

Stochastic Programming with Equilibrium Constraints

A. SHAPIRO¹

Communicated by P. M. Pardalos

Abstract. In this paper, we discuss here-and-now type stochastic programs with equilibrium constraints. We give a general formulation of such problems and study their basic properties such as measurability and continuity of the corresponding integrand functions. We discuss also the consistency and rate of convergence of sample average approximations of such stochastic problems.

Key Words. Equilibrium constraints, two-stage stochastic programming, variational inequalities, complementarity conditions, statistical inference, exponential rates.

1. Introduction

In this paper, we discuss optimization problems where the objective function is given as the expectation $\mathbb{E}[F(x, y(x, \omega), \omega)]$, with $y(x, \omega)$ being a random vector defined as a solution of a second stage optimization or complementarity problem. By analogy with stochastic programming with recourse, we can view such problems as here-and-now type problems where a decision should be made before a realization of random data becomes available (see e.g. Ref. 1). We refer to such problems as stochastic mathematical programming with equilibrium constraints (SMPEC) problems. The basic difference between SMPEC problems and classical two-stage stochastic programming problems is that the function $F(x, y(x, \omega), \omega)$ does not have to be the optimal value of the corresponding second-stage problem.

SMPEC problems of here-and-now type were discussed in Refs. 2–4 for example. The main subject of Refs. 2 and 3 is an investigation of the existence of the optimal solutions of SMPEC problems, while in Ref. 4 an

¹Professor, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.

example of a variant of the newsboy (vendor) problem is discussed and an approach to the numerical solution of a class of SMPEC problems is suggested. A convergence theory of discrete approximations of a type of SMPEC problems different from here-and-now was discussed in Ref. 5.

Stochastic Stackelberg-Nash-Cournot equilibrium problem were formulated in the form of stochastic mathematical programs with complementarity constraints (SMPCC) and studied in Ref. 6; a general form of SMPCC is discussed in Ref. 7. There is also an extensive literature on deterministic MPEC problems. In that respect, we may refer to Refs. 8–9 and references therein.

We use the following notation and terminology. For a number $a \in \mathbb{R}$, we denote

$$[a]_+ := \max\{0, a\}.$$

By

$$\|x\| := (x^T x)^{1/2},$$

we denote the Euclidean norm of the vector $x \in \mathbb{R}^n$. For a set $S \subset \mathbb{R}^n$, we denote by

$$\text{dist}(x, S) := \inf_{y \in S} \|x - y\|$$

the distance from a point x to S and by $\text{conv}(S)$ the convex hull of S . For a convex set S , we denote by $T_S(x)$ the tangent cone to S at $x \in S$. For a cone $Q \subset \mathbb{R}^s$, we denote by

$$Q^* := \{z \in \mathbb{R}^s : z^T y \leq 0, \forall y \in Q\}$$

its polar (negative dual) cone. For a mapping $G(x, y)$ from $\mathbb{R}^n \times \mathbb{R}^m$ to \mathbb{R}^s , we denote by $\nabla_y G(x, y)$ the matrix of its partial derivative, i.e., the $s \times m$ Jacobian matrix with respect to the vector y . For a function $g: \mathbb{R}^m \rightarrow \mathbb{R}$, we denote by

$$\text{epi}(g) := \{(x, \alpha) : g(x) \leq \alpha\}$$

its epigraph and by $\nabla^2 g(y)$ its $m \times m$ Hessian matrix of second-order partial derivatives. For a mapping $G(y) = (g_1(y), \dots, g_s(y)) : \mathbb{R}^m \rightarrow \mathbb{R}^s$ and a vector $d \in \mathbb{R}^m$, we denote

$$\nabla^2 G(y)d := ([\nabla^2 g_1(y)]d, \dots, [\nabla^2 g_s(y)]d)^T.$$

For $\varepsilon \geq 0$, it is said that \bar{x} is an ε -optimal solution of the problem of minimization of a function $f(x)$ over a set X if $\bar{x} \in X$ and $f(\bar{x}) \leq \inf_{x \in X} f(x) + \varepsilon$. For $a, b \in \mathbb{R}$, we define

$$a \vee b := \max\{a, b\}.$$

2. Model

In this section, we give a precise definition of the considered class of SMPEC problems.

Consider the following optimization problem:

$$\text{Min}_{x \in X} \{f(x) := \mathbb{E}[\vartheta(x, \omega)]\}, \tag{1}$$

where X is a nonempty closed subset of \mathbb{R}^n ,

$$\vartheta(x, \omega) := \inf_{y \in \mathbb{S}(x, \omega)} F(x, y, \omega), \tag{2}$$

and $\mathbb{S}(x, \omega)$ is the set of solutions of the variational inequality

$$H(x, y, \omega) \in N_{K(x, \omega)}(y); \tag{3}$$

i.e., $y \in \mathbb{S}(x, \omega)$ iff (3) holds. Here, (Ω, \mathcal{F}, P) is a probability space, $F: \mathbb{R}^n \times \mathbb{R}^m \times \Omega \rightarrow \mathbb{R}$ is a real-valued function, $K: \mathbb{R}^n \times \Omega \rightrightarrows \mathbb{R}^m$ is a set-valued mapping (multifunction), $H: \mathbb{R}^n \times \mathbb{R}^m \times \Omega \rightarrow \mathbb{R}^m$, and $N_K(y)$ denotes a normal cone to a set $K \subset \mathbb{R}^m$ at y . Unless stated otherwise, we make all probabilistic statements with respect to the probability measure P . In particular, the expectation $\mathbb{E}[\cdot]$ is taken with respect to P . We refer to (1)–(3) as a stochastic mathematical programming with equilibrium constraints (SMPEC) problem.

For a convex and closed set $K \subset \mathbb{R}^m$, the corresponding normal cone is defined in the standard way,

$$N_K(y) := \{z \in \mathbb{R}^m : z^T(y' - y) \leq 0, \forall y' \in K\}, \text{ if } y \in K, \tag{4a}$$

$$N_K(y) = \emptyset, \text{ if } y \notin K. \tag{4b}$$

For a nonconvex set K , there are several possible concepts of normal cones. We use the following construction. We assume that $K(x, \omega)$ is given in the form

$$K(x, \omega) := \{y \in \mathbb{R}^m : G(x, y, \omega) \in Q\}, \tag{5}$$

where $Q \subset \mathbb{R}^s$ is a closed convex cone and $G: \mathbb{R}^n \times \mathbb{R}^m \times \Omega \rightarrow \mathbb{R}^s$ is a continuously differentiable in (x, y) mapping. Then, by definition,

$$N_{K(x,\omega)}(y) := \{\nabla_y G(x, y, \omega)^T \lambda : \lambda \in N_Q(G(x, y, \omega))\}, \tag{6}$$

where $N_Q(z)$ is the (standard) normal cone to Q at z . It can be noted that $N_{K(x,\omega)}(y)$, defined in (6), coincides with the polar of the tangent cone to $K(x, \omega)$ at $y \in K(x, \omega)$, provided that a constraint qualification holds at the point y . Moreover, if the set $K(x, \omega)$ is convex, then $N_{K(x,\omega)}(y)$ coincides with the standard normal cone.

Let us remark at this point that, in order to ensure that the above problem is well defined, we need to verify that the expectation $\mathbb{E}[\vartheta(x, \omega)]$ is well defined. We will discuss this later. Note also that, by definition, $\vartheta(x, \omega) = +\infty$ if the set $\mathbb{S}(x, \omega)$ is empty. Therefore, for a given $x \in X$, if the set $\mathbb{S}(x, \omega)$ is empty with a positive probability, then $f(x) = +\infty$. That is, implicitly the optimization in (1) is performed over such $x \in X$ that the set $\mathbb{S}(x, \omega)$ is nonempty for almost every (a.e.) $\omega \in \Omega$.

For $N_{K(x,\omega)}(y)$ defined in (6), the variational inequality (3) takes the form of the following so-called generalized equations:

$$-H(x, y, \omega) + \nabla_y G(x, y, \omega)^T \lambda = 0, \quad \lambda \in N_Q(G(x, y, \omega)). \tag{7}$$

Since Q is a convex cone, we have that, for any $z \in Q$,

$$N_Q(z) = \{z^* \in Q^* : z^T z^* = 0\}, \tag{8}$$

where Q^* is the polar of the cone Q . Therefore, $\lambda \in N_Q(G(x, y, \omega))$ iff

$$\lambda \in Q^*, \quad G(x, y, \omega) \in Q, \quad \lambda^T G(x, y, \omega) = 0, \tag{9}$$

or equivalently iff

$$G(x, y, \omega) \in N_{Q^*}(\lambda).$$

It follows that we can write the generalized equations (7) as the variational inequality (cf. Ref. 10):

$$\mathcal{H}(x, \zeta, \omega) \in N_Q(\zeta), \tag{10}$$

where $\zeta := (y, \lambda) \in \mathbb{R}^{m+s}$ and

$$\mathcal{H}(x, \zeta, \omega) := \begin{bmatrix} -H(x, y, \omega) + \nabla_y G(x, y, \omega)^T \lambda \\ G(x, y, \omega) \end{bmatrix}, \quad Q := \mathbb{R}^m \times Q^*. \tag{11}$$

It can be observed that the set Q is a closed convex cone which does not depend on x and ω . The price of this simplification is that the variational inequality (10) is to be solved jointly in y and λ . We denote by $\overline{S}(x, \omega) \subset \mathbb{R}^{m+s}$ the set of solutions of the variational inequality (10); for $y \in \mathbb{R}^m$, we denote by $\Lambda(x, y, \omega)$ the set of all λ satisfying equations (7). Note that $\Lambda(x, y, \omega)$ is nonempty iff $y \in S(x, \omega)$.

By the above discussion, $\overline{S}(x, \omega)$ is also the set of solutions of the generalized equations (7) and the set $S(x, \omega)$ is obtained by the projection of the set $\overline{S}(x, \omega)$ onto \mathbb{R}^m . Therefore, $\vartheta(x, \omega)$ coincides with the optimal value of the following problem:

$$\text{Min}_{y \in \mathbb{R}^m, \lambda \in \mathbb{R}^s} F(x, y, \omega), \tag{12a}$$

$$\text{s.t.} \quad -H(x, y, \omega) + \nabla_y G(x, y, \omega)\lambda = 0, \tag{12b}$$

$$\lambda \in Q^*, G(x, y, \omega) \in Q, \lambda^T G(x, y, \omega) = 0. \tag{12c}$$

Moreover, if

$$H(x, y, \omega) = -\nabla_y h(x, y, \omega),$$

where $h: \mathbb{R}^m \times \mathbb{R}^m \times \Omega \rightarrow \mathbb{R}$ is a continuously differentiable in (x, y) function, then the generalized equations (7) represent the first-order (KKT) optimality conditions for the optimization problem

$$\text{Min}_{y \in \mathbb{R}^m} h(x, y, \omega), \quad \text{s.t. } G(x, y, \omega) \in Q. \tag{13}$$

We refer to the set $S(x, \omega)$ of solutions of the corresponding variational inequality as the set of stationary points of (13). That is, y is a stationary point of (13) if there exists $\lambda \in \mathbb{R}^s$ such that

$$\nabla_y h(x, y, \omega) + \nabla_y G(x, y, \omega)^T \lambda = 0, \quad \lambda \in N_Q(G(x, y, \omega)). \tag{14}$$

The set $\Lambda(x, y, \omega)$ of all λ satisfying (14) represents the set of Lagrange multipliers associated with y . It is well known that, under a constraint qualification, a locally optimal solution of (13) is also its stationary point. In the case where the set $S(x, \omega)$ is given by the set of stationary points of an optimization problem of the form (13), we refer to (1) as a stochastic bilevel mathematical programming (SBMP) problem.

Example 2.1. Let $Q := -\mathbb{R}_+^r \times \{0\} \subset \mathbb{R}^s$, where \mathbb{R}_+^r is the nonnegative orthant of \mathbb{R}^r , $\{0\}$ is the null space of \mathbb{R}^{r-s} , and

$$G(x, y, \omega) = (g_1(x, y, \omega), \dots, g_s(x, y, \omega)).$$

Here, the set

$$K(x, \omega) = \{y : g_i(x, y, \omega) \leq 0, \quad i = 1, \dots, r, g_i(x, y, \omega) = 0, i = r + 1, \dots, s\} \tag{15}$$

is defined by a finite number of inequality and equality constraints. Also, the normal cone to Q at $z \in Q$ can be written as follows:

$$N_Q(z) = \{z^* \in \mathbb{R}^s : z_i^* \geq 0, i = 1, \dots, r, z_i^* z_i = 0, i = 1, \dots, r\} \tag{16}$$

and

$$\nabla_y G(x, y, \omega)^T \lambda = \sum_{i=1}^s \lambda_i \nabla_y g_i(x, y, \omega),$$

where $\lambda \in \mathbb{R}^s$; hence, the generalized equations (7) take the form

$$-H(x, y, \omega) + \sum_{i=1}^s \lambda_i \nabla_y g_i(x, y, \omega) = 0, \tag{17a}$$

$$g_i(x, y, \omega) \leq 0, \quad i = 1, \dots, r, \tag{17b}$$

$$g_i(x, y, \omega) = 0, \quad i = r + 1, \dots, s, \tag{17c}$$

$$\lambda_i \geq 0, \quad \lambda_i g_i(x, y, \omega) = 0, \quad i = 1, \dots, r. \tag{17d}$$

Example 2.2. Suppose that, for $x \in X$, the set $\mathbb{S}(x, \omega)$ is given by the set of optimal solutions of the following optimization problem:

$$\text{Min}_{y \in \mathbb{R}^m} q(\omega)^T y, \quad \text{s.t. } T(\omega)x + W(\omega)y = h(\omega), \quad y \geq 0, \tag{18}$$

where $q(\omega)$ and $h(\omega)$ are vector-valued and $T(\omega)$ and $W(\omega)$ are matrix-valued random variables defined on the probability space (Ω, \mathcal{F}, P) . The above problem (18) is a linear programming problem. Moreover, if

$$F(x, y, \omega) := c^T x + q(\omega)^T y,$$

then the corresponding problem (1) becomes a two-stage stochastic linear program with recourse.

We have that $y(x, \omega)$ is an optimal solution of (18) iff $y(x, \omega)$ is a solution of the variational inequality

$$-q(\omega) \in N_{K(x, \omega)}(y), \tag{19}$$

where

$$K(x, \omega) := \{y \in \mathbb{R}^m : T(\omega)x + W(\omega)y = h(\omega), y \geq 0\}. \tag{20}$$

Here, for every $(x, \omega) \in \mathbb{R}^n \times \Omega$, the set $K(x, \omega)$ is convex polyhedral and the corresponding normal cone is defined in the standard way. The set $\mathbb{S}(x, \omega)$ of solutions of problem (18), or equivalently the variational inequality (19), can be empty, contain a single extreme point of $K(x, \omega)$, or can be formed by a nontrivial face of $K(x, \omega)$. Let us remark that the set $\mathbb{S}(x, \omega)$ can be empty for two somewhat different reasons. Namely, it may happen that the feasible set $K(x, \omega)$ of problem (18) is empty, in which case its optimal value is $+\infty$, or it may happen that problem (18) is unbounded from below, i.e., its optimal value is $-\infty$. Note that, in our framework, in both cases we assign the value $+\infty$ to the corresponding function $\vartheta(x, \omega)$.

The dual of problem (18) is

$$\text{Max}_{\lambda} \lambda^T (h(\omega) - T(\omega)x), \quad \text{s.t. } W(\omega)^T \lambda \leq q(\omega). \tag{21}$$

The optimality conditions (14) take here the form

$$y \geq 0, \quad W(\omega)^T \lambda - q(\omega) \leq 0, \tag{22a}$$

$$y^T (W(\omega)^T \lambda - q(\omega)) = 0, \quad T(\omega)x + W(\omega)y - h(\omega) = 0, \tag{22b}$$

and can be written as the variational inequality (10) with $\zeta = (y, \lambda)$ and

$$\mathcal{H}(x, \zeta, \omega) := \begin{bmatrix} W(\omega)^T \lambda - q(\omega) \\ T(\omega)x + W(\omega)y - h(\omega) \end{bmatrix}, \quad \mathcal{Q} := \mathbb{R}_+^m \times \mathbb{R}^s. \tag{23}$$

3. Properties of SMPEC Problems

In this section, we discuss some basic properties of the SMPEC problem (1) with the function $\vartheta(x, \omega)$ given by the optimal value of the problem (12). To some extent, such SMPEC problem can be studied in the framework of two-stage stochastic programming problems with recourse (see Chapters 1 and 2 in Ref. 11).

As it was mentioned in the previous section, we need to ensure that the expectation $\mathbb{E}[\vartheta(x, \omega)]$ is well defined for any $x \in X$. That is, we need to verify that

$$\vartheta_x(\omega) := \vartheta(x, \omega)$$

is measurable and

$$\text{either } \mathbb{E}[(\vartheta_x(\omega))_+] < +\infty \quad \text{or } \mathbb{E}[(-\vartheta_x(\omega))_+] < +\infty.$$

For a thorough discussion of the following measurability concepts, we may refer to Refs. 12 and 13. When considering the space $\mathbb{R}^n \times \Omega$, we equip it always with the sigma algebra given by the product of the Borel sigma algebra of \mathbb{R}^n and \mathcal{F} . A multifunction $\mathcal{G}: \Omega \rightrightarrows \mathbb{R}^n$ is said to be closed valued if $\mathcal{G}(\omega)$ is a closed subset of \mathbb{R}^n for every $\omega \in \Omega$. A closed valued multifunction \mathcal{G} is said to be measurable if $\mathcal{G}^{-1}(A) \in \mathcal{F}$ for every closed set $A \subset \mathbb{R}^n$. A function $h: \mathbb{R}^n \times \Omega \rightarrow \overline{\mathbb{R}}$ is said to be random lower semicontinuous if the epigraphical mapping $\omega \mapsto \text{epi } h(\cdot, \omega)$ is closed valued and measurable (random lower semicontinuous functions are called normal integrands in Ref. 13). It is said that a mapping $G: \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}^m$ is a Carathéodory mapping if $G(x, \omega)$ is continuous in x for every ω and is measurable in ω for every x . Note that a Carathéodory function $g(x, \omega)$ is random lower semicontinuous. We have the following result (e.g. Theorem 14.37 in Ref. 13).

Theorem 3.1. Let $g: \mathbb{R}^n \times \Omega \rightarrow \overline{\mathbb{R}}$ be a random lower semicontinuous function. Then, the min function $\phi(\omega) := \inf_{x \in \mathbb{R}^n} g(x, \omega)$ is measurable and the multifunction $\mathcal{G}(\omega) := \arg \min_{x \in \mathbb{R}^n} g(x, \omega)$ is closed valued and measurable.

We make the following assumption throughout the paper.

- (A1) The mappings $H(x, y, \omega)$, $G(x, y, \omega)$, and the function $F(x, y, \omega)$ are Carathéodory mappings, i.e., continuous in (x, y) and measurable in ω .

Consider the variational inequality (10). By Assumption (A1), we have that the mapping $\mathcal{H}(x, \zeta, \omega)$, defined in (11), is a Carathéodory mapping. The variational inequality (10) can be transformed into an optimization problem as follows. Consider the following (regularized) gap function:

$$\gamma(x, \zeta, \omega) := \sup_{\zeta' \in \mathcal{Q}} \{\mathcal{H}(x, \zeta, \omega)^T (\zeta' - \zeta) - (1/2) \|\zeta - \zeta'\|^2\}. \quad (24)$$

Regularized gap functions were introduced in Refs. 14 and 15 (see also Section 10.2.1 in Ref. 16 for a discussion of gap functions). We have that, for given x and ω , a point $\bar{\zeta} \in \mathcal{Q}$ is a solution of (10), i.e.,

$$\bar{\zeta} \in \overline{\mathcal{S}}(x, \omega), \quad \text{iff } \gamma(x, \bar{\zeta}, \omega) = 0.$$

Since

$$\gamma(x, \zeta, \omega) \geq 0, \quad \text{for any } \zeta \in \mathcal{Q},$$

it follows that $\bar{\zeta} \in \bar{\mathbb{S}}(x, \omega)$ iff $\bar{\zeta}$ is an optimal solution of the problem

$$\text{Min}_{\zeta \in \mathcal{Q}} \gamma(x, \zeta, \omega), \quad \text{s.t. } \gamma(x, \zeta, \omega) = 0. \tag{25}$$

We have also that the gap function $\gamma(x, \zeta, \omega)$ is continuous in (x, ζ) , measurable in ω , and hence is a Carathéodory function. Consequently, we obtain that the multifunction $(x, \omega) \mapsto \bar{\mathbb{S}}(x, \omega)$ is closed valued and measurable.

The function $\vartheta_x(\omega)$ can be written also in the form

$$\vartheta_x(\omega) = \inf_{(y, \lambda) \in \bar{\mathbb{S}}(x, \omega)} F(x, y, \omega). \tag{26}$$

For a given x , since the multifunction $\bar{\mathbb{S}}(x, \cdot)$ is closed valued and measurable and since by (A1) the function $F(x, \cdot, \cdot)$ is random lower semicontinuous, we obtain that $\vartheta_x(\omega)$ is measurable. This settles the question of measurability of the integrand function inside the expectation in (1).

It is said that $\zeta(\omega) = (y(\omega), \lambda(\omega))$ is a measurable selection of $\bar{\mathbb{S}}(x, \omega)$ if $\zeta(\omega) \in \bar{\mathbb{S}}(x, \omega)$, for a.e. $\omega \in \Omega$, and $\zeta(\omega)$ is measurable. Consider now the function

$$\tilde{f}(x) := \inf_{\zeta(\cdot) \in \bar{\mathbb{S}}(x, \cdot)} \mathbb{E}[F(x, y(\omega), \omega)], \tag{27}$$

where $\zeta(\cdot) \in \bar{\mathbb{S}}(x, \cdot)$ means that the optimization is performed over all measurable selections $\zeta(\omega)$ of $\bar{\mathbb{S}}(x, \omega)$. By the definition, $\tilde{f}(x) := +\infty$ if the set $\bar{\mathbb{S}}(x, \omega)$ is empty with positive probability. We have then that

$$f(x) = \tilde{f}(x), \quad \text{for every } x \in X.$$

Indeed, assume that $\bar{\mathbb{S}}(x, \omega)$ is nonempty for a.e. $\omega \in \Omega$ [otherwise, $f(x) = \tilde{f}(x) = +\infty$]. For any measurable selection $\zeta(\omega) \in \bar{\mathbb{S}}(x, \omega)$, we have that $F(x, y(\cdot), \cdot) \geq \vartheta_x(\cdot)$ and hence $\tilde{f}(x) \geq f(x)$. Conversely, by the Castaing representation theorem (Theorem 14.5 in Ref. 13), for any $\varepsilon > 0$, there exists a measurable selection $\zeta(\omega) \in \bar{\mathbb{S}}(x, \omega)$ such that

$$F(x, y(\cdot), \cdot) \leq \vartheta_x(\cdot) + \varepsilon.$$

It follows that

$$\tilde{f}(x) \leq f(x) + \varepsilon,$$

and hence $\tilde{f}(x) = f(x)$. We obtain the following result.

Lemma 3.1. Problem (1) is equivalent to the problem

$$\text{Min}_{x \in X, \zeta(\cdot) \in \overline{\mathbb{S}}(x, \cdot)} \mathbb{E}[F(x, y(\omega), \omega)], \tag{28}$$

where the notation $\zeta(\cdot) \in \overline{\mathbb{S}}(x, \cdot)$ means that the optimization is performed over all measurable selections $\zeta(\omega)$ of $\overline{\mathbb{S}}(x, \omega)$ and such $x \in X$ that $\overline{\mathbb{S}}(x, \omega)$ is nonempty for a.e. $\omega \in \Omega$.

In other words, the optimization in (28) is performed over $x \in X$ and measurable $\zeta(\omega) = (y(\omega), \lambda(\omega))$ which satisfy the feasibility constraints of problem (12). In particular, suppose that $\Omega = \{\omega_1, \dots, \omega_K\}$ is finite with respective probabilities p_1, \dots, p_K . Then, we can write the problem (1) in the following equivalent form:

$$\text{Min}_{x \in X, y_1, \dots, y_K} \sum_{k=1}^K p_k F(x, y_k, \omega_k), \tag{29a}$$

$$\text{s.t.} \quad -H(x, y_k, \omega_k) + \nabla_y G(x, y_k, \omega_k) \lambda_k = 0, \quad k = 1, \dots, K, \tag{29b}$$

$$\lambda_k \in Q^*, \quad G(x, y_k, \omega_k) \in Q,$$

$$\lambda_k^T G(x, y_k, \omega_k) = 0, \quad k = 1, \dots, K. \tag{29c}$$

4. Properties of the Expectation Function

In this section, we discuss the continuity and differentiability properties of the expectation function $f(x) := \mathbb{E}[\vartheta(x, \omega)]$, where $\vartheta(x, \omega)$ is given by the optimal value of problem (12).

Consider a point $\bar{x} \in X$ and suppose that there exists a P -integrable function $\psi(\omega)$ such that $\vartheta(x, \omega) \geq \psi(\omega)$ for all $\omega \in \Omega$ and all x in a neighborhood of \bar{x} . It is said that a measurable function $\psi : \Omega \rightarrow \mathbb{R}$ is P -integrable if $\mathbb{E}[|\psi|] < +\infty$. Then, we have by the Fatou lemma that

$$\liminf_{x \rightarrow \bar{x}} \mathbb{E}[\vartheta(x, \omega)] \geq \mathbb{E}[\liminf_{x \rightarrow \bar{x}} \vartheta(x, \omega)]. \tag{30}$$

It follows that $f(\cdot)$ is lower semicontinuous at \bar{x} if $\vartheta(\cdot, \omega)$ is lower semicontinuous at \bar{x} for a.e. $\omega \in \Omega$. Therefore, the question of the lower semicontinuity of $f(\cdot)$ is reduced to studying the lower semicontinuity of $\vartheta(\cdot, \omega)$.

It is not difficult to give an example of a min-function of a parametric family of continuous functions which is not lower semicontinuous (e.g. Example 4.1 in Ref. 17). Therefore, we need some type of a boundedness condition. We consider a point $\bar{x} \in X$ and make the following assumptions. Recall that $\overline{\mathbb{S}}(x, \omega)$ denotes the set of solutions of the generalized equations (7) and that $\overline{\mathbb{S}}(x, \omega)$ coincides with the feasible set of problem (12).

- (A2) The mappings $H(\cdot, \cdot, \omega)$, $\nabla_y G(\cdot, \cdot, \omega)$ and function $F(\cdot, \cdot, \omega)$ are continuous.
- (A3) For $x = \bar{x}$, the set $\overline{\mathbb{S}}(\bar{x}, \omega)$ is nonempty and the (possibly empty) sets $\overline{\mathbb{S}}(x, \omega)$ are uniformly bounded for all x in a neighborhood of \bar{x} .

Let us observe that, because of Assumption (A2) and since cone Q and cone Q^* are closed, the multifunction $x \mapsto \overline{\mathbb{S}}(x, \omega)$ is closed. That is, if $x_n \rightarrow \bar{x}$, $\zeta_n \in \mathbb{S}(x_n, \omega)$ and $\zeta_n \rightarrow \bar{\zeta}$, then $\bar{\zeta} \in \mathbb{S}(\bar{x}, \omega)$. The following result follows easily by the compactness arguments and closedness of the multifunction $\overline{\mathbb{S}}(\cdot, \omega)$; it is quite well known.

Lemma 4.1. Suppose that Assumptions (A2) and (A3) hold. Then, the min function $\vartheta(\cdot, \omega)$ is lower semicontinuous at \bar{x} .

By the above discussion, we have the following result.

Proposition 4.1. Suppose that Assumptions (A2) and (A3) hold for a.e. $\omega \in \Omega$ and there exists a P -integrable function $\psi(\omega)$ such that $\vartheta(x, \omega) \geq \psi(\omega)$ for all $\omega \in \Omega$ and all x in a neighborhood of \bar{x} . Then, the expectation function $f(x)$ is lower semicontinuous at \bar{x} .

The above proposition shows that, under mild boundedness conditions, the expectation function is lower semicontinuous. Then, it follows that problem (1) has an optimal solution provided it has a nonempty and bounded level set,

$$\text{lev}_\alpha f := \{x \in X : f(x) \leq \alpha\},$$

for some $\alpha \in \mathbb{R}$; we refer to such condition as the inf-compactness condition.

In order to ensure the continuity of $f(x)$, one needs to investigate the continuity properties of $\vartheta(\cdot, \omega)$. That is, by using the Lebesgue dominated convergence theorem, it is straightforward to show that, if there exists a P -integrable function $\psi(\omega)$ such that $|\vartheta(x, \omega)| \leq \psi(\omega)$, for a.e. $\omega \in \Omega$ and all x in a neighborhood of \bar{x} , then $f(\cdot)$ is continuous at \bar{x} if $\vartheta(\cdot, \omega)$ is continuous at \bar{x} for a.e. $\omega \in \Omega$ (e.g. Ref. 11, page 66).

Consider the following assumption associated with a point $\bar{y} \in \mathbb{S}(\bar{x}, \omega)$.

- (A4) To every x in a neighborhood of \bar{x} , there corresponds a $y_\omega(x) \in \mathbb{S}(x, \omega)$ such that $y_\omega(x)$ tends to \bar{y} as $x \rightarrow \bar{x}$.

The above assumption implies that the set $\mathbb{S}(x, \omega)$, hence the set $\overline{\mathbb{S}}(x, \omega)$, is nonempty for all x in a neighborhood of \bar{x} . Then, we have the following result (e.g. Proposition 4.4 in Ref. 17).

Lemma 4.2. Suppose that Assumptions (A2) and (A3) are satisfied and that Assumption (A4) holds with respect to a point $\bar{y} \in \arg \min_{y \in \mathbb{S}(\bar{x}, \omega)} F(\bar{x}, y, \omega)$. Then, the min function $\vartheta(\cdot, \omega)$ is continuous at \bar{x} .

The above discussion implies the following proposition.

Proposition 4.2. Suppose that Assumptions (A2)–(A4) hold for a.e. $\omega \in \Omega$ and there exist a P -integrable function $\psi(\omega)$ such that $|\vartheta(x, \omega)| \leq \psi(\omega)$ for all $\omega \in \Omega$ and all x in a neighborhood of \bar{x} . Then, the expectation function $f(x)$ is continuous at \bar{x} .

The above Assumption (A4) is formed actually by two parts, namely:

- (A4a) The set $\mathbb{S}(x, \omega)$, equivalently $\bar{\mathbb{S}}(x, \omega)$, is nonempty for all x in a neighborhood of \bar{x} .
- (A4b) There exists a selection $y_\omega(x) \in \mathbb{S}(x, \omega)$ converging to \bar{y} as $x \rightarrow \bar{x}$.

The nonemptiness of $\mathbb{S}(x, \omega)$ can be ensured by various conditions. For example, consider the optimization problem (13). It has an optimal solution $\bar{y}_\omega(x)$ provided the corresponding inf compactness condition holds. Moreover, if a constraint qualification holds at $\bar{y}_\omega(x)$, then $\bar{y}_\omega(x)$ is a stationary point and hence $\mathbb{S}(x, \omega)$ is nonempty. For a general discussion of existence of solutions of variational inequalities, we may refer to Ref. 18 and Section 2.2 of Ref. 16. For example, if the set $K = K(x, \omega)$ is convex, then the variational inequality (3) has a solution if there exists $y^* \in K$ such that the set

$$\{y \in K : H(x, y, \omega)^T (y - y^*) > 0\}$$

is bounded (see Proposition 2.2.3 in Ref. 16). Note that the set $K(x, \omega)$ is closed, since it is assumed that $G(x, \cdot, \omega)$ is continuous and Q is closed. In particular, this condition holds if the set K is convex and bounded.

A way to ensure the above condition (A4b) is to verify the local uniqueness of the solution \bar{y} . Indeed, let \bar{y} be an isolated point of $\mathbb{S}(\bar{x}, \omega)$; i.e., there is a neighborhood $V \subset \mathbb{R}^m$ of \bar{y} such that $\mathbb{S}(\bar{x}, \omega) \cap V = \{\bar{y}\}$. Of course, we can always choose the neighborhood V to be bounded. In that case, under the boundedness assumption (A3), we have by compactness arguments that, if $y_\omega(x) \in \mathbb{S}(x, \omega) \cap V$, then $y_\omega(x)$ tends to \bar{y} as $x \rightarrow \bar{x}$. It is also of interest to estimate the rate at which $y_\omega(x)$ converges to \bar{y} . We say that the multifunction $x \mapsto \mathbb{S}(x, \omega)$ is locally upper Hölder of degree γ at \bar{x} for \bar{y} , if there exist neighborhoods W and V of \bar{x} and \bar{y} , respectively, and a constant $c = c(\omega)$ such that

$$\|y_\omega(x) - \bar{y}\| \leq c \|x - \bar{x}\|^\gamma, \tag{31}$$

for all $x \in W$ and any $y_\omega(x) \in \mathbb{S}(x, \omega) \cap V$. In particular, if this holds with $\gamma = 1$, we say that $\mathbb{S}(\cdot, \omega)$ is locally upper Lipschitz at \bar{x} for \bar{y} .

It is possible to give various conditions ensuring the local uniqueness of a solution $\bar{y} \in \mathbb{S}(\bar{x}, \omega)$. For fixed $\bar{x} \in X$ and $\omega \in \Omega$, we use the notation $H(\cdot) = H(\bar{x}, \cdot, \omega)$ and $G(\cdot) = G(\bar{x}, \cdot, \omega)$, etc. It is said that the Robinson constraint qualification for the system $G(y) \in Q$ holds at \bar{y} if

$$[\nabla G(\bar{y})]\mathbb{R}^m + T_Q(G(\bar{y})) = \mathbb{R}^s. \tag{32}$$

Note that, if (32) holds, then the set $\Lambda(\bar{y}) = \Lambda(\bar{x}, \bar{y}, \omega)$ of the corresponding Lagrange multipliers is bounded. Noted also that, if the system is defined by a finite number of constraints, as in Example 2.1, then the Robinson constraint qualification coincides with the Mangasarian-Fromovitz constraint qualification.

Recall that the cone

$$C(\bar{y}) := \{d \in \mathbb{R}^m : H(\bar{y})^T d = 0, \nabla G(\bar{y})d \in T_Q(G(\bar{y}))\}$$

is called the critical cone. For a vector $d \in \mathbb{R}^m$ and $G(y) = (g_1(y), \dots, g_s(y))$, consider the set of Lagrange multipliers

$$\Lambda^*(\bar{y}, d) := \arg \max_{\lambda \in \Lambda(\bar{y})} \sum_{i=1}^s \lambda_i d^T \nabla^2 g_i(\bar{y}) d.$$

We have the following result (see Theorems 3.1, 3.2, 4.3, and 5.1 in Ref. 19).

Proposition 4.3. Consider a point $\bar{y} \in \mathbb{S}(\bar{x}, \omega)$. Suppose that $H(y)$ is continuously differentiable, $G(y)$ is twice continuously differentiable, the Robinson constraint qualification (32) holds, and the cone Q is polyhedral. Then, \bar{y} is an isolated point of $\mathbb{S}(\bar{x}, \omega)$ if the following condition holds for vectors $d \in \mathbb{R}^m$:

$$\nabla H(\bar{y})d \in \text{conv} \{ \cup_{\lambda \in \Lambda^*(\bar{y}, d)} [\nabla^2 G(\bar{y})d]^T \lambda \} + N_{C(\bar{y})}(d) \text{ implies } d = 0. \tag{33}$$

Also, if the above condition (33) holds and if $H(x, y, \omega)$ is continuously differentiable and $G(x, y, \omega)$ is twice continuously differentiable jointly in x and y , then $\mathbb{S}(\cdot, \omega)$ is locally upper Hölder of degree $\gamma = 1/2$, at \bar{x} for \bar{y} . Moreover, if the mapping $G(x, y, \omega)$ and hence the set $K(x, \omega)$ do not depend on x , then $\mathbb{S}(\cdot, \omega)$ is locally upper Lipschitz at \bar{x} for \bar{y} .

It is possible to extend the above condition (33) to situations where the cone Q is not polyhedral (see Ref. 19), although the analysis becomes more involved. Of course, if the mapping $G(y)$ is affine, then

$$\nabla^2 G(y) = 0;$$

hence, the system in the left-hand side of (33) reduces to

$$\nabla H(\bar{y})d \in N_{C(\bar{y})}(d).$$

Condition (33) is equivalent to the quadratic growth condition, for a scaled regularized gap function, to hold at the point \bar{y} . In case \bar{y} is a locally optimal solution of the optimization problem (13), condition (33) is equivalent (under the assumptions of the above proposition) to the following quadratic growth condition: there exists $c > 0$ and a neighborhood V of \bar{y} such that

$$h(y) \geq h(\bar{y}) + c\|y - \bar{y}\|^2, \quad \forall y \in K(\bar{x}, \omega) \cap V. \quad (34)$$

In the general case of $G(x, y, \omega)$ depending on x , in order to ensure the locally upper Lipschitz continuity of $\mathbb{S}(\cdot, \omega)$ at \bar{x} for \bar{y} , one needs conditions which are considerably stronger than assumption (33). For optimization problems, for example, the quadratic growth condition (34) does not imply the locally upper Lipschitz continuity of $\mathbb{S}(\cdot, \omega)$ and a strong form of the second-order optimality conditions is needed (see Ref. 20).

Of course, for any $y_\omega(x) \in \arg \min_{y \in \mathbb{S}(x, \omega)} F(x, y, \omega)$, we have that

$$\vartheta(x, \omega) = F(x, y_\omega(x), \omega).$$

Consequently, if $F(\cdot, \cdot, \omega)$ is continuously differentiable and hence is locally Lipschitz continuous, we can write

$$|\vartheta(x, \omega) - \vartheta(\bar{x}, \omega)| = |F(x, y_\omega(x), \omega) - F(\bar{x}, \bar{y}, \omega)| \leq \ell(\omega)\|y_\omega(x) - \bar{y}\|,$$

where $x \in W$ and $\ell(\omega)$ is a corresponding Lipschitz constant (independent of x). Then, it follows from (31) that, for $x \in W$,

$$|\vartheta(x, \omega) - \vartheta(\bar{x}, \omega)| \leq \kappa(\omega)\|x - \bar{x}\|^\gamma, \quad (35)$$

where

$$\kappa(\omega) := \ell(\omega)c(\omega).$$

In turn, the above inequality (35) implies that, for $x \in W$,

$$|f(x) - f(\bar{x})| \leq L\|x - \bar{x}\|^\gamma, \quad (36)$$

where

$$L := \mathbb{E}[\kappa(\omega)].$$

Of course, the estimate (36) is meaningful only if the expectation $\mathbb{E}[\kappa(\omega)]$, hence the constant L , is finite.

The Lipschitz continuity and differentiability properties of the optimal value function $\vartheta(\cdot, \omega)$ were also studied also in Ref. 21 in the case when the set $K(x, \omega)$ is defined by a finite number of constraints, as in Example 2.1.

Let us discuss now the differentiability properties of $f(x)$. By $f'(\bar{x}, p)$ and $\vartheta'_\omega(\bar{x}, p)$, we denote the directional derivatives of $f(\cdot)$ and $\vartheta(\cdot, \omega)$ respectively at \bar{x} in the direction p . It is said that $f(\cdot)$ is directionally differentiable at \bar{x} if $f'(\bar{x}, p)$ exists for all $p \in \mathbb{R}^n$. By using the Lebesgue dominated convergence theorem, it is not difficult to show the following (see Chapter 2, Proposition 2 in Ref. 11).

Lemma 4.3. Suppose that $\vartheta(\bar{x}, \cdot)$ is P -integrable, $\vartheta(\cdot, \omega)$ is directionally differentiable at $\bar{x} \in X$ for a.e. $\omega \in \Omega$, and there exists a P -integrable function $\kappa : \Omega \rightarrow \mathbb{R}_+$ such that (35) holds with $\gamma = 1$, i.e.,

$$|\vartheta(x, \omega) - \vartheta(\bar{x}, \omega)| \leq \kappa(\omega) \|x - \bar{x}\|, \tag{37}$$

for all x in a neighborhood of \bar{x} and a.e. $\omega \in \Omega$. Then, $f(\cdot)$ is directionally differentiable at \bar{x} and

$$f'(\bar{x}, p) = \mathbb{E}[\vartheta'_\omega(\bar{x}, p)], \quad \forall p \in \mathbb{R}^n. \tag{38}$$

Of course, (36) follows from (37) with $\gamma = 1$; hence, the above assumptions imply that $f(x)$ is finite valued for all x near \bar{x} .

Suppose now that, for all x near \bar{x} , there exists

$$y_\omega(x) \in \arg \min_{y \in \mathbb{S}(x, \omega)} F(x, y, \omega)$$

converging to a point \bar{y} as $x \rightarrow \bar{x}$. Then, since

$$\vartheta(x, \omega) = F(x, y_\omega(x), \omega),$$

we obtain by the chain rule that

$$\vartheta'_\omega(\bar{x}, p) = \nabla_y F(\bar{x}, \bar{y}, \omega)^T y'_\omega(\bar{x}, p), \tag{39}$$

provided that $F(\bar{x}, \cdot, \omega)$ is continuously differentiable and the directional derivative $y'_\omega(\bar{x}, p)$ does exist. Concerning the directional differentiability of $y_\omega(x)$, we have the following result (Corollary 4.1 in Ref. 19).

Proposition 4.4. Suppose that the assumptions of Proposition 4.3 and Assumption (A4) holds and that: (i) the mapping $G(x, y, \omega)$ does not depend on x ; (ii) for a given $\omega \in \Omega$ and any $p \in \mathbb{R}^n$, the system

$$\begin{aligned} & \nabla_x H(\bar{x}, \bar{y}, \omega)p + \nabla_y H(\bar{x}, \bar{y}, \omega)d \\ & \in \text{conv} \{ \cup_{\lambda \in \Lambda^*(\bar{y}, d)} [\nabla^2 G(\bar{y})d]^T \lambda \} + N_{C(\bar{y})}(d) \end{aligned} \tag{40}$$

has the unique solution $\bar{d} = \bar{d}(p)$. Then, $y_\omega(\cdot)$ is directionally differentiable at \bar{x} and $y'_\omega(\bar{x}, p) = \bar{d}(p)$.

Note that, by setting $p=0$ in (40), we obtain that condition (ii) in the above proposition implies condition (33) of Proposition 4.3.

5. Statistical Inference

In order to solve SMPEC problems numerically, one needs to discretize (possibly continuous) distributions of the involved random variables. In this section, we discuss briefly the Monte Carlo sampling approach to such a discretization. Assume that all the involved random data depend on a random vector $\xi(\omega)$, where $\xi(\cdot): \Omega \rightarrow \mathbb{R}^d$ is a measurable mapping. Denote by $\Xi \subset \mathbb{R}^d$ the support of the distribution of $\xi(\omega)$. Let $F(x, y, \xi)$ be a real-valued function, $F: \mathbb{R}^n \times \mathbb{R}^m \times \Xi \rightarrow \mathbb{R}$, and suppose that the objective function in (2) can be written (with a certain abuse of notation) as $F(x, y, \xi(\omega))$, and similarly $H(x, y, \xi(\omega))$ and $G(x, y, \xi(\omega))$. Let ξ^1, \dots, ξ^N be a random sample of $\xi(\omega)$. Define

$$\hat{f}_N(x) := (1/N) \sum_{j=1}^N \vartheta(x, \xi^j), \tag{41}$$

where $\vartheta(x, \xi)$ is the optimal value of the following problem:

$$\text{Min}_{y \in \mathbb{R}^m, \lambda \in \mathbb{R}^s} F(x, y, \xi), \tag{42a}$$

$$\text{s.t.} \quad -H(x, y, \xi) + \nabla_y G(x, y, \xi)\lambda = 0, \tag{42b}$$

$$\lambda \in Q^*, G(x, y, \xi) \in Q, \lambda^T G(x, y, \xi) = 0. \tag{42c}$$

Consequently, the true (expected value) optimization problem (1) is approximated by the following so-called sample average approximation (SAA) problem:

$$\text{Min}_{x \in X} \hat{f}_N(x). \tag{43}$$

Let us denote by \hat{v}_N and \hat{S}_N the optimal value and the set of optimal solutions respectively of the SAA problem (43); let us denote by v^* and S^* the optimal value and the set of optimal solutions, respectively, of the true problem (1). Note that here the above optimal value function depends on $\omega \in \Omega$ through the random vector $\xi(\omega)$; therefore, we write $\vartheta(x, \xi)$ when we view it as a function of the two vector variables $x \in \mathbb{R}^n$ and $\xi \in \mathbb{R}^m$; we write $\vartheta(x, \xi(\omega))$ when we view it as a random function.

It is possible to apply an available statistical inference to the SAA problem (43) in a more or less straightforward way. By the (strong) law of large numbers (LLN), we have that, under standard conditions, e.g. if the sample is iid (independent identically distributed), $\hat{f}_N(x)$ converges pointwise (i.e., for every fixed $x \in X$) with probability one (w.p.1) to $f(x)$. Moreover, the following uniform convergence result holds (e.g. Proposition 7 in Ref. 22).

Proposition 5.1. Let V be a nonempty and compact subset of \mathbb{R}^n . Suppose that:

- (i) For a.e. ω , the function $\vartheta(\cdot, \xi(\omega))$ is continuous on V .
- (ii) There is a P -integrable function $\psi(\omega)$ such that $|\vartheta(x, \xi(\omega))| \leq \psi(\omega)$ for all $x \in V$ and $\omega \in \Omega$.
- (iii) The random sample ξ^1, \dots, ξ^N is iid.

Then, $f(x)$ is continuous on V and $\hat{f}_N(x)$ converges $f(x)$ w.p.1 uniformly on V , i.e.,

$$\sup_{x \in V} |\hat{f}_N(x) - f(x)| \rightarrow 0, \quad \text{w.p.1, as } N \rightarrow \infty.$$

We have the following consistency result.

Proposition 5.2. Suppose that there is a compact set $V \subset X$ such that the above assumptions (i)–(iii) hold and moreover that the set \hat{S}_N is nonempty and is contained in V w.p.1 for N large enough. Then, \hat{v}_N converges w.p.1 to v^* and $\sup_{x \in \hat{S}_N} \text{dist}(x, S^*) \rightarrow 0$ w.p.1 as $N \rightarrow \infty$.

In order to verify the above assumption (i), we can use the analysis of the previous section. By ad hoc methods, we may verify, for a.e. $\omega \in \Omega$, the existence of a selection

$$y_\omega(x) \in \mathbb{S}(x, \xi(\omega)), \quad \text{for all } x \in X.$$

Furthermore, making sure that $y_\omega(x)$ is locally unique, i.e., $y_\omega(x)$ is an isolated point of $\mathbb{S}(x, \xi(\omega))$, we obtain under mild boundedness conditions that $y_\omega(\cdot)$ is continuous at x . Finally, we would have to verify that $y_\omega(x)$ is a minimizer of $F(x, \cdot, \omega)$ over $\mathbb{S}(x, \xi(\omega))$. Of course, this holds automatically if $\mathbb{S}(x, \xi(\omega)) = \{y_\omega(x)\}$ is a singleton.

5.1. Exponential Rates of Convergence. By using the large deviations (LD) theory, it is also possible to give an estimate of the sample size N which is required in order to solve the true problem with a given accuracy. Assume that the sample is iid. For constants $\varepsilon > \delta \geq 0$, consider the set of ε -optimal solutions of the true problem and the set of δ -optimal solutions of the SAA problem (43). Let us make the following assumptions. Recall that the moment generating function of a random variable Z is defined as

$$M(t) := \mathbb{E} \left[e^{tZ} \right].$$

(B1) There exists a constant $\sigma > 0$ such that, for any $x', x \in X$, the moment generating function $M_{x',x}(t)$ of $\vartheta(x', \omega) - \vartheta(x, \omega) - \mathbb{E}[\vartheta(x', \omega) - \vartheta(x, \omega)]$ satisfies

$$M_{x',x}(t) \leq \exp \left[(1/2)\sigma^2 t^2 \right], \quad \forall t \in \mathbb{R}. \tag{44}$$

For fixed x', x , if the random variable

$$Z(\omega) := \vartheta(x', \omega) - \vartheta(x, \omega)$$

has a normal distribution, then the above Assumption (B1) holds with σ^2 being the variance of $Z(\omega)$. In general, condition (B1) means that the tail probabilities $\text{Prob}(|Z(\omega)| > t)$ are bounded from above by $O(1) \exp(-t^2/2\sigma^2)$; here $O(1)$, denotes a generic constant.

Further, suppose that X is a bounded subset of \mathbb{R}^n of diameter

$$D := \sup_{x', x \in X} \|x' - x\|$$

and that there exists a (measurable) function $\kappa : \Xi \rightarrow \mathbb{R}_+$ and $\gamma > 0$ such that the following relation holds:

$$|\vartheta(x, \xi) - \vartheta(\bar{x}, \xi)| \leq \kappa(\xi) \|x - \bar{x}\|^\gamma, \tag{45}$$

for all $x, \bar{x} \in X$ and all $\xi \in \Xi$. It follows by (45) that

$$|\hat{f}_N(x) - \hat{f}_N(\bar{x})| \leq N^{-1} \sum_{j=1}^N |\vartheta(x, \xi^j) - \vartheta(\bar{x}, \xi^j)| \leq \hat{\kappa}_N \|x - \bar{x}\|^\gamma, \tag{46}$$

where

$$\hat{\kappa}_N := N^{-1} \sum_{j=1}^N \kappa(\xi^j).$$

Let us make also the following assumption.

(B2) The moment generating function $M_\kappa(t) := \mathbb{E}[e^{t\kappa(\omega)}]$ of $\kappa(\xi)$ is finite valued for all t in a neighborhood of 0.

It follows that the expectation $L := \mathbb{E}[\kappa(\omega)]$ is finite; moreover, by the Cramér LD theorem, for any $L' > L$, there exists a positive constant $\beta = \beta(L')$ such that

$$P(\hat{\kappa}_N > L') \leq e^{-N\beta}. \tag{47}$$

Then, we have the following estimate:

$$N \geq \left[4\sigma^2/(\varepsilon - \delta)^2 \right] \left[n \left\{ \log D + \gamma^{-1} \log [2L'/(\varepsilon - \delta)] \right\} + \log (O(1)/\alpha) \right] \sqrt{\left[\beta^{-1} \log(2/\alpha) \right]}, \tag{48}$$

for the sample size which is required to solve the true problem with accuracy $\varepsilon > 0$ by solving the SAA problem with accuracy $\delta \in [0, \varepsilon)$. That is, if $\alpha \in (0, 1)$ is a given significance level and N satisfies (48), then with probability at least $1 - \alpha$ any δ -optimal solution of the SAA problem is an ε -optimal solution of the true problem. The distinctive feature of the estimate (48) is that the dimension n of the first stage problem enters it linearly. For a detailed derivation of this result and a further discussion of complexity of stochastic programs, we refer to Refs. 22 and 23.

References

1. BIRGE, J. R., and LOUVEAUX, F., *Introduction to Stochastic Programming*, Springer, New York, NY, 1997.
2. PATRIKSSON, M., and WYNTER, L., *Stochastic Mathematical Programs with Equilibrium Constraints*, Operations Research Letters, Vol. 25, pp. 159–167, 1999.

3. EVGRAFOV, A., and PATRIKSSON M., *On the Existence of Solutions to Stochastic Mathematical Programs with Equilibrium Constraints*, Journal of Optimization Theory and Applications, Vol. 121, pp. 65–76, 2004.
4. LIN, G. H., and FUKUSHIMA, M., *A Class of Mathematical Programs with Complementarity Constraints: Reformulations and Algorithms*, Technical Report 2003-10, Department of Applied Mathematics and Physics, Kyoto University, Kyoto, Japan, 2003.
5. BIRBIL, S. I., GÜRKAN, G., and LISTES, O., *Simulation-Based Solution of Stochastic Mathematical Programs with Complementarity Constraints: Sample-Path Analysis*, Research Report ERS-2004-016-LIS, Erasmus University, Rotterdam, Holland, 2004.
6. XU, H., *An Approach for Stochastic Stackelberg-Nash-Cournot Equilibrium*, Optimization (to appear).
7. XU, H., *An Implicit Programming Approach for a Class of Stochastic Mathematical Programs with Complementarity Constraints*, Preprint, 2005.
8. LUO, Z. Q., PANG, J. S., and RALPH, D., *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, UK, 1996.
9. SCHEEL, H., and SCHOLTES, S., *Mathematical Programs with Complementarity Constraints: Stationarity, Optimality, and Sensitivity*, Mathematics of Operations Research, Vol. 25, pp. 1–22, 2000.
10. ROBINSON, S. M., *Strongly Regular Generalized Equations*, Mathematics of Operations Research, Vol. 5, pp. 43–62, 1980.
11. RUSZCZYŃSKI, A., and SHAPIRO, A., Editors, *Stochastic Programming*, Handbooks in Operations Research and Management Science, North-Holland Publishing Company, Amsterdam, Holland, Vol. 10, 2003.
12. CASTAING, G., and VALADIER, M., *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Mathematics, Springer, Berlin, Germany, Vol. 580, 1977.
13. ROCKAFELLAR, R. T., and WETS, R. J. B., *Variational Analysis*, Springer Verlag, New York, NY, 1998.
14. AUCHMUTY, G., *Variational Principles for Variational Inequalities*, Numerical Functional Analysis and Optimization, Vol. 10, pp. 863–874, 1989.
15. FUKUSHIMA, M., *Equivalent Differentiable Optimization Problems and Descent Methods for Asymmetric Variational Inequality Problems*, Mathematical Programming, Vol. 53, pp. 99–110, 1992.
16. FACCHINEI, F., and PANG, J. S., *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer, New York, NY, 2003.
17. BONNANS, J. F., and SHAPIRO, A., *Perturbation Analysis of Optimization Problems*, Springer, New York, NY, 2000.
18. KINDERLEHRER, D., and STAMPACCHIA, G., *An Introduction to Variational Inequalities and Their Applications*, Academic Press, New York, NY, 1980.
19. SHAPIRO, A., *Sensitivity Analysis of Parametrized Variational Inequalities*, Mathematics of Operations Research, Vol. 30, pp. 76–91, 2005.
20. ROBINSON, S. M., *Generalized Equations and Their Solutions, Part 2: Applications to Nonlinear Programming*, Mathematical Programming Study, Vol. 19, pp. 200–221, 1982.

21. HU, X., and RALPH, D., *A Note on Sensitivity of Value Functions of Mathematical Programs with Complementarity Constraints*, *Mathematical Programming*, Vol. 93A, pp. 265–279, 2002.
22. SHAPIRO, A., *Monte Carlo Sampling Methods*, *Stochastic Programming*, Edited by A. Ruszczyński and A. Shapiro, *Handbooks in Operations Research and Management Science*, North-Holland Publishing Company, Amsterdam, Holland, Vol. 10, 2003.
23. SHAPIRO, A., and NEMIROVSKI, A., *On Complexity of Stochastic Programming Problems*; see [http:// www.optimization-online.org](http://www.optimization-online.org), 2004.