Math. Meth. Oper. Res. (2006) 64: 255–269
DOI 10.1007/s00186-006-0076-2

ORIGINAL ARTICLE

Hongxia Yin · Jianzhong Zhang

Global convergence of a smooth approximation method for mathematical programs with complementarity constraints

Received: 29 May 2004 / Accepted: 4 September 2005 / Published online: 14 July 2006

Abstract A new smoothing approach was given for solving the mathematical programs with complementarity constraints (MPCC) by using the aggregation technique. As the smoothing parameter tends to zero, if the KKT point sequence generated from the smoothed problems satisfies the second-order necessary condition, then any accumulation point of the sequence is a B-stationary point of MPCC if the linear independence constraint qualification (LICQ) and the upper level strict complementarity (ULSC) condition hold at the limit point. The ULSC condition is weaker than the lower level strict complementarity (LLSC) condition generally used in the literatures. Moreover, the method can be easily extended to the mathematical programs with general vertical complementarity constraints.

Keywords MPCC · LICQ · ULSC condition · B-stationary point · Global convergence

H. Yin (✉)
Chinese Academy of Sciences Research Center on Data Technology and Knowledge Economy, Department of Mathematics, Graduate University of the Chinese Academy of Sciences, Beijing 100049, People's Republic of China
E-mail: hxyin@gucas.ac.cn

J. Zhang
Department of System Engineering and Engineering Management, The Chinese University of Hong Kong, Hong Kong
E-mail: mazhang@cityu.edu.hk

1 Introduction

We consider the mathematical programs with complementarity constraints (MPCC):

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & G_1(z) \geq 0,\ G_2(z) \geq 0, \\ & G_1(z)^T G_2(z) = 0, \\ & g(z) \leq 0, \\ & h(z) = 0, \end{array} \tag{1}$$

where $z \in \mathcal{R}^n$, $f : \mathcal{R}^n \to \mathcal{R}$, $G_j : \mathcal{R}^n \to \mathcal{R}^m (j = 1, 2)$, $g : \mathcal{R}^n \to \mathcal{R}^p$, and $h : \mathcal{R}^n \to \mathcal{R}^q$ are twice continuously differentiable functions. As an important subclass of mathematical programs with equilibrium constrains (MPEC), MPCC is very useful for the study of bilevel programming problems and a large number of engineering design problems, see Luo et al. (1996a) and the references therein. In the past two decades, optimal conditions and algorithms for MPEC were deeply studied by some authors, and various stationarity concepts were introduced, see Fukushima and Pang (1999), Luo et al. (1996a,b), Outrata et al. (1998), Pang and Fukushima (1999), Scheel and Scholtes (2000), Scholtes (2001), Scholtes and Stühr (1999, 2001), Ye and Ye (1997) and Ye (1999, 2000) for example. Among them, Bouligand stationarity conditions (B-stationary) introduced by Luo et al. (1996a) provides the strongest first-order optimality conditions for a local minimizer of MPCC. Scheel and Scholtes (2000) made an excellent clarification on these concepts and elucidated their connections. On the other hand, many efforts have also been made to produce the robust algorithms for MPEC. As we known, most of the methods are based on the non-smooth approach or the nonlinear program relaxation approach (Fukushima and Tseng 2002; Liu and Sun 2004; Luo et al. 1996a,b; Outrata et al. 1998; Outrata and Zowe 1995; Raghunathan and Biegler 2005; Scholtes and Stühr 1999). Recently, the smoothing continuous function methods, which have been widely used for solving complementarity and variational inequality problems, were also proposed for solving MPCC with some special complementarity constraints, see Faccinei et al. (1999), Fukushima and Pang (1999), Fukushima et al. (1998) and Jiang and Ralph (2000). However, some of these methods require the nondegeneracy (that is, the lower level strict complementarity) condition at a limit point so as to guarantee the methods converge to a B-stationary point, and some can only converge to a point satisfying some weak stationarity conditions such as the Clarke stationarity condition. Fukushima and Pang (1999) gave a smoothing continuation method for solving MPCC by using the perturbed Fischer–Burmeister function. They proved that the second-order stationary point sequence of the smoothed problems converges to a B-stationary point of the MPCC as the smoothing parameter tends to zero. The basic assumptions they need are the linear independence constraint qualification (LICQ) and the asymptotic weak nondegeneracy at a limit point. The later condition means that for some $i \in \{1, \ldots, m\}$, if both G_{1i} and G_{2i} are equal to zero at the limit point, then $G_{1i}(z)$ and $G_{2i}(z)$ tend to zero in the same order of magnitude. Jiang and Ralph (2000) proposed two smooth SQP methods for MPCC. Global convergence of the methods depend on the lower level strict complementarity condition amongst some other conditions, such as the LICQ or the Mangasarian-Fromovitz constrained qualification (MFCQ).

In this paper, we propose a new smoothing approximation method for MPCC by using the aggregation technique. Aggregation function is a well known smoothing function for max-type functions. Let $w : \mathcal{R}^n \to \mathcal{R}$, $w(x) = \max\{w_1(x), w_2(x), \dots, w_m(x)\}$, where $w_i, i = 1, \dots, m$, are continuously differentiable functions, it is clear that $w(\cdot)$ is continuous in $\mathcal{R}^n$ but not differentiable everywhere. For any $t > 0$, *the aggregation function* of $w(x)$, noted as $w(t, x) : \mathcal{R}^{n+1} \to \mathcal{R}$, is defined by

$$w(t, x) := t \ln \sum_{i=1}^{m} \exp(w_i(x)/t). \tag{2}$$

Function (2), viewed as the exponential penalty function, has been studied and employed as the multiplier method for nonlinear programming (Goldstein 1997; Kachiyan 1996; Tseng and Bertsekas 1993). According to Goldstein (1997), Chang (1980) first introduced the function (2). Independently, Li (1991, 1992) studied (2) and named it as the aggregation function. Recently, some authors investigated the differentiable properties of the function, and used it to propose smoothing methods for generalized (extended) linear complementarity problems and some nonlinear complementarity problems (Peng and Lin 1999; Qi and Liao 1999; Qi et al. 2000).

Noticing that $G_1(z) \geq 0$, $G_2(z) \geq 0$ and $G_1(z)^T G_2(z) = 0$ if and only if $G(z) = \min\{G_1(z), G_2(z)\} = 0$, we define

$$G(t, z) = \begin{pmatrix} -t \ln(\exp(-G_{11}(z)/t) + \exp(-G_{21}(z)/t)) \\ -t \ln(\exp(-G_{12}(z)/t) + \exp(-G_{22}(z))/t)) \\ \vdots \\ -t \ln(\exp(-G_{1m}(z)/t) + \exp(-G_{2m}(z)/t)) \end{pmatrix} \tag{3}$$

for $t > 0$. It is easy to see that $G(t, z)$ is continuously differentiable with respect to t for all $t > 0$, and $\lim_{t \downarrow 0} G(t, z) = G(z)$. As such, it is natural to define $G(0, z) = G(z)$. Then we produce the following parametric nonlinear programming problems $(\mathbf{P_t})$:

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & G(t, z) = 0, \\ & g(z) \leq 0, \\ & h(z) = 0. \end{array} \tag{4}$$

For $t > 0$, (4) is said to be *the smoothing approximation* of MPCC (1). Obviously, the solution of $(\mathbf{P_t})$ tends to the solution of MPCC (1) as the smoothing parameter t tends to 0. Under the LICQ and a condition called upper level strict complementarity (ULSC)(see section 2), we prove that any limit point of the sequence generated from solving the smooth approximation problem (4) (as $t \to 0$) must be a B-stationary point of MPCC (1), if the sequence satisfies the second-order necessary conditions. It is shown that ULSC condition is weaker than the generally used lower level strict complementarity (LLSC) condition. Therefore, the convergence condition of this paper is weaker than that of Jiang and Ralph (2000), and in some sense, weaker than that of Fukushima and Pang (1999).

The paper is organized as follows. In section 2 we recall some concepts and propositions about the MPCC and investigate the properties of the aggregation function $G(t, z)$ for establishing the relationship of a B-stationary point of MPCC (1) with the solution of $(\mathbf{P_t})$. Section 3 is devoted to the convergence analysis of our smooth approximation method. Further remarks on the extension of the method will be drawn in section 4.

Some notations: denote $\mathcal{R}_+$ as the set of all nonnegative real numbers, $\mathcal{R}_{++}$ the set of all positive real numbers. Let $\mathcal{I}$ and $\mathcal{J}$ be the subindex set of the inequality constraints $g(z) \le 0$ and the subindex set of the equality constraints $h(z) = 0$, respectively. $|\mathcal{I}|$ is the number of elements in the index set $\mathcal{I}$. For a vector $v \in R^m$ and an index set $\alpha \subseteq \{1, 2, \ldots, m\}$, v_α denotes the subvector of v with components v_i, $i \in \alpha$.

2 Preliminaries

A feasible point $\bar{z}$ of MPCC is a *B-stationary point* if $\nabla f(\bar{z})^T d \ge 0$ for every d satisfying

$$\begin{array}{l} \min\{\nabla G_{ji}(\bar{z})^T d \mid j : G_j(\bar{z}) = 0, \quad j = 1, 2\} = 0, \ \ i = 1, 2, \ldots, m. \\ \nabla g_r(\bar{z})^T d \le 0, \ \ r \in \{r : g_r(\bar{z}) = 0, r \in I\}, \\ \nabla h(\bar{z})^T d = 0. \end{array} \tag{5}$$

Clearly, any local optimal solution of MPCC (1) is a B-stationary point. Furthermore, we define the following index sets at $\bar{z}$:

$$\begin{array}{l} \bar{\alpha} = \{i \mid G_{1i}(\bar{z}) = 0 < G_{2i}(\bar{z})\}, \\ \bar{\beta} = \{i \mid G_{1i}(\bar{z}) = 0 = G_{2i}(\bar{z})\}, \\ \bar{\gamma} = \{i \mid G_{1i}(\bar{z}) > 0 = G_{2i}(\bar{z})\}, \\ \bar{\mathcal{I}} = \{r \mid g_r(\bar{z}) = 0\}. \end{array} \tag{6}$$

If the vectors $\{\nabla G_{1i}(\bar{z}) | i \in \bar{\alpha} \cup \bar{\beta}\} \cup \{\nabla G_{2i}(\bar{z}) | i \in \bar{\beta} \cup \bar{\gamma}\} \cup \{\nabla g_r(\bar{z}) | r \in \bar{\mathcal{I}}\} \cup \{\nabla h_l(\bar{z}) | l \in \mathcal{J}\}$ are linearly independent, we say the LICQ holds at $\bar{z}$. In particular, we say that $\bar{z}$ satisfies the lower level strict complementarity (LLSC) condition if $\bar{\beta} = \emptyset$, i.e., $G_{1i}(\bar{z}) + G_{2i}(\bar{z}) > 0$ for all $i = 1, \ldots, m$. In this case $\bar{z}$ is also said to be nondegenerate.

Proposition 2.1 Fukushima and Pang (1999) *and* Luo et al. (1996a) *assume $\bar{z}$ is a feasible point of MPCC, and the LICQ holds at $\bar{z}$. If there exist vectors $\bar{\nu} \in \mathcal{R}^m$, $\bar{\omega} \in \mathcal{R}^m$, $\bar{\lambda} \in \mathcal{R}^{\bar{\mathcal{I}}}$, $\bar{\mu} \in \mathcal{R}^{J}$ such that $\bar{z}$ satisfies the conditions:*

$$\begin{array}{l} \nabla f(\bar{z}) - \sum_{i \in \bar{\alpha}} \bar{\nu}_i \nabla G_{1i}(\bar{z}) - \sum_{i \in \bar{\beta}} \bar{\nu}_i \nabla G_{1i}(\bar{z}) - \sum_{i \in \bar{\beta}} \bar{\omega}_i \nabla G_{2i}(\bar{z}) \\ \quad - \sum_{i \in \bar{\gamma}} \bar{\omega}_i \nabla G_{2i}(\bar{z}) + \sum_{r \in \bar{\mathcal{I}}} \bar{\lambda}_r \nabla g_r(\bar{z}) + \sum_{l \in \mathcal{J}} \bar{\mu}_l \nabla h_l(\bar{z}) = 0, \\ G_{1\bar{\alpha}}(\bar{z}) = 0, \quad G_{2\bar{\gamma}}(\bar{z}) = 0, \quad h(\bar{z}) = 0, \\ \bar{\nu}_{\bar{\beta}} \ge 0, \quad G_{1\bar{\beta}}(\bar{z}) \ge 0, \quad \bar{\nu}_{\bar{\beta}}^T G_{1\bar{\beta}}(\bar{z}) = 0, \\ \bar{\omega}_{\bar{\beta}} \ge 0, \quad G_{2\bar{\beta}}(\bar{z}) \ge 0, \quad \bar{\omega}_{\bar{\beta}}^T G_{2\bar{\beta}}(\bar{z}) = 0, \\ \bar{\lambda} \ge 0, \quad g(\bar{z}) \le 0, \quad \bar{\lambda}^T g(\bar{z}) = 0, \end{array} \tag{7}$$

then $\bar{z}$ is a B-stationary point of MPCC.

LICQ is a generic constraint qualification for MPCC, many optimal theories and algorithms for MPCC are established under this condition. The point satisfying system (7) is called the Karush–Kuhn–Tucher (KKT) point of MPCC (1). Moreover, according to Scheel and Scholtes (2000), a point $\bar{z}$ is said to satisfy the upper level strict complementarity condition (ULSC) if ν_i and ω_i, the multipliers correspondence to G_{1i} and G_{2i} respectively, satisfy $\nu_i \omega_i \neq 0$ for all $i \in \bar{\beta}$. It is well known that a point $\bar{z}$ satisfies the lower level strict complementarity condition (LLSC) if $G_{1i} + G_{2i} > 0$ hold for all $i \in \{1, 2, \ldots, m\}$. We can see from the following example that the ULSC condition is considerably weaker than the LLSC condition, and in practice, it may make more sense than the later one.

Example 2.1 Scheel and Scholtes (2000)

$$\begin{array}{ll} \min & (z_1 - t)^2 + (z_2 - t)^2 \\ \text{s.t.} & z_1 z_2 = 0, \\ & z_1 \geq 0, \\ & z_2 \geq 0, \end{array} \tag{8}$$

where $z_1, z_2 \in \mathcal{R}$, $t \leq 0$ is a parameter.

Obviously, (0, 0) is the optimal solution of (8). The LLSC condition is violated at (0, 0) for every $t \leq 0$, while the ULSC condition is only violated at (0, 0) for $t = 0$. From the KKT system for problem (8), it is not difficult to see that, for $(z_1, z_2) = (0, 0)$ and $t < 0$, we may take the multipliers ν and ω in (7) as $\nu = \omega = -2t$. Hence, $\nu\omega = 4t^2$, which means that the ULSC condition holds for any $t < 0$.

In the next part of this section, we investigate the properties of aggregation function $G(t, z)$ in (3). From Proposition 3.2 in Qi and Liao (1999) and by simple calculation, we can get the following properties of $G(t, z)$ with respect to t and z, respectively.

Lemma 2.1 (i) *$G(t, z)$ is a locally Lipschitz continuous function on $\mathcal{R}_+ \times \mathcal{R}^n$.*

(ii) *$(G(t, z))_i$ is a continuously differentiable decreasing and convex function with respect to t ($t > 0$), and for any $i = 1, 2, \ldots, m$,*

$$(G(z))_i - t \ln 2 \leq (G(t, z))_i \leq (G(z))_i, \tag{9}$$

$$\nabla_t (G(t, z))_i = -\ln\left(\frac{\exp(-G_{1i}(z))}{t}\right) + \exp\left(\frac{-G_{2i}(z)}{t}\right) - \frac{1}{t}[\eta_{1i}(t, z) G_{1i}(z) + \eta_{2i}(t, z) G_{2i}(z)], \tag{10}$$

where

$$\eta_{1i}(t, z) = \frac{\exp(-G_{1i}(z)/t)}{\exp(-G_{1i}(z)/t) + \exp(-G_{2i}(z)/t)} \in (0, 1),$$
$$\eta_{2i}(t, z) = \frac{\exp(-G_{2i}(z)/t)}{\exp(-G_{1i}(z)/t) + \exp(-G_{2i}(z)/t)} \in (0, 1), \tag{11}$$

with $\eta_{1i}(t, z) + \eta_{2i}(t, z) = 1$.

(iii) *If $G_1(z)$ and $G_2(z)$ are twice continuously differentiable functions, then $G(t,z)$ is at least twice continuously differentiable on $\mathcal{R}_{++} \times \mathcal{R}^n$, and for all $i = 1, 2, \ldots, m$,*

$$\nabla_z (G(t,z))_i = \eta_{1i}(t,z) \nabla G_{1i}(z) + \eta_{2i}(t,z) \nabla G_{2i}(z),$$

$$\begin{aligned} \nabla_z^2 (G(t,z))_i = {} & \eta_{1i}(t,z) \nabla^2 G_{1i}(z) + \eta_{2i}(t,z) \nabla^2 G_{2i}(z) \\ & -\frac{1}{t}\eta_{1i}(t,z)\eta_{2i}(t,z)\Big[\nabla G_{1i}(z)(\nabla G_{1i}(z))^T \\ & + \nabla G_{2i}(z)(\nabla G_{2i}(z))^T\Big] \\ & +\frac{1}{t}\eta_{1i}(t,z)\eta_{2i}(t,z)\Big[\nabla G_{1i}(z)(\nabla G_{2i}(z))^T \\ & + \nabla G_{2i}(z)(\nabla G_{1i}(z))^T\Big], \end{aligned} \tag{12}$$

and

$$\lim_{t\to 0} \nabla_z (G(t,z))_i = \frac{1}{|B_i(z)|} \sum_{j\in B_i(z)} \nabla G_{ji}(z), \tag{13}$$

where $B_i(z) = \{j | G_{ji}(z) = G_i(z),\ \ j = 1, 2\}$, $|B_i(z)|$ is the element number of the index set $B_i(z)$.

Since $G(0,z) = G(z)$, we have the following lemma which is useful for establishing our convergence result in Section 3.

Lemma 2.2 *Let $\bar{z} \in \mathcal{R}^n$, $z_t \to \bar{z}$ as $t \to 0$. Then*

(i)

$$B_i(z_t) \subseteq B_i(\bar{z}) \quad (i = 1, 2, \ldots, m),$$

(ii)

$$\mathcal{I}(z_t) \subseteq \bar{\mathcal{I}}, \quad \mathcal{I}(z_t) = \{r |\ g_r(z_t) = 0\}$$

holds for all t sufficiently small.

Proof (i) We assume, by contradiction, that there exists a sequence $\{t^k\} \in \mathcal{R}_{++}$ with $t^k \to 0$ as $k \to \infty$ and an index $j(k) \in B_i(z_{t^k})$ but $j(k) \notin B_i(\bar{z})$ for some $i \in \{1, 2, \ldots, m\}$. From the definition of $B_i(z)$, there is at most two elements in each $B_i(z)$. We suppose that $j(k) = j$ for all k without loss of generality. Then we have $G_{ji}(z_{t^k}) = G_i(z_{t^k})$. Noticing the continuity of G_{ji} and G_i, we obtain that $G_{ji}(\bar{z}) = G_i(\bar{z})$ from the fact that $z_{t^k} \to \bar{z}$ as $t^k \to 0$. Therefore, $j \in B_i(\bar{z})$. This contradiction establishes (i).

(ii) Similar to the proof of (i). □

We now recall the Clarke generalized gradient/Jacobian.

Definition 2.1 *Let the operator $G : \mathcal{R}^n \to \mathcal{R}^m$ be Lipschitz continuous near $z_0 \in \mathcal{R}^n$ and let σ_G denote the set of points at which G fails to be differentiable. The generalized Jacobian of G at z_0 is*

$$\partial G(z_0) = \text{conv}\,\{\lim_{k\to\infty} \nabla G(z^k) | z^k \to z_0,\ z^k \notin \sigma_G\}. \tag{14}$$

For $m = 1$, $\partial G(z_0)$ is termed the generalized gradient of G at z_0.

The following theorem is important for deriving the global convergence of our smoothing method for MPCC.

Theorem 2.1 *Assume $G_1(z)$, $G_2(z)$ are continuously differentiable functions. Let $\{t^k\}$ be a sequence of positive scalars tending to zero, $\{z^k\}$ be a sequence tending to $\bar{z}$. Then for all $i = 1, 2, \ldots, m$,*

$$\lim_{\substack{z^k \to \bar{z} \\ t^k \to 0}} \operatorname{dist}\{\nabla_z (G(t^k, z^k))_i,\ \partial (G(\bar{z}))_i\} = 0. \tag{15}$$

Proof From the definition of $G(\cdot)$ and $\partial G(\cdot)$, we have that, for all $i = 1, 2, \ldots, m$,

$$\begin{aligned} \partial (G(\bar{z}))_i &= \operatorname{conv}\{\lim_{k\to\infty} \nabla (G(z^k))_i | z^k \to \bar{z}, G_{1i}(z^k) \neq G_{2i}(z^k)\} \\ &= \big\{\bar{\eta}_{1i} \nabla G_{1i}(\bar{z}) + \bar{\eta}_{2i} \nabla G_{2i}(\bar{z}) |\ \bar{\eta}_{1i}, \bar{\eta}_{2i} \in [0, 1], \\ &\qquad \text{and} \quad \bar{\eta}_{1i} + \bar{\eta}_{2i} = 1\big\}. \end{aligned} \tag{16}$$

Since $t^k > 0$, $G(t, z)$ is continuously differentiable at (t^k, z^k), and

$$\nabla_z (G(t^k, z^k))_i = \eta_{1i}(t^k, z^k) \nabla G_{1i}(z^k) + \eta_{2i}(t^k, z^k) \nabla G_{2i}(z^k), \tag{17}$$

where

$$\begin{aligned} \eta_{1i}(t^k, z^k) &= \frac{\exp(-G_{1i}(z^k)/t^k)}{\exp(-G_{1i}(z^k)/t^k) + \exp(-G_{2i}(z^k)/t^k)} \\ &= \frac{\exp((G(z^k))_i - G_{1i}(z^k)/t^k)}{\exp((G(z^k))_i - G_{1i}(z^k)/t^k) + \exp((G(z^k))_i - G_{2i}(z^k)/t^k)} \in (0, 1), \\ \eta_{2i}(t^k, z^k) &= \frac{\exp(-G_{2i}(z^k)/t^k)}{\exp(-G_{1i}(z^k)/t^k) + \exp(-G_{2i}(z^k)/t^k)} \\ &= \frac{\exp((G(z^k))_i - G_{2i}(z^k)/t^k)}{\exp((G(z^k))_i - G_{1i}(z^k)/t^k) + \exp((G(z^k))_i - G_{2i}(z^k)/t^k)} \in (0, 1) \end{aligned} \tag{18}$$

and

$$\eta_{1i}(t^k, z^k) + \eta_{2i}(t^k, z^k) = 1. \tag{19}$$

Furthermore, for $j \in \{1, 2\}$,

$$\begin{aligned} &\exp((G(z^k))_i - G_{ji}(z^k)/t^k) = 1, \quad \text{if } j \in B(z^k), \\ &\lim_{t^k \to 0} \exp((G(z^k))_i - G_{ji}(z^k)/t^k) = 0, \quad \text{if } j \notin B(z^k). \end{aligned} \tag{20}$$

If $G_{1i}(\bar{z}) \neq G_{2i}(\bar{z})$, then $(G(z))_i = G_{1i}(z)$ (or $G_{2i}(z)$) near $\bar{z}$. Hence, $(G(z))_i$ is continuously differentiable near $\bar{z}$, which implies that $\lim_{z^k \to \bar{z}} \nabla (G(0, z^k))_i = \nabla (G(\bar{z}))_i$. Since $\|B_i(\bar{z})\| = 1$, by Lemma 2.2 we have that $B_i(z^k) \subset B_i(\bar{z})$ for all k big enough. Hence, from (13), $\lim_{t^k\to 0} \nabla_z(G(t^k, z^k))_i = \nabla G_{ji}(z^k)$ for all $j \in B_i(z^k)$. Therefore,

$$\begin{aligned}
&\lim_{\substack{z^k\to\bar{z}\\ t^k\to 0}} \operatorname{dist}\{\nabla_z(G(t^k, z^k))_i - \partial(G(\bar{z}))_i\} = \lim_{\substack{z^k\to\bar{z}\\ t^k\to 0}} \| \nabla_z(G(t^k, z^k))_i - \nabla(G(\bar{z}))_i\| \\
&\leq \lim_{\substack{z^k\to\bar{z}\\ t^k\to 0}} \{\| \nabla_z(G(t^k, z^k))_i - \nabla(G(z^k))_i\| + \| \nabla (G(z^k))_i - \nabla(G(\bar{z}))_i\|\} \\
&= 0.
\end{aligned} \tag{21}$$

If $G_{1i}(\bar{z}) = G_{2i}(\bar{z})$, from (17), (18), (19) and Lemma 2.2, we have that

$$\lim_{\substack{z^k\to\bar{z}\\ t^k\to 0}} \nabla(G(t^k, z^k))_i \in \partial(G(\bar{z}))_i. \tag{22}$$

Therefore, (15) is true for all $i = 1, 2, \dots, m$. The proof is complete. □

3 Global convergence of smoothing method for MPCC

Let z_t be a local optimal solution of problem ($\mathbf{P_t}$). Since ($\mathbf{P_t}$) is an ordinary nonlinear programming problem, under some constraint qualifications, there exist Lagrange multipliers $\nu_t \in \mathcal{R}^m, \lambda_t \in \mathcal{R}^p$, and $\mu_t \in \mathcal{R}^q$ such that the vector $(z_t, \nu_t, \lambda_t, \mu_t)$ satisfies the following KKT conditions for problem ($\mathbf{P_t}$):

$$\begin{aligned}
&\nabla f(z_t) - \sum_{i=1}^{m}(\nu_t)_i \nabla (G(t, z_t))_i + \sum_{r\in\mathcal{I}_t}(\lambda_t)_r \nabla g_r(z_t) \\
&\quad + \sum_{l\in\mathcal{J}}(\mu_t)_l \nabla h_l(z_t) = 0, \\
&G(t, z_t) = 0, \quad h(z_t) = 0, \\
&(\lambda_t)_r \geq 0, \quad g_r(z_t) \leq 0, \quad (\lambda_t)_r g_r(z_t) = 0.
\end{aligned} \tag{23}$$

Moreover, $(z_t, \nu_t, \lambda_t, \mu_t)$ satisfies the inequalities

$$d_t^T \nabla_z^2 L_t(z_t, \nu_t, \lambda_t, \mu_t)d_t \geq 0 \quad \text{for all } d_t \in \mathcal{T}_t(z_t), \tag{24}$$

where

$$L_t(z, \nu, \lambda, \mu) = f(z) + \sum_{i=1}^{m} \nu_i(G(t, z))_i + \sum_{r\in\mathcal{I}_t}\lambda_r g_r(z) + \sum_{l\in\mathcal{J}}\mu_l h_l(z), \tag{25}$$

with $\mathcal{I}_t = \{r|g_r(z_t) = 0\}$ and $\mathcal{T}_t(z) = \{d \in \mathcal{R}^n| \nabla G(t, z)^T d = 0, \nabla g_{\mathcal{I}_t}(z)^T d = 0, \nabla h(z)^T d = 0\}$.

Conditions (23) and (24) are called the *second order necessary conditions* of problem ($\mathbf{P_t}$). The next lemma shows the relationship of the feasibility between MPCC (1) and problem ($\mathbf{P_t}$) when $t > 0$ is sufficiently small, which is crucial for proving our main result.

Lemma 3.1 *For each $t > 0$, let z_t be a feasible point of problem* ($\mathbf{P_t}$). *Suppose that $z_t \to \bar{z}$ as $t \to 0$, then $\bar{z}$ is a feasible point of MPCC (1). Moreover, if LICQ holds at $\bar{z}$, then the gradients*

$$\{\nabla_z(G(t, z_t))_i | i = 1, 2, \ldots, m\} \cup \{\nabla g_r(z_t) | r \in \mathcal{I}_t\} \cup \{\nabla h_l(z_t) | l \in \mathcal{J}\} \tag{26}$$

are linearly independent for all $t > 0$ small enough.

Proof From (9), (10) and the continuity of $G(t, z)$, $g(z)$ and $h(z)$, it is easy to see that $\bar{z}$ is a feasible point of MPCC (1).

We now prove the second part of the lemma. By Lemma 2.1 (iii), $\nabla_z(G(t, z))_i$ is continuously differentiable with respect to z for each $t > 0$ and any $i \in \{1, 2, \ldots, m\}$, so there exists a constant L such that

$$\| \nabla_z(G(t, z_t))_i - \nabla_z(G(t, \bar{z}))_i \| \leq L \|z_t - \bar{z}\| \tag{27}$$

holds for all $t > 0$ and $i \in \{1, 2, \ldots, m\}$. Since $z_t \to \bar{z}$ as $t \to 0$, we have

$$\lim_{t \to 0} \| \nabla_z(G(t, z_t))_i - \nabla_z(G(t, \bar{z}))_i \| = 0. \tag{28}$$

By Lemma 2.1(iii) we obtain that

$$\lim_{t \to 0} \| \nabla_z(G(t, \bar{z}))_i - \sum_{j \in B_i(\bar{z})} \frac{\nabla G_{ji}(\bar{z})}{|B_i(\bar{z})|} \| = 0. \tag{29}$$

From (6), $G_{1i}(\bar{z}) = 0 < G_{2i}(\bar{z})$ for $i \in \bar{\alpha}$, which implies that $G_i(\bar{z}) = G_{1i}(\bar{z})$, $B_i(\bar{z}) = \{1\}$ and $|B_i(\bar{z})| = 1$. Hence, (28) and (29) imply that

$$\begin{aligned} &\lim_{t \to 0} \| \nabla_z(G(t, z_t))_i - \nabla G_{1i}(\bar{z}) \| \\ &\leq \lim_{t \to 0} \| \nabla_z(G(t, z_t))_i - \nabla_z(G(t, \bar{z}))_i \| + \lim_{t \to 0} \| \nabla_z(G(t, \bar{z}))_i - \nabla G_{1i}(\bar{z}) \| \\ &= 0, \end{aligned} \tag{30}$$

Similarly, we can prove that for $i \in \bar{\gamma}$,

$$\lim_{t \to 0} \nabla_z(G(t, z_t))_i = \nabla G_{2i}(\bar{z}). \tag{31}$$

For $i \in \bar{\beta}$, by Theorem 2.1 we have

$$\lim_{t \to 0} \text{dist}\{\nabla_z(G(t, z_t))_i, \partial(G(\bar{z}))_i\} = 0. \tag{32}$$

From the continuous differentiability of functions $g(z)$ and $h(z)$, it is clear that

$$\lim_{t \to 0} \nabla g_r(z_t) = \nabla g_r(\bar{z}), \quad \text{for } r \in \bar{\mathcal{I}}, \tag{33}$$

and

$$\lim_{t \to 0} \nabla h_l(z_t) = \nabla h_l(\bar{z}), \quad \text{for } l \in \mathcal{J}. \tag{34}$$

Therefore, from (16) and the equations (30), (31), (31), (33) and (34) above we have

$$\begin{aligned}&\sum_{i\in\bar{\alpha}}\bar{\nu}_i \bigtriangledown G_{1i}(\bar{z})+\sum_{i\in\bar{\beta}}\bar{\delta}_i[\bar{\eta}_{1i}\bigtriangledown G_{1i}(\bar{z})+\bar{\eta}_{2i}\bigtriangledown G_{2i}(\bar{z})]\\&+\sum_{i\in\bar{\gamma}}\bar{\omega}_i\bigtriangledown G_{2i}(\bar{z})+\sum_{r\in\bar{\mathcal{I}}}\bar{\lambda}_r\bigtriangledown g_r(\bar{z})+\sum_{l\in\mathcal{J}}\bar{\mu}_l\bigtriangledown h_l(\bar{z})=0,\end{aligned}\tag{35}$$

Since LICQ holds at $\bar{z}$, (35) implies that $\bar{\nu}_i=\bar{\omega}_i=\bar{\delta}_i\bar{\eta}_{1i}=\bar{\delta}_i\bar{\eta}_{2i}=\bar{\lambda}_r=\bar{\mu}_l=0$. From $\bar{\eta}_{1i}\geq 0,\ \bar{\eta}_{2i}\geq 0$ and $\bar{\eta}_{1i}+\bar{\eta}_{2i}=1$, we have that $\bar{\delta}_i=0$. Therefore, $\{\nabla_z(G(0,\bar{z}))_i\,|\,i=1,2,\dots,m\}\cup\{\nabla g_r(\bar{z})\,|\,r\in\bar{\mathcal{I}}\}\cup\{\nabla h_l(\bar{z})\,|\,l\in\mathcal{J}\}$ are linearly independent. By Lemma 2.2, $\bar{\mathcal{I}}_t\subseteq\bar{\mathcal{I}}$ for all $t>0$ small enough. Furthermore, from the twice continuous differentiability of $G(t,z),\ g(z)$ and $h(z)$, we deduce that the gradients $\{\nabla_z(G(t,z_t))_i|i=1,2,\dots,m\}\cup\{\nabla g_r(z_t)|r\in\mathcal{I}_t\}\cup\{\nabla h_l(z_t)|l\in\mathcal{J}\}$ are linearly independent for all $t>0$ small enough. □

We now prove the global convergence of the smoothing method for MPCC (1).

Theorem 3.1 *Let $\{t^k\}$ be a sequence of positive scalars tending to zero, $(z^k,\nu^k,\lambda^k,\mu^k)$ satisfy the second-order necessary conditions (23) and (24) for problem $(\mathbf{P}_{\mathbf{t^k}})$. Suppose that $(z^k,\nu^k,\lambda^k,\mu^k)\to(\bar{z},\bar{\nu},\bar{\lambda},\bar{\mu})$ as $t^k\to 0$. If LICQ and ULSC hold at $\bar{z}$, then $\bar{z}$ is a B-stationary point of MPCC (1).*

Proof Since $(z^k,\nu^k,\lambda^k,\mu^k)$ satisfy the second-order necessary conditions (23) and (24) for problem $(\mathbf{P}_{\mathbf{t^k}})$, then we have

$$\begin{aligned}&\nabla f(z^k)-\sum_{i=1}^m\nu_i^k\bigtriangledown(G(t^k,z^k))_i+\sum_{r\in\mathcal{I}^k}\lambda_r^k\bigtriangledown g_r(z^k)\\&\quad+\sum_{l\in\mathcal{J}}\mu_l^k\bigtriangledown h_l(z^k)=0,\\&G(t^k,z^k)=0,\quad h(z^k)=0,\\&\lambda_r^k\geq 0,\quad g_r(z^k)\leq 0,\quad \lambda_r^{k^T}g_r(z^k)=0.\end{aligned}\tag{36}$$

and

$$d^{kT}\bigtriangledown_z^2L_{t^k}(z^k,\nu^k,\lambda^k,\mu^k)d^k\geq 0\quad\text{for all } d^k\in\mathcal{T}_{t^k}(z^k).\tag{37}$$

By Lemma 2.2 we know that $B_i(z^k)\subseteq B_i(\bar{z})$ $(i=1,2,\dots,m)$ and $\mathcal{I}(z^k)\subseteq\bar{\mathcal{I}}$ for all k sufficiently large. It is easy to verify that $\bar{\lambda}_r\geq 0$ for $r\in\bar{\mathcal{I}}$ and $\bar{\lambda}_r=0$ for $r\notin\bar{\mathcal{I}}$.

From the proof of Lemma 3.1 we have that

(i) for $i\in\bar{\alpha}$, $\lim_{\substack{z^k\to\bar{z}\\ t^k\to 0}}\nabla(G(t^k,z^k))_i=\nabla G_{1i}(\bar{z})$.

(ii) for $i\in\bar{\gamma}$, $\lim_{\substack{z^k\to\bar{z}\\ t^k\to 0}}\nabla(G(t^k,z^k))_i=\nabla G_{2i}(\bar{z})$.

(iii) for $i\in\bar{\beta}$, Let $\bar{u}_i=\lim_{\substack{z^k\to\bar{z}\\ t^k\to 0}}\nu^k\eta_{1i}(t^k,z^k),\ \bar{w}_i=\lim_{\substack{z^k\to\bar{z}\\ t^k\to 0}}\nu^k\eta_{2i}(t^k,z^k)$.

Then from (i), (ii), (iii), equality (17), by taking limit on system (36), we get

$$\begin{aligned}&\nabla f(\bar{z})-\sum_{i\in\bar{\alpha}}\bar{\nu}_i\bigtriangledown G_{1i}(\bar{z})-\sum_{i\in\bar{\gamma}}\bar{\nu}_i\bigtriangledown G_{2i}(\bar{z})-\sum_{i\in\bar{\beta}}\bar{u}_i\bigtriangledown G_{1i}(\bar{z})\\&-\sum_{i\in\bar{\beta}}\bar{w}_i\bigtriangledown G_{2i}(\bar{z})+\sum_{r\in\bar{\mathcal{I}}}\bar{\lambda}_r\bigtriangledown g_r(\bar{z})+\sum_{l\in\mathcal{J}}\bar{\mu}_l\bigtriangledown h_l(\bar{z})=0,\\&G(\bar{z})=0,\quad h(\bar{z})=0,\\&\bar{\lambda}_r\geq 0,\quad g(\bar{z})\leq 0,\quad \bar{\lambda}^{rT}g(\bar{z})=0.\end{aligned}\tag{38}$$

Now we need to prove $\bar{u}_i \geq 0$, $\bar{w}_i \geq 0$ for $i \in \bar{\beta}$. By contradiction, we assume that $\bar{u}^j < 0$ for some $j \in \bar{\beta}$. Since LICQ holds at $\bar{z}$, from Lemma 3.1 we may choose a vector $d^k \in \mathcal{R}^n$ such that, for k big enough,

$$\begin{array}{l} \triangledown G_{1j}(z^k)^T d^k = \eta_{2j}(t^k, z^k), \\ \triangledown G_{1i}(z^k)^T d^k = 0, \;\; i \in \bar{\alpha} \cup \bar{\beta} \setminus \{j\}, \\ \triangledown G_{2j}(z^k)^T d^k = -\eta_{1j}(t^k, z^k), \\ \triangledown G_{2i}(z^k)^T d^k = 0, \;\; i \in \bar{\gamma} \cup \bar{\beta} \setminus \{j\}, \\ \triangledown g_r(z^k)^T d^k = 0, \;\; r \in \mathcal{I}^k, \\ \triangledown h_l(z^k)^T d^k = 0, \;\; l \in \mathcal{J}. \end{array} \tag{39}$$

Therefore, $\triangledown(G(z^k))_j{}^T d^k = \eta_{1j}(t^k, z^k) \triangledown G_{1j}(z^k)^T d^k + \eta_{2j}(t^k, z^k) \triangledown G_{2j}(z^k)^T d^k = 0$, which implies that $d^k \in \mathcal{T}_{t^k}(z^k)$. Furthermore,

$$\begin{aligned} & d^{kT} \triangledown_z^2 L(z^k, \nu^k, \lambda^k, \mu^k) d^k \\ & = d^{kT}[\triangledown^2 f(z^k) + \sum_{r \in \mathcal{I}(z^k)} \lambda_r^k \triangledown^2 g(z^k) + \sum_{l \in \mathcal{J}} \mu_l^k \triangledown^2 h(z^k) \\ & \qquad - \sum_{i=1}^m \nu_i^k \triangledown_z^2 (G(t^k, z^k))_i] d^k. \end{aligned} \tag{40}$$

It follows from (12) that

$$\begin{aligned} & d^{kT} \triangledown_z^2 (G(t^k, z^k))_i d^k \\ & = \eta_{1i}(t^k, z^k) d^{kT} \triangledown^2 G_{1i}(z^k) d^k + \eta_{2i}(t^k, z^k) d^{kT} \triangledown^2 G_{2i}(z^k) d^k \\ & \quad - \frac{1}{t^k} \eta_{1i}(t^k, z^k) \eta_{2i}(t^k, z^k) \big[d^{kT} \triangledown G_{1i}(z^k) (\triangledown G_{1i}(z^k))^T d^k \\ & \qquad + d^{kT} \triangledown G_{2i}(z^k) (\triangledown G_{2i}(z^k))^T d^k \big] \\ & \quad + \frac{1}{t^k} \eta_{1i}(t^k, z^k) \eta_{2i}(t^k, z^k) \big[d^{kT} \triangledown G_{1i}(z^k) (\triangledown G_{2i}(z^k))^T d^k \\ & \quad + d^{kT} \triangledown G_{2i}(z^k) (\triangledown G_{1i}(z^k))^T d^k \big]. \end{aligned} \tag{41}$$

Hence, for $i = j$, from (39), (41) and Lemma 2.1 we have that

$$\begin{aligned} & d^{kT} \triangledown_z^2 (G(t^k, z^k))_j d^k \\ & = \lambda_{1j}(t^k, z^k) d^{kT} \triangledown^2 G_{1j}(z^k) d^k + \eta_{2j}(t^k, z^k) d^{kT} \triangledown^2 G_{2j}(z^k) d^k \\ & \quad - \frac{1}{t^k} \eta_{1j}(t^k, z^k) \eta_{2j}(t^k, z^k) [\eta_{2j}^2(t^k, z^k) + (-\eta_{1j}(t^k, z^k))^2] \\ & \quad + \frac{1}{t^k} \eta_{1j}(t^k, z^k) \eta_{2j}(t^k, z^k) \Big[(-\eta_{1j}(t^k, z^k)) \eta_{2j}(t^k, z^k) \\ & \qquad + \eta_{2j}(t^k, z^k)(-\lambda_{1j}(t^k, z^k)) \Big] \end{aligned} \tag{42}$$

$$= \eta_{1j}(t^k, z^k) d^{kT} \nabla^2 G_{1j}(z^k) d^k + \eta_{2j}(t^k, z^k) d^{kT} \nabla^2 G_{2j}(z^k) d^k$$
$$- \frac{1}{t^k} \eta_{1j}(t^k, z^k) \eta_{2j}(t^k, z^k).$$

Similarly, for $i \neq j$, we have

$$d^{kT} \nabla_z^2 (G(t^k, z^k))_i d^k = \eta_{1i}(t^k, z^k) d^{kT} \nabla^2 G_{1i}(z^k) d^k + \eta_{2i}(t^k, z^k) d^{kT} \nabla^2 G_{2i}(z^k) d^k. \quad (43)$$

Equations (40), (42) and (43) imply that

$$\begin{aligned} & d^{kT} \nabla_z^2 L(z^k, \nu^k, \lambda^k, \mu^k) d^k \\ &= d^{kT} \nabla^2 f(z^k) d^k + \sum_{r \in \mathcal{I}^k} \lambda_r^k d^{kT} \nabla^2 g(z^k) d^k + \sum_{l \in \mathcal{J}} \mu_l^k d^{kT} \nabla^2 h(z^k) d^k \\ & \quad - \sum_{i=1}^m \nu_i^k [\eta_{1i}(t^k, z^k) d^{kT} \nabla^2 G_{1i}(z^k) d^k + \eta_{2i}(t^k, z^k) d^{kT} \nabla^2 G_{2i}(z^k) d^k] \\ & \quad + \frac{1}{t^k} \nu_j^k \eta_{1j}(t^k, z^k) \eta_{2j}(t^k, z^k). \end{aligned} \quad (44)$$

Since ULSC condition holds at $\bar{z}$, i.e., $\bar{u}_i \bar{w}_i \neq 0$ for all $i \in \bar{\beta}$, we have that $\lim_{t^k \to 0} \eta_{1j}(t^k, z^k) \eta_{2j}(t^k, z^k) > 0$. Moreover, $\bar{u}_j < 0$ implies that $\lim_{t^k \to 0} \nu_j^k = \bar{\nu}_j < 0$. Hence

$$\frac{1}{t^k} \nu_j^k \eta_{1j}(t^k, z^k) \eta_{2j}(t^k, z^k) \to -\infty \quad \text{as } t^k \to 0.$$

Noticing the boundedness of d^k, $\eta_{1i}(\cdot)$ and $\eta_{2i}(\cdot)(i = 1, \ldots, m)$, we have $d^{kT} \nabla^2 f(z^k) d^k$, $\lambda_r^k d^{kT} \nabla^2 g(z^k) d^k$, $\mu_l^k d^{kT} \nabla^2 h(z^k) d^k$, and $\nu_i^k [\eta_{1i}(t^k, z^k) d^{kT} \nabla^2 G_{1i}(z^k) d^k + \eta_{2i}(t^k, z^k) d^{kT} \nabla^2 G_{2i}(z^k) d^k]$ $(i = 1, \ldots, m)$ are bounded as $t^k \to 0$. It follows that

$$d^{kT} \nabla_z^2 L_{t^k}(z^k, \nu^k, \lambda^k, \mu^k) d^k \to -\infty \quad \text{as } t^k \to 0, \quad (45)$$

which contradicts the assumption that $(z^k, \nu^k, \lambda^k, \mu^k)$ satisfies the second-order necessary conditions. Hence we proved that $\bar{u}_i \geq 0$ holds for all $i \in \bar{\beta}$.

Similarly, $w_i \geq 0$ holds for all $i \in \bar{\beta}$. By taking $\nu = \begin{cases} \bar{\nu}_i, & i \in \bar{\alpha}, \\ \bar{u}_i, & i \in \bar{\beta}, \end{cases}$ $\omega = \begin{cases} \bar{w}_i, & i \in \bar{\beta}, \\ \bar{\nu}_i, & i \in \bar{\gamma}, \end{cases}$ $\lambda_r = \bar{\lambda}_r$ for $r \in \bar{\mathcal{I}}$ and $\mu_l = \bar{\mu}_l$ for $l \in \bar{\mathcal{J}}$, we have the equation (7) hold at $\bar{z}$. That is, $\bar{z}$ is a B-stationary point of MPCC (1). □

From (22) in the proof of Theorem 2.1, we know that, for each $i \in \bar{\beta}$, any accumulation point of $\{\nabla(G(t^k, z^k))_i\}$ (as $z^k \to \bar{z}$, $t^k \to 0$) belongs to $\partial G(\bar{z})_i$ and has the form of

$$\bar{r} = \bar{\eta}_{1i} \nabla G_{1i}(\bar{z}) + \bar{\eta}_{2i} \nabla G_{2i}(\bar{z}), \quad (46)$$

where $\bar{\eta}_{1i}, \bar{\eta}_{2i} \in [0, 1]$, and $\bar{\eta}_{1i} + \bar{\eta}_{2i} = 1$. Furthermore, from the last part of the proof of Theorem 3.1, we can see that the ULSC condition can be replaced by a much weaker condition.

Corollary 3.1 *Let $\{t^k\}$ be a sequence of positive scalars tending to zero, $(z^k, \nu^k, \lambda^k, \mu^k)$ satisfy the second-order necessary conditions (23) and (24) for problem* $(\mathbf{P}_{\mathbf{t^k}})$. *Suppose that $(z^k, \nu^k, \lambda^k, \mu^k) \to (\bar{z}, \bar{\nu}, \bar{\lambda}, \bar{\mu})$ as $t^k \to 0$. If LICQ hold at $\bar{z}$, and for $i \in \bar{\beta}$, the limitations of $\eta_{i1}(t^k, z^k)$ and $\eta_{i1}(t^k, z^k)$ in (18), denoted as $\bar{\eta}_{1i}$ and $\bar{\eta}_{2i}$ respectively, satisfy $\bar{\eta}_{1i}\bar{\eta}_{2i} \neq 0$, then $\bar{z}$ is a B-stationary point of MPCC (1).*

$\{z^k\}$ is saied to be asymptotically weakly nondegenerate if $\bar{\eta}_{1i}\bar{\eta}_{2i} \neq 0$ for $i \in \bar{\beta}$ in Fukushima and Pang (1999). For the smoothing technique used in Fukushima and Pang (1999), this condition means that $G_{1i}(z^k)$ and $G_{2i}(z^k)$ approach to zero in the same order of magnitude. However, from (18) we can easy to see that, by using the aggregation technique, it is not necessary to request $G_{1i}(z^k)$ and $G_{2i}(z^k)$ approaching to zero in the same order of magnitude for guaranteeing $\bar{\eta}_{1i}\bar{\eta}_{2i} \neq 0$ for $i \in \bar{\beta}$. For example, let $G_1(z) = z$ and $G_2(z) = z^2$, and suppose that $z^k = t^k$, then as $t^k \to 0$, $G_1(z^k) \to 0$ and $G_2(z^k) \to 0$, but in different order of magnitude. And in this case, $\eta_1(z^k) \to \exp(-1)/(\exp(-1)+1)$ and $\eta_2(z^k) \to 1/(\exp(-1)+1)$, which implies that $\bar{\eta}_1\bar{\eta}_2 \neq 0$. In this sense, the convergence condition in Corollary 3.1 is weaker than that in Fukushima and Pang (1999).

4 Further remarks

In this paper, we give a new smoothing method for solving MPCC. By using the aggregation technique, we approximate the MPCC (1) with a sequence of parametric smooth nonlinear programming problem. Under some mild conditions, the sequence generated from solving the smooth problems converges to a B-stationary point of MPCC.

By defining

$$(G(t,z))_i = -t\ln\Big(\sum_{j=1}^{k}\exp(-G_{ij}(z)/t)\Big), \quad \text{for } i = 1, 2, \ldots, m,$$

the method in this paper can be easily extended to solve the mathematical programming with general vertical complementarity constraints (Scheel and Scholtes 2000):

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{s.t.} & \min\{G_{i1}(z), G_{i2}(z), \ldots, G_{ik}(z)\} = 0, \;\; i = 1, 2, \ldots, m, \\ & g(z) \leq 0, \\ & h(z) = 0. \end{array} \tag{47}$$

The concepts for MPCC in section 2, such as the B-stationary point, KKT point, LICQ, ULSC condition etc., could be extended to problem (47) but would be much complicated, see Scheel and Scholtes (2000) for reference. We also noticed that the Fischer–Burmeister smoothing methods for MPCC Fukushima and Pang (1999) is much difficult to be extended to handle this kind of problems.

Acknowledgements The authors are very grateful to the associate editor and anonymous referees for their valuable comments, which have helped us to improve the quality of the paper. H. Yin was partially supported by the Hong Kong Research Grant Council and the National Natural Science Foundation of China NSF70472074. J. Zhang is supported by the Hong Kong Research Council.

References

Chang PL (1980) A minimax approach to nonlinear programming, Doctoral Dissertation, Department of Mathematics, University of Washington, Seattle

Clarke FH (1983) Optimization and nonsmooth analysis. Wiley, Vhichester

Cottle RW, Dantzig GB (1970) A generalization of the linear complementarity problem. J Comb Theory 8:79–90

Faccinei F, Jiang H, Qi L (1999) A smoothing methods for mathematical programs with equilibrium constraints. Math Program 85:107–135

Fukushima M, Pang J-S (1999) Convergence of a smoothing continuation method for mathematical programs with complementarity constraints. Ill-Posed variational problems and regularization techniques (Trier, 1998) pp 99–110, Lecture notes in economics and mathematical systems, vol 477, Théra M, Tichatschke R (eds). Springer, Berlin Heidelberg New York

Fukushima M, Luo Z-Q, Pang J-S (1998) A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints. Comput Optim Appl 13:1–13

Fukushima M, Tseng P (2002) An implementable active-set algorithm for computing a B-stationary point of a mathematical program with linear complimentarity constraints. SIAM J Optim 12:724–739

Goldstein AA (1997) Chebyshew approximation and linear inequalities via exponentials. Report, Department od Mathematics, University of Washington

Jiang H, Ralph D (2000) Smooth SQP methods for mathematical programs with nonlinear complementarity constraints SIAM J Optim 10:779–808

Kachiyan K (1996) Coordination complexity of parallel price-directive decomposition. Math Oper Res 21:321–340

Li XS (1991) An aggregate function method for nonlinear programming. Sci China (Ser A) 34:1467–1473

Li XS (1992) An entropy-based aggregate method for minimax optimization. Engineering. Optimization 18:277–185

Liu X, Sun J (2004) Generalized stationary points and an interior-point method for mathematical programs with equilibrium constraints. Math Program 101:231–261

Luo Z-Q, Pang J-S, Ralph D (1996a) Mathematical programs with equilibrium constraints. Cambridge University Press

Luo Z-Q, Pang J-S, Ralph D, Wu SQ (1996b) Exact penalization and stationary conditions of mathematical programs with equilibrium constraints. Math Program 76:19–76

Outrata JV, Kočvara M, Zowe J (1998) Nonsmooth approach to optimization problems with equilibrium constraints: theory, applications and numerical results, Nonconvex optimization and its applications, vol 28. Kluwer, Dordrecht

Outrata J, Zowe J (1995) A numerical approach to optimization problems with variational inequality constraints. Math program 68:105–130

Pang J-S, Fukushima M (1999) Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints. Comput Optim Appl 13:111–136

Peng J, Lin Z (1999) A non-interior continuation method for generalized liner complementarity problems. Math Program 86:533–563

Qi H, Liao L (1999) A smoothing Newton method for extended vertical linear complementarity problems. SIAM J Matrix Anal Appl 21:45–66

Qi L, Sun D, Zhou G (2000) A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities. Math Program 87:1–35

Raghunathan AU, Biegler LT (2005) An interior point method for mathematical programs with complementarity constraints (MPCCs). SIAM J Optim 15:720–750

Scheel H, Scholtes S (2000) Mathematical programs with equilibrium constraints: stationarity, optimality and sensitivity. Math Oper Res 25:1–22

Scholtes S (2001) Convergence properties of a regularisation scheme for mathematical programs with complementarity constraints? SIAM J Optim 11:918–936

Scholtes S, Stühr M (1999) Exact penalization of mathematical programs with complementarity constraints. SIAM J Control Optim 37:617–652

Scholtes S, Stühr M (2001) How stringent is the linear independence assumption for mathematical programs with complementarity constraints. Math Oper Res 26:851–863

Tseng P, Bertsekas DP (1993) On the convergence of the exponential multiplier method for convex programming. Math Program 60:1–19
Ye JJ, Ye XY (1997) Necessary optimality conditions for optimization problems with variational inequality constraints. Math Oper Res 22:977–997
Ye JJ (1999) Optimality conditions for optimization problems with complementarity constraints. SIAM J Optim 9:374–387
Ye JJ (2000) Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints. SIAM J Optim 10:943–962