SEMISMOOTH METHODS FOR LINEAR AND NONLINEAR SECOND-ORDER CONE PROGRAMS[1]

Christian Kanzow[2] and Masao Fukushima[3]

April 19, 2006

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.

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

---

[1]This work was supported in part by the Scientific Research Grant-in-Aid from Japan Society for the Promotion of Science.

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

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

1 Introduction

We consider both the linear second-order cone program

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

and the nonlinear second-order cone program

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

where $f : \mathbb{R}^n \to \mathbb{R}$ 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

$$\mathcal{K}_i := \left\{ z_i = (z_{i0}, \bar{z}_i) \in \mathbb{R} \times \mathbb{R}^{n_i - 1} \,\middle|\, z_{i0} \geq \|\bar{z}_i\| \right\},$$

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 second-order cone program has been investigated in many previous works, and we refer the interested reader to the two survey papers [16, 1] and the books [2, 4] for many important applications and theoretical properties. Software for the solution of linear second-order cone programs is also available, see, for example, [15, 24, 20, 23]. In many cases, the linear second-order cone program is viewed as a special case of a (linear) semidefinite program (see [1] for a suitable reformulation). However, we feel that the second-order cone program 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 [13]) are able to deal with second-order cone constraints separately.

The treatment of the nonlinear second-order cone program is much more recent, and, in the moment, the number of publications is rather limited, see [3, 5, 6, 7, 8, 11, 12, 14, 21, 22, 25]. 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 second-order cone program.

The method of choice for the solution of (at least) the linear second-order cone program 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, 12, 22]. 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 second-order cone programs. 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.

The paper is organized as follows: Section 2 states a number of preliminary results for a projection mapping onto a second-order cone, which will later be used in order to reformulate the optimality conditions of the second-order cone program as a system of equations. Section 3 then investigates the nonlinear second-order cone program, whereas its linear counterpart is discussed in Section 4. We close with some final remarks in Section 5.

Most of our notation is standard. For a differentiable mapping $G : \mathbb{R}^n \to \mathbb{R}^m$, 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

$$\partial_B G(z) := \left\{ H \in \mathbb{R}^{m \times n} \,\middle|\, \exists \{z^k\} \subseteq D_G : z^k \to z, G'(z^k) \to H \right\}$$

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 [19, 18, 17, 10] for details.

2 Projection Mapping onto Second-Order Cone

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

$$\mathcal{K} := \left\{ z = (z_0, \bar{z}) \in \mathbb{R} \times \mathbb{R}^n \,\middle|\, z_0 \geq \|\bar{z}\| \right\}.$$

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 $\mathcal{K}^* := \{d \in \mathbb{R} \times \mathbb{R}^n \,|\, z^T d \geq 0 \ \forall z \in \mathcal{K}\}$ denotes the dual cone of $\mathcal{K}$, cf. [1, Lemma 1]. Hence the following result holds, see, e.g., [11, 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 [11, Proposition 3.3].

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

$$P_{\mathcal{K}}(z) = \max\{0, \lambda_1\} u^{(1)} + \max\{0, \lambda_2\} u^{(2)},$$

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

$$\lambda_1 := z_0 - \|\bar{z}\|, \qquad \lambda_2 := z_0 + \|\bar{z}\|,$$

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

where $\bar{w}$ is any vector in $\mathbb{R}^n$ 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 [12, 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 $z = (z_0, \bar{z}) \in \mathbb{R} \times \mathbb{R}^n$ 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.

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

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

where

$$\bar{w} := \frac{\bar{z}}{\|\bar{z}\|}, \quad H := \Big(\frac{z_0}{\|\bar{z}\|} + 1\Big) I_n - \frac{z_0}{\|\bar{z}\|}\bar{w}\bar{w}^T.$$

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

Proof. First assume that $\bar{z} \neq 0$ holds.

Consider the case $z_0 < -\|\bar{z}\|$. This implies not only $z_0 + \|\bar{z}\| < 0$ but also $z_0 - \|\bar{z}\| < 0$. Therefore, in view of Lemma 2.2, it follows that $P_{\mathcal{K}}$ is the zero mapping in a neighborhood of the given point z. Hence the Jacobian at z is the zero matrix.

Next consider the case $z_0 > +\|\bar{z}\|$. Then z belongs to the interior of the cone $\mathcal{K}$ and, therefore, $P_{\mathcal{K}}$ is the identity mapping in a small neighborhood of z (this can also be verified by direct calculation using Lemma 2.2 once again). Hence the Jacobian at z is the identity matrix.

Now let $z_0 \in (-\|\bar{z}\|, +\|\bar{z}\|)$. Then $z_0 + \|\bar{z}\| > 0$ and $z_0 - \|\bar{z}\| < 0$. Consequently, using Lemma 2.2 again, we have (locally)

$$P_{\mathcal{K}}(z) = \frac{1}{2}\big(z_0 + \|\bar{z}\|\big) \begin{pmatrix} 1 \\ \frac{\bar{z}}{\|\bar{z}\|} \end{pmatrix}.$$

Let us write

$$\theta(\bar{z}) := \frac{\bar{z}}{\|\bar{z}\|}.$$

Since $\bar{z} \neq 0$, it is easy to see that θ is differentiable at $\bar{z}$ with Jacobian

$$\theta'(\bar{z}) = \frac{1}{\|\bar{z}\|^2}\Big(\|\bar{z}\| I_n - \frac{1}{\|\bar{z}\|}\bar{z}\bar{z}^T\Big).$$

Hence we obtain

$$\begin{aligned} P'_{\mathcal{K}}(z) &= \frac{1}{2}\big(z_0 + \|\bar{z}\|\big) \begin{pmatrix} 0 & 0 \\ 0 & \theta'(\bar{z}) \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ \frac{\bar{z}}{\|\bar{z}\|} \end{pmatrix} \Big(1, \frac{\bar{z}^T}{\|\bar{z}\|}\Big) \\ &= \frac{1}{2}\big(z_0 + \|\bar{z}\|\big) \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{\|\bar{z}\|^2}\big(\|\bar{z}\| I_n - \frac{1}{\|\bar{z}\|}\bar{z}\bar{z}^T\big) \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 & \frac{\bar{z}^T}{\|\bar{z}\|} \\ \frac{\bar{z}}{\|\bar{z}\|} & \frac{\bar{z}\bar{z}^T}{\|\bar{z}\|^2} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & \frac{\bar{z}^T}{\|\bar{z}\|} \\ \frac{\bar{z}}{\|\bar{z}\|} & \big(\frac{z_0}{\|\bar{z}\|} + 1\big) I_n - \frac{z_0}{\|\bar{z}\|}\frac{\bar{z}\bar{z}^T}{\|\bar{z}\|^2} \end{pmatrix}, \end{aligned}$$

which is precisely the statement in this case.

Finally, let us consider the case where $\bar{z} = 0$. Then the statement is a special case of [11, Proposition 5.2]. However, since $P_{\mathcal{K}}$ is continuously differentiable at $(z_0, 0)$ for any $z_0 \neq 0$ by Lemma 2.4, the Jacobian at $(z_0, \bar{z})$ can also be derived by taking a limit of a sequence $\{z^k\} = \{(z_0^k, \bar{z}^k)\}$ converging to $z = (z_0, 0)$ and satisfying $\bar{z}^k \neq 0$. □

Based on the previous results, we are now in a position to give an expression for the elements of the B-subdifferential $\partial_B P_{\mathcal{K}}(z)$ at an arbitary point z. A similar representation of the elements of the Clarke generalized Jacobian $\partial P_{\mathcal{K}}(z)$ is given in [12, Proposition 4.8], without discussion on conditions for the nonsingularity of those matrices. Here we prefer to deal with the smaller set $\partial_B P_{\mathcal{K}}(z)$ 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 hold under weaker assumptions than the nonsingularity of all elements of the corresponding Clarke generalized Jacobian.

Lemma 2.6 *Given a general point $z = (z_0, \bar{z}) \in \mathbb{R} \times \mathbb{R}^n$, 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

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

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

(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 $V = I_{n+1}$ or V belongs to the set

$$\left\{ \frac{1}{2}\begin{pmatrix} 1 & \bar{w}^T \\ \bar{w} & H \end{pmatrix} \,\middle|\, H = (w_0 + 1)I_n - w_0\bar{w}\bar{w}^T \text{ for some } |w_0| \leq 1 \text{ and } \|\bar{w}\| = 1 \right\}.$$

Proof. Throughout the proof, we denote by D the set of points where the mapping $P_{\mathcal{K}}$ is differentiable. Recall that this set is characterized in Lemma 2.4.

(a) Under the stated assumptions on $(z_0, \bar{z})$, it follows from Lemma 2.4 that $P_{\mathcal{K}}$ is continuously differentiable at z. Hence the B-subdifferential consists of one element only, namely the Jacobian $P'_{\mathcal{K}}(z)$. This observation gives the first statement.

(b) Let $\bar{z} \neq 0$ and $z_0 = +\|\bar{z}\|$. Furthermore, let $\{z^k\} \subseteq D$ be an arbitrary sequence converging to z. Writing $z^k = (z_0^k, \bar{z}^k)$, we can assume without loss of generality that, for each $k \in \mathbb{N}$, we have either $-\|\bar{z}^k\| < z_0^k < \|\bar{z}^k\|$ or $z_0^k > \|\bar{z}^k\|$. If $z_0^k > \|\bar{z}^k\|$ for all $k \in \mathbb{N}$, we have $P'_{\mathcal{K}}(z^k) = I_{n+1}$ by Lemma 2.5 and, therefore, $P'_{\mathcal{K}}(z^k) \to I_{n+1}$ as $k \to \infty$. On the other hand, if $-\|\bar{z}^k\| < z_0^k < \|\bar{z}^k\|$ for all $k \in \mathbb{N}$, Lemma 2.5 shows that

$$P'_{\mathcal{K}}(z^k) = \frac{1}{2}\begin{pmatrix} 1 & (\bar{w}^k)^T \\ \bar{w}^k & H_k \end{pmatrix}$$

for all $k \in \mathbb{N}$, where

$$\bar{w}^k := \frac{\bar{z}^k}{\|\bar{z}^k\|} \quad \text{and} \quad H_k := \Big(\frac{z_0^k}{\|\bar{z}^k\|} + 1\Big) I_n - \frac{z_0^k}{\|\bar{z}^k\|}\bar{w}^k(\bar{w}^k)^T. \tag{1}$$

Taking the limit $k \to \infty$, we therefore obtain

$$P'_{\mathcal{K}}(z^k) \to \frac{1}{2}\begin{pmatrix} 1 & \bar{w}^T \\ \bar{w} & H \end{pmatrix} \quad \text{with } \bar{w} := \frac{\bar{z}}{\|\bar{z}\|}, \;\; H := 2I_n - \bar{w}\bar{w}^T.$$

Finally, if the sequence $\{z^k\}$ is such that $z_0^k > \|\bar{z}^k\|$ for some $k \in \mathbb{N}$ and $-\|\bar{z}^k\| < z_0^k < \|\bar{z}^k\|$ for the remaining $k \in \mathbb{N}$, we do not get any other limiting elements for $P'_{\mathcal{K}}(z^k)$.

(c) Let $\bar{z} \neq 0$ and $z_0 = -\|\bar{z}\|$. Moreover, let $\{z^k\} \subseteq D$ be any sequence converging to z. Writing $z^k = (z_0^k, \bar{z}^k)$ for each $k \in \mathbb{N}$, we may assume without loss of generality that, for each $k \in \mathbb{N}$, we have either $z_0^k < -\|\bar{z}^k\|$ or $-\|\bar{z}^k\| < z_0^k < \|\bar{z}^k\|$. By the same reason as in the proof of part (b), it suffices to consider the two cases: (i) $z_0^k < -\|\bar{z}^k\|$ for all $k \in \mathbb{N}$, and (ii) $-\|\bar{z}^k\| < z_0^k < \|\bar{z}^k\|$ for all $k \in \mathbb{N}$. In the former case, it follows from Lemma 2.5 that $P'_{\mathcal{K}}(z^k) = 0$ for all $k \in \mathbb{N}$, hence $P'_{\mathcal{K}}(z^k) \to 0$ as $k \to \infty$. In the latter case, it also follows from Lemma 2.5 that

$$P'_{\mathcal{K}}(z^k) = \frac{1}{2}\begin{pmatrix} 1 & (\bar{w}^k)^T \\ \bar{w}^k & H_k \end{pmatrix}$$

with $\bar{w}^k, H_k$ given by (1). Taking the limit $k \to \infty$ gives

$$P'_{\mathcal{K}}(z^k) \to \frac{1}{2}\begin{pmatrix} 1 & \bar{w}^T \\ \bar{w} & H \end{pmatrix} \quad \text{with } \bar{w} := \frac{\bar{z}}{\|\bar{z}\|}, \;\; H := \bar{w}\bar{w}^T.$$

(d) Let $\bar{z} = 0$ and $z_0 = 0$. Let $\{z^k\} = \{(z_0^k, \bar{z}^k)\} \subseteq D$ be any sequence converging to $z = (z_0, \bar{z}) = (0, 0)$. Then there are three possibilities: (i) $z_0^k < -\|\bar{z}^k\|$, in which case we have $P'_{\mathcal{K}}(z^k) = 0$ by Lemma 2.5, (ii) $z_0^k > \|\bar{z}^k\|$, in which case we have $P'_{\mathcal{K}}(z^k) = I_{n+1}$ by Lemma 2.5, and (iii) $-\|\bar{z}^k\| < z_0^k < \|\bar{z}^k\|$, in which case we have

$$P'_{\mathcal{K}}(z^k) = \frac{1}{2}\begin{pmatrix} 1 & (\bar{w}^k)^T \\ \bar{w}^k & H_k \end{pmatrix}$$

with $\bar{w}^k, H_k$ given by (1), again by Lemma 2.5. Taking the limit (possibly on a subsequence) in each of these cases, we have either $P'_{\mathcal{K}}(z^k) \to 0$ or $P'_{\mathcal{K}}(z^k) \to I_{n+1}$ or

$$P'_{\mathcal{K}}(z^k) \to \frac{1}{2}\begin{pmatrix} 1 & \bar{w}^T \\ \bar{w} & H \end{pmatrix} \quad \text{with } H := (w_0 + 1)I_n - w_0\bar{w}\bar{w}^T$$

for some vector $w = (w_0, \bar{w})$ satisfying $|w_0| \leq 1$ and $\|\bar{w}\| = 1$. This completes the proof. $\square$

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

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

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

The eigenvalues 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 $z = (z_0, \bar{z}) \in \mathbb{R} \times \mathbb{R}^n$ and $V \in \partial_B P_{\mathcal{K}}(z)$. Assume that $V \notin \{0, I_{n+1}\}$ so that V has the third representation in (2) with $H = (1+\alpha)I_n - \alpha\bar{w}\bar{w}^T$ for some scalar $\alpha \in [-1, +1]$ and some vector $\bar{w} \in \mathbb{R}^n$ satisfying $\|\bar{w}\| = 1$. Then V has the two single eigenvalues $\lambda = 0$ and $\lambda = 1$ as well as the eigenvalue $\lambda = \frac{1}{2}(\alpha + 1)$ with multiplicity $n - 1$ (unless $\alpha = \pm 1$, where the multiplicities change in an obvious way). Moreover, the corresponding eigenvectors of V are given by*

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

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

Proof. By assumption, we have

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

for some $\alpha \in [-1, +1]$ and some vector $\bar{w}$ satisfying $\|\bar{w}\| = 1$. Now take an arbitrary vector $\bar{v} \in \mathbb{R}^n$ orthogonal to $\bar{w}$, and let $u = (0, \bar{v}^T)^T$. Then an elementary calculation shows that $Vu = \lambda u$ holds for $\lambda = \frac{1}{2}(\alpha + 1)$. Hence this λ is an eigenvalue of V with multiplicity $n - 1$ since we can choose $n - 1$ linearly independent vectors $\bar{v} \in \mathbb{R}^n$ such that $\bar{v}^T\bar{w} = 0$. On the other hand, if $\lambda = 0$, it is easy to see that $Vu = \lambda u$ holds with $u = (-1, \bar{w}^T)^T$, whereas for $\lambda = 1$ we have $Vu = \lambda u$ by taking $u = (1, \bar{w}^T)^T$. This completes the proof. □

Note that Lemma 2.7 particularly implies $\lambda \in [0,1]$ for all eigenvalues λ of V. This fact 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

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

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

$$\begin{pmatrix} -1 \\ \frac{\bar{z}}{\|\bar{z}\|} \end{pmatrix}, \quad \begin{pmatrix} 1 \\ \frac{\bar{z}}{\|\bar{z}\|} \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 0 \\ \bar{v}_j \end{pmatrix}, \quad j = 1, \dots, n-1, \tag{3}$$

where $\bar{v}_1, \dots, \bar{v}_{n-1}$ comprise an orthogonal basis of the linear subspace $\{\bar{v} \in \mathbb{R}^n \mid \bar{v}^T \bar{z} = 0\}$. However, an elementary calculation shows that these are also the eigenvectors of the arrow matrix $\mathbf{Arw}(z)$, with corresponding single eigenvalues $\lambda_1 = z_0 - \|\bar{z}\|, \lambda_2 = z_0 + \|\bar{z}\|$ and the multiple eigenvalues $\lambda_i = z_0$, $i = 3, \dots, n+1$. 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 Nonlinear Second-Order Cone Programs

In this section, we consider the nonlinear second-order cone program (nonlinear SOCP for short)

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

where $f : \mathbb{R}^n \to \mathbb{R}$ 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 $\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_r$ is the Cartesian product of second-order cones $\mathcal{K}_i \subseteq \mathbb{R}^{n_i}$ with $n_1 + \cdots + n_r = n$. Here we are particularly interested in the case where the objective function f is nonlinear. The special case where f is a linear function will be discussed in more detail in the subsequent section.

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

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

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

$$M(z) := M(x, \mu, \lambda) := \begin{pmatrix} \nabla f(x) - A^T\mu - \lambda \\ Ax - b \\ x_1 - P_{\mathcal{K}_1}(x_1 - \lambda_1) \\ \vdots \\ x_r - P_{\mathcal{K}_r}(x_r - \lambda_r) \end{pmatrix}. \tag{4}$$

Then we can apply the nonsmooth Newton method [18, 19, 17]

$$z^{k+1} := z^k - W_k^{-1} M(z^k), \ \ W_k \in \partial_B M(z^k), \quad k = 0, 1, 2, \dots, \tag{5}$$

to the system of equations $M(z) = 0$ in order to solve the nonlinear 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 [19, 18], we have to guarantee that, on 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* (4) *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. □

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 $D^a = \operatorname{diag}\big(a_1, \dots, a_n\big)$ and $D^b = \operatorname{diag}\big(b_1, \dots, b_n\big)$ 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*

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

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. Assume that the following two conditions hold:

(a) The matrix H is positive semidefinite on the subspace $\mathcal{S}_\alpha := \{d \in \mathbb{R}^n \,|\, Ad = 0,\ Q_\alpha^T d = 0\}$, and positive definite on the subspace $\mathcal{S}_{\alpha,\beta} := \{d \in \mathbb{R}^n \,|\, Ad = 0,\ Q_\alpha^T d = 0,\ Q_\beta^T d = 0\} = \mathcal{S}_\alpha \cap \{d \in \mathbb{R}^n \,|\, Q_\beta^T d = 0\}$.

(b) The matrix $\big(AQ_\beta, AQ_\gamma\big)$ has full row rank.

Then the matrix

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

is nonsingular.

Proof. Let $y = \big(y^{(1)}, y^{(2)}, y^{(3)}\big) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ be any vector such that $Wy = 0$. Then

$$Hy^{(1)} - A^T y^{(2)} - y^{(3)} = 0, \tag{6}$$

$$Ay^{(1)} = 0, \tag{7}$$

$$V^a y^{(1)} + V^b y^{(3)} = 0. \tag{8}$$

Using the spectral decompositions of V^a and V^b, equation (8) can be rewritten as

$$D^a \tilde{y}^{(1)} + D^b \tilde{y}^{(3)} = 0 \quad \text{with} \quad \tilde{y}^{(1)} := Q^T y^{(1)},\ \tilde{y}^{(3)} := Q^T y^{(3)}.$$

In view of the definitions of the index sets α, β, and γ, this implies

$$\tilde{y}_\alpha^{(1)} = 0, \quad \tilde{y}_\gamma^{(3)} = 0, \quad \text{and} \quad D_\beta^a \tilde{y}_\beta^{(1)} + D_\beta^b \tilde{y}_\beta^{(3)} = 0. \tag{9}$$

This shows that

$$(y^{(1)})^T y^{(3)} = (y^{(1)})^T Q Q^T y^{(3)} = (\tilde{y}^{(1)})^T \tilde{y}^{(3)} = (\tilde{y}_\beta^{(1)})^T \tilde{y}_\beta^{(3)} = -(\tilde{y}_\beta^{(1)})^T \big(D_\beta^b\big)^{-1} D_\beta^a \tilde{y}_\beta^{(1)}.$$

Therefore, premultiplying (6) by $(y^{(1)})^T$, we obtain from (7) that

$$(y^{(1)})^T H y^{(1)} = (y^{(1)})^T y^{(3)} = -(\tilde{y}_\beta^{(1)})^T \big(D_\beta^b\big)^{-1} D_\beta^a \tilde{y}_\beta^{(1)} \leq 0, \tag{10}$$

and the inequality is strict whenever $\tilde{y}_\beta^{(1)} \neq 0$. However, in view of (7) and $0 = \tilde{y}_\alpha^{(1)} = Q_\alpha^T y^{(1)}$, cf. (9), we obtain from the positive semidefiniteness of H on the subspace $\mathcal{S}_\alpha$ that $\tilde{y}_\beta^{(1)} = 0$, i.e., $Q_\beta^T y^{(1)} = 0$. This in turn gives $\tilde{y}_\beta^{(3)} = 0$ by (9). Using (10) and the fact that $y^{(1)} \in \mathcal{S}_{\alpha,\beta}$, we then obtain $y^{(1)} = 0$ from the assumed positive definiteness of H on the subspace $\mathcal{S}_{\alpha,\beta}$. Now premultiplying (6) by Q^T and writing down the corresponding block equations for the β and γ blocks, we obtain from $y^{(1)} = 0, \tilde{y}_\beta^{(3)} = 0$, and $\tilde{y}_\gamma^{(3)} = 0$ that

$$Q_\beta^T A^T y^{(2)} = 0 \quad \text{and} \quad Q_\gamma^T A^T y^{(2)} = 0.$$

Hence $y^{(2)} = 0$ follows from Assumption (b). Finally, by (6), we have $y^{(3)} = 0$. □

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:

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

We also call a KKT point $z^* = (x^*, \mu^*, \lambda^*)$ of the nonlinear SOCP *strictly complementary* if $x_i^* + \lambda_i^* \in \text{int}\mathcal{K}_i$ holds for all block components $i = 1, \ldots, r$. 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 nonlinear SOCP. Then precisely one of the following six cases holds for each block pair (x_i^*, λ_i^*), $i = 1, \ldots, r$:*

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

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 nonlinear 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 $P'_{\mathcal{K}_i}(s_i) = I$.

(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 $s_i := x_i^* - \lambda_i^*$.

Proof. (a) We have $s_i = x_i^* - \lambda_i^* = x_i^* \in \text{int}\mathcal{K}_i$. Hence, locally around s_i, the projection mapping $P_{\mathcal{K}_i}$ is the identity mapping, so that its Jacobian equals the identity matrix.

(b) We have $s_i = x_i^* - \lambda_i^* = -\lambda_i^*$ with $\lambda_i^* \in \text{int}\mathcal{K}_i$. Hence, locally around s_i, the projection $P_{\mathcal{K}_i}$ maps everything onto the zero vector. This implies $P'_{\mathcal{K}_i}(s_i) = 0$.

(c) Write $x_i^* = (x_{i0}^*, \bar{x}_i^*)$, $\lambda_i^* = (\lambda_{i0}^*, \bar{\lambda}_i^*)$. By assumption, we have $x_{i0}^* = \|\bar{x}_i^*\| \neq 0$ and $\lambda_{i0}^* = \|\bar{\lambda}_i^*\| \neq 0$. Let $s_i := x_i^* - \lambda_i^*$ and write $s_i = (s_{i0}, \bar{s}_i)$. In view of Lemma 2.4, it suffices to show that $s_{i0} \neq \pm\|\bar{s}_i\|$. First suppose that $s_{i0} = \|\bar{s}_i\|$. Then $x_{i0}^* - \lambda_{i0}^* = \|\bar{x}_i^* - \bar{\lambda}_i^*\|$. However, since $x_i^* \neq 0$ and $\lambda_i^* \neq 0$, it follows from [1, Lemma 15] that there is a constant $\alpha > 0$ such that $\lambda_{i0}^* = \alpha x_{i0}^*$ and $\bar{\lambda}_i^* = -\alpha \bar{x}_i^*$. Consequently, we obtain $(1-\alpha)x_{i0}^* = (1+\alpha)\|\bar{x}_i^*\| = (1+\alpha)x_{i0}^*$, which implies $\alpha = 0$, a contradiction to $\alpha > 0$. In a similar way, one gets a contradiction for the case $s_{i0} = -\|\bar{s}_i\|$. $\square$

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

$$\begin{aligned} J_I &:= \big\{ i \,\big|\, x_i^* \in \mathrm{int}\mathcal{K}_i, \lambda_i^* = 0 \big\}, \\ J_B &:= \big\{ i \,\big|\, x_i^* \in \mathrm{bd}^+\mathcal{K}_i, \lambda_i^* \in \mathrm{bd}^+\mathcal{K}_i \big\}, \\ J_0 &:= \big\{ i \,\big|\, x_i^* = 0, \lambda_i^* \in \mathrm{int}\mathcal{K}_i \big\} \end{aligned} \tag{11}$$

form a partition of the block indices $i = 1, \ldots, r$. 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

$$V_i = I \;\; \forall i \in J_I \quad \text{and} \quad V_i = 0 \;\; \forall i \in J_0. \tag{12}$$

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 Lemma 2.7 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 assume that the eigenvalues of V_i are ordered in such a way that

$$D_i = \mathrm{diag}\big(0, *, \ldots, *, 1\big) \quad \forall i \in J_B, \tag{13}$$

where $*$ denotes the remaining eigenvalues that lie in the open interval $(0, 1)$. Correspondingly we also partition the orthogonal matrices Q_i as

$$Q_i = \big(q_i, \, \hat{Q}_i, \, q_i'\big) \quad \forall i \in J_B, \tag{14}$$

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:

$$Q_i = \big(q_i, \, \bar{Q}_i\big) = \big(\tilde{Q}_i, q_i'\big) \quad \forall i \in J_B, \tag{15}$$

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 (3), the vectors q_i and q_i' are actually given by

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

(Note that we have $\bar{x}_i^* - \bar{\lambda}_i^* \neq 0$ since $x_i^{*T}\lambda_i^* = 0, x_i^* \in \text{bd}^+\mathcal{K}_i, \lambda_i^* \in \text{bd}^+\mathcal{K}_i$ [1, Lemma 15].)

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 nonlinear SOCP and let the (block) index sets J_I, J_B, J_0 be defined by* (11). *Assume the following conditions:*

(a) The Hessian $\nabla^2 f(x^*)$ is positive semidefinite on the subspace $\mathcal{S} := \{d \in \mathbb{R}^n \mid Ad = 0, d_i = 0 \, \forall i \in J_0, \, q_i^T d_i = 0 \, \forall i \in J_B\}$, and positive definite on the subspace $\hat{\mathcal{S}} := \{d \in \mathbb{R}^n \mid Ad = 0, \, d_i = 0 \, \forall i \in J_0, \, \tilde{Q}_i^T d_i = 0 \, \forall i \in J_B\} = \mathcal{S} \cap \{d \in \mathbb{R}^n \mid \hat{Q}_i^T d_i = 0 \, \forall i \in J_B\}$, where a vector $d \in \mathbb{R}^n$ is partitioned into $d = (d_1, \dots, d_r)^T$ with block components $d_i \in \mathbb{R}^{n_i}$.

(b) The matrix $\big(A_i(i \in J_I), \, A_i\bar{Q}_i(i \in J_B)\big)$ has linearly independent rows, where A_i is the submatrix of A consisting of those columns which correspond to the block index i.

Then the Jacobian $M'(z^*)$ exists and is nonsingular.

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

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

for the block diagonal matrix $V = \text{diag}(V_1, \dots, V_r)$ with $V_i = P'_{\mathcal{K}_i}(x_i^* - \lambda_i^*)$. Therefore, taking into account 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 $V^a := I - V$ and $V^b := V$) as soon as we have identified the index sets $\alpha, \beta, \gamma \subseteq \{1, \dots, n\}$.

For each $i \in J_I$, we have $V_i = I$ (see (12)) and, therefore, $Q_i = I$ and $D_i = I$. 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 (12)), and this corresponds to $Q_i = I$ 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 $D_i = \operatorname{diag}(0, *, \dots, *, 1)$ 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 these considerations into account, we immediately see that our assumptions (a) and (b) correspond precisely to the assumptions (a) and (b) in Proposition 3.2. □

We now want to extend Theorem 3.5 to the case where strict complementarity is violated. Let $z^* = (x^*, \mu^*, \lambda^*)$ be an arbitary KKT point of the nonlinear SOCP, and let J_I, J_B, J_0 denote the index sets defined by (11). In view of Lemma 3.3, in addition to these sets, we also need to consider the three index sets

$$\begin{aligned} J_{B0} &:= \{ i \mid x_i^* \in \operatorname{bd}^+ \mathcal{K}_i, \lambda_i^* = 0 \}, \\ J_{0B} &:= \{ i \mid x_i^* = 0, \lambda_i^* \in \operatorname{bd}^+ \mathcal{K}_i \}, \\ J_{00} &:= \{ i \mid x_i^* = 0, \lambda_i^* = 0 \}, \end{aligned} \tag{16}$$

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 a counterpart of Lemma 3.4 in the general case.

Lemma 3.6 *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 $V_i = I_{n+1}$ or $V_i = Q_i D_i Q_i^T$ with $D_i = \operatorname{diag}(0, 1, \dots, 1)$ 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 $D_i = \operatorname{diag}(0, \dots, 0, 1)$ and $Q_i = (\tilde{Q}_i, q_i')$.

(c) If $i \in J_{00}$, then we have $V_i = I_{n+1}$ or $V_i = 0$ or $V_i = Q_i D_i Q_i^T$ with D_i and Q_i given by (13) and (14), respectively, or by $D_i = \operatorname{diag}(0, 1, \dots, 1)$ and $Q_i = (q_i, \bar{Q}_i)$, or by $D_i = \operatorname{diag}(0, \dots, 0, 1)$ 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 w_0 in Lemma 2.6 (d) (namely $w_0 \in (-1, 1), w_0 = 1$, and $w_0 = -1$). □

The previous result 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 (13)–(15) for all $i \in J_B$, as well as those specified in Lemma 3.6 for all indices $i \in J_{B0} \cup J_{0B} \cup J_{00}$.

Theorem 3.7 *Let $z^* = (x^*, \mu^*, \lambda^*)$ be a (not necessarily strictly complementary) KKT point of the nonlinear SOCP and let the (block) index sets $J_I, J_B, J_0, J_{B0}, J_{0B}, J_{00}$ be defined by (11) and (16). Suppose that 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 hold:*

(a) The Hessian $\nabla^2 f(x^*)$ is positive semidefinite on the subspace

$$
\begin{aligned}
\mathcal{S} \;:=\; \big\{ d \in \mathbb{R}^n \,\big|\, & Ad = 0, \\
& d_i = 0 \quad \forall i \in J_0 \cup J_{0B}^1 \cup J_{00}^2, \\
& q_i^T d_i = 0 \quad \forall i \in J_B \cup J_{B0}^2 \cup J_{00}^3 \cup J_{00}^4, \\
& \tilde{Q}_i^T d_i = 0 \quad \forall i \in J_{0B}^2 \cup J_{00}^5 \big\},
\end{aligned}
$$

and positive definite on the subspace

$$
\begin{aligned}
\hat{\mathcal{S}} \;:=\; \big\{ d \in \mathbb{R}^n \,\big|\, & Ad = 0, \\
& d_i = 0 \quad \forall i \in J_0 \cup J_{0B}^1 \cup J_{00}^2, \\
& q_i^T d_i = 0 \quad \forall i \in J_{B0}^2 \cup J_{00}^4, \\
& \tilde{Q}_i^T d_i = 0 \quad \forall i \in J_B \cup J_{0B}^2 \cup J_{00}^3 \cup J_{00}^5 \big\} \\
= \;\; \mathcal{S} \cap \big\{ d \in \mathbb{R}^n \,\big|\, & \hat{Q}_i^T d_i = 0 \quad \forall i \in J_B \cup J_{00}^3 \big\}.
\end{aligned}
$$

(b) The matrix

$$
\big(A_i \, (i \in J_I \cup J_{B0}^1 \cup J_{00}^1), \;\; A_i \bar{Q}_i \, (i \in J_B \cup J_{B0}^2 \cup J_{00}^3 \cup J_{00}^4), \;\; A_i q_i' \, (i \in J_{0B}^2 \cup J_{00}^5) \big)
$$

has linearly independent rows.

Then all matrices $W \in \partial_B M(z^*)$ are nonsingular.

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

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

for a suitable block diagonal matrix $V = \operatorname{diag}(V_1, \dots, V_r)$ with $V_i \in \partial_B P_{\mathcal{K}_i}(x_i^* - \lambda_i^*)$. In principle, the proof is now 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. 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 new index sets J_{B0}, J_{0B}, and J_{00}. In view of Lemma 3.6, 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

$$\begin{aligned} &J_{B0}^1 := \{i \,|\, V_i = I_{n+1}\}, \quad J_{B0}^2 := J_{B0} \setminus J_{B0}^1, \\ &J_{0B}^1 := \{i \,|\, V_i = 0\}, \qquad\quad J_{0B}^2 := J_{0B} \setminus J_{0B}^1 \end{aligned}$$

and

$$\begin{aligned} J_{00}^1 &:= \{i \,|\, V_i = I_{n+1}\}, \\ J_{00}^2 &:= \{i \,|\, V_i = 0\}, \\ J_{00}^3 &:= \{i \,|\, V_i = Q_i D_i Q_i^T \text{ with } D_i \text{ and } Q_i \text{ given by (13) and (14), respectively}\}, \\ J_{00}^4 &:= \{i \,|\, V_i = Q_i D_i Q_i^T \text{ with } D_i = \operatorname{diag}(0, 1, \dots, 1) \text{ and } Q_i = (q_i, \bar{Q}_i)\}, \\ J_{00}^5 &:= \{i \,|\, V_i = Q_i D_i Q_i^T \text{ with } D_i = \operatorname{diag}(0, \dots, 0, 1) \text{ and } Q_i = (\tilde{Q}_i, q_i')\}. \end{aligned}$$

Using these definitions and Lemmas 3.4 and 3.6, 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 $Q_i = I$ 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 3.6, 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 3.6 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 $Q_i = I$.
- 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.

From these observations, it follows that the condition $Q_\alpha^T d = 0$ is equivalent to

$$\begin{aligned} d_i &= 0 \quad \forall i \in J_0 \cup J_{0B}^1 \cup J_{00}^2, \\ q_i^T d_i &= 0 \quad \forall i \in J_B \cup J_{B0}^2 \cup J_{00}^3 \cup J_{00}^4, \\ \tilde{Q}_i^T d_i &= 0 \quad \forall i \in J_{0B}^2 \cup J_{00}^5. \end{aligned}$$

Furthermore, the requirement $Q_\beta^T d = 0$ can be rewritten as

$$\hat{Q}_i^T d_i = 0 \quad \forall i \in J_B \cup J_{00}^3.$$

Since $\tilde{Q}_i = (q_i, \hat{Q}_i) \in \mathbb{R}^{n_i \times (n_i - 1)}$, the two conditions $Q_\alpha^T d = 0$ and $Q_\beta^T d = 0$ together become

$$\begin{aligned} d_i &= 0 \quad \forall i \in J_0 \cup J_{0B}^1 \cup J_{00}^2, \\ q_i^T d_i &= 0 \quad \forall i \in J_{B0}^2 \cup J_{00}^4, \\ \tilde{Q}_i^T d_i &= 0 \quad \forall i \in J_B \cup J_{0B}^2 \cup J_{00}^3 \cup J_{00}^5. \end{aligned}$$

This shows that assumption (a) of this theorem coincides with the corresponding assumption in Proposition 3.2.

In a similar way, we can verify the equivalence between assumption (b) in this theorem and the corresponding condition in Proposition 3.2. In fact, the above identifications of the index sets β and γ show that the matrix $\big(AQ_\beta,\, AQ_\gamma\big)$ has linearly independent rows if and only if the matrix

$$\big(A_i\hat{Q}_i\,(i\in J_B\cup J_{00}^3),\;\; A_i\,(i\in J_I\cup J_{B0}^1\cup J_{00}^1),\;\; A_iq_i'\,(i\in J_B\cup J_{0B}^2\cup J_{00}^3\cup J_{00}^5),\;\; A_i\bar{Q}_i\,(i\in J_{B0}^2\cup J_{00}^4)\big)$$

has linearly independent rows. Since $\bar{Q}_i = (\hat{Q}_i, q_i') \in \mathbb{R}^{n_i\times(n_i-1)}$, it is then easy to see that this condition is equivalent to our assumption (b). Hence the assertion follows from Proposition 3.2. □

Note that, in the case of a strictly complementary KKT point, Theorem 3.7 reduces to Theorem 3.5. Using Theorems 3.1 and 3.7 along with [18], we get the following result.

Theorem 3.8 *Let $z^* = (x^*, \mu^*, \lambda^*)$ be a (not necessarily strictly complementary) KKT point of the nonlinear SOCP, and suppose that the assumptions of Theorem 3.7 hold at this KKT point. Then the nonsmooth Newton method* (5) *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 Linear Second-Order Cone Programs

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

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

where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$ are the given data, and $\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_r$ is the Cartesian product of second-order cones $\mathcal{K}_i \subseteq \mathbb{R}^{n_i}$ with $n_1 + \cdots + n_r = n$. Clearly the linear SOCP is a special case of the nonlinear SOCP discussed in the previous section. In particular, we may rewrite the corresponding KKT conditions as a system of equations

$M_0(z) = 0$, where $M_0 : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ is defined by

$$M_0(z) := M_0(x, \mu, \lambda) := \begin{pmatrix} c - A^T\mu - \lambda \\ Ax - b \\ x_1 - P_{\mathcal{K}_1}(x_1 - \lambda_1) \\ \vdots \\ x_r - P_{\mathcal{K}_r}(x_r - \lambda_r) \end{pmatrix}. \tag{17}$$

An immediate consequence of Theorem 3.1 is the fact that the mapping M_0 is strongly semismooth. On the other hand, Theorems 3.5 and 3.7 do not apply to the linear SOCP, since the second-order conditions do not hold. We therefore need another suitable condition that guarantees the required nonsingularity of the (generalized) Jacobian of M_0 at a KKT point $z^* = (x^*, \mu^*, \lambda^*)$. Before presenting such a result, we follow the presentation of the previous section and first state a general proposition.

Proposition 4.1 *Let $A \in \mathbb{R}^{m\times n}$. 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 $D^a = \operatorname{diag}(a_1, \dots, a_n)$, $D^b = \operatorname{diag}(b_1, \dots, b_n)$ satisfying $V^a = QD^aQ^T$, $V^b = QD^bQ^T$ and $a_j \geq 0, b_j \geq 0$, $a_j + b_j > 0$ for all $j = 1, \dots, n$. Let the index set $\{1, \dots, n\}$ be partitioned as $\{1, \dots, n\} = \alpha \cup \beta \cup \gamma$ with*

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

Assume that the following two conditions hold:

(a) The matrix AQ_γ has full column rank.

(b) The matrix $\big(AQ_\beta, AQ_\gamma\big)$ has full row rank.

Then the matrix

$$W_0 := \begin{pmatrix} 0 & -A^T & -I \\ A & 0 & 0 \\ V^a & 0 & V^b \end{pmatrix}$$

is nonsingular.

Proof. Let $y = \big(y^{(1)}, y^{(2)}, y^{(3)}\big) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ be an arbitrary vector such that $W_0 y = 0$. Then

$$\begin{aligned} A^T y^{(2)} + y^{(3)} &= 0, \qquad (18)\\ A y^{(1)} &= 0, \qquad (19)\\ V^a y^{(1)} + V^b y^{(3)} &= 0. \qquad (20) \end{aligned}$$

Using the spectral decompositions of V^a and V^b, equation (20) can be rewritten as

$$D^a \tilde{y}^{(1)} + D^b \tilde{y}^{(3)} = 0 \quad \text{with} \quad \tilde{y}^{(1)} := Q^T y^{(1)},\ \tilde{y}^{(3)} := Q^T y^{(3)}.$$

Taking into account the definitions of the index sets α, β, and γ, we obtain

$$\tilde{y}^{(1)}_\alpha = 0, \quad \tilde{y}^{(3)}_\gamma = 0, \quad \text{and} \quad D^a_\beta \tilde{y}^{(1)}_\beta + D^b_\beta \tilde{y}^{(3)}_\beta = 0. \qquad (21)$$

Premultiplying (18) by Q^T, using the definitions of $\tilde{y}^{(1)}, \tilde{y}^{(3)}$, and writing down the resulting equations componentwise for each block corresponding to the indices in α, β, and γ, we get

$$\begin{aligned} Q^T_\alpha A^T y^{(2)} + \tilde{y}^{(3)}_\alpha &= 0, \qquad (22)\\ Q^T_\beta A^T y^{(2)} + \tilde{y}^{(3)}_\beta &= 0, \qquad (23)\\ Q^T_\gamma A^T y^{(2)} &= 0. \qquad (24) \end{aligned}$$

Furthermore, equation (19) can be written as

$$0 = A y^{(1)} = A Q Q^T y^{(1)} = A Q \tilde{y}^{(1)} = A Q_\beta \tilde{y}^{(1)}_\beta + A Q_\gamma \tilde{y}^{(1)}_\gamma. \qquad (25)$$

Using the nonsingularity of the diagonal submatrices D^a_β, D^b_β, we obtain from (21) and (23) that

$$\tilde{y}^{(1)}_\beta = -(D^a_\beta)^{-1} D^b_\beta \tilde{y}^{(3)}_\beta = (D^a_\beta)^{-1} D^b_\beta Q^T_\beta A^T y^{(2)}.$$

Substituting this expression for $\tilde{q}^{(1)}_\beta$ in (25) yields

$$A Q_\beta (D^a_\beta)^{-1} D^b_\beta Q^T_\beta A^T y^{(2)} + A Q_\gamma \tilde{y}^{(1)}_\gamma = 0.$$

Premultiplying by $(y^{(2)})^T$ and using (24), we have

$$(y^{(2)})^T A Q_\beta (D^a_\beta)^{-1} D^b_\beta Q^T_\beta A^T y^{(2)} = 0.$$

However, by the definition of the index set β, the matrix $(D^a_\beta)^{-1} D^b_\beta$ is positive definite. Consequently, we obtain

$$Q^T_\beta A^T y^{(2)} = 0. \qquad (26)$$

Then (23) implies $\tilde{y}^{(3)}_\beta = 0$, and (21) yield $\tilde{y}^{(1)}_\beta = 0$. This, in turn, gives $AQ_\gamma \tilde{y}^{(1)}_\gamma = 0$ by (25). Assumption (a) then gives $\tilde{y}^{(1)}_\gamma = 0$. Furthermore, (26) and (24) together with assumption (b) yields $y^{(2)} = 0$, which also implies $\tilde{y}^{(3)}_\alpha = 0$ by (22). Thus we have $\tilde{y}^{(1)} = 0$ and $\tilde{y}^{(3)} = 0$. Then it is easy to deduce that $y = \big(y^{(1)}, y^{(2)}, y^{(3)}\big) = (0,0,0)$, i.e., W_0 is nonsingular. □

We now come back to the linear SOCP. First consider the case where $z^* = (x^*, \mu^*, \lambda^*)$ is a strictly complementary KKT point, and use the index sets J_I, J_B, and J_0 defined by (11) in order to partition any given vector in $\mathbb{R}^n$ in a suitable way. Let $V_i := P'_{\mathcal{K}_i}(x^*_i - \lambda^*_i)$. Similarly to the discussion in the previous section, we have

$$V_i = I \ \ \forall i \in J_I \quad \text{and} \quad V_i = 0 \ \ \forall i \in J_0. \tag{27}$$

Moreover, for each $i \in J_B$, let $V_i = Q_i D_i Q_i^T$ be the spectral decomposition of the matrix V_i. For the matrices D_i and $Q_i, i \in J_B$, we use the same decomposition as in (13)–(15).

Using this notation, we are now in a position to state the following nonsingularity result which may be viewed as a counterpart of the corresponding result for interior-point methods given in [1, Theorem 28]. Recall that A_i denotes the submatrix of A consisting of those columns which correspond to the block index i.

Theorem 4.2 *Let $z^* = (x^*, \mu^*, \lambda^*)$ be a strictly complementary KKT point of the linear SOCP. Suppose further that the following conditions hold:*

(a) The matrix $\big(A_i\,(i \in J_I),\ \ A_i q'_i\,(i \in J_B)\big)$ has linearly independent columns.

(b) The matrix $\big(A_i\,(i \in J_I),\ \ A_i \bar{Q}_i\,(i \in J_B)\big)$ has linearly independent rows.

Then the Jacobian $M'_0(z^*)$ exists and is nonsingular.

Proof. We have

$$M'_0(z^*) = \begin{pmatrix} 0 & -A^T & -I \\ A & 0 & 0 \\ I - V & 0 & V \end{pmatrix}$$

for a block diagonal matrix $V = \text{diag}\big(V_1, \dots, V_r\big)$ with $V_i := P'_{\mathcal{K}_i}(x^*_i - \lambda^*_i)$. Let $V^a := I - V$ and $V^b := V$. It then follows that V^a, V^b are both symmetric positive semidefinite with a positive definite sum $V^a + V^b$, and these two matrices obviously have a common basis of eigenvectors. Hence we are in the situation of Proposition 4.1. In order to apply this result, we need to take a closer look at the orthogonal matrix and the special structure of the eigenvalues in the case that is currently under consideration.

We have

$$V^a = Q(I - D)Q^T \quad \text{and} \quad V^b = QDQ^T$$

with the block matrices

$$Q := \operatorname{diag}\big(Q_1, \ldots, Q_r\big) \quad \text{and} \quad D := \operatorname{diag}\big(D_1, \ldots, D_r\big),$$

where the orthogonal blocks Q_i are identity matrices for all indices $i \in J_I \cup J_0$, cf. (27), whereas for all $i \in J_B$, these orthogonal matrices are given by (14), or (15), and the corresponding diagonal matrices D_i are given by (13).

In order to apply Proposition 4.1, we need to identify the index sets β and γ. From (27) and (13), we see that the following indices comprise the index set β:

- all middle indices j belonging to a block index $i \in J_B$, with $\hat{Q}_i$ consisting of the middle $n_i - 2$ columns of the corresponding orthogonal matrix Q_i.

Similarly, we see that the following indices comprise the index set γ:

- all indices j belonging to a block index $i \in J_I$, with $Q_i = I$ being the corresponding orthogonal matrix.
- the last component j of each block index $i \in J_B$, with q_i' being the last column of the corresponding orthogonal matrix Q_i.

Using these identifications, it is easy to see that assumptions (a) and (b) in Proposition 4.1 are identical to the two conditions in this theorem. Consequently the assertion follows from Proposition 4.1. □

We next generalize Theorem 4.2 to the case where strict complementarity is not satisfied. In addition to J_I, J_B, J_0, we use the index sets J_{B0}, J_{0B}, J_{00} defined by (16).

Theorem 4.3 *Let $z^* = (x^*, \mu^*, \lambda^*)$ be a (not necessarily strictly complementary) KKT point of the linear SOCP. Suppose that for any partitioning $J_{B0} = J_{B0}^1 \cup J_{B0}^2$, any subset $J_{0B}^2 \subseteq J_{0B}$, and any mutually disjoint subsets $J_{00}^1, J_{00}^3, J_{00}^4, J_{00}^5 \subseteq J_{00}$, the following two conditions hold:*

(a) The matrix

$$\big(A_i\,(i\in J_I\cup J_{B0}^1\cup J_{00}^1),\;\;A\bar{Q}_i\,(i\in J_{B0}^2\cup J_{00}^4),\;\;A_iq_i'\,(i\in J_B\cup J_{0B}^2\cup J_{00}^3\cup J_{00}^5)\big)$$

has linearly independent columns.

(b) The matrix

$$\big(A_i\,(i\in J_I\cup J_{B0}^1\cup J_{00}^1),\;\;A_i\bar{Q}_i\,(i\in J_B\cup J_{B0}^2\cup J_{00}^3\cup J_{00}^4),\;\;A_iq_i'\,(i\in J_{0B}^2\cup J_{00}^5)\big)$$

has linearly independent rows.

Then all matrices $W_0\in\partial_B M_0(z^*)$ are nonsingular.

Proof. Choose $W_0\in\partial_B M_0(z^*)$ arbitrarily. Then

$$W_0=\begin{pmatrix}0&-A^T&-I\\A&0&0\\I-V&0&V\end{pmatrix}$$

for a block diagonal matrix $V=\operatorname{diag}\big(V_1,\dots,V_r\big)$ with $V_i\in\partial_B P_{\mathcal{K}_i}(x_i^*-\lambda_i^*)$. Following the proof of Theorem 4.2, we define $V^a:=I-V$ and $V^b:=V$ and see that these two matrices satisfy the assumptions of Proposition 4.1. Similarly to the proof of Theorem 4.2, we also have

$$V^a=Q(I-D)Q^T\quad\text{and}\quad V^b=QDQ^T$$

with the block matrices

$$Q:=\operatorname{diag}\big(Q_1,\dots,Q_r\big)\quad\text{and}\quad D:=\operatorname{diag}\big(D_1,\dots,D_r\big).$$

Then we need to identify the index sets α,β and γ in Proposition 4.1. To this end, we follow the proof of Theorem 3.7 and partition the index sets J_{B0}, J_{0B}, and J_{00} further into certain subsets in a similar manner. Then, as in the proof of Theorem 3.7, we are able to show which indices j belong to the sets α,β, and γ. Moreover, it is not difficult to see that the two assumptions (a) and (b) in this theorem correspond precisely to the two requirements (a) and (b) in Proposition 4.1. We leave the details to the reader (note that assumption (b) coincides with the corresponding assumption in Theorem 3.7, so there is nothing to verify for this part). □

Note that a corresponding result for interior-point methods does not hold since the Jacobian matrices arising in that context are singular whenever strict complementarity does not hold.

As a consequence of Theorem 4.3 and [19], we obtain the following theorem.

Theorem 4.4 *Let $z^* = (x^*, \mu^*, \lambda^*)$ be a (not necessarily strictly complementary) KKT point of the linear SOCP, and suppose that the assumptions of Theorem 4.3 hold at this KKT point. Then the nonsmooth Newton method $z^{k+1} := z^k - W_k^{-1} M_0(z^k)$, with $W_k \in \partial_B M_0(z^k)$, applied to the system of equations $M_0(z) = 0$ is locally quadratically convergent.*

5 Final Remarks

We have investigated the local properties of a semismooth equation reformulation of both the linear and the nonlinear second-order cone programs. 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.

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. Chen, 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. Fukushima, Z.-Q. Luo, and P. Tseng: *Smoothing functions for second-order-cone complementarity problems.* SIAM Journal on Optimization 12, 2001, pp. 436–460.

[12] 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.

[13] C. Helmberg: `http://www-user.tu-chemnitz.de/~helmberg/semidef.html`.

[14] H. Kato and M. Fukushima: *An SQP-type algorithm for nonlinear second-order cone programs.* Technical Report 2006-004, Department of Applied Mathematics and Physics, Kyoto University, March 2006.

[15] 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.

[16] 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.

[17] J.-S. Pang and L. Qi: *Nonsmooth equations: motivation and algorithms.* SIAM Journal on Optimization 3, 1993, pp. 443–465.

[18] L. Qi: *Convergence analysis of some algorithms for solving nonsmooth equations.* Mathematics of Operations Research 18, 1993, pp. 227–244.

[19] L. Qi and J. Sun: *A nonsmooth version of Newton's method.* Mathematical Programming 58, 1993, pp. 353–367.

[20] 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/`

[21] D. Sun and J. Sun: *Strong semismoothness of the Fischer-Burmeister SDC and SOC complementarity functions.* Mathematical Programming 103, 2005, pp. 575–581.

[22] P. Tseng: *Smoothing methods for second-order cone programs/complementarity problems.* Talk presented at the SIAM Conference on Optimization, Stockholm, May 2005.

[23] 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`

[24] R.J. Vanderbei and H. Yuritan: *Using LOQO to solve second-order cone programming problems.* Technical Report, Statistics and Operations Research, Princeton University, 1998.

[25] 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).