A SENSITIVITY RESULT FOR QUADRATIC SECOND-ORDER CONE PROGRAMMING AND ITS APPLICATION

Qi Zhao, Zhenjiang, Wenhao Fu, Zhongwen Chen, Suzhou

Received October 19, 2019. Published online December 14, 2020.

Abstract. In this paper, we present a sensitivity result for quadratic second-order cone programming under the weak form of second-order sufficient condition. Based on this result, we analyze the local convergence of an SQP-type method for nonlinear second-order cone programming. The subproblems of this method at each iteration are quadratic second-order cone programming problems. Compared with the local convergence analysis done before, we do not need the assumption that the Hessian matrix of the Lagrangian function is positive definite. Besides, the iteration sequence which is proved to be superlinearly convergent does not contain the Lagrangian multiplier.

Keywords: sensitivity; quadratic second-order cone programming; nonlinear second-order cone programming; local convergence

MSC 2020: 90C22, 90C31, 90C20

1. Introduction

In this paper, we denote a vector z as $(z_{(1)}; \overline{z})$, where $z_{(1)}$ is the first entry of z and $\overline{z}$ is the subvector that consists of the remaining entries. The m_i-dimensional second-order cone $\mathcal{K}_i$ $(i = 1, 2, \ldots, l)$ is defined as

$$\mathcal{K}_i = \begin{cases} \{z \in \mathbb{R} \mid z \geqslant 0\} & \text{if } m_i = 1, \\ \{(z_{(1)}; \overline{z}) \in \mathbb{R} \times \mathbb{R}^{m_i - 1} \mid z_{(1)} \geqslant \|\overline{z}\|\} & \text{if } m_i \geqslant 2, \end{cases}$$

where $\|\cdot\|$ denotes the Euclidean norm, $m_1 + m_2 + \ldots + m_l = m$.

The work was supported by Chinese NSF grant 11871362 and Overseas Study Fund and Start-up Fund for doctoral research by Jiangsu University of Science and Technology.

DOI: 10.21136/AM.2020.0278-19

The following problem is a nonlinear second-order cone programming (NSOCP) problem:

$$
\begin{aligned}
&\min_{x\in\mathbb{R}^n} f(x) \\
&\text{ s.t. } h(x)=0, \\
&\qquad g_i(x)\in\mathcal{K}_i,\ \mathcal{K}_i\subset\mathbb{R}^{m_i},\quad i=1,2,\ldots,l,
\end{aligned}
\tag{1.1}
$$

where $f\colon \mathbb{R}^n\to\mathbb{R}$ and $h\colon \mathbb{R}^n\to\mathbb{R}^p$, $g_i\colon \mathbb{R}^n\to\mathbb{R}^{m_i}$ are smooth functions. The NSOCP problem is a particular case of nonlinear semidefinite programming (see [1]) and has many applications (see [10]). Theoretical properties of NSOCP have been studied in [2], [3], [6]. Kato and Fukushima [9] proposed an SQP-type algorithm for (1.1). At each iteration, the algorithm solves a subproblem in which the constraints involve linear approximations of the constraint functions in the original problem and the objective function is a convex quadratic function, i.e.

$$
\begin{aligned}
&\min_{d\in\mathbb{R}^n} \nabla f(x_k)^\top d+\frac{1}{2}d^\top B_k d \\
&\text{ s.t. } h(x_k)+Dh(x_k)d=0, \\
&\qquad g_i(x_k)+Dg_i(x_k)d\in\mathcal{K}_i,\quad i=1,2,\ldots,l,
\end{aligned}
\tag{1.2}
$$

where x_k is the current iteration point, $\nabla f(x)$ is the gradient of $f(x)$, $Dh(x)$ and $Dg_i(x)$ are the Jacobian matrices of $h(x)$ and $g_i(x)$, respectively. The matrix B_k is symmetric and positive definite containing the second-order information of problem (1.1). The solution d_k (if it exists) is defined as a search direction, its corresponding Lagrangian multiplier is $(\lambda_{k+1},\mu_{k+1})\in\mathbb{R}^p\times\mathbb{R}^m$,

$$
\mu_{k+1}=(\mu_{k+1,1};\mu_{k+1,2};\ldots;\mu_{k+1,i};\ldots;\mu_{k+1,l}),\quad \mu_{k+1,i}\in\mathbb{R}^{m_i}.
$$

The trial step is $x_{k+1}=x_k+\alpha d_k$ and the l_1-penalty function is used as a merit function to determine the step size α. Problem (1.2) has the form of the quadratic second-order cone programming (QSOCP) below:

$$
\begin{aligned}
&\min_{d\in\mathbb{R}^n} b^\top d+\frac{1}{2}d^\top Hd \\
&\text{ s.t. } Cd+c=0, \\
&\qquad A_id+a_i\in\mathcal{K}_i,\quad i=1,2,\ldots,l,
\end{aligned}
\tag{1.3}
$$

where $C\in\mathbb{R}^{p\times n}$, $c\in\mathbb{R}^p$, $A_i\in\mathbb{R}^{m_i\times n}$, $a_i\in\mathbb{R}^{m_i}$, $b\in\mathbb{R}^n$, $H\in\mathbb{R}^{n\times n}$ (for a general QSOCP, H may not be positive definite). Problem (1.3) is a generalization of the quadratic programming (QP) in nonlinear programming (NLP).

Kanzow et al. [8] discuss the local convergence of semismooth Newton methods for linear and nonlinear second-order cone programs without strict complementarity. The local convergence of the algorithm presented in [9] is analyzed by Wang et al. [14], where B_k is taken as the Hessian matrix, and the local superlinearly convergent sequence contains the Lagrangian multiplier. However, for nonlinear programming, it has been proved that the iteration sequence without the Lagrangian multiplier is superlinearly convergent [8] by introducing the projection matrix and the active set. Motivated by this fact, we aim to prove a similar result for (1.1). To this end, we first present a sensitivity result for certain local optimal solutions of the general, possibly nonconvex QSOCP. We then apply this result to show that the sequence $\{x_k\}$ generated by the algorithm in [9] converges to its local solution x^* with superlinear convergence rate under the nondegeneracy, strict complementarity and the weak second-order sufficient conditions.

The paper is organized as follows. In Section 2, we introduce some notations and preliminaries. In Section 3, we give the sensitivity result for (1.3). We apply this result to (1.1) to analyse local convergence in Section 4 and conclude with final remarks in Section 5.

2. Notations and preliminaries

For any vectors $a, b \in \mathbb{R}^{\widehat{m}}$, let $a = (a_{(1)}; \bar{a}) \in \mathbb{R} \times \mathbb{R}^{\widehat{m}-1}$, $b = (b_{(1)}; \bar{b}) \in \mathbb{R} \times \mathbb{R}^{\widehat{m}-1}$, their Jordan product (see [1], Section 4) is given by

$$a \circ b = (a^\top b; a_{(1)}\bar{b} + b_{(1)}\bar{a}).$$

For a closed convex and pointed cone $\mathcal{K}$, its interior, boundary, and boundary excluding the origin are denoted by $\operatorname{int}(\mathcal{K})$, $\operatorname{bd}(\mathcal{K})$, and $\operatorname{bd}^+(\mathcal{K})$, respectively. Denote by $O(t_k)$ a sequence $\{v_k\}$ satisfying $\|v_k\| \leqslant \beta t_k$ for some constant β independent of k, and by $o(t_k)$ a sequence $\{v_k\}$ satisfying $\|v_k\| \leqslant \beta_k t_k$ for some positive sequence $\{\beta_k\}$ with $\lim\limits_{k\to\infty} \beta_k = 0$.

We introduce some useful properties which will be used later. The first and second statement of Proposition 2.1 can be easily proved by the definition of the second-order cone, Jordan product and Cauchy-Schwarz inequality, so the proof is omitted.

Proposition 2.1. *If $\mathcal{K}$ is a second-order cone, then the following results hold.*

(1) *If $a, b \in \mathcal{K}$ and $a \circ b = 0$, then one of the following three cases occurs:*
 (i) $a = 0$,
 (ii) $b = 0$,
 (iii) $a, b \in \operatorname{bd}^+(\mathcal{K})$, *and* $(a_{(1)}; \bar{a}) = \kappa(b_{(1)}; -\bar{b})$, *where* $\kappa > 0$ *is a constant.*

(2) *If* $a \circ b = 0$ *and* $a \in \mathrm{bd}^+(\mathcal{K})$, *then there exists a constant* κ *such that* $b = \kappa(a_{(1)}; -\bar{a})$.

(3) *If* $a + b \in \mathrm{int}(\mathcal{K})$ *and* $a \circ b = 0$, *then* $a, b \in \mathcal{K}$.

(4) *For* $a = (a_{(1)}; \bar{a}) \in \mathrm{bd}^+(\mathcal{K}) \subseteq \mathbb{R}^{\widehat{m}}$, $b = \kappa(a_{(1)}; -\bar{a})$, $\kappa > 0$, *if there exist* $c, d \in \mathbb{R}^{\widehat{m}}$ *such that* $a \circ c + b \circ d = 0$, *then we have that*

$$a^\top c = 0, \quad b^\top d = 0, \quad c_{(1)}^2 - \|\bar{c}\|^2 = \kappa c^\top d.$$

(5) *If* $a \in \mathrm{bd}^+(\mathcal{K})$, $b^\top a = 0$, *then* $b_{(1)}^2 - \|\bar{b}\|^2 \leqslant 0$;

(6) *If* $a \circ b = 0$ *and* $a \in \mathrm{int}(\mathcal{K})$, *then* $b = 0$.

Proof. (3) By $a + b \in \mathrm{int}(\mathcal{K})$, we have that $a_{(1)} + b_{(1)} > 0$ and $(a_{(1)} + b_{(1)})^2 > \|\bar{a} + \bar{b}\|^2$. We can assume, without loss of generality, that $a_{(1)} > 0$. By $a \circ b = 0$, we have that $a_{(1)}\bar{b} + b_{(1)}\bar{a} = 0$, $a_{(1)}b_{(1)} + \bar{a}^\top\bar{b} = 0$, whence it follows that

$$b_{(1)}(a_{(1)}^2 - \bar{a}^\top\bar{a}) = 0.$$

Next we consider two cases.

If $b_{(1)} = 0$, then $\bar{b} = -b_{(1)}\bar{a}/a_{(1)} = 0$, $a \in \mathrm{int}(\mathcal{K})$. If $b_{(1)} \neq 0$, then $a_{(1)}^2 - \|\bar{a}\|^2 = 0$, so $a \in \mathrm{bd}^+(\mathcal{K})$. By Proposition 2.1(2), we have $|b_{(1)}| = \|\bar{b}\|$. It follows from

$$(a_{(1)} + b_{(1)})^2 > \|\bar{a} + \bar{b}\|^2 \quad \text{and} \quad a_{(1)}b_{(1)} + \bar{a}^\top\bar{b} = 0$$

that

$$(a_{(1)} + b_{(1)})^2 - \|\bar{a} + \bar{b}\|^2 = a_{(1)}^2 + 2a_{(1)}b_{(1)} + b_{(1)}^2 - \|\bar{a}\|^2 - \|\bar{b}\|^2 - 2\bar{a}^\top\bar{b} = 4a_{(1)}b_{(1)} > 0,$$

which implies that $b_{(1)} > 0$, so $b \in \mathrm{bd}^+(\mathcal{K})$. In both cases we have that $a, b \in \mathcal{K}$.

(4) Let $c = (c_{(1)}; \bar{c})$, $d = (d_{(1)}; \bar{d})$. By the definition of Jordan product, we have that

$$a_{(1)}c_{(1)} + \bar{a}^\top\bar{c} + \kappa a_{(1)}d_{(1)} - \kappa\bar{a}^\top\bar{d} = 0, \tag{2.1}$$

$$a_{(1)}\bar{c} + c_{(1)}\bar{a} + \kappa a_{(1)}\bar{d} - \kappa d_{(1)}\bar{a} = 0. \tag{2.2}$$

Let $p = a_{(1)}c_{(1)} + \bar{a}^\top\bar{c}$, $q = \kappa a_{(1)}d_{(1)} - \kappa\bar{a}^\top\bar{d}$. By (2.2), we have that

$$a_{(1)}(\bar{c} + \kappa\bar{d}) + (c_{(1)} - \kappa d_{(1)})\bar{a} = 0. \tag{2.3}$$

It follows from (2.3) and $a_{(1)} = \|\bar{a}\|$ that

$$p - q = a_{(1)}(c_{(1)} - \kappa d_{(1)}) + \bar{a}^\top(\bar{c} + \kappa\bar{d}) = 0. \tag{2.4}$$

Since (2.1) holds, $p+q=0$. It follows from (2.4) that $p=q=0$, i.e., $a^\top c=0$, $b^\top d=0$. Note that (2.3) and (2.4) imply that $a\circ(c_{(1)}-\kappa d_{(1)})=0$, so from Proposition 2.1(2), there exists a constant ϱ such that

$$\varrho a_{(1)}=c_{(1)}-\kappa d_{(1)},\quad \varrho\bar a=-(\bar c+\kappa\bar d). \tag{2.5}$$

Multiplying the two equalities of (2.5) by $c_{(1)}$ and $\bar c^\top$, respectively, we have that

$$\varrho a_{(1)}c_{(1)}=(c_{(1)}^2-\kappa c_{(1)}d_{(1)}),\quad \varrho\bar c^\top\bar a=-(\bar c^\top\bar c+\kappa\bar c^\top\bar d). \tag{2.6}$$

Adding the two equations in (2.6) together, it follows from $a^\top c=0$ that $c_{(1)}^2-\|\bar c\|^2=\kappa c^\top d$.

(5) Since $a\in\mathrm{bd}^+(\mathcal{K})$ and $b^\top a=0$, we have that $b_{(1)}=-\bar a^\top\bar b/a_{(1)}$ and $a_{(1)}^2=\bar a^\top\bar a$. By the Cauchy-Schwarz inequality, we have that

$$b_{(1)}^2-\|\bar b\|^2=\frac{(\bar a^\top\bar b)^2}{a_{(1)}^2}-\|\bar b\|^2\leqslant\frac{(\bar a^\top\bar a)(\bar b^\top\bar b)}{a_{(1)}^2}-\|\bar b\|^2=0.$$

(6) If $a\in\mathrm{int}(\mathcal{K})$, then by $a_{(1)}b_{(1)}+\bar a^\top\bar b=0$ and the Cauchy-Schwarz inequality,

$$|a_{(1)}b_{(1)}|=|\bar a^\top\bar b|\leqslant\|\bar a\|\|\bar b\|.$$

Note that $a\circ b=0$, $|b_{(1)}|=a_{(1)}\|\bar b\|/\|\bar a\|$, which implies that $a_{(1)}^2\|\bar b\|/\|\bar a\|\leqslant\|\bar a\|\|\bar b\|$. By $a_{(1)}>\|\bar a\|$, we have that $\|b\|=0$, which yields that $b=0$. □

3. Sensitivity result for QSOCP

Let us now consider (1.3), which is a general, possibly nonconvex, quadratic second-order cone problem. This problem is described by the data

$$\mathcal{D}:=[b,H,C,c,A,a],$$

where $A=[A_1,A_2,\dots,A_l]$, $a=[a_1,a_2,\dots,a_l]$. Let $s_i=A_id+a_i\in\mathbb{R}^{m_i}$, $\mu_i\in\mathbb{R}^{m_i}$,

$$\mu=(\mu_1;\mu_2;\dots;\mu_i;\dots;\mu_l)\in\mathbb{R}^m,\ s=(s_1;s_2;\dots;s_i;\dots;s_l)\in\mathbb{R}^m,\quad i=1,2,\dots,l.$$

The Lagrangian function of (1.3) is

$$L(d,\lambda,\mu)=b^\top d+\frac12 d^\top Hd-\lambda^\top(Cd+c)-\sum_{i=1}^{l}\mu_i^\top(A_id+a_i).$$

Denote by d^* the local solution of (1.3), $s_i^* = A_i d^* + a_i \in \mathbb{R}^{m_i}$, $\mu_i^* \in \mathbb{R}^{m_i}$,

$\mu^* = (\mu_1^*; \mu_2^*; \ldots; \mu_i^*; \ldots; \mu_l^*) \in \mathbb{R}^m$, $s^* = (s_1^*; s_2^*; \ldots; s_i^* \ldots; s_l^*) \in \mathbb{R}^m$, $i = 1, 2, \ldots, l$.

If the MFCQ constraint qualification [2] holds at d^*, i.e.,

$$(3.1) \qquad \begin{cases} C^\top \text{ is of full column rank,} \\ \exists\, \hat{d} \in \mathbb{R}^n \text{ such that } C\hat{d} = 0,\ A_i\hat{d} + a_i \in \operatorname{int}(\mathcal{K}_i),\ i = 1, 2, \ldots, l, \end{cases}$$

then there exist $\lambda^* \in \mathbb{R}^p$ and $\mu^* \in \mathbb{R}^m$ such that

$$(3.2) \qquad \begin{cases} A_i d^* + a_i = s_i^*, \\ Cd^* + c = 0, \\ b + Hd^* - C^\top \lambda^* = \sum_{i=1}^{l} A_i^\top \mu_i^*, \\ \mu_i^* \circ s_i^* = 0, \quad \mu_i^*, s_i^* \in \mathcal{K}_i \end{cases}$$

and $(d^*, \lambda^*, \mu^*, s^*)$ is called a stationary point of (1.3).

In the theorem below, we will present a sensitivity result for the solution of (1.3) when the data $\mathcal{D}$ is changed to $\mathcal{D} + \Delta\mathcal{D}$, where

$$(3.3) \qquad \Delta\mathcal{D} := [\Delta b, \Delta H, \Delta C, \Delta c, \Delta A, \Delta a]$$

is a sufficiently small perturbation. To this end, we need the following assumptions.

Assumptions (A)

(A1) There exists a local minimizer d^* of (1.3) where the MFCQ condition holds. The Lagrangian multiplier pair (λ^*, μ^*) corresponding to x^* is unique.

(A2) The second-order sufficient condition holds at d^* (see Definition 3.1 in [5]).

$$q^\top Hq + q^\top \mathcal{H}(d^*, \mu^*)q > 0 \quad \forall\, q \in \mathcal{C}(d^*) \setminus \{0\},$$

where

$$\mathcal{C}(d^*) = \{q \in \mathbb{R}^n \mid q^\top b = 0,\ Cq = 0,\ A_i q \in T_{\mathcal{K}_i}(s_i^*),\ i = 1, 2, \ldots, l\},$$

$T_{\mathcal{K}_i}(s_i^*)$ is the tangent cone of $\mathcal{K}_i$ at s_i^*, $\mathcal{H}(d^*, \mu^*) = \sum_{i=1}^{l} \mathcal{H}^i(d^*, \mu_i^*)$,

$$\mathcal{H}^i(d^*, \mu_i^*) = \begin{cases} -\dfrac{(\mu_i^*)_{(1)}}{(s_i^*)_{(1)}} A_i^\top R_i A_i, & s_i^* \in \operatorname{bd}^+(\mathcal{K}_i), \\ 0, & \text{otherwise,} \end{cases} \qquad R_i = \begin{pmatrix} 1 & 0^\top \\ 0 & -I_{m_i-1} \end{pmatrix},$$

where $(\mu_i^*)_{(1)}$ and $(s_i^*)_{(1)}$ denote the first entry of μ_i^* and s_i^*, respectively (see the definition in Section 1).

(A3) The following strict complementarity condition holds:

$$\mu_i^* + s_i^* \in \operatorname{int}(\mathcal{K}_i), \quad i = 1, 2, \ldots, l.$$

R e m a r k 3.1. $\mathcal{H}(d^*, \mu^*)$ is similar to the "sigma term" in nonlinear semidefinite programming [3].

$$\mathcal{C}(d^*) = \left\{ q \in \mathbb{R}^n \mid Cq = 0, \begin{cases} A_i q \in \mathcal{K}_i, & s_i^* = 0,\ \mu_i^* = 0, \\ \langle \overline{A_i q}, \overline{s_i^*} \rangle - \langle (A_i q)_{(1)}, (s_i^*)_{(1)} \rangle \leqslant 0, & s_i^* \in \operatorname{bd}^+(\mathcal{K}_i),\ \mu_i^* = 0, \\ A_i q = 0, & s_i^* = 0,\ \mu_i^* \in \operatorname{int}(\mathcal{K}_i), \\ A_i q \in \mathbb{R}_+((\mu_i^*)_{(1)}; -\overline{\mu_i^*}), & s_i^* = 0,\ \mu_i^* \in \operatorname{bd}^+(\mathcal{K}_i), \\ \langle A_i q, \mu_i^* \rangle = 0, & s_i^* \in \operatorname{bd}^+(\mathcal{K}_i),\ \mu_i^* \in \operatorname{bd}^+(\mathcal{K}_i) \end{cases} \right\}$$

where $A_i q = ((A_i q)_{(1)}; \overline{A_i q})$. For $q \in \mathcal{C}(d^*)$,

$$\begin{aligned} q^\top \mathcal{H}(d^*, \mu^*) q &= \sum_{s_i^* \in \operatorname{bd}^+(\mathcal{K}_i)} q^\top \mathcal{H}^i(d^*, \mu_i^*) q \\ &= \sum_{s_i^*, \mu_i^* \in \operatorname{bd}^+(\mathcal{K}_i)} -\frac{(\mu_i^*)_{(1)}}{(s_i^*)_{(1)}} (((A_i q)_{(1)})^2 - \|\overline{(A_i q)}\|^2), \end{aligned}$$

since $\langle A_i q, \mu_i^* \rangle = 0$ holds, it follows by Proposition 2.1(5) that $q^\top \mathcal{H}(d^*, \mu^*) q \geqslant 0$, which implies that the second-order sufficient condition with $\mathcal{H}(d^*, \mu^*)$ is weak. Besides, under Assumption (A3), from Proposition 2.1(1) we can also give the simple form of $\mathcal{C}(d^*)$ as follows.

$$\mathcal{C}(d^*) = \left\{ q \in \mathbb{R}^n \mid Cq = 0, \begin{cases} A_i q = 0, & s_i^* = 0,\ \mu_i^* \in \operatorname{int}(\mathcal{K}_i); \\ s_i^{*\top} R_i A_i q = 0, & s_i^* \in \operatorname{bd}^+(\mathcal{K}_i),\ \mu_i^* \in \operatorname{bd}^+(\mathcal{K}_i). \end{cases} \right\}$$

The main result of this section can be stated as follows.

Theorem 3.1. *Under Assumptions* (A)*, let* $(d^*, \lambda^*, \mu^*, s^*)$ *be the stationary point of* (1.3) *with the data* $\mathcal{D}$*, then for all sufficiently small perturbations* $\Delta\mathcal{D}$*, there exists a local stationary point* $(d(\Delta\mathcal{D}), \lambda(\Delta\mathcal{D}), \mu(\Delta\mathcal{D}), s(\Delta\mathcal{D}))$ *of the perturbed program* (1.3) *with the data* $\mathcal{D} + \Delta\mathcal{D}$*. Moreover, the point* $(d(\Delta\mathcal{D}), \lambda(\Delta\mathcal{D}), \mu(\Delta\mathcal{D}), s(\Delta\mathcal{D}))$ *is a differentiable function of the perturbation* $\Delta\mathcal{D}$ *and we have* $(d(\Delta\mathcal{D}), \lambda(\Delta\mathcal{D}), \mu(\Delta\mathcal{D}),$

$s(\Delta\mathcal{D})) = (d^*, \lambda^*, \mu^*, s^*)$ for $\Delta\mathcal{D} = 0$. The derivative $D_{\mathcal{D}}(d^*, \lambda^*, \mu^*, s^*)$ is characterized by the directional derivatives

$$(\dot{d}, \dot{\lambda}, \dot{\mu}, \dot{s}) = D_{\mathcal{D}}(d^*, \lambda^*, \mu^*, s^*)\Delta\mathcal{D}$$

for any $\Delta\mathcal{D}$. Here $(\dot{d}, \dot{\lambda}, \dot{\mu}, \dot{s})$ is the unique solution of the system of linear equations

$$\begin{cases} A_i\dot{d} = -\Delta A_i d^* - \Delta a_i + \dot{s}_i, \quad i = 1, 2, \ldots, l, \\ C\dot{d} = -\Delta C d^* - \Delta c, \\ H\dot{d} - C^\top\dot{\lambda} - \sum\limits_{i=1}^{l} A_i^\top \dot{\mu} = -\Delta H d^* + \Delta C^\top \lambda^* + \sum\limits_{i=1}^{l} \Delta A_i^\top \mu_i^*, \\ \mu_i^* \circ \dot{s}_i + s_i^* \circ \dot{\mu}_i = 0, \quad i = 1, 2, \ldots, l \end{cases} \tag{3.4}$$

for the unknowns $(\dot{d}, \dot{\lambda}, \dot{\mu}, \dot{s}) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m \times \mathbb{R}^m$, where

$$\dot{s} = (\dot{s}_1; \dot{s}_2; \ldots; \dot{s}_i; \ldots; \dot{s}_l),\ \dot{\mu} = (\dot{\mu}_1; \dot{\mu}_2; \ldots; \dot{\mu}_i; \ldots; \dot{\mu}_l), \quad \dot{s}_i \in \mathbb{R}^{m_i},\ \dot{\mu}_i \in \mathbb{R}^{m_i}.$$

Finally, the second-order sufficient condition holds at $(d(\Delta\mathcal{D}), \lambda(\Delta\mathcal{D}), \mu(\Delta\mathcal{D}), s(\Delta\mathcal{D}))$ if $\Delta\mathcal{D}$ is sufficiently small.

Remark 3.2. The sensitivity result for quadratic semidefinite programming (QSDP) is analyzed by Freund et al. [4], where the second-order sufficient condition is of the form without the sigma term. Grace et al. [7] give the generalization under the weak second-order sufficient conditions, i.e., the sigma term is taken into consideration. We give the proof of Theorem 3.1 by a similar proof technique. However, since our constraints contain second-order cones, the details are quite different from [4], [7].

Remark 3.3. Assumptions (A1) and (A2) imply that d^* is a strict local minimizer of (1.3). The MFCQ condition and the uniqueness of the Lagrangian multiplier can be replaced by a stronger condition, i.e., the nondegeneracy condition, which corresponds to the LICQ condition in nonlinear programming (see [2], Section 4).

Proof of Theorem 3.1. The proof is divided into four parts.

Part 1. First, we establish the following result: if the perturbed program has a local solution that is a differentiable function of the perturbation, then the derivative is indeed a solution of (3.4).

Suppose that there exists a solution $(d^* + \Delta d, s^* + \Delta s)$ of the perturbed problem near (d^*, s^*), where $\Delta s = (\Delta s_1; \Delta s_2; \ldots; \Delta s_i; \ldots; \Delta s_l) \in \mathbb{R}^m$, $\Delta s_i \in \mathbb{R}^{m_i}$. Since the MFCQ condition (3.1) is invariant under small perturbations of the problem data, there exist $\Delta\mu = (\Delta\mu_1; \Delta\mu_2; \dots; \Delta\mu_i; \dots \Delta\mu_l) \in \mathbb{R}^m (\Delta\mu_i \in \mathbb{R}^{m_i})$ and $\Delta\lambda \in \mathbb{R}^p$ such that $\mu_i^* + \Delta\mu_i \in \mathcal{K}_i, s_i^* + \Delta s_i \in \mathcal{K}_i$ and

$$(3.5) \qquad \begin{cases} (A_i + \Delta A_i)(d^* + \Delta d) + (a_i + \Delta a_i) = s_i^* + \Delta s_i, \quad i = 1, 2, \dots, l, \\ (C + \Delta C)(d^* + \Delta d) + (c + \Delta c) = 0, \\ (b + \Delta b) + (H + \Delta H)(d^* + \Delta d) - (C^\top + \Delta C^\top)(\lambda^* + \Delta\lambda) \\ \qquad\qquad = \sum_{i=1}^{l} (A_i^\top + \Delta A_i^\top)(\mu_i^* + \Delta\mu_i), \\ (\mu_i^* + \Delta\mu_i) \circ (s_i^* + \Delta s_i) = 0, \quad i = 1, 2, \dots, l. \end{cases}$$

Neglecting the second-order term in (3.5) and using the result of (3.2), we obtain the result in (3.4).

Part 2. Now we prove that the system of linear equations (3.4) has a unique solution. To this end, we will show that the homogeneous version of (3.4) only has a trivial solution, i.e., the system

$$(3.6) \qquad A_i \dot{d} = \dot{s}_i, \quad i = 1, 2, \dots, l,$$

$$(3.7) \qquad C\dot{d} = 0,$$

$$(3.8) \qquad H\dot{d} - C^\top \dot{\lambda} - \sum_{i=1}^{l} A_i^\top \dot{\mu}_i = 0,$$

$$(3.9) \qquad \mu_i^* \circ \dot{s}_i + s_i^* \circ \dot{\mu}_i = 0, \quad i = 1, 2, \dots, l,$$

only has the trivial solution $(\dot{d}, \dot{\lambda}, \dot{\mu}, \dot{s}) = (0, 0, 0, 0) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m \times \mathbb{R}^m$.

Multiplying (3.8) by $\dot{d}^\top$, by (3.6) and (3.7), we get

$$\dot{d}^\top H\dot{d} = \dot{d}^\top C^\top \lambda + \sum_{i=1}^{l} \dot{d}^\top A_i^\top \dot{\mu}_i = \sum_{i=1}^{l} \dot{\mu}_i^\top \dot{s}_i.$$

Since $s_i^* \circ \mu_i^* = 0$, if $s_i^* \in \text{int}(\mathcal{K}_i)$, then it follows from Proposition 2.1(6) that $\mu_i^* = 0$. By (3.9), we have that $\dot{\mu}_i \circ s_i^* = 0$, whence it follows by Proposition 2.1(6) that $\dot{\mu}_i = 0$; If $s_i^* = 0$, by Assumption (A3), we have that $\mu_i^* \in \text{int}(\mathcal{K}_i)$. By Proposition 2.1(6) and (3.9), we have that $\dot{s}_i = 0$. From the discussion above, we have that

$$(3.10) \qquad \dot{d}^\top H\dot{d} = \sum_{s_i^* \in \text{bd}^+(\mathcal{K}_i)} \dot{\mu}_i^\top \dot{s}_i.$$

By Assumption (A3), we know that if $s_i^* \in \text{bd}^+(\mathcal{K}_i)$, then $\mu_i^* \in \text{bd}^+(\mathcal{K}_i)$. Let $\dot{s}_i = ((\dot{s}_i)_{(1)}; \bar{\dot{s}}_i)$, $\kappa_i = (\mu_i^*)_{(1)}/(s_i^*)_{(1)}$. It follows from Proposition 2.1(1), (4) and (3.9) that

$$(3.11) \qquad (1/\kappa_i)\dot{\mu}_i^\top \dot{s}_i = (\dot{s}_i)_{(1)} - \|\bar{\dot{s}}_i\|^2.$$

Using the definition of $\mathcal{H}(d^*, \mu^*)$ and (3.6), (3.11), we have that

$$
\begin{aligned}
(3.12) \qquad \dot{d}^\top \mathcal{H}(d^*, \mu^*)\dot{d} &= \dot{d}^\top \sum_{s_i^* \in \mathrm{bd}^+(\mathcal{K}_i)} \mathcal{H}^i(x^*, \mu_i^*)\dot{d} \\
&= \sum_{s_i^* \in \mathrm{bd}^+(\mathcal{K}_i)} -\kappa_i (A_i\dot{d})^\top R_i (A_i\dot{d}) \\
&\overset{(3.6)}{=} -\kappa_i \sum_{s_i^* \in \mathrm{bd}^+(\mathcal{K}_i)} (\dot{s}_i)_{(1)} - \|\bar{\dot{s}}_i\|^2) \\
&\overset{(3.11)}{=} - \sum_{s_i^* \in \mathrm{bd}^+(\mathcal{K}_i)} \dot{\mu}_i^\top \dot{s}_i,
\end{aligned}
$$

which gives, using (3.10), that

$$
(3.13) \qquad \dot{d}^\top (H + \mathcal{H}(d^*, \mu^*))\dot{d} = 0.
$$

Next we will show that $\dot{d} \in \mathcal{C}(d^*)$ which is defined in Remark 3.1. It is clear that $C\dot{d} = 0$ holds because of (3.7). If $s_i^* = 0$, then, by (3.6), we have $A_i\dot{d} = 0$; If $s_i^* \in \mathrm{bd}^+(\mathcal{K}_i)$, then, by (3.9), Proposition 2.1(1) and (4), we have that ${\mu_i^*}^\top \dot{s}_i = 0$ and $\mu_i^* = \kappa_i R_i s_i^*$ (where $\kappa_i = (\mu_i^*)_{(1)}/(s_i^*)_{(1)}$), whence it follows by (3.6) that ${s_i^*}^\top R_i A_i \dot{d} = 0$. Thus we have that $\dot{d} \in \mathcal{C}(d^*)$. It follows from (3.13) and Assumption (A2) that $\dot{d} = 0$, which implies immediately that $\dot{s} = 0$ because of (3.6).

At last, we will show that $\dot{\mu} = 0$. Note that if $s_i^* \in \mathrm{int}(\mathcal{K}_i)$, then, by the proof above, we have already shown that $\dot{\mu}_i = 0$. Therefore, we only consider $s_i^* = 0$ and $s_i^* \in \mathrm{bd}^+(\mathcal{K}_i)$. If there exists $s_i^* = 0$ such that $\dot{\mu}_i \neq 0$, then from Assumption (A3), $\mu_i^* \in \mathrm{int}(\mathcal{K}_i)$. We can define $\mu_i^\tau = \mu_i^* + \tau\dot{\mu}_i$, where $\tau > 0$ is chosen to be a sufficiently small constant such that $\mu_i^\tau \in \mathcal{K}_i$. If there exists $s_i^* \in \mathrm{bd}^+(\mathcal{K}_i)$ such that $\dot{\mu}_i \neq 0$, since $\dot{s}_i = 0$, by (3.9), we have $\dot{\mu}_i \circ s_i^* = 0$. It follows from Proposition 2.1(2) that $\dot{\mu}_i = \kappa\mu_i^*$, where $\kappa \neq 0$ is a constant. Similarly, we can define $\mu_i^\tau = \mu_i^* + \tau\dot{\mu}_i$ such that $\mu_i^\tau \in \mathcal{K}_i$ when τ is sufficiently small. From the discussions above, we also have that $\mu_i^\tau \circ s_i^* = 0$ holds for $i = 1, 2, \ldots, l$. Thus, let $\mu^\tau = (\mu_1^\tau; \mu_2^\tau; \ldots; \mu_l^\tau)$ and

$$
\lambda^\tau = \lambda^* + \tau\dot{\lambda}, \ \mu^\tau = \mu^* + \tau\dot{\mu}.
$$

By (3.8), (3.9) and $\dot{d} = 0$, it is easy to verify that all the relations in (3.2) still hold for $(d^*, \lambda^\tau, \mu^\tau, s^*)$. So $(d^*, \lambda^\tau, \mu^\tau, s^*)$ is a stationary point for (1.3) and (λ^τ, μ^τ) is the corresponding Lagrangian multiplier, which contradicts Assumption (A1) that the Lagrangian multiplier is unique. Thus, we have that $\dot{\lambda} = 0$ and $\dot{\mu} = 0$.

Part 3. We will show that the following nonlinear system (3.14) has a solution which depends smoothly on the perturbation $\Delta\mathcal{D}$. Using the result in Part 2, we can now apply the implicit function theorem to the system

$$
\begin{cases}
A_i d + a_i = s_i, & i = 1, 2, \ldots, l, \\
Cd + c = 0, & \\
b + Hd - C^\top \lambda = \sum_{i=1}^{l} A_i^\top \mu_i, & \\
\mu_i \circ s_i = 0, & i = 1, 2, \ldots, l.
\end{cases} \tag{3.14}
$$

Since the linearization of (3.14) at the point $(d^*, \lambda^*, \mu^*, s^*)$ is nonsingular, system (3.14) has a differentiable and locally unique solution $(d(\Delta\mathcal{D}), \lambda(\Delta\mathcal{D}), \mu(\Delta\mathcal{D}), s(\Delta\mathcal{D}))$. By Assumption (A3), we have that $\mu_i^* + s_i^* \in \mathrm{int}(\mathcal{K}_i)$ $(i = 1, 2, \ldots, l)$. By the continuity of $\mu(\Delta\mathcal{D}), s(\Delta\mathcal{D})$, we have that $\lim\limits_{\Delta\mathcal{D}\to 0}(\mu(\Delta\mathcal{D}), s(\Delta\mathcal{D})) = (\mu^*, s^*)$. Therefore,

$$\mu_i(\Delta\mathcal{D}) + s_i(\Delta\mathcal{D}) \in \mathrm{int}(\mathcal{K}_i), \quad i = 1, 2, \ldots, l,$$

which implies by $\mu_i(\Delta\mathcal{D}) \circ s_i(\Delta\mathcal{D}) = 0$ and Proposition 2.1(3) that $\mu_i(\Delta\mathcal{D})$, $s_i(\Delta\mathcal{D}) \in \mathcal{K}_i$ $(i = 1, 2, \ldots, l)$. This implies that the local solutions of system (3.14) are actually stationary points when the perturbation $\Delta\mathcal{D}$ is sufficiently small.

Part 4. Finally, we will prove that the second-order sufficient condition is satisfied at the perturbed solution $(d(\Delta\mathcal{D}), \lambda(\Delta\mathcal{D}), \mu(\Delta\mathcal{D}), s(\Delta\mathcal{D}))$. Assume, by contradiction, that the statement is not true. Let $\{\Delta\mathcal{D}_k\}$ be a sequence of perturbations with $\Delta\mathcal{D}_k$ tending to zero and $q_k \in \mathcal{C}(d(\Delta\mathcal{D}_k)) \setminus \{0\}$. Then there exists a subset K such that $\lim\limits_{k\to\infty, k\in K} q_k = \hat{q}$, $\|\hat{q}\| = 1$ and

$$q_k^\top((H + \Delta H_k) + \mathcal{H}(d(\Delta\mathcal{D}_k), \mu(\Delta\mathcal{D}_k))q_k \leqslant 0. \tag{3.15}$$

Let $k \to \infty$, $k \in K$, by the continuity and Assumption (A3), we have that

$$\lim_{k\to\infty, k\in K} \mathcal{C}(d(\Delta\mathcal{D}_k)) \subseteq \mathcal{C}(d^*), \qquad \lim_{k\to\infty, k\in K} \mathcal{H}(d(\Delta\mathcal{D}_k), \mu(\Delta\mathcal{D}_k)) = \mathcal{H}(d^*, \mu^*).$$

By (3.15), there exists $\hat{q} \in \mathcal{C}(d^*)$ such that

$$\hat{q}(H + \mathcal{H}(d^*, \mu^*))\hat{q} \leqslant 0,$$

which is a contradiction to Assumption (A2). □

4. Local convergence of an SQP-type method for NSOCP

In this section, we apply the sensitivity result of QSOCP to the local convergence of the SQP-type method for NSOCP in [9]. Denote by $h_i(x)$ the ith component of $h(x)$, by $\nabla h_i(x)$ the gradient of $h_i(x)$ and by $\nabla^2 h_i(x)$ the Hessian matrix of $h_i(x)$. Denote by $Dh(x)$ the Jacobian matrix of $h(x)$. Denote by $g_{i,j}(x)$ the jth component of $g_i(x)$, its gradient and Hessian matrix are denoted by $\nabla g_{i,j}(x)$ and $\nabla^2 g_{i,j}(x)$, respectively. We need the following Assumptions (B) for (1.1), which are for general NSOCP and are different from Assumptions (A) for QSOCP.

Assumptions (B)

(B1) The functions $f(x), h(x), g_i(x)$ are twice continuously differentiable.

(B2) If $\{x_k\}$ is an infinite sequence generated by an SQP-type algorithm, then $\lim\limits_{k\to\infty} x_k = x^*$ and x^* is a KKT point of (1.1).

(B3) The constrained nondegeneracy condition (see [2]) is satisfied at x^*, i.e., the vectors

$$\begin{cases} \nabla h_k(x^*), & k = 1,2,\dots,p, \\ Dg_i(x^*)^\top R_i g_i(x^*),\ g_i(x^*) \in \mathrm{bd}^+(\mathcal{K}_i), & \\ \nabla g_{i,j}(x^*), & j = 1,2,\dots,m_i, \\ g_i(x^*) = 0, & i = 1,2,\dots,l \end{cases}$$

are linearly independent, where

$$R_i = \begin{pmatrix} 1 & 0^\top \\ 0 & -I_{m_i-1} \end{pmatrix}.$$

(B4) The second-order sufficient condition holds at x^*, i.e.,

$$d^\top \nabla^2_{xx} L(x^*, \lambda^*, \mu^*) d + d^\top \widehat{\mathcal{H}}(x^*, \mu^*) d > 0 \quad \forall\, d \in \widehat{\mathcal{C}}(x^*) \setminus \{0\},$$

where $\widehat{\mathcal{H}}(x^*, \mu^*) = \sum\limits_{i=1}^{l} \widehat{\mathcal{H}}^i(x^*, \mu_i^*)$, $\mu^*_{i,j}$ is the jth component of μ_i^*, λ_i^* is the ith component of λ^*,

$$\nabla^2_{xx} L(x^*, \lambda^*, \mu^*) = \nabla^2 f(x^*) - \sum_{i=1}^{p} \lambda_i^* \nabla^2 h_i(x^*) - \sum_{i=1}^{l} \sum_{j=1}^{m_i} \mu^*_{i,j} \nabla^2 g_{i,j}(x^*),$$

$$\widehat{\mathcal{C}}(x^*) = \{d \in \mathbb{R}^n \mid \nabla f(x^*)^\top d = 0,\ Dh(x^*)d = 0,\ Dg_i(x^*)d \in \mathcal{T}_{\mathcal{K}_i} g_i(x^*),\ i = 1, 2\dots, l\},$$

$$\widehat{\mathcal{H}}^i(x^*, \mu_i^*) = \begin{cases} -\dfrac{(\mu_i^*)_{(1)}}{(g_i(x^*))_{(1)}} Dg_i(x^*)^\top R_i Dg_i(x^*), & g_i(x^*) \in \mathrm{bd}^+(\mathcal{K}_i), \\ 0, & \text{otherwise.} \end{cases}$$

(B5) The strict complementarity condition holds, i.e., $g_i(x^*) + \mu_i^* \in \mathrm{int}(\mathcal{K}_i)$ for $i = 1, 2\dots, l$.

(B6) The Hessian approximation B_k satisfies

$$\|P_k(W_k - B_k)d_k\| = o(\|d_k\|),$$

where $P_k(y)$ is the orthogonal projection of a vector y onto $\operatorname{Ker} V_k$, V_k is a matrix whose column vectors are

$$\begin{cases} \nabla h_k(x_k, & k = 1, 2, \ldots, p, \\ Dg_i(x_k)^\top R_i g_i(x_k), \ g_i(x_k) \in \operatorname{bd}^+(\mathcal{K}_i), & \\ \nabla g_{i,j}(x_k), & j = 1, 2, \ldots, m_i, \\ g_i(x_k) = 0, & i = 1, 2, \ldots, l \end{cases}$$

and $W_k = \nabla^2_{xx} L(x_k, \lambda_{k+1}, \mu_{k+1})$.

Remark 4.1. The nondegeneracy condition (B3) is similar to the LICQ condition in nonlinear programming, which is stronger than the MFCQ condition stated as (3.1). Denote by V^* a matrix whose columns are formed by the vectors stated in Assumption (B3) above, then (B3) is equivalent to the condition that V^* is of full column rank. For simplicity, we still use the notation (λ^*, μ^*) for the Lagrangian multiplier corresponding with x^*. If the constrained nondegeneracy condition holds at x^*, then the corresponding Lagrangian multiplier (λ^*, μ^*) is unique (see [2], [3]).

Remark 4.2. If Assumption (B5) holds, then $\widehat{\mathcal{C}}(x^*)$ has the following form:

$$\widehat{\mathcal{C}}(x^*) = \left\{ d \in \mathbb{R}^n \mid Dh(x^*)d = 0, \ \begin{cases} Dg_i(x^*)d = 0, & g_i(x^*) = 0, \\ g_i(x^*)^\top R_i Dg_i(x^*)d = 0, & g_i(x^*) \in \operatorname{bd}^+(\mathcal{K}_i) \end{cases} \right\},$$

that is, $V^{*\top} d = 0$. Furthermore, for $d \in \widehat{\mathcal{C}}(x^*)$, we have that

$$\begin{aligned} d^\top \widehat{\mathcal{H}}(x^*, \mu^*)d &= \sum_{g_i(x^*) \in \operatorname{bd}^+(\mathcal{K}_i)} d^\top \widehat{\mathcal{H}}^i(x^*, \mu_i^*)d \\ &= \sum_{g_i(x^*) \in \operatorname{bd}^+(\mathcal{K}_i)} -\frac{(\mu_i^*)_{(1)}}{(g_i(x^*))_{(1)}} (((Dg_i(x^*)d)_{(1)})^2 - \|\overline{Dg_i(x^*)d}\|^2), \end{aligned}$$

which yields by Proposition 2.1(5) that $d^\top \widehat{\mathcal{H}}(x^*, \mu^*)d \geqslant 0$.

Remark 4.3. In [9], [14], B_k is taken as W_k, which is stronger than Assumption (B6).

Lemma 4.1. *Under Assumptions* (B), *we have that* $\lim_{k \to \infty} (d_k, \lambda_{k+1}, \mu_{k+1}) = (0, \lambda^*, \mu^*)$.

P r o o f. The proof of Lemma 4.1 requires the sensitivity result from Theorem 3.1. Suppose, by contradiction, that there exists a subset K such that for $k \in K$, $\lim\limits_{k\to\infty,k\in K} d_k \neq 0$ and $\lim\limits_{k\to\infty,k\in K} B_k = B^*$.

Consider the following subproblem:

$$\begin{aligned} \min_{d\in\mathbb{R}^n} \ & \nabla f(x^*)^\top d + \frac{1}{2} d^\top B^* d \\ \text{s.t. } & h(x^*) + Dh(x^*)d = 0, \\ & g_i(x^*) + Dg_i(x^*)d \in \mathcal{K}_i, \quad i = 1,2,\ldots,l. \end{aligned} \tag{4.1}$$

Denote

$$g(x) = (g_1(x), g_2(x), \ldots, g_l(x)), \quad Dg(x) = (Dg_1(x), Dg_2(x), \ldots, Dg_l(x)).$$

Then problem (4.1) is described by the data

$$\mathcal{D} = (\nabla f(x^*), B^*, Dh(x^*), h(x^*), Dg(x^*), g(x^*)).$$

Comparing (4.1) with (1.3), we can take $b = \nabla f(x^*)$, $H = B^*$, $C = Dh(x^*)$, $c = h(x^*)$, $A_i = Dg_i(x^*)$, $a_i = g_i(x^*)$ in order to use the result in Theorem 3.1. First, we have to show that the Assumptions (A1)–(A3) hold for (4.1) at $d = 0$. By Assumptions (B1), (B2) and (B3), x^* is a local minimizer of (1.1) and (λ^*, μ^*) is the unique Lagrangian multiplier. Using the KKT conditions of (1.1), it is easy to verify that $(0, \lambda^*, \mu^*)$ is the solution of (4.1). Besides, by simple calculation, we know that the nondegeneracy condition (which implies the MFCQ condition) and strict complementarity condition for (4.1) also hold at $d = 0$ because of Assumptions (B3) and (B5). By Remark 4.2, we know that for $q \in \widehat{\mathcal{C}}(x^*)$, $q^\top \widehat{\mathcal{H}}(x^*, \mu^*) q > 0$ holds. Since $\mathcal{C}(0) = \widehat{\mathcal{C}}(x^*)$, $\mathcal{H}(0, \mu^*) = \widehat{\mathcal{H}}(x^*, \mu^*)$ and B_k is positive definite, we have that

$$q^\top (B^* + \mathcal{H}(0, \mu^*)) q > 0 \tag{4.2}$$

holds for all $q \in \mathcal{C}(0) \setminus \{0\}$, which implies that the second-order sufficient condition for (4.1) holds at $d = 0$.

Now let $(d_k, \lambda_{k+1}, \mu_{k+1})$ be the solution of the perturbed problem

$$\begin{aligned} \min_{d\in\mathbb{R}^n} \ & \nabla f(x_k)^\top d + \frac{1}{2} d^\top B_k d \\ \text{s.t. } & h(x_k) + Dh(x_k)d = 0, \\ & g_i(x_k) + Dg_i(x_k) \in \mathcal{K}_i, \quad i = 1,2,\ldots,l, \end{aligned} \tag{4.3}$$

where $\mu_{k+1} = (\mu_{k+1,1}; \ldots; \mu_{k+1,i}; \ldots; \mu_{k+1,l})$, $\mu_{k+1,i} \in \mathcal{K}_i$. The perturbation is $\Delta\mathcal{D}_k$, i.e.,

$$(\nabla f(x_k) - \nabla f(x^*), B_k - B^*, Dh(x_k) - Dh(x^*),$$
$$h(x_k) - h(x^*), Dg(x_k) - Dg(x^*), g(x_k) - g(x^*)).$$

By Theorem 3.1, we have that $\lim\limits_{k\to\infty, k\in K} d_k = 0$, which is a contradiction. Therefore, $(\lambda_{k+1}, \mu_{k+1}) \to (\lambda^*, \mu^*)$ can be proved by a similar technique. □

Remark 4.4. For k sufficiently large, if $g_i(x^*) \in \operatorname{int}(\mathcal{K}_i)$, by Lemma 4.1, we have that $g_i(x_k) + Dg_i(x_k)d_k \in \operatorname{int}(\mathcal{K}_i)$ holds. Note that x^* is a KKT point of (1.1) and d_k is the KKT point of (4.3), it follows from the complementarity conditions that

$$\mu_i^* \circ g_i(x^*) = 0, \quad \mu_{k+1,i} \circ (g_i(x_k) + Dg_i(x_k)d_k) = 0,$$

so it follows from Proposition 2.1(6) that $\mu_{k+1,i} = \mu_i^* = 0$. As a result, we do not have to take such constraints into consideration. Define the index sets

$$I_1^* = \{i \mid g_i(x^*) = 0\}, \quad I_2^* = \{i \mid g_i(x^*) \in \operatorname{bd}^+(\mathcal{K}_i)\},$$
$$I_1^k = \{i \mid g_i(x_k) + Dg_i(x_k)d_k = 0\}, \quad I_2^k = \{i \mid g_i(x_k) + Dg_i(x_k)d_k \in \operatorname{bd}^+(\mathcal{K}_i)\}.$$

We show the relations of the index sets in the following lemma.

Lemma 4.2. *Under Assumptions* (B1)–(B5), *for k sufficiently large, we have that*

$$I_1^k = I_1^*, \quad I_2^k = I_2^*.$$

Proof. The results $I_1^k \subseteq I_1^*$ and $I_2^k \subseteq I_2^*$ follow directly by the continuity and $\lim\limits_{k\to\infty} d_k = 0$. If $i \in I_1^*$, then, by the strict complementarity condition in Assumption (B5), we have $\mu_i^* \in \operatorname{int}(\mathcal{K}_i)$. It follows from the continuity and Lemma 4.1 that $\mu_{k+1,i} \in \operatorname{int}(\mathcal{K}_i)$ for k sufficiently large. By $\mu_{k+1,i} \circ (g_i(x_k) + Dg_i(x_k)d_k) = 0$ and Proposition 2.1(1), we have that $i \in I_1^k$. Therefore, $I_1^k \supseteq I_1^*$. The inclusion $I_2^k \supseteq I_2^*$ can be proved by a similar technique. □

By Lemma 4.2, for k sufficiently large, we can write $I_1^k = I_1^* = I_1, I_2^k = I_2^* = I_2$ for simplicity. Thus, when k is sufficiently large and Assumptions (B1)–(B5) hold, (1.2) turns into the following form:

$$\begin{aligned}
(4.4) \qquad & \min_{d\in\mathbb{R}^n} g_k^\top d + \frac{1}{2} d^\top B_k d \\
& \text{s.t. } h_k + Dh(x_k)d = 0, \\
& \quad g_i(x_k) + Dg_i(x_k)d = 0, \qquad i \in I_1, \\
& \quad g_i(x_k) + Dg_i(x_k)d \in \operatorname{bd}^+(\mathcal{K}_i), \quad i \in I_2.
\end{aligned}$$

We state the main result of the local convergence of the SQP-type method [9] for (1.1) in the following theorem.

Theorem 4.1. *Under Assumptions* (B), *if* $x_{k+1} = x_k + d_k$ *holds for* k *sufficiently large, then* $\|x_k + d_k - x^*\| = o(\|x_k - x^*\|)$.

R e m a r k 4.5. Kato et al. [9] give the local convergence of the SQP-type method for second-order cone programming when $B_k = W_k$, which is proved by Wang et al. [14]. The result in Theorem 4.1, compared with the result in [9], does not need the assumption that W_k is positive definite when k is sufficiently large. Instead, we replace the assumption with Assumptions (B5) and (B6). Besides, Theorem 4.3 shows the convergence rate of $\{x_k\}$ without the Lagrangian multiplier.

We need some lemmas in order to prove Theorem 4.1.

Lemma 4.3. *Under Assumptions* (B1)–(B5), *for* k *sufficiently large, we have that*

$$\|\lambda_{k+1} - \lambda^*\| = O(\|d_k\|) + O(\|x_k - x^*\|), \quad \|\mu_{k+1} - \mu^*\| = O(\|d_k\|) + O(\|x_k - x^*\|).$$

P r o o f. By the KKT conditions of (1.1) and (1.2), we have that

$$\nabla f(x^*) - Dh(x^*)^\top \lambda^* - \sum_{i \in I_1 \cup I_2} Dg_i(x^*)^\top \mu_i^* = 0, \tag{4.5}$$

$$\nabla f(x_k) + B_k d_k - Dh(x_k)^\top \lambda_{k+1} - \sum_{i \in I_1 \cup I_2} Dg_i(x_k)^\top \mu_{k+1,i} = 0. \tag{4.6}$$

For $i \in I_2$, by Proposition 2.1(1), we can denote

$$\mu_i^* = \kappa_i^* R_i g_i(x^*), \quad \mu_{k+1,i} = \kappa_{k+1,i} R_i (g_i(x_k) + Dg_i(x_k) d_k), \tag{4.7}$$

where

$$\kappa_i^* = \frac{(\kappa_i{}^*)_{(1)}}{(g_i(x^*))_{(1)}}, \quad \kappa_{k+1,i} = \frac{(\kappa_{k+1,i})_{(1)}}{(g_i(x_k) + Dg_i(x_k) d_k)_{(1)}}, \quad R_i = \begin{pmatrix} 1 & 0^\top \\ 0 & -I_{m_i - 1} \end{pmatrix}.$$

By the continuity and (4.5)–(4.7), we have that

$$\begin{aligned} Dh(x^*)(\lambda_{k+1} - \lambda^*) &+ \sum_{i \in I_1} Dg_i(x^*)^\top (\mu_{k+1,i} - \mu_i^*) \\ &+ \sum_{i \in I_2} Dg_i(x^*)^\top R_i g_i(x^*)(\kappa_{k+1,i} - \kappa_i^*) \\ &= O(\|d_k\|) + O(\|x_k - x^*\|), \end{aligned}$$

whence it follows by Assumption (B2) that $\|\lambda_{k+1} - \lambda^*\| = O(\|d_k\|) + O(\|x_k - x^*\|)$ and

$$\|\mu_{k+1,i} - \mu_i^*\| = O(\|d_k\|) + O(\|x_k - x^*\|), \quad i \in I_1,$$
$$\|\kappa_{k+1,i} - \kappa_i^*\| = O(\|d_k\|) + O(\|x_k - x^*\|), \quad i \in I_2.$$

For $i \in I_2$, it follows from the continuity and (4.7) that

$$\|\mu_{k+1,i} - \mu_i^*\| = O(\|d_k\|) + O(\|x_k - x^*\|), \quad i \in I_2.$$

Thus the statement is true. □

We still need the following auxiliary problem to analyze local convergence.

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \ & f(x) \\ \text{s.t. } & h(x) = 0, \\ & F_i(x) = 0, \quad i \in I_1 \cup I_2, \end{aligned} \tag{4.8}$$

where $F_i(x)$ is defined as

$$F_i(x) = \begin{cases} g_i(x), & i \in I_1, \\ \frac{1}{2}(g_i(x)_{(1)})^2 - \frac{1}{2}\|\overline{g_i(x)}\|^2, & i \in I_2. \end{cases}$$

Denote

$$\widehat{\mu}_{k+1,i} = \begin{cases} \mu_{k+1,i}, & i \in I_1, \\ \dfrac{(\mu_{k+1,i})_{(1)}}{(g_i(x_k) + Dg_i(x_k)d_k)_{(1)}}, & i \in I_2, \end{cases}$$

$$\widehat{\mu}_i^* = \begin{cases} \mu_i^*, & i \in I_1, \\ \dfrac{(\mu_i^*)_{(1)}}{(g_i(x^*))_{(1)}}, & i \in I_2. \end{cases}$$

The Lagrangian function of (4.8) is

$$\widehat{L}(x, \lambda, \mu) = f(x) - Dh(x)^\top \lambda - \sum_{i\in I_1\cup I_2} DF_i(x)^\top \mu_i.$$

By simple calculation, we have that

$$\begin{aligned} \nabla_x \widehat{L}(x^*, \lambda^*, \widehat{\mu}^*) &= \nabla f(x^*) - Dh(x^*)^\top \lambda^* - \sum_{i\in I_1\cup I_2} DF_i(x^*)^\top \widehat{\mu}_i^* \\ &= \nabla f(x^*) - Dh(x^*)^\top \lambda^* - \sum_{i\in I_1\cup I_2} Dg_i(x^*)^\top \mu_i^* = 0, \end{aligned} \tag{4.9}$$

where the ith component of $\widehat{\mu}^*$ is $\widehat{\mu}_i^*$, and

$$\begin{aligned}(4.10)\quad \nabla_{xx}\widehat{L}(x^*,\lambda^*,\widehat{\mu}^*) &= \nabla^2 f(x^*) - \sum_{i=1}^{p}\lambda_i^*\nabla^2 h_i(x^*) - \sum_{i\in I_1\cup I_2}\sum_{j=1}^{m_i}\widehat{\mu}_{i,j}^*\nabla^2 F_{i,j}(x^*)\\ &= \nabla^2 f(x^*) - \sum_{i=1}^{p}\lambda_i^*\nabla^2 h_i(x^*) - \sum_{i\in I_1\cup I_2}\sum_{j=1}^{m_i}\mu_{i,j}^*\nabla^2 g_{i,j}(x^*)\\ &\quad - \sum_{i\in I_2}\widehat{\mu}_i^* Dg_i(x^*)^\top R_i Dg_i(x^*)\\ &= \nabla_{xx}^2 L(x^*,\lambda^*,\mu^*) + \sum_{i\in I_2}\widehat{\mathcal{H}}^i(x^*,\mu_i^*),\end{aligned}$$

where $\mu_{i,j}^*$ denotes the jth component of μ_i^*. By the definition of P_k in Assumption (B6), we have that

$$(4.11)\qquad \begin{aligned} P_k DF_i(x_k)^\top &= 0, \quad i\in I_1\cup I_2;\\ P_k Dh(x_k)^\top &= 0.\end{aligned}$$

Lemma 4.4. *Under Assumptions* (B), *for k sufficiently large, if $x_{k+1} = x_k + d_k$, then*

$$(4.12)\qquad \begin{aligned} &Dg_i(x_k)^\top \mu_{k+1,i} - DF_i(x_k)^\top \widehat{\mu}_{k+1,i}\\ &\qquad = -\widehat{\mathcal{H}}^i(x_k,\mu_{k+1,i})d_k + O(\|d_k\|^2), \quad i\in I_2.\end{aligned}$$

Proof. By (4.7) and the definition of $\widehat{\mu}_{k+1,i}$, for $i\in I_2$, we have that

$$(4.13)\qquad \begin{aligned} Dg_i(x_k)^\top\mu_{k+1,i} &= Dg_i(x_k)^\top R_i(g_i(x_k)\\ &\quad + Dg_i(x_k)d_k)\frac{(\mu_{k+1,i})_{(1)}}{(g_i(x_k)+Dg_i(x_k)d_k)_{(1)}},\end{aligned}$$

$$(4.14)\qquad DF_i(x_k)^\top\widehat{\mu}_{k+1,i} = Dg_i(x_k)^\top R_i g_i(x_k)\frac{(\mu_{k+1,i})_{(1)}}{(g_i(x_k)+Dg_i(x_k)d_k)_{(1)}}.$$

It follows from (4.13) and (4.14) that

$$\begin{aligned} &Dg_i(x_k)^\top\mu_{k+1,i} - DF_i(x_k)^\top\widehat{\mu}_{k+1,i}\\ &\qquad = \frac{(\mu_{k+1,i})_{(1)}}{(g_i(x_k)+Dg_i(x_k)d_k)_{(1)}} Dg_i(x_k)^\top R_i Dg_i(x_k)d_k.\end{aligned}$$

Since

$$\widehat{\mathcal{H}}^i(x_k,\mu_{k+1,i}) = -\frac{(\mu_{k+1,i})_{(1)}}{(g_i(x_k))_{(1)}} Dg_i(x_k)^\top R_i Dg_i(x_k),$$

we have that

$$\begin{aligned} &Dg_i(x_k)^\top \mu_{k+1,i} - DF_i(x_k)^\top \widehat{\mu}_{k+1,i} \\ &\quad= \frac{(\mu_{k+1,i})_{(1)}}{(g_i(x_k))_{(1)}} Dg_i(x_k)^\top R_i Dg_i(x_k) d_k \\ &\qquad + \Big(\frac{(\mu_{k+1,i})_{(1)}}{(g_i(x_k)+Dg_i(x_k)d_k)_{(1)}} - \frac{(\mu_{k+1,i})_{(1)}}{(g_i(x_k))_{(1)}}\Big) \\ &\qquad \times Dg_i(x_k)^\top R_i Dg_i(x_k) d_k \\ &\quad= -\widehat{\mathcal{H}}^i(x_k,\mu_{k+1,i}) + O(\|d_k\|^2). \end{aligned}$$

Thus the statement is true. □

Lemma 4.5. *Under Assumptions* (B1)–(B4), *for k sufficiently large, the matrix*

$$\begin{pmatrix} P(x^*)\nabla^2_{xx}\widehat{L}(x^*,\lambda^*,\widehat{\mu}^*) \\ (V^*)^\top \end{pmatrix}$$

is of full column rank, where $P(x^*) = I - V^*({V^*}^\top V^*)^{-1}{V^*}^\top$, V^* is the matrix defined in Remark 4.1.

P r o o f. We only need to show that the system

$$P(x^*)\nabla^2_{xx}\widehat{L}(x^*,\lambda^*,\widehat{\mu}^*)d = 0, \tag{4.15}$$
$$(V^*)^\top d = 0 \tag{4.16}$$

has only the trivial solution $d=0$. By the definition of $P(x^*)$ and (4.16), we have $P(x^*)d = d$. Multiplying both sides of (4.15) by $d^\top$, we get

$$d^\top P(x^*)\nabla^2_{xx}\widehat{L}(x^*,\lambda^*,\widehat{\mu}^*)d = d^\top\nabla^2_{xx}\widehat{L}(x^*,\lambda^*,\widehat{\mu}^*)d = 0. \tag{4.17}$$

By Assumption (B4), (4.10) and (4.17), we have $d=0$. Therefore, the matrix is a full column rank matrix. □

Finally, we give the proof of Theorem 4.1.

P r o o f of Theorem 4.3. First, we have that

$$
\begin{aligned}
&(4.18)\ P_k(B_k - \nabla^2_{xx} L(x^*, \lambda^*, \mu^*))d_k \\
&\overset{(4.6)}{=} P_k\Big(- \nabla f(x_k) + Dh(x_k)^\top \lambda_{k+1} + \sum_{i \in I_1 \cup I_2} Dg_i(x_k)\mu_{k+1,i} - \nabla^2_{xx} L(x^*, \lambda^*, \mu^*)d_k\Big) \\
&\overset{(4.11)}{=} P_k\Big(- \nabla f(x_k) + Dh(x_k)^\top \lambda^* + \sum_{i \in I_1 \cup I_2} DF_i(x_k)^\top \widehat{\mu}_i^* - \nabla^2_{xx} \widehat{L}(x^*, \lambda^*, \widehat{\mu}^*)d_k\Big) \\
&\qquad + P_k\Big(Dh(x_k)^\top \lambda_{k+1} + \sum_{i \in I_1 \cup I_2} Dg_i(x_k)^\top \mu_{k+1,i} \\
&\qquad\qquad - \nabla^2_{xx} L(x^*, \lambda^*, \mu^*)d_k + \nabla^2_{xx} \widehat{L}(x^*, \lambda^*, \widehat{\mu}^*)d_k\Big) \\
&\overset{(4.9),(4.11)}{=} - P_k(\nabla_x \widehat{L}(x_k, \lambda^*, \widehat{\mu}^*) - \nabla_x \widehat{L}(x^*, \lambda^*, \widehat{\mu}^*) + \nabla^2_{xx} L(x^*, \lambda^*, \mu^*)d_k) \\
&\qquad + P_k\Big(\sum_{i \in I_1 \cup I_2} Dg_i(x_k)^\top \mu_{k+1,i} - \sum_{i \in I} DF_i(x_k)^\top \widehat{\mu}_{k+1,i}\Big) \\
&\qquad + P_k(-\nabla^2_{xx} L(x^*, \lambda^*, \mu^*)d_k + \nabla^2_{xx} \widehat{L}(x^*, \lambda^*, \widehat{\mu}^*)d_k) \\
&\overset{(4.10)}{=} - P_k \nabla^2_{xx} \widehat{L}(x^*, \lambda^*, \widehat{\mu}^*)(x_k + d_k - x^*) + o(\|x_k - x^*\|) \\
&\qquad + P_k\Big(\sum_{i \in I_1 \cup I_2} Dg_i(x_k)^\top \mu_{k+1,i} - \sum_{i \in I} DF_i(x_k)^\top \widehat{\mu}_{k+1,i}\Big) \\
&\qquad + P_k \sum_{i \in I_2} \mathcal{H}^i(x^*, \mu_i^*)d_k \\
&\overset{(4.12)}{=} - P_k \nabla^2_{xx} \widehat{L}(x^*, \lambda^*, \widehat{\mu}^*)(x_k + d_k - x^*) + o(\|x_k - x^*\|) \\
&\qquad + P_k\Big(-\sum_{i \in I_2} \widehat{\mathcal{H}}^i(x_k, \mu_{k+1,i})d_k + \sum_{i \in I_2} \widehat{\mathcal{H}}^i(x^*, \mu_i^*)d_k\Big) + o(\|d_k\|).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&P_k(B_k - \nabla^2_{xx} L(x^*, \lambda^*, \mu^*))d_k \\
&\quad = - P_k \nabla^2_{xx} \widehat{L}(x^*, \lambda^*, \widehat{\mu}^*)(x_k + d_k - x^*) + o(\|x_k - x^*\|) + o(\|d_k\|),
\end{aligned}
$$

which implies by Assumption (B6) that

$$
(4.20) \qquad P_k \nabla^2_{xx} \widehat{L}(x^*, \lambda^*, \widehat{\mu}^*)(x_k + d_k - x^*) = o(\|x_k - x^*\|) + o(\|d_k\|).
$$

By the Taylor expansion,

$$
h(x_k) = h(x_k) - h(x^*) = Dh(x_k)(x_k - x^*) + O(\|x_k - x^*\|^2),
$$

which implies by $Dh(x_k)d_k + h(x_k) = 0$ that

$$Dh(x^*)(x_k + d_k - x^*) = o(\|x_k - x^*\|). \tag{4.21}$$

For $i \in I_1$, $g_i(x^*) = 0$ and $g_i(x_k) + Dg_i(x_k)d_k = 0$. Similarly to the proof above, we have that

$$Dg_i(x^*)(x_k + d_k - x^*) = o(\|x_k - x^*\|), \quad i \in I_1. \tag{4.22}$$

For $i \in I_2$, denote by $\prod_{\mathcal{K}_i} u$ the merit projector of $u \in \mathbb{R}^{m_i}$ to the second-order cone $\mathcal{K}_i$. By this notation, $g_i(x^*) \in \mathcal{K}_i$, $\mu_i^* \in \mathcal{K}_i$ and $g_i(x^*) \circ \mu_i^* = 0$ is equivalent to

$$\prod_{\mathcal{K}_i}(g_i(x^*) - \mu_i^*) = g_i(x^*), \quad i \in I_2. \tag{4.23}$$

Similarly, we have that, for $i \in I_2$,

$$g_i(x_k) + Dg_i(x_k)d_k \in \mathcal{K}_i,\ \mu_{k+1,i} \in \mathcal{K}_i, \quad (g_i(x_k) + Dg_i(x_k)d_k) \circ \mu_{k+1,i} = 0$$

is equivalent to

$$\prod_{\mathcal{K}_i}(g_i(x_k) + Dg_i(x_k)d_k - \mu_{k+1,i}) = g_i(x_k) + Dg_i(x_k)d_k, \quad i \in I_2. \tag{4.24}$$

As the projection operator $\prod_{\mathcal{K}_i} u$ is strongly semismooth [12], there exists $U_i \in \partial_B \prod_{\mathcal{K}_i}(g_i(x^*) - \mu_i^*)$ such that (∂_B is the B-subdifferential [13])

$$\begin{aligned}\prod_{\mathcal{K}_i}(g_i(x^*) - \mu_i^*) = &\prod_{\mathcal{K}_i}(g_i(x_k) + Dg_i(x_k)d_k - \mu_{k+1,i})\\ &+ U_i(g_i(x^*) - \mu_i^* - g_i(x_k) - Dg_i(x_k)d_k + \mu_{k+1,i})\\ &+ o(\|x_k - x^*\|) + o(\|d_k\|) + o(\|\mu_{k+1,i} - \mu_i^*\|),\end{aligned}$$

which gives, using (4.22) and (4.23), that

$$\begin{aligned}g_i(x^*) = &\, g_i(x_k) + Dg_i(x_k)d_k + U_i(g_i(x^*) - \mu_i^* - g_i(x_k) - Dg_i(x_k)d_k + \mu_{k+1,i})\\ &+ o(\|x_k - x^*\|) + o(\|d_k\|) + o(\|\mu_{k+1,i} - \mu_i^*\|).\end{aligned}$$

It follows from the Taylor expansion that

$$\begin{aligned}Dg_i(x^*)(x_k + d_k - x^*) = &\, U_i(g_i(x^*) - \mu_i^* - g_i(x_k) - Dg_i(x_k)d_k + \mu_{k+1,i})\\ &+ o(\|x_k - x^*\|) + o(\|d_k\|) + o(\|\mu_{k+1,i} - \mu_i^*\|).\end{aligned} \tag{4.25}$$

Multiplying both sides of (4.24) by $g_i(x^*)^\top R_i$, we have that

$$\begin{aligned} & g_i(x^*)^\top R_i Dg_i(x^*)(x_k + d_k - x^*) \\ & \qquad = g_i(x^*)^\top R_i U_i(g_i(x^*) - \mu_i^* - g_i(x_k) - Dg_i(x_k)d_k + \mu_{k+1,i}) \\ & \qquad\quad + o(\|x_k - x^*\|) + o(\|d_k\|) + o(\|\mu_{k+1,i} - \mu_i^*\|). \end{aligned}$$

For $g_i(x^*) \in \mathrm{bd}^+(\mathcal{K}_i)$, it follows from Assumption (B5) that $\mu_i^* \in \mathrm{bd}^+(\mathcal{K}_i)$. Using the formulas for the subdifferential of the projector operator $\prod_{\mathcal{K}_i} u$ (see Lemma 2.4 in [14]), we have that $g_i(x^*)^\top R_i U_i = 0$ (the details are given in the appendix part). This fact, along with Lemma 4.4, shows that, for $i \in I_2$,

$$g_i(x^*)^\top R_i Dg_i(x^*)(x_k + d_k - x^*) = o(\|x_k - x^*\|) + o(\|d_k\|). \tag{4.26}$$

By (4.20), (4.21) and (4.25), we have that

$$(V^*)^\top (x_k + d_k - x^*) = o(\|x_k - x^*\|) + o(\|d_k\|). \tag{4.27}$$

By (4.19), (4.26) and $\|d_k\| = \|(x_{k+1} - x^*) - (x_k - x^*)\| \leqslant \|x_{k+1} - x^*\| + \|x_k - x^*\|$, we have that

$$\begin{pmatrix} P_k \nabla^2_{xx} \widehat{L}(x^*, \lambda^*, \widehat{\mu}^*) \\ (V^*)^\top \end{pmatrix} (x_k + d_k - x^*) = o(\|x_{k+1} - x^*\|) + o(\|x_k - x^*\|). \tag{4.28}$$

By Lemma 4.6 and the continuity, the matrix on the left-side of (4.27) is nonsingular, which implies that $(x_k + d_k - x^*) = o(\|x_k - x^*\|)$. □

5. Final remarks

In this paper, we analyze the sensitivity of the quadratic second-order cone programming under the weak second-order sufficient condition. By the sensitivity result, we give the superlinear convergence rate of an SQP-type method for NSOCP. The result is different from other recent work, because we do not need the assumption that the Hessian matrix is positive definite near the solution. Furthermore, we analyze the local convergence rate of the iterates $\{x_k\}$ instead of the iterates $\{x_k, \lambda_k, \mu_k\}$.

6. Appendix: the proof details in Theorem 4.1

Lemma 6.1. *Let $\mathcal{K} \subseteq \mathbb{R}^{\widehat{m}}$ be a second-order cone, then for $u = (u_{(1)}; \bar{u}) \in \mathbb{R} \times \mathbb{R}^{\widehat{m}-1}$ and $|u_{(1)}| < \|\bar{u}\|$, the subdifferential of the projector $\prod_{\mathcal{K}} u$ is*

$$\partial_B \prod_{\mathcal{K}} u = \frac{1}{2}\begin{pmatrix} 1 & \dfrac{\bar{u}^\top}{\|\bar{u}\|} \\ \dfrac{\bar{u}}{\|\bar{u}\|} & I_{\widehat{m}-1} + \dfrac{u_{(1)}}{\|\bar{u}\|} I_{\widehat{m}-1} - \dfrac{u_{(1)}}{\|\bar{u}\|}\dfrac{\bar{u}\bar{u}^\top}{\|\bar{u}\|^2} \end{pmatrix}.$$

Furthermore, if $a = (a_{(1)}; \bar{a}) \in \mathrm{bd}^+(\mathcal{K}) \subseteq \mathbb{R} \times \mathbb{R}^{\widehat{m}-1}$, $b = \kappa(a_{(1)}; -\bar{a})$, where $\kappa > 0$ is a constant, $\xi \in \partial_B \prod_{\mathcal{K}}(a-b)$, then $b^\top \xi = 0$.

Proof. The first result can be found in Pang et al. [11]. We only show $b^\top \xi = 0$ by calculation,

$$\kappa a_{(1)} + \kappa(-\bar{a}^\top)\frac{(1+\kappa)\bar{a}}{(1+\kappa)\|\bar{a}\|} = \kappa(a_{(1)} - \|\bar{a}\|) = 0,$$

$$\kappa a_{(1)}\frac{(1+\kappa)\bar{a}^\top}{(1+\kappa)\|\bar{a}\|} - \kappa\bar{a}^\top - \kappa\bar{a}^\top\frac{(1-\kappa)a_{(1)}}{(1+\kappa)\|\bar{a}\|} + \kappa\bar{a}^\top\frac{(1-\kappa)a_{(1)}}{(1+\kappa)\|\bar{a}\|} = 0.$$

□

Let $a = g_i(x^*) \in \mathrm{bd}^+(\mathcal{K}_i)$, $b = \mu_i^* = \kappa_i^* R_i g_i(x^*)$, $\xi = U_i \in \partial_B \prod_{\mathcal{K}_i}(g_i(x^*) - \mu_i^*)$. From the above result we have $g_i(x^*)^\top R_i U_i = 0$. (This result is used in Theorem 4.3.)

Acknowledgments. The paper was written when the first author was at the University of Portsmouth as an academic visitor (February 2019–January 2020). The first author wishes to express his sincere thanks to Dr Chee Khian Sim for his advice and help. We would also like to thank the editor and the anonymous referees for their valuable and helpful comments that have improved the quality of this paper greatly.

References

[1] *F. Alizadeh, D. Goldfarb*: Second-order cone programming. Math. Program. *95* (2003), 3–51. zbl MR doi
[2] *J. F. Bonnans, C. H. Ramírez*: Perturbation analysis of second-order cone programming problems. Math. Program. *104* (2005), 205–207. zbl MR doi
[3] *J. F. Bonnans, A. Shapiro*: Perturbation Analysis of Optimization Problems. Springer Series in Operations Research. Springer, New York, 2000. zbl MR doi
[4] *R. W. Freund, F. Jarre, C. H. Vogelbusch*: Nonlinear semidefinite programming: Sensitivity,convergence, and an application in passive reduced-order modeling. Math. Program. *109* (2007), 581–611. zbl MR doi

[5] *E. H. Fukuda, M. Fukushima*: The use of squared slack variables in nonlinear second-order cone programming. J. Optim. Theory Appl. *170* (2016), 394–418. zbl MR doi
[6] *E. H. Fukuda, P. J. S. Silva, M. Fukushima*: Differentiable exact penalty functions for nonlinear second-order cone programs. SIAM J. Optim. *22* (2012), 1607–1633. zbl MR doi
[7] *R. Garcés, W. Gómez, F. Jarre*: A sensitivity result for quadratic semidefinite programs with an application to a sequential quadratic semidefinite programming algorithm. Comput. Appl. Math. *31* (2012), 205–218. zbl MR doi
[8] *C. Kanzow, I. Ferenczi, M. Fukushima*: On the local convergence of semismooth Newton methods for linear and nonlinear second-order cone programs without strict complementarity. SIAM J. Optim. *20* (2009), 297–320. zbl MR doi
[9] *H. Kato, M. Fukushima*: An SQP-type algorithm for nonlinear second-order cone programs. Optim. Lett. *1* (2007), 129–144. zbl MR doi
[10] *M. S. Lobo, L. Vandenberghe, S. Boyd, H. Lebret*: Applications of second-order cone programming. Linear Algebra Appl. *284* (1998), 193–228. zbl MR doi
[11] *J.-S. Pang, D. Sun, J. Sun*: Semismooth homeomorphisms and strong stability of semidefinite and Lorentz cone complementarity problems. Math. Oper. Res. *28* (2003), 39–63. zbl MR doi
[12] *L. Qi, J. Sun*: A nonsmooth version of Newton's method. Math. Program. *58* (1993), 353–367. zbl MR doi
[13] *D. Sun*: The strong second-order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications. Math. Oper. Res. *31* (2006), 761–776. zbl MR doi
[14] *Y. Wang, L. Zhang*: Properties of equation reformulation of the Karush-Kuhn-Tucker condition for nonlinear second order cone optimization problems. Math. Meth. Oper. Res. *70* (2009), 195–218. zbl MR doi

Authors' addresses: *Qi Zhao*, Jiangsu University of Science and Technology, Zhenjiang, 212001, P.R. China, e-mail: `johnzqzq@163.com`; *Wenhao Fu*, School of Mathematical Sciences, Soochow University, Suzhou, 215006, P.R. China, e-mail: `wenhfu@163.com`; *Zhongwen Chen* (corresponding author), School of Mathematical Sciences, Soochow University, Suzhou, 215006, P.R. China, e-mail: `zwchen@suda.edu.cn`.