A Class of Stochastic Mathematical Programs with Complementarity Constraints: Reformulations and Algorithms *

Gui-Hua Lin† and Masao Fukushima‡

July 18, 2003

Abstract

We consider a class of stochastic mathematical programs with equilibrium constraints (SMPECs), in which all decisions are required to be made here-and-now, before a random event is observed. We show that many problems can be formulated as this kind of SMPECs. In general, SMPECs are more difficult to deal with than MPECs. In order to develop an effective approach, we first give a number of reformulations of the problem and then, based on these reformulations, we propose some smoothed penalty methods for solving SMPECs. A comprehensive convergence theory is also included.

Key Words. stochastic mathematical program with equilibrium constraints, stochastic linear complementarity problem, here-and-now decision problem, convex programming, stationarity, subdifferential, convergence.

1 Introduction

Mathematical programs with equilibrium constraints (MPECs) play an important role in many fields such as engineering design, economic equilibrium, multilevel game, and mathematical programming itself, and they therefore have been receiving much attention in the optimization world, see the monograph [12] and some attached references. Recently, Lin, Chen and Fukushima [8] discuss MPECs under uncertainty; they study stochastic mathematical programs with equilibrium constraints (SMPECs), which can be thought as generalizations of MPECs. It has been shown that many decision problems can be formulated as SMPECs in practice. Moreover, SMPECs have also close connection with the well-known two-stage stochastic programs with recourse, see [8] for details.

*This work was supported in part by the Scientific Research Grant-in-Aid from the Ministry of Education, Science, Sports, and Culture of Japan.

†Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan. /Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, China. E-mail: ghlin@amp.i.kyoto-u.ac.jp.

‡Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan. E-mail: fuku@amp.i.kyoto-u.ac.jp.

There are two kinds of SMPECs discussed in [8]: One is the lower-level wait-and-see model, in which the upper-level decision is made at once and a lower-level decision may be made after a random event is observed, and the other is the here-and-now model that requires us to make all decisions before a random event is observed. In particular, the latter problem can be stated as follows:

$$\begin{array}{ll} \text{minimize} & E_\omega[c(x,y,\omega)+d^T z(\omega)] \\ \text{subject to} & x \in X, \\ & y \geq 0,\ F(x,y,\omega)+z(\omega) \geq 0, \\ & y^T(F(x,y,\omega)+z(\omega))=0, \\ & z(\omega)\geq 0, \quad \omega\in\Omega. \end{array} \tag{1.1}$$

Here, X is a subset of $\Re^n$, $c:\Re^{n+m}\times\Omega\to\Re$ and $F:\Re^{n+m}\times\Omega\to\Re^m$ are mappings, E_ω means expectation with respect to the random variable $\omega\in\Omega$, $z(\omega)$ is a recourse variable, and $d\in\Re^m$ is a constant vector with positive elements. In this problem, x denotes the upper-level decision, y represents the lower-level decision, and both the decisions x and y need to be made at once, before ω is observed. In [8], a smoothing implicit programming method incorporating a penalty technique has been proposed for solving problem (1.1) with ω being a discrete random variable and $F(\cdot,\cdot,\omega)$ being affine for each ω.

In this paper, we consider the following here-and-now problem:

$$\begin{array}{ll} \text{minimize} & f(x,y)+E_\omega[d^T z(\omega)] \\ \text{subject to} & g(x,y)\leq 0,\ h(x,y)=0, \\ & y\geq 0,\ N(\omega)x+M(\omega)y+q(\omega)+z(\omega)\geq 0, \\ & y^T(N(\omega)x+M(\omega)y+q(\omega)+z(\omega))=0, \\ & z(\omega)\geq 0, \quad \omega\in\Omega, \end{array} \tag{1.2}$$

where $f:\Re^{n+m}\to\Re, g:\Re^{n+m}\to\Re^{s_1}$, and $h:\Re^{n+m}\to\Re^{s_2}$ are all continuously differentiable, the data $N(\omega)\in\Re^{m\times n}$, $M(\omega)\in\Re^{m\times m}$, and $q(\omega)\in\Re^m$ are random variables, $z(\omega)$ is the corresponding recourse variable, and the vector $d\in\Re^m$ is a constant vector with positive elements. Note that g and h are functions of both the upper-level and the lower-level variables, unlike in [8] where they are functions of the variable x only.

Many problems can be formulated as this kind of models, see Section 2 for some examples. However, SMPECs are more difficult to solve than MPECs in general. In order to develop an effective approach, we will give a number of reformulations of problem (1.2) in Section 3. Based on these reformulations, we will propose some smoothed penalty methods for solving (1.2) in Section 4. A comprehensive convergence theory will also be included.

The following notations will be used in this paper: All vectors are column vectors, and $w[i]$ stands for the ith coordinate of vector $w\in\Re^s$, whereas for a matrix M, $M[i]$ denotes the column vector whose elements comprise the ith row of M. For any vectors u and v of the same dimension, $u \perp v$ means $u^T v = 0$. Given a function $F : \Re^s \to \Re^{s'}$ and a vector $w \in \Re^s$, $\nabla F(w) \in \Re^{s \times s'}$ is the transposed Jacobian of F at w and $\mathcal{I}_F(w) := \{i \mid F_i(w) = 0\}$ stands for the active index set of F at w. For two vectors $u, v \in \Re^s$, $\min(u, v)$ is understood to be taken componentwise, i.e.,

$$\min(u, v) = (\min\{u[1], v[1]\}, \cdots, \min\{u[s], v[s]\})^T.$$

In addition, e_i denotes the unit vector with $e_i[i] = 1$; I and O denote the identity matrix and the zero matrix with suitable dimension, respectively.

2 Examples

In this section, we give some examples to show that problem (1.2) plays an important role in practice. The first example was originally given in [8].

Example 2.1 `The Picnic-Vendor Decision Problem:` There are a food company who makes picnic lunches and a vendor who sells lunches to hikers on every Sunday. The company and the vendor have the following contract: (C1) The vendor buys lunches from the company at the price $x \in [a, b]$ determined by the company, where a and b are two positive constants. (C2) The vendor decides the amount s of lunches that he buys from the company, where s must be no less than the minimum amount $c > 0$. (C3) The vendor pays the company for the whole lunches he buys, i.e., the vendor pays sx to the company. (C4) The vendor sells lunches to hikers at the price $2x$ and get the proceeds for the total number of lunches actually sold. (C5) Even if there are any unsold lunches, the vendor cannot return them to the company but he can dispose of the unsold lunches with no cost.

We suppose that the demand of lunches depends on the price and the weather on that day. Since the weather is uncertain, we may treat it as a random variable. More specifically, we suppose that the demand is given by the function

$$\phi(x, \omega) := \beta(\omega) - \gamma(\omega)x, \quad \omega \in \Omega,$$

where $\beta(\omega) \geq 0$ and $\gamma(\omega) \geq 0$ are random variables. Therefore, the actual amount of lunches sold is given by $\min(s, \phi(x, \omega))$, which also depends on the weather on that day.

The decisions by the company and the vendor are x and s, respectively. The company's objective is to maximize its total profit sx, while the vendor's objective is to maximize its total profit $2x \min(s, \phi(x, \omega)) - sx$. The latter problem may be written as

$$\begin{array}{ll} \text{maximize}_{s,t} & x(2t - s) \\ \text{subject to} & s \geq c, \;\; s - t \geq 0, \\ & \beta(\omega) - \gamma(\omega)x - t \geq 0, \end{array}$$

whose optimality conditions are stated as

$$\begin{pmatrix} x \\ -2x \end{pmatrix} - u\begin{pmatrix} 1 \\ 0 \end{pmatrix} - v\begin{pmatrix} 1 \\ -1 \end{pmatrix} - w\begin{pmatrix} 0 \\ -1 \end{pmatrix} = 0, \tag{2.1}$$

$$0 \le u \perp (s-c) \ge 0,$$

$$0 \le v \perp (s-t) \ge 0,$$

$$0 \le w \perp (\beta(\omega) - \gamma(\omega)x - t) \ge 0. \tag{2.2}$$

It follows from (2.1) that

$$u = x - v, \quad w = 2x - v.$$

This implies that $w = x + u \ge a > 0$, which together with (2.2) yields $t = \beta(\omega) - \gamma(\omega)x$. Thus the above optimality conditions may further be rewritten as

$$\begin{aligned} &0 \le (x-v) \perp (s-c) \ge 0, \\ &0 \le v \perp (s - \beta(\omega) + \gamma(\omega)x) \ge 0. \end{aligned} \tag{2.3}$$

Then we may formulate the company's problem as the following stochastic MPEC:

$$\begin{aligned} \text{minimize}_{x,s} \quad & -sx \\ \text{subject to} \quad & a \le x \le b, \\ & 0 \le (x-v) \perp (s-c) \ge 0, \\ & 0 \le v \perp (s - \beta(\omega) + \gamma(\omega)x) \ge 0, \end{aligned}$$

or equivalently, by letting $y = (s-c, v)^T$, we have

$$\begin{aligned} \text{minimize} \quad & -x(y[1]+c) \\ \text{subject to} \quad & a \le x \le b, \quad \omega \in \Omega, \\ & 0 \le y \perp (N(\omega)x + My + q(\omega)) \ge 0, \end{aligned}$$

where

$$N(\omega) = \begin{pmatrix} 1 \\ \gamma(\omega) \end{pmatrix}, \quad M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad q(\omega) = \begin{pmatrix} 0 \\ c - \beta(\omega) \end{pmatrix}.$$

Furthermore, suppose that both the company and the vendor have to make decision on Saturday, without knowing the weather of Sunday. In this case, there is no (x, v) satisfying (2.3) for all $\omega \in \Omega$ in general. So, by introducing the recourse variables $z(\omega), \omega \in \Omega$, the company's problem is represented as the following model:

$$\begin{aligned} \text{minimize} \quad & -x(y[1]+c) + E_\omega[\, d^T z(\omega)\,] \\ \text{subject to} \quad & a \le x \le b, \\ & 0 \le y \perp (N(\omega)x + My + q(\omega) + z(\omega)) \ge 0, \\ & z(\omega) \ge 0, \quad \omega \in \Omega, \end{aligned}$$

where $d \in \Re^2$ is a constant vector with positive elements.

The second example is related to the linear complementarity problem.

Example 2.2 An ordinary linear complementarity problem is to find a vector $y \in \Re^m$ such that

$$y \geq 0, \quad My + q \geq 0, \quad y^T(My + q) = 0,$$

where $M \in \Re^{m \times m}$ and $q \in \Re^m$. In many practical problems, some elements may involve uncertain data, which can be characterized by random variables. Therefore, it is meaningful to consider the following stochastic linear complementarity problem:

$$y \geq 0, \ M(\omega)y + q(\omega) \geq 0, \ y^T(M(\omega)y + q(\omega)) = 0, \quad \forall \omega \in \Omega. \tag{2.4}$$

In general, there may not exist a vector y satisfying these complementarity conditions for all $\omega \in \Omega$. In order to get a reasonable resolution, we may introduce recourse variables $z(\omega) \geq 0$ to the inequality $M(\omega)y + q(\omega) \geq 0$ and try to find a vector $y \geq 0$ that minimizes the total recourse. Thus, we obtain the following problem

$$\begin{array}{ll} \text{minimize} & E_\omega[d^T z(\omega)] \\ \text{subject to} & 0 \leq y \perp (M(\omega)y + q(\omega) + z(\omega)) \geq 0, \\ & z(\omega) \geq 0, \ \omega \in \Omega, \end{array} \tag{2.5}$$

where d is a constant vector with positive elements.

3 Reformulations

Because of the existence of the complementarity constraints, problem (1.2) does not satisfy a standard constraint qualification such as the linear independence constraint qualification or the Mangasarian-Fromovitz constraint qualification at any feasible point [2]. Therefore, the conventional theory and algorithms in nonlinear programming cannot be applied to this problem directly. In order to look for some effective methods for solving problem (1.2), we give some reformulations of (1.2) in the following.

At first, we define the function $Q : \Re^n \times \Re^m \times \Omega \to [0, +\infty]$ by

$$Q(x, y, \omega) := \sup \Big\{ -(u + ty)^T (N(\omega)x + M(\omega)y + q(\omega)) \ \Big| \ u + ty \leq d, \ \ u \geq 0 \Big\}$$

and consider the problem

$$\begin{array}{ll} \text{minimize} & f(x, y) + E_\omega[Q(x, y, \omega)] \\ \text{subject to} & g(x, y) \leq 0, \ h(x, y) = 0, \ \ y \geq 0. \end{array} \tag{3.1}$$

We will show that problem (3.1) is equivalent to problem (1.2). To this end, we will give some properties of the function $Q(x, y, \omega)$. In what follows, we let $\mathcal{F}_1$ and $\mathcal{F}_2$ denote the feasible regions of problems (1.2) and (3.1), respectively. We assume both problems (1.2) and (3.1) have an optimal solution. Moreover, we suppose f is bounded from below on $\mathcal{F}_2$.

Theorem 3.1 *For any $x \in \Re^n, y \in \Re^m$, and $\omega \in \Omega$, $Q(x,y,\omega) < +\infty$ if and only if the set*

$$Z(x,y,\omega) := \left\{ z(\omega) \;\middle|\; \begin{array}{l} y^T(N(\omega)x + M(\omega)y + q(\omega) + z(\omega)) = 0 \\ N(\omega)x + M(\omega)y + q(\omega) + z(\omega) \geq 0, \; z(\omega) \geq 0 \end{array} \right\}$$

is nonempty; moreover, we have

$$Q(x,y,\omega) = \min\Big\{ d^T z(\omega) \;\Big|\; z(\omega) \in Z(x,y,\omega) \Big\}. \tag{3.2}$$

We can prove the above theorem using the duality theorem in linear programming immediately. On the other hand, we may write

$$\begin{aligned} Z(x,y,\omega) &= \left\{ z(\omega) \;\middle|\; \begin{array}{l} y^T(N(\omega)x + M(\omega)y + q(\omega) + z(\omega)) \leq 0 \\ N(\omega)x + M(\omega)y + q(\omega) + z(\omega) \geq 0, \; z(\omega) \geq 0 \end{array} \right\} \\ &= \left\{ z(\omega) \;\middle|\; \begin{array}{l} y^T z(\omega) \leq -y^T(N(\omega)x + M(\omega)y + q(\omega)) \\ z(\omega) \geq -(N(\omega)x + M(\omega)y + q(\omega)), \; z(\omega) \geq 0 \end{array} \right\}. \end{aligned}$$

Therefore, when $Q(x,y,\omega) < +\infty$, we have from Theorem 3.1 and the duality theorem that

$$Q(x,y,\omega) = \max\Big\{ -(u+ty)^T(N(\omega)x + M(\omega)y + q(\omega)) \;\Big|\; u + ty \leq d, \;\; u \geq 0, \;\; t \leq 0 \Big\}.$$

Furthermore, we have the the following result, which will be used in the subsequent analysis.

Theorem 3.2 *Let $g(x,y) \leq 0, h(x,y) = 0, y \geq 0$, and $\omega \in \Omega$. Then $Q(x,y,\omega) = +\infty$ if and only if there exists an index i such that*

$$\Big(N(\omega)x + M(\omega)y + q(\omega)\Big)[i] > 0, \quad y[i] > 0. \tag{3.3}$$

Proof: " if ": Suppose that there exists an index i satisfying (3.3). Let t be a real number and $u(t) \in \Re^m$ be defined by

$$u(t) := ty[i]e_i - ty.$$

Then, for any $t \leq 0$, we have

$$u(t) \geq 0, \quad u(t) + ty \leq d.$$

It follows from the definition of Q and (3.3) that

$$\begin{aligned} Q(x,y,\omega) &\geq \sup\Big\{ -(u(t)+ty)^T(N(\omega)x + M(\omega)y + q(\omega)) \;\Big|\; t \leq 0 \Big\} \\ &= \sup\Big\{ -ty[i](N(\omega)x + M(\omega)y + q(\omega))[i] \;\Big|\; t \leq 0 \Big\} \\ &= +\infty, \end{aligned}$$

from which we obtain $Q(x,y,\omega) = +\infty$ immediately.

" only if ": First of all, we define

$$\begin{aligned}
\mathcal{I}_1 &:= \Big\{ i \,\Big|\, \; y[i] = 0 \Big\}, \\
\mathcal{I}_2 &:= \Big\{ i \,\Big|\, \; y[i] > 0, \quad (N(\omega)x + M(\omega)y + q(\omega))[i] = 0 \Big\}, \\
\mathcal{I}_3 &:= \Big\{ i \,\Big|\, \; y[i] > 0, \quad (N(\omega)x + M(\omega)y + q(\omega))[i] < 0 \Big\}, \\
\mathcal{I}_4 &:= \Big\{ i \,\Big|\, \; y[i] > 0, \quad (N(\omega)x + M(\omega)y + q(\omega))[i] > 0 \Big\}.
\end{aligned}$$

These index sets are mutually disjoint and, since $y \geq 0$, we see that

$$\mathcal{I}_1 \cup \mathcal{I}_2 \cup \mathcal{I}_3 \cup \mathcal{I}_4 = \{1, 2, \cdots, n\}. \tag{3.4}$$

Denote $C := \{ \ (u,t) \ | \ \ u + ty \leq d, \ u \geq 0\}$. Then, we have

$$Q(x,y,\omega) = \sup \Big\{ -\sum_{i=1}^{m} (u+ty)[i](N(\omega)x + M(\omega)y + q(\omega))[i] \ \Big| \ (u,t) \in C \Big\}. \tag{3.5}$$

(a) For any $i \in \mathcal{I}_1$ and any $(u,t) \in C$, we have $0 \leq u[i] \leq d[i]$. It follows that the term

$$-\sum_{i \in \mathcal{I}_1} (u+ty)[i](N(\omega)x + M(\omega)y + q(\omega))[i] = -\sum_{i \in \mathcal{I}_1} u[i](N(\omega)x + M(\omega)y + q(\omega))[i]$$

is bounded on C.

(b) It is obvious that

$$-\sum_{i \in \mathcal{I}_2} (u+ty)[i](N(\omega)x + M(\omega)y + q(\omega))[i] = 0.$$

(c) Let $i \in \mathcal{I}_3$. For any $(u,t) \in C$, we have

$$-\sum_{i \in \mathcal{I}_3} (u+ty)[i](N(\omega)x + M(\omega)y + q(\omega))[i] \leq -\sum_{i \in \mathcal{I}_3} d[i](N(\omega)x + M(\omega)y + q(\omega))[i]$$

and so the term on the left-hand side is bounded above on C.

The facts (a)–(c) together with (3.4) and (3.5) indicate that, if $Q(x,y,\omega) = +\infty$, we must have $\mathcal{I}_4 \neq \emptyset$, from which we see that (3.3) must hold for some index i and hence the proof is completed. ∎

From Theorem 3.2, we have the following result immediately.

Corollary 3.1 *Let $g(x,y) \leq 0, h(x,y) = 0, y \geq 0$, and $\omega \in \Omega$. Then $Q(x,y,\omega) < +\infty$ if and only if*

$$y[i](N(\omega)x + M(\omega)y + q(\omega))[i] \leq 0 \tag{3.6}$$

holds for every i.

This corollary, together with Theorem 3.1, yields the next result.

Corollary 3.2 *Let $g(x,y) \leq 0, h(x,y) = 0, y \geq 0$, and $\omega \in \Omega$. If $Q(x,y,\omega) < +\infty$, then*

$$Q(x,y,\omega) = d^T z(x,y,\omega), \tag{3.7}$$

where

$$z(x,y,\omega) := \max\{-(N(\omega)x + M(\omega)y + q(\omega)),\ 0\}. \tag{3.8}$$

Proof: Noticing that (3.6) together with $y \geq 0$ holds for each i, we can easily prove that $z(x,y,\omega) \in Z(x,y,\omega)$. On the other hand, for any $z(\omega) \in Z(x,y,\omega)$, since

$$z(\omega) \geq -(N(\omega)x + M(\omega)y + q(\omega)), \qquad z(\omega) \geq 0,$$

we have from the definition (3.8) that

$$z(\omega) - z(x,y,\omega) \geq 0,$$

which implies that

$$d^T z(\omega) \geq d^T z(x,y,\omega).$$

Since $z(\omega) \in Z(x,y,\omega)$ is arbitrary, (3.7) follows form (3.2) at once. ∎

We next show the equivalence between problems (1.2) and (3.1).

3.1 Discrete case

Suppose that $\Omega = \{\omega_1, \omega_2, \cdots, \omega_L\}$ and, for each $\ell = 1, 2, \cdots, L$, the probability p_ℓ of the random event ω_ℓ is positive. Then, problems (1.2) and (3.1) reduce to

$$\begin{array}{ll} \text{minimize} & f(x,y) + \sum_{\ell=1}^{L} p_\ell d^T z_\ell \\ \text{subject to} & g(x,y) \leq 0,\ h(x,y) = 0, \\ & y \geq 0,\ N_\ell x + M_\ell y + q_\ell + z_\ell \geq 0, \\ & y^T(N_\ell x + M_\ell y + q_\ell + z_\ell) = 0, \\ & z_\ell \geq 0, \quad \ell = 1, 2, \cdots, L \end{array} \tag{3.9}$$

and

$$\begin{array}{ll} \text{minimize} & f(x,y) + \sum_{\ell=1}^{L} p_\ell Q(x,y,\omega_\ell) \\ \text{subject to} & g(x,y) \leq 0,\ h(x,y) = 0, \quad y \geq 0, \end{array} \tag{3.10}$$

respectively. The following result shows that problem (3.10) is equivalent to the MPEC (3.9).

Theorem 3.3 *If (x^*, y^*) solves problem (3.10), then there exist $z_\ell^* \in Z(x^*, y^*, \omega_\ell), \ell = 1, 2, \cdots, L$, such that $(x^*, y^*, z_1^*, \cdots, z_L^*)$ solves the MPEC (3.9). Conversely, if $(x^*, y^*, z_1^*, \cdots, z_L^*)$ solves the MPEC (3.9), then (x^*, y^*) solves problem (3.10).*

Proof: (a) Suppose that (x^*, y^*) solves (3.10). By assumption, problem (3.10) has a finite optimal value and furthermore, taking into account that each p_ℓ is positive, we see that

$$Q(x^*, y^*, \omega_\ell) < +\infty, \qquad \ell = 1, 2, \cdots, L.$$

Then, by Corollary 3.2, there exist $z_\ell^* \in Z(x^*, y^*, \omega_\ell), \ell = 1, 2, \cdots, L$, such that $(x^*, y^*, z_1^*, \cdots, z_L^*)$ is feasible to problem (3.9) and

$$Q(x^*, y^*, \omega_\ell) = d^T z_\ell^*, \qquad \ell = 1, 2, \cdots, L.$$

Here, we use the lower boundedness of the function $d^T z_\ell$ on $Z(x^*, y^*, \omega_\ell)$ and the polyhedral convexity of the set $Z(x^*, y^*, \omega_\ell)$. We next prove that $(x^*, y^*, z_1^*, \cdots, z_L^*)$ solves problem (3.9).

Let $(x, y, z_1, \cdots, z_L) \in \mathcal{F}_1$. This implies that $z_\ell \in Z(x, y, \omega_\ell)$ for each ℓ and hence, by Theorem 3.1, we have

$$Q(x, y, \omega_\ell) \le d^T z_\ell, \qquad \ell = 1, 2, \cdots, L.$$

It then follows that

$$\begin{aligned} f(x^*, y^*) + \sum_{\ell=1}^{L} p_\ell d^T z_\ell^* &= f(x^*, y^*) + \sum_{\ell=1}^{L} p_\ell Q(x^*, y^*, \omega_\ell) \\ &\le f(x, y) + \sum_{\ell=1}^{L} p_\ell Q(x, y, \omega_\ell) \\ &\le f(x, y) + \sum_{\ell=1}^{L} p_\ell d^T z_\ell, \end{aligned}$$

which means that $(x^*, y^*, z_1^*, \cdots, z_L^*)$ is an optimal solution of problem (3.9).

(b) Suppose that $(x^*, y^*, z_1^*, \cdots, z_L^*)$ solves problem (3.9). Note that, $z_\ell^* \in Z(x^*, y^*, \omega_\ell), \ell = 1, 2, \cdots, L$. It then follows from Theorem 3.1 that

$$Q(x^*, y^*, \omega_\ell) \le d^T z_\ell^*, \qquad \ell = 1, 2, \cdots, L. \tag{3.11}$$

Let $(x, y) \in \mathcal{F}_2$. First we suppose that $Q(x, y, \omega_\ell) < +\infty$ for every ℓ. This means that $Z(x, y, \omega_\ell)$ is nonempty for each ℓ. For any $z_\ell \in Z(x, y, \omega_\ell), \ell = 1, 2, \cdots, L$, we have $(x, y, z_1, \cdots, z_L) \in \mathcal{F}_1$ and then, by (3.11) and the optimality of $(x^*, y^*, z_1^*, \cdots, z_L^*)$,

$$\begin{aligned} f(x^*, y^*) + \sum_{\ell=1}^{L} p_\ell Q(x^*, y^*, \omega_\ell) &\le f(x^*, y^*) + \sum_{\ell=1}^{L} p_\ell d^T z_\ell^* \\ &\le f(x, y) + \sum_{\ell=1}^{L} p_\ell d^T z_\ell. \end{aligned}$$

Noticing that $z_\ell \in Z(x, y, \omega_\ell)$ is arbitrary for every ℓ and taking (3.2) into account, we obtain

$$f(x^*, y^*) + \sum_{\ell=1}^{L} p_\ell Q(x^*, y^*, \omega_\ell) \le f(x, y) + \sum_{\ell=1}^{L} p_\ell Q(x, y, \omega_\ell) \tag{3.12}$$

for any $(x, y) \in \mathcal{F}_2$. If $Q(x, y, \omega_\ell) = +\infty$ for some index ℓ, (3.12) holds immediately from the positivity of p_l. This means that (x^*, y^*) is an optimal solution of problem (3.10). ∎

3.2 Continuous case

Let ω be a continuous random variable and $\phi(\omega)$ represent the probability density function of ω. Suppose that the probability of any nonempty open set in Ω is positive and $(N(\omega), M(\omega), q(\omega))$ is continuous with respect to ω. Analogous to the discrete case, we have the following result:

Theorem 3.4 *If (x^*, y^*) solves problem (3.1), then there exist $z^*(\omega) \in Z(x^*, y^*, \omega)$, $\omega \in \Omega$, such that $(x^*, y^*, z^*(\omega))_{\omega\in\Omega}$ solves problem (1.2). Conversely, if $(x^*, y^*, z^*(\omega))_{\omega\in\Omega}$ solves problem (1.2), then (x^*, y^*) solves (3.1).*

Proof: We first suppose that (x^*, y^*) solves (3.1). Then we claim that

$$Q(x^*, y^*, \omega) < +\infty, \qquad \forall \omega \in \Omega. \tag{3.13}$$

In fact, if $Q(x^*, y^*, \bar\omega) = +\infty$ for some $\bar\omega \in \Omega$, then, by Theorem 3.2, there exists an index i such that

$$\Big(N(\bar\omega)x^* + M(\bar\omega)y^* + q(\bar\omega)\Big)[i] > 0, \quad y^*[i] > 0.$$

It follows from the continuity of $(N(\omega), M(\omega), q(\omega))$ that there is a neighborhood $U(\bar\omega)$ of $\bar\omega$ such that

$$\Big(N(\omega)x^* + M(\omega)y^* + q(\omega)\Big)[i] > 0, \quad y^*[i] > 0$$

hold for any $\omega \in U(\bar\omega)$. We then have from Theorem 3.2 that

$$Q(x^*, y^*, \omega) = +\infty, \qquad \forall \omega \in U(\bar\omega).$$

Therefore,

$$f(x^*, y^*) + E_\omega[Q(x^*, y^*, \omega)] \ge f(x^*, y^*) + \int_{U(\bar\omega)} Q(x^*, y^*, \omega)\phi(\omega)d\omega = +\infty.$$

This is a contradiction and hence (3.13) must hold. As a result, we have from Theorem 3.1 and Corollary 3.2 that there exists $z^*(\omega) \in Z(x^*, y^*, \omega)$ such that

$$Q(x^*, y^*, \omega) = \min\Big\{d^T z(\omega) \;\Big|\; z(\omega) \in Z(x^*, y^*, \omega)\Big\} = d^T z^*(\omega).$$

It then follows that, for any $(x, y, z(\omega))_{\omega\in\Omega} \in \mathcal{F}_1$,

$$\begin{aligned} f(x^*, y^*) + E_\omega[d^T z^*(\omega)] &= f(x^*, y^*) + E_\omega[Q(x^*, y^*, \omega)] \\ &\leq f(x, y) + E_\omega[Q(x, y, \omega)] \\ &\leq f(x, y) + E_\omega[d^T z(\omega)], \end{aligned} \tag{3.14}$$

where the first inequality follows from the optimality of (x^*, y^*) and the last inequality follows from Theorem 3.1 and the fact that $(x, y, z(\omega))_{\omega\in\Omega} \in \mathcal{F}_1$ implies $Z(x, y, \omega)$ is nonempty (and hence there holds (3.2)) for any $\omega \in \Omega$. The inequality (3.14) means that $(x^*, y^*, z^*(\omega))_{\omega\in\Omega}$ is an optimal solution of problem (1.2).

Conversely, if $(x^*, y^*, z^*(\omega))_{\omega\in\Omega}$ is an optimal solution of problem (1.2), in a similar way to Part (b) of the proof of Theorem 3.3 and making use of the fact that, if $Q(x, y, \hat{\omega}) = +\infty$ for some $\hat{\omega} \in \Omega$, there exists a neighborhood $U(\hat{\omega})$ of $\hat{\omega}$ such that

$$Q(x, y, \omega) = +\infty, \qquad \forall \omega \in U(\hat{\omega}),$$

(this fact can be shown from Theorem 3.2 as stated above), we can show that (x^*, y^*) solves problem (3.1). This completes the proof of (ii). ∎

4 Smoothed Penalty Methods for Discrete Random Variables

We will develop some effective methods for solving problem (3.9) in this section. At first, we see from Theorem 3.3 that problem (3.9) is equivalent to problem (3.10) and furthermore, by Corollary 3.1 and the positivity of every p_ℓ, problem (3.10) is equivalent to the problem

$$\begin{array}{ll} \text{minimize} & f(x, y) + \displaystyle\sum_{\ell=1}^{L} p_\ell Q(x, y, \omega_\ell) \\ \text{subject to} & g(x, y) \leq 0,\ h(x, y) = 0, \ \ y \geq 0, \\ & y[i](N_\ell x + M_\ell y + q_\ell)[i] \leq 0, \\ & i = 1, \cdots, m,\ \ell = 1, \cdots, L. \end{array} \tag{4.1}$$

Let $\mathcal{F}_3$ denote the feasible region of problem (4.1). We then have from Corollary 3.1 that

$$Q(x, y, \omega_\ell) < +\infty, \qquad \forall (x, y) \in \mathcal{F}_3,\ \forall \ell.$$

By Corollary 3.2, we see that, for any $(x, y) \in \mathcal{F}_3$ and any ℓ,

$$Q(x, y, \omega_\ell) = d^T z_\ell(x, y)$$

with $z_\ell(x, y) = \max(-(N_\ell x + M_\ell y + q_\ell),\ 0)$. Let

$$\theta(x, y) := \sum_{\ell=1}^{L} p_\ell d^T \max\Big(-(N_\ell x + M_\ell y + q_\ell),\ 0\Big). \tag{4.2}$$

Then problem (4.1) may be rewritten as

$$\begin{array}{ll} \text{minimize} & f(x,y)+\theta(x,y) \\ \text{subject to} & g(x,y)\le 0,\ h(x,y)=0,\quad y\ge 0, \\ & y[i](N_\ell x+M_\ell y+q_\ell)[i]\le 0, \\ & i=1,\cdots,m,\ \ell=1,\cdots,L, \end{array} \tag{4.3}$$

which is no longer an SMPEC.

Example 4.1 Consider the SMPEC

$$\begin{array}{ll} \text{minimize} & p(z_1[1]+z_1[2])+(1-p)(z_2[1]+z_2[2]) \\ \text{subject to} & z_1[1]\ge 0,\ z_1[2]\ge 0,\ z_2[1]\ge 0,\ z_2[2]\ge 0, \\ & 0\le \begin{pmatrix} y[1] \\ y[2] \end{pmatrix} \perp \begin{pmatrix} 2y[1]-3+z_1[1] \\ y[2]-5+z_1[2] \end{pmatrix} \ge 0, \\ & 0\le \begin{pmatrix} y[1] \\ y[2] \end{pmatrix} \perp \begin{pmatrix} 2y[2]-7+z_2[1] \\ y[1]-2+z_2[2] \end{pmatrix} \ge 0, \end{array} \tag{4.4}$$

where p is a constant such that $0<p<1$. For this problem, problem (4.3) becomes

$$\begin{array}{ll} \text{minimize} & p\Big(\max\{3-2y[1],\ 0\}+\max\{5-y[2],\ 0\}\Big) \\ & \quad +(1-p)\Big(\max\{7-2y[2],\ 0\}+\max\{2-y[1],\ 0\}\Big) \\ \text{subject to} & y[1]\ge 0,\quad y[2]\ge 0, \\ & y[1](2y[1]-3)\le 0,\quad y[2](y[2]-5)\le 0, \\ & y[1](2y[2]-7)\le 0,\quad y[2](y[1]-2)\le 0. \end{array} \tag{4.5}$$

Problem (4.5) has a unique solution $y^*=(\frac{3}{2},\frac{7}{2})$ for any $p\in(0,1)$ and it is easy to verify that the linear independence constraint qualification holds at y^*. This indicates that problem (4.5) has ordinary constraints, unlike the original SMPEC (4.4) that does not satisfy any standard constraint qualification at any feasible point.

Problem (4.3) possesses the same optimal solution set as problem (3.9) in the sense of Theorem 3.3. However, it may not be easy to obtain a global optimal solution of an optimization problem in practice, whereas computation of stationary points may be relatively easy. Therefore, it is necessary to study the relation between the stationary points of problems (4.3) and (3.9). Recall that, for each ℓ, $N_\ell[i]$ and $M_\ell[i]$ denote the column vectors whose elements comprise the ith row of the matrices N_ℓ and M_ℓ, respectively.

Since θ is a nonsmooth convex function, we will use the following standard definition of stationarity for problem (4.3):

Definition 4.1 A point $(x^*,y^*)\in\Re^{n+m}$ is said to be *stationary* to problem (4.3) if it is feasible to (4.3) and there exist Lagrange multiplier vectors λ^*,μ^*,ν^*, and ξ^*_ℓ, $\ell=1,\cdots,L$, such that

$$0\in\nabla f(x^*,y^*)+\partial\theta(x^*,y^*)+\nabla g(x^*,y^*)\lambda^*+\nabla h(x^*,y^*)\mu^*-\begin{pmatrix} O \\ I \end{pmatrix}\nu^*$$

$$+\sum_{\ell=1}^{L}\sum_{i=1}^{m}\xi_\ell^*[i]\Big((N_\ell x^*+M_\ell y^*+q_\ell)[i]\begin{pmatrix}0\\ e_i\end{pmatrix}+y^*[i]\begin{pmatrix}N_\ell[i]\\ M_\ell[i]\end{pmatrix}\Big), \tag{4.6}$$

$$0\le\lambda^*\perp(-g(x^*,y^*))\ge 0, \tag{4.7}$$

$$0\le\nu^*\perp\quad y^*\ge 0, \tag{4.8}$$

$$0\le\xi_\ell^*[i]\perp(-y^*[i](N_\ell x^*+M_\ell y^*+q_\ell)[i])\ge 0. \tag{4.9}$$

Here, $\partial\theta(x^*,y^*)$ stands for the subdifferential [15] of θ at the point (x^*,y^*).

For each ℓ and i, we let $\theta_{\ell,i}(x,y):=\max\{-(N_\ell x+M_\ell y+q_\ell)[i],0\}$. It then follows that

$$\partial\theta_{\ell,i}(x^*,y^*)=\begin{cases}\mathrm{co}\Big\{\begin{pmatrix}-N_\ell[i]\\ -M_\ell[i]\end{pmatrix},0\Big\}, & (N_\ell x^*+M_\ell y^*+q_\ell)[i]=0\\ \Big\{\begin{pmatrix}-N_\ell[i]\\ -M_\ell[i]\end{pmatrix}\Big\}, & (N_\ell x^*+M_\ell y^*+q_\ell)[i]<0\\ \Big\{0\Big\}, & (N_\ell x^*+M_\ell y^*+q_\ell)[i]>0\end{cases} \tag{4.10}$$

and

$$\partial\theta(x^*,y^*)=\sum_{\ell=1}^{L}\sum_{i=1}^{m}\ p_\ell\ d[i]\ \partial\theta_{\ell,i}(x^*,y^*), \tag{4.11}$$

where co stands for the convex hull.

On the other hand, it is easy to see that problem (3.9) is equivalent to the following ordinary MPEC:

$$\begin{array}{ll}\text{minimize} & f(x,y)+\mathbf{d}^T\mathbf{z}\\ \text{subject to} & g(x,y)\le 0,\ h(x,y)=0,\\ & \mathbf{y}-Dy=0,\ \mathbf{z}\ge 0,\\ & \mathbf{y}\ge 0,\ Nx+M\mathbf{y}+q+\mathbf{z}\ge 0,\\ & \mathbf{y}^T(Nx+M\mathbf{y}+q+\mathbf{z})=0,\end{array} \tag{4.12}$$

where $\mathbf{y}\in\Re^{mL},\mathbf{z}\in\Re^{mL}$, and

$$\mathbf{d}=\begin{pmatrix}p_1d\\ \vdots\\ p_Ld\end{pmatrix},\quad D=\begin{pmatrix}I\\ \vdots\\ I\end{pmatrix},\quad N=\begin{pmatrix}N_1\\ \vdots\\ N_L\end{pmatrix},\quad M=\begin{pmatrix}M_1 & & O\\ & \ddots & \\ O & & M_L\end{pmatrix},\quad q=\begin{pmatrix}q_1\\ \vdots\\ q_L\end{pmatrix}.$$

Therefore, we may employ the same terminologies as in the literature on MPECs. Suppose that $(x^*,y^*,\mathbf{y}^*,\mathbf{z}^*)$ is a feasible point of problem (4.12).

Definition 4.2 We say the MPEC-*linear independence constraint qualification* (MPEC-LICQ) to hold at $(x^*,y^*,\mathbf{y}^*,\mathbf{z}^*)$ if the system

$$\begin{array}{l}g(x,y)\le 0,\ h(x,y)=0,\\ \mathbf{y}-Dy=0,\ \mathbf{z}\ge 0,\\ \mathbf{y}\ge 0,\ Nx+M\mathbf{y}+q+\mathbf{z}\ge 0\end{array}$$

satisfies the linear independence constraint qualification (LICQ) at $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$.

This condition is not particularly stringent and has been assumed often in the literature on MPECs, see [5, 8, 9, 10, 11, 17]. Note that this definition is different from the standard definition of LICQ in nonlinear programming theory for problem (4.12) that would require the equality constraint $\mathbf{y}^T(Nx + M\mathbf{y} + q + \mathbf{z}) = 0$ to join the above system; however, the resulting system fails to satisfy the LICQ in any case actually [2].

Definition 4.3 We say $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ to be a *Bouligand or B-stationary* point of the MPEC (4.12) if

$$\mathbf{v}^T \begin{pmatrix} \nabla f(x^*, y^*) \\ 0 \\ \mathbf{d} \end{pmatrix} \geq 0, \qquad \forall \mathbf{v} \in \mathcal{T}(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*),$$

where $\mathcal{T}(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ stands for the tangent cone of the feasible region of problem (4.12) at $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$.

It is well-known [16] that, under the MPEC-LICQ, $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ is a B-stationary point of (4.12) if and only if there exist multiplier vectors $\lambda, \mu, \nu, \alpha, \beta$, and γ such that

$$\begin{pmatrix} \nabla_x f(x^*, y^*) \\ \nabla_y f(x^*, y^*) \\ 0 \\ \mathbf{d} \end{pmatrix} + \begin{pmatrix} \nabla_x g(x^*, y^*) \\ \nabla_y g(x^*, y^*) \\ 0 \\ 0 \end{pmatrix} \lambda + \begin{pmatrix} \nabla_x h(x^*, y^*) \\ \nabla_y h(x^*, y^*) \\ 0 \\ 0 \end{pmatrix} \mu$$

$$+ \begin{pmatrix} O \\ -D^T \\ I \\ O \end{pmatrix} \nu - \begin{pmatrix} O \\ O \\ O \\ I \end{pmatrix} \alpha - \begin{pmatrix} O \\ O \\ I \\ O \end{pmatrix} \beta - \begin{pmatrix} N^T \\ O \\ M^T \\ I \end{pmatrix} \gamma = 0, \tag{4.13}$$

$$0 \leq \lambda \perp (-g(x^*, y^*)) \geq 0, \tag{4.14}$$

$$0 \leq \alpha \perp \mathbf{z}^* \geq 0, \tag{4.15}$$

$$\beta \perp \mathbf{y}^* \geq 0, \tag{4.16}$$

$$\gamma \perp (Nx^* + M\mathbf{y}^* + q + \mathbf{z}^*) \geq 0, \tag{4.17}$$

$$\beta[i] \geq 0, \ \gamma[i] \geq 0, \quad \forall i \in \mathcal{I}^* := \{i \mid \ \mathbf{y}^*[i] = (Nx^* + M\mathbf{y}^* + q + \mathbf{z}^*)[i] = 0\}. \tag{4.18}$$

Moreover, for any feasible point (x, y) of problem (4.3), the point $(x, y, z_1, \cdots, z_L)$ with

$$z_\ell = \max(-(N_\ell x + M_\ell y + q_\ell), \ 0), \qquad \ell = 1, \cdots, L$$

is a feasible point of problem (4.12) and this point is eligible in the sense that, if $(x, y, \hat{z}_1, \cdots, \hat{z}_L)$ is another feasible point of (4.12), then

$$f(x, y) + \sum_{\ell=1}^{L} p_\ell d^T z_\ell \leq f(x, y) + \sum_{\ell=1}^{L} p_\ell d^T \hat{z}_\ell.$$

We further have the following theorem. Since the proof is somewhat lengthy, we give it in the appendix.

Theorem 4.1 *Suppose that (x^*, y^*) is a feasible point of problem (4.3). Let*

$$z_\ell^* = \max(-(N_\ell x^* + M_\ell y^* + q_\ell),\ 0), \qquad \ell = 1, \cdots, L \tag{4.19}$$

and

$$\mathbf{y}^* = ((y^*)^T, \cdots, (y^*)^T)^T, \qquad \mathbf{z}^* = ((z_1^*)^T, \cdots, (z_L^*)^T)^T. \tag{4.20}$$

If $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ is a B-stationary point of the MPEC (4.12) and the MPEC-LICQ holds at $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$, then (x^*, y^*) is a stationary point of problem (4.3). Conversely, if (x^*, y^*) is a stationary point of problem (4.3) and the MPEC-LICQ holds at $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$, then $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ is a B-stationary point of the MPEC (4.12).

In the rest of this section, we will propose some methods for solving problem (4.3). Note that, although (4.3) is no longer an MPEC and the function θ is convex, this problem may not be easy to deal with because firstly, the objective function is not differentiable everywhere, and secondly, since L is usually very large in practice, problem (4.3) has a great many constraints.

The following functions will be used later on: Let ϵ be a nonnegative constant. The functions $\phi_\epsilon : \Re \to (0, +\infty)$ and $\psi_\epsilon : \Re \to (0, +\infty)$ are defined by

$$\phi_\epsilon(t) := \frac{\sqrt{t^2 + \epsilon^2} + t}{2}$$

and

$$\psi_\epsilon(t) := \sqrt{t^2 + \epsilon^2},$$

respectively. It is obvious that, for any fixed t,

$$\lim_{\epsilon \downarrow 0} \phi_\epsilon(t) = \phi_0(t) = \max\{t, 0\}, \qquad \lim_{\epsilon \downarrow 0} \psi_\epsilon(t) = \psi_0(t) = |t|.$$

However, unlike the limit functions ϕ_0 and ψ_0, both ϕ_ϵ and ψ_ϵ are differentiable everywhere for each $\epsilon > 0$.

4.1 Smoothed penalty method — I

By means of the differentiable function ϕ_ϵ and with the help of a smoothed penalty technique, we obtain the following smooth approximation of problem (4.3):

$$\begin{array}{ll} \text{minimize} & f(x,y) + \vartheta_\epsilon(x,y) + \rho\delta_\epsilon(x,y) \\ \text{subject to} & g(x,y) \le 0,\ h(x,y) = 0,\ \ y \ge 0, \end{array} \tag{4.21}$$

where $\rho > 0$ is a penalty parameter and

$$\vartheta_\epsilon(x,y) := \sum_{\ell=1}^{L} \sum_{i=1}^{m} p_\ell d[i] \phi_\epsilon(-(N_\ell x + M_\ell y + q_\ell)[i]), \tag{4.22}$$

$$\delta_\epsilon(x,y) := \sum_{\ell=1}^{L} \sum_{i=1}^{m} \phi_\epsilon(y[i](N_\ell x + M_\ell y + q_\ell)[i]). \tag{4.23}$$

Note that, for any (x, y),

$$\begin{aligned} \vartheta_0(x,y) &= \theta(x,y), \\ \delta_0(x,y) &= \sum_{\ell=1}^{L}\sum_{i=1}^{m} \max\{y[i](N_\ell x + M_\ell y + q_\ell)[i],\ 0\}. \end{aligned}$$

Hence, when $\epsilon = 0$ and $\rho = \bar{\rho} > 0$, problem (4.21) reduces to

$$\begin{array}{ll} \text{minimize} & f(x,y) + \theta(x,y) + \bar{\rho}\delta_0(x,y) \\ \text{subject to} & g(x,y) \le 0,\ h(x,y) = 0,\quad y \ge 0. \end{array} \tag{4.24}$$

For problem (4.24), we will use a similar definition of stationarity to problem (4.3): We say $(x^*, y^*) \in \Re^{n+m}$ to be *stationary* to problem (4.24) if it is feasible and there exist Lagrange multiplier vectors λ, μ, and ν such that

$$0 \in \nabla f(x^*,y^*) + \partial\theta(x^*,y^*) + \bar{\rho}\partial\delta_0(x^*,y^*) + \nabla g(x^*,y^*)\lambda + \nabla h(x^*,y^*)\mu - \begin{pmatrix} O \\ I \end{pmatrix}\nu, \tag{4.25}$$

$$0 \le \lambda \perp (-g(x^*,y^*)) \ge 0, \tag{4.26}$$

$$0 \le \nu \perp \quad y^* \ge 0, \tag{4.27}$$

where $\partial\delta_0$ denotes Clarke subdifferential operator [3]. Let

$$\delta_{\ell,i}(x,y) := \max\left\{y[i](N_\ell x + M_\ell y + q_\ell)[i],\ 0\right\}, \qquad \forall\ell,\ \forall i.$$

Then, since the functions $\delta_{\ell,i}$ are Clarke regular [3], we have

$$\begin{aligned} &\partial\delta_{\ell,i}(x,y) \\ &= \begin{cases} \mathrm{co}\left\{(N_\ell x + M_\ell y + q_\ell)[i]\begin{pmatrix} 0 \\ e_i \end{pmatrix} + y[i]\begin{pmatrix} N_\ell[i] \\ M_\ell[i] \end{pmatrix}, 0\right\}, & y[i](N_\ell x + M_\ell y + q_\ell)[i] = 0 \\ \left\{(N_\ell x + M_\ell y + q_\ell)[i]\begin{pmatrix} 0 \\ e_i \end{pmatrix} + y[i]\begin{pmatrix} N_\ell[i] \\ M_\ell[i] \end{pmatrix}\right\}, & y[i](N_\ell x + M_\ell y + q_\ell)[i] > 0 \\ \left\{0\right\}, & y[i](N_\ell x + M_\ell y + q_\ell)[i] < 0 \end{cases} \end{aligned} \tag{4.28}$$

and

$$\partial\delta_0(x,y) = \sum_{\ell=1}^{L}\sum_{i=1}^{m} \partial\delta_{\ell,i}(x,y), \tag{4.29}$$

where $\partial\delta_{\ell,i}$ denotes Clarke subdifferential operator of $\delta_{\ell,i}$.

Theorem 4.2 *Let (x^*, y^*) be a stationary point of problem (4.3). Then, (x^*, y^*) is a stationary point of problem (4.24) for any $\bar{\rho}$ sufficiently large. Conversely, if (x^*, y^*) is a stationary point of problem (4.24), and $\delta_0(x^*, y^*) = 0$, i.e., $(x^*, y^*) \in \mathcal{F}_3$, then (x^*, y^*) is stationary to (4.3).*

Proof: (a) Suppose (x^*, y^*) is a stationary point of problem (4.3). We will show that, when $\bar{\rho}$ is sufficiently large, (x^*, y^*) is a stationary point of problem (4.24). By the stationarity of

(x^*, y^*) to problem (4.3), there exist multiplier vectors λ^*, μ^*, ν^*, and ξ_ℓ^*, $\ell = 1, \cdots, L$, satisfying conditions (4.6)–(4.9). Let

$$\lambda := \lambda^*, \qquad \mu := \mu^*, \qquad \nu := \nu^*.$$

Then (4.26) and (4.27) follow from (4.7) and (4.8) immediately. Comparing (4.25) with (4.6), in order to finish the proof, we only need to show that, when $\bar{\rho}$ is sufficiently large,

$$\sum_{\ell=1}^{L}\sum_{i=1}^{m} \xi_\ell^*[i]\Big((N_\ell x^* + M_\ell y^* + q_\ell)[i]\begin{pmatrix} 0 \\ e_i \end{pmatrix} + y^*[i]\begin{pmatrix} N_\ell[i] \\ M_\ell[i] \end{pmatrix}\Big) \in \bar{\rho}\partial\delta_0(x^*, y^*).$$

By (4.28) and (4.29), it is sufficient to show that, when $\bar{\rho}$ is sufficiently large,

$$\forall \ell,\ \forall i, \qquad \begin{cases} \xi_\ell^*[i] \in [0, \bar{\rho}], & y^*[i](N_\ell x^* + M_\ell y^* + q_\ell)[i] = 0, \\ \xi_\ell^*[i] = \bar{\rho}\ , & y^*[i](N_\ell x^* + M_\ell y^* + q_\ell)[i] > 0, \\ \xi_\ell^*[i] = 0\ , & y^*[i](N_\ell x^* + M_\ell y^* + q_\ell)[i] < 0. \end{cases}$$

Indeed, this follows from (4.9) and the fact that $y^*[i](N_\ell x^* + M_\ell y^* + q_\ell)[i] \leq 0$ for each ℓ and i.

(b) Suppose (x^*, y^*) is a stationary point of problem (4.24) and $\delta_0(x^*, y^*) = 0$. Then, there exist multiplier vectors λ, μ, and ν satisfying (4.25)–(4.27). Note that $(x^*, y^*) \in \mathcal{F}_3$ implies

$$y^*[i](N_\ell x^* + M_\ell y^* + q_\ell)[i] \leq 0, \qquad \forall \ell,\ \forall i. \tag{4.30}$$

It then follows that

$$\begin{aligned} &\partial\delta_{\ell,i}(x^*, y^*) \\ = &\begin{cases} \mathrm{co}\Big\{(N_\ell x^* + M_\ell y^* + q_\ell)[i]\begin{pmatrix} 0 \\ e_i \end{pmatrix} + y^*[i]\begin{pmatrix} N_\ell[i] \\ M_\ell[i] \end{pmatrix}, 0\Big\}, & y^*[i](N_\ell x^* + M_\ell y^* + q_\ell)[i] = 0 \\ \qquad\Big\{\, 0 \,\Big\}\ , & y^*[i](N_\ell x^* + M_\ell y^* + q_\ell)[i] < 0 \end{cases} \end{aligned}$$

for any ℓ and i, and

$$\partial\delta_0(x^*, y^*) = \sum_{\ell=1}^{L}\sum_{i=1}^{m} \partial\delta_{\ell,i}(x^*, y^*).$$

Condition (4.25) means that there exist multiplier vectors ξ_ℓ, $\ell = 1, \cdots, L$, such that

$$\forall \ell,\ \forall i, \qquad \begin{cases} \xi_\ell[i] \in [0, \bar{\rho}], & y^*[i](N_\ell x^* + M_\ell y^* + q_\ell)[i] = 0 \\ \xi_\ell[i] = 0\ , & y^*[i](N_\ell x^* + M_\ell y^* + q_\ell)[i] < 0 \end{cases} \tag{4.31}$$

and

$$\begin{aligned} 0 \in \nabla f(x^*, y^*) + \partial\theta(x^*, y^*) + \nabla g(x^*, y^*)\lambda + \nabla h(x^*, y^*)\mu - \begin{pmatrix} O \\ I \end{pmatrix}\nu \\ + \sum_{\ell=1}^{L}\sum_{i=1}^{m} \xi_\ell[i]\Big((N_\ell x^* + M_\ell y^* + q_\ell)[i]\begin{pmatrix} 0 \\ e_i \end{pmatrix} + y^*[i]\begin{pmatrix} N_\ell[i] \\ M_\ell[i] \end{pmatrix}\Big). \end{aligned} \tag{4.32}$$

Let

$$\lambda^* := \lambda, \quad \mu^* := \mu, \quad \nu^* := \nu$$

and

$$\xi_\ell^* := \xi_\ell, \qquad \ell = 1, \cdots, L.$$

Then (4.6)–(4.8) follow from (4.32) and (4.26)–(4.27), and (4.9) follows from (4.30)–(4.31). Therefore, (x^*, y^*) is a stationary point of problem (4.3). ∎

We then have the following algorithm for problem (4.3).

Algorithm SP-I:

Step 1: Choose $\epsilon_0 > 0$ and $\rho_0 > 0$. Set $k := 0$.

Step 2: Solve problem (4.21) with $\epsilon = \epsilon_k$ and $\rho = \rho_k$ to get a stationary point (x^k, y^k) and go to Step 3.

Step 3: If a stopping rule is satisfied, then terminate. Otherwise, choose $\epsilon_{k+1} \in (0, \epsilon_k)$ and $\rho_{k+1} \geq \rho_k$. Go to Step 2 with $k := k + 1$.

In what follows, we suppose that the sequences $\{\epsilon_k\}$ and $\{\rho_k\}$ satisfy

$$\lim_{k\to\infty} \epsilon_k = 0, \qquad \lim_{k\to\infty} \rho_k = \bar{\rho}, \tag{4.33}$$

where $\bar{\rho}$ is a sufficiently large constant. Recall that $\mathcal{F}_3$ denotes the feasible region of problem (4.1), which is the same as that of problem (4.3).

We next investigate the limiting behavior of the sequence $\{(x^k, y^k)\}$ generated by Algorithm SP-I. The convergence result can be stated as follows.

Theorem 4.3 *Suppose that Algorithm SP-I generates a sequence $\{(x^k, y^k)\}$ of stationary points of (4.21) with $\epsilon = \epsilon_k$ and $\rho = \rho_k$. For any accumulation point (x^*, y^*) of the sequence $\{(x^k, y^k)\}$, if the system*

$$g(x, y) \leq 0, \quad h(x, y) = 0, \quad y \geq 0 \tag{4.34}$$

satisfies the Mangasarian-Fromovitz constraint qualification (MFCQ) at (x^*, y^*), then (x^*, y^*) is a stationary point of problem (4.24). Furthermore, if $\delta_0(x^*, y^*) = 0$, then (x^*, y^*) is a stationary point of problem (4.3).

Proof: Assume without loss of generality that $\lim_{k\to\infty}(x^k, y^k) = (x^*, y^*)$. We will show that (x^*, y^*) is stationary to problem (4.24), i.e., there exist multiplier vectors λ, μ, and ν such that (4.25)–(4.27) hold.

First of all, by the stationarity of (x^k, y^k) for problem (4.21) with $\epsilon = \epsilon_k$ and $\rho = \rho_k$, there exist Lagrange multiplier vectors λ^k, μ^k, and ν^k such that

$$\begin{aligned} &\nabla f(x^k,y^k) + \nabla\vartheta_{\epsilon_k}(x^k,y^k) + \rho_k \nabla\delta_{\epsilon_k}(x^k,y^k) \\ &\qquad\qquad + \nabla g(x^k,y^k)\lambda^k + \nabla h(x^k,y^k)\mu^k - \begin{pmatrix} O \\ I \end{pmatrix} \nu^k = 0, \qquad (4.35) \\ &0 \le \lambda^k \perp (-g(x^k,y^k)) \ge 0, \qquad (4.36) \\ &0 \le \nu^k \perp \quad y^k \ge 0. \qquad (4.37) \end{aligned}$$

Note that, by (4.22) and (4.23),

$$\begin{aligned} \nabla\vartheta_{\epsilon_k}(x^k,y^k) &= \sum_{\ell=1}^{L}\sum_{i=1}^{m} p_\ell d[i] \phi'_{\epsilon_k}(-(N_\ell x^k + M_\ell y^k + q_\ell)[i]) \begin{pmatrix} -N_\ell[i] \\ -M_\ell[i] \end{pmatrix}, \qquad (4.38) \\ \nabla\delta_{\epsilon_k}(x^k,y^k) &= \sum_{\ell=1}^{L}\sum_{i=1}^{m} \phi'_{\epsilon_k}(y^k[i](N_\ell x^k + M_\ell y^k + q_\ell)[i]) \\ &\qquad \Big((N_\ell x^k + M_\ell y^k + q_\ell)[i] \begin{pmatrix} 0 \\ e_i \end{pmatrix} + y^k[i] \begin{pmatrix} N_\ell[i] \\ M_\ell[i] \end{pmatrix} \Big), \qquad (4.39) \end{aligned}$$

where

$$\phi'_{\epsilon_k}(t) = \frac{1}{2}\Big(\frac{t}{\sqrt{t^2+\epsilon_k^2}} + 1 \Big), \qquad \forall\, t \in \Re. \qquad (4.40)$$

We can then rewrite (4.35) as

$$\begin{aligned} &-\nabla f(x^k,y^k) - \nabla g(x^k,y^k)\lambda^k - \nabla h(x^k,y^k)\mu^k + \begin{pmatrix} O \\ I \end{pmatrix} \nu^k \\ = &\sum_{\ell=1}^{L}\sum_{i=1}^{m} \xi_\ell^k[i] \Big((N_\ell x^k + M_\ell y^k + q_\ell)[i] \begin{pmatrix} 0 \\ e_i \end{pmatrix} + y^k[i] \begin{pmatrix} N_\ell[i] \\ M_\ell[i] \end{pmatrix} \Big) \\ &+ \sum_{\ell=1}^{L}\sum_{i=1}^{m} p_\ell d[i] \sigma_\ell^k[i] \begin{pmatrix} -N_\ell[i] \\ -M_\ell[i] \end{pmatrix}, \qquad (4.41) \end{aligned}$$

where

$$\begin{aligned} \xi_\ell^k[i] &:= \rho_k \phi'_{\epsilon_k}(y^k[i](N_\ell x^k + M_\ell y^k + q_\ell)[i]), \qquad (4.42) \\ \sigma_\ell^k[i] &:= \phi'_{\epsilon_k}(-(N_\ell x^k + M_\ell y^k + q_\ell)[i]). \qquad (4.43) \end{aligned}$$

We next prove that the sequences $\{\lambda^k\}$, $\{\mu^k\}$, and $\{\nu^k\}$ are bounded. For the contradiction purpose, let

$$\tau_k := \sum_{i=1}^{m} (\lambda^k[i] + |\mu^k[i]| + \nu^k[i]), \qquad (4.44)$$

and suppose $\lim_{k\to\infty} \tau_k = +\infty$. Taking a subsequence if necessary, we may assume that the limits

$$\lambda'[i] := \lim_{k\to\infty} \frac{\lambda^k[i]}{\tau_k}, \quad \mu'[i] := \lim_{k\to\infty} \frac{\mu^k[i]}{\tau_k}, \quad \nu'[i] := \lim_{k\to\infty} \frac{\nu^k[i]}{\tau_k}, \qquad i = 1, \cdots, m$$

exist. It is clear from (4.44) that

$$\sum_{i=1}^{m}(\lambda'[i] + |\mu'[i]| + \nu'[i]) = 1.$$

It follows from (4.33), (4.40), and (4.42)–(4.43) that both $\{\xi_\ell^k[i]\}$ and $\{\sigma_\ell^k[i]\}$ are bounded for each ℓ and each i. Thus, dividing (4.41) by τ_k and taking a limit, we get

$$-\nabla g(x^*, y^*)\lambda' - \nabla h(x^*, y^*)\mu' + \begin{pmatrix} O \\ I \end{pmatrix} \nu' = 0.$$

Furthermore, taking (4.36) and (4.37) into account, we obtain $\lambda' \geq 0, \nu' \geq 0$, and

$$\begin{aligned} \lambda'[i] &= 0, \qquad i \notin \mathcal{I}_g(x^*, y^*), \\ \nu'[i] &= 0, \qquad i \notin \mathcal{I}_\pi(x^*, y^*), \end{aligned}$$

where $\pi : \Re^{n+m} \to \Re^m$ is given by $\pi(x, y) = y$. Thus we have

$$\begin{aligned} \sum_{i \in \mathcal{I}_g(x^*,y^*)} \lambda'[i] + \sum_{i=1}^{s_2} \mu'[i] + \sum_{i \in \mathcal{I}_\pi(x^*,y^*)} \nu'[i] = 1, \\ \lambda'[i] \geq 0, \qquad i \in \mathcal{I}_g(x^*, y^*), \\ \nu'[i] \geq 0, \qquad i \in \mathcal{I}_\pi(x^*, y^*), \end{aligned}$$

and

$$-\sum_{i \in \mathcal{I}_g(x^*,y^*)} \lambda'[i]\nabla g_i(x^*, y^*) - \sum_{i=1}^{s_2} \mu'[i]\nabla h_i(x^*, y^*) + \sum_{i \in \mathcal{I}_\pi(x^*,y^*)} \nu'[i] \begin{pmatrix} 0 \\ e_i \end{pmatrix} = 0.$$

This contradicts the assumption that the system (4.34) satisfies the MFCQ at (x^*, y^*). Hence all the sequences $\{\lambda^k\}$, $\{\mu^k\}$, and $\{\nu^k\}$ are bounded. Since $\{\xi_\ell^k[i]\}$ and $\{\sigma_\ell^k[i]\}$ are bounded for any ℓ and i, we may assume, without loss of generality, that the following limits exist:

$$\lambda := \lim_{k\to\infty} \lambda^k, \quad \mu := \lim_{k\to\infty} \mu^k, \quad \nu := \lim_{k\to\infty} \nu^k,$$

and

$$\xi_\ell := \lim_{k\to\infty} \xi_\ell^k, \qquad \sigma_\ell := \lim_{k\to\infty} \sigma_\ell^k, \qquad \forall \ell.$$

Taking a limit in (4.36), (4.37), and (4.41), we obtain (4.26), (4.27), and

$$\begin{aligned} & -\nabla f(x^*, y^*) - \nabla g(x^*, y^*)\lambda - \nabla h(x^*, y^*)\mu + \begin{pmatrix} O \\ I \end{pmatrix} \nu \\ = \ & \sum_{\ell=1}^{L}\sum_{i=1}^{m} \xi_\ell[i]\Big((N_\ell x^* + M_\ell y^* + q_\ell)[i] \begin{pmatrix} 0 \\ e_i \end{pmatrix} + y^*[i] \begin{pmatrix} N_\ell[i] \\ M_\ell[i] \end{pmatrix}\Big) \\ & + \sum_{\ell=1}^{L}\sum_{i=1}^{m} p_\ell d[i]\sigma_\ell[i] \begin{pmatrix} -N_\ell[i] \\ -M_\ell[i] \end{pmatrix}. \end{aligned} \qquad (4.45)$$

Thus, in order to show that (x^*, y^*) is a stationary point of problem (4.24), we only need to prove that the vector on the right-hand side of (4.45) belongs to the set $\bar{\rho}\partial\delta_0(x^*, y^*) + \partial\theta(x^*, y^*)$.

(i) We first prove that

$$\sum_{\ell=1}^{L}\sum_{i=1}^{m} \xi_\ell[i]\Big((N_\ell x^* + M_\ell y^* + q_\ell)[i] \begin{pmatrix} 0 \\ e_i \end{pmatrix} + y^*[i] \begin{pmatrix} N_\ell[i] \\ M_\ell[i] \end{pmatrix}\Big) \in \bar{\rho}\partial\delta_0(x^*, y^*). \tag{4.46}$$

By (4.29), it is sufficient to show that, for any ℓ and i,

$$\xi_\ell[i]\Big((N_\ell x^* + M_\ell y^* + q_\ell)[i] \begin{pmatrix} 0 \\ e_i \end{pmatrix} + y^*[i] \begin{pmatrix} N_\ell[i] \\ M_\ell[i] \end{pmatrix}\Big) \in \bar{\rho}\partial\delta_{\ell,i}(x^*, y^*),$$

which, by (4.28), is equivalent to

$$\begin{cases} \xi_\ell[i] \in [0, \bar{\rho}], & y^*[i](N_\ell x^* + M_\ell y^* + q_\ell)[i] = 0, \\ \xi_\ell[i] = \bar{\rho}\ , & y^*[i](N_\ell x^* + M_\ell y^* + q_\ell)[i] > 0, \\ \xi_\ell[i] = 0\ , & y^*[i](N_\ell x^* + M_\ell y^* + q_\ell)[i] < 0. \end{cases} \tag{4.47}$$

In fact, we can obtain (4.47) immediately from (4.33) and the facts that

$$\lim_{k\to\infty} (x^k, y^k) = (x^*, y^*)$$

and

$$\xi_\ell^k[i] = \frac{\rho_k}{2}\Big(\frac{y^k[i](N_\ell x^k + M_\ell y^k + q_\ell)[i]}{\sqrt{(y^k[i](N_\ell x^k + M_\ell y^k + q_\ell)[i])^2 + \epsilon_k^2}} + 1\Big),$$

and hence (4.46) must hold.

(ii) We next prove that

$$\sum_{\ell=1}^{L}\sum_{i=1}^{m} p_\ell d[i]\sigma_\ell[i] \begin{pmatrix} -N_\ell[i] \\ -M_\ell[i] \end{pmatrix} \in \partial\theta(x^*, y^*).$$

By (4.11), it is enough to show

$$\sigma_\ell[i] \begin{pmatrix} -N_\ell[i] \\ -M_\ell[i] \end{pmatrix} \in \partial\theta_{\ell,i}(x^*, y^*), \qquad \forall \ell,\ \forall i. \tag{4.48}$$

There are three cases:

(iia) Suppose $(N_\ell x^* + M_\ell y^* + q_\ell)[i] = 0$. We then have from (4.40) and (4.43) that

$$0 \le \sigma_\ell^k[i] \le 1, \qquad \forall k.$$

Passing to the limit yields $0 \le \sigma_\ell[i] \le 1$ and hence

$$\sigma_\ell[i] \begin{pmatrix} -N_\ell[i] \\ -M_\ell[i] \end{pmatrix} \in \mathrm{co}\Big\{ \begin{pmatrix} -N_\ell[i] \\ -M_\ell[i] \end{pmatrix}, 0 \Big\} = \partial\theta_{\ell,i}(x^*, y^*),$$

where the equality follows from (**??**).

(iib) Suppose $(N_\ell x^* + M_\ell y^* + q_\ell)[i] < 0$. Note that

$$\sigma_\ell^k[i] = \frac{1}{2}\Big(\frac{-(N_\ell x^k + M_\ell y^k + q_\ell)[i]}{\sqrt{((N_\ell x^k + M_\ell y^k + q_\ell)[i])^2 + \epsilon_k^2}} + 1\Big), \qquad \forall k.$$

Taking a limit in the above equality, we obtain $\sigma_\ell[i] = 1$ immediately and so

$$\sigma_\ell[i]\begin{pmatrix} -N_\ell[i] \\ -M_\ell[i] \end{pmatrix} = \begin{pmatrix} -N_\ell[i] \\ -M_\ell[i] \end{pmatrix} \in \partial\theta_{\ell,i}(x^*, y^*).$$

(iic) Suppose $(N_\ell x^* + M_\ell y^* + q_\ell)[i] > 0$. It is easy to show that, for any k,

$$0 \le \sigma_\ell^k[i] \le \frac{\epsilon_k}{2\Big(\sqrt{((N_\ell x^k + M_\ell y^k + q_\ell)[i])^2 + \epsilon_k^2} + (N_\ell x^k + M_\ell y^k + q_\ell)[i]\Big)}.$$

Letting $k \to \infty$, we see that $\sigma_\ell[i] = 0$ and so

$$\sigma_\ell[i]\begin{pmatrix} -N_\ell[i] \\ -M_\ell[i] \end{pmatrix} = 0 \in \partial\theta_{\ell,i}(x^*, y^*).$$

Consequently, (4.48) holds in each case. (i) and (ii) indicate that the vector on the right-hand side of (4.45) belongs to the set $\bar\rho\partial\delta_0(x^*, y^*) + \partial\theta(x^*, y^*)$. This completes the proof of the first part of the theorem. The second half readily follows from Theorem 4.2. ∎

4.2 Smoothed penalty method — II

In the last subsection, making use of the function ϕ_ϵ, we obtained the smooth problem (4.21). By applying a similar smoothed penalty technique to the constraints of problem (4.21), we may further get the following unconstrained smooth problem:

$$\text{minimize} \qquad f(x,y) + \vartheta_{\epsilon_k}(x,y) + \rho_k\delta_{\epsilon_k}(x,y) + \rho_k r_{\epsilon_k}(x,y), \tag{4.49}$$

where

$$r_{\epsilon_k}(x,y) := \sum_{i=1}^{s_1}\phi_{\epsilon_k}(g_i(x,y)) + \sum_{i=1}^{s_2}\psi_{\epsilon_k}(h_i(x,y)) + \sum_{i=1}^{m}\phi_{\epsilon_k}(-y[i]). \tag{4.50}$$

Corresponding to problem (4.24), we have a nonsmooth approximation of problem (4.3):

$$\text{minimize} \qquad f(x,y) + \theta(x,y) + \bar\rho\delta_0(x,y) + \bar\rho r_0(x,y), \tag{4.51}$$

where

$$r_0(x,y) := \sum_{i=1}^{s_1}\max\{(g_i(x,y)),\ 0\} + \sum_{i=1}^{s_2}|h_i(x,y)| + \sum_{i=1}^{m}\max\{-y[i],\ 0\}.$$

We say $(x^*, y^*) \in \Re^{n+m}$ to be *stationary* to problem (4.51) if

$$0 \in \nabla f(x^*, y^*) + \partial\theta(x^*, y^*) + \bar\rho\partial\delta_0(x^*, y^*) + \bar\rho\partial r_0(x^*, y^*). \tag{4.52}$$

Theorem 4.4 *Let (x^*,y^*) be a stationary point of problem (4.3). Then, (x^*,y^*) is a stationary point of problem (4.51) for any $\bar\rho$ sufficiently large. Conversely, if (x^*,y^*) is a stationary point of problem (4.51), and $\delta_0(x^*,y^*)+r_0(x^*,y^*)=0$, i.e., $(x^*,y^*)\in\mathcal{F}_3$, then (x^*,y^*) is stationary to problem (4.3).*

The proof of this theorem is similar to that of Theorem 4.2 and so it is omitted here.

A new algorithm for solving problem (4.3) can be stated as follows.

Algorithm SP-II:

Step 1: Choose $\epsilon_0>0$ and $\rho_0>0$. Set $k:=0$.

Step 2: Solve problem (4.49) to get a stationary point, say (x^k,y^k), and go to Step 3.

Step 3: If a stopping rule is satisfied, then terminate. Otherwise, choose $\epsilon_{k+1}\in(0,\epsilon_k)$ and $\rho_{k+1}\ge\rho_k$. Go to Step 2 with $k:=k+1$.

We suppose that the sequences $\{\epsilon_k\}$ and $\{\rho_k\}$ also satisfy condition (4.33). Then, we have the following convergence result for Algorithm SP-II.

Theorem 4.5 *Suppose that Algorithm SP-II generates a sequence $\{(x^k,y^k)\}$ of stationary points of (4.49). Then any accumulation point (x^*,y^*) of the sequence $\{(x^k,y^k)\}$ is a stationary point of problem (4.51). Furthermore, if $\delta_0(x^*,y^*)+r_0(x^*,y^*)=0$, then (x^*,y^*) is a stationary point of problem (4.3).*

Proof: Assume without loss of generality that $\lim_{k\to\infty}(x^k,y^k)=(x^*,y^*)$. Since the second part of the theorem follows from the first part and Theorem 4.4 directly, we only prove the first part of the theorem, namely, we will show (4.52).

Firstly, it follows from the stationarity of (x^k,y^k) that

$$\nabla f(x^k,y^k)+\nabla\vartheta_{\epsilon_k}(x^k,y^k)+\rho_k\nabla\delta_{\epsilon_k}(x^k,y^k)+\rho_k\nabla r_{\epsilon_k}(x^k,y^k)=0. \tag{4.53}$$

Note that, by (4.50),

$$\begin{aligned}\nabla r_{\epsilon_k}(x^k,y^k) &= \sum_{i=1}^{s_1}\phi'_{\epsilon_k}(g_i(x^k,y^k))\nabla g_i(x^k,y^k)\\ &\quad+\sum_{i=1}^{s_2}\psi'_{\epsilon_k}(h_i(x^k,y^k))\nabla h_i(x^k,y^k)-\sum_{i=1}^{m}\phi'_{\epsilon_k}(-y^k[i])\begin{pmatrix}0\\ e_i\end{pmatrix},\end{aligned} \tag{4.54}$$

where ϕ'_{ϵ_k} is given by (4.40) and

$$\psi'_{\epsilon_k}(t)=\frac{t}{\sqrt{t^2+\epsilon_k^2}},\qquad\forall\ t\in\Re. \tag{4.55}$$

Substituting (4.38), (4.39), and (4.54) into (4.53), we obtain

$$-\nabla f(x^k,y^k) = \nabla g(x^k,y^k)\lambda^k+\nabla h(x^k,y^k)\mu^k-\begin{pmatrix}O\\ I\end{pmatrix}\nu^k$$

$$
\begin{aligned}
&+\sum_{\ell=1}^{L}\sum_{i=1}^{m}\xi_\ell^k[i]\Big((N_\ell x^k+M_\ell y^k+q_\ell)[i]\begin{pmatrix}0\\ e_i\end{pmatrix}+y^k[i]\begin{pmatrix}N_\ell[i]\\ M_\ell[i]\end{pmatrix}\Big)\\
&+\sum_{\ell=1}^{L}\sum_{i=1}^{m}p_\ell d[i]\sigma_\ell^k[i]\begin{pmatrix}-N_\ell[i]\\ -M_\ell[i]\end{pmatrix},
\end{aligned} \tag{4.56}
$$

where the multiplier vectors $\lambda^k,\mu^k,\nu^k,\xi_\ell^k$, and σ_ℓ^k are given by

$$
\begin{aligned}
\lambda^k[i] &:= \rho_k\phi'_{\epsilon_k}(g_i(x^k,y^k)), && i=1,\cdots,s_1,\\
\mu^k[i] &:= \rho_k\psi'_{\epsilon_k}(h_i(x^k,y^k)), && i=1,\cdots,s_2,\\
\nu^k[i] &:= \rho_k\phi'_{\epsilon_k}(-y^k[i]), && i=1,\cdots,m,\\
\xi_\ell^k[i] &:= \rho_k\phi'_{\epsilon_k}(y^k[i](N_\ell x^k+M_\ell y^k+q_\ell)[i]), && i=1,\cdots,m,\ \ell=1,\cdots,L,\\
\sigma_\ell^k[i] &:= \phi'_{\epsilon_k}(-(N_\ell x^k+M_\ell y^k+q_\ell)[i]), && i=1,\cdots,m,\ \ell=1,\cdots,L,
\end{aligned}
$$

respectively. Since $\{\rho_k\}$ is bounded, it follows from (4.40) and (4.55) that all the multiplier vectors are bounded. Therefore, we may assume without loss of generality that the following limits exist:

$$
\lambda:=\lim_{k\to\infty}\lambda^k,\quad \mu:=\lim_{k\to\infty}\mu^k,\quad \nu:=\lim_{k\to\infty}\nu^k,
$$

and

$$
\xi_\ell:=\lim_{k\to\infty}\xi_\ell^k,\qquad \sigma_\ell:=\lim_{k\to\infty}\sigma_\ell^k,\qquad \forall\ell.
$$

Taking a limit in (4.56) yields

$$
\begin{aligned}
-\nabla f(x^*,y^*) \;=\;& \nabla g(x^*,y^*)\lambda+\nabla h(x^*,y^*)\mu-\begin{pmatrix}O\\ I\end{pmatrix}\nu\\
&+\sum_{\ell=1}^{L}\sum_{i=1}^{m}\xi_\ell[i]\Big((N_\ell x^*+M_\ell y^*+q_\ell)[i]\begin{pmatrix}0\\ e_i\end{pmatrix}+y^*[i]\begin{pmatrix}N_\ell[i]\\ M_\ell[i]\end{pmatrix}\Big)\\
&+\sum_{\ell=1}^{L}\sum_{i=1}^{m}p_\ell d[i]\sigma_\ell[i]\begin{pmatrix}-N_\ell[i]\\ -M_\ell[i]\end{pmatrix}.
\end{aligned} \tag{4.57}
$$

Thus, in order to show that (x^*,y^*) is a stationary point of problem (4.51), we only need to prove that the vector on the right-hand side of (4.57) belongs to the set

$$
\bar\rho\partial r_0(x^*,y^*)+\bar\rho\partial\delta_0(x^*,y^*)+\partial\theta(x^*,y^*).
$$

In a similar way to the proof of Theorem 4.3, we can show that

$$
\sum_{\ell=1}^{L}\sum_{i=1}^{m}\xi_\ell[i]\Big((N_\ell x^*+M_\ell y^*+q_\ell)[i]\begin{pmatrix}0\\ e_i\end{pmatrix}+y^*[i]\begin{pmatrix}N_\ell[i]\\ M_\ell[i]\end{pmatrix}\Big)\in\bar\rho\partial\delta_0(x^*,y^*)
$$

and

$$
\sum_{\ell=1}^{L}\sum_{i=1}^{m}p_\ell d[i]\sigma_\ell[i]\begin{pmatrix}-N_\ell[i]\\ -M_\ell[i]\end{pmatrix}\in\partial\theta(x^*,y^*).
$$

Now let us prove that

$$\nabla g(x^*, y^*)\lambda + \nabla h(x^*, y^*)\mu - \begin{pmatrix} O \\ I \end{pmatrix} \nu \in \bar{\rho}\partial r_0(x^*, y^*). \tag{4.58}$$

To this end, it is enough to show

$$i = 1, \cdots, s_1, \quad \begin{cases} \lambda[i] \in [0, \bar{\rho}], & g_i(x^*, y^*) = 0 \\ \lambda[i] = \bar{\rho}\ , & g_i(x^*, y^*) > 0 \\ \lambda[i] = 0\ , & g_i(x^*, y^*) < 0, \end{cases} \tag{4.59}$$

$$i = 1, \cdots, s_2, \quad \begin{cases} \mu[i] \in [-\bar{\rho}, \bar{\rho}], & h_i(x^*, y^*) = 0 \\ \mu[i] = \bar{\rho}\ , & h_i(x^*, y^*) > 0 \\ \mu[i] = -\bar{\rho}\ , & h_i(x^*, y^*) < 0, \end{cases} \tag{4.60}$$

and

$$i = 1, \cdots, m, \quad \begin{cases} \nu[i] \in [0, \bar{\rho}], & y^*[i] = 0 \\ \nu[i] = \bar{\rho}\ , & y^*[i] < 0 \\ \nu[i] = 0\ , & y^*[i] > 0. \end{cases} \tag{4.61}$$

In fact, (4.59)–(4.61) follow from (4.33) and the facts that

$$\lim_{k\to\infty} (x^k, y^k) = (x^*, y^*)$$

and

$$\begin{aligned} \lambda^k[i] &= \frac{\rho_k}{2}\Big(\frac{g_i(x^k, y^k)}{\sqrt{(g_i(x^k, y^k))^2 + \epsilon_k^2}} + 1\Big), \\ \mu^k[i] &= \frac{\rho_k h_i(x^k, y^k)}{\sqrt{(h_i(x^k, y^k))^2 + \epsilon_k^2}}, \\ \nu^k[i] &= \frac{\rho_k}{2}\Big(\frac{-y^k[i]}{\sqrt{(-y^k[i])^2 + \epsilon_k^2}} + 1\Big), \end{aligned}$$

and hence (4.58) must hold. This completes the proof. ∎

4.3 Numerical results for Example 4.1

We have tested the proposed methods on Example 4.1. In our experiments, we employed the MATLAB 6.5 built-in solver function *fmincon* to solve the constrained subproblems (4.21) and used *fminunc* to solve the unconstrained subproblems (4.49). For both SP-I and SP-II, we set $\epsilon_0 = 10^{-2}$, $\rho_0 = 10^3$, and updated these parameters by $\epsilon_{k+1} = 10^{-1}\epsilon_k$ and $\rho_{k+1} = \min\{10\rho_k,\ \bar{\rho}\}$, respectively. Moreover, the constants p and $\bar{\rho}$ are set to be 0.25 and 10^5, respectively. In addition, the initial point is chosen to be $y^0 = (0,0)$ and the computed solution y^k at the kth iteration is used as the starting point for the next iteration.

The computational results for Example 4.1 by SP-I and SP-II are reported in Tables 1 and 2, respectively. In the tables, *Ite* stands for the number of iterations spent by *fmincon* or *fminunc* to solve the subproblems. The results shown in the tables reveal that the proposed methods are able to solve Example 4.1 successfully.

Table 1: Computational Results by SP-I

ϵ_k	ρ_k	y^k	Ite
10^{-2}	10^3	(1.4206,3.4292)	22
10^{-3}	10^4	(1.4744,3.4780)	6
10^{-4}	10^5	(1.4919,3.4931)	7
10^{-5}	10^5	(1.4992,3.4993)	12
10^{-6}	10^5	(1.4999,3.4999)	6
10^{-7}	10^5	(1.5000,3.5000)	6

Table 2: Computational Results by SP-II

ϵ_k	ρ_k	y^k	Ite
10^{-2}	10^3	(1.4171,3.4339)	34
10^{-3}	10^4	(1.4744,3.4780)	9
10^{-4}	10^5	(1.4919,3.4931)	11
10^{-5}	10^5	(1.4992,3.4993)	12
10^{-6}	10^5	(1.4999,3.4999)	7
10^{-7}	10^5	(1.5000,3.5000)	20

5 Conclusions

A class of stochastic mathematical programs with linear complementarity constraints, called the here-and-now model, has been dealt with in this paper. We have presented a number of reformulations of the problem and then, based on these reformulations, proposed two smoothed penalty methods. Comprehensive convergence theory for the two methods have been established.

References

[1] J.R. Birge and F. Louveaux, Introduction to Stochastic Programming, Springer, New York, 1997.

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

[3] F.H. Clarke, Optimization and Nonsmooth Analysis, SIAM, Philadelphia, 1990.

[4] R.W. Cottle, J.S. Pang, and R.E. Stone, The Linear Complementarity Problem, Academic Press, New York, NY, 1992.

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

[6] M. Fukushima and P. Tseng, *An implementable active-set algorithm for computing a B-stationary point of the mathematical program with linear complementarity constraints*, SIAM Journal on Optimization, 12 (2002), 724-739.

[7] P. Kall and S.W. Wallace, Stochastic Programming, John Wiley & Sons, Chichester, 1994.

[8] G.H. Lin, X. Chen and M. Fukushima, *Smoothing implicit programming approaches for stochastic mathematical programs with linear complementarity constraints*, Technical Report 2003-006, Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto, Japan, 2003.

[9] G.H. Lin and M. Fukushima, *A modified relaxation scheme for mathematical programs with complementarity constraints*, Annals of Operations Research, to appear.

[10] G.H. Lin and M. Fukushima, *A new relaxation method for mathematical programs with complementarity constraints*, Journal of Optimization Theory and Applications, 118 (2003), 81–116.

[11] G.H. Lin and M. Fukushima, *Hybrid algorithms with active set identification for mathematical programs with complementarity constraints*, Technical Report 2002-008, Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto, Japan, 2002.

[12] Z.Q. Luo, J.S. Pang, and D. Ralph, Mathematical Programs with Equilibrium Constraints, Cambridge University Press, Cambridge, United Kingdom, 1996.

[13] J.S. Pang and M. Fukushima, *Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints*, Computational Optimization and Applications, 13 (1999), 111–136.

[14] M. Patriksson and L. Wynter, *Stochastic mathematical programs with equilibrium constraints*, Operations Research Letters, 25 (1999), 159–167.

[15] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, NJ, 1970.

[16] H.S. Scheel and S. Scholtes, *Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity*, Mathematics of Operations Research, 25 (2000), 1–22.

[17] S. Scholtes, *Convergence properties of a regularization scheme for mathematical programs with complementarity constraints*, SIAM Journal on Optimization, 11 (2001), 918–936.

Appendix: Proof of Theorem 4.1

(a) Suppose that $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ is a B-stationary point of problem (4.12) and the MPEC-LICQ holds at $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$. Then there exist multiplier vectors $\lambda, \mu, \nu, \alpha, \beta$, and γ satisfying (4.13)–(4.18). We will show that there exist multiplier vectors λ^*, μ^*, ν^*, and $\xi^*_\ell,\ \ell = 1, \cdots, L$, such that conditions (4.6)–(4.9) hold. To this end, we denote

$$\nu := \begin{pmatrix} \nu_1 \\ \vdots \\ \nu_L \end{pmatrix}, \quad \alpha := \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_L \end{pmatrix}, \quad \beta := \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_L \end{pmatrix}, \quad \gamma := \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_L \end{pmatrix},$$

and let

$$\lambda^* := \lambda, \qquad \mu^* := \mu, \tag{A.1}$$

and, for each ℓ and i,

$$\xi^*_\ell[i] := \begin{cases} \frac{\alpha_\ell[i]}{y^*[i]}, & y^*[i] > 0,\ (N_\ell x^* + M_\ell y^* + q_\ell)[i] = 0, \\ \frac{|\sum_{\ell=1}^L \beta_\ell[i]|}{(N_\ell x^* + M_\ell y^* + q_\ell)[i]}, & y^*[i] = 0,\ (N_\ell x^* + M_\ell y^* + q_\ell)[i] > 0, \\ 0, & \text{otherwise}, \end{cases} \tag{A.2}$$

$$\sigma^*_\ell[i] := \begin{cases} 1 - \frac{\alpha_\ell[i]}{p_\ell d[i]}, & y^*[i] = 0, \\ 1, & y^*[i] > 0, \end{cases} \tag{A.3}$$

$$\nu^* := \sum_{\ell=1}^L \beta_\ell + \begin{pmatrix} \sum_{\ell=1}^L \xi^*_\ell[1](N_\ell x^* + M_\ell y^* + q_\ell)[1] \\ \vdots \\ \sum_{\ell=1}^L \xi^*_\ell[m](N_\ell x^* + M_\ell y^* + q_\ell)[m] \end{pmatrix}. \tag{A.4}$$

Note that (4.13) can be rewritten as

$$\nabla f(x^*, y^*) + \nabla g(x^*, y^*)\lambda + \nabla h(x^*, y^*)\mu - \sum_{\ell=1}^L \begin{pmatrix} O \\ I \end{pmatrix} \nu_\ell - \sum_{\ell=1}^L \begin{pmatrix} N_\ell^T \\ O \end{pmatrix} \gamma_\ell = 0, \tag{A.5}$$

$$\nu_\ell - \beta_\ell - M_\ell^T \gamma_\ell = 0, \qquad \ell = 1, \cdots, L, \tag{A.6}$$

$$p_\ell d - \alpha_\ell - \gamma_\ell = 0, \qquad \ell = 1, \cdots, L. \tag{A.7}$$

Substituting (A.6) into (A.5), we obtain

$$\nabla f(x^*, y^*) + \nabla g(x^*, y^*)\lambda + \nabla h(x^*, y^*)\mu - \begin{pmatrix} O \\ I \end{pmatrix} \sum_{\ell=1}^L \beta_\ell - \sum_{\ell=1}^L \begin{pmatrix} N_\ell^T \\ M_\ell^T \end{pmatrix} \gamma_\ell = 0,$$

or equivalently,

$$\nabla f(x^*, y^*) + \nabla g(x^*, y^*)\lambda + \nabla h(x^*, y^*)\mu - \begin{pmatrix} O \\ I \end{pmatrix} \sum_{\ell=1}^L \beta_\ell - \sum_{\ell=1}^L \sum_{i=1}^m \gamma_\ell[i] \begin{pmatrix} N_\ell[i] \\ M_\ell[i] \end{pmatrix} = 0. \tag{A.8}$$

We next prove that

$$\gamma_\ell[i] = p_\ell d[i] \sigma^*_\ell[i] - \xi^*_\ell[i] y^*[i], \qquad \forall \ell,\ \forall i. \tag{A.9}$$

Since (x^*, y^*) is feasible to problem (4.3), we only need to consider three cases (a1)–(a3):

(a1) Suppose $y^*[i] > 0$ and $(N_\ell x^* + M_\ell y^* + q_\ell)[i] = 0$. It then follows from (A.2), (A.3), and (A.7) that

$$p_\ell d[i] \sigma^*_\ell[i] - \xi^*_\ell[i] y^*[i] = p_\ell d[i] - \alpha_\ell[i] = \gamma_\ell[i].$$

Namely, (A.9) holds in this case.

(a2) Suppose that $y^*[i] > 0$ and $(N_\ell x^* + M_\ell y^* + q_\ell)[i] < 0$. We then have from (4.19) that $z_\ell^*[i] \geq -(N_\ell x^* + M_\ell y^* + q_\ell)[i] > 0$. Since $\mathbf{z}^*[(\ell-1)m+i] = z_\ell^*[i] > 0$, by (4.15), we see that $\alpha_\ell[i] = \alpha[(\ell-1)m+i] = 0$. This, together with (A.2)–(A.3) and (A.7) yields

$$p_\ell d[i]\sigma_\ell^*[i] - \xi_\ell^*[i]y^*[i] = p_\ell d[i] = p_\ell d[i] - \alpha_\ell[i] = \gamma_\ell[i].$$

(a3) Suppose that $y^*[i] = 0$. It then follows from (A.2), (A.3), and (A.7) that

$$p_\ell d[i]\sigma_\ell^*[i] - \xi_\ell^*[i]y^*[i] = p_\ell d[i] - \alpha_\ell[i] = \gamma_\ell[i].$$

Therefore, (A.9) holds in any case. Substituting (A.1), (A.4), and (A.9) into (A.8), we obtain

$$\begin{aligned} 0 \;=\;& \nabla f(x^*,y^*) + \nabla g(x^*,y^*)\lambda^* + \nabla h(x^*,y^*)\mu^* - \begin{pmatrix} O \\ I \end{pmatrix} \nu^* \\ &+ \sum_{\ell=1}^{L}\sum_{i=1}^{m} \xi_\ell^*[i]\Big((N_\ell x^* + M_\ell y^* + q_\ell)[i]\begin{pmatrix} 0 \\ e_i \end{pmatrix} + y^*[i]\begin{pmatrix} N_\ell[i] \\ M_\ell[i] \end{pmatrix}\Big) \\ &+ \sum_{\ell=1}^{L}\sum_{i=1}^{m} p_\ell d[i]\sigma_\ell^*[i]\begin{pmatrix} -N_\ell[i] \\ -M_\ell[i] \end{pmatrix}. \end{aligned}$$

So, in order to prove (4.6), we only need to show

$$\sum_{\ell=1}^{L}\sum_{i=1}^{m} p_\ell d[i]\sigma_\ell^*[i]\begin{pmatrix} -N_\ell[i] \\ -M_\ell[i] \end{pmatrix} \in \partial\theta(x^*,y^*),$$

which, by (4.10) and (4.11), is equivalent to

$$\forall\ell,\ \forall i,\quad \begin{cases} \sigma_\ell^*[i] \in [0,1], & (N_\ell x^* + M_\ell y^* + q_\ell)[i] = 0, \\ \sigma_\ell^*[i] = 1\ , & (N_\ell x^* + M_\ell y^* + q_\ell)[i] < 0, \\ \sigma_\ell^*[i] = 0\ , & (N_\ell x^* + M_\ell y^* + q_\ell)[i] > 0. \end{cases} \tag{A.10}$$

(a4) Suppose $(N_\ell x^* + M_\ell y^* + q_\ell)[i] = 0$. Then, we have from (4.19) that

$$(N_\ell x^* + M_\ell y^* + q_\ell + z_\ell^*)[i] = 0.$$

If $y^*[i] > 0$, then $\sigma_\ell^*[i] = 1 \in [0,1]$ by (A.3). If $y^*[i] = 0$, we have from (4.20) that

$$\mathbf{y}^*[(\ell-1)m+i] = y^*[i] = 0$$

and

$$(Nx^* + M\mathbf{y}^* + q + \mathbf{z}^*)[(\ell-1)m+i] = (N_\ell x^* + M_\ell y^* + q_\ell + z_\ell^*)[i] = 0.$$

Thus, by (4.18), $\gamma_\ell[i] = \gamma[(\ell-1)m+i] \geq 0$ and hence, by (4.15) and (A.7),

$$0 \leq \alpha_\ell[i] = p_\ell d[i] - \gamma_\ell[i] \leq p_\ell d[i].$$

As a result, by (A.3), we obtain

$$\sigma_\ell^*[i] = 1 - \frac{\alpha_\ell[i]}{p_\ell d[i]} \in [0,1].$$

(a5) Suppose $(N_\ell x^* + M_\ell y^* + q_\ell)[i] < 0$. In a similar way to (a2), we can show that $\alpha_\ell[i] = 0$ and hence, by the definition (A.3) of $\sigma_\ell^*[i]$, we see that $\sigma_\ell^*[i] = 1$.

(a6) Suppose $(N_\ell x^* + M_\ell y^* + q_\ell)[i] > 0$. It then follows from the feasibility of (x^*, y^*) to problem (4.3) that $y^*[i] = 0$ and furthermore, by (4.19),

$$\begin{aligned}(Nx^* + M\mathbf{y}^* + q + \mathbf{z}^*)[(\ell-1)m+i] &= (N_\ell x^* + M_\ell y^* + q_\ell + z_\ell^*)[i] \\ &= (N_\ell x^* + M_\ell y^* + q_\ell)[i] \\ &> 0.\end{aligned}$$

From (4.17), we see that $\gamma_\ell[i] = \gamma[(\ell-1)m+i] = 0$. Thus, by (A.7), we obtain $\alpha_\ell[i] = p_\ell d[i]$ and hence, by the definition (A.3) of $\sigma_\ell^*[i]$,

$$\sigma_\ell^*[i] = 1 - \frac{\alpha_\ell[i]}{p_\ell d[i]} = 0.$$

This completes the proof of (A.10) and hence (4.6) holds. Note that condition (4.7) follows from (4.14) and (A.1) immediately. In addition, the definition of ξ_ℓ^* implies

$$\xi_\ell^*[i] \geq 0, \quad \xi_\ell^*[i] y^*[i] (N_\ell x^* + M_\ell y^* + q_\ell)[i] = 0, \qquad \forall \ell,\ \forall i. \tag{A.11}$$

Since $(x^*, y^*, z_1^*, \cdots, z_L^*)$ is feasible to problem (3.9), it then follows that, for any ℓ and i,

$$-y^*[i](N_\ell x^* + M_\ell y^* + q_\ell)[i] = y^*[i] z_\ell^*[i] \geq 0,$$

which, together with (A.11) indicates that (4.9) holds. We next show (4.8).

At first, for every i, we have from (4.16), (A.11), and the definition (A.4) of ν^* that

$$\nu^*[i] y^*[i] = \sum_{\ell=1}^{L} \beta[(\ell-1)m+i] \mathbf{y}^*[(\ell-1)m+i] + \sum_{\ell=1}^{L} \xi_\ell^*[i] y^*[i] (N_\ell x^* + M_\ell y^* + q_\ell)[i] = 0. \tag{A.12}$$

The rest is to prove $\nu^* \geq 0$. If $y^*[i] > 0$, we see from (A.12) that $\nu^*[i] = 0$. If $y^*[i] = 0$, we let

$$\mathcal{I}(i) := \{\ell \mid (N_\ell x^* + M_\ell y^* + q_\ell)[i] > 0\}.$$

It follows from (A.2) and (A.4) that

$$\begin{aligned}\nu^*[i] &= \sum_{\ell=1}^{L} \beta_\ell[i] + \sum_{\ell \in \mathcal{I}(i)} \xi_\ell^*[i] (N_\ell x^* + M_\ell y^* + q_\ell)[i] \\ &= \sum_{\ell=1}^{L} \beta_\ell[i] + n_1 \Big| \sum_{\ell=1}^{L} \beta_\ell[i] \Big|,\end{aligned} \tag{A.13}$$

where n_1 denotes the cardinality of the index set $\mathcal{I}(i)$. If $\mathcal{I}(i) \neq \emptyset$, then $\nu^*[i] \geq 0$ by (A.13). If $\mathcal{I}(i) = \emptyset$, it follows that

$$(N_\ell x^* + M_\ell y^* + q_\ell)[i] \leq 0, \qquad \forall \ell$$

and then, by (4.19) and (4.20), there hold

$$\mathbf{y}^*[(\ell-1)m+i] = y^*[i] = 0$$

and

$$(Nx^* + M\mathbf{y}^* + q + \mathbf{z}^*)[(\ell-1)m+i] = (N_\ell x^* + M_\ell y^* + q_\ell + z_\ell^*)[i] = 0$$

for every ℓ. From (4.18) and (A.13), we obtain $\nu^*[i] = \sum_{\ell=1}^L \beta_\ell[i] \geq 0$. Thus, we have $\nu^* \geq 0$ in any case. This completes the proof of (4.8) and hence (x^*, y^*) is a stationary point of problem (4.3).

(b) Suppose that (x^*, y^*) is a stationary point of problem (4.3) and the MPEC-LICQ holds at the point $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$. By the stationarity of (x^*, y^*), there exist multiplier vectors λ^*, μ^*, ν^*, and ξ_ℓ^*, $\ell = 1, \cdots, L$, satisfying conditions (4.6)–(4.9). We will prove that $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ is a B-stationary point of problem (4.12), i.e., there exist multiplier vectors $\lambda, \mu, \nu, \alpha, \beta$, and γ such that (4.13)–(4.18) hold.

By (4.10) and (4.11), condition (4.6) means that there exist multiplier vectors σ_ℓ^*, $\ell = 1, \cdots, L$, satisfying (A.10) and

$$\begin{aligned} 0 \;=\; & \nabla f(x^*, y^*) + \nabla g(x^*, y^*)\lambda^* + \nabla h(x^*, y^*)\mu^* - \begin{pmatrix} O \\ I \end{pmatrix} \nu^* \\ & + \sum_{\ell=1}^L \sum_{i=1}^m \xi_\ell^*[i] \Big((N_\ell x^* + M_\ell y^* + q_\ell)[i] \begin{pmatrix} 0 \\ e_i \end{pmatrix} + y^*[i] \begin{pmatrix} N_\ell[i] \\ M_\ell[i] \end{pmatrix} \Big) \\ & + \sum_{\ell=1}^L \sum_{i=1}^m p_\ell d[i] \sigma_\ell^*[i] \begin{pmatrix} -N_\ell[i] \\ -M_\ell[i] \end{pmatrix}. \end{aligned} \tag{A.14}$$

Let

$$\lambda := \lambda^*, \tag{A.15}$$

$$\mu := \mu^*, \tag{A.16}$$

$$\alpha_\ell[i] := p_\ell d[i](1 - \sigma_\ell^*[i]) + \xi_\ell^*[i] y^*[i], \tag{A.17}$$

$$\beta_\ell[i] := \frac{1}{L}\nu^*[i] - \xi_\ell^*[i](N_\ell x^* + M_\ell y^* + q_\ell)[i], \tag{A.18}$$

$$\gamma_\ell[i] := p_\ell d[i] \sigma_\ell^*[i] - \xi_\ell^*[i] y^*[i], \tag{A.19}$$

$$\nu_\ell := \beta_\ell + M_\ell^T \gamma_\ell, \tag{A.20}$$

and

$$\alpha := \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_L \end{pmatrix}, \quad \beta := \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_L \end{pmatrix}, \quad \gamma := \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_L \end{pmatrix}, \quad \nu := \begin{pmatrix} \nu_1 \\ \vdots \\ \nu_L \end{pmatrix}.$$

It is easy to verify (A.6)–(A.8) from (A.14)–(A.19) and so we obtain (4.13). In addition, condition (4.14) follows from (4.7) and (A.15) immediately. We next show (4.15)–(4.18). Note that, by (4.9) and (A.10), we have $\alpha \geq 0$. Moreover, by (4.19) and (4.20), we have $\mathbf{y}^* \geq 0$, $\mathbf{z}^* \geq 0$, and $N_\ell x^* + M_\ell y^* + q_\ell + z_\ell^* \geq 0$ for any ℓ, which in turn implies $Nx^* + M\mathbf{y}^* + q + \mathbf{z}^* \geq 0$. Take an arbitrary index j with $1 \leq j \leq mL$. Then, there exist ℓ and i such that

$$1 \leq \ell \leq L, \quad 1 \leq i \leq m, \quad j = (\ell - 1)m + i.$$

(b1) Suppose $\mathbf{z}^*[j] > 0$. This means $z_\ell^*[i] > 0$. It then follows from (4.19) that $(N_\ell x^* + M_\ell y^* + q_\ell)[i] < 0$ and hence we have

$$\sigma_\ell^*[i] = 1, \qquad \xi_\ell^*[i] y^*[i] = 0$$

from (A.10) and (4.9), respectively. This indicates $\alpha[j] = \alpha_\ell[i] = 0$ and so (4.15) holds.

(b2) It is easy to see from (4.8) and (4.9) that

$$\beta[j]\mathbf{y}[j] = \beta_\ell[i] y^*[i] = \frac{1}{L}\nu^*[i] y^*[i] - \xi_\ell^*[i] y^*[i] (N_\ell x^* + M_\ell y^* + q_\ell)[i] = 0,$$

which indicates that (4.16) holds.

(b3) Suppose $(Nx^* + M\mathbf{y}^* + q + \mathbf{z}^*)[j] > 0$. This means that $(N_\ell x^* + M_\ell y^* + q_\ell + z_\ell^*)[i] > 0$. It then follows from (4.19) that $(N_\ell x^* + M_\ell y^* + q_\ell)[i] > 0$ and hence $y^*[i] = 0$. This indicates $\gamma[j] = \gamma_\ell[i] = 0$ and therefore (4.17) holds.

(b4) Let $\mathcal{I}^*$ be defined as in (4.18) and suppose $j \in \mathcal{I}^*$. It is obvious that

$$\gamma[j] = p_\ell d[i] \sigma_\ell^*[i] \geq 0.$$

On the other hand, since $(Nx^* + M\mathbf{y}^* + q + \mathbf{z}^*)[j] = 0$ implies that $(N_\ell x^* + M_\ell y^* + q_\ell + z_\ell^*)[i] = 0$, we see from (4.19) that $(N_\ell x^* + M_\ell y^* + q_\ell)[i] \leq 0$. Therefore,

$$\beta[j] = \beta_\ell[i] = \frac{1}{L}\nu^*[i] - \xi_\ell^*[i](N_\ell x^* + M_\ell y^* + q_\ell)[i] \geq \frac{1}{L}\nu^*[i] \geq 0.$$

This indicates that (4.18) holds.

Thus, the multiplier vectors defined by (A.15)–(A.19) satisfy conditions (4.13)–(4.18) and hence $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ is a B-stationary point of problem (4.12). ∎