# A PERTURBATION APPROACH FOR AN INVERSE QUADRATIC PROGRAMMING PROBLEM OVER SECOND-ORDER CONES

### YI ZHANG, LIWEI ZHANG, JIA WU, AND JIANZHONG ZHANG

ABSTRACT. This paper is devoted to studying a type of inverse second-order cone quadratic programming problems, in which the parameters in both the objective function and the constraint set of a given second-order cone quadratic programming problem need to be adjusted as little as possible so that a known feasible solution becomes optimal. This inverse problem can be written as a minimization problem with second-order cone complementarity constraints and a positive semidefinite cone constraint. Applying the duality theory, we reformulate this problem as a linear second-order cone complementarity constrained optimization problem with a semismoothly differentiable objective function, which has fewer variables than the original one. A perturbed problem is proposed with the help of the projection operator over second-order cones, whose feasible set and optimal solution set are demonstrated to be continuous and outer semicontinuous, respectively, as the parameter decreases to zero. A smoothing Newton method is constructed to solve the perturbed problem and its global convergence and local quadratic convergence rate are shown. Finally, the numerical results are reported to show the effectiveness for the smoothing Newton method to solve the inverse second-order cone quadratic programming problem.

### 1. INTRODUCTION

In an optimization problem, all parameters of the model are given, and we need to find from among all feasible solutions an optimal solution for a specified objective function. However, in an inverse optimization problem, the situation is reversed and we try to find values of parameters in an optimization model which make a given feasible solution optimal and which differ from the given estimates as little as possible.

Burton and Toint [3] first investigated an inverse shortest path problem motivated by two interesting examples in traffic flow and seismic tomography. Since then, a variety of inverse combinatorial optimization problems have been studied;

Received by the editor March 18, 2011 and, in revised form, May 13, 2013.

<sup>2010</sup> Mathematics Subject Classification. Primary 90C26, 90C33; Secondary 49M15, 49M20. Key words and phrases. Inverse optimization, second-order cone quadratic programming, perturbation approach, smoothing Newton method.

The first author was supported by the Fundamental Research Funds for the Central Universities under project no. 222201314037 and the China Postdoctoral Science Foundation funded project no. 2013M541479.

The second author was supported by the National Natural Science Foundation of China under project nos. 11071029, 91130007 and 91330206.

The third author was supported by the National Natural Science Foundation of China under project no. 11301049 and the China Postdoctoral Science Foundation funded project no. 2013M541217.

see for instance, [1,4,31]. Recently, several special inverse continuous optimization problems have been investigated. Zhang and Liu [29, 30] first studied the inverse linear programming, and Iyengar and Kang [12] discussed inverse conic programming models and their applications in portfolio optimization. After that, Zhang and Zhang [32], Zhang, Zhang and Xiao [33] and Xiao and Zhang [26] studied numerical methods for two different types of inverse quadratic programming problems. In [26] and [32], parameters in the objective function of a quadratic programming problem need to be estimated and in [33] parameters in both the objective function and the constraint set are required to be adjusted. Numerical algorithms were developed for a type of inverse semidefinite quadratic programming problems in [27] and [28], in which only parameters in the objective function were required to be estimated. In this paper, we focus on an inverse optimization problem with secondorder cone constraints in which parameters in both the objective function and the constraint set are required to be adjusted. Just like the inverse quadratic programming problem addressed in [33], the problem considered in this paper is essentially a mathematical program with second-order cone complementarity constraints, and we propose a perturbation approach for finding an approximation solution.

1.1. Motivation. We take an interesting facility location problem, which is an extension of Weber problem in [8,14], to show that the study of inverse programming over second-order cones is significant in practice.

A manager wants to find the best location for the warehouse of a company, in such a way that the total transportation cost to serve the customers is minimum. Suppose that there are m customers needing to be served and different customers may have different demands, to be translated as weight  $w_i$  for customer i,  $i = 1, 2, \dots, m$ . Let the location of customer i be  $a'_i$ . Denote the desired location of the warehouse to be x' and the additional transportation cost to be a quadratic term with respect to x', say  $\frac{1}{2}x'^T G'x'$ , here G' is a symmetric matrix. Then, the optimization problem is

(1.1) 
$$\min_{x' \in \mathbb{R}^2} \frac{1}{2} x'^T G' x' + \sum_{i=1}^m w_i \|x' - a'_i\|$$

with  $\|\cdot\|$  being the Euclidean norm. In particular, if  $G' = O_2$ , problem (1.1) is reduced to

$$\min_{x' \in \mathbb{R}^2} \sum_{i=1}^m w_i \|x' - a_i'\|,$$

which is the famous Weber problem introduced by German economist Alfred Weber in 1909.

It follows from [2, Section 2.2] that (1.1) can be formulated as a second-order cone programming of the form

(1.2) 
$$\min \quad \frac{1}{2} x'^T G' x' + \sum_{i=1}^m w_i t_i$$
  
s. t.  $\begin{pmatrix} t_i \\ x' - a'_i \end{pmatrix} \in Q_3, i = 1, 2, \cdots, m$ 

where  $Q_3 := \{(s_1; s_2) \in \mathbb{R} \times \mathbb{R}^2 | ||s_2|| \le s_1\}$  is a 3-dimensional second-order cone. We rewrite problem (1.2) as

where  $x = (t_1; \cdots; t_m; x') \in \mathbb{R}^{m+2}, w = (w_1; \cdots; w_m; 0_2) \in \mathbb{R}^{m+2}, a_i = (0; a'_i) \in \mathbb{R}^3$ ,

$$G = \begin{pmatrix} O_m \\ & G' \end{pmatrix} \in \mathbb{R}^{(m+2)\times(m+2)}, \quad A_i = \begin{pmatrix} e_i^T \\ & I_2 \end{pmatrix} \in \mathbb{R}^{3\times(m+2)},$$

 $O_m$  is the zero matrix in  $\mathbb{R}^{m \times m}$ ,  $I_2$  is the identity matrix in  $\mathbb{R}^{2 \times 2}$  and  $e_i$  is the *i*-th unit element in  $\mathbb{R}^m$ ,  $i = 1, 2, \cdots, m$ .

Let  $(G, \hat{w}, \hat{a}_1, \dots, \hat{a}_m)$  be a current estimate of  $(G, w, a_1, \dots, a_m)$  and  $\hat{x}$  an optimal solution of problem (1.3) with  $(G, w, a_1, \dots, a_m) = (\hat{G}, \hat{w}, \hat{a}_1, \dots, \hat{a}_m)$ . As information (e.g. the price of petrol, the capacity and living level of the consumers, target customer segments) changes dynamically, the manager has to judge the efficiency of location  $\hat{x}$  after a period of time. Suppose that the additional transportation cost, the customer locations and the customer demands are changed, and  $(\bar{G}, \bar{w}, \bar{a}_1, \dots, \bar{a}_m)$  is an estimate of  $(G, w, a_1, \dots, a_m)$ , which is obtained by observations, investigations or experience. The manager faces the question of whether the current location  $\hat{x}$  would result in a loss. On the other hand, a new efficient location  $\tilde{x}$  is likely to have a high return. Thus the manager needs to balance these two factors.

Let  $\Phi(\hat{x})$  denote the set of all  $(G, w, a_1, \dots, a_m)$  which make  $\hat{x}$  optimal to problem (1.3), i.e.,

 $\Phi(\hat{x}) = \{ (G, w, a_1, \cdots, a_m) | \hat{x} \text{ is optimal to problem } (1.3) \}.$ 

We consider the following inverse optimization problem:

(1.4) min 
$$\sigma(G, w, a_1, \cdots, a_m) = \|(G, w, a_1, \cdots, a_m) - (\bar{G}, \bar{w}, \bar{a}_1, \cdots, \bar{a}_m)\|^2$$
  
(1.4) s. t.  $(G, w, a_1, \cdots, a_m) \in \Phi(\hat{x}),$ 

where  $\|\cdot\|$  is defined by

$$||(G, w, a_1, \cdots, a_m)|| = \sqrt{\operatorname{Tr}(G^T G) + w^T w + a_1^T a_1 + \cdots + a_m^T a_m}.$$

Let  $\sigma^*$  denote the optimal value of (1.4). It is clear that if  $(\bar{G}, \bar{w}, \bar{a}_1, \dots, \bar{a}_m) \in \Phi(\hat{x})$ , then  $\sigma^* = 0$ , and  $\hat{x}$  is still an optimal location to the new estimate  $(\bar{G}, \bar{w}, \bar{a}_1, \dots, \bar{a}_m)$ , and changes will not be made on  $\hat{x}$ . If  $(\bar{G}, \bar{w}, \bar{a}_1, \dots, \bar{a}_m) \notin \Phi(\hat{x})$ , then  $\sigma^* > 0$  and the strategy to manage the trade-off between the moving costs and higher return depends on the closeness of the new estimate  $(\bar{G}, \bar{w}, \bar{a}_1, \dots, \bar{a}_m)$  to the set  $\Phi(\hat{x})$ , that is, adjustments will be made on the warehouse location  $\hat{x}$  if  $\sigma^* > \alpha$ , where the value of  $\alpha$  is determined by the manager's threshold. In any case, we need to solve problem (1.4), which is a typical inverse quadratic programming problem over second-order cones. 1.2. The problem description and paper organization. We consider a quadratic programming problem over second-order cones of the form

(1.5) (QSOCP) 
$$\min_{\substack{x \in \mathbb{R}^n \\ \text{s. t.}}} f(x) = \frac{1}{2} x^T G x + c^T x$$
  
s. t.  $g_j(x) = A_j x - b_j \in Q_{m_j+1}, j = 1, 2, \cdots, J,$ 

where  $G \in \mathbb{R}^{n \times n}$  is a symmetric matrix,  $c \in \mathbb{R}^n$ ,  $A_j \in \mathbb{R}^{(m_j+1) \times n}$ ,  $b_j \in \mathbb{R}^{m_j+1}$  and  $Q_{m_j+1}$  is a  $m_j + 1$ -dimensional second-order cone defined by

$$Q_{m_j+1} := \{ s = (s_1; s_2) \in \mathbb{R} \times \mathbb{R}^{m_j} | \| s_2 \| \le s_1 \},\$$

with  $\|\cdot\|$  being the Euclidean norm. Let  $S^n$  denote the space of  $n \times n$  symmetric matrices, and let SOL(P) be the set of optimal solutions to a problem (P).

Let  $A_j$  be given for  $j = 1, 2, \dots, J$ . Given a feasible point  $x_0$  to (1.5), which is required to be an optimal solution and a set  $(G^0, c^0, b^0_{m_1+1}, \dots, b^0_{m_J+1}) \in \mathcal{S}^n \times \mathbb{R}^n \times \mathbb{R}^{m_1+1} \times \dots \times \mathbb{R}^{m_J+1}$  which is an estimate to  $(G, c, b_{m_1+1}, \dots, b_{m_J+1})$ . The inverse problem considered in this paper is to find  $(G, c, b_{m_1+1}, \dots, b_{m_J+1}) \in \mathcal{S}^n_+ \times \mathbb{R}^n \times \mathbb{R}^{m_1+1} \times \dots \times \mathbb{R}^{m_J+1}$  to solve the problem:

where  $S^n_+$  is the cone of positive semidefinite symmetric matrices in  $S^n$  and  $\|\cdot\|$  is defined by

$$\| (G', c', b'_{m_1+1}, \cdots, b'_{m_J+1}) \| := \sqrt{\mathrm{Tr}(G'^T G') + c'^T c' + b'_{m_1+1} b'_{m_1+1} + \cdots + b'_{m_J+1} b'_{m_J+1}}$$
  
for  $(G', c', b'_{m_1+1}, \cdots, b'_{m_J+1}) \in \mathcal{S}^n \times \mathbb{R}^n \times \mathbb{R}^{m_1+1} \times \cdots \times \mathbb{R}^{m_J+1}.$ 

Problem (1.6) is a cone-constrained optimization problem with a quadratic objective function, whose scale will be quite large when n is large as the number of decision variables is  $n+n(n+1)/2+\sum_{j=1}^{J}(m_j+1)$ . With the help of duality theory, we formulate (1.6) as a linear second-order cone complementarity constrained optimization problem (3.10) (see Section 3) with semismoothly differentiable objective function which is essentially a mathematical program with second-order cone complementarity constraints and has fewer decision variables than the original inverse problem.

The mathematical program with second-order cone complementarity constraints includes the mathematical program with (vector) complementarity constraints (denoted as MPCC for short) as a special case. For MPCCs, there have been proposed many algorithms such as the sequential quadratic programming approach, the implicit programming approach, the penalty function approach, the reformulation approach, and so on; see [6,9,13,15,20,21] for references. However, there are only a few studies on the mathematical program with second-order cone complementarity constraints. Zhang, Zhang and Wu [34] approximated the second-order cone complementarity constraints by the smoothing projection operator over second-order cones and introduced a smoothing approach. Wu, Zhang and Zhang [25] employed a smoothing Newton method for the mathematical programs governed by secondorder cone constrained parameterized generalized equations. However, the analysis in these two papers requires twice differentiability of the objective function. In this paper we will propose a perturbation approach to solve the mathematical program with second-order cone complementarity constraints under a weaker condition, the objective function is not twice differentiable, but semismoothly differentiable. Furthermore, a smoothing Newton method is employed to solve the perturbed problems and numerical results reported show that the approach we adopted is quite effective.

The organization of this paper is as follows. Section 2 reviews several basic results in nonsmooth analysis and some properties of a second-order cone. Section 3 is devoted to reformulating the inverse quadratic programming problem over second-order cones as a linear second-order cone complementarity constrained optimization problem with a semismoothly differentiable objective function. In Section 4, we use a smoothing function to approximate the complementarity relation and demonstrate the convergence behavior of the perturbed problem. In Section 5, a smoothing Newton method is constructed to solve the perturbed problem and its global convergence and local quadratic convergence rate is shown. Finally, we report the numerical results of the smoothing Newton method for solving the perturbed problem in Section 6.

The following notations are used throughout the paper. For any vectors  $x, y \in \mathbb{R}^n$  and matrices  $A, B \in \mathbb{R}^{n \times n}$ , the inner product  $\langle x, y \rangle := x^T y, \langle A, B \rangle = \operatorname{Tr}(A^T B)$ and the norm  $||x|| := \sqrt{x^T x}, ||A||_F := \sqrt{\operatorname{Tr}(A^T A)}$ . We denote  $I_n$  and  $O_n$  as the identity matrix and zero matrix in  $\mathbb{R}^{n \times n}$ . We use ";" for adjoining vectors in a column, for example, vector  $d = (d_1; d_2) = (d_1^T d_2^T)^T \in \mathbb{R}^{m_1+m_2}$ , with  $d_1 \in \mathbb{R}^{m_1}$ and  $d_2 \in \mathbb{R}^{m_2}$ . diag  $(B_1, B_2)$  is denoted as the block diagonal matrix whose block diagonal entries are symmetric matrices  $B_1$  and  $B_2$ . For two matrices A and B, we write  $A \succeq B(A \succ B)$  to mean that A - B is positive semidefinite (positive definite). For a differentiable mapping  $F : \mathbb{R}^n \to \mathbb{R}^m$  and a vector  $z \in \mathbb{R}^n$ , we denote by  $\mathcal{J}F(z)$  the Jacobian matrix of F at z and  $\nabla F(z) := \mathcal{J}F(z)^T$ .

### 2. Preliminaries

2.1. Background in nonsmooth analysis and variational analysis. Now we give some basic concepts on nonsmooth analysis. The first one is the general Jacobian of a Lipschitz continuous function.

**Definition 2.1.** Let X and Y be two finite dimensional real vector spaces. Let  $\mathcal{O}$  be an open set in X and  $\Phi : \mathcal{O} \subseteq X \to Y$  be a locally Lipschitz continuous function on the open set  $\mathcal{O}$ . By Rademacher's theorem,  $\Phi$  is almost everywhere Fréchet-differentiable in  $\mathcal{O}$ . We denote by  $\mathcal{D}_{\Phi}$  the set of Fréchet-differentiable points of  $\Phi$  in  $\mathcal{O}$ . Then, the Bouligand-subdifferential of  $\Phi$  at  $x \in \mathcal{O}$ , denoted as  $\partial_B \Phi(x)$ , is

$$\partial_B \Phi(x) := \left\{ \lim_{k \to \infty} \mathcal{J} \Phi(x^k) | x^k \in \mathcal{D}_\Phi, x^k \to x \right\},$$

where  $\mathcal{J}\Phi(x^k)$  is the Jacobian of  $\Phi$  at  $x^k$ . The generalized Jacobian in the sense of Clarke is the convex hull of  $\partial_B \Phi(x)$ , i.e.,

$$\partial \Phi(x) = \operatorname{conv} \{ \partial_B \Phi(x) \}.$$

The following concept of semismoothness was extended by Qi and Sun [17] from real valued functions to vector valued functions.

**Definition 2.2.** Let X and Y be two finite dimensional real vector spaces. Let  $\Phi : \mathcal{O} \subseteq X \to Y$  be a locally Lipschitz continuous function on the open set  $\mathcal{O}$ . We say that  $\Phi$  is semismooth at a point  $x \in \mathcal{O}$  if:

- (i)  $\Phi$  is directionally differentiable at x; and
- (ii) for any  $\Delta x \in X$  and  $V \in \partial \Phi(x + \Delta x)$  with  $\Delta x \to 0$ ,

$$\Phi(x + \Delta x) - \Phi(x) - V(\Delta x) = o(\|\Delta x\|)$$

Furthermore,  $\Phi$  is said to be strongly semismooth at  $x \in \mathcal{O}$  if  $\Phi$  is semismooth at x and for any  $\Delta x \in X$  and  $V \in \partial \Phi(x + \Delta x)$  with  $\Delta x \to 0$ ,

$$\Phi(x + \Delta x) - \Phi(x) - V(\Delta x) = O(||\Delta x||^2).$$

The definition of  $SC^1$  function in the sense of [10] is given below, which will be used in Section 3.

**Definition 2.3.** A function  $\tau : \mathbb{R}^n \to \mathbb{R}$  is said to be a  $SC^1$  (semismoothly differentiable) function on an open set  $\mathcal{O} \subseteq \mathbb{R}^n$  if  $\tau$  is continuously differentiable on  $\mathcal{O}$  and  $\nabla \tau$  is semismooth on  $\mathcal{O}$ . The generalized Hessian of  $\tau$  at x is defined to be the set  $\partial^2 \tau(x)$  of  $n \times n$  matrices by

$$\partial^2 \tau(x) = \operatorname{conv}\{\partial_B[\nabla \tau](x)\}$$

where  $\partial_B[\nabla \tau](x)$  is the B-differential of  $\nabla \tau$  at x, which can be expressed as

$$\partial_B[\nabla \tau](x) = \{ H \in \mathbb{R}^{n \times n} \mid \exists \ x^k \to x \text{ with } \nabla \tau \text{ differentiable at } x^k \text{ and } \nabla^2 \tau(x^k) \to H \}.$$

For a  $SC^1$  function  $\tau$  on an open set  $\mathcal{O} \subseteq \mathbb{R}^n$ , we have the following second-order Taylor-like formula and mean value formula; see [7]. For a segment  $[y, y'] \subset \mathcal{O}$ , there is a matrix  $H \in \partial^2 \tau(w)$  at some  $w \in (y, y')$  such that

$$\tau(y') = \tau(y) + \nabla \tau(y)^T (y' - y) + \frac{1}{2} (y' - y)^T H(y' - y)$$

and there exist  $l(l \ge n)$  points  $w^1, \ldots, w^l \in (y, y')$  and  $H^i \in \partial^2 \tau(w^i)$  for  $i = 1, \ldots, l$  such that

(2.1) 
$$\nabla \tau(y') = \nabla \tau(y) + \sum_{i=1}^{l} t_i H^i(y'-y),$$

where  $t_i \ge 0, i = 1, ..., l$ , and  $t_1 + \dots + t_l = 1$ .

Let Y be a finite dimensional real vector space and K a closed convex set in Y. For instance, the convex set K will be chosen as the convex cone  $S^n_+$ ,  $Q_{m+1}$  or  $\mathbb{R}^n_+$  in the following sections. It is well known that the metric projector  $\Pi_K(\cdot)$  is Lipschitz continuous with the Lipschitz constant 1 and semismooth everywhere. Then for any  $y \in Y$ ,  $\partial \Pi_K(y)$  is well defined. Below is a lemma on some general properties of  $\partial \Pi_K(\cdot)$ .

**Lemma 2.1** ([16, Proposition 1]). Let  $K \subseteq Y$  be a closed convex set. Then, for any  $y \in Y$  and  $V \in \partial \Pi_K(y)$ , it holds that: (i) V is self-adjoint, (ii)  $\langle d, Vd \rangle \ge 0$ ,  $\forall d \in Y$ ,

(iii)  $\langle Vd, d-Vd \rangle \ge 0, \quad \forall d \in Y.$ 

The concepts of the semicontinuity and continuity of a set-valued mapping in [19, Definition 5.4] and the epi-continuity properties of a function-valued mapping in [19, Definition 7.1] are given below.

**Definition 2.4.** A set-valued mapping  $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  is outer semicontinuous at  $\bar{x}$  if  $\limsup_{x \to \bar{x}} S(x) \subset S(\bar{x})$ , but inner semicontinuous at  $\bar{x}$  if  $\liminf_{x \to \bar{x}} S(x) \supset S(\bar{x})$ , where

$$\limsup_{x \to \bar{x}} S(x) := \{ u | \exists x^k \to x, \exists u^k \to u \quad \text{with} \quad u^k \in S(x^k) \},$$
$$\liminf_{x \to \bar{x}} S(x) := \{ u | \forall x^k \to x, \exists u^k \to u \quad \text{with} \quad u^k \in S(x^k) \}.$$

It is called continuous at  $\bar{x}$  if both conditions hold, i.e., if  $\lim_{x \to \bar{x}} S(x) = S(\bar{x})$ .

**Definition 2.5.** For a function  $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ , the function-valued mapping  $\varepsilon \to f(\cdot, \varepsilon)$  is said to be epi-continuous at  $\overline{\varepsilon}$ , denoted as  $e - \lim_{\varepsilon \to \overline{\varepsilon}} f(\cdot, \varepsilon) = f(\cdot, \overline{\varepsilon})$ , if

$$\lim_{\varepsilon \to \bar{\varepsilon}} \operatorname{epi} f(\cdot, \varepsilon) = \operatorname{epi} f(\cdot, \bar{\varepsilon}).$$

where  $\operatorname{epi} f(\cdot, \varepsilon) := \{ (x, \alpha) \in \mathbb{R}^n \times \mathbb{R} | f(x, \varepsilon) \le \alpha \}.$ 

2.2. Basic results for a second-order cone. In this subsection, we review some basic results for a second-order cone.

Let  $Q_{m+1}$  be a second-order cone of dimension m + 1. The topological interior part and the boundary of  $Q_{m+1}$  denoted by  $\operatorname{int} Q_{m+1}$  and  $\operatorname{bd} Q_{m+1}$ , respectively, are given by

int
$$Q_{m+1} := \{ s = (s_1; s_2) \in \mathbb{R} \times \mathbb{R}^m \mid ||s_2|| < s_1 \},\$$
  
bd $Q_{m+1} := \{ s = (s_1; s_2) \in \mathbb{R} \times \mathbb{R}^m \mid ||s_2|| = s_1 \}.$ 

For any  $x = (x_1; x_2) \in \mathbb{R} \times \mathbb{R}^m$  and  $y = (y_1; y_2) \in \mathbb{R} \times \mathbb{R}^m$ , we define their Jordan product as

$$x \circ y = (x^T y; y_1 x_2 + x_1 y_2).$$

The identity element under this product is  $e := (1; 0; \dots; 0) \in \mathbb{R}^{m+1}$ . We write  $x^2$  to mean  $x \circ x$  and write x + y to mean the usual componentwise addition of vectors. It is known that  $x^2 \in Q_{m+1}$  for all  $x \in \mathbb{R}^{m+1}$ . Moreover, if  $x \in Q_{m+1}$ , there exists a unique vector in  $Q_{m+1}$  denoted by  $\sqrt{x}$ , such that  $(\sqrt{x})^2 = \sqrt{x} \circ \sqrt{x} = x$ . For  $x \in \mathbb{R}^{m+1}$ , we define its determinant as  $\det(x) = x_1^2 - \|x_2\|^2$ .

We introduce the spectral factorization of vectors in  $\mathbb{R}^{m+1}$  associated with  $Q_{m+1}$ . Let  $x = (x_1; x_2) \in Q_{m+1}$ , then x can be decomposed as

$$x = \lambda_1 c_1 + \lambda_2 c_2,$$

where  $\lambda_1$ ,  $\lambda_2$  and  $c_1$ ,  $c_2$  are the spectral values and the associated spectral vectors of x given by

$$\begin{split} \lambda_i &= x_1 + (-1)^i \|x_2\|, \\ c_i &= \begin{cases} & \frac{1}{2} \left( 1, (-1)^i \frac{x_2}{\|x_2\|} \right), & \text{if } x_2 \neq 0, \\ & \frac{1}{2} \left( 1, (-1)^i w \right), & \text{if } x_2 = 0, \end{cases} \end{split}$$

for i = 1, 2, with w being any vector in  $\mathbb{R}^m$  satisfying ||w|| = 1. By [11], we also know that

$$\sqrt{x} = \sqrt{\lambda_1}c_1 + \sqrt{\lambda_2}c_2.$$

For any  $x = (x_1; x_2) \in \mathbb{R} \times \mathbb{R}^m$ , we define the linear mapping  $L_x$  from  $\mathbb{R}^{m+1}$  to  $\mathbb{R}^{m+1}$  as

$$L_x y := \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I_m \end{bmatrix} y.$$

It can be easily verified that  $x \circ y = L_x y$ ,  $\forall y \in \mathbb{R}^{m+1}$ , and  $L_x$  is positive definite (and hence invertible) if and only if  $x \in intQ_{m+1}$ . Also, we have

$$L_x^{-1} = \frac{1}{\det(x)} \begin{bmatrix} x_1 & -x_2^T \\ -x_2 & \frac{\det(x)}{x_1} I_m + \frac{x_2 x_2^T}{x_1} \end{bmatrix}.$$

More interesting properties about the spectral factorization of x, as well as  $x^2$ ,  $\sqrt{x}$  and the matrix  $L_x$ , can be found in [11].

The following consequence from [2, Lemma 15] will be used in Section 4.

**Lemma 2.2.** Let  $x = (x_1; x_2) \in Q_{m+1}$ ,  $y = (y_1; y_2) \in Q_{m+1}$  and  $x \circ y = 0$ . Then we have either x = 0 or y = 0, or there exists  $\sigma > 0$  such that  $x = \sigma(y_1; -y_2)$ .

### 3. PROBLEM REFORMULATION

If  $G \in \mathcal{S}^n_+$ , then  $x_0 \in \text{SOL}(\text{QSOCP})$  if and only if there exist  $u_j \in Q_{m_j+1}$  such that

(3.1) 
$$\begin{aligned} c + Gx_0 - \sum_{j=1}^J A_j^T u_j &= 0, \\ u_j \circ (A_j x_0 - b_j) &= 0, \ j = 1, 2, \cdots, J. \end{aligned}$$

Let Q be the Cartesian product of second-order cones, that is,  $Q := Q_{m_1+1} \times Q_{m_2+1} \times \cdots \times Q_{m_J+1}$  and  $q = \sum_{j=1}^{J} (m_j + 1)$ . Denote

$$A := (A_1; A_2; \cdots; A_J) \in \mathbb{R}^{q \times n},$$
  

$$v := (v_1; v_2; \cdots; v_J) \in Q, \ v_j := A_j x^0 - b_j \in Q_{m_j+1},$$
  

$$v^0 := (v_1^0; v_2^0; \cdots; v_J^0) \in Q, \ v_j^0 := A_j x^0 - b_j^0 \in Q_{m_j+1},$$
  

$$u := (u_1; u_2; \cdots; u_J) \in Q,$$
  

$$u \circ v = ((u_1 \circ v_1); (u_2 \circ v_2); \cdots; (u_J \circ v_J)).$$

Then combining with (3.1), we have an equivalent formulation of (1.6) as follows:

(3.2)  

$$\min \quad \frac{1}{2} \| (G - G^0, c - c^0, v - v^0) \|^2$$
s. t.  $c + Gx_0 - A^T u = 0,$   
 $u \circ v = 0,$   
 $(G, c, u, v) \in \mathcal{S}^n_+ \times \mathbb{R}^n \times Q \times Q$ 

First, we consider a subproblem of (3.2) below:

(3.3) 
$$\min \ \frac{1}{2} \| (G - G^0, c - c^0) \|^2$$
  
s. t.  $c + Gx_0 - A^T u = 0,$   
 $(G, c) \in \mathcal{S}^n_+ \times \mathbb{R}^n,$ 

which is a convex programming parameterized by  $u \in Q$ . Let its optimal value be  $f_0(u)$ , i.e.,

(3.4)

$$f_0(u) := \frac{1}{2} \inf \left\{ \|G - G^0\|_F^2 + \|c - c^0\|^2 \|c + Gx_0 - A^T u = 0, (G, c) \in \mathcal{S}^n_+ \times \mathbb{R}^n \right\}.$$

Define

$$f(u,v) := \frac{1}{2} \|v - v^0\|^2 + f_0(u).$$

For any  $u \in Q$ , it is obvious that the generalized Slater constraint qualification holds for problem (3.3). Therefore, by the classical duality theory for convex programming, there is no duality gap between problem (3.3) and its dual. Let  $L: S^n_+ \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$  be the Lagrange function of (3.3), defined by

$$L(G, c, x) := \frac{1}{2} \|G - G^0\|_F^2 + \frac{1}{2} \|c - c^0\|^2 + \langle x, c + Gx_0 - A^T u \rangle.$$

The Lagrange dual problem of (3.3) is

(3.5) 
$$\max_{x \in \mathbb{R}^n} \nu(x) := \inf_{(G,c) \in \mathcal{S}^n_+ \times \mathbb{R}^n} L(G,c,x).$$

Let  $\mathcal{B}: \mathbb{R}^n \to \mathcal{S}^n$  defined by

$$\mathcal{B}x := \frac{xx_0^T + x_0^T x}{2},$$

the adjoint of  $\mathcal{B}$  is denoted by  $\mathcal{B}^*$ . Define

$$\bar{G}(x) := G^0 - \mathcal{B}x,$$
  

$$\Psi(x) := \|\Pi_{\mathcal{S}^n_+}(\bar{G}(x))\|_F^2 = \operatorname{Tr}\left(\Pi_{\mathcal{S}^n_+}(\bar{G}(x))^T \Pi_{\mathcal{S}^n_+}(\bar{G}(x))\right)$$

The Moreau envelop function and proximal mapping of  $\Psi$  are defined by Morea-Yosida regularization (see [19, Definition 1.22]), i.e.,

$$e_{\lambda}\Psi(x) := \inf_{\omega} \left\{ \Psi(\omega) + \frac{1}{2\lambda} \|\omega - x\|^2 \right\}$$

and

$$P_{\lambda}\Psi(x) := \operatorname{Argmin}_{\omega} \left\{ \Psi(\omega) + \frac{1}{2\lambda} \|\omega - x\|^2 \right\},$$

respectively. If  $\Psi$  is lower semicontinuous, proper and convex, by [19, Theorem 2.26], we get that  $e_{\lambda}\Psi$  is convex, and continuously differentiable with

(3.6) 
$$\nabla e_{\lambda}\Psi(x) = \frac{1}{\lambda}[x - P_{\lambda}\Psi(x)].$$

Furthermore,  $\nabla e_{\lambda} \Psi$  is Lipschitz continuous.

**Lemma 3.1.** The function  $\nu(x)$  defined in (3.5) and the function  $f_0(u)$  defined in (3.4) can be expressed by

(3.7) 
$$\nu(x) = -\frac{1}{2} \|x\|^2 + \langle c^0 - A^T u, x \rangle - \frac{1}{2} \|\Pi_{\mathcal{S}^n_+}(\bar{G}(x))\|_F^2 + \frac{1}{2} \|G^0\|_F^2$$

and

$$f_0(u) = \frac{1}{2} \|G^0\|_F^2 + \frac{1}{2} \|c^0 - A^T u\|^2 - \frac{1}{2} e_{1/2} \Psi(c^0 - A^T u),$$

respectively.

*Proof.* From the definition of  $\nu$  , we have

$$\nu(x) = \inf_{(G,c)\in S^n_+\times\mathbb{R}^n} \left\{ \frac{1}{2} \|c - c^0\|^2 + \langle c, x \rangle + \frac{1}{2} \|G - G^0\|_F^2 + \langle Gx_0, x \rangle - \langle A^T u, x \rangle \right\}$$
  
=  $-\langle A^T u, x \rangle + \inf_{c\in\mathbb{R}^n} \left\{ \frac{1}{2} \|c - c^0\|^2 + \langle c, x \rangle \right\} + \inf_{G\in S^n_+} \left\{ \frac{1}{2} \|G - G^0\|_F^2 + \langle Gx_0, x \rangle \right\}.$ 

As the unconstrained quadratic programming problem

$$\min_{c \in \mathbb{R}^n} \frac{1}{2} \|c - c^0\|^2 + \langle c, x \rangle$$

takes its minimum at

(3.8) 
$$c^*(x) = c^0 - x,$$

we have that

$$\inf_{c \in \mathbb{R}^n} \left\{ \frac{1}{2} \|c - c^0\|^2 + \langle c, x \rangle \right\} = -\frac{1}{2} \|x\|^2 + c^{0T} x.$$

From the expression

$$\inf_{G \in \mathcal{S}^{n}_{+}} \left\{ \frac{1}{2} \| G - G^{0} \|_{F}^{2} + \langle Gx_{0}, x \rangle \right\}$$
$$= \inf_{G \in \mathcal{S}^{n}_{+}} \left\{ \frac{1}{2} [\| G - \bar{G}(x) \|_{F}^{2} + 2 \langle G^{0}, \mathcal{B}x \rangle - \| \mathcal{B}x \|_{F}^{2}] \right\},$$

we know that the minimum value is reached at

(3.9) 
$$G^*(x) = \prod_{\mathcal{S}^n_+} (\bar{G}(x)),$$

and thus

$$\inf_{G \in \mathcal{S}_{+}^{n}} \left\{ \frac{1}{2} \| G - G^{0} \|_{F}^{2} + \langle Gx_{0}, x \rangle \right\} 
= \frac{1}{2} \left[ \| \bar{G}(x) - \Pi_{\mathcal{S}_{+}^{n}} \left( \bar{G}(x) \right) \|_{F}^{2} - \| \bar{G}(x) \|_{F}^{2} + \| G^{0} \|_{F}^{2} \right], 
= -\frac{1}{2} \| \Pi_{\mathcal{S}_{+}^{n}} \left( \bar{G}(x) \right) \|_{F}^{2} + \frac{1}{2} \| G^{0} \|_{F}^{2}.$$

Therefore, the function  $\nu(x)$  has the expression (3.7).

It follows from the zero duality gap property between (3.4) and (3.5) that

$$\begin{split} f_{0}(u) &= \max_{x} \nu(x) \\ &= \max_{x} \left\{ -\frac{1}{2} \|x\|^{2} + \langle c^{0} - A^{T}u, x \rangle - \frac{1}{2} \|\Pi_{\mathcal{S}^{n}_{+}} \left(\bar{G}(x)\right)\|_{F}^{2} + \frac{1}{2} \|G^{0}\|_{F}^{2} \right\} \\ &= \frac{1}{2} \|G^{0}\|_{F}^{2} + \frac{1}{2} \|c^{0} - A^{T}u\|^{2} - \frac{1}{2} \min_{x} \left\{ \|\Pi_{\mathcal{S}^{n}_{+}} \left(\bar{G}(x)\right)\|_{F}^{2} + \|x - (c^{0} - A^{T}u)\|^{2} \right\} \\ &= \frac{1}{2} \|G^{0}\|_{F}^{2} + \frac{1}{2} \|c^{0} - A^{T}u\|^{2} - \frac{1}{2} \min_{x} \left\{ \Psi(x) + \|x - (c^{0} - A^{T}u)\|^{2} \right\} \\ &= \frac{1}{2} \|G^{0}\|_{F}^{2} + \frac{1}{2} \|c^{0} - A^{T}u\|^{2} - \frac{1}{2} e_{1/2} \Psi(c^{0} - A^{T}u). \end{split}$$

The proof is completed.

In view of the above lemma, (3.2) can be equivalently expressed as

(3.10) 
$$\min_{\substack{s. t. u \in Q, v \in Q, \\ u \in Q, v \in Q, \\ }} \frac{\min_{\substack{s. t. u \in Q, v \in Q, \\ u \in Q, v \in Q, \\ }} f(u,v) = \frac{1}{2} \|v - v^0\|^2 + \frac{1}{2} \|G^0\|_F^2 + \frac{1}{2} \|c^0 - A^T u\|^2 - \frac{1}{2} e_{1/2} \Psi(c^0 - A^T u)$$

which is a mathematical program with linear second-order cone complementarity constraints.

Note that the dimension of the above problem is O(q), much smaller than that of problem (3.2) when n is large, which is  $O(n^2)$ . Now we focus on discussing how to solve problem (3.10).

Let

$$\begin{aligned} x(u) &:= P_{1/2} \Psi(c^0 - A^T u) \\ &= \operatorname{Argmin}_x \left\{ \| \Pi_{\mathcal{S}^n_+} \left( \bar{G}(x) \right) \|_F^2 + \| x - (c^0 - A^T u) \|^2 \right\}, \end{aligned}$$

so that

$$e_{1/2}\Psi(c^0 - A^T u) = \|\Pi_{\mathcal{S}^n_+} \left( \bar{G}(x(u)) \right) \|_F^2 + \|x(u) - (c^0 - A^T u)\|^2.$$

If (3.10) has an optimal solution  $(u^*, v^*)$ , then we know from (3.8) and (3.9) that

$$(G^*, c^*, b^*) = \left(\Pi_{\mathcal{S}^n_+}(\bar{G}(x(u^*))), c^0 - x(u^*), Ax_0 - v^*\right)$$

is an optimal solution to the original problem (1.6).

**Lemma 3.2.** The function  $\Psi(x)$  is convex and the mapping x(u) is strongly semismooth. Furthermore,

 $\partial_B x(u) \Delta u \subset \left\{ -\left[\mathcal{B}^* W \mathcal{B} + I_n\right]^{-1} A^T \Delta u | W \in \partial_B \Pi_{\mathcal{S}^n_+} \left( \bar{G} \left( x(u) \right) \right) \right\}, \quad \forall \ \Delta u \in \mathbb{R}^q,$ where  $\mathcal{B}^*$  is the adjoint operator of  $\mathcal{B}$ .

*Proof.* Let  $g(x) = \prod_{\mathcal{S}^n_+} (\bar{G}(x))$ , then  $\Psi(x) = ||g(x)||_F^2$ . Function  $\Psi$  is a SC<sup>1</sup> function with

$$\nabla \Psi(x) = 2\mathcal{J}\bar{G}(x)^*g(x) = -2\mathcal{B}^*g(x)$$

It follows from [24, Lemma 3] that

$$\partial^2 \Psi(x) = -2\mathcal{B}^* \partial g(x).$$

From [23, Lemma 2.1], we have

$$\partial g(x) \subset \partial \Pi_{\mathcal{S}^n_+} \left( \bar{G}(x) \right) \mathcal{J} \bar{G}(x) = -\partial \Pi_{\mathcal{S}^n_+} \left( \bar{G}(x) \right) \mathcal{B},$$

and therefore

$$\partial^2 \Psi(x) \subset 2\mathcal{B}^* \partial \Pi_{\mathcal{S}^n_+} \left( \bar{G}(x) \right) \mathcal{B}, \quad \forall x \in \mathbb{R}^n$$

For any  $x, x' \in \mathbb{R}^n$ , it follows from (2.1) that there exist  $l \ge n$  points  $w^1, \dots, w^l \in (x, x')$  and  $H^i \in \partial^2 \Psi(w^i)$  for  $i = 1, \dots, l$  such that

(3.11) 
$$\nabla \Psi(x') - \nabla \Psi(x) = \sum_{i=1}^{l} t_i H^i (x' - x).$$

where  $t_i \ge 0$ ,  $i = 1, \dots, l$  and  $t_1 + \dots + t_l = 1$ . In view of (ii) of Definition 2.2, we know from (3.11) that each  $H^i$  is positively semidefinite, which implies that

$$\langle \nabla \Psi(x') - \nabla \Psi(x), x' - x \rangle \ge 0, \quad \forall x', x \in \mathbb{R}^n,$$

namely the gradient mapping  $\nabla \Psi$  is a monotone operator. Thus from [19, Theorem 12.17], we obtain that  $\Psi$  is a convex function.

From the definition of x(u), we know that x(u) is the unique solution of the following strongly semismooth system

(3.12) 
$$-2\mathcal{B}^*\Pi_{\mathcal{S}^n_+}(\bar{G}(x)) + 2[x - (c^0 - A^T u)] = 0.$$

Since the generalized Jacobian of the left-hand side with respect to x is contained in  $2(I + \mathcal{B}^* \partial \Pi_{\mathcal{S}^n_+} (\bar{G}(x)) \mathcal{B})$ , and its every element is nonsingular, by [24, Lemma 1], we know that x(u) is strongly semismooth and

(3.13) 
$$-\mathcal{B}^*\Pi_{\mathcal{S}^n_+}\left(\bar{G}\left(x(u)\right)\right) + x(u) - (c^0 - A^T u) = 0$$

Let  $\mathcal{D}_x$  denote the set of  $u' \in \mathbb{R}^q$  such that  $\mathcal{J}x(u')$  exists. Let  $Z \in \partial_B x(u)$ . Then there is a sequence  $u^k \subset \mathcal{D}_x$  with  $u_k \to u$ , such that  $\mathcal{J}x(u^k) \to Z$ . For any  $\Delta u \in \mathbb{R}^q$ and each k, we have from (3.13) that

$$-\mathcal{B}^*\Pi'_{\mathcal{S}^n_+}\left(\bar{G}(x(u_k), -\mathcal{B}\mathcal{J}x(u_k)\Delta u\right) + \mathcal{J}x(u^k)\Delta u + A^T\Delta u = 0.$$

As  $\Pi_{\mathcal{S}^n_+}(\cdot)$  is strongly semismooth, we know that there is  $W^k_{\Delta u} \in \partial_B \Pi_{\mathcal{S}^n_+}(\bar{G}(x(u_k)))$ such that  $\Pi'_{\mathcal{S}^n_+}(\bar{G}(x(u_k), -\mathcal{BJ}x(u_k)\Delta u) = W^k_{\Delta u}[-\mathcal{BJ}x(u_k)\Delta u]$ . Therefore,

$$\mathcal{J}x(u_k)\Delta u = -\left[\mathcal{B}^* W_{\Delta u}^k \mathcal{B} + I_n\right]^{-1} A^T \Delta u_k$$

which leads to

$$\mathcal{I}x(u_k)\Delta u \in \left\{-\left[\mathcal{B}^*W_{\Delta u}^k \mathcal{B} + I_n\right]^{-1} A^T \Delta u | W_{\Delta u}^k \in \partial_B \Pi_{\mathcal{S}^n_+} \left(\bar{G}\left(x(u_k)\right)\right)\right\}.$$

As  $\partial_B \Pi_{\mathcal{S}^n_+}(\cdot)$  is outer semicontinuous, taking  $k \to \infty$ , we obtain

$$Z\Delta u \in \left\{ -\left[\mathcal{B}^*W\mathcal{B} + I_n\right]^{-1} A^T \Delta u | W \in \partial_B \Pi_{\mathcal{S}^n_+} \left(\bar{G}\left(x(u)\right)\right) \right\},\$$

which yields the inclusion.

**Lemma 3.3.** The function f is convex with

(3.14) 
$$\nabla f(u,v) = \begin{bmatrix} -Ax(u) \\ v - v^0 \end{bmatrix}$$

and (3.15)

$$\partial_{B}[\nabla f](u,v)\left(\Delta u;\Delta v\right) \subset \left\{ \begin{bmatrix} A[\mathcal{B}^{*}W\mathcal{B}+I_{n}]^{-1}A^{T}\Delta u\\ \Delta v \end{bmatrix} | W \in \partial_{B}\Pi_{\mathcal{S}^{n}_{+}}\left(\bar{G}(x(u))\right) \right\},$$
(3.16)

$$\partial [\nabla f](u,v) \left(\Delta u; \Delta v\right) \subset \left\{ \left[ \begin{array}{c} A[\mathcal{B}^*W\mathcal{B} + I_n]^{-1}A^T\Delta u \\ \Delta v \end{array} \right] | W \in \partial \Pi_{\mathcal{S}^n_+} \left( \bar{G}(x(u)) \right) \right\}.$$

Furthermore, if A has full rank, then f is strongly convex on any compact set  $\Gamma \subset \mathbb{R}^q \times \mathbb{R}^q$ . Namely, there is some constant  $\sigma > 0$  such that

$$\begin{split} f[(1-t)(u',v')+t(u,v)] &\leq (1-t)f(u',v')+tf(u,v)-\frac{1}{2}\sigma t(1-t)\|(u'-u,v'-v)\|^2,\\ when \ t \in (0,1) \ and \ (u',v'), (u,v) \in \Gamma. \end{split}$$

*Proof.* It follows from (3.6) that

$$\nabla_u \mathbf{e}_{1/2} \Psi(c^0 - A^T u) = -2A[c^0 - A^T u - x(u)].$$

Hence we obtain the expression (3.14) for  $\nabla f(u, v)$  from (3.10). By (3.14) and Lemma 3.2, we obtain

$$\partial_B \left[ \nabla f \right](u, v)(\Delta u; \Delta v) = \left\{ \begin{bmatrix} -AZ\Delta u \\ \Delta v \end{bmatrix} | Z \in \partial_B x(u) \right\}$$
$$\subset \left\{ \begin{bmatrix} A [\mathcal{B}^*W\mathcal{B} + I_n]^{-1}A^T\Delta u \\ \Delta v \end{bmatrix} | W \in \partial_B \Pi_{\mathcal{S}^n_+} \left( \bar{G}(x(u)) \right) \right\}.$$

The inclusion (3.16) comes from (3.15).

Similar to the proof for the function  $\Psi$  in Lemma 3.2, we can use (2.1), (3.16) and (ii) of Definition 2.2 to prove that  $\nabla f$  is a monotone operator, and hence f is a convex function.

For a compact set  $\Gamma \subset \mathbb{R}^q \times \mathbb{R}^q$ , define

$$\gamma_0 := \sup_{u \in \Pi_u(\Gamma)} \left\{ \|\mathcal{B}^* W \mathcal{B}\| | W \in \partial \Pi_{\mathcal{S}^n_+} \left( \bar{G}(x(u)) \right) \right\},\$$

where  $\Pi_u(\Gamma)$  is the projection of  $\Gamma$  onto *u*-space, then  $\gamma_0$  is well defined and finite. Let  $\underline{\gamma} := \min\{1, (\gamma_0 + 1)^{-1}\lambda_{\min}(AA^T)\}$ . Then we have, from the assumption that A is of full row rank and by the mean value formula, that  $\gamma > 0$  and

$$\langle \nabla f(u',v') - \nabla f(u,v), (u'-u,v'-v) \rangle \ge \underline{\gamma} \| (u'-u,v'-v) \|^2.$$

Namely  $\nabla f$  is strongly monotone and hence f is strongly convex with constant  $\underline{\gamma}$ .

#### 4. Perturbation Approach and convergence behavior

Now, we discuss how to solve (3.10). Due to the second-order cone complementarity constraints in (3.10), Robinson's constraint qualifications do not hold at any feasible point (see [35]), then the KKT conditions may fail at any local minimizer. To overcome this difficulty, we choose a smoothing function  $\varphi_{\mu}(u, v) = 0$  to approximate the second-order cone complementarity relation  $u \circ v = 0$ ,  $u \in Q_{m+1}$ ,  $v \in Q_{m+1}$ , where  $\varphi_{\mu}(u, v) = 0$  is defined by

$$\varphi_{\mu}(u,v) = u + v - \sqrt{(v-u)^2 + 4\mu^2 e}$$

with e the identity element in  $\mathbb{R}^{m+1}$  and  $\mu > 0$  a parameter. It is easily verified that  $\lim_{\mu \searrow 0} \varphi_{\mu}(u, v) = \varphi_{0}(u, v)$  and  $\varphi_{0}(u, v) = 0$  if and only if  $u \circ v = 0$ ,  $u \in Q_{m+1}$ ,  $v \in Q_{m+1}$ .

Define function  $\Phi_{\mu} : \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}^q$  as:

$$\Phi_{\mu}(u,v) = \begin{pmatrix} \varphi_{1\mu}(u_1,v_1) \\ \varphi_{2\mu}(u_2,v_2) \\ \vdots \\ \varphi_{J\mu}(u_J,v_J) \end{pmatrix},$$

where  $\varphi_{j\mu}(u_j, v_j) = u_j + v_j - \sqrt{(v_j - u_j)^2 + 4\mu^2 e_j}$  and  $e_j = (1, 0, \dots, 0) \in \mathbb{R}^{m_j + 1}$ ,  $j = 1, 2, \dots, J$ . Then, we construct a perturbation problem of (3.10) with parameter  $\mu > 0$  as follows:

$$(\mathbf{P}_{\mu}) \qquad \qquad \min_{\mathbf{s. t.}} \begin{array}{l} f(u,v) \\ \mathbf{s. t.} \quad \Phi_{\mu}(u,v) = 0. \end{array}$$

Obviously, function  $\Phi_{\mu}(u, v)$  is continuously differentiable with respect to u and v when  $\mu > 0$ .

Now we consider how close is the optimal solution set of  $(P_{\mu})$  from the optimal solution set of (3.10) as  $\mu \searrow 0$ . Let us introduce some notations:

$$\begin{split} \Omega_{j}^{0} &:= \{(u_{j}, v_{j}) \in Q_{m_{j}+1} \times Q_{m_{j}+1} | \ u_{j} \circ v_{j} = 0\}, \ j = 1, 2, \cdots, J, \\ \Omega^{0} &:= \{(u, v) \in Q \times Q | \ u \circ v = 0\} = \Omega_{1}^{0} \times \Omega_{2}^{0} \times \cdots \times \Omega_{J}^{0}, \\ \Omega_{j}(\mu) &:= \{(u_{j}, v_{j}) \in Q_{m_{j}+1} \times Q_{m_{j}+1} | \ \varphi_{j\mu}(u_{j}, v_{j}) = 0\}, \ j = 1, 2, \cdots, J, \\ \Omega(\mu) &:= \{(u, v) \in Q \times Q | \ \Phi_{\mu}(u, v) = 0\} = \Omega_{1}(\mu) \times \Omega_{2}(\mu) \times \cdots \times \Omega_{J}(\mu), \\ \bar{f}(u, v, \mu) &:= \begin{cases} \ f(u, v), \ (u, v) \in \Omega(\mu), \\ \infty, \ otherwise, \end{cases} \\ \kappa(\mu) &:= \inf\{f(u, v) | \ (u, v) \in \Omega(\mu)\}, \\ S(\mu) &:= \arg\min\{f(u, v) | \ (u, v) \in \Omega(\mu)\}. \end{split}$$

We will give a conclusion about the convergence of the set-value mapping  $\Omega(\mu)$  at  $\mu = 0$  with respect to  $\mathbb{R}_+$  in the following proposition.

**Proposition 4.1.** Let  $\Omega(\mu)$  and  $\Omega^0$  be defined as above. Then we have

$$\lim_{\mu\searrow 0}\Omega(\mu)=\Omega^0.$$

*Proof.* Noting that both  $\Omega(\mu)$  and  $\Omega^0$  are the Cartesian product of finite sets, without loss of generality, we only consider the case J = 1.

From the definition of  $\Omega^0$  and  $\Omega(\mu)$ , by Definition 2.1, one has

$$\Omega^{0} = \left\{ [(u_{1}; u_{2}), (v_{1}; v_{2})] \in Q_{m+1} \times Q_{m+1} \mid u_{1}v_{1} + u_{2}^{T}v_{2} = 0, u_{1}v_{2} + v_{1}u_{2} = 0 \right\}$$
  
=  $\Omega_{1} \cup \Omega_{2} \cup \Omega_{3} \cup \Omega_{4} \cup \Omega_{5} \cup \Omega_{6}$ 

with

$$\begin{split} \Omega_1 &:= \{(u,v) | u = 0, v = 0\},\\ \Omega_2 &:= \{(u,v) | u = 0, v \in \mathrm{bd}Q \setminus \{0\}\},\\ \Omega_3 &:= \{(u,v) | u \in \mathrm{bd}Q \setminus \{0\}, v = 0\},\\ \Omega_4 &:= \{(u,v) | u = 0, v \in \mathrm{int}Q\},\\ \Omega_5 &:= \{(u,v) | u \in \mathrm{int}Q, v = 0\},\\ \Omega_6 &:= \left\{(u,v) | u = (u_1; -\frac{u_1}{\|v_2\|}v_2) \in \mathrm{bd}Q \setminus \{0\}, v = (\|v_2\|; v_2) \in \mathrm{bd}Q \setminus \{0\}\right\}, \end{split}$$

and

(4.1) 
$$\Omega(\mu) = \{ [(u_1^{\mu}; u_2^{\mu}), (v_1^{\mu}; v_2^{\mu})] \in Q_{m+1} \times Q_{m+1} | u_1^{\mu} v_1^{\mu} + u_2^{\mu T} v_2^{\mu} \\ = \mu^2, u_1^{\mu} v_2^{\mu} + v_1^{\mu} u_2^{\mu} = 0 \}.$$

It follows from Chapter 4 of [19] that  $\limsup_{\mu \searrow 0} \Omega(\mu) \subset \Omega^0$ , then we only need to establish the inclusion

(4.2) 
$$\Omega^0 \subset \liminf_{\mu \searrow 0} \Omega(\mu),$$

namely, for all  $(u, v) \in \Omega^0$ , we can find  $(u^{\mu}, v^{\mu}) \in \Omega(\mu)$  with  $(u^{\mu}, v^{\mu}) \to (u, v)$  when  $\mu \searrow 0$ . We consider the following six cases:

**Case 1:**  $(u, v) = (0, 0) \in \Omega_1$ .

Taking  $(u^{\mu}, v^{\mu}) = ((\mu; 0), (\mu; 0))$ , obviously,  $(u^{\mu}, v^{\mu}) \in \Omega(\mu)$  and  $(u^{\mu}, v^{\mu}) \rightarrow 0$ (u, v) = (0, 0) when  $\mu \searrow 0$ .

**Case 2:**  $(u, v) = (0, (||v_2||; v_2)) \in \Omega_2.$ 

Taking

$$u^{\mu} = \left(\mu; -\frac{2\mu v_2}{\mu + \sqrt{\mu^2 + 4\|v_2\|^2}}\right), \quad v^{\mu} = \left(\frac{\mu + \sqrt{\mu^2 + 4\|v_2\|^2}}{2}; v_2\right),$$

we have that  $(u^{\mu}, v^{\mu}) \in \Omega(\mu)$  with  $(u^{\mu}, v^{\mu}) \rightarrow (u, v) = (0, (||v_2||; v_2))$  when  $\mu \searrow 0$ . **Case 3:**  $(u, v) = ((||u_2||; u_2), 0) \in \Omega_3.$ 

Similar to Case 2, we omit the detail.

**Case 4:**  $(u, v) = (0, (v_1; v_2)) \in \Omega_4$ .

Let

$$u_1^{\mu} = \frac{\mu^2 v_1}{v_1^2 - \|v_2\|^2}, \ u_2^{\mu} = -\frac{\mu^2 v_2}{v_1^2 - \|v_2\|^2}, \ v_1^{\mu} = v_1, \ v_2^{\mu} = v_2,$$

then we see that  $u^{\mu} = (u_1^{\mu}; u_2^{\mu}), v^{\mu} = (v_1^{\mu}; v_2^{\mu})$  satisfies the equation in (4.1), hence  $(u^{\mu}, v^{\mu}) \in \Omega(\mu)$ . When  $\mu \searrow 0$ , we have that  $(u^{\mu}, v^{\mu}) \rightarrow (u, v) = (0, (v_1; v_2))$ . **Case 5:**  $(u, v) = ((u_1; u_2), 0) \in \Omega_5$ .

Similar to Case 4, we omit the details, too.

**Case 6:** 
$$(u, v) = \left( (u_1; -\frac{u_1}{\|v_2\|}v_2), (\|v_2\|; v_2) \right) \in \Omega_6.$$
  
Let

$$u^{\mu} = \left(u_1; -\frac{2u_1^2 v_2}{\mu^2 + \sqrt{\mu^4 + 4u_1^2 \|v_2\|^2}}\right), \quad v^{\mu} = \left(\frac{\mu^2 + \sqrt{\mu^4 + 4u_1^2 \|v_2\|^2}}{2u_1}; v_2\right),$$

then  $(u^{\mu}, v^{\mu})$  satisfies the equation in (4.1) and  $(u^{\mu}, v^{\mu}) \in \Omega(\mu)$ . Since

$$\frac{2u_1^2 v_2}{\mu^2 + \sqrt{\mu^4 + 4u_1^2 \|v_2\|^2}} \to \frac{u_1}{\|v_2\|^2} v_2, \quad \frac{\mu^2 + \sqrt{\mu^4 + 4u_1^2 \|v_2\|^2}}{2u_1} \to \|v_2\|^2 \text{ as } \mu \searrow 0,$$

we have that  $(u^{\mu}, v^{\mu}) \to (u, v) = \left( (u_1; -\frac{u_1}{\|v_2\|}v_2), (\|v_2\|; v_2) \right)$  when  $\mu \searrow 0$ .

As in each of the six cases,  $(u^{\mu}, v^{\mu})$  can be constructed to satisfy  $(u^{\mu}, v^{\mu}) \in \Omega(\mu)$ and  $(u^{\mu}, v^{\mu}) \rightarrow (u, v)$ . Therefore, (4.2) holds and the proof is completed.  $\square$ 

The following lemma shows the relationship between the convergence of a setvalued mapping and the epi-convergence of the indicator function of the set-valued mapping.

**Lemma 4.1.** For a set-valued mapping  $\mathcal{T}: [0,\infty) \to \mathbb{R}^{2q}$ , the following equivalence holds:

$$\lim_{\mu \searrow 0} \mathcal{T}(\mu) = \mathcal{T}(0) \Longleftrightarrow e - \lim_{\mu \searrow 0} \delta_{\mathcal{T}(\mu)}(\cdot) = \delta_{\mathcal{T}(0)}(\cdot).$$

*Proof.* Noting that

$$\operatorname{epi}[\delta_{\mathcal{T}(\mu)}(\cdot)] = \{(z,\alpha) \mid z \in \mathcal{T}(\mu), \alpha \ge 0\} = \mathcal{T}(\mu) \times \mathbb{R}.$$

and

$$\lim_{\mu \searrow 0} \mathcal{T}(\mu) \times \mathbb{R}_{+} = \mathcal{T}(0) \times \mathbb{R}_{+} = \operatorname{epi}[\delta_{\mathcal{T}(0)}(\cdot)],$$

we obtain the equivalence.

The following lemma discusses the epi-convergence of the sum of a functionvalued mapping and a continuous function.

**Lemma 4.2.** Let  $\psi(\cdot)$  satisfy

$$e - \lim_{\mu \searrow 0} \psi(\mu)(\cdot) = \psi(0)(\cdot)$$

and  $\psi_1$  be a continuous function, then

$$e - \lim_{\mu \searrow 0} [\psi(\mu)(\cdot) + \psi_1(\cdot)] = \psi(0)(\cdot) + \psi_1(\cdot).$$

*Proof.* The conclusion is obvious by the definition of epi-convergence.

Then the convergence behavior of the feasible set and optimal solution set of  $(P_{\mu})$  is obtained in the following two theorems, which are the main results in this paper.

**Theorem 4.1.** The function-valued mapping  $\mu \to \overline{f}(\cdot, \cdot, \mu)$  is epi-continuous at  $\mu = 0$  with respect to  $\mathbb{R}_+$ .

*Proof.* Since f(u, v) is continuous and from Proposition 4.1 that  $\Omega(\mu)$  is continuous at  $\mu = 0$  with respect to  $\mathbb{R}_+$ , the result comes from Lemmas 4.1 and Lemma 4.2.  $\Box$ 

**Theorem 4.2.** Assume A is of full row rank. Then the function  $\kappa(\mu)$  is continuous at  $\mu = 0$  with respect to  $\mathbb{R}_+$  and the set-valued mapping  $S(\mu)$  is outer semicontinuous at  $\mu = 0$  with respect to  $\mathbb{R}_+$ .

Proof. As A is of full row rank, we have from Lemma 3.3 that f is strongly convex. Then for a given  $\bar{\alpha}$  such that  $\operatorname{lev}_{\leq \bar{\alpha}} f \neq \emptyset$ , we have that  $\operatorname{lev}_{\leq \bar{\alpha}} f$  is a nonempty compact set. Hence there is an M > 0 such that the level set  $\operatorname{lev}_{\leq \bar{\alpha}} f \subset [-M \mathbb{1}_{2q}, M \mathbb{1}_{2q}]$ , where  $\mathbb{1}_{2q}$  is a vector in  $\mathbb{R}^{2q}$  with all entries being ones. Let  $\bar{\mu} > 0$  be a given number, and define  $\mathcal{E} = [0, \bar{\mu})$ . As for any  $\alpha \leq \bar{\alpha}$ ,  $\operatorname{lev}_{\leq \alpha} \bar{f}(\cdot, \cdot, \mu) = \operatorname{lev}_{\leq \alpha} f \bigcap \Omega(\mu) \subset [-M \mathbb{1}_{2q}, M \mathbb{1}_{2q}]$ , the level set is uniformly bounded for all  $\mu \in \mathcal{E}$ , i.e.,  $\bar{f}$  is level-bounded in (u, v) uniformly for all  $\mu \in \mathcal{E}$ . Hence we have, from [19, Theorem 7.41], that  $\kappa(\mu)$  is continuous at  $\mu = 0$  and  $S(\mu)$  is outer semi-continuous at  $\mu = 0$  with respect to  $\mathbb{R}_+$ .

### 5. Smoothing Newton method for solving $P_{\mu}$

We obtain that the optimal solution set of  $(P_{\mu})$  is outer semicontinuous at  $\mu = 0$  if we set  $(P_0)$  as problem (3.10) by Theorem 4.2. Therefore, we focus on the issue of solving  $(P_{\mu})$  with a sufficiently small  $\mu > 0$ .

Let  $w_j := \sqrt{(v_j - u_j)^2 + 4\mu^2 e_j}$ ,  $w := (w_1; \dots; w_J) \in \mathbb{R}^q$ ,  $L_u := \text{diag}(L_{u_1}, \dots, L_{u_J})$ ,  $L_v := \text{diag}(L_{v_1}, \dots, L_{v_J})$  and  $L_w := \text{diag}(L_{w_1}, \dots, L_{w_J})$ . The following lemma implies the linear independent constraint qualification (LICQ) holds for  $(P_{\mu})$  automatically, then we can get its KKT conditions at local optimal solutions.

**Lemma 5.1.**  $\mathcal{J}_{(u,v)}\Phi_{\mu}(u,v)$  is of full row rank.

224

*Proof.* For any  $u_j \in Q_{m_j+1}$ ,  $v_j \in Q_{m_j+1}$ , we have  $w_j^2 = (v_j - u_j)^2 + 4\mu^2 e_j \in intQ_{m_j+1}$ , hence  $w_j \in intQ_{m_j+1}$  by [11, Prosition 3.4]. Applying [11, Lemma 5], it holds that

$$L_{w_j} - L_{v_j - u_j} \succ 0, \ L_{w_j} + L_{v_j - u_j} \succ 0, \ L_{w_j} \succ 0$$

Similar to the proof in [11, Corollary 5.4], for any j,

$$\mathcal{J}_{(u_j,v_j)}\varphi_{j\mu}(u_j,v_j) = \begin{bmatrix} I_{m_j+1} + L_{w_j}^{-1}L_{v_j-u_j} & I_{m_j+1} - L_{w_j}^{-1}L_{v_j-u_j} \end{bmatrix}.$$

Suppose  $\mathcal{J}_{(u_j,v_j)}\varphi_{j\mu}(u_j,v_j)^T d_j = 0$  with  $d_j \in \mathbb{R}^{m_j+1}$ , namely,

$$\begin{bmatrix} I_{m_j+1} + L_{v_j-u_j} L_{w_j}^{-1} \\ I_{m_j+1} - L_{v_j-u_j} L_{w_j}^{-1} \end{bmatrix} d_j = 0$$

or, equivalently,

$$(L_{w_j} + L_{v_j - u_j})L_{w_j}^{-1}d_j = 0$$
 and  $(L_{w_j} - L_{v_j - u_j})L_{w_j}^{-1}d_j = 0.$ 

As the matrices  $L_{w_j} + L_{v_j - u_j}$ ,  $L_{w_j} - L_{v_j - u_j}$  and  $L_{w_j}$  are invertible, we get  $d_j = 0$  for  $j = 1, 2, \dots, J$ . Similarly, suppose

$$\mathcal{J}_{(u,v)}\Phi_{\mu}(u,v)^{T}d = 0,$$

with  $d = (d_1; d_2; \cdots; d_J), d_j \in \mathbb{R}^{m_j+1}, j = 1, 2, \cdots, J$ . We have

$$\mathcal{J}_{(u,v)}\Phi_{\mu}(u,v)^{T}d = \begin{bmatrix} I_{q} + L_{v-u}L_{w}^{-1} \\ I_{q} - L_{v-u}L_{w}^{-1} \end{bmatrix} d = 0,$$

which implies d = 0. Then  $\mathcal{J}_{(u,v)}\Phi_{\mu}(u,v)$  is of full row rank. The proof is completed.

Let  $\mathcal{L} : \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}$  be the Lagrange function for  $(P_\mu)$ :

$$\mathcal{L}(u,v,\lambda) = f(u,v) + \lambda^T \Phi_\mu(u,v) = f(u,v) + \sum_{j=1}^J \lambda_j^T \varphi_{j\mu}(u_j,v_j),$$

where  $\lambda = (\lambda_1; \lambda_2; \cdots; \lambda_J) \in \mathbb{R}^q, \lambda_j \in \mathbb{R}^{m_j+1}, \ j = 1, 2, \cdots, J$ . Define  $F : \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}^{3q}$ :

$$F(u, v, \lambda) := \begin{bmatrix} \nabla_u \mathcal{L}(u, v, \lambda) \\ \nabla_v \mathcal{L}(u, v, \lambda) \\ \Phi_\mu(u, v) \end{bmatrix}.$$

As the LICQ holds at any local minimizer  $(\bar{u}, \bar{v})$  of  $(P_{\mu})$ , there is a unique Lagrange multiplier  $\bar{\lambda} \in \mathbb{R}^q$  such that the KKT conditions are satisfied at  $(\bar{u}, \bar{v}, \bar{\lambda})$ , hence  $F(\bar{u}, \bar{v}, \bar{\lambda}) = 0_{3q}$ . Now we consider the second-order sufficient conditions of  $(P_{\mu})$  at  $(\bar{u}, \bar{v}, \bar{\lambda})$ .

**Lemma 5.2.** Denote  $s_j := L_{w_j}^{-1} \lambda_j \in \mathbb{R}^{m_j+1}$ , then for any  $M \in \partial_{(u,v)}^2 \mathcal{L}(u,v,\lambda)$ there exists  $Z \in \partial x(u)$  such that

$$M = \begin{bmatrix} -AZ - L_s + L_{v-u}L_w^{-1}L_sL_w^{-1}L_{v-u} & L_s - L_{v-u}L_w^{-1}L_sL_w^{-1}L_{v-u} \\ L_s - L_{v-u}L_w^{-1}L_sL_w^{-1}L_{v-u} & I_q - L_s + L_{v-u}L_w^{-1}L_sL_w^{-1}L_{v-u} \end{bmatrix},$$
  
where  $L_s = \text{diag}(L_{s_1}, L_{s_2}, \cdots, L_{s_J}).$ 

*Proof.* It follows from the proof in Lemma 5.1, we have

$$\nabla_u \Phi_\mu(u, v) = I_q + L_{v-u} L_w^{-1},$$
  
$$\nabla_v \Phi_\mu(u, v) = I_q - L_{v-u} L_w^{-1},$$

together with (3.14), it holds that

(5.2) 
$$\nabla_u \mathcal{L}(u, v, \lambda) = -Ax(u) + (I_q + L_{v-u}L_w^{-1})\lambda,$$

(5.3) 
$$\nabla_{v}\mathcal{L}(u,v,\lambda) = v - v_0 + (I_q - L_{v-u}L_w^{-1})\lambda.$$

Since  $s_j = L_{w_j}^{-1} \lambda_j \in \mathbb{R}^{m_j+1}$ , we have  $L_{w_j} s_j = \lambda_j$ , then

(5.4) 
$$w_j \circ s_j = \lambda_j.$$

From [5, Lemma 3.1], taking the Jacobian matrix of both sides in equation (5.4) with respect to  $(u_j, v_j)$  yields

$$(5.5) L_{s_j}\mathcal{J}_{u_j}w_j + L_{w_j}\mathcal{J}_{u_j}s_j = 0,$$

$$(5.6) L_{s_j} \mathcal{J}_{v_j} w_j + L_{w_j} \mathcal{J}_{v_j} s_j = 0$$

As  $\mathcal{J}_{u_j} w_j = -L_{w_j}^{-1} L_{v_j-u_j}$  and  $\mathcal{J}_{v_j} w_j = L_{w_j}^{-1} L_{v_j-u_j}$ , combining with (5.5) and (5.6), we have that

$$\begin{aligned} \mathcal{J}_{u_j} s_j &= -L_{w_j}^{-1} L_{s_j} \mathcal{J}_{u_j} w_j = L_{w_j}^{-1} L_{s_j} L_{w_j}^{-1} L_{v_j - u_j}, \\ \mathcal{J}_{v_j} s_j &= -L_{w_j}^{-1} L_{s_j} \mathcal{J}_{v_j} w_j = -L_{w_j}^{-1} L_{s_j} L_{w_j}^{-1} L_{v_j - u_j}. \end{aligned}$$

Therefore,

$$\begin{aligned} \mathcal{J}_{u_j} [(I_{m_j+1} + L_{v_j - u_j} L_{w_j}^{-1})\lambda_j] &= \mathcal{J}_{u_j} (L_{v_j - u_j} L_{w_j}^{-1}\lambda_j) \\ &= \mathcal{J}_{u_j} (L_{v_j - u_j} s_j) \\ &= -L_{s_j} + L_{v_j - u_j} \mathcal{J}_{u_j} s_j \\ &= -L_{s_j} + L_{v_j - u_j} L_{w_j}^{-1} L_{s_j} L_{w_j}^{-1} L_{v_j - u_j}, \\ \mathcal{J}_{v_j} [(I_{m_j+1} + L_{v_j - u_j} L_{w_j}^{-1})\lambda_j] &= \mathcal{J}_{v_j} (L_{v_j - u_j} L_{w_j}^{-1}\lambda_j) \\ &= \mathcal{J}_{v_j} (L_{v_j - u_j} S_j) \\ &= L_{s_j} + L_{v_j - u_j} \mathcal{J}_{v_j} s_j \\ &= L_{s_j} - L_{v_j - u_j} L_{w_j}^{-1} L_{s_j} L_{w_j}^{-1} L_{v_j - u_j}; \end{aligned}$$

similarly,

$$\begin{aligned} \mathcal{J}_{u_j}[(I_{m_j+1} - L_{v_j - u_j}L_{w_j}^{-1})\lambda_j] = & L_{s_j} - L_{v_j - u_j}L_{w_j}^{-1}L_{s_j}L_{w_j}^{-1}L_{v_j - u_j}, \\ \mathcal{J}_{v_j}[(I_{m_j+1} - L_{v_j - u_j}L_{w_j}^{-1})\lambda_j] = & -L_{s_j} + L_{v_j - u_j}L_{w_j}^{-1}L_{s_j}L_{w_j}^{-1}L_{v_j - u_j}. \end{aligned}$$

In a compact form, one has

$$\begin{aligned} \mathcal{J}_{u}[(I_{q}+L_{v-u}L_{w}^{-1})\lambda] &= -L_{s}+L_{v-u}L_{w}^{-1}L_{s}L_{w}^{-1}L_{v-u},\\ \mathcal{J}_{v}[(I_{q}+L_{v-u}L_{w}^{-1})\lambda] &= L_{s}-L_{v-u}L_{w}^{-1}L_{s}L_{w}^{-1}L_{v-u},\\ \mathcal{J}_{u}[(I_{q}-L_{v-u}L_{w}^{-1})\lambda] &= L_{s}-L_{v-u}L_{w}^{-1}L_{s}L_{w}^{-1}L_{v-u},\\ \mathcal{J}_{v}[(I_{q}-L_{v-u}L_{w}^{-1})\lambda] &= -L_{s}+L_{v-u}L_{w}^{-1}L_{s}L_{w}^{-1}L_{v-u}.\end{aligned}$$

Then, for any  $M \in \partial^2_{(u,v)} \mathcal{L}(u,v,\lambda)$ , there exists  $Z \in \partial x(u)$  such that

$$M = \begin{bmatrix} -AZ - L_s + L_{v-u}L_w^{-1}L_sL_w^{-1}L_{v-u} & L_s - L_{v-u}L_w^{-1}L_sL_w^{-1}L_{v-u} \\ L_s - L_{v-u}L_w^{-1}L_sL_w^{-1}L_{v-u} & I_q - L_s + L_{v-u}L_w^{-1}L_sL_w^{-1}L_{v-u} \end{bmatrix}.$$
  
e proof is completed.

The proof is completed.

Assumption 5.1. Let  $(\bar{u}, \bar{v}, \bar{\lambda})$  be a KKT point of  $(P_{\mu})$ . Matrices  $L_{\bar{u}}, L_{\bar{v}}, L_{\bar{w}}$  and  $L_{\bar{s}}$  satisfy

$$(L_{\bar{w}} + L_{\bar{v}-\bar{u}}) (L_{\bar{v}-\bar{u}} - L_{\bar{w}})^{-1} L_{\bar{s}} + L_{\bar{s}} (L_{\bar{v}-\bar{u}} - L_{\bar{w}})^{-1} (L_{\bar{w}} + L_{\bar{v}-\bar{u}}) \succeq 0.$$

**Lemma 5.3.** If Assumption 5.1 holds, then the second-order sufficient condition of  $(P_{\mu})$  holds at  $(\bar{u}, \bar{v}, \bar{\lambda})$ , namely, for any  $\bar{M} \in \partial^2_{(u,v)} \mathcal{L}(\bar{u}, \bar{v}, \bar{\lambda})$ ,

$$d^T \overline{M} d > 0, \ \forall d \in \ker \mathcal{J}_{(u,v)} \Phi_\mu(\overline{u}, \overline{v}) \setminus \{0\},$$

where kerC denotes the kernel of operator C.

*Proof.* For any vector  $d = (d_1; d_2) \in \ker \mathcal{J}\Phi_\mu(\bar{u}, \bar{v}) \setminus \{0\}$  with  $d_i \in \mathbb{R}^q$ , i = 1, 2, we have

$$\mathcal{J}\Phi_{\mu}(\bar{u},\bar{v})d=0,$$

namely,

$$\left(I_q + L_{\bar{w}}^{-1} L_{\bar{v}-\bar{u}}\right) d_1 + \left(I_q - L_{\bar{w}}^{-1} L_{\bar{v}-\bar{u}}\right) d_2 = 0,$$

hence

(5.7) 
$$L_{\bar{w}}^{-1}L_{\bar{v}-\bar{u}}(d_2-d_1) = d_1+d_2$$

As  $\left(I_q - L_{\bar{w}}^{-1}L_{\bar{v}-\bar{u}}\right)$  is invertible, we have

(5.8) 
$$d_{2} = -\left(I_{q} - L_{\bar{w}}^{-1}L_{\bar{v}-\bar{u}}\right)^{-1}\left(I_{q} + L_{\bar{w}}^{-1}L_{\bar{v}-\bar{u}}\right)d_{1} = \left(L_{\bar{v}-\bar{u}} - L_{\bar{w}}\right)^{-1}\left(L_{\bar{w}} + L_{\bar{v}-\bar{u}}\right)d_{1}.$$

Therefore, under Assumption 5.1 and together with (5.7) and (5.8), it holds that

$$(5.9) \quad \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}^T \begin{bmatrix} -L_{\bar{s}} + L_{\bar{v}-\bar{u}} L_{\bar{w}}^{-1} L_{\bar{s}} L_{\bar{w}}^{-1} L_{\bar{v}-\bar{u}} & L_{\bar{s}} - L_{\bar{v}-\bar{u}} L_{\bar{w}}^{-1} L_{\bar{s}} L_{\bar{w}-\bar{u}}^{-1} L_{\bar{v}-\bar{u}} \end{bmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \\ = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}^T \left\{ \begin{bmatrix} L_{\bar{v}-\bar{u}} L_{\bar{w}}^{-1} L_{\bar{s}} L_{\bar{w}}^{-1} L_{\bar{v}-\bar{u}} & -L_{\bar{v}-\bar{u}} L_{\bar{w}}^{-1} L_{\bar{s}} L_{\bar{w}}^{-1} L_{\bar{v}-\bar{u}} \\ -L_{\bar{v}-\bar{u}} L_{\bar{w}}^{-1} L_{\bar{s}} L_{\bar{w}}^{-1} L_{\bar{v}-\bar{u}} & L_{\bar{v}-\bar{u}} L_{\bar{w}}^{-1} L_{\bar{s}-\bar{u}} \end{bmatrix} \\ + \begin{bmatrix} -L_{\bar{s}} & L_{\bar{s}} \\ L_{\bar{s}} & -L_{\bar{s}} \end{bmatrix} \right\} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \\ = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}^T \begin{bmatrix} L_{\bar{v}-\bar{u}} L_{\bar{w}}^{-1} \\ -L_{\bar{v}-\bar{u}} L_{\bar{w}}^{-1} \end{bmatrix} L_{\bar{s}} \begin{bmatrix} L_{\bar{w}}^{-1} L_{\bar{v}-\bar{u}} & -L_{\bar{w}}^{-1} L_{\bar{v}-\bar{u}} \end{bmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \\ + \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}^T \begin{bmatrix} L_{\bar{v}-\bar{u}} L_{\bar{w}}^{-1} \\ -L_{\bar{v}-\bar{u}} L_{\bar{w}}^{-1} \end{bmatrix} L_{\bar{s}} \begin{bmatrix} d_1 \\ d_2 \end{pmatrix} \\ = (d_1 - d_2)^T L_{\bar{v}-\bar{u}} L_{\bar{w}}^{-1} L_{\bar{s}} L_{\bar{w}}^{-1} L_{\bar{v}-\bar{u}} (d_1 - d_2) - d_1^T L_{\bar{s}} d_1 - d_2^T L_{\bar{s}} d_2 + 2d_1^T L_{\bar{s}} d_2 \\ = (d_1^T + d_2^T) L_{\bar{s}} (d_1 + d_2) - d_1^T L_{\bar{s}} d_1 - d_2^T L_{\bar{s}} d_2 + 2d_1^T L_{\bar{s}} d_2 \\ = (d_1^T + d_2^T) L_{\bar{s}} (L_{\bar{v}-\bar{u}} - L_{\bar{w}})^{-1} (L_{\bar{w}} + L_{\bar{v}-\bar{u}}) d_1 \\ \ge 0. \end{cases}$$

From Lemma 5.2, we obtain that for any  $\overline{M} \in \partial^2_{(u,v)} \mathcal{L}(\overline{u}, \overline{v}, \overline{\lambda})$ , there exists  $\overline{Z} \in \partial x(\overline{u})$ , such that

(5.10)

$$\bar{M} = \begin{bmatrix} -A\bar{Z} - L_{\bar{s}} + L_{\bar{v}-\bar{u}}L_{\bar{w}}^{-1}L_{\bar{s}}L_{\bar{w}}^{-1}L_{\bar{v}-\bar{u}} & L_{\bar{s}} - L_{\bar{v}-\bar{u}}L_{\bar{w}}^{-1}L_{\bar{s}}L_{\bar{w}}^{-1}L_{\bar{v}-\bar{u}} \\ L_{\bar{s}} - L_{\bar{v}-\bar{u}}L_{\bar{w}}^{-1}L_{\bar{s}}L_{\bar{w}}^{-1}L_{\bar{v}-\bar{u}} & I_q - L_{\bar{s}} + L_{\bar{v}-\bar{u}}L_{\bar{w}}^{-1}L_{\bar{s}}L_{\bar{w}}^{-1}L_{\bar{v}-\bar{u}} \end{bmatrix}.$$

It follows from Lemma 3.2 that  $-A\overline{Z}$  is positive definite for any  $\overline{Z} \in \partial x(\overline{u})$ , then together with (5.9) and (5.10), we have

$$d^{T} M d = \begin{pmatrix} d_{1} \\ d_{2} \end{pmatrix}^{T} \begin{bmatrix} -A\bar{Z} & 0 \\ 0 & I_{q} \end{bmatrix} \begin{pmatrix} d_{1} \\ d_{2} \end{pmatrix}$$

$$+ \begin{pmatrix} d_{1} \\ d_{2} \end{pmatrix}^{T} \begin{bmatrix} -L_{\bar{s}} + L_{\bar{v}-\bar{u}}L_{\bar{w}}^{-1}L_{\bar{s}}L_{\bar{w}}^{-1}L_{\bar{v}-\bar{u}} & L_{\bar{s}} - L_{\bar{v}-\bar{u}}L_{\bar{w}}^{-1}L_{\bar{s}}L_{\bar{v}-\bar{u}} \\ L_{\bar{s}} - L_{\bar{v}-\bar{u}}L_{\bar{w}}^{-1}L_{\bar{s}}L_{\bar{w}}^{-1}L_{\bar{v}-\bar{u}} & -L_{\bar{s}} + L_{\bar{v}-\bar{u}}L_{\bar{w}}^{-1}L_{\bar{s}}L_{\bar{w}}^{-1}L_{\bar{v}-\bar{u}} \end{bmatrix} \begin{pmatrix} d_{1} \\ d_{2} \end{pmatrix}$$

$$> 0$$

for any  $d = (d_1; d_2) \in \ker \mathcal{J}\Phi_{\mu}(u, v) \setminus \{0\}$ , namely the second-order sufficient condition of  $(\mathbf{P}_{\mu})$  holds at  $(\bar{u}, \bar{v}, \bar{\lambda})$ .

If the second-order sufficient condition holds at  $(\bar{u}, \bar{v}, \bar{\lambda})$ , then  $(\bar{u}, \bar{v}, \bar{\lambda})$  is a local minimizer of  $(P_{\mu})$ . Therefore we turn to solve the equation  $F(u, v, \lambda) = 0$ . Define  $\hat{F} : \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}^{3q+n}$ :

$$\hat{F}(x, u, v, \lambda) := \begin{bmatrix} -Ax + (I_q + L_{v-u}L_w^{-1})\lambda \\ v - v_0 + (I_q - L_{v-u}L_w^{-1})\lambda \\ \Phi_{\mu}(u, v) \\ x - (c^0 - A^T u) - \mathcal{B}^*\Pi_{\mathcal{S}^n_+}(\bar{G}(x)) \end{bmatrix}$$

Noting that  $\|\Pi_{\mathcal{S}^n_+}(\bar{G}(x))\|^2 + \|x - (c^0 - A^T u)\|^2$  is strongly convex with respect to x and x(u) is the unique solution of (3.12), then together with (5.2) and (5.3), we have that solving  $F(u, v, \lambda) = 0$  is equivalent to solving  $\hat{F}(x, u, v, \lambda) = 0$ .

Now we are in a position to state a smoothing Newton method for solving  $\hat{F}(x, u, v, \lambda) = 0$ . Considering a square smoothing function  $\psi : \mathbb{R} \times S^n \to S^n$ ; see [22], defined by

$$\psi(\varepsilon, X) := \sqrt{X^2 + \varepsilon^2 I_n}, \quad \forall \ (\varepsilon, X) \in \mathbb{R} \times \mathcal{S}^n$$

Then,  $\psi$  is continuously differentiable at  $(\varepsilon, X)$  unless  $\varepsilon = 0$ . For  $X \in S^n$ , we also have the Lyapunov operator  $\mathbb{L}_X$ :

$$\mathbb{L}_X(Y) := XY + YX, \quad \forall \quad Y \in \mathcal{S}^n,$$

with  $\mathbb{L}_X^{-1}$  being its inverse (if it exists at all), i.e., for any  $Y \in \mathcal{S}^n$ ,  $\mathbb{L}_X^{-1}(Y)$  is the unique  $Z \in \mathcal{S}^n$  satisfying XZ + ZX = Y.

Let  $T: \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^n$  be a smoothing approximation mapping defined by

$$T(\varepsilon, x, u, v, \lambda) := \begin{bmatrix} -Ax + \left(I_q + L_{v-u}L_w^{-1}\right)\lambda \\ v - v_0 + \left(I_q - L_{v-u}L_w^{-1}\right)\lambda \\ \Phi_{\mu}(u, v) \\ A^Tu - c^0 + x - \frac{1}{2}\mathcal{B}^*\left(\bar{G}(x) + \sqrt{\bar{G}(x)^2 + \varepsilon^2 I_n}\right) \end{bmatrix}$$

Obviously,  $\lim_{\varepsilon \to 0} T(\varepsilon, x, u, v, \lambda) = \hat{F}(x, u, v, \lambda)$ . The smoothing Newton method is based on solving

(5.11) 
$$E(\varepsilon, x, u, v, \lambda) := \begin{bmatrix} \varepsilon \\ T(\varepsilon, x, u, v, \lambda) \end{bmatrix} = 0$$

and uses the merit function  $\phi(z) := \|E(\varepsilon, x, u, v, \lambda)\|^2$  for the line search, where  $z := (\varepsilon, x, u, v, \lambda)$ . Let  $\hat{\varepsilon} > 0$  and  $\eta \in (0, 1)$  be such that  $\eta \hat{\varepsilon} < 1$ . Define an auxiliary point  $\hat{z}$  by  $\hat{z} := (\hat{\varepsilon}, 0, 0, 0, 0) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^q$  and  $\theta : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}_+$  by  $\theta(z) := \eta \min\{1, \phi(z)\}$ . The smoothing Newton method, proposed by [18] can be described as follows:

## Algorithm 5.1 (Smoothing Newton method)

**step 1:** Select constants  $\delta \in (0, 1)$  and  $\sigma \in (0, 1/2)$ . Let  $\varepsilon^0 := \hat{\varepsilon}$ ,  $(x^0, u^0, v^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^q$  be an arbitrary point. Then let the initial point  $z^0 = (\varepsilon^0, x^0, u^0, v^0, \lambda^0)$  and k = 0.

step 2: If  $E(z^k) = 0$ , then stop; otherwise, let  $\theta_k := \theta(z^k)$ .

step 3: Compute  $\Delta z^k := (\Delta \varepsilon^k, \Delta x^k, \Delta u^k, \Delta v^k, \Delta \lambda^k) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^q$  by

(5.12) 
$$E(z^k) + \mathcal{J}E(z^k)(\Delta z^k) = \theta_k \hat{z}.$$

step 4: Let  $l_k$  be the smallest nonnegative integer l satisfying

 $\phi(z^k + \delta^l \Delta z^k) \le (1 - 2\sigma(1 - \eta\hat{\varepsilon})\delta^l)\phi(z^k).$ 

Define  $z^{k+1} = z^k + \delta^{l_k} \Delta z^k$ .

**step 5:** k := k + 1, go to Step 2.

From [18, Theorem 8], the key conditions for quadratic convergence of Algorithm 5.1 are: (a) the strong semismoothness of the smoothing function E and (b) the nonsingularity of all matrices  $H \in \partial_B E(\bar{z})(\bar{z} = (0, \bar{x}, \bar{u}, \bar{v}, \bar{\lambda})$  is a solution of E(z) = 0). Here (a) is naturally implied by the strong semismoothness of  $\psi$ , then we focus on proving the nonsingularity of the elements in  $\partial_B E(\bar{z})$ .

**Lemma 5.4.** Let  $\bar{z} = (0, \bar{x}, \bar{u}, \bar{v}, \bar{\lambda}) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^q$  be a solution of E(z) = 0. If A is of full row rank and Assumption 5.1 holds at  $(\bar{u}, \bar{v}, \bar{\lambda})$ , then for any  $H \in \partial_B E(\bar{z})$ , we have H is nonsingular.

*Proof.* Suppose that there exists  $(\Delta \varepsilon, \Delta x, \Delta u, \Delta v, \Delta \lambda) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^q$  such that  $H(\Delta \varepsilon, \Delta x, \Delta u, \Delta v, \Delta \lambda) = 0$ , namely,

$$(5.13) \qquad \qquad \Delta \varepsilon = 0,$$

(5.14) 
$$-A\Delta x + \bar{N}\Delta u - \bar{N}\Delta v + \left(I_q + L_{\bar{v}-\bar{u}}L_{\bar{w}}^{-1}\right)\Delta\lambda = 0,$$

(5.15) 
$$-\bar{N}\Delta u + \Delta v + \bar{N}\Delta v + \left(I_q - L_{\bar{v}-\bar{u}}L_{\bar{w}}^{-1}\right)\Delta\lambda = 0,$$

(5.16) 
$$(I_q + L_{\bar{w}}^{-1} L_{\bar{v}-\bar{u}}) \Delta u + (I_q - L_{\bar{w}}^{-1} L_{\bar{v}-\bar{u}}) \Delta v = 0,$$

(5.17) 
$$\Delta x + \frac{1}{2}\mathcal{B}^*\mathcal{B}\Delta x + \frac{1}{2}\mathcal{B}^*V(\Delta\varepsilon,\mathcal{B}\Delta x) + A^T\Delta u = 0,$$

with  $\overline{N} = -L_{\overline{s}} + L_{\overline{v}-\overline{u}}L_{\overline{w}}^{-1}L_{\overline{s}}L_{\overline{w}}^{-1}L_{\overline{v}-\overline{u}}$  and  $V \in \partial_B \psi(0, \overline{G}(\overline{x}))$ . Together with (5.13), (5.17) implies that

(5.18) 
$$\Delta x + \frac{1}{2} \mathcal{B}^* \mathcal{B} \Delta x + \frac{1}{2} \mathcal{B}^* V(0, \mathcal{B} \Delta x) + A^T \Delta u = 0.$$

To prove that  $\mathcal{J}E(z)$  is nonsingular we only need to show  $\Delta z = 0$ . By premultiplying  $\Delta u^T$  to both sides of (5.14) and  $\Delta v^T$  to both sides of (5.15), together with the equality (5.16), we obtain that

(5.19) 
$$\Delta v^T \Delta v - \Delta u^T A \Delta x + \Delta u^T \bar{N} \Delta u + \Delta v^T \bar{N} \Delta v - 2\Delta u^T \bar{N} \Delta v = 0.$$

By (5.18), submitting  $A^T \Delta u = -\left(\Delta x + \frac{1}{2}\mathcal{B}^*\mathcal{B}\Delta x + \frac{1}{2}\mathcal{B}^*V(0,\mathcal{B}\Delta x)\right)$  to (5.19), then

(5.20) 
$$\Delta u^T \bar{N} \Delta u + \Delta v^T \bar{N} \Delta v - 2\Delta u^T \bar{N} \Delta v$$
$$= -\Delta x^T \Delta x - \frac{1}{2} \Delta x^T \left( \mathcal{B}^* \mathcal{B} \Delta x + \mathcal{B}^* V(0, \mathcal{B} \Delta x) \right) - \Delta v^T \Delta v.$$

Suppose that  $\bar{G}(\bar{x})$  has the spectral decomposition  $\bar{G}(\bar{x}) = \bar{P}\bar{\Sigma}\bar{P}^T$ , where  $\bar{\Sigma}$  is the diagonal matrix of eigenvalues of  $\bar{G}(\bar{x})$  and  $\bar{P}$  is a corresponding orthogonal matrix of the orthogonal eigenvectors. Then from [22, Proposition 3.1], we have that

$$V(0, \mathcal{B}\Delta x) = \bar{P} \left[ \Omega \bullet \left( \bar{P}^T (\mathcal{B}\Delta x) \bar{P} \right) \right] \bar{P}^T,$$

where  $\Omega \in S^n_+$  with all entries belonging to [-1, 1],  $\bullet$  denotes the Hadamard product of two matrices and

$$\Delta x^{T} \left( \mathcal{B}^{*} \mathcal{B} \Delta x + \mathcal{B}^{*} V(0, \mathcal{B} \Delta x) \right)$$
  
= $\left\langle \mathcal{B} \Delta x, \mathcal{B} \Delta x + \bar{P} \left[ \Omega \bullet \left( \bar{P}^{T} (\mathcal{B} \Delta x) \bar{P} \right) \right] \bar{P}^{T} \right\rangle$   
= $\left\langle \bar{P}^{T} (\mathcal{B} \Delta x) \bar{P}, \bar{P}^{T} (\mathcal{B} \Delta x) \bar{P} + \Omega \bullet \left( \bar{P}^{T} (\mathcal{B} \Delta x) \bar{P} \right) \right\rangle$   
 $\geq 0.$ 

Hence the right-hand side of equation (5.20) is less than zero. On the other hand, from (5.16), we see that  $(\Delta u; \Delta v) \in \ker \mathcal{J}_{(u,v)} \Phi_{\mu}(\bar{u}, \bar{v})$ . As Assumption 5.1 holds, by (5.9), we have  $\Delta u^T \bar{N} \Delta u + \Delta v^T \bar{N} \Delta v - 2\Delta u^T \bar{N} \Delta v \ge 0$ . Then it is not difficult to deduce from (5.20) that  $\Delta x = 0$ ,  $\Delta v = 0$ . By (5.18), since A is of full row rank, we have  $\Delta u = 0$ , which in turn implies  $\Delta \lambda = 0$ . Consequently, H is nonsingular.

We now state the convergence behavior and quadratic convergence rate of Algorithm 5.1 in the following theorem, which is directly from [18, Theorem 4 and Theorem 8].

**Theorem 5.1.** Let a sequence  $\{z^k\}$  be generated by Algorithm 5.1 and let  $\bar{z} := (\bar{z}, \bar{x}, \bar{u}, \bar{v}, \bar{\lambda})$  be an accumulation point  $\{z^k\}$ . Suppose that for every k,  $\mathcal{J}E(z^k)$  is nonsingular, matrix A is of full row rank and Assumption 5.1 holds at  $(\bar{u}, \bar{v}, \bar{\lambda})$ . Then  $\bar{z}$  is a solution of E(z) = 0 and the sequence  $\{z^k\}$  converges to  $\bar{z}$  quadratically.

### 6. Numerical results

In this section, we report the numerical experiments conducted for testing the efficiency of Algorithm 5.1. The main task of Algorithm 5.1 for solving  $(P_{\mu})$ , at the k-th iterate, is to compute the direction  $\Delta z^k$  at a given  $z^k$  from the linear system (5.12). For writing convenience, we omit the iterate index k in the following discussion.

230

It follows from [22, Lemma 2.3] that  $\mathcal{J}E(z)(\Delta z)$  in linear system (5.12) has the expression

(6.1) 
$$\mathcal{J}E(z)(\Delta z) \qquad \qquad \Delta \varepsilon \\ = \begin{bmatrix} -A\Delta x + N\Delta u - N\Delta v + (I_q + L_{v-u}L_w^{-1})\Delta \lambda \\ -N\Delta u + N\Delta v + \Delta v + (I_q - L_{v-u}L_w^{-1})\Delta \lambda \\ (I_q + L_w^{-1}L_{v-u})\Delta u + (I_q - L_w^{-1}L_{v-u})\Delta v \\ \Delta x + A^T\Delta u + \frac{1}{2}\mathcal{B}^*\mathcal{B}\Delta x - \frac{1}{2}\mathcal{B}^*\mathbb{L}^{-1}_{\sqrt{\overline{G}(x)^2 + \varepsilon^2 I_n}}\left[\mathbb{L}_{\overline{G}(x)}(-\mathcal{B}\Delta x) + 2\varepsilon\Delta\varepsilon I_n\right] \end{bmatrix},$$

where  $N = -L_s + L_{v-u}L_w^{-1}L_sL_w^{-1}L_{v-u}$ . Next, we compute the term  $\mathcal{B}^*\mathbb{L}_{\sqrt{\bar{G}(x)^2 + \epsilon^2 I_n}}^{-1}\mathbb{L}_{\bar{G}(x)}(-\mathcal{B}\Delta x)$  in (6.1). Since  $\bar{G}(x) \in$  $\mathcal{S}^n$ , suppose that  $\bar{G}(x)$  has the spectral decomposition  $\bar{G}(x) = P\Sigma P^T$ , where  $\Sigma = \text{diag}(\sigma_1, \cdots, \sigma_n)$  is the diagonal matrix of eigenvalues of  $\overline{G}(x)$  and P is a corresponding orthogonal matrix of the orthogonal eigenvectors, then  $\sqrt{\bar{G}(x)^2 + \varepsilon^2 I_n} =$  $P\sqrt{\Sigma^2 + \varepsilon^2 I_n}P^T.$ 

Denote 
$$X := \mathcal{B}\Delta x \in \mathcal{S}^n_+, Y := \mathbb{L}^{-1}_{\sqrt{\bar{G}(x)^2 + \varepsilon^2 I_n}} \mathbb{L}_{\bar{G}(x)}(\mathcal{B}\Delta x) \in \mathcal{S}^n_+$$
, we have

(6.2) 
$$\sqrt{\bar{G}(x)^2 + \varepsilon^2 I_n} Y + Y \sqrt{\bar{G}(x)^2 + \varepsilon^2 I_n} = \bar{G}(x) X + X \bar{G}(x).$$

Then (6.2) can be written as

$$\sqrt{\Sigma^2 + \varepsilon^2 I_n} P^T Y P + \sqrt{\Sigma^2 + \varepsilon^2 I_n} P^T Y P = \Sigma P^T X P + P^T X P \Sigma,$$

hence

(6.3)  

$$\left(\sqrt{\sigma_i^2 + \varepsilon^2} + \sqrt{\sigma_j^2 + \varepsilon^2}\right) [P^T Y P]_{ij} = (\sigma_i + \sigma_j) [P^T X P]_{ij}, \quad i, j = 1, 2, \cdots, n.$$

Then

$$\mathbb{L}_{\sqrt{\bar{G}(x)^2 + \varepsilon^2 I_n}}^{-1} \mathbb{L}_{\bar{G}(x)}(-\mathcal{B}\Delta x) = -P(\Xi \bullet (P^T \mathcal{B}\Delta x P))P^T,$$

where "•" denotes the Hadamard product and  $\Xi \in \mathcal{S}^n_+$  is defined by

$$[\Xi]_{ij} = \frac{\sigma_i + \sigma_j}{\sqrt{\sigma_i^2 + \varepsilon^2} + \sqrt{\sigma_j^2 + \varepsilon^2}},$$

hence

$$\mathcal{B}^* \mathbb{L}_{\sqrt{\bar{G}(x)^2 + \varepsilon^2 I}}^{-1} \mathbb{L}_{\bar{G}(x)}(-\mathcal{B}\Delta x) = -P(\Xi \bullet (P^T \mathcal{B}\Delta x P))P^T x_0.$$

Since  $\mathcal{J}E(z)$  in (5.12) is a linear operator on  $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^q$ , it is difficult to express it explicitly without any vector. This implies that it is impractical to use direct methods to solve linear system (5.12). Given the fact that the operator  $\mathcal{J}E(z)$  is nonsymmetric, it is natural to choose the BiCGStab method as our iterative solver for solving the linear system (5.12), as  $\mathcal{J}E(z)(\Delta z)$  has the expression (6.1).

We report the numerical results of Algorithm 5.1 for solving  $(P_{\mu})$  with a sufficiently small perturbed parameter  $\mu$  in MATLAB 7.8 running on a PC Intel Core 2 of 2.40 GHz CPU and 4GB of RAM. The stopping criterion chosen for Algorithm 5.1 is

$$res_k := \|E(\varepsilon^k, x^k, u^k, v^k, \lambda^k)\| \le 10^{-5}.$$

The maximum number of BiCGStab step at each iteration is set as 300. The other parameters used in Algorithm 5.1 are set as  $\hat{\varepsilon} = 3.0$ ,  $\eta = 0.02$ ,  $\delta = 0.5$ ,  $\sigma = 0.3$ . The testing examples for the inverse problem (3.2) we considered are randomly generated, which are described below.

**Example 6.1.** The matrix  $A = (A_1; \dots; A_J)$  is a randomly generated  $q \times n$  full row rank matrix with entries in [-1,1] by MATLAB code:

$$A = 2.0 * rand(q, n) - ones(q, n).$$

The parameter set  $(x_0, G^0, c^0, b_{m_1+1}^0, \cdots, b_{m_J+1}^0)$  required in (1.6) is generated as follows: Let  $x \in \text{SOL}(\text{QSOCP})$  with  $G \in \mathcal{S}^n_+, c \in \mathbb{R}^n, b = (b_{m_1+1}; \cdots; b_{m_J+1}) \in \mathbb{R}^q$ be randomly generated. Then  $(x_0, G^0, c^0, b_{m_1+1}^0, \cdots, b_{m_J+1}^0)$  is obtained by perturbing  $(x, G, c, b_{m_1+1}, \cdots, b_{m_J+1})$ . We consider the inverse problem (3.2) with a single second-order cone complementarity constraint, that is, J = 1, and solve the corresponding perturbed problem  $(P_{\mu})$  with  $\mu = 1.0e - 5, 1.0e - 6, 1.0e - 7$ , and 1.0e - 8, respectively.

We take the initial  $u^0$ ,  $v^0$  and  $\lambda^0$  being the random  $q \times 1$  vectors and the initial point  $x^0$  being zero in  $\mathbb{R}^n$  in Algorithm 5.1. The random problems of each size are generated 50 times. The obtained results are shown in Table 1 and Table 2, and each column represents the following:

- iter.: the average number of iterations among the test problems with the same dimension.
- infea.: the average infeasibility among the test problems with the same dimension. The infeasibility of inverse problem (3.2) denotes the absolute value  $|\langle u^*, v^* \rangle|$  at the final iteration of Algorithm 5.1, where  $(u^*, v^*)$  is the finally iterative value when Algorithm 5.1 terminates.
- res.: the average residual norm of  $E(\cdot)$  at the final iteration among the test problems with the same dimension.
- time: the average cpu time among the test problems with the same dimension.

From Table 1 and Table 2, we can see that Algorithm 5.1 is able to achieve high accuracy as the infeasibility is less than  $5 \times 10^{-8}$  for all the cases. The number of iteration and the accuracy for the random problems with the same dimension does not change so much when the parameter  $\mu$  changes from 1.0e - 5 to 1.0e - 8. However, for a group of problems with the same dimension, Algorithm 5.1 takes a little longer cpu time to solve problem (P<sub>µ</sub>) with  $\mu = 1.0e - 5$  than with the other three parameters. When  $\mu \leq 1.0e - 8$ , by our numerical observation, the matrix  $L_{w^k}$  generated in some iterations may be close to be singular for n more than 200.

Therefore, the performance of Algorithm 5.1 for solving  $P_{\mu}$  with  $\mu = 1.0e - 6$ and  $\mu = 1.0e - 7$  is better than that with  $\mu = 1.0e - 5$  and  $\mu = 1.0e - 8$ . **Example 6.2.** All the data are the same as in Example 6.1. We consider the inverse problem (3.2) with multiple second-order cone complementarity constraints, and solve the corresponding perturbed problem  $(P_{\mu})$  with  $\mu \equiv 1.0e - 6$ .

The random problems of each size for Example 6.2 are generated 10 times, and the initial  $u^0, v^0, \lambda^0 \in \mathbb{R}^q$  are randomly generated and the initial point  $x^0$  is chosen to be zero in  $\mathbb{R}^n$ . The obtained results are listed in Table 3 with "Q" being the structure of second-order cones and "iter.", "infea.", "res." representing the same as which in Table 1 and Table 2. From the numerical results reported in Table 3, we can see that the cpu time is relevant to not only the variable dimensions m, q but also the number of second-order cone complementarity constraints in the inverse problem (3.2). For cases with the same dimensions m and q, Algorithm 5.1 for solving the corresponding perturbed problem ( $P_{\mu}$ ) with less second-order cone blocks in the largest numerical example we tested in this paper is n = 1000, q = 400. In this case, there are roughly 1,000,000 unknowns in the primal problem. Regarding the scales of problems solved, cpu time spent and the accuracy achieved in Table 3, the smoothing Newton method we adopted to solve ( $P_{\mu}$ ) is quite effective.

TABLE 1. Numerical results for Example 6.1 with  $\mu = 1.0e - 5$ and  $\mu = 1.0e - 6$ 

n	q	$\mu = 1.0e - 5$							
		iter.	infea.	res.	time	iter.	infea.	res.	time
50	20	26.7	2.3e-10	6.66e-06	1.6s	26.0	1.9e-09	7.51e-06	1.3s
100	40	26.5	2.2e-09	6.27e-06	4.9s	25.5	6.9e-09	7.96e-06	3.5s
200	80	29.0	1.1e-08	6.35e-06	25.2s	26.7	9.6e-10	7.14e-06	19.9s
300	120	32.0	3.9e-08	5.98e-06	2m49.7s	28.3	6.7e-09	7.32e-06	$1\mathrm{m}43.5\mathrm{s}$
500	200	29.0	1.9e-08	7.15e-06	$10\mathrm{m}49.3\mathrm{s}$	28.7	2.4e-09	5.91e-06	8m12.9s
800	320	32.1	4.1e-08	7.66e-06	49m54.1s	29.9	3.0e-09	7.30e-06	38m1.0s

TABLE 2. Numerical results for Example 6.1 with  $\mu = 1.0e - 7$ and  $\mu = 1.0e - 8$ 

	~		$\mu = 1.0e - 7$	7			$\mu = 1.0e - 3$	8	
$\pi$	q	iter.	infea.	res.	time	iter.	infea.	res.	time
50	20	25.7	5.4e-09	6.78e-06	1.4s	26.5	2.0e-09	7.00e-06	1.5s
100	40	26.1	1.2e-08	6.80e-06	3.8s	27.1	5.9e-10	7.71e-06	3.8s
200	80	27.7	2.5e-09	6.64e-06	21.8s	27.3	1.6e-08	6.95e-06	22.9s
300	120	26.8	2.3e-08	7.35e-06	1m44.3s	28.0	1.1e-08	6.77e-06	1m58.8s
500	200	28.5	4.6e-09	7.45e-06	8m21.8s	27.7	1.3e-10	5.60e-06	8m31.0s
800	320	29.0	1.6e-08	5.07e-06	38m41.3s	27.1	1.0e-08	6.40e-06	$40\mathrm{m}12.7\mathrm{s}$

n	q	Q	iter.	infea.	res.	time
50	20	$Q_{10}  imes Q_{10}$	33.3	9.2e-10	6.66e-06	3.2s
50	20	$Q_5  imes Q_5  imes Q_5  imes Q_5$	36.5	5.8e-09	6.14e-06	4.9s
100	40	$Q_{20}  imes Q_{20}$	33.8	4.2e-09	7.24e-06	7.6s
100	40	$Q_{10} \times Q_{10} \times Q_{10} \times Q_{10}$	35.9	1.8e-08	8.13e-06	11.7s
100	40	$Q_5  imes Q_5  imes Q_5  imes Q_5  imes Q_5  imes Q_5  imes Q_5  imes Q_5$	48.7	7.9e-09	1.92e-06	33.4s
200	80	$Q_{40}  imes Q_{40}$	34.1	7.3e-08	4.16e-06	34.6s
200	80	$Q_{20} \times Q_{20} \times Q_{20} \times Q_{20}$	37.5	6.5e-09	7.72e-06	59.4s
200	80	$Q_{16} \times Q_{16} \times Q_{16} \times Q_{16} \times Q_{16}$	35.8	1.5e-09	8.43e-06	1m7.8s
200	80	$Q_{10} \times Q_{10} \times Q_{10} \times Q_{10} \times Q_{10} \times Q_{10} \times Q_{10} \times Q_{10}$	45.4	9.8e-09	6.61e-06	1m26.3s
300	120	$Q_{60}  imes Q_{60}$	32.7	1.2e-08	7.11e-06	1m57.8s
300	120	$Q_{40} \times Q_{40} \times Q_{40}$	32.3	1.0e-08	8.37e-06	2m10.3s
300	120	$Q_{30} \times Q_{30} \times Q_{30} \times Q_{30}$	38.8	1.6e-09	9.31e-06	3m16.9s
300	120	$Q_{20} \times Q_{20} \times Q_{20} \times Q_{20} \times Q_{20} \times Q_{20}$	39.5	2.7e-10	7.21e-06	3m45.7s
400	160	$Q_{80}  imes Q_{80}$	29.7	4.7e-08	8.64e-06	4m42.2s
400	160	$Q_{40}  imes Q_{40}  imes Q_{40}  imes Q_{40}$	37.6	1.8e-08	7.82e-06	5m45.8s
400	160	$Q_{20} \times Q_{20} \times Q_{20} \times Q_{20} \times Q_{20} \times Q_{20} \times Q_{20} \times Q_{20}$	38.0	2.5e-09	7.07e-06	9m15.3s
500	200	$Q_{100}  imes Q_{100}$	28.0	3.9e-09	7.83e-06	8m15.5s
500	200	$Q_{50} imes Q_{50} imes Q_{50} imes Q_{50}$	37.1	2.8e-08	7.72e-06	11m34.7s
500	200	$Q_{25} \times Q_{25} \times Q_{25} \times Q_{25} \times Q_{25} \times Q_{25} \times Q_{25} \times Q_{25}$	40.2	9.5e-10	8.55e-06	15m47.5s
600	240	$Q_{120} \times Q_{120}$	36.0	2.6e-08	7.07e-06	17m28.3s
600	240	$Q_{60}  imes Q_{60}  imes Q_{60}  imes Q_{60}$	38.5	6.4e-09	$8.04\mathrm{e}{\text{-}06}$	22m15.7s
600	240	$Q_{40}  imes Q_{40}  imes Q_{40}  imes Q_{40}  imes Q_{40}  imes Q_{40}$	35.2	8.9e-09	8.03e-06	24m52.5s
600	240	$Q_{30} \times Q_{30} \times Q_{30} \times Q_{30} \times Q_{30} \times Q_{30} \times Q_{30} \times Q_{30}$	47.4	5.8e-10	5.65e-06	46m3.9s
800	320	$Q_{160} \times Q_{160}$	32.3	1.3e-08	6.06e-06	41m40.6s
800	320	$Q_{80} \times Q_{80} \times Q_{80} \times Q_{80}$	34.4	1.4e-09	8.25e-06	49m1.1s
800	320	$Q_{40} \times Q_{40} \times Q_{40} \times Q_{40} \times Q_{40} \times Q_{40} \times Q_{40} \times Q_{40}$	35.7	2.2e-09	8.28e-06	1h7m7.9s
1000	400	$Q_{200} \times Q_{200}$	33.4	8.3e-10	8.21e-06	1h2m10.0s
1000	400	$Q_{100} \times Q_{100} \times Q_{100} \times Q_{100}$	34.0	1.1e-08	1.44e-06	1h14m9.3s
1000	400	$Q_{50} \times Q_{50} \times Q_{50} \times Q_{50} \times Q_{50} \times Q_{50} \times Q_{50} \times Q_{50}$	39.2	2.4e-09	8.56e-06	1h38m21.4s

TABLE 3. Numerical results for Example 6.2

## 7. CONCLUSION

In this paper, we study an inverse second-order cone quadratic programming problem in which the parameters in both the objective function and the constraint set need to be adjusted. By the duality theory, we formulate the inverse problem as a linear second-order cone complementarity constrained optimization problem with semismoothly differentiable objective function involving a mapping x(u) which is defined by a Yosida-regularization. We use a smoothing function to approximate the complementarity relation and demonstrate the outer semicontinuous of the optimal solution set of the perturbed problem  $(P_{\mu})$ . The smoothing Newton method with Armijo line search is employed to solve  $(P_{\mu})$  and the numerical results reported show that our algorithm is quite effective.

The methodology in this paper can be extended to solve other inverse conic optimization problems in which both the objective function and the constraint set need to be adjusted.

### Acknowledgement

The authors are grateful to the referees for their valuable comments and suggestions on improving the quality of this paper.

#### References

- Ravindra K. Ahuja and James B. Orlin, *Inverse optimization*, Oper. Res. 49 (2001), no. 5, 771–783, DOI 10.1287/opre.49.5.771.10607. MR1860425 (2002i:90097)
- F. Alizadeh and D. Goldfarb, Second-order cone programming, Math. Program. 95 (2003), no. 1, Ser. B, 3–51, DOI 10.1007/s10107-002-0339-5. ISMP 2000, Part 3 (Atlanta, GA). MR1971381 (2004j:90060)
- [3] D. Burton and Ph. L. Toint, On an instance of the inverse shortest paths problem, Math. Programming 53 (1992), no. 1, Ser. A, 45–61, DOI 10.1007/BF01585693. MR1151764 (93e:90063)
- M. C. Cai, X. G. Yang, and J. Z. Zhang, The complexity analysis of the inverse center location problem, J. Global Optim. 15 (1999), no. 2, 213–218, DOI 10.1023/A:1008360312607. MR1714066 (2000e:90029)
- [5] Jein-Shan Chen, Defeng Sun, and Jie Sun, The SC<sup>1</sup> property of the squared norm of the SOC Fischer-Burmeister function, Oper. Res. Lett. 36 (2008), no. 3, 385–392, DOI 10.1016/j.orl.2007.08.005. MR2424468 (2009g:90123)
- [6] Xiaojun Chen and Masao Fukushima, A smoothing method for a mathematical program with P-matrix linear complementarity constraints, Comput. Optim. Appl. 27 (2004), no. 3, 223– 246, DOI 10.1023/B:COAP.0000013057.54647.6d. MR2035714 (2005a:90159)
- [7] Frank H. Clarke, Optimization and Nonsmooth Analysis, Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley & Sons Inc., New York, 1983. A Wiley-Interscience Publication. MR709590 (85m:49002)
- [8] Zvi Drezner and Horst W. Hamacher (eds.), Facility location, Springer-Verlag, Berlin, 2002. Applications and theory. MR1933965 (2003e:90003)
- [9] Francisco Facchinei, Houyuan Jiang, and Liqun Qi, A smoothing method for mathematical programs with equilibrium constraints, Math. Program. 85 (1999), no. 1, Ser. A, 107–134, DOI 10.1007/s101070050048. MR1689366 (2000f:90065)
- [10] Francisco Facchinei and Jong-Shi Pang, Finite-dimensional Variational Inequalities and Complementarity Problems. Vol. I, Springer Series in Operations Research, Springer-Verlag, New York, 2003. MR1955648 (2004g:90003a)
- [11] Masao Fukushima, Zhi-Quan Luo, and Paul Tseng, Smoothing functions for second-ordercone complementarity problems, SIAM J. Optim. 12 (2001/02), no. 2, 436–460 (electronic), DOI 10.1137/S1052623400380365. MR1885570 (2002m:90084)
- [12] Garud Iyengar and Wanmo Kang, Inverse conic programming with applications, Oper. Res. Lett. 33 (2005), no. 3, 319–330, DOI 10.1016/j.orl.2004.04.007. MR2108284 (2005g:90101)
- [13] Houyuan Jiang and Daniel Ralph, Smooth SQP methods for mathematical programs with nonlinear complementarity constraints, SIAM J. Optim. 10 (2000), no. 3, 779–808 (electronic), DOI 10.1137/S1052623497332329. MR1774773 (2001f:90073)
- [14] Jianlin Jiang and Xiaoming Yuan, A Barzilai-Borwein-based heuristic algorithm for locating multiple facilities with regional demand, Comput. Optim. Appl. 51 (2012), no. 3, 1275–1295, DOI 10.1007/s10589-010-9392-9. MR2891938
- [15] Zhi-Quan Luo, Jong-Shi Pang, and Daniel Ralph, Mathematical Programs with Equilibrium Constraints, Cambridge University Press, Cambridge, 1996. MR1419501 (97j:90002)
- [16] Fanwen Meng, Defeng Sun, and Gongyun Zhao, Semismoothness of solutions to generalized equations and the Moreau-Yosida regularization, Math. Program. 104 (2005), no. 2-3, Ser. B, 561–581, DOI 10.1007/s10107-005-0629-9. MR2179251 (2006m:90173)
- [17] Li Qun Qi and Jie Sun, A nonsmooth version of Newton's method, Math. Programming 58 (1993), no. 3, Ser. A, 353–367, DOI 10.1007/BF01581275. MR1216791 (94b:90077)
- [18] Liqun Qi, Defeng Sun, and Guanglu Zhou, A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities, Math. Program. 87 (2000), no. 1, Ser. A, 1–35. MR1734657 (2001e:90113)
- [19] R. Tyrrell Rockafellar and Roger J.-B. Wets, Variational Analysis, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 317, Springer-Verlag, Berlin, 1998. MR1491362 (98m:49001)
- [20] Stefan Scholtes and Michael Stöhr, Exact penalization of mathematical programs with equilibrium constraints, SIAM J. Control Optim. 37 (1999), no. 2, 617–652 (electronic), DOI 10.1137/S0363012996306121. MR1670641 (99k:90179)
- [21] Stefan Scholtes, Convergence properties of a regularization scheme for mathematical programs with complementarity constraints, SIAM J. Optim. 11 (2001), no. 4, 918–936 (electronic), DOI 10.1137/S1052623499361233. MR1855214 (2002g:90138)

- [22] Jie Sun, Defeng Sun, and Liqun Qi, A squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems, SIAM J. Optim. 14 (2003), no. 3, 783–806, DOI 10.1137/S1052623400379620. MR2085943 (2005g:90098)
- [23] Defeng Sun, The strong second-order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications, Math. Oper. Res. **31** (2006), no. 4, 761–776, DOI 10.1287/moor.1060.0195. MR2281228 (2007m:90060)
- [24] Defeng Sun, Jie Sun, and Liwei Zhang, The rate of convergence of the augmented Lagrangian method for nonlinear semidefinite programming, Math. Program. 114 (2008), no. 2, Ser. A, 349–391, DOI 10.1007/s10107-007-0105-9. MR2393047 (2009c:90077)
- [25] Jia Wu, Liwei Zhang, and Yi Zhang, A smoothing Newton method for mathematical programs governed by second-order cone constrained generalized equations, J. Global Optim. 55 (2013), no. 2, 359–385, DOI 10.1007/s10898-012-9880-9. MR3015981
- [26] Xiantao Xiao and Liwei Zhang, Solving a class of inverse QP problems by a smoothing Newton method, J. Comput. Math. 27 (2009), no. 6, 787–801, DOI 10.4208/jcm.2009.09m2674. MR2583394 (2011a:90079)
- [27] Xiantao Xiao, Liwei Zhang, and Jianzhong Zhang, A smoothing Newton method for a type of inverse semi-definite quadratic programming problem, J. Comput. Appl. Math. 223 (2009), no. 1, 485–498, DOI 10.1016/j.cam.2008.01.028. MR2463131 (2010a:90067)
- [28] X. Xiao, L.Zhang and J. Zhang, On convergence of augmented Lagrange method for inverse semidefinite quadratic programming problems, Journal of Industrial and Management Optimization 52(2009), 319-339.
- [29] Jianzhong Zhang and Zhenhong Liu, Calculating some inverse linear programming problems, J. Comput. Appl. Math. **72** (1996), no. 2, 261–273, DOI 10.1016/0377-0427(95)00277-4. MR1406213 (97d:90056)
- [30] Jianzhong Zhang and Zhenhong Liu, A further study on inverse linear programming problems,
   J. Comput. Appl. Math. 106 (1999), no. 2, 345–359, DOI 10.1016/S0377-0427(99)00080-1.
   MR1696416 (2000c:90047)
- [31] Jianzhong Zhang, Zhenhong Liu, and Zhongfan Ma, Some reverse location problems, European J. Oper. Res. **124** (2000), no. 1, 77–88, DOI 10.1016/S0377-2217(99)00122-8. MR1778308 (2001j:90039)
- [32] Jianzhong Zhang and Liwei Zhang, An augmented Lagrangian method for a class of inverse quadratic programming problems, Appl. Math. Optim. 61 (2010), no. 1, 57–83, DOI 10.1007/s00245-009-9075-z. MR2575314 (2011b:90155)
- [33] Jianzhong Zhang, Liwei Zhang, and Xiantao Xiao, A perturbation approach for an inverse quadratic programming problem, Math. Methods Oper. Res. 72 (2010), no. 3, 379–404, DOI 10.1007/s00186-010-0323-4. MR2739050 (2011j:90119)
- [34] Yi Zhang, Liwei Zhang, and Jia Wu, Convergence properties of a smoothing approach for mathematical programs with second-order cone complementarity constraints, Set-Valued Var. Anal. 19 (2011), no. 4, 609–646, DOI 10.1007/s11228-011-0190-z. MR2836713 (2012j:90180)
- [35] Y. Zhang, J. Wu and L. Zhang, First order necessary optimality conditions for mathematical programs with second-order cone complementarity constraints, submitted.

Department of Mathematics, School of Science, East China University of Science and Technology, Shanghai, 200237, China.

### E-mail address: zhangyi8407@163.com

School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China.

E-mail address: lwzhang@dlut.edu.cn

School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China.

E-mail address: wujia@dlut.edu.cn

Division of Science and Technology, Beijing Normal University-Hong Kong Baptist University United International College, Zhuhai, 519085, China.

E-mail address: jzzhang@uic.edu.hk