A Smoothing Method for Mathematical Programs with Second-Order Cone Complementarity Constraints

Guidance

Professor Masao FUKUSHIMA

Takeshi EJIRI

2005 Graduate Course

in

Department of Applied Mathematics and Physics

Graduate School of Informatics

Kyoto University

February 2007

Abstract

Real-world problems are subject to various uncertainties. To cope with difficulties arising in those situations, there has been increasing attention to optimization problems under uncertainty in recent years. However, just a few works have dealt with Mathematical Programs with Equilibrium Constraints (MPEC) under uncertainty. In particular, there have been very few studies on a robust optimization approach to those problems.

In this paper, we formulate an MPEC under uncertainty as a Mathematical Program with Second-Order Cone Complementarity Constraints (MPSOCCC). Since the MPSOCCC is difficult to deal with directly, we use a smoothing function to approximate the problem and solve a sequence of smooth approximation problems to get a solution of the original MPSOCCC. The main theoretical result of this work is to establish that the proposed smoothing method has the global convergence property to a stationary point of the MPSOCCC.

Contents

1 Introduction

Optimization problems whose constraints contain complementarity conditions or variational inequalities are called Mathematical Programs with Equilibrium Constraints (MPECs). Although MPECs are very difficult problems to deal with because of their nonconvexity and nondifferentiability, they have high ability of formulating a wide range of practical problems. Therefore researchers have paid much attention to them and actively studied them in recent years. In particular, a bilevel program is a special case of MPEC, which includes many important problems such as a leader-follower decision-making problem. The bilevel program can be reformulated as a Mathematical Program with Complementarity Constraints (MPCC), by replacing the lower-level optimization problem by its KKT conditions. For MPECs, MPCCs and bilevel programs, there are many research results about optimality conditions and effective algorithms, such as branch-and-bound methods [2, 14, 18, 22], complementary pivoting algorithm [4] and smoothing methods [9, 12, 20].

While optimization can be applied to a wide variety of real problems and systems, a majority of the methods dealing with those problems treat the data as exactly known, that is, those methods get a solution by formulating an optimization problem with deterministic data. In problems we really face, however, it is in fact a rare case that the related information is perfectly known, but rather there are uncertain data almost always. To overcome these issues, there have recently been proposed a number of solution methods which even include uncertainty into problem formulation, instead of replacing uncertain data by some fixed data. For example, some methods use the expected values of stochastic data [13], or their variances [8], while other methods called robust optimization [3] seek a solution in the worst possible scenario. Especially, the worst case optimization is essential in dealing with actual problems with uncertain information, in view of the importance of accident prevention in dangerous operations which might lead to disasters like explosions in a large scale chemical plant, or risk avoidance in the finance field which has grown massively with derivatives such as futures trading. In a robust optimization method, if the data are known to belong to some ellipsoidal set, an optimization problem with uncertain data can be represented as a Second-Order Cone Programming (SOCP) problem [1, 3].

On the other hand, there have been just a few works to deal with MPECs under uncertainty [17, 21]. Particularly, very few studies have been done with regard to a robust optimization method for such problems, in spite of its significance. When there is uncertainty in the lower level optimization problem of a bilevel program, it can be formulated as a bilevel program having SOCP as its lower level problem. This paper discusses Mathematical Programs with Second-Order Cone Complementarity Constraints (MPSOCCC), which arise when the lower level SOCP in the bilevel problem is replaced by its KKT conditions. SOCP includes Linear Programming and Quadratically Constrained Quadratic Programming. Moreover, Second-Order Cone Complementarity Conditions (SOCCC) include the ordinary Complementarity Conditions as a special case. Hence MPSOCCC is an upper class of MPCC.

In this paper, we propose a smoothing method for solving MPSOCCC, and discuss its convergence properties. Similarly to MPCC and MPLCC, MPSOCCC is a nonconvex and nondifferentiable problem in general, and therefore it is difficult to deal with directly. Using a smoothing function [11], the proposed method constructs a smooth approximation of the nonsmooth problem. The main result of this paper is that the proposed method is guaranteed to have the global convergence property to a stationary point of the MPSOCCC under some assumptions.

The remainder of this paper is organized as follows: In Section 2, we formulate the MPSOCCC and review the Jordan algebra associated with the Second-Order Cones. In Section 3, we consider a smoothing function and propose a smoothing method for the MPSOCCC. The optimality conditions and convergence properties are discussed in Section 4. Lastly in Section 5, we make some remarks to conclude this paper.

2 Preliminaries

2.1 Bilevel program and MPSOCCC

We consider the following bilevel programming problem:

$$\begin{aligned} \mathrm{P}: \quad \min_{x,y} \quad & f(x,y) \\ \text{s.t.} \quad & x \in X, \\ & y \in Y(x), \end{aligned}$$

where $f : \Re^{n_1 \times n_2} \to \Re$ is a continuously differentiable function, $X \subseteq \Re^{n_1}$, and $Y(x) \subseteq \Re^{n_2}$ represents the solution set of the following Second-Order Cone Program (SOCP) with parameter x:

$$\begin{aligned} \mathrm{SOCP}(x): \quad \min_{y} \quad & g(x,y) \\ \text{s.t.} \quad & h(x,y) := A(x)y + b(x) = 0, \\ & y \in \mathcal{K}, \end{aligned}$$

where function $g : \Re^{n_1 \times n_2} \to \Re$ is continuously differentiable, $A(x) = [A_1(x) \ \cdots \ A_s(x)] \in \Re^{m \times n_2}$ with $A_i(x) \in \Re^{m \times l_i}$, $b(x) \in \Re^m$, and the set $\mathcal{K}$ is the Cartesian product of second-order cones $\mathcal{K}^{l_i} = \{y = (y_1, y_2)^\top \in \Re \times \Re^{l_i - 1} \mid y_1 \geq \|y_2\|\}$ $(i = 1, \cdots, s)$, i.e., $\mathcal{K} = \mathcal{K}^{l_1} \times \cdots \times \mathcal{K}^{l_s}$ with $n_2 = l_1 + \cdots + l_s$.

First we consider the reformulation of P. For this purpose, we make the following assumptions on problems P and SOCP(x):

A 1. *$X \subseteq \Re^{n_1}$ is nonempty and compact.*

A 2. *For any $x \in X$, the feasible set of SOCP(x) is nonempty.*

A 3. *For any $x \in X$, $g(x, \cdot)$ is strongly convex.*

Since the second-order cone $\mathcal{K}^{l_i}$ is a convex set, the problem SOCP(x) is a convex programming problem under assumption A3. Moreover, from assumptions A2 and A3, the problem SOCP(x) has a unique solution for every $x \in X$. However the upper level problem P is generally not a convex programming problem since the feasible set of P is not a convex set in general, even when SOCP(x) has a unique solution for all $x \in X$.

Let $\mathcal{I}(y)$ denote the index set such that $\mathcal{I}(y) := \{i \mid y^i \in \mathrm{bd}\mathcal{K}^{l_i}\}$ for any $y \in \Re^{n_2}$, where $y^i \in \Re^{l_i}$ is the block component of y and $\mathrm{bd}\mathcal{K}^{l_i}$ denotes the boundary of $\mathcal{K}^{l_i}$. Furthermore, for the matrix $A(x) = [A_1(x) \ \cdots \ A_s(x)]$, let $[A_i(x)]_{i\in\mathcal{I}(y)} \in \Re^{m\times\sum_{i\in\mathcal{I}} l_i}$ denote the submatrix consisting of $A_i(x)$, $i \in \mathcal{I}(y)$. Similarly, we denote the submatrix with components $A_i(x)$, $i \notin \mathcal{I}(y)$, by $[A_i(x)]_{i\notin\mathcal{I}(y)} \in \Re^{m\times\sum_{i\notin\mathcal{I}} l_i}$. With these definitions, we introduce the following assumption.

A 4. *For any $x \in X$ and any optimal solution y of SOCP(x), there exists an index i such that $i \notin \mathcal{I}(y)$. Furthermore, $[A_i(x)]_{i\notin\mathcal{I}(y)}$ has full row rank.*

The KKT conditions for SOCP(x) are given by

$$\begin{array}{c} \nabla_y g(x,y) - \eta - A(x)^\top \zeta = 0, \\ A(x)y + b(x) = 0, \\ y \in \mathcal{K},\ \eta \in \mathcal{K},\ y^\top \eta = 0, \end{array} \tag{1}$$

where $\eta \in \Re^{n_2}$ and $\zeta \in \Re^m$ are Lagrange multiplier vectors. The last conditions in the above conditions (1) are called Second-Order Cone Complementarity Constraints (SOCCC). Since SOCP(x) is a convex program under our assumptions, the above KKT conditions (1) become the necessary and sufficient optimality condition for this problem. Thus, by replacing SOCP(x) with its KKT conditions, we can reformulate problem P into the nonlinear programming problem involving SOCCC:

$$\begin{array}{lll} \text{MPSOCCC}: & \min & f(x,y) \\ & \text{s.t.} & x \in X, \\ & & \nabla_y g(x,y) - \eta - A(x)^\top \zeta = 0, \\ & & A(x)y + b(x) = 0, \\ & & y \in \mathcal{K},\ \eta \in \mathcal{K},\ y^\top \eta = 0. \end{array}$$

2.2 Jordan algebra associated with the SOC

In this subsection, we introduce Jordan algebra, which provides a useful tool for dealing with SOCs.

For two n-dimensional vectors $x = (x_1, x_2) \in \Re \times \Re^{n-1}$ and $y = (y_1, y_2) \in \Re \times \Re^{n-1}$, we define the Jordan product by

$$x \circ y := (x^\top y,\ x_1 y_2 + y_1 x_2)^\top.$$

Note that the Jordan product generates an n-dimensional vector from two n-dimensional vectors. For convenience, we denote $x \circ x$ by x^2, and define the vector e by $e := (1, 0, \cdots, 0)^\top$. Then Jordan product has the following properties [10]:

Property 1. *For any $x, y, z \in \Re^n$, the following properties hold.*

1. $e \circ x = x$,
2. $x \circ y = y \circ x$,

3. $(x + y) \circ z = x \circ z + y \circ z$,

4. $x \circ (x^2 + y) = x^2 \circ (x \circ y)$.

Property 1 shows that vector e plays the role of the unit vector. Properties 2 and 3 represent commutativity and distributivity, respectively. Notice that the Jordan product is not associative, that is, $x \circ (y \circ z) \neq (x \circ y) \circ z$ in general.

The next proposition shows that the second-order cone complementarity condition can be represented in another form by use of the Jordan product [11]:

Proposition 1. *For any two vectors x and y in $\Re^n$, the relation*

$$x \in \mathcal{K}^n,\ y \in \mathcal{K}^n,\ x^\top y = 0 \iff x \in \mathcal{K}^n,\ y \in \mathcal{K}^n,\ x \circ y = 0$$

holds.

Next we introduce the spectral factorization of vectors in $\Re^n$ associated with the second-order cone $\mathcal{K}^n$. Let $z = (z_1, z_2) \in \Re \times \Re^{n-1}$. Then we can decompose z as

$$z = \lambda_1 u^{\{1\}} + \lambda_2 u^{\{2\}},$$

where λ_1 and λ_2 are called the spectral values of z defined by

$$\lambda_i = z_1 + (-1)^i \|z_2\|,\ i = 1, 2, \tag{2}$$

and $u^{\{1\}}$ and $u^{\{2\}}$ are called the spectral vectors of z defined by

$$u^{\{i\}} = \begin{cases} \frac{1}{2}\left(1, (-1)^i \frac{z_2}{\|z_2\|}\right) & \text{if } z_2 \neq 0, \\ \frac{1}{2}\left(1, (-1)^i w\right) & \text{if } z_2 = 0, \end{cases} \qquad i = 1, 2, \tag{3}$$

where $w \in \Re^{n-1}$ is any vector satisfying $\|w\| = 1$. Some of the notable features of spectral values and vectors are the following [11, 15].

Property 2. *For any vector $z \in \Re^n$, the spectral values λ_1 and λ_2 given by (2) and spectral vectors $u^{\{1\}}$ and $u^{\{2\}}$ given by (3) possess the following properties:*

1. $u^{\{1\}} \circ u^{\{2\}} = 0,\ \|u^{\{1\}}\| = \|u^{\{2\}}\| = 1/\sqrt{2}$,

2. $u^{\{i\}} \in \mathrm{bd}\mathcal{K}^n,\ u^{\{i\}} \circ u^{\{i\}} = u^{\{i\}},\ i = 1, 2$,

3. $\lambda_1 \leq \lambda_2, \quad \lambda_1 \geq 0 \iff z \in \mathcal{K}^n$.

The projection mapping $P_{\mathcal{K}^n}$ onto the second-order cone $\mathcal{K}^n$ is defined by

$$P_{\mathcal{K}^n}(z) := \arg \min_{z' \in \mathcal{K}^n} \|z' - z\|.$$

By use of the spectral factorization, we can express this function explicitly as follows [11]:

$$P_{\mathcal{K}^n}(z) = \max\{0, \lambda_1\} u^{\{1\}} + \max\{0, \lambda_2\} u^{\{2\}}. \tag{4}$$

2.3 Reformulation of SOCCC into equation

The SOCCC on $\mathcal{K} = \mathcal{K}^{l_1} \times \cdots \times \mathcal{K}^{l_s}$,

$$y \in \mathcal{K},\ \ \eta \in \mathcal{K},\ \ y^\top \eta = 0, \tag{5}$$

can be written componentwise as follows [11, 15]:

$$y^i \in \mathcal{K}^{l_i},\ \ \eta^i \in \mathcal{K}^{l_i},\ \ (y^i)^\top \eta^i = 0 \ \ (i = 1, \cdots, s), \tag{6}$$

where $y = (y^1, \cdots, y^s)^\top \in \Re^{l_1} \times \cdots \times \Re^{l_s}$, $\eta = (\eta^1, \cdots, \eta^s)^\top \in \Re^{l_1} \times \cdots \times \Re^{l_s}$. For the decomposed SOCCC (6), we introduce functions $\varphi^i : \Re^{l_i} \times \Re^{l_i} \to \Re^{l_i}$ $(i = 1, \cdots, s)$ satisfying the following relation:

$$\varphi^i(y^i, \eta^i) = 0 \iff y^i \in \mathcal{K}^{l_i},\ \ \eta^i \in \mathcal{K}^{l_i},\ \ (y^i)^\top \eta^i = 0. \tag{7}$$

By using such functions, we can then define the equivalent equation to SOCCC such that

$$\Phi(y, \eta) = 0 \iff y \in \mathcal{K},\ \ \eta \in \mathcal{K},\ \ y^\top \eta = 0,$$

where $\Phi : \Re^{n_2} \times \Re^{n_2} \to \Re^{n_2}$ is a vector-valued function having functions φ^i $(i = 1, \cdots, s)$ as block components, i.e.,

$$\Phi(y, \eta) := \begin{pmatrix} \varphi^1(y^1, \eta^1) \\ \vdots \\ \varphi^s(y^s, \eta^s) \end{pmatrix}.$$

In [11], it is showed that the following function $\varphi^i_{\rm NR} : \Re^{l_i} \times \Re^{l_i} \to \Re^{l_i}$ $(i = 1, \cdots, s)$ defined by using the projection mapping $P_{\mathcal{K}^{l_i}}$ defined by (4) satisfies the relation (7):

$$\varphi^i_{\rm NR}(y^i, \eta^i) := y^i - P_{\mathcal{K}^{l_i}}(y^i - \eta^i).$$

This is called the natural residual function. Note that $P_{\mathcal{K}^{l_i}}$ is not differentiable. Using the functions $\varphi^i_{\rm NR}$, we define function $\Phi_{\rm NR} : \Re^{n_2} \times \Re^{n_2} \to \Re^{n_2}$ by

$$\Phi_{\rm NR}(y, \eta) := \begin{pmatrix} \varphi^1_{\rm NR}(y^1, \eta^1) \\ \vdots \\ \varphi^s_{\rm NR}(y^s, \eta^s) \end{pmatrix},$$

and $H_0 : \Re^{n_1} \times \Re^{n_2} \times \Re^{n_2} \times \Re^m \to \Re^{n_2} \times \Re^m \times \Re^{n_2}$ by

$$H_0(x, y, \eta, \zeta) := \begin{pmatrix} \nabla_y g(x, y) - \eta - A(x)^\top \zeta \\ A(x)y + b(x) \\ \Phi_{\rm NR}(y, \eta) \end{pmatrix}.$$

Then we can rewrite the KKT conditions (1) of SOCP(x) as

$$H_0(x, w) := H_0(x, y, \eta, \zeta) = 0.$$

Using the above relation, instead of MPSOCCC, we can consider the following equivalent nonlinear optimization problem:

$$\begin{aligned} \mathrm{P}_0: \quad \min \quad & f(x,y) \\ \text{s.t.} \quad & x \in X, \\ & H_0(x,y,\eta,\zeta) = 0. \end{aligned}$$

For any $x \in X$ and any optimal solution y of SOCP(x), let $(\eta, \zeta) \in \Re^{n_2} \times \Re^m$ denote a solution of the equation $H_0(x,y,\eta,\zeta) = 0$. Then for any $i \notin \mathcal{I}(y)$, we have from (7) that $\eta^i = 0$. Hence the first condition of the KKT conditions (1) is rewritten as

$$\begin{aligned} \eta^i + A_i(x)^\top \zeta &= \nabla_{y^i} g(x,y), \quad (i \in \mathcal{I}(y)), \\ \eta^i = 0, \qquad A_i(x)^\top \zeta &= \nabla_{y^i} g(x,y), \quad (i \notin \mathcal{I}(y)). \end{aligned} \tag{8}$$

Since matrix $[A_i(x)]_{i \notin \mathcal{I}(y)}$ has full row rank from assumption A4, the matrix

$$\begin{bmatrix} I_{\sum_{i\in\mathcal{I}(y)} l_i} & [A_i(x)]_{i\in\mathcal{I}(y)}^\top \\ 0 & [A_i(x)]_{i\notin\mathcal{I}(y)}^\top \end{bmatrix}$$

has full column rank. Therefore the solution of (8) exists uniquely.

The next theorem shows the equivalence of P and P_0.

Theorem 1. *Assume that A1–A4 hold. Then the next two statements are equivalent.*

1. (x^*, y^*) *is a global (local) optimal solution of problem P.*

2. *There exists a vector* (η^*, ζ^*) *such that* $(x^*, y^*, \eta^*, \zeta^*)$ *is a global (local) optimal solution of problem* P_0.

Proof. First, we assume that (x^*, y^*) is a local minimizer of problem P, and show the statement 2. Then from assumption A4, there exists a unique point (η^*, ζ^*) such that $(x^*, y^*, \eta^*, \zeta^*)$ is a feasible point of P_0. Suppose that $(x^*, y^*, \eta^*, \zeta^*)$ is not a local minimizer of P_0. Then there exists a sequence $\{(x^k, y^k, \eta^k, \zeta^k)\}$ of feasible points of P_0 converging to $(x^*, y^*, \eta^*, \zeta^*)$ and satisfying $f(x^k, y^k) < f(x^*, y^*)$ for all k. For such points $(x^k, y^k, \eta^k, \zeta^k)$, however, points (x^k, y^k) are feasible to P for all k, which contradicts the fact that (x^*, y^*) is a local optimal solution of problem P.

Next, assume that (x^*, y^*) is a global optimal solution of problem P. By what we have just mentioned, there is a unique vector (η^*, ζ^*) such that $(x^*, y^*, \eta^*, \zeta^*)$ is a local minimizer of P_0. If $(x^*, y^*, \eta^*, \zeta^*)$ is not a global minimizer of P_0, then there is another point $(\bar{x}, \bar{y}, \bar{\eta}, \bar{\zeta})$ satisfying $f(\bar{x}, \bar{y}) < f(x^*, y^*)$. The point $(\bar{x}, \bar{y})$ is, however, a feasible point of problem P, which therefore contradicts the assumption that (x^*, y^*) is a global optimal solution of problem P.

Conversely, assume now that $(x^*, y^*, \eta^*, \zeta^*)$ is a local minimizer of P_0. Since $y^* \in Y(x^*)$, (x^*, y^*) is a feasible solution of P. Suppose that (x^*, y^*) is not a local minimizer of P. Then there exists a sequence $\{(x^k, y^k)\}$ of feasible points of problem P which converges to (x^*, y^*) and satisfies $f(x^k, y^k) < f(x^*, y^*)$ for all k. From assumption A4, there exists a unique vector (η^k, ζ^k) such that $(x^k, y^k, \eta^k, \zeta^k)$ is a feasible point of problem P_0 for all k. If (η^k, ζ^k) converges to (η^*, ζ^*), we have a contradiction with the local optimality of $(x^*, y^*, \eta^*, \zeta^*)$. We will assume that (η^k, ζ^k) does not converge to (η^*, ζ^*), and show contradiction.

To begin with, we assume that the sequence $\{(\eta^k, \zeta^k)\}$ is unbounded. Since $(x^k, y^k, \eta^k, \zeta^k)$ is feasible to problem P_0 for all k, it follows that

$$\nabla_y g(x^k, y^k) - \eta^k - A(x^k)^\top \zeta^k = 0.$$

Dividing both sides of this equation by $\|(\eta^k, \zeta^k)^\top\|$, and passing to the limit $k \to \infty$, we have

$$(I \;\; A(x^*)^\top) \begin{pmatrix} \tilde{\eta} \\ \tilde{\zeta} \end{pmatrix} = 0, \tag{9}$$

where we assume, without loss of generality, $\lim_{k\to\infty} \frac{(\eta^k,\zeta^k)^\top}{\|(\eta^k,\zeta^k)^\top\|} = (\tilde{\eta}, \tilde{\zeta})^\top \neq 0$. From SOCCC and the continuity of functions, we obtain $\tilde{\eta}^i = 0$ for index $i \notin \mathcal{I}(y^*)$. Therefore using a similar reasoning for the uniqueness of vector (η, ζ) satisfying (8), the vector $(\tilde{\eta}, \tilde{\zeta})$ satisfying (9) is uniquely determined as $(\tilde{\eta}, \tilde{\zeta}) = 0$. But this contradicts $(\tilde{\eta}, \tilde{\zeta}) \neq 0$. Hence the sequence $\{(\eta^k, \zeta^k)\}$ is bonded. Then, without loss of generality, we can assume that there exists a point $(\bar{\eta}, \bar{\zeta})$ such that $(\eta^k, \zeta^k) \to (\bar{\eta}, \bar{\zeta})$. By the continuity of the functions, we see that the point $(x^*, y^*, \bar{\eta}, \bar{\zeta})$ is feasible to problem P_0 with $(x^*, y^*, \bar{\eta}, \bar{\zeta}) \neq (x^*, y^*, \eta^*, \zeta^*)$. From assumption A4, however, there is a unique vector (η, ζ) such that (x^*, y^*, η, ζ) is feasible to P_0. This is a contradiction.

Finally, suppose that $(x^*, y^*, \eta^*, \zeta^*)$ is a global optimal solution of problem P_0. If (x^*, y^*) is not a global optimal solution of problem P, then there exists a feasible solution $(\bar{x}, \bar{y})$ of P such that $f(\bar{x}, \bar{y}) < f(x^*, y^*)$. From assumption A4, there exists a unique vector $(\bar{\eta}, \bar{\zeta})$ such that $(\bar{x}, \bar{y}, \bar{\eta}, \bar{\zeta})$ is a feasible solution of problem P_0, which contradicts the global optimality of $(x^*, y^*, \eta^*, \zeta^*)$. Hence the point (x^*, y^*) is a global optimal solution of problem P. □

3 Smoothing method for problem P_0

In the previous section we have shown that problem P_0 is equivalent to problem P. Problem P_0 is no longer a bilevel program or an MPEC, but an ordinary nonlinear programming problem. On the other hand, owing to the nondifferentiability of function H_0, we cannot apply algorithms using derivatives of constraint functions, such as Newton-like methods. So in this section, we will discuss a smooth approximation of problem P_0, which is obtained by smoothing the function H_0.

For a nondifferentiable function $\psi : \Re^n \to \Re^m$, we consider a function $\psi_\mu : \Re^n \to \Re^m$ having the following properties with a parameter $\mu > 0$:

1. For any $\mu > 0$, ψ_μ is differentiable.
2. For any $x \in \Re^n$, $\lim_{\mu\to+0} \psi_\mu(x) = \psi(x)$.

Any function having these properties is called a smoothing function of ψ.

We will now discuss smoothing functions of φ^i_{NR} defined in the previous section. For this purpose, we first consider a continuously differentiable convex function $\hat{g} : \Re \to \Re$ having the properties

$$\lim_{\alpha\to-\infty} \hat{g}(\alpha) = 0, \;\; \lim_{\alpha\to\infty} (\hat{g}(\alpha) - \alpha) = 0, \;\; 0 < \hat{g}'(\alpha) < 1. \tag{10}$$

For example, the functions

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

and

$$\hat{g}(\alpha) = \log\left(\exp(\alpha) + 1\right)$$

have the above properties [7]. With the use of such $\hat{g}$ and a parameter $\mu > 0$, we can define $P_{\mathcal{K}^n,\mu} : \Re^n \to \Re^n$ by

$$P_{\mathcal{K}^n,\mu}(z) := \mu\hat{g}(\frac{\lambda_1}{\mu})u^{\{1\}} + \mu\hat{g}(\frac{\lambda_2}{\mu})u^{\{2\}},$$

where λ_1 and λ_2 are the spectral values of z defined by (2), and $u^{\{1\}}$ and $u^{\{2\}}$ are the spectral vectors of z defined by (3). Then this function $P_{\mathcal{K}^n,\mu}$ becomes a smoothing function of the projection mapping $P_{\mathcal{K}^n}$ onto the SOC $\mathcal{K}^n$ [11]. Hence using this function, we can define a smoothing function $\varphi^i_\mu : \Re^{l_i} \times \Re^{l_i} \to \Re^{l_i}$ of φ^i_{NR} by

$$\varphi^i_\mu(y^i, \eta^i) := y^i - P_{\mathcal{K}^{l_i},\mu}(y^i - \eta^i).$$

In addition, let $\Phi_\mu : \Re^{n_2} \times \Re^{n_2} \to \Re^{n_2}$ and $H_\mu : \Re^{n_1} \times \Re^{n_2} \times \Re^{n_2} \times \Re^m \to \Re^{n_2} \times \Re^m \times \Re^{n_2}$ be defined by

$$\Phi_\mu(y, \eta) := \begin{pmatrix} \varphi^1_\mu(y^1, \eta^1) \\ \vdots \\ \varphi^s_\mu(y^s, \eta^s) \end{pmatrix}$$

and

$$H_\mu(x, w) := H_\mu(x, y, \eta, \zeta) := \begin{pmatrix} \nabla_y g(x, y) - \eta - A(x)^\top \zeta \\ A(x)y + b(x) \\ \Phi_\mu(y, \eta) \end{pmatrix}.$$

Then for any $\mu > 0$, the function H_μ is differentiable and also has the properties of a smoothing function of H_0.

According to the above consideration, we can formulate a nonlinear programming problem with parameter $\mu > 0$ as follows:

$$\begin{array}{rll} \mathrm{P}_\mu : & \min & f(x, y) \\ & \text{s.t.} & x \in X, \\ & & H_\mu(x, w) = 0. \end{array}$$

This problem can be seen as an approximation of problem P_0 having a parameter $\mu > 0$. In particular, P_μ is a smooth optimization problem for any $\mu > 0$.

We now make new assumptions on function H_0 and the smoothing function H_μ:

A 5. *For any $\bar{x} \in X$ and $\bar{w} \in \Re^{2n_2+m}$ such that $H_0(\bar{x}, \bar{w}) = 0$, all the matrices belonging to the generalized Jacobian $\partial_w H_0(\bar{x}, \bar{w})$ are nonsingular.*

A 6. *For any $\bar{x} \in X$ and $\mu > 0$, there exists the inverse matrix $\nabla_w H_\mu(\bar{x}, w)^{-1} \in \Re^{(2n_2+m)\times(2n_2+m)}$ of the Jacobian of $H_\mu(\bar{x}, w)$ for all $w \in \Re^{2n_2+m}$. Furthermore, there exists a constant $\Gamma > 0$ such that $\|\nabla_w H_\mu(\bar{x}, w)^{-1}\| \le \Gamma$ for all $w \in \Re^{2n_2+m}$.*

Under assumption A6, Hadmard Theorem [19] ensures that $H_\mu(\bar{x}, \cdot)$ is homeomorphism of $\Re^{2n_2+m}$ onto $\Re^{2n_2+m}$ for any $\mu > 0$ and $\bar{x} \in X$, that is, the function $H_\mu(\bar{x}, \cdot)$ is one-to-one and $H_\mu(\bar{x}, \cdot)$ and $H_\mu(\bar{x}, \cdot)^{-1}$ are continuous. Especially, there exists a unique solution of the equation $H_\mu(\bar{x}, \cdot) = 0$.

For $\mu > 0$, let F_μ represent the feasible set of P_μ, i.e., $F_\mu := \{(x, w) \mid x \in X, H_\mu(x, w) = 0\}$. The following theorem implies that there exists an accumulation point of a sequence generated by the smoothing method.

Theorem 2. *Assume that A1–A6 hold, and let $\bar{\mu} > 0$ be any fixed constant. Then F_μ is nonempty and uniformly bounded on $[0, \bar{\mu}]$.*

Proof. From the definition of F_μ, assumption A1, and the fact that the equation $H_\mu(x, w) = 0$ $(\mu \ge 0)$ has at least one solution for any $x \in X$ (by assumption A4 or A6), there is $(x, w) \in F_\mu$ for any $\mu \in [0, \bar{\mu}]$. As X is compact, F_μ is bounded with respect to the x-component. Assuming that F_μ is not bounded for the w-component, we will derive a contradiction. Under this assumption, there are sequences $\{\mu^k\}$, $\{x^k\}$, and $\{w^k\}$ such that $(x^k, w^k) \in F_{\mu^k}$ and $\mu^k \to \tilde{\mu} \in [0, \bar{\mu}]$, $\|w^k\| \to \infty$ as $k \to \infty$. We can now assume without loss of generality that $\{x^k\}$ has a limit point $\tilde{x} \in X$. By assumptions A2–A4 and A6, there exists a unique $\tilde{w}$ satisfying $\tilde{w} \in F_{\tilde{\mu}}(\tilde{x})$, namely $H_{\tilde{\mu}}(\tilde{x}, \tilde{w}) = 0$ holds. On the other hand, from assumption A5 or A6, the (generalized) Jacobian of $H_{\tilde{\mu}}(\tilde{x}, \tilde{w})$ is nonsingular. Hence by the implicit function theorem there exists a continuous function $w(\mu, x)$ satisfying $H_\mu(x, w(\mu, x)) = 0$ in a neighborhood of the point $(\tilde{\mu}, \tilde{x})$. Since the solution of equation $H_\mu(x, w) = 0$ is also determined uniquely in a neighborhood of $(\tilde{\mu}, \tilde{x})$ by assumption A6, it eventually holds for sufficiently large k that $w^k = w(\mu^k, x^k)$. Since the function $w(\cdot, \cdot)$ is continuous, and since the sequences $\{\mu^k\}$ and $\{x^k\}$ are both bounded, there exists a constant $\beta > 0$ such that $\lim_{k\to\infty} \|w(\mu^k, x^k)\| \le \beta$, which contradicts the unboundedness of the sequence $\{w^k\}$. □

4 Optimality conditions

In this section, we discuss optimality conditions for problems P_0 and P_μ. In the case of $\mu > 0$, P_μ is a smooth nonlinear program and hence the optimality conditions can be written as KKT conditions. Our purpose is to obtain an optimal solution of MPSOCCC or P_0 from a KKT point of P_μ. To this end, we first refer to the optimality conditions for problem P_0.

In the following, we suppose that the set X is specified as $X := \{x \mid \tilde{g}(x) \le 0, \tilde{h}(x) = 0\}$, where $\tilde{g} : \Re^{n_1} \to \Re^p$ and $\tilde{h} : \Re^{n_1} \to \Re^q$ are both continuously differentiable functions.

Since P_0 is a nonsmooth problem, we cannot consider the ordinary KKT conditions. In [9], by using C-subdifferential of nondifferential functions, necessary conditions for optimality are represented as the following Fritz-John conditions: If the point (x, y, η, ζ) is optimal for

P_0, then there exist vectors $\delta \in \Re_+,\ \theta \in \Re^{n_2},\ \rho \in \Re^m,\ \sigma \in \Re^{n_2},\ \nu \in \Re^p_+,\ \xi \in \Re^q$ such that $(\delta, \theta, \rho, \sigma, \nu, \xi) \neq 0$ and

$$\left\{\begin{array}{l} 0 \in \delta \begin{pmatrix} \nabla_x f(x,y) \\ \nabla_y f(x,y) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \nabla_x L(x,y,\eta,\zeta) \\ \nabla_y L(x,y,\eta,\zeta) \\ \nabla_\eta L(x,y,\eta,\zeta) \\ \nabla_\zeta L(x,y,\eta,\zeta) \end{pmatrix} \theta + \begin{pmatrix} \nabla_x (A(x)y) + \nabla b(x) \\ A(x)^\top \\ 0 \\ 0 \end{pmatrix} \rho \\ \quad + \begin{pmatrix} 0 \\ \begin{bmatrix} \partial_{y^1} \varphi^1_{\mathrm{NR}}(y^1,\eta^1) & & 0 \\ & \ddots & \\ 0 & & \partial_{y^s} \varphi^s_{\mathrm{NR}}(y^s,\eta^s) \end{bmatrix} \\ \begin{bmatrix} \partial_{\eta^1} \varphi^1_{\mathrm{NR}}(y^1,\eta^1) & & 0 \\ & \ddots & \\ 0 & & \partial_{\eta^s} \varphi^s_{\mathrm{NR}}(y^s,\eta^s) \end{bmatrix} \\ 0 \end{pmatrix} \sigma + \begin{pmatrix} \nabla_x \tilde{g}(x) \\ 0 \\ 0 \\ 0 \end{pmatrix} \nu + \begin{pmatrix} \nabla_x \tilde{h}(x) \\ 0 \\ 0 \\ 0 \end{pmatrix} \xi, \\ L(x,y,\eta,\zeta) = 0, \\ A(x)y + b(x) = 0, \\ \varphi^i_{\mathrm{NR}}(y^i,\eta^i) = 0,\ (i = 1, \cdots, s), \\ \tilde{h}(x) = 0, \\ \tilde{g}(x) \leq 0,\ \nu \geq 0,\ \nu^\top \tilde{g}(x) = 0, \end{array}\right. \tag{11}$$

where $L : \Re^{n_1} \times \Re^{n_2} \times \Re^{n_2} \times \Re^m \to \Re^{n_2}$ is defined by $L(x,y,\eta,\zeta) := \nabla_y g(x,y) - \eta - A(x)^\top \zeta$. According to [16], by letting $z^i := y^i - \eta^i$, the C-subdifferential of the function $\varphi^i_{\mathrm{NR}}(y^i,\eta^i)$ is calculated as

$$\begin{aligned} \partial_{y^i} \varphi^i_{\mathrm{NR}}(y^i,\eta^i) &= I_{l_i} - \partial P_{\mathcal{K}^{l_i}}(z^i), \\ \partial_{\eta^i} \varphi^i_{\mathrm{NR}}(y^i,\eta^i) &= \partial P_{\mathcal{K}^{l_i}}(z^i), \end{aligned} \tag{12}$$

where in the case of $z^i_1 \neq \pm ||z^i_2||$, $P_{\mathcal{K}^{l_i}}$ is differentiable and its derivative is given as

$$\nabla P_{\mathcal{K}^{l_i}}(z^i) = \begin{cases} 0, & (z^i_1 < -\|z^i_2\|) \\ I_{l_i}, & (z^i_1 > \|z^i_2\|) \\ \dfrac{1}{2} \begin{pmatrix} 1 & (v^i)^\top \\ v^i & M \end{pmatrix}, & (-\|z^i_2\| < z^i_1 < \|z^i_2\|) \end{cases} \tag{13}$$
$$\text{with} \ \ v^i := \frac{z^i_2}{\|z^i_2\|},\ M := \left(\frac{z^i_1}{\|z^i_2\|} + 1 \right) I_{l_i - 1} - \frac{z^i_1}{\|z^i_2\|} v^i (v^i)^\top,$$

while, in the case of $z^i_1 = \pm ||z^i_2||$, $P_{\mathcal{K}^{l_i}}$ is not differentiable and its C-subdifferential is given as follows: if $z^i_2 \neq 0$ and $z^i_1 = \|z^i_2\|$, then

$$\partial P_{\mathcal{K}^{l_i}}(z^i) = \mathrm{co} \left\{ I_{l_i},\ \frac{1}{2} \begin{pmatrix} 1 & (v^i)^\top \\ v^i & M \end{pmatrix} \right\},\ \text{with} \ \ v^i := \frac{z^i_2}{\|z^i_2\|},\ M := 2I_{l_i - 1} - v^i (v^i)^\top, \tag{14}$$

if $z^i_2 \neq 0$ and $z^i_1 = -\|z^i_2\|$, then

$$\partial P_{\mathcal{K}^{l_i}}(z^i) = \mathrm{co} \left\{ 0,\ \frac{1}{2} \begin{pmatrix} 1 & (v^i)^\top \\ v^i & M \end{pmatrix} \right\},\ \text{with} \ \ v^i := \frac{z^i_2}{\|z^i_2\|},\ M := v^i (v^i)^\top, \tag{15}$$

finally if $z^i = 0$, then

$$\partial P_{\mathcal{K}^{l_i}}(z^i) = \mathrm{co}\left\{ \{0\} \cup \{I_{l_i}\} \cup \left\{ \frac{1}{2}\begin{pmatrix} 1 & (v^i)^\top \\ v^i & M \end{pmatrix} \;\middle|\; \begin{array}{c} M = (v_0^i + 1)I_{l_i - 1} - v_0^i v^i (v^i)^\top \\ \text{for some } |v_0^i| \leq 1 \text{ and } \|v^i\| = 1 \end{array} \right\} \right\}, \quad (16)$$

where co denotes the convex hull.

If $\delta \neq 0$ holds in the Fritz-John conditions (11), then we can assume without loss of generality that $\delta = 1$, and conditions (11) become the KKT conditions for problem P_0:

$$\left\{ \begin{array}{l} 0 \in \begin{pmatrix} \nabla_x f(x,y) \\ \nabla_y f(x,y) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \nabla_x L(x,y,\eta,\zeta) \\ \nabla_y L(x,y,\eta,\zeta) \\ \nabla_\eta L(x,y,\eta,\zeta) \\ \nabla_\zeta L(x,y,\eta,\zeta) \end{pmatrix} \theta + \begin{pmatrix} \nabla_x (A(x)y) + \nabla b(x) \\ A(x)^\top \\ 0 \\ 0 \end{pmatrix} \rho \\ \quad + \begin{pmatrix} 0 \\ \begin{bmatrix} \partial_{y^1} \varphi_{\mathrm{NR}}^1(y^1,\eta^1) & & 0 \\ & \ddots & \\ 0 & & \partial_{y^s} \varphi_{\mathrm{NR}}^s(y^s,\eta^s) \end{bmatrix} \\ \begin{bmatrix} \partial_{\eta^1} \varphi_{\mathrm{NR}}^1(y^1,\eta^1) & & 0 \\ & \ddots & \\ 0 & & \partial_{\eta^s} \varphi_{\mathrm{NR}}^s(y^s,\eta^s) \end{bmatrix} \\ 0 \end{pmatrix} \sigma + \begin{pmatrix} \nabla_x \tilde{g}(x) \\ 0 \\ 0 \\ 0 \end{pmatrix} \nu + \begin{pmatrix} \nabla_x \tilde{h}(x) \\ 0 \\ 0 \\ 0 \end{pmatrix} \xi, \\ L(x,y,\eta,\zeta) = 0, \\ A(x)y + b(x) = 0, \\ \varphi_{\mathrm{NR}}^i(y^i,\eta^i) = 0, \ (i = 1, \cdots, s), \\ \tilde{h}(x) = 0, \\ \tilde{g}(x) \leq 0, \ \nu \geq 0, \ \nu^\top \tilde{g}(x) = 0. \end{array} \right. \quad (17)$$

Now let $\tilde{\mathcal{I}}(x)$ denote the index set of the active constraints in $\tilde{g}_i(x) \leq 0$, i.e., $\tilde{\mathcal{I}}(x) := \{i \mid \tilde{g}_i(x) = 0\}$. We introduce the following linear independence assumption.

A 7. *For any point (x,y,η,ζ) satisfying the Fritz-John conditions (11), vectors $\nabla \tilde{g}_i(x)$ $(i \in \tilde{\mathcal{I}}(x))$, $\nabla \tilde{h}_j(x)$ $(j = 1, \cdots, q)$ are linearly independent.*

Under this assumption, it can be shown that every Fritz-John point satisfies the KKT conditions (17).

Theorem 3. *Assume that A1, A2, A5 and A7 hold. If (x,y,η,ζ) satisfies Fritz-John conditions (11) with multipliers $(\delta,\theta,\rho,\sigma,\nu,\xi)$, then δ is not equal to zero.*

Proof. Suppose that (x,y,η,ζ) satisfies the Fritz-John conditions (11) with $\delta = 0$. Then there exist $(\theta, \rho, \sigma, \nu, \xi) \neq 0$ satisfying the following system:

$$0 \in \begin{pmatrix} \nabla_x L(x,y,\eta,\zeta) \\ \nabla_y L(x,y,\eta,\zeta) \\ \nabla_\eta L(x,y,\eta,\zeta) \\ \nabla_\zeta L(x,y,\eta,\zeta) \end{pmatrix} \theta + \begin{pmatrix} \nabla_x (A(x)y) + \nabla b(x) \\ A(x)^\top \\ 0 \\ 0 \end{pmatrix} \rho$$

$$+ \begin{pmatrix} 0 \\ \begin{bmatrix} \partial_{y^1} \varphi^1_{\mathrm{NR}}(y^1,\eta^1) & & 0 \\ & \ddots & \\ 0 & & \partial_{y^s} \varphi^s_{\mathrm{NR}}(y^s,\eta^s) \end{bmatrix} \\ \begin{bmatrix} \partial_{\eta^1} \varphi^1_{\mathrm{NR}}(y^1,\eta^1) & & 0 \\ & \ddots & \\ 0 & & \partial_{\eta^s} \varphi^s_{\mathrm{NR}}(y^s,\eta^s) \end{bmatrix} \\ 0 \end{pmatrix} \sigma + \begin{pmatrix} \nabla_x \tilde{g}(x) \\ 0 \\ 0 \\ 0 \end{pmatrix} \nu + \begin{pmatrix} \nabla_x \tilde{h}(x) \\ 0 \\ 0 \\ 0 \end{pmatrix} \xi, \tag{18}$$

$$\begin{aligned} & L(x,y,\eta,\zeta) = 0, \\ & A(x)y + b(x) = 0, \\ & \varphi^i_{\mathrm{NR}}(y^i,\eta^i) = 0, \ (i = 1, \cdots, s), \\ & \tilde{h}(x) = 0, \\ & \tilde{g}(x) \leq 0, \ \nu \geq 0, \ \nu^\top \tilde{g}(x) = 0. \end{aligned} \tag{19}$$

Note that (18) can be rewritten as

$$0 = \left[\nabla_x L(x,y,\eta,\zeta) \ \nabla_x h(x,y) \ \nabla_x \tilde{g}(x) \ \nabla_x \tilde{h}(x) \right] \begin{bmatrix} \theta \\ \rho \\ \nu \\ \xi \end{bmatrix}, \tag{20}$$

$$0 \in \begin{bmatrix} \nabla_y L(x,y,\eta,\zeta) & A(x)^\top & \begin{bmatrix} \partial_{y^1} \varphi^1_{\mathrm{NR}}(y^1,\eta^1) & & 0 \\ & \ddots & \\ 0 & & \partial_{y^s} \varphi^s_{\mathrm{NR}}(y^s,\eta^s) \end{bmatrix} \\ \nabla_\eta L(x,y,\eta,\zeta) & 0 & \begin{bmatrix} \partial_{\eta^1} \varphi^1_{\mathrm{NR}}(y^1,\eta^1) & & 0 \\ & \ddots & \\ 0 & & \partial_{\eta^s} \varphi^s_{\mathrm{NR}}(y^s,\eta^s) \end{bmatrix} \\ \nabla_\zeta L(x,y,\eta,\zeta) & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta \\ \rho \\ \sigma \end{bmatrix}. \tag{21}$$

Since the matrix on the right-hand side of (21) represents the generalized Jacobian $\partial_w H_0(x,w)$, it is nonsingular by assumption A5. Hence (21) implies that $(\theta, \rho, \sigma) = 0$. Furthermore it follows from (20) together with the complementarity condition (19) and assumption A7 that $(\nu, \xi) = 0$, which contradicts $(\theta, \rho, \sigma, \nu, \xi) \neq 0$. □

Next we consider optimality conditions for problem P_μ. Since problem P_μ is a smooth optimization problem, its KKT conditions state that there exist Lagrange multipliers $\theta \in \Re^{n_2}$,

$\rho \in \Re^m$, $\sigma \in \Re^{n_2}$, $\nu \in \Re^p_+$ and $\xi \in \Re^q$ satisfying

$$\left\{\begin{array}{l}
0 \in \begin{pmatrix} \nabla_x f(x,y) \\ \nabla_y f(x,y) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \nabla_x L(x,y,\eta,\zeta) \\ \nabla_y L(x,y,\eta,\zeta) \\ \nabla_\eta L(x,y,\eta,\zeta) \\ \nabla_\zeta L(x,y,\eta,\zeta) \end{pmatrix} \theta + \begin{pmatrix} \nabla_x (A(x)y) + \nabla b(x) \\ A(x)^\top \\ 0 \\ 0 \end{pmatrix} \rho \\
\quad + \begin{pmatrix} 0 \\ \begin{bmatrix} \nabla_{y^1}\varphi^1_\mu(y^1,\eta^1) & & 0 \\ & \ddots & \\ 0 & & \nabla_{y^s}\varphi^s_\mu(y^s,\eta^s) \end{bmatrix} \\ \begin{bmatrix} \nabla_{\eta^1}\varphi^1_\mu(y^1,\eta^1) & & 0 \\ & \ddots & \\ 0 & & \nabla_{\eta^s}\varphi^s_\mu(y^s,\eta^s) \end{bmatrix} \\ 0 \end{pmatrix} \sigma + \begin{pmatrix} \nabla_x \tilde{g}(x) \\ 0 \\ 0 \\ 0 \end{pmatrix} \nu + \begin{pmatrix} \nabla_x \tilde{h}(x) \\ 0 \\ 0 \\ 0 \end{pmatrix} \xi, \\
L(x,y,\eta,\zeta) = 0, \\
A(x)y + b(x) = 0, \\
\varphi^i_\mu(y^i,\eta^i) = 0, \ (i = 1,\cdots,s), \\
\tilde{h}(x) = 0, \\
\tilde{g}(x) \le 0, \ \nu \ge 0, \ \nu^\top \tilde{g}(x) = 0,
\end{array}\right. \quad (22)$$

where $\nabla_{y^i}\varphi^i_\mu$ and $\nabla_{\eta^i}\varphi^i_\mu$ are calculated by letting $z^i := y^i - \eta^i$ as follows [11]:

$$\begin{aligned} \nabla_{y^i}\varphi^i_\mu(y^i,\eta^i) &= I_{l_i} - \nabla P_{\mathcal{K}^{l_i},\mu}(z^i), \\ \nabla_{\eta^i}\varphi^i_\mu(y^i,\eta^i) &= \nabla P_{\mathcal{K}^{l_i},\mu}(z^i), \end{aligned} \quad (23)$$

$$\nabla P_{\mathcal{K}^{l_i},\mu}(z^i) = \left\{\begin{array}{ll} \begin{bmatrix} b_i & c_i(v^i)^\top \\ c_i v^i & a_i I_{l_i-1} + (b_i - a_i)v^i(v^i)^\top \end{bmatrix}, & (z_2 \neq 0) \\ \hat{g}'(\frac{z_1}{\mu}) I_{l_i}, & (z_2 = 0), \end{array}\right. \quad (24)$$

where

$$v^i := \frac{z^i_2}{\|z^i_2\|},$$

$$a_i := \frac{\mu\left(\hat{g}(\frac{\lambda^i_2}{\mu}) - \hat{g}(\frac{\lambda^i_1}{\mu})\right)}{\lambda^i_2 - \lambda^i_1}, \ b_i := \frac{1}{2}\left(\hat{g}'(\frac{\lambda^i_2}{\mu}) + \hat{g}'(\frac{\lambda^i_1}{\mu})\right), \ c_i := \frac{1}{2}\left(\hat{g}'(\frac{\lambda^i_2}{\mu}) - \hat{g}'(\frac{\lambda^i_1}{\mu})\right), \quad (25)$$

$$\lambda^i_j := z^i_1 + (-1)^j \|z^i_2\|, \ j = 1,2. \quad (26)$$

The differences between these two optimality conditions (17) and (22) lie in the function φ^i_{NR} in (17) and φ^i_μ in (22), and the C-subdifferential and the ordinary differential for these functions. A limit of KKT points for P_μ may be expected to satisfy the optimality conditions (17) for problem P_0 under suitable assumptions, as shown in the next theorem. In fact, to get a solution of MPSOCCC, we may employ a smoothing method which sequentially computes KKT points for P_μ with $\mu > 0$ tending to zero.

Theorem 4. *Assume that A1, A2 and A7 hold. For any sequence* $\{\mu^k\}$ *such that* $\mu^k > 0$ *and* $\mu^k \to 0$, *let* $(x^k, y^k, \eta^k, \zeta^k)$ *satisfy the KKT conditions for* P_{μ^k} *with Lagrange multipliers* $(\theta^k, \rho^k, \sigma^k, \nu^k)$. *Assume that the sequence* $\{(x^k, y^k, \eta^k, \zeta^k, \theta^k, \rho^k, \sigma^k, \nu^k)\}$ *has an accumulation point. Then every accumulation point* $(x^*, y^*, \eta^*, \zeta^*, \theta^*, \rho^*, \sigma^*, \nu^*)$ *satisfies the optimality conditions (17) for* P_0.

Proof. Without loss of generality, we assume that the sequence $\{(x^k, y^k, \eta^k, \zeta^k, \theta^k, \rho^k, \sigma^k, \nu^k)\}$ converges to $(x^*, y^*, \eta^*, \zeta^*, \theta^*, \rho^*, \sigma^*, \nu^*)$ as $k \to \infty$. From the assumptions of the theorem, $(x^k, y^k, \eta^k, \zeta^k, \theta^k, \rho^k, \sigma^k, \nu^k)$ satisfies conditions (22) for all k. For the continuity of the functions involved, $\nabla f(x^k, y^k)$, $\nabla L(x^k, y^k, \eta^k, \zeta^k)\theta^k$, $\nabla h(x^k, y^k)\rho^k$, $\nabla \tilde{g}(x^k)\nu^k$, $\nabla \tilde{h}(x^k)\xi^k$, $L(x^k, y^k, \eta^k, \zeta^k)\theta^k$, $A(x^k)y^k + b(x^k)$, $\varphi^i_{\mu^k}(y^k, \eta^k)$, $(i = 1, \cdots, s)$, $\tilde{h}(x^k)$, $\tilde{g}(x^k)$ and $(\nu^k)^\top \tilde{g}(x^k)$ converge to $\nabla f(x^*, y^*)$, $\nabla L(x^*, y^*, \eta^*, \zeta^*)\theta^*$, $\nabla h(x^*, y^*)\rho^*$, $\nabla \tilde{g}(x^*)\nu^*$, $\nabla \tilde{h}(x^*)\xi^*$, $L(x^*, y^*, \eta^*, \zeta^*)\theta^*$, $A(x^*)y^* + b(x^*)$, $\varphi^i_{\mathrm{NR}}(y^*, \eta^*)$, $(i = 1, \cdots, s)$, $\tilde{h}(x^*)$, $\tilde{g}(x^*)$ and $(\nu^*)^\top \tilde{g}(x^*)$, respectively. Hence it is enough to prove that

$$\lim_{k\to\infty} \nabla \varphi^i_{\mu^k}((y^i)^k, (\eta^i)^k) \in \partial \varphi^i_{\mathrm{NR}}((y^i)^*, (\eta^i)^*) \tag{27}$$

holds true for every $i \in \{1, \cdots, s\}$. Moreover, from the relation between (12) and (23), it is sufficient for (27) to show that

$$\lim_{k\to\infty} \nabla P_{\mathcal{K}^{l_i}, \mu^k}((z^i)^k) \in \partial P_{\mathcal{K}^{l_i}}((z^i)^*), \tag{28}$$

where $(z^i)^k := (y^i)^k - (\eta^i)^k$.

For any k and $i \in \{1, \cdots, s\}$, define a_i^k, b_i^k, c_i^k and $(\lambda^i_j)^k$ $(j = 1, 2)$ by (25) and (26). When $\lim_{k\to\infty}(\lambda^i_j)^k \neq 0$, since $\mu^k \to 0$ as $k \to \infty$, we have $\frac{|(\lambda^i_j)^k|}{\mu^k} \to \infty$. Thus from the properties (10) of function $\hat{g}$, the following relations hold:

$$\begin{array}{lll} \text{if} \ \lim\limits_{k\to\infty} (\lambda^i_j)^k < 0 & \text{then} \ \lim\limits_{k\to\infty} \hat{g}(\frac{(\lambda^i_j)^k}{\mu^k}) = 0, & \lim\limits_{k\to\infty} \hat{g}'(\frac{(\lambda^i_j)^k}{\mu^k}) = 0, \\ \text{if} \ \lim\limits_{k\to\infty} (\lambda^i_j)^k > 0 & \text{then} \ \lim\limits_{k\to\infty} \left(\hat{g}(\frac{(\lambda^i_j)^k}{\mu^k}) - \frac{(\lambda^i_j)^k}{\mu^k}\right) = 0, & \lim\limits_{k\to\infty} \hat{g}'(\frac{(\lambda^i_j)^k}{\mu^k}) = 1. \end{array} \tag{29}$$

On the other hand, when $\lim_{k\to\infty}(\lambda^i_j)^k = 0$, if $\lim_{k\to\infty} \frac{(\lambda^i_j)^k}{\mu^k} = -\infty$ hold, then

$$\lim_{k\to\infty} \hat{g}(\frac{(\lambda^i_j)^k}{\mu^k}) = 0,$$

if $\lim_{k\to\infty} \frac{(\lambda^i_j)^k}{\mu^k} = \infty$ hold, then

$$\lim_{k\to\infty} \left(\hat{g}(\frac{(\lambda^i_j)^k}{\mu^k}) - \frac{(\lambda^i_j)^k}{\mu^k}\right) = 0,$$

and finally if $\limsup_{k\to\infty} \left|\frac{(\lambda^i_j)^k}{\mu^k}\right| < \infty$, then by the continuity of function $\hat{g}$, we have

$$\limsup_{k\to\infty} \left|\hat{g}(\frac{(\lambda^i_j)^k}{\mu^k})\right| < \infty.$$

As a result, in each case we have

$$\lim_{k\to\infty}\left|\mu^k\hat{g}(\frac{(\lambda_j^i)^k}{\mu^k})\right| = 0. \tag{30}$$

Using the above relations, we will show that a limit of $\nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k)$ belongs to the set $\partial P_{\mathcal{K}^{l_i}}((z^i)^*)$ represented by (13)–(16) in each of the following cases:

1. $(z_1^i)^* < -\|(z_2^i)^*\|$,
2. $(z_1^i)^* > \|(z_2^i)^*\|$,
3. $-\|(z_2^i)^*\| < (z_1^i)^* < \|(z_2^i)^*\|$,
4. $(z_1^i)^* = \|(z_2^i)^*\| \neq 0$,
5. $(z_1^i)^* = -\|(z_2^i)^*\| \neq 0$,
6. $(z^i)^* = 0$,

where $(z^i)^* := \lim_{k\to\infty}(z^i)^k$. Similarly, we define $a_i^* := \lim_{k\to\infty} a_i^k$, $b_i^* := \lim_{k\to\infty} b_i^k$, $c_i^* := \lim_{k\to\infty} c_i^k$ and $(\lambda_j^i)^* := \lim_{k\to\infty}(\lambda_j^i)^k$ $(j = 1, 2)$.

Case 1. $(z_1^i)^* < -\|(z_2^i)^*\|$

By taking an appropriate subsequence if necessary, it is sufficient to consider the following two cases: (i) $\|(z_2^i)^k\| = 0$ holds for all k, (ii) $\|(z_2^i)^k\| \neq 0$ holds for all k.

First, we consider the case that $\|(z_2^i)^k\| = 0$ holds for all k. Since $\lim_{k\to\infty}\frac{(z_1^i)^k}{\mu^k} = -\infty$ holds from $(z_1^i)^* < 0$, it follows from (24) and (10) that

$$\lim_{k\to\infty}\nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k) = \lim_{k\to\infty}\hat{g}'(\frac{(z_1^i)^k}{\mu^k})I_{l_i} = 0.$$

We next consider the case that $\|(z_2^i)^k\| \neq 0$ holds for all k. It follows from the assumption $(z_1^i)^* < -\|(z_2^i)^*\|$ along with (26) that $\lim_{k\to\infty}(\lambda_1^i)^k \leq \lim_{k\to\infty}(\lambda_2^i)^k < 0$. Therefore from the definition (25) of b_i^k and c_i^k and (29), we obtain

$$b_i^* = c_i^* = 0.$$

Since $(\lambda_1^i)^k < (\lambda_2^i)^k$ holds for all k, we have from the mean value theorem that there exist $\alpha_i^k \in \left[\frac{(\lambda_1^i)^k}{\mu^k}, \frac{(\lambda_2^i)^k}{\mu^k}\right]$ for all k such that

$$\hat{g}'(\alpha_i^k) = a_i^k = \frac{\hat{g}(\frac{(\lambda_2^i)^k}{\mu^k}) - \hat{g}(\frac{(\lambda_1^i)^k}{\mu^k})}{\frac{(\lambda_2^i)^k}{\mu^k} - \frac{(\lambda_1^i)^k}{\mu^k}}.$$

As shown above, we have $\lim_{k\to\infty}(\lambda_1^i)^k \leq \lim_{k\to\infty}(\lambda_2^i)^k < 0$, which implies $\lim_{k\to\infty}\alpha_i^k = -\infty$. Moreover, we have

$$a_i^* = \lim_{k\to\infty} a_i^k = \lim_{k\to\infty}\hat{g}'(a_i^k) = 0.$$

Hence we also obtain in this case

$$\lim_{k\to\infty} \nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k) = 0.$$

Therefore in each case we have

$$\lim_{k\to\infty} \nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k) = \nabla P_{\mathcal{K}^{l_i}}((z^i)^*) = 0.$$

Case 2. $(z_1^i)^* > \|(z_2^i)^*\|$

Similarly to Case 1, it is sufficient to consider the two cases: (i) $\|(z_2^i)^k\| = 0$ holds for all k, (ii) $\|(z_2^i)^k\| \neq 0$ holds for all k.

Assume that $\|(z_2^i)^k\| = 0$ holds for all k. Since $\lim_{k\to\infty} \frac{(z_1^i)^k}{\mu^k} = \infty$ by $(z_1^i)^* > 0$, it follows from (24) and (10) that

$$\lim_{k\to\infty} \nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k) = \lim_{k\to\infty} \hat{g}'(\frac{(z_1^i)^k}{\mu^k}) I_{l_i} = I_{l_i}.$$

We next consider the case that $\|(z_2^i)^k\| \neq 0$ for all k. Since we have $\lim_{k\to\infty}(\lambda_2^i)^k \geq \lim_{k\to\infty}(\lambda_1^i)^k > 0$ from the assumption $(z_1^i)^* > \|(z_2^i)^*\|$ and (26), it follows from the definition (25) of b_i^k and c_i^k and (29) that

$$b_i^* = 1, \quad c_i^* = 0.$$

Since $(\lambda_1^i)^k < (\lambda_2^i)^k$ for all k, from the mean value theorem, there exists $\alpha_i^k \in \left[\frac{(\lambda_1^i)^k}{\mu^k}, \frac{(\lambda_2^i)^k}{\mu^k}\right]$ for all k such that

$$\hat{g}'(\alpha_i^k) = a_i^k = \frac{\hat{g}(\frac{(\lambda_2^i)^k}{\mu^k}) - \hat{g}(\frac{(\lambda_1^i)^k}{\mu^k})}{\frac{(\lambda_2^i)^k}{\mu^k} - \frac{(\lambda_1^i)^k}{\mu^k}}.$$

As we have $\lim_{k\to\infty} \alpha_i^k = \infty$ from $\lim_{k\to\infty}(\lambda_2^i)^k \geq \lim_{k\to\infty}(\lambda_1^i)^k > 0$, it follows that

$$a_i^* = \lim_{k\to\infty} a_i^k = \lim_{k\to\infty} \hat{g}'(a_i^k) = 1.$$

Hence we obtain

$$\lim_{k\to\infty} \nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k) = I_{l_i}$$

from (24), which yields in any case that

$$\lim_{k\to\infty} \nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k) = \nabla P_{\mathcal{K}^{l_i}}((z^i)^*) = I_{l_i}.$$

Case 3. $-\|(z_2^i)^*\| < (z_1^i)^* < \|(z_2^i)^*\|$

Notice that in this case $\|(z_2^i)^*\|$ is not equal to zero.

Since $(\lambda_1^i)^* < 0 < (\lambda_2^i)^*$ holds by (26), a_i^k, b_i^k and c_i^k are calculated from their definitions (25) and (29) as

$$a_i^* = \lim_{k\to\infty} \frac{\mu^k\left(\hat{g}(\frac{(\lambda_2^i)^k}{\mu^k}) - \hat{g}(\frac{(\lambda_1^i)^k}{\mu^k})\right)}{(\lambda_2^i)^k - (\lambda_1^i)^k} = \lim_{k\to\infty} \frac{\mu^k\left(\frac{(\lambda_2^i)^k}{\mu^k} - \hat{g}(\frac{(\lambda_1^i)^k}{\mu^k})\right)}{2\|(z_2^i)^k\|} = \frac{1}{2}\left(1 + \frac{(z_1^i)^*}{\|(z_2^i)^*\|}\right),$$

$$b_i^* = \frac{1}{2}, \quad c_i^* = \frac{1}{2}.$$

Then by (13) and (24), we have

$$\lim_{k\to\infty} \nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k) = \nabla P_{\mathcal{K}^{l_i}}((z^i)^*).$$

Case 4. $(z_1^i)^* = \|(z_2^i)^*\| \neq 0$

We have $0 = (\lambda_1^i)^* < (\lambda_2^i)^*$ from (26). Since $\lim_{k\to\infty}(\lambda_1^i)^k = 0$, by using (29) and (30), we have

$$a_i^* = \lim_{k\to\infty} \frac{\mu^k\left(\hat{g}(\frac{(\lambda_2^i)^k}{\mu^k}) - \hat{g}(\frac{(\lambda_1^i)^k}{\mu^k})\right)}{(\lambda_2^i)^k - (\lambda_1^i)^k} = \lim_{k\to\infty} \frac{\mu^k\left(\frac{(\lambda_2^i)^k}{\mu^k} - \hat{g}(\frac{(\lambda_1^i)^k}{\mu^k})\right)}{(\lambda_2^i)^k} = \lim_{k\to\infty} \frac{(\lambda_2^i)^k}{(\lambda_2^i)^k} = 1.$$

Since $\left\{\hat{g}'(\frac{(\lambda_1^i)^k}{\mu^k})\right\} \subset (0,1)$ from the properties (10) of $\hat{g}'$, by taking an appropriate subsequence if necessary, we can assume $\lim_{k\to\infty} \hat{g}'(\frac{(\lambda_1^i)^k}{\mu^k}) = \gamma \in [0,1]$ without loss of generality. Then b_i^* and c_i^* are calculated as

$$b_i^* = \lim_{k\to\infty} \frac{1}{2}\left(\hat{g}'(\frac{(\lambda_2^i)^k}{\mu^k}) + \hat{g}'(\frac{(\lambda_1^i)^k}{\mu^k})\right) = \frac{1}{2}(1+\gamma),$$
$$c_i^* = \lim_{k\to\infty} \frac{1}{2}\left(\hat{g}'(\frac{(\lambda_2^i)^k}{\mu^k}) - \hat{g}'(\frac{(\lambda_1^i)^k}{\mu^k})\right) = \frac{1}{2}(1-\gamma).$$

Therefore from (24), we obtain

$$\begin{aligned}\lim_{k\to\infty} \nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k) &= \begin{bmatrix} \frac{1}{2}(1+\gamma) & \frac{1}{2}(1-\gamma)(v^i)^\top \\ \frac{1}{2}(1-\gamma)v^i & I_{l_i-1} + \left(\frac{1}{2}(1+\gamma) - 1\right) v^i(v^i)^\top \end{bmatrix} \\ &= \gamma I_{l_i} + \frac{1-\gamma}{2}\begin{bmatrix} 1 & (v^i)^\top \\ v^i & 2I_{l_i-1} - v^i(v^i)^\top \end{bmatrix},\end{aligned}$$

where without loss of generality we assume that $v^i = \lim_{k\to\infty} \frac{(z_2^i)^k}{\|(z_2^i)^k\|}$. Since $\gamma \in [0,1]$, this matrix belongs to the convex set $\partial P_{\mathcal{K}^{l_i}}((z^i)^*)$ given by (14), i.e.,

$$\lim_{k\to\infty} \nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k) \in \partial P_{\mathcal{K}^{l_i}}((z^i)^*).$$

Case 5. $(z_1^i)^* = -\|(z_2^i)^*\| \neq 0$

From (26), we have $(\lambda_1^i)^* < (\lambda_2^i)^* = 0$. Similarly to Case 4, it follows that

$$\begin{aligned}
a_i^* &= \lim_{k\to\infty} \frac{\mu^k \left(\hat{g}(\frac{(\lambda_2^i)^k}{\mu^k}) - \hat{g}(\frac{(\lambda_1^i)^k}{\mu^k}) \right)}{(\lambda_2^i)^k - (\lambda_1^i)^k} = \lim_{k\to\infty} \frac{\mu^k \hat{g}(\frac{(\lambda_2^i)^k}{\mu^k})}{-(\lambda_1^i)^k} = 0, \\
b_i^* &= \lim_{k\to\infty} \frac{1}{2} \left(\hat{g}'(\frac{(\lambda_2^i)^k}{\mu^k}) + \hat{g}'(\frac{(\lambda_1^i)^k)}{\mu^k}) \right) = \frac{1}{2}\gamma, \\
c_i^* &= \lim_{k\to\infty} \frac{1}{2} \left(\hat{g}'(\frac{(\lambda_2^i)^k}{\mu^k}) - \hat{g}'(\frac{(\lambda_1^i)^k)}{\mu^k}) \right) = \frac{1}{2}\gamma,
\end{aligned}$$

where we assume without loss of generality that $\gamma = \lim_{k\to\infty} \hat{g}'(\frac{(\lambda_2^i)^k}{\mu^k}) \in [0,1]$. Hence from (24), we have

$$\begin{aligned}
\lim_{k\to\infty} \nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k) &= \begin{bmatrix} \frac{\gamma}{2} & \frac{\gamma}{2}(v^i)^\top \\ \frac{\gamma}{2}v^i & \frac{\gamma}{2}v^i(v^i)^\top \end{bmatrix} \\
&= \frac{\gamma}{2} \begin{bmatrix} 1 & (v^i)^\top \\ v^i & v^i(v^i)^\top \end{bmatrix},
\end{aligned}$$

where without loss of generality we assume $v^i = \lim_{k\to\infty} \frac{(z_2^i)^k}{\|(z_2^i)^k\|}$. Since $\gamma \in [0,1]$, it follows from (15) that

$$\lim_{k\to\infty} \nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k) \in \partial P_{\mathcal{K}^{l_i}}((z^i)^*).$$

Case 6. $(z^i)^* = 0$

It is sufficient to consider the following two cases: (i) $\|(z_2^i)^k\| = 0$ holds for all k, (ii) $\|(z_2^i)^k\| \neq 0$ holds for all k.

First we consider the case that $\|(z_2^i)^k\| = 0$ holds for all k. From (24), we have

$$\lim_{k\to\infty} \nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k) = \lim_{k\to\infty} \hat{g}'(\frac{(z_1^i)^k}{\mu^k}) I_{l_i}.$$

Since $0 < \hat{g}'(\alpha) < 1$ holds by the properties (10) of function $\hat{g}$, we can assume without loss of generality that $\lim_{k\to\infty} \hat{g}'(\frac{(z_1^i)^k}{\mu^k}) = \gamma \in [0,1]$. Therefore we obtain

$$\lim_{k\to\infty} \nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k) = \gamma I_{l_i} \in \partial P_{\mathcal{K}^{l_i}}((z^i)^*).$$

We next consider the case that $\|(z_2^i)^k\| \neq 0$ holds for all k and $(z_2^i)^k$ converges to zero as $k \to \infty$. From (26) and the assumption that $(z^i)^* = 0$, we now have $(\lambda_1^i)^* = (\lambda_2^i)^* = 0$. Since $0 < \hat{g}'(\alpha) < 1$ from (10), we can suppose without loss of generality that the limit points of $\hat{g}'(\frac{(\lambda_j^i)^k}{\mu^k})$ $(j = 1,2)$ exist in $[0,1]$. Moreover, since it follows from the assumption $\|(z_2^i)^k\| \neq 0$ that $(\lambda_2^i)^k - (\lambda_1^i)^k = 2\|(z_2^i)^k\| > 0$ holds for all k, and since function $\hat{g}'$ is monotonically nondecreasing by the convexity of $\hat{g}$, the limit points of $\hat{g}'(\frac{(\lambda_j^i)^k}{\mu^k}),\ (j=1,2)$ can be represented as constants $\gamma_1, \gamma_2 \in [0,1]$ such that $\gamma_1 \le \gamma_2$, i.e.,

$$\lim_{k\to\infty} \hat{g}'(\frac{(\lambda_1^i)^k}{\mu^k}) = \gamma_1,$$
$$\lim_{k\to\infty} \hat{g}'(\frac{(\lambda_2^i)^k}{\mu^k}) = \gamma_2.$$

Consequently b_i^* and c_i^* are calculated as

$$b_i^* = \lim_{k\to\infty} \frac{1}{2}\left(\hat{g}'(\frac{(\lambda_2^i)^k}{\mu^k}) + \hat{g}'(\frac{(\lambda_1^i)^k}{\mu^k})\right) = \frac{1}{2}(\gamma_2+\gamma_1),$$
$$c_i^* = \lim_{k\to\infty} \frac{1}{2}\left(\hat{g}'(\frac{(\lambda_2^i)^k}{\mu^k}) - \hat{g}'(\frac{(\lambda_1^i)^k}{\mu^k})\right) = \frac{1}{2}(\gamma_2-\gamma_1).$$

Moreover by the mean value theorem, there exists $\alpha_i^k \in \left[\frac{(\lambda_1^i)^k}{\mu^k},\ \frac{(\lambda_2^i)^k}{\mu^k}\right]$ such that

$$\hat{g}'(\alpha_i^k) = a_i^k = \frac{\hat{g}(\frac{(\lambda_2^i)^k}{\mu^k}) - \hat{g}(\frac{(\lambda_1^i)^k}{\mu^k})}{\frac{(\lambda_2^i)^k}{\mu^k} - \frac{(\lambda_1^i)^k}{\mu^k}}.$$

Then from the monotonically nondecreasing property of $\hat{g}'$ again, we have

$$\hat{g}'(\frac{(\lambda_1^i)^k}{\mu^k}) \ \le\ \hat{g}'(\alpha_i^k) \ \le\ \hat{g}'(\frac{(\lambda_2^i)^k}{\mu^k}).$$

Since there exists $\gamma^k \in [-1,1]$ satisfying

$$a_i^k = \hat{g}'(\alpha_i^k) = \frac{1}{2}\left((1-\gamma^k)\hat{g}(\frac{(\lambda_1^i)^k}{\mu^k}) + (1+\gamma^k)\hat{g}(\frac{(\lambda_2^i)^k}{\mu^k})\right),$$

the limit of a_i^k can be represented as

$$a_i^* = \lim_{k\to\infty} a_i^k = \frac{1}{2}\left((1-\gamma^*)\gamma_1 + (1+\gamma^*)\gamma_2\right),$$

where we assume $\gamma^* := \lim_{k\to\infty}\gamma^k \in [-1,1]$ without loss of generality.

In addition, any accumulation point of the sequence $\{\frac{(z_2^i)^k}{\|(z_2^i)^k\|}\}$ can be represented by a vector $v^i \in \Re^{l_i-1}$ such that $\|v^i\| = 1$.

As a result, by (24), the limit of $\nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k)$ is obtained as

$$\begin{aligned}\lim_{k\to\infty} \nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k) &= \begin{bmatrix} b_i^* & c_i^*(v^i)^\top \\ c_i^* v^i & a_i^* I_{l_i-1} + (b_i^*-a_i^*)v^i(v^i)^\top \end{bmatrix} \\ &= \gamma_1 I_{l_i} + \frac{\gamma_2-\gamma_1}{2}\begin{bmatrix} 1 & (v^i)^\top \\ v^i & (\gamma^*+1)I_{l_i-1} - \gamma^* v^i (v^i)^\top \end{bmatrix}.\end{aligned}$$

Since $\gamma_1,\gamma_2 \in [0,1], \gamma_2-\gamma_1 \ge 0$ and $\gamma^* \in [-1,1]$ hold, the above expression implies that $\lim_{k\to\infty}\nabla P_{\mathcal{K}^{l_i},\mu^k}((z^i)^k)$ belongs to the convex set $\partial P_{\mathcal{K}^{l_i}}((z^i)^*)$ given by (16). □

5 Concluding remarks

In this paper, we have proposed a smoothing method for the MPSOCCC, which is an extension of the MPCC. We have proved the global convergence property of the proposed method to a stationary point of the MPSOCCC.

Some of the assumptions in this paper, however, might be somewhat restrictive. In particular, we have made assumptions A5 and A6 to simplify our discussions. From the theoretical viewpoint, it may be interesting to replace those assumptions by conditions directly based on the optimality conditions of SOCP or SOCCP studied in [1, 5, 10]. This remains as a further research subject.

This work has focused on the theoretical convergence property of a somewhat conceptual smoothing method. It is certainly important to develop a more concrete practical algorithm. In fact, practical smoothing methods have been applied to solve various problems such as variational inequalities [6] and MPECs [9, 12, 20]. We hope that this work contributes to future development of efficient algorithms for MPSOCCC.

Acknowledgments

The author wishes to express his sincerest thanks and appreciation to Professor Masao Fukushima for his kind guidance and direction in this study, invaluable discussions, constructive criticisms in the writing of the manuscripts, extreme patience, and encouragement throughout the course of this work. The author wishes to tender his acknowledgments to Associate Professor Nobuo Yamashita for his constructive comments and kind guidance. The author also wishes to express his thanks to Assistant Professor Shunsuke Hayashi for his invaluable advice, helpful guidance especially on theories of Second-Order Cones, and for lectures on ramen. The author greatly appreciates the help of all members of System Optimization Laboratory. Finally, but not the least, precious thanks are due to his family, for their strong support and affectionate encouragement throughout the course of this study.

References

[1] F. Alizadeh and D. Goldfarb, *Second-order cone programming*, Mathematical Programming, 95 (2003), pp. 3–51.

[2] J. F. Bard and J. T. Moore, *A branch and bound algorithm for the bilevel programming problem*, SIAM Journal on Scientific and Statistical Computing, 11 (1990), pp. 281–292.

[3] A. Ben-Tal and A. Nemirovski, *Robust convex optimization*, Mathematics of Operations Research, 23 (1998), pp. 769–805.

[4] W. F. Bialas and M. H. Karwan, *Two-level linear programming*, Management Science, 30 (1984), pp. 1004–1020.

[5] J. F. Bonnans and H. C. Ramírez, *Perturbation analysis of second-order cone programming problems*, Mathematical Programming, 104 (2005), pp. 205–227.

[6] B. Chen and P. T. Harker, *A continuation method for monotone variational inequalities*, Mathematical Programming, 69 (1995), pp. 237–253.

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

[8] X. Chen and M. Fukushima, *Expected residual minimization method for stochastic linear complementarity problems*, Mathematics of Operations Research, 30 (2005), pp. 1022–1038.

[9] F. Facchinei, H. Jiang, and L. Qi, *A smoothing method for mathematical programs with equilibrium constraints*, Mathematical Programming, 85 (1999), pp. 107–134.

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

[11] M. Fukushima, Z.-Q. Luo, and P. Tseng, *Smoothing functions for second-order-cone complementarity problems*, SIAM Journal on Optimization, 12 (2001), pp. 436–460.

[12] M. Fukushima and J.-S. Pang, *Convergence of a smoothing continuation method for mathematical programs with complementarity constraints*, SIAM Journal on Scientific and Statistical Computing, 13 (1992), pp. 1194–1217.

[13] G. Gürkan, A. Y. Özge, and S. M. Robinson, *Sample-path solution of stochastic variational inequalities*, Mathematical Programming, 84 (1999), pp. 313–333.

[14] P. Hansen, B. Jaumard, and G. Savard, *New branch-and-bound rules for linear bilevel programming*, SIAM Journal on Scientific and Statistical Computing, 13 (1992), pp. 1194–1217.

[15] S. Hayashi, *Studies on Second-Order Cone Complementarity Problems*, PhD thesis, Graduate School of Informatics, Kyoto University, 2004.

[16] C. Kanzow, I. Ferenczi, and M. Fukushima, *Semismooth methods for linear and nonlinear second-order cone programs*, Tech. Rep. 2006–005, Department of Applied Mathematics and Physics, Kyoto University, April 2006. revised January 2007.

[17] G.-H. Lin and M. Fukushima, *A class of stochastic mathematical programs with complementarity constraints: reformulations and algorithms*, Journal of Industrial and Management Optimization, 1 (2005), pp. 99–122.

[18] J. T. Moore and J. F. Bard, *The mixed integer linear bilevel programming problem*, Operations Research, 38 (1990), pp. 911–921.

[19] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[20] S. Scholtes, *Convergence properties of a regularization scheme for mathematical programs with complementarity constraints*, SIAM Journal on Optimization, 11 (2001), pp. 917–936.

[21] A. Shapiro, *Stochastic programming with equilibrium constraints*, Journal of Optimization Theory and Applications, 128 (2006), pp. 223–243.

[22] N. V. Thoai, Y. Yamamoto, and A. Yoshise, *Global optimization method for solving mathematical programs with linear complemetarity constraints*, Journal of Optimization Theory and Applications, 124 (2005), pp. 467–490.