

SEMISMOOTH METHODS FOR LINEAR AND NONLINEAR SECOND-ORDER CONE PROGRAMS¹

Christian Kanzow², Izabella Ferenczi³, and Masao Fukushima⁴

April 19, 2006 (revised January 10, 2007)

Abstract

The optimality conditions of a nonlinear second-order cone program can be reformulated as a nonsmooth system of equations using a projection mapping. This allows the application of nonsmooth Newton methods for the solution of the nonlinear second-order cone program. Conditions for the local quadratic convergence of these nonsmooth Newton methods are investigated. Related conditions are also given for the special case of a linear second-order cone program. An interesting and important feature of these conditions is that they do not require strict complementarity of the solution. Some numerical results are included in order to illustrate the theoretical considerations.

Key Words: Linear second-order cone program, nonlinear second-order cone program, semismooth function, nonsmooth Newton method, quadratic convergence

1 Introduction

We consider both the linear second-order cone program (linear SOCP)

$$\min c^T x \quad \text{s.t.} \quad Ax = b, \quad x \in \mathcal{K}, \quad (1)$$

¹This work was supported in part by the international doctorate program “Identification, Optimization and Control with Applications in Modern Technologies” of the Elite Network of Bavaria, Germany, and by the Scientific Research Grant-in-Aid from Japan Society for the Promotion of Science.

²Institute of Mathematics, University of Würzburg, Am Hubland, 97074 Würzburg, Germany, e-mail: kanzow@mathematik.uni-wuerzburg.de

³Institute of Mathematics, University of Würzburg, Am Hubland, 97074 Würzburg, Germany, e-mail: ferenczi@mathematik.uni-wuerzburg.de

⁴Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan, e-mail: fuku@amp.i.kyoto-u.ac.jp

and the nonlinear second-order cone program (nonlinear SOCP)

$$\min f(x) \quad \text{s.t.} \quad Ax = b, \quad x \in \mathcal{K}, \quad (2)$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a twice continuously differentiable function, $A \in \mathbb{R}^{m \times n}$ is a given matrix, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ are given vectors, and

$$\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_r$$

is a Cartesian product of second-order cones $\mathcal{K}_i \subseteq \mathbb{R}^{n_i}$, $n_1 + \cdots + n_r = n$. Recall that the second-order cone (or ice-cream cone or Lorentz cone) of dimension n_i is defined by

$$\mathcal{K}_i := \{x_i = (x_{i0}, \bar{x}_i) \in \mathbb{R} \times \mathbb{R}^{n_i-1} \mid x_{i0} \geq \|\bar{x}_i\|\},$$

where $\|\cdot\|$ denotes the Euclidean norm. Observe the special notation that is used in the definition of \mathcal{K}_i and that will be applied throughout this manuscript: For a given vector $z \in \mathbb{R}^\ell$ for some $\ell \geq 1$, we write $z = (z_0, \bar{z})$, where z_0 is the first component of the vector z , and \bar{z} consists of the remaining $\ell - 1$ components of z .

The linear SOCP has been investigated in many previous works, and we refer the interested reader to the two survey papers [18, 1] and the books [2, 4] for many important applications and theoretical properties. Software for the solution of linear SOCPs is also available, see, for example, [17, 28, 24, 27]. In many cases, the linear SOCP may be viewed as a special case of a (linear) semidefinite program (see [1] for a suitable reformulation). However, we feel that the SOCP should be treated directly since the reformulation of a second-order cone constraint as a semidefinite constraint increases the dimension of the problem significantly and, therefore, decreases the efficiency of any solver. In fact, many solvers for semidefinite programs (see, for example, the list given on Helmberg's homepage [14]) are able to deal with second-order cone constraints separately.

The treatment of the nonlinear SOCP is much more recent, and, in the moment, the number of publications is rather limited, see [3, 5, 6, 7, 8, 12, 13, 16, 25, 26, 29]. These papers deal with different topics; some of them investigate different kinds of solution methods (interior-point methods, smoothing methods, SQP-type methods, or methods based on unconstrained optimization), while some of them consider certain theoretical properties or suitable reformulations of the SOCP.

The method of choice for the solution of (at least) the linear SOCP is currently an interior-point method. However, some recent preliminary tests indicate that the class of smoothing or semismooth methods is sometimes superior to the class of interior-point methods, especially for nonlinear problems, see [8, 13, 26]. On the other hand, the theoretical properties of interior-point methods are much better understood than those of the smoothing and semismooth methods.

The aim of this paper is to provide some results which help to understand the theoretical properties of semismooth methods being applied to both linear and nonlinear SOCPs. The investigation here is of local nature, and we provide sufficient conditions for those methods

to be locally quadratically convergent. An interesting and important feature of those sufficient conditions is that they do not require strict complementarity of the solution. This is an advantage compared to interior-point methods where singular Jacobians occur if strict complementarity is not satisfied. Similar results were recently obtained in [15] (see also [11]) for linear semidefinite programs. In principle, these results can also be applied to linear SOCPs, but this requires a reformulation of the SOCP as a semidefinite program which, as mentioned above, is not necessarily the best approach, and therefore motivates a direct treatment of SOCPs.

The paper is organized as follows: Section 2 states a number of preliminary results for the projection mapping onto a second-order cone, which will later be used in order to reformulate the optimality conditions of the SOCP as a system of equations. Section 3 then investigates conditions that ensure the nonsingularity of the generalized Jacobian of this system, so that the nonsmooth Newton method is locally quadratically convergent. Some preliminary numerical examples illustrating the local convergence properties of the method are given in Section 4. We close with some final remarks in Section 5.

Most of our notation is standard. For a differentiable mapping $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$, we denote by $G'(z) \in \mathbb{R}^{m \times n}$ the Jacobian of G at z . If G is locally Lipschitz continuous, the set

$$\partial_B G(z) := \{H \in \mathbb{R}^{m \times n} \mid \exists \{z^k\} \subseteq D_G : z^k \rightarrow z, G'(z^k) \rightarrow H\}$$

is nonempty and called the B-subdifferential of G at z , where $D_G \subseteq \mathbb{R}^n$ denotes the set of points at which G is differentiable. The convex hull $\partial G(z) := \text{conv} \partial_B G(z)$ is the generalized Jacobian of Clarke [9]. We assume that the reader is familiar with the concepts of (strongly) semismooth functions, and refer to [23, 22, 20, 10] for details. The identity matrix of order n is denoted by I_n .

2 Projection Mapping onto Second-Order Cone

Throughout this section, let \mathcal{K} be the single second-order cone

$$\mathcal{K} := \{z = (z_0, \bar{z}) \in \mathbb{R} \times \mathbb{R}^{n-1} \mid z_0 \geq \|\bar{z}\|\}.$$

In the subsequent sections, \mathcal{K} will be the Cartesian product of second-order cones. The results of this section will later be applied componentwise to each of the second-order cones \mathcal{K}_i in the Cartesian product.

Recall that the second-order cone \mathcal{K} is self-dual, i.e. $\mathcal{K}^* = \mathcal{K}$, where $\mathcal{K}^* := \{d \in \mathbb{R} \times \mathbb{R}^{n-1} \mid z^T d \geq 0 \forall z \in \mathcal{K}\}$ denotes the dual cone of \mathcal{K} , cf. [1, Lemma 1]. Hence the following result holds, see, e.g., [12, Proposition 4.1].

Lemma 2.1 *The following equivalence holds:*

$$x \in \mathcal{K}, y \in \mathcal{K}, x^T y = 0 \iff x - P_{\mathcal{K}}(x - y) = 0,$$

where $P_{\mathcal{K}}(z)$ denotes the (Euclidean) projection of a vector z on \mathcal{K} .

An explicit representation of the projection $P_{\mathcal{K}}(z)$ is given in the following result, see [12, Proposition 3.3].

Lemma 2.2 *For any given $z = (z_0, \bar{z}) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we have*

$$P_{\mathcal{K}}(z) = \max\{0, \eta_1\}u^{(1)} + \max\{0, \eta_2\}u^{(2)},$$

where η_1, η_2 are the spectral values and $u^{(1)}, u^{(2)}$ are the spectral vectors of z , respectively, given by

$$\begin{aligned} \eta_1 &:= z_0 - \|\bar{z}\|, & \eta_2 &:= z_0 + \|\bar{z}\|, \\ u^{(1)} &:= \begin{cases} \frac{1}{2} \begin{pmatrix} 1 \\ -\frac{\bar{z}}{\|\bar{z}\|} \end{pmatrix} & \text{if } \bar{z} \neq 0, \\ \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w} \end{pmatrix} & \text{if } \bar{z} = 0, \end{cases} & u^{(2)} &:= \begin{cases} \frac{1}{2} \begin{pmatrix} 1 \\ \frac{\bar{z}}{\|\bar{z}\|} \end{pmatrix} & \text{if } \bar{z} \neq 0, \\ \frac{1}{2} \begin{pmatrix} 1 \\ \bar{w} \end{pmatrix} & \text{if } \bar{z} = 0, \end{cases} \end{aligned}$$

where \bar{w} is any vector in \mathbb{R}^{n-1} with $\|\bar{w}\| = 1$.

It is well-known that the projection mapping onto an arbitrary closed convex set is non-expansive and hence is Lipschitz continuous. When the set is the second-order cone \mathcal{K} , a stronger smoothness property can be shown, see [5, Proposition 4.3], [7, Proposition 7], or [13, Proposition 4.5].

Lemma 2.3 *The projection mapping $P_{\mathcal{K}}$ is strongly semismooth.*

We next characterize the points at which the projection mapping $P_{\mathcal{K}}$ is differentiable.

Lemma 2.4 *The projection mapping $P_{\mathcal{K}}$ is differentiable at a point $z = (z_0, \bar{z}) \in \mathbb{R} \times \mathbb{R}^{n-1}$ if and only if $z_0 \neq \pm\|\bar{z}\|$ holds. In fact, the projection mapping is continuously differentiable at every z such that $z_0 \neq \pm\|\bar{z}\|$.*

Proof. The statement can be derived directly from the representation of $P_{\mathcal{K}}(z)$ given in Lemma 2.2. Alternatively, it can be derived as a special case of more general results stated in [7], see, in particular, Propositions 4 and 5 in that reference. \square

We next calculate the Jacobian of the projection mapping $P_{\mathcal{K}}$ at a point where it is differentiable. The proof is not difficult and therefore omitted.

Lemma 2.5 *The Jacobian of $P_{\mathcal{K}}$ at a point $z = (z_0, \bar{z}) \in \mathbb{R} \times \mathbb{R}^{n-1}$ with $z_0 \neq \pm\|\bar{z}\|$ is given by*

$$P'_{\mathcal{K}}(z) = \begin{cases} 0, & \text{if } z_0 < -\|\bar{z}\|, \\ I_n, & \text{if } z_0 > +\|\bar{z}\|, \\ \frac{1}{2} \begin{pmatrix} 1 & \bar{w}^T \\ \bar{w} & H \end{pmatrix}, & \text{if } -\|\bar{z}\| < z_0 < +\|\bar{z}\|, \end{cases}$$

where

$$\bar{w} := \frac{\bar{z}}{\|\bar{z}\|}, \quad H := \left(1 + \frac{z_0}{\|\bar{z}\|}\right)I_{n-1} - \frac{z_0}{\|\bar{z}\|}\bar{w}\bar{w}^T.$$

(Note that the denominator is automatically nonzero in this case.)

Based on the above results, we give in the next lemma an expression for the elements of the B-subdifferential $\partial_B P_{\mathcal{K}}(z)$ at an arbitrary point z . A similar representation of the elements of the Clarke generalized Jacobian $\partial P_{\mathcal{K}}(z)$ is given in [13, Proposition 4.8] (see also [19, Lemma 14] and [7, Lemma 4]), and hence we omit the proof of the lemma. Note that we deal with the smaller set $\partial_B P_{\mathcal{K}}(z)$ here, since this will simplify our subsequent analysis to give sufficient conditions for the nonsingularity of all elements in $\partial_B P_{\mathcal{K}}(z)$. In fact, the nonsingularity of all elements of the B-subdifferential usually holds under weaker assumptions than the nonsingularity of all elements of the corresponding Clarke generalized Jacobian.

Lemma 2.6 *Given a general point $z = (z_0, \bar{z}) \in \mathbb{R} \times \mathbb{R}^{n-1}$, each element $V \in \partial_B P_{\mathcal{K}}(z)$ has the following representation:*

- (a) *If $z_0 \neq \pm\|\bar{z}\|$, then $P_{\mathcal{K}}$ is continuously differentiable at z and $V = P'_{\mathcal{K}}(z)$ with the Jacobian $P'_{\mathcal{K}}(z)$ given in Lemma 2.5.*
- (b) *If $\bar{z} \neq 0$ and $z_0 = +\|\bar{z}\|$, then*

$$V \in \left\{ I_n, \frac{1}{2} \begin{pmatrix} 1 & \bar{w}^T \\ \bar{w} & H \end{pmatrix} \right\},$$

where $\bar{w} := \frac{\bar{z}}{\|\bar{z}\|}$ and $H := 2I_{n-1} - \bar{w}\bar{w}^T$.

- (c) *If $\bar{z} \neq 0$ and $z_0 = -\|\bar{z}\|$, then*

$$V \in \left\{ 0, \frac{1}{2} \begin{pmatrix} 1 & \bar{w}^T \\ \bar{w} & H \end{pmatrix} \right\},$$

where $\bar{w} := \frac{\bar{z}}{\|\bar{z}\|}$ and $H := \bar{w}\bar{w}^T$.

- (d) *If $\bar{z} = 0$ and $z_0 = 0$, then either $V = 0$ or $V = I_n$ or V belongs to the set*

$$\left\{ \frac{1}{2} \begin{pmatrix} 1 & \bar{w}^T \\ \bar{w} & H \end{pmatrix} \mid H = (1 + \rho)I_{n-1} - \rho\bar{w}\bar{w}^T \text{ for some } |\rho| \leq 1 \text{ and } \|\bar{w}\| = 1 \right\}.$$

We can summarize Lemma 2.6 as follows: Any element $V \in \partial_B P_{\mathcal{K}}(z)$ is equal to

$$V = 0 \quad \text{or} \quad V = I_n \quad \text{or} \quad V = \frac{1}{2} \begin{pmatrix} 1 & \bar{w}^T \\ \bar{w} & H \end{pmatrix} \tag{3}$$

for some vector $\bar{w} \in \mathbb{R}^{n-1}$ with $\|\bar{w}\| = 1$ and some matrix $H \in \mathbb{R}^{(n-1) \times (n-1)}$ of the form $H = (1 + \rho)I_{n-1} - \rho\bar{w}\bar{w}^T$ with some scalar $\rho \in \mathbb{R}$ satisfying $|\rho| \leq 1$. Specifically, in cases (a)–(c), we have $\bar{w} = \bar{z}/\|\bar{z}\|$, whereas in case (d), \bar{w} can be any vector of length one. Moreover, we have $\rho = z_0/\|\bar{z}\|$ in case (a), $\rho = 1$ in case (b), $\rho = -1$ in case (c), whereas there is no further specification of ρ in case (d) (here the two simple cases $V = 0$ and $V = I_n$ are always excluded).

The eigenvalues and eigenvectors of any matrix $V \in \partial_B P_{\mathcal{K}}(z)$ can be given explicitly, as shown in the following result.

Lemma 2.7 *Let $z = (z_0, \bar{z}) \in \mathbb{R} \times \mathbb{R}^{n-1}$ and $V \in \partial_B P_{\mathcal{K}}(z)$. Assume that $V \notin \{0, I_n\}$ so that V has the third representation in (3) with $H = (1 + \rho)I_{n-1} - \rho\bar{w}\bar{w}^T$ for some scalar $\rho \in [-1, +1]$ and some vector $\bar{w} \in \mathbb{R}^{n-1}$ satisfying $\|\bar{w}\| = 1$. Then V has the two single eigenvalues $\eta = 0$ and $\eta = 1$ as well as the eigenvalue $\eta = \frac{1}{2}(1 + \rho)$ with multiplicity $n - 2$ (unless $\rho = \pm 1$, where the multiplicities change in an obvious way). In particular, when $P'_{\mathcal{K}}(z)$ exists, i.e., in case (a) of Lemma 2.6, the multiple eigenvalue is given by $\eta = \frac{1}{2}(1 + \frac{z_0}{\|\bar{z}\|})$. Moreover, the eigenvectors of V are given by*

$$\begin{pmatrix} -1 \\ \bar{w} \end{pmatrix}, \begin{pmatrix} 1 \\ \bar{w} \end{pmatrix}, \text{ and } \begin{pmatrix} 0 \\ \bar{v}_j \end{pmatrix}, j = 1, \dots, n - 2,$$

where $\bar{v}_1, \dots, \bar{v}_{n-2}$ are arbitrary vectors that span the linear subspace $\{\bar{v} \in \mathbb{R}^{n-1} \mid \bar{v}^T \bar{w} = 0\}$.

Proof. By assumption, we have

$$V = \frac{1}{2} \begin{pmatrix} 1 & \bar{w}^T \\ \bar{w} & H \end{pmatrix} \quad \text{with} \quad H = (1 + \rho)I_{n-1} - \rho\bar{w}\bar{w}^T$$

for some $\rho \in [-1, +1]$ and some vector \bar{w} satisfying $\|\bar{w}\| = 1$. Now take an arbitrary vector $\bar{v} \in \mathbb{R}^{n-1}$ orthogonal to \bar{w} , and let $u = (0, \bar{v}^T)^T$. Then an elementary calculation shows that $Vu = \eta u$ holds for $\eta = \frac{1}{2}(1 + \rho)$. Hence this η is an eigenvalue of V with multiplicity $n - 2$ since we can choose $n - 2$ linearly independent vectors $\bar{v} \in \mathbb{R}^{n-1}$ such that $\bar{v}^T \bar{w} = 0$. On the other hand, if $\eta = 0$, it is easy to see that $Vu = \eta u$ holds with $u = (-1, \bar{w}^T)^T$, whereas for $\eta = 1$ we have $Vu = \eta u$ by taking $u = (1, \bar{w}^T)^T$. The multiple eigenvalue of $P'_{\mathcal{K}}(z)$ (in the differentiable case) can be checked directly from the formula given in Lemma 2.5. This completes the proof. \square

Note that Lemma 2.7 particularly implies $\eta \in [0, 1]$ for all eigenvalues η of V . This observation can alternatively be derived from the fact that $P_{\mathcal{K}}$ is a projection mapping, without referring to the explicit representation of V as given in Lemma 2.6.

We close this section by pointing out an interesting relation between the matrix $V \in \partial_B P_{\mathcal{K}}(z)$ and the so-called arrow matrix

$$\mathbf{Arw}(z) := \begin{pmatrix} z_0 & \bar{z}^T \\ \bar{z} & z_0 I_{n-1} \end{pmatrix} \in \mathbb{R}^{n \times n}$$

associated with $z = (z_0, \bar{z}) \in \mathbb{R} \times \mathbb{R}^{n-1}$, which frequently occurs in the context of interior-point methods and in the analysis of SOCPs, see, e.g., [1]. To this end, consider the case where $P_{\mathcal{K}}$ is differentiable at z , excluding the two trivial cases where $P'_{\mathcal{K}}(z) = 0$ or $P'_{\mathcal{K}}(z) = I_n$, cf. Lemma 2.5. Then by Lemma 2.7, the eigenvalues of the matrix $V = P'_{\mathcal{K}}(z)$ are given by $\eta = 0$, $\eta = 1$, and $\eta = \frac{1}{2}(1 + \frac{z_0}{\|\bar{z}\|})$ with multiplicity $n - 2$, and the corresponding eigenvectors are given by

$$\begin{pmatrix} -1 \\ \frac{\bar{z}}{\|\bar{z}\|} \end{pmatrix}, \begin{pmatrix} 1 \\ \frac{\bar{z}}{\|\bar{z}\|} \end{pmatrix}, \text{ and } \begin{pmatrix} 0 \\ \bar{v}_j \end{pmatrix}, \quad j = 1, \dots, n - 2, \quad (4)$$

where $\bar{v}_1, \dots, \bar{v}_{n-2}$ comprise an orthogonal basis of the linear subspace $\{\bar{v} \in \mathbb{R}^{n-1} \mid \bar{v}^T \bar{z} = 0\}$. However, an elementary calculation shows that these are also the eigenvectors of the arrow matrix $\mathbf{Arw}(z)$, with corresponding single eigenvalues $\hat{\eta}_1 = z_0 - \|\bar{z}\|$, $\hat{\eta}_2 = z_0 + \|\bar{z}\|$ and the multiple eigenvalues $\hat{\eta}_i = z_0$, $i = 3, \dots, n$. Therefore, although the eigenvalues of $V = P'_{\mathcal{K}}(z)$ and $\mathbf{Arw}(z)$ are different, both matrices have the same set of eigenvectors.

3 Second-Order Cone Programs

In this section, we consider the SOCP

$$\min f(x) \quad \text{s.t.} \quad Ax = b, \quad x \in \mathcal{K},$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a twice continuously differentiable function, $A \in \mathbb{R}^{m \times n}$ is a given matrix, $b \in \mathbb{R}^m$ is a given vector, and $\mathcal{K} = \mathcal{K}_1 \times \dots \times \mathcal{K}_r$ is the Cartesian product of second-order cones $\mathcal{K}_i \subseteq \mathbb{R}^{n_i}$ with $n_1 + \dots + n_r = n$. The vector x and the matrix A are partitioned as $x = (x_1, \dots, x_r)$ and $A = (A_1, \dots, A_r)$, respectively, where $x_i = (x_{i0}, \bar{x}_i) \in \mathbb{R} \times \mathbb{R}^{n_i-1}$ and $A \in \mathbb{R}^{m \times n_i}$, $i = 1, \dots, r$. Although the objective function f is supposed to be nonlinear in general, we will particularly discuss the linear case as well.

Under certain conditions like convexity of f and a Slater-type constraint qualification [4], solving the SOCP is equivalent to solving the corresponding KKT conditions, which can be written as follows:

$$\begin{aligned} \nabla f(x) - A^T \mu - \lambda &= 0, \\ Ax &= b, \\ x_i \in \mathcal{K}_i, \quad \lambda_i \in \mathcal{K}_i, \quad x_i^T \lambda_i &= 0, \quad i = 1, \dots, r. \end{aligned}$$

Using Lemma 2.1, it follows that these KKT conditions are equivalent to the system of equations $M(z) = 0$, where $M : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ is defined by

$$M(z) := M(x, \mu, \lambda) := \begin{pmatrix} \nabla f(x) - A^T \mu - \lambda \\ Ax - b \\ x_1 - P_{\mathcal{K}_1}(x_1 - \lambda_1) \\ \vdots \\ x_r - P_{\mathcal{K}_r}(x_r - \lambda_r) \end{pmatrix}. \quad (5)$$

Then we can apply the nonsmooth Newton method [22, 23, 20]

$$z^{k+1} := z^k - W_k^{-1}M(z^k), \quad W_k \in \partial_B M(z^k), \quad k = 0, 1, 2, \dots, \quad (6)$$

to the system of equations $M(z) = 0$ in order to solve the SOCP or, at least, the corresponding KKT conditions. Our aim is to show fast local convergence of this iterative method. In view of the results in [23, 22], we have to guarantee that, on the one hand, the mapping M , though not differentiable everywhere, is still sufficiently ‘smooth’, and, on the other hand, it satisfies a local nonsingularity condition under suitable assumptions.

The required smoothness property of M is summarized in the following result.

Theorem 3.1 *The mapping M defined by (5) is semismooth. Moreover, if the Hessian $\nabla^2 f$ is locally Lipschitz continuous, then the mapping M is strongly semismooth.*

Proof. Note that a continuously differentiable mapping is semismooth. Moreover, if the Jacobian of a differentiable mapping is locally Lipschitz continuous, then this mapping is strongly semismooth. Now Lemma 2.3 and the fact that a given mapping is (strongly) semismooth if and only if all component functions are (strongly) semismooth yield the desired result. \square

Our next step is to provide suitable conditions which guarantee the nonsingularity of all elements of the B-subdifferential of M at a KKT point. This requires some more work, and we begin with the following general result.

Proposition 3.2 *Let $H \in \mathbb{R}^{n \times n}$ be symmetric, and $A \in \mathbb{R}^{m \times n}$. Let $V^a, V^b \in \mathbb{R}^{n \times n}$ be two symmetric positive semidefinite matrices such that their sum $V^a + V^b$ is positive definite and V^a and V^b have a common basis of eigenvectors, so that there exist an orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ and diagonal matrices $D^a = \text{diag}(a_1, \dots, a_n)$ and $D^b = \text{diag}(b_1, \dots, b_n)$ satisfying $V^a = QD^aQ^T, V^b = QD^bQ^T$ as well as $a_j \geq 0, b_j \geq 0$ and $a_j + b_j > 0$ for all $j = 1, \dots, n$. Let the index set $\{1, \dots, n\}$ be partitioned as $\{1, \dots, n\} = \alpha \cup \beta \cup \gamma$, where*

$$\begin{aligned} \alpha &:= \{j \mid a_j > 0, b_j = 0\}, \\ \beta &:= \{j \mid a_j > 0, b_j > 0\}, \\ \gamma &:= \{j \mid a_j = 0, b_j > 0\}, \end{aligned}$$

and let Q_α, Q_β , and Q_γ denote the submatrices of Q consisting of the columns from Q corresponding to the index sets α, β , and γ , respectively. Let us also partition the diagonal matrices D^a and D^b into $D^a = \text{diag}(D_\alpha^a, D_\beta^a, D_\gamma^a)$ and $D^b = \text{diag}(D_\alpha^b, D_\beta^b, D_\gamma^b)$, respectively, and let

$$D_\beta := (D_\beta^b)^{-1}D_\beta^a. \quad (7)$$

Assume that the following two conditions hold:

- (a) The matrix $(AQ_\beta, AQ_\gamma) \in \mathbb{R}^{m \times (|\beta| + |\gamma|)}$ has full row rank.

(b) The matrix

$$\begin{pmatrix} Q_\beta^T H Q_\beta + D_\beta & Q_\beta^T H Q_\gamma \\ Q_\gamma^T H Q_\beta & Q_\gamma^T H Q_\gamma \end{pmatrix} \in \mathbb{R}^{(|\beta|+|\gamma|) \times (|\beta|+|\gamma|)}$$

is positive definite on the subspace $\mathcal{V} := \left\{ \begin{pmatrix} d_\beta \\ d_\gamma \end{pmatrix} \in \mathbb{R}^{|\beta|+|\gamma|} \mid (A Q_\beta, A Q_\gamma) \begin{pmatrix} d_\beta \\ d_\gamma \end{pmatrix} = 0 \right\}$.

Then the matrix

$$W := \begin{pmatrix} H & -A^T & -I_n \\ A & 0 & 0 \\ V^a & 0 & V^b \end{pmatrix}$$

is nonsingular. In particular, when $H = 0$, the matrix W is nonsingular if the following condition holds together with (a):

(c) The matrix $A Q_\gamma$ has full column rank.

Proof. An elementary calculation shows that the matrix W is nonsingular if and only if the matrix

$$W' := \begin{pmatrix} Q^T H Q & -(A Q)^T & -I_n \\ A Q & 0 & 0 \\ D^a & 0 & D^b \end{pmatrix}$$

is nonsingular. Taking into account the definition of the three index sets α, β, γ , we obtain

$$\begin{aligned} D^a &= \text{diag}(D_\alpha^a, D_\beta^a, D_\gamma^a) = \text{diag}(D_\alpha^a, D_\beta^a, 0), \\ D^b &= \text{diag}(D_\alpha^b, D_\beta^b, D_\gamma^b) = \text{diag}(0, D_\beta^b, D_\gamma^b). \end{aligned}$$

Using this structure and premultiplying the matrix W' by

$$\begin{pmatrix} I_n & & \\ & I_m & \\ & & D \end{pmatrix} \quad \text{with} \quad D := \text{diag}((D_\alpha^a)^{-1}, (D_\beta^a)^{-1}, I_{|\gamma|}),$$

we see that the matrix W' is nonsingular if and only if

$$W'' := \begin{pmatrix} Q^T H Q & -(A Q)^T & -I_n \\ A Q & 0 & 0 \\ \tilde{D}^a & 0 & \tilde{D}^b \end{pmatrix}$$

is nonsingular, where \tilde{D}^a and \tilde{D}^b are diagonal matrices given by

$$\tilde{D}^a := \text{diag}(I_{|\alpha|}, I_{|\beta|}, 0) \quad \text{and} \quad \tilde{D}^b := \text{diag}(0, D_\beta^{-1}, D_\gamma^b).$$

Note that the matrix D_β defined by (7) is a positive definite diagonal matrix. It then follows that the matrix W'' is a block upper triangular matrix with its lower right block

D_γ^b being a nonsingular diagonal matrix. Therefore the matrix W'' is nonsingular if and only if its upper left block

$$\tilde{W} := \begin{pmatrix} Q^T H Q & -(A Q)^T & -I_\alpha & -I_\beta \\ A Q & 0 & 0 & 0 \\ I_\alpha^T & 0 & 0 & 0 \\ I_\beta^T & 0 & 0 & D_\beta^{-1} \end{pmatrix} \quad (8)$$

is nonsingular, where I_α, I_β denote the matrices in $\mathbb{R}^{n \times |\alpha|}, \mathbb{R}^{n \times |\beta|}$ consisting of all columns of the identity matrix corresponding to the index sets $i \in \alpha, i \in \beta$, respectively. (Note the difference between I_α, I_β and the square matrices $I_{|\alpha|}, I_{|\beta|}$.) In other words, the matrix W is nonsingular if and only if \tilde{W} is nonsingular.

In order to show the nonsingularity of \tilde{W} , let $\tilde{W}y = 0$ for a suitably partitioned vector $y = (d, p, q_\alpha, q_\beta)^T \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{|\alpha|} \times \mathbb{R}^{|\beta|}$. We will see that $y = 0$ under assumptions (a) and (b). Using (8), we may write $\tilde{W}y = 0$ as

$$Q^T H Q d - Q^T A^T p - \begin{pmatrix} q_\alpha \\ q_\beta \\ 0 \end{pmatrix} = 0, \quad (9)$$

$$A Q d = 0, \quad (10)$$

$$d_\alpha = 0, \quad (11)$$

$$d_\beta + D_\beta^{-1} q_\beta = 0. \quad (12)$$

Premultiplying (9) by d^T and taking into account (10) and (11), we obtain

$$\begin{pmatrix} d_\beta \\ d_\gamma \end{pmatrix}^T (Q_\beta, Q_\gamma)^T H (Q_\beta, Q_\gamma) \begin{pmatrix} d_\beta \\ d_\gamma \end{pmatrix} - d_\beta^T q_\beta = 0,$$

which along with (12) yields

$$\begin{pmatrix} d_\beta \\ d_\gamma \end{pmatrix}^T \begin{pmatrix} Q_\beta^T H Q_\beta + D_\beta & Q_\beta^T H Q_\gamma \\ Q_\gamma^T H Q_\beta & Q_\gamma^T H Q_\gamma \end{pmatrix} \begin{pmatrix} d_\beta \\ d_\gamma \end{pmatrix} = 0.$$

On the other hand, from (10) and (11), we have

$$(A Q_\beta, A Q_\gamma) \begin{pmatrix} d_\beta \\ d_\gamma \end{pmatrix} = 0. \quad (13)$$

Then, by assumption (b), we obtain $d_\beta = 0$ and $d_\gamma = 0$, which together with (12) implies $q_\beta = 0$. Now it follows from (9) that

$$-Q_\alpha^T A^T p - q_\alpha = 0 \quad (14)$$

and

$$-\begin{pmatrix} Q_\beta^T A^T \\ Q_\gamma^T A^T \end{pmatrix} p = 0. \quad (15)$$

By assumption (a), (15) yields $p = 0$, which in turn implies $q_\alpha = 0$ from (14). Consequently, we have $y = 0$. This shows \tilde{W} , and hence W , is nonsingular.

When $H = 0$, we obtain from (9)–(12)

$$d_\beta^T D_\beta d_\beta = -d_\beta^T q_\beta = 0.$$

Since D_β is positive definite, this implies $d_\beta = 0$. Then by assumption (c), it follows from (13) that $d_\gamma = 0$. The rest of the proof goes in the same manner as above. \square

The two central assumptions (a) and (b) of Proposition 3.2 can also be formulated in a different way: Using some elementary calculations, it is not difficult to see that assumption (a) is equivalent to

(a') The matrix $(Q^T A^T, I_\alpha) \in \mathbb{R}^{n \times (m+|\alpha|)}$ has full column rank;

whereas assumption (b) is equivalent to

(b') $H + Q_\beta D_\beta Q_\beta^T$ is positive definite on the subspace $\mathcal{S} := \{v \in \mathbb{R}^n \mid Av = 0, Q_\alpha^T v = 0\}$.

Now let us go back to the mapping M defined by (5). In order to apply Proposition 3.2 to the (generalized) Jacobian of the mapping M at a KKT point, we first introduce some more notation:

$$\begin{aligned} \text{int}\mathcal{K}_i &:= \{x_i \mid x_{i0} > \|\bar{x}_i\|\} && \text{denotes the interior of } \mathcal{K}_i, \\ \text{bd}\mathcal{K}_i &:= \{x_i \mid x_{i0} = \|\bar{x}_i\|\} && \text{denotes the boundary of } \mathcal{K}_i, \text{ and} \\ \text{bd}^+\mathcal{K}_i &:= \text{bd}\mathcal{K}_i \setminus \{0\} && \text{is the boundary of } \mathcal{K}_i \text{ excluding the origin.} \end{aligned}$$

We also call a KKT point $z^* = (x^*, \mu^*, \lambda^*)$ of the SOCP *strictly complementary* if $x_i^* + \lambda_i^* \in \text{int}\mathcal{K}_i$ holds for all block components $i = 1, \dots, r$. This notation enables us to restate the following result from [1].

Lemma 3.3 *Let $z^* = (x^*, \mu^*, \lambda^*)$ be a KKT point of the SOCP. Then precisely one of the following six cases holds for each block pair (x_i^*, λ_i^*) , $i = 1, \dots, r$:*

x_i^*	λ_i^*	<i>SC</i>
$x_i^* \in \text{int}\mathcal{K}_i$	$\lambda_i^* = 0$	<i>yes</i>
$x_i^* = 0$	$\lambda_i^* \in \text{int}\mathcal{K}_i$	<i>yes</i>
$x_i^* \in \text{bd}^+\mathcal{K}_i$	$\lambda_i^* \in \text{bd}^+\mathcal{K}_i$	<i>yes</i>
$x_i^* \in \text{bd}^+\mathcal{K}_i$	$\lambda_i^* = 0$	<i>no</i>
$x_i^* = 0$	$\lambda_i^* \in \text{bd}^+\mathcal{K}_i$	<i>no</i>
$x_i^* = 0$	$\lambda_i^* = 0$	<i>no</i>

The last column in the table indicates whether or not strict complementarity (SC) holds.

We also need the following simple result which, in particular, shows that the projection mapping $P_{\mathcal{K}_i}$ involved in the definition of the mapping M is continuously differentiable at $s_i := x_i^* - \lambda_i^*$ for any block component i satisfying strict complementarity.

Lemma 3.4 *Let $z^* = (x^*, \mu^*, \lambda^*)$ be a KKT point of the SOCP. Then the following statements hold for each block pair (x_i^*, λ_i^*) :*

- (a) *If $x_i^* \in \text{int}\mathcal{K}_i$ and $\lambda_i^* = 0$, then $P_{\mathcal{K}_i}$ is continuously differentiable at $s_i := x_i^* - \lambda_i^*$ with $P'_{\mathcal{K}_i}(s_i) = I_{n_i}$.*
- (b) *If $x_i^* = 0$ and $\lambda_i^* \in \text{int}\mathcal{K}_i$, then $P_{\mathcal{K}_i}$ is continuously differentiable at $s_i := x_i^* - \lambda_i^*$ with $P'_{\mathcal{K}_i}(s_i) = 0$.*
- (c) *If $x_i^* \in \text{bd}^+\mathcal{K}_i$ and $\lambda_i^* \in \text{bd}^+\mathcal{K}_i$, then $P_{\mathcal{K}_i}$ is continuously differentiable at $s_i := x_i^* - \lambda_i^*$ with $P'_{\mathcal{K}_i}(s_i) = \frac{1}{2} \begin{pmatrix} 1 & \bar{w}_i^T \\ \bar{w}_i & H_i \end{pmatrix}$, where $\bar{w}_i = \frac{\bar{s}_i}{\|\bar{s}_i\|}$ and $H_i = (1 + \frac{s_{i0}}{\|\bar{s}_i\|})I_{n_i-1} - \frac{s_{i0}}{\|\bar{s}_i\|}\bar{w}_i\bar{w}_i^T$.*

Proof. Parts (a) and (b) immediately follow from Lemma 2.5. To prove part (c), write $x_i^* = (x_{i0}^*, \bar{x}_i^*)$, $\lambda_i^* = (\lambda_{i0}^*, \bar{\lambda}_i^*)$, and $s_i = (s_{i0}, \bar{s}_i) := x_i^* - \lambda_i^* = (x_{i0}^* - \lambda_{i0}^*, \bar{x}_i^* - \bar{\lambda}_i^*)$. Since $x_i^* \neq 0$ and $\lambda_i^* \neq 0$, we see from [1, Lemma 15] that there is a constant $\rho > 0$ such that $\lambda_{i0}^* = \rho x_{i0}^*$ and $\bar{\lambda}_i^* = -\rho \bar{x}_i^*$, implying $s_{i0} = (1 - \rho)x_{i0}^*$ and $\|\bar{s}_i\| = (1 + \rho)\|\bar{x}_i^*\|$. Since $x_{i0}^* = \|\bar{x}_i^*\| \neq 0$ by assumption, we have $s_{i0} = \frac{1-\rho}{1+\rho}\|\bar{s}_i\|$. Hence we obtain $s_{i0} = \|\bar{s}_i\| - \frac{2\rho}{1+\rho}\|\bar{s}_i\| < \|\bar{s}_i\|$ and $s_{i0} = \frac{2}{1+\rho}\|\bar{s}_i\| - \|\bar{s}_i\| > -\|\bar{s}_i\|$. The desired result then follows from Lemma 2.5. \square

We are now almost in a position to apply Proposition 3.2 to the Jacobian of the mapping M at a KKT point $z^* = (x^*, \mu^*, \lambda^*)$ provided that this KKT point satisfies strict complementarity. This strict complementarity assumption will be removed later, but for the moment it is quite convenient to assume this condition. For example, it then follows from Lemma 3.3 that the three index sets

$$\begin{aligned} J_I &:= \{i \mid x_i^* \in \text{int}\mathcal{K}_i, \lambda_i^* = 0\}, \\ J_B &:= \{i \mid x_i^* \in \text{bd}^+\mathcal{K}_i, \lambda_i^* \in \text{bd}^+\mathcal{K}_i\}, \\ J_0 &:= \{i \mid x_i^* = 0, \lambda_i^* \in \text{int}\mathcal{K}_i\} \end{aligned} \tag{16}$$

form a partition of the block indices $i = 1, \dots, r$. Here, the subscripts I, B and 0 indicate whether the block component x_i^* belongs to the interior of the cone \mathcal{K}_i , or x_i^* belongs to the boundary of \mathcal{K}_i (excluding the zero vector), or x_i^* is the zero vector.

Let $V_i := P'_{\mathcal{K}_i}(x_i^* - \lambda_i^*)$. Then Lemma 3.4 implies that

$$V_i = I_{n_i} \quad \forall i \in J_I \quad \text{and} \quad V_i = 0 \quad \forall i \in J_0. \tag{17}$$

To get a similar representation for indices $i \in J_B$, we need the spectral decompositions $V_i = Q_i D_i Q_i^T$ of the matrices V_i . Since strict complementarity holds, it follows from

Lemmas 2.7 and 3.4 that each V_i has precisely one eigenvalue equal to zero and precisely one eigenvalue equal to one, whereas all other eigenvalues are strictly between zero and one. Without loss of generality, we can therefore assume that the eigenvalues of V_i are ordered in such a way that

$$D_i = \text{diag}(0, \eta_i, \dots, \eta_i, 1) \quad \forall i \in J_B, \quad (18)$$

where η_i denotes the multiple eigenvalue that lies in the open interval $(0, 1)$. Correspondingly we also partition the orthogonal matrices Q_i as

$$Q_i = (q_i, \hat{Q}_i, q'_i) \quad \forall i \in J_B, \quad (19)$$

where $q_i \in \mathbb{R}^{n_i}$ denotes the first column of Q_i , $q'_i \in \mathbb{R}^{n_i}$ is the last column of Q_i , and $\hat{Q}_i \in \mathbb{R}^{n_i \times (n_i - 2)}$ contains the remaining $n_i - 2$ middle columns of Q_i . We also use the following partitionings of the matrices Q_i :

$$Q_i = (q_i, \bar{Q}_i) = (\tilde{Q}_i, q'_i) \quad \forall i \in J_B, \quad (20)$$

where, again, $q_i \in \mathbb{R}^{n_i}$ and $q'_i \in \mathbb{R}^{n_i}$ are the first and the last columns of Q_i , respectively, and $\bar{Q}_i \in \mathbb{R}^{n_i \times (n_i - 1)}$ and $\tilde{Q}_i \in \mathbb{R}^{n_i \times (n_i - 1)}$ contain the remaining $n_i - 1$ columns of Q_i . It is worth noticing that, by (4), the vectors q_i and q'_i are actually given by

$$q_i = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ \frac{\bar{x}_i^* - \bar{\lambda}_i^*}{\|\bar{x}_i^* - \bar{\lambda}_i^*\|} \end{pmatrix} \quad \text{and} \quad q'_i = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \frac{\bar{x}_i^* - \bar{\lambda}_i^*}{\|\bar{x}_i^* - \bar{\lambda}_i^*\|} \end{pmatrix},$$

where $1/\sqrt{2}$ is the normalizing coefficient. Also, by Lemma 2.7, the eigenvalue η_i in (18) is given by

$$\eta_i = \frac{1}{2} \left(1 + \frac{x_{i0}^* - \lambda_{i0}^*}{\|\bar{x}_i^* - \bar{\lambda}_i^*\|} \right). \quad (21)$$

(From [1, Lemma 15], we may easily deduce $\bar{x}_i^* - \bar{\lambda}_i^* \neq 0$ whenever $x_i^{*T} \lambda_i^* = 0$, $x_i^* \in \text{bd}^+ \mathcal{K}_i$, $\lambda_i^* \in \text{bd}^+ \mathcal{K}_i$.)

Consider the matrix D_β defined by (7). In the SOCP under consideration, for each $j \in \beta$, a_j and b_j are given by

$$a_j = \frac{1}{2} \left(1 - \frac{s_{i0}}{\|\bar{s}_i\|} \right), \quad b_j = \frac{1}{2} \left(1 + \frac{s_{i0}}{\|\bar{s}_i\|} \right)$$

with $s_i := x_i^* - \lambda_i^*$ corresponding to some index i belonging to J_B (cf. the proof of Theorem 3.5 below). For any such pair (x_i^*, λ_i^*) , $i \in J_B$, we have

$$x_{i0}^* = \|\bar{x}_i^*\|, \quad \lambda_{i0}^* = \|\bar{\lambda}_i^*\|$$

and

$$x_i^* = \rho_i R_i \lambda_i^*,$$

where $\rho_i = x_{i0}^*/\lambda_{i0}^*$ and $R_i = \begin{pmatrix} 1 & 0 \\ 0 & -I_{n_i-1} \end{pmatrix}$, see [1, Lemma 15]. Hence we have

$$s_i = (\rho_i R_i - I_{n_i}) \lambda_i^* = - \begin{pmatrix} (1 - \rho_i) \lambda_{i0}^* \\ (1 + \rho_i) \bar{\lambda}_i^* \end{pmatrix},$$

which implies $s_{i0}/\|\bar{s}_i\| = (1 - \rho_i)/(1 + \rho_i)$. Therefore we obtain

$$a_j = \frac{\rho_i}{1 + \rho_i}, \quad b_j = \frac{1}{1 + \rho_i}, \quad \frac{a_j}{b_j} = \rho_i \left(= \frac{x_{i0}^*}{\lambda_{i0}^*} \right).$$

This indicates that $D_\beta = (D_\beta^b)^{-1} D_\beta^a$ is a block diagonal matrix with block components of the form $\rho_i I$, where ρ_i and the size of the identity matrix I vary with blocks. The matrix D_β contains the curvature information of the second-order cone at a boundary surface and $\rho_i = x_{i0}^*/\lambda_{i0}^*$ corresponds to the quantity that appears in the second-order condition given by Bonnans and Ramírez [3, eq.(43)]. In fact, we may regard the conditions given in this paper as a dual counterpart of those given in [3], since the problem studied in the present paper corresponds to the primal problem and that of [3] corresponds to the dual problem in the sense of [1].

We are now able to prove the following nonsingularity result under the assumption that the given KKT point satisfies strict complementarity.

Theorem 3.5 *Let $z^* = (x^*, \mu^*, \lambda^*)$ be a strictly complementary KKT point of the SOCP (2), and let the (block) index sets J_I, J_B, J_0 be defined by (16). Let*

$$|\beta| := \sum_{i \in J_B} (n_i - 2) = \sum_{i \in J_B} n_i - 2|J_B|, \quad |\gamma| := \sum_{i \in J_I} n_i + |J_B|,$$

and

$$D_\beta := \text{diag} \left(\rho_i I_{n_i-2} \ (i \in J_B) \right) \in \mathbb{R}^{|\beta| \times |\beta|} \quad \text{with} \quad \rho_i := \frac{x_{i0}^*}{\lambda_{i0}^*} \ (i \in J_B).$$

Then the Jacobian $M'(z^*)$ exists and is nonsingular if the following conditions hold:

(a) *The matrix*

$$(AQ_\beta, AQ_\gamma) := \left(A_i \hat{Q}_i \ (i \in J_B), A_i \ (i \in J_I), A_i q'_i \ (i \in J_B) \right) \in \mathbb{R}^{m \times (|\beta| + |\gamma|)}$$

has full row rank.

(b) *The matrix*

$$\begin{pmatrix} C_1 + D_\beta & C_2 \\ C_2^T & C_3 \end{pmatrix} \in \mathbb{R}^{(|\beta| + |\gamma|) \times (|\beta| + |\gamma|)}$$

is positive definite on the subspace $\mathcal{V} := \left\{ \begin{pmatrix} d_\beta \\ d_\gamma \end{pmatrix} \in \mathbb{R}^{|\beta| + |\gamma|} \mid (AQ_\beta, AQ_\gamma) \begin{pmatrix} d_\beta \\ d_\gamma \end{pmatrix} = 0 \right\}$,

where

$$C_1 := \left(\hat{Q}_i^T \nabla_{x_i x_j}^2 f(x^*) \hat{Q}_j \ (i, j \in J_B) \right) \in \mathbb{R}^{|\beta| \times |\beta|},$$

$$\begin{aligned}
C_2 &:= \left(\hat{Q}_i^T \nabla_{x_i x_j}^2 f(x^*) (i \in J_B, j \in J_I), \hat{Q}_i^T \nabla_{x_i x_j}^2 f(x^*) q'_j (i \in J_B, j \in J_B) \right) \in \mathbb{R}^{|\beta| \times |\gamma|}, \\
C_3 &:= \begin{pmatrix} \nabla_{x_i x_j}^2 f(x^*) (i \in J_I, j \in J_I) & \nabla_{x_i x_j}^2 f(x^*) q'_j (i \in J_I, j \in J_B) \\ q_i^T \nabla_{x_i x_j}^2 f(x^*) (i \in J_B, j \in J_I) & q_i^T \nabla_{x_i x_j}^2 f(x^*) q'_j (i \in J_B, j \in J_B) \end{pmatrix} \in \mathbb{R}^{|\gamma| \times |\gamma|}.
\end{aligned}$$

For the linear SOCP (1), the assertion holds with condition (b) replaced by the following condition:

(c) The matrix

$$AQ_\gamma := \left(A_i (i \in J_I), A_i q'_i (i \in J_B) \right) \in \mathbb{R}^{m \times |\gamma|}$$

has full column rank.

Proof. The existence of the Jacobian $M'(z^*)$ follows immediately from the assumed strict complementarity of the given KKT point together with Lemma 3.4. A simple calculation shows that

$$M'(z^*) = \begin{pmatrix} \nabla^2 f(x^*) & -A^T & -I_n \\ A & 0 & 0 \\ I_n - V & 0 & V \end{pmatrix},$$

where V is the block diagonal matrix $\text{diag}(V_1, \dots, V_r)$ with $V_i := P'_{\mathcal{K}_i}(x_i^* - \lambda_i^*)$. Therefore, taking into account the fact that all eigenvalues of the matrix V belong to the interval $[0, 1]$ by Lemma 2.7, we are able to apply Proposition 3.2 (with $V^a := I_n - V$ and $V^b := V$) as soon as we identify the index sets $\alpha, \beta, \gamma \subseteq \{1, \dots, n\}$ and the structure of the matrices Q and D from that result.

For each $i \in J_I$, we have $V_i = I_{n_i}$ (see (17)) and, therefore, $Q_i = I_{n_i}$ and $D_i = I_{n_i}$. Hence all components j from the block components $i \in J_I$ belong to the index set γ .

On the other hand, for each $i \in J_0$, we have $V_i = 0$ (see (17)), and this corresponds to $Q_i = I_{n_i}$ and $D_i = 0$. Hence all components j from the block components $i \in J_0$ belong to the index set α .

Finally, let $i \in J_B$. Then $V_i = Q_i D_i Q_i^T$ with $D_i = \text{diag}(0, \eta_i, \dots, \eta_i, 1)$, where $\eta_i \in (0, 1)$ is given by (21), and $Q_i = (q_i, \hat{Q}_i, q'_i)$. Hence the first component for each block index $i \in J_B$ is an element of the index set α , the last component for each block index $i \in J_B$ belongs to the index set γ , and all the remaining middle components belong to the index set β .

Taking into account that $Q = \text{diag}(Q_1, \dots, Q_r)$ and $D = \text{diag}(D_1, \dots, D_r)$ with Q_i, D_i as specified above, and using the partitioning

$$\begin{pmatrix} \nabla_{x_i x_j}^2 f(x^*) (i \in J_I, j \in J_I) & \nabla_{x_i x_j}^2 f(x^*) (i \in J_I, j \in J_B) & \nabla_{x_i x_j}^2 f(x^*) (i \in J_I, j \in J_0) \\ \nabla_{x_i x_j}^2 f(x^*) (i \in J_B, j \in J_I) & \nabla_{x_i x_j}^2 f(x^*) (i \in J_B, j \in J_B) & \nabla_{x_i x_j}^2 f(x^*) (i \in J_B, j \in J_0) \\ \nabla_{x_i x_j}^2 f(x^*) (i \in J_0, j \in J_I) & \nabla_{x_i x_j}^2 f(x^*) (i \in J_0, j \in J_B) & \nabla_{x_i x_j}^2 f(x^*) (i \in J_0, j \in J_0) \end{pmatrix}$$

of the Hessian $\nabla^2 f(x^*)$, it follows immediately from the above observations that conditions (a), (b), and (c) correspond precisely to conditions (a), (b), and (c), respectively, in

Proposition 3.2. □

The following simple example illustrates the conditions in the above theorem.

Example 3.6 Consider the nonlinear SOCP

$$\begin{aligned} \min \quad & \frac{1}{2}x_1^2 + \frac{1}{2}(x_2 - 2)^2 - \frac{\varepsilon}{2}x_3^2 \\ \text{s.t.} \quad & x \in \mathcal{K}^3, \end{aligned}$$

where \mathcal{K}^3 denotes the second-order cone in \mathbb{R}^3 and ε is a scalar parameter. This problem contains only one second-order cone constraint. (Here, unlike the rest of this section, x_i denotes the i th (scalar) component of the vector x .) Note that the objective function is nonconvex for any $\varepsilon > 0$. It is easy to see that the solution of this problem is given by $x^* = (1, 1, 0)^T \in \text{bd}^+\mathcal{K}^3$ together with the multiplier vector $\lambda^* = (1, -1, 0)^T \in \text{bd}^+\mathcal{K}^3$, which satisfies strict complementarity. Furthermore, we have

$$V = P'_{\mathcal{K}^3}(x^* - \lambda^*) = QDQ^T,$$

where $D = \text{diag}(0, \frac{1}{2}, 1)$ and

$$Q = (q, \hat{Q}, q') = \begin{pmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix}.$$

Since there is no equality constraint, condition (a) in Theorem 3.5 is automatically satisfied. Moreover, by direct calculation, we have $C_1 = -\varepsilon$, $C_2 = 0$, $C_3 = 1$, $D_\beta = 1$, and hence

$$\begin{pmatrix} C_1 + D_\beta & C_2 \\ C_2^T & C_3 \end{pmatrix} = \begin{pmatrix} -\varepsilon + 1 & 0 \\ 0 & 1 \end{pmatrix},$$

for which condition (b) holds as long as $\varepsilon < 1$, since $\mathcal{V} = \mathbb{R}^2$. This example shows that condition (b) may be secured with the aid of the curvature term D_β even if the Hessian of the objective function fails to be positive definite in itself. ◇

We now want to extend Theorem 3.5 to the case where strict complementarity is violated. Let $z^* = (x^*, \mu^*, \lambda^*)$ be an arbitrary KKT point of the SOCP, and let J_I, J_B, J_0 denote the index sets defined by (16). In view of Lemma 3.3, in addition to these sets, we also need to consider the three index sets

$$\begin{aligned} J_{B0} &:= \{i \mid x_i^* \in \text{bd}^+\mathcal{K}_i, \lambda_i^* = 0\}, \\ J_{0B} &:= \{i \mid x_i^* = 0, \lambda_i^* \in \text{bd}^+\mathcal{K}_i\}, \\ J_{00} &:= \{i \mid x_i^* = 0, \lambda_i^* = 0\}, \end{aligned} \tag{22}$$

which correspond to the block indices where strict complementarity is violated. Note that these index sets have double subscripts; the first (resp. second) subscript indicates

whether x_i^* (resp. λ_i^*) is on the boundary of \mathcal{K}_i (excluding zero) or equal to the zero vector. The following result summarizes the structure of the matrices $V_i \in \partial_B P_{\mathcal{K}_i}(x_i^* - \lambda_i^*)$ for $i \in J_{B0} \cup J_{0B} \cup J_{00}$. Hence it is the counterpart of Lemma 3.4 in the general case.

Lemma 3.7 *Let $i \in J_{B0} \cup J_{0B} \cup J_{00}$ and $V_i \in \partial_B P_{\mathcal{K}_i}(x_i^* - \lambda_i^*)$. Then the following statements hold:*

- (a) *If $i \in J_{B0}$, then we have either $V_i = I_{n_i}$ or $V_i = Q_i D_i Q_i^T$ with $D_i = \text{diag}(0, 1, \dots, 1)$ and $Q_i = (q_i, \bar{Q}_i)$.*
- (b) *If $i \in J_{0B}$, then we have either $V_i = 0$ or $V_i = Q_i D_i Q_i^T$ with $D_i = \text{diag}(0, \dots, 0, 1)$ and $Q_i = (\tilde{Q}_i, q_i')$.*
- (c) *If $i \in J_{00}$, then we have $V_i = I_{n_i}$ or $V_i = 0$ or $V_i = Q_i D_i Q_i^T$ with D_i and Q_i given by $D_i = \text{diag}(0, \eta_i, \dots, \eta_i, 1)$ for some $\eta_i \in (0, 1)$ and $Q_i = (q_i, \hat{Q}_i, q_i')$, or by $D_i = \text{diag}(0, 1, \dots, 1)$ and $Q_i = (q_i, \bar{Q}_i)$, or by $D_i = \text{diag}(0, \dots, 0, 1)$ and $Q_i = (\tilde{Q}_i, q_i')$.*

Proof. First let $i \in J_{B0}$. Then $s_i := x_i^* - \lambda_i^* = x_i^* \in \text{bd}^+ \mathcal{K}_i$. Therefore, if we write $s_i = (s_{i0}, \bar{s}_i)$, it follows that $s_{i0} = \|\bar{s}_i\|$ and $\bar{s}_i \neq 0$. Statement (a) then follows immediately from Lemma 2.6 (b) in combination with Lemma 2.7.

In a similar way, the other two statements can be derived by using Lemma 2.6 (c) and (d), respectively, together with Lemma 2.7 in order to get the eigenvalues. Here the five possible choices in statement (c) depend, in particular, on the value of the scalar ρ in Lemma 2.6 (d) (namely $\rho \in (-1, 1)$, $\rho = 1$, and $\rho = -1$). \square

Lemma 3.7 enables us to generalize Theorem 3.5 to the case where strict complementarity does not hold. Note that, from now on, we use the spectral decompositions $V_i = Q_i D_i Q_i^T$ and the associated partitionings (18)–(20) for all $i \in J_B$, as well as those specified in Lemma 3.7 for all indices $i \in J_{B0} \cup J_{0B} \cup J_{00}$.

Theorem 3.8 *Let $z^* = (x^*, \mu^*, \lambda^*)$ be a (not necessarily strictly complementary) KKT point of the SOCP (2), and let the (block) index sets $J_I, J_B, J_0, J_{B0}, J_{0B}, J_{00}$ be defined by (16) and (22). Then all matrices $W \in \partial_B M(z^*)$ are nonsingular if, for any partitioning $J_{B0} = J_{B0}^1 \cup J_{B0}^2$, any partitioning $J_{0B} = J_{0B}^1 \cup J_{0B}^2$, and any partitioning $J_{00} = J_{00}^1 \cup J_{00}^2 \cup J_{00}^3 \cup J_{00}^4 \cup J_{00}^5$, the following two conditions (a) and (b) hold with*

$$\begin{aligned}
|\beta| &:= \sum_{i \in J_B \cup J_{00}^3} (n_i - 2), \\
|\gamma| &:= \sum_{i \in J_I \cup J_{B0}^1 \cup J_{00}^1} n_i + |J_B \cup J_{0B}^2 \cup J_{00}^3 \cup J_{00}^5| + \sum_{i \in J_{B0}^2 \cup J_{00}^4} (n_i - 1), \\
D_\beta &:= \text{diag}\left(\rho_i I_{n_i-2} \ (i \in J_B \cup J_{00}^3)\right) \in \mathbb{R}^{|\beta| \times |\beta|} \quad \text{with} \quad \rho_i = \frac{x_{i0}^*}{\lambda_{i0}^*} \ (i \in J_B), \ \rho_i > 0 \ (i \in J_{00}^3) :
\end{aligned}$$

(a) *The matrix*

$$(AQ_\beta, AQ_\gamma) := \begin{pmatrix} A_i \hat{Q}_i (i \in J_B \cup J_{00}^3), A_i (i \in J_I \cup J_{B0}^1 \cup J_{00}^1), \\ A_i q'_i (i \in J_B \cup J_{0B}^2 \cup J_{00}^3 \cup J_{00}^5), A_i \bar{Q}_i (i \in J_{B0}^2 \cup J_{00}^4) \end{pmatrix} \in \mathbb{R}^{m \times (|\beta| + |\gamma|)}$$

has full row rank.

(b) *The matrix*

$$\begin{pmatrix} C_1 + D_\beta & C_2 \\ C_2^T & C_3 \end{pmatrix} \in \mathbb{R}^{(|\beta| + |\gamma|) \times (|\beta| + |\gamma|)}$$

is positive definite on the subspace $\mathcal{V} := \left\{ \begin{pmatrix} d_\beta \\ d_\gamma \end{pmatrix} \in \mathbb{R}^{|\beta| + |\gamma|} \mid (AQ_\beta, AQ_\gamma) \begin{pmatrix} d_\beta \\ d_\gamma \end{pmatrix} = 0 \right\}$, where

$$\begin{aligned} C_1 &:= \left(\hat{Q}_i^T \nabla_{x_i x_j}^2 f(x) \hat{Q}_j (i, j \in J_B \cup J_{00}^3) \right) \in \mathbb{R}^{|\beta| \times |\beta|}, \\ C_2 &:= (C_2^1, C_2^2, C_2^3) \in \mathbb{R}^{|\beta| \times |\gamma|}, \\ C_3 &:= \begin{pmatrix} C_3^{11} & C_3^{12} & C_3^{13} \\ (C_3^{12})^T & C_3^{22} & C_3^{23} \\ (C_3^{13})^T & (C_3^{23})^T & C_3^{33} \end{pmatrix} \in \mathbb{R}^{|\gamma| \times |\gamma|} \end{aligned}$$

with the submatrices

$$\begin{aligned} C_2^1 &:= \left(\hat{Q}_i^T \nabla_{x_i x_j}^2 f(x^*) (i \in J_B \cup J_{00}^3, j \in J_I \cup J_{B0}^1 \cup J_{00}^1) \right), \\ C_2^2 &:= \left(\hat{Q}_i^T \nabla_{x_i x_j}^2 f(x^*) q'_j (i \in J_B \cup J_{00}^3, j \in J_B \cup J_{0B}^2 \cup J_{00}^3 \cup J_{00}^5) \right), \\ C_2^3 &:= \left(\hat{Q}_i^T \nabla_{x_i x_j}^2 f(x^*) \bar{Q}_j (i \in J_B \cup J_{00}^3, j \in J_{B0}^2 \cup J_{00}^4) \right) \end{aligned}$$

and

$$\begin{aligned} C_3^{11} &:= \left(\nabla_{x_i x_j}^2 f(x^*) (i \in J_I \cup J_{B0}^1 \cup J_{00}^1, j \in J_I \cup J_{B0}^1 \cup J_{00}^1) \right), \\ C_3^{12} &:= \left(\nabla_{x_i x_j}^2 f(x^*) q'_j (i \in J_I \cup J_{B0}^1 \cup J_{00}^1, j \in J_B \cup J_{0B}^2 \cup J_{00}^3 \cup J_{00}^5) \right), \\ C_3^{13} &:= \left(\nabla_{x_i x_j}^2 f(x^*) \bar{Q}_j (i \in J_I \cup J_{B0}^1 \cup J_{00}^1, j \in J_{B0}^2 \cup J_{00}^4) \right), \\ C_3^{22} &:= \left(q_i'^T \nabla_{x_i x_j}^2 f(x^*) q'_j (i \in J_B \cup J_{0B}^2 \cup J_{00}^3 \cup J_{00}^5, j \in J_B \cup J_{0B}^2 \cup J_{00}^3 \cup J_{00}^5) \right), \\ C_3^{23} &:= \left(q_i'^T \nabla_{x_i x_j}^2 f(x^*) \bar{Q}_j (i \in J_B \cup J_{0B}^2 \cup J_{00}^3 \cup J_{00}^5, j \in J_{B0}^2 \cup J_{00}^4) \right), \\ C_3^{33} &:= \left(\bar{Q}_i^T \nabla_{x_i x_j}^2 f(x^*) \bar{Q}_j (i \in J_{B0}^2 \cup J_{00}^4, j \in J_{B0}^2 \cup J_{00}^4) \right). \end{aligned}$$

For the linear SOCP (1), the assertion holds with condition (b) replaced by the following condition:

(c) *The matrix*

$$AQ_\gamma := \begin{pmatrix} A_i (i \in J_I \cup J_{B0}^1 \cup J_{00}^1), A_i q'_i (i \in J_B \cup J_{0B}^2 \cup J_{00}^3 \cup J_{00}^5), \\ A_i \bar{Q}_i (i \in J_{B0}^2 \cup J_{00}^4) \end{pmatrix} \in \mathbb{R}^{m \times |\gamma|}$$

has full column rank.

Proof. Choose $W \in \partial_B M(z^*)$ arbitrarily. Then a simple calculation shows that

$$W = \begin{pmatrix} \nabla^2 f(x^*) & -A^T & -I_n \\ A & 0 & 0 \\ I_n - V & 0 & V \end{pmatrix}$$

for a suitable block diagonal matrix $V = \text{diag}(V_1, \dots, V_r)$ with $V_i \in \partial_B P_{\mathcal{K}_i}(x_i^* - \lambda_i^*)$. In principle, the proof is similar to the one of Theorem 3.5: We want to apply Proposition 3.2 (with $V^a := I - V$ and $V^b := V$) by identifying the index sets α, β, γ (and the matrices Q, D). The situation is, however, more complicated here, since these index sets may depend on the particular element W chosen from the B-subdifferential $\partial_B M(z^*)$. To this end, we take a closer look especially at the index sets J_{B0}, J_{0B} , and J_{00} . In view of Lemma 3.7, we further partition these index sets into

$$\begin{aligned} J_{B0} &= J_{B0}^1 \cup J_{B0}^2, \\ J_{0B} &= J_{0B}^1 \cup J_{0B}^2, \\ J_{00} &= J_{00}^1 \cup J_{00}^2 \cup J_{00}^3 \cup J_{00}^4 \cup J_{00}^5 \end{aligned}$$

with

$$\begin{aligned} J_{B0}^1 &:= \{i \mid V_i = I_{n_i}\}, & J_{B0}^2 &:= J_{B0} \setminus J_{B0}^1, \\ J_{0B}^1 &:= \{i \mid V_i = 0\}, & J_{0B}^2 &:= J_{0B} \setminus J_{0B}^1, \end{aligned}$$

and

$$\begin{aligned} J_{00}^1 &:= \{i \mid V_i = I_{n_i}\}, \\ J_{00}^2 &:= \{i \mid V_i = 0\}, \\ J_{00}^3 &:= \{i \mid V_i = Q_i D_i Q_i^T \text{ with } D_i \text{ and } Q_i \text{ given by (18) and (19), respectively}\}, \\ J_{00}^4 &:= \{i \mid V_i = Q_i D_i Q_i^T \text{ with } D_i = \text{diag}(0, 1, \dots, 1) \text{ and } Q_i = (q_i, \bar{Q}_i)\}, \\ J_{00}^5 &:= \{i \mid V_i = Q_i D_i Q_i^T \text{ with } D_i = \text{diag}(0, \dots, 0, 1) \text{ and } Q_i = (\bar{Q}_i, q'_i)\}. \end{aligned}$$

Using these definitions and Lemmas 3.4 and 3.7, we see that the following indices j belong to the index set α in Proposition 3.2:

- All indices j belonging to one of the block indices $i \in J_0 \cup J_{0B}^1 \cup J_{00}^2$, with $Q_i = I_{n_i}$ being the corresponding orthogonal matrix.
- The first component belonging to a block index $i \in J_B \cup J_{B0}^2 \cup J_{00}^3 \cup J_{00}^4$, with q_i being the first column of the corresponding orthogonal matrix Q_i .

- The first $n_i - 1$ components belonging to each block index $i \in J_{0B}^2 \cup J_{00}^5$, with \tilde{Q}_i consisting of the first $n_i - 1$ columns of the corresponding orthogonal matrix Q_i .

We next consider the index set β in Proposition 3.2. In view of Lemmas 3.4 and 3.7, the following indices j belong to this set:

- All middle indices belonging to a block index $i \in J_B \cup J_{00}^3$, with \hat{Q}_i consisting of the middle $n_i - 2$ columns of the corresponding orthogonal matrix Q_i .

Using Lemmas 3.4 and 3.7 again, we finally see that the following indices j belong to the index set γ in Proposition 3.2:

- All indices j belonging to one of the block indices $i \in J_I \cup J_{B0}^1 \cup J_{00}^1$. The corresponding orthogonal matrix is $Q_i = I_{n_i}$.
- The last index of each block index $i \in J_B \cup J_{0B}^2 \cup J_{00}^3 \cup J_{00}^5$, with q'_i being the last column of the corresponding orthogonal matrix Q_i .
- The last $n_i - 1$ indices j belonging to a block index $i \in J_{B0}^2 \cup J_{00}^4$, with \bar{Q}_i consisting of the last $n_i - 1$ columns of the corresponding orthogonal matrix Q_i .

The theorem then follows from Proposition 3.2 in a way similar to the proof of Theorem 3.5. \square

Note that, in the case of a strictly complementary KKT point, Theorem 3.8 reduces to Theorem 3.5. It may be worth noticing that, for interior-point methods of the linear SOCP, we cannot expect to have a result corresponding to Theorem 3.8, since the Jacobian matrices arising in that context are singular whenever the strict complementarity fails to hold. Using Theorems 3.1 and 3.8 along with [22], we get the following result.

Theorem 3.9 *Let $z^* = (x^*, \mu^*, \lambda^*)$ be a (not necessarily strictly complementary) KKT point of the SOCP (1), and suppose that the assumptions of Theorem 3.8 hold at this KKT point. Then the nonsmooth Newton method (6) applied to the system of equations $M(z) = 0$ is locally superlinearly convergent. If, in addition, f has a locally Lipschitz continuous Hessian, then it is locally quadratically convergent.*

4 Numerical Examples

In this section, we show some preliminary numerical results with the nonsmooth Newton method tested on linear and nonlinear SOCPs. The main aim of our numerical experiments is to demonstrate the theoretical results established in the previous section by examining the local behaviour of the method, rather than making a comparison with existing solvers. Note that the usage of symbols such as x and x_i in this section is different from the previous

k	$\ M(z^k)\ $	x_1^k	x_2^k	x_3^k	$\ \nabla f(x^k) - \lambda^k\ $
0	1.020784e+02	2.000000e+00	2.000000e+00	2.000000e+00	3.464102e+00
1	1.414214e+00	0.000000e+00	2.000000e+00	0.000000e+00	0.000000e+00
2	0.000000e+00	1.000000e+00	1.000000e+00	0.000000e+00	0.000000e+00

Table 1: Numerical results for the nonconvex SOCP of Example 4.1

sections. However there should be no confusion since the meaning will be clear from the context.

Example 4.1 We first consider the nonconvex SOCP of Example 3.6. Letting $\varepsilon := \frac{1}{2}$ and using the starting point $x^0 := (2, 2, 2)^T$ together with the multipliers $\lambda^0 := (2, 2, 2)^T$, we obtain the results shown in Table 1. Here we have very fast convergence in just two iterations. \diamond

Our next example is taken from [13].

Example 4.2 Consider the following nonlinear (convex) SOCP:

$$\begin{aligned} \min \quad & \exp(x_1 - x_3) + 3(2x_1 - x_2)^4 + \sqrt{1 + (3x_2 + 5x_3)^2} \\ \text{s.t.} \quad & \begin{pmatrix} 4 & 6 & 3 \\ -1 & 7 & -5 \end{pmatrix} x + \begin{pmatrix} -1 \\ 2 \end{pmatrix} \in \mathcal{K}^2, \quad x \in \mathcal{K}^3, \end{aligned}$$

where \mathcal{K}^r denotes the second-order cone in \mathbb{R}^r . This problem can be transformed into the standard form

$$\min f(w) \quad \text{s.t.} \quad Aw = b, \quad w \in \mathcal{K}$$

with $f(w) := \exp(x_1 - x_3) + 3(2x_1 - x_2)^4 + \sqrt{1 + (3x_2 + 5x_3)^2}$ and

$$w := (x^T, y^T)^T \in \mathbb{R}^3 \times \mathbb{R}^2, \quad A := \left(\begin{array}{ccc|cc} 4 & 6 & 3 & -1 & 0 \\ -1 & 7 & -5 & 0 & -1 \end{array} \right), \quad b := \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad \mathcal{K} := \mathcal{K}^3 \times \mathcal{K}^2.$$

Table 2 shows a sequence generated by the nonsmooth Newton method with a starting point randomly chosen from the box $[0, 1]^5 \subset \mathbb{R}^5$. We may observe a typical feature of the local quadratic convergence. \diamond

We next consider a particular example of the linear SOCP, for which the assumptions for the local quadratic convergence of the nonsmooth Newton method hold.

Example 4.3 Consider the problem of minimizing the maximum distance to $N + 1$ points b_i ($i = 0, 1, \dots, N$) in the Euclidean space \mathbb{R}^n :

$$\min_{t,x} t \quad \text{s.t.} \quad \|x - b_i\| \leq t, \quad i = 0, 1, \dots, N.$$

k	$\ M(z^k)\ $	x_1^k	x_2^k	x_3^k	$\ Aw^k - b\ $
0	1.273197e+02	9.501293e-01	2.311385e-01	6.068426e-01	5.663307e+00
1	3.765549e+01	3.019551e-01	-5.312774e-01	1.198684e-01	6.280370e-16
2	3.158146e+01	2.331042e-01	-9.730924e-02	2.171462e-01	1.110223e-15
3	3.259677e+00	1.196822e-01	-9.886805e-02	3.688092e-02	0.000000e+00
4	1.675676e+00	1.973609e-01	-8.539481e-02	2.409751e-01	4.440892e-16
5	3.516159e-01	2.357895e-01	-9.820433e-02	2.153560e-01	1.110223e-16
6	4.875888e-02	2.325429e-01	-7.451132e-02	2.203468e-01	0.000000e+00
7	1.511531e-04	2.324026e-01	-7.308263e-02	2.206131e-01	2.220446e-16
8	7.295537e-10	2.324025e-01	-7.307928e-02	2.206135e-01	1.110223e-16
9	1.102302e-15	2.324025e-01	-7.307928e-02	2.206135e-01	4.440892e-16

Table 2: Numerical results for the nonlinear (convex) SOCP of Example 4.2

By translating the axes if necessary, we assume without loss of generality that $b_0 = 0$. Then this problem can be rewritten as

$$\min_{t, x, y_1, \dots, y_N} t \quad \text{s.t.} \quad \begin{pmatrix} t \\ x \end{pmatrix} - y_i = \begin{pmatrix} 0 \\ b_i \end{pmatrix}, \quad \begin{pmatrix} t \\ x \end{pmatrix} \in \mathcal{K}^{n+1}, \quad y_i \in \mathcal{K}^{n+1}, \quad i = 1, \dots, N,$$

where \mathcal{K}^{n+1} denotes the second-order cone in \mathbb{R}^{n+1} . This is a linear SOCP of the standard form

$$\min f(w) \quad \text{s.t.} \quad Aw = b, \quad w \in \mathcal{K},$$

with the objective function $f(w) := c^T w$, the variables

$$w := (t, x^T, y_1^T, \dots, y_N^T)^T \in \mathbb{R}^p, \quad p := 1 + n + (n+1)N,$$

and the data

$$\begin{aligned} c &:= (1, 0, \dots, 0)^T \in \mathbb{R}^p, \\ b &:= (0, b_1^T, 0, b_2^T, \dots, 0, b_N^T)^T \in \mathbb{R}^{(n+1)N}, \\ A &:= \begin{pmatrix} I_{n+1} & -I_{n+1} & & & 0 \\ I_{n+1} & & -I_{n+1} & & \\ \vdots & & & \ddots & \\ I_{n+1} & 0 & & & -I_{n+1} \end{pmatrix} \in \mathbb{R}^{(n+1)N \times p}, \\ \mathcal{K} &:= \underbrace{\mathcal{K}^{n+1} \times \dots \times \mathcal{K}^{n+1}}_{(N+1)\text{-times}}. \end{aligned}$$

To be more specific, let us consider the particular instance with $n = 2$, $N = 2$, and $b_1 = (4, 0)^T$, $b_2 = (4, 4)^T$. The solution of this problem is given by $x^* := (2, 2)$ with optimal distance $t^* = 2\sqrt{2}$ and the vectors $y_1^* = (2\sqrt{2}, -2, 2)^T$ and $y_2^* = (2\sqrt{2}, -2, -2)^T$. An elementary calculation shows that the corresponding optimal multipliers are given by $\mu^* = (0, 0, 0, \frac{1}{2}, \frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}})^T$ and $\lambda^* = (\frac{1}{2}, \frac{-1}{2\sqrt{2}}, \frac{-1}{2\sqrt{2}}, 0, 0, 0, \frac{1}{2}, \frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}})^T$.

k	$\ M(z^k)\ $	x_1^k	x_2^k	$\ Aw^k - b\ $	$\ \phi(w^k, \lambda^k)\ $
0	7.029663e+00	8.121259e-01	9.082626e-01	6.393857e+00	1.769382e+00
1	4.816071e+01	4.094855e+00	2.944570e+00	1.297632e-13	4.816071e+01
2	3.107185e+01	1.971163e+00	1.659028e+00	5.043708e-12	3.107185e+01
3	2.109201e+00	3.278852e+00	9.489322e-01	9.485085e-11	2.109201e+00
4	9.107635e-01	2.232357e+00	1.816440e+00	1.719429e-11	9.107635e-01
5	4.255234e-02	2.032107e+00	1.967839e+00	5.819005e-10	4.255234e-02
6	5.825558e-04	2.000563e+00	1.999989e+00	1.239050e-10	5.825558e-04
7	4.474272e-08	2.000000e+00	2.000000e+00	1.379164e-11	4.474272e-08
8	7.675809e-15	2.000000e+00	2.000000e+00	6.616780e-15	3.890536e-15

Table 3: Numerical results for the linear SOCP of Example 4.3

k	$\ M(z^k)\ $	$\ Aw^k - b\ $	$\ \phi(w^k, \lambda^k)\ $
34	2.397717e-03	7.130277e-13	2.397717e-03
35	6.252936e-07	9.029000e-13	6.252936e-07
36	1.470491e-09	5.177835e-13	1.470491e-09
37	4.781069e-12	6.815003e-13	4.732249e-12

Table 4: Numerical results for the linear SOCP `nb_L1` from the DIMACS collection

We find that the solution violates the strict complementarity condition (namely in its second block component). Nevertheless, it is not difficult to see that the assumptions of Theorem 3.8 (for the linear SOCP) are satisfied for this particular example (here we have $J_B = \{1, 3\}$ and $J_{B_0} = \{2\}$, so there exist only two possible partitionings of the index set J_{B_0}). In fact, taking into account the special structure of the matrix A in this application of SOCP, there is a good chance that these assumptions hold for many instances of this problem.

We applied the nonsmooth Newton method to this example. The results are shown in Table 3, where the function ϕ in the last column is defined by $\phi(w, \lambda) := w - P_{\mathcal{K}}(w - \lambda)$. We observe that the method is just a local one: The residual $\|M(z^k)\|$ increases in the beginning. Fortunately, after a few steps, $\|M(z^k)\|$ starts to decrease, and eventually exhibits nice local quadratic convergence. \diamond

We also applied the nonsmooth Newton method to the three SOCPs in the DIMACS library, see [21]. Due to its local nature, the method sometimes failed to converge depending on the choice of a starting point. Nevertheless, the asymptotic behaviour of the method applied to problem `nb_L1` from the DIMACS collection, as shown in Table 4, indicates that the rate of convergence is at least superlinear for this problem. Whether the non-quadratic convergence has to do with the fact that our assumptions are violated, or it is simply due to the finite precision arithmetic of the computer, is currently not clear to us.

5 Final Remarks

We have investigated the local properties of a semismooth equation reformulation of both the linear and the nonlinear SOCPs. In particular, we have shown nonsingularity results that provide basic conditions for local quadratic convergence of a nonsmooth Newton method. Strict complementarity of a solution is not needed in our nonsingularity results. Apart from these local properties, it is certainly of interest to see how the local Newton method can be globalized in a suitable way. We leave it as a future research topic.

Acknowledgment. The authors are grateful to the referees for their critical comments that have led to a significant improvement of the paper.

References

- [1] F. ALIZADEH AND D. GOLDFARB: *Second-order cone programming*. Mathematical Programming 95, 2003, pp. 3–51.
- [2] A. BEN-TAL AND A. NEMIROVSKI: *Lectures on Modern Convex Optimization*. MPS-SIAM Series on Optimization, SIAM, Philadelphia, PA, 2001.
- [3] J.F. BONNANS AND H. RAMÍREZ C.: *Perturbation analysis of second-order cone programming problems*. Mathematical Programming 104, 2005, pp. 205–227.
- [4] S. BOYD AND L. VANDENBERGHE: *Convex Optimization*. Cambridge University Press, Cambridge, 2004.
- [5] X.D. CHEN, D. SUN, AND J. SUN: *Complementarity functions and numerical experiments for second-order cone complementarity problems*. Computational Optimization and Applications 25, 2003, pp. 39–56.
- [6] J.-S. CHEN: *Alternative proofs for some results of vector-valued functions associated with second-order cones*. Journal of Nonlinear and Convex Analysis 6, 2005, pp. 297–325.
- [7] J.-S. CHEN, X. CHEN, AND P. TSENG: *Analysis of nonsmooth vector-valued functions associated with second-order cones*. Mathematical Programming 101, 2004, pp. 95–117.
- [8] J.-S. CHEN AND P. TSENG: *An unconstrained smooth minimization reformulation of the second-order cone complementarity problem*. Mathematical Programming 104, 2005, pp. 293–327.
- [9] F.H. CLARKE: *Optimization and Nonsmooth Analysis*. John Wiley & Sons, New York, NY, 1983 (reprinted by SIAM, Philadelphia, PA, 1990).

- [10] F. FACCHINEI AND J.S. PANG: *Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume II*. Springer, New York, NY, 2003.
- [11] M.L. FLEGEL AND C. KANZOW: *Equivalence of two nondegeneracy conditions for semidefinite programs*. Journal of Optimization Theory and Applications, to appear.
- [12] M. FUKUSHIMA, Z.-Q. LUO, AND P. TSENG: *Smoothing functions for second-order cone complementarity problems*. SIAM Journal on Optimization 12, 2001, pp. 436–460.
- [13] S. HAYASHI, N. YAMASHITA, AND M. FUKUSHIMA: *A combined smoothing and regularization method for monotone second-order cone complementarity problems*. SIAM Journal on Optimization 15, 2005, pp. 593–615.
- [14] C. HELMBERG: <http://www-user.tu-chemnitz.de/~helmborg/semidef.html>.
- [15] C. KANZOW AND C. NAGEL: *Quadratic convergence of a nonsmooth Newton-type method for semidefinite programs without strict complementarity*. SIAM Journal on Optimization 15, 2005, pp. 654–672.
- [16] H. KATO AND M. FUKUSHIMA: *An SQP-type algorithm for nonlinear second-order cone programs*. Optimization Letters, to appear.
- [17] M.S. LOBO, L. VANDENBERGHE, AND S. BOYD: *SOCP – Software for second-order cone programming. User’s guide*. Technical Report, Department of Electrical Engineering, Stanford University, April 1997.
- [18] M.S. LOBO, L. VANDENBERGHE, S. BOYD, AND H. LEBRET: *Applications of second-order cone programming*. Linear Algebra and Its Applications 284, 1998, pp. 193–228.
- [19] J.-S. PANG, D. SUN, AND J. SUN: *Semismooth homeomorphisms and strong stability of semidefinite and Lorentz complementarity problems*. Mathematics of Operations Research 28, 2003, pp. 39–63.
- [20] J.-S. PANG AND L. QI: *Nonsmooth equations: Motivation and algorithms*. SIAM Journal on Optimization 3, 1993, pp. 443–465.
- [21] G. PATAKI AND S. SCHMIETA: *The DIMACS library of semidefinite-quadratic-linear programs*. Preliminary draft, Computational Optimization Research Center, Columbia University, New York, NY, July 2002.
- [22] L. QI: *Convergence analysis of some algorithms for solving nonsmooth equations*. Mathematics of Operations Research 18, 1993, pp. 227–244.
- [23] L. QI AND J. SUN: *A nonsmooth version of Newton’s method*. Mathematical Programming 58, 1993, pp. 353–367.

- [24] J.F. STURM: *Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones*. Optimization Methods and Software 11/12, 1999, pp. 625–653. <http://sedumi.mcmaster.ca/>
- [25] D. SUN AND J. SUN: *Strong semismoothness of the Fischer-Burmeister SDC and SOC complementarity functions*. Mathematical Programming 103, 2005, pp. 575–581.
- [26] P. TSENG: *Smoothing methods for second-order cone programs/complementarity problems*. Talk presented at the SIAM Conference on Optimization, Stockholm, May 2005.
- [27] R.H. TÜTÜNCÜ, K.C. TOH, AND M.J. TODD: *Solving semidefinite-quadratic-linear programs using SDPT3*. Mathematical Programming 95, 2003, pp. 189–217. <http://www.math.nus.edu.sg/~mattohkc/sdpt3.html>
- [28] R.J. VANDERBEI AND H. YURITAN: *Using LOQO to solve second-order cone programming problems*. Technical Report, Statistics and Operations Research, Princeton University, 1998.
- [29] H. YAMASHITA AND H. YABE: *A primal-dual interior point method for nonlinear optimization over second order cones*. Technical Report, Mathematical Systems Inc., Tokyo, May 2005 (revised February 2006).