the second-order cone via Jordan algebra. To ease the discussion, in this technical section we assume the cone K is defined with $N = 1$. First we observe that, closely associated with the cone K is a matrix

$$Q = \mathrm{diag}\,(1, -1, \cdots, -1).$$

We refer to Q as the representation matrix of the cone K since there holds simply

$$K = \{x \in \Re^n : x^T Q x \geq 0,\ x_1 \geq 0\}.$$

Obviously one has $Q^2 = E$.

2.1 Jordan Algebra

The Euclidean Jordan algebra for the second-order cone K is defined by the bilinear operator

$$x \circ s = (x^T s, x_1 s_2 + s_1 x_2, \cdots, x_1 s_n + s_1 x_n)^T = (x^T s, x_1 s_{2:n} + s_1 x_{2:n})^T, \tag{4}$$

where $x, s \in \Re^n$. Obviously, the Jordan product $\circ$ is commutative, i.e., $x \circ s = s \circ x$. It is also easy to verify that for any $x, s \in \Re^n$, one has

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

It may be worthwhile to point out that the cone K is not closed under the Jordan product. For example, if $n = 3$, then $x = (1.5, 1, 1)^T \in K$ and $s = (1.5, 1, -1)^T \in K$, but $x \circ s = (2.25, 3, 0)^T \notin K$. Another interesting, easy to verify and important fact is that for any $z \in \mathbb{R}^n$ one has $z \circ z \in K$. In other words, the Jordan square of any vector belongs to K. On the other hand, for every nonzero $x \in K$, the equation $z \circ z = x$ has two solutions z with $z_1 \geq 0$, namely

$$z_1 = \sqrt{\tfrac{1}{2}(x_1 + \lambda)}, \quad z_i = \frac{x_i}{2z_1}, \quad 2 \leq i \leq n, \tag{5}$$

where $\lambda = \pm\sqrt{x_1^2 - \sum_{i \geq 2} x_i^2}$. One may easily verify that $z \in K$ if $\lambda \geq 0$, whereas $z \notin K$ if $\lambda < 0$. We conclude from this that for every $x \in K$, the equation $z \circ z = x$ has a unique solution z in K. Later on we will denote this solution as $x^{\frac{1}{2}}$, or as $\sqrt{x}$, and we will consider more general functions on K. Before doing so, we need to introduce the concepts of trace and determinant with respect to K.

2.2 Eigenvalues, Trace and Determinant Associated with Second-Order Cone

We denote by $\lambda_{\max}(x)$ and $\lambda_{\min}(x)$ the maximal and minimal eigenvalues of the matrix $\mathrm{mat}\,(x)$, respectively. Namely

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

The trace and the determinant of a vector $x \in \Re^n$ associated with K can be defined as follows[6].

Definition 2.1 [2] *For any $x \in \Re^n$, the trace of x associated with K is defined by*

$$\mathbf{Tr}\,(x) = \lambda_{\max}(x) + \lambda_{\min}(x) = 2x_1; \tag{7}$$

and the determinant of x associated with K is given by

$$\det(x) = \lambda_{\max}(x)\lambda_{\min}(x) = x_1^2 - \|x_{2:n}\|^2. \tag{8}$$

From the above definitions, one can easily see that for any $x, s \in \Re^n$, one has

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

Moreover, in light of the definitions (7) and (8), it is straightforward to verify the following result.

Lemma 2.2 *Suppose the trace and determinant of a vector $x \in \Re^n$ are defined by (7) and (8), respectively. Then $x \succeq_K 0$ (or $x \succ_K 0$) if and only if it has nonnegative (or positive) trace and determinant, or in other words $\mathbf{Tr}\,(x) \geq 0, \det(x) \geq 0$ (or $\mathbf{Tr}\,(x) > 0, \det(x) > 0$).*

[2] These definitions can be viewed as variants of the trace and determinant of general matrices.

Our next lemma collects several elementary results about the behavior of the trace and determinant of the Jordan product of two vectors. These results demonstrate the differences between the determinant and trace for elements of the second-order cone K, and those notions as usually defined for matrices[3] .

Lemma 2.3 *Suppose that x and s are two vectors in K. Then we have*

$$\lambda_{\max}(x)\lambda_{\min}(s) + \lambda_{\min}(x)\lambda_{\max}(s) \leq \mathbf{Tr}\,(x \circ s) \leq \lambda_{\max}(x)\lambda_{\max}(s) + \lambda_{\min}(x)\lambda_{\min}(s); \tag{9}$$

and

$$\det(x \circ s) \leq \det(x)\det(s). \tag{10}$$

Furthermore, equality holds in (10) if and only if there exist two constants $\beta_1, \beta_2 \in \Re$ with $|\beta_1| + |\beta_2| > 0$ such that $\beta_1 x_{2:n} = \beta_2 s_{2:n}$.

Proof: We first consider the relation (9). Using the notations $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ given by (6) and the well-known Cauchy-Schwartz inequality, since both x and s belong to K, one has

$$\begin{aligned} 0 &\leq \lambda_{\max}(x)\lambda_{\min}(s) + \lambda_{\min}(x)\lambda_{\max}(s) = 2\left(x_1 s_1 - \|x_{2:n}\|\,\|s_{2:n}\|\right) \\ &\leq 2x^T s = \mathbf{Tr}\,(x \circ s) \leq 2\left(x_1 s_1 + \|x_{2:n}\|\,\|s_{2:n}\|\right) \\ &= \lambda_{\max}(x)\lambda_{\max}(s) + \lambda_{\min}(x)\lambda_{\min}(s), \end{aligned}$$

which gives (9).

To prove (10), we note that, by making use of the definition (8), one gets

$$\begin{aligned} \det(x \circ s) &= (x^T s)^2 - \|x_1 s_{2:n} + s_1 x_{2:n}\|^2 = (x_1 s_1)^2 + (x_{2:n} s_{2:n}^T)^2 - (x_1)^2\,\|s_{2:n}\|^2 - (s_1)^2\,\|x_{2:n}\|^2 \\ &\leq (x_1 s_1)^2 + \|x_{2:n}\|^2\,\|s_{2:n}\|^2 - (x_1)^2\,\|s_{2:n}\|^2 - (s_1)^2\,\|x_{2:n}\|^2 \\ &= \left((x_1)^2 - \|x_{2:n}\|^2\right)\left((s_1)^2 - \|s_{2:n}\|^2\right) = \det(x)\det(s); \end{aligned}$$

and the equality holds if and only if $\left|x_{2:n} s_{2:n}^T\right| = \|x_{2:n}\|\,\|s_{2:n}\|$. This means the equality holds only when the vectors $x_{2:n}$ and $s_{2:n}$ are linearly dependent. The proof of the lemma is completed. □

2.3 Functions Associated with Second-Order Cone and Their Derivatives

Note that if $n = 1$, then $K = \Re_+$, and if $n \geq 1$, then $\Re_+ \subseteq K$. Our aim in this section is to show that any function mapping $\Re_+$ into $\Re_+$, naturally can be extended to a function that maps K into itself. We start with a more general definition with respect to functions associated with the second-order cone K.

Definition 2.4 *Suppose that $\psi(t)$ is a function from $\Re$ to $\Re$ and $x \in \Re^n$. Then the function $\psi(x) : \Re^n \to \Re^n$ associated with the second-order cone K is defined as follows*[4]

$$\psi(x) = \begin{cases} \left(\frac{1}{2}\left(\psi(\lambda_{\max}(x)) + \psi(\lambda_{\min}(x))\right), \frac{\psi(\lambda_{\max}(x)) - \psi(\lambda_{\min}(x))}{2\|x_{2:n}\|} x_{2:n}\right)^T, & \text{if} \quad x_{2:n} \neq 0; \\ (\psi(\lambda_{\max}(x)), 0, \cdots, 0)^T, & \text{if} \quad x_{2:n} = 0. \end{cases} \tag{11}$$

[3]This distinction can be expected by noticing that $\mathbf{Tr}\,(\mathrm{mat}\,(x)) = n x_1 = \frac{n}{2}\mathbf{Tr}\,(x)$ and that

$$\det(\mathrm{mat}\,(x)) = x_1^{n-2}\left(x_1^2 - \|x_{2:n}\|^2\right) = x_1^{n-2}\det(x).$$

[4]We note that recently, Fukushima, Luo and Tseng [9] also defined functions associated with second-order cone. Their definition is slightly different from our definition (11). However, the definition given by (11) is clearer and more direct.

It can easily be verified that if $\psi(t) \geq 0$ for any $t \geq 0$ and $x \in K$, then the above definition implies that $\psi(x) \in K$. Thus it becomes clear that every nonnegative (positive) function on the nonnegative (positive) axis naturally extends to a function that maps (the interior of) K into itself. Likewise the LO and SDO cases, the function $\psi(t)$ is called the kernel function of $\psi(x)$. As a consequence of the above definition we have a big source of functions mapping K into itself. For instance, we may write x^p, where p is any number in $\Re$ and $x \in K$. Let us consider some special cases, for example, $p = -1$. In this case Definition 11 yields

$$x^{-1} = \frac{1}{\det(x)}(x_1, -x_2, -x_3, \cdots, -x_n)^T, \qquad \forall x \succ_K,$$

and one may easily see that $x \circ x^{-1} = \tilde{e}$. One can also easily verify that if $p = \frac{1}{2}$ then $\sqrt{x}$, as defined by Definition 11 is given by (5). Moreover, it is also clear that any analytical functions like $\exp(x), \sin(x)$ are now well-defined. Similarly we can also define the function $\psi'(x)$ by (11) whose kernel function is $\psi'(t)$.

However, this raises a question. For example, with $\psi(t) = t^2$, can it be justified to write $\psi(x) = x^2$? The answer is affirmative due to the fact that in that case $\psi(x) = x \circ x$, as easily can be verified. More generally, when composing functions by taking products, the behaviour with respect to the Jordan product is nice, as the following lemma shows.

Lemma 2.5 *Suppose that $\psi_1(t)$ and $\psi_2(t)$ are two functions from $\Re$ into $\Re$ and that $\psi_1(x)$ and $\psi_2(x)$ are two associated functions defined by (11). If $\psi_0(t) = \psi_1(t)\psi_2(t)$, then $\psi_0(x) = \psi_1(x) \circ \psi_2(x)$ holds for any $x \in \Re^n$.*

Proof: We consider first the case that $\|x_{2:n}\| > 0$. From (11) we deduce

$$\psi_j(x) = \left(\frac{1}{2}\left(\psi_j(\lambda_{\max}(x)) + \psi_j(\lambda_{\min}(x))\right), \frac{\psi_j(\lambda_{\max}(x)) - \psi_j(\lambda_{\min}(x))}{2\|x_{2:n}\|} x_{2:n}\right)^T, \quad j = 0, 1, 2;$$

Hence, by simple computation, one obtains

$$\begin{aligned}
(\psi_1(x) \circ \psi_2(x))_1 &= \frac{1}{2}\left(\psi_1(\lambda_{\max}(x))\psi_2(\lambda_{\max}(x)) + \psi_1(\lambda_{\min}(x))\psi_2(\lambda_{\min}(x))\right); \\
(\psi_1(x) \circ \psi_2(x))_i &= \frac{x_i}{2\|x_{2:n}\|}\left(\psi_1(\lambda_{\max}(x))\psi_2(\lambda_{\max}(x)) - \psi_1(\lambda_{\min}(x))\psi_2(\lambda_{\min}(x))\right), \quad i = 2, \cdots, n.
\end{aligned}$$

The statement of the lemma follows immediately from the above two relations. It is straightforward to verify the conclusion of the lemma when $\|x_{2:n}\| = 0$. □

It is trivial to verify the following result about general functions associated with the second-order cone defined by (11).

Lemma 2.6 *Suppose that the function $\psi(x) : \Re^n \to \Re^n$ be defined by (11). Then there hold*

$$\begin{aligned}
\|\psi(x)\| &= \frac{\sqrt{2}}{2}\sqrt{\psi^2(\lambda_{\max}(x)) + \psi^2(\lambda_{\min}(x))}; \\
\mathbf{Tr}(\psi(x)) &= \psi(\lambda_{\max}(x)) + \psi(\lambda_{\min}(x)); \\
\det(\psi(x)) &= \psi(\lambda_{\max}(x))\psi(\lambda_{\min}(x)).
\end{aligned}$$

Let us elaborate a little more on the relations among the eigenvalues of two vectors in the second-order cone and those of their Jordan product. Note that from Lemma 2.3 we know that for any

$x, s \succ_K 0$, the determinant $\det(x \circ s) = \det(x)\det(s)$ if and only if the vectors $x_{2:n}$ and $s_{2:n}$ are linearly dependent. Without loss of generality, we may assume that $x_{2:n} \neq 0$ and $s_{2:n} = \beta x_{2:n}$ for some $\beta \in \Re$. In the following discussions we will show this further implies that the vector s can be represented as a function of x and that the vector $x \circ s \in K_+$.

Lemma 2.7 *Suppose that x and s are two vectors belong to K_+ with $x_{2:n} \neq 0$. If*

$$\det(x \circ s) = \det(x)\det(s),$$

then there exists a function $\psi(t) : \Re_+ \to \Re_+$ such that $s = \psi(x)$ and the Jordan product $x \circ s \in K_+$. Moreover there holds

$$\lambda_{\min}(x)\lambda_{\min}(s) \leq \lambda_{\min}(x \circ s) \leq \lambda_{\max}(x \circ s) \leq \lambda_{\max}(x)\lambda_{\max}(s). \tag{12}$$

Proof: By making use of the last conclusion of Lemma 2.3, since $\det(x \circ s) = \det(x)\det(s)$, one can claim that

$$s_{2:n} = \frac{\beta \|s_{2:n}\|}{\|x_{2:n}\|} x_{2:n},$$

where β equals to 1 or -1. Let $\psi(t)$ be a univariate function from $\Re_{++}$ into $\Re_{++}$ satisfying the following equalities

$$\frac{1}{2}\left(\psi\left(\lambda_{\max}(x)\right) + \psi\left(\lambda_{\min}(x)\right)\right) = s_1, \quad \frac{1}{2}\left(\psi\left(\lambda_{\max}(x)\right) - \psi\left(\lambda_{\min}(x)\right)\right) = \beta\left\|s_{2:n}\right\|.$$

Such a function exists since $s \in K_+$. Thus, by definition (11) we can claim $s = \psi(x)$. Now invoking Lemma 2.5, we can write $x \circ s = \psi_1(x)$ where the kernel function $\psi_1(t) = t\psi(t)$. Therefore, since $\psi_1(t) > 0$ for any $t > 0$ and $x_{2:n} \neq 0$, from its basic definition (11) it follows $x \circ s = \psi_1(x) \succ_K 0$.

It remains to prove (12). Since $x \circ s = \psi_1(x)$, thus we have

$$\begin{aligned} \lambda_{\max}(x \circ s) &= \max\{\psi_1(\lambda_{\max}(x)), \psi_1(\lambda_{\min}(x))\} \\ &\leq \lambda_{\max}(x)\max\{\psi(\lambda_{\max}(x)), \psi(\lambda_{\min}(x))\} = \lambda_{\max}(x)\lambda_{\max}(s), \end{aligned}$$

which, combining with the assumption in the lemma that $\det(x \circ s) = \det(x)\det(s)$, further yields

$$\lambda_{\min}(x \circ s) \geq \lambda_{\min}(x)\lambda_{\min}(s).$$

The proof of the lemma is completed. □

It is worthwhile to compare the above lemma with its analogue in matrix theory. Suppose that X and S are both symmetric and positive definite. Let $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ denote the maximal and minimal eigenvalues of the corresponding matrix. Then we have

$$\lambda_{\min}(X)\lambda_{\min}(S) \leq \lambda_{\min}(XS) \leq \lambda_{\max}(XS) \leq \lambda_{\max}(X)\lambda_{\max}(S). \tag{13}$$

Note that for any $X, S \in \Re^{n \times n}$, there holds trivially $\det(XS) = \det(X)\det(S)$. The results presented in Lemma 2.7 are very helpful in our later discussion about the features of *self-regular* functions associated with the second-order cone which can be viewed as a direct extension of univariate *self-regular* functions introduced in [19].

Definition 2.8 *A function $\psi(t) \in \mathcal{C}^2 : \Re_{++} \to \Re_+$ is self-regular if it satisfies the following conditions*

C.1 $\psi(t)$ *is strictly convex with respect to* $t > 0$ *and vanishes at its global minimal point* $t = 1$, *i.e.,* $\psi(1) = \psi'(1) = 0$. *Further, there exist positive constants* $\nu_2 \geq \nu_1 > 0$ *and* $p \geq 1, q \geq 1$ *such that*

$$\nu_1(t^{p-1} + t^{-1-q}) \leq \psi''(t) \leq \nu_2(t^{p-1} + t^{-1-q}), \quad \forall t \in (0, \infty); \tag{14}$$

C.2 For any $t_1, t_2 > 0$,

$$\psi(t_1^r t_2^{1-r}) \leq r\psi(t_1) + (1-r)\psi(t_2), \quad \forall r \in [0, 1]. \tag{15}$$

We call parameter q the *barrier degree* and p the *growth degree* of the function $\psi(t)$ if it is *self-regular*. A typical family of *self-regular* functions is given by

$$\Upsilon_{p,q}(t) = \frac{1}{p(p+1)}\left(t^{p+1} - 1\right) + \frac{1}{q(q-1)}\left(t^{1-q} - 1\right) + \frac{p-q}{pq}(t-1), \quad p, q \geq 1. \tag{16}$$

It is worth to mention that the function $\Upsilon_{p,q}(t)$ satisfies condition C.1 with $\nu_1 = \nu_2 = 1$.

The definition of *self-regular* function for the second-order cone K is recorded as follows.

Definition 2.9 *A function* $\psi(x)$ *associated with the second-order cone* K *given by (11) is said to be self-regular if its kernel function* $\psi(t)$ *is self-regular.*

We denote by $\Psi(x)$ the trace of the function $\psi(x)$, i.e.,

$$\Psi(x) = \mathbf{Tr}\left(\psi(x)\right) = \psi(\lambda_{\max}(x)) + \psi(\lambda_{\min}(x)). \tag{17}$$

Our next proposition characterizes several important properties of a *self-regular* function for the second-order cone K.

Proposition 2.10 *Let the functions* $\psi(x) : K_+ \to K$ *and* $\Psi(x) : K_+ \to \Re_+$ *be defined by (11) and (17) respectively. If the function* $\psi(x)$ *is self-regular, then the following statements hold.*

(i) $\Psi(x)$ *is strictly convex with respect to* $x \in K_+$ *and vanishes at its global minimal point* $x = \tilde{e}$, *i.e.,* $\Psi(\tilde{e}) = 0, \psi(\tilde{e}) = \psi'(\tilde{e}) = 0$. *Further, there exist two positive constants* $\nu_1, \nu_2 > 0$ *and* $p, q \geq 1$ *such that*

$$\nu_1(x^{p-1} + x^{-1-q}) \preceq_K \psi''(x) \preceq_K \nu_2(x^{p-1} + x^{-1-q}); \tag{18}$$

(ii) Suppose x *and* s *are two vectors in* K_+. *If* $v \in K_+$ *satisfies*

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

then

$$\Psi(v) \leq \frac{1}{2}\left(\Psi(x) + \Psi(s)\right). \tag{19}$$

Proof: To show that the first claim of the proposition is true, we need to show that $\Psi(x)$ is strictly convex for $x \succ_K 0$, that is for any $x, s \succ_K 0$ and $x \neq s$, there holds

$$\Psi\left(\frac{x+s}{2}\right) < \frac{1}{2}\left(\Psi(x) + \Psi(s)\right).$$

Since $x, s \in K_+$, through simple calculus one can prove that

$$\lambda_{\max}(\frac{x+s}{2}) = \frac{x_1+s_1}{2} + \frac{1}{2}\|x_{2:n}+s_{2:n}\| \leq \frac{1}{2}\left(\lambda_{\max}(x)+\lambda_{\max}(s)\right);$$

and similarly

$$\lambda_{\min}(\frac{x+s}{2}) = \frac{x_1+s_1}{2} - \frac{1}{2}\|x_{2:n}+s_{2:n}\| \geq \frac{1}{2}\left(\lambda_{\min}(x)+\lambda_{\min}(s)\right).$$

Recalling the definitions of $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$, it follows trivially

$$\lambda_{\max}(\frac{x+s}{2}) + \lambda_{\min}(\frac{x+s}{2}) = x_1+s_1 = \frac{1}{2}\left(\lambda_{\max}(x)+\lambda_{\min}(x)+\lambda_{\max}(s)+\lambda_{\min}(s)\right).$$

Thus, from the above three relations we can conclude that there exist two constants $\beta_1 \geq 0$ and $\beta_2 \geq 0$ with $\beta_1+\beta_2 = 1$ such that

$$\begin{aligned}
\lambda_{\max}(\frac{x+s}{2}) &= \frac{\beta_1}{2}\left(\lambda_{\min}(x)+\lambda_{\min}(s)\right) + \frac{\beta_2}{2}\left(\lambda_{\max}(x)+\lambda_{\max}(s)\right);\\
\lambda_{\min}(\frac{x+s}{2}) &= \frac{\beta_2}{2}\left(\lambda_{\min}(x)+\lambda_{\min}(s)\right) + \frac{\beta_1}{2}\left(\lambda_{\max}(x)+\lambda_{\max}(s)\right).
\end{aligned}$$

Now, making use of the strict convexity of the function $\psi(t)$ twice, one obtains

$$\begin{aligned}
\Psi\left(\frac{x+s}{2}\right) &= \psi\left(\lambda_{\max}(\frac{x+s}{2})\right) + \psi\left(\lambda_{\min}(\frac{x+s}{2})\right)\\
&= \psi\left(\frac{\beta_1}{2}\left(\lambda_{\min}(x)+\lambda_{\min}(s)\right) + \frac{\beta_2}{2}\left(\lambda_{\max}(x)+\lambda_{\max}(s)\right)\right)\\
&\quad + \psi\left(\frac{\beta_2}{2}\left(\lambda_{\min}(x)+\lambda_{\min}(s)\right) + \frac{\beta_1}{2}\left(\lambda_{\max}(x)+\lambda_{\max}(s)\right)\right)\\
&\leq \psi\left(\frac{\lambda_{\min}(x)+\lambda_{\min}(s)}{2}\right) + \psi\left(\frac{\lambda_{\max}(x)+\lambda_{\max}(s)}{2}\right)\\
&\leq \frac{1}{2}\left(\psi\left(\lambda_{\max}(x)\right) + \psi\left(\lambda_{\min}(x)\right) + \psi\left(\lambda_{\max}(s)\right) + \psi\left(\lambda_{\min}(s)\right)\right)\\
&= \frac{1}{2}\left(\Psi(x)+\Psi(s)\right).
\end{aligned}$$

Note that since $x \neq s$, at least one of the two inequalities in the above proof holds strictly. This proves the strict convexity of $\Psi(x)$. The remaining terms in the first statement can be verified through direct calculus.

It remains to prove the second statement of the proposition. For this we first observe that, since $v \in K_+$, there hold

$$\det(v) = \det\left(v^2\right)^{\frac{1}{2}} = \det(x)^{\frac{1}{2}}\det(s)^{\frac{1}{2}} = \left(\lambda_{\min}(x)\lambda_{\min}(s)\right)^{\frac{1}{2}}\left(\lambda_{\max}(x)\lambda_{\max}(s)\right)^{\frac{1}{2}}; \tag{20}$$

and

$$\begin{aligned}
\mathbf{Tr}(v) &= \lambda_{\max}(v)+\lambda_{\min}(v) = \left(\lambda_{\max}(v)^2+\lambda_{\min}(v)^2+2\lambda_{\max}(v)\lambda_{\min}(v)\right)^{\frac{1}{2}}\\
&= \left(\mathbf{Tr}\left(v^2\right)+2\det(v)\right)^{\frac{1}{2}} = \left(\mathbf{Tr}(x\circ s)+2\left(\det(x)\det(s)\right)^{\frac{1}{2}}\right)^{\frac{1}{2}}\\
&\leq \left(\lambda_{\min}(x)\lambda_{\min}(s)+\lambda_{\max}(x)\lambda_{\max}(s)+2\left(\lambda_{\min}(x)\lambda_{\min}(s)\lambda_{\max}(x)\lambda_{\max}(s)\right)^{\frac{1}{2}}\right)^{\frac{1}{2}}\\
&= \sqrt{\lambda_{\min}(x)\lambda_{\min}(s)}+\sqrt{\lambda_{\max}(x)\lambda_{\max}(s)},
\end{aligned} \tag{21}$$

where the inequality follows from (9). Therefore, by making use of (20) and (21), we obtain

$$\begin{aligned}\lambda_{\max}(v) &= \frac{1}{2}\left(\mathbf{Tr}\,(v) + \lambda_{\max}(v) - \lambda_{\min}(v)\right) = \frac{1}{2}\mathbf{Tr}\,(v) + \frac{1}{2}\sqrt{\mathbf{Tr}\,(v)^2 - 4\det\,(v^2)} \\ &\leq \frac{1}{2}\mathbf{Tr}\,(v) + \frac{1}{2}\left(\sqrt{\lambda_{\max}(x)\lambda_{\max}(s)} - \sqrt{\lambda_{\min}(x)\lambda_{\min}(s)}\right) \\ &\leq \lambda_{\max}(x)^{\frac{1}{2}}\lambda_{\max}(s)^{\frac{1}{2}},\end{aligned}$$

where all inequalities follow from (20) and (21). Now, invoking (20), we can further claim

$$\lambda_{\min}(v) \geq \lambda_{\min}(x)^{\frac{1}{2}}\lambda_{\min}(s)^{\frac{1}{2}}.$$

From the above discussions we can easily verify that there exists a constant $r \in [\frac{1}{2}, 1)$ such that

$$\begin{aligned}\lambda_{\min}(v) &= \lambda_{\min}(x)^{\frac{r}{2}}\lambda_{\min}(s)^{\frac{r}{2}}\lambda_{\max}(x)^{\frac{1-r}{2}}\lambda_{\max}(s)^{\frac{1-r}{2}}; \\ \lambda_{\max}(v) &= \lambda_{\max}(x)^{\frac{r}{2}}\lambda_{\max}(s)^{\frac{r}{2}}\lambda_{\min}(x)^{\frac{1-r}{2}}\lambda_{\min}(s)^{\frac{1-r}{2}}.\end{aligned}$$

By applying condition C.2 twice, we deduce

$$\begin{aligned}\Psi(v) &= \psi\left(\lambda_{\min}(v)\right) + \psi\left(\lambda_{\max}(v)\right) \\ &= \psi\left(\lambda_{\min}(x)^{\frac{r}{2}}\lambda_{\min}(s)^{\frac{r}{2}}\lambda_{\max}(x)^{\frac{1-r}{2}}\lambda_{\max}(s)^{\frac{1-r}{2}}\right) \\ &\quad + \psi\left(\lambda_{\max}(x)^{\frac{r}{2}}\lambda_{\max}(s)^{\frac{r}{2}}\lambda_{\min}(x)^{\frac{1-r}{2}}\lambda_{\min}(s)^{\frac{1-r}{2}}\right) \\ &\leq \psi\left(\lambda_{\min}(x)^{\frac{1}{2}}\lambda_{\min}(s)^{\frac{1}{2}}\right) + \psi\left(\lambda_{\max}(x)^{\frac{1}{2}}\lambda_{\max}(s)^{\frac{1}{2}}\right) \\ &\leq \frac{1}{2}\left(\psi(\lambda_{\min}(x)) + \psi(\lambda_{\max}(x)) + \psi(\lambda_{\max}(s)) + \psi(\lambda_{\min}(s))\right) \\ &= \frac{1}{2}\left(\Psi(x) + \Psi(s)\right).\end{aligned}$$

This completes the proof of the proposition. □

It is of interest to compare Proposition 2.10 with its SDO analogue Proposition 4.4 in [19]. First we find that statement (ii) in the present paper is slightly different from condition C.4 which is required in [19]. Actually, one can easily see that the matrix used in condition C.4 of [19] satisfies certain conditions similar to one posed in Proposition 2.10. However, the choice of the vector v allowing such conditions in second order cone is much more strict than that for SDO. One possible reason for this phenomena is that, for general $x, s \in K_+$, the Jordan product $x \circ s$ might not belong to K. For instance, let us consider an example in $K \in \Re^3$ with $x = (2+t, 1, 1)^T, s = (2+t, 1, -1)^T \in K_+$ where t is some small positive number. Obviously one has $x \circ s = ((2+t)^2, 4+2t, 0)^T \in K_+$. Moreover, it is trivial to see $\lambda_{\min}(x \circ s) = 2t + t^2$. Thus when t reduces to zero, the function $\Psi(x \circ s)$ goes to infinity. However, one can readily verify that for sufficiently small $t > 0$, both the functions $\Psi(x)$ and $\Psi(s)$ are bounded above. This examples shows that for $x, s \in K_+$, if $\det(x \circ s) \neq \det(x)\det(s)$, then the relation

$$\Psi(x \circ s) \leq \frac{1}{2}\left(\Psi(x^2) + \Psi(s^2)\right)$$

might fail. We also mention that, in the SDO case, for any positive definite matrices X and S, since the matrix XS is diagonaliziable and has positive eigenvalues, the function $\Psi(XS)$ is well-defined (see [19]).

Likewise what we have observed in the LO and SDO cases [19], in order to establish the complexity of the algorithm, we need to bound the derivatives of certain proximity functions in

suitable spaces. With specification to SOCO, this requires us to discuss the derivatives of the functions $\psi(x(t))$ and $\Psi(x(t))$ where

$$x(t) = (x_1(t), \cdots, x_n(t))^T$$

is a mapping from $\Re$ into $\Re^n$. Our next result for functions in the second-order cone K resembles Lemma 4.9 for matrix function in [19]. However, our proof for the SOCO case here is much simpler and direct. First, for simplification of expression, let us denote by $x'(t)$ the derivative of $x(t)$ with respect to t such that

$$x'(t) = (x_1'(t), \cdots, x_n'(t))^T.$$

Our result provides means to measure the first-order directional derivative of a general function $\Psi(x(t))$ and bound its second-order derivative with respect to the variable t. Recall that, by (11), we can define the function $\psi'(x)$ as the function whose kernel function is $\psi'(t)$.

Lemma 2.11 *Suppose that $x(t)$ is a mapping from $\Re$ into $\Re^n$. If $x(t)$ is twice differentiable with respect to t for all $t \in (l_t, u_t)$ and $\psi(t)$ is also a twice continuously differentiable function in a suitable domain which contains $\lambda_{\max}(x(t))$ and $\lambda_{\min}(x(t))$, then*

$$\frac{d}{dt}\mathbf{Tr}\left(\psi(x(t))\right) = \mathbf{Tr}\left(\psi'(x(t)) \circ x'(t)\right), \forall t \in (l_t, u_t);$$

and

$$\frac{d^2}{dt^2}\mathbf{Tr}\left(\psi(x(t))\right) \leq \varpi \mathbf{Tr}\left(x'(t) \circ x'(t)\right) + \mathbf{Tr}\left(\psi'(x(t)) \circ x''(t)\right) \tag{22}$$

where

$$\varpi = \max\{\left|\psi''(\lambda_{\max}(x(t)))\right|, \left|\psi''(\lambda_{\min}(x(t)))\right|, \frac{\left|\psi'(\lambda_{\max}(x(t))) - \psi'(\lambda_{\min}(x(t)))\right|}{2\left\|x_{2:n}(t)\right\|}\}.$$

Proof: Without loss of generality, we assume that $\|x_{2:n}\| > 0$. From Lemma 2.6 we obtain

$$\mathbf{Tr}\left(\psi(x(t))\right) = \psi(\lambda_{\max}(x(t))) + \psi(\lambda_{\min}(x(t))).$$

It follows

$$\begin{aligned}\frac{d}{dt}\mathbf{Tr}\left(\psi(x(t))\right) &= \psi'(\lambda_{\max}(x(t)))\left(x_1'(t) + \frac{1}{\|x_{2:n}(t)\|}\sum_{i=2}^{n} x_i(t)x_i'(t)\right) \\ &\quad + \psi'(\lambda_{\min}(x(t)))\left(x_1'(t) - \frac{1}{\|x_{2:n}(t)\|}\sum_{i=2}^{n} x_i(t)x_i'(t)\right).\end{aligned}$$

Now recalling Definition (11), we obtain

$$\psi'(x(t)) = \left(\frac{1}{2}\left(\psi'\left(\lambda_{\max}(x(t))\right) + \psi'\left(\lambda_{\min}(x(t))\right)\right), \frac{\psi'\left(\lambda_{\max}(x(t))\right) - \psi'\left(\lambda_{\min}(x(t))\right)}{2\left\|x_{2:n}(t)\right\|}x_{2:n}(t)\right)^T.$$

By simple calculus, from the above two equalities one can readily check that

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

This proves the first statement of the lemma.

To prove the second statement of the lemma, we first observe that

$$\frac{d^2}{dt^2}\Psi(x(t)) = \mathbf{Tr}\left(\frac{d}{dt}\psi'(x(t)) \circ x'(t)\right) + \mathbf{Tr}\left(\psi'(t) \circ x''(t)\right).$$

It is straightforward to check that

$$\frac{d}{dt}\psi'(x(t)) \ = v1 + v2 + v3$$

where

$$v1 = \frac{\psi''\left(\lambda_{\max}(x(t))\right)\left(x_1'(t) + \frac{\sum_{i=2}^n x_i(t)x_i'(t)}{\|x_{2:n}(t)\|}\right)}{2}\left(1, \frac{x_{2:n}(t)}{\|x_{2:n}(t)\|}\right)^T;$$

$$v2 = \frac{\psi''\left(\lambda_{\min}(x(t))\right)\left(x_1'(t) - \frac{\sum_{i=2}^n x_i(t)x_i'(t)}{\|x_{2:n}(t)\|}\right)}{2}\left(1, -\frac{x_{2:n}(t)}{\|x_{2:n}(t)\|}\right)^T;$$

$$v3 = \frac{\psi'\left(\lambda_{\max}(x(t))\right) - \psi'\left(\lambda_{\min}(x(t))\right)}{2\|x_{2:n}(t)\|}\left(0, x_{2:n}'(t) - \frac{\sum_{i=2}^n x_i(t)x_i'(t)}{\|x_{2:n}(t)\|^2}x_{2:n}(t)\right)^T.$$

Through using the well-known Cauchy-Schwarz inequality, we deduce

$$\frac{\sum_{i=2}^n x_i(t)x_i'(t)}{\|x_{2:n}(t)\|} \le \|x_{2:n}'(t)\|.$$

This relation, together with the definition of ϖ further implies

$$\begin{aligned}\mathbf{Tr}\left((v1+v2)\circ x'(t)\right) \ &= \psi''\left(\lambda_{\max}(x(t))\right)\left(x_1'(t) + \frac{\sum_{i=2}^n x_i(t)x_i'(t)}{\|x_{2:n}(t)\|}\right)^2 \\ &\quad + \psi''\left(\lambda_{\min}(x(t))\right)\left(x_1'(t) - \frac{\sum_{i=2}^n x_i(t)x_i'(t)}{\|x_{2:n}(t)\|}\right)^2 \\ &\le 2\varpi\left((x_1'(t))^2 + \left(\frac{\sum_{i=2}^n x_i(t)x_i'(t)}{\|x_{2:n}(t)\|}\right)^2\right).\end{aligned}$$

On the other hand, through simple calculus, one has

$$\begin{aligned}\mathbf{Tr}\left(v3\circ x'(t)\right) \ &= \frac{\psi'\left(\lambda_{\max}(x(t))\right) - \psi'\left(\lambda_{\min}(x(t))\right)}{\|x_{2:n}(t)\|}\left(\|x_{2:n}'(t)\|^2 - \left(\frac{\sum_{i=2}^n x_i(t)x_i'(t)}{\|x_{2:n}(t)\|}\right)^2\right) \\ &\le 2\varpi\left(\|x_{2:n}'(t)\|^2 - \left(\frac{\sum_{i=2}^n x_i(t)x_i'(t)}{\|x_{2:n}(t)\|}\right)^2\right).\end{aligned}$$

Adding the above two inequalities together, we obtain the desired relation (22). This completes the proof of the lemma. □

It is worthwhile to consider the special case that $K \in \Re^2$ where we can also cast a SOCO problem as an SDO problem. Note that for the SDO case, one has

$$X(t) = \begin{pmatrix} x_1(t) & x_2(t) \\ x_2(t) & x_1(t) \end{pmatrix}.$$

In this situation, the equalities

$$\|X'(t)\|^2 = 2\|x'(t)\|^2 = \mathbf{Tr}\left(x'(t)\circ x'(t)\right)$$

hold trivially. Recalling the difference between the definitions of $\Psi(x)$ by (17) and $\Psi(X)$ in [19], one can easily verify that the estimations given in Lemma 2.11 are precisely the same as that presented in its SDO analogue Lemma 4.9 of [19].

3 Self-regular Proximities and New Search Directions for SOCO

3.1 Preliminaries

In the present section we consider diverse search directions used in IPMs for solving SOCO and introduce some new search directions based on *self-regular* proximities in the second-order cone.

Most IPMs for solving SOCO employ different search directions together with suitable strategies to follow the central path appropriately. Similar to the SDO case, the search directions for SOCO are usually derived from certain Newton systems in various scaled spaces. To address this issue more clearly, we need to go into a little more details. First we note that for the Jordan algebra $\circ$, x and s commute, i.e., $x \circ s = s \circ x$. This is different from the SDO situation where the matrix pair X and S do not commute in general. From this fact, one might be encouraged to guess that the Newton system for SOCO is well-defined if both $x \succ_K 0$ and $s \succ_K 0$ are feasible for SOCO. In the sequel it will be shown that this conjecture might not be true[5].

Consider the following linearized Newton system for (3):

$$\begin{pmatrix} A & 0 & 0 \\ 0 & E_n & A^T \\ \mathrm{mat}\,(s) & \mathrm{mat}\,(x) & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta s \\ \Delta y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \mu\tilde{e} - \mathrm{mat}\,(x)\, s \end{pmatrix}, \quad x, s \succ_K 0. \tag{23}$$

It follows readily that the system (23) is well defined if and only if the Jacobian matrix in it is nonsingular. For simplification of expression, let us denote temporarily $X = \mathrm{mat}\,(x)$ and $S = \mathrm{mat}\,(s)$. We also denote by E_m and E_n the identity matrices in the suitable spaces. Now by using simple multiplications of matrices, one has

$$\begin{pmatrix} A & 0 & 0 \\ 0 & E_n & A^T \\ S & X & 0 \end{pmatrix} \begin{pmatrix} E_n & -S^{-1}X & 0 \\ 0 & E_n & 0 \\ 0 & 0 & E_m \end{pmatrix} \begin{pmatrix} E_n & 0 & 0 \\ 0 & E_n & -A^T \\ 0 & 0 & E_m \end{pmatrix} = \begin{pmatrix} A & -AS^{-1}X & AS^{-1}XA^T \\ 0 & E_n & 0 \\ S & 0 & 0 \end{pmatrix}.$$

This relation implies immediately

Lemma 3.1 *The system (23) is well defined if and only if the matrix*

$$A\ S^{-1}XA^T$$

is nonsingular.

Recall that in the introduction of this paper, we have assumed that the matrix A is of full row rank. As consequence, if the matrix $S^{-1}X$ is positive definite, then the matrix $A\ S^{-1}XA^T$ is also positive definite and thus nonsingular. For LO, if the primal-dual pair x and s are both strictly feasible, then the matrix $S^{-1}X$ is diagonal and positive. Hence the Newton system for LO is well defined. In case of SOCO, however, the matrix $A\ S^{-1}XA^T$ may fail to be nonsingular.

[5]One possible explanation for this phenomenon is that, although the Jordan algebra $\circ$ commute, it indeed does not satisfies the associative law. For example, let $x = (3,1,2)^T, s = (3,-1,-2)^T$ and $z = (1,2,1)^T$, one can easily check that $(x \circ s) \circ z \neq x \circ (s \circ z)$. On the other hand, the general multiplication operator for matrices is not commutative but associative, i.e., for any $X, S, Z \in \Re^{n\times n}$, there holds $(XS)Z = X(SZ)$. We also mention that, as it has been shown in our discussions after Proposition 2.10 in the last section, for some $x, s \succ_K 0$, the Jordan product $x \circ s$ might not be in K.

The reason is that if Z is a matrix with positive eigenvalues, then the symmetric matrix $Z + Z^T$ may fail to be positive definite. We show this by an example. Consider the following SOCO problem with

$$A = (0, \sqrt{3.69} + 0.7, 1), \quad K = \{x \in \Re^3, x_1 \geq \sqrt{x_2^2 + x_3^2}\}.$$

Assuming we have a feasible primal-dual pair (x, s) with $x = (1, 0.8, 0.5)^T, s = (1, 0.7, 0.7)^T$. Obviously there holds $x, s \succ_K 0$. However, by simple computation, one can easily verify that $A\ S^{-1}XA^T = 0$. This illustrates that the system (23) is not well defined for this specific example.

3.2 Scaling Schemes

As we have observed in the aforementioned example, the linearized Newton system might not be well defined. To obtain a Newton-type system that has a unique solution, people usually refer to some scaling schemes. In the sequel we will introduce certain variants of such scaling schemes for SOCO, as first proposed and studied by Tsuchiya [22, 23]. In the rest of this section, we consider the more general case where $N > 1$. In this situation, the definitions $\psi(x)$, $\Psi(x)$ and Jordan algebra should also be modified accordingly as follows.

$$\psi(x) \ = (\psi(x^1), \psi(x^2), \cdots, \psi(x^N))^T, \quad \Psi(x) = \sum_{i=1}^{N} \Psi(x^i); \tag{24}$$

$$x \circ s \ = \left(\left(x^1 \circ s^1\right)^T, \left(x^2 \circ x^2\right)^T, \cdots, \left(x^N \circ s^N\right)^T \right)^T. \tag{25}$$

Before we move to discuss these scaling techniques, recall from Section 2 that closely associated with each cone K^j are the matrices

$$E_{n_j} := \mathrm{diag}\,(1, 1, \cdots, 1); \quad Q^j = \mathrm{diag}\,(1, -1, \cdots, -1)$$

where the first one E_{n_j} denotes the identity matrix in space $\Re^{n_j \times n_j}$ and the later Q^j is the representation matrix of the cone K^j since

$$K^j = \{x^j \in \Re^{n_j} : (x^j)^T Q^j x^j \geq 0,\ x_1^j \geq 0\}.$$

It is trivial to see $\left(Q^j\right)^2 = E_{n_j}$.

Now we are ready to give the definition of a scaling matrix for general second-order cones.

Definition 3.2 *A matrix $W^j \in \Re^{n_j \times n_j}$ is a scaling matrix for the cone K^j if it satisfies the following condition*

$$W^j Q^j W^j = Q^j, \quad W^j \succ 0.$$

We remind the reader that here $W^j \succ 0$ means that W^j is positive definite and symmetric. In view of this definition, if W^j is a scaling matrix for the cone K^j, so is $\left(W^j\right)^{-1}$.

A scaled pair $(\tilde{x}, \tilde{s})$ is obtained by the transformation

$$\tilde{x} = UWx, \qquad \tilde{s} := (UW)^{-1} s,$$

where

$$W := \mathrm{diag}\left(W^1, W^2, \cdots, W^N\right), \quad U := \mathrm{diag}\left(u_1 E_{n_1}, u_2 E_{n_2}, \cdots, u_N E_{n_N}\right), \quad u_1, \cdots, u_N > 0.$$

Several elementary properties of such a transformation are summarized in the following proposition.

Proposition 3.3 *For any $j \in \{1, 2, \cdots, N\}$, we have*
(i): $\mathbf{Tr}\left(x^j \circ s^j\right) = \mathbf{Tr}\left(\tilde{x}^j \circ \tilde{s}^j\right)$;
(ii): $u_j^2 \det\left(x^j\right) = \det\left(\tilde{x}^j\right), \qquad u_j^{-2} \det\left(s^j\right) = \det\left(\tilde{s}^j\right)$;
(iii): $x \succ_K 0$ *(or* $x \succeq_K 0$*) if and only if* $\tilde{x} \succ_K 0$ *(or* $\tilde{x} \succeq_K 0$*).*

Proof: The proof follows directly from the definition of scaling matrices. We refer the details to [4, 22]. □

Letting us define

$$\widetilde{A} = A\left(UW\right)^{-1}, \quad \tilde{c} = \left(UW\right)^{-T} c.$$

One can rewrite the system (3) in the scaled space as

$$\begin{cases} \widetilde{A}\tilde{x} & = b, \\ \widetilde{A}^T y + \tilde{s} & = \tilde{c}, \\ \mathrm{mat}\left(\tilde{x}\right)\tilde{s} & = \mu\tilde{e}, \quad \tilde{x}, \tilde{s} \succeq_K 0, \end{cases} \tag{26}$$

If both x and s are strictly feasible for (SOCO) and (SOCD), so are the vectors $\tilde{x}$ and $\tilde{s}$ for the new SOCO problem in the scaled space. In this case, the linearized Newton system for (26) amounts to solve the following equation system

$$\begin{cases} \widetilde{A}\tilde{d}_x & = & 0, \\ \widetilde{A}^T d_y + \tilde{d}_s & = & 0, \\ \mathrm{mat}\left(\tilde{x}\right)\tilde{d}_s + \mathrm{mat}\left(\tilde{s}\right)\tilde{d}_x & = \mu\tilde{e} - \mathrm{mat}\left(\tilde{x}\right)\tilde{s}, & \tilde{x}, \tilde{s} \succeq_K 0, \end{cases} \tag{27}$$

There are several popular choices for the scaling matrices W^j and the constants u_j. For instance, if UW is the identity matrix and the system (27) is well-defined, then the solution of (27) yields the so-called A.H.O. search direction [1]; if UW is chosen such that $\tilde{s} = \tilde{e}$ (or $\tilde{x} = \tilde{e}$), then we obtain the primal (or dual) H.K.M. direction; if UW is chosen such that $\tilde{x} = \tilde{s}$, then we have the NT search direction [22, 23]. Note that when the NT-scaling is used, since $\tilde{x} = \tilde{s}$, the matrix in the scaled space $\widetilde{A}\mathrm{mat}\left(\tilde{s}\right)^{-1}\mathrm{mat}\left(\tilde{s}\right)\widetilde{A}^T = \widetilde{A}\widetilde{A}^T$ is positive definite, thus the system (27) is well defined. In [13], Monteiro and Tsuchiya studied some other search directions for SOCO as well.

3.3 A Variational Principle for Scaling

When we choose the scaling scheme, the desirability of choosing a scaling such that the corresponding proximity takes a minimal value in the scaled space is obvious. In the sequel we will address the issue which kind of scaling scheme is 'optimal' under certain variational principle. For this we consider the primal SOCO problem in the scaled space

$$\begin{array}{ll} \text{(Scaled SOCO)} & \min \tilde{c}^T\tilde{x} \\ & \widetilde{A}\tilde{x} = b, \qquad \tilde{x} \succeq_K 0, \end{array}$$

and its dual problem

$$\begin{array}{ll} \text{(Scaled SOCD)} & \max b^T y \\ & \widetilde{A}^T y + \tilde{s} = \tilde{c}, \quad \tilde{s} \succeq_K 0. \end{array}$$

We assume that a certain barrier method is employed to solve both the scaled primal and dual problems, namely we minimize a specific potential function $\tilde{c}^T\tilde{x} + \Psi(\tilde{x})$ and maximize $b^T y - \Psi(\tilde{s})$

where $\Psi(\cdot)$ is a barrier function for the second-order cone K. The question arises for which kinds of scaling matrix W and matrix U, the function

$$\tilde{c}^T\tilde{x} - b^T y + \Psi(\tilde{x}) + \Psi(\tilde{s})$$

attains its global minimal value. In the sequel we solve this issue under the assumption that the corresponding kernel function $\psi(\cdot)$ satisfies condition C.2 strictly, which is equivalent to $\psi'(t) + t\psi''(t) > 0$ as shown in [19].

Proposition 3.4 *Suppose the functions $\psi(x)$ and $\Psi(\cdot)$ are defined by (11) and (24). If the function $\psi(t)$ satisfies condition C.2 strictly, then the function $\tilde{c}^T\tilde{x} - b^T y + \Psi(\tilde{x}) + \Psi(\tilde{s})$ attains its global minimal value with matrices W and U such that $\tilde{x} = \tilde{s}$.*

Proof: First we observe that the inner product $\tilde{c}^T\tilde{x}$ is invariant for any nonsingular matrices W and U. Thus we need only to prove that $\Psi(\tilde{x}) + \Psi(\tilde{s})$ has a global minimizer when the matrices W and U are chosen so that $\tilde{x} = \tilde{s}$. We start with a discussion about the existence of such matrices W and U. For any $x, s \succ_K 0$ and $j = 1, 2, \cdots, N$, let us define

$$u_j := \left(\frac{\det(x^j)}{\det(s^j)}\right)^{\frac{1}{4}}; \tag{28}$$

$$w^j := \frac{u_j^{-1}s^j + u_j Q^j x^j}{\sqrt{2}\sqrt{\mathbf{Tr}(x^j \circ s^j) + \sqrt{\det(x^j)\det(s^j)}}}; \tag{29}$$

and

$$W_{NT}^j = \begin{pmatrix} w_1^j & (w_{2:n_j}^j)^T \\ w_{2:n_j}^j & E_{n_j-1} + \frac{1}{1+w_1^j} w_{2:n_j}^j (w_{2:n_j}^j)^T \end{pmatrix} = -Q^j + \frac{1}{1+w_1^j}(\tilde{e}^j + w^j)(\tilde{e}^j + w^j)^T. \tag{30}$$

Note that when $x, s \succ_K 0$, the denominator in the expression of w^j is positive. For the above choices, there holds $\tilde{x}^j = u_j W_{NT}^j x^j = u_j^{-1}\left(W_{NT}^j\right)^{-1} s^j = \tilde{s}^j$ (see [4, 22, 23]). This proves the existence of such a scaling matrix W_{NT} and the matrix $U_{NT} = \mathrm{diag}(u_1 E_{n_1}, \cdots, u_N E_{n_N})$. Hence it remains to show that for these specific choices of W_{NT} and U_{NT}, the value of the function $\Psi(\tilde{x}) + \Psi(\tilde{s})$ is optimal. To distinguish the NT-scaling scheme from many other scaling schemes, we denote by

$$v = \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^N \end{pmatrix} = \tilde{x}_{NT} = \tilde{s}_{NT} = \begin{pmatrix} u_1 W_{NT}^1 x^1 \\ u_2 W_{NT}^2 x^2 \\ \vdots \\ u_N W_{NT}^N x^N \end{pmatrix} = \begin{pmatrix} u_1^{-1}\left(W_{NT}^1\right)^{-1} s^1 \\ u_2^{-1}\left(W_{NT}^2\right)^{-1} s^2 \\ \vdots \\ u_N^{-1}\left(W_{NT}^N\right)^{-1} s^N \end{pmatrix},$$

the scaled vector based on the NT-scaling while $\tilde{x}$ and $\tilde{s}$ denote the scaled vectors using general scaling techniques. It follows

$$\Psi(v) = \frac{1}{2}\sum_{j=1}^{N}\left(\psi\left(\lambda_{\max}(v^j)\right) + \psi\left(\lambda_{\min}(v^j)\right)\right).$$

Thus the proof will be finished if we can show that for any $j = 1, \cdots, N$,

$$\begin{aligned}\psi\left(\lambda_{\max}(v^j)\right) + \psi\left(\lambda_{\min}(v^j)\right) \; &\leq \frac{1}{2}\left(\psi\left(\lambda_{\max}(\tilde{x}^j)\right) + \psi\left(\lambda_{\min}(\tilde{x}^j)\right)\right) \\ &+\frac{1}{2}\left(\psi\left(\lambda_{\max}(\tilde{s}^j)\right) + \psi\left(\lambda_{\min}(\tilde{s}^j)\right)\right) \end{aligned} \tag{31}$$

and the equality is true if and only if $\tilde{x}^j = \tilde{s}^j$. Now by recalling the definitions of the scaling matrices W_{NT} and U_{NT}, we can conclude that for any $j = 1, \cdots, N$, there holds

$$\mathbf{Tr}\left(x^j \circ s^j\right) = \mathbf{Tr}\left(\tilde{x}^j \circ \tilde{s}^j\right) = \mathbf{Tr}\left(v^j \circ v^j\right) = \mathbf{Tr}\left([v^j]^2\right); \tag{32}$$

$$\det\left(v^j\right) = u_j^2 \det\left(x^j\right) = \sqrt{\det\left(x^j\right)\det\left(s^j\right)} = \sqrt{\det\left(\tilde{x}^j\right)\det\left(\tilde{s}^j\right)}. \tag{33}$$

Thus the vector v^j satisfies the requirements in the second statement of Proposition 2.10 where x and s are replaced by $\tilde{x}^j, \tilde{s}^j$, respectively. Progressing in a similar vein as we have done in the proof of the second statement of Proposition 2.10, we can obtain the desired relation (31), which further concludes the proof of the proposition. □

We remark that, as observed by Tsuchiya [23], a large-update IPM for SOCO based on the NT search direction always has a theoretically lower iteration bound than the large-update IPMs relying on other search directions.

3.4 New Proximities and Search Directions

As a prelude to our new search direction, we next proceed to describe the NT-direction for SOCO. Suppose the current pair (x, s) are strictly feasible for both the primal problem (SOCO) and its dual (SOCD), the so-called NT direction for SOCO can be defined as follows. Let us denote

$$\bar{A} := \frac{1}{\sqrt{\mu}} A\left(U_{NT}W_{NT}\right)^{-1}, \qquad v = \frac{1}{\sqrt{\mu}} U_{NT}W_{NT}x; \tag{34}$$

$$d_x = \frac{1}{\sqrt{\mu}} U_{NT}W_{NT}\Delta x, \qquad d_s = \frac{1}{\sqrt{\mu}}\left(U_{NT}W_{NT}\right)^{-1}\Delta s. \tag{35}$$

Obviously there holds $v \succ_K 0$. The NT search direction for SOCO is defined as the unique solution of the following system [23, 4]

$$\begin{cases} \bar{A}d_x & = & 0, \\ \bar{A}^T\Delta y + d_s & = & 0, \\ d_s + d_x & = & v^{-1} - v. \end{cases} \tag{36}$$

Before introducing the new search direction for SOCO, let us define the proximity measure used in our new IPM for SOCO first. Similar to the cases of LO and SDO, the new proximity measures for SOCO is given by

$$\Psi(x, s, \mu) := \Psi(v) = \mathbf{Tr}\left(\psi(v)\right), \tag{37}$$

where $\psi(\cdot)$ is a univariate *self-regular* function.

The new search direction that we propose for SOCO is a slight modification of the NT direction defined by the solution of the following system

$$\begin{aligned} \bar{A}d_x &= 0, \\ -\bar{A}^T\Delta y &= d_s, \\ d_x + d_s &= -\psi'(v). \end{aligned} \tag{38}$$

Once we get d_x and d_s, then one can compute Δx and Δs via (35). In view of the orthogonality of Δx and Δs, one can easily verify that

$$d_x^T d_s = 0. \tag{39}$$

We proceed to discuss versatile properties of these *self-regular* proximities for SOCO. For this we first introduce some notations. Let us denote by σ^2 the trace of the vector $\psi'(v)^2$ where $\psi'(v)$ is the vector at the right-hand side in the last equation of the system (38). Therefore we have

$$\sigma^2 = \mathbf{Tr}\left(\psi'(v) \circ \psi'(v)\right) = 2\left\|\psi'(v)\right\|^2; \quad \sigma = \sqrt{2}\left\|\psi'(v)\right\|. \tag{40}$$

To facilitate the foregoing analysis, we also define

$$\lambda_{\max}(v) = \max\{\lambda_{\max}(v^j) : j = 1, \cdots, N\}; \quad \lambda_{\min}(v) = \min\{\lambda_{\min}(v^j) : j = 1, \cdots, N\}. \tag{41}$$

Recall that by Lemma 2.6, we can deduce

$$\sigma^2 = \sum_{j=1}^{N} \left(\left(\psi'(\lambda_{\max}(v^j))\right)^2 + \left(\psi'(\lambda_{\min}(v^j))\right)^2 \right).$$

By using this relation and taking a similar chain of reasoning as in the proof of Proposition 3.3 of [19], we can prove the following results which include several features of the proximity. These properties are naturally shared by general *self-regular* functions in the second-order cone K.

Proposition 3.5 *Let the proximity $\Psi(v)$ be defined by (37) and σ by (40). If the kernel function $\psi(\cdot)$ used in the proximity satisfies condition C.1, then we have*

$$\Psi(v) \leq \frac{\sigma^2}{2\nu_1}, \tag{42}$$

$$\lambda_{\min}(v) \geq \left(1 + \frac{q\sigma}{\nu_1}\right)^{-\frac{1}{q}}, \tag{43}$$

and

$$\lambda_{\max}(v) \leq \left(1 + \frac{p\sigma}{\nu_1}\right)^{\frac{1}{p}}. \tag{44}$$

If $\lambda_{\max}(v) > 1$ and $\lambda_{\min}(v) < 1$, then

$$\sigma \geq \nu_1 \left(\frac{(\lambda_{\max}(v)^p - 1)^2}{p^2} + \frac{(\lambda_{\min}(v)^{-q} - 1)^2}{q^2} \right)^{\frac{1}{2}}. \tag{45}$$

For any $\vartheta > 1$,

$$\Psi(\vartheta v) \ \leq \frac{\nu_2}{\nu_1}\left(\vartheta^{p+1}\Psi(v) + 2\vartheta\Upsilon'_{p,q}(\vartheta)\sqrt{N\nu_1\Psi(v)} + 2N\nu_1\Upsilon_{p,q}(\vartheta)\right). \tag{46}$$

Particularly there exist two constant ν_3 and ν_4 depending only on the kernel function $\psi(t)$ such that for any $\vartheta \in (1, 1+\nu_3]$, we have

$$\Psi(\vartheta v) \leq \frac{\nu_2\nu_4}{\nu_1}\Psi(v) + \frac{2\nu_2\nu_4\sqrt{N\nu_1\Psi(v)}}{\nu_1}(\vartheta - 1) + 2N\nu_2\nu_4(\vartheta - 1)^2. \tag{47}$$

We close this section by discussing the relations between the duality gap and the proximity. By following a similar chain of reasoning as in the proof of Lemma 2.12 in [19], we can easily deduce

$$\begin{aligned}\frac{\Psi(v)}{\nu_1} \ &\geq \frac{1}{2}\sum_{j=1}^{N}\left(\left(\lambda_{\max}(v^j) - 1\right)^2 + \left(\lambda_{\min}(v^j) - 1\right)^2\right) \\ &= \|v\|^2 - \sum_{j=1}^{N}\left(\lambda_{\max}(v^j) + \lambda_{\min}(v^j)\right) + N \geq \|v\|^2 - 2\sqrt{N}\,\|v\| + N.\end{aligned}$$

The above relation means

$$\|v\| \leq \sqrt{N} + \sqrt{\frac{\Psi(v)}{\nu_1}}.$$

It readily follows

$$\mathbf{Tr}\,(x \circ s) = 2\mu\,\|v\|^2 \leq 2N\mu + 4\mu\sqrt{\frac{N\Psi(v)}{\nu_1}} + \frac{2\Psi(v)}{\nu_1}\mu. \tag{48}$$

Therefore, $\mathbf{Tr}\,(x \circ s) \leq \mathcal{O}\,(N\mu)$ holds whenever $\Psi(v) \leq \mathcal{O}(N)$. In such a situation, the proximity plays the role of a potential function for minimizing the duality gap.

4 Polynomial Primal-Dual Algorithms for SOCO

4.1 The Algorithm

The present section commits to describe the new primal-dual algorithm for solving SOCO. At the start we would like to mention that, by using the NT-scaling, one can rewrite the centrality condition for SOCO as $v = \tilde{e}$. Consequently the neighborhood of the central path used in our new algorithm is also dependent on v. Denote

$$\mathcal{F}_{SOCO} = \{(x, s) \in K \times K : Ax = b; A^T y + s = c\}.$$

We define the neighborhood of the central path as below:

$$\mathcal{N}(\tau, \mu) = \{(x, s) : (x, s) \in \mathcal{F}_{SOCO}, \Psi(x, s, \mu) = \Psi(v) \leq \tau\}. \tag{49}$$

Assuming that a starting point in a certain neighborhood of the central path is available (actually by using the so-called self-dual embedding model, one can further get as an initial point the point on the central path corresponding to $\mu = 1$, see [11, 24]), we can set out from this point. By reducing μ properly and solving the system (38), one obtains a search direction. Then the iterate

can be updated by means of line search. If the current iterate goes outside the neighborhood of the targeted point $\mu\tilde{e}$, then we will utilize the inner iterations to get a new iterate in the neighborhood. Otherwise we progress with the outer iteration and update μ by a fixed factor. The algorithm will stop when the duality gap, bounded by a multiple of μ, is sufficiently small and hence an approximate solution of the undertaking problem is presented. The procedure of the new algorithm is outlined as follows.

Primal-Dual Algorithm for SOCO

Input:
A proximity parameter $\tau \geq \nu_1^{-1}$;
an accuracy parameter $\varepsilon > 0$;
a fixed barrier update parameter $\theta \in (0, 1)$;
a strictly feasible (x, s) and $\mu = 1$ such that $\Psi(x, s, \mu) \leq \tau$.

begin
while $N\mu \geq \varepsilon$ **do**
begin
$\mu := (1 - \theta)\mu$;
while $\Psi(x, s, \mu) \geq \tau$ **do**
begin
Solve the system (38),
Compute a step size α;
$x := x + \alpha\Delta x$;
$s := s + \alpha\Delta s$;
$y := y + \alpha\Delta y$;
end
end
end

Remark 4.1 *The algorithm will stop when an iterate satisfies $N\mu \leq \varepsilon$ and $\Psi(x, s, \mu) \leq \tau$. By recalling (48) we can claim*

$$x^T s = \frac{1}{2}\mathbf{Tr}\,(x \circ s) \leq N\mu + 2\mu\sqrt{\frac{N\tau}{\nu_1}} + \mu\frac{\tau}{\nu_1}.$$

For instance, let us choose the parameter $\tau = N$ and the proximity satisfying condition C.1 with $\nu_1 = 1$. In such a case, the algorithm works indeed in a large neighborhood of the central path. One can easily verify that the algorithm will finally report a solution satisfying $x^T s \leq 4\varepsilon$.

4.2 Complexity of the Algorithm

Having stated the algorithm in the previous section, we are going to establish the polynomial complexity of the algorithm in the present section. As we have already observed in [19] for LO and SDO, a crucial step in the estimate of the algorithm's complexity is to evaluate how fast we can reduce the value of the proximity for a feasible step size along the search direction.

Note that once the search direction $(\Delta x, \Delta s)$ is obtained, we need to decide how far we can go along this direction while staying in the feasible region, this amounts to estimate the maximal

feasible step size. It should be noticed that for any step size α, the primal-dual pair $(x + \alpha\Delta x, s + \alpha\Delta s)$ is feasible if and only if the scaled primal-dual pair $(v + \alpha d_x, v + \alpha d_s)$ (see (iii) of Proposition 3.3) is feasible. In the sequel we give a certain sufficient condition for a step size to be strictly feasible and thus provide a lower bound for the maximal step size. To facilitate the analysis, for any $x^j \in \Re^{n_j}, j = 1, \cdots, N$, we define

$$\lambda_{\max}(\left|x^j\right|) = \left|x_1^j\right| + \left\|x_{2:n_j}^j\right\|, \qquad \lambda_{\min}(\left|x^j\right|) = \left|x_1^j\right| - \left\|x_{2:n_j}^j\right\|;$$

and

$$\lambda_{\max}(|x|) = \max\{\lambda_{\max}(\left|x^j\right|) : j = 1, \cdots, N\}, \qquad \lambda_{\min}(|x|) = \min\{\lambda_{\min}(\left|x^j\right|) : j = 1, \cdots, N\}.$$

A direct consequence of the above definitions is

$$\frac{1}{2}\sum_{j=1}^{N}\left(\lambda_{\max}(\left|x^j\right|)^2 + \lambda_{\min}(\left|x^j\right|)^2\right) = \|x\|^2, \quad x = (x^1, \cdots, x^N)^T. \tag{50}$$

Now we have

Lemma 4.2 *Let $\alpha_{\max}$ be the maximal feasible step-size and*

$$\bar{\alpha} = \lambda_{\min}(v)\sigma^{-1}. \tag{51}$$

Then we have

$$\alpha_{\max} \geq \bar{\alpha} \geq \sigma^{-1}\left(1 + \frac{q\sigma}{\nu_1}\right)^{-\frac{1}{q}}.$$

Proof: First we observe that the primal-dual pair $(v + \alpha d_x, v + \alpha d_s)$ is strictly feasible if and only if for any $j = 1, \cdots, N$,

$$v^j + \alpha d_x^j \succeq_{K^j} 0; \quad v^j + \alpha d_s^j \succeq_{K^j} 0.$$

For any fixed j, it is trivial to see that

$$\left(v^j + \alpha d_x^j\right)_1 - \left\|\left(v^j + \alpha d_x^j\right)_{2:n_j}\right\| \geq v_1^j - \left\|v_{2:n_j}^j\right\| - \alpha\lambda_{\max}(\left|d_x^j\right|) \tag{52}$$

$$\geq \lambda_{\min}(v) - \alpha\lambda_{\max}(|d_x|) \geq \lambda_{\min}(v) - \sqrt{2}\alpha\|d_x\|, \tag{53}$$

where the last inequality follows from (50). Now recalling the orthogonality relation (39), we can deduce

$$\|d_x\| \leq \|(d_x, d_s)\| = \sqrt{\|d_x\|^2 + \|d_s\|^2} = \frac{1}{\sqrt{2}}\sigma.$$

From the inequality (43) in Proposition 3.5 we know that for any $\alpha \in [0, \bar{\alpha}]$, there holds

$$\lambda_{\min}(v^j + \alpha d_x) \geq 0,$$

which is equivalent to $v^j + \alpha d_x^j \succeq_{K^j} 0$. Similarly one can show that for any $\alpha \in [0, \bar{\alpha}]$, there holds $v^j + \alpha d_s \succeq_{K^j} 0$. This completes the proof of the lemma. □

In view of Lemma 4.2, it is clear that we can use any $\alpha \in (0, \bar{\alpha})$ as a step size. Note that, after such a step, we get a new primal-dual pair $(x+\alpha\Delta x, s+\alpha\Delta s)$ or the scaled pair $(v+\alpha d_x, v+\alpha d_s)$ and then we need to use the NT-scaling scheme to transform the primal and dual vectors to the same vector, which we denote by v^+. On the other hand, according to (37), the proximity after

this step is defined as $\Psi(v^+)$. Let us denote the gap between the proximity before and after one step as a function of the step-size, that is

$$g(\alpha) = \Psi(v^+) - \Psi(v). \tag{54}$$

The main task in the rest of this section is to study the decreasing behavior of $g(\alpha)$ for $\alpha \in [0, \bar{\alpha})$. Since v^+ is the the scaled vector by using the NT-scaling, from Proposition 3.3 we conclude that

$$\det\left(((v^+)^j)^2\right) = \det(x + \alpha\Delta x)\det(s + \alpha\Delta s) = \det\left(v^j + \alpha d_x^j\right)\det\left(v^j + \alpha d_s^j\right), \quad j = 1, \cdots, N,$$

and

$$\mathbf{Tr}\left(((v^+)^j)^2\right) = \mathbf{Tr}\left((x + \alpha\Delta x) \circ (s + \alpha\Delta s)\right) = \mathbf{Tr}\left((v^j + \alpha d_x^j) \circ (v^j + \alpha d_s^j)\right), \quad j = 1, \cdots, N.$$

Thus for any $j = 1, \cdots, N$, the vectors $(v^+)^j, v^j + \alpha d_x^j$ and $v^j + \alpha d_x^j$ satisfy the requirement in the second statement of Proposition 2.10. Therefore, when the kernel function $\psi(\cdot)$ in (37) is self-regular, it follows readily from the second statement of Proposition 2.10 that

$$g(\alpha) = \Psi(v^+) - \Psi(v) \leq \frac{1}{2}\left(\Psi\left(v + \alpha d_x\right) + \Psi(v + \alpha d_s)\right) - \Psi(v) =: g_1(\alpha).$$

In the sequel we estimate the decrement of the function $g_1(\alpha)$ for $\alpha \in [0, \bar{\alpha})$. For our specific purpose, we will first estimate the first and second derivatives of $g_1(\alpha)$. From Lemma 2.11 it follows

$$g_1'(\alpha) = \frac{1}{2}\mathbf{Tr}\left(\psi'(v + \alpha d_x) \circ d_x + \psi'(v + \alpha d_s) \circ d_s\right), \tag{55}$$

and

$$g_1''(\alpha) = \frac{1}{2}\frac{d^2}{d\alpha^2}\mathbf{Tr}\left(\psi(v + \alpha d_x) + \psi(v + \alpha d_s)\right). \tag{56}$$

The next result presents an upper bound for the second-order derivatives of $g_1(\alpha)$. This result plays a crucial role in establishing the polynomial complexity of the algorithm.

Lemma 4.3 *Suppose that the kernel function $\psi(\cdot)$ used in (37) is self-regular. Then for any $\alpha \in (0, \bar{\alpha})$, there holds*

$$g_1''(\alpha) \leq \frac{1}{2}\nu_2\sigma^2\left((\lambda_{\max}(v) + \alpha\sigma)^{p-1} + (\lambda_{\min}(v) - \alpha\sigma)^{-q-1}\right).$$

Proof: From Lemma 4.2 we know that the step-size used in the lemma is strictly feasible. For any $j = 1, \cdots, N$, let us denote by $v_{x+}^j = v^j + \alpha d_x^j, v_{s+}^j = v^j + \alpha d_s^j$ respectively. By means of using Lemma 2.11 one gets

$$g_1''(\alpha) \leq \varpi_1 \|d_x\|^2 + \varpi_2 \|d_s\|^2,$$

where

$$\varpi_1 = \max_{1\leq j\leq N}\left\{\left|\psi''\left(\lambda_{\max}(v_{x+}^j)\right)\right|, \left|\psi''\left(\lambda_{\min}(v_{x+}^j)\right)\right|, \frac{\left|\psi'\left(\lambda_{\max}(v_{x+}^j)\right) - \psi'\left(\lambda_{\max}(v_{x+}^j)\right)\right|}{2\left\|\left(v_{x+}^j\right)_{2:n}\right\|}\right\};$$

$$\varpi_2 = \max_{1\leq j\leq N}\left\{\left|\psi''\left(\lambda_{\max}(v_{s+}^j)\right)\right|, \left|\psi''\left(\lambda_{\min}(v_{s+}^j)\right)\right|, \frac{\left|\psi'\left(\lambda_{\max}(v_{s+}^j)\right) - \psi'\left(\lambda_{\max}(v_{s+}^j)\right)\right|}{2\left\|\left(v_{s+}^j\right)_{2:n}\right\|}\right\}.$$

Recall that $\sigma^2 = 2\left(\|d_x\|^2 + \|d_s\|^2\right)$, thus the proof is done if one can show that

$$\max\left(\varpi_1, \varpi_2\right) \leq \nu_2(\lambda_{\max}(v) + \alpha\sigma)^{p-1} + (\lambda_{\min}(v) - \alpha\sigma)^{-q-1}. \tag{57}$$

Combining the choice of ϖ_1 with the mean value theorem [20], one can conclude that there exists a constant $\zeta_* \in [\lambda_{\min}(v), \lambda_{\max}(v)]$ such that

$$\varpi_1 = \left|\psi''(\zeta^*)\right|.$$

By using condition C.1 we immediately obtain

$$\varpi_1 \leq \nu_2\left(\zeta_*^{p-1} + \zeta_*^{-q-1}\right). \tag{58}$$

Now recalling the assumption $\alpha \in [0, \bar{\alpha}]$ and taking a similar chain of reasoning as in the proof of Lemma 4.2, we can conclude that for any $j \in \{1, 2, \cdots, N\}$,

$$\lambda_{\min}(v^j) - \alpha\sigma \leq \lambda_{\min}(v^j) - \alpha\lambda_{\max}(|d_x|) \leq \lambda_{\min}(v^j + \alpha d_x^j);$$

and

$$\lambda_{\max}(v^j + \alpha d_x^j) \leq \lambda_{\max}(v^j) + \alpha\lambda_{\max}(|d_x|) \leq \lambda_{\max}(v) + \alpha\sigma.$$

It follows immediately

$$\lambda_{\min}(v) - \alpha\sigma \leq \zeta_* \leq \lambda_{\max}(v) + \alpha\sigma;$$

which, together with (58), yields

$$\varpi_1 \leq \nu_2\left((\lambda_{\max}(v) + \alpha\sigma)^{p-1} + (\lambda_{\min}(v) - \alpha\sigma)^{-q-1}\right).$$

In an analogous vein one can deduce

$$\varpi_2 \leq \nu_2\left((\lambda_{\max}(v) + \alpha\sigma)^{p-1} + (\lambda_{\min}(v) - \alpha\sigma)^{-q-1}\right).$$

The above two inequalities give (57), which further gives the statement of the lemma. □

The remaining discussions in this section follow a very similar procedure as in the LO and SDO cases. First we observe that by applying Lemma 2.11 to the function $g(\alpha)$, we readily claim

$$g'(0) = g_1'(0) = -\frac{\sigma^2}{2}.$$

From Lemma 4.3 it follows that

$$g(\alpha) \leq g_1(\alpha) \leq -\frac{\sigma^2\alpha}{2} + \frac{1}{2}\nu_2\sigma^2\int_0^\alpha\int_0^\xi\left((\lambda_{\max}(v) + \zeta\sigma)^{p-1} + (\lambda_{\min}(v) - \zeta\sigma)^{-q-1}\right)d\zeta d\xi,$$

which is essentially the same as its LO analogue (the relation (45) in [19]) where the variables $v_{\max}, v_{\min}$ are replaced by $\lambda_{\max}(v)$ and $\lambda_{\min}(v)$, respectively. Let us denote by

$$g_2(\alpha) := -\frac{\sigma^2\alpha}{2} + \frac{1}{2}\nu_2\sigma^2\int_0^\alpha\int_0^\xi\left((\lambda_{\max}(v) + \zeta\sigma)^{p-1} + (\lambda_{\min}(v) - \zeta\sigma)^{-q-1}\right)d\zeta d\xi.$$

It is straightforward to verify that $g_2(\alpha)$ is strictly convex and twice differentiable for all $\alpha \in [0, \bar{\alpha})$. Let α^* be the unique global minimizer of $g_2(\alpha)$ in the interval $[0, \bar{\alpha})$, namely

$$\alpha^* = \arg\min_{0 \leq \alpha < \bar{\alpha}} g_2(\alpha), \tag{59}$$

or equivalently α^* is the unique solution of the following equation

$$-\sigma + \frac{\nu_2}{p}\left(\lambda_{\max}(v) + \alpha^*\sigma\right)^p - \lambda_{\max}(v)^p) + \frac{\nu_2}{q}\left(\lambda_{\min}(v) - \alpha^*\sigma\right)^{-q} - \lambda_{\min}(v)^{-q}) = 0. \tag{60}$$

For this choice of α^*, by applying Lemma 3.12 of [19], we can readily claim that

$$g(\alpha^*) \leq g_2(\alpha^*) \leq \frac{1}{2}g_2'(0)\alpha^* = \frac{1}{2}g'(0)\alpha^*. \tag{61}$$

Thus it remains to estimate the value of α^*. One has

Lemma 4.4 *Let the constant α^* be defined by (59). Suppose that $\Psi(v) \geq \nu_1^{-1}$ and $v_{\max} > 1$ and let*

$$\nu_5 = \min\{\frac{\nu_1}{2\nu_1\nu_2 + p(\nu_1 + 2\nu_2)}, \frac{\nu_1^2}{(1+\nu_1)\left(2\nu_2(\nu_1+q)+\nu_1 q\right)}\}. \tag{62}$$

Then

$$\alpha^* \geq \nu_5\sigma^{-\frac{q+1}{q}} \tag{63}$$

holds. In the special case where $\psi(t) = \Upsilon_{p,q}(t)$ is given by (16) with $\nu_1 = \nu_2 = 1$, the above bound simplifies to

$$\alpha^* \geq \min\left(\frac{1}{3p+2}, \frac{1}{4+6q}\right)\sigma^{-\frac{q+1}{q}}. \tag{64}$$

Proof: See the proof of Lemma 3.11 of [19]. □

Our following result estimates the decreasing value of the proximity in case that the step-size α is given by α^* (59) or $\alpha = \nu_5\sigma^{\frac{q-1}{q}}$. The proof of the theorem is analogous to that of its LO counterpart, thus the details are omitted here.

Theorem 4.5 *Let the function $g(\alpha)$ be defined by (54) with $\Psi(v) \geq \nu_1^{-1}$. Then the step-size given by $\alpha = \alpha^*$ (59) or $\alpha = \nu_5\sigma^{\frac{q-1}{q}}$ is strictly feasible. Moreover there holds*

$$g(\alpha) \leq \frac{1}{2}g'(0)\alpha \leq -\frac{\nu_5\nu_1^{\frac{q-1}{2q}}}{4}\Psi(v)^{\frac{q-1}{2q}}.$$

In the special case where $\psi(t) = \Upsilon_{p,q}(t)$ with $\nu_1 = \nu_2 = 1$, the above bound simplifies to

$$g(\alpha) \leq -\min\left(\frac{1}{12p+8}, \frac{1}{24q+16}\right)\Psi(v)^{\frac{q-1}{2q}}.$$

To get the total complexity result of the algorithm, we still need to describe the growth behavior of the proximity $\Psi(v)$. Suppose that the current point is in the neighborhood $\mathcal{N}(\mu,\tau)$ given by (49) and thus the inequality $\Psi(v) \leq \tau$ holds at the present iterate, then we update μ to $(1-\theta)\mu$ for some $\theta \in (0,1)$. By making use of the relation (46) in Proposition 3.5, one can show that after the update of μ, the proximity is still bounded above by the number $\psi_0(\theta,\tau,2N)$. Here $\psi_0(\theta,\tau,2N)$ denotes the expression at the right-hand side of (46) where $\psi(v)$ and ϑ are replaced by τ and $\frac{1}{\sqrt{1-\theta}}$ respectively, i.e.,

$$\psi_0(\theta,\tau,2N) := \frac{\nu_2\tau}{\nu_1(1-\theta)^{\frac{p+1}{2}}} + 2\nu_2\Upsilon_{p,q}'\left((1-\theta)^{-\frac{1}{2}}\right)\sqrt{\frac{N\tau}{\nu_1(1-\theta)}} + 2N\nu_2\Upsilon_{p,q}\left((1-\theta)^{-\frac{1}{2}}\right).$$

The following lemma is an immediate consequence of Lemma 3.15 of [19].

Lemma 4.6 *Let $\Psi(x,s,\mu) \leq \tau$ and $\tau \geq \nu_1^{-1}$. Then after an update of the barrier parameter no more than*

$$\left\lceil \frac{8q\nu_1^{-\frac{q-1}{2q}}}{\nu_5\,(q+1)} \left(\psi_0(\theta,\tau,2N)\right)^{\frac{q+1}{2q}} \right\rceil$$

iterations are needed to recenter. In the special case where $\psi(t) = \Upsilon_{p,q}(t)$ is given by (16) with $\nu_1 = \nu_2 = 1$, at most

$$\left\lceil \frac{8q\max\left(3p+2,6q+4\right)}{q+1} \left(\psi_0(\theta,\tau,2N)\right)^{\frac{q+1}{2q}} \right\rceil$$

inner iterations are needed to recenter.

Thus the total complexity of the algorithm can be estimated as follows.

Theorem 4.7 *If $\tau \geq \nu_1^{-1}$, the total number of iterations required by the primal-dual Newton algorithm is not more than*

$$\left\lceil \frac{8q\nu_1^{-\frac{q-1}{2q}}}{\nu_5\,(q+1)} \left(\psi_0(\theta,\tau,2N)\right)^{\frac{1+q}{2q}} \right\rceil \left\lceil \frac{1}{\theta}\log\frac{N}{\varepsilon} \right\rceil.$$

In the special case where $\psi(t) = \Upsilon_{p,q}(t)$ is given by (16) with $\nu_1 = \nu_2 = 1$, then the total number of iterations required by the primal-dual Newton algorithm is less than or equal to

$$\left\lceil \frac{8q\max\left(3p+2,6q+4\right)}{q+1} \left(\psi_0(\theta,\tau,2N)\right)^{\frac{q+1}{2q}} \right\rceil \left\lceil \frac{1}{\theta}\log\frac{N}{\varepsilon} \right\rceil.$$

Neglecting the influence of the constants in the expression in Theorem 4.7, one can safely conclude that for any fixed $\theta \in (0,1)$ with constants $p, q \geq 1$, the algorithm with large-update for SOCO in the present section has an $\mathcal{O}\left(N^{\frac{q+1}{2q}}\log\frac{N}{\varepsilon}\right)$ iterations bound, while the algorithm with small-update ($\theta = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$) still stays with the complexity of $\mathcal{O}(\sqrt{N}\log\frac{N}{\varepsilon})$ iterations bound. Furthermore, from Theorem 4.7 one can readily see that if p is a constant and $q = \log N$, then the new large-update algorithm has an $\mathcal{O}\left(\sqrt{N}\log N\log\frac{N}{\varepsilon}\right)$ iterations bound.

In closing this paper we would like to point out that some constants (note that we replace n by $2N$ in the estimation of the total inner iterations) in the expression for the iteration bound of the algorithm for SOCO are slightly larger than presented in the case for LO and SDO. One might wonder how this comes. By ways of illustration, let us consider simply a SOCO problem with only one 2-dimensional second-order cone constraint $K = \{(x_1,x_2)^T : x_1 \geq |x_2|\}$. Note that this constraint can be rewritten as linear inequalities by adding artificial variables as follows.

$$x_1 - x_2 - x_3 = 0, x_1 + x_2 - x_4 = 0, x_1, x_3, x_4 \geq 0.$$

Thus we get an LO problem with four variables and three of them are nonnegative while the original SOCO problem has only one conic constraint. In other words, the size of the reformulated LO problem is four times as large as that of the original SOCO problem. Taking into account this point, the larger constants in the estimation of the complexity of the algorithm for SOCO seem to be reasonable. On the other hand, as we mentioned in the introduction of this paper, a SOCO problem can also be solved by casting it as a SDO problem. In such a situation, the algorithm posed in [19] for solving the reformulated SDO problem has an iteration complexity bound as $\mathcal{O}\left(n^{\frac{q+1}{2q}}\log\frac{n}{\varepsilon}\right)$. If $K^j \in \Re^2$ for any $j = 1,\cdots,N$, then one can see that the

iteration complexity of the algorithm for SOCO is precisely the same as that of its counterpart for SDO. However, the cost per iteration is significantly lower in the SOCO case than that in the SDO reformulation. Further, it is straightforward to see that when $2N < n$, the algorithm working directly on the original SOCO problem has a better iteration bound. The improvement is significant if $N \ll n$. Finally we remark that although theoretically our new large-update IPMs have better iteration bound than classical large-update IPMs, it still needs much work to test the practical efficiency of the new approach. With regarding to this point, the considerable flexibility in choosing the parameters p and q might help us to find new IPMs which are efficient both in theory and practice.

References

[1] I. Adler and F. Alizadeh. Primal-dual interior point algorithms for convex quadratically constrained and semidefinite optimization problems. Technical Report RRR-111-95, Rutcor, Rutgers Center for Operations Research, P.O.Box 5062, New Brunswick, New Jersey, 1995.

[2] F. Alizadeh, J.P. Haeberly and M.L. Overton. A new primal-dual interior-point method for semidefinite programming. In J.G. Lewis, Editor, *Proceedings of the fifth SIAM Conference on Applied Linear Algebra,* 113-117, SIAM, 1994.

[3] F. Alizadeh and S. Schmieta. Symmetric cones, potential reduction methods and word-by-word extensions. In H. Wolkowicz, R. Saigal and L. Vandenberghe Eds., *Handbook of Semidefinite Programming (Theory, Algorithms and Applications)*, pp. 195-234, Kluwer Academic Publishers, Boston, 2000.

[4] E.D. Andersen, C. Roos and T. Terlaky. On implementing a primal-dual interior-point method for conic quadratic optimization. Technical Report, Faculty of Information Technology and System, Delft University of Technology, The Netherlands, May, 2000.

[5] A. Ben-Tal and A. Nemirovskii. Convex Optimization in Engineering: Modeling, Analysis, Algorithms. Working note, Faculty of Industrial Engineering and Management, Technion-City, Haifa 32000, Israel. To be pubished by SIAM.

[6] J. Faraut and A. Korányi. *Analysis on Symmetric Cones.* Oxford University Press, NY, USA, 1994.

[7] L. Faybusovich. Euclidean Jordan algebras and interior-point algorithms. *Positivity*, 1(4), 331-357, 1997.

[8] L. Faybusovich. A Jordan-algebraic approach to potential-reduction algorithms. Technical Report, Department of Mathematics, University of Notre Dame, Notre Dame, IN 46555-5683, 1998.

[9] M. Fukushima, Z.Q. Luo and P. Tseng. Smoothing functions for second-order-cone complementarity problems. Preprint, Department of Mathematics, University of Washington, Seattle, WA 98195, USA, November, 2000.

[10] N.K. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

[11] E. de Klerk. *Interior Point Methods for Semidefinite Programming.* Ph.D. Thesis, Faculty of ITS/TWI, Delft University of Technology, The Netherlands, 1997.

[12] M.S. Lobo, L. Vandenberghe, S. Boyd and H. Lebret. Applications of second-order cone programming. *Linear Algebra and its Applications*, Vol(284),193-228, 1998.

[13] R.D.C. Monteiro and T. Tsuchiya. Polynomial convergence of primal-dual algorithms for the second-order cone program based on the MZ-family of directions. *Mathematical Programming*, Vol(88), 61-83, 2000.

[14] A.S. Nemirovskii and K. Scheinberg. Extension of Karmarkar's algorithm onto convex quadratically constrained quadratic programming. *Mathematical Programming*, Vol(72), 273-289, 1996.

[15] Y.E. Nesterov and A.S. Nemirovskii. Interior point polynomial algorithms in convex programming. *SIAM Studies in Applied Mathematics*, Vol(13), SIAM, Philadelphia, USA, 1994.

[16] 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.

[17] Y.E. Nesterov and M.J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM J. Optimization*, Vol(8), 324–362, 1998.

[18] J. Peng, C. Roos and T. Terlaky. A new class of polynomial primal-dual methods for linear and semidefinite optimization. Technical Report, Faculty of Technical Mathematics and Informatics, Delft University of Technology, The Netherlands, 1999.

[19] J. Peng, C. Roos and T. Terlaky. Self-regular proximities and new search directions for linear and semidefinite optimization. Technical Report, Faculty of Technical Mathematics and Informatics, Delft University of Technology, The Netherlands, 2000.

[20] W. Rudin, *Principles of Mathematical Analysis*. Mac-Graw Hill Company, New York, 1978.

[21] J. Sturm. *Theory and Algorithms of Semidefinite Programming.* In H. Frenk, C. Roos, T. Terlaky and S. Zhang Editors, *High Performance Optimization*, pp. 3-194, Kluwer Academic Publishers, Dordrecht, 1999.

[22] T. Tsuchiya. A polynomial primal-dual path-following algorithm for second-order cone programming. Technical Report 649, The Institute of Statistical Mathematics, Tokyo, Japan, December, 1997.

[23] T. Tsuchiya. A convergent analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming. *Optimization Methods and Software*, Vol(11/12), 141-182, 1999.

[24] H. Wolkowicz, R. Saigal and L. Vandenberghe. *Handbook of Semidefinite Programming (Theory, Algorithms and Applications).* Kluwer Academic Publishers, Boston, 2000.