Regularization Method for Stochastic Mathematical Programs with Complementarity Constraints *

Gui-Hua Lin† and Masao Fukushima‡

September 6, 2003

Abstract

In this paper, we consider a class of stochastic mathematical programs with equilibrium constraints (SMPECs) that has been discussed by Lin and Fukushima (2003). Based on a reformulation given therein, we propose a regularization method for solving the problems. We show that, under a weak condition, an accumulation point of the generated sequence is a feasible point of the original problem. We also show that such an accumulation point is S-stationary to the problem under additional assumptions.

Key Words. stochastic mathematical program with equilibrium constraints, S-stationarity, Mangasarian-Fromovitz constraint qualification.

AMS Subject classification. 90C30, 90C33

1 Introduction

Mathematical program with equilibrium constraints (MPEC) plays an important role in many fields such as engineering design, economic equilibrium, multilevel game, and mathematical programming itself, and it therefore has been receiving much attention in the optimization world, see the monograph [7]. Recently, a class of more practical problems, called stochastic mathematical programs with equilibrium constraints (SMPECs), has

*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. Tel: (+81)75-753-4922. Fax: (+81)75-753-4756.

‡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. Tel: (+81)75-753-5519. Fax: (+81)75-753-4756.

been studied [5, 6, 8]. 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 [5, 6] for more details.

In this paper, assuming that the underlying sample space Ω is discrete and finite, i.e., $\Omega = \{\omega_1, \omega_2, \cdots, \omega_L\}$ for some integer $L > 0$, we consider the here-and-now model [5, 6] of SMPECs:

$$\begin{array}{ll} \text{minimize} & f(x, y) + \sum_{\ell=1}^{L} p_\ell d^T z_\ell \\ \text{subject to} & g(x, y) \leq 0, \;\; h(x, y) = 0, \\ & y \geq 0, \;\; N_\ell x + M_\ell y + q_\ell + z_\ell \geq 0, \\ & y^T (N_\ell x + M_\ell y + q_\ell + z_\ell) = 0, \\ & z_\ell \geq 0, \quad \ell = 1, 2, \cdots, L. \end{array} \tag{1.1}$$

Here, the functions $f : \Re^{n+m} \to \Re$, $g : \Re^{n+m} \to \Re^{s_1}$, $h : \Re^{n+m} \to \Re^{s_2}$ are all continuously differentiable, $d \in \Re^m$ is a constant vector with positive elements and, for each ℓ, $N_\ell := N(\omega_\ell) \in \Re^{m\times n}$, $M_\ell := M(\omega_\ell) \in \Re^{m\times m}$, and $q_\ell := q(\omega_\ell) \in \Re^m$ are given matrices and vectors associated with the random event ω_ℓ, $z_\ell \in \Re^m$ is a recourse variable, p_ℓ denotes the probability of ω_ℓ and is assumed to be positive throughout.

In the here-and-now model (1.1), x denotes the upper-level decision variable, y represents the lower-level decision variable, and both the decisions x and y need to be made at once, before a random event is observed. This is in contrast with the wait-and-see model as studied in [8], where the lower-level decision is made after a random event is observed. It has been shown that many problems including the stochastic linear complementarity problem can be formulated as this kind of here-and-now models [6]. However, it is well-known that problem (1.1) does not satisfy a standard constraint qualification such as the linear independence constraint qualification (LICQ) or the Mangasarian-Fromovitz constraint qualification (MFCQ) at any feasible point [2], so that the conventional theory of nonlinear programming cannot be applied to this problem directly.

Based on some reformulations, two penalty methods have been proposed for solving problem (1.1) in [6]. In addition, a smoothing implicit programming method incorporating a penalty technique has been suggested for solving a similar problem in [5]. However, like the penalty methods in standard nonlinear programming, the methods suggested in [5, 6] cannot ensure the feasibility of a limit point of a generated sequence in general. In this paper, we will present a regularization method for problem (1.1) and show that, under a quite weak condition, an accumulation point of the generated sequence is a feasible point of the original problem. We will also establish global convergence to an S-stationary point of the problem under additional assumptions.

The following notations will be used in this paper: All vectors are regarded as column vectors and $w[i]$ stands for the ith element of vector $w \in \Re^s$, whereas for a matrix M, we denote by $M[i]$ the column vector whose elements consist of 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 $w \in \Re^s$, $\nabla F(w)$ denotes 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. In addition, I and O denote the identity matrix and the zero matrix of suitable dimension, respectively.

2 Preliminaries and Regularization Method

In this section, we propose a regularization method for problem (1.1). We first recall some basic concepts. Since problem (1.1) is equivalent to the following ordinary MPEC (2.1), we will employ the same stationary concepts as in the literature on MPECs:

$$\begin{array}{ll} \text{minimize} & f(x,y) + \mathbf{d}^T\mathbf{z} \\ \text{subject to} & g(x,y) \le 0,\ h(x,y) = 0, \\ & \mathbf{y} - Dy = 0,\ \mathbf{z} \ge 0, \\ & \mathbf{y} \ge 0,\ Nx + M\mathbf{y} + \mathbf{q} + \mathbf{z} \ge 0, \\ & \mathbf{y}^T(Nx + M\mathbf{y} + \mathbf{q} + \mathbf{z}) = 0, \end{array} \tag{2.1}$$

where $\mathbf{y} := (y^T, \cdots, y^T)^T \in \Re^{mL}, \mathbf{z} := (z_1^T, \cdots, z_L^T)^T \in \Re^{mL}$, and

$$\mathbf{d} := \begin{pmatrix} p_1 d \\ \vdots \\ p_L d \end{pmatrix},\ D := \begin{pmatrix} I \\ \vdots \\ I \end{pmatrix},\ N := \begin{pmatrix} N_1 \\ \vdots \\ N_L \end{pmatrix},\ M := \begin{pmatrix} M_1 & & O \\ & \ddots & \\ O & & M_L \end{pmatrix},\ \mathbf{q} := \begin{pmatrix} q_1 \\ \vdots \\ q_L \end{pmatrix}. \tag{2.2}$$

Suppose that $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ is a feasible point of problem (2.1).

Definition 2.1 We say $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ to be a *Bouligand or B-stationary* point of the MPEC (2.1) if

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

where $\mathcal{T}(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ stands for the tangent cone of the feasible region of problem (2.1) at $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$.

Definition 2.2 We say $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ to be a *strongly or S-stationary* point of (2.1) if there exist multiplier vectors λ, μ, $\boldsymbol{\nu}$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ such that

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

$$+\begin{pmatrix} O \\ -D^T \\ I \\ O \end{pmatrix}\boldsymbol{\nu} - \begin{pmatrix} O \\ O \\ O \\ I \end{pmatrix}\boldsymbol{\alpha} - \begin{pmatrix} O \\ O \\ I \\ O \end{pmatrix}\boldsymbol{\beta} - \begin{pmatrix} N^T \\ O \\ M^T \\ I \end{pmatrix}\boldsymbol{\gamma} = 0, \tag{2.3}$$

$$0 \le \lambda \perp (-g(x^*, y^*)) \ge 0, \tag{2.4}$$

$$0 \le \boldsymbol{\alpha} \perp \mathbf{z}^* \ge 0, \tag{2.5}$$

$$\mathbf{y}^* \ge 0, \tag{2.6}$$

$$\mathbf{y}^*[i] > 0 \Rightarrow \boldsymbol{\beta}[i] = 0, \tag{2.7}$$

$$(Nx^* + M\mathbf{y}^* + \mathbf{q} + \mathbf{z}^*) \ge 0, \tag{2.8}$$

$$(Nx^* + M\mathbf{y}^* + \mathbf{q} + \mathbf{z}^*)[i] > 0 \Rightarrow \boldsymbol{\gamma}[i] = 0, \tag{2.9}$$

$$\boldsymbol{\beta}[i] \ge 0,\ \boldsymbol{\gamma}[i] \ge 0, \quad \forall i \in \mathcal{I}^* := \{i \mid\ \mathbf{y}^*[i] = (Nx^* + M\mathbf{y}^* + \mathbf{q} + \mathbf{z}^*)[i] = 0\}. \tag{2.10}$$

It is well-known [9] that any S-stationary point of (2.1) must be a B-stationary point of problem (2.1). In order to look for some effective methods for solving problem (1.1), some equivalent reformulations of problem (1.1) have been introduced recently [6]. In particular, for any $x \in \Re^n, y \in \Re^m$, and each ℓ, it has been shown [6] that the set

$$Z_\ell(x,y) := \left\{ z_\ell \;\middle|\; \begin{array}{c} y^T(N_\ell x + M_\ell y + q_\ell + z_\ell) = 0 \\ N_\ell x + M_\ell y + q_\ell + z_\ell \ge 0,\ z_\ell \ge 0 \end{array} \right\}$$

is nonempty if and only if

$$Q_\ell(x,y) := \sup\Big\{ -(u+ty)^T(N_\ell x + M_\ell y + q_\ell) \;\Big|\; u + ty \le d, \quad u \ge 0, \quad t \le 0 \Big\}$$

is finite. Based on this observation, we obtain the model [6]

$$\begin{array}{ll} \text{minimize} & f(x,y) + \displaystyle\sum_{\ell=1}^{L} p_\ell Q_\ell(x,y) \\ \text{subject to} & g(x,y) \le 0,\ h(x,y) = 0, \quad y \ge 0. \end{array} \tag{2.11}$$

Furthermore, we have the following result.

Theorem 2.1 [6] *If (x^*, y^*) solves problem (2.11), then there exist $z_\ell^*, \ell = 1, 2, \cdots, L$, such that $(x^*, y^*, z_1^*, \cdots, z_L^*)$ solves problem (1.1). Conversely, if $(x^*, y^*, z_1^*, \cdots, z_L^*)$ solves problem (1.1), then the point (x^*, y^*) solves problem (2.11).*

In what follows, we denote by $\mathcal{F}_1$ and $\mathcal{F}_2$ the feasible regions of problems (1.1) and (2.11), respectively. The next result will be used later on.

Theorem 2.2 [6] *Let $g(x,y) \le 0, h(x,y) = 0$, and $y \ge 0$. Then the following statements are equivalent:*

(i) $Q_\ell(x,y) < +\infty$ *for every* $\ell = 1, 2, \cdots, L$.

(ii) *For any* i *and any* ℓ*, there holds*

$$y[i](N_\ell x + M_\ell y + q_\ell)[i] \le 0. \tag{2.12}$$

(iii) *The point* $(x, y, z_1, \cdots, z_L)$ *with* $z_\ell := \max(-(N_\ell x + M_\ell y + q_\ell),\ 0), \ell = 1, \cdots, L$*, is a feasible point of problem (1.1).*

On the other hand, we note that, for every ℓ, the function Q_ℓ may not be finite-valued in general. We next introduce a smooth approximation of this function: Let ϵ be a positive parameter. For each ℓ, we define the function $Q_\ell^\epsilon : \Re^n \times \Re^m \to [0, +\infty)$ as follows:

$$\begin{aligned} Q_\ell^\epsilon(x,y) \;:=\; \max\Big\{ -(u+ty)^T(N_\ell x + M_\ell y + q_\ell) - \frac{\epsilon}{2}(t^2 + \|u\|^2) \;\Big|\; \\ u + ty \le d, \quad u \ge 0, \quad t \le 0 \Big\}. \end{aligned}$$

By the convex programming theory, we see that any Karush-Kuhn-Tucker point of the problem

$$\begin{array}{ll} \text{maximize} & -(u+ty)^T(N_\ell x + M_\ell y + q_\ell) - \frac{\epsilon}{2}(t^2 + \|u\|^2) \\ \text{subject to} & u + ty \le d, \quad u \ge 0, \quad t \le 0 \end{array} \tag{2.13}$$

must be an optimal solution and, since $\epsilon > 0$, problem (2.13) indeed has a unique optimal solution (u_ℓ, t_ℓ). This implies that the function Q_ℓ^ϵ is well-defined for each ℓ. Moreover, Q_ℓ^ϵ is differentiable everywhere and

$$\nabla Q_\ell^\epsilon(x,y) = \begin{pmatrix} -N_\ell^T(u_\ell + t_\ell y) \\ -M_\ell^T(u_\ell + t_\ell y) - t_\ell(N_\ell x + M_\ell y + q_\ell) \end{pmatrix}. \tag{2.14}$$

As a result, we obtain the problem

$$\begin{array}{ll} \text{minimize} & f(x,y) + \sum_{\ell=1}^{L} p_\ell Q_\ell^\epsilon(x,y) \\ \text{subject to} & g(x,y) \le 0,\ h(x,y) = 0, \quad y \ge 0. \end{array} \tag{2.15}$$

This is a smooth problem, unlike problem (2.11) whose objective function may not be differentiable everywhere. We then have the following algorithm.

Algorithm RA:

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

Step 2: Solve problem (2.15) with $\epsilon = \epsilon_k$ 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 an $\epsilon_{k+1} \in (0, \epsilon_k)$ and return to Step 2 with $k := k + 1$.

In what follows, we suppose that the sequence $\{\epsilon_k\}$ is convergent to 0 and, for simplicity, we denote $Q_\ell^{\epsilon_k}$ by Q_ℓ^k for each k and ℓ. Recall that $\mathcal{F}_2$ denotes the feasible region of problem (2.11), which is the same as that of problem (2.15).

3 Convergence Analysis

We will investigate the limiting behavior of the sequence generated by Algorithm RA in this section. Our first result is concerned with the feasibility of the limit point of the generated sequence, which can be stated as follows.

Theorem 3.1 *Let $\{(x^k, y^k)\}$ be a sequence generated by Algorithm RA and suppose that $\{Q_\ell^k(x^k, y^k)\}$ is bounded for each ℓ. Then, for any accumulation point (x^*, y^*) of the sequence $\{(x^k, y^k)\}$, the vector $(x^*, y^*, z_1^*, \cdots, z_L^*)$ is feasible to problem (1.1), where*

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

Proof: Assume without loss of generality that $\lim_{k\to\infty}(x^k, y^k) = (x^*, y^*)$. It is obvious from the continuity of the functions g and h that

$$g(x^*, y^*) \le 0, \quad h(x^*, y^*) = 0, \quad y^* \ge 0.$$

Suppose that the assertion of the theorem does not hold. Then, by Theorem 2.2, there exist some ℓ and i such that

$$y^*[i] > 0, \quad (N_\ell x^* + M_\ell y^* + q_\ell)[i] > 0.$$

Therefore, we can find a constant $\eta > 0$ and an integer $k_0 > 0$ such that

$$y^k[i] > \eta, \quad (N_\ell x^k + M_\ell y^k + q_\ell)[i] > \eta, \qquad \forall k \ge k_0. \tag{3.1}$$

For any $t \le 0$, we define $u(t) := ty^k[i]e_i - ty^k$. It is easy to see that

$$u(t) \ge 0, \quad u(t) + ty^k \le d, \qquad \forall t \le 0.$$

It then follows from the definition of Q_ℓ^k that

$$\begin{aligned} Q_\ell^k(x^k, y^k) &\ge \sup\Big\{ -(u(t) + ty^k)^T(N_\ell x^k + M_\ell y^k + q_\ell) - \frac{\epsilon_k}{2}(t^2 + \|u(t)\|^2) \ \Big|\ t \le 0\Big\} \\ &= \sup\Big\{ -ty^k[i](N_\ell x^k + M_\ell y^k + q_\ell)[i] - \frac{\epsilon_k}{2}t^2(1 + \|y^k[i]e_i - y^k\|^2) \ \Big|\ t \le 0\Big\}. \end{aligned}$$

By straightforward calculus, we can show that, for any $k \geq k_0$,

$$\begin{aligned} Q_\ell^k(x^k, y^k) &\geq \frac{(y^k[i](N_\ell x^k + M_\ell y^k + q_\ell)[i])^2}{2\epsilon_k(1 + \|y^k[i]e_i - y^k\|^2)} \\ &\geq \frac{\eta^4}{2\epsilon_k(1 + \|y^k[i]e_i - y^k\|^2)}, \end{aligned} \tag{3.2}$$

where the second inequality follows from (3.1). Taking into account the fact that

$$\lim_{k\to\infty} \|y^k[i]e_i - y^k\| = \|y^*[i]e_i - y^*\|, \quad \lim_{k\to\infty} \epsilon_k = 0,$$

we see from (3.2) that the sequence $\{Q_\ell^k(x^k, y^k)\}$ is unbounded. This is a contradiction and hence there must be some vectors $z_\ell^*, \ell = 1, 2, \cdots, L$, such that $(x^*, y^*, z_1^*, \cdots, z_L^*)$ is feasible to problem (1.1). This completes the proof. ∎

The main convergence result can be stated as follows.

Theorem 3.2 *Suppose that Algorithm RA generates a sequence $\{(x^k, y^k)\}$ of stationary points of problems (2.15) and, for each k, (u_ℓ^k, t_ℓ^k) is the corresponding unique optimal solution of problem (2.13) with $\epsilon := \epsilon_k$. Assume that, for each ℓ, both $\{Q_\ell^k(x^k, y^k)\}$ and $\{t_\ell^k\}$ are bounded. Moreover, suppose that (x^*, y^*) is an accumulation point of the sequence $\{(x^k, y^k)\}$ such that the system*

$$g(x, y) \leq 0, \quad h(x, y) = 0, \quad y \geq 0 \tag{3.3}$$

satisfies the MFCQ at (x^*, y^*), and let

$$z_\ell^* := \max(-(N_\ell x^* + M_\ell y^* + q_\ell),\ 0), \qquad \ell = 1, \cdots, L \tag{3.4}$$

and

$$\mathbf{y}^* := ((y^*)^T, \cdots, (y^*)^T)^T, \qquad \mathbf{z}^* := ((z_1^*)^T, \cdots, (z_L^*)^T)^T. \tag{3.5}$$

Then $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ is an S-stationary point of problem (2.1).

Proof: Assume without loss of generality that $\lim_{k\to\infty}(x^k, y^k) = (x^*, y^*)$. From Theorem 3.1, we see $(x^*, y^*, z_1^*, \cdots, z_L^*) \in \mathcal{F}_1$ and hence $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ is a feasible point of (2.1). We next show that $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ is S-stationary to problem (2.1), that is, there exist multiplier vectors $\lambda, \mu, \boldsymbol{\nu}, \boldsymbol{\alpha}, \boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ such that (2.3)–(2.10) hold.

First of all, by the stationarity of (x^k, y^k) to (2.15), there exist Lagrange multiplier vectors $a^k \in \Re^{s_1}, b^k \in \Re^{s_2}$, and $c^k \in \Re^m$ such that

$$\nabla f(x^k, y^k) + \sum_{\ell=1}^{L} p_\ell \nabla Q_\ell^k(x^k, y^k) + \nabla g(x^k, y^k)a^k + \nabla h(x^k, y^k)b^k - \begin{pmatrix} O \\ I \end{pmatrix} c^k = 0, \tag{3.6}$$

$$0 \leq a^k \perp (-g(x^k, y^k)) \geq 0, \tag{3.7}$$

$$0 \leq c^k \perp \quad y^k \geq 0. \tag{3.8}$$

From (2.14), we can rewrite (3.6) as

$$\nabla f(x^k, y^k) + \sum_{\ell=1}^{L} p_\ell \begin{pmatrix} -N_\ell^T(u_\ell^k + t_\ell^k y^k) \\ -M_\ell^T(u_\ell^k + t_\ell^k y^k) - t_\ell^k(N_\ell x^k + M_\ell y^k + q_\ell) \end{pmatrix} + \nabla g(x^k, y^k)a^k + \nabla h(x^k, y^k)b^k - \begin{pmatrix} O \\ I \end{pmatrix} c^k = 0,$$

which is equivalent to

$$\begin{aligned} 0 \;=\; & \nabla f(x^k, y^k) + \nabla g(x^k, y^k)a^k + \nabla h(x^k, y^k)b^k - \begin{pmatrix} O \\ I \end{pmatrix} c^k \\ & - \sum_{\ell=1}^{L} p_\ell t_\ell^k \begin{pmatrix} O \\ I \end{pmatrix} (N_\ell x^k + M_\ell y^k + q_\ell) - \sum_{\ell=1}^{L} p_\ell \begin{pmatrix} N_\ell^T \\ M_\ell^T \end{pmatrix} (u_\ell^k + t_\ell^k y^k). \end{aligned} \tag{3.9}$$

We next prove that the sequences $\{a^k\}$, $\{b^k\}$ and $\{c^k\}$ are bounded. To this end, let

$$\rho_k := \sum_{i=1}^{s_1} a^k[i] + \sum_{i=1}^{s_2} |b^k[i]| + \sum_{i=1}^{m} c^k[i]. \tag{3.10}$$

Suppose that at least one of the sequences $\{a^k\}$, $\{b^k\}$ and $\{c^k\}$ is unbounded. We then have $\lim_{k\to\infty} \rho_k = +\infty$ and, taking a subsequence if necessary, we may assume that the limits

$$\begin{aligned} \bar{a}[i] &:= \lim_{k\to\infty} \frac{a^k[i]}{\rho_k}, \qquad i = 1, \cdots, s_1, \\ \bar{b}[i] &:= \lim_{k\to\infty} \frac{b^k[i]}{\rho_k}, \qquad i = 1, \cdots, s_2, \\ \bar{c}[i] &:= \lim_{k\to\infty} \frac{c^k[i]}{\rho_k}, \qquad i = 1, \cdots, m \end{aligned}$$

exist. It is clear from (3.10) that

$$\sum_{i=1}^{s_1} \bar{a}[i] + \sum_{i=1}^{s_2} |\bar{b}[i]| + \sum_{i=1}^{m} \bar{c}[i] = 1.$$

For each ℓ, since $\{t_\ell^k\}$ is bounded and

$$0 \le u_\ell^k \le d - t_\ell^k y^k, \qquad \forall k,$$

$\{u_\ell^k\}$ is also bounded. Thus, dividing (3.9) by ρ_k and taking a limit, we get

$$\nabla g(x^*, y^*)\bar{a} + \nabla h(x^*, y^*)\bar{b} - \begin{pmatrix} O \\ I \end{pmatrix} \bar{c} = 0.$$

Furthermore, taking (3.7) and (3.8) into account, we obtain $\bar{a} \ge 0, \bar{c} \ge 0$, and

$$\begin{aligned} \bar{a}[i] = 0, &\quad i \notin \mathcal{I}_g(x^*, y^*), \\ \bar{c}[i] = 0, &\quad i \notin \mathcal{I}_\pi(x^*, y^*), \end{aligned}$$

where $\pi : \Re^{n+m} \to \Re^m$ is given by $\pi(x,y) := y$. It follows that

$$\begin{array}{ll} \sum_{i\in\mathcal{I}_g(x^*,y^*)} \bar{a}[i] + \sum_{i=1}^{s_2} |\bar{b}[i]| + \sum_{i\in\mathcal{I}_\pi(x^*,y^*)} \bar{c}[i] = 1, & \\ \bar{a}[i] \geq 0, \qquad i \in \mathcal{I}_g(x^*,y^*), & \\ \bar{c}[i] \geq 0, \qquad i \in \mathcal{I}_\pi(x^*,y^*), & \end{array}$$

and

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

This contradicts the assumption that the system (3.3) satisfies the MFCQ at (x^*,y^*) and hence all the sequences $\{a^k\}$, $\{b^k\}$, and $\{c^k\}$ are bounded. Recall that $\{t_\ell^k\}$ and $\{u_\ell^k\}$ are also bounded for each ℓ.

Now let us proceed to showing (2.3)–(2.10) step by step. First we show (2.3) and (2.4). Let

$$\lambda^k := a^k, \tag{3.11}$$

$$\mu^k := b^k, \tag{3.12}$$

$$\alpha_\ell^k[i] := p_\ell(d[i] - u_\ell^k[i] - t_\ell^k y^k[i]), \tag{3.13}$$

$$\beta_\ell^k[i] := \frac{1}{L}c^k[i] + p_\ell t_\ell^k(N_\ell x^k + M_\ell y^k + q_\ell)[i], \tag{3.14}$$

$$\gamma_\ell^k[i] := p_\ell(u_\ell^k[i] + t_\ell^k y^k[i]), \tag{3.15}$$

$$\nu_\ell^k := \beta_\ell^k + M_\ell^T\gamma_\ell^k, \tag{3.16}$$

and

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

Then (3.9) can be rewritten as

$$\begin{aligned} 0 \;=\; & \begin{pmatrix} \nabla_x f(x^k,y^k) \\ \nabla_y f(x^k,y^k) \\ 0 \\ \mathbf{d} \end{pmatrix} + \begin{pmatrix} \nabla_x g(x^k,y^k) \\ \nabla_y g(x^k,y^k) \\ 0 \\ 0 \end{pmatrix} \lambda^k + \begin{pmatrix} \nabla_x h(x^k,y^k) \\ \nabla_y h(x^k,y^k) \\ 0 \\ 0 \end{pmatrix} \mu^k \\ & + \begin{pmatrix} O \\ -D^T \\ I \\ O \end{pmatrix} \boldsymbol{\nu}^k - \begin{pmatrix} O \\ O \\ O \\ I \end{pmatrix} \boldsymbol{\alpha}^k - \begin{pmatrix} O \\ O \\ I \\ O \end{pmatrix} \boldsymbol{\beta}^k - \begin{pmatrix} N^T \\ O \\ M^T \\ I \end{pmatrix} \boldsymbol{\gamma}^k, \end{aligned} \tag{3.17}$$

where $\mathbf{d}, D, N$, and M are defined as in (2.2). Since all the multiplier vectors are bounded, without loss of generality, we may assume that

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

and

$$\boldsymbol{\alpha} := \lim_{k\to\infty} \boldsymbol{\alpha}^k, \qquad \boldsymbol{\beta} := \lim_{k\to\infty} \boldsymbol{\beta}^k, \qquad \boldsymbol{\gamma} := \lim_{k\to\infty} \boldsymbol{\gamma}^k.$$

Taking a limit in (3.17), we obtain (2.3) immediately. Moreover, we have (2.4) from (3.7) and (3.11) by letting $k \to \infty$.

We next prove (2.5)–(2.10). To this end, we let

$$z_\ell^k := \max(-(N_\ell x^k + M_\ell y^k + q_\ell),\ 0), \qquad \ell = 1, \cdots, L \tag{3.18}$$

and

$$\mathbf{y}^k := ((y^k)^T, \cdots, (y^k)^T)^T, \qquad \mathbf{z}^k := ((z_1^k)^T, \cdots, (z_L^k)^T)^T. \tag{3.19}$$

Recall that, for each ℓ and each k, (u_ℓ^k, t_ℓ^k) is a Karush-Kuhn-Tucker point of problem (2.13) with $\epsilon = \epsilon_k$. Thus, there exist Lagrange multiplier vectors $\zeta_\ell^k, \eta_\ell^k$, and ξ_ℓ^k such that

$$(N_\ell x^k + M_\ell y^k + q_\ell) + \epsilon_k u_\ell^k + \zeta_\ell^k - \eta_\ell^k = 0, \tag{3.20}$$
$$(y^k)^T(N_\ell x^k + M_\ell y^k + q_\ell) + \epsilon_k t_\ell^k + (y^k)^T\zeta_\ell^k + \xi_\ell^k = 0, \tag{3.21}$$
$$0 \le \zeta_\ell^k \perp (d - u_\ell^k - t_\ell^k y^k) \ge 0, \tag{3.22}$$
$$0 \le \eta_\ell^k \perp \quad u_\ell^k \ge 0, \tag{3.23}$$
$$0 \le \xi_\ell^k \perp (-t_\ell^k) \ge 0. \tag{3.24}$$

Moreover, for any index j with $1 \le j \le mL$, there exist ℓ and i such that

$$1 \le \ell \le L, \quad 1 \le i \le m, \quad j = (\ell - 1)m + i. \tag{3.25}$$

It is obvious that $\boldsymbol{\alpha}^k \ge 0$ and $\mathbf{z}^k \ge 0$ from the definitions (3.13), (3.18) and (3.19) for every k. Taking a limit, we obtain $\boldsymbol{\alpha} \ge 0$ and $\mathbf{z}^* \ge 0$. Suppose $\mathbf{z}^*[j] > 0$ and let ℓ and i satisfy (3.25). Then, since $z_\ell^*[i] > 0$, it follows from (3.4) that $(N_\ell x^* + M_\ell y^* + q_\ell)[i] < 0$. We then have from (3.20) and (3.23) that

$$\begin{aligned}
\zeta_\ell^k[i] &\ge \zeta_\ell^k[i] - \eta_\ell^k[i] \\
&= -(N_\ell x^k + M_\ell y^k + q_\ell)[i] - \epsilon_k u_\ell^k[i] \\
&\to -(N_\ell x^* + M_\ell y^* + q_\ell)[i] \\
&> 0.
\end{aligned}$$

This implies that $\zeta_\ell^k[i] > 0$ when k is large sufficiently and so, by (3.22),

$$u_\ell^k[i] + t_\ell^k y^k[i] = d[i].$$

Therefore, we have from the definition (3.13) that $\boldsymbol{\alpha}^k[j] = \alpha_\ell^k[i] = 0$ for all k sufficiently large. By taking a limit, we have $\boldsymbol{\alpha}[j] = 0$ and hence (2.5) holds.

It is easy to see that $\mathbf{y}^* \geq 0$. Suppose $\mathbf{y}^*[j] > 0$ and let ℓ and i satisfy (3.25). Then, since $\mathbf{y}^*[j] = y^*[i] > 0$, by Theorem 2.2, there must hold $(N_\ell x^* + M_\ell y^* + q_\ell)[i] = 0$. Moreover, we have $y^k[i] > 0$ when k is large sufficiently. Thus, it follows from (3.8) that $c^k[i] = 0$ for all k sufficiently large. Taking a limit in (3.14) and noting that $\{t_\ell^k\}$ is bounded, we obtain

$$\boldsymbol{\beta}[j] = \lim_{k\to\infty} \boldsymbol{\beta}^k[j] = \lim_{k\to\infty} \beta_\ell^k[i] = 0.$$

This shows (2.6) and (2.7).

It is easy to see that $Nx^* + M\mathbf{y}^* + \mathbf{q} + \mathbf{z}^* \geq 0$. Suppose $(Nx^* + M\mathbf{y}^* + \mathbf{q} + \mathbf{z}^*)[j] > 0$ and let ℓ and i satisfy (3.25). Then, since $(N_\ell x^* + M_\ell y^* + q_\ell + z_\ell^*)[i] > 0$, it follows from (3.4) that $(N_\ell x^* + M_\ell y^* + q_\ell)[i] > 0$ and hence, by Theorem 2.2, $y^*[i] = 0$. Moreover, we have from (3.20) and (3.22)–(3.23) that

$$\begin{array}{rcl} \eta_\ell^k[i] & = & (N_\ell x^k + M_\ell y^k + q_\ell)[i] + \epsilon_k u_\ell^k[i] + \zeta_\ell^k[i] \\ & \geq & (N_\ell x^k + M_\ell y^k + q_\ell)[i] \\ & \to & (N_\ell x^* + M_\ell y^* + q_\ell)[i] \\ & > & 0. \end{array}$$

In consequence, $\eta_\ell^k[i] > 0$ when k is large sufficiently and then, by (3.23), we have $u_\ell^k[i] = 0$. Taking a limit in (3.15) and noting that $\{t_\ell^k\}$ is bounded, we obtain

$$\boldsymbol{\gamma}[j] = \lim_{k\to\infty} \boldsymbol{\gamma}^k[j] = \lim_{k\to\infty} \gamma_\ell^k[i] = 0$$

and hence (2.8) and (2.9) hold.

Let $\mathcal{I}^*$ be defined as in (2.10) and suppose $j \in \mathcal{I}^*$. Note that

$$u_\ell^k[i] \geq 0, \quad t_\ell^k \leq 0, \quad c^k[i] \geq 0, \qquad \forall k$$

and $(Nx^* + M\mathbf{y}^* + \mathbf{q} + \mathbf{z}^*)[j] = 0$ implies $(N_\ell x^* + M_\ell y^* + q_\ell + z_\ell^*)[i] = 0$, where ℓ and i satisfy (3.25), and hence $(N_\ell x^* + M_\ell y^* + q_\ell)[i] \leq 0$ by (3.4). We then have from (3.14)–(3.15) that

$$\boldsymbol{\beta}[j] = \lim_{k\to\infty} \beta_\ell^k[i] \geq \lim_{k\to\infty} p_\ell t_\ell^k (N_\ell x^k + M_\ell y^k + q_\ell)[i] \geq 0$$

and

$$\boldsymbol{\gamma}[j] = \lim_{k\to\infty} \gamma_\ell^k[i] \geq \lim_{k\to\infty} p_\ell t_\ell^k y^k[i] = 0.$$

This indicates that (2.10) holds.

Therefore, the multiplier vectors $\lambda, \mu, \boldsymbol{\nu}, \boldsymbol{\alpha}, \boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ indeed satisfy conditions (2.3)–(2.10) and so $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ is an S-stationary point of problem (2.1). This completes the proof of the theorem. ∎

4 Conclusions

The SMPEC (1.1) has been discussed in [6] and two penalty methods have been proposed there. The main difficulty with the two methods consists in the feasibility of a limit point of the generated sequence, which has not been addressed completely. In this paper, based on a reformulation given in [6], we propose a regularization method for solving the SMPEC (1.1). It has been shown that, under a weak condition, an accumulation point of the generated sequence is a feasible point of the original problem. Global convergence to an S-stationary point of the problem has also 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] R.W. Cottle, J.S. Pang, and R.E. Stone, *The Linear Complementarity Problem*, Academic Press, New York, NY (1992).

[4] P. Kall and S.W. Wallace, *Stochastic Programming*, John Wiley & Sons, Chichester (1994).

[5] 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).

[6] G.H. Lin and M. Fukushima, A class of stochastic mathematical programs with complementarity constraints: Reformulations and algorithms, Technical Report 2003-010, Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto, Japan (2003).

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

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

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