New Reformulations for Stochastic Nonlinear Complementarity Problems [*]

Gui-Hua Lin[†] and Masao Fukushima[‡]

January 28, 2005

Abstract

We consider the stochastic nonlinear complementarity problem (SNCP). We first formulate the problem as a stochastic mathematical program with equilibrium constraints and then, in order to develop efficient algorithms, we give some reformulations of the problem. Furthermore, based on the reformulations, we propose a smoothed penalty method for solving SNCP. A rigorous convergence analysis is also given.

Key Words. stochastic nonlinear complementarity problem, stochastic mathematical program with equilibrium constraints, stationarity, subdifferential, convergence.

1 Introduction

The nonlinear complementarity problem (NCP) is to find a vector $x \in \Re^n$ such that

$$x \geq 0, \quad F(x) \geq 0, \quad x^T F(x) = 0,$$

where $F : \Re^n \to \Re^n$ is a mapping. This problem is one of the fundamental problems in the optimization theory and its applications can be found in many fields, see [4] for details. Since in many practical problems, some elements may involve uncertain data, the stochastic nonlinear complementarity problem and stochastic variational inequality problem have been receiving much attention in the recent literature. In particular, Gürkan et al. [5] consider the following stochastic variational inequality problem $\mathrm{VI}(f, C)$:

$$(x - x^*)^T f(x^*) \geq 0, \quad \forall x \in C, \tag{1.1}$$

where $f : \Re^n \to \Re^n$ is an (unobservable) expectation or steady-state function and $C \subseteq \Re^n$ is a convex polyhedron. Since the function f is usually difficult to evaluate exactly, the authors assume that a sequence $\{f_k\}$, converging to the function f in a certain sense, can be observed. Then a solution of the problem $VI(f, C)$ may be obtained by solving the sequence of problems $VI(f_k, C)$. Haurie and Moresino [6] deal with a stochastic oligopoly problem under the S-adapted information structure and formulate its optimality conditions as a large scale variational inequality problem (VIP). A sampling technique to reduce the size of the VIP is discussed as well. More recently, Chen and Fukushima [2] consider the following stochastic linear complementarity problem (SLCP):

$$x \geq 0, \ M(\omega)x + q(\omega) \geq 0, \ x^T(M(\omega)x + q(\omega)) = 0, \quad \text{a.e. } \omega \in \Omega,$$

where Ω is the underlying sample space and, for each ω, $M(\omega) \in \Re^{n \times n}$ and $q(\omega) \in \Re^n$. The authors formulate the SLCP as a problem of minimizing an expected residual defined by an NCP function. Then, they employ a quasi-Monte Carlo method and give some convergence results under suitable assumptions on the involved matrices.

In this paper, for a given mapping $F : \Re^n \times \Omega \to \Re^n$, we consider the following stochastic nonlinear complementarity problem (SNCP):

$$x \geq 0, \ F(x, \omega) \geq 0, \ x^T F(x, \omega) = 0, \quad \text{a.e. } \omega \in \Omega. \tag{1.2}$$

We study problem (1.2) from another point of view. Note that, in general, there may not exist a vector x satisfying the complementarity conditions for (almost) all $\omega \in \Omega$. In order to get a reasonable resolution, we may introduce recourse variables $z(\omega) \geq 0$ to the inequality $F(x, \omega) \geq 0$ and try to find a vector $x \geq 0$ that minimizes the total recourse. Thus, we obtain the following problem:

$$
\begin{aligned}
\min \quad & E_\omega[d^T z(\omega)] \\
\text{s.t.} \quad & x \geq 0, \ F(x, \omega) + z(\omega) \geq 0, \\
& x^T(F(x, \omega) + z(\omega)) = 0, \\
& z(\omega) \geq 0, \quad \omega \in \Omega,
\end{aligned}
\tag{1.3}
$$

where E_ω indicates the expectation with respect to the random variable $\omega \in \Omega$ and d is a constant vector with positive elements. Throughout, we assume that F is continuously differentiable with respect to x and, if ω is a continuous random variable, F is continuous with respect to ω. We still call problem (1.3) SNCP, although it is actually a stochastic mathematical program with equilibrium constraints (SMPEC).

Mathematical programs with equilibrium constraints (MPECs) play an important role in many fields such as engineering design, economic equilibrium, and multilevel games, see the monograph [13] and some attached references. To go one step further, the stochastic mathematical programs with equilibrium constraints have been studied [8, 9, 12, 15, 16, 17] recently. There are two kinds of SMPECs discussed in the literature: One is the lower-level

wait-and-see model, in which the upper-level decision is made at once while a lower-level decision can be made after a random event is observed. The other is the here-and-now model that requires us to make all decisions before a random event is observed. Therefore, problem (1.3) is actually a special here-and-now SMPEC. Some approaches have been proposed for solving here-and-now problems [8, 9, 12]. In particular, the authors [9] presented a smoothing penalty method for the linear case. In this paper, we extend the approach to the nonlinear case.

The following notations will be used later on: All vectors are column vectors, and $x[i]$ stands for the ith coordinate of vector $x \in \Re^n$. For any vectors x and y of the same dimension, $x \perp y$ means $x^T y = 0$. Given a function $H : \Re^n \to \Re^m$ and a vector $x \in \Re^n$, $\nabla H(x) \in \Re^{n \times m}$ is the transposed Jacobian of H at x.

2 Reformulations

It is well known that SMPECs are very difficult to deal with. In order to look for effective algorithms for solving (1.3), we try to reformulate the problem in this section. To this end, we define $Q : \Re^n \times \Omega \to [0, +\infty]$ by

$$Q(x, \omega) := \sup\{-(u + tx)^T F(x, \omega) \mid u + tx \leq d, \ u \geq 0, \ t \leq 0\}$$

and let $\mathcal{F}$ denote the feasible region of (1.3).

2.1 Properties of the function Q

First of all, we have from the duality theorem in linear programming that, for any $x \in \Re^n$ and $\omega \in \Omega$, $Q(x, \omega) < +\infty$ if and only if the set

$$Z(x, \omega) := \left\{ z(\omega) \ \middle| \ \begin{array}{l} x^T(F(x, \omega) + z(\omega)) \leq 0 \\ F(x, \omega) + z(\omega) \geq 0, \ z(\omega) \geq 0 \end{array} \right\}$$

is nonempty and, if $Z(x, \omega)$ is nonempty, there holds

$$Q(x, \omega) = \inf\{d^T z(\omega) \mid z(\omega) \in Z(x, \omega)\}. \tag{2.1}$$

Theorem 2.1 *Let $x \geq 0$ and $\omega \in \Omega$. Then $Q(x, \omega) < +\infty$ if and only if $x[i]F_i(x, \omega) \leq 0$ for every i.*

Proof: (a) Suppose $Q(x, \omega) < +\infty$. We claim that $x[i]F_i(x, \omega) \leq 0$ holds for every i. Otherwise, there must exist an index i such that both $F_i(x, \omega)$ and $x[i]$ are positive. Let t be a real number and $u(t) \in \Re^n$ be defined by $u(t) := tx[i]e_i - tx$. Then, for any $t \leq 0$, we have

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

3

It follows from the definition of Q that

$$
\begin{aligned}
Q(x,\omega) &\geq \sup\{-(u(t)+tx)^T F(x,\omega) \mid t \leq 0\} \\
&= \sup\{-tx[i]F_i(x,\omega) \mid t \leq 0\} \\
&= +\infty,
\end{aligned}
$$

which is a contradiction.

(b) Suppose that $x[i]F_i(x,\omega) \leq 0$ holds for every i. Let

$$
\begin{aligned}
\mathcal{I}_1 &:= \{ i \mid x[i] = 0\}, \\
\mathcal{I}_2 &:= \{ i \mid x[i] > 0,\ F_i(x,\omega) = 0\}, \\
\mathcal{I}_3 &:= \{ i \mid x[i] > 0,\ F_i(x,\omega) < 0\}.
\end{aligned}
$$

It is easy to see that these sets are mutually disjoint and $\mathcal{I}_1 \cup \mathcal{I}_2 \cup \mathcal{I}_3 = \{1, \cdots, n\}$. Denote $C := \{ (u,t) \in \Re^{n+1} \mid u + tx \leq d,\ u \geq 0,\ t \leq 0\}$. Then, we have

$$
Q(x,\omega) = \sup\{-\sum_{i=1}^m (u+tx)[i]F_i(x,\omega) \mid (u,t) \in C\}. \tag{2.2}
$$

(b1) For any $i \in \mathcal{I}_1$ and $(u,t) \in C$, we have $0 \leq u[i] \leq d[i]$. It follows that the term

$$
-\sum_{i \in \mathcal{I}_1} (u+tx)[i]F_i(x,\omega) = -\sum_{i \in \mathcal{I}_1} u[i]F_i(x,\omega)
$$

is bounded on C.

(b2) It is obvious that $\quad -\sum_{i \in \mathcal{I}_2} (u+tx)[i]F_i(x,\omega) = 0.$

(b3) Let $i \in \mathcal{I}_3$. For any $(u,t) \in C$, we have

$$
-\sum_{i \in \mathcal{I}_3} (u+tx)[i]F_i(x,\omega) \leq -\sum_{i \in \mathcal{I}_3} d[i]F_i(x,\omega)
$$

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

We then have from (b1)–(b3) and (2.2) that $Q(x,\omega) < +\infty$. ∎

Corollary 2.1 *Let $x \geq 0$ and $\omega \in \Omega$. If $Q(x,\omega) < +\infty$, we have*

$$
Q(x,\omega) = d^T z(x,\omega), \tag{2.3}
$$

where $z(x,\omega) := \max\{-F(x,\omega),\ 0\}$.

Proof: It is obvious by the definition of $z(x,\omega)$ that

$$
F(x,\omega) + z(x,\omega) \geq 0, \qquad z(x,\omega) \geq 0.
$$

Moreover, noting that $(F(x,\omega) + z(x,\omega))[i] > 0$ implies $F_i(x,\omega) > 0$, we have from Theorem 2.1 that $x[i] = 0$. It follows that

$$
x^T(F(x,\omega) + z(x,\omega)) = 0.
$$

4

Thus, we have $z(x, \omega) \in Z(x, \omega)$. On the other hand, for any $z(\omega) \in Z(x, \omega)$, since

$$z(\omega) \geq -F(x, \omega), \qquad z(\omega) \geq 0,$$

we see from the definition of $z(x, \omega)$ that $z(\omega) - z(x, \omega) \geq 0$, which implies $d^T z(\omega) \geq d^T z(x, \omega)$. Since $z(\omega) \in Z(x, \omega)$ is arbitrary, we obtain (2.3) form (2.1) immediately. ∎

We next show the equivalence between (1.3) and the following problem

$$\min_{x \geq 0} \ E_\omega[Q(x, \omega)]. \tag{2.4}$$

2.2 Continuous case

Let ω be a continuous random variable and $p(\omega)$ represent the probability density function of ω. Suppose that the probability of any nonempty open set in Ω is positive and problem (2.4) has a finite optimal value. We then have the following result.

Theorem 2.2 *If x^* solves problem (2.4), then there exists $z^*(\omega) \in Z(x^*, \omega)$ for each $\omega \in \Omega$ such that $(x^*, z^*(\omega))_{\omega \in \Omega}$ solves problem (1.3). Conversely, if $(x^*, z^*(\omega))_{\omega \in \Omega}$ solves problem (1.3), then x^* solves problem (2.4).*

Proof: (a) Suppose that x^* solves (2.4). Then we claim that

$$Q(x^*, \omega) < +\infty, \qquad \forall \omega \in \Omega. \tag{2.5}$$

In fact, if $Q(x^*, \bar{\omega}) = +\infty$ for some $\bar{\omega} \in \Omega$, we then have from Theorem 2.1 that there exists an index i such that

$$F_i(x^*, \bar{\omega}) > 0, \quad x^*[i] > 0.$$

It follows from the continuity of $F(x^*, \cdot)$ that there is a neighborhood $U(\bar{\omega})$ of $\bar{\omega}$ such that

$$F_i(x^*, \omega) > 0, \quad x^*[i] > 0$$

hold for any $\omega \in U(\bar{\omega})$. Applying Theorem 2.1 again, we see that

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

Therefore,

$$E_\omega[Q(x^*, \omega)] \geq \int_{U(\bar{\omega})} Q(x^*, \omega) p(\omega) d\omega = +\infty.$$

This contradicts the fact that problem (2.4) has a finite optimal value and hence (2.5) must hold. As a result, we have from (2.1) and Corollary 2.1 that, for any $\omega \in \Omega$, there exists $z^*(\omega) \in Z(x^*, \omega)$ such that

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

5

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

$$E_\omega[d^T z^*(\omega)] = E_\omega[Q(x^*, \omega)] \le E_\omega[Q(x, \omega)] \le E_\omega[d^T z(\omega)], \qquad (2.6)$$

where the first inequality follows from the optimality of x^* and the last inequality follows from (2.1) and the fact that $(x, z(\omega))_{\omega \in \Omega} \in \mathcal{F}$ implies $Z(x, \omega)$ is nonempty (and hence there holds (2.1)) for any $\omega \in \Omega$. The inequality (2.6) means that $(x^*, z^*(\omega))_{\omega \in \Omega}$ is an optimal solution of problem (1.3).

(b) Let $(x^*, z^*(\omega))_{\omega \in \Omega}$ be an optimal solution of problem (1.3). Note that $z^*(\omega) \in Z(x^*, \omega)$ for any $\omega \in \Omega$. It then follows from (2.1) that

$$Q(x^*, \omega) \le d^T z^*(\omega), \qquad \omega \in \Omega. \qquad (2.7)$$

We next show that x^* is a global optimal solution of problem (2.4), namely, for any $x \ge 0$,

$$E_\omega[Q(x^*, \omega)] \le E_\omega[Q(x, \omega)]. \qquad (2.8)$$

Let $x \ge 0$.

(b1) Suppose that $Q(x, \omega) < +\infty$ for every $\omega \in \Omega$ and let $z(x, \omega)$ be the vector defined in Corollary 2.1. By the same corollary, we see that $z(x, \omega) \in Z(x, \omega_\ell)$ and $Q(x, \omega) = d^T z(x, \omega)$. It is not difficult to see that $(x, z(x, \omega))_{\omega \in \Omega} \in \mathcal{F}$ and

$$E_\omega[Q(x^*, \omega)] \le E_\omega[d^T z^*(\omega)] \le E_\omega[d^T z(x, \omega)] = E_\omega[Q(x, \omega)],$$

where the first inequality follows from (2.7) and the second inequality follows from the optimality of $(x^*, z^*(\omega))_{\omega \in \Omega}$ to problem (1.3). So, (2.8) is valid in this case.

(b2) If $Q(x, \hat{\omega}) = +\infty$ for some $\hat{\omega} \in \Omega$, in a similar way to (a), we can show that there exists a neighborhood $U(\hat{\omega})$ of $\hat{\omega}$ such that

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

It follows that $E_\omega[Q(x, \omega)] = +\infty$, which implies that (2.8) remains true.

Therefore, x^* is a global optimal solution of problem (2.4) and hence the proof of the theorem is completed. ∎

2.3 Discrete case

Suppose that $\Omega = \{\omega_1, \omega_2, \cdots, \omega_L\}$ and, for each $\ell = 1, \cdots, L$, the probability p_ℓ of the random event ω_ℓ is positive. Also, for simplicity, we denote the functions $F(\cdot, \omega_\ell)$ and $Q(\cdot, \omega_\ell)$ by $F_\ell(\cdot)$ and $Q_\ell(\cdot)$, respectively. Then, problems (1.3) and (2.4) reduce to

$$\begin{aligned} \min \quad & \sum_{\ell=1}^{L} p_\ell d^T z_\ell \\ \text{s.t.} \quad & x \ge 0, \; F_\ell(x) + z_\ell \ge 0, \\ & x^T(F_\ell(x) + z_\ell) = 0, \\ & z_\ell \ge 0, \;\; \ell = 1, \cdots, L \end{aligned} \qquad (2.9)$$

6

and

$$\min_{x \geq 0} \ \theta(x) := \sum_{\ell=1}^{L} p_\ell Q_\ell(x), \tag{2.10}$$

respectively. In a similar way to Theorem 2.2, we can prove the following result.

Theorem 2.3 *If x^* solves problem (2.10), then there exist $z_\ell^* \in Z(x^*, \omega_\ell), \ell = 1, \cdots, L,$ such that $(x^*, z_1^*, \cdots, z_L^*)$ solves the SMPEC (2.9). Conversely, if $(x^*, z_1^*, \cdots, z_L^*)$ solves the SMPEC (2.9), then x^* solves problem (2.10).*

3 Smoothed Penalty Method for Discrete Problems

We continue to discuss the discrete problem (2.9) in this section. From Theorem 2.1 and the positivity of every p_ℓ, we see that problem (2.10) (and hence (2.9) by Theorem 2.3) is equivalent to the following problem:

$$\begin{aligned}
\min \quad & \theta(x) \\
\text{s.t.} \quad & x \geq 0, \quad x[i]F_{\ell,i}(x) \leq 0, \\
& i = 1, \cdots, n, \ \ell = 1, \cdots, L.
\end{aligned} \tag{3.1}$$

It follows from Corollary 2.1 that, for any $x \in \mathcal{X}$,

$$\theta(x) = \sum_{\ell=1}^{L} p_\ell d^T \max(-F_\ell(x), \ 0). \tag{3.2}$$

In the following, we denote by $\mathcal{X}$ the feasible region of (3.1).

Note that, although (3.1) is no longer an SMPEC, 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 (3.1) has a great many constraints. As a remedy of these difficulties, we propose a smoothed penalty method for solving (3.1) in this section.

3.1 Algorithm and convergence

Let ϵ be a nonnegative constant and the function $\phi_\epsilon : \Re \rightarrow [0, +\infty)$ be defined by

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

Obviously, ϕ_ϵ is differentiable everywhere for each $\epsilon > 0$. In the following, we always assume $\epsilon > 0$. Then, by means of this differentiable function ϕ_ϵ and with the help of a smoothed penalty technique, we obtain the following smooth approximation of (3.1):

$$\min_{x \geq 0} \ \vartheta_\epsilon(x) + \rho\delta_\epsilon(x), \tag{3.3}$$

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

$$\vartheta_\epsilon(x) \quad := \quad \sum_{\ell=1}^{L} \sum_{i=1}^{n} p_\ell d[i] \phi_\epsilon(-F_{\ell,i}(x)), \tag{3.4}$$

$$\delta_\epsilon(x) \quad := \quad \sum_{\ell=1}^{L} \sum_{i=1}^{n} \phi_\epsilon(x[i] F_{\ell,i}(x)). \tag{3.5}$$

Let $\bar\rho > 0$ be a sufficiently large constant. When $\epsilon = 0$ and $\rho = \bar\rho$, (3.3) reduces to

$$\min_{x \geq 0} \ \theta(x) + \bar\rho \delta_0(x), \tag{3.6}$$

where $\delta_0(x) = \sum_{\ell=1}^{L} \sum_{i=1}^{n} \max\{x[i] F_{\ell,i}(x), \ 0\}$.

Definition 3.1 A point $x^* \in \mathcal{X}$ is said to be *stationary* to problem (3.1) if there exist Lagrange multiplier vectors λ^* and μ_ℓ^*, $\ell = 1, \cdots, L$, such that

$$0 \in \partial\theta(x^*) - \lambda^* + \sum_{\ell=1}^{L} \sum_{i=1}^{n} \mu_\ell^*[i](F_{\ell,i}(x^*)e_i + x^*[i]\nabla F_{\ell,i}(x^*)), \tag{3.7}$$

$$0 \leq \lambda^* \perp \quad x^* \geq 0, \tag{3.8}$$

$$0 \leq \mu_\ell^*[i] \perp (-x^*[i] F_{\ell,i}(x^*)) \geq 0, \qquad \forall i, \ \forall \ell. \tag{3.9}$$

Here and later, $\partial\theta$ denotes the Clarke subdifferential operator [3] of θ.

For each ℓ and i, we let $\theta_{\ell,i}(x) := \max\{-F_{\ell,i}(x), 0\}$. It then follows that

$$\partial\theta_{\ell,i}(x^*) = \begin{cases} \text{co}\{-\nabla F_\ell(x^*), 0\}, & F_{\ell,i}(x^*) = 0 \\ \{-\nabla F_\ell(x^*)\}, & F_{\ell,i}(x^*) < 0 \\ \{0\}, & F_{\ell,i}(x^*) > 0 \end{cases} \tag{3.10}$$

and

$$\partial\theta(x^*) = \sum_{\ell=1}^{L} \sum_{i=1}^{n} p_\ell \ d[i] \ \partial\theta_{\ell,i}(x^*), \tag{3.11}$$

where co stands for the convex hull.

Definition 3.2 We say $x^* \geq 0$ to be *stationary* to problem (3.6) if there exists a Lagrange multiplier vector λ such that

$$0 \in \partial\theta(x^*) + \bar\rho\partial\delta_0(x^*) - \lambda, \tag{3.12}$$

$$0 \leq \lambda \perp x^* \geq 0. \tag{3.13}$$

For any ℓ and i, let $\delta_{\ell,i}(x) := \max\{x[i] F_{\ell,i}(x), \ 0\}$. Then, since the functions $\delta_{\ell,i}$ are Clarke regular [3], we have

$$\partial\delta_{\ell,i}(x) = \begin{cases} \text{co}\{F_{\ell,i}(x)e_i + x[i]\nabla F_{\ell,i}(x), 0\}, & x[i] F_{\ell,i}(x) = 0 \\ \{F_{\ell,i}(x)e_i + x[i]\nabla F_{\ell,i}(x)\}, & x[i] F_{\ell,i}(x) > 0 \\ \{0\}, & x[i] F_{\ell,i}(x) < 0 \end{cases} \tag{3.14}$$

8

and

$$\partial\delta_0(x) = \sum_{\ell=1}^{L}\sum_{i=1}^{n} \partial\delta_{\ell,i}(x). \tag{3.15}$$

Theorem 3.1 *Let x^* be a stationary point of problem (3.1). Then, x^* is a stationary point of problem (3.6) for any $\bar\rho$ sufficiently large. Conversely, if x^* is a stationary point of problem (3.6), and $\delta_0(x^*) = 0$, i.e., $x^* \in \mathcal{X}$, then x^* is stationary to (3.1).*

Proof: (a) Suppose x^* is a stationary point of problem (3.1). We will show that, when $\bar\rho$ is sufficiently large, x^* is a stationary point of problem (3.6). By the stationarity of x^* to problem (3.1), there exist multiplier vectors λ^* and μ_ℓ^*, $\ell = 1, \cdots, L$, satisfying conditions (3.7)–(3.9).

Let $\lambda := \lambda^*$. Then (3.13) follows from (3.8) immediately. Comparing (3.12) with (3.7), in order to complete the proof of the first part of the theorem, we only need to show that, when $\bar\rho$ is sufficiently large,

$$\sum_{\ell=1}^{L}\sum_{i=1}^{n} \mu_\ell^*[i](F_{\ell,i}(x^*)e_i + x^*[i]\nabla F_{\ell,i}(x^*)) \in \bar\rho\partial\delta_0(x^*).$$

By (3.14) and (3.15), it is sufficient to show that, when $\bar\rho$ is sufficiently large,

$$\forall\ell,\ \forall i, \quad \begin{cases} \mu_\ell^*[i] \in [0, \bar\rho], & x^*[i]F_{\ell,i}(x^*) = 0, \\ \mu_\ell^*[i] = \bar\rho, & x^*[i]F_{\ell,i}(x^*) > 0, \\ \mu_\ell^*[i] = 0, & x^*[i]F_{\ell,i}(x^*) < 0. \end{cases}$$

Indeed, as long as the constant $\bar\rho$ is larger than all $\mu_\ell^*[i]$ and taking into account the fact that $x^*[i]F_{\ell,i}(x^*) \leq 0$ for each ℓ and i, we can readily obtain the above results from (3.9).

(b) Suppose x^* is a stationary point of problem (3.6) and $\delta_0(x^*) = 0$. Then, there exists a multiplier vector λ satisfying (3.12)–(3.13). Note that $x^* \in \mathcal{X}$ implies

$$x^*[i]F_{\ell,i}(x^*) \leq 0, \qquad \forall\ell,\ \forall i. \tag{3.16}$$

It then follows that

$$\partial\delta_{\ell,i}(x^*) = \begin{cases} \mathrm{co}\{F_{\ell,i}(x^*)e_i + x^*[i]\nabla F_{\ell,i}(x^*), 0\}, & x^*[i]F_{\ell,i}(x^*) = 0 \\ \{0\}, & x^*[i]F_{\ell,i}(x^*) < 0 \end{cases}$$

for any ℓ and i, and

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

Condition (3.12) means that there exist multiplier vectors μ_ℓ, $\ell = 1, \cdots, L$, such that

$$\forall\ell,\ \forall i, \quad \begin{cases} \mu_\ell[i] \in [0, \bar\rho], & x^*[i]F_{\ell,i}(x^*) = 0 \\ \mu_\ell[i] = 0, & x^*[i]F_{\ell,i}(x^*) < 0 \end{cases} \tag{3.17}$$

9

and

$$0 \in \partial\theta(x^*) - \lambda + \sum_{\ell=1}^{L}\sum_{i=1}^{n}\mu_\ell[i](F_{\ell,i}(x^*)e_i + x^*[i]\nabla F_{\ell,i}(x^*)). \tag{3.18}$$

Let $\lambda^* := \lambda$ and

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

Then (3.7)–(3.8) follow from (3.18) and (3.13), and (3.9) follows from (3.16)–(3.17). Therefore, x^* is a stationary point of problem (3.1). ∎

We then have the following algorithm for problem (3.1).

Algorithm SP:

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

Step 2: Solve problem (3.3) with $\epsilon = \epsilon_k$ and $\rho = \rho_k$ to get a stationary point x^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

$$\lim_{k\to\infty}\epsilon_k = 0, \qquad \lim_{k\to\infty}\rho_k = \bar{\rho}, \tag{3.19}$$

where $\bar{\rho}$ is a sufficiently large constant. We next investigate the limiting behavior of the sequence $\{x^k\}$ generated by Algorithm SP. The convergence result can be stated as follows.

theorem Suppose that Algorithm SP generates a sequence $\{x^k\}$ of stationary points of (3.3) with $\epsilon = \epsilon_k$ and $\rho = \rho_k$. Then any accumulation point x^* of the sequence $\{x^k\}$ must be a stationary point of problem (3.6). Furthermore, if $\delta_0(x^*) = 0$, then x^* is a stationary point of problem (3.1).

Proof: Assume without loss of generality that $\lim_{k\to\infty} x^k = x^*$. We will show that x^* is stationary to problem (3.6), i.e., there exists a multiplier vector λ such that (3.12)–(3.13) hold.

First of all, by the stationarity of x^k for problem (3.3) with $\epsilon = \epsilon_k$ and $\rho = \rho_k$, there exists some Lagrange multiplier vector λ^k such that

$$\nabla\vartheta_{\epsilon_k}(x^k) + \rho_k\nabla\delta_{\epsilon_k}(x^k) - \lambda^k = 0, \tag{3.20}$$

$$0 \leq \lambda^k \perp x^k \geq 0. \tag{3.21}$$

Note that, by (3.4) and (3.5),

$$\nabla\vartheta_{\epsilon_k}(x^k) = -\sum_{\ell=1}^{L}\sum_{i=1}^{n}p_\ell d[i]\phi'_{\epsilon_k}(-F_{\ell,i}(x^k))\nabla F_{\ell,i}(x^k), \tag{3.22}$$

$$\nabla\delta_{\epsilon_k}(x^k) = \sum_{\ell=1}^{L}\sum_{i=1}^{n}\phi'_{\epsilon_k}(x^k[i]F_{\ell,i}(x^k))(F_{\ell,i}(x^k)e_i + x^k[i]\nabla F_{\ell,i}(x^k)), \tag{3.23}$$

10

where

$$\phi'_{\epsilon_k}(t) = \frac{1}{2}\left(\frac{t}{\sqrt{t^2 + \epsilon_k^2}} + 1\right), \qquad \forall\, t \in \Re. \tag{3.24}$$

We can then rewrite (3.20) as

$$\lambda^k = \sum_{\ell=1}^{L}\sum_{i=1}^{n}\mu_\ell^k[i](F_{\ell,i}(x^k)e_i + x^k[i]\nabla F_{\ell,i}(x^k)) - \sum_{\ell=1}^{L}\sum_{i=1}^{n}p_\ell d[i]\sigma_\ell^k[i]\nabla F_{\ell,i}(x^k), \tag{3.25}$$

where, for any ℓ and any i,

$$\mu_\ell^k[i] \;\; := \;\; \rho_k \phi'_{\epsilon_k}(x^k[i]F_{\ell,i}(x^k)), \tag{3.26}$$
$$\sigma_\ell^k[i] \;\; := \;\; \phi'_{\epsilon_k}(-F_{\ell,i}(x^k)). \tag{3.27}$$

Since both $\{\rho_k\}$ and $\{x^k\}$ are bounded, we see that $\{\mu_\ell^k\}$ and $\{\sigma_\ell^k\}$ are bounded for every ℓ. It further follows from (3.25) that the sequence $\{\lambda^k\}$ is also bounded. Without loss of generality, we may assume that the following limits exist:

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

Taking a limit in (3.21) and (3.25), we obtain (3.13) and

$$\lambda = \sum_{\ell=1}^{L}\sum_{i=1}^{n}\mu_\ell[i](F_{\ell,i}(x^*)e_i + x^*[i]\nabla F_{\ell,i}(x^*)) - \sum_{\ell=1}^{L}\sum_{i=1}^{n}p_\ell d[i]\sigma_\ell[i]\nabla F_\ell(x^*). \tag{3.28}$$

Thus, in order to show that x^* is a stationary point of problem (3.6), we only need to prove that the vector on the right-hand side of (3.28) belongs to the set $\bar{\rho}\partial\delta_0(x^*) + \partial\theta(x^*)$.

(i) We first prove that

$$\sum_{\ell=1}^{L}\sum_{i=1}^{n}\mu_\ell[i](F_{\ell,i}(x^*)e_i + x^*[i]\nabla F_{\ell,i}(x^*)) \in \bar{\rho}\partial\delta_0(x^*). \tag{3.29}$$

By (3.15), it is sufficient to show that, for any ℓ and i,

$$\mu_\ell[i](F_{\ell,i}(x^*)e_i + x^*[i]\nabla F_{\ell,i}(x^*)) \in \bar{\rho}\partial\delta_{\ell,i}(x^*),$$

which, by (3.14), is equivalent to showing

$$\begin{cases} \mu_\ell[i] \in [0,\bar{\rho}], & x^*[i]F_{\ell,i}(x^*) = 0, \\ \mu_\ell[i] = \bar{\rho}\,, & x^*[i]F_{\ell,i}(x^*) > 0, \\ \mu_\ell[i] = 0\,, & x^*[i]F_{\ell,i}(x^*) < 0. \end{cases} \tag{3.30}$$

In fact, we can obtain (3.30) immediately from (3.19) and the facts that

$$\lim_{k\to\infty} x^k = x^*$$

and

$$\mu_\ell^k[i] = \frac{\rho_k}{2}\left(\frac{x^k[i]F_{\ell,i}(x^k)}{\sqrt{(x^k[i]F_{\ell,i}(x^k))^2 + \epsilon_k^2}} + 1\right).$$

This completes the proof of (3.29).

(ii) We next prove that

$$-\sum_{\ell=1}^{L}\sum_{i=1}^{n} p_\ell d[i]\sigma_\ell[i]\nabla F_{\ell,i}(x^*) \in \partial\theta(x^*).$$

By (3.11), it is enough to show

$$-\sigma_\ell[i]\nabla F_{\ell,i}(x^*) \in \partial\theta_{\ell,i}(x^*), \qquad \forall\ell, \ \forall i. \tag{3.31}$$

There are three cases:

(iia) Suppose $F_{\ell,i}(x^*) = 0$. We then have from (3.24) and (3.27) that

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

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

$$-\sigma_\ell[i]\nabla F_{\ell,i}(x^*) \in \mathrm{co}\{-\nabla F_{\ell,i}(x^*), 0\} = \partial\theta_{\ell,i}(x^*),$$

where the equality follows from (3.10).

(iib) Suppose $F_{\ell,i}(x^*) < 0$. Note that

$$\sigma_\ell^k[i] = \frac{1}{2}\left(\frac{-F_{\ell,i}(x^k)}{\sqrt{(F_{\ell,i}(x^k))^2 + \epsilon_k^2}} + 1\right), \qquad \forall k.$$

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

$$-\sigma_\ell[i]\nabla F_{\ell,i}(x^*) = -\nabla F_{\ell,i}(x^*) \in \partial\theta_{\ell,i}(x^*).$$

(iic) Suppose $F_{\ell,i}(x^*) > 0$. It is easy to show that, for any k,

$$0 \leq \sigma_\ell^k[i] \leq \frac{\epsilon_k}{2(\sqrt{(F_{\ell,i}(x^k))^2 + \epsilon_k^2} + F_{\ell,i}(x^k))}.$$

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

$$-\sigma_\ell[i]\nabla F_{\ell,i}(x^*) = 0 \in \partial\theta_{\ell,i}(x^*).$$

Consequently, (3.31) holds in each case. (i) and (ii) indicate that the vector on the right-hand side of (3.28) belongs to the set $\bar\rho\partial\delta_0(x^*) + \partial\theta(x^*)$. This completes the proof of the first part of the theorem. The second half readily follows from Theorem 3.1. ∎

3.2 Preliminary numerical results

We have tested the proposed method on the following example.

Example 3.1 Consider the following SMPEC formulation of an SLCP:

$$
\begin{aligned}
\min \quad & p(z_1[1] + z_1[2]) + (1-p)(z_2[1] + z_2[2]) \\
\text{s.t.} \quad & z_1[1] \geq 0,\ z_1[2] \geq 0,\ z_2[1] \geq 0,\ z_2[2] \geq 0, \\
& 0 \leq \begin{pmatrix} x[1] \\ x[2] \end{pmatrix} \perp \begin{pmatrix} 2x[2] - 5 + z_1[1] \\ x[1] + x[2] - 5 + z_1[2] \end{pmatrix} \geq 0, \\
& 0 \leq \begin{pmatrix} x[1] \\ x[2] \end{pmatrix} \perp \begin{pmatrix} 3x[1] + 2x[2] - 18 + z_2[1] \\ x[1] - 4 + z_2[2] \end{pmatrix} \geq 0,
\end{aligned} \tag{3.32}
$$

where p is a constant such that $0 < p < 1$. For this problem, problem (3.1) becomes

$$
\begin{aligned}
\min \quad & p(\max\{5 - 2x[2],\ 0\} + \max\{5 - x[1] - x[2],\ 0\}) \\
& + (1-p)(\max\{18 - 3x[1] - 2x[2],\ 0\} + \max\{4 - x[1],\ 0\}) \\
\text{s.t.} \quad & x[1] \geq 0,\quad x[2] \geq 0, \\
& x[1](2x[2] - 5) \leq 0,\quad x[2](x[1] + x[2] - 5) \leq 0, \\
& x[1](3x[1] + 2x[2] - 18) \leq 0,\quad x[2](x[1] - 4) \leq 0.
\end{aligned} \tag{3.33}
$$

We tested two cases: $p = 0.2$ and $p = 0.8$. The corresponding solutions are $x = (4, 1)$ and $x = (2.5, 2.5)$, respectively. It is easy to verify that the linear independence constraint qualification holds at the solutions. This indicates that problem (3.33) has ordinary constraints, unlike the original SMPEC (3.32) that does not satisfy any standard constraint qualification at any feasible point.

In our experiments, we employed the MATLAB 6.5 built-in solver *fmincon* to solve the subproblems (3.3). We set $\rho_0 = 10^3$, $\bar\rho = 10^5$, and updated the parameter by $\rho_{k+1} = \min\{10\rho_k,\ \bar\rho\}$. In addition, the initial point is chosen to be $x^0 = (0, 0)$ and the computed solution x^k at the kth iteration is used as the starting point in the next iteration.

The computational results are reported in Tables 1 and 2. In the table, *Ite* stands for the number of iterations spent by *fmincon* to solve the subproblems. The results shown in the tables reveal that the proposed method is able to solve the example successfully.

Table 1: Computational Results for $p = 0.2$

ϵ_k	ρ_k	x^k	Ite
10^{-2}	10^3	(3.8338,1.0710)	26
10^{-3}	10^4	(3.9521,1.0153)	7
10^{-4}	10^5	(3.9853,1.0041)	10
10^{-6}	10^5	(3.9999,1.0000)	21
10^{-8}	10^5	(4.0000,1.0000)	12

Table 2: Computational Results for $p = 0.8$

ϵ_k	ρ_k	x^k	Ite
10^{-2}	10^3	(2.4847,2.4359)	15
10^{-3}	10^4	(2.4954,2.4796)	6
10^{-4}	10^5	(2.4986,2.4935)	6
10^{-6}	10^5	(2.5000,2.4999)	9
10^{-8}	10^5	(2.5000,2.5000)	7

4 Conclusions

We have shown that the stochastic nonlinear complementarity problem can be formulated as a stochastic mathematical program with equilibrium constraints. In order to develop some efficient methods for solving the problem, some reformulations of the problem have been given. Based on these reformulations, a smoothed penalty method has been proposed and a rigorous convergence analysis has also been presented.

References

[1] J.R. BIRGE AND F. LOUVEAUX, Introduction to Stochastic Programming, Springer, New York, 1997.

[2] X. CHEN AND M. FUKUSHIMA, *Expected Residual Minimization Method for stochastic linear complementarity problems*, Mathematics of Operations Research, to appear.

[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] G. GÜRKAN, A.Y. ÖZGE AND S.M. ROBINSON, *Sample-path solution of stochastic variational inequalities*, Mathematical Programming, 84 (1999), 313–333.

[6] A. HAURIE AND F. MORESINO, *S-adapted oligopoly equilibria and approximations in stochastic variational inequalities, with application to option pricing*, Annals of Operations Research, 114 (2002), 183–201.

[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 class of stochastic mathematical programs with complementarity constraints: Reformulations and algorithms*, Journal of Industrial and Management Optimization, 1 (2005), 99–122.

[10] G.H. LIN AND M. FUKUSHIMA, *A modified relaxation scheme for mathematical programs with complementarity constraints*, Annals of Operations Research, 133 (2005), 63–84.

[11] G.H. LIN AND M. FUKUSHIMA, *New relaxation method for mathematical programs with complementarity constraints*, Journal of Optimization Theory and Applications, 118 (2003), 81–116.

[12] G.H. LIN AND M. FUKUSHIMA, *Regularization method for stochastic mathematical programs with complementarity constraints*, European Series of Applied and Industrial Mathematics (ESAIM): Control, Optimisation and Calculus of Variations, 11 (2005), to appear.

[13] Z.Q. LUO, J.S. PANG, AND D. RALPH, Mathematical Programs with Equilibrium Constraints, Cambridge University Press, Cambridge, United Kingdom, 1996.

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

[15] A. SHAPIRO, *Stochastic mathematical programs with equilibrium constraints*, Preprint, School of Industrial and System Engineering, Georgia Institute of Technology, Antalanta, Georgia, USA, 2004.

[16] A. SHAPIRO AND H. XU, *Stochastic mathematical programs with equilibrium constraints, modeling and sample average approximation*, Preprint, School of Industrial and System Engineering, Georgia Institute of Technology, Antalanta, Georgia, USA, 2005.

[17] H. XU, *An implicit programming approach for a class of stochastic mathematical programs with linear complementarity constraints*, Preprint, School of Mathematics, University of Southampton, 2004.