Technical Report 2009-008
Department of Applied Mathematics and Physics
Graduate School of Informatics
Kyoto University

Stochastic Equilibrium Problems and Stochastic Mathematical Programs with Equilibrium Constraints: A Survey[1]

Gui-Hua Lin[2] and Masao Fukushima[3]

March 30, 2009

Abstract. In the recent optimization research community, various equilibrium problems and related problems under uncertainty have drawn increasing attention. Novel formulations and numerical methods have been proposed to deal with those problems. This paper provides a brief review of the recent developments in the topics including stochastic variational inequality problems, stochastic complementarity problems and stochastic mathematical programs with equilibrium constraints.

Keywords. Stochastic variational inequality problem, stochastic complementarity problem, stochastic mathematical program with equilibrium constraints.

2000 Mathematics Subject Classification. 65K10, 90C30, 90C33, 90C15.

[1]The first author's work was supported in part by NSFC Grant #10771025 and SRFDP Grant #20070141063. The second author's work was supported in part by the Scientific Research Grant-in-Aid from Japan Society for the Promotion of Science.

[2]Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, China. E-mail: lin_g_h@yahoo.com.cn.

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

1 Introduction

Let a nonempty closed convex subset K of $\Re^n$ and a mapping $\mathbf{F} : \Re^n \to \Re^n$ be given. The classical variational inequality problem (VIP), denoted by VI($K, \mathbf{F}$), is to find a vector $x^* \in K$ such that

$$(x - x^*)^T \mathbf{F}(x^*) \geq 0, \qquad \forall x \in K.$$

When $K = \Re^n_+$, the variational inequality problem reduces to the following nonlinear complementarity problem (NCP), denoted by NCP($\mathbf{F}$): Find a vector x^* such that

$$x^* \geq 0, \quad \mathbf{F}(x^*) \geq 0, \quad (x^*)^T \mathbf{F}(x^*) = 0.$$

In particular, the problem is called the linear complementarity problem (LCP), denoted by LCP(M, q), if the function $\mathbf{F}$ is given by $\mathbf{F}(x) = Mx + q$ with $M \in \Re^{n \times n}$ and $q \in \Re^n$. NCP and LCP are generally called the complementarity problem (CP). The systematic study of VIP and CP began in the middle 1960s. As a result of more than four decades of research, the subject has developed into a well-established and fruitful discipline within mathematical programming. Several monographs [11, 13, 24, 25] and survey papers [19, 43] have documented the basic theory, effective algorithms, and important applications of VIP and CP in engineering, economics, and the optimization theory itself.

On the other hand, variational inequalities and complementarity conditions often appear within constraints of an optimization problem. This kind of problems is generally called a mathematical program with equilibrium constraints (MPEC), which can be stated as

$$\begin{array}{ll} \min & f(x,y) \\ \text{s.t.} & (x,y) \in Z, \\ & y \text{ solves VI}(C(x), F(x,\cdot)) \end{array} \tag{1.1}$$

or

$$\begin{array}{ll} \min & f(x,y) \\ \text{s.t.} & (x,y) \in Z, \\ & y \geq 0,\ F(x,y) \geq 0,\ y^T F(x,y) = 0. \end{array} \tag{1.2}$$

Here, Z is a nonempty subset of $\Re^{n+m}$ and $f : \Re^{n+m} \to \Re, F : \Re^{n+m} \to \Re^m$, $C : \Re^n \to 2^{\Re^m}$ are all mappings. Problem (1.2) is also called a mathematical program with complementarity constraints (MPCC). MPEC is generally regarded as a generalization of the bilevel programming problem and it plays a very important role in many fields such as engineering design, economic equilibrium, transportation science and game theory. However, MPEC may not be treated as an ordinary nonlinear programming problem because its constraints fail to satisfy the standard constraint qualifications including the linear independence constraint qualification (LICQ) and the Mangasarian-Fromovitz constraint qualification (MFCQ) at any feasible point [10]. This means that the well-developed nonlinear programming theory cannot be applied to MPEC directly. See the monographs [39, 42] and the survey paper [17] for more details about the MPEC theory, algorithms, and applications.

Since some elements may involve uncertain data in many practical problems, the stochastic versions of the above problems have drawn much attention in the recent literature. The purpose of this paper is to review some recent developments in these topics, which include stochastic variational inequality problems (SVIP), stochastic complementarity problems (SCP), or more specifically, stochastic linear and nonlinear complementarity problems (SLCP, SNCP), and stochastic mathematical programs with equilibrium constraints (SMPEC).

2 Examples

In this section, we give some examples to show how uncertainty arises in practice. The first example is considered in [9].

Example 2.1. Consider a transportation network with a set of nodes and a set of paths between origin-destination (OD) pairs. Denote

ξ: vector of path flows;

τ: vector of minimum transportation costs between OD pairs;

d: vector of travel demand between OD pairs;

$A\xi + b$: users' travel cost function with A being a positive definite matrix;

B: path-OD pair incidence matrix.

By the Wardrop's user equilibrium principle, the traffic equilibrium conditions can be formulated as an LCP

$$\begin{aligned} &\xi \geq 0, \quad A\xi + b - B^T\tau \geq 0, \quad \xi^T(A\xi + b - B^T\tau) = 0, \\ &\tau \geq 0, \qquad B\xi - d \geq 0, \qquad \tau^T(B\xi - d) = 0. \end{aligned}$$

In practice, the travel demand is not always constant, but may vary depending on the weather etc. The travel cost function may also vary because of some unpredictable reasons. In such uncertain situations, we may regard the demand and the cost coefficients as random variables $d(\omega)$ and $(A(\omega), b(\omega))$ with $\omega \in \Omega$, where Ω is a sample space. Thus the traffic equilibrium conditions can be modeled as an SLCP

$$\begin{aligned}
&\xi \geq 0, \ A(\omega)\xi + b(\omega) - B^T\tau \geq 0, \ \xi^T(A(\omega)\xi + b(\omega) - B^T\tau) = 0, \\
&\tau \geq 0, \qquad\quad B\xi - d(\omega) \geq 0, \qquad\quad \tau^T(B\xi - d(\omega)) = 0.
\end{aligned}$$

Note that there is in general no vector (ξ, τ) satisfying the above conditions for all ω in Ω simultaneously. Nevertheless one may want to estimate a traffic flow ξ along with the corresponding travel cost τ that is most likely to occur on the whole. How to construct a model that produces reasonable solutions of such problems is the main topic of Section 3.

We next give an example from [30], which is used to illustrate models of SMPEC.

Example 2.2. Consider a food company that wholesales picnic lunches to m venders who sell lunches at different spots. Denote

$x \in [a, b]$:	selling price of the company with $b > a > 0$;
$\kappa_i x$:	selling price of the ith vender with $\kappa_i > 1$;
$s_i \in [c, +\infty)$:	amount of lunches booked by the ith vender with $c > 0$;
d_i:	demand at the ith spot.

Here a, b, c and κ_i's are given constants. Suppose that, even if there are unsold lunches, the venders cannot return them to the company but they can dispose of the unsold lunches with no cost. In general, the demands of lunches depend on the price and some uncertainty, say the weather of the day. We treat the demands as random variables and denote by $d_i(x, \omega)$ the demand at the ith spot. Since the ith vender's objective is to maximize its total earnings $\kappa_i x \min(s_i, d_i(x, \omega)) - xs_i$, it is not difficult to show that his decision is given by $s_i = \max\{d_i(x, \omega), c\}$. Thus, by letting $y_i = s_i - c$ for each i, we can formulate the company's problem as an SMPEC

$$\begin{array}{ll}
\min & -\sum_{i=1}^{m} x(y_i + c) \\
\text{s.t.} & a \leq x \leq b, \\
& y_i \geq 0, \ -d_i(x, \omega) + y_i + c \geq 0, \\
& y_i(-d_i(x, \omega) + y_i + c) = 0, \quad i = 1, \ldots, m.
\end{array}$$

This problem can be viewed as a bilevel, or leader-follower, problem with an upper level decision x by the company (leader) and lower level decisions $y_1, \ldots, y_m$ by the venders (followers).

Now suppose that the company and the vendors have to decide their prices x and $y_1, \ldots, y_m$ of lunches for sale on Sunday. Note that there are two cases concerning the times when the decisions of the company and the venders are made; (i) both the company and the venders make decisions on Saturday, without knowing the weather of Sunday, and (ii) the company makes a decision on Saturday, while the venders make decisions on Sunday, after knowing the actual weather of that day. These two cases lead to the models called 'here-and-now' and 'lower-level wait-and-see', respectively. More details of these models will be discussed in Section 4.

3 Stochastic Equilibrium Problems

Let $K \subseteq \Re^n$ be a nonempty closed convex set, $(\Omega, \mathcal{F}, \mathcal{P})$ be the underlying probability space, and $F : K \times \Omega \to \Re^n$ be a given mapping. The SVIP is to find $x^* \in K$ such that

$$\mathcal{P}\{\omega \in \Omega \mid (x - x^*)^T F(x^*, \omega) \geq 0,\ \forall\, x \in K\} = 1$$

or equivalently

$$(x - x^*)^T F(x^*, \omega) \geq 0, \qquad \forall\, x \in K, \quad \omega \in \Omega \text{ a.s.}, \tag{3.1}$$

where "a.s." is the abbreviation for "almost surely" under the given probability measure. Similarly, the SCP is to find a vector x^* such that

$$\mathcal{P}\{\omega \in \Omega \mid x^* \geq 0,\ F(x^*, \omega) \geq 0,\ (x^*)^T F(x^*, \omega) = 0\} = 1,$$

which is equivalent to

$$x^* \geq 0,\ F(x^*, \omega) \geq 0,\ (x^*)^T F(x^*, \omega) = 0, \qquad \omega \in \Omega \text{ a.s.} \tag{3.2}$$

Because of the existence of a random element ω, however, we cannot generally expect that there exists a vector x^* satisfying (3.1) or (3.2) for almost all $\omega \in \Omega$. That is, both (3.1) and (3.2) may not have a solution in general. Therefore, an important issue in the study of SCP and SVIP is to present an appropriate deterministic formulation of the considered problem. In what follows, we review some existing formulations and algorithms for SCP and SVIP.

3.1 Stochastic CP

In the study of SCP, there have been proposed three types of formulations; the expected value (EV) formulation, the expected residual minimization (ERM) formulation, and the SMPEC formulation.

3.1.1 EV formulation. This model is studied by Gürkan et al. [18]. The problem considered in [18] is actually an SVIP, which can be traced back to King and Rockafellar [26]. When applied to the SCP (3.2), the EV model can be stated as follows:

$$x^* \geq 0,\ \mathbb{E}[F(x^*, \omega)] \geq 0,\ (x^*)^T \mathbb{E}[F(x^*, \omega)] = 0, \tag{3.3}$$

where $\mathbb{E}$ means expectation with respect to ω. Since the expectation function $\mathbb{E}[F(\cdot,\omega)]$ is usually difficult to evaluate exactly, it is assumed that a sequence $\{F^k\}$ of deterministic functions, converging to the function $\mathbb{E}[F(\cdot,\omega)]$ in a certain sense, can be observed. Then a solution of problem (3.3) may be obtained by solving a sequence of deterministic complementarity problems. It is shown that, under some regularity conditions on F and some (unknown) solution x^* of (3.3), the approximation problem NCP(F^k) has a solution close to x^* if F^k is sufficiently close to $\mathbb{E}[F(\cdot,\omega)]$.

Since many parameter estimation problems can be transformed to a root-seeking problem for an unknown function, stochastic approximation methods have been extensively studied and applied to solving various stochastic problems; see for instance [7, 21, 45] and the references therein. Jiang and Xu [20] propose several stochastic approximation approaches for SVIP and SCP. Among others, a stochastic approximation method for (3.3) is presented by utilizing the well known approach in NCP that a deterministic NCP can be transformed into a nonlinear equation by using the so-called NCP functions. As usual, a function $\phi : \Re^2 \to \Re$ is called an NCP function if it satisfies

$$\phi(a,b)=0 \quad \Longleftrightarrow \quad a\geq 0,\ b\geq 0,\ ab=0.$$

See [13] for more details about NCP functions. A popular NCP function is the Fischer-Burmeister function $\phi_{\rm FB} : \Re^2 \to \Re$ defined by

$$\phi_{\rm FB}(a,b) := a+b-\sqrt{a^2+b^2}. \tag{3.4}$$

Using the function $\phi_{\rm FB}$, (3.3) can be transformed equivalently into a system of equations

$$\Phi_{\rm FB}(x) = 0, \tag{3.5}$$

where the function $\Phi_{\rm FB} : \Re^n \to \Re^n$ is defined by

$$\Phi_{\rm FB}(x) := \begin{pmatrix} \phi_{\rm FB}(x_1, \mathbb{E}[F_1(x,\omega)]) \\ \vdots \\ \phi_{\rm FB}(x_n, \mathbb{E}[F_n(x,\omega)]) \end{pmatrix}.$$

The equation (3.5) is then reformulated as the optimization problem

$$\min_{x\in\Re^n} \ \frac{1}{2}\|\Phi_{\rm FB}(x)\|^2$$

and the following stochastic approximation method is proposed:

$$x^{k+1} = x^k - \tau_k \zeta^k, \qquad k=0,1,2,\ldots, \tag{3.6}$$

where $\tau_k > 0$ is a pre-determined step-length and ζ^k is a stochastic approximation of the vector

$$d^k := \begin{pmatrix} \phi_{\rm FB}(x_1^k, \mathbb{E}[F_1(x^k,\omega)])\, \partial_b \phi_{\rm FB}(x_1^k, \mathbb{E}[F_1(x^k,\omega)]) \\ \vdots \\ \phi_{\rm FB}(x_n^k, \mathbb{E}[F_n(x^k,\omega)])\, \partial_b \phi_{\rm FB}(x_n^k, \mathbb{E}[F_n(x^k,\omega)]) \end{pmatrix}.$$

Here $\partial_b \phi_{\rm FB}(a,b)$ denotes a subgradient of the function $\phi_{\rm FB}$ with respect to the second argument. It is shown in [20] that, under suitable assumptions, the sequence $\{x^k\}$ generated by (3.6) is almost surely convergent to a solution of (3.3) for any initial point $x^0 \in \Re^n$.

3.1.2 ERM formulation. This model is presented by Chen and Fukushima [8] for the SLCP

$$x \ge 0,\ \ M(\omega)x + q(\omega) \ge 0,\ \ x^T(M(\omega)x + q(\omega)) = 0, \quad \omega \in \Omega \text{ a.s.}, \tag{3.7}$$

where $M : \Omega \to \Re^{n\times n}$ and $q : \Omega \to \Re^n$. By employing an NCP function ϕ, the SLCP (3.7) is transformed equivalently into the stochastic equations

$$\Phi(x,\omega) = 0, \quad \omega \in \Omega \text{ a.s.},$$

where $\Phi : \Re^n \times \Omega \to \Re^n$ is defined by

$$\Phi(x,\omega) := \begin{pmatrix} \phi(x_1, [M(\omega)x + q(\omega)]_1) \\ \vdots \\ \phi(x_n, [M(\omega)x + q(\omega)]_n) \end{pmatrix}.$$

Then the following deterministic optimization formulation for (3.7) is proposed:

$$\min_{x\in\Re^n_+} \ \theta(x) := \mathbb{E}[\|\Phi(x,\omega)\|^2]. \tag{3.8}$$

Problem (3.8) is called an ERM formulation for the SLCP (3.7). The NCP functions employed in [8] include the Fischer-Burmeister function $\phi_{\rm FB}$ defined by (3.4) and the "min" function

$$\phi_{\min}(a,b) := \min(a,b). \tag{3.9}$$

As observed in [8], the ERM formulations with different NCP functions may have different properties. Various existence theorems and quasi-Monte Carlo methods are given in [8].

The ERM formulation for SLCP has subsequently been studied in [14, 9, 60, 58]. Fang et al. [14] introduce a new concept of stochastic matrices: $M(\cdot)$ is called a stochastic R_0 matrix if

$$\mathcal{P}\{\omega \in \Omega \mid x \ge 0, M(\omega)x \ge 0, x^T M(\omega)x = 0\} = 1 \quad \Longrightarrow \quad x = 0.$$

This is an extension of an R_0 matrix that plays an important role in the LCP theory [11]. It is shown that $M(\cdot)$ being a stochastic R_0 matrix is a necessary and sufficient condition for the solution set of the ERM model (3.8) being nonempty and bounded for any random vector $q(\cdot)$.

Furthermore, a number of necessary and/or sufficient conditions for stochastic R_0 matrices are given. Local and global error bounds for (3.8) are also studied in [14].

Chen et al. [9] consider the SLCP (3.7) in which the expectation matrix $\bar{M} := \mathbb{E}[M(\omega)]$ is positive semi-definite. Such an SLCP is called a monotone SLCP. The ERM model (3.8) studied in [9] uses the min function $\phi_{\min}$ defined above and the penalized Fischer-Burmeister function [4] defined by

$$\phi_{\mathrm{PFB}}(a,b) := \lambda(a+b-\sqrt{a^2+b^2}) + (1-\lambda)a_+b_+,$$

where $\lambda \in (0,1)$ is a given scalar and z_+ denotes $\max(z,0)$. Firstly, for the ERM formulation (3.8), some results on boundedness of the level sets

$$\mathcal{L}(c) := \{x \mid \theta(x) \leq c\}$$

are obtained for the above-mentioned two NCP functions. Then, a regularization method for the monotone SLCP is suggested and its convergence is also discussed. Moreover, an error bound property of the ERM formulation for the original SLCP is investigated. Specifically, when the sample space Ω is a finite set, if $M(\omega)$ is a positive semi-definite matrix and $\mathrm{LCP}(M(\omega), q(\omega))$ has a nonempty solution set $\mathrm{SOL}(M(\omega), q(\omega))$ for each $\omega \in \Omega$, then there exists a constant $\beta > 0$ such that

$$\mathbb{E}[\mathrm{dist}(x, \mathrm{SOL}(M(\omega), q(\omega)))] \;\leq\; \beta\sqrt{\theta_{\mathrm{PFB}}(x)}, \qquad \forall x \in \Re^n_+,$$

where θ_{PFB} is the objective function of (3.8) defined by means of the penalized Fischer-Burmeister function ϕ_{PFB} and $\mathrm{dist}(x, \mathrm{SOL}(M(\omega), q(\omega)))$ denotes the distance from a point x to the set $\mathrm{SOL}(M(\omega), q(\omega))$. This result particularly shows that a solution x^* of the ERM formulation (3.8) satisfies the inequality

$$\mathbb{E}[\mathrm{dist}(x^*, \mathrm{SOL}(M(\omega), q(\omega)))] \;\leq\; \beta\sqrt{\theta_{\mathrm{PFB}}(x^*)}.$$

Unlike an error bound for the deterministic LCP, the left-hand side of this inequality does not vanish in general. Nevertheless, the inequality suggests that the expected distance to the solution set $\mathrm{SOL}(M(\omega), q(\omega))$ of $\mathrm{LCP}(M(\omega), q(\omega))$ for $\omega \in \Omega$ is also likely to be small at x^*. In other words, one may expect that a solution of the ERM formulation (3.8) has a minimum sensitivity with respect to random parameter variation in SLCP. In this sense, solutions of (3.8) can be regarded as robust solutions for the SLCP (3.7).

Zhou and Cacceta [60] consider the SLCP (3.7) in which the sample space Ω has only finite realizations $\{\omega_1, \omega_2, \ldots, \omega_L\}$. By introducing slack variables $y = (y^1, \ldots, y^L) \in \Re^{nL}$ and making use of the penalized Fischer-Burmeister function ϕ_{PFB}, the SLCP is reformulated as the following nonsmooth equations with nonnegative constraints:

$$\hat{\Phi}(x, y) = 0, \quad y \geq 0, \tag{3.10}$$

where

$$\hat{\Phi}(x, y) := \begin{pmatrix} \phi_{\text{PFB}}(x_1, [\bar{M}x + \bar{q}]_1) \\ \vdots \\ \phi_{\text{PFB}}(x_n, [\bar{M}x + \bar{q}]_n) \\ M(\omega_1)x + q(\omega_1) - y^1 \\ \vdots \\ M(\omega_L)x + q(\omega_L) - y^L \end{pmatrix} \in \Re^{n+nL}$$

and $\bar{M} := \mathbb{E}[M(\omega)]$, $\bar{q} := \mathbb{E}[q(\omega)]$. The system (3.10) has $n(1 + L)$ equations with $n(1 + L)$ unknowns and the function $\hat{\Phi}$ is shown to be strongly semismooth; see [46] for details about semismoothness. Since (3.10) may have no solution, it is further transformed into the following optimization problem:

$$\min_{(x,y) \in \Re_+^{n+nL}} \frac{1}{2} \|\hat{\Phi}(x, y)\|^2. \tag{3.11}$$

Then, a feasible semismooth Newton method is proposed for solving problem (3.11) and its convergence properties are investigated. Results of numerical experiments are also reported in [60].

Zhang and Chen [58] consider the SLCP (3.7) and its ERM formulation (3.8) with the function $\phi_{\min}$ defined by (3.9). Let $\rho : \Re \to [0, +\infty)$ be a piecewise continuous density function satisfying $\phi(s) = \rho(-s)$ and $\int_{-\infty}^{\infty} |s| \, \rho(s) ds < +\infty$. Then the following function serves as a smooth approximation to $\phi_{\min}$ [6]:

$$\phi_{\min}^{\mu}(a, b) := a - \int_{-\infty}^{\infty} (a - b - \mu s)_+ \, \rho(s) ds,$$

where $\mu > 0$ is a smoothing parameter. By using this function, a smooth approximation of the objective function in problem (3.8) is constructed as

$$\theta_{\min}^{\mu}(x) := \sum_{i=1}^{n} \mathbb{E}[{\phi_{\min}^{\mu}}^2(x_i, [M(\omega)x + q(\omega)]_i)].$$

In [58], differentiability properties of the function $\theta_{\min}^{\mu}$ are investigated. Moreover, a smoothing projected gradient method with the above smoothing technique is proposed and its convergence analysis is given in [58].

The ERM formulation for the SNCP (3.2) is discussed in [29, 59, 35]. Based on the observation that the ERM formulation (3.8) retains the nonnegative constraints $x \geq 0$, Lin et al. [29]

use a restricted NCP function $\psi : \Re^2 \to \Re$, which is nonnegative-valued everywhere and satisfies the property

$$\psi(a,b) = 0,\ a \geq 0 \quad \Longleftrightarrow \quad a \geq 0,\ b \geq 0,\ ab = 0, \tag{3.12}$$

to formulate the SNCP (3.2) as the following ERM problem:

$$\min_{x \in \Re^n_+} \ \theta_{\mathrm{R}}(x) := \sum_{i=1}^{n} \mathbb{E}[\psi(x_i, F_i(x,\omega))]. \tag{3.13}$$

The restricted NCP functions are studied in [56] for deterministic NCPs. In [29], the following three new restricted NCP functions are mainly considered:

(i) $\psi_1(a,b) := (ab)_+ + (-b)_+$, which can be rewritten as $\psi_1(a,b) = \max(ab, -b)$ when $a \geq 0$;

(ii) $\psi_2(a,b) := (ab)_+^2 + (-b)_+^2$, which is a smoothed modification of ψ_1;

(iii) $\psi_3(a,b) := a^2b^2 + (-b)_+^2$, which may also be regarded as a smoothed modification of ψ_1.

It is shown in [29] that the level sets

$$\mathcal{L}^+(c) := \{x \in \Re^n_+ \mid \theta_{\mathrm{R}}(x) \leq c\}$$

defined by the above restricted NCP functions are bounded for any nonnegative scalar c, if one of the following conditions holds: (a) the expectation function $\bar{F}(x) := \mathbb{E}[F(x,\omega)]$ is monotone and the problem NCP($\bar{F}$) has a nonempty bounded solution set; (b) the function $\bar{F}$ is an R_0 function on $\Re^n_+$ in the sense of [5]. Error bound conditions are also investigated. Some counterexamples are constructed to show that the new restricted NCP functions indeed have some favorable properties that the NCP functions $\phi_{\min}$ and ϕ_{FB} do not have in dealing with SNCP.

Zhang and Chen [59] introduce a new concept of stochastic R_0 function, which can be regarded as a generalization of the deterministic R_0 function given in [5] and the stochastic R_0 matrix given in [14]. It is shown that, under suitable assumptions, the objective function of the ERM problem (3.13) with ψ replaced by either $\phi_{\min}$ or ϕ_{FB} is coercive if and only if the function F is a stochastic R_0 function. Furthermore, a traffic equilibrium problem under uncertainty is modeled as an SNCP and the objective function in the corresponding ERM formulation is shown to be a stochastic R_0 function. Numerical results reported in [59] indicate that the ERM approach has various desirable properties.

Ling et al. [35] mainly discuss properties of the objective function θ_{R} of problem (3.13) in which ψ is replaced by the Fischer-Burmeister function ϕ_{FB}. Their main result is that, under some mild conditions, θ_{R} possesses the so-called SC^1 property, that is, θ_{R} is continuously differentiable and its gradient is semismooth.

3.1.3 SMPEC formulation. Lin and Fukushima [33] study the SNCP (3.2) from another point of view. Recall that there may not exist a vector x satisfying the complementarity conditions for almost all $\omega \in \Omega$. In order to get a reasonable resolution, recourse variables $z(\omega) \geq 0$ are introduced for each inequality $F(x, \omega) \geq 0$ and the total recourse is minimized. Thus, one obtains the following formulation for (3.2):

$$\begin{aligned} \min \quad & \mathbb{E}[d^T z(\omega)] \\ \text{s.t.} \quad & x \geq 0, \;\; F(x,\omega) + z(\omega) \geq 0, \\ & x^T(F(x,\omega) + z(\omega)) = 0, \\ & z(\omega) \geq 0, \quad \omega \in \Omega \text{ a.s.}, \end{aligned} \tag{3.14}$$

where d is a weight vector with positive components. Problem (3.14) is actually a special here-and-now model of SMPEC, which will formally be introduced in Section 4.

It is well known that SMPECs are very difficult to deal with. In order to develop effective methods for solving (3.14), a new function $Q : \Re^n \times \Omega \to [0, +\infty]$ is defined in [33] by

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

It is easy to see from the duality theorem in linear programming that, for any fixed $x \in \Re^n$ and $\omega \in \Omega$, $Q(x, \omega)$ is finite if and only if the set

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

is nonempty and, if $\mathcal{Z}(x, \omega)$ is nonempty, one can actually write

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

It can be shown that, for any $x \in \Re^n_+$ and $\omega \in \Omega$, $Q(x, \omega) < +\infty$ if and only if $x_i F_i(x, \omega) \leq 0$ for all $i = 1, \ldots, n$ and, when $Q(x, \omega) < +\infty$, one has

$$Q(x,\omega) = d^T(-F(x,\omega))_+,$$

where $(-F(x, \omega))_+$ denotes the vector with components $(-F_i(x, \omega))_+$, $i = 1, \ldots, m$. Based on these results, it is not difficult to show the equivalence between the SMPEC formulation (3.14) and the optimization problem

$$\min_{x \geq 0} \;\; \mathbb{E}[Q(x,\omega)]$$

or

$$\begin{aligned} \min \quad & \mathbb{E}[d^T(-F(x,\omega))_+] \\ \text{s.t.} \quad & x \geq 0, \\ & x \circ F(x,\omega) \leq 0, \quad \omega \in \Omega \text{ a.s.}, \end{aligned} \tag{3.15}$$

where $\circ$ denotes the Hadamard product, i.e., $x \circ F(x, \omega) = (x_1 F_1(x, \omega), \ldots, x_n F_n(x, \omega))^T$. For the case where Ω is a finite set, a smoothed penalty method based on (3.15) is presented in [33].

Subsequently, Lin [27] proposes a new SMPEC formulation for the SNCP (3.2):

$$
\begin{aligned}
\min \quad & \mathbb{E}[\|z(\omega)\|^2] \\
\text{s.t.} \quad & x \geq 0,\ F(x, \omega) + z(\omega) \geq 0, \\
& x^T(F(x, \omega) + z(\omega)) = 0, \\
& z(\omega) \geq 0, \quad \omega \in \Omega \text{ a.s.}
\end{aligned}
\tag{3.16}
$$

Since both $d^T z(\omega)$ and $\|z(\omega)\|^2$ serve as a penalty term for the possible violation of the complementarity constraints, the objectives of problems (3.14) and (3.16) are essentially the same. However, the quadratic penalty $\|z(\omega)\|^2$ in (3.16) yields the equivalent problem

$$
\begin{aligned}
\min \quad & \mathbb{E}[\|(-F(x, \omega))_+\|^2] \\
\text{s.t.} \quad & x \geq 0, \\
& x \circ F(x, \omega) \leq 0, \quad \omega \in \Omega \ \text{ a.s.},
\end{aligned}
\tag{3.17}
$$

which has a differentiable objective function, while the linear penalty $d^T z(\omega)$ yields the equivalent problem (3.15) whose objective function is not differentiable everywhere. Note that both the recourse variables and the complementarity constraints vanish in problem (3.17). However, (3.17) is actually a semi-infinite programming problem with a large number of complementarity-like constraints. Moreover, it involves an expectation in the objective function. Therefore, problem (3.17) is generally more difficult to handle than an ordinary semi-infinite programming problem. In [27], the case of a compact sample space Ω is considered first. For this case, the Monte Carlo methods [47, Chapter 6] are employed to approximate the expectation function involved in (3.17) and a penalty technique is used to deal with the complementarity-like constraints. Furthermore, for the case where the sample space is unbounded, a compact approximation approach is suggested.

Besides the above three formulations, one may use stochastic programming techniques to study SNCP from other points of view. For instance, Wang et al. [51] treat the SMPEC formulation (3.14) as a two-stage stochastic program

$$
\min_{x \geq 0} \ \mathbb{E}[Q(x, \omega)],
$$

where $Q(x, \omega)$ is defined by

$$
Q(x, \omega) := \inf \{ d^T z(\omega) \,|\, z(\omega) \geq 0,\ F(x, \omega) + z(\omega) \geq 0,\ x^T(F(x, \omega) + z(\omega)) = 0 \},
$$

and propose penalty-based sampling approximation methods to solve it. Moreover, applications of SNCP in supply chain network equilibria are discussed.

3.2 Stochastic VIP

There have been presented two types of formulations for SVIP; the EV formulation and the ERM formulation.

3.2.1 EV formulation. Suppose that the expectation function $\mathbb{E}[F(x,\omega)]$ is well-defined. Then the EV formulation for SVIP (3.1) is defined as

$$(x - x^*)^T \mathbb{E}[F(x^*,\omega)] \geq 0, \qquad \forall x \in K. \tag{3.18}$$

A class of problems that contains problem (3.18) as a special case is the EV formulation of stochastic generalized equations, which is to find a vector x^* such that

$$0 \in \mathbb{E}[F(x^*,\omega)] + \mathcal{N}(x^*), \tag{3.19}$$

where $\mathcal{N} : \Re^n \to 2^{\Re^n}$ is a set-valued function. When $\mathcal{N}$ is the normal cone operator associated with the closed convex set K, the generalized equation (3.19) boils down to the EV formutaion (3.18) of SVIP.

As mentioned in Subsection 3.1, the EV formulation for SVIP may be attributed to King and Rockafellar [26] and Gürkan et al. [18]. In particular, for the EV formutaion (3.18), Gürkan et al. [18] consider a sequence $\{F^k\}$ of deterministic functions converging to the function $\mathbb{E}[F(\cdot,\omega)]$ in a certain sense, and show that a solution of problem (3.18) can be obtained by solving the sequence of variational inequality problems VI(K, F^k). King and Rockafellar [26] consider the generalized equation (3.19) and study the asymptotic behavior of a sequence of solutions to generalized equations (3.19) in which $\mathbb{E}[F(x^*,\omega)]$ is replaced by some F^k. See also [47, Chapter 6] for stochastic generalized equations.

Recently, Jiang and Xu [20] consider two types of the stochastic approximation methods for solving the EV formulation (3.18). One is based on the following well-known result in the deterministic VIP theory: x^* solves (3.18) if and only if

$$x^* = \text{Proj}_{K,G}(x^* - \tau G^{-1}\mathbb{E}[F(x^*,\omega)]),$$

where G is an $n \times n$ symmetric positive definite matrix, $\tau > 0$ is a constant, and $\text{Proj}_{K,G}$ denotes the projection operator onto K under G-norm. The corresponding iterative scheme is given by

$$x^{k+1} = \text{Proj}_{K,G}(x^k - \tau_k G^{-1}\zeta^k), \qquad k = 0, 1, 2, \ldots, \tag{3.20}$$

where τ_k is a stepsize and ζ^k is a stochastic approximation of $\mathbb{E}[F(x^k,\omega)]$. The formula (3.20) can be regarded as a Robbins-Monro type stochastic approximation method for solving (3.18). For the case where $\mathbb{E}[F(\cdot,\omega)]$ is strongly monotone on K, it is shown that, under some standard conditions, the sequence $\{x^k\}$ generated by (3.20) is convergent to the unique solution of (3.18) with probability one for any initial point $x^0 \in \Re^n$.

The second approach considered in [20] is based on reformulation techniques with merit functions [16] that have been used as popular tools for dealing with deterministic VIPs. Specifically, in [20], the regularized gap function

$$g_\alpha(x) := \max_{y \in K} \left\{ (x-y)^T \mathbb{E}[F(x,\omega)] - \frac{\alpha}{2}\|x-y\|^2 \right\} \tag{3.21}$$

introduced in [15] is used to transform (3.18) into a constrained optimization problem

$$\min_{x \in K} \; g_\alpha(x), \tag{3.22}$$

and the D-gap function

$$g_{\alpha\beta}(x) := g_\alpha(x) - g_\beta(x)$$

introduced in [57] is employed to transform (3.18) into an unconstrained optimization problem

$$\min_{x \in \Re^n} \; g_{\alpha\beta}(x), \tag{3.23}$$

where the parameters are chosen to satisfy $\beta > \alpha > 0$. Both the regularized gap function and the D-gap function are proved to be continuously differentiable if $\mathbb{E}[F(\cdot,\omega)]$ is continuously differentiable. In the case where $\mathbb{E}[F(\cdot,\omega)]$ is strongly monotone on K, problem (3.18) has a unique solution, which is also the unique solution of problems (3.22) and (3.23). By applying the stochastic approximation methods [45] for stochastic optimization problems, the iterative scheme based on (3.22) is given by

$$x^{k+1} = \mathrm{Proj}_K(x^k - \tau_k \zeta_\alpha^k), \qquad k = 0, 1, 2, \ldots, \tag{3.24}$$

where τ_k is a stepsize, ζ_α^k is a stochastic approximation of $\nabla g_\alpha(x^k)$, and Proj_K is the Euclidean projection operator onto K, while the iterative scheme based on (3.23) is given by

$$x^{k+1} = x^k - \tau_k \zeta_{\alpha\beta}^k, \qquad k = 0, 1, 2, \ldots, \tag{3.25}$$

where τ_k is a stepsize and $\zeta_{\alpha\beta}^k$ is a stochastic approximation of $\nabla g_{\alpha\beta}(x^k)$. Assuming further that the function $\mathbb{E}[F(\cdot,\omega)]$ is affine, it is shown that, under suitable assumptions, the sequence $\{x^k\}$ generated by (3.24) or (3.25) is almost surely convergent to the unique solution of (3.18) for any initial point $x^0 \in \Re^n$. Furthermore, since the iterative schemes (3.24) and (3.25) require to evaluate the Jacobians of the functions involved, a derivative-free iterative scheme is presented, which is based on the one proposed in [57], and is given by

$$x^{k+1} = x^k - \tau_k \zeta^k, \qquad k = 0, 1, 2, \ldots, \tag{3.26}$$

where τ_k is a stepsize and ζ^k is a stochastic approximation of the vector

$$d^k := H_\alpha(x^k) - H_\beta(x^k) + \rho[\alpha(x^k - H_\alpha(x^k)) - \beta(x^k - H_\beta(x^k))],$$

where $H_\alpha(x) := \mathrm{Proj}_K(x - \alpha^{-1}\mathbb{E}[F(x,\omega)])$ is the unique solution of the maximization problem in (3.21), and $\rho > 0$ is chosen sufficiently small so that d^k can serve as a descent direction of the merit function $g_{\alpha\beta}$ at x^k. Global convergence of the iterative method (3.26) is shown in [20].

Wang et al. [52] propose a sample average approximation method [47, Chapter 6] for problem (3.22). By taking independently and identically distributed samples $\{\omega_1, \omega_2, \ldots, \omega_k\} \subseteq \Omega$, the following sample average approximation of (3.22) is obtained:

$$\begin{aligned} \min \quad & g_\alpha^k(x) := \max_{y\in K} \{(x-y)^T \big[\frac{1}{k}\sum_{j=1}^k F(x,\omega_j)\big] - \frac{\alpha}{2}\|x-y\|_G^2\} \\ \text{s.t.} \quad & x \in K, \end{aligned} \tag{3.27}$$

whose objective function is differentiable everywhere [15]. By solving problems (3.27) with increasing k, a sequence of approximate solutions of (3.22) can be obtained. Comprehensive convergence analysis is established under suitable assumptions.

Xu and Zhang [55] consider a stochastic Nash equilibrium problem with N non-cooperative players. In particular, the first-order equilibrium conditions of the problem are formulated as the following generalized equations:

$$0 \in \mathbb{E}[\partial_{x^\nu} f_\nu(x^\nu, x^{-\nu}, \omega)] + \mathcal{N}_{K_\nu}(x^\nu), \qquad \nu = 1, 2, \cdots, N, \tag{3.28}$$

where x^ν, K_ν and f_ν denote the strategy, the strategy set, and the cost function of the νth player, respectively, $x^{-\nu}$ denotes the vector $(x^1, \cdots, x^{\nu-1}, x^{\nu+1}, \cdots, x^N)$, $\mathcal{N}_{K_\nu}(x^\nu)$ denotes the normal cone of K_ν at x^ν, and ∂_{x^ν} denotes the Clarke subdifferential operator. To compute a solution $x^* = (x^{1,*}, \ldots, x^{N,*}) \in K_1 \times \cdots \times K_N$ satisfying (3.28), the sample average approximation approach is considered in [55]. Specifically, by taking independently and identically distributed samples $\{\omega_1, \omega_2, \ldots, \omega_k\}$ from Ω, the sample average approximation generalized equations

$$0 \in \frac{1}{k}\sum_{j=1}^k \partial_{x^\nu} f_\nu(x^\nu, x^{-\nu}, \omega_j) + \mathcal{N}_{K_\nu}(x^\nu), \qquad \nu = 1, 2, \cdots, N,$$

are solved to get a solution $x^k = (x^{1,k}, \ldots, x^{N,k}) \in K_1 \times \cdots \times K_N$. It is shown that, under suitable conditions, the sequence $\{x^k\}$ converges to a solution of problem (3.28) with probability approaching one at exponential rate as the sample size tends to infinity.

3.2.2 ERM formulation. As a natural extension of Chen and Fukushima [8], Luo and Lin [36] present an ERM formulation for the SVIP (3.1). In a similar manner to [15], the regularized gap function $g_\alpha : \mathbb{R}^n \times \Omega \to [0, \infty)$ for (3.1) is defined by

$$g_\alpha(x, \omega) := \max_{y \in K} \big\{(x-y)^T F(x, \omega) - \frac{\alpha}{2}\|x-y\|_G^2\big\}, \tag{3.29}$$

where α is a positive parameter and G is an $n \times n$ symmetric positive definite matrix. Then the ERM formulation for (3.1) is given as the following optimization problem:

$$\begin{array}{ll} \min & \theta(x) := \mathbb{E}[g_\alpha(x, \omega)] = \displaystyle\int_\Omega g_\alpha(x, \omega)\rho(\omega)d\omega \\ \text{s.t.} & x \in K, \end{array} \tag{3.30}$$

where $\rho : \Omega \to [0, \infty)$ stands for the probability density function. In [36], the function $F(x, \omega)$ is assumed to be affine with respect to x. It is shown that, under very mild conditions, the level sets

$$\mathcal{L}^K(c) := \{x \in K \mid \theta(x) \le c\}$$

are bounded for any $c \ge 0$, which implies the existence of a solution of the ERM problem (3.30). Furthermore, since an expectation function is generally difficult to evaluate exactly, a quasi-Monte Carlo method for numerical integration [41] is suggested to solve (3.30). Specifically, by generating a sample set $\{\omega_1, \ldots, \omega_k\}$ from Ω by a quasi-Monte Carlo method, one may construct, for each k, an approximation to problem (3.30) as follows:

$$\begin{array}{ll} \min & \displaystyle\frac{1}{k}\sum_{i=1}^{k} g_\alpha(x, \omega_i)\rho(\omega_i) \\ \text{s.t.} & x \in K. \end{array} \tag{3.31}$$

It is shown [36] that, under some standard conditions, every accumulation point of the sequence of optimal solutions (or stationary points) of the approximation problems (3.31) is an optimal solution (or a stationary point) of the ERM problem (3.30).

Subsequently, Luo and Lin [37] generalize the above approach to the case where the function $F(x, \omega)$ is nonlinear with respect to x. First, assuming that the underlying sample space is a compact set, it is shown that the conclusions established in [36] remain valid for the nonlinear case. For the case where the sample space Ω is unbounded, the following approximation to problem (3.30) is considered:

$$\begin{array}{ll} \min & \theta_\nu(x) := \displaystyle\int_{\Omega_\nu} g_\alpha(x, \omega)\rho(\omega)d\omega \\ \text{s.t.} & x \in K, \end{array} \tag{3.32}$$

where $\nu > 0$ is a parameter and

$$\Omega_\nu := \{\omega \in \Omega \mid \|\omega\| \le \nu\}$$

is a compact approximation of Ω. Since problem (3.32) has a compact sample space, one may employ the above-mentioned method to solve it. It is shown that, for any sequence $\{\nu_k\}$ tending to infinity, every accumulation point of the sequence of optimal solutions to the corresponding problems (3.32) is an optimal solution of the ERM problem (3.30) under some conditions.

In general, an optimization problem obtained through the ERM formulation of SVIP is not necessarily convex, and hence it may be difficult to obtain a global optimal solution. In this respect, it is important and interesting to investigate conditions that guarantee the convexity of ERM problems. Agdeppa et al. [1] consider an affine SVIP, which is to find $x^* \in K(\omega)$ such that

$$(x - x^*)^T F(x^*, \omega) \geq 0, \qquad \forall\, x \in K(\omega), \tag{3.33}$$

where $F(x,\omega) := M(\omega)x + q(\omega)$ with $M : \Omega \to \Re^{n\times n}$ and $q : \Omega \to \Re^n$, and $K(\omega) := \{y \in \Re^n \mid A(\omega)y = b(\omega),\, y \geq 0\}$ with $A : \Omega \to \Re^{m\times n}$ and $b : \Omega \to \Re^m$. It is shown that the regularized gap function g_α defined by (3.29) and the D-gap function of the form

$$\hat{g}_\alpha(x,\omega) = g_{1/\alpha}(x,\omega) - g_\alpha(x,\omega),$$

where $\alpha > 1$, are (strongly) convex under suitable conditions on $M(\omega)$ and α. In consequence, the corresponding ERM formulation defined by g_α or $\hat{g}_\alpha$ is a convex program and hence a global optimal solution may be obtained by using standard optimization methods. Applications of the affine SVIP (3.33) to traffic equilibrium problems under uncertainty are also illustrated in [1].

Luo and Lin [38] consider an SVIP with additional constraints, which is to find a vector $x^* \in K$ satisfying

$$(x - x^*)^T F(x^*, \omega) \geq 0, \qquad \forall\, x \in K, \;\; \omega \in \Omega \text{ a.s.}, \tag{3.34}$$

and

$$u(x^*) \leq 0, \qquad v(x^*) = 0, \tag{3.35}$$

where $u : \Re^n \to \Re^p$ and $v : \Re^n \to \Re^q$ are smooth but not necessarily convex. Since the VIP is a generalization of a system of equations, the deterministic counterpart of (3.34)–(3.35) may be regarded as a generalization of constrained nonlinear equations, see for example [23]. Similarly to Luo and Lin [36, 37], since the system (3.34)–(3.35) may have no solution in general, the regularized gap function g_α defined by (3.29) is used to formulate the following ERM problem associated with (3.34)–(3.35):

$$\begin{array}{ll} \min & \mathbb{E}[g_\alpha(x,\omega)] \\ \text{s.t.} & x \in K, \\ & u(x) \leq 0,\; v(x) = 0. \end{array} \tag{3.36}$$

Quasi-Monte Carlo methods for solving (3.36) are considered in [38]. In addition, as an application of the SVIP with additional constraints, a supply chain network with random demands treated in [12] is further considered in [38]. The main difference between [12] and [38] is that Luo and Lin [38] discuss the case where the outputs are limited by the production scale or raw materials supply and hence the equilibrium conditions for the model in [38] are formulated as an SVIP in the form of (3.34)–(3.35), whereas the SVIP formulation established in [12] is essentially an SNCP.

4 Stochastic Mathematical Programs with Equilibrium Constraints

An SMPEC can be formulated as

$$\begin{aligned} \min \quad & \mathbb{E}[f(x,y,\omega)] \\ \text{s.t.} \quad & (x,y) \in Z \subseteq \Re^{n+m}, \\ & y \text{ solves } \mathrm{VI}(C(x,\omega), F(x,\cdot,\omega)),\ \omega \in \Omega \text{ a.s.} \end{aligned} \tag{4.1}$$

corresponding to (1.1), or

$$\begin{aligned} \min \quad & \mathbb{E}[f(x,y,\omega)] \\ \text{s.t.} \quad & (x,y) \in Z \subseteq \Re^{n+m}, \\ & y \geq 0,\ F(x,y,\omega) \geq 0,\ y^T F(x,y,\omega) = 0,\ \omega \in \Omega \text{ a.s.} \end{aligned} \tag{4.2}$$

corresponding to (1.2). In what follows, we restrict ourselves to the SMPEC of the form (4.2). Problem (4.1) may be dealt with in a similar manner by replacing the SVIP constraints by its Karush-Kuhn-Tucker representation [13]. Since an MPEC is already very hard to handle, SMPECs are generally more difficult to deal with because the number of elements in the sample space is usually very large or even infinite in practice.

Note that problem (4.2) has no feasible solution in general, and hence it is not a well-defined problem yet. Therefore, as in the study of SVIP and SNCP, the first task is to develop an appropriate deterministic formulation for (4.2). Two kinds of SMPECs are introduced in the unpublished paper [28] (see also [30]): One is called a "here-and-now" model, in which both the upper-level decision x and the lower-level decision y are required to be made before a random event is observed. The other is an "upper-level here-and-now and lower-level wait-and-see" model, in which the upper-level decision x is made at once and the lower-level decision y may be made after a random event is observed. For brevity, this model will simply be called a "lower-level wait-and-see" model. The difference between a here-and-now model and a lower-level wait-and see model is illustrated at the end of Example 2.2 in Section 2. We next review some recent developments for these models individually.

4.1 Here-and-now model with recourse

In the here-and-now case, for a given upper-level decision x, the lower-level decision y is required to satisfy the complementarity constraints for almost all ω in Ω. However, such a vector y does not exist in general. In order to get a reasonable resolution, Lin et al. [28, 30] introduce recourse variables $z(\omega) \geq 0$, $\omega \in \Omega$, and present a here-and-now model with recourse as follows:

$$\begin{array}{ll} \min\limits_{x,y,z(\cdot)} & \mathbb{E}[f(x,y,\omega) + d^T z(\omega)] \\ \text{s.t.} & (x,y) \in Z, \\ & y \geq 0,\ \ F(x,y,\omega) + z(\omega) \geq 0, \\ & y^T(F(x,y,\omega) + z(\omega)) = 0, \\ & z(\omega) \geq 0, \quad \omega \in \Omega \text{ a.s.}, \end{array} \tag{4.3}$$

where $d \in \Re^m$ is a weight vector with positive components. The SMPEC formulation (3.14) for SNCP is obviously a special case of the here-and-now model with recourse. Note that, although the feasibility issue is resolved, problem (4.3) is still not easy to deal with, because of the recourse variables in addition to the stochastic equilibrium constraints.

Lin et al. [30] focus on the here-and-now model with stochastic linear complementarity constraints

$$\begin{array}{ll} \min\limits_{x,y,z(\cdot)} & \mathbb{E}[f(x,y,\omega) + d^T z(\omega)] \\ \text{s.t.} & g(x) \leq 0,\ \ h(x) = 0, \\ & y \geq 0,\ \ N(\omega)x + M(\omega)y + q(\omega) + z(\omega) \geq 0, \\ & y^T(N(\omega)x + M(\omega)y + q(\omega) + z(\omega)) = 0, \\ & z(\omega) \geq 0, \quad \omega \in \Omega \text{ a.s.} \end{array} \tag{4.4}$$

First, suppose that the sample space is a finite set, i.e., $\Omega = \{\omega_1, \omega_2, \ldots, \omega_L\}$, and that, for each $\omega_l \in \Omega$, its probability p_l is known and positive. Then, by letting

$$N := \sum_{l=1}^{L} N(\omega_l), \quad M := \sum_{l=1}^{L} M(\omega_l), \quad q := \sum_{l=1}^{L} q(\omega_l),$$

problem (4.4) is rewritten as

$$\begin{aligned}
\min_{x,y,z(\omega_1),\ldots,z(\omega_L)} \quad & \sum_{l=1}^{L} p_l\Big(f(x,y,\omega_l) + d^T z(\omega_l)\Big) \\
\text{s.t.} \quad & g(x) \le 0,\ h(x) = 0,\ z(\omega_l) \ge 0, \\
& N(\omega_l)x + M(\omega_l)y + q(\omega_l) + z(\omega_l) \ge 0, \quad l = 1,\ldots,L, \\
& u = Nx + My + q + \textstyle\sum_{j=1}^{L} z(\omega_j), \\
& \phi_{\mathrm{FB}}(y_i,u_i) = 0, \quad i = 1,\ldots,m,
\end{aligned} \tag{4.5}$$

where ϕ_{FB} is the Fischer-Burmeister function defined in (3.4). As pointed out in [30], problem (4.5) fails to satisfy the MPEC-linear independence constraint qualification (MPEC-LICQ), which is often assumed in the literature on MPEC. Hence, besides the intractability due to a large number of variables, problem (4.5) is in some sense more difficult than an ordinary MPEC. A combined smoothing implicit programming and penalty method presented in [30] for solving the ill-posed MPEC (4.5) is described as follows: Suppose that M is a P_0 matrix, that is, all the principal minors are nonnegative. Then, for any $\mu > 0$, the matrix $M + \mu I$ is a P matrix, that is, all the principal minors are positive; see [11, 13] for details about P_0 and P matrices. As a result, for any $\mu > 0$ and any fixed $(x, z(\omega_1),\ldots,z(\omega_L))$, the following system of equations has a unique solution:

$$u = Nx + (M + \mu I)y + q + \textstyle\sum_{j=1}^{L} z(\omega_j), \tag{4.6}$$

$$\phi^{\mu}_{\mathrm{FB}}(y_i,u_i) = 0, \quad i = 1,\ldots,m, \tag{4.7}$$

where ϕ^{μ}_{FB} denotes the smoothed Fischer-Burmeister function defined by

$$\phi^{\mu}_{\mathrm{FB}}(a,b) := a + b - \sqrt{a^2 + b^2 + 2\mu^2}. \tag{4.8}$$

Let $(y_\mu(x, \sum_{j=1}^{L} z(\omega_j)), u_\mu(x, \sum_{j=1}^{L} z(\omega_j)))$ be the unique solution of (4.6)–(4.7). Note that it is continuously differentiable with respect to $(x, z(\omega_1),\ldots,z(\omega_L))$. Applying a penalty technique, we obtain the following approximation problem of (4.5):

$$\begin{aligned}
\min_{x,z(\omega_1),\ldots,z(\omega_L)} \quad & \sum_{l=1}^{L} p_l\Big(f(x, y_\mu(x, \textstyle\sum_{j=1}^{L} z(\omega_j)), \omega_l) + d^T z(\omega_l)\Big) \\
& + \rho \sum_{l=1}^{L} \varphi(-(N(\omega_l)x + M(\omega_l)y_\mu(x, \textstyle\sum_{j=1}^{L} z(\omega_j)) + q(\omega_l) + z(\omega_l))) \\
\text{s.t.} \quad & g(x) \le 0,\ h(x) = 0,\ z(\omega_l) \ge 0, \quad l = 1,\ldots,L,
\end{aligned} \tag{4.9}$$

where ρ is a positive parameter and $\varphi : \Re^m \to [0, +\infty)$ is a smooth penalty function. Problem (4.9) is a standard smooth nonlinear programming problem and, under some suitable conditions, it is a convex program. Therefore, problem (4.9) may be relatively easy to solve, provided that the evaluation of the implicit function $y_\mu(x, \sum_{j=1}^{L} z(\omega_j))$ is not very expensive. It is shown in [30] that, under suitable conditions, every accumulation point of the sequence of local optimal solutions (or stationary points) of the approximation problems (4.9) is a local optimal solution (or a C-stationary point) of the MPEC (4.5).

In [30], problem (4.4) with continuous random variables is also studied. For this case, the quasi-Monte Carlo methods for numerical integration [41] are employed to discretize (4.4). Comprehensive convergence analysis is established as well.

Lin and Fukushima [31, 32] consider the following here-and-now model with recourse:

$$\begin{array}{cl} \min\limits_{x,y,z(\cdot)} & f(x,y) + \mathbb{E}[d^T z(\omega)] \\ \text{s.t.} & g(x,y) \le 0,\ h(x,y) = 0, \\ & y \ge 0,\ N(\omega)x + M(\omega)y + q(\omega) + z(\omega) \ge 0, \\ & y^T(N(\omega)x + M(\omega)y + q(\omega) + z(\omega)) = 0, \\ & z(\omega) \ge 0, \quad \omega \in \Omega \text{ a.s.} \end{array} \tag{4.10}$$

In [31], by making use of the duality theorem in nonlinear programming theory, the recourse variables are removed from problem (4.10). As a result, the following problem equivalent to (4.10) is obtained:

$$\begin{array}{cl} \min\limits_{x,y} & f(x,y) + \mathbb{E}[Q(x,y,\omega)] \\ \text{s.t.} & g(x,y) \le 0,\ h(x,y) = 0,\ y \ge 0, \end{array} \tag{4.11}$$

where $Q : \Re^{n+m} \times \Omega \to [0, +\infty]$ is defined by

$$Q(x,y,\omega) := \sup_{u+ty \le d,\, u \ge 0,\, t \le 0} -(u+ty)^T(N(\omega)x + M(\omega)y + q(\omega)).$$

Since the function Q may not be finite-valued everywhere, it is further shown that (4.11) is equivalent to

$$\begin{array}{cl} \min\limits_{x,y} & f(x,y) + \mathbb{E}[d^T(-F(x,y,\omega))_+] \\ \text{s.t.} & g(x,y) \le 0,\ h(x,y) = 0, \quad y \ge 0, \\ & y \circ F(x,y,\omega) \le 0, \quad \omega \in \Omega \text{ a.s.} \end{array} \tag{4.12}$$

In [31], the sample space Ω is assumed to be a finite set $\{\omega_1, \omega_2, \ldots, \omega_L\}$. Noticing that the objective function in (4.12) is not differentiable everywhere and the problem has a great many constraints because L is usually very large in practice, the function $\frac{1}{2}(\sqrt{t^2+\epsilon^2}+t)$ with a small $\epsilon > 0$ is used to approximate the function $(t)_+ := \max(t, 0)$ and a penalty technique is employed to deal with the constraints. Then, a smoothing penalty method is presented for solving (4.12). See [31] for details about convergence of the method.

Subsequently, Lin and Fukushima [32] study the equivalent problem (4.11) of (4.10) from another viewpoint. Note that, for any fixed $\omega \in \Omega$, the function $Q(\cdot, \cdot, \omega)$ may be neither finite-valued nor differentiable everywhere. To construct an appropriate smooth approximation of this function, one may define the function $Q^\epsilon(\cdot, \cdot, \omega) : \Re^n \times \Re^m \to [0, +\infty)$ for a given positive parameter ϵ and fixed $\omega \in \Omega$ as follows:

$$Q^\epsilon(x, y, \omega) := \max_{u+ty\le d,\, u\ge 0,\, t\le 0} \Big\{ -(u + ty)^T (N(\omega)x + M(\omega)y + q(\omega)) - \frac{\epsilon}{2}(t^2 + \|u\|^2) \Big\}.$$

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

$$\begin{array}{ll} \min\limits_{u,t} & (u + ty)^T (N(\omega)x + M(\omega)y + q(\omega)) + \dfrac{\epsilon}{2}(t^2 + \|u\|^2) \\ \text{s.t.} & u + ty \le d, \quad u \ge 0, \quad t \le 0 \end{array} \tag{4.13}$$

must be an optimal solution and, since $\epsilon > 0$, problem (4.13) indeed has a unique optimal solution. This indicates that the function $Q^\epsilon(\cdot, \cdot, \omega)$ is well-defined for each $\omega \in \Omega$. It is further shown that $Q^\epsilon(\cdot, \cdot, \omega)$ is differentiable everywhere and

$$\begin{aligned} \nabla_{(x,y)} Q^\epsilon(x, y, \omega) \;=\; & -[\, N(\omega), M(\omega)\,]^T (u^\epsilon(x, y, \omega) + t^\epsilon(x, y, \omega)y) \\ & - t^\epsilon(x, y, \omega) \begin{pmatrix} 0 \\ N(\omega)x + M(\omega)y + q(\omega) + \zeta^\epsilon(x, y, \omega) \end{pmatrix}, \end{aligned}$$

where $(u^\epsilon(x, y, \omega), t^\epsilon(x, y, \omega))$ is the unique solution of problem (4.13) and $\zeta^\epsilon(x, y, \omega)$ is a corresponding Lagrangian multiplier vector. Thus, the problem

$$\begin{array}{ll} \min\limits_{x,y} & f(x, y) + \mathbb{E}[Q^\epsilon(x, y, \omega)] \\ \text{s.t.} & g(x, y) \le 0,\ h(x, y) = 0,\ y \ge 0 \end{array}$$

is a smooth approximation of problem (4.11). A regularization method based on the above idea is proposed. See [32] for more details.

4.2 Here-and-now models without recourse

Birbil et al. [2] present another here-and-now model

$$\begin{array}{ll} \min\limits_{x,y} & \mathbb{E}[f(x, y, \omega)] \\ \text{s.t.} & (x, y) \in Z, \\ & y \ge 0,\ \mathbb{E}[F(x, y, \omega)] \ge 0,\ y^T \mathbb{E}[F(x, y, \omega)] = 0, \end{array} \tag{4.14}$$

in which the lower-level decisions y are made so as to satisfy the complementarity constraints 'on average'. A sample-path method similar to the one given in [18] for dealing with SVIP is suggested to solve problem (4.14). Specifically, possibly by using Monte Carlo methods, some sequences $\{f^k\}$ and $\{F^k\}$ of functions are generated so that they converge almost surely to the functions $\mathbb{E}[f(\cdot,\cdot,\omega)]$ and $\mathbb{E}[F(\cdot,\cdot,\omega)]$, respectively. Then, a sequence of deterministic problems

$$\begin{array}{ll} \min & f^k(x,y) \\ \text{s.t.} & (x,y) \in Z, \\ & y \geq 0,\ F^k(x,y) \geq 0,\ y^T F^k(x,y) = 0 \end{array} \tag{4.15}$$

is solved to get an estimate of the true solution of problem (4.14). As usual, sensitivity analysis related to the deterministic approximation problems plays an important role in convergence analysis of the sample-path methods. In [2], in order to provide theoretical support for the convergence of the proposed sample-path method, some sufficient conditions are given under which the approximation problems (4.15) have solutions almost surely and the corresponding approximate solutions are close to the true solution of (4.14) when the simulation run length is sufficiently large. An application related to toll pricing in a transportation network is discussed as well.

Meng and Xu [40] consider the following problem:

$$\begin{array}{ll} \min & \mathbb{E}[f(z,\omega)] \\ \text{s.t.} & z \in Z, \\ & \mathbb{E}[F(z,\omega)] \geq 0,\ \mathbb{E}[G(z,\omega)] \geq 0,\ \mathbb{E}[F(z,\omega)]^T \mathbb{E}[G(z,\omega)] = 0, \end{array} \tag{4.16}$$

which is a generalization of the here-and-now model (4.14). A sample average approximation method, which is similar to the sample-path method given in [2], is proposed to solve problem (4.16). Specifically, taking independently and identically distributed random samples $\{\omega_1, \omega_2, \ldots, \omega_k\}$ from Ω, the problem

$$\begin{array}{ll} \min & \displaystyle\frac{1}{k}\sum_{j=1}^{k} f(z,\omega_j) \\ \text{s.t.} & z \in Z, \\ & \displaystyle\frac{1}{k}\sum_{j=1}^{k} F(z,\omega_j) \geq 0,\ \frac{1}{k}\sum_{j=1}^{k} G(z,\omega_j) \geq 0,\ \left[\frac{1}{k}\sum_{j=1}^{k} F(z,\omega_j)\right]^T \left[\frac{1}{k}\sum_{j=1}^{k} G(z,\omega_j)\right] = 0 \end{array} \tag{4.17}$$

is solved to get a weak stationary point; see [48] for various stationarity concepts for MPEC. It is shown in [40] that, under suitable conditions, the generated sequence converges to a weak stationary point of (4.16) with probability approaching one at exponential rate as the sample size tends to infinity.

Lin et al. [34] study the penalty-based Monte Carlo approach for the here-and-now model (4.14). The smoothed Fischer-Burmeister function defined in (4.8) and a penalty technique are used to deal with the complementarity constraints. Then a Monte Carlo method is employed to approximate the expectations. The approximation problem is given by

$$\begin{aligned} \min \quad & \frac{1}{k}\sum_{j=1}^{k} f(x,y,\omega_j) + \rho_k \sum_{i=1}^{m} \phi_{\mathrm{FB}}^{\mu_k}\Big(y_i, \frac{1}{k}\sum_{j=1}^{k} F_i(x,y,\omega_j)\Big) \\ \text{s.t.} \quad & (x,y) \in Z, \end{aligned} \tag{4.18}$$

where ρ_k and μ_k are positive penalty and smoothing parameters, respectively. Note that problem (4.18) does no longer contain any equilibrium constraint, whereas the approximation problems (4.15) and (4.17) are still MPECs. Strategies for selection of the parameters ρ_k and μ_k in (4.18) are discussed for the case where F is affine with respect to (x, y). Convergence of the proposed approach is established in [34]. Furthermore, since quasi-Monte Carlo methods are generally faster than Monte Carlo methods, a penalty-based quasi-Monte Carlo approach is considered as well.

Lin et al. [29] deal with SMPEC (4.2) in a different manner. Let ψ be a given nonnegative-valued restricted NCP function defined as in (3.12). Then problem (4.2) becomes

$$\begin{aligned} \min \quad & \mathbb{E}[f(x,y,\omega)] \\ \text{s.t.} \quad & (x,y) \in Z,\ y \geq 0, \\ & \psi(y_i, F_i(x,y,\omega)) = 0,\ \omega \in \Omega \text{ a.s.},\ i = 1,\dots,m. \end{aligned}$$

Recall that the above problem is not well-defined because its feasible region is generally empty. Therefore, a penalty technique is employed to present a deterministic formulation as follows:

$$\begin{aligned} \min \quad & \mathbb{E}\Big[f(x,y,\omega) + \rho\sum_{i=1}^{m} \psi(y_i, F_i(x,y,\omega))\Big] \\ \text{s.t.} \quad & (x,y) \in Z,\ y \geq 0, \end{aligned} \tag{4.19}$$

where $\rho > 0$ is a penalty parameter. Unlike the here-and-now models presented in [28, 30] and [2], problem (4.19) is no longer an MPEC and hence it may be relatively easy to deal with. Algorithms based on sample average approximations can be developed to solve (4.19). Moreover, when x and f are not present and Z is the whole space, problem (4.19) reduces to the ERM formulation (3.13) for SNCP.

4.3 Lower-level wait-and-see model

The lower-level wait-and-see model can be formulated as follows:

$$\begin{aligned}
\min_{x,y(\cdot)} \quad & \mathbb{E}[f(x,y(\omega),\omega)] \\
\text{s.t.} \quad & (x,y(\omega)) \in Z, \\
& y(\omega) \geq 0,\ \ F(x,y(\omega),\omega) \geq 0, \\
& y(\omega)^T F(x,y(\omega),\omega) = 0, \quad \omega \in \Omega \text{ a.s.}
\end{aligned} \tag{4.20}$$

This model is discussed in Patriksson and Wynter [44], where the existence of a solution of the model (4.20) is studied and some results on convexity and directional differentiability of the implicit upper-level objective function are given. The links between (4.20) and two-stage stochastic programs with recourse are discussed and some suggestions on algorithms for solving (4.20) are made.

For the case of a finite sample space $\Omega = \{\omega_1, \ldots, \omega_L\}$, Lin et al. [28] deal with the problem

$$\begin{aligned}
\min_{x,y(\omega_1),\ldots,y(\omega_L)} \quad & \sum_{l=1}^{L} p_l f(x,y(\omega_l),\omega_l) \\
\text{s.t.} \quad & g(x) \leq 0,\ \ h(x) = 0, \\
& y(\omega_l) \geq 0,\ \ N(\omega_l)x + M(\omega_l)y + q(\omega_l) \geq 0, \\
& y(\omega_l)^T (N(\omega_l)x + M(\omega_l)y + q(\omega_l)) = 0, \quad l = 1, \ldots, L,
\end{aligned} \tag{4.21}$$

in a similar manner to [30] for the here-and-now model (4.4) described in Subsection 4.1. Specifically, by using the Fischer-Burmeister function ϕ_{FB}, problem (4.21) is rewritten as

$$\begin{aligned}
\min_{x,y(\omega_1),\ldots,y(\omega_L)} \quad & \sum_{l=1}^{L} p_l f(x,y(\omega_l),\omega_l) \\
\text{s.t.} \quad & g(x) \leq 0,\ \ h(x) = 0, \\
& \phi_{\mathrm{FB}}(y_i(\omega_l), [N(\omega_l)x + M(\omega_l)y(\omega_l) + q(\omega_l)]_i) = 0, \\
& i = 1, \ldots, m, \quad l = 1, \ldots, L.
\end{aligned} \tag{4.22}$$

Suppose that p_l is positive and $M(\omega_l)$ is a P_0-matrix for each l, and let $\mu > 0$ be a smoothing parameter. Then the system of equations

$$\phi_{\mathrm{FB}}^{\mu}(y_i(\omega_l), [N(\omega_l)x + (M(\omega_l) + \mu I)y(\omega_l) + q(\omega_l)]_i) = 0, \quad i = 1, \ldots, m,$$

where ϕ_{FB}^{μ} is the smoothed Fischer-Burmeister function defined in (4.8), has a unique solution $y_\mu^l(x)$ for each fixed x. Moreover, $y_\mu^l(\cdot)$ is a continuously differentiable function. Applying to

(4.22), one obtains the following approximation problem:

$$\begin{aligned} \min_{x} \quad & \sum_{l=1}^{L} p_l f(x, y_\mu^l(x), \omega_l) \\ \text{s.t.} \quad & g(x) \leq 0,\ h(x) = 0. \end{aligned} \tag{4.23}$$

Note that, unlike the method given in [30] for problem (4.4), penalty techniques are not required in dealing with problem (4.21). Convergence of the proposed method is established.

Xu [53] studies a supply side oligopoly market consisting of a leader and several followers with random demands. Such a game is known as a stochastic Stackelberg-Nash-Cournot game. Suppose that the demand is not realized at the time when the leader makes his decision. Then the leader's purpose is to maximize his expected profit based on his knowledge of the distribution of demand and the followers' reactions in each scenario. The Stackelberg-Nash-Cournot equilibrium, in which the leader chooses an optimal supply to maximize his expected profit and the followers react to the leader's choice to reach a Nash-Cournot equilibrium, of the market can be formulated as a stochastic bilevel programming problem. By reformulating the lower-level problem as a stochastic complementarity problem, the stochastic Stackelberg-Nash-Cournot equilibrium problem is transformed into an SMPEC in the form of (4.20). Properties of the followers' Nash-Cournot equilibrium and existence of the stochastic Stackelberg-Nash-Cournot equilibrium are investigated. Furthermore, for the discrete case, a smoothing approach based on the smoothed Fischer-Burmeister function is suggested and, for the continuous case, a discretization approach based on implicit numerical integration is proposed. In addition, a stochastic Stackelberg-Nash-Cournot equilibrium problem with two leaders is discussed as well.

Xu [54] considers the following special case of the model (4.20):

$$\begin{aligned} \min_{x, y(\cdot)} \quad & \mathbb{E}[f(x, y(\omega), \omega)] \\ \text{s.t.} \quad & x \in X, \\ & y(\omega) \geq 0,\ F(x, y(\omega), \omega) \geq 0, \\ & y(\omega)^T F(x, y(\omega), \omega) = 0, \quad \omega \in \Omega \text{ a.s.} \end{aligned} \tag{4.24}$$

The approach suggested in [54] may be regarded as an extension of [28] from the discrete and linear case to the continuous and nonlinear case. The Fischer-Burmeister function ϕ_{FB} is used to rewrite problem (4.24) as follows:

$$\begin{aligned} \min_{x, y(\cdot)} \quad & \mathbb{E}[f(x, y(\omega), \omega)] \\ \text{s.t.} \quad & x \in X, \\ & \phi_{\mathrm{FB}}(y_i(\omega), F_i(x, y(\omega), \omega)) = 0, \quad i = 1, \ldots, m. \end{aligned} \tag{4.25}$$

Under the assumption that the function $F(x, y, \omega)$ is uniformly strongly monotone in y and uniformly locally Lipschitz continuous in x, it is shown that the system of equations

$$\phi_{\text{FB}}(y_i(\omega), F_i(x, y(\omega), \omega)) = 0, \quad i = 1, \ldots, m,$$

defines an implicit function $y(x, \omega)$ that is piecewise smooth in x, whereas the system of equations

$$\phi^{\mu}_{\text{FB}}(y_i(\omega), F_i(x, y(\omega), \omega)) = 0, \quad i = 1, \ldots, m,$$

defines an implicit function $y_\mu(x, \omega)$ that is smooth in x. Here, $\mu > 0$ and ϕ^{μ}_{FB} is the smoothed Fischer-Burmeister function. Based on the above results, two implicit programming models

$$\begin{array}{ll} \min\limits_{x} & \mathbb{E}[f(x, y(x, \omega), \omega)] \\ \text{s.t.} & x \in X, \end{array} \tag{4.26}$$

which is equivalent to problem (4.25), and

$$\begin{array}{ll} \min\limits_{x} & \mathbb{E}[f(x, y_\mu(x, \omega), \omega)] \\ \text{s.t.} & x \in X, \end{array}$$

which is a smooth approximation of problem (4.25), are investigated. In addition, a deterministic discretization method, in which a set of grid points is chosen from the support set of ω and the stochastic complementarity constraints are replaced by a set of complementarity constraints with ω taking values at these grid points, is presented. Limiting behavior of Clarke stationary points of the approximation problems is also discussed.

Shapiro [49] considers a class of SMPECs of the form[4]

$$\begin{array}{ll} \min\limits_{x} & \mathbb{E}[\vartheta(x, \omega)] \\ \text{s.t.} & x \in X. \end{array} \tag{4.27}$$

Here $\vartheta(x, \omega)$ is the optimal value of the following optimization problem with x and ω fixed:

$$\begin{array}{ll} \min\limits_{y} & f(x, y, \omega) \\ \text{s.t.} & 0 \in H(x, y, \omega) + \mathcal{N}_{C(x,\omega)}(y), \end{array} \tag{4.28}$$

where $f : \Re^n \times \Re^m \times \Omega \to \Re$, $H : \Re^n \times \Re^m \times \Omega \to \Re^s$, $C : \Re^n \times \Omega \to 2^{\Re^m}$, and $\mathcal{N}_{C(x,\omega)}(y)$ denotes the normal cone to the set $C(x, \omega)$ at y. The constraints of (4.28) is a so-called generalized equation, which is essentially equivalent to a VIP, provided that $C(x, \omega) \equiv K$ with K being a

[4] In [49], problem (4.27) is referred to as a 'here-and-now' type problem. According to the categorization in the present paper, however, this problem is actually a 'lower-level wait-and-see' model, since the lower-level decision y is determined as a solution of problem (4.28) that depends not only x but also ω.

closed convex set. In [49], the measurability of the function $\vartheta(x,\cdot)$ and the continuity of the objective function of (4.27) are studied.

Even a crude discretization of the sample space yields an exponential number of scenarios with respect to its dimension, and hence the discretized problems can easily become unmanageable when the number of random parameters increases. Shapiro and Xu [50] study the sample average approximation approach, which uses the Monte Carlo sampling techniques to reduce the set of considered scenarios to a manageable level, for solving problem (4.27)–(4.28) with $C(x,\omega) \equiv K$ for a closed convex set K. Under the assumption that the optimization problem (4.28) has a unique solution $y(x,\omega)$ for every x in X and almost all ω in Ω, problem (4.27)–(4.28) is rewritten as

$$\begin{array}{ll} \min\limits_{x} & \mathbb{E}[f(x, y(x,\omega), \omega)] \\ \text{s.t.} & x \in X. \end{array} \tag{4.29}$$

Then, by taking independently and identically distributed random samples $\{\omega_1, \omega_2, \dots, \omega_k\} \subseteq \Omega$, the following sample average approximation of problem (4.29) is obtained:

$$\begin{array}{ll} \min\limits_{x} & \dfrac{1}{k}\displaystyle\sum_{j=1}^{k} f(x, y(x,\omega_j), \omega_j) \\ \text{s.t.} & x \in X. \end{array} \tag{4.30}$$

Convergence properties of the sample average approximation method is investigated in [50]. In particular, it is shown that, under some reasonable assumptions, stationary points of the approximation problems (4.30) almost surely converge to stationary points of problem (4.29). Furthermore, based on the large deviation theory, uniform exponential convergence of the sample average approximations is shown under additional assumptions.

5 Concluding Remarks

In principle, any problem involving uncertain or stochastic data can hardly be of practical use unless it is formulated as a deterministic problem, and it is where a variety of ideas come into play. In stochastic programming [3, 22], two-stage stochastic programming with recourse and chance-constrained programming are, among others, popular vehicles for constructing deterministic models, for which rich theory and many practical algorithms have been developed for half a century. For CP and VIP, the study on their stochastic counterparts is much more recent and much less abundant than stochastic programming. Moreover, although SMPEC certainly constitutes a subclass of stochastic programming problems, it is much more complicated structurally than the standard stochastic programming problem since its deterministic counterpart is already of bilevel nature. As we have seen in the previous sections, several formulations and solution methods have been proposed to deal with those problems. However, the results obtained so far are by no means satisfactory and there remains much room to be done. It should be emphasized that, in general, different deterministic formulations may yield different solutions and hence have different implications. Therefore, it would be particularly important to clarify the characteristics of each formulation from various aspects, so that one can find the most suitable formulation for a practical application at hand.

Acknowledgements. We would like to thank Professor Xiaojun Chen and Dr Huifu Xu for their helpful comments and suggestions.

References

[1] R.P. Agdeppa, N. Yamashita and M. Fukushima, Convex expected residual models for stochastic affine variational inequality problems and its application to the traffic equilibrium problem, *Pacific Journal of Optimization*, to appear.

[2] S.I. Birbil, G. Gürkan and O. Listes, Solving stochastic mathematical programs with complementarity constraints using simulation, *Mathematics of Operations Research*, 31 (2006), 739–760.

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

[4] B. Chen, X. Chen and C. Kanzow, A penalized Fischer-Burmeister NCP-function, *Mathematical Programming*, 88 (2000), 211–216.

[5] B. Chen and P.T. Harker, Smooth approximations to nonlinear complementarity problems, *SIAM Journal on Optimization*, 7 (1997), 403–420.

[6] C. Chen and O.L. Mangasarian, A class of smoothing functions for nonlinear and mixed complementarity problems, *Computational Optimization and Applications*, 5 (1996), 97–138.

[7] H.F. Chen, *Stochastic Approximation and Its Applications*, Kluwer Academic Publishers, Dordrecht, 2003.

[8] X. Chen and M. Fukushima, Expected residual minimization method for stochastic linear complementarity problems, *Mathematics of Operations Research*, 30 (2005), 1022–1038

[9] X. Chen, C. Zhang and M. Fukushima, Robust solution of monotone stochastic linear complementarity problems, *Mathematical Programming*, 117 (2009), 51–80.

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

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

[12] J. Dong, D. Zhang and A. Nagurney, A supply chain network equilibrium model with random demands, *European Journal of Operational Research*, 156 (2004), 194–212.

[13] F. Facchinei and J.S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer-Verlag, New York, 2003.

[14] H. Fang, X. Chen and M. Fukushima, Stochastic R_0 matrix linear complementarity problems, *SIAM Journal on Optimization*, 18 (2007), 482–506.

[15] M. Fukushima, Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems, *Mathematical Programming*, 53 (1992), 99–110.

[16] M. Fukushima, Merit functions for variational inequality and complementarity problems, in *Nonlinear Optimization and Applications*, G. Di Pillo and F. Giannessi (eds.), Plenum Press, New York, 1996, pp. 155–170.

[17] M. Fukushima and G.H. Lin, Smoothing methods for mathematical programs with equilibrium constraints, *Proceedings of the ICKS'04, IEEE Computer Society*, 2004, pp. 206–213.

[18] G. Gürkan, A.Y. Özge and S.M. Robinson, Sample-path solution of stochastic variational inequalities, *Mathematical Programming*, 84 (1999), 313–333.

[19] P.T. Harker and J.S. Pang, Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications, *Mathematical Programming*, 48 (1990), 161–220.

[20] H. Jiang and H. Xu, Stochastic approximation approaches to the stochastic variational inequality problem, *IEEE Transactions on Automatic Control*, 53 (2008), 1462–1475.

[21] A. Juditsky, G. Lany, A. Nemirovski and A. Shapiro, Stochastic approximation approach to stochastic programming, http://www.optimization-online.org/DB_HTML/2007/09/1787.html.

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

[23] C. Kanzow, N. Yamashita and M. Fukushima, Levenberg-Marquardt methods for constrained nonlinear equations with strong local convergence properties, *Journal of Computational and Applied Mathematics*, 172 (2004), 375–397.

[24] I.V. Konnov, *Equilibrium Models and Variational Inequalities*, Elsevier, Amsterdam, 2007.

[25] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*, Academic Press, New York, 1980.

[26] A.J. King and R.T. Rochafellar, Asymptotic theory for solutions in statistical estimation and stochastic programming, *Mathematics of Operations Research*, 18 (1993), 148–162.

[27] G.H. Lin, Combined Monte Carlo sampling and penalty method for stochastic nonlinear complementarity problems, *Mathematics of Computation*, to appear (DOI: 10.1090/S0025-5718-09-02206-6).

[28] 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.

[29] G.H. Lin, X. Chen and M. Fukushima, New restricted NCP function and their applications to stochastic NCP and stochastic MPEC, *Optimization*, 56 (2007), 641–753.

[30] G.H. Lin, X. Chen and M. Fukushima, Solving stochastic mathematical programs with equilibrium constraints via approximation and smoothing implicit programming with penalization, *Mathematical Programming*, 116 (2009), 343–368.

[31] 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.

[32] 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.

[33] G.H. Lin and M. Fukushima, New reformulations for stochastic complementarity problems, *Optimization Methods and Software*, 21 (2006), 551–564.

[34] G.H. Lin, H. Xu and M. Fukushima, Monte Carlo and quasi-Monte Carlo sampling methods for a class of stochastic mathematical programs with equilibrium constraints, *Mathematical Methods of Operations Research*, 67 (2008), 423–441.

[35] C. Ling, L. Qi, G. Zhou and L. Caccetta, The SC^1 property of an expected residual function arising from stochastic complementarity problems, *Operations Research Letters*, 36 (2008), 456–460.

[36] M.J. Luo and G.H. Lin, Expected residual minimization method for stochastic variational inequality problems, *Journal of Optimization Theory and Application*, 140 (2009), 103–116.

[37] M.J. Luo and G.H. Lin, Convergence results of the ERM method for nonlinear stochastic variational inequality problems, *Journal of Optimization Theory and Application*, to appear (DOI: 10.1007/s10957-009-9534-3).

[38] M.J. Luo and G.H. Lin, Stochastic variational inequality problems with additional constraints and their applications in supply chain network equilibria, preprint, 2009.

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

[40] F. Meng and H. Xu, Exponential convergence of sample average approximation methods for a class of stochastic mathematical programs with complementarity constraints, *Journal of Computational Mathematics*, 24 (2006), 733–748.

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

[42] J.V. Outrata, M. Kocvara and J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results*, Kluwer Academic Publishers, Boston, 1998.

[43] J.S. Pang, Complementarity problems, in *Handbook in Global Optimization*, R. Horst and P. Pardalos, (eds.), Kluwer Academic Publishers, Boston, 1994.

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

[45] G.C. Pflug, *Optimization of Stochastic Models: The Interface between Simulation and Optimization*, Kluwer Academic Publishers, Boston, 1996.

[46] L. Qi and J. Sun, A nonsmooth version of Newton's method, *Mathematical Programming*, 58 (1993), 353–367.

[47] A. Ruszczynski and A. Shapiro, *Stochastic Programming, Handbooks in Operations Research and Management Science*, Elsevier, 2003.

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

[49] A. Shapiro, Stochastic mathematical programs with equilibrium constraints, *Journal of Optimization Theory and Applications*, 128 (2006), 223–243.

[50] A. Shapiro and H. Xu, Stochastic mathematical programs with equilibrium constraints, modeling and sample average approximation, *Optimization*, 57 (2008), 395–418.

[51] M.Z. Wang, M.M. Ali and G.H. Lin, Smoothing-based sample average approximation method for stochastic nonlinear complementarity problems with applications to supply chain super-networks, preprint, 2007.

[52] M.Z. Wang, G.H. Lin, Y. Gao and M.M. Ali, Sample average approximation method for a class of stochastic variational inequality problems, preprint, 2008.

[53] H. Xu, An MPCC approach for stochastic Stackelberg-Nash-Cournot equilibrium, *Optimization*, 54 (2005), 27–57.

[54] H. Xu, An implicit programming approach for a class of stochastic mathematical programs with linear complementarity constraints, *SIAM Journal on Optimization*, 16 (2006), 670–696.

[55] H. Xu and D. Zhang, Stochastic Nash equilibrium problems: Sample average approximation and applications, preprint, 2008.

[56] N. Yamashita, Properties of restricted NCP functions for nonlinear complementarity problems, *Journal of Optimization Theory and Applications*, 98 (1998), 701–717.

[57] N. Yamashita, K. Taji and M. Fukushima, Unconstrained optimization reformulations of variational inequality problems, *Journal of Optimization Theory and Applications*, 92 (1997), 439–456.

[58] C. Zhang and X. Chen, Smoothing projected gradient method and its application to stochastic linear complementarity problems, *SIAM Journal on Optimization*, to appear.

[59] C. Zhang and X. Chen, Stochastic nonlinear complementarity problem and applications to traffic equilibrium under uncertainty, *Journal of Optimization Theory and Applications*, 137 (2008), 277–295.

[60] G.L. Zhou and L. Caccetta, Feasible semismooth Newton method for a class of stochastic linear complementarity problems, *Journal of Optimization Theory and Applications*, 139 (2008), 379–392.