November 1, 2000

Smoothing Functions for Second-Order-Cone Complementarity Problems*

Masao Fukushima[1], Zhi-Quan Luo[2] and Paul Tseng[3]

ABSTRACT

Smoothing functions have been much studied in the solution of optimization and complementarity problems with nonnegativity constraints. In this paper, we extend smoothing functions to problems where the nonnegative orthant is replaced by the direct product of second-order cones. These smoothing functions include the Chen-Mangasarian class and the smoothed Fischer-Burmeister function. We study the Lipschitzian and differential properties of these functions and, in particular, we derive computable formulas for these functions and their Jacobians. These properties and formulas can then be used to develop and analyze non-interior continuation methods for solving the corresponding optimization and complementarity problems. In particular, we establish existence and uniqueness of the Newton direction when the underlying mapping is monotone.

---

* This research is supported by a Grant-in-Aid for Scientific Research (B) from the Ministry of Education, Science, Sports and Culture of Japan. The third author is also supported by the National Science Foundation Grant CCR-9731273.

[1] Department of Applied Mathematics and Physics, School of Informatics, Kyoto University, Kyoto 606-8501, Japan.

[2] Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, L8S 4L7, Canada.

[3] Department of Mathematics, University of Washington, Seattle, WA 98195, USA.

1 Introduction

The *second-order cone* (SOC) in $\Re^n$ $(n \geq 1)$, also called Lorentz cone, is defined to be

$$\mathcal{K}^n = \{(x_1, x_2) \in \Re \times \Re^{n-1} \mid \|x_2\| \leq x_1\},$$

where $\|\cdot\|$ denotes the Euclidean norm. If $n = 1$, $\mathcal{K}^n$ is the set of nonnegative reals $\Re_+$. We are interested in optimization and complementarity problems whose constraints involve the direct product of second-order cones. In particular, we wish to find vectors $x, y \in \Re^n$ and $\zeta \in \Re^\ell$ satisfying

$$\langle x, y\rangle = 0, \quad x \in \mathcal{K}, \quad y \in \mathcal{K}, \quad F(x, y, \zeta) = 0, \tag{1.1}$$

where $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product, $F : \Re^n \times \Re^n \times \Re^\ell \to \Re^n \times \Re^\ell$ is a continuously differentiable mapping, and

$$\mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_m}, \tag{1.2}$$

with $\ell \geq 0$, $m, n_1, \ldots, n_m \geq 1$ and $n_1 + \cdots + n_m = n$. We will refer to (1.1), (1.2) as the *second-order-cone complementarity problem* (SOCCP). This problem has wide applications and, in particular, includes a large class of quadratically constrained problems as special cases [14]. It also includes as a special case the well-known nonlinear complementarity problem (NCP), corresponding to $n_i = 1$ for all i, i.e., $\mathcal{K}$ is the nonnegative orthant $\Re^n_+$. In particular, when $\ell = 0$ and the mapping F has the form

$$F(x, y, \zeta) = F_0(x) - y \tag{1.3}$$

for some $F_0 : \Re^n \to \Re^n$, the SOCCP (1.1) becomes

$$\langle x, F_0(x)\rangle = 0, \quad x \in \mathcal{K}, \quad F_0(x) \in \mathcal{K},$$

which is a natural generalization of the ordinary NCP corresponding to $\mathcal{K} = \Re^n_+$.

Optimization problems with SOC constraints have been the focus of several recent studies. It is known that $\mathcal{K}^n$, like $\Re^n_+$ and the cone $\mathcal{S}^n$ of $n \times n$ real symmetric positive semidefinite matrices, belongs to the class of symmetric cones, to which a Jordan algebra may be associated [9]. Using this connection, interior-point methods have been developed for solving linear programs with SOC constraints [14, 15, 22] and, more generally, linear programs with symmetric cone constraints [1, 17]. An alternative approach based on reformulating SOC constraints as smooth convex constraints was studied in [23]. It is also known that $\mathcal{K}^n$ may be viewed as a linear section of $\mathcal{S}^n$. In particular, it is easily verified that

$$(x_1, x_2) \in \mathcal{K}^n \quad \Longleftrightarrow \quad \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{bmatrix} \in \mathcal{S}^n, \tag{1.4}$$

where I denotes the identity matrix and superscript T denotes transpose. Thus, $\mathcal{K}^n = \mathcal{S}^n \cap \mathcal{A}$ for some subspace $\mathcal{A}$ of $\Re^{n\times n}$. Using this fact, we can reformulate any problem with SOC constraints as

a problem with semidefinite cone constraints plus an additional subspace constraint. However, this increases the problem dimension from n to $n(n+1)/2$ and it is not known whether this increase can be mitigated by, say, exploiting the structure of $\mathcal{A}$ when carrying out the linear algebra in computation.

In this paper, we consider a (non-interior) smoothing approach to solving problems with SOC constraints. In this approach, we choose a continuously differentiable function $\phi_\mu : \Re^n \times \Re^n \to \Re^n$, parameterized by $\mu > 0$, such that the pointwise limit $\phi_0(x,y) = \lim_{\mu \to 0^+} \phi_\mu(x,y)$ satisfies

$$\langle x, y \rangle = 0, \quad x \in \mathcal{K}, \quad y \in \mathcal{K} \quad \Longleftrightarrow \quad \phi_0(x,y) = 0. \tag{1.5}$$

We note that ϕ_0 is typically nonsmooth. The smoothing function ϕ_μ leads to the following non-interior continuation method for solving (1.1): Starting with an initial $\mu > 0$, we solve approximately the smooth equations

$$\phi_\mu(x,y) = 0, \qquad F(x,y,\zeta) = 0 \tag{1.6}$$

by, say, applying a few Newton steps; then we decrease μ and repeat the iteration. This non-interior approach offers an attractive alternative to the interior approach and it has been much studied in the case of NCP. In this case, one popular choice of ϕ_μ is a smooth approximation of the "min" function suggested by Chen and Harker [3], Kanzow [13], Smale [18], and further generalized by Chen and Mangasarian [4, 5] (also see [2, 7, 8, 10, 11, 21] and references therein). Another popular choice is a smooth approximation of the Fischer-Burmeister function suggested by Kanzow [13]. Recently, these smoothing functions and their nonsmooth analogs were extended to the setting of semidefinite complementarity problems (SDCP) [6, 19, 20]. Analogous to the semidefinite setting, our definition of smoothing functions is based on the spectral factorization of vectors in $\Re^n$, as specified by the Jordan algebra associated with SOC. This appears to be the most natural and convenient way to extend a function defined for NCP to one defined for SOCCP, while preserving essential differential and Lipschitzian properties. The latter properties are needed in analyzing the convergence properties of the corresponding continuation method; see, e.g., [2, 7, 8, 21] for such analysis in the case of NCP and [6] for an analysis in the setting of SDCP. Unlike the semidefinite setting, the simpler structure of SOC allows for a more direct analysis of its differential properties. In particular, we do not need to invoke a local upper Lipschitzian property of the spectral factorization, as is done in [6].

In what follows, $\Re^n$ ($n \geq 1$) denotes the space of n-dimensional real column vectors, $\Re^{n_1} \times \cdots \times \Re^{n_m}$ is identified with $\Re^{n_1+\cdots+n_m}$. Thus, $(x_1, \ldots, x_m) \in \Re^{n_1} \times \cdots \times \Re^{n_m}$ is viewed as a column vector in $\Re^{n_1+\cdots+n_m}$. For any $x, y \in \Re^n$ we write $x \succeq_{\mathcal{K}} y$ or $y \preceq_{\mathcal{K}} x$ (respectively, $x \succ_{\mathcal{K}} y$ or $y \prec_{\mathcal{K}} x$) if $x - y$ is in $\mathcal{K}$ (respectively, the interior of $\mathcal{K}$, denoted by int $\mathcal{K}$). For any square matrices $A, B \in \Re^{n \times n}$, we write $A \succ B$ (respectively, $A \succeq B$) if the symmetric part of $A - B$, namely $(A - B + A^T - B^T)/2$, is positive definite (respectively, positive semidefinite). Also, $\Re_+$ and $\Re_{++}$ denote the nonnegative and

positive reals. For any $x, y \in \Re^n$, the Euclidean inner product and norm are denoted $\langle x, y\rangle = x^T y$ and $\|x\| = \sqrt{x^T x}$. For any Fréchet-differentiable mapping $G : \Re^n \to \Re^m$, we denote its (transposed) Jacobian at $x \in \Re^n$ by $\nabla G(x) \in \Re^{n\times m}$, i.e., $(G(x+u) - G(x) - \nabla G(x)^T u)/\|u\| \to 0$ as $u \to 0$.

2 Jordan Algebra Associated with SOC

For any $x = (x_1, x_2) \in \Re \times \Re^{n-1}$ and $y = (y_1, y_2) \in \Re \times \Re^{n-1}$, we define their *Jordan product* as

$$x \cdot y = \begin{pmatrix} x^T y, & y_1 x_2 + x_1 y_2 \end{pmatrix}. \tag{2.1}$$

We will write x^2 to mean $x \cdot x$ and write $x + y$ to mean the usual componentwise addition of vectors. Then, $\cdot$, $+$, together with

$$e = (1, 0, \ldots, 0) \in \Re^n,$$

have the following basic properties [9, Chap. II].

Property 2.1

1. $e \cdot x = x \quad \forall x \in \Re^n.$ (*Identity*)
2. $x \cdot y = y \cdot x \quad \forall x, y \in \Re^n.$ (*Commutativity 1*)
3. $x \cdot (x^2 \cdot y) = x^2 \cdot (x \cdot y) \quad \forall x, y \in \Re^n.$ (*Commutativity 2*)
4. $(x + y) \cdot z = x \cdot z + y \cdot z \quad \forall x, y, z \in \Re^n.$ (*Distributivity*)

Notice that the Jordan product is *not associative* in general. For example, for $n = 3$ and $x = (1, -1, 1)$, $y = z = (1, 0, 1)$, we have

$$(x \cdot y) \cdot z = (4, -1, 4) \quad \neq \quad x \cdot (y \cdot z) = (4, -2, 4).$$

In fact, this example shows that $(x \cdot y) \cdot y \neq x \cdot y^2$ in general. On the other hand, it can be verified that associativity holds under the inner product in the sense that

$$\langle x, y \cdot z\rangle = \langle y, z \cdot x\rangle = \langle z, x \cdot y\rangle \quad \forall x, y, z \in \Re^n.$$

The Jordan product (2.1) is associated with the SOC $\mathcal{K}^n$ via the following useful facts.

1. For each $x = (x_1, x_2) \in \Re \times \Re^{n-1}$, the *determinant* and the *trace* of x are defined by

$$\det(x) = x_1^2 - \|x_2\|^2 \quad \text{and} \quad \mathrm{tr}(x) = 2x_1,$$

respectively. Unlike matrices, we have in general $\det(x \cdot y) \neq \det(x)\det(y)$ unless $x_2 = y_2$.

2. A vector $x = (x_1, x_2) \in \Re \times \Re^{n-1}$ is said to be *invertible* if $\det(x) \neq 0$. If x is invertible, then there exists a unique $y = (y_1, y_2) \in \Re \times \Re^{n-1}$ satisfying $x \cdot y = e$. We shall call this y the *inverse* of x and denote it by x^{-1}. Direct calculation yields

$$x^{-1} = \frac{1}{x_1^2 - \|x_2\|^2}(x_1, -x_2) = \frac{\mathrm{tr}(x)\, e - x}{\det(x)}. \tag{2.2}$$

   From (2.2), it is clear that $x \in \mathrm{int}\, \mathcal{K}^n$ if and only if $x^{-1} \in \mathrm{int}\, \mathcal{K}^n$

3. If $x \in \mathcal{K}^n$, then there exists a unique vector in $\mathcal{K}^n$, which we denote by $x^{1/2}$, such that $(x^{1/2})^2 = x^{1/2} \cdot x^{1/2} = x$. Direct calculation yields

$$x^{1/2} = \left(s, \frac{x_2}{2s}\right), \quad \text{where} \quad s = \sqrt{\left(x_1 + \sqrt{x_1^2 - \|x_2\|^2}\right)/2}. \tag{2.3}$$

   In the above formula, the term x_2/s is defined to be the zero vector if $x_2 = 0$ and $s = 0$, i.e., $x = 0$.

4. For any $x \in \Re^n$, we have $x^2 \in \mathcal{K}^n$. Hence there exists a unique vector $(x^2)^{1/2} \in \mathcal{K}^n$, which we denote by $|x|$. Clearly we have $x^2 = |x|^2$.

Notice that the second-order cone $\mathcal{K}^n$ is *not closed* under Jordan product. For example,

$$x = \left(\sqrt{2}, 1, 1\right) \in \mathcal{K}^3, \quad y = \left(\sqrt{2}, 1, -1\right) \in \mathcal{K}^3, \qquad \text{but} \qquad x \cdot y = \left(2, 2\sqrt{2}, 0\right) \notin \mathcal{K}^3.$$

For any $x = (x_1, x_2) \in \Re \times \Re^{n-1}$, we define the symmetric matrix

$$L_x = \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{bmatrix}, \tag{2.4}$$

viewed as a linear mapping from $\Re^n$ to $\Re^n$. It is easily verified that

$$L_x y = x \cdot y \quad \forall y \in \Re^n.$$

Moreover, L_x is positive definite (and hence invertible) if and only if $x \in \mathrm{int}\, \mathcal{K}^n$ (see (1.4)). Notice, however, that in general $L_x^{-1} y \neq x^{-1} \cdot y$ for $x \in \mathrm{int}\, \mathcal{K}^n$ and $y \in \Re^n$.

We next introduce the spectral factorization of vectors in $\Re^n$ associated with $\mathcal{K}^n$. Let $x = (x_1, x_2) \in \Re \times \Re^{n-1}$. Then x can be decomposed as

$$x = \lambda_1 u^{(1)} + \lambda_2 u^{(2)}, \tag{2.5}$$

where λ_1, λ_2 and $u^{(1)}$, $u^{(2)}$ are the spectral values and the associated spectral vectors of x given by

$$\lambda_i = x_1 + (-1)^i \|x_2\|, \tag{2.6}$$

$$u^{(i)} = \begin{cases} \frac{1}{2}\left(1, (-1)^i \dfrac{x_2}{\|x_2\|}\right), & \text{if} \quad x_2 \neq 0, \\ \frac{1}{2}\left(1, (-1)^i w\right), & \text{if} \quad x_2 = 0, \end{cases} \tag{2.7}$$

for $i = 1, 2$, with w being any vector in $\Re^{n-1}$ satisfying $\|w\| = 1$.If $x_2 \neq 0$, the decomposition (2.5) is unique.

Some interesting properties of λ_1, λ_2 and $u^{(1)}$, $u^{(2)}$ are summarized below. Notice that the identity element e is uniquely identified by its two spectral values which are exactly equal to 1.

Property 2.2 *For any $x = (x_1, x_2) \in \Re \times \Re^{n-1}$, the spectral values λ_1, λ_2 and spectral vectors $u^{(1)}$, $u^{(2)}$, as given by (2.6) and (2.7), have the following properties:*

1. *$u^{(1)}$ and $u^{(2)}$ are orthogonal under Jordan product and have length $1/\sqrt{2}$, i.e.,*

$$u^{(1)} \cdot u^{(2)} = 0, \qquad \|u^{(1)}\| = \|u^{(2)}\| = 1/\sqrt{2}.$$

2. *$u^{(1)}$ and $u^{(2)}$ are idempotent under Jordan product, i.e.,*

$$u^{(i)} \cdot u^{(i)} = u^{(i)}, \quad i = 1, 2.$$

3. *λ_1 and λ_2 are nonnegative (respectively, positive) if and only if $x \in \mathcal{K}^n$ (respectively, $x \in$ int $\mathcal{K}^n$).*

4. *The determinant, the trace and the Euclidean norm of x can all be represented in terms of λ_1 and λ_2:*

$$\det(x) = \lambda_1 \lambda_2, \quad \mathrm{tr}(x) = \lambda_1 + \lambda_2, \quad 2\|x\|^2 = \lambda_1^2 + \lambda_2^2.$$

It can be verified that λ_1, λ_2 are in fact the eigenvalues of the $n \times n$ matrix L_x (see (2.4)) with $u^{(1)}$, $u^{(2)}$ being the corresponding eigenvectors. The remaining $n-2$ eigenvalues of this matrix are identically x_1, with corresponding eigenvectors of the form $(0, v)$, where v lies the subspace of $\Re^{n-1}$ orthogonal to x_2. A corollary of this is the equivalence (1.4). It can also be verified that λ_1, λ_2 are Lipschitz continuous in x, with Lipschitz constant $\sqrt{2}$, and $u^{(1)}, u^{(2)}$ have a local upper Lipschitzian property. However, we will not need these properties in the subsequent analysis.

The next proposition shows that, in the case of $\mathcal{K} = \mathcal{K}^n$, the SOCCP (1.1) can be equivalently stated as

$$x \cdot y = 0, \quad x \in \mathcal{K}, \quad y \in \mathcal{K}, \quad F(x, y, \zeta) = 0.$$

This equivalence also extends to the general case where $\mathcal{K}$ has the direct product structure (1.2), provided that the Jordan product is defined according to this structure (see Section 6).

Proposition 2.1 *For any x and y in $\Re^n$, we have*

$$\langle x, y \rangle = 0, \ x \in \mathcal{K}^n, \ y \in \mathcal{K}^n \quad \Longleftrightarrow \quad x \cdot y = 0, \ x \in \mathcal{K}^n, \ y \in \mathcal{K}^n.$$

Proof. The '$\Leftarrow$' direction is obvious from the definition of Jordan product. To prove the '$\Rightarrow$' direction, suppose $\langle x, y\rangle = 0,\ x \in \mathcal{K}^n,\ y \in \mathcal{K}^n$. Then,

$$\langle x, y\rangle = x_1 y_1 + x_2^T y_2 = 0, \quad x_1 \geq \|x_2\|, \quad y_1 \geq \|y_2\|,$$

implying $-x_2^T y_2 = x_1 y_1 \geq \|x_2\|\|y_2\|$. Since $-x_2^T y_2 \leq \|x_2\|\|y_2\|$, this shows that x_2, y_2 make an angle of 180° and, moreover, $x_1 = \|x_2\|$, $y_1 = \|y_2\|$. Thus,

$$x \cdot y = (\langle x, y\rangle, x_1 y_2 + y_1 x_2) = (0, \|x_2\| y_2 + \|y_2\| x_2) = 0.$$

This completes the proof. □

3 Functions Associated with SOC

For any function $\hat{g} : \Re \to \Re$, we define a function on $\Re^n$ associated with $\mathcal{K}^n$ $(n \geq 1)$ by

$$g(x) = \hat{g}(\lambda_1) u^{(1)} + \hat{g}(\lambda_2) u^{(2)} \quad \forall x = (x_1, x_2) \in \Re \times \Re^{n-1}, \tag{3.1}$$

where $\lambda_1, \lambda_2, u^{(1)}, u^{(2)}$ are the spectral values and vectors of x (see (2.5), (2.6), (2.7)). If $\hat{g}$ is defined only on a subset of $\Re$, then g is defined on the corresponding subset of $\Re^n$. The definition (3.1) is clearly unambiguous when $x_2 \neq 0$ since $u^{(1)}, u^{(2)}$ are unique. When $x_2 = 0$, we see from (2.6), (2.7) that $g(x) = \hat{g}(x_1)e$ so the definition is again unambiguous. The cases of $g(x) = x^{1/2},\ x^2,\ e^x$ are discussed in the book of Faraut and Korányi [9]. The above definition (3.1) is analogous to one associated with the semidefinite cone $\mathcal{S}^n$, see [6, 19]. In fact, if a function $\tilde{g}$ associated with $\mathcal{S}^n$ has the property that it maps the subspace $\mathcal{A}$ of matrices of the form (2.4) into $\mathcal{A}$, then $g = L^{-1} \circ \tilde{g} \circ L$ is a function associated with $\mathcal{K}^n$, where $L : \Re^n \to \mathcal{A}$ is the linear mapping such that $L(x)$ is the matrix L_x given by (2.4) for all $(x_1, x_2) \in \Re \times \Re^{n-1}$. However, the functions considered in [6, 19] do not have this property.

For any $x \in \Re^n$ and integer $k \geq 1$, we can take the Jordan product of x with itself k times to obtain the kth power of x, denoted by x^k. We define $x^0 = e$. If $x \in \text{int}\, \mathcal{K}^n$, then $x^{-k} = (x^k)^{-1}$ is also well defined. The spectral factorization (2.5)–(2.7) and its properties (see Property 2.2) provide a very useful tool for evaluating functions defined using powers. For example, for any $x \in \Re^n$, the spectral factorization of x yields

$$\begin{aligned} x^2 &= (\lambda_1 u^{(1)} + \lambda_2 u^{(2)}) \cdot (\lambda_1 u^{(1)} + \lambda_2 u^{(2)}) \\ &= \lambda_1^2 (u^{(1)})^2 + \lambda_2^2 (u^{(2)})^2 + 2\lambda_1 \lambda_2 u^{(1)} \cdot u^{(2)} \\ &= \lambda_1^2 u^{(1)} + \lambda_2^2 u^{(2)}, \end{aligned} \tag{3.2}$$

where we used the orthogonal and idempotent properties of $u^{(1)}$ and $u^{(2)}$ (see Property 2.2). Notice that this is consistent with (3.1). The above formula shows that squaring a vector is the same as

squaring the spectral values in its spectral factorization. A corollary of this formula is that, for any $x \in \Re^n$, the spectral values of x^2 are nonnegative and hence $x^2 \in \mathcal{K}^n$. Conversely, when $x \in \mathcal{K}^n$, we have from Property 2.2 that $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ so that $w = \sqrt{\lambda_1}u^{(1)} + \sqrt{\lambda_2}u^{(2)}$ is defined and in $\mathcal{K}^n$. Moreover, analogous to (3.2), $w^2 = x$. Thus, $x^{1/2} = w$, i.e.,

$$x^{1/2} = \sqrt{\lambda_1}u^{(1)} + \sqrt{\lambda_2}u^{(2)}. \tag{3.3}$$

The above derivation can be extended beyond powers to any function admitting a power series expansion. This is shown in the proposition below.

Proposition 3.1 *Suppose $\hat{g} : \Re \to \Re$ admits a power series expansion $\hat{g}(\alpha) = \sum_{k=0}^{\infty} a_k \alpha^k$ for some real coefficients $a_0, a_1, \ldots$. Then the function $g : \Re^n \to \Re^n$ given by* (3.1) *satisfies*

$$g(x) \;=\; \sum_{k=0}^{\infty} a_k(\lambda_1^k u^{(1)} + \lambda_2^k u^{(2)}) \;=\; \sum_{k=0}^{\infty} a_k x^k \qquad \forall\, x \in \Re^n,$$

where λ_1, λ_2, $u^{(1)}$, $u^{(2)}$ are the spectral values and vectors of x given by (2.6), (2.7).

Proof. Using Property 2.2, we have for all $k \geq 1$,

$$(u^{(i)})^k = u^{(i)}, \qquad u^{(i)} \cdot u^{(j)} = 0,$$

where $i, j = 1, 2$, and $i \neq j$. Using the binomial expansion, this yields

$$x^k = (\lambda_1 u^{(1)} + \lambda_2 u^{(2)})^k = \lambda_1^k u^{(1)} + \lambda_2^k u^{(2)}.$$

Substituting this into the definition of $g(x)$ yields

$$\begin{aligned} g(x) &= \hat{g}(\lambda_1)u^{(1)} + \hat{g}(\lambda_2)u^{(2)} \\ &= \left(\sum_{k=0}^{\infty} a_k \lambda_1^k\right) u^{(1)} + \left(\sum_{k=0}^{\infty} a_k \lambda_2^k\right) u^{(2)} \\ &= \sum_{k=0}^{\infty} a_k(\lambda_1^k u^{(1)} + \lambda_2^k u^{(2)}) \\ &= \sum_{k=0}^{\infty} a_k x^k. \end{aligned}$$

□

Define

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \qquad \forall x \in \Re^n.$$

Also, define $\ln(x)$ to be the unique $w \in \Re^n$ satisfying $e^w = x$ for each $x \in \text{int } \mathcal{K}^n$. In the proposition below, we use Proposition 3.1 to derive a simple formula for e^x. This in turn is used to show that $\ln(x)$ is well defined and has a simple formula.

Proposition 3.2

(a) *For any* $x = (x_1, x_2) \in \Re \times \Re^{n-1}$,

$$e^x = \begin{cases} \frac{1}{2} e^{x_1} \left(\cosh(\|x_2\|), \sinh(\|x_2\|) \frac{x_2}{\|x_2\|} \right), & \text{if } x_2 \neq 0, \\ e^{x_1} (1, 0), & \text{if } x_2 = 0, \end{cases}$$

where $\cosh(\alpha) = (e^\alpha + e^{-\alpha})/2$ *and* $\sinh(\alpha) = (e^\alpha - e^{-\alpha})/2$ *for* $\alpha \in \Re$.

(b) *For any* $x = (x_1, x_2) \in \operatorname{int} \mathcal{K}^n$, $\ln(x)$ *is well defined and*

$$\ln(x) = \begin{cases} \frac{1}{2} \left(\ln \left(x_1^2 - \|x_2\|^2 \right), \ln \left(\frac{x_1 + \|x_2\|}{x_1 - \|x_2\|} \right) \frac{x_2}{\|x_2\|} \right), & \text{if } x_2 \neq 0, \\ \ln(x_1) (1, 0), & \text{if } x_2 = 0. \end{cases}$$

Proof. (a) By Proposition 3.1, $e^x = g(x)$ with $\hat{g}(\alpha) = e^\alpha$. Using (3.1) with λ_1, λ_2 and $u^{(1)}, u^{(2)}$ given by (2.6) and (2.7), we have

$$\begin{aligned} e^x &= e^{\lambda_1} u^{(1)} + e^{\lambda_2} u^{(2)} \\ &= e^{x_1 - \|x_2\|} u^{(1)} + e^{x_1 + \|x_2\|} u^{(2)} \\ &= e^{x_1} \left(e^{-\|x_2\|} u^{(1)} + e^{\|x_2\|} u^{(2)} \right). \end{aligned}$$

We consider only the case of $x_2 \neq 0$. The case of $x_2 = 0$ can be argued analogously. Using (2.7), we have

$$u^{(i)} = \frac{1}{2} \left(1, \ \frac{(-1)^i x_2}{\|x_2\|} \right), \qquad i = 1, 2,$$

the previous expression simplifies to

$$e^x = \frac{1}{2} e^{x_1} \left(e^{\|x_2\|} + e^{-\|x_2\|}, \ (e^{\|x_2\|} - e^{-\|x_2\|}) \cdot \frac{x_2}{\|x_2\|} \right).$$

(b) Fix any $x = (x_1, x_2) \in \operatorname{int} \mathcal{K}^n$, i.e., $\|x_2\| < x_1$. We will prove by construction that there is a unique $w \in \Re^n$ satisfying $e^w = x$, which will show that $\ln(x)$ is well defined. We have from part (a) that $w = (w_1, w_2) \in \Re \times \Re^{n-1}$ satisfies $e^w = x$ if and only if

$$x_1 = \frac{1}{2} e^{w_1} \left(e^{\|w_2\|} + e^{-\|w_2\|} \right), \qquad x_2 = \frac{1}{2} e^{w_1} \left(e^{\|w_2\|} - e^{-\|w_2\|} \right) \frac{w_2}{\|w_2\|}. \tag{3.4}$$

Letting $a = e^{w_1}$ and $b = e^{\|w_2\|}$, we obtain from (3.4) that

$$x_1 = \frac{a}{2}(b + b^{-1}), \qquad \|x_2\| = \frac{a}{2}(b - b^{-1}).$$

We can solve these equations uniquely for a and b to yield

$$b = \sqrt{\frac{x_1 + \|x_2\|}{x_1 - \|x_2\|}}, \qquad a = \frac{2x_1}{b + b^{-1}} = \sqrt{x_1^2 - \|x_2\|^2}.$$

Suppose $x_2 \neq 0$, so that $b \neq 1$. We obtain $w_1 = \ln(a)$, $\|w_2\| = \ln(b)$, and we have from (3.4) that $w_2 = 2a^{-1}\|w_2\|x_2/(b - b^{-1})$. Thus,

$$w = (w_1, w_2) = \left(\ln(a), 2a^{-1}\ln(b)x_2/(b - b^{-1})\right).$$

Using the above formulas for a and b and then simplifying expressions, we obtain the desired formula for w. Suppose instead $x_2 = 0$, so that $a = x_1$ and $b = 1$. We then obtain $w_1 = \ln(x_1)$, $\|w_2\| = \ln(1) = 0$ so that $w_2 = 0$. □

It can be verified that $\ln(x)$ is alternatively given by (3.1) with $\hat{g}(\alpha) = \ln(\alpha)$. For any $x \in \Re^n$, we define $[x]_+$ to be the nearest-point (in the Euclidean norm) projection of x onto $\mathcal{K}^n$. For $\alpha \in \Re$, let $[\alpha]_+ = \max\{0, \alpha\}$. The following proposition shows that $|x|$ and $[x]_+$ have the form (3.1) and are related to each other as in the cases of nonnegative orthant $\Re^n_+$ and positive semidefinite cone $\mathcal{S}^n$.

Proposition 3.3 *For any $x = (x_1, x_2) \in \Re \times \Re^{n-1}$, let λ_1, λ_2, $u^{(1)}, u^{(2)}$ denote the spectral values and vectors of x, i.e., (2.5)–(2.7). Then the following results hold.*

(a) $|x| = (x^2)^{1/2} = |\lambda_1|u^{(1)} + |\lambda_2|u^{(2)}$.

(b) *The projection onto the second-order cone $\mathcal{K}^n$ can be written as*

$$[x]_+ = [\lambda_1]_+ u^{(1)} + [\lambda_2]_+ u^{(2)} = (x + |x|)/2. \tag{3.5}$$

Proof. (a) By using (3.2) and (3.3), we have

$$\begin{aligned} (x^2)^{1/2} &= \left(\lambda_1^2 u^{(1)} + \lambda_2^2 u^{(2)}\right)^{1/2} \\ &= (\lambda_1^2)^{1/2} u^{(1)} + (\lambda_2^2)^{1/2} u^{(2)} \\ &= |\lambda_1| u^{(1)} + |\lambda_2| u^{(2)}. \end{aligned}$$

(b) Since $[\alpha]_+ = (\alpha + |\alpha|)/2$ for all $\alpha \in \Re$, the second equality in (3.5) follows from part (a) and (2.5). We now prove the first equality in (3.5). First, consider the case of $x \in \mathcal{K}^n$. Then $[x]_+ = x$. Also, $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$, so $[\lambda_1]_+ = \lambda_1$ and $[\lambda_2]_+ = \lambda_2$. Thus the first equality in (3.5) holds. Next, consider the case of $x \notin \mathcal{K}^n$. Then, $x_1 < \|x_2\|$ so that $[\lambda_1]_+ = [x_1 - \|x_2\|]_+ = 0$. Consider the projection problem:

$$\min \frac{1}{2}\|y - x\|^2 \quad \text{subject to} \quad \|y_2\| - y_1 \leq 0.$$

This problem is convex and its unique optimal solution $y = [x]_+$ satisfies the inequality constraint with equality. Hence this optimal solution satisfies the following necessary conditions for optimality:

$$y - x + \mu(-1, \nu) = 0, \quad \|y_2\| = y_1, \tag{3.6}$$

for some nonnegative $\mu \in \Re$ (μ is a Lagrange multiplier) and some subgradient $\nu \in \Re^{n-1}$ of $\|\cdot\|$ at y_2, i.e.,

$$\nu = y_2/\|y_2\| \quad \text{if } y_2 \neq 0, \qquad \|\nu\| \leq 1 \quad \text{if } y_2 = 0. \tag{3.7}$$

Suppose $y_2 \neq 0$. Then, solving (3.6) and (3.7) for y and μ and using $x_1 < \|x_2\|$ yield

$$y_1 = \frac{1}{2}[x_1 + \|x_2\|]_+ \qquad \text{and} \qquad y_2 = \begin{cases} \frac{1}{2}[x_1 + \|x_2\|]_+ x_2/\|x_2\|, & \text{if } x_2 \neq 0, \\ 0, & \text{if } x_2 = 0. \end{cases}$$

If $x_2 = 0$, then $x_1 < 0$ so $[x_1 + \|x_2\|]_+ = 0$ and hence $y_2 = 0$, a contradiction of $y_2 \neq 0$. Thus $x_2 \neq 0$, and we obtain

$$(y_1, y_2) \;=\; \frac{1}{2}[x_1 + \|x_2\|]_+ \left(1, \frac{x_2}{\|x_2\|}\right) \;=\; [\lambda_2]_+ u^{(2)}.$$

Since $[\lambda_1]_+ = 0$, this proves the first equality in (3.5). Suppose instead $y_2 = 0$. Then, (3.6) implies $y_1 = 0$ and $x_1 = -\mu$, $x_2 = \mu\nu$. Since $\mu \geq 0$ and $\|\nu\| \leq 1$, then $\|x_2\| = \mu\|\nu\| \leq \mu = -x_1$. Thus $x_1 + \|x_2\| \leq 0$, implying $[\lambda_2]_+ = 0$. Since $[\lambda_1]_+ = 0$, this proves the first equality in (3.5). □

A corollary of Proposition 3.3 is the third part of Property 2.2. Another corollary is

$$|x| \succeq_{\mathcal{K}^n} x \qquad \forall x \in \Re^n.$$

Using this corollary, we now prove the following proposition which establishes some order preserving properties of x^2 and L_x^2 relative to $\mathcal{K}^n$ and $\mathcal{S}^n$, respectively. These properties, though difficult to prove, are crucial to some of our subsequent analyses. We speculate that these properties may be useful in other contexts.

Proposition 3.4 *For any x, y in $\Re^n$ and any $w \succ_{\mathcal{K}^n} 0$, we have*

$$w^2 \succ_{\mathcal{K}^n} x^2 + y^2 \qquad \Longrightarrow \qquad L_w^2 \succ L_x^2 + L_y^2, \tag{3.8}$$

$$w^2 \succ_{\mathcal{K}^n} x^2 \qquad \Longrightarrow \qquad w \succ_{\mathcal{K}^n} x. \tag{3.9}$$

Moreover, the implications (3.8) *and* (3.9) *remain true when "$\succ$" is replaced by "$\succeq$" everywhere.*

Proof. First, we prove (3.8) for the case where w has the form

$$w = (x^2 + y^2 + \delta e)^{1/2} \tag{3.10}$$

for some $\delta > 0$. Fix any $x = (x_1, x_2)$, $y = (y_1, y_2)$ in $\Re \times \Re^{n-1}$ and any $\delta > 0$. Let $w = (x^2 + y^2 + \delta e)^{1/2}$. Since $w^2 = x^2 + y^2 + \delta e \succ_{\mathcal{K}^n} 0$, then $w \succ_{\mathcal{K}^n} 0$. Moreover, (2.3) yields

$$w = (s, w_2) \quad \text{with} \quad w_2 = (x_1 x_2 + y_1 y_2)/s, \tag{3.11}$$

where $s = \sqrt{\|x\|^2 + \|y\|^2 + \delta + \sqrt{(\|x\|^2 + \|y\|^2 + \delta)^2 - 4\|x_1 x_2 + y_1 y_2\|^2}}/\sqrt{2}$. Since $\delta > 0$, we have $s > 0$. Since $w \succ_{\mathcal{K}^n} 0$, we have from (3.11) that $\|w_2\| < s$. Since $w^2 - x^2 - y^2 = \delta e \succ_{\mathcal{K}^n} 0$, we have

$$L_{w^2} - L_{x^2} - L_{y^2} \succ 0.$$

Direct calculation using (3.11) shows that $L_{w^2} - L_{x^2} - L_{y^2} = (\|w\|^2 - \|x\|^2 - \|y\|^2)I$, so it must be that

$$\|x\|^2 + \|y\|^2 < \|w\|^2. \tag{3.12}$$

Using (3.11), the inequality (3.12) expands out to

$$x_1^2 + \|x_2\|^2 + y_1^2 + \|y_2\|^2 < s^2 + x_1^2\|x_2\|^2/s^2 + y_1^2\|y_2\|^2/s^2 + 2x_1 y_1 x_2^T y_2/s^2,$$

which, upon multiplying both sides by s^2 and rearranging terms, can be rewritten equivalently as

$$\|y_1 x_2 - x_1 y_2\|^2 < (s^2 - x_1^2 - y_1^2)(s^2 - \|x_2\|^2 - \|y_2\|^2).$$

Thus, either both $s^2 - x_1^2 - y_1^2$ and $s^2 - \|x_2\|^2 - \|y_2\|^2$ are positive or both are negative. If both are negative, then we would have $x_1^2 + y_1^2 > s^2$ and $\|x_2\|^2 + \|y_2\|^2 > s^2$, contradicting the fact that $\|x\|^2 + \|y\|^2 < \|w\|^2 = s^2 + \|w_2\|^2 < 2s^2$ (recall (3.12) and $\|w_2\| < s$). Thus, we must instead have

$$x_1^2 + y_1^2 < s^2 \quad \text{and} \quad \|x_2\|^2 + \|y_2\|^2 < s^2. \tag{3.13}$$

By direct calculation using (3.11), we have

$$L_w^2 - L_x^2 - L_y^2 = \text{diag}\left\{\|w\|^2 - \|x\|^2 - \|y\|^2, (s^2 - x_1^2 - y_1^2)I + w_2 w_2^T - x_2 x_2^T - y_2 y_2^T\right\}.$$

Our goal is to show that this matrix is positive definite. In view of (3.12), it suffices to prove

$$(s^2 - x_1^2 - y_1^2)I + w_2 w_2^T - x_2 x_2^T - y_2 y_2^T \succ 0. \tag{3.14}$$

For any $d \in \Re^{n-1}$, we have from (3.11) that

$$\begin{aligned}
& d^T(w_2 w_2^T - x_2 x_2^T - y_2 y_2^T)d \\
= \;& (w_2^T d)^2 - (x_2^T d)^2 - (y_2^T d)^2 \\
= \;& (x_1 x_2^T d + y_1 y_2^T d)^2/s^2 - (x_2^T d)^2 - (y_2^T d)^2 \\
= \;& \left(x_1^2 (x_2^T d)^2 + y_1^2 (y_2^T d)^2 + 2(y_1 x_2^T d)(x_1 y_2^T d)\right)/s^2 - (x_2^T d)^2 - (y_2^T d)^2 \\
\leq \;& \left(x_1^2 (x_2^T d)^2 + y_1^2 (y_2^T d)^2 + (y_1 x_2^T d)^2 + (x_1 y_2^T d)^2\right)/s^2 - (x_2^T d)^2 - (y_2^T d)^2 \\
= \;& \left(x_1^2 + y_1^2 - s^2\right)\left((x_2^T d)^2 + (y_2^T d)^2\right)/s^2,
\end{aligned}$$

where the inequality uses the fact $2\alpha\beta \leq \alpha^2 + \beta^2$ for $\alpha, \beta \in \Re$. By (3.13), the right-hand side is always non-positive, which implies that $0 \succeq w_2 w_2^T - x_2 x_2^T - y_2 y_2^T$ and hence

$$w_2 w_2^T - x_2 x_2^T - y_2 y_2^T \succeq \mathrm{tr}(w_2 w_2^T - x_2 x_2^T - y_2 y_2^T) I = (\|w_2\|^2 - \|x_2\|^2 - \|y_2\|^2) I.$$

This in turn implies

$$\begin{aligned} (s^2 - x_1^2 - y_1^2) I + w_2 w_2^T - x_2 x_2^T - y_2 y_2^T &\succeq (s^2 - x_1^2 - y_1^2 + \|w_2\|^2 - \|x_2\|^2 - \|y_2\|^2) I \\ &= (\|w\|^2 - \|x\|^2 - \|y\|^2) I. \end{aligned}$$

Using (3.12), this proves (3.14), as desired.

We have proven that the implication (3.8) holds true for any x, y in $\Re^n$ and any w of the form (3.10) for some $\delta > 0$. We will now use this result to prove (3.8) for any x, y in $\Re^n$ and any $w \succ_{\mathcal{K}^n} 0$. Suppose $w^2 \succ_{\mathcal{K}^n} x^2 + y^2$. Then, there exists some $\delta > 0$ such that

$$w^2 \succ_{\mathcal{K}^n} x^2 + y^2 + 2\delta e.$$

Let $\bar{w} = (x^2 + y^2 + \delta e)^{1/2}$. Then $w^2 \succ_{\mathcal{K}^n} \bar{w}^2 + \delta e$, so $z = (w^2 - \bar{w}^2 - \delta e)^{1/2}$ is defined and $w = (\bar{w}^2 + z^2 + \delta e)^{1/2}$. Thus,

$$\bar{w}^2 \succ_{\mathcal{K}^n} x^2 + y^2, \qquad w^2 \succ_{\mathcal{K}^n} \bar{w}^2 + z^2.$$

Since $\bar{w}$ has the form (3.10) with "w" replaced by "$\bar{w}$", then (3.8) holds true with this replacement. Similarly, since w has the form (3.10) with "x", "y" replaced by "$\bar{w}$", "z", respectively, then (3.8) holds true with this replacement. Thus, the above relations imply

$$L_{\bar{w}}^2 \succ L_x^2 + L_y^2, \qquad L_w^2 \succ L_{\bar{w}}^2 + L_z^2.$$

These two relations combine to yield $L_w^2 \succ L_x^2 + L_y^2$. Thus (3.8) holds true.

We now use (3.8) to prove (3.9) for any x, y in $\Re^n$ and any $w \succ_{\mathcal{K}^n} 0$. Suppose $w^2 \succ_{\mathcal{K}^n} x^2$. Let $y = 0$. Then $w^2 \succ_{\mathcal{K}^n} x^2 + y^2 = |x|^2 + y^2$ and, by (3.8), we have $L_w^2 \succ L_{|x|}^2 + L_y^2$ and hence $L_w^2 - L_{|x|}^2 \succ 0$. Since $w \succ_{\mathcal{K}^n} 0$ and $|x| \succeq_{\mathcal{K}^n} 0$ so that $L_w \succ 0$ and $L_{|x|} \succeq 0$, we also have $L_w + L_{|x|} \succ 0$ and, in particular, $(L_w + L_{|x|})^{1/2}$ is defined and invertible. Finally, we have

$$(L_w - L_{|x|})(L_w + L_{|x|}) = (L_w^2 - L_{|x|}^2) + (L_w L_{|x|} - L_{|x|} L_w) \succ 0, \tag{3.15}$$

where the last step follows from $L_w^2 - L_{|x|}^2 \succ 0$ and the fact that $L_w L_{|x|} - L_{|x|} L_w$ is anti-symmetric. Thus, the matrix $(L_w - L_{|x|})(L_w + L_{|x|})$ is positive definite, though not necessarily symmetric. Since $(L_w + L_{|x|})^{1/2}$ is invertible, the symmetric matrix $(L_w + L_{|x|})^{1/2}(L_w - L_{|x|})(L_w + L_{|x|})^{1/2}$ has the same eigenvalues as the matrix $(L_w - L_{|x|})(L_w + L_{|x|})$. It follows that the eigenvalues of the latter matrix must be real, and in view of (3.15), positive. As a result, we have $(L_w + L_{|x|})^{1/2}(L_w -$

$L_{|x|})(L_w + L_{|x|})^{1/2} \succ 0$ and hence $L_w - L_{|x|} \succ 0$ or, equivalently, $w \succ_{\mathcal{K}^n} |x|$. Since $|x| \succeq_{\mathcal{K}^n} x$, this yields $w \succ_{\mathcal{K}^n} x$. Thus (3.9) holds true.

Finally, it can be shown by a standard continuity argument that the implications (3.8) and (3.9) remain true when "$\succ$" is replaced by "$\succeq$" everywhere. □

The converses of both (3.8) and (3.9) are false. In particular, take $n = 3$ and let

$$x = (\sqrt{2.5}, 1, 1), \quad y = (0, 0, 0), \quad w = (2\sqrt{2.5}, 2, 0).$$

Then, it can be checked that $w \succ_{\mathcal{K}^n} x \succ_{\mathcal{K}^n} 0$ and $L_w^2 \succ L_x^2 = L_x^2 + L_y^2$ while $w^2 \not\succ_{\mathcal{K}^n} x^2 = x^2 + y^2$. The property that $w \succ_{\mathcal{K}^n} x$ whenever $w \succ_{\mathcal{K}^n} 0$ and $w^2 \succ_{\mathcal{K}^n} x^2$ has a matrix counterpart. In particular, it can be shown that $W \succ X$ whenever $W \succ 0$ and $W^2 \succ X^2$, where W, X are $n \times n$ real symmetric matrices [20, Lemma 6.1(c)]. The converse of this implication is false, much like its vector counterpart.

By using Proposition 3.4, we have the following lemma on the positive definite property of certain matrices. This lemma, analogous to a result in the semidefinite setting [6, proof of Lemma 6], will be needed to establish the existence of Newton direction when a smoothing approach based on (4.6) is applied to solve SOCCP (see Proposition 6.2).

Lemma 3.1 *For any x, y in $\Re^n$ and any $w \succ_{\mathcal{K}^n} 0$, we have*

$$w^2 \succ_{\mathcal{K}^n} x^2 + y^2 \quad \Longrightarrow \quad (L_w - L_x)(L_w - L_y) \succ 0, \quad L_w - L_x \succ 0, \quad L_w - L_y \succ 0. \tag{3.16}$$

Moreover, (3.16) *remains true when "$\succ$" is replaced by "$\succeq$" everywhere.*

Proof. Fix any x, y in $\Re^n$ and any $w \succ_{\mathcal{K}^n} 0$. Suppose $w^2 \succ_{\mathcal{K}^n} x^2 + y^2$. Then, $w^2 \succ_{\mathcal{K}^n} x^2$ and $w^2 \succ_{\mathcal{K}^n} y^2$, so it follows from Proposition 3.4 that

$$L_w^2 \succ L_x^2 + L_y^2, \quad L_w \succ L_x, \quad L_w \succ L_y.$$

Thus, it suffices to prove that $(L_w - L_x)(L_w - L_y) \succ 0$. Since the matrix $(L_w - L_x)(L_w - L_y)$ is not necessarily symmetric, we need to prove that its symmetric part is positive definite. Let S denote the symmetric part of $(L_w - L_x)(L_w - L_y)$. Then

$$\begin{aligned} S &= L_w^2 - \frac{L_x + L_y}{2} L_w - L_w \frac{L_x + L_y}{2} + \frac{L_x L_y + L_y L_x}{2} \\ &= \left(\frac{1}{2} L_w^2 + \frac{1}{2} L_x^2 + \frac{1}{2} L_y^2 - \frac{L_x + L_y}{2} L_w - L_w \frac{L_x + L_y}{2} + \frac{L_x L_y + L_y L_x}{2} \right) + \frac{1}{2} \left(L_w^2 - L_x^2 - L_y^2 \right) \\ &= \frac{1}{2} \left(L_w - L_x - L_y \right)^2 + \frac{1}{2} \left(L_w^2 - L_x^2 - L_y^2 \right) \quad \succ \quad 0, \end{aligned}$$

where the third equality follows from completing the square, and the last step is due to the fact that $L_w^2 - L_x^2 - L_y^2 \succ 0$.

An analogous argument shows that the implication (3.16) remains true when "$\succ$" is replaced by "$\succeq$" everywhere. □

We note that $(L_w - L_x)(L_w - L_y) \succ 0$ is false if w is replaced more generally by any vector w such that $L_w - L_x \succ 0$ and $L_w - L_y \succ 0$. In other words, the product of two positive definite matrices of the form (2.4) may not be positive definite. Thus, the specific form of w (as it relates to x and y) appears to be crucial.

4 Smoothing Functions for SOCCP

In this and the next section, we use the results of the previous section to define and analyze smoothing functions for SOCCP (1.1). The first class of smoothing functions that we consider is a natural generalization of a proposal of Chen and Mangasarian [4, 5] in the case of NCP:

$$\phi_\mu(x, y) = x - \mu g((x - y)/\mu), \tag{4.1}$$

where $\mu > 0$ and $g \in \mathcal{CM}$. Here $\mathcal{CM}$ denotes the class of functions $g : \Re^n \to \Re^n$ defined by (3.1) with $\hat{g} : \Re \to \Re_+$ a continuously differentiable convex function satisfying $\lim_{\alpha \to -\infty} \hat{g}(\alpha) = 0$, $\lim_{\alpha \to \infty}(\hat{g}(\alpha) - \alpha) = 0$ and $0 < \hat{g}'(\alpha) < 1$ for all $\alpha \in \Re$ (see [21]). A special case of this smoothing function is a proposal of Chen and Harker [3], Kanzow [13], and Smale [18], corresponding to

$$\hat{g}(\alpha) = (\sqrt{\alpha^2 + 4} + \alpha)/2.$$

Another choice of $\hat{g}$ is obtained by integrating the sigmoid function $\alpha \mapsto 1/(1 + e^{-\alpha})$ used in neural networks [4]:

$$\hat{g}(\alpha) = \ln(e^\alpha + 1).$$

Using Proposition 3.3, we derive below simple formulas for ϕ_μ and ϕ_0, and we show that ϕ_0 satisfies (1.5).

Proposition 4.1 *For any $g \in \mathcal{CM}$ and $\mu > 0$, let $\phi_\mu : \Re^n \times \Re^n \to \Re^n$ be the smoothing function given by (4.1). Then the following results hold.*

(a) *For any $x = (x_1, x_2), y = (y_1, y_2) \in \Re \times \Re^{n-1}$, $\phi_\mu(x, y)$ has the formula*

$$\phi_\mu(x, y) = x - \mu(\hat{g}(\lambda_1/\mu)u^{(1)} + \hat{g}(\lambda_2/\mu)u^{(2)}), \tag{4.2}$$

where, for $i = 1, 2$,

$$\lambda_i \;=\; x_1 - y_1 + (-1)^i \|x_2 - y_2\|, \tag{4.3}$$

$$u^{(i)} \;=\; \begin{cases} \frac{1}{2}\Big(1, (-1)^i \dfrac{x_2 - y_2}{\|x_2 - y_2\|}\Big), & \text{if } \; x_2 \neq y_2, \\ \frac{1}{2}\left(1, (-1)^i w\right), & \text{if } \; x_2 = y_2, \end{cases} \tag{4.4}$$

with $w \in \Re^{n-1}$ *being an arbitrary vector satisfying* $\|w\| = 1$.

(b) *The pointwise limit* $\phi_0 = \lim_{\mu \to 0^+} \phi_\mu$ *has the formula*

$$\phi_0(x, y) \;=\; x - ([\lambda_1]_+ u^{(1)} + [\lambda_2]_+ u^{(2)}) \;=\; x - [x - y]_+, \tag{4.5}$$

where, for $i = 1, 2$, λ_i *and* $u^{(i)}$ *are given by* (4.3) and (4.4), *respectively, with* $w \in \Re^{n-1}$ *being an arbitrary vector satisfying* $\|w\| = 1$. *Moreover,* ϕ_0 *satisfies* (1.5) *with* $\mathcal{K} = \mathcal{K}^n$.

Proof. Part (a) follows from (4.1), (3.1) and (2.5)–(2.7). We prove part (b) below.

Since $\lim_{\alpha \to -\infty} \hat{g}(\alpha) = 0$ and $\lim_{\alpha \to \infty} (\hat{g}(\alpha) - \alpha) = 0$, we have $\lim_{\mu \to 0^+} \mu \hat{g}(\lambda_i/\mu) = [\lambda_i]_+$ for $i = 1, 2$. It then follows from (4.2) that

$$\begin{aligned} \lim_{\mu \to 0^+} \phi_\mu(x, y) \;&=\; x - \lim_{\mu \to 0^+} \mu \left(\hat{g}(\lambda_1/\mu) u^{(1)} + \hat{g}(\lambda_2/\mu) u^{(2)}\right) \\ &=\; x - \left([\lambda_1]_+ u^{(1)} + [\lambda_2]_+ u^{(2)}\right) \\ &=\; x - [x - y]_+, \end{aligned}$$

where λ_i, $i = 1, 2$, are given by (4.3), and the last equality follows from (3.5). This proves (4.5).

Finally, $\mathcal{K}^n$ is a closed convex cone that is self-dual, i.e., $\mathcal{K}^n = \{x \in \Re^n | \langle x, y \rangle \geq 0 \;\; \forall y \in \mathcal{K}^n\}$. Then, it is well known that two vectors x, y in $\Re^n$ satisfies $x = [x - y]_+$ if and only if $x \in \mathcal{K}^n$, $y \in \mathcal{K}^n$, $\langle x, y \rangle = 0$. This together with (4.5) proves (1.5). □

A second type of smoothing function that we consider is a generalization of a proposal of Kanzow [4] in the case of NCP, based on smoothing the Fischer-Burmeister function:

$$\phi_\mu(x, y) \;=\; x + y - (x^2 + y^2 + 2\mu^2 e)^{1/2}. \tag{4.6}$$

Notice that $\phi_\mu(x, y) = 0$ if and only if $x \in \text{int } \mathcal{K}^n$, $y \in \text{int } \mathcal{K}^n$ and $x \cdot y = \mu^2 e$. Analogous to Proposition 4.1, we derive below simple formulas for ϕ_μ and the corresponding ϕ_0, and we show that ϕ_0 satisfies (1.5).

Proposition 4.2 *For any* $\mu > 0$, *let* $\phi_\mu : \Re^n \times \Re^n \to \Re^n$ *be given by* (4.6). *Then the following results hold.*

(a) *For any* $x = (x_1, x_2), y = (y_1, y_2) \in \Re \times \Re^{n-1}$, $\phi_\mu(x, y)$ *has the formula*

$$\phi_\mu(x, y) = x + y - \left(\sqrt{\lambda_1} u^{(1)} + \sqrt{\lambda_2} u^{(2)}\right), \tag{4.7}$$

where, for $i = 1, 2$,

$$\lambda_i = \|x\|^2 + \|y\|^2 + 2\mu^2 + 2(-1)^i \|x_1 x_2 + y_1 y_2\|, \tag{4.8}$$

$$u^{(i)} = \begin{cases} \frac{1}{2}\left(1, (-1)^i \dfrac{x_1 x_2 + y_1 y_2}{\|x_1 x_2 + y_1 y_2\|}\right), & \text{if } x_1 x_2 + y_1 y_2 \neq 0, \\ \frac{1}{2}\left(1, (-1)^i w\right), & \text{if } x_1 x_2 + y_1 y_2 = 0, \end{cases} \tag{4.9}$$

with $w \in \Re^{n-1}$ *being an arbitrary vector satisfying* $\|w\| = 1$.

(b) *The pointwise limit* $\phi_0 = \lim_{\mu \to 0^+} \phi_\mu$ *has the formula*

$$\phi_0(x, y) = x + y - \left(\sqrt{\lambda_1} u^{(1)} + \sqrt{\lambda_2} u^{(2)}\right) = x + y - (x^2 + y^2)^{1/2}, \tag{4.10}$$

where, for $i = 1, 2$,

$$\lambda_i = \|x\|^2 + \|y\|^2 + 2(-1)^i \|x_1 x_2 + y_1 y_2\|, \tag{4.11}$$

and $u^{(i)}$ *is given by* (4.9), *with* $w \in \Re^{n-1}$ *being an arbitrary vector satisfying* $\|w\| = 1$. *Moreover,* ϕ_0 *satisfies* (1.5) *with* $\mathcal{K} = \mathcal{K}^n$.

Proof. Part (a) follows from (4.6), (3.3), and the observation that λ_1, λ_2 given by (4.8) are the spectral values of $x^2 + y^2 + 2\mu^2 e$. We prove part (b) below.

Fix any $x = (x_1, x_2), y = (y_1, y_2) \in \Re \times \Re^{n-1}$. For $i = 1, 2$, we have that λ_i given by (4.8) converges to λ_i given by (4.11) as $\mu \to 0^+$. Since $u^{(i)}$ given by (4.9) is independent of μ, then $\phi_\mu(x, y)$ given by (4.7) converges to the second term in (4.10), thus proving the first equality in (4.10). Using (3.3) and the observation that λ_1, λ_2 given by (4.11) are the spectral values of $x^2 + y^2$, we obtain the second equality in (4.10).

To show that (1.5) holds, let us first suppose $\langle x, y \rangle = 0$, $x \in \mathcal{K}^n$, $y \in \mathcal{K}^n$. Then, by Proposition 2.1, we have $x \cdot y = 0$, implying $(x + y)^2 = x^2 + y^2$ and hence $x + y = (x^2 + y^2)^{1/2}$, i.e., $\phi_0(x, y) = 0$. Conversely, suppose x and y satisfy $\phi_0(x, y) = 0$. Then, $x + y = (x^2 + y^2)^{1/2}$, so that, upon squaring both sides, we obtain $x \cdot y = 0$. Let $w = (x^2 + y^2)^{1/2}$. Then, $w \succeq_{\mathcal{K}^n} 0$ and $w^2 = x^2 + y^2$, implying $w^2 \succeq_{\mathcal{K}^n} x^2$ and $w^2 \succeq_{\mathcal{K}^n} y^2$. Thus, it follows from Proposition 3.4 that $w \succeq_{\mathcal{K}^n} x$ and $w \succeq_{\mathcal{K}^n} y$. As a result, $x = w - y \succeq_{\mathcal{K}^n} 0$ and $y = w - x \succeq_{\mathcal{K}^n} 0$, or equivalently, $x, y \in \mathcal{K}^n$. Since $x \cdot y = 0$ as shown above, this together with Proposition 2.1 completes the proof. □

5 Differential and Lipschitzian Properties of Smoothing Functions

In this section we study differential and Lipschitzian properties of the smoothing functions (4.1) and (4.6) introduced in the previous section.

Proposition 5.1 *For any $\mu > 0$, let $\phi_\mu : \Re^n \times \Re^n \to \Re^n$ be given by either (4.1) with $g \in \mathcal{CM}$ or (4.6). Then, for any x, $y \in \Re^n$ and any $\mu > \nu > 0$, we have*

$$\rho(\mu - \nu)e \succeq_{\mathcal{K}^n} \phi_\nu(x,y) - \phi_\mu(x,y) \succ_{\mathcal{K}^n} 0, \tag{5.1}$$

$$\rho\mu e \succeq_{\mathcal{K}^n} \phi_0(x,y) - \phi_\mu(x,y) \succ_{\mathcal{K}^n} 0, \tag{5.2}$$

where $\rho = \hat{g}(0)$ if ϕ_μ is given by (4.1) and $\rho = \sqrt{2}$ if ϕ_μ is given by (4.6).

Proof. Fix any $x = (x_1, x_2)$, $y = (y_1, y_2)$ in $\Re \times \Re^{n-1}$.

Suppose ϕ_μ is given by (4.1) with $g \in \mathcal{CM}$. Then, for any $\mu > \nu > 0$, Proposition 4.1 (a) yields

$$\phi_\nu(x,y) - \phi_\mu(x,y) = (\mu\hat{g}(\lambda_1/\mu) - \nu\hat{g}(\lambda_1/\nu))u^{(1)} + (\mu\hat{g}(\lambda_2/\mu) - \nu\hat{g}(\lambda_2/\nu))u^{(2)}, \tag{5.3}$$

where λ_i and $u^{(i)}$ are given by (4.3) and (4.4), for $i = 1, 2$. By [21, Lemma 3.1], we have $0 < \mu\hat{g}(\lambda_i/\mu) - \nu\hat{g}(\lambda_i/\nu) \leq \hat{g}(0)(\mu - \nu)$ for $i = 1, 2$. This together with (5.3) and Property 2.2 and $u^{(i)} \succeq_{\mathcal{K}^n} 0$ for $i = 1, 2$, yields

$$\begin{array}{rcl} 0 & \preceq_{\mathcal{K}^n} & \phi_\nu(x,y) - \phi_\mu(x,y) \\ & \preceq_{\mathcal{K}^n} & \hat{g}(0)(\mu - \nu)u^{(1)} + \hat{g}(0)(\mu - \nu)u^{(2)} \\ & = & \hat{g}(0)(\mu - \nu)e. \end{array}$$

This proves (5.1) with $\rho = \hat{g}(0)$.

Suppose ϕ_μ is given by (4.6). Then, for any $\mu > \nu > 0$, Proposition 4.2 (a) yields that

$$\begin{array}{rcl} \phi_\nu(x,y) - \phi_\mu(x,y) & = & \left(\sqrt{\lambda_1(\mu)} - \sqrt{\lambda_1(\nu)}\right)u^{(1)} + \left(\sqrt{\lambda_2(\mu)} - \sqrt{\lambda_2(\nu)}\right)u^{(2)} \\ & = & \dfrac{2(\mu^2 - \nu^2)}{\sqrt{\lambda_1(\mu)} + \sqrt{\lambda_1(\nu)}}u^{(1)} + \dfrac{2(\mu^2 - \nu^2)}{\sqrt{\lambda_2(\mu)} + \sqrt{\lambda_2(\nu)}}u^{(2)}, \end{array} \tag{5.4}$$

where we denote

$$\lambda_i(\mu) = \|x\|^2 + \|y\|^2 + 2\mu^2 + 2(-1)^i\|x_1x_2 + y_1y_2\|,$$

and $u^{(i)}$ is given by (4.9), for $i = 1, 2$. Since $\mu > \nu > 0$ so that $\mu^2 - \nu^2 > 0$, we obtain from (5.4) and $u^{(i)} \succeq_{\mathcal{K}^n} 0$ for $i = 1, 2$, that

$$\phi_\nu(x,y) - \phi_\mu(x,y) \succ_{\mathcal{K}^n} 0.$$

Moreover, since

$$\lambda_i(\mu) \geq 2\mu^2, \quad \lambda_i(\nu) \geq 2\nu^2, \quad i = 1, 2,$$

we have

$$\frac{2(\mu^2 - \nu^2)}{\sqrt{\lambda_i(\mu)} + \sqrt{\lambda_i(\nu)}} \leq \sqrt{2}(\mu - \nu), \quad i = 1, 2.$$

Then, we obtain from (5.4) and Property 2.2 that

$$\begin{array}{rcl} \phi_\nu(x,y) - \phi_\mu(x,y) & \preceq_{\mathcal{K}^n} & \sqrt{2}(\mu-\nu)u^{(1)} + \sqrt{2}(\mu-\nu)u^{(2)} \\ & = & \sqrt{2}(\mu-\nu)e. \end{array}$$

This proves (5.1) with $\rho = \sqrt{2}$.

We have shown for either choice of ϕ_μ that (5.1) holds for any $\mu > \nu > 0$. This implies that $-\phi_\nu$ is monotone in $\nu > 0$ with respect to the partial ordering $\succ_{\mathcal{K}^n}$. Then, taking $\nu \to 0^+$ in (5.1) yields

$$\rho\mu e \succeq_{\mathcal{K}^n} \lim_{\nu \to 0^+} \phi_\nu(x,y) - \phi_\mu(x,y) = \phi_0(x,y) - \phi_\mu(x,y) \succ_{\mathcal{K}^n} 0.$$

This proves (5.2). □

Next, we analyze the differential properties of the smoothing functions given by (4.1) and (4.6). We begin with the following key lemma showing that the smoothness property of $\hat{g}$ is inherited by the corresponding function g associated with SOC.

Proposition 5.2 *For any $\hat{g} : \Re \to \Re$ that is Fréchet-differentiable (respectively, continuously differentiable), the function $g : \Re^n \to \Re^n$ defined by (3.1) is Fréchet-differentiable (respectively, continuously differentiable) and its Jacobian at $z = (z_1, z_2) \in \Re \times \Re^{n-1}$ is given by the formula $\nabla g(z) = \hat{g}'(z_1)I$ if $z_2 = 0$, and otherwise is given by*

$$\nabla g(z) = \begin{bmatrix} b & c\ z_2^T/\|z_2\| \\ c\ z_2/\|z_2\| & aI + (b-a)z_2z_2^T/\|z_2\|^2 \end{bmatrix}, \tag{5.5}$$

where

$$a = \frac{\hat{g}(\lambda_2) - \hat{g}(\lambda_1)}{\lambda_2 - \lambda_1}, \quad b = \frac{1}{2}\left(\hat{g}'(\lambda_2) + \hat{g}'(\lambda_1)\right), \quad c = \frac{1}{2}\left(\hat{g}'(\lambda_2) - \hat{g}'(\lambda_1)\right) \tag{5.6}$$

with $\lambda_i = z_1 + (-1)^i\|z_2\|$, $i = 1, 2$. If $\hat{g}'(\alpha) > 0$ for all $\alpha \in \Re$, then $\nabla g(z)$ is positive definite for all $z \in \Re^n$.

Proof. Assume $\hat{g}$ is Fréchet-differentiable. Fix any $z = (z_1, z_2) \in \Re \times \Re^{n-1}$. First, we consider the case of $z_2 \neq 0$. By (3.1), we have

$$g(z) = \hat{g}(\lambda_1)u^{(1)} + \hat{g}(\lambda_2)u^{(2)}, \tag{5.7}$$

where

$$\lambda_i = z_1 + (-1)^i \|z_2\|, \qquad u^{(i)} = \frac{1}{2}\left(1, (-1)^i \frac{z_2}{\|z_2\|}\right), \qquad i = 1, 2.$$

Using the fact

$$\nabla_w \left(\frac{w}{\|w\|}\right) = \frac{1}{\|w\|}\left(I - \frac{ww^T}{\|w\|^2}\right), \qquad \forall w \neq 0,$$

we obtain that $u^{(i)}$ is Fréchet-differentiable with respect to z, with

$$\nabla_z u^{(i)} = \frac{(-1)^i}{2\|z_2\|}\begin{bmatrix} 0 & 0 \\ 0 & I - z_2 z_2^T/\|z_2\|^2 \end{bmatrix}, \qquad i = 1, 2. \tag{5.8}$$

Similarly, λ_i is Fréchet-differentiable with respect to z, with

$$\nabla_z \lambda_i = 2u^{(i)}, \qquad i = 1, 2. \tag{5.9}$$

Since $\hat{g}$ is Fréchet-differentiable, then using the chain rule and product rule for differentiation, we obtain from (5.7) that

$$\begin{aligned}
\nabla g(z) &= \hat{g}(\lambda_1)\nabla_z u^{(1)} + u^{(1)}(\nabla_z \hat{g}(\lambda_1))^T + \hat{g}(\lambda_2)\nabla_z u^{(2)} + u^{(2)}(\nabla_z \hat{g}(\lambda_2))^T \\
&= \hat{g}(\lambda_1)\nabla_z u^{(1)} + \hat{g}(\lambda_2)\nabla_z u^{(2)} + u^{(1)}\hat{g}'(\lambda_1)(\nabla_z \lambda_1)^T + u^{(2)}\hat{g}'(\lambda_2)(\nabla_z \lambda_2)^T \\
&= \frac{\hat{g}(\lambda_2) - \hat{g}(\lambda_1)}{2\|z_2\|}\begin{bmatrix} 0 & 0 \\ 0 & I - z_2 z_2^T/\|z_2\|^2 \end{bmatrix} + 2\hat{g}'(\lambda_1)u^{(1)}(u^{(1)})^T + 2\hat{g}'(\lambda_2)u^{(2)}(u^{(2)})^T \\
&= a\begin{bmatrix} 0 & 0 \\ 0 & I - z_2 z_2^T/\|z_2\|^2 \end{bmatrix} + 2\hat{g}'(\lambda_1)u^{(1)}(u^{(1)})^T + 2\hat{g}'(\lambda_2)u^{(2)}(u^{(2)})^T,
\end{aligned}$$

where the third equality uses (5.8) and (5.9), and the last equality uses the definition of a and the fact $\lambda_2 - \lambda_1 = 2\|z_2\|$. Notice that

$$u^{(i)}(u^{(i)})^T = \frac{1}{4}\begin{bmatrix} 1 & (-1)^i z_2^T/\|z_2\| \\ (-1)^i z_2/\|z_2\| & z_2 z_2^T/\|z_2\|^2 \end{bmatrix}, \quad i = 1, 2.$$

Substituting this into the previous expression and simplifying yield

$$\begin{aligned}
\nabla g(z) &= a\begin{bmatrix} 0 & 0 \\ 0 & I - z_2 z_2^T/\|z_2\|^2 \end{bmatrix} + \begin{bmatrix} b & c\ z_2^T/\|z_2\| \\ c\ z_2/\|z_2\| & b\ z_2 z_2{}^T/\|z_2\|^2 \end{bmatrix} \\
&= \begin{bmatrix} b & c\ z_2^T/\|z_2\| \\ c\ z_2/\|z_2\| & aI + (b-a)z_2 z_2^T/\|z_2\|^2 \end{bmatrix},
\end{aligned}$$

where we have used the definitions of b and c. This proves the proposition for the case of $z_2 \neq 0$.

Next we consider the case of $z_2 = 0$. We calculate the Jacobian matrix ∇g at z by perturbing z by $\Delta z = (\Delta z_1, \Delta z_2) \in \Re \times \Re^{n-1}$ and considering the resulting variation in g: First consider the case of $\Delta z_2 \neq 0$. We have from (3.1) that

$$g(z + \Delta z) = \hat{g}(\lambda + \Delta\lambda_1)u^{(1)} + \hat{g}(\lambda + \Delta\lambda_2)u^{(2)}, \qquad g(z) = \hat{g}(\lambda)u^{(1)} + \hat{g}(\lambda)u^{(2)},$$

where we denote

$$\lambda = z_1, \quad \Delta\lambda_i = \Delta z_1 + (-1)^i \|\Delta z_2\|, \quad u^{(i)} = \frac{1}{2}\left(1, (-1)^i \frac{\Delta z_2}{\|\Delta z_2\|}\right), \quad i = 1, 2.$$

We also have from the Taylor expansion of $\hat{g}$ at λ that

$$\hat{g}(\lambda + \Delta\lambda_i) - \hat{g}(\lambda) = \hat{g}'(\lambda)\Delta\lambda_i + o(\Delta\lambda_i) = \hat{g}'(\lambda)\Delta\lambda_i + o(\|\Delta z\|),$$

where $o(\cdot)$ is the usual "little o" notation, i.e., $\beta = o(\alpha)$ means $\beta/\alpha \to 0$ as $\alpha \to 0$. Combining the preceding two relations, we obtain

$$\begin{aligned}
g(z + \Delta z) - g(z) &= \left(\hat{g}(\lambda + \Delta\lambda_1)u^{(1)} + \hat{g}(\lambda + \Delta\lambda_2)u^{(2)}\right) - \left(\hat{g}(\lambda)u^{(1)} + \hat{g}(\lambda)u^{(2)}\right) \\
&= (\hat{g}(\lambda + \Delta\lambda_1) - \hat{g}(\lambda))u^{(1)} + (\hat{g}(\lambda + \Delta\lambda_2) - \hat{g}(\lambda))u^{(2)} \\
&= \hat{g}'(\lambda)\Delta\lambda_1 u^{(1)} + \hat{g}'(\lambda)\Delta\lambda_2 u^{(2)} + o(\|\Delta z\|) \\
&= \frac{\hat{g}'(\lambda)}{2}\left(\Delta\lambda_2 + \Delta\lambda_1, (\Delta\lambda_2 - \Delta\lambda_1)\frac{\Delta z_2}{\|\Delta z_2\|}\right) + o(\|\Delta z\|) \\
&= \hat{g}'(\lambda)\,(\Delta z_1, \Delta z_2) + o(\|\Delta z\|) \\
&= \hat{g}'(\lambda)\Delta z + o(\|\Delta z\|),
\end{aligned}$$

where the fourth equality follows from the definition of $u^{(i)}$, and the fifth equality uses the fact

$$\Delta\lambda_2 + \Delta\lambda_1 = 2\Delta z_1, \qquad \Delta\lambda_2 - \Delta\lambda_1 = 2\|\Delta z_2\|.$$

In the case of $\Delta z_2 = 0$, the same argument applies except $\Delta z_2/\|\Delta z_2\|$ is replaced by any $w \in \Re^{n-1}$ satisfying $\|w\| = 1$. This shows that g is Fréchet-differentiable at z and

$$\nabla g(z) = \hat{g}'(\lambda)I.$$

Assume $\hat{g}$ is continuously differentiable. It is readily seen from (5.5) and (5.6) that ∇g is continuous (entrywise) at every z with $z_2 \neq 0$. To argue the continuity of ∇g at every z with $z_2 = 0$, fix any $x = (x_1, 0) \in \Re^n$ and consider the limit of $\nabla g(z)$ as $z = (z_1, z_2)$ approaches x. Let $\lambda_i = z_1 + (-1)^i\|z_2\|$, $i = 1, 2$, and a, b, c be given by (5.6). Then, since $\lim_{z\to x} \lambda_1 = \lim_{z\to x} \lambda_2 = x_1$, we have

$$\lim_{z\to x} a = \lim_{z\to x} \frac{\hat{g}(\lambda_2) - \hat{g}(\lambda_1)}{\lambda_2 - \lambda_1} = \hat{g}'(x_1), \quad \lim_{z\to x} b = \hat{g}'(x_1), \quad \lim_{z\to x} c = 0.$$

Thus, taking the limit in (5.5) as $z \to x$ yields

$$\lim_{z \to x} \nabla g(z) = \hat{g}'(x_1) I = \nabla g(x).$$

Lastly, assume $\hat{g}'(\alpha) > 0$ for all $\alpha \in \Re$. Fix any $z = (z_1, z_2)$. We will show that $\nabla g(z)$ is positive definite. If $z_2 \neq 0$, then $\nabla g(z)$ is given by (5.5) with a, b, c given by (5.6) and $\lambda_i = z_1 + (-1)^i \|z_2\|$, $i = 1, 2$. Since $b > 0$, it suffices to show that the Schur complement of b in $\nabla g(z)$ is positive definite. This Schur complement has the form

$$aI + (b-a)\frac{z_2 z_2^T}{\|z_2\|^2} - \frac{c^2 z_2 z_2^T}{b\|z_2\|^2} = a\left(I - \frac{z_2 z_2^T}{\|z_2\|^2}\right) + b\left(1 - \frac{c^2}{b^2}\right)\frac{z_2 z_2^T}{\|z_2\|^2}.$$

Since $a > 0$, $b > 0$ and $b > c \geq 0$, the right-hand side is a linear positive combination of the matrices $I - z_2 z_2^T/\|z_2\|^2$ and $z_2 z_2^T/\|z_2\|^2$, so it is positive definite. If $z_2 = 0$, then $\nabla g(z) = \hat{g}'(z_1) I$, which is positive definite due to $\hat{g}'(z_1) > 0$. □

Notice that the Jacobian $\nabla g(z)$ is symmetric. Moreover, it can be seen that Proposition 5.2 still holds if "$\hat{g} : \Re \to \Re$" is replaced by "$\hat{g} : \Re_{++} \to \Re$" and "$z \in \Re^n$" is replaced by "$z \in \text{int } \mathcal{K}^n$." By using Proposition 5.2 and this observation, we obtain the following differentiability results for the smoothing functions given by (4.1) and (4.6).

Corollary 5.1 *For any $\mu > 0$ and any $g \in \mathcal{CM}$, the Chen-Mangasarian smoothing function $\phi_\mu : \Re^n \times \Re^n \to \Re^n$ given by (4.1) is continuously differentiable and its Jacobian is given by*

$$\nabla \phi_\mu(x, y) = \begin{bmatrix} I - \nabla g(z) \\ \nabla g(z) \end{bmatrix},$$

where $z = (x - y)/\mu$ and $\nabla g(z)$ has the formula in Proposition 5.2.

Corollary 5.2 *For any $\mu > 0$, the smoothed Fischer-Burmeister function $\phi_\mu : \Re^n \times \Re^n \to \Re^n$ given by (4.6) is continuously differentiable and its Jacobian can be written as*

$$\nabla \phi_\mu(x, y) = \begin{bmatrix} I - 2L_x \nabla g(z) \\ I - 2L_y \nabla g(z) \end{bmatrix} = \begin{bmatrix} I - L_x L_w^{-1} \\ I - L_y L_w^{-1} \end{bmatrix},$$

where $z = x^2 + y^2 + 2\mu^2 e$, $w = z^{1/2}$, and $\nabla g(z)$ has the formula in Proposition 5.2 with $\hat{g}(\alpha) = \alpha^{1/2}$ for all $\alpha \in \Re_{++}$.

Proof. Fix any $\mu > 0$. Define the mapping $Z : \Re^n \times \Re^n \to \Re^n$ by $Z(x, y) = x^2 + y^2 + 2\mu^2 e$. Direct calculation yields

$$Z(x, y) = (\|x\|^2 + \|y\|^2 + 2\mu^2, 2(x_1 x_2 + y_1 y_2)),$$

from which we see that Z is continuously differentiable and that

$$\nabla_x Z(x,y) = 2\begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{bmatrix} = 2L_x, \qquad \nabla_y Z(x,y) = 2\begin{bmatrix} y_1 & y_2^T \\ y_2 & y_1 I \end{bmatrix} = 2L_y. \tag{5.10}$$

Also, since $x^2 \succeq_{\mathcal{K}^n} 0$, $y^2 \succeq_{\mathcal{K}^n} 0$, $e \succ_{\mathcal{K}^n} 0$, we have that $Z(x,y) \succ_{\mathcal{K}^n} 0$ for all $(x,y) \in \Re^n \times \Re^n$, i.e., Z maps $\Re^n \times \Re^n$ into int $\mathcal{K}^n$. Since $\hat{g}(\cdot) = (\cdot)^{1/2}$ is continuously differentiable on $\Re_{++}$, it follows from Proposition 5.2 and the subsequent remark that g is continuously differentiable on int $\mathcal{K}^n$. Thus, by the chain rule for differentiation, the composite mapping $\psi_\mu = g \circ Z$ is continuously differentiable and

$$\nabla_x \psi_\mu(x,y) = \nabla_x Z(x,y) \nabla g(Z(x,y)) \ , \qquad \nabla_y \psi_\mu(x,y) = \nabla_y Z(x,y) \nabla g(Z(x,y)) \ .$$

This together with (5.10) and the following fact

$$\nabla g(z) = \frac{1}{2} L_w^{-1} \qquad \text{with} \qquad w = z^{1/2} \tag{5.11}$$

yields the desired Jacobian formula.

It remains to verify (5.11). Although this can be verified by using (5.5), (5.6) and $\hat{g}(\alpha) = \alpha^{1/2}$ to calculate $\nabla g(z)^{-1}$ explicitly and comparing it with L_w, the algebra becomes quite involved. Instead we will use the following simpler argument: Since $z \succ_{\mathcal{K}^n} 0$, there exists scalar $\delta > 0$ such that $z + d \succ_{\mathcal{K}^n} 0$ for all $d \in \Re^n$ with $\|d\| < \delta$. For any such d, we have upon letting $u = (z+d)^{1/2} - w$ and using $w = z^{1/2}$ that

$$d = (w+u)^2 - w^2 = 2w \cdot u + u^2 = 2L_w u + u^2.$$

Since $w \succ_{\mathcal{K}^n} 0$, then L_w is invertible so that applying L_w^{-1} to both sides yields

$$L_w^{-1} d = 2u + L_w^{-1} u^2.$$

This shows that $\|u\|$ is in the order of $\|d\|$ whenever $\|d\|$ is sufficiently small. Moreover,

$$g(z+d) - g(z) = u = \frac{1}{2} L_w^{-1} d - \frac{1}{2} L_w^{-1} u^2,$$

and it follows that $\nabla g(z) = \frac{1}{2} L_w^{-1}$. □

For simplicity, we have considered in this section only the case of a single SOC. Extension of the results in this section to the case of a direct product of SOCs, i.e., (1.2) is straightforward and is omitted.

6 A Smoothing Approach to Solving SOCCP

In this section, we study the use of smoothing functions in a smoothing approach to solving the SOCCP (1.1), where $F : \Re^n \times \Re^n \times \Re^\ell \to \Re^n \times \Re^\ell$ is continuously differentiable and $\mathcal{K}$ is given by

(1.2) with $m, n_1, \ldots, n_m \geq 1$ and $n_1 + \cdots + n_m = n$. Our study bears some similarity with that given in [6] for the SDCP, though our analysis is quite different due to the different structures of the SOC.

In what follows, we write

$$x = (x_1, \ldots, x_m) \in \Re^n$$

to implicitly mean $x_i \in \Re^{n_i}$, $i = 1, \ldots, m$. Then, using the direct product structure of (1.2), the SOCCP (1.1) may be written equivalently as

$$\langle x_i, y_i \rangle = 0, \quad x_i \in \mathcal{K}^{n_i}, \quad y_i \in \mathcal{K}^{n_i}, \quad i = 1, \ldots, m, \qquad F(x, y, \zeta) = 0.$$

We choose, for each $i \in \{1, \ldots, m\}$, a continuously differentiable function $\phi_\mu^i : \Re^{n_i} \times \Re^{n_i} \to \Re^{n_i}$, parameterized by $\mu > 0$, such that the pointwise limit $\phi_0^i(x_i, y_i) = \lim_{\mu \to 0^+} \phi_\mu^i(x_i, y_i)$ satisfies

$$\langle x_i, y_i \rangle = 0, \quad x_i \in \mathcal{K}^{n_i}, \quad y_i \in \mathcal{K}^{n_i} \quad \Longleftrightarrow \quad \phi_0^i(x_i, y_i) = 0.$$

Then the function $\phi_\mu : \Re^n \times \Re^n \to \Re^n$ defined by

$$\phi_\mu(x, y) = (\phi_\mu^1(x_1, y_1), \ldots, \phi_\mu^m(x_m, y_m)) \quad \forall x, y \in \Re^n \tag{6.1}$$

is continuously differentiable and satisfies (1.5). The system of smooth equations (1.6) decomposes accordingly.

In the smoothing approach to solving (1.1), we fix $\mu > 0$ and solve (1.6) approximately by applying a few Newton steps; then we decrease μ and repeat this iteration. For the Newton direction to be well defined and unique, it is essential that the Jacobian matrix of the left-hand side of (1.6), viewed as a mapping from $\Re^{2n+\ell}$ to $\Re^{2n+\ell}$, be invertible. To this end, consider the following rank and monotonicity assumptions on $\nabla F(x, y, \zeta)$:

$$\operatorname{rank} \nabla_\zeta F(x, y, \zeta) = \ell, \tag{6.2}$$

$$(u, v, \Delta) \in \Re^n \times \Re^n \times \Re^\ell, \quad \nabla F(x, y, \zeta)^T (u, v, \Delta) = 0 \quad \Longrightarrow \quad u^T v \geq 0. \tag{6.3}$$

These assumptions are reasonable and, in particular, hold for convex programs with linear and SOC constraints (see next section). When $F(x, y, \zeta)$ has the form (1.3), the above assumptions reduce to the assumption that $\nabla F_0(x)$ is positive semidefinite. The following proposition shows that if (6.2) and (6.3) hold and each ϕ_μ^i is given by (4.1), then the aforementioned Jacobian is invertible for any $\mu > 0$, so the corresponding Newton direction is well defined.

Proposition 6.1 *Assume $F : \Re^n \times \Re^n \times \Re^\ell \to \Re^n \times \Re^\ell$ is continuously differentiable. Let $\phi_\mu : \Re^n \times \Re^n \to \Re^n$ be defined by* (6.1) *with*

$$\phi_\mu^i(x_i, y_i) = x_i - \mu g((x_i - y_i)/\mu) \quad \forall x_i, y_i \in \Re^{n_i},$$

and $g \in \mathcal{CM}$. *Then, for each* $\mu > 0$ *and* $(x, y, \zeta) \in \Re^n \times \Re^n \times \Re^\ell$ *satisfying* (6.2) *and* (6.3), *the matrix* $J_\mu(x, y, \zeta)$ *is invertible, where*

$$J_\mu(x,y,\zeta) = \begin{bmatrix} I - \mathrm{diag}\{\nabla g(z_i)\}_{i=1}^m & \mathrm{diag}\{\nabla g(z_i)\}_{i=1}^m & 0 \\ \nabla_x F(x,y,\zeta)^T & \nabla_y F(x,y,\zeta)^T & \nabla_\zeta F(x,y,\zeta)^T \end{bmatrix}, \tag{6.4}$$

$z_i = (x_i - y_i)/\mu$ *and* $\nabla g(z_i) \in \Re^{n_i \times n_i}$ *is specified as in Proposition 5.2.*

Proof. Fix any $\mu > 0$ and any $(x, y, \zeta) \in \Re^n \times \Re^n \times \Re^\ell$ satisfying (6.2) and (6.3). By using (6.1) and Corollary 5.1, it is readily verified that $J_\mu(x, y, \zeta)$ given by (6.4) is indeed the Jacobian matrix for the left-hand side of (1.6) with ϕ_μ as defined. We show below that $J_\mu(x, y, \zeta)$ is invertible.

Let $(u, v, \Delta) \in \Re^n \times \Re^n \times \Re^\ell$ be a vector in the null space of $J_\mu(x, y, \zeta)$. We will show that $u = v = 0$, $\Delta = 0$. By (6.4),

$$(I - B)\, u + Bv = 0, \qquad \nabla F(x,y,\zeta)^T (u, v, \Delta) = 0, \tag{6.5}$$

where for simplicity we denote $B = \mathrm{diag}\{B_1, B_2, \ldots, B_m\}$, and $B_i = \nabla g(z_i)$, with $z_i = (x_i - y_i)/\mu$. In what follows, we write $z_i = (z_{i1}, z_{i2}) \in \Re \times \Re^{n_i - 1}$ for $i = 1, \ldots, m$.

We claim that $I \succ B \succ 0$. To see this, fix any $i \in \{1, \ldots, m\}$. Suppose $z_{i2} = 0$. Then Proposition 5.2 yields $B_i = \hat{g}'(z_{i1})I$. Since $g \in \mathcal{CM}$ so that $1 > \hat{g}'(z_{i1}) > 0$, this implies $I \succ B_i \succ 0$. Suppose $z_{i2} \neq 0$. Then Proposition 5.2 yields that B_i is symmetric and

$$I - B_i = \begin{bmatrix} 1 - b_i & -\dfrac{c_i {z_{i2}}^T}{\|z_{i2}\|} \\ -\dfrac{c_i z_{i2}}{\|z_{i2}\|} & (1 - a_i)I + (a_i - b_i)\dfrac{z_{i2}{z_{i2}}^T}{\|z_{i2}\|^2} \end{bmatrix},$$

where

$$a_i = \frac{\hat{g}(\lambda_{i2}) - \hat{g}(\lambda_{i1})}{\lambda_{i2} - \lambda_{i1}}, \quad b_i = \frac{1}{2}(\hat{g}'(\lambda_{i2}) + \hat{g}'(\lambda_{i1})), \quad c_i = \frac{1}{2}(\hat{g}'(\lambda_{i2}) - \hat{g}'(\lambda_{i1})),$$

with $\lambda_{ij} = z_{i1} + (-1)^j \|z_{i2}\|$ for $j = 1, 2$. Since $g \in \mathcal{CM}$, we have $1 > \hat{g}'(\alpha) > 0$ for all $\alpha \in \Re$. Also, $\lambda_{i2} > \lambda_{i1}$ and convexity of $\hat{g}$ imply $\hat{g}'(\lambda_{i2}) \geq \hat{g}'(\lambda_{i1})$. Thus,

$$1 > a_i > 0, \quad b_i > c_i \geq 0, \quad 1 - b_i > c_i \geq 0. \tag{6.6}$$

Since $0 < b_i < 1$, it suffices to show that the Schur complement of b_i in $I - B_i$ is positive definite. This Schur complement has the form

$$\begin{aligned} & (1 - a_i)I + (a_i - b_i)\frac{z_{i2}{z_{i2}}^T}{\|z_{i2}\|^2} - \frac{c_i^2 z_{i2}{z_{i2}}^T}{(1 - b_i)\|z_{i2}\|^2} \\ = \; & (1 - a_i)\left(I - \frac{z_{i2}{z_{i2}}^T}{\|z_{i2}\|^2}\right) + (1 - b_i)\left(1 - \frac{c_i^2}{(1 - b_i)^2}\right)\frac{z_{i2}{z_{i2}}^T}{\|z_{i2}\|^2}, \end{aligned}$$

which, by (6.6), is a linear positive combination of $I - z_{i2}z_{i2}{}^T/\|z_{i2}\|^2$ and $z_{i2}z_{i2}{}^T/\|z_{i2}\|^2$ and hence is positive definite. Thus, $I \succ B_i$. Also, Proposition 5.2 shows that $B_i \succ 0$. Thus, we have shown that $I \succ B_i \succ 0$ for all i or, equivalently, $I \succ B \succ 0$.

Since B is invertible, multiplying both sides of the first equation in (6.5) by $u^T B^{-1}$ yields

$$u^T(B^{-1} - I)u + u^T v = 0. \tag{6.7}$$

Since B is symmetric and $I \succ B \succ 0$, It follows that B^{-1} is symmetric and $B^{-1} \succ I$, so that $u^T(B^{-1} - I)u \geq 0$. Also, by the assumption (6.3), we have from the second equation in (6.5) that $u^T v \geq 0$. Thus, (6.7) implies $u^T(B^{-1} - I)u = 0$ and hence $u = 0$. Since B is invertible, the first equation in (6.5) then implies $v = 0$. Since $\nabla_\zeta F(x, y, \zeta)$ has rank ℓ, the second equation in (6.5) then implies $\Delta = 0$. Thus, the null space of $J_\mu(x, y, \zeta)$ comprises only the origin, so $J_\mu(x, y, \zeta)$ is invertible. □

The next proposition shows that, when the smoothed Fischer-Burmeister function is used, the Jacobian matrix for the left-hand side of the equation (1.6) is invertible. The proof uses Corollary 5.2 and Lemma 3.1.

Proposition 6.2 *Assume $F : \Re^n \times \Re^n \times \Re^\ell \to \Re^n \times \Re^\ell$ is continuously differentiable. Let $\phi_\mu : \Re^n \times \Re^n \to \Re^n$ be defined by* (6.1) *with*

$$\phi_\mu^i(x_i, y_i) = x_i + y_i - (x_i^2 + y_i^2 + 2\mu^2 e)^{1/2} \quad \forall x_i, y_i \in \Re^{n_i}.$$

Then, for each $\mu > 0$ and $(x, y, \zeta) \in \Re^n \times \Re^n \times \Re^\ell$ satisfying (6.2) *and* (6.3)*, the matrix $J_\mu(x, y, \zeta)$ is invertible, where*

$$J_\mu(x, y, \zeta) = \begin{bmatrix} \mathrm{diag}\{\nabla_{x_i}\phi_\mu^i(x_i, y_i)^T\}_{i=1}^m & \mathrm{diag}\{\nabla_{y_i}\phi_\mu^i(x_i, y_i)^T\}_{i=1}^m & 0 \\ \nabla_x F(x, y, \zeta)^T & \nabla_y F(x, y, \zeta)^T & \nabla_\zeta F(x, y, \zeta)^T \end{bmatrix}, \tag{6.8}$$

$\nabla_{x_i}\phi_\mu^i(x_i, y_i)$ and $\nabla_{y_i}\phi_\mu^i(x_i, y_i)$ are given as in Corollary 5.2 with x, y, w, z replaced by x_i, y_i, w_i, z_i, respectively, and $\nabla g(z_i) \in \Re^{n_i \times n_i}$ is specified as in Proposition 5.2 with $\hat{g}(\alpha) = \alpha^{1/2}$ for all $\alpha \in \Re_{++}$.

Proof. Fix any $\mu > 0$ and $(x, y, \zeta) \in \Re^n \times \Re^n \times \Re^\ell$ satisfying (6.2) and (6.3). From (6.1), it is easily seen that $J_\mu(x, y, \zeta)$ given by (6.8) is the Jacobian matrix of the left-hand side of the equation (1.6) with ϕ_μ as defined. We show below that $J_\mu(x, y, \zeta)$ is invertible.

Let $(u, v, \Delta) \in \Re^n \times \Re^n \times \Re^\ell$ be a vector in the null space of $J_\mu(x, y, \zeta)$. We will show that $u = v = 0$, $\Delta = 0$. By Corollary 5.2, we have $\nabla_{x_i}\phi_\mu^i(x_i, y_i)^T = I - L_{w_i}^{-1}L_{x_i}$ and $\nabla_{y_i}\phi_\mu^i(x_i, y_i)^T = I - L_{w_i}^{-1}L_{y_i}$, where $w_i = z_i^{1/2}$ and $z_i = x_i^2 + y_i^2 + 2\mu^2 e$. Then, by (6.8),

$$(I - L_w^{-1}L_x)u + (I - L_w^{-1}L_y)v = 0, \qquad \nabla F(x, y, \zeta)^T(u, v, \Delta) = 0, \tag{6.9}$$

where for simplicity we denote

$$L_w = \mathrm{diag}\,\{L_{w_1}, \ldots, L_{w_m}\}, \quad L_x = \mathrm{diag}\,\{L_{x_1}, \ldots, L_{x_m}\}, \quad L_y = \mathrm{diag}\,\{L_{y_1}, \ldots, L_{y_m}\}.$$

Applying L_w to both sides of the first equation in (6.9) yields

$$(L_w - L_x)u + (L_w - L_y)v = 0.$$

Since $w_i^2 = x_i^2 + y_i^2 + 2\mu^2 e \succ_{\mathcal{K}^{n_i}} x_i^2 + y_i^2$, Lemma 3.1 shows that $L_{w_i} - L_{y_i}$ is invertible for all i and hence $L_w - L_y$ is invertible. Then, multiplying both sides on the left by $u^T(L_w - L_y)^{-1}$ yields

$$u^T(L_w - L_y)^{-1}(L_w - L_x)u + u^T v = 0. \tag{6.10}$$

Lemma 3.1 shows that $(L_{w_i} - L_{x_i})(L_{w_i} - L_{y_i}) \succ 0$ for all i and hence $(L_w - L_x)(L_w - L_y) \succ 0$. Then $u^T(L_w - L_y)^{-1}(L_w - L_x)u = d^T(L_w - L_x)(L_w - L_y)d \geq 0$, where we let $d = (L_w - L_y)^{-1}u$. Also, by the assumption (6.3), we have from the second equation in (6.9) that $u^T v \geq 0$. Thus, (6.10) implies $d^T(L_w - L_x)(L_w - L_y)d = 0$ and hence $d = 0$. Then $u = 0$ and since $I - L_w^{-1}L_y$ is invertible, the first equation in (6.9) implies $v = 0$. Since $\nabla_\zeta F(x, y, \zeta)$ has rank ℓ, the second equation in (6.9) then implies $\Delta = 0$. Thus, the null space of $J_\mu(x, y, \zeta)$ comprises only the origin, so $J_\mu(x, y, \zeta)$ is invertible. □

Thus far we have not specified an algorithm for solving SOCCP. This is on purpose since algorithm design and analysis is not a focus of this paper. In some cases and respects, a smoothing algorithm designed for solving NCP can be extended to solve SOCCP in a relatively straightforward manner. For example, the algorithm in [21], which was extended in [6] to solve SDCP, can also be extended to solve SOCCP. Its global convergence analysis [21, Proposition 3.1] then extends accordingly using Propositions 6.1 and 6.2.

7 Applications to Optimization with SOC Constraints

Consider the following optimization problem with SOC constraints:

$$\begin{array}{ll} \text{minimize} & f(\zeta) \\ \text{subject to} & h(\zeta) - x = 0, \quad x \in \mathcal{K}, \end{array} \tag{7.1}$$

where $f : \Re^\ell \to \Re$ and $h : \Re^\ell \to \Re^n$ are twice continuously differentiable functions and $\mathcal{K}$ has the form (1.2). In the special case where both f and g are affine, this problem reduces to that studied in [14]. We discuss below how results from the previous section can be applied to solve this problem.

By attaching Lagrange multiplier vector $y \in \Re^n$ to the equality constraints in (7.1) and introducing the Lagrangian

$$l(x, y, \zeta) = f(\zeta) - h(\zeta)^T y + x^T y,$$

we can write the first-order necessary optimality conditions for (7.1) in the form of SOCCP (1.1) with

$$F(x,y,\zeta) = \begin{pmatrix} \nabla_y l(x,y,\zeta) \\ \nabla_\zeta l(x,y,\zeta) \end{pmatrix} = \begin{pmatrix} x - h(\zeta) \\ \nabla f(\zeta) - \nabla h(\zeta) y \end{pmatrix}. \tag{7.2}$$

Notice that F is continuously differentiable. Moreover, it is straightforward to verify that $\nabla F(x,y,\zeta)$ satisfies the rank and monotonicity assumptions (6.2) and (6.3) if

$$\operatorname{rank} \nabla h(\zeta) = \ell \qquad \text{and} \qquad \nabla^2_{\zeta\zeta} l(x,y,\zeta) \succeq 0.$$

In particular, this is satisfied if f is convex and h is affine with ∇h having linearly independent columns.

Using (7.2), it can be seen that the Newton equation associated with (1.6) has the form

$$\begin{aligned} \nabla_x \phi_\mu(x,y)^T u + \nabla_y \phi_\mu(x,y)^T v &= -\phi_\mu(x,y), \\ u - A^T \Delta &= h(\zeta) - x, \\ Q\Delta - Av &= -\nabla_\zeta l(x,y,\zeta), \end{aligned}$$

where for simplicity we denote $A = \nabla h(\zeta) \in \Re^{\ell \times n}$ and $Q = \nabla^2_{\zeta\zeta} l(x,y,\zeta) \in \Re^{\ell \times \ell}$. Upon eliminating u and v from these equations, we obtain the reduced system

$$\Big(Q + A(\nabla_y \phi_\mu(x,y)^T)^{-1} \nabla_x \phi_\mu(x,y)^T A^T\Big)\Delta = r,$$

where $r = -\nabla_\zeta l(x,y,\zeta) - A(\nabla_y \phi_\mu(x,y)^T)^{-1} \Big(\phi_\mu(x,y) + \nabla_x \phi_\mu(x,y)^T (h(\zeta) - x)\Big)$. Upon writing $A = [A_1 \ \cdots \ A_m]$, with $A_i \in \Re^{\ell \times n_i}$ for $i = 1, \ldots, m$, and using the product structure (6.1), we can rewrite the above reduced system as

$$\left(Q + \sum_{i=1}^m A_i D_i A_i^T\right) \Delta = r,$$

where for simplicity we denote

$$D_i = (\nabla_{y_i} \phi^i_\mu(x_i,y_i)^T)^{-1} \nabla_{x_i} \phi^i_\mu(x_i,y_i)^T.$$

Here and throughout, our notations are as in the previous section. We show below that the left-hand matrix can be efficiently computed for the choices of ϕ^i_μ specified in Proposition 6.1 or 6.2. In particular, we show that D_i can be efficiently computed and has a sparsity structure.

For ϕ^i_μ as specified in Proposition 6.1, we have

$$D_i = \nabla g(z_i)^{-1} - I$$

with $z_i = (x_i - y_i)/\mu$. If $z_{i2} = 0$, then Proposition 5.2 yields $\nabla g(z_i)^{-1} = \hat{g}'(z_{i1})^{-1} I$. If $z_{i2} \neq 0$, then Proposition 5.2 together with $a_i b_i = b_i^2 - c_i^2$ yields

$$\nabla g(z_i)^{-1} = \frac{1}{a_i} \begin{bmatrix} 1 & -c_i z_{i2}^T/(b_i \|z_{i2}\|) \\ -c_i z_{i2}/(b_i \|z_{i2}\|) & I \end{bmatrix},$$

where a_i, b_i, c_i are as defined in the proof of Proposition 6.1. Thus D_i is almost diagonal, except for a dense first row and first column.

For ϕ_μ^i as specified in Proposition 6.2, we have by Corollary 5.2

$$\begin{aligned} D_i &= (I - L_{w_i}^{-1} L_{y_i})^{-1} (I - L_{w_i}^{-1} L_{x_i}) \\ &= (L_{w_i} - L_{y_i})^{-1} (L_{w_i} - L_{x_i}) \\ &= (L_{w_i - y_i})^{-1} L_{w_i - x_i}, \end{aligned}$$

where $w_i = (x_i^2 + y_i^2 + 2\mu^2 e)^{1/2}$. When calculating $L_{w_i - x_i}$, we only need to compute and store the vector $w_i - x_i$ since it can be used to easily determine $L_{w_i - x_i}$. The main computational effort therefore lies in computing and storing the matrix $L_{w_i - y_i}^{-1}$. Since $L_{w_i - y_i}$ is the sum of the matrix $(w_{i1} - y_{i1})I$, whose inverse is easily obtained, and a symmetric rank two matrix, we can store only the vector $w_i - y_i$ and use the well-known Sherman-Morrison-Woodbury updating formula to evaluate $L_{w_i - y_i}^{-1}$. The total number of such updates to compute all the D_i's is then proportional to m, and the total storage requirement is proportional to n.

References

[1] F. Alizadeh and S. Schmieta, *Symmetric cones, potential reduction methods*, in Handbook of Semidefinite Programming, H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds., Kluwer, Boston, 2000, pp. 195–233.

[2] J. Burke and S. Xu, *The global linear convergence of a non-interior path-following algorithm for linear complementarity problems*, Math. Oper. Res., 23 (1998), pp. 719–734.

[3] B. Chen and P. T. Harker, *A non-interior-point continuation method for linear complementarity problems*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 1168–1190.

[4] C. Chen and O. L. Mangasarian, *Smoothing methods for convex inequalities and linear complementarity problems*, Math. Programming, 71 (1995), pp. 51–69.

[5] C. Chen and O. L. Mangasarian, *A class of smoothing functions for nonlinear and mixed complementarity problems*, Comput. Optim. Appl., 5 (1996), pp. 97–138.

[6] X. Chen and P. Tseng, *Non-interior continuation methods for solving semidefinite complementarity problems,* Manuscript, Department of Mathematics, University of Washington, Seattle, October 1999.

[7] B. Chen and N. Xiu, *A global linear and local quadratic non-interior continuation method for nonlinear complementarity problems based on Chen-Mangasarian smoothing function*, SIAM J. Optim., 9 (1999), pp. 605–623.

[8] X. Chen, L. Qi, and D. Sun, *Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities*, Math. Comp., 67 (1998), pp. 519–540.

[9] U. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford Mathematical Monographs, Oxford University Press, New York, 1994.

[10] M. C. Ferris and J.-S. Pang, eds., *Complementarity and Variational Problems: State of the Art*, SIAM, Philadelphia, 1997.

[11] M. Fukushima and L. Qi, eds., *Reformulation – Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, Kluwer, Boston, 1999.

[12] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[13] C. Kanzow, *Some noninterior continuation methods for linear complementarity problems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 851–868.

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

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

[16] L. Qi and H. Jiang, *Semismooth Karush-Kuhn-Tucker equations and convergence analysis of Newton and quasi-Newton methods for solving these equations*, Math. Oper. Res., 22 (1997), pp. 301–325.

[17] S. Schmieta and F. Alizadeh, *Associative algebras, symmetric cones and polynomial time interior point algorithms*, Report RRR 17-98, RUTCOR, Rutgers University, Piscataway, 1998.

[18] S. Smale, *Algorithms for solving equations*, in Proceedings of the International Congress of Mathematicians, A. M. Gleason, ed., American Mathematical Society, Providence, 1987, pp. 172–195.

[19] D. Sun and J. Sun, *Semismooth matrix valued functions*, Manuscript, School of Mathematics, The University of New South Wales, Sydney, January 2000.

[20] P. Tseng, *Merit functions for semi-definite complementarity problems*, Math. Programming, 83 (1998), pp. 159–185.

[21] P. Tseng, *Analysis of a non-interior continuation methods based on Chen-Mangasarian smoothing functions for complementarity problems,* in Reformulation – Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, M. Fukushima and L. Qi, eds., Kluwer, Boston, 1999, pp. 381–404.

[22] T. Tsuchiya, *A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming*, Optim. Methods Softw., to appear.

[23] R. J. Vanderbei and H. Y. Benson, *On formulating semidefinite programming problems as smooth convex nonlinear optimization problems*, ORFE 99-01, Department of Operations Research and Financial Engineering, Princeton University, Princeton NJ, 1999.

[24] N. Yamashita and M. Fukushima, *Modified Newton methods for solving semismooth reformulations of monotone complementarity problems*, Math. Programming, 76 (1997), pp. 469–491.