Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 16, No. 4, pp. 525–545, 2020

A SMOOTHING METHOD FOR MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY CONSTRAINTS*

Suxiang He, Tingting Shen and Jie Zhang†

Abstract: This paper studies a smoothing method for mathematical programs with complementarity constraints (MPCC) based on the integral of the sigmoid function, in which the original MPCC is reformulated as a standard smooth approximation minimization problem with a smooth parameter and the approximate solution of the MPCC is obtained by solving a series of the smooth subproblems where the smooth parameter approaches zero. It is proven that the accumulation point of the sequence of KKT solutions to the smooth subproblems is a C-stationary point of the MPCC under the MPCC-MFCQ, without the upper level strict complementarity and the asymptotically weakly nondegenerate condition. Furthermore, the accumulation point can be proven to be an S-stationary point under the weak second-order necessary condition. Moreover, the characterizations of the linear independence constraints qualification, the KKT condition and the second-order sufficient condition for the smooth approximation problem are established under several assumptions on the original MPCC, which ensures the existence of KKT solutions to the smooth subproblems. At last, the numerical experiments are implemented to test the performance of the smoothing method by solving some typical problems in MacMPEC database. The reported numerical results show that the smoothing method is promising by comparing with those by the other typical methods in the existent references.

Key words: *mathematical programs with complementarity constraints, smooth approximation problem, KKT solution, C-stationary point, S-stationary point*

Mathematics Subject Classification: *90C30, 90C33*

1 Introduction

In this paper, we consider the mathematical programs with complementarity constraints (MPCC) of the form:

$$\begin{aligned} &\min f(x) \\ &\text{s.t. } g(x) \le 0, h(x) = 0, \\ &\qquad G(x) \ge 0, H(x) \ge 0, \\ &\qquad G(x)^T H(x) = 0, \end{aligned} \tag{1.1}$$

where $f : \Re^n \to \Re$, $g : \Re^n \to \Re^m$, $h : \Re^n \to \Re^p$, and G, $H : \Re^n \to \Re^l$ are all twice continuously differentiable functions, the inequality constraint $g(x) \le 0$ and the equality constraint

*The authors are supported by the NSFC (No.11671183), the Fundamental Research Funds for the Central Universities (No.2018IB016, No.2019IB010, No. 2020IB010), the Equipment Pre-Research Ministry of Education Joint Fund (No.6141A02033703) and the Natural Science Foundation of Liaoning Province of China, Program for Liaoning Excellent Talents in University (No. LR2017049).

†Corresponding author.

$h(x) = 0$ are called general constraints, and $G(x) \geq 0$, $H(x) \geq 0$ and $G(x)^T H(x) = 0$ are called complementarity constraints.

Problem (1.1) is an important subclass of mathematical programs with equilibrium constraints (MPEC), which has wide applications in economics, engineering, transportation, equilibrium game and so on (see [10, 18, 25, 30]). Therefore, problem (1.1) has been attracting many researchers to investigate it profoundly. However, it is difficult to solve problem (1.1) directly by the existing algorithms for the standard constrained optimization due to the existence of complementarity constraints, which makes any feasible point of problem (1.1) do not satisfy the constraints qualification of the standard constrained optimization, such as the linear independence constraints qualification (LICQ) and the Mangasarian-Fromovitz constraints qualification (MFCQ) (see [4] and [28]). Moreover, the feasible region of problem (1.1) is more complex than that of the standard constrained optimization, that is, its feasible region has no favorable properties such as connectivity, closure and convexity. In view of these reasons, many characteristic methods have been explored for solving problem (1.1), such as the smoothing method (see [2, 5–7, 9, 14, 19, 20, 27, 29]), the sequential quadratic programming method (see [1, 26]), the relaxation method (see [11, 13, 15, 17, 23, 24]), the generalized project metric method (see [8]), the efficient method for solving the stationary point for problem (1.1) (see [16]), and the inexact log-exponential regularization method (see [21]).

Among them, the smoothing method is one kind of typical approaches to solving problem (1.1), in which a smooth function with a parameter is employed to approximate the complementarity constraints in problem (1.1), and then the original problem (1.1) is reformulated into a standard smooth approximate optimization problem. Thus, a sufficiently approximate solution or some type of stationary point of problem (1.1) can be obtained by solving a series of the constructed smooth subproblems. The existing smoothing method mainly applies smooth functions to approximate complementarity constraints based on Fischer-Burmeister function and maximum or minimal function. For example, Fischer-Burmeister perturbation function

$$\phi_\varepsilon(a, b) = a + b - \sqrt{a^2 + b^2 + \varepsilon}$$

is explored in [9], where a and b are constants, and ε is a smooth parameter. The authors use the function based constraints to approximate the complementarity constraints and it is proved that any sequence of stationary points converges to B-stationary point of problem (1.1) under the MPCC-LICQ, the asymptotically weakly nondegenerate condition and the weak second-order necessary condition.

In [7], MPEC problem with strongly monotone variational inequalities is transformed into an equivalent one-level nonsmooth optimization problem, whose form can be regarded as problem (1.1) to some extent. Furthermore, by means of the smooth Fischer-Burmeister perturbation function below

$$\varphi_\mu(a, b) = \sqrt{(a-b)^2 + 4\mu^2} - (a + b),$$

where μ is a smooth parameter, a sequence of smooth and regular problems that progressively approximates the nonsmooth problem is obtained, it is shown that the stationary points of the approximate problems converge to a solution of the original problem, and the numerical results are reported, which show the viability of the approach.

Moreover, a smooth approximation function is constructed in [6] by using the following function

$$\Psi_\varepsilon(t) = \frac{2t}{\pi} \arctan\left(\frac{t}{\varepsilon}\right),$$

where t is a variable, ε is a smooth parameter, based on which only the equality $G(z)^T H(z) = 0$ is approximated by a smooth inequality. It is proved that the arbitrary accumulation point of the sequence of solutions is a C-stationary point under MFCQ. If some additional assumptions are coupled with, the accumulation point is an M-stationary point, or even an S-stationary point. And the numerical results obtained demonstrate that the method is much more promising than the similar ones in [7] under the same initial conditions.

A smoothing regularization method for problem (1.1) is presented based on the construction of a new smoothing and regular function in [5], where the complementarity constrains are replaced by

$$\varphi(a, b) = \frac{1}{2}(|a| - a + |b| - b + |ab| + ab) = 0$$

and then problem (1.1) is reformulated as a standard smooth nonlinear program by replacing $|\cdot|$ with $\sqrt{|\cdot|^2 + \varepsilon^2}$ in the above function, where ε is a smooth parameter. It is proved that any accumulation point of the sequence of solutions to the corresponding approximate problem is an S-stationary point of the problem (1.1) and numerical results show the efficiency of the proposed approach by comparing with the method in [7].

Motivated by the above works, this paper focuses on investigating a smoothing method for problem (1.1) by using the simple smooth integral of the Sigmoid function developed in [3], which has many different interesting properties from the above functions. For example, it has favorable smoothness. And the problem (1.1) is reformulated as a standard smooth optimization problem with simple form by transforming the complementarity constraints into the equality constraints by means of the smooth integral with a smooth parameter. In addition, we find that the upper level strict complementarity, the asymptotically weakly nondegenerate condition and the other assumptions are needed in some smoothing methods, such as [9,14,29]. But it is interesting that the smooth integral of the Sigmoid function with a smooth parameter has good properties such that it does not require these special conditions in the convergence analysis of the corresponding smoothing method.

Furthermore, notice that the sequence of KKT (Karush-Kuhn-Tucker) solutions to the smooth approximation problem is generally supposed to be existent directly in the above mentioned smoothing methods, but they do not explore the problems that whether or not the KKT solutions to the smooth approximation problem exist, and if so, what assumptions on the original problem (1.1) should be made. In view of this reason, this paper aims to discuss the linearly independent constraint qualification, the KKT condition and the second-order sufficient condition of smooth approximation problem under some given assumptions on problem (1.1) after the convergence of the studied method is analyzed. This argument guarantees the existence of KKT solutions to the smooth approximation subproblem and convince us that the KKT solutions can be available by the existing optimization methods.

This paper is organized as follows. We recall some concepts of problem (1.1) and present the smooth approximate optimization formulation of problem (1.1) by using the smooth integral of the Sigmoid function (see [3]) in Section 2. In Section 3, we prove that any accumulation point of the sequence of KKT solutions obtained by solving a series of the approximate subproblems is a C-stationary point as the smooth parameter tends to zero under MPCC-MFCQ. Moreover, the accumulation point can be proven to be an S-stationary point under the weak second-order necessary condition. Section 4 further makes discussion on the existence of KKT solutions to the smooth approximate subproblem by characterizing the linearly independent constraint qualification, the KKT condition and the second-order sufficient condition under some mild assumptions on problem (1.1). Moreover, although the theoretical convergence analyses are comprehensive in some references, the numerical experiments are not implemented to show the computational efficiency of their smoothing method, such as Refs. [9, 14, 27, 29]. Hence, we carry out numerical experiments based on the smooth approximate problem, in which some typical problems with different dimensions in MacMPEC database (see [12]) are used to test the performance of the smooth method in Section 5. Meanwhile, considering that the methods in [5, 7] have the better behaviors, we make comparisons with them under the same experimental conditions. The reported numerical results indicate that the proposed smoothing method is promising. In the last section, some conclusions are drawn.

2 Preliminaries and the Smoothing Method

This section firstly recalls some preliminaries of problem (1.1), and then presents a smoothing method for problem (1.1) based on the integral of the sigmoid function in Ref. [3].

Denote by $\mathcal{F}$ the feasible set of problem (1.1). Let $x \in \mathcal{F}$, and we define the following index sets:

$$\begin{aligned} I_{0+}(x) &= \{i|G_i(x) = 0, H_i(x) > 0, i = 1, 2, \ldots, l\}, \\ I_{00}(x) &= \{i|G_i(x) = 0, H_i(x) = 0, i = 1, 2, \ldots, l\}, \\ I_{+0}(x) &= \{i|G_i(x) > 0, H_i(x) = 0, i = 1, 2, \ldots, l\}, \\ I_g(x) &= \{i|g_i(x) = 0, i = 1, 2, \ldots, m\}. \end{aligned}$$

The following definitions refer to Ref. [18] and [22].

Definition 2.1. A point $\bar{x}$ in $\mathcal{F}$ is said to satisfy MPCC-MFCQ if and only if the vectors of $\{\nabla h_i(\bar{x}), i = 1, 2, \ldots, p\}$, $\{\nabla G_i(\bar{x}), i \in I_{00}(\bar{x}) \cup I_{0+}(\bar{x})\}$ and $\{\nabla H_i(\bar{x}), i \in I_{00}(\bar{x}) \cup I_{+0}(\bar{x})\}$ are linearly independent, and there exists a vector $d \in \Re^n$ such that $\nabla g_i(\bar{x})^T d < 0, i \in I_g(\bar{x})$, $\nabla h_i(\bar{x})^T d = 0,\ i = 1, 2, \ldots, p$, $\nabla G_i(\bar{x})^T d = 0,\ i \in I_{00}(\bar{x}) \cup I_{0+}(\bar{x})$, and $\nabla H_i(\bar{x})^T d = 0,\ i \in I_{00}(\bar{x}) \cup I_{+0}(\bar{x})$.

Definition 2.2. A point $\bar{x}$ in $\mathcal{F}$ is said to satisfy MPCC-LICQ if the vectors of $\{\nabla g_i(\bar{x}), i \in I_g(\bar{x})\}$, $\{\nabla h_i(\bar{x}), i = 1, 2, \ldots, p\}$, $\{\nabla G_i(\bar{x}), i \in I_{00}(\bar{x}) \cup I_{0+}(\bar{x})\}$, and $\{\nabla H_i(\bar{x}), i \in I_{00}(\bar{x}) \cup I_{+0}(\bar{x})\}$ are linearly independent.

Definition 2.3. The Lagrange function of problem (1.1) at x is defined as follows:

$$L(x, \lambda, \mu, u, v) = f(x) + \lambda^T g(x) + \mu^T h(x) - u^T G(x) - v^T H(x),$$

where $\lambda \in \Re^m, \mu \in \Re^p, u \in \Re^l$ and $v \in \Re^l$ are Lagrange multipliers.

Then one has

$$\nabla_x L(x, \lambda, \mu, u, v) = \nabla f(x) + \lambda^T \nabla g(x) + \mu^T \nabla h(x) - u^T \nabla G(x) - v^T \nabla H(x),$$

and

$$\begin{aligned} \nabla_x^2 L(x, \lambda, \mu, u, v) &= \nabla^2 f(x) + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x) + \sum_{i=1}^{p} \mu_i \nabla^2 h_i(x) \\ &\quad - \sum_{i=1}^{l} u_i \nabla^2 G_i(x) - \sum_{i=1}^{l} v_i \nabla^2 H_i(x). \end{aligned}$$

Definition 2.4. Let $\tilde{x} \in \mathcal{F}$. Then

(1) $\tilde{x}$ is said to be W-stationary, if there exist multiplier vectors $\tilde{\lambda} \in \Re^m$, $\tilde{\mu} \in \Re^p$, $\tilde{u} \in \Re^l$, and $\tilde{v} \in \Re^l$ such that

$$\nabla_x L(\tilde{x}, \tilde{\lambda}, \tilde{\mu}, \tilde{u}, \tilde{v}) = 0,$$
$$\tilde{\lambda} \geq 0, \tilde{\lambda}^T g(\tilde{x}) = 0,$$
$$\tilde{u}_i = 0, i \in I_{+0}(\tilde{x}),$$
$$\tilde{v}_i = 0, i \in I_{0+}(\tilde{x});$$

(2) $\tilde{x}$ is said to be C-stationary, if it is W-stationary and

$$\tilde{u}_i \tilde{v}_i \geq 0 \ for \ i \in I_{00}(\tilde{x});$$

(3) $\tilde{x}$ is said to be M-stationary, if it is W-stationary and

$$\tilde{u}_i > 0, \tilde{v}_i > 0, \ or \ \tilde{u}_i \tilde{v}_i = 0 \ for \ i \in I_{00}(\tilde{x});$$

(4) $\bar{x}$ is said to be S-stationary, if it is W-stationary and

$$\tilde{u}_i \geq 0, \tilde{v}_i \geq 0 \ for \ i \in I_{00}(\tilde{x}).$$

Definition 2.5. A finite set of vectors $\{a_i, i \in I_1\} \cup \{b_i, i \in I_2\}$ is said to be positive-linearly dependent if there exists $(\alpha, \beta) \in \Re^{|I_1|+|I_2|} \neq 0$ such that for all $i \in I_1$, $\alpha_i \geq 0$, and

$$\sum_{i \in I_1} \alpha_i a_i + \sum_{i \in I_2} \beta_i b_i = 0. \tag{2.1}$$

Conversely, if (2.1) holds if and only if $(\alpha, \beta) = 0$, then the group of vectors is said to be positive-linearly independent.

The following Lemma will be used in the forthcoming theoretical analysis (see [11]).

Lemma 2.6. *A point $\bar{x}$ in $\mathcal{F}$ satisfies MPCC-MFCQ if and only if the vectors of* $\{\nabla g_i(\bar{x}), i \in I_g(\bar{x})\}$, $\{\nabla h_i(\bar{x}), i = 1, 2, \ldots, p\}$, $\{\nabla G_i(\bar{x}), i \in I_{00}(\bar{x}) \cup I_{0+}(\bar{x})\}$, *and* $\{\nabla H_i(\bar{x}), i \in I_{00}(\bar{x}) \cup I_{+0}(\bar{x})\}$ *are positive-linearly independent.*

Next, we shall explore a smoothing method with a simple form for problem (1.1). Notice that the complementarity constraints

$$a \geq 0, b \geq 0, ab = 0 \tag{2.2}$$

is equivalent to

$$\min\{a, b\} = 0. \tag{2.3}$$

Furthermore, it can be shown that (2.3) is equivalent to

$$a - [a - b]_+ = 0, \tag{2.4}$$

where $[a - b]_+ = \max\{a - b, 0\}$.

And it follows from [3], the nonsmooth function $[x]_+ = \max\{x, 0\}$ can be approximated by the following smooth function:

$$p(x, \alpha) = x + \alpha \ln(1 + e^{-\frac{x}{\alpha}}), \tag{2.5}$$

where $x \in \Re$, $\alpha > 0$ is a smooth parameter, and $p(x, \alpha)$ is just the integral function of Sigmoid function $\frac{1}{1+e^{-\frac{x}{\alpha}}}$, which is commonly used in neural networks.

The following Lemma unfolds some nice properties of $p(x, \alpha)$ (see Ref. [3]).

Lemma 2.7. *For $p(x,\alpha)$ defined by (2.5), it holds that*

(1) *$p(x,\alpha)$ is increasing with respect to α, and $[x]_+ \le p(x,\alpha) \le [x]_+ + \alpha \ln 2$;*

(2) *$[x]_+ \approx p(x,\alpha)$ as $\alpha \to 0$.*

Thus, it follows from Lemma 2.7 that for $\alpha > 0$ being small enough, (2.4) can be approximated by the following formulation:

$$a - \max\{a-b, 0\} = 0 \approx a - (a-b) - \alpha \ln(1 + e^{-\frac{a-b}{\alpha}}) = 0,$$

which means that the complementary constraint condition (2.2) can be approximately reformulated as

$$b - \alpha \ln(1 + e^{(-\frac{1}{\alpha})(a-b)}) = 0.$$

Define function $\Phi(x,\alpha) : \Re^n \times \Re^+ \to \Re^l$ as

$$\Phi(x,\alpha) = \begin{pmatrix} \Phi_1(x,\alpha) \\ \vdots \\ \Phi_l(x,\alpha) \end{pmatrix},$$

where

$$\Phi_i(x,\alpha) = H_i(x) - \alpha \ln(1 + e^{-\frac{1}{\alpha}(G_i(x) - H_i(x))}, \quad \text{for } i = 1,2,\ldots,l. \tag{2.6}$$

Then, $G(x) \ge 0$, $H(x) \ge 0$, and $G(x)^T H(x) = 0$ in problem (1.1) can be approximated by equality $\Phi(x,\alpha) = 0$ for $\alpha > 0$ being small enough. Thus, for $\alpha > 0$ being small enough, problem (1.1) can be reformulated as the smooth approximation problem of the form:

$$\begin{aligned} &\min f(x) \\ &\text{s.t. } g(x) \le 0, \\ &\qquad h(x) = 0, \\ &\qquad \Phi(x,\alpha) = 0. \end{aligned} \tag{2.7}$$

Notice that an entropic regularization approach is presented in Ref. [2] for mathematical programs with equilibrium constraints, in which the the complementarity constraints is approximated by the equality constraints based on the smoothing function by entropic regularization. And the smoothing function by entropic regularization in Ref. [2] is a smooth approximation for the minimum function in (2.3), which is different from the smooth function in (2.5). Furthermore, notice that the smooth function (2.5) is mentioned as a special case in Ref. [27]. But the different assumptions are adopted in the theoretical analysis. For example, the multiplier with respect to the approximate smooth equality constraint is assumed to be bounded in Ref. [27], which is not needed in the subsequent discussion in this paper. Hence it is desirable to investigate the smoothing method for problem (1.1) based on the approximation problem (2.7) from the different points of view.

Next, we will investigate the properties of $\Phi(x,\alpha)$. Denote the feasible set of problem (2.7) by $\mathcal{F}_\alpha$.

Lemma 2.8. *Let $\Phi_i(x,\alpha)$ $(i = 1,2,\ldots,l)$ be defined by (2.6). Then we have the following conclusions.*

(1) *For any $\alpha > 0$, let $x^\alpha \in \mathcal{F}_\alpha$. Suppose that $x^\alpha \to \bar{x}$ as $\alpha \to 0$. Then $\bar{x} \in \mathcal{F}$.*

(2) *For any $\alpha > 0$, the gradient and Hessian of $\Phi_i(x,\alpha)$ with respect to x are calculated by*

$$\begin{aligned}\nabla_x\Phi_i(x,\alpha) &= \xi_i(x,\alpha)\nabla G_i(x) + \eta_i(x,\alpha)\nabla H_i(x),\\ \nabla_x^2\Phi_i(x,\alpha) &= \xi_i(x,\alpha)\nabla^2 G_i(x) + \eta_i(x,\alpha)\nabla^2 H_i(x)\\ &\quad - \frac{1}{\alpha}\xi_i(x,\alpha)\eta_i(x,\alpha)(\nabla G_i(x)\nabla G_i(x)^T - \nabla H_i(x)\nabla G_i(x)^T)\\ &\quad + \frac{1}{\alpha}\xi_i(x,\alpha)\eta_i(x,\alpha)(\nabla G_i(x)\nabla H_i(x)^T - \nabla H_i(x)\nabla H_i(x)^T),\end{aligned}$$

where

$$\xi_i(x,\alpha) = \frac{e^{(-\frac{1}{\alpha})(G_i(x)-H_i(x))}}{1+e^{(-\frac{1}{\alpha})(G_i(x)-H_i(x))}}, \eta_i(x,\alpha) = \frac{1}{1+e^{(-\frac{1}{\alpha})(G_i(x)-H_i(x))}}, \tag{2.8}$$

$\xi_i(x,\alpha) \in (0,1)$, $\eta_i(x,\alpha) \in (0,1)$, and $\xi_i(x,\alpha) + \eta_i(x,\alpha) = 1$. And for $i = 1,2,\ldots,l$, $\xi_i(x,\alpha)$ and $\eta_i(x,\alpha)$ are continuously differentiable with respect to x for any $\alpha > 0$.

(3) *Suppose that $\{x_k\}$ is a sequence in $\Re^n$ such that $x_k \to \bar{x} \in \mathcal{F}$ as $k \to \infty$. Then there exists a sequence $\{\alpha_k\}$ approaching zero in $\Re_+$ such that*

$$\lim_{\substack{x_k\to\bar{x}\\ \alpha_k\to 0}} \nabla_x\Phi_i(x_k,\alpha_k) = \bar{\xi}_i(\bar{x})\nabla G_i(\bar{x}) + \bar{\eta}_i(\bar{x})\nabla H_i(\bar{x}),$$

where $\bar{\xi}_i(\bar{x}) = \lim_{\substack{x_k\to\bar{x}\\ \alpha_k\to 0}} \xi_i(x_k,\alpha_k) \in [0,1]$, $\bar{\eta}_i(\bar{x}) = \lim_{\substack{x_k\to\bar{x}\\ \alpha_k\to 0}} \eta_i(x_k,\alpha_k) \in [0,1]$, $\bar{\xi}_i(\bar{x}) + \bar{\eta}_i(\bar{x}) = 1$, $\nabla G_i(\bar{x}) = \lim_{x_k\to\bar{x}} \nabla G_i(x_k)$ and $\nabla H_i(\bar{x}) = \lim_{x_k\to\bar{x}} \nabla H_i(x_k)$. In particular, if $i \in I_{0+}(\bar{x})$, then $\bar{\xi}_i(\bar{x}) = 1$, and $\bar{\eta}_i(\bar{x}) = 0$; If $i \in I_{+0}(\bar{x})$, then $\bar{\xi}_i(\bar{x}) = 0$, and $\bar{\eta}_i(\bar{x}) = 1$; And if $i \in I_{00}(\bar{x})$, then $\bar{\xi}_i(\bar{x}) \in (0,1)$, and $\bar{\eta}_i(\bar{x}) \in (0,1)$.

Proof. (1) By Lemma 2.7, and the definition and continuity of $\Phi(x,\alpha)$, the first conclusion is true.

(2) According to (2.6), it follows that

$$\begin{aligned}\nabla_x\Phi_i(x,\alpha) &= \nabla H_i(x) - \alpha\cdot\frac{e^{-\frac{1}{\alpha}(G_i(x)-H_i(x))}}{1+e^{-\frac{1}{\alpha}(G_i(x)-H_i(x))}}\cdot\left(-\frac{1}{\alpha}\right)(\nabla G_i(x) - \nabla H_i(x))\\ &= \frac{e^{-\frac{1}{\alpha}(G_i(x)-H_i(x))}}{1+e^{-\frac{1}{\alpha}(G_i(x)-H_i(x))}}\nabla G_i(x) + \left(1 - \frac{e^{-\frac{1}{\alpha}(G_i(x)-H_i(x))}}{1+e^{-\frac{1}{\alpha}(G_i(x)-H_i(x))}}\right)\nabla H_i(x)\\ &= \xi_i(x,\alpha)\nabla G_i(x) + \eta_i(x,\alpha)\nabla H_i(x),\end{aligned}$$

where $\xi_i(x,\alpha) = \frac{e^{(-\frac{1}{\alpha})(G_i(x)-H_i(x))}}{1+e^{(-\frac{1}{\alpha})(G_i(x)-H_i(x))}}$, $\eta_i(x,\alpha) = 1 - \frac{e^{(-\frac{1}{\alpha})(G_i(x)-H_i(x))}}{1+e^{(-\frac{1}{\alpha})(G_i(x)-H_i(x))}}$, and we have that $\xi_i(x,\alpha) \in (0,1)$, $\eta_i(x,\alpha) \in (0,1)$, and $\xi_i(x,\alpha) + \eta_i(x,\alpha) = 1$.

Moreover, we have that

$$\begin{aligned}\nabla_x^2\Phi_i(x,\alpha) &= \nabla\xi_i(x,\alpha)\nabla G_i(x)^T + \xi_i(x,\alpha)\nabla^2 G_i(x)\\ &\quad + \nabla\eta_i(x,\alpha)\nabla H_i(x)^T + \eta_i(x,\alpha)\nabla^2 H_i(x),\end{aligned}$$

where $\nabla\xi_i(x,\alpha) = -\frac{1}{\alpha}\xi_i(x,\alpha)\eta_i(x,\alpha)(\nabla G_i(x) - \nabla H_i(x))$, $\nabla\eta_i(x,\alpha) = \frac{1}{\alpha}\xi_i(x,\alpha)\eta_i(x,\alpha)$ $(\nabla G_i(x) - \nabla H_i(x))$. That is,

$$\begin{aligned}\nabla_x^2\Phi_i(x,\alpha) &= \xi_i(x,\alpha)\nabla^2 G_i(x) + \eta_i(x,\alpha)\nabla^2 H_i(x)\\ &\quad - \frac{1}{\alpha}\xi_i(x,\alpha)\eta_i(x,\alpha)(\nabla G_i(x)\nabla G_i(x)^T - \nabla H_i(x)\nabla G_i(x)^T)\\ &\quad + \frac{1}{\alpha}\xi_i(x,\alpha)\eta_i(x,\alpha)(\nabla G_i(x)\nabla H_i(x) - \nabla H_i(x)\nabla H_i(x)^T).\end{aligned}$$

(3) It follows from the conclusion (2) that

$$\xi_i(x_k,\alpha) = \frac{e^{(-\frac{1}{\alpha})(G_i(x_k)-H_i(x_k))}}{1+e^{(-\frac{1}{\alpha})(G_i(x_k)-H_i(x_k))}} = \frac{1}{\frac{1}{e^{(-\frac{1}{\alpha})(G_i(x_k)-H_i(x_k))}}+1},$$

$$\eta_i(x_k,\alpha) = 1-\frac{e^{(-\frac{1}{\alpha})(G_i(x_k)-H_i(x_k))}}{1+e^{(-\frac{1}{\alpha})(G_i(x_k)-H_i(x_k))}} = \frac{1}{1+e^{(-\frac{1}{\alpha})(G_i(x_k)-H_i(x_k))}}.$$

If $i \in I_{0+}(\bar{x})$, that is, $G_i(\bar{x}) = 0$, $H_i(\bar{x}) > 0$, we conclude that $G_i(x_k) < H_i(x_k)$ for all k sufficiently large. Thus, for any sequence $\{\alpha_k\}$ approaching zero in $\Re_+$, it holds that $\lim_{\substack{x_k\to\bar{x}\\ \alpha_k\to 0}} \xi_i(x_k,\alpha_k) = 1$ and $\lim_{\substack{x_k\to\bar{x}\\ \alpha_k\to 0}} \eta_i(x_k,\alpha_k) = 0$. Let $\bar{\xi}_i(\bar{x}) = \lim_{\substack{x_k\to\bar{x}\\ \alpha_k\to 0}} \xi_i(x_k,\alpha_k)$ and $\bar{\eta}_i(\bar{x}) = \lim_{\substack{x_k\to\bar{x}\\ \alpha_k\to 0}} \eta_i(x_k,\alpha_k)$. Hence the conclusion holds. Similarly, it can be proven that the conclusion is true for $i \in I_{+0}(\bar{x})$.

Since $\xi_i(x_k,\alpha) \in (0,1)$, $\eta_i(x_k,\alpha) \in (0,1)$ and $\xi_i(x_k,\alpha)+\eta_i(x_k,\alpha) = 1$ for $i \in I_{00}(\bar{x})$, there exists a sequence $\{\alpha_k\}$ approaching zero in $\Re_+$ such that $\lim_{\substack{x_k\to\bar{x}\\ \alpha_k\to 0}} \xi_i(x_k,\alpha_k) \in (0,1)$, $\lim_{\substack{x_k\to\bar{x}\\ \alpha_k\to 0}} \eta_i(x_k,\alpha_k) \in (0,1)$ and $\lim_{\substack{x_k\to\bar{x}\\ \alpha_k\to 0}} \xi_i(x_k,\alpha_k) + \lim_{\substack{x_k\to\bar{x}\\ \alpha_k\to 0}} \eta_i(x_k,\alpha_k) = 1$. Let $\bar{\xi}_i(\bar{x}) = \lim_{\substack{x_k\to\bar{x}\\ \alpha_k\to 0}} \xi_i(x_k,\alpha_k)$ and $\bar{\eta}_i(\bar{x}) = \lim_{\substack{x_k\to\bar{x}\\ \alpha_k\to 0}} \eta_i(x_k,\alpha_k)$. The proof is completed. □

Remark 2.9. Lemma 2.8 plays a significant role in the subsequent theoretical analysis, and the the upper level strict complementarity and the asymptotically weakly nondegenerate condition are not needed in the proof of Lemma 2.8.

3 Convergence Analysis

In this section, we discuss the convergence of the smoothing method, in which it is proven that the sequence of the KKT solutions to problem (2.7) converges to a C-stationary point of problem (1.1) as the smooth parameter tends to zero under the mild assumptions. And, the accumulation point is further proven to be an S-stationary point under the weak second-order necessary condition.

Theorem 3.1. *Let $\alpha = \alpha_k > 0$ in problem (2.7) with $\alpha_k \to 0$ $(k \to \infty)$. Suppose that $\{x_k\}$ is a sequence of KKT solutions to problem (2.7). If $\bar{x} \in \mathcal{F}$ is an accumulation point of the sequence $\{x_k\}$ and MPCC-MFCQ holds at $\bar{x}$, then $\bar{x}$ is a C-stationary point of problem (1.1) as α_k tends to zero.*

Proof. Since it follows from the assumption that the sequence $\{x_k\}$ is a KKT solution to problem (2.7), there exist Lagrange multipliers of λ^k, μ^k, and γ^k such that

$$\nabla f(x_k) + \sum_{i=1}^{m}\lambda_i^k\nabla g_i(x_k) + \sum_{i=1}^{p}\mu_i^k\nabla h_i(x_k) + \sum_{i=1}^{l}\gamma_i^k\nabla_x\Phi_i(x_k,\alpha_k) = 0, \tag{3.1}$$

$$\lambda_i^k \geq 0, \lambda_i^k g_i(x_k) = 0, i = 1,2,\ldots,m, \tag{3.2}$$

$$g_i(x_k) \leq 0, i = 1,\ldots,m, h_i(x_k) = 0, i = 1,\ldots,p, \Phi_i(x_k,\alpha_k) = 0, i = 1,\ldots,l. \tag{3.3}$$

Moreover, (3.1) can be rewritten as

$$\begin{aligned}-\nabla f(x_k) &= \sum_{i=1}^{m}\lambda_i^k\nabla g_i(x_k)+\sum_{i=1}^{p}\mu_i^k\nabla h_i(x_k)+\sum_{i=1}^{l}\gamma_i^k\nabla_x\Phi_i(x_k,\alpha_k)\\ &= \sum_{i=1}^{m}\lambda_i^k\nabla g_i(x_k)+\sum_{i=1}^{p}\mu_i^k\nabla h_i(x_k)+\sum_{i=1}^{l}\gamma_i^k\xi_i(x_k,\alpha_k)\nabla G_i(x_k) \qquad (3.4)\\ &\quad+\sum_{i=1}^{l}\gamma_i^k\eta_i(x_k,\alpha_k)\nabla H_i(x_k).\end{aligned}$$

Define $u_i^k=-\gamma_i^k\xi_y(x_k,\alpha_k)$, $v_i^k=-\gamma_i^k\eta_i(x_k,\alpha_k)$, then (3.4) is equivalent to

$$-\nabla f(x_k)=\sum_{i=1}^{m}\lambda_i^k\nabla g_i(x_k)+\sum_{i=1}^{p}\mu_i^k\nabla h_i(x_k)-\sum_{i=1}^{l}u_i^k\nabla G_i(x_k)-\sum_{i=1}^{l}v_i^k\nabla H_i(x_k). \qquad (3.5)$$

Next we prove that the sequence $\{(\lambda_i^k,\mu_i^k,u_i^k,v_i^k)\}$ associated with (3.5) is bounded. Suppose that this sequence is unbounded, then there exists a subset K such that for any $k\in K$, it holds that

$$\frac{(\lambda^k,\mu^k,u^k,v^k)}{\|(\lambda^k,\mu^k,u^k,v^k)\|}\to(\lambda',\mu',u',v')\ (k\to\infty).$$

If $i\notin I_g(\bar{x})$, which implies $g_i(\bar{x})<0$, by $x_k\to\bar{x}$ as $k\to\infty$ and the continuity of $g(x)$, we get $g_i(x_k)<0$ as k is sufficiently large. And according to (3.2), we know that $\lambda_i^k=0$ for $i\notin I_g(\bar{x})$. Thus, $\lambda_i'=0$. And $\lambda_i'\geq 0$ for $i\in I_g(\bar{x})$. Moreover, by the definitions of u_i^k and v_i^k, and conclusion (3) of Lemma 2.8, ${u_i}'=0$ for $i\in I_{+0}(\bar{x})$, and ${v_i}'=0$ for $i\in I_{0+}(\bar{x})$. We divide both sides of (3.5)) by $\left\|(\lambda^k,\mu^k,u^k,v^k)\right\|$, and let $k\to\infty$, then we have

$$\begin{aligned}&\sum_{i\in I_g(\bar{x})}{\lambda_i}'\nabla g_i(\bar{x})+\sum_{i=1}^{p}{\mu_i}'\nabla h_i(\bar{x})\\ &-\sum_{i\in I_{00}(\bar{x})\cup I_{0+}(\bar{x})}{u_i}'\nabla G_i(\bar{x})-\sum_{i\in I_{00}(\bar{x})\cup I_{+0}(\bar{x})}{v_i}'\nabla H_i(\bar{x})=0,\end{aligned}$$

in which $(\lambda',\mu',u',v')\neq 0$, contradicting Lemma 2.6 from the fact that MPCC-MFCQ holds at $\bar{x}$. Thus, the sequence of $\{(\lambda_i^k,\mu_i^k,u_i^k,v_i^k)\}$ is bounded.

Without loss of generality, we now suppose that the sequence $(\lambda^k,\mu^k,u^k,v^k)$ converges to $(\bar{\lambda},\bar{\mu},\bar{u},\bar{v})$ as $k\to\infty$. By (3.2), we have $\bar{\lambda}_i\geq 0$, $\bar{\lambda}_i g_i(\bar{x})=0$, $i=1,2,\dots,m$, which means that $\bar{\lambda}_i=0$ for $i\notin I_g(\bar{x})$. As $i\in I_{+0}(\bar{x})$, it follows from the conclusion (3) of Lemma 2.8 that $\lim_{k\to\infty}\xi_i(x_k,\alpha_k)=0$, which means $\bar{u}_i=0$. Similarly, as $i\in I_{0+}(\bar{x})$, we obtain $\bar{v}_i=0$. Since $g(x)$, $h(x)$, $G(x)$, and $H(x)$ are continuously differentiable functions, setting $k\to\infty$ in (3.5), we have

$$\begin{aligned}&\nabla f(\bar{x})+\sum_{i\in I_g(\bar{x})}\bar{\lambda}_i\nabla g_i(\bar{x})+\sum_{i=1}^{p}\bar{\mu}_i\nabla h_i(\bar{x})-\sum_{i\in I_{00}(\bar{x})\cup I_{0+}(\bar{x})}\bar{u}_i\nabla G_i(\bar{x})\\ &-\sum_{i\in I_{00}(\bar{x})\cup I_{+0}(\bar{x})}\bar{v}_i\nabla H_i(\bar{x})=0.\end{aligned}$$

Since $u_i^k v_i^k = (\gamma_i^k)^2 \xi_i(x_k, \alpha_k)\eta_i(x_k, \alpha_k)$, and taking into account the conclusion (3) of Lemma 2.8 that $\bar{\xi}_i(\bar{x}) \in (0,1)$ and $\bar{\eta}_i(\bar{x}) \in (0,1)$ for $i \in I_{00}(\bar{x})$, we obtain

$$\bar{u}_i \bar{v}_i = \bar{\gamma}_i^2 \bar{\xi}_i(\bar{x})\bar{\eta}_i(\bar{x}) \geq 0, \quad i \in I_{00}(\bar{x}).$$

Therefore, $\bar{x}$ is a C-stationary point of problem (1.1). □

We now further explore the property of the accumulation point $\bar{x}$ in Theorem 3.1. Let the Lagrange function of problem (2.7) be defined below:

$$\bar{L}_\alpha(x, \lambda, \mu, \gamma) = f(x) + \lambda^T g(x) + \mu^T h(x) + \gamma^T \Phi(x, \alpha),$$

where $\lambda \in \Re^m$, $\mu \in \Re^p$, and $\gamma \in \Re^l$ are Lagrange multipliers. By Lemma 2.8, we get

$$\begin{aligned} \nabla_x \bar{L}_\alpha(x, \lambda, \mu, \gamma) = {} & \nabla f(x) + \lambda^T \nabla g(x) + \mu^T \nabla h(x) \\ & + \sum_{i=1}^{l} \gamma_i \xi_i(x, \alpha) \nabla G_i(x) + \sum_{i=1}^{l} \gamma_i \eta_i(x, \alpha) \nabla H_i(x), \end{aligned} \tag{3.6}$$

$$\begin{aligned} \nabla_x^2 \bar{L}_\alpha(x, \lambda, \mu, \gamma) = {} & \nabla^2 f(x) + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x)^T + \sum_{i=1}^{p} \mu_i \nabla^2 h_i(x) \\ & + \sum_{i=1}^{l} \gamma_i \xi_i(x, \alpha) \nabla^2 G_i(x) + \sum_{i=1}^{l} \gamma_i \eta_i(x, \alpha) \nabla^2 H_i(x) \\ & - \frac{1}{\alpha} \sum_{i=1}^{l} \gamma_i \xi_i(x, \alpha) \eta_i(x, \alpha) M_i(x) M_i(x)^T, \end{aligned} \tag{3.7}$$

where $\xi_i(x, \alpha)$ and $\eta_i(x, \alpha)$ $(i = 1, 2, \ldots, l)$ are defined in Lemma 2.8, and $M_i(x) = \nabla G_i(x) - \nabla H_i(x)$, $i = 1, 2, \ldots, l$.

Definition 3.2. Let $\tilde{x}$ be a local optimal solution of the approximation problem (2.7). The approximate problem (2.7) is said to satisfy the weak second-order necessary condition at $\tilde{x}$ if there exist Lagrange multipliers $\lambda \in \Re_+^m$, $\mu \in \Re^p$, and $\gamma \in \Re^l$ such that $\nabla_x \bar{L}_\alpha(\tilde{x}, \lambda, \mu, \gamma) = 0$ and $d^T \nabla_x^2 \bar{L}_\alpha(\tilde{x}, \lambda, \mu, \gamma) d \geq 0$ for $d \in C_\alpha(\hat{x})$, where $C_\alpha(\tilde{x}) = \{d \in \Re^n | \nabla g_i(\tilde{x})^T d = 0, i \in I_g(\tilde{x}), \nabla h_i(\tilde{x})^T d = 0, i = 1, 2, \ldots, p, \nabla_x \Phi_i(\tilde{x}, \alpha)^T d = 0, i = 1, 2, \ldots, l\}$.

Theorem 3.3. *Let $\alpha = \alpha_k > 0$ in problem (2.7) with $\alpha_k \to 0$ $(k \to \infty)$. Suppose that $\{x_k\}$ is a sequence of KKT solutions to problem (2.7), and the problem (2.7) satisfies the weak second-order necessary condition at x_k. If $\bar{x} \in \mathcal{F}$ is an accumulation point of the sequence $\{x_k\}$ and MPCC-LICQ holds at $\bar{x}$, then $\bar{x}$ is an S-stationary point of problem (1.1) as α_k tends to* 0.

Proof. It follows from Theorem 3.1 that $\bar{x}$ is a C-stationary point of problem (1.1) since MPCC-LICQ implies MPCC-MFCQ. Assume that $\bar{x}$ is not an S-stationary point of problem (1.1), there must exist $i_0 \in I_{00}(\bar{x})$ such that

$$\bar{u}_{i_0} = \lim_{k \to \infty} (-\gamma_{i_0}^k \xi_{i_0}(x_k, \alpha_k)) < 0,$$

$$\bar{v}_{i_0} = \lim_{k \to \infty} (-\gamma_{i_0}^k \eta_{i_0}(x_k, \alpha_k)) < 0.$$

Since MPCC-LICQ holds at $\bar{x}$, it follows from the continuity of $\nabla g_i(x)$ $(i = 1, 2, \ldots, m)$, $\nabla h_i(x)$ $(i = 1, 2, \ldots, p)$, $\nabla G_i(x)$ $(i = 1, 2, \ldots, l)$, and $\nabla H_i(x)$ $(i = 1, 2, \ldots, l)$ that the vectors of $\{\nabla g_i(x_k), i \in I_g(x_k)\}$, $\{\nabla h_i(x_k), i = 1, 2, \ldots, p\}$, $\{\nabla G_i(x_k), i \in I_{00}(\bar{x}) \cup I_{0+}(\bar{x})\}$ and $\{\nabla H_i(x_k), i \in I_{00}(\bar{x}) \cup I_{+0}(\bar{x})\}$ are linear independent for k being large enough, where $I_g(x_k) \subseteq I_g(\bar{x})$. Thus, there exists a bounded sequence $\{d_k\}$ such that

$$\begin{aligned}
&\nabla g_i(x_k)^T d_k = 0, i \in I_g(x_k),\\
&\nabla h_i(x_k)^T d_k = 0, i = 1, 2, \ldots, p,\\
&\nabla G_i(x_k)^T d_k = 0, i \in I_{00}(\bar{x}) \cup I_{0+}(\bar{x}) \backslash \{i_0\},\\
&\nabla H_i(x_k)^T d_k = 0, i \in I_{00}(\bar{x}) \cup I_{+0}(\bar{x}) \backslash \{i_0\},\\
&\nabla G_{i_0}(x_k)^T d_k = \eta_{i_0}(x_k, \alpha_k),\\
&\nabla H_{i_0}(x_k)^T d_k = -\xi_{i_0}(x_k, \alpha_k).
\end{aligned}$$

Next we prove that $d_k \in C_{\alpha_k}(x_k)$, where,

$$\begin{aligned}
C_{\alpha_k}(x_k) = \{&d_k \in \Re^n | \nabla g_i(x_k)^T d_k = 0, i \in I_g(x_k), \nabla h_i(x_k)^T d_k = 0, i = 1, 2, \ldots, p,\\
&\nabla_x \Phi_i(x_k, \alpha_k)^T d_k = 0, i = 1, 2, \ldots, l\}.
\end{aligned}$$

From Lemma 2.8, we have

$$\nabla_x \Phi_i(x_k, \alpha_k)^T d_k = \xi_i(x_k, \alpha_k) \nabla G_i(x_k)^T d_k + \eta_i(x_k, \alpha_k) \nabla H_i(x_k)^T d_k.$$

Hence,

$$\begin{aligned}
\nabla_x \Phi_{i_0}(x_k, \alpha_k)^T d_k &= \xi_{i_0}(x_k, \alpha_k)\eta_{i_0}(x_k, \alpha_k) - \eta_{i_0}(x_k, \alpha_k)\xi_{i_0}(x_k, \alpha_k)\\
&= 0.
\end{aligned}$$

Moreover, we have $\nabla_x \Phi_i(x_k, \alpha_k)^T d_k = 0, i = 1, 2, \ldots, l$ $(i \neq i_0)$, $\nabla g_i(x_k)^T d_k = 0, i \in I_g(x_k)$, and $\nabla h_i(x_k)^T d_k = 0, i = 1, 2, \ldots, p$ for k being large enough. Hence $d_k \in C_{\alpha_k}(x_k)$ for k being large enough. Then from the weak second-order necessary condition, we obtain

$$d_k^T \nabla_x^2 \bar{L}_{\alpha_k}(x_k, \lambda_k, \mu_k, \gamma_k) d_k \geq 0. \tag{3.8}$$

That is,

$$\begin{aligned}
&d_k^T \nabla^2 f(x_k) d_k + \sum_{i \in I_g(x_k)} \lambda_i^k d_k^T \nabla^2 g_i(x_k) d_k + \sum_{i=1}^{p} \mu_i^k d_k^T \nabla^2 h_i(x_k) d_k\\
&+ \sum_{i=1}^{l} \gamma_i^k d_k^T \nabla_x^2 \Phi_i(x_k, \alpha_k) d_k \geq 0.
\end{aligned}$$

From Lemma 2.8, we have

$$\begin{aligned}
d_k^T \nabla_x^2 \Phi_i(x_k, \alpha_k) d_k =& \xi_i(x_k, \alpha_k) d_k^T \nabla^2 G_i(x_k) d_k + \eta_i(x_k, \alpha_k) d_k^T \nabla^2 H_i(x) d_k\\
&- \frac{1}{\alpha_k} \xi_i(x_k, \alpha_k)\eta_i(x_k, \alpha_k) d_k^T M_i(x_k) M_i(x_k)^T d_k.
\end{aligned}$$

For $i \neq i_0$,

$$d_k^T \nabla_x^2 \Phi_i(x_k, \alpha_k) d_k = \xi_i(x_k, \alpha_k) d_k^T \nabla^2 G_i(x_k) d_k + \eta_i(x_k, \alpha_k) d_k^T \nabla^2 H_i(x) d_k.$$

For $i = i_0$,

$$d_k^T \nabla_x^2 \Phi_i(x_k, \alpha_k) d_k = \xi_{i_0}(x_k, \alpha_k) d_k^T \nabla^2 G_{i_0}(x_k) d_k + \eta_{i_0}(x_k, \alpha_k) d_k^T \nabla^2 H_{i_0}(x) d_k - \frac{1}{\alpha_k} \xi_{i_0}(x_k, \alpha_k) \eta_{i_0}(x_k, \alpha_k).$$

Thus,

$$\begin{aligned} & d_k^T \nabla_x^2 \bar{L}_{\alpha_k}(x_k, \lambda_k, \mu_k, \gamma_k) d_k = \\ & d_k^T \nabla^2 f(x_k) d_k + \sum_{i \in I_g(x_k)} \lambda_i^k d_k^T \nabla^2 g_i(x_k) d_k + \sum_{i=1}^{p} \mu_i^k d_k^T \nabla^2 h_i(x_k) d_k \\ & + \sum_{i=1}^{l} \gamma_i^k \left(\xi_i(x_k, \alpha_k) d_k^T \nabla^2 G_i(x_k) d_k + \eta_i(x_k, \alpha_k) d_k^T \nabla^2 H_i(x_k) d_k \right) \\ & - \frac{1}{\alpha_k} \gamma_{i_0}^k \xi_{i_0}(x_k, \alpha_k) \eta_{i_0}(x_k, \alpha_k). \end{aligned} \tag{3.9}$$

Since d_k is bounded, $f(x)$, $g(x)$, $h(x)$, $G(x)$ and $H(x)$ are twice continuously differentiable, and the Lagrange multipliers are bounded (see the proof in Theorem 3.1), the first five terms of (3.9) are bounded for k being large enough. Since $\lim_{k\to\infty} -\gamma_{i_0}^k \xi_{i_0}(x_k, \alpha_k) = \lim_{k\to\infty} u_{i_0}^k = \bar{u}_{i_0} < 0$, and $\xi_{i_0}(x_k, \alpha_k) \in (0, 1)$, one has that $\lim_{k\to\infty} \gamma_{i_0}^k = \bar{\gamma}_{i_0} > 0$. According to the conclusion (iii) of Lemma 2.8, $\lim_{k\to\infty} \xi_{i_0}(x_k, \alpha_k) = \bar{\xi}_{i_0}(\bar{x})$, $\lim_{k\to\infty} \eta_{i_0}(x_k, \alpha_k) = \bar{\eta}_{i_0}(\bar{x})$, and $\bar{\xi}_{i_0}(\bar{x}), \bar{\eta}_{i_0}(\bar{x}) \in (0, 1)$, it follows that

$$-\frac{1}{\alpha_k} \gamma_{i_0}^k \xi_{i_0}(x_k, \alpha_k) \eta_{i_0}(x_k, \alpha_k) \to -\infty,$$

which is a contradiction with (3.8). Then we get $\bar{u}_{i_0} \geq 0$. Similarly, it follows that $\bar{v}_{i_0} \geq 0$. Therefore $\bar{x}$ is an S-stationary point of problem (1.1). □

Remark 3.4. From Theorem 3.1 and Theorem 3.3, we know that the KKT solution to the smoothing problem (2.7) is vital in establishing the convergence of the smoothing method. Hence it is important to analyze the existence of KKT solution to problem (2.7) under some mild assumptions on the original problem (1.1), on which we will make detailed discussion in the next section.

4 Discussion of KKT Solutions to Smooth Approximation Problem

Although the convergence analyses of some current smoothing methods are discussed, the assumption is directly made that the sequence of the KKT solutions to their corresponding smooth approximation problems is existent. A natural question is that whether or not the assumption is reasonable. That is, the relationship between the assumption and the original problem (1.1) should be elaborated. For this consideration, this section mainly focuses on exploring the existence of KKT solutions to the smooth approximation problem (2.7) under some given assumptions on problem (1.1). In fact, under these assumptions, we characterize the linearly independent constraint qualification, the KKT condition and the second-order sufficient condition of smooth approximation problem, which guarantee the existence of the KKT solution to the smooth approximation problem (2.7).

Suppose that $\hat{x} \in \mathcal{F}$ and the further assumption on problem (1.1) at $\hat{x}$ is stated as follows.

(A1) For any $z \in C(\hat{x})$, it holds that

$$z^T \nabla_x^2 L(\hat{x}, \hat{\lambda}, \hat{\mu}, \hat{u}, \hat{v}) z > 0,$$

where $C(\hat{x}) = \{z \in R^n | \nabla g_i(\hat{x})^T z = 0, i \in I_g(\hat{x}), \nabla h_i(\hat{x})^T z = 0, i = 1, 2, \ldots, p, \nabla G_i(\hat{x})^T z = 0, i \in I_{0+}(\hat{x}) \cup I_{00}(\hat{x}), \nabla H_i(\hat{x})^T z = 0, i \in I_{+0}(\hat{x}) \cup I_{00}(\hat{x})\}$.

Next, we will investigate the LICQ, the KKT condition and the second-order sufficient condition of the smooth approximation problem (2.7).

Theorem 4.1. *Suppose that problem (1.1) satisfies MPCC-LICQ and assumption (A1) at $\hat{x}$, and $\hat{x}$ is a W-stationary point of problem (1.1) associated with Lagrange multipliers of $\hat{\lambda}$, $\hat{\mu}$, $\hat{u}$, and $\hat{v}$. Let $x^\alpha \in \mathcal{F}_\alpha$ and $x^\alpha \to \hat{x}$ as $\alpha \to 0$. Then there exists a sufficiently small constant $\hat{\alpha} > 0$ such that for any $x^\alpha \in \mathcal{F}_\alpha$ whenever $\alpha \in (0, \hat{\alpha})$, the following conclusions are true.*

(i) *The vectors of $\{\nabla g_i(x^\alpha),\ i \in I_g(x^\alpha)\}$, $\{\nabla h_i(x^\alpha),\ i = 1, 2, \ldots, p)\}$, and $\{\nabla_x \Phi_i(x^\alpha, \alpha),\ i = 1, 2, \ldots, l\}$ are linearly independent vectors;*

(ii) *There exist $\lambda^\alpha \in \Re^m$, $\mu^\alpha \in \Re^p$ and $\gamma^\alpha \in \Re^l$ such that $\nabla_x \bar{L}_\alpha(x^\alpha, \lambda^\alpha, \mu^\alpha, \gamma^\alpha) = 0$;*

(iii) *Furthermore, suppose that for any $i \in I_{00}(\hat{x})$, $\hat{u}_i \geq 0$ or $\hat{v}_i \geq 0$, and there exists an index $i \in I_{00}(\hat{x})$ such that $\hat{u}_i > 0$ or $\hat{v}_i > 0$. Then for any $0 \neq z^\alpha \in \hat{C}_\alpha(x^\alpha)$, the following inequality holds*

$$(z^\alpha)^T \nabla_x^2 \bar{L}_\alpha(x^\alpha, \lambda^\alpha, \mu^\alpha, \gamma^\alpha) z^\alpha > 0,$$

where $\hat{C}_\alpha(x^\alpha) = \{z \in \Re^n | \nabla g_i(x^\alpha)^T z = 0,\ i \in I_g(x^\alpha),\ \nabla h_i(x^\alpha)^T z = 0,\ i = 1, 2, \ldots, p, \nabla_x \Phi_i(x^\alpha, \alpha)^T z = 0,\ i = 1, 2, \ldots, l\}$.

Proof. We now prove conclusion (i). Firstly, we prove that there exists a constant $\hat{\alpha} > 0$ such that for any $x^\alpha \in \mathcal{F}_\alpha$ whenever $\alpha \in (0, \hat{\alpha})$, the vectors of $\{\nabla g_i(x^\alpha),\ i \in I_g(\hat{x})\}$, $\{\nabla h_i(x^\alpha),\ i = 1, 2, \ldots, p\}$, and $\{\nabla_x \Phi_i(x^\alpha, \alpha),\ i = 1, 2, \ldots, l\}$ are linearly independent. Suppose that such a constant $\hat{\alpha} > 0$ does not exist, then there is a sequence $x^{\alpha_k} \in \mathcal{F}_{\alpha_k}$ with $\alpha_k \to 0$ and $x^{\alpha_k} \to \hat{x}$ as $k \to \infty$ and $\|(a^k, b^k, c^k)\| \equiv 1$ such that

$$\sum_{i \in I_g(\hat{x})} a_i^k \nabla g_i(x^{\alpha_k}) + \sum_{i=1}^{p} b_i^k \nabla h_i(x^{\alpha_k}) + \sum_{i=1}^{l} c_i^k \nabla_x \Phi_i(x^{\alpha_k}, \alpha_k) = 0.$$

That is,

$$\sum_{i \in I_g(\hat{x})} a_i^k \nabla g_i(x^{\alpha_k}) + \sum_{i=1}^{p} b_i^k \nabla h_i(x^{\alpha_k}) + \sum_{i=1}^{l} c_i^k \xi_i(x^{\alpha_k}, \alpha_k) \nabla G_i(x^{\alpha_k}) + \sum_{i=1}^{l} c_i^k \eta_i(x^{\alpha_k}, \alpha_k) \nabla H_i(x^{\alpha_k}) = 0.$$

From the definitions and properties of $\xi_i(x, \alpha)$ and $\eta_i(x, \alpha)$ in Lemma 2.8, we obtain $\lim_{k\to\infty} \xi_i(x^{\alpha_k}, \alpha_k) = 0$ and $\lim_{k\to\infty} \eta_i(x^{\alpha_k}, \alpha_k) = 1$ for $i \in I_{+0}(\hat{x})$; $\lim_{k\to\infty} \eta_i(x^{\alpha_k}, \alpha_k) = 0$ and $\lim_{k\to\infty} \xi_i(x^{\alpha_k}, \alpha_k) = 1$ for $i \in I_{0+}(\hat{x})$; and for $i \in I_{00}(\hat{x})$, $\lim_{k\to\infty} \xi_i(x^{\alpha_k}, \alpha_k) + \lim_{k\to\infty} \eta_i(x^{\alpha_k}, \alpha_k) = 1$.

Without loss of generality, we may assume that $a^k \to a^*$, $b^k \to b^*$, and $c^k \to c^*$ as $k \to \infty$. Thus, we get

$$\sum_{i\in I_g(\hat{x})} a_i^*\nabla g_i(\hat{x}) + \sum_{i=1}^{p} b_i^*\nabla h_i(\hat{x}) + \sum_{i\in I_{0+}(\hat{x})} c_i^*\nabla G_i(\hat{x}) + \sum_{i\in I_{00}(\hat{x})} c_i^*\bar{\xi}_i(\hat{x})\nabla G_i(\hat{x}) + \sum_{i\in I_{+0}(\hat{x})} c_i^*\nabla H_i(\hat{x}) + \sum_{i\in I_{00}(\hat{x})} c_i^*\bar{\eta}_i(\hat{x})\nabla H_i(\hat{x}) = 0,$$

and $\|(a^*, b^*, c^*)\| \equiv 1$, where $\bar{\xi}_i(\hat{x}) = \lim_{k\to\infty} \xi_i(x^{\alpha_k}, \alpha_k)$ and $\bar{\eta}_i(\hat{x}) = \lim_{k\to\infty} \eta_i(x^{\alpha_k}, \alpha_k)$ for $i \in I_{00}(\hat{x})$.

It follows from MPCC-LICQ that $a_i^* = 0$ for $i \in I_g(\hat{x})$, $b_i^* = 0$ for $i = 1, 2, \ldots, p$, $c_i^* = 0$ for $i \in I_{0+}(\hat{x})$, $c_i^* = 0$ for $i \in I_{+0}(\hat{x})$, and $c_i^*\bar{\xi}_i(x) = 0$ and $c_i^*\bar{\eta}_i(x) = 0$ for $i \in I_{00}(\hat{x})$. Since for $i \in I_{00}(\hat{x})$, $\bar{\xi}_i(\hat{x}) + \bar{\eta}_i(\hat{x}) = 1$, it is true that $c_i^* = c_i^*\bar{\xi}_i(x) + c_i^*\bar{\eta}_i(x) = 0$. That is, $a^* = 0$, $b^* = 0$, and $c^* = 0$, which contradicts with $\|(a^*, b^*, c^*)\| \equiv 1$. Hence, there exists a constant $\hat{\alpha} > 0$ such that for any $x^\alpha \in \mathcal{F}_\alpha$ whenever $\alpha \in (0, \hat{\alpha})$, the vectors of $\{\nabla g_i(x^\alpha),\ i \in I_g(\hat{x})\}$, $\{\nabla h_i(x^\alpha),\ i = 1, 2, \ldots, p\}$, and $\{\nabla_x \Phi_i(x^\alpha, \alpha),\ i = 1, 2, \ldots, l\}$ are linearly independent. Moreover, since $I_g(x^\alpha) \subseteq I_g(\hat{x})$, the conclusion (i) holds.

Next we prove conclusion (ii). We now prove that there exist $\bar{\lambda} \in \Re^m$, $\bar{\mu} \in \Re^p$, and $\bar{\gamma} \in \Re^l$ such that $\lim_{\alpha\to 0} \nabla_x \bar{L}_\alpha(x^\alpha, \bar{\lambda}, \bar{\mu}, \bar{\gamma}) = 0$ for any $x^\alpha \in \mathcal{F}_\alpha$. According to (3.6) and Lemma 2.8, for any $(\bar{\lambda}, \bar{\mu}, \bar{\gamma}) \in \Re^m \times \Re^p \times \Re^l$, we obtain

$$\begin{aligned}\lim_{\alpha\to 0} \nabla_x \bar{L}_\alpha(x^\alpha, \bar{\lambda}, \bar{\mu}, \bar{\gamma}) = \nabla f(\hat{x}) &+ \sum_{i=1}^{m} \bar{\lambda}_i \nabla g_i(\hat{x}) + \sum_{i=1}^{p} \bar{\mu}_i \nabla h_i(\hat{x}) \\ &+ \sum_{i\in I_{+0}(\hat{x})} \bar{\gamma}_i\bar{\xi}_i(\hat{x})\nabla G_i(\hat{x}) + \sum_{i\in I_{0+}(\hat{x})} \bar{\gamma}_i\bar{\xi}_i(\hat{x})\nabla G_i(\hat{x}) \\ &+ \sum_{i\in I_{00}(\hat{x})} \bar{\gamma}_i\bar{\xi}_i(\hat{x})\nabla G_i(\hat{x}) + \sum_{i\in I_{+0}(\hat{x})} \bar{\gamma}_i\bar{\eta}_i(\hat{x})\nabla H_i(\hat{x}) \\ &+ \sum_{i\in I_{0+}(\hat{x})} \bar{\gamma}_i\bar{\eta}_i(\hat{x})\nabla H_i(\hat{x}) + \sum_{i\in I_{00}(\hat{x})} \bar{\gamma}_i\bar{\eta}_i(\hat{x})\nabla H_i(\hat{x}),\end{aligned} \tag{4.1}$$

where $\bar{\xi}_i(.)$ and $\bar{\eta}_i(.)$ are defined in Lemma 2.8. By the assumption that $\hat{x}$ is a W-stationary point in which $\hat{\lambda}$, $\hat{\mu}$, $\hat{u}$ and $\hat{v}$ are corresponding Lagrange multipliers , set $\bar{\lambda}_i = \hat{\lambda}_i$ for $i = 1, 2, \ldots, m$, which implies that $\bar{\lambda}_i g_i(\hat{x}) = 0$ and $\bar{\lambda}_i \geq 0$. Similarly, set $\bar{\mu}_i = \hat{\mu}_i$ for $i = 1, 2, \ldots, p$.

For $i \in I_{+0}(\hat{x})$, set $\bar{\gamma}_i = -\hat{v}_i$, which means that $\bar{\gamma}_i\bar{\xi}_i(\hat{x}) = 0$ and $\bar{\gamma}_i\bar{\eta}_i(\hat{x}) = -\hat{v}_i$, since $\bar{\xi}_i(\hat{x}) = 0$ and $\bar{\eta}_i(\hat{x}) = 1$ for $i \in I_{+0}(\hat{x})$. For $i \in I_{0+}(\hat{x})$, set $\bar{\gamma}_i = -\hat{u}_i$, which means that $\bar{\gamma}_i\bar{\xi}_i(\hat{x}) = -\hat{u}_i$ and $\bar{\gamma}_i\bar{\eta}_i(\hat{x}) = 0$, since $\bar{\xi}_i(\hat{x}) = 1$ and $\bar{\eta}_i(\hat{x}) = 0$ for $i \in I_{0+}(\hat{x})$. For $i \in I_{00}(\hat{x})$, set $\bar{\gamma}_i = -\hat{u}_i - \hat{v}_i$, which means that $\bar{\gamma}_i\bar{\xi}_i(\hat{x}) = -\hat{u}_i$ and $\bar{\gamma}_i\bar{\eta}_i(\hat{x}) = -\hat{v}_i$ if $\hat{u}_i\bar{\eta}_i(\hat{x}) = \hat{v}_i\bar{\xi}_i(\hat{x})$, since $\bar{\xi}_i(\hat{x}) + \bar{\eta}_i(\hat{x}) = 1$, $\bar{\xi}_i(\hat{x}) \in (0, 1)$, and $\bar{\eta}_i(\hat{x}) \in (0, 1)$ for $i \in I_{00}(\hat{x})$. Thus, by the assumption that $\hat{x}$ is a W-stationary point of problem (1.1)and (4.1), there exist $\bar{\lambda} \in \Re^m$, $\bar{\mu} \in \Re^p$, and $\bar{\gamma} \in \Re^l$ such that the following holds

$$\begin{aligned}\lim_{\alpha\to 0} \nabla_x \bar{L}_\alpha(x^\alpha, \bar{\lambda}, \bar{\mu}, \bar{\gamma}) &= \nabla f(\hat{x}) + \sum_{i\in I_g(\hat{x})} \hat{\lambda}_i \nabla g_i(\hat{x}) + \sum_{i=1}^{p} \hat{\mu}_i \nabla h_i(\hat{x}) \\ &\quad - \sum_{i\in I_{0+}(\hat{x})\cup I_{00}(\hat{x})} \hat{u}_i \nabla G_i(\hat{x}) - \sum_{i\in I_{+0}(\hat{x})\cup I_{00}(\hat{x})} \hat{v}_i \nabla H_i(\hat{x}) \\ &= \nabla_x L(\hat{x}, \hat{\lambda}, \hat{\mu}, \hat{u}, \hat{v}) = 0.\end{aligned} \tag{4.2}$$

According to (4.2) and MPCC-LICQ, it can be proven that there exists $(\lambda^\alpha, \mu^\alpha, \gamma^\alpha) \in \Re^m \times \Re^p \times \Re^l$ with $(\lambda^\alpha, \mu^\alpha, \gamma^\alpha) \to (\bar{\lambda}, \bar{\mu}, \bar{\gamma})$ as $\alpha \to 0$ such that the conclusion (ii) holds for x^α whenever $\alpha \in (0, \hat{\alpha})$ if $\hat{\alpha}$ is chosen to be small enough.

At last, we prove (iii). From (3.7) and the definition of Lagrange function of problem (1.1), notice that

$$\begin{aligned}\nabla_x^2 \bar{L}_\alpha(x, \lambda, \mu, \gamma) = {} & \nabla_x^2 L(x, \lambda, \mu, -\gamma\xi(x, \alpha), -\gamma\eta(x, \alpha)) \\ & + \frac{1}{\alpha}\sum_{i=1}^{l}(-\gamma_i)\xi_i(x, \alpha)\eta_i(x, \alpha)M_i(x)M_i(x)^T,\end{aligned} \tag{4.3}$$

where

$$\begin{aligned}\gamma\xi(x, \alpha) &= (\gamma_1\xi_1(x, \alpha), \ldots, \gamma_l\xi_l(x, \alpha))^T, \\ \gamma\eta(x, \alpha) &= (\gamma_1\eta_1(x, \alpha), \ldots, \gamma_l\eta_l(x, \alpha))^T.\end{aligned}$$

By the proof of conclusion (ii), taking into account Lemma 2.8, we have

$$\lim_{\alpha\to 0} \nabla_x^2 L(x^\alpha, \lambda^\alpha, \mu^\alpha, -\gamma^\alpha\xi(x, \alpha), -\gamma^\alpha\eta(x, \alpha)) = \nabla_x^2 L(\hat{x}, \hat{\lambda}, \hat{\mu}, \hat{u}, \hat{v}), \tag{4.4}$$

and

$$\begin{aligned}&\lim_{\alpha\to 0}\sum_{i=1}^{l}(-\gamma_i^\alpha)\xi_i(x, \alpha)\eta_i(x, \alpha)M_i(x)M_i(x)^T \\ &= \sum_{i\in I_{00}(\hat{x})}(-\bar{\gamma}_i)\bar{\xi}_i(\hat{x})\bar{\eta}_i(\hat{x})M_i(\hat{x})M_i(\hat{x})^T \\ &= \sum_{i\in I_{00}(\hat{x})}\hat{u}_i\bar{\eta}_i(\hat{x})M_i(\hat{x})M_i(\hat{x})^T(\text{or} \sum_{i\in I_{00}(\hat{x})}\hat{v}_i\bar{\xi}_i(\hat{x})M_i(\hat{x})M_i(\hat{x})^T).\end{aligned} \tag{4.5}$$

By Lemma 2.8, $\lim_{\alpha\to 0}\nabla_x\Phi_i(x^\alpha, \alpha) = \bar{\xi}_i(\hat{x})\nabla G_i(\hat{x}) + \bar{\eta}_i(\hat{x})\nabla H_i(\hat{x})$, where $\bar{\xi}_i(\hat{x})$, and $\bar{\eta}_i(\hat{x})$ are defined by Lemma 2.8. And let $\nabla_x\hat{\Phi}_i(\hat{x}) = \lim_{\alpha\to 0}\nabla_x\Phi_i(x^\alpha, \alpha)$ for $i = 1, 2, \ldots, l$. Then it follows from the properties of $\bar{\xi}_i(\hat{x})$ and $\bar{\eta}_i(\hat{x})$ that

$$\nabla_x\hat{\Phi}_i(\hat{x}) = \nabla G_i(\hat{x}), \text{ for } i \in I_{0+}(\hat{x}), \tag{4.6}$$

$$\nabla_x\hat{\Phi}_i(\hat{x}) = \nabla H_i(\hat{x}), \text{ for } i \in I_{+0}(\hat{x}), \tag{4.7}$$

$$\nabla_x\hat{\Phi}_i(\hat{x}) = \bar{\xi}_i(\hat{x})\nabla G_i(\hat{x}) + \bar{\eta}_i(\hat{x})\nabla H_i(\hat{x}), \text{ for } i \in I_{00}(\hat{x}). \tag{4.8}$$

By conclusion (i), there exists a bounded vector $z^\alpha \in \hat{C}_\alpha(x^\alpha)$, whose components are not all zeros. And let $z^\alpha \to \hat{z}$ as $\alpha \to 0$. For $\nabla_x\Phi_i(x^\alpha, \alpha)^T z^\alpha = 0$ $(i = 1, 2, \ldots, l)$, we conclude from (4.6)-(4.8) that

$$\begin{aligned}&\lim_{\alpha\to 0}\nabla_x\Phi_i(x^\alpha, \alpha)^T z^\alpha = \nabla G_i(\hat{x})^T\hat{z} = 0, \text{ for } i \in I_{0+}(\hat{x}), \\ &\lim_{\alpha\to 0}\nabla_x\Phi_i(x^\alpha, \alpha)^T z^\alpha = \nabla H_i(\hat{x})^T\hat{z} = 0, \text{ for } i \in I_{+0}(\hat{x}), \\ &\lim_{\alpha\to 0}\nabla_x\Phi_i(x^\alpha, \alpha)^T z^\alpha = \bar{\xi}_i(\hat{x})\nabla G_i(\hat{x})^T\hat{z} + \bar{\eta}_i(\hat{x})\nabla H_i(\hat{x})^T\hat{z} = 0, \text{ for } i \in I_{00}(\hat{x}).\end{aligned} \tag{4.9}$$

By (4.9) and $\bar{\xi}_i(\hat{x}) + \bar{\eta}_i(\hat{x}) = 1$, it yields $\nabla G_i(\hat{x})^T\hat{z} = \bar{\eta}_i(\hat{x})M_i(\hat{x})^T\hat{z}$ or $\nabla H_i(\hat{x})^T\hat{z} = -\bar{\xi}_i(\hat{x})M_i(\hat{x})^T\hat{z}$ for $i \in I_{00}(\hat{x})$.

If $\nabla G_i(\hat{x})^T\hat{z} = 0$ for $i \in I_{00}(\hat{x})$, combining with $\bar{\eta}_i(\hat{x}) \in (0,1)$, we obtain $M_i(\hat{x})^T\hat{z} = 0$. Thus, $\nabla H_i(\hat{x})^T\hat{z} = 0$. Hence it follows from (4.4), (4.5) and assumption (A1) that for any $z^\alpha \in \hat{C}_\alpha(x^\alpha)$ with $\nabla G_i(\hat{x})^T\hat{z} = 0$, it holds that

$$\lim_{\alpha\to 0}(z^\alpha)^T\nabla_x^2\bar{L}_\alpha(x^\alpha,\lambda^\alpha,\mu^\alpha,\gamma^\alpha)z^\alpha = \hat{z}^T\nabla_x^2 L(\hat{x},\hat{\lambda},\hat{\mu},\hat{u},\hat{v})\hat{z} > 0.$$

If $\nabla G_i(\hat{x})^T\hat{z} \neq 0$ for $i \in I_{00}(\hat{x})$, then $\bar{\eta}_i(\hat{x})M_i(\hat{x})^T\hat{z} \neq 0$, which implies $M_i(\hat{x})^T\hat{z} \neq 0$, $\nabla H_i(\hat{x})^T\hat{z} = -\bar{\xi}_i(\hat{x})M_i(\hat{x})^T\hat{z} \neq 0$ for $i \in I_{00}(\hat{x})$. Then by (4.4), and assumption of $\hat{u}_i \geq 0, i \in I_{00}(\hat{x})$ and for some $i \in I_{00}(\hat{x})$ such that $\hat{u}_i > 0$, as $\alpha \to 0$, for any $z^\alpha \in \hat{C}_\alpha(x^\alpha)$, one has that

$$\sum_{i\in I_{00}(\hat{x})}\hat{u}_i\hat{\eta}_i z^T M_i(\hat{x})M_i(\hat{x})^T\hat{z} > 0. \tag{4.10}$$

And by (4.4) and the boundedness of $z^\alpha \in \hat{C}_\alpha(x^\alpha)$, $\hat{z}^T\nabla_x^2 L(\hat{x},\hat{\lambda},\hat{\mu},\hat{u},\hat{v})\hat{z}$ is bounded. According to (4.3) and (4.10), we have

$$\lim_{\alpha\to 0}(z^\alpha)^T\nabla_x^2\bar{L}_\alpha(x^\alpha,\lambda^\alpha,\mu^\alpha,\gamma^\alpha)z^\alpha \to +\infty.$$

Based on the above discussion, we can choose $\hat{\alpha}$ to be small enough such that $(z^\alpha)^T\nabla_x^2\bar{L}_\alpha(x^\alpha,\lambda^\alpha,\mu^\alpha,\gamma^\alpha)z^\alpha > 0$ for $0 \neq z^\alpha \in \hat{C}_\alpha(x^\alpha)$ whenever $\alpha \in (0,\hat{\alpha})$. That is, the conclusion (iii) is true. The proof is completed. □

5 Numerical Results

We develop an implementable algorithm to obtain the approximate solution of problem (1.1) by solving problem (2.7) based on the aforementioned theoretical analysis in this section and test the computational efficiency of the discussed smoothing method by solving some typical problems in MacMPEC database (see [12]).

We now present the following algorithm for solving problem (2.7).

Algorithm 5.1.

Step 1 Given an initial point $x_1 \in R^n$. Choose $\alpha_1 > 0, \varepsilon_{\text{stop}}, \beta \in (0,1)$. And set $k := 1$.

Step 2 Solve problem (2.7) with α_k being the current smooth parameter, and obtain its optimum solution x_k.

Step 3 Compute l_k. If $l_k < \varepsilon_{\text{stop}}$, then the algorithm terminates. Otherwise, set $\alpha_{k+1} := \beta\alpha_k$, $x_{k+1} := x_k$, $k := k+1$, and return to Step 2.

Note. In Step 3, $l_k = \max\{\|\max\{g(x_k),0\}\|, \|h(x_k)\|, \|\min\{G(x_k),H(x_k)\}\|\}$. In Algorithm 5.1, l_k is regarded as the termination condition to measure the accuracy of solution (see [6]). If $l_k = 0$, then x_k is a feasible point of problem (1.1), which is also a stationary point of problem (2.7) under some assumption conditions. Furthermore, by Theorem 3.1, x_k must be an approximate optimal solution of problem (1.1), and the closer l_k approaches zero, the higher the accuracy of the solution becomes.

Next, we test the numerical performance of the smoothing method by solving some test problems including the high dimension problems in MacMPEC database [12]. For the sake of comparison, we implement Algorithm 5.1 (named by Algorithm 1), the typical algorithm developed in [7] (named by Algorithm 2) and the typical algorithm developed in [5] (named by Algorithm 3) to solve the same test problems under the same initial conditions. The corresponding computer procedures run in Matlab R2014a with the computer environment of 2.13GHz CPU and 2.00GB memory based operation system of Windows 7. We use the built-in function fmincon to solve the smooth subproblem.

The initial smooth parameter is set as $\alpha_1 = 1$, and $\beta = 0.1$ for the reduction of the smooth parameter. We set the initial parameter $\alpha_1 = 0.001$ for problem bilevel1 in order to get the desired values efficiently. For all the test problems, the initial values of the same kind of problems are the same. The obtained numerical results for low dimensional problems are reported in Table 5.1, in which the name of test problem, the dimension (n, m, l) of test problem, the initial value x_1, the algorithm that is adopted, the optimal value f^* of objective function, the optimal solution x^*, the number k_1 of inner iteration and the number k_2 of outer iterations, and the termination condition l_{k_1} are listed.

Remark 5.2. From Table 5.1, the same optimal values and optimal solutions are obtained by using Algorithm 1, Algorithm 2 and Algorithm 3. But Algorithm 1 requires a less number of iterations for all the test problems than Algorithm 2, and it also requires a less number of iterations for the test problems than Algorithm 3 except for problem desilva. In addition, the values of the termination condition l_k obtained by Algorithm 1 for these test problems are closer to zero than those by Algorithm 3, and the values are also closer to zero than those by Algorithm 2 except for problem bilevel1 and problem gnash16. Hence the numerical results show that the smoothing method explored in this paper is reliable.

Considering that the dimensions of all the solved test problems in Table 5.1 are not more than 20, we attempt to test problems with more than 80 dimensions from MacMPEC database. The numerical results of nine test problems are reported in Table 5.2, which are compared with those by Algorithm 2 and Algorithm 3.
Note. In Table 5.2, repmat(1/12,12,1) in (repmat(1/12,12,1); zeros(75,1)) means that each of the first 12 elements of x_1 is 1/12.

Remark 5.3. Table 5.2 shows that Algorithm 1, Algorithm 2 and Algorithm 3 can obtain the optimal solutions of these test problems. And Algorithm 1 requires a less number of iterations than Algorithm 2 and Algorithm 3 except for problem portf1-i-1 and problem flp4-3. Moreover, the optimal values f^* of these test problems by Algorithm 1 approach zeros well than those by Algorithm 2 and Algorithm 3 except for problem portfl-i-3. Hence the numerical results show that the smoothing method explored in this paper is feasible and promising.

6 Conclusions

In this paper, the original problem (1.1) is reformulated as a standard smooth approximation optimization model based on the smooth integral of the Sigmoid function discussed in [3], in which the complementarity constraints in problem (1.1) are transformed into the equality constraints with a smooth parameter by means of the smooth integral. Then we show that the accumulation point of the sequence of KKT solutions to the smooth approximation problem is a C-stationary point of the original problem as the smooth parameter tends to zero under the weaker condition of MPCC-MFCQ. Furthermore, it is proven that the accumulation point is an S-stationary point under some mild assumptions. That whether the sequence of KKT solutions to the smooth approximation problem exists is further studied, in which we explore the linear independence constraints qualification, the KKT solutions condition and the second order sufficient condition for the smooth approximation problem

Table 5.1: The comparison of numerical results for problems with low dimension.

$Prob$	(n,m,l)	x_1	Al	f^*	x^*	k_1	k_2	l_{k_1}
bilevel1	(2,2,6)	(50,50)	1	-4.9994	(25.0013,30.0000)	4	78	6.2398e-07
bilevel1	(2,2,6)	(50,50)	2	-4.9994	(25.0013,30.0000)	7	122	3.5300e-07
bilevel1	(2,2,6)	(50,50)	3	-4.9994	(25.0013,30.0000)	5	93	6.6632e-07
bilevel2	(4,4,12)	(5,5, 15,15)	1	-6.6000e+03	(7.1421,3.2132, 11.8579,17.7868)	3	23	2.6370e-11
bilevel2	(4,4,12)	(5,5, 15,15)	2	-6.6000e+03	(7.3019,3.4529, 11.6981,17.5471)	5	42	1.4552e-8
bilevel2	(4,4,12)	(5,5, 15,15)	3	-6.6000e+03	(7.1867,3.2030, 11.8734,17.3622)	4	29	4.8678e-10
bilevel3	(2,6,4)	(0,2)	1	-12.6787	(0,2.0000)	3	18	1.3270e-15
bilevel3	(2,6,4)	(0,2)	2	-12.6787	(0,2.0000)	5	25	1.4832e-08
bilevel3	(2,6,4)	(0,2)	3	-12.6787	(0,2.0000)	4	29	3.8627e-09
desilva	(2,2,2)	(0,0)	1	-1	(0.5005,0.5005)	5	74	6.1610e-07
desilva	(2,2,2)	(0,0)	2	-1	(0.5002,0.5002)	5	79	6.2953e-07
desilva	(2,2,2)	(0,0)	3	-1	(0.5005,0.5005)	5	70	6.1825e-07
gnashl0	(1,4,8)	75	1	-343.3453	55.5513	1	9	5.8785e-11
gnashl0	(1,4,8)	75	2	-343.3453	55.5513	4	19	4.6387e-08
gnashl0	(1,4,8)	75	3	-343.3453	55.5513	3	15	2.3865e-09
gnash11	(1,4,8)	75	1	-203.1551	42.5382	1	9	1.1472e-11
gnash11	(1,4,8)	75	2	-203.1551	42.5382	4	21	5.1908e-07
gnash11	(1,4,8)	75	3	-203.1551	42.5382	5	12	2.8142e-08
gnash12	(1,4,8)	75	1	-68.1357	24.1451	1	9	9.5024e-13
gnash12	(1,4,8)	75	2	-68.1357	24.1451	4	20	6.6333e-08
gnash12	(1,4,8)	75	3	-68.1357	24.1451	6	18	4.3019e-09
gnash13	(1,4,8)	75	1	-19.1541	12.3727	1	9	9.1090e-09
gnash13	(1,4,8)	75	2	-19.1541	12.3727	4	19	8.6584e-08
gnash13	(1,4,8)	75	3	-19.1541	12.3727	2	16	5.7542e-08
gnash14	(1,4,8)	75	1	-3.1612	4.7535	1	11	1.4344e-14
gnash14	(1,4,8)	75	2	-3.1612	4.7535	4	19	1.1683e-07
gnash14	(1,4,8)	75	3	-3.1612	4.7535	3	13	2.5867e-08
gnashl5	(1,4,8)	25	1	-346.8932	50.0000	2	20	1.1781e-14
gnashl5	(1,4,8)	25	2	-346.8932	50.0000	5	21	4.4957e-08
gnashl5	(1,4,8)	25	3	-346.8932	50.0000	6	25	6.7326e-07
gnash16	(1,4,8)	20	1	-224.0372	39.7914	3	20	1.4225e-07
gnash16	(1,4,8)	20	2	-224.0372	39.7914	5	28	3.6478e-08
gnash16	(1,4,8)	20	3	-224.0372	39.7914	4	23	2.3386e-07
gnash17	(1,4,8)	15	1	-80.7860	24.2571	3	21	4.6002e-12
gnash17	(1,4,8)	15	2	-80.7860	24.2571	5	29	1.8518e-08
gnash17	(1,4,8)	15	3	-80.7860	24.2571	4	26	3.8645e-08
gnash18	(1,4,8)	12.5	1	-22.8371	13.0197	3	21	7.5351e-11
gnash18	(1,4,8)	12.5	2	-22.8371	13.0197	5	26	5.4911e-08
gnash18	(1,4,8)	12.5	3	-22.8371	13.0197	4	24	3.8654e-08
gnash19	(1,4,8)	10	1	-5.3491	6.0023	3	19	5.1876e-12
gnash19	(1,4,8)	10	2	-5.3491	6.0023	5	25	1.7337e-08
gnash19	(1,4,8)	10	3	-5.3491	6.0023	4	23	2.3421e-09
outrata31(a)	(1,6,4)	0	1	3.2077	4.0604	3	36	3.2817e-07
outrata31(a)	(1,6,4)	0	2	3.2077	4.0604	7	81	4.5672e-07
outrata31(a)	(1,6,4)	0	3	3.2077	4.0604	7	86	5.3729e-07
outrata31(b)	(1,6,4)	10	1	3.2077	4.0604	3	34	3.5208e-07
outrata31(b)	(1,6,4)	10	2	3.2077	4.0604	7	79	6.7321e-07
outrata31(b)	(1,6,4)	10	3	3.2077	4.0604	7	89	5.0325e-07
outrata32(a)	(1,6,4)	0	1	3.4494	5.1536	3	24	1.6914e-09
outrata32(a)	(1,6,4)	0	2	3.4494	5.1536	5	30	2.0042e-08
outrata32(a)	(1,6,4)	0	3	3.4494	5.1536	6	37	3.6451e-08
outrata32(b)	(1,6,4)	10	1	3.4494	5.1536	3	19	1.6892e-09
outrata32(b)	(1,6,4)	10	2	3.4494	5.1536	5	38	2.0143e-08
outrata32(b)	(1,6,4)	10	3	3.4494	5.1536	6	39	2.7963e-08
outrata33(a)	(1,6,4)	0	1	4.6043	2.3894	3	21	2.8738e-08
outrata33(a)	(1,6,4)	0	2	4.6043	2.3894	5	33	3.5255e-08
outrata33(a)	(1,6,4)	0	3	4.6043	2.3894	8	36	2.9673e-08
outrata33(b)	(1,6,4)	10	1	4.6043	2.3894	3	23	2.8740e-08
outrata33(b)	(1,6,4)	10	2	4.6043	2.3894	5	33	3.1782e-08
outrata33(b)	(1,6,4)	10	3	4.6043	2.3894	6	29	3.0654e-08
outrata34(a)	(1,6,4)	0	1	6.5927	1.3731	3	21	1.5233e-09
outrata34(a)	(1,6,4)	0	2	6.5927	1.3731	3	24	5.2677e-08
outrata34(a)	(1,6,4)	0	3	6.5927	1.3731	5	28	4.5671e-08
outrata34(b)	(1,6,4)	10	1	6.5927	1.3731	3	20	1.5217e-09
outrata34(b)	(1,6,4)	10	2	6.5927	1.3731	3	25	5.2677e-08
outrata34(b)	(1,6,4)	10	3	6.5927	1.3731	6	30	4.1729e-08
stackelberg1	(1,1,1)	0	1	-3.2667e+03	93.3333	1	4	3.1875e-12
stackelberg1	(1,1,1)	0	2	-3.2667e+03	93.3333	4	12	3.7500e-08
stackelberg1	(1,1,1)	0	3	-3.2667e+03	93.3333	2	5	5.7829e-11

under several assumptions on problem (1.1). The reported numerical results indicate that the proposed smoothing method is promising. However, many works still deserve us to investigate. For example, whether we can obtain the same convergence results under the weaker conditions such as MPCC-ACQ. And it is also needed to test the much more higher dimensional problems by using the smoothing method.

Table 5.2: The comparison of numerical results for problems with high dimension.

$Prob$	Dim	x_1	Al	f^*	k_1	k_2	l_{k_1}
flp4-1	80	(ones(1,50),zeros(1,30))	1	5.3012e-14	4	46	2.3845e-07
flp4-1	80	(ones(1,50),zeros(1,30))	2	3.1080e-12	5	50	3.8910e-07
flp4-1	80	(ones(1,50),zeros(1,30))	3	1.4521e-13	6	57	1.2378e-07
portfl-i-1	87	(repmat(1/12,12,1);zeros(75,1))	1	2.3418e-07	6	128	3.12950e-13
portfl-i-1	87	(repmat(1/12,12,1);zeros(75,1))	2	6.0419e-06	6	113	1.0010e-07
portfl-i-1	87	(repmat(1/12,12,1);zeros(75,1))	3	3.3081e-07	6	120	8.3325e-09
portfl-i-2	87	(repmat(1/12,12,1);zeros(75,1))	1	7.0563e-06	5	62	3.3749e-12
portfl-i-2	87	(repmat(1/12,12,1);zeros(75,1))	2	8.3172e-06	6	67	1.0209e-07
portfl-i-2	87	(repmat(1/12,12,1);zeros(75,1))	3	7.3418e-06	7	87	5.6217e-13
portfl-i-3	87	(repmat(1/12,12,1);zeros(75,1))	1	3.4030e-06	5	59	9.6247e-16
portfl-i-3	87	(repmat(1/12,12,1);zeros(75,1))	2	3.4029e-06	6	120	4.4092e-08
portfl-i-3	87	(repmat(1/12,12,1);zeros(75,1))	3	1.4315e-08	8	190	1.7852e-07
portfl-i-4	87	zeros(87,1)	1	9.6926e-07	5	92	5.5133e-08
portfl-i-4	87	zeros(87,1)	2	9.6925e-07	6	110	1.4374e-07
portfl-i-4	87	zeros(87,1)	3	2.1671e-06	8	201	6.7038e-08
portfl-i-5	87	(repmat(1/12,12,1);zeros(75,1))	1	1.1759e-06	5	80	4.9659e-12
portfl-i-5	87	(repmat(1/12,12,1);zeros(75,1))	2	1.1759e-06	6	119	1.0343e-07
portfl-i-5	87	(repmat(1/12,12,1);zeros(75,1))	3	3.0683e-06	8	197	7.9031e-08
flp4-2	110	(ones(1,50),zeros(1,60))	1	6.3379e-13	4	30	1.5456e-06
flp4-2	110	(ones(1,50),zeros(1,60))	2	4.3520e-09	5	32	1.9267e-07
flp4-2	110	(ones(1,50),zeros(1,60))	3	5.3012e-11	4	29	1.1029e-07
flp4-3	140	(ones(1,70),zeros(1,70))	1	3.2171e-13	4	38	1.1939e-07
flp4-3	140	(ones(1,70),zeros(1,70))	2	4.3801e-11	6	40	1.8641e-08
flp4-3	140	(ones(1,70),zeros(1,70))	3	3.6953e-12	3	28	9.5137e-15
flp4-4	200	(ones(1,100),zeros(1,100))	1	6.8162e-13	4	43	3.7037e-08
flp4-4	200	(ones(1,100),zeros(1,100))	2	1.3405e-08	5	36	2.9584e-07
flp4-4	200	(ones(1,100),zeros(1,100))	3	5.3165e-12	4	46	6.0962e-08

Acknowledgements

The authors are indebted to the Associate Editor and anonymous referees for their careful reading of the manuscript and many valuable comments and suggestions which helped to improve the presentation of this paper.

References

[1] M. Benko and H. Gfrerer, An SQP method for mathematical programs with complementarity constraints with strong convergence properties, *Kybernetika* 52 (2016) 169–208.

[2] I. Birbil, S. Fang and J. Han, An entropic regularization approach for mathematical programs with equilibrium constraints, *Comput. Oper. Res.* 31 (2004) 2249–2262.

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

[4] Y. Chen and M. Florian, *The nonlinear bilevel programming problem: formulations, regularity and optimality conditions*, *Optimization* 32 (1995) 193-209.

[5] Y. Chen, S. Huang and Z. Wan, A strong convergent smoothing regularization method for mathematical programs with complementarity constraints, *Pacific J. Optim.* 12 (2016) 497–519.

[6] Y. Chen and Z. Wan, A locally smoothing method for mathematical programs with complementarity constraints, *ANZIAM J.* 56 (2015) 299–315.

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

[8] M. Fang and Z. Zhu, A generalized project metric algorithm for mathematical programs with equilibrium constraints, *J. Comput. Appl. Math.* 288 (2015) 99–115.

[9] M. Fukushima and J.S. Pang, Convergence of a smoothing continuation method for mathematical programs with complementarity constraints, in: Lecture Notes in Econom. and Math. Systems, vol. 477, 1999, pp. 99–110.

[10] T. Guo and J.T. Allison, On the use of mathematical programs with complementarity constraints in combined topological and parametric design of biochemical enzyme networks, *Eng. Optim.* 49 (2017) 345–364.

[11] T. Hoheisel, C. Kanzow and A.Schwartz, Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints, *Math. Program.* 137 (2013) 257–288.

[12] S. Leyffer, MacMPEC: AMPL collection of MPECs. http://wiki. msc. anl. gov/leyffer/index. php/MacMPEC.

[13] J. Li, X. Huang and J. Jian, A new relaxation method for mathematical programs with nonlinear complementarity constraints, *J. Comput. Anal. Appl.* 20 (2016) 548–565.

[14] Y. Y. Li, T. Tan and X.S. Li, A log-exponential smoothing method for mathematical programs with complementarity constraints, *Appl. Math. Comput.* 218 (2012) 5900–5909.

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

[16] G.H. Lin, L. Guo and J.J. Ye, Solving mathematical programs with equilibrium constraints, *J. Optimiz. Theory App.* 166 (2015) 234–256.

[17] S.X. Liu and G.Q. Chen, A multiplier relaxation method for solving mathematical programs with complementarity constraints, *OR. Transactions.* 18 (2014) 119–130.

[18] Z.Q. Luo, J. Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, 1996.

[19] T. Melo, J. Matias and T. Monteiro, Numerical optimization experiments using the hyperbolic smoothing strategy to solve MPCC, *CMMSE* 3 (2013) 1029–1036.

[20] F. Meng and F. Xu, A regularized sample average approximation method for stochastic mathematical programs with nonsmooth equality constraints, *SIAM J. Optim.* 17 (2006) 891–919.

[21] L.P. Pang, N. Xu and J. Lv, The inexact log-ponential regularization method for mathematical programs with vertical complemenarity constraints, *J. Ind. Manag. Optim.* 15 (2019) 59–79.

[22] E. Polak, *Optimization: Algorithms and Consistent Approximations*, Springer-Verlag, New York, 1997.

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

[24] S. Steffensen and M. Ulbrich, A new relaxation scheme for mathematical programs with equilibrium constraints, *SIAM J. Optim.* 20 (2010) 2504–2539.

[25] H. Xu, J. Pang, F. Ordóñez and M. Dessouky, Complementarity models for traffic equilibrium with ridesharing, *Transport. Res. B-Meth.* 81 (2015) 161–182.

[26] H. Yamamura, T. Okuno, S. Hayashi and M. Fukushima, A smoothing SQP method for mathematical programs with linear second-order cone complementarity constraints, *Pacific J. Optim.* 9 (2013) 345–372.

[27] T. Yan, A class of smoothing methods for mathematical programs with complementarity constraints, *Appl. Math. Comput.* 186 (2007) 1–9.

[28] J. J. Ye, D.L. Zhu, Q.J. Zhu, Exact penalization and necessary optimality conditions for generalized bilevel programming problems, *SIAM J. Optim.* 7 (1997) 481–507.

[29] H.X. Yin and J.Z. Zhang, Global convergence of a smooth approximation method for mathematical programs with complementarity constraints, *Math. Method. Oper. Res.* 64 (2006) 255–269.

[30] Y. Zhang and N.V. Sahinidis, Global optimization of mathematical programs with complementarity constraints and application to clean energy deployment, *Optim. Lett.* 10 (2016) 325–340.

Manuscript received 27 December 2018
revised 1 November 2019
accepted for publication 2 January 2020

Suxiang He
School of Science/Center for Mathematical Sciences
Wuhan University of Technology, Wuhan 430070, P.R. China
E-mail address: hesux@whut.edu.cn

Tingting Shen
School of Science/Center for Mathematical Sciences
Wuhan University of Technology, Wuhan 430070, P.R. China
E-mail address: 963288657@qq.com

Jie Zhang
School of Mathematics, Liaoning Normal University
Dalian 116029, P.R. China
E-mail address: jzhang158@gmail.com