Open access • Journal Article • DOI:10.1007/S10957-005-7566-X

Stochastic Programming with Equilibrium Constraints — Source link

Alexander Shapiro

Institutions: Georgia Institute of Technology

Published on: 01 Jan 2006 - Journal of Optimization Theory and Applications (Kluwer Academic Publishers-Plenum Publishers)

Topics: Stochastic optimization, Stochastic programming, Stochastic approximation, Constrained optimization and Rate of convergence

Stochastic Programming with Equilibrium Constraints

Alexander Shapiro

January 17, 2005, revised

School of Industrial and Systems Engineering,
Georgia Institute of Technology,
Atlanta, Georgia 30332-0205, USA
e-mail: ashapiro@isye.gatech.edu

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 also discuss consistency and rates 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., [3, section 4.3]). 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.

The SMPEC problems of "here-and-now" type were discussed in [13],[6] and [11], for example. The main subject of [13] and [6] is an investigation of existence of optimal solutions of SMPEC problems, while in [11] an example of a variant of the newsboy (vendor) problem is discussed and an approach to a numerical solution of a class SMPEC problems is suggested. A convergence theory of discrete approximations of a different, from "here-and-now", type of SMPEC problems was discussed in [2]. Stochastic Stackelberg-Nash-Cournot equilibrium problems were formulated in a form of stochastic mathematical programs with complementarity constraints (SMPCC) and studied in [22], and a general form of SMPCC is discussed in [23]. There is also an extensive literature on deterministic MPEC problems. In that respect we may refer to [12] and [18], 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 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

$$Q^* := \left\{ z \in \mathbb{R}^s : z^T y \le 0,\ \forall\, y \in Q \right\}.$$

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 $\nabla_y G(x, y)$ its partial derivative, i.e., $s \times m$ Jacobian matrix, with respect to y. For a function $g : \mathbb{R}^m \to \mathbb{R}$ we denote by $\text{epi}(g) := \{(x, \alpha) : g(x) \le \alpha\}$ its epigraph, by $\nabla^2 g(y)$ its $m \times m$ Hessian matrix of second order partial derivatives. For a mapping $G(y) = (g_1(y), ..., g_s(y)) : \mathbb{R}^m \to \mathbb{R}^s$ and vector $d \in \mathbb{R}^m$, we denote

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

For $\varepsilon \ge 0$ 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 $f(\bar{x}) \le \inf_{x \in X} f(x) + \varepsilon$. For $a, b \in \mathbb{R}$, $a \vee b := \max\{a, b\}$.

2 The model

In this section we give a precise definition of the considered class of SMPEC problems. Consider the following optimization problem

$$\operatorname*{Min}_{x\in X} \big\{ f(x) := \mathbb{E}[\vartheta(x,\omega)] \big\}, \tag{2.1}$$

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

$$\vartheta(x,\omega) := \inf_{y\in\mathbb{S}(x,\omega)} F(x,y,\omega) \tag{2.2}$$

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

$$H(x,y,\omega) \in N_{K(x,\omega)}(y), \tag{2.3}$$

i.e., $y \in \mathbb{S}(x,\omega)$ iff (2.3) holds. Here $(\Omega, \mathcal{F}, P)$ is a probability space, $F : \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}$ is a real valued function, $K : \mathbb{R}^n \times \Omega \rightrightarrows \mathbb{R}^m$ is a set-valued mapping (multifunction), $H : \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}^m$, 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 (2.1)–(2.3) 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:

$$N_K(y) := \left\{ z \in \mathbb{R}^m : z^T(y'-y) \le 0, \ \ \forall\, y' \in K \right\}, \ \text{ if } y \in K, \tag{2.4}$$

and $N_K(y) = \emptyset$ if $y \notin K$. 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

$$K(x,\omega) := \big\{ y \in \mathbb{R}^m : G(x,y,\omega) \in Q \big\}, \tag{2.5}$$

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

$$N_{K(x,\omega)}(y) := \left\{ \nabla_y G(x,y,\omega)^T \lambda : \lambda \in N_Q(G(x,y,\omega)) \right\}, \tag{2.6}$$

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 (2.6), 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. If, moreover, 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 the definition, $\vartheta(x,\omega) = +\infty$ if the set $\mathbb{S}(x,\omega)$ is empty. Therefore, if for a given $x \in X$ the set $\mathbb{S}(x,\omega)$ is empty with a positive probability, then $f(x) = +\infty$. That is, implicitly the optimization in (2.1) 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 (2.6), variational inequality (2.3) takes the form of the following so-called generalized equations

$$-H(x,y,\omega) + \nabla_y G(x,y,\omega)^T \lambda = 0, \quad \lambda \in N_Q(G(x,y,\omega)). \tag{2.7}$$

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

$$N_Q(z) = \left\{ z^* \in Q^* : z^T z^* = 0 \right\}, \tag{2.8}$$

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

$$\lambda \in Q^*, \;\; G(x,y,\omega) \in Q \;\text{ and }\; \lambda^T G(x,y,\omega) = 0, \tag{2.9}$$

or equivalently iff $G(x,y,\omega) \in N_{Q^*}(\lambda)$. It follows that we can write generalized equations (2.7) as the variational inequality (cf., [14]):

$$\mathcal{H}(x,\zeta,\omega) \in N_{\mathcal{Q}}(\zeta), \tag{2.10}$$

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

$$\mathcal{H}(x,\zeta,\omega) := \left[\begin{array}{c} -H(x,y,\omega) + \nabla_y G(x,y,\omega)^T \lambda \\ G(x,y,\omega) \end{array} \right] \text{and } \; \mathcal{Q} := \mathbb{R}^m \times Q^*. \tag{2.11}$$

It can be observed that the set $\mathcal{Q}$ is a closed convex cone which does not depend on x and ω. The price of this simplification is that variational inequality (2.10) should be solved jointly in y and λ. We denote by $\overline{\mathbb{S}}(x,\omega) \subset \mathbb{R}^{m+s}$ the set of solutions of variational inequality (2.10), and for $y \in \mathbb{R}^m$ we denote by $\Lambda(x,y,\omega)$ the set of all λ satisfying equations (2.7). Note that $\Lambda(x,y,\omega)$ is nonempty iff $y \in \mathbb{S}(x,\omega)$.

By the above discussion, $\overline{\mathbb{S}}(x,\omega)$ is also the set of solutions of generalized equations (2.7), and the set $\mathbb{S}(x,\omega)$ is obtained by the projection of the set $\overline{\mathbb{S}}(x,\omega)$ onto $\mathbb{R}^m$. Therefore, $\vartheta(x,\omega)$ coincides with the optimal value of the following problem

$$\begin{array}{cl} \text{Min}_{y\in\mathbb{R}^m,\, \lambda\in\mathbb{R}^s} & F(x,y,\omega) \\ \text{subject to} & -H(x,y,\omega) + \nabla_y G(x,y,\omega)\lambda = 0, \\ & \lambda \in Q^*, \; G(x,y,\omega) \in Q, \; \lambda^T G(x,y,\omega) = 0. \end{array} \tag{2.12}$$

If, moreover, $H(x,y,\omega) = -\nabla_y h(x,y,\omega)$, where $h : \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}$ is a continuously differentiable in (x,y) function, then generalized equations (2.7) represent first-order (KKT) optimality conditions for the optimization problem

$$\underset{y\in\mathbb{R}^m}{\text{Min}}\, h(x,y,\omega) \;\text{ subject to } G(x,y,\omega) \in Q. \tag{2.13}$$

We refer to the set $\mathbb{S}(x,\omega)$ of solutions of the corresponding variational inequality as the set of stationary points of (2.13). That is, y is a stationary point of (2.13) if there exists $\lambda \in \mathbb{R}^s$ such that

$$\nabla_y h(x,y,\omega) + \nabla_y G(x,y,\omega)^T \lambda = 0, \quad \lambda \in N_Q(G(x,y,\omega)). \tag{2.14}$$

The set $\Lambda(x,y,\omega)$, of all λ satisfying (2.14), represents the set of Lagrange multipliers associated with y. It is well known that, under a constraint qualification, a locally optimal solution of (2.13) is also its stationary point. In case the set $\mathbb{S}(x,\omega)$ is given by the set of stationary points of an optimization problem, of the form (2.13), we refer to (2.1) as a Stochastic Bilevel Mathematical Programming (SBMP) problem.

Example 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 $G(x,y,\omega) = (g_1(x,y,\omega),...,g_s(x,y,\omega))$. Here the set

$$K(x,\omega) = \left\{ y : g_i(x,y,\omega) \le 0,\ i = 1,...,r,\ g_i(x,y,\omega) = 0,\ i = r+1,...,s \right\} \tag{2.15}$$

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

$$N_Q(z) = \{z^* \in \mathbb{R}^s : z_i^* \ge 0,\ i = 1,...,r,\ z_i^* z_i = 0,\ i = 1,...,r\}, \tag{2.16}$$

and $\nabla_y G(x,y,\omega)^T \lambda = \sum_{i=1}^s \lambda_i \nabla_y g_i(x,y,\omega)$, where $\lambda \in \mathbb{R}^s$, and hence generalized equations (2.7) take the form:

$$\begin{array}{l} -H(x,y,\omega) + \sum_{i=1}^s \lambda_i \nabla_y g_i(x,y,\omega) = 0, \\ g_i(x,y,\omega) \le 0,\ i = 1,...,r,\ g_i(x,y,\omega) = 0,\ i = r+1,...,s, \\ \lambda_i \ge 0,\ \lambda_i g_i(x,y,\omega) = 0,\ i = 1,...,r. \end{array} \tag{2.17}$$

Example 2 Suppose that for an $x \in X$, the set $\mathbb{S}(x,\omega)$ is given by the set of optimal solutions of the following optimization problem

$$\operatorname*{Min}_{y \in \mathbb{R}^m} q(\omega)^T y \ \text{ subject to } T(\omega)x + W(\omega)y = h(\omega),\ y \ge 0, \tag{2.18}$$

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 $(\Omega, \mathcal{F}, P)$. The above problem (2.18) is a linear programming problem. If, moreover, $F(x,y,\omega) := c^T x + q(\omega)^T y$, then the corresponding problem (2.1) becomes a two-stage stochastic linear program with recourse.

We have that $y(x,\omega)$ is an optimal solution of (2.18) iff $y(x,\omega)$ is a solution of the variational inequality

$$-q(\omega) \in N_{K(x,\omega)}(y), \tag{2.19}$$

where

$$K(x,\omega) := \big\{y \in \mathbb{R}^m : T(\omega)x + W(\omega)y = h(\omega),\ y \geq 0\big\}. \tag{2.20}$$

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 (2.18), or equivalently variational inequality (2.19), can be: empty, contain a single extreme point of $K(x,\omega)$, or 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 (2.18) is empty, in which case its optimal value is $+\infty$, or it may happen that problem (2.18) is unbounded from below, i.e., its optimal value is $-\infty$. Note that in our framework in both cases we assign value $+\infty$ to the corresponding function $\vartheta(x,\omega)$.

The dual of problem (2.18) is

$$\operatorname*{Max}_{\lambda} \lambda^T \left(h(\omega) - T(\omega)x\right) \ \text{ subject to } W(\omega)^T\lambda \leq q(\omega). \tag{2.21}$$

The optimality conditions (2.14) take here the form:

$$y \geq 0,\ W(\omega)^T\lambda - q(\omega) \leq 0,\ y^T\left(W(\omega)^T\lambda - q(\omega)\right) = 0,\ T(\omega)x + W(\omega)y - h(\omega) = 0, \tag{2.22}$$

and can be written as variational inequality (2.10) with $\zeta = (y,\lambda)$ and

$$\mathcal{H}(x,\zeta,\omega) := \begin{bmatrix} W(\omega)^T\lambda - q(\omega) \\ T(\omega)x + W(\omega)y - h(\omega) \end{bmatrix} \text{ and } \mathcal{Q} := \mathbb{R}^m_+ \times \mathbb{R}^s. \tag{2.23}$$

3 Properties of SMPEC problems

In this section we discuss some basic properties of SMPEC problem (2.1) with the function $\vartheta(x,\omega)$ given by the optimal value of the problem (2.12). To some extent such SMPEC problems can be studied in the framework of two-stage stochastic programming problems with recourse (cf., [17, Chapters 1 and 2]).

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 either $\mathbb{E}\left[(\vartheta_x(\omega))_+\right] < +\infty$ or $\mathbb{E}\left[(-\vartheta_x(\omega))_+\right] < +\infty$. For a thorough discussion of the following measurability concepts we may refer to [5] and [16]. When considering space $\mathbb{R}^n \times \Omega$ we always equip it 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 $h : \mathbb{R}^n \times \Omega \to \overline{\mathbb{R}}$ is said to be *random lower semicontinuous* if the epigraphical mapping $\omega \mapsto \operatorname{epi} h(\cdot,\omega)$ is closed valued and measurable. (Random lower semicontinuous functions are called *normal integrands* in [16].) It is said that a mapping $G : \mathbb{R}^n \times \Omega \to \mathbb{R}^m$ 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., [16, Theorem 14.37]).

Theorem 1 *Let $g : \mathbb{R}^n \times \Omega \to \overline{\mathbb{R}}$ 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 function $F(x, y, \omega)$ are Carathéodory mappings, i.e., continuous in (x, y) and measurable in ω.

Consider variational inequality (2.10). By assumption (A1) we have that mapping $\mathcal{H}(x, \zeta, \omega)$, defined in (2.11), is a Carathéodory mapping. Variational inequality (2.10) can be transformed into an optimization problem as follows. Consider the following (regularized) gap function:

$$\gamma(x, \zeta, \omega) := \sup_{\zeta' \in \mathcal{Q}} \left\{ \mathcal{H}(x, \zeta, \omega)^T (\zeta' - \zeta) - \tfrac{1}{2} \|\zeta - \zeta'\|^2 \right\}. \tag{3.1}$$

Regularized gap functions were introduced in [1] and [8] (see also [7, section 10.2.1] 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 (2.10), i.e., $\bar{\zeta} \in \overline{\mathbb{S}}(x, \omega)$, iff $\gamma(x, \bar{\zeta}, \omega) = 0$. Since $\gamma(x, \zeta, \omega) \geq 0$ for any $\zeta \in \mathcal{Q}$, it follows that $\bar{\zeta} \in \overline{\mathbb{S}}(x, \omega)$ iff $\bar{\zeta}$ is an optimal solution of the problem

$$\operatorname*{Min}_{\zeta \in \mathcal{Q}} \gamma(x, \zeta, \omega) \text{ subject to } \gamma(x, \zeta, \omega) = 0. \tag{3.2}$$

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

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

$$\vartheta_x(\omega) = \inf_{(y,\lambda) \in \overline{\mathbb{S}}(x,\omega)} F(x, y, \omega). \tag{3.3}$$

Since, for a given x, the multifunction $\overline{\mathbb{S}}(x, \cdot)$ is closed valued and measurable and 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 (2.1).

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

$$\tilde{f}(x) := \inf_{\zeta(\cdot) \in \overline{\mathbb{S}}(x,\cdot)} \mathbb{E}[F(x, y(\omega), \omega)], \tag{3.4}$$

where $\zeta(\cdot) \in \overline{\mathbb{S}}(x, \cdot)$ means that the optimization is performed over all measurable selections $\zeta(\omega)$ of $\overline{\mathbb{S}}(x, \omega)$. By the definition, $\tilde{f}(x) := +\infty$ if the set $\overline{\mathbb{S}}(x, \omega)$ is empty with positive probability. We have then that $f(x) = \tilde{f}(x)$ for every $x \in X$. Indeed, assume that $\overline{\mathbb{S}}(x, \omega)$ is nonempty for a.e. $\omega \in \Omega$ (otherwise $f(x) = \tilde{f}(x) = +\infty$). For any measurable selection $\zeta(\omega) \in \overline{\mathbb{S}}(x, \omega)$ 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 ([16, Theorem 14.5]) for any $\varepsilon > 0$ there exists a measurable selection $\zeta(\omega) \in \overline{\mathbb{S}}(x, \omega)$ 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.

Problem (2.1) is equivalent to the problem

$$\operatorname*{Min}_{x \in X,\, \zeta(\cdot) \in \overline{\mathbb{S}}(x,\cdot)} \mathbb{E}[F(x, y(\omega), \omega)], \tag{3.5}$$

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 optimization in (3.5) is performed over $x \in X$ and measurable $\zeta(\omega) = (y(\omega), \lambda(\omega))$ which satisfy the feasibility constraints of problem (2.12). In particular, suppose that $\Omega = \{\omega_1, ..., \omega_K\}$ is finite with respective probabilities $p_1, ..., p_K$. Then we can write problem (2.1) in the following equivalent form

$$\begin{array}{ll} \operatorname*{Min}_{x \in X,\, y_1, \dots, y_K} & \sum_{k=1}^K p_k F(x, y_k, \omega_k) \\ \text{subject to} & -H(x, y_k, \omega_k) + \nabla_y G(x, y_k, \omega_k)\lambda_k = 0,\ k = 1, ..., K, \\ & \lambda_k \in Q^*,\ G(x, y_k, \omega_k) \in Q,\ \lambda_k^T G(x, y_k, \omega_k) = 0,\ k = 1, ..., K. \end{array} \tag{3.6}$$

4 Properties of the expectation function

In this section we discuss 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 (2.12).

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 $\psi : \Omega \to \mathbb{R}$ is P-integrable if $\mathbb{E}[|\psi|] < +\infty$.) Then we have by Fatou's lemma that

$$\liminf_{x \to \bar{x}} \mathbb{E}\big[\vartheta(x, \omega)\big] \geq \mathbb{E}\big[\liminf_{x \to \bar{x}} \vartheta(x, \omega)\big]. \tag{4.1}$$

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 lower semicontinuity of $f(\cdot)$ is reduced to studying 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., [4, example 4.1]). Therefore we will 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 generalized equations (2.7) and that $\overline{\mathbb{S}}(x,\omega)$ coincides with the feasible set of problem (2.12).

(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 cones Q and Q^* are closed, the multifunction $x \mapsto \overline{\mathbb{S}}(x,\omega)$ is *closed*. That is, if $x_n \to \bar{x}$, $\zeta_n \in \overline{\mathbb{S}}(x_n,\omega)$ and $\zeta_n \to \bar{\zeta}$, then $\bar{\zeta} \in \overline{\mathbb{S}}(\bar{x},\omega)$. The following result follows easily by compactness arguments and closedness of the multifunction $\overline{\mathbb{S}}(\cdot,\omega)$, and is quite well known.

> *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 then the following result.

Proposition 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. It follows then that problem (2.1) has an optimal solution provided it has a nonempty and bounded level set

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

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

In order to ensure continuity of $f(x)$ one needs to investigate 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., [17, p.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 $\bar{x}$, corresponds a $y_\omega(x) \in \mathbb{S}(x,\omega)$ such that $y_\omega(x)$ tends to $\bar{y}$ as $x \to \bar{x}$.

The above assumption implies that the set $\mathbb{S}(x,\omega)$, and hence the set $\overline{\mathbb{S}}(x,\omega)$, is nonempty for all x in a neighborhood of $\bar{x}$. We have then the following result (e.g., [4, Proposition 4.4]).

Suppose that assumptions (A2) *and* (A3) *are satisfied and assumptions* (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 2 *Suppose that assumptions* (A2)–(A4) *hold for a.e.* $\omega \in \Omega$, *and there exists 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) actually is formed by two parts, namely:

(A4a) The set $\mathbb{S}(x, \omega)$, or equivalently of $\overline{\mathbb{S}}(x, \omega)$, 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 $x \to \bar{x}$.

Nonemptiness of $\mathbb{S}(x, \omega)$ can be ensured by various conditions. For example, consider the optimization problem (2.13). It has an optimal solution $\bar{y}_\omega(x)$ provided the corresponding inf-compactness condition holds. If, moreover, 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 [10] and [7, section 2.2]. For example, if the set $K = K(x, \omega)$ is convex, then variational inequality (2.3) has a solution if there exists $y^* \in K$ such that the set

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

is bounded (cf., [7, Proposition 2.2.3]). (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 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 $x \to \bar{x}$. It is also of interest to estimate a 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

$$\|y_\omega(x) - \bar{y}\| \leq c \|x - \bar{x}\|^\gamma \tag{4.2}$$

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 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 notation $H(\cdot) = H(\bar{x}, \cdot, \omega)$ and

$G(\cdot) = G(\bar{x}, \cdot, \omega)$, etc. It is said that Robinson's constraint qualification, for the system $G(y) \in Q$, holds at $\bar{y}$ if

$$[\nabla G(\bar{y})]\mathbb{R}^m + T_Q(G(\bar{y})) = \mathbb{R}^s. \tag{4.3}$$

Note that if (4.3) holds, then the set $\Lambda(\bar{y}) = \Lambda(\bar{x}, \bar{y}, \omega)$, of the corresponding Lagrange multipliers, is bounded. Note also that if the system is defined by a finite number of constraints as in example 1, then Robinson's constraint qualification coincides with the Mangasarian-Fromovitz constraint qualification.

Recall that the cone

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

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

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

of Lagrange multipliers. We have the following result (cf., [20, Theorems 3.1,3.2,4.3 and 5.1]).

Proposition 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, Robinson's constraint qualification* (4.3) *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 $d \in \mathbb{R}^n$):*

$$\nabla H(\bar{y}) d \in \operatorname{conv}\left\{\cup_{\lambda \in \Lambda^*(\bar{y}, d)} \left[\nabla^2 G(\bar{y}) d\right]^T \lambda\right\} + N_{C(\bar{y})}(d) \ \text{ implies } d = 0. \tag{4.4}$$

Also if the above condition (4.4) *holds and $H(x, y, \omega)$ 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 in addition 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 (4.4) to situations where the cone Q is not polyhedral (see [20]), although the analysis becomes more involved. Of course, if the mapping $G(y)$ is affine, then $\nabla^2 G(y) = 0$, and hence the system in the left hand side of (4.4) reduces to $\nabla H(\bar{y}) d \in N_{C(\bar{y})}(d)$. Condition (4.4) 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 (2.13), condition (4.4) is equivalent (under the assumptions of the above proposition) to the quadratic growth condition: there exists $c > 0$ and a neighborhood V of $\bar{y}$ such that

$$h(y) \geq h(\bar{y}) + c\|y - \bar{y}\|^2, \ \ \forall\, y \in K(\bar{x}, \omega) \cap V. \tag{4.5}$$

In the general case of $G(x, y, \omega)$ depending on x, in order to ensure locally upper *Lipschitz* continuity of $\mathbb{S}(\cdot, \omega)$, at $\bar{x}$ for $\bar{y}$, one needs conditions which are considerably stronger than assumption (4.4). For optimization problems, for example, the quadratic growth condition (4.5) does not imply locally upper Lipschitz continuity of $\mathbb{S}(\cdot, \omega)$, and a strong form of second order optimality conditions is needed (cf., [15]).

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). It follows then from (4.2) that, for $x \in W$,

$$|\vartheta(x, \omega) - \vartheta(\bar{x}, \omega)| \leq \kappa(\omega) \|x - \bar{x}\|^\gamma, \tag{4.6}$$

where $\kappa(\omega) := \ell(\omega) c(\omega)$. The above inequality (4.6), in turn, implies that, for $x \in W$,

$$|f(x) - f(\bar{x})| \leq L \|x - \bar{x}\|^\gamma, \tag{4.7}$$

where $L := \mathbb{E}[\kappa(\omega)]$. Of course, the estimate (4.7) is meaningful only if the expectation $\mathbb{E}[\kappa(\omega)]$, i.e., the constant L, is finite.

Lipschitz continuity and differentiability properties of the optimal value function $\vartheta(\cdot, \omega)$ were also studied in [9] in the case when the set $K(x, \omega)$ is defined by a finite number of constraints, as in example 1.

Let us discuss now 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 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 (e.g., [17, Chapter 2, Proposition 2]):

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 $\kappa : \Omega \to \mathbb{R}_+$ such that (4.6) holds with $\gamma = 1$, i.e.,

$$|\vartheta(x, \omega) - \vartheta(\bar{x}, \omega)| \leq \kappa(\omega) \|x - \bar{x}\| \tag{4.8}$$

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

$$f'(\bar{x}, p) = \mathbb{E}\left[\vartheta'_\omega(\bar{x}, p)\right], \quad \forall p \in \mathbb{R}^n. \tag{4.9}$$

Of course, (4.7) follows from (4.8), with $\gamma = 1$, and 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 $x \to \bar{x}$. Then since $\vartheta(x,\omega) = F(x, y_\omega(x), \omega)$, we obtain by the chain rule that

$$\vartheta'_\omega(\bar{x}, p) = \nabla_y F(\bar{x}, \bar{y}, \omega)^T y'_\omega(\bar{x}, p), \tag{4.10}$$

provided that $F(\bar{x}, \cdot, \omega)$ is continuously differentiable and the directional derivative $y'_\omega(\bar{x}, p)$ does exist. Concerning directional differentiability of $y_\omega(x)$ we have the following result ([20, Corollary 4.1]).

Proposition 4 *Suppose that the assumptions of Proposition* 3 *and the assumption* (A4) *hold 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*

$$\nabla_x H(\bar{x}, \bar{y}, \omega)p + \nabla_y H(\bar{x}, \bar{y}, \omega)d \in \operatorname{conv}\left\{\cup_{\lambda\in\Lambda^*(\bar{y},d)} \left[\nabla^2 G(\bar{y})d\right]^T \lambda\right\} + N_{C(\bar{y})}(d) \tag{4.11}$$

has 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 (4.11), we obtain that condition (ii), in the above proposition, implies condition (4.4) of Proposition 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 briefly discuss the Monte Carlo sampling approach to such a discretization. Assume that all involved random data depend on a random vector $\xi(\omega)$, where $\xi(\cdot) : \Omega \to \mathbb{R}^d$ 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, $F : \mathbb{R}^n \times \mathbb{R}^m \times \Xi \to \mathbb{R}$, and suppose that the objective function in (2.2) 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 $\xi^1, ..., \xi^N$ be a random sample of $\xi(\omega)$. Define

$$\hat{f}_N(x) := \frac{1}{N}\sum_{j=1}^{N} \vartheta(x, \xi^j), \tag{5.1}$$

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

$$\begin{array}{rl} \operatorname{Min}_{y\in\mathbb{R}^m,\, \lambda\in\mathbb{R}^s} & F(x, y, \xi) \\ \text{subject to} & -H(x, y, \xi) + \nabla_y G(x, y, \xi)\lambda = 0, \\ & \lambda \in Q^*,\ G(x, y, \xi) \in Q,\ \lambda^T G(x, y, \xi) = 0. \end{array} \tag{5.2}$$

Consequently, the "true" (expected value) optimization problem (2.1) is approximated by the following, so-called sample average approximation (SAA), problem:

$$\operatorname*{Min}_{x \in X} \hat{f}_N(x). \tag{5.3}$$

Let us denote by $\hat{v}_N$ and $\widehat{S}_N$ the optimal value and the set of optimal solutions, respectively, of the SAA problem (5.3), and by v^* and S^* the optimal value and the set of optimal solutions, respectively, of the true problem (2.1). Note that here the above optimal value function depends on $\omega \in \Omega$ through the random vector $\xi(\omega)$, and therefore we write $\vartheta(x, \xi)$ when we view it as a function of two vector variables $x \in \mathbb{R}^n$ and $\xi \in \mathbb{R}^m$, and 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 (5.3) 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., [19, Proposition 7]).

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 $|\vartheta(x, \xi(\omega))| \le \psi(\omega)$ for all $x \in V$ and $\omega \in \Omega$,*

(iii) *The random sample $\xi^1, ..., \xi^N$ is iid.*

Then $f(x)$ is continuous on V, and $\hat{f}_N(x)$ converges to $f(x)$ w.p.1 uniformly on V, i.e., $\sup_{x \in V} |\hat{f}_N(x) - f(x)| \to 0$ w.p.1 as $N \to \infty$.

We have then the following consistency result.

Proposition 5 *Suppose that there is a compact set $V \subset X$ such that the above assumptions* (i)–(iii) *hold and, moreover, the set $\widehat{S}_N$ 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* $\sup_{x \in \widehat{S}_N} \operatorname{dist}(x, S^*) \to 0$ *w.p.1 as $N \to \infty$.*

In order to verify the above assumption (i) we can use analysis of the previous section. By ad hoc methods we may verify, for a.e. $\omega \in \Omega$, existence of a selection $y_\omega(x) \in \mathbb{S}(x, \xi(\omega))$ 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 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 (5.3). 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 constant $\sigma > 0$ such that for any $x', x \in X$, the moment generating function $M_{x',x}(t)$ of $\vartheta(x', \omega) - \vartheta(x, \omega) - \mathbb{E}\big[\vartheta(x', \omega) - \vartheta(x, \omega)\big]$ satisfies:

$$M_{x',x}(t) \leq \exp\left(\tfrac{1}{2}\sigma^2 t^2\right), \quad \forall\, t \in \mathbb{R}. \tag{5.4}$$

If, for fixed x', x, 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 tail probabilities $\text{Prob}\big(|Z(\omega)| > t\big)$ are bounded from above by $O(1)\exp\left(-\frac{t^2}{2\sigma^2}\right)$ (here $O(1)$ denotes a generic constant).

Suppose, further, that X is a bounded subset of $\mathbb{R}^n$ of diameter

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

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

$$|\vartheta(x, \xi) - \vartheta(\bar{x}, \xi)| \leq \kappa(\xi) \|x - \bar{x}\|^{\gamma} \tag{5.5}$$

holds for all $x, \bar{x} \in X$ and all $\xi \in \Xi$. It follows by (5.5) that

$$\big|\hat{f}_N(x) - \hat{f}_N(\bar{x})\big| \leq N^{-1} \sum_{j=1}^{N} \big|\vartheta(x, \xi^j) - \vartheta(\bar{x}, \xi^j)\big| \leq \hat{\kappa}_N \|x - \bar{x}\|^{\gamma}, \tag{5.6}$$

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 $M_\kappa(t) := \mathbb{E}\left[e^{t\kappa(\omega)}\right]$ 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, and moreover, by Cramér's LD Theorem that for any $L' > L$ there exists a positive constant $\beta = \beta(L')$ such that

$$P\left(\hat{\kappa}_N > L'\right) \leq e^{-N\beta}. \tag{5.7}$$

We have then the following estimate

$$N \geq \frac{4\sigma^2}{(\varepsilon - \delta)^2} \left[n \left(\log D + \gamma^{-1} \log \frac{2L'}{\varepsilon - \delta} \right) + \log \left(\frac{O(1)}{\alpha} \right) \right] \bigvee \left[\beta^{-1} \log \left(\frac{2}{\alpha} \right) \right] \tag{5.8}$$

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 (5.8), 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 (5.8) 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 [19, 21].

References

[1] G. Auchmuty, Variational principles for variational inequalities, *Numerical Functional Analysis and Optimization*, 10 (1989), 863–874.

[2] S.I. Birbil, G. Gürkan and O. Listes, Simulation-based solution of stochastic mathematical programs with complementarity constraints: sample-path analysis, Research Report ERS-2004-016-LIS, Erasmus University, 2004.

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

[4] J.F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[5] C. Castaing and M. Valadier, *Convex Analysis and Measurable Multifunctions* , Lecture Notes in Mathematics, 580, Springer, Berlin, 1977.

[6] A. Evgrafov and M. Patriksson, On the existence of solutions to stochastic mathematical programs with equilibrium constraints, *Journal of Optimization Theory and Applications*, 121 (2004), 65-76.

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

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

[9] X. Hu and D. Ralph, A note on sensitivity of value functions of mathematical programs with complementarity constraints, *Mathematical Programming, Ser. A*, 93 (2002), 265–279.

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

[11] G.-H. Lin and M. Fukushima, A class of stochastic mathematical programs with complementarity constraints: reformulations and algorithms, Technical Report 2003-10, Department of Applied Mathematics and Physics, Kyoto University, 2003.

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

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

[14] S.M. Robinson, Strongly regular generalized equations, *Mathematics of Operations Research*, 5 (1980), 43-62.

[15] S.M. Robinson, Generalized equations and their solutions, Part II: Applications to nonlinear programming, *Mathematical Programming Study*, 19 (1982), 200-221.

[16] R.T. Rockafellar and R.J.-B. Wets, *Variational Analysis*, Springer-Verlag, New York, 1998.

[17] A. Ruszczyński and A. Shapiro, Eds., *Stochastic Programming*, in: Handbooks in OR & MS, Vol. 10, North-Holland Publishing Company, Amsterdam, 2003.

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

[19] A. Shapiro, Monte Carlo Sampling Methods, in A. Rusczyński and A.Shapiro, editors, *Stochastic Programming*, Handbooks in OR & MS, Vol. 10, North-Holland Publishing Company, Amsterdam, 2003.

[20] A. Shapiro, Sensitivity analysis of parameterized variational inequalitis, *Mathematics of Operations Research*, 30 (2005), 76-91.

[21] A. Shapiro and A. Nemirovski, On complexity of stochastic programming problems, *E-print available at:* `http://www.optimization-online.org`, 2004.

[22] H. Xu, An approach for stochastic Stackelberg-Nash-Cournot equilibrium, *Optimization*, to appear.

[23] H. Xu, An implicit programming approach for a class of stochastic mathematical programs with complementarity constraints, *preprint*.