Smoothing method for mathematical programs with symmetric cone complementarity constraints [*] [1]

Tao Yan [a†] [2] and Masao Fukushima [b‡3]

[a] *Department of Information and Computational Sciences, College of Science, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China.*

[b] *Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan.*

September 16, 2009

Abstract

In this paper, we consider the mathematical program with symmetric cone complementarity constraints (MPSCCC) in a general form. It includes the mathematical program with second-order-cone complementarity constraints (MPSOCCC) and the mathematical program with complementarity constraints (MPCC). We present a smoothing method which approximates the primal MPSCCC by means of the Chen-Mangasarian class of smoothing functions. We show that a sequence of stationary points of the approximate programs converges to a C(larke)-stationary point of the primal MPSCCC under suitable conditions.

Keywords: Mathematical program with symmetric cone complementarity constraints; Chen-Mangasarian smoothing functions; stationary point

AMS Subject Classification: 90C33; 65K05

1 Introduction

Let J be an n-dimensional real Euclidean space, $\mathcal{A} = (J, \langle\cdot,\cdot\rangle, \circ)$ be a Euclidean Jordan algebra, and K be a symmetric cone in $\mathcal{A}$. In this paper, we consider the following mathematical program with symmetric cone complementarity constraints (MPSCCC):

$$\begin{array}{ll} \min & f(x,y) \\ \text{s.t.} & x \in X, \\ & H(x,y) \in K, \quad G(x,y) \in K, \\ & \langle H(x,y), G(x,y)\rangle = 0, \end{array} \tag{1.1}$$

where $X \subseteq R^n$, and $f : R^n \times J \to R,\ H : R^n \times J \to J,\ G : R^n \times J \to J$ are continuously differentiable functions.

When the lower level problem of a bilevel optimization problem includes the symmetric cone constraints, we can formulate the problem as an MPSCCC under some suitable conditions. We consider the following example:

$$\begin{array}{ll} \min & \psi(x,y) \\ \text{s.t.} & x \in X, \\ & y \in Y(x), \end{array} \tag{1.2}$$

where $X \subseteq R^n$, $\psi : R^n \times J \to R$ is a continuously differentiable function, $Y(x) \subseteq J$ is the solution set of the following symmetric cone program with parameter x:

$$\begin{array}{ll} \min & \varphi(x,y) \\ \text{s.t.} & A(x)y + b(x) = 0, \\ & y \in K, \end{array} \tag{1.3}$$

where the function $\varphi : R^n \times J \to R$ is continuously differentiable, $A(x) : J \to R^m$ is a linear operator and $b(x) \in R^m$. If $\varphi(x, y) = \langle s, y\rangle$ with $s \in J$, problem (1.3) is a linear program over symmetric cone. Under some suitable conditions including the convexity of $\varphi(x, \cdot)$, problem (1.3) has an equivalent representation as follows:

$$\begin{gathered}
\nabla_y \varphi(x, y) - A(x)^*\nu - \xi = 0,\\
A(x)y + b(x) = 0,\\
y \in K,\ \xi \in K,\ \langle y, \xi\rangle = 0,
\end{gathered}$$

where $A(x)^*$ is the adjoint of $A(x)$, $\nu \in R^m$ and $\xi \in J$ are Lagrange multipliers. Then problem (1.2) can be reformulated as the MPSCCC:

$$\begin{array}{ll}
\min & \psi(x, y)\\
\text{s.t.} & x \in X,\\
& \nabla_y \varphi(x, y) - A(x)^*\nu - \xi = 0,\\
& A(x)y + b(x) = 0,\\
& y \in K, \xi \in K, \langle y, \xi\rangle = 0.
\end{array}$$

The mathematical program with equilibrium constraints (MPEC) is an optimization problem where the essential constraints are defined by a parametric variational inequality or complementarity system [17]. It is a generalization of a bilevel program. In addition, when the variational inequality in the lower level is rewritten as the KKT conditions, the MPEC is formulated as the mathematical program with complementarity constraints (MPCC), which is also an important subclass of the MPEC. Both problems are difficult because of their unusual nature of the constraints. Nevertheless, they play an important role in many practical problems, such as engineering design and economic modeling [17]. There have been many approaches for solving these problems, such as sequential quadratic programming approach [8, 12, 17], penalty function approach [11, 17, 21], implicit programming approach [17], and reformulation approach [6, 9, 14, 15, 22].

In practice, we often face optimization problems involving uncertain data. Hence, there have been many approaches that bring uncertainty into problem formulation. Stochastic optimization starts by assuming that the uncertainty has a probabilistic description. For example, in [13, 16], the authors studied stochastic mathematical programs with equilibrium constraints. Another more recent approach to optimization under uncertainty is robust optimization, in which the uncertainty model is not stochastic, but concerned with the worst-case scenario. Moreover, if the uncertain data set is ellipsoidal, then the robust optimization problem can be represented as a second-order cone programming (SOCP) problem [1]. Furthermore, if a bilevel program contains an SOCP as its lower level problem, and the SOCP can be replaced by its KKT conditions, then it yields the mathematical program with second-order cone complementarity constraints (MPSOCCC) [5].

It is well known that when $J = R^n$ with $\langle x, y\rangle = x^T y$, and $K = R^n_+$, the symmetric cone complementarity problem (SCCP) reduces to the nonlinear complementarity problem (NCP). Therefore the MPSCCC contains the MPCC as a special case. On the other hand, if K is the Cartesian product of second-order cones, i.e., $K = K^{n_1} \times \cdots \times K^{n_l}$, where $K^{n_i} = \{(z_1, z_2) \in R \times R^{n_i - 1} : \|z_2\| \le z_1\}$, with $l, n_1, \cdots, n_l \ge 1$ and $\sum_{i=1}^{l} n_i = n$, then the symmetric cone complementarity problem (SCCP) reduces to the second-order cone complementarity problem (SOCCP), and hence the MPSCCC contains the MPSOCCC as a special case. In consideration of these facts, the MPSCCC is a natural extension of the MPCC and the MPSOCCC.

The aim of this paper is to study a smoothing method for the MPSCCC. We particularly use the Chen-Mangasarian class of smoothing functions for a general form of MPSCCC. We will show that the smoothing approximations to the MPSCCC are well defined, and a sequence of stationary points of those problems converges to a C(larke)-stationary point of the original MPSCCC under some conditions.

The analysis of the method will be based on knowledge of the Euclidean Jordan algebra. So in Section 2, we review some basic concepts of the Euclidean Jordan algebra. In Section 3, we derive a smoothing approximation of the MPSCCC and analyze the properties of the feasible set. An algorithm based on smoothing approximation and its convergence analysis are presented in Section 4. Section 5 concludes the paper.

In what follows, for vectors $x, y \in J$, we write $x \succeq_K y$ (respectively, $x \succ_K y$) to mean $x - y \in K$ (respectively, $x - y \in \text{intK}$). Also, we write $B \succ C$ (respectively, $B \succeq C$) to mean $B - C$ being positive definite (respectively, positive semidefinite) for linear operators B and C from J into itself. For a differentiable mapping $G : J \to J$ and a vector $z \in J$, we denote by $G'(z)$ the Jocobian of G at z, and $\nabla G(z) = G'(z)^*$. Moreover, for a differentiable mapping H defined on $R^n \times J$, we denote by $H'_y(x, y) = \nabla_y H(x, y)^*$ the partial Jacobian with respect to the second argument y. If G is locally Lipschitz continuous, the Clarke generalized Jacobian [4] is defined by $\partial G(z) = \text{conv}\{\partial_B G(z)\}$, where $\partial_B G(z) = \{ \lim_{\substack{z_k \to z \\ z_k \in D_G}} G'(z_k)\}$ with D_G being the set of points at which G is differentiable, and conv denotes the convex hull. In parallel, we denote $\partial G(z)^* = \{V^* : V \in \partial G(z)\}$. The gradient of a real-valued function $g : J \to R$ at z is denoted by $\nabla g(z) = g'(z)^*$. For a given $z \in J$ and a set $S \subseteq J$, we denote $d(z, S) = \min\{\|z - y\| : y \in S\}$ for all $z \in J$, where $\| \cdot \|$ is the norm on J induced by the inner product $\langle \cdot, \cdot \rangle$, i.e., $\|x\| = \sqrt{\langle x, x \rangle}$.

2 Euclidean Jordan algebras and Löwner operators

In this section, we briefly describe the Euclidean Jordan algebra. For comprehensive details, see [7].

Let J be an n-dimensional vector space over the field of real numbers, and $(x, y) \mapsto x \circ y$ be a bilinear mapping. Then $(J, \circ)$ is called a Jordan algebra if

(1) $x \circ y = y \circ x$ for all $x, y \in J$,

(2) $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$ for all $x, y \in J$, where $x^2 = x \circ x$.

A Jordan algebra $(J, \circ)$ is said to be Euclidean and denoted as $\mathcal{A} = (J, \circ, \langle \cdot, \cdot \rangle)$, if an inner product $\langle \cdot, \cdot \rangle$ is defined and satisfies

(3) $\langle x \circ y, z \rangle = \langle x, y \circ x \rangle$ for all $x, y, z \in J$.

In general, a Jordan algebra is not associative, but it is power associative. A Jordan algebra has the identity element if there exists a unique element $e \in J$ such that $x \circ e = e \circ x = x$ for all $x \in J$.

For every $x \in J$, the Lyapunov transformation $L(x) : J \to J$ is defined by $L(x)y = x \circ y$ for all $y \in J$. It is a symmetric operator such that $\langle L(x)y, z \rangle = \langle y, L(x)z \rangle$ holds for all $y, z \in J$. Especially, $L(x)e = x$ and $L(x)x = x^2$ hold for all $x \in J$. By means of the Lyapunov transformation, the quadratic representation of $x \in J$ is defined as $Q(x) = 2L^2(x) - L(x^2)$.

Given a Euclidean Jordan algebra $\mathcal{A}$, the set $K = \{x^2 : x \in J\}$ is called the cone of squares of $\mathcal{A}$. Then K is a symmetric cone from Theorem III 2.1 in [7]. That is to say, K is a self-dual closed convex cone and its automorphism group acts transitively on its interior, and for each $x, y \in \mathrm{int}(K)$, the interior of K, there is a linear transformation $\mathcal{T}$ such that $\mathcal{T}(x) = y$ and $\mathcal{T}(K) = K$.

Before stating the spectral decomposition theorem, we recall some concepts related to Jordan algebra.

1. For any $x \in J$, let $m(x)$ be the degree of x, which is defined as $m(x) = \min\{k : \{e, x, x^2, \cdots, x^k\}$ are linearly dependent$\}$. Then the rank of $\mathcal{A}$ is defined by $r = \max\{m(x) : x \in J\}$, which obviously does not exceed the dimension of J.

2. A nonzero element $c \in J$ is called an idempotent if $c^2 = c$. An idempotent c is said to be primitive if it cannot be expressed as the sum of two other idempotents. A set of primitive idempotents $\{c_1, c_2, \cdots, c_k\}$ is called a Jordan frame if $c_i \circ c_j = 0$ for all $i \neq j$ and $\sum_{i=1}^{k} c_i = e$.

Theorem 2.1 (Spectral Decomposition Theorem [7, Theorem III.1.2]) *Let $(J, \circ, \langle \cdot, \cdot \rangle)$ be a Euclidean Jordan algebra with rank r. Then for any $x \in J$, there exist a Jordan frame $\{c_1, c_2, \cdots, c_r\}$ and real numbers $\lambda_1(x), \lambda_2(x), \cdots, \lambda_r(x)$ such that $x = \sum_{i=1}^{r} \lambda_i(x) c_i$. The numbers $\lambda_i(x), i = 1, 2, \cdots, r$, are called the eigenvalues of x, which are uniquely determined by x.*

Note that the Jordan frame $\{c_1, c_2, \cdots, c_r\}$ in the above theorem depends on x, but we omit the dependence in order to simplify the notation. An eigenvalue $\lambda_i(x)$ is continuous with respect to x. The trace of x is defined as $\mathrm{Tr}(x) = \sum_{i=1}^{r} \lambda_i(x)$ and the determinant of x is defined as $\mathrm{Det}(x) = \prod_{i=1}^{r} \lambda_i(x)$. Moreover, $x \in K$ (respectively, $x \in \mathrm{int}K$) if and only if $\lambda_i(x) \geq 0$ (respectively, $\lambda_i(x) > 0$) for all $i = 1, 2, \cdots, r$.

We now recall the definition and some properties of Löwner operator [24].

Definition 2.1 *Let $x = \sum_{i=1}^{r} \lambda_i(x) c_i$ and $\gamma : R \to R$. The Löwner operator $\Gamma : J \to J$ associated with γ is defined as*

$$\Gamma(x) = \sum_{i=1}^{r} \gamma(\lambda_i(x)) c_i = \gamma(\lambda_1(x)) c_1 + \gamma(\lambda_2(x)) c_2 + \cdots + \gamma(\lambda_r(x)) c_r. \tag{2.4}$$

Suppose that $\gamma : R \to R$ is differentiable at $\tau_i, i = 1, 2, \cdots, r$. Define the first divided difference $\gamma^{[1]}(\tau)$ of γ at $\tau = (\tau_1, \tau_2, \cdots, \tau_r)^T$ as the $r \times r$ symmetric matrix with the (i, j)th entry given by

$$[\gamma^{[1]}(\tau)]_{ij} = [\tau_i, \tau_j]_\gamma = \begin{cases} \frac{\gamma(\tau_i) - \gamma(\tau_j)}{\tau_i - \tau_j} & \text{if} \quad \tau_i \neq \tau_j \\ \gamma'(\tau_i) & \text{if} \quad \tau_i = \tau_j \end{cases} \quad i, j = 1, 2, \cdots, r. \tag{2.5}$$

Then we have the following theorem [19, 24].

Theorem 2.2 *Let $x = \sum_{i=1}^{r} \lambda_i(x) c_i$. Then the Löwner operator Γ associated with $\gamma : R \to R$ is (continuously) differentiable at x if and only if γ is (continuously) differentiable at $\lambda_i(x)$ for each $i \in \{1, 2, \cdots, r\}$. In this case, the Jacobian of Γ is given by*

$$\Gamma'(x) = 2 \sum_{\substack{i,j=1 \\ i \neq j}}^{r} [\lambda_i(x), \lambda_j(x)]_\gamma L(c_i) L(c_j) + \sum_{i=1}^{r} \gamma'(\lambda_i(x)) Q(c_i).$$

Moreover, $\Gamma'(x)$ is a linear and symmetric operator from J into itself.

3 Smoothing formulation and the feasible set

Let us introduce two new variables u and v which will be equal to $H(x, y)$ and $G(x, y)$, respectively, in (1.1). Corresponding to the symmetric cone K, let P_K denote the metric projection onto K. Then by Proposition 6 in [10], we have

$$v - P_K(v - u) = 0 \iff u \in K,\ v \in K,\ \langle u, v\rangle = 0.$$

Define the function $\Psi : R^n \times J \times J \times J \to J \times J \times J$ by

$$\Psi(x, y, u, v) = \begin{pmatrix} H(x, y) - u \\ G(x, y) - v \\ v - P_K(v - u) \end{pmatrix}. \tag{3.6}$$

Then MPSCCC (1.1) can be reformulated equivalently as

$$\begin{array}{ll} \min & f(x, y) \\ \text{s.t.} & x \in X, \\ & \Psi(x, y, u, v) = 0. \end{array} \tag{3.7}$$

We will deal with this equivalent problem in the rest of the paper. Throughout we make the following assumption on problem (3.7).

Assumption 1. (a) $X \subseteq R^n$ is compact. (b) The feasible set of problem (3.7), denoted by $\mathcal{F}$, is nonempty. (c) For each $x \in X$, there exists a solution (y, u, v) of $\Psi(x, y, u, v) = 0$.

We now introduce a smoothing approximation to problem (3.7). Let $p : R \to R$ be the function defined by $p(t) = (t)_+ = \max\{t, 0\}$. Then, a continuously differentiable convex function $\phi : R \to R_+$ satisfying

$$\lim_{t\to-\infty} \phi(t) = 0, \quad \lim_{t\to\infty} (\phi(t) - t) = 0 \ \text{ and } \ 0 < \phi'(t) < 1 \quad \text{for all} \ \ t \in R \tag{3.8}$$

is called the Chen-Mangasarian smoothing function of p [3, 25]. A special case of this smoothing function is the CHKS function $\phi(t) = \frac{\sqrt{t^2+4}+t}{2}$, which is proposed by Chen and Harker [2], Kanzow [18], and Smale [23]. Another special case is the sigmoid function $\phi(t) = \ln(e^t + 1)$ used in neural networks [3]. Using the Chen-Mangasarian smoothing function $\phi : R \to R_+$, we define a smoothing function of the projection operator $P_K : J \to J$ as the Löwner operator $\Phi : J \to J$ associated with ϕ. The class of such functions Φ is denoted by $\mathcal{CM}$. Now let $\mu \neq 0$ be a real parameter and define the function $\omega : J \times J \to J$ by

$$\omega_\mu(u, v) = v - |\mu|\Phi(\frac{v - u}{|\mu|}), \tag{3.9}$$

where $\Phi \in \mathcal{CM}$. Let $v - u = \sum_{i=1}^{r} \lambda_i(v - u)c_i$, where $\{c_1, c_2, \cdots, c_r\}$ is a Jordan frame at $v - u$. Then from (2.4), (3.9) can be rewritten as

$$\omega_\mu(u, v) = v - |\mu| \sum_{i=1}^{r} \phi(\frac{\lambda_i(v - u)}{|\mu|})c_i. \tag{3.10}$$

Recall [24] that the projection operator P_K can be written as

$$P_K(v - u) = \sum_{i=1}^{r} (\lambda_i(v - u))_+ c_i. \tag{3.11}$$

Since the function ϕ in (3.8) is continuously differentiable, Löwner operator Φ is also continuously differentiable by Theorem 2.2. Moreover, it is not difficult to deduce from (3.8) that

$$\lim_{\mu\to 0}|\mu|\phi(\frac{\lambda_i(v-u)}{|\mu|}) = (\lambda_i(v-u))_+. \tag{3.12}$$

It yields the following proposition immediately.

Proposition 3.1 *Let $\omega_\mu(u,v)$ be given by (3.9). Then for every $\mu \neq 0$, the function ω_μ is continuously differentiable on $J \times J$. Moreover, we have $\lim\limits_{\mu\to 0}\omega_\mu(u,v) = v - P_K(v-u)$.*

Let us introduce another function $\Psi_\mu : R^n \times J \times J \times J \to J \times J \times J$ defined by

$$\Psi_\mu(x,y,u,v) = \begin{pmatrix} H(x,y)-u \\ G(x,y)-v \\ v-|\mu|\Phi(\frac{v-u}{|\mu|}) \end{pmatrix}. \tag{3.13}$$

Combining the continuous differentiability of functions H and G with Proposition 3.1, we can deduce the continuous differentiability of the function $\Psi_\mu(x,y,u,v)$. Let us consider the following optimization problem:

$$\begin{array}{ll} \min & f(x,y) \\ \text{s.t.} & x \in X, \\ & \Psi_\mu(x,y,u,v) = 0. \end{array} \tag{3.14}$$

Then problem (3.14) is regarded as a smoothing approximation of problem (3.7).

For the sake of convenience, we define the function $\Lambda : R \times J \to J$ by

$$\Lambda(\mu,z) = \begin{cases} P_K(z) & \mu = 0, \\ |\mu|\Phi(\frac{z}{|\mu|}) & \mu \neq 0. \end{cases} \tag{3.15}$$

Then problem (3.7) and problem (3.14) can be unified as the following problem with $\mu \in R$ treated as a parameter:

$$\begin{array}{ll} \min & f(x,y) \\ \text{s.t.} & x \in X, \\ & F(\mu,x,y,u,v) = 0, \end{array} \tag{3.16}$$

where

$$F(\mu,x,y,u,v) = \begin{pmatrix} H(x,y)-u \\ G(x,y)-v \\ v-\Lambda(\mu,v-u) \end{pmatrix}. \tag{3.17}$$

Note that by (3.6) and (3.13), we have

$$F(\mu,x,y,u,v) = \begin{cases} \Psi(x,y,u,v) & \mu = 0, \\ \Psi_\mu(x,y,u,v) & \mu \neq 0. \end{cases}$$

We denote the feasible set of (3.16) by $\mathcal{F}_\mu$. Before discussing function Λ defined by (3.15), we give the following lemma.

Lemma 3.1 *Let $x = \sum\limits_{i=1}^{r}\lambda_i(x)c_i$ and α be a constant. If $\alpha e \succeq_K x \succeq_K \succeq 0$ holds, then we have $\|\alpha e\| \geq \|x\|$.*

Proof: Since $x = \sum_{i=1}^{r} \lambda_i(x) c_i$ and $e = \sum_{i=1}^{r} c_i$, we have $\alpha e - x = \sum_{i=1}^{r} (\alpha - \lambda_i(x)) c_i$. In addition, it follows from $\alpha e \succeq_K x \succeq_K \succeq 0$ that $\alpha e - x \in K$ and $x \in K$. This implies $\alpha - \lambda_i(x) \geq 0$ and $\lambda_i(x) \geq 0$ for all $i = 1, \cdots, r$. Then $\langle \alpha e - x, x \rangle = \sum_{i=1}^{r} (\alpha - \lambda_i(x)) \lambda_i(x) \|c_i\|^2 \geq 0$. Consequently, we obtain

$$\|\alpha e\|^2 - \|x\|^2 = \langle \alpha e, \alpha e \rangle - \langle x, x \rangle = \langle \alpha e - x, \alpha e - x \rangle + \langle \alpha e - x, x \rangle + \langle x, \alpha e - x \rangle \geq 0.$$

This completes the proof. $\square$

Lemma 3.2 *The function $\Lambda(\mu, z)$ defined by (3.15) is locally Lipschitz with respect to (μ, z).*

Proof: Let (μ_1, z_1) and (μ_2, z_2) be chosen arbitrarily. First, suppose $\mu_1 > \mu_2 \geq 0$. Note that

$$\|\Lambda(\mu_1, z_1) - \Lambda(\mu_2, z_2)\| \leq \|\Lambda(\mu_1, z_1) - \Lambda(\mu_1, z_2)\| + \|\Lambda(\mu_1, z_2) - \Lambda(\mu_2, z_2)\|. \tag{3.18}$$

Since $\Phi(\cdot)$ is continuously differentiable, we have

$$\|\Lambda(\mu_1, z_1) - \Lambda(\mu_1, z_2)\| \leq \sup_{0 \leq t \leq 1} \|\Phi'(\frac{z_1}{\mu_1} + t(\frac{z_1}{\mu_1} - \frac{z_2}{\mu_1}))\| \|z_1 - z_2\|. \tag{3.19}$$

By Lemma 3.1 and Proposition 4.3 in [19], we obtain

$$\|\Lambda(\mu_1, z_2) - \Lambda(\mu_2, z_2)\| \leq |\phi(0)| \|e\| |\mu_1 - \mu_2|. \tag{3.20}$$

It then follows from (3.18),(3.19) and (3.20) that

$$\|\Lambda(\mu_1, z_1) - \Lambda(\mu_2, z_2)\| \leq (M_1 + |\phi(0)| \|e\|) \|(\mu_1, z_1) - (\mu_2, z_2)\|,$$

where $M_1 = \sup_{0 \leq t \leq 1} \|\Phi'(\frac{z_1}{\mu_1} + t(\frac{z_1}{\mu_1} - \frac{z_2}{\mu_1}))\|$. By symmetry, the result also holds when $0 \geq \mu_1 > \mu_2$.

Next let $\mu_1 = \mu_2 = 0$. Since $P_K(z)$ is locally Lipschitz, there exists a constant $M_2 > 0$ depending on z_1 and z_2 such that $\|\Lambda(0, z_1) - \Lambda(0, z_2)\| \leq M_2 \|z_1 - z_2\|$.

Finally, let $\mu_1 > 0 > \mu_2$. By Lemma 3.1 and Proposition 4.3 in [19], we have

$$\begin{aligned}
& \|\Lambda(\mu_1, z_1) - \Lambda(\mu_2, z_2)\| \\
\leq\ & \|\Lambda(\mu_1, z_1) - \Lambda(0, z_1)\| + \|\Lambda(0, z_1) - \Lambda(0, z_2)\| + \|\Lambda(0, z_2) - \Lambda(\mu_2, z_2)\| \\
\leq\ & |\phi(0)| \|e\| |\mu_1| + M_2 \|z_1 - z_2\| + |\phi(0)| \|e\| |\mu_2| \\
\leq\ & (2|\phi(0)| \|e\| + M_2) \|(\mu_1, z_1) - (\mu_2, z_2)\|.
\end{aligned}$$

Hence, we complete the proof. $\square$

Lemma 3.3 *Assume that $G'_y(x, y)^* H'_y(x, y)$ is positive definite for any feasible point of (3.16). Then the partial (Clark generalized) Jacobian of F with respect to (y, u, v) is nonsingular.*

Proof: If $\mu \neq 0$, then $F(\mu, x, y, u, v) = \Psi_\mu(x, y, u, v)$. Since the function $\Psi_\mu(x, y, u, v)$ is continuously differentiable, its partial Jacobian with respect to (y, u, v) is given by

$$W_1 = \begin{pmatrix} H'_y(x, y) & -I & 0 \\ G'_y(x, y) & 0 & -I \\ 0 & \Phi'(\frac{v-u}{|\mu|}) & I - \Phi'(\frac{v-u}{|\mu|}) \end{pmatrix}. \tag{3.21}$$

In order to prove the nonsingularity of W_1, we will show that $W_1 q = 0$ implies $q = 0$, where $q = (q_1, q_2, q_3) \in J \times J \times J$. Using (3.21), we have

$$H'_y(x, y)q_1 - q_2 = 0, \tag{3.22}$$

$$G'_y(x, y)q_1 - q_3 = 0, \tag{3.23}$$

$$\Phi'(\frac{v-u}{|\mu|})q_2 + (I - \Phi'(\frac{v-u}{|\mu|}))q_3 = 0. \tag{3.24}$$

By (3.22) and (3.23), we have

$$\langle q_3, q_2 \rangle = \langle q_1, G'_y(x, y)^* H'_y(x, y)q_1 \rangle. \tag{3.25}$$

From Proposition 4.4 in [19] and Theorem 2.2, $\Phi'(\frac{v-u}{|\mu|})$ is symmetric positive definite and $0 \prec \Phi'(\frac{v-u}{|\mu|}) \prec I$. Hence the matrix $(\Phi'(\frac{v-u}{|\mu|}))^{-1}(I - \Phi'(\frac{v-u}{|\mu|}))$ is positive definite. Moreover, by (3.24), we have

$$\langle q_3, q_2 \rangle = -\langle q_3, (\Phi'(\frac{v-u}{|\mu|}))^{-1}(I - \Phi'(\frac{v-u}{|\mu|}))q_3 \rangle. \tag{3.26}$$

Combining (3.25) and (3.26) yields

$$\langle q_1, G'_y(x, y)^* H'_y(x, y)q_1 \rangle = -\langle q_3, (\Phi'(\frac{v-u}{|\mu|}))^{-1}(I - \Phi'(\frac{v-u}{|\mu|}))q_3 \rangle. \tag{3.27}$$

Then it is not difficult to deduce from (3.27) that $q_1 = 0$ and $q_3 = 0$, which in turn implies $q_2 = 0$. Consequently, W_1 is nonsingular.

If $\mu = 0$, then $F(\mu, x, y, u, v) = \Psi(x, y, u, v)$. An element of the Clarke generalized Jacobian of Ψ with respect to (y, u, v) is given by the matrix

$$W_2 = \begin{pmatrix} H'_y(x, y) & -I & 0 \\ G'_y(x, y) & 0 & -I \\ 0 & V & I - V \end{pmatrix},$$

where $V \in \partial P_K(v - u)$ with the property that V is symmetric positive semidefinite and $0 \preceq V \preceq I$ [19].

For $V \in \partial P_K(v - u)$, there exists an orthogonal matrix Q and a diagonal matrix D with diagonal element $a_i \in [0, 1]$ such that $V = QDQ^*$. Let

$$\tilde{W_2} = \begin{pmatrix} I & & \\ & I & \\ & & Q^* \end{pmatrix} W_2 \begin{pmatrix} I & & \\ & I & \\ & & Q \end{pmatrix} = \begin{pmatrix} H'_y(x, y) & -I & 0 \\ G'_y(x, y) & 0 & -I \\ 0 & D & I - D \end{pmatrix}.$$

Then it is easy to show that W_2 is nonsingular if and only if $\tilde{W_2}$ is nonsingular. Assume that $\tilde{W_2} q = 0$ for some vector $q = (q_1, q_2, q_3) \in J \times J \times J$. Then

$$H'_y(x, y)q_1 - q_2 = 0, \tag{3.28}$$

$$G'_y(x, y)q_1 - q_3 = 0, \tag{3.29}$$

$$Dq_2 + (I - D)q_3 = 0. \tag{3.30}$$

By (3.28) and (3.29), we have

$$\langle q_3, q_2 \rangle = \langle q_1, G'_y(x, y)^* H'_y(x, y)q_1 \rangle. \tag{3.31}$$

Recall that the elements a_i of the diagonal matrix D are in $[0, 1]$. From (3.30), we have $a_i q_{2i} + (1 - a_i) q_{3i} = 0$ for each i. Then it is easy to see that $q_{3i} q_{2i} \leq 0$ for all i, which implies $\langle q_3, q_2 \rangle \leq 0$. Since $G'_y(x, y)^* H'_y(x, y)$ is positive definite, it follows from (3.31) that $q_1 = 0$. This in turn implies $q_2 = 0$ from (3.28) and $q_3 = 0$ from (3.29). Thus $\tilde{W}_2$ is nonsingular, and hence W_2 is nonsingular as desired. □

For a function $M : J \times J \to J$, let $\pi_y \partial M(x, y)$ denote the set of all matrices P such that, for some matrix U, the matrix (U, P) belongs to $\partial M(x, y)$. In order to apply the implicit function theorem [4] in the following result, $\pi_{(y,u,v)} \partial \Psi(x, y, u, v)$ needs to consist of nonsingular matrices. However, this property seems to be difficult to verify. On the other hand, the nonsingularity of $\partial_{(y,u,v)} \Psi(x, y, u, v)$ can be verified under some conditions, as shown in Lemma 3.3. Hence, we make the following assumption.

Assumption 2. $\pi_{(y,u,v)} \partial \Psi(x, y, u, v) \subseteq \partial_{(y,u,v)} \Psi(x, y, u, v)$ for any $(x, y, u, v) \in \mathcal{F}$.

Based on the foregoing arguments, we can obtain the result on the following feasible set of (3.16).

Lemma 3.4 *Assume that $G'_y(x, y)^* H'_y(x, y)$ is positive definite for any feasible point of (3.7). Let $\bar{x} \in X$ be given, and let Assumption 1(c) and Assumption 2 hold. Then for every (μ, x) in a neighborhood of $(0, \bar{x})$, there exists a unique vector (y, u, v) such that $F(\mu, x, y, u, v) = 0$.*

Proof: First notice that $\Lambda(\mu, v - u)$ is locally Lipschitz with respect to (μ, u, v) by Lemma 3.2. Hence, it follows from the continuous differentiability of H and G that $F(\mu, x, y, u, v)$ is locally Lipschitz with respect to (μ, x, y, u, v). In addition, since $F(0, x, y, u, v) = \Psi(x, y, u, v)$, by Assumption 1(c), there exists a vector $(\bar{y}, \bar{u}, \bar{v})$ such that $F(0, \bar{x}, \bar{y}, \bar{u}, \bar{v}) = 0$. Moreover, from Assumption 2, we have $\pi_{(y,u,v)} \partial F(0, \bar{x}, \bar{y}, \bar{u}, \bar{v}) \subseteq \partial_{(y,u,v)} F(0, \bar{x}, \bar{y}, \bar{u}, \bar{v})$, which implies the nonsingularity of $\pi_{(y,u,v)} \partial F(0, \bar{x}, \bar{y}, \bar{u}, \bar{v})$ by Lemma 3.3. Therefore, the result follows from Clarke's implicit function theorem [4]. □

Theorem 3.3 *Assume that $G'_y(x, y)^* H'_y(x, y)$ is positive definite for any feasible point of (3.7). If Assumption 1 and Assumption 2 hold, then the feasible set $\mathcal{F}_\mu$ of (3.16) is nonempty and uniformly compact for $\mu \in [-\bar{\mu}, \bar{\mu}]$, where $\bar{\mu} > 0$ is a sufficiently small constant.*

Proof: The nonemptiness follows from Lemma 3.4. We now prove that the feasible set is uniformly compact. By Assumption 1(a), the x-component of any feasible solution of (3.16) belongs to the compact set X. Then it is sufficient to prove that the (y, u, v)-component is contained in a bounded set. Suppose this is not true, namely, there exist sequences $\{(x_k, y_k, u_k, v_k)\}$ and $\{\mu_k\}$ with $\mu_k \in [-\bar{\mu}, \bar{\mu}]$ such that $(x_k, y_k, u_k, v_k) \in \mathcal{F}_{\mu_k}$ for all k, and $\lim_{k \to \infty} \|(y_k, u_k, v_k)\| = \infty$. Without loss of generality, we assume that $\mu_k \to \tilde{\mu} \in [-\bar{\mu}, \bar{\mu}]$ and $x_k \to \tilde{x}$. Then it follows from Lemma 3.4 that there exists a unique vector $(\tilde{y}, \tilde{u}, \tilde{v})$ such that $F(\tilde{\mu}, \tilde{x}, \tilde{y}, \tilde{u}, \tilde{v}) = 0$. On the other hand, by Lemma 3.3, the partial (Clark generalized) Jacobian of $F(\mu, x, y, u, v)$ with respect to (y, u, v) is nonsingular at $(\tilde{\mu}, \tilde{x}, \tilde{y}, \tilde{u}, \tilde{v})$. Combining this property with the regularity of the continuously differentiable function Ψ_μ when $\mu \neq 0$ and Assumption 2 when $\mu = 0$, we obtain that every element of $\pi_{(y,u,v)} \partial F(\mu, x, y, u, v)$ is nonsingular at $(\tilde{\mu}, \tilde{x}, \tilde{y}, \tilde{u}, \tilde{v})$. Therefore, by the implicit function theorem [4, 20], there exists a neighborhood Ω of $(\tilde{x}, \tilde{\mu})$ and a continuous function $(y(\cdot, \cdot), u(\cdot, \cdot), v(\cdot, \cdot)) : \Omega \to J \times J \times J$ such that, for each $(x, \mu) \in \Omega$,

$$F(\mu, x, y(x, \mu), u(x, \mu), v(x, \mu)) = 0.$$

Moreover, by the uniqueness, we have $(y_k, u_k, v_k) = (y(x_k, \mu_k), u(x_k, \mu_k), v(x_k, \mu_k))$. Since functions $y(\cdot,\cdot), u(\cdot,\cdot), v(\cdot,\cdot)$ are continuous and $(\mu_k, x_k) \to (\tilde{\mu}, \tilde{x})$, we must have that (y_k, u_k, v_k) converges to $(\tilde{y}, \tilde{u}, \tilde{v})$. This is a contradiction. Hence, $\mathcal{F}_\mu$ is uniformly bounded for $\mu \in [-\overline{\mu}, \overline{\mu}]$. Moreover, by the fact that $F(\mu, x, y, u, v)$ is locally Lipschitz, $\mathcal{F}_\mu$ is closed for each $\mu \in [-\overline{\mu}, \overline{\mu}]$. The proof is complete. □

4 Algorithm and convergence analysis

In this section, we first present an algorithm that sequentially solves the smooth approximation problems (3.14), and then consider the limiting behavior of a sequence of points generated by the algorithm.

We define functions $\mathcal{H} : R^n \times J \times J \times J \to J, \mathcal{G} : R^n \times J \times J \times J \to J, \mathcal{P} : R^n \times J \times J \times J \to J$, and $\Xi_\mu : R^n \times J \times J \times J \to J$ for any $\mu \neq 0$ by

$$\mathcal{H}(x, y, u, v) = H(x, y) - u, \qquad \mathcal{G}(x, y, u, v) = G(x, y) - v,$$

$$\mathcal{P}(x, y, u, v) = v - P_K(v - u), \quad \Xi_\mu(x, y, u, v) = v - |\mu|\Phi(\tfrac{v-u}{|\mu|}).$$

Suppose that the set X is given by $X = \{x : g(x) \leq 0, h(x) = 0\}$, where $g : R^n \to R^m$ and $h : R^n \to R^l$ are continuously differentiable, and the following assumption holds.

Assumption 3. The set $\{g_i'(x), h_j'(x) : i \in \mathcal{I}_g,\ j = 1, \cdots, l\}$, where $\mathcal{I}_g = \{i : g_i(x) = 0\}$, is linearly independent for any $x \in X$.

We try to solve problem (3.7) by solving a sequence of smooth approximation problems (3.14) with $\mu \to 0$. However, for the latter problems, we can only find their stationary points by a standard optimization method. Taking into account this fact, we propose the following algorithm.

Algorithm 1

Step 1. Choose an initial point $w_1 = (x_1, y_1, u_1, v_1) \in R^n \times J \times J \times J$, and a nonzero sequence $\{\mu_k\}$ such that $\mu_k \to 0$. Set the stopping tolerance $\varepsilon \geq 0$ and $k = 1$.

Step 2. Find a stationary point $w_k = (x_k, y_k, u_k, v_k)$ of problem (3.14) with $\mu = \mu_k$.

Step 3. If $d(w_k, \mathcal{F}) \leq \varepsilon$, then stop; otherwise set $k = k + 1$, and go to step 2.

Let us consider the limiting behavior of a generated sequence $\{w_k\}$. Unlike problem (3.14) with $\mu \neq 0$, problem (3.7) is a nonsmooth optimization problem. Hence, the ordinary KKT conditions cannot be used directly. Here, we consider the Fritz-John conditions shown in [4]. Namely, a necessary condition for a point (x, y, u, v) to be a local optimal solution of problem (3.7) is that there exists a nonzero multiplier vector

$(\eta, \theta, \delta, \sigma, \zeta, \tau) \in R \times J \times J \times J \times R^m \times R^l$ with $\eta \geq 0, \zeta \geq 0$ such that

$$0 \in \quad \eta \begin{pmatrix} \nabla_x f(x,y) \\ \nabla_y f(x,y) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \nabla_x H(x,y) \\ \nabla_y H(x,y) \\ -I \\ 0 \end{pmatrix} \theta + \begin{pmatrix} \nabla_x G(x,y) \\ \nabla_y G(x,y) \\ 0 \\ -I \end{pmatrix} \delta$$

$$+ \partial \mathcal{P}(x,y,u,v)^* \sigma + \begin{pmatrix} \nabla g(x) \\ 0 \\ 0 \\ 0 \end{pmatrix} \zeta + \begin{pmatrix} \nabla h(x) \\ 0 \\ 0 \\ 0 \end{pmatrix} \tau, \tag{4.32}$$

$$\begin{aligned} \Psi(x,y,u,v) = 0, \\ h(x) = 0, \\ g(x) \leq 0, g(x)^T \zeta = 0. \end{aligned}$$

In particular, if $\eta \neq 0$, in which case we can assume $\eta = 1$ without loss of generality, then conditions (4.32) become the KKT conditions for problem (3.7). In the following theorem, we show that η is actually not zero under some conditions.

Theorem 4.4 *Assume that $w = (x, y, u, v)$ is a Fritz-John stationary point of problem (3.7), i.e, it together with some $(\eta, \theta, \delta, \sigma, \zeta, \tau)$ satisfies conditions (4.32). If $G_y'(x,y)^* H_y'(x,y)$ is positive definite for any feasible point of (3.7) and Assumption 3 holds, then $\eta \neq 0$ and w is a KKT point of problem (3.7).*

Proof: We prove the theorem by contradiction. Suppose that $\eta = 0$. Then (4.32) implies

$$0 = \begin{pmatrix} \nabla_x H(x,y) & \nabla_x G(x,y) & 0 & \nabla g(x) & \nabla h(x) \\ \nabla_y H(x,y) & \nabla_y G(x,y) & 0 & 0 & 0 \\ -I & 0 & V & 0 & 0 \\ 0 & -I & I - V & 0 & 0 \end{pmatrix} \begin{pmatrix} \theta \\ \delta \\ \sigma \\ \zeta \\ \tau \end{pmatrix} \tag{4.33}$$

for some symmetric matrix $V \in \partial P_K(v - u)$ and $0 \neq (\theta, \delta, \sigma, \zeta, \tau)$ with $\zeta \geq 0$ and $g(x)^T \zeta = 0$. From the last three rows of equation (4.33), we obtain

$$0 = \begin{pmatrix} \nabla_y H(x,y) & \nabla_y G(x,y) & 0 \\ -I & 0 & V \\ 0 & -I & I - V \end{pmatrix} \begin{pmatrix} \theta \\ \delta \\ \sigma \end{pmatrix}, \tag{4.34}$$

which yields $(\theta, \delta, \sigma) = 0$ by Lemma 3.3. Again by the first row of equation (4.33) and $g(x)^T \zeta = 0$, we have $(\zeta, \tau) = 0$ from Assumption 3. This contradicts $(\eta, \theta, \delta, \sigma, \zeta, \tau) \neq 0$, and hence we must have $\eta \neq 0$. □

Lemma 4.5 *Let $\{z_k\} \subseteq J$ be convergent to $\bar{z}$ and the nonzero sequence $\{\mu_k\}$ be convergent to 0. Then we have $\lim\limits_{k \to \infty} \Phi'(\frac{z_k}{|\mu_k|}) \in \partial P_K(\bar{z})$.*

Proof: Let $z_k = \sum\limits_{i=1}^{r} \lambda_i(z_k) c_i(z_k)$. Then from Theorem 2.2,

$$\Phi'(\frac{z_k}{|\mu_k|}) = 2 \sum_{\substack{i,j=1 \\ i \neq j}}^{r} a_{ij}(z_k) L(c_i(z_k)) L(c_j(z_k)) + \sum_{i=1}^{r} a_{ii}(z_k) Q(c_i(z_k)),$$

where $a_{ij}(z_k)$ are given by (2.5), i.e.,

$$a_{ij}(z_k) = \begin{cases} \frac{\phi(\lambda_i(z_k)/|\mu_k|) - \phi(\lambda_j(z_k)/|\mu_k|)}{\lambda_i(z_k)/|\mu_k| - \lambda_j(z_k)/|\mu_k|} & \text{if} \quad \lambda_i(z_k) \neq \lambda_j(z_k), \\ \phi'(\lambda_i(z_k)/|\mu_k|) & \text{if} \quad \lambda_i(z_k) = \lambda_j(z_k). \end{cases}$$

By the continuity of $\lambda_i(\cdot)$ and the property (3.12), $\{a_{ij}(z_k)\}$ are convergent to

$$b_{ij} = \begin{cases} \frac{\max\{\lambda_i(\bar{z}),0\} - \max\{\lambda_j(\bar{z}),0\}}{\lambda_i(\bar{z}) - \lambda_j(\bar{z})} & \text{if} \quad \lambda_i(\bar{z}) \neq \lambda_j(\bar{z}), \\ \phi^-(\lambda_i(\bar{z})) & \text{if} \quad \lambda_i(\bar{z}) = \lambda_j(\bar{z}), \end{cases}$$

where

$$\phi^-(\lambda_i(\bar{z})) = \begin{cases} 0 & \text{if} \quad \lambda_i(\bar{z}) < 0, \\ \phi'(0) \in (0,1) & \text{if} \quad \lambda_i(\bar{z}) = 0, \\ 1 & \text{if} \quad \lambda_i(\bar{z}) > 0. \end{cases}$$

Moreover, since $L(\cdot)$ and $Q(\cdot)$ are continuous, we have

$$\lim_{k\to\infty} \Phi'(\frac{z_k}{|\mu_k|}) = 2 \sum_{\substack{i,j=1 \\ i\neq j}}^{r} b_{ij} L(\bar{c}_i) L(\bar{c}_j) + \sum_{i=1}^{r} b_{ii} Q(\bar{c}_i),$$

where $\{\bar{c}_1, \bar{c}_2, \cdots, \bar{c}_r\}$ is a Jordan frame at $\bar{z}$. By Theorem 2.4 in [19], we then obtain $\lim\limits_{k\to\infty} \Phi'(\frac{z_k}{|\mu_k|}) \in \partial P_K(\bar{z})$. □

Theorem 4.5 *Let the sequence $\{w_k\} \subseteq R^n \times J \times J \times J$ be generated by Algorithm 1, where $w_k = (x_k, y_k, u_k, v_k)$. Assume that $\bar{\mu} = \sup\limits_k |\mu_k|$ is sufficiently small and $G'_y(x,y)^* H'_y(x,y)$ is positive definite for any feasible point of (3.7). Moreover, let Assumption 1, Assumption 2 and Assumption 3 hold. Then we have the following statements:*

(a) The sequence $\{w_k\}$ is bounded.

(b) If $\varepsilon = 0$ and $\bar{w} = (\bar{x}, \bar{y}, \bar{u}, \bar{v})$ is an accumulation point of the sequence $\{w_k\}$, then $\bar{w}$ is a KKT point of problem (3.7), i.e., $(\bar{x}, \bar{y})$ is a C-stationary point of the MPSCCC.

Proof: (a) Since $\bar{\mu} = \sup\limits_k |\mu_k|$ is sufficiently small, the sequence $\{w_k\}$ is contained in a compact set from Theorem 3.3.

(b) Suppose the stopping tolerance ε is zero and the algorithm generates an infinite sequence $\{w_k\}$. Since w_k is a KKT point of problem (3.14), there exist multipliers $(\theta_k, \delta_k, \sigma_k, \zeta_k, \tau_k) \in J \times J \times J \times R^m \times R^l$ with $\zeta_k \geq 0$, $g(x_k)^T \zeta_k = 0$ such that

$$\begin{aligned} 0 = & \begin{pmatrix} \nabla_x f(x_k, y_k) \\ \nabla_y f(x_k, y_k) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \nabla_x H(x_k, y_k) \\ \nabla_y H(x_k, y_k) \\ -I \\ 0 \end{pmatrix} \theta_k + \begin{pmatrix} \nabla_x G(x_k, y_k) \\ \nabla_y G(x_k, y_k) \\ 0 \\ -I \end{pmatrix} \delta_k \\ & + \begin{pmatrix} \nabla_x \Xi_{\mu_k}(x_k, y_k, u_k, v_k) \\ \nabla_y \Xi_{\mu_k}(x_k, y_k, u_k, v_k) \\ \nabla_u \Xi_{\mu_k}(x_k, y_k, u_k, v_k) \\ \nabla_v \Xi_{\mu_k}(x_k, y_k, u_k, v_k) \end{pmatrix} \sigma_k + \begin{pmatrix} \nabla g(x_k) \\ 0 \\ 0 \\ 0 \end{pmatrix} \zeta_k + \begin{pmatrix} \nabla h(x_k) \\ 0 \\ 0 \\ 0 \end{pmatrix} \tau_k. \end{aligned} \tag{4.35}$$

By the continuity, we have $\nabla f(x_k, y_k) \to \nabla f(\bar{x}, \bar{y}), \nabla g(x_k) \to \nabla g(\bar{x}), \nabla h(x_k) \to \nabla h(\bar{x})$, $\nabla \mathcal{H}(x_k, y_k, u_k, v_k) \to \nabla \mathcal{H}(\bar{x}, \bar{y}, \bar{u}, \bar{v})$ and $\nabla \mathcal{G}(x_k, y_k, u_k, v_k) \to \nabla \mathcal{G}(\bar{x}, \bar{y}, \bar{u}, \bar{v})$. Moreover, $\Phi'(\cdot)$ is symmetric by Theorem 2.2, we have

$$\begin{array}{ll} \nabla_x \Xi_{\mu_k}(x_k, y_k, u_k, v_k) = 0, & \nabla_y \Xi_{\mu_k}(x_k, y_k, u_k, v_k) = 0 \\ \nabla_u \Xi_{\mu_k}(x_k, y_k, u_k, v_k) = \Phi'(\frac{v_k - u_k}{|\mu_k|}), & \nabla_v \Xi_{\mu_k}(x_k, y_k, u_k, v_k) = I - \Phi'(\frac{v_k - u_k}{|\mu_k|}). \end{array}$$

It follows from Lemma 4.5 that $\nabla \Xi_{\mu_k}(x_k, y_k, u_k, v_k)$ converges to an element in $\partial \mathcal{P}(\bar{x}, \bar{y}, \bar{u}, \bar{v})^*$. Next, we prove that $\{(\theta_k, \delta_k, \sigma_k, \zeta_k, \tau_k)\}$ is bounded. Suppose it is not true. We divide (4.35) by $\|(\theta_k, \delta_k, \sigma_k, \zeta_k, \tau_k)\|$ and let $(\hat{\theta}_k, \hat{\delta}_k, \hat{\sigma}_k, \hat{\zeta}_k, \hat{\tau}_k)$ be the normalized vector of multipliers. Then, by taking a subsequence if necessary, we may assume that the latter sequence converges to $(\tilde{\theta}, \tilde{\delta}, \tilde{\sigma}, \tilde{\zeta}, \tilde{\tau})$ and we obtain

$$\begin{aligned} 0 \in & \begin{pmatrix} \nabla_x H(\bar{x}, \bar{y}) \\ \nabla_y H(\bar{x}, \bar{y}) \\ -I \\ 0 \end{pmatrix} \tilde{\theta} + \begin{pmatrix} \nabla_x G(\bar{x}, \bar{y}) \\ \nabla_y G(\bar{x}, \bar{y}) \\ 0 \\ -I \end{pmatrix} \tilde{\delta} \\ & + \partial \mathcal{P}(\bar{x}, \bar{y}, \bar{u}, \bar{v})^* \tilde{\sigma} + \begin{pmatrix} \nabla g(\bar{x}) \\ 0 \\ 0 \\ 0 \end{pmatrix} \tilde{\zeta} + \begin{pmatrix} \nabla h(\bar{x}) \\ 0 \\ 0 \\ 0 \end{pmatrix} \tilde{\tau}, \end{aligned}$$

which means that $\bar{w}$ is a Fritz-John point with zero multiplier for the objective function $f(x, y)$. However, it contradicts Theorem 4.4. So $\{(\theta_k, \delta_k, \sigma_k, \zeta_k, \tau_k)\}$ is bounded. Without loss of generality, we assume $\{(\theta_k, \delta_k, \sigma_k, \zeta_k, \tau_k)\}$ converges to $(\bar{\theta}, \bar{\delta}, \bar{\sigma}, \bar{\zeta}, \bar{\tau})$ and $\{w_k\}$ converges to $\bar{w}$. Then passing to the limit in (4.35), we have $\bar{\zeta} \geq 0, g(\bar{x})^T \bar{\zeta} = 0$ and

$$\begin{aligned} 0 \in & \begin{pmatrix} \nabla_x f(\bar{x}, \bar{y}) \\ \nabla_y f(\bar{x}, \bar{y}) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \nabla_x H(\bar{x}, \bar{y}) \\ \nabla_y H(\bar{x}, \bar{y}) \\ -I \\ 0 \end{pmatrix} \bar{\theta} + \begin{pmatrix} \nabla_x G(\bar{x}, \bar{y}) \\ \nabla_y G(\bar{x}, \bar{y}) \\ 0 \\ -I \end{pmatrix} \bar{\delta} \\ & + \partial \mathcal{P}(\bar{x}, \bar{y}, \bar{u}, \bar{v})^* \bar{\sigma} + \begin{pmatrix} \nabla g(\bar{x}) \\ 0 \\ 0 \\ 0 \end{pmatrix} \bar{\zeta} + \begin{pmatrix} \nabla h(\bar{x}) \\ 0 \\ 0 \\ 0 \end{pmatrix} \bar{\tau}. \end{aligned}$$

Moreover, since $\mu_k \to 0$, by the continuity of the constraint function of problem (3.14) and Proposition 3.1, we have $\bar{w} \in \mathcal{F}$. Consequently, $\bar{w}$ is a KKT point of problem (3.7), i.e., $(\bar{x}, \bar{y})$ is a C-stationary point of the MPSCCC. □

5 Conclusion

In this paper, we have presented a smoothing method for mathematical program with symmetric cone complementarity constraints (MPSCCC). We have also discussed the convergence of the method and shown that an accumulation point of the sequence generated by the smoothing method is a C(larke)-stationary point of the original MPSCCC under some conditions.

References

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

[2] B. Chen and P. Harker, *A non-interior-point continuation method for linear complementarity problems,* SIAM J. Matrix Anal. Appl., 14 (1993), pp. 1168-1190.

[3] C. Chen and O. Mangasarian, *Smoothing methods for convex inequalities and linear complementarity problems*, Math. Program., 71 (1995), pp. 51-69.

[4] F. H. Clarke, *Optimization and Nonsmooth Analysis,* Wiley, New York, 1983.

[5] T. Ejiri, *A smoothing method for mathematical programmming programs with second-order-cone complementarity constraints,* Master thesis, Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto, Japan. 2007.

[6] F. Facchinei, H. Jiang and L. Qi, *A smoothing method for mathematical programs with equilibrium constrains,* Math. Program., 85 (1999), pp. 107-134.

[7] J. Faraut and A. Korányi, *Analysis on Symmetric Cones,* Oxford University Press, Oxford, UK, 1994.

[8] M. Fukushima, Z.Q. Luo and J.S. Pang, *A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints,* Comput. Optim. Appl., 10 (1998), pp. 5-34.

[9] M. Fukushima and J.S. Pang, *Convergence of a smoothing continuation method for mathematical problem with complementarity constraints,* Ill Posed Variational Problems and Regularization Techniques, Lecture Notes in Economics and Mathematical Systems, Vol 477, M. Théra and R. Tichatschke (eds.), Springer-Verlag, Berlin/Heidelberg, 1999, pp. 105-116.

[10] M. S. Gowda, R. Sznajder, and J. Tao, *Some P-properties for linear transformations on Euclidean Jordan algebras,* Linear Algebra Appl., 393 (2004), pp. 203-232.

[11] X. Hu and D. Ralph, *Convergence of a penalty method for mathematical programming with complementarity constraints,* J. Optim. Theory Appl., 123 (2004), pp. 365-390.

[12] H. Jiang and D. Ralph, *Smooth SQP methods for mathematical programs with nonlinear complementarity constraints,* SIAM J. Optim., 10 (2000), pp. 779-808.

[13] G.H. Lin, X. Chen and M. Fukushima, *Solving stochastic mathematical programs with equilibrium constraints via approximation and smoothing implicit programming with penalization,* Math. Program., 116 (2009), pp. 343-368.

[14] G.H. Lin and M. Fukushima, *A new relaxation method for mathematical programs with complementarity constraints,* J. Optim. Theory Appl., 118 (2003), pp. 81-116.

[15] G.H. Lin and M. Fukushima, *A modified relaxation scheme for mathematical programs with complementarity constrains,* Annals Oper. Res., 133 (2005), pp. 63-84.

[16] G.H. Lin, H. Xu and M. Fukushima, *Monte Carlo and quasi-Monte Carlo sampling methods for a class of stochastic mathematical programs with equilibrium constraints,* Math. Meth. Oper. Res., 67 (2008), pp. 423-441.

[17] Z.Q. Luo, J-S. Pang and D. Ralph, *Mathematical Programs with Equilibrium Constrains*, Cambridge University Press, Cambridge, 1996.

[18] C. Kanzow, *Some noninterior continuation methods for linear complementarity problems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 851-868.

[19] L. Kong, J. Sun, and N. Xiu, *A regularized smoothing Newton method for symmetric cone complementarity problems,* SIAM J. Optim., 19 (2008) pp. 1028-1047.

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

[21] S. Scholtes and M. Stöhr, *Exact penalization of mathematical programs with complementarity constraints*, SIAM J. Control Optim., 37 (1999), pp. 617-652.

[22] S. Scholtes, *Convergence properties of a regularization scheme for mathematical programs with complementarity constraints,* SIAM J. Optim., 11 (2001), pp. 918-936.

[23] S. Smale, *Algorithms for solving equations*, in Proceedings of the International Congress of Mathematicians, American Mathematical Society, Providence, RI, 1987, pp. 172-195.

[24] D. Sun and J. Sun, *Löwner's operator and spectral functions in Euclidean Jordan algebras,* Math. Oper. Res., 33 (2008), pp. 421-445.

[25] P. Tseng, *Analysis of a non-interior continuation method based on Chen-Mangasarian smoothing functions for complementarity problems,* Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, M. Fukushima and L. Qi (eds.), Kluwer, Dordrecht, 1999, pp. 381-404.