Solving Stochastic Mathematical Programs with Equilibrium Constraints via Approximation and Smoothing Implicit Programming with Penalization *

Gui-Hua Lin[†], Xiaojun Chen[‡] and Masao Fukushima[§]

May 10, 2005

Abstract In this paper, we consider the stochastic mathematical programs with equilibrium constraints, which includes two kinds of models called here-and-now and lower-level wait-and-see problems. We present a combined smoothing implicit programming and penalty method for the problems with a finite sample space. Then, we suggest a quasi-Monte Carlo approximation method for solving a problem with continuous random variables. A comprehensive convergence theory is included as well. We further report numerical results with the so-called picnic vender decision problem.

Key words. Stochastic mathematical program with equilibrium constraints, wait-and-see, here-and-now, smoothing implicit programming, penalty method, quasi-Monte Carlo method.

1 Introduction

Mathematical program with equilibrium constraints (MPEC) is a constrained optimization problem in which the essential constraints are defined by a parametric variational inequality. This problem can be regarded as a generalization of a bilevel programming problem and it therefore plays an important role in many fields such as engineering design, economic equilibrium, multilevel game, and mathematical programming itself. For more details, see the monograph of Luo et al [12] and the references therein.

In [9], the authors considered *stochastic* mathematical programs with equilibrium constraints (SMPECs). As the bilevel nature of MPECs allows the uncertainty to enter at different levels, the authors give two formulations of SMPECs in [9]. In the first formulation, only the upper-level decision is made under an uncertain circumstance, and the lower-level decision is made after the random event ω is observed. This results in the following problem, which is called the *lower-level wait-and-see* model:

$$\begin{array}{ll} \text{minimize} & E_\omega[f(x,y(\omega),\omega)] \\ \text{subject to} & x \in X, \\ & y(\omega) \text{ solves } \mathrm{VI}(F(x,\cdot,\omega),C(x,\omega)), \quad \omega \in \Omega \quad \text{a.e.,} \end{array} \tag{1.1}$$

*This work was supported in part by the Scientific Research Grant-in-Aid from Japan Society for the Promotion of Science.

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

‡Department of Mathematical System Science, Hirosaki University, Hirosaki 036-8560, Japan. E-mail: chen@cc.hirosaki-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.

where $X \subseteq \Re^n, f : \Re^{n+m} \times \Omega \to \Re, F : \Re^{n+m} \times \Omega \to \Re^m,\ C : \Re^n \times \Omega \to 2^{\Re^m},\ E_\omega$ means expectation with respect to the random variable $\omega \in \Omega$, "a.e." is the abbreviation for "almost everywhere", and $\mathrm{VI}(F(x,\cdot,\omega), C(x,\omega))$ denotes the variational inequality defined by the pair $(F(x,\cdot,\omega), C(x,\omega))$. Note that the wait-and-see model [18] in the classical stochastic programming study is not an optimization problem. However, the lower-level wait-and-see model (1.1) is an optimization problem in which essential variables consist of the upper-level decision x.

When $C(x,\omega) \equiv \Re^m_+$ for any $x \in X$ and almost every $\omega \in \Omega$ in problem (1.1), the variational inequality constraints reduce to the complementarity constraints and problem (1.1) is equivalent to the following stochastic mathematical program with complementarity constraints:

$$\begin{array}{ll} \text{minimize} & E_\omega[f(x, y(\omega), \omega)] \\ \text{subject to} & x \in X, \\ & y(\omega) \geq 0,\ F(x, y(\omega), \omega) \geq 0, \\ & y(\omega)^T F(x, y(\omega), \omega) = 0, \quad \omega \in \Omega \ \text{ a.e.} \end{array} \tag{1.2}$$

On the other hand, if the set-valued function C in problem (1.1) is defined by $C(x,\omega) := \{y \in \Re^m |\ c(x,y,\omega) \leq 0\}$, where $c(\cdot,\cdot,\omega)$ is continuously differentiable, then, under some suitable conditions, $\mathrm{VI}(F(x,\cdot,\omega), C(x,\omega))$ has an equivalent Karush-Kuhn-Tucker representation

$$\begin{array}{l} F(x, y(\omega), \omega) + \nabla_y c(x, y(\omega), \omega)\lambda(x,\omega) = 0, \\ \lambda(x,\omega) \geq 0, \quad c(x, y(\omega), \omega) \leq 0, \quad \lambda(x,\omega)^T c(x, y(\omega), \omega) = 0, \end{array}$$

where $\lambda(x,\omega)$ is the Lagrange multiplier vector [14]. As a result, problem (1.1) can be reformulated as a program like (1.2) under some conditions, see [12] for more details. Hence, problem (1.2) constitutes an important subclass of SMPECs.

Another formulation that we are particularly interested in is the following problem that requires us to make all decisions at once, before ω is observed:

$$\begin{array}{ll} \text{minimize} & E_\omega[f(x, y, \omega) + d^T z(\omega)] \\ \text{subject to} & x \in X, \\ & y \geq 0,\ F(x, y, \omega) + z(\omega) \geq 0, \\ & y^T(F(x, y, \omega) + z(\omega)) = 0, \\ & z(\omega) \geq 0, \quad \omega \in \Omega \ \text{ a.e.} \end{array} \tag{1.3}$$

Here, both the decisions x and y are independent of the random variable ω, $z(\omega)$ is called a recourse variable, and $d \in \Re^m$ is a vector with positive elements. We call (1.3) a *here-and-now* model. Compared with the lower-level wait-and-see model (1.2), the here-and-now model (1.3) involves more variables and hence seems more difficult to deal with. Moreover, a feasible vector y in (1.3) is required to satisfy the complementarity condition for almost all $\omega \in \Omega$, which is different from the ordinary complementarity condition if Ω has more than one realization. Because of this restriction, some results for MPECs cannot be applied to (1.3) directly. Special new treatment has to be developed.

In [9], the authors proposed a smoothing implicit programming approach for solving the SMPECs with a finite sample space. Subsequently, there have been a number of attempts [2, 10, 11, 16, 17, 19] to deal with various models of SMPECs. In particular, Lin and Fukushima [10, 11] suggested a smoothing penalty method and a regularization method, respectively, for a special class of here-and-now problems. Shapiro and Xu [16, 17, 19] discussed the sample average approximation and implicit programming approaches for the lower-level wait-and-see problems. In addition, Birbil et al [2] considered an SMPEC in which both the objective and constraints involve expectations.

In this paper, we will mainly consider the here-and-now model (1.3). Especially, unlike our past work [9, 10, 11], we will also deal with an SMPEC with continuous random variables. In Section 3, we describe the combined smoothing implicit programming and penalty method proposed in [9] for the discrete SMPECs and, in Section 4, we suggest a quasi-Monte Carlo method to discretize the here-and-now problem with continuous random variables. Comprehensive convergence theory is established as well. In Section 5, we give some numerical experiments with the so-called picnic-vender decision problem. This appears to be the first attempt to report numerical results for SMPECs in the literature.

Notation used in the paper: Throughout, all vectors are thought as column vectors and $x[i]$ stands for the ith coordinate of $x \in \Re^n$. For a matrix M and an index set $\mathcal{K}$, we let $M[\mathcal{K}]$ be the principal submatrix of M whose elements consist of those of M indexed by $\mathcal{K}$. For any vectors u and v of the same dimension, we denote $u \perp v$ to mean $u^T v = 0$. For a given function $F : \Re^n \to \Re^m$ and a vector $x \in \Re^n$, $\nabla F(x)$ is the transposed Jacobian of F at x and $\mathcal{I}_F(x) := \{i \mid F_i(x) = 0\}$ stands for the active index set of F at x. In addition, I and O denote the identity matrix and the zero matrix with suitable dimension, respectively.

2 Preliminaries

In this section, we recall some basic concepts and properties that will be used later on. First we consider the standard smooth nonlinear programming problem:

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & c_i(z) \leq 0, \quad i = 1, \cdots, t, \\ & c_i(z) = 0, \quad i = t+1, \cdots, \nu. \end{array} \tag{2.1}$$

We will use the standard definition of stationarity, i.e., a feasible point z is said to be *stationary* to (2.1) if there exists a Lagrange multiplier vector $\lambda \in \Re^\nu$ satisfying the Karush-Kuhn-Tucker conditions

$$\begin{array}{l} \nabla f(z) + \nabla c(z)\lambda = 0, \\ \lambda[i] \geq 0, \ \lambda[i] c_i(z) = 0, \quad \ i = 1, \cdots, t. \end{array}$$

We next consider the mathematical program with complementarity constraints:

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & g(z) \leq 0, \ h(z) = 0, \\ & G(z) \geq 0, \ H(z) \geq 0, \\ & G(z)^T H(z) = 0, \end{array} \tag{2.2}$$

where $f : \Re^s \to \Re, g : \Re^s \to \Re^{s_1}, h : \Re^s \to \Re^{s_2}$, and $G, H : \Re^s \to \Re^{s_3}$ are all continuously differentiable functions. Let $\mathcal{Z}$ denote the feasible region of the MPEC (2.2).

It is well-known that the MPEC (2.2) fails to satisfy a standard constraint qualification (CQ) at any feasible point [5], which causes a difficulty in dealing with MPECs by a conventional nonlinear programming approach. The following special CQ turns out to be useful in the study of MPECs.

Definition 2.1 The MPEC-*linear independence constraint qualification* (MPEC-LICQ) is said to hold at $\bar{z} \in \mathcal{Z}$ if the set of vectors

$$\left\{ \nabla g_l(\bar{z}), \nabla h_r(\bar{z}), \nabla G_i(\bar{z}), \nabla H_j(\bar{z}) \;\middle|\; l \in \mathcal{I}_g(\bar{z}),\ r = 1, \cdots, s_2,\ i \in \mathcal{I}_G(\bar{z}),\ j \in \mathcal{I}_H(\bar{z}) \right\}$$

is linearly independent.

Definition 2.2 [15] (1) $\bar{z} \in \mathcal{Z}$ is called a *Clarke or C-stationary* point of problem (2.2) if there exist multiplier vectors $\bar{\lambda} \in \Re^{s_1}$, $\bar{\mu} \in \Re^{s_2}$, and $\bar{u}, \bar{v} \in \Re^{s_3}$ such that $\bar{\lambda} \geq 0$ and

$$\nabla f(\bar{z}) + \sum_{i \in \mathcal{I}_g(\bar{z})} \bar{\lambda}[i] \nabla g_i(\bar{z}) + \sum_{i=1}^{s_2} \bar{\mu}[i] \nabla h_i(\bar{z}) - \sum_{i \in \mathcal{I}_G(\bar{z})} \bar{u}[i] \nabla G_i(\bar{z}) - \sum_{i \in \mathcal{I}_H(\bar{z})} \bar{v}[i] \nabla H_i(\bar{z}) = 0, \quad (2.3)$$

$$\bar{u}[i]\bar{v}[i] \geq 0, \quad i \in \mathcal{I}_G(\bar{z}) \cap \mathcal{I}_H(\bar{z}). \quad (2.4)$$

(2) $\bar{z} \in \mathcal{Z}$ is called a *strongly or S-stationary* point of problem (2.2) if there exist multiplier vectors $\bar{\lambda}, \bar{\mu}, \bar{u}$, and $\bar{v}$ such that (2.3) holds with

$$\bar{u}[i] \geq 0, \quad \bar{v}[i] \geq 0, \quad i \in \mathcal{I}_G(\bar{z}) \cap \mathcal{I}_H(\bar{z}).$$

It is easy to see that S-stationarity implies C-stationarity. Moreover, under the strict complementarity condition (namely, $\mathcal{I}_G(\bar{z}) \cap \mathcal{I}_H(\bar{z}) = \emptyset$), they are equivalent.

Definition 2.3 [6] Suppose that M is an $m \times m$ matrix.

(1) We call M a *P-matrix* if all the principal minors of M are positive, or equivalently,

$$\max_{1 \leq i \leq m} y[i](My)[i] > 0, \qquad 0 \neq \forall y \in \Re^m,$$

and we call M a P_0*-matrix* if all the principal minors of M are nonnegative, or equivalently,

$$\max_{1 \leq i \leq m} y[i](My)[i] \geq 0, \qquad \forall y \in \Re^m.$$

(2) We call M is a *nondegenerate* matrix if all of its principal submatrices are nonsingular.

(3) We call M an R_0*-matrix* if

$$y \geq 0,\ My \geq 0,\ y^T My = 0 \quad \Longrightarrow \quad y = 0.$$

Obviously a P-matrix is a P_0-matrix and a nondegenerate matrix. Moreover, it is easy to see that a P-matrix is an R_0-matrix. If M is a P_0-matrix and μ is any positive number, then the matrix $M + \mu I$ is a P-matrix.

For given $N \in \Re^{m \times n}$, $M \in \Re^{m \times m}$, $q \in \Re^m$, and two positive scalars ϵ and μ, we define

$$\Phi_{\epsilon,\mu}(x, y, w; N, M, q) := \begin{pmatrix} Nx + (M + \epsilon I)y + q - w \\ \phi_\mu(y[1], w[1]) \\ \vdots \\ \phi_\mu(y[m], w[m]) \end{pmatrix}, \quad (2.5)$$

where $\phi_\mu : \Re^2 \to \Re$ is the perturbed Fischer-Burmeister function defined by $\phi_\mu(a, b) := a + b - \sqrt{a^2 + b^2 + 2\mu^2}$. Then we have the following well-known result [4].

Theorem 2.1 *Suppose that M is a P_0-matrix. Then, for given $x \in \Re^n$, $\epsilon > 0$, and $\mu > 0$, we have the following statements:*

(i) *The function $\Phi_{\epsilon,\mu}$ defined by (2.5) is continuously differentiable with respect to (y, w) and the Jacobian matrix $\nabla_{(y,w)}\Phi_{\epsilon,\mu}(x, y, w; N, M, q)$ is nonsingular everywhere;*

(ii) *The equation $\Phi_{\epsilon,\mu}(x, y, w; N, M, q) = 0$ has a unique solution $(y(x, \epsilon, \mu), w(x, \epsilon, \mu))$, which is continuously differentiable with respect to x and satisfies*

$$\begin{aligned} &y(x, \epsilon, \mu) > 0, \quad w(x, \epsilon, \mu) > 0, \\ &y(x, \epsilon, \mu)[i]w(x, \epsilon, \mu)[i] = \mu^2, \quad i = 1, \cdots, m. \end{aligned}$$

In the rest of the paper, to mitigate the notational complication, we assume $\epsilon = \mu$ and denote $\Phi_{\epsilon,\mu}$, $y(x, \epsilon, \mu)$, and $w(x, \epsilon, \mu)$ by Φ_μ, $y(x, \mu)$, and $w(x, \mu)$, respectively. Our analysis will remain valid, however, even though the two parameters are treated independently.

Suppose that M is a P_0-matrix and $\mu > 0$. Theorem 2.1 indicates that the smooth equation

$$\Phi_\mu(x, y, w; N, M, q) = 0 \tag{2.6}$$

gives two smooth functions $y(\cdot, \mu)$ and $w(\cdot, \mu)$. Note that

$$\phi_\mu(a, b) = 0 \quad \Longleftrightarrow \quad a \geq 0, b \geq 0, ab = \mu^2.$$

As a result, the equation (2.6) is equivalent to the system

$$\begin{aligned} &y \geq 0, \; Nx + (M + \mu I)y + q \geq 0, \\ &y[i]\Big(Nx + (M + \mu I)y + q\Big)[i] = \mu^2, \quad i = 1, \cdots, m \end{aligned} \tag{2.7}$$

in the sense that $y(x, \mu)$ solves (2.7) if and only if

$$\Phi_\mu(x, y(x, \mu), w(x, \mu); N, M, q) = 0,$$

where $w(x, \mu) := Nx + (M + \mu I)y(x, \mu) + q$. Since (2.7) with $\mu = 0$ reduces to the linear complementarity problem

$$y \geq 0, \quad Nx + My + q \geq 0, \quad y^T(Nx + My + q) = 0, \tag{2.8}$$

we see that $y(x, \mu)$ tends to a solution of (2.8) as $\mu \to 0$, provided that it is convergent.

In our analysis, we will assume that $y(x, \mu)$ is bounded as $\mu \to 0$. In particular, if M is a P-matrix, then (2.8) has a unique solution for any x and it can be shown that $y(x, \mu)$ actually converges to it as $\mu \to 0$, even without using the regularization term μI in (2.7), see [4].

3 Combined Smoothing Implicit Programming and Penalty Method for Discrete Here-and-Now Problems

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

$$\begin{aligned} \text{minimize} \quad & \sum_{\ell=1}^{L} p_\ell\Big(f(x, y, \omega_\ell) + d^T z_\ell\Big) \\ \text{subject to} \quad & g(x) \leq 0, \; h(x) = 0, \\ & y \geq 0, \; N_\ell x + M_\ell y + q_\ell + z_\ell \geq 0, \\ & y^T(N_\ell x + M_\ell y + q_\ell + z_\ell) = 0, \\ & z_\ell \geq 0, \; \ell = 1, \cdots, L, \end{aligned} \tag{3.1}$$

which corresponds to the discrete case where $\Omega := \{\omega_1, \omega_2, \cdots, \omega_L\}$. The problem with continuous random variables will be considered in the next section. In (3.1), p_ℓ denotes the probability of the random event $\omega_\ell \in \Omega$, the functions $f : \Re^{n+m} \to \Re, g : \Re^n \to \Re^{s_1}$, $h : \Re^n \to \Re^{s_2}$ are all continuously differentiable, $N_\ell \in \Re^{m\times n}$, $M_\ell \in \Re^{m\times m}$, $q_\ell \in \Re^m$ are realizations of the random coefficients, d is a constant vector with positive elements, and z_ℓ is the recourse variable corresponding to ω_ℓ. Throughout we assume $p_\ell > 0$ for all $\ell = 1, \cdots, L$.

It is easy to see that problem (3.1) can be rewritten as

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{\ell=1}^{L} p_\ell\Big(f(x,y,\omega_\ell) + d^T z_\ell\Big) \\ \text{subject to} & g(x) \le 0,\ h(x) = 0,\ z_\ell \ge 0, \\ & N_\ell x + M_\ell y + q_\ell + z_\ell \ge 0,\ \ell = 1, \cdots, L, \\ & y \ge 0,\ Nx + My + q + \sum_{l=1}^{L} z_l \ge 0, \\ & y^T(Nx + My + q + \sum_{l=1}^{L} z_l) = 0 \end{array} \tag{3.2}$$

with $N := \sum_{l=1}^{L} N_l, M := \sum_{l=1}^{L} M_l$, and $q := \sum_{l=1}^{L} q_l$, or equivalently,

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{\ell=1}^{L} p_\ell f(x,y,\omega_\ell) + \mathbf{d}^T\mathbf{z} \\ \text{subject to} & g(x) \le 0,\ h(x) = 0, \\ & \mathbf{y} - \mathbf{D}y = 0,\ \mathbf{z} \ge 0, \\ & \mathbf{y} \ge 0,\ \mathbf{N}x + \mathbf{M}\mathbf{y} + \mathbf{q} + \mathbf{z} \ge 0, \\ & \mathbf{y}^T(\mathbf{N}x + \mathbf{M}\mathbf{y} + \mathbf{q} + \mathbf{z}) = 0, \end{array} \tag{3.3}$$

where

$$\mathbf{y} := \begin{pmatrix} y_1 \\ \vdots \\ y_L \end{pmatrix}, \quad \mathbf{z} := \begin{pmatrix} z_1 \\ \vdots \\ z_L \end{pmatrix}, \quad \mathbf{d} := \begin{pmatrix} p_1 d \\ \vdots \\ p_L d \end{pmatrix}, \quad \mathbf{D} := \begin{pmatrix} I \\ \vdots \\ I \end{pmatrix}, \tag{3.4}$$

and

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

Note that both problems (3.2) and (3.3) are different from ordinary MPECs, because they require $y_1 = y_2 = \cdots = y_L$. This restriction makes the problems harder to deal with than ordinary MPECs. In particular, for any feasible point $(x, y, z_1, \cdots, z_L)$ of problem (3.2), $(Nx + My + q + \sum_{l=1}^{L} z_l)[i] = 0$ implies that $(N_\ell x + M_\ell y + q_\ell + z_\ell)[i] = 0$ holds for every ℓ. This indicates that the MPEC-LICQ does not hold for problem (3.2) in general. On the other hand, since L is usually very large in practice, problem (3.3) is a large-scale program with variables $(x, y, \mathbf{y}, \mathbf{z}) \in \Re^{n+(1+2L)m}$ so that some methods for MPECs may cause more computational difficulties.

In this section, we describe a combined smoothing implicit programming and penalty method for solving the ill-posed MPEC (3.2) directly. This method was originally presented in an unpublished paper [9]. For a complete analysis of the method, we give a somewhat detailed presentation of the method in this paper. It is worth mentioning that a similar smoothing method for ordinary MPECs with linear complementarity constraints has been considered in [4]. However, several differences should be emphasized here: (a) In [4], the matrix $\mathbf{M}$ is assumed to be a P-matrix, whereas in this paper, it is assumed to be a P_0-matrix only; (b) In order to make the new method applicable, in addition to smoothing, we employ a regularization technique and a penalty technique.

As mentioned above, the MPEC-LICQ does not hold for problem (3.2) in general. From now on, the MPEC-LICQ means the one for problem (3.3). On the other hand, because the complementarity constraints in problem (3.2) are lower dimensional, we use them to generate the subproblems.

3.1 SIPP method

Suppose that the matrix M in problem (3.2) is a P_0-matrix. We denote by Λ the matrix $(I, \cdots, I) \in \Re^{m \times mL}$. For each $(x, \mathbf{z})$ and $\mu_k > 0$, let $y(x, \Lambda\mathbf{z}, \mu_k)$ and $w(x, \Lambda\mathbf{z}, \mu_k)$ solve

$$\Phi_{\mu_k}\Big(x, y(x, \Lambda\mathbf{z}, \mu_k), w(x, \Lambda\mathbf{z}, \mu_k); N, M, q + \Lambda\mathbf{z}\Big) = 0. \tag{3.5}$$

The existence and differentiability of the above implicit functions follow from Theorem 2.1. Note that the implicit functions are denoted by $y(x, \Lambda\mathbf{z}, \mu_k)$ and $w(x, \Lambda\mathbf{z}, \mu_k)$, rather than $y(x, \mathbf{z}, \mu_k)$ and $w(x, \mathbf{z}, \mu_k)$, respectively. We then obtain an approximation of problem (3.2)

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{\ell=1}^{L} p_\ell \Big(f(x, y(x, \Lambda\mathbf{z}, \mu_k), \omega_\ell) + d^T z_\ell \Big) \\ \text{subject to} & g(x) \le 0, \ h(x) = 0, \\ & N_\ell x + M_\ell y(x, \Lambda\mathbf{z}, \mu_k) + q_\ell + z_\ell \ge 0, \\ & z_\ell \ge 0, \ \ell = 1, \cdots, L. \end{array} \tag{3.6}$$

Since the feasible region of problem (3.6) is dependent on μ_k, (3.6) may not be easy to solve. Therefore, we apply a penalty technique to this problem and have the following approximation:

$$\begin{array}{ll} \text{minimize} & \theta_k(x, \mathbf{z}) \\ \text{subject to} & g(x) \le 0, \ h(x) = 0, \ \mathbf{z} \ge 0, \end{array} \tag{3.7}$$

where

$$\theta_k(x, \mathbf{z}) \ := \ \sum_{\ell=1}^{L} p_\ell f(x, y(x, \Lambda\mathbf{z}, \mu_k), \omega_\ell) + \mathbf{d}^T\mathbf{z} + \rho_k \sum_{\ell=1}^{L} \psi\Big(-(N_\ell x + M_\ell y(x, \Lambda\mathbf{z}, \mu_k) + q_\ell + z_\ell)\Big),$$

ρ_k is a positive parameter, $\psi : \Re^m \to [0, +\infty)$ is a smooth penalty function, and $z_\ell := (\mathbf{z}[(\ell - 1)m + 1], \cdots, \mathbf{z}[\ell m])^T$ for each ℓ. Some specific penalty functions will be given later. Note that, unlike problem (3.6), the feasible region of problem (3.7) is common for all k.

Now we present our method, called the *combined smoothing implicit programming and penalty method* (SIPP), for problem (3.2): Choose two sequences $\{\mu_k\}$ and $\{\rho_k\}$ of positive numbers satisfying

$$\lim_{k\to\infty} \mu_k = 0, \quad \lim_{k\to\infty} \rho_k = +\infty, \quad \lim_{k\to\infty} \mu_k\rho_k = 0. \tag{3.8}$$

We then solve the problems (3.7) to get a sequence $\{(x^{(k)}, \mathbf{z}^{(k)})\}$ and let

$$y^{(k)} := y(x^{(k)}, \Lambda\mathbf{z}^{(k)}, \mu_k).$$

Note that, by Theorem 2.1, problem (3.7) is a smooth mathematical program. Moreover, under some suitable conditions, (3.7) is a convex program, see [4] for details. Therefore, we may expect that problem (3.7) may be relatively easy to deal with, provided the evaluation of the function $y_\ell(x, \mu_k)$ is not very expensive.

In what follows, we denote by $\mathcal{F}$ and $\mathcal{X}$ the feasible regions of problems (3.2) and (3.7), respectively. Moreover, particular sequences generated by the method will be denoted by $\{x^{(k)}\}, \{y^{(k)}\}$, etc., while general sequences will be denoted by $\{x^k\}, \{y^k\}$, etc. Also, we use (3.4) to generate some related vectors such as $\mathbf{y}^{(k)}, \mathbf{y}^*, \mathbf{z}^{(k)}, \mathbf{z}^*$, and so on.

3.2 Convergence results

We investigate the limiting behavior of a sequence generated by SIPP in this subsection. The following lemma will be used later.

Lemma 3.1 [9] *Suppose the matrix M in (3.2) is a P_0-matrix and, for any bounded sequence $\{(x^k, \mathbf{z}^k)\}$ in $\mathcal{X}$, $\{y(x^k, \Lambda\mathbf{z}^k, \mu_k)\}$ is bounded. If $(x^*, y^*, \mathbf{z}^*) \in \mathcal{F}$ and the submatrix $M[\mathcal{K}^*]$ is nondegenerate, where $\mathcal{K}^* := \{\ i \mid (Nx^* + My^* + q + \Lambda\mathbf{z}^*)[i] = 0\}$, then there exist a neighborhood U^* of $(x^*, y^*, \mathbf{z}^*)$ and a positive constant π^* such that*

$$\|y(x, \Lambda\mathbf{z}, \mu_k) - y\| \le \mu_k \pi^* (\|y\| + \sqrt{m}) \tag{3.9}$$

holds for any $(x, y, \mathbf{z}) \in U^* \cap \mathcal{F}$ and any k.

We first discuss the limiting behavior of local optimal solutions of problems (3.7).

Theorem 3.1 *Let the matrix M in (3.2) be a P_0-matrix, $\psi : \Re^m \to [0, +\infty)$ be a continuously differentiable function satisfying*

$$\psi(0) = 0, \quad \psi(u) \le \psi(u') \ \textit{ for any } u' \ge u \textit{ in } \Re^m, \tag{3.10}$$

and, for each bounded sequence $\{(x^k, \mathbf{z}^k)\}$ in $\mathcal{X}$, $\{y(x^k, \Lambda\mathbf{z}^k, \mu_k)\}$ be bounded. Suppose that the sequence $\{(x^{(k)}, y^{(k)}, \mathbf{z}^{(k)})\}$ generated by SIPP with $(x^{(k)}, \mathbf{z}^{(k)})$ being a local optimal solution of problem (3.7) is convergent to $(x^*, y^*, \mathbf{z}^*) \in \mathcal{F}$. If there exists a neighborhood V^* of $(x^*, y^*, \mathbf{z}^*)$ such that $(x^{(k)}, \mathbf{z}^{(k)})$ minimizes θ_k over $V^*|_{\mathcal{X}} := \{(x, \mathbf{z}) \in \mathcal{X} \mid \exists y \text{ s.t. } (x, y, \mathbf{z}) \in V^*\}$ for all k large enough and the submatrix $M[\mathcal{K}^*]$ is nondegenerate with $\mathcal{K}^*$ being the same as in Lemma 3.1, then $(x^*, y^*, \mathbf{z}^*)$ is a local optimal solution of problem (3.2).

Proof. By Lemma 3.1, there exist a closed sphere $\mathcal{B} \subseteq V^*$ centered at the point $(x^*, y^*, \mathbf{z}^*)$ with positive radius and a positive number π^* such that (3.9) holds for any $(x, y, \mathbf{z}) \in \mathcal{F} \cap \mathcal{B}$ and every k. Since $\mathcal{F} \cap \mathcal{B}$ is a nonempty compact set, the problem

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{\ell=1}^{L} p_\ell f(x, y, \omega_\ell) + \mathbf{d}^T\mathbf{z} \\ \text{subject to} & (x, y, \mathbf{z}) \in \mathcal{F} \cap \mathcal{B} \end{array} \tag{3.11}$$

has an optimal solution, say $(\bar{x}, \bar{y}, \bar{\mathbf{z}})$.

Suppose $(x, y, \mathbf{z}) \in \mathcal{F} \cap \mathcal{B}$. We then have from the mean-value theorem that

$$\begin{aligned} \theta_k(x, \mathbf{z}) \ = \ & \sum_{\ell=1}^{L} p_\ell \Big(f(x, y, \omega_\ell) + (y(x, \Lambda\mathbf{z}, \mu_k) - y)^T \nabla_y f(x, (1-t)y(x, \Lambda\mathbf{z}, \mu_k) + ty, \omega_\ell) \Big) \\ & + \mathbf{d}^T\mathbf{z} + \rho_k \sum_{\ell=1}^{L} \psi\Big(-(N_\ell x + M_\ell y(x, \Lambda\mathbf{z}, \mu_k) + q_\ell + z_\ell) \Big), \end{aligned} \tag{3.12}$$

where $t \in [0,1]$. Note that, by (3.9),

$$\begin{aligned}\|(1-t)y(x,\Lambda\mathbf{z},\mu_k)+ty\| &= \|(1-t)(y(x,\Lambda\mathbf{z},\mu_k)-y)+y\| \\ &\le \|y(x,\Lambda\mathbf{z},\mu_k)-y\|+\|y\| \\ &\le \mu_k\pi^*(\|y\|+\sqrt{m})+\|y\|.\end{aligned}$$

This indicates that the set

$$\{(x,(1-t)y(x,\Lambda\mathbf{z},\mu_k)+ty)\ \big|\ (x,y,\mathbf{z})\in\mathcal{F}\cap\mathcal{B}, t\in[0,1], k=1,2,\cdots\}$$

is bounded. Similarly, we see that

$$\{(x,tM_\ell(y-y(x,\Lambda\mathbf{z},\mu_k)))\ \big|\ (x,y,\mathbf{z})\in\mathcal{F}\cap\mathcal{B}, \ell=1,\cdots,L, t\in[0,1], k=1,2,\cdots\}$$

is also bounded. Then, by the continuous differentiability of both f and ψ, there exists a constant $\tau>0$ such that, for $\ell=1,\cdots,L$,

$$\|\nabla_y f(x,(1-t)y(x,\Lambda\mathbf{z},\mu_k)+ty,\omega_\ell)\| \le \tau, \tag{3.13}$$
$$\|\nabla\psi\Big(tM_\ell(y-y(x,\Lambda\mathbf{z},\mu_k))\Big)\| \le \tau \tag{3.14}$$

hold for any $(x,y,\mathbf{z})\in\mathcal{F}\cap\mathcal{B}, t\in[0,1]$, and every k. Noticing that $(x,y,\mathbf{z})\in\mathcal{F}\cap\mathcal{B}$ implies $N_\ell x+M_\ell y+q_\ell+z_\ell\ge 0$ for each ℓ, we have from (3.10) and (3.14) that

$$\begin{aligned}\psi\Big(-(N_\ell x+M_\ell y(x,\Lambda\mathbf{z},\mu_k)+q_\ell+z_\ell)\Big) &\le \psi\Big(M_\ell(y-y(x,\Lambda\mathbf{z},\mu_k))\Big) \\ &= \psi\Big(M_\ell(y-y(x,\Lambda\mathbf{z},\mu_k))\Big)-\psi(0) \\ &= \nabla\psi\Big(t'M_\ell(y-y(x,\Lambda\mathbf{z},\mu_k))\Big)^T M_\ell\Big(y-y(x,\Lambda\mathbf{z},\mu_k)\Big) \\ &\le \tau\|M_\ell\|\ \|y-y(x,\Lambda\mathbf{z},\mu_k)\|,\end{aligned}$$

where $t'\in[0,1]$ and the second equality follows from the mean-value theorem. This, together with (3.12)–(3.13) and (3.9), yields

$$\begin{aligned}\Big|\theta_k(x,\mathbf{z})-\sum_{\ell=1}^L p_\ell f(x,y,\omega_\ell)-\mathbf{d}^T\mathbf{z}\Big| &\le \tau\|y(x,\Lambda\mathbf{z},\mu_k)-y\|+\Big(\tau\rho_k\sum_{\ell=1}^L\|M_\ell\|\Big)\|y-y(x,\Lambda\mathbf{z},\mu_k)\| \\ &\le \pi^*\tau\Big(\mu_k+\mu_k\rho_k\sum_{\ell=1}^L\|M_\ell\|\Big)(\|y\|+\sqrt{m})\end{aligned}$$

for any $(x,y,\mathbf{z})\in\mathcal{F}\cap\mathcal{B}$ and k. In particular,

$$\Big|\theta_k(\bar{x},\bar{\mathbf{z}})-\sum_{\ell=1}^L p_\ell f(\bar{x},\bar{y},\omega_\ell)-\mathbf{d}^T\bar{\mathbf{z}}\Big| \le \pi^*\tau\Big(\mu_k+\mu_k\rho_k\sum_{\ell=1}^L\|M_\ell\|\Big)(\|\bar{y}\|+\sqrt{m}). \tag{3.15}$$

Moreover, since ψ is always nonnegative, we have from the continuity of f that

$$\begin{aligned}\lim_{k\to\infty}\theta_k(x^{(k)},\mathbf{z}^{(k)}) &\ge \lim_{k\to\infty}\Big(\sum_{\ell=1}^L p_\ell f(x^{(k)},y^{(k)},\omega_\ell)+\mathbf{d}^T\mathbf{z}^{(k)}\Big) \\ &= \sum_{\ell=1}^L p_\ell f(x^*,y^*,\omega_\ell)+\mathbf{d}^T\mathbf{z}^*. \end{aligned}\tag{3.16}$$

Note that, by the fact that $\mathcal{F} \cap \mathcal{B} \subseteq V^*$, $(x^{(k)}, \mathbf{z}^{(k)})$ is an optimal solution of the problem

$$\begin{array}{ll} \text{minimize} & \theta_k(x, \mathbf{z}) \\ \text{subject to} & (x, \mathbf{z}) \in \mathcal{X}_1 := \{\, (x, \mathbf{z}) \in \mathcal{X} \mid \exists\, y \text{ s.t. } (x, y, \mathbf{z}) \in \mathcal{F} \cap \mathcal{B} \,\}, \end{array}$$

provided k is large enough, and $(\bar{x}, \bar{\mathbf{z}})$ is a feasible point of this problem. We then have from (3.15) that, for every k sufficiently large,

$$\begin{aligned} \theta_k(x^{(k)}, \mathbf{z}^{(k)}) &\leq \theta_k(\bar{x}, \bar{\mathbf{z}}) \\ &\leq \sum_{\ell=1}^{L} p_\ell f(\bar{x}, \bar{y}, \omega_\ell) + \mathbf{d}^T \bar{\mathbf{z}} + \pi^* \tau \Big(\mu_k + \mu_k \rho_k \sum_{\ell=1}^{L} \|M_\ell\| \Big) (\|\bar{y}\| + \sqrt{m}). \end{aligned} \tag{3.17}$$

Therefore, taking into account the equality (3.16) and the assumption (3.8), we have by letting $k \to \infty$ in (3.17) that

$$\sum_{\ell=1}^{L} p_\ell f(x^*, y^*, \omega_\ell) + \mathbf{d}^T \mathbf{z}^* \leq \sum_{\ell=1}^{L} p_\ell f(\bar{x}, \bar{y}, \omega_\ell) + \mathbf{d}^T \bar{\mathbf{z}},$$

while the converse inequality immediately follows from the fact that $(\bar{x}, \bar{y}, \bar{\mathbf{z}})$ is an optimal solution of problem (3.11). As a result, we have

$$\sum_{\ell=1}^{L} p_\ell f(x^*, y^*, \omega_\ell) + \mathbf{d}^T \mathbf{z}^* = \sum_{\ell=1}^{L} p_\ell f(\bar{x}, \bar{y}, \omega_\ell) + \mathbf{d}^T \bar{\mathbf{z}},$$

namely, $(x^*, y^*, \mathbf{z}^*)$ is an optimal solution of problem (3.11) and hence it is a local optimal solution of problem (3.2). This completes the proof. $\square$

It is not difficult to see that the function

$$\psi(u) := \sum_{i=1}^{m} \Big(\max(u[i], 0) \Big)^{\sigma}, \tag{3.18}$$

where $\sigma \geq 2$ is a positive integer, satisfies the conditions assumed in Theorem 3.1. This function is often employed for solving constrained optimization problems. For more details, see [1].

Note that, in practice, it may not be easy to obtain an optimal solution, whereas computation of stationary points may be relatively easy. Therefore, it is necessary to study the limiting behavior of stationary points of subproblems (3.7).

Theorem 3.2 *Suppose the matrix M in (3.2) is a P_0-matrix, the function $\psi : \Re^m \to [0, +\infty)$ is given by (3.18) with $\sigma = 2$, and $(x^{(k)}, \mathbf{z}^{(k)})$ is a stationary point of (3.7) for each k. Let $(x^*, y^*, \mathbf{z}^*) \in \mathcal{F}$ be an accumulation point of the sequence $\{(x^{(k)}, y^{(k)}, \mathbf{z}^{(k)})\}$ generated by SIPP. If the MPEC-LICQ is satisfied at $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ in the MPEC (3.3), then $(x^*, y^*, \mathbf{z}^*)$ is a C-stationary point of problem (3.2). Furthermore, if y^* satisfies the strict complementarity condition, then $(x^*, y^*, \mathbf{z}^*)$ is S-stationary to (3.2).*

Although the results established in this theorem are interesting and important, its proof is somewhat lengthy and technical. To avoid disturbing the readability, we refer the readers to see [9] for a detailed proof of the theorem.

Remark 3.1 For the lower-level wait-and-see problems, we may consider a similar but somewhat simpler approach. In particular, for the discrete model

$$\begin{array}{ll} \text{minimize} & \sum_{\ell=1}^{L} p_\ell f(x, y_\ell) \\ \text{subject to} & g(x) \le 0,\ h(x) = 0, \\ & y_\ell \ge 0,\ N_\ell x + M_\ell y_\ell + q_\ell \ge 0, \\ & y_\ell^T (N_\ell x + M_\ell y_\ell + q_\ell) = 0,\ \ell = 1, \cdots, L \end{array} \tag{3.19}$$

with p_ℓ, N_ℓ, M_ℓ, and q_ℓ being the same as in (3.1), the subproblem corresponding to (3.7) becomes

$$\begin{array}{ll} \text{minimize} & \sum_{\ell=1}^{L} p_\ell f(x, y_\ell(x, \mu_k)) \\ \text{subject to} & g(x) \le 0,\ h(x) = 0, \end{array}$$

where $y_\ell(x, \mu)$ satisfies the equation $\Phi_\mu(x, y_\ell(x, \mu), w_\ell(x, \mu); N_\ell, M_\ell, q_\ell) = 0$ with $w_\ell(x, \mu) = N_\ell x + (M_\ell + \mu I) y_\ell(x, \mu) + q_\ell$ for each ℓ. Therefore, we do not need the penalty steps for problem (3.19). See [9] for more details.

4 Discretization of Here-and-Now Problems with Continuous Random Variable

In this section, we consider the here-and-now problem

$$\begin{array}{ll} \text{minimize} & E_\omega[f(x, y, \omega) + d^T z(\omega)] \\ \text{subject to} & g(x) \le 0,\ h(x) = 0, \\ & 0 \le y \perp (N(\omega)x + M(\omega)y + q(\omega) + z(\omega)) \ge 0, \\ & z(\omega) \ge 0, \quad \forall \omega \in \Omega, \\ & x \in \Re^n,\ y \in \Re^m,\ z(\cdot) \in \mathcal{C}(\Omega), \end{array} \tag{4.1}$$

where $\Omega := [a_1, b_1] \times \cdots \times [a_\nu, b_\nu] \subset \Re^\nu$, g, h, d are the same as in Section 3, the functions $f : \Re^{n+m} \times \Omega \to \Re$, $N : \Omega \to \Re^{m \times n}$, $M : \Omega \to \Re^{m \times m}$, and $q : \Omega \to \Re^m$ are all continuous. In addition, $\mathcal{C}(\Omega)$ denotes the family of continuous functions from Ω into $\Re^m$. Without loss of generality, we assume that $\Omega := [0, 1]^\nu$. Let $\zeta : \Omega \to [0, +\infty)$ be the continuous probability density function of ω. Then we have

$$E_\omega[f(x, y, \omega) + d^T z(\omega)] = \int_\Omega \Big(f(x, y, \omega) + d^T z(\omega) \Big) \zeta(\omega) d\omega.$$

We next employ a quasi-Monte Carlo method [13] for numerical integration to discretize problem (4.1). Roughly speaking, given a function $\phi : \Omega \to \Re$, the quasi-Monte Carlo estimate for $E_\omega[\phi(\omega)]$ is obtained by taking a uniformly distributed sample set $\Omega_L := \{\omega_1, \cdots, \omega_L\}$ from Ω and letting $E_\omega[\phi(\omega)] \approx \frac{1}{L} \sum_{\omega \in \Omega_L} \phi(\omega)$. Therefore, the following problem is an appropriate discrete approximation of problem (4.1):

$$\begin{array}{ll} \text{minimize} & \frac{1}{L} \sum_{\omega \in \Omega_L} \zeta(\omega) \Big(f(x, y, \omega) + d^T z(\omega) \Big) \\ \text{subject to} & g(x) \le 0,\ h(x) = 0, \\ & 0 \le y \perp (N(\omega)x + M(\omega)y + q(\omega) + z(\omega)) \ge 0, \\ & z(\omega) \ge 0,\ \omega \in \Omega_L. \end{array} \tag{4.2}$$

This problem has been discussed in the last section. Note that the sample set Ω_L is chosen to be asymptotically dense in Ω.

In order to prove our convergence result, we first give some lemmas.

Lemma 4.1 *Suppose the function $\xi : \Omega \to \Re$ is continuous. Then we have*

$$\lim_{L\to\infty} \frac{1}{L} \sum_{\omega\in\Omega_L} \xi(\omega)\zeta(\omega) = \int_\Omega \xi(\omega)\zeta(\omega)d\omega.$$

It is not difficult to prove this lemma by the results given in Chapter 2 of [13]. We then have from Lemma 4.1 immediately that, for any $z(\cdot) \in \mathcal{C}(\Omega)$,

$$\lim_{L\to\infty} \frac{1}{L} \sum_{\omega\in\Omega_L} \zeta(\omega)\Big(f(x,y,\omega) + d^T z(\omega)\Big) = \int_\Omega \Big(f(x,y,\omega) + d^T z(\omega)\Big)\zeta(\omega)d\omega \tag{4.3}$$

and particularly,

$$\lim_{L\to\infty} \frac{1}{L} \sum_{\omega\in\Omega_L} \zeta(\omega) = \int_\Omega \zeta(\omega)d\omega = 1. \tag{4.4}$$

Lemma 4.2 *Let $\bar{M} \in \Re^{m\times m}$ and $\{M_L\} \subset \Re^{m\times m}$ be convergent to $\bar{M}$. Suppose $\bar{M}$ is an R_0-matrix. Then, there exists an integer $L_0 > 0$ such that M_L is an R_0-matrix for every $L \geq L_0$.*

Proof. Suppose the conclusion is not true. Taking a subsequence if necessary, we may assume that $\{M_L\}$ is not an R_0-matrix for every L. From Definition 2.3, there exists a vector $y^L \in \Re^m$ such that

$$0 \leq y^L \perp M_L y^L \geq 0, \quad \|y^L\| = 1. \tag{4.5}$$

We may further assume that the sequence $\{y^L\}$ is convergent to a vector $\bar{y}$. Letting $L \to +\infty$ in (4.5), we get

$$0 \leq \bar{y} \perp \bar{M}\bar{y} \geq 0, \quad \|\bar{y}\| = 1.$$

This contradicts the fact that $\bar{M}$ is an R_0-matrix and hence the conclusion is valid. □

Theorem 4.1 *Let the set $X := \{x \in \Re^n \mid g(x) \leq 0, h(x) = 0\}$ be nonempty and bounded, and the function f be bounded and uniformly continuous with respect to (x, y, ω). Let*

$$\bar{M} := \int_\Omega M(\omega)\zeta(\omega)d\omega$$

be an R_0-matrix. Then, the following statements are true.

(i) *Problem (4.2) has at least one optimal solution when L is large enough.*

(ii) *Let $(x^L, y^L, z^L(\omega))_{\omega\in\Omega_L}$ be a solution of (4.2) for each L large enough. Then the sequence $\{(x^L, y^L)\}$ is bounded.*

(iii) *Let (x^*, y^*) be an accumulation point of the sequence $\{(x^L, y^L)\}$ and $z^*(\cdot)$ be defined by*

$$z^*(\omega) := \max\Big\{ -(N(\omega)x^* + M(\omega)y^* + q(\omega)),\ 0\Big\}, \quad \omega \in \Omega. \tag{4.6}$$

Then $(x^*, y^*, z^*(\cdot))$ is an optimal solution of problem (4.1).

Proof. (i) For each L, let $M_L := \frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)M(\omega)$. It then follows from Lemma 4.1 that $\bar{M} = \lim_{L\to\infty} M_L$. Since $\bar{M}$ is an R_0-matrix, by Lemma 4.2, there exists an integer $L_0 > 0$ such that M_L is an R_0-matrix for every $L \geq L_0$.

Let $L \geq L_0$ be fixed. We denote by $\mathcal{F}_L$ the feasible region of problem (4.2). It is easy to see that $\mathcal{F}_L$ is a nonempty and closed set and the objective function of problem (4.2) is bounded below on $\mathcal{F}_L$. Then, there exists a sequence $\{(x^k, y^k, z^k(\omega))_{\omega\in\Omega_L}\} \subseteq \mathcal{F}_L$ such that

$$\lim_{k\to\infty}\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\Big(f(x^k,y^k,\omega)+d^Tz^k(\omega)\Big) \quad = \inf_{(x,y,z(\omega))_{\omega\in\Omega_L}\in\mathcal{F}_L}\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\Big(f(x,y,\omega)+d^Tz(\omega)\Big). \tag{4.7}$$

Since the function f is bounded, it follows that the sequence $\left\{\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)d^Tz^k(\omega)\right\}$ is bounded. Note that the elements of d are positive. Thus, the sequence $\left\{\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)z^k(\omega)\right\}$ is also bounded. Moreover, we have from the boundedness of X that the sequence $\{x^k\}$ is bounded.

On the other hand, noting that $(x^k, y^k, z^k(\omega))_{\omega\in\Omega_L} \in \mathcal{F}_L$ for each k, we have

$$0 \leq y^k \perp \Big(\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)N(\omega)x^k + M_Ly^k + \frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)q(\omega) + \frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)z^k(\omega)\Big) \geq 0. \tag{4.8}$$

Suppose the sequence $\{y^k\}$ is unbounded. Taking a subsequence if necessary, we assume that

$$\lim_{k\to\infty}\|y^k\| = +\infty, \quad \lim_{k\to\infty}\frac{y^k}{\|y^k\|} = \bar{y}, \quad \|\bar{y}\| = 1. \tag{4.9}$$

Then, dividing (4.8) by $\|y^k\|$ and letting $k \to +\infty$, we obtain $0 \leq \bar{y} \perp M_L\,\bar{y} \geq 0$. Since M_L is an R_0-matrix, by Definition 2.3, we have $\bar{y} = 0$. This contradicts (4.9) and hence $\{y^k\}$ is bounded.

The boundedness of $\left\{\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)z^k(\omega)\right\}$ implies that the sequence $\{z^k(\omega)\}$ is bounded for each $\omega \in \Omega_L$ with $\zeta(\omega) > 0$. For any $\omega \in \Omega_L$ with $\zeta(\omega) = 0$, we re-define $z^k(\omega)$ by

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

In consequence, the sequence $\left\{(x^k, y^k, z^k(\omega))_{\omega\in\Omega_L}\right\}$ is bounded and (4.7) remains valid. Since $\mathcal{F}_L$ is closed, any accumulation point of $\left\{(x^k, y^k, z^k(\omega))_{\omega\in\Omega_L}\right\}$ must be an optimal solution of problem (4.2). This completes the proof of (i).

(ii) Let $(x^L, y^L, z^L(\omega))_{\omega\in\Omega_L}$ be a solution of (4.2) for each sufficiently large L. The boundedness of $\{x^L\}$ follows from the boundedness of the set X immediately. We next prove that $\{y^L\}$ is also bounded. To this end, we let $\bar{x} \in X$ and define

$$\bar{z}(\omega) := \max\Big\{-(N(\omega)\bar{x} + q(\omega)),\ 0\Big\}, \quad \omega \in \Omega.$$

Then, $(\bar{x}, 0, \bar{z}(\omega))_{\omega\in\Omega_L}$ is feasible to problem (4.2). Since $(x^L, y^L, z^L(\omega))_{\omega\in\Omega_L}$ is an optimal solution of (4.2), we have

$$\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\Big(f(x^L,y^L,\omega)+d^Tz^L(\omega)\Big) \leq \frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\Big(f(\bar{x},0,\omega)+d^T\bar{z}(\omega)\Big) \tag{4.10}$$

and

$$0 \leq y^L \ \perp\ \Big(N(\omega)x^L + M(\omega)y^L + q(\omega) + z^L(\omega)\Big) \geq 0, \qquad \omega \in \Omega_L. \tag{4.11}$$

It follows from (4.10) that

$$\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)d^Tz^L(\omega)\le\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\Big(f(\bar{x},0,\omega)-f(x^L,y^L,\omega)\Big)+\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)d^T\bar{z}(\omega).$$

Note that, from (4.4) and the boundedness of f, the sequence $\Big\{\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\Big(f(\bar{x},0,\omega)-f(x^L,y^L,\omega)\Big)\Big\}$ is bounded and, by Lemma 4.1,

$$\lim_{L\to\infty}\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)d^T\bar{z}(\omega)=\int_\Omega\zeta(\omega)d^T\bar{z}(\omega)d\omega.$$

In consequence, the sequence $\Big\{\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)d^Tz^L(\omega)\Big\}$ is bounded. Since the elements of d are positive, the sequence $\Big\{\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)z^L(\omega)\Big\}$ is bounded. Moreover, we have from (4.11) that

$$0\le y^L\perp\Big(\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)N(\omega)x^L+\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)M(\omega)y^L+\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)q(\omega)+\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)z^L(\omega)\Big)\ge 0.$$

Note that both $\{x^L\}$ and $\Big\{\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)z^L(\omega)\Big\}$ are bounded and

$$\begin{aligned}&\lim_{L\to\infty}\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)M(\omega)=\bar{M},\\&\lim_{L\to\infty}\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)N(\omega)=\int_\Omega N(\omega)\zeta(\omega)d\omega,\\&\lim_{L\to\infty}\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)q(\omega)=\int_\Omega q(\omega)\zeta(\omega)d\omega.\end{aligned}$$

In a similar way to (i), we can show that $\{y^L\}$ is bounded.

(iii) By the assumptions, the sequence $\{(x^L,y^L)\}$ contains a subsequence converging to (x^*,y^*). Without loss of generality, we suppose $\lim_{L\to\infty}(x^L,y^L)=(x^*,y^*)$.

(iiia) We first prove that $(x^*,y^*,z^*(\cdot))$ is feasible to problem (4.1). To this end, we define

$$\tilde{z}^L(\omega):=\max\Big\{-(N(\omega)x^L+M(\omega)y^L+q(\omega)),\ 0\Big\},\qquad\omega\in\Omega_L.\tag{4.12}$$

It is obvious that $(x^L,y^L,\tilde{z}^L(\omega))_{\omega\in\Omega_L}$ is feasible in problem (4.2) for each L. Since $z^*(\cdot)\in\mathcal{C}(\Omega)$ and $N(\omega)x^*+M(\omega)y^*+q(\omega)+z^*(\omega)\ge 0$ by the definition (4.6), it is sufficient to show that

$$(y^*)^T\Big(N(\omega)x^*+M(\omega)y^*+q(\omega)+z^*(\omega)\Big)=0,\qquad\omega\in\Omega.\tag{4.13}$$

Let $\bar{\omega}\in\Omega$ be fixed. Since the sample set Ω_L is chosen to be asymptotically dense in Ω, there exists a sequence $\{\omega_L\}$ of samples such that $\omega_L\in\Omega_L$ for each L and $\lim_{L\to\infty}\omega_L=\bar{\omega}$. We then have

$$(y^L)^T\Big(N(\omega_L)x^L+M(\omega_L)y^L+q(\omega_L)+\tilde{z}^L(\omega_L)\Big)=0,\qquad L=1,2,\cdots.$$

Letting $L\to+\infty$ and taking the continuity of the functions $N(\cdot),M(\cdot),q(\cdot)$ on the compact set Ω into account, we obtain

$$(y^*)^T\Big(N(\bar{\omega})x^*+M(\bar{\omega})y^*+q(\bar{\omega})+z^*(\bar{\omega})\Big)=0.\tag{4.14}$$

By the arbitrariness of $\bar{\omega}$ in Ω, we have (4.13) immediately. This completes the proof of the feasibility of $(x^*, y^*, z^*(\cdot))$ in (4.1).

(iiib) Let $(x, y, z(\cdot))$ be an arbitrary feasible solution of (4.1). It is obvious that $(x, y, z(\omega))_{\omega\in\Omega_L}$ is feasible to problem (4.2) for any L. Moreover, we have

$$\begin{aligned} &\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\Big(f(x^L, y^L, \omega) + d^T z^L(\omega)\Big) - \frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\Big(f(x^L, y^L, \omega) + d^T \tilde{z}^L(\omega)\Big) \\ = \;& \frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)d^T \min\Big\{N(\omega)x^L + M(\omega)y^L + q(\omega) + z^L(\omega),\ z^L(\omega)\Big\} \;\;\geq\;\; 0, \end{aligned}$$

where the equality follows from (4.12) and the inequality follows from the feasibility of $(x^L, y^L, z^L(\omega))_{\omega\in\Omega_L}$ in (4.2). Thus, $(x^L, y^L, \tilde{z}^L(\omega))_{\omega\in\Omega_L}$ is also an optimal solution of problem (4.2). We then have

$$\begin{aligned} &\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\Big(f(x^*, y^*, \omega) + d^T z^*(\omega)\Big) - \frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\Big(f(x, y, \omega) + d^T z(\omega)\Big) \\ \leq \;& \frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\Big(f(x^*, y^*, \omega) + d^T z^*(\omega)\Big) - \frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\Big(f(x^L, y^L, \omega) + d^T \tilde{z}^L(\omega)\Big) \\ \leq \;& \frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\bigg(\Big|f(x^*, y^*, \omega) - f(x^L, y^L, \omega)\Big| + \Big|d^T\Big(z^*(\omega) - \tilde{z}^L(\omega)\Big)\Big|\bigg). \end{aligned} \tag{4.15}$$

Note that f is uniformly continuous with respect to (x, y, ω) and, by (4.4), the sequence $\{\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\}$ is bounded. This yields

$$\lim_{L\to\infty}\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\Big|f(x^*, y^*, \omega) - f(x^L, y^L, \omega)\Big| = 0. \tag{4.16}$$

On the other hand, it is easy to see from the definitions (4.6) and (4.12) that

$$\Big|d^T\Big(z^*(\omega) - \tilde{z}^L(\omega)\Big)\Big| \leq \Big|d^T\Big(N(\omega)(x^* - x^L) + M(\omega)(y^* - y^L)\Big)\Big|, \qquad \omega\in\Omega_L.$$

By the boundedness of the sequence $\{\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\}$ and the functions $N(\cdot)$ and $M(\cdot)$ on Ω, we have

$$\lim_{L\to\infty}\frac{1}{L}\sum_{\omega\in\Omega_L}\zeta(\omega)\Big|d^T\Big(z^*(\omega) - \tilde{z}^L(\omega)\Big)\Big| = 0. \tag{4.17}$$

Thus, by letting $L\to+\infty$ in (4.15) and taking (4.3) and (4.16)–(4.17) into account, we obtain

$$\int_\Omega\Big(f(x^*, y^*, \omega) + d^T z^*(\omega)\Big)\zeta(\omega)d\omega \leq \int_\Omega\Big(f(x, y, \omega) + d^T z(\omega)\Big)\zeta(\omega)d\omega, \tag{4.18}$$

which implies that (x^*, y^*) together with $z^*(\cdot)$ constitutes an optimal solution of (4.1). □

5 Numerical Examples

The following example illustrates the here-and-now and lower-level wait-and-see models.

Example 5.1 [9] There are a food company who makes picnic lunches and several venders who sell lunches to hikers on every Sunday at different spots. The company and the venders have the following contract:

C1: The venders buy 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 ith vender decides the amount s_i of lunches that he buys from the company. Every vender must buy no less than the minimum amount $c > 0$.

C3: Every vender pays the company for the whole lunches he buys, i.e., the ith vender pays xs_i to the company.

C4: The ith vender sells lunches to hikers at the price $\kappa_i x$ and gets the proceeds for the total number of lunches actually sold, where $\kappa_i > 1$ is a constant.

C5: Even if there are any unsold lunches, the venders cannot return them to the company but they can dispose of the unsold lunches with no cost.

We suppose that the demands of lunches depend on the price and the weather on that day. Since the weather is uncertain, we may treat it as a random variable. Suppose there are m venders located at different spots. Assume the demand at the ith spot is given by the function $d_i(x, \omega)$. Then, the actual amount of lunches sold at the ith spot is given by $\min(s_i, d_i(x, \omega))$, which also depends on the weather on that day.

The decisions by the company and the ith vender are x and s_i, respectively. The company's objective is to maximize its total earnings $\sum_{i=1}^m xs_i$, while the ith vender's objective is to maximize its total earnings $\kappa_i x \min(s_i, d_i(x, \omega)) - xs_i$. We first consider the latter problem:

$$\begin{array}{ll} \text{maximize}_{s_i} & \kappa_i x \min(s_i, d_i(x, \omega)) - xs_i \\ \text{subject to} & s_i \geq c. \end{array}$$

It is not difficult to show that its solution is $s_i = \max\{d_i(x, \omega), c\}$, irrespective of the value of $\kappa_i > 1$. Therefore, by letting $y_i = s_i - c$ for each i, we may formulate the company's problem as the following stochastic MPEC:

$$\begin{array}{ll} \text{minimize} & -\sum_{i=1}^m x(y_i + c) \\ \text{subject to} & a \leq x \leq b, \\ & 0 \leq y_i \perp (-d_i(x, \omega) + y_i + c) \geq 0, \\ & i = 1, \ldots, m, \quad \text{a.e. } \omega \in \Omega. \end{array}$$

Now there are two cases.

Here-and-now model: Suppose that both the company and the venders have to make decisions on Saturday, without knowing the weather of Sunday. In this case, there may be no y_i satisfying the complementarity constraints for almost all $\omega \in \Omega$ in general. So, by introducing the recourse variables, the company's problem is represented as the following model:

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^m (-x(y_i + c) + \tau E_\omega[\, z_i(\omega)\,]) \\ \text{subject to} & a \leq x \leq b, \quad z_i(\omega) \geq 0, \\ & 0 \leq y_i \perp (-d_i(x, \omega) + y_i + c + z_i(\omega)) \geq 0, \\ & i = 1, \ldots, m, \quad \text{a.e. } \omega \in \Omega, \end{array}$$

where $\tau > 0$ is a weight constant.

Lower-level wait-and-see model: Suppose that the company makes a decision on Saturday, but the venders can make their decisions on Sunday morning after knowing the weather of that day. In this case, the decisions of the venders may depend on the observation of ω, which are given by $y_i(\omega), i = 1, \cdots, m$, that satisfies

$$0 \leq y_i(\omega) \perp (-d_i(x, \omega) + y_i(\omega) + c) \geq 0$$

for each $\omega \in \Omega$. Therefore the company's problem is represented as the following model:

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^m E_\omega[\,-x(y_i(\omega) + c)\,] \\ \text{subject to} & a \leq x \leq b, \\ & 0 \leq y_i(\omega) \perp (-d_i(x, \omega) + y_i(\omega) + c) \geq 0, \\ & i = 1, \ldots, m, \quad \text{a.e. } \omega \in \Omega. \end{array}$$

Below we report our numerical experience with these two models. We consider the case where $m = 4$ and assume that the weather parameter ω is normally distributed with $\mathcal{N}(0, 1)$ and the demand function for the ith vender is given by

$$d_i(x, \omega) := u_i(\omega) - v_i(\omega)x,$$

where $u_i(\omega)$ and $v_i(\omega)$ are random variables. Moreover, we assume that $u_i(\omega)$ and $v_i(\omega)$ are linear functions of ω such that

$$u_i(\omega) := u_{i0} + u_{i1}\omega, \qquad v_i(\omega) := v_{i0} + v_{i1}\omega$$

with constants $(u_{i0}, u_{i1}, v_{i0}, v_{i1}), i = 1, \cdots, 4$, given as in Table 1.

Table 1: Data for the Demand Functions

	u_{i0}	u_{i1}	v_{i0}	v_{i1}
$i = 1$	165	20	12	3
$i = 2$	218	13.5	18	2
$i = 3$	131	12	8	1.75
$i = 4$	195	13	9	2

In our implementation, we used the classical constructions method in [13] to approximate the continuous distributions by discrete ones.

- Generate $\omega_k, k = 1, \cdots, K$, from the 99% confidence interval $\mathcal{I} := [-3, 3]$ with sample size $K = 10^6$.
- Divide $\mathcal{I}$ into L subintervals with equal length, which represent different conditions of weather such as `bad`, `fair`, `good`, and so on.
- For each subinterval $\mathcal{I}_\ell$, estimate the probability by the relative frequency $p_\ell = k_\ell / K$, where k_ℓ is the number of samples contained in $\mathcal{I}_\ell$.
- For every subinterval $\mathcal{I}_\ell$ and every vender i, calculate

$$u_{i\ell} = \frac{1}{k_\ell} \sum_{\omega_k \in \mathcal{I}_\ell} u_i(\omega_k), \quad v_{i\ell} = \frac{1}{k_\ell} \sum_{\omega_k \in \mathcal{I}_\ell} v_i(\omega_k).$$

The data for the testing problem with $L = 3$ are listed in Table 2.

Table 2: Data for Testing Problem with $L = 3$

	p	u_1	v_1	u_2	v_2	u_3	v_3	u_4	v_4
good	0.1589	195.5213	16.5782	238.6019	21.0521	149.3128	10.6706	214.8388	12.0521
fair	0.6821	165.0092	12.0014	218.0062	18.0009	131.0055	8.0008	195.0060	9.0009
bad	0.1590	134.4962	7.4244	197.4100	14.9496	112.6977	5.3309	175.1726	5.9496

We set $a = 1, b = 14, c = 15, \tau = 1$, and employed the MATLAB 6.5 built-in solver *fmincon* to solve the subproblems. When we solved (3.7), we used the same penalty function ψ as in Theorem 3.2. Moreover, we set $\mu_0 = 10^{-2}$, $\rho_0 = 10^2$, and updated the parameters by $\mu_{k+1} = 10^{-2}\mu_k$ and $\rho_{k+1} = 10\rho_k$, respectively. In addition, the initial point is chosen to be $(x, \mathbf{z}) = (6, \cdots, 6)$ in the here-and-now problems and $x = 6$ in the lower-level wait-and-see problems, respectively, and the computed solution at the kth iteration is used as the starting point in the next $(k+1)$th iteration. The computational results for the here-and-now and lower-level wait-and-see cases with $L = 3$ are shown in Tables 3 and 4, respectively.

Table 3: Computational Results for Here-and-Now Case with $L = 3$

μ_k	ρ_k	$x^{(k)}$	$(y_1^{(k)}, y_2^{(k)}, y_3^{(k)}, y_4^{(k)})$	Obj	Res	Ite
10^{-2}	10^2	6.8603	(66.5927,78.9393,60.9143,116.7923)	2.6291e+003	3.2324	20
10^{-4}	10^3	6.8502	(67.9593,79.3917,61.1821,117.2787)	2.6360e+003	0.0397	17
10^{-6}	10^4	6.8497	(66.9665,79.4023,61.1834,117.2863)	2.6360e+003	0.0062	17
10^{-8}	10^5	6.8499	(66.9630,79.3981,61.1819,117.2837)	2.6360e+003	6.4550e-004	20

Table 4: Computational Results for Lower-Level Wait-and-See Case with $L = 3$

μ_k	$x^{(k)}$	$(\bar{y}_1^{(k)}, \bar{y}_2^{(k)}, \bar{y}_3^{(k)}, \bar{y}_4^{(k)})$	Obj	Res	Ite
10^{-2}	7.5489	(58.8243,66.4546,55.0579,110.9502)	2.6518e+003	8.7385	3
10^{-4}	7.5426	(59.4825,67.2262,55.6535,112.1053)	2.6736e+003	0.0883	2
10^{-6}	7.5426	(59.4891,67.2339,55.6595,112.1170)	2.6738e+003	8.8349e-004	2

In Tables 3 and 4, `Obj` means the company's earnings, `Ite` stands for the number of iterations spent by *fmincon* to solve the subproblems, and `Res` denotes the residual at the current point defined by

$$\mathrm{Res}(x^{(k)}, y^{(k)}, \mathbf{z}^{(k)}) := \sum_{i=1}^{m}\sum_{\ell=1}^{L}\left|\min\left(y_i^{(k)},\ -d_{i\ell}^{(k)} + y_i^{(k)} + c + z_{i\ell}^{(k)}\right)\right|$$

for the here-and-now case, or

$$\mathrm{Res}(x^{(k)}, \mathbf{y}^{(k)}) := \sum_{i=1}^{m}\sum_{\ell=1}^{L}\left|\min\left(y_{i\ell}^{(k)},\ -d_{i\ell}^{(k)} + y_{i\ell}^{(k)} + c\right)\right|$$

for the lower-level wait-and-see case, where $d_{i\ell}^{(k)} := u_{i\ell} - v_{i\ell}\, x^{(k)}$. Moreover, in Table 4, we denote $\bar{y}_i^{(k)} := \sum_{\ell=1}^{L} p_\ell\, y_{i\ell}^{(k)}$ for each i.

Table 5: Values of price x^L

	here-and-now	wait-and-see
$L = 3$	6.8499	7.5426
$L = 5$	6.7500	7.5454
$L = 7$	6.7500	7.5452
$L = 9$	6.7500	7.5453
$L = 11$	6.7500	7.5452

We have also computed the solutions of the two models with various values of L. Table 5 shows the values of the price x^L for $L = 3, 5, 7, 9, 11$. As shown in Table 5, the prices set by the company in the lower-level wait-and-see model are consistently higher than the ones in the here-and-now model. This seems reasonable because, in the lower-level wait-and-see case, the company has to take higher risk.

6 Concluding Remarks

We have presented a combined smoothing implicit programming and penalty method for an SMPEC with a finite sample space and suggested a quasi-Monte Carlo method to discretize an SMPEC with continuous random variables. We may extend the approaches to the lower-level wait-and-see problems. Recall that SMPECs contain the ordinary MPECs as a special subclass. In consequence, the conclusions given in Section 3 remain true for standard MPECs. Comparing with the results given in the literature, the assumptions employed in Section 3 are relatively weak.

References

[1] D.P. Bertsekas, Constrained Optimization and Lagrange Multiplier Methods, Academic Press, New York, 1982.

[2] S.I. Birbil, G. Gürkan, and O. Listes, *Simulation-based solution of stochastic mathematical programs with complementarity constraints: Sample-path analysis*, Preprint, 2004.

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

[4] X. Chen and M. Fukushima, *A smoothing method for a mathematical program with P-matrix linear complementarity constraints*, Computational Optimization and Applications, 27 (2004), 223–246.

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

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

[7] A. Fischer, *A special Newton-type optimization method*, Optimization, 24 (1992), 269–284.

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

[9] 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.

[10] 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.

[11] 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), 252–265.

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

[13] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, SIAM, Philadelphia, 1992.

[14] J.S. Pang, *Complementarity problems*, Handbook on Global Optimization, R. Horst and P. Pardalos (eds.), Kluwer Academic Publishers, B.V. Dordrecht, 1994, 271–338.

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

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

[17] 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.

[18] S. Vajda, Probabilistic Programming, Academic Press, New York, 1972.

[19] 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.