Optimality Conditions and Combined Monte Carlo Sampling and Penalty Method for Stochastic Mathematical Programs with Complementarity Constraints and Recourse*

Gui-Hua Lin† and Masao Fukushima‡

June 2005, Revised March 2006

Abstract. In this paper, we consider a new formulation for stochastic mathematical programs with complementarity constraints and recourse. We show that the new formulation is equivalent to a smooth semi-infinite program that does no longer contain recourse variables. Optimality conditions for the problem are deduced and connections among the conditions are investigated. Then, we propose a combined Monte Carlo sampling and penalty method for solving the problem, and examine the limiting behavior of optimal solutions and stationary points of the approximation problems.

Key words. Stochastic mathematical program with complementarity constraints, here-and-now, recourse, semi-infinite programming, optimality conditions, Monte Carlo sampling, penalty method.

1 Introduction

Recently, stochastic mathematical programs with equilibrium constraints (SMPECs) have been receiving much attention in the optimization world [1,13–15,17,21,25–27]. In particular, Lin et al. [13] introduced two kinds of SMPECs: One is the *lower-level wait-and-see* model, in which the upper-level decision is made before a random event is observed, while a lower-level decision is made after a random event is observed. The other is the *here-and-now* model that requires us to make all decisions before a random event is observed. Lin and Fukushima [14,15,17] suggested a smoothing penalty method and a regularization method, respectively, for a special class of here-and-now problems. Shapiro and Xu [25–27] discussed the sample average approximation and

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

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

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

implicit programming approaches for the lower-level wait-and-see problems. In addition, Birbil et al. [1] considered an SMPEC in which both the objective and constraints involve expectations.

In [13], the here-and-now problem is formulated as follows:

$$\begin{aligned} \min \quad & \mathbb{E}[f(x,y,\omega) + d^T z(\omega)] \\ \text{s.t.} \quad & g(x) \le 0, \ \ h(x) = 0, \\ & 0 \le y \perp (F(x,y,\omega) + z(\omega)) \ge 0, \\ & z(\omega) \ge 0, \quad \omega \in \Omega \ \ \text{a.s.}, \end{aligned} \tag{1.1}$$

where $f : \Re^{n+m} \times \Omega \to \Re$, $g : \Re^n \to \Re^{s_1}$, $h : \Re^n \to \Re^{s_2}$, and $F : \Re^{n+m} \times \Omega \to \Re^m$ are functions, $\mathbb{E}$ means expectation with respect to the random variable $\omega \in \Omega$, the symbol $\perp$ means the two vectors are perpendicular to each other, "a.s." is the abbreviation for "almost surely" under the given probability measure, $z(\omega)$ is a recourse variable, and $d \in \Re^m$ is a constant vector with positive elements. Moreover, x denotes the upper-level decision, y represents the lower-level decision, and both the decisions x and y need to be made at once, before ω is observed. We suppose that all functions involved are continuous and, particularly, f and F are continuously differentiable with respect to (x, y), g and h are continuously differentiable with respect to x.

Lin et al. [13, 15, 17] considered the case where the function F is affine and the underlying sample space Ω is discrete and finite. In this paper, we consider a general case, i.e., F is nonlinear and Ω is a compact subset of $\Re^l$. A general strategy for SMPECs with infinitely many samples is to discretize the problem by some kind of sampling selection methods, which means the approximation problems are still MPECs [14]. The strategy of this paper is, in contrast, to solve some standard nonlinear programs as approximations of the original SMPEC.

The main contributions of the paper can be stated as follows. We note that problem (1.1) has the following difficulties:

- The problem contains recourse, which is a function of ω, and an expectation. Both of them may cause computational difficulty in general.
- Because of the presense of complementarity constraints, problem (1.1) fails to satisfy a standard constraint qualification at any feasible point [5].

We will get rid of the recourse variables. To this end, we first consider the following formulation of SMPECs with recourse, which slightly differs from (1.1):

$$\begin{aligned} \min \quad & \mathbb{E}[f(x,y,\omega) + \sigma \|z(\omega)\|^2] \\ \text{s.t.} \quad & g(x) \le 0, \ \ h(x) = 0, \\ & 0 \le y \perp (F(x,y,\omega) + z(\omega)) \ge 0, \\ & z(\omega) \ge 0, \quad \omega \in \Omega \ \ \text{a.s.}, \end{aligned} \tag{1.2}$$

where $\sigma > 0$ is a weight constant. We can show that problem (1.2) is equivalent to

$$\begin{aligned} \min \quad & \mathbb{E}[f(x,y,\omega) + \sigma\|u(x,y,\omega)\|^2] \\ \text{s.t.} \quad & g(x) \le 0,\ h(x) = 0,\ y \ge 0, \\ & y \circ F(x,y,\omega) \le 0, \quad \omega \in \Omega \ \text{ a.s.}, \end{aligned} \tag{1.3}$$

where $u : \Re^{n+m} \times \Omega \to \Re^m$ is defined by

$$u(x,y,\omega) := \max\{-F(x,y,\omega),\ 0\} \tag{1.4}$$

and $\circ$ denotes the Hadamard product, i.e., $y \circ F(x,y,\omega) := (y_1 F_1(x,y,\omega), \cdots, y_m F_m(x,y,\omega))^T$. See the appendix for a proof of the equivalence between (1.2) and (1.3). Problem (1.3) does no longer contain recourse variables. The reasons we consider (1.2) instead of (1.1) are stated as follows:

- Since both $d^T z(\omega)$ and $\sigma\|z(\omega)\|^2$ serve as a penalty term for the possible violation of the complementarity constraint $0 \le y \perp F(x,y,\omega) \ge 0$, problems (1.1) and (1.2) are essentially the same.
- The quadratic penalty $\sigma\|z(\omega)\|^2$ yields the equivalent problem (1.3) that has a differentiable objective function, but the linear penalty $d^T z(\omega)$ does not.

Note that problem (1.3) is actually a semi-infinite programming problem with a large number of complementarity-like constraints and it also involves an expectation in the objective function. Therefore, problem (1.3) is generally more difficult to handle than an ordinary semi-infinite programming problem. Firstly, we discuss the optimality conditions for the problems and investigate their connections. Then, we make use of a Monte Carlo sampling method to handle the expectation and propose a penalty technique to deal with the complementarity-like constraints. We also examine the limiting behavior of optimal solutions and stationary points of the approximation problems.

The following notations are used in the paper. For any vectors a and b of the same dimension, both $\max\{a,b\}$ and $\min\{a,b\}$ are understood to be taken componentwise. For a given function $c : \Re^s \to \Re^{s'}$ and a vector $t \in \Re^s$, $\nabla c(t)$ is the transposed Jacobian of c at t and $\mathcal{I}_c(t) := \{i \mid c_i(t) = 0\}$ stands for the active index set of c at t. In addition, e_i denotes the unit vector whose ith element is one.

2 Optimality Conditions

We first consider the semi-infinite programming problem (1.3). In the literature on semi-infinite programming, it is often assumed that there are a finite number of active constraints at a solution (see, e.g., a survey paper [11]). However, the above assumption does not hold in problem (1.3) in general. For example, if $y_i = 0$ for some index i, there must be infinitely many active constraints at the point. This indicates that problem (1.3) is more difficult to deal with than an ordinary semi-infinite programming problem. We define the stationarity for problem (1.3) as follows.

Let (x^*, y^*) be a local optimal solution of problem (1.3) and $\bar{\Omega}$ be the largest subset of Ω such that (x^*, y^*) is feasible to

$$\begin{aligned} & g(x) \le 0,\ \ h(x) = 0,\ \ y \ge 0, \\ & y \circ F(x, y, \omega) \le 0, \quad \forall \omega \in \bar{\Omega}. \end{aligned}$$

Let p denote the probability measure on Ω. It follows from the feasibility of (x^*, y^*) in (1.3) that $p(\Omega \setminus \bar{\Omega}) = 0$. This indicates that, for any integrable function ξ defined on Ω, there must hold

$$\int_{\Omega} \xi(\omega)\, dp = \int_{\bar{\Omega}} \xi(\omega)\, dp$$

and conversely, for any integrable function ξ defined on $\bar{\Omega}$, we can extend its definition to Ω such that the above condition holds. For simplicity, we suppose $\bar{\Omega} = \Omega$ in the following.

Let B be an arbitrary measurable subset of $\bar{\Omega}$ with null probability measure. We consider the problem

$$\begin{aligned} \min \quad & \mathbb{E}[f(x, y, \omega) + \sigma \|u(x, y, \omega)\|^2] \\ \text{s.t.} \quad & g(x) \le 0,\ \ h(x) = 0,\ \ y \ge 0, \\ & y \circ F(x, y, \omega) \le 0, \quad \forall \omega \in \Omega \setminus B. \end{aligned} \tag{2.1}$$

Since any feasible solution of problem (2.1) must be feasible to (1.3), the point (x^*, y^*) is also a local optimal solution of problem (2.1).

We denote the feasible region of problem (2.1) by $\mathscr{F}_B$ and, in addition, we use $\mathcal{N}_{\mathscr{F}_B}(x^*, y^*)$ and $\mathcal{T}_{\mathscr{F}_B}(x^*, y^*)$ to stand for the *normal cone* and the *tangent cone* of the set $\mathscr{F}_B$ at the point (x^*, y^*), respectively. Then we have

$$-\nabla_{(x,y)} \mathbb{E}\Big[f(x^*, y^*, \omega) + \sigma \|u(x^*, y^*, \omega)\|^2\Big] \ \in\ \mathcal{N}_{\mathscr{F}_B}(x^*, y^*). \tag{2.2}$$

Suppose that there exists an integrable function $\eta : \Omega \to [0, +\infty]$ such that

$$\Big\|\nabla_{(x,y)} f(x, y, \omega) - 2\sigma \nabla_{(x,y)} F(x, y, \omega) u(x, y, \omega)\Big\| \ \le\ \eta(\omega) \tag{2.3}$$

holds for any (x, y, ω). It then follows from (1.4) along with the corollary in §7-4 of [3] that

$$\nabla_{(x,y)} \mathbb{E}\Big[f(x, y, \omega) + \sigma \|u(x, y, \omega)\|^2\Big] = \mathbb{E}\Big[\nabla_{(x,y)} f(x, y, \omega) - 2\sigma \nabla_{(x,y)} F(x, y, \omega) u(x, y, \omega)\Big].$$

Note that $\mathcal{N}_{\mathscr{F}_B}(x^*, y^*)$ is the dual cone of $\mathcal{T}_{\mathscr{F}_B}(x^*, y^*)$ [22], that is, $\mathcal{N}_{\mathscr{F}_B}(x^*, y^*) = [\mathcal{T}_{\mathscr{F}_B}(x^*, y^*)]^*$. We then have from (2.2) and the arbitrariness of B that

$$-\mathbb{E}\Big[\nabla_{(x,y)} f(x^*, y^*, \omega) - 2\sigma\nabla_{(x,y)} F(x^*, y^*, \omega) u(x^*, y^*, \omega)\Big] \in \bigcap_{\substack{B\subset\Omega \\ p(B)=0}} \Big[\mathcal{T}_{\mathscr{F}_B}(x^*, y^*)\Big]^*. \tag{2.4}$$

In what follows, we denote $\mathcal{I}_Y^* := \{i \mid y_i^* = 0\}$ and $\Omega_i := \{\omega \in \Omega \mid F_i(x^*, y^*, \omega) = 0\}$ for each $i \notin \mathcal{I}_Y^*$. Consider the cone

$$\mathcal{C}_{\mathscr{F}_B}(x^*, y^*) := \left\{ \begin{pmatrix} dx \\ dy \end{pmatrix} \in \Re^{n+m} \;\middle|\; \begin{array}{ll} (\nabla g_i(x^*))^T dx \le 0 & (i \in \mathcal{I}_g(x^*)); \\ (\nabla h_i(x^*))^T dx = 0 & (i = 1, \cdots, s_2); \\ e_i^T dy \ge 0 & (i \in \mathcal{I}_Y^*); \\ (\nabla_{(x,y)}[y_i^* F_i(x^*, y^*, \omega)])^T \binom{dx}{dy} \le 0 & (i \in \mathcal{I}_Y^*, \forall \omega \in \Omega \setminus B); \\ (\nabla_{(x,y)}[y_i^* F_i(x^*, y^*, \omega)])^T \binom{dx}{dy} \le 0 & (i \notin \mathcal{I}_Y^*, \forall \omega \in \Omega_i \setminus B), \end{array} \right\}$$

which is known as the *linearizing cone* of the feasible region of problem (2.1) at the point (x^*, y^*) and satisfies $\mathcal{T}_{\mathscr{F}_B}(x^*, y^*) \subseteq \mathcal{C}_{\mathscr{F}_B}(x^*, y^*)$. Then, by the nonhomogeneous Farkas lemma for linear semi-infinite systems [8,12], the dual cone of $\mathcal{C}_{\mathscr{F}_B}(x^*, y^*)$ can be written as

$$\Big[\mathcal{C}_{\mathscr{F}_B}(x^*, y^*)\Big]^* = \left\{ \begin{pmatrix} dx \\ dy \end{pmatrix} \in \Re^{n+m} \;\middle|\; \begin{array}{l} \binom{dx}{dy} \in \sum\limits_{i\in\mathcal{I}_g(x^*)} \alpha_i \binom{\nabla g_i(x^*)}{0} + \sum\limits_{i=1}^{s_2} \beta_i \binom{\nabla h_i(x^*)}{0} - \sum\limits_{i\in\mathcal{I}_Y^*} \gamma_i \binom{0}{e_i} \\ \quad + \sum\limits_{i\in\mathcal{I}_Y^*} \overline{\text{cone}}\Big\{\nabla_{(x,y)}[y_i^* F_i(x^*, y^*, \omega)] \;\Big|\; \omega \in \Omega \setminus B\Big\} \\ \quad + \sum\limits_{i\notin\mathcal{I}_Y^*} \overline{\text{cone}}\left\{\nabla_{(x,y)}[y_i^* F_i(x^*, y^*, \omega)] \mid \omega \in \Omega_i \setminus B\right\}, \\ \alpha_i \ge 0\ (i \in \mathcal{I}_g(x^*)), \beta_i \text{ is free } (i = 1, \cdots, s_2), \gamma_i \ge 0\ (i \in \mathcal{I}_Y^*) \end{array} \right\},$$

where $\overline{\text{cone}}$ means the closed convex conical hull.

Let $\mathcal{M}^+(A)$ denote the set of nonnegative regular Borel measures on a σ-algebra σ(A) with $A \subseteq \Omega$. For each $\nu \in \mathcal{M}^+(A)$, we let $\tilde{\nu} \in \mathcal{M}^+(\Omega)$ be a measure satisfying the following conditions:

- $\tilde{\nu}\,|_{\sigma(A)} \equiv \nu$;
- $\tilde{\nu}(\Omega \setminus A) = 0$;
- $\int_\Omega \phi(\omega)\, d\tilde{\nu} = \int_A \phi(\omega)\, d\nu$ holds for any function ϕ defined on Ω and integrable with respect to ν on A.

It then follows from Lemma 6.3 in [11] that

$$\begin{aligned} \overline{\text{cone}}\Big\{\nabla_{(x,y)}[y_i^* F_i(x^*, y^*, \omega)] \;\Big|\; \omega \in \Omega \setminus B\Big\} &= \left\{ \int_{\Omega\setminus B} \nabla_{(x,y)}[y_i^* F_i(x^*, y^*, \omega)] d\nu \;\middle|\; \nu \in \mathcal{M}^+(\Omega \setminus B) \right\} \\ &= \left\{ \int_{\Omega} \nabla_{(x,y)}[y_i^* F_i(x^*, y^*, \omega)] d\tilde{\nu} \;\middle|\; \tilde{\nu} \in \mathcal{M}^+(\Omega) \right\} \end{aligned}$$

for $i \in \mathcal{I}_Y^*$ and

$$\overline{\text{cone}}\Big\{\nabla_{(x,y)}[y_i^* F_i(x^*,y^*,\omega)] \,\Big|\, \omega \in \Omega_i \setminus B\Big\} = \Big\{\int_{\Omega_i\setminus B} \nabla_{(x,y)}[y_i^* F_i(x^*,y^*,\omega)]d\nu \,\Big|\, \nu \in \mathcal{M}^+(\Omega_i\setminus B)\Big\}$$
$$= \Big\{\int_{\Omega_i} \nabla_{(x,y)}[y_i^* F_i(x^*,y^*,\omega)]d\tilde{\nu} \,\Big|\, \tilde{\nu} \in \mathcal{M}^+(\Omega_i)\Big\}$$

for $i \notin \mathcal{I}_Y^*$.

Recall that a measure ν is said to be *absolutely continuous* with respect to another measure ς, denoted $\nu \ll \varsigma$ as usual, if $\nu(B) = 0$ whenever $\varsigma(B) = 0$. Suppose that the following constraint qualification holds:

$$\bigcup_{\substack{B\subset\Omega\\ p(B)=0}} \mathcal{C}_{\mathscr{F}_B}(x^*,y^*) \subseteq \bigcup_{\substack{B\subset\Omega\\ p(B)=0}} \mathcal{T}_{\mathscr{F}_B}(x^*,y^*). \tag{2.5}$$

It then follows that

$$\bigcap_{\substack{B\subset\Omega\\ p(B)=0}} \Big[\mathcal{T}_{\mathscr{F}_B}(x^*,y^*)\Big]^* \subseteq \bigcap_{\substack{B\subset\Omega\\ p(B)=0}} \Big[\mathcal{C}_{\mathscr{F}_B}(x^*,y^*)\Big]^*.$$

Thus, we have from (2.4) that

$$\begin{aligned}
&-\mathbb{E}\Big[\nabla_{(x,y)} f(x^*,y^*,\omega) - 2\sigma\nabla_{(x,y)}F(x^*,y^*,\omega)u(x^*,y^*,\omega)\Big] \\
\in\ & \bigcap_{\substack{B\subset\Omega\\ p(B)=0}} \Big[\mathcal{C}_{\mathscr{F}_B}(x^*,y^*)\Big]^* \\
=\ & \left\{ \begin{pmatrix} dx \\ dy \end{pmatrix} \in \Re^{n+m} \,\middle|\, \begin{array}{l} \begin{pmatrix} dx \\ dy \end{pmatrix} \in \sum\limits_{i\in\mathcal{I}_g(x^*)} \alpha_i \begin{pmatrix} \nabla g_i(x^*) \\ 0 \end{pmatrix} + \sum\limits_{i=1}^{s_2} \beta_i \begin{pmatrix} \nabla h_i(x^*) \\ 0 \end{pmatrix} - \sum\limits_{i\in\mathcal{I}_Y^*} \gamma_i \begin{pmatrix} 0 \\ e_i \end{pmatrix} \\ \quad + \sum\limits_{i\in\mathcal{I}_Y^*} \bigcap\limits_{\substack{B\subset\Omega\\ p(B)=0}} \overline{\text{cone}}\Big\{\nabla_{(x,y)}[y_i^* F_i(x^*,y^*,\omega)] \,\Big|\, \omega \in \Omega\setminus B\Big\} \\ \quad + \sum\limits_{i\notin\mathcal{I}_Y^*} \bigcap\limits_{\substack{B\subset\Omega\\ p(B)=0}} \overline{\text{cone}}\left\{\nabla_{(x,y)}[y_i^* F_i(x^*,y^*,\omega)] \mid \omega \in \Omega_i\setminus B\right\}, \\ \alpha_i \geq 0\ (i\in\mathcal{I}_g(x^*)), \beta_i \text{ is free } (i=1,\cdots,s_2), \gamma_i \geq 0\ (i\in\mathcal{I}_Y^*) \end{array} \right\} \\
=\ & \left\{ \begin{pmatrix} dx \\ dy \end{pmatrix} \in \Re^{n+m} \,\middle|\, \begin{array}{l} \begin{pmatrix} dx \\ dy \end{pmatrix} \in \sum\limits_{i\in\mathcal{I}_g(x^*)} \alpha_i \begin{pmatrix} \nabla g_i(x^*) \\ 0 \end{pmatrix} + \sum\limits_{i=1}^{s_2} \beta_i \begin{pmatrix} \nabla h_i(x^*) \\ 0 \end{pmatrix} - \sum\limits_{i\in\mathcal{I}_Y^*} \gamma_i \begin{pmatrix} 0 \\ e_i \end{pmatrix} \\ \quad + \sum\limits_{i\in\mathcal{I}_Y^*} \left\{ \int_\Omega \nabla_{(x,y)}[y_i^* F_i(x^*,y^*,\omega)]d\tilde{\nu} \,\middle|\, \substack{\tilde{\nu}\in\mathcal{M}^+(\Omega)\\ \tilde{\nu}\ll p} \right\} \\ \quad + \sum\limits_{i\notin\mathcal{I}_Y^*} \left\{ \int_{\Omega_i} \nabla_{(x,y)}[y_i^* F_i(x^*,y^*,\omega)]d\tilde{\nu} \,\middle|\, \substack{\tilde{\nu}\in\mathcal{M}^+(\Omega_i)\\ \tilde{\nu}\ll p} \right\}, \\ \alpha_i \geq 0\ (i\in\mathcal{I}_g(x^*)), \beta_i \text{ is free } (i=1,\cdots,s_2), \gamma_i \geq 0\ (i\in\mathcal{I}_Y^*) \end{array} \right\}.
\end{aligned}$$

As a result, there exist multipliers

$$\begin{aligned}
\alpha_i^* &\geq 0, && (i\in\mathcal{I}_g(x^*)); \\
\beta_i^* &\text{ is free}, && (i=1,\cdots,s_2); \\
\gamma_i^* &\geq 0, && (i\in\mathcal{I}_Y^*);
\end{aligned}$$

and nonnegative regular Borel measures

$$\tilde{\nu}_i^* \in \mathcal{M}^+(\Omega),\ \tilde{\nu}_i^* \ll p, \qquad (i \in \mathcal{I}_Y^*);$$
$$\tilde{\nu}_i^* \in \mathcal{M}^+(\Omega_i),\ \tilde{\nu}_i^* \ll p, \qquad (i \notin \mathcal{I}_Y^*)$$

such that

$$\begin{aligned} &-\mathbb{E}\Big[\nabla_{(x,y)} f(x^*, y^*, \omega) - 2\sigma \nabla_{(x,y)} F(x^*, y^*, \omega) u(x^*, y^*, \omega)\Big] \\ = \quad & \sum_{i \in \mathcal{I}_g(x^*)} \alpha_i^* \begin{pmatrix} \nabla g_i(x^*) \\ 0 \end{pmatrix} + \sum_{i=1}^{s_2} \beta_i^* \begin{pmatrix} \nabla h_i(x^*) \\ 0 \end{pmatrix} - \sum_{i \in \mathcal{I}_Y^*} \gamma_i^* \begin{pmatrix} 0 \\ e_i \end{pmatrix} \\ & + \sum_{i \in \mathcal{I}_Y^*} \int_\Omega \nabla_{(x,y)} [y_i^* F_i(x^*, y^*, \omega)] d\tilde{\nu}_i^* + \sum_{i \notin \mathcal{I}_Y^*} \int_{\Omega_i} \nabla_{(x,y)} [y_i^* F_i(x^*, y^*, \omega)] d\tilde{\nu}_i^*. \end{aligned}$$

By Theorem 7-7B (Radon-Nikodym theorem) and Exercise 7-53 in [3], there are some finite-valued nonnegative measurable functions $\delta_i^*, i = 1, \cdots, m$, defined on Ω or Ω_i such that

$$\int_\Omega \nabla_{(x,y)} [y_i^* F_i(x^*, y^*, \omega)] d\tilde{\nu}_i^* = \int_\Omega \nabla_{(x,y)} [y_i^* F_i(x^*, y^*, \omega)] \delta_i^*(\omega) dp, \quad i \in \mathcal{I}_Y^*$$

and

$$\int_{\Omega_i} \nabla_{(x,y)} [y_i^* F_i(x^*, y^*, \omega)] d\tilde{\nu}_i^* = \int_{\Omega_i} \nabla_{(x,y)} [y_i^* F_i(x^*, y^*, \omega)] \delta_i^*(\omega) dp, \quad i \notin \mathcal{I}_Y^*.$$

For each $i \notin \mathcal{I}_Y^*$, we set $\delta_i^*(\omega) := 0$ for any $\omega \in \Omega \setminus \Omega_i$. Then, for each $i = 1, \cdots, m$, there holds

$$\int_{\Omega_i} \nabla_{(x,y)} [y_i^* F_i(x^*, y^*, \omega)] d\tilde{\nu}_i^* = \mathbb{E}\Big[\nabla_{(x,y)} [y_i^* F_i(x^*, y^*, \omega)] \delta_i^*(\omega)\Big].$$

In consequence, by letting $\alpha_i^* := 0$ for every $i \notin \mathcal{I}_g(x^*)$ and $\gamma_i^* := 0$ for every $i \notin \mathcal{I}_Y^*$, we obtain the following result.

Theorem 2.1 *Suppose that (x^*, y^*) is a local optimal solution of problem (1.3). Assume that there exists an integrable function $\eta : \Omega \to [0, +\infty]$ satisfying condition (2.3) and there holds the constraint qualification (2.5). Then, there exist some multiplier vectors $\alpha^* \in \Re^{s_1}, \beta^* \in \Re^{s_2}, \gamma^* \in \Re^m$, and a multiplier function $\delta^* : \Omega \to \Re^m$ such that*

$$\begin{aligned} 0 \ &= \ \mathbb{E}[\nabla_x f(x^*, y^*, \omega) - 2\sigma \nabla_x F(x^*, y^*, \omega) u(x^*, y^*, \omega)] \\ &\qquad + \nabla g(x^*) \alpha^* + \nabla h(x^*) \beta^* + \mathbb{E}[\nabla_x (y^* \circ F(x^*, y^*, \omega)) \delta^*(\omega)], \end{aligned} \tag{2.6}$$

$$\begin{aligned} 0 \ &= \ \mathbb{E}[\nabla_y f(x^*, y^*, \omega) - 2\sigma \nabla_y F(x^*, y^*, \omega) u(x^*, y^*, \omega)] \\ &\qquad - \gamma^* + \mathbb{E}[\nabla_y (y^* \circ F(x^*, y^*, \omega)) \delta^*(\omega)], \end{aligned} \tag{2.7}$$

$$0 \ \leq \ \alpha^* \ \perp \ -g(x^*) \ \geq \ 0, \tag{2.8}$$

$$\beta^* \ : \ \text{free}, \qquad h(x^*) = 0, \tag{2.9}$$

$$0 \ \leq \ \gamma^* \ \perp \ y^* \ \geq \ 0, \tag{2.10}$$

$$0 \ \leq \ \delta^*(\omega) \ \perp \ -y^* \circ F(x^*, y^*, \omega) \ \geq \ 0, \quad \omega \in \Omega \ \text{ a.s.} \tag{2.11}$$

The above result naturally yields the following definition of stationarity for problem (1.3).

Definition 2.1 *We say (x^*, y^*) is stationary to (1.3) if there exist Lagrangian multiplier vectors $\alpha^* \in \Re^{s_1}, \beta^* \in \Re^{s_2}, \gamma^* \in \Re^m$, and a Lagrangian multiplier function $\delta^* : \Omega \to \Re^m$ such that conditions (2.6)–(2.11) hold.*

Similarly, we define the strong stationarity for problem (1.2) as follows.

Definition 2.2 *We say $(x^*, y^*, z^*(\cdot))$ is strongly stationary to problem (1.2) if it is feasible in (1.2) and there exist Lagrangian multiplier vectors $\bar{\alpha} \in \Re^{s_1}, \bar{\beta} \in \Re^{s_2}, \bar{\gamma} \in \Re^m$, and Lagrangian multiplier functions $\bar{\lambda}, \bar{\mu} : \Omega \to \Re^m$ such that*

$$
\begin{aligned}
0 &= \mathbb{E}[\nabla_x f(x^*, y^*, \omega)] + \nabla g(x^*)\bar{\alpha} + \nabla h(x^*)\bar{\beta} - \mathbb{E}[\nabla_x F(x^*, y^*, \omega)\bar{\lambda}(\omega)] && (2.12)\\
0 &= \mathbb{E}[\nabla_y f(x^*, y^*, \omega)] - \bar{\gamma} - \mathbb{E}[\nabla_y F(x^*, y^*, \omega)\bar{\lambda}(\omega)], && (2.13)\\
0 &= \mathbb{E}[2\sigma z^*(\omega)] - \mathbb{E}[\bar{\lambda}(\omega)] - \mathbb{E}[\bar{\mu}(\omega)], && (2.14)\\
0 &\le \bar{\alpha} \quad \perp \quad -g(x^*) \quad \ge \quad 0, && (2.15)\\
\bar{\beta} &: \text{free}, \qquad h(x^*) = 0, && (2.16)\\
\bar{\gamma}_i &: \text{zero if } i \notin \mathcal{I}_Y^*;\ \text{free if } i \notin \mathcal{I}_W^*;\ \text{nonnegative if } i \in \mathcal{I}_Y^* \cap \mathcal{I}_W^*, && (2.17)\\
\bar{\lambda}_i(\omega) &: \text{free if } i \notin \mathcal{I}_Y^*;\ \text{zero if } i \notin \mathcal{I}_W^*;\ \text{nonnegative if } i \in \mathcal{I}_Y^* \cap \mathcal{I}_W^*, \quad \omega \in \Omega \ \text{ a.s.}, && (2.18)\\
0 &\le \bar{\mu}(\omega) \quad \perp \quad z^*(\omega) \quad \ge \quad 0, \quad \omega \in \Omega \ \text{ a.s.} && (2.19)
\end{aligned}
$$

where $\mathcal{I}_W^* := \{i \mid F_i(x^*, y^*, \omega) + z_i^*(\omega) = 0,\ \omega \in \Omega \ \text{ a.s.}\}$.

This definition can obviously be regarded as a generalization of the strong stationarity in the literature on MPEC [23]. The connections between the above concepts can be stated as follows.

Theorem 2.2 *If (x^*, y^*) is a stationary point of problem (1.3), then $(x^*, y^*, u(x^*, y^*, \cdot))$ is a strongly stationary point of problem (1.2), where u is defined by (1.4).*

Proof. Let $z^*(\omega) := u(x^*, y^*, \omega)$ for any $\omega \in \Omega$. Then $(x^*, y^*, z^*(\cdot))$ is feasible to problem (1.2) (cf. Appendix). We next show that there exist Lagrangian multiplier vectors $\bar{\alpha} \in \Re^{s_1}, \bar{\beta} \in \Re^{s_2}, \bar{\gamma} \in \Re^m$, and Lagrangian multiplier functions $\bar{\lambda}, \bar{\mu} : \Omega \to \Re^m$ satisfying (2.12)–(2.19). Since (x^*, y^*) is stationary to (1.3), there must be Lagrangian multiplier vectors $\alpha^* \in \Re^{s_1}, \beta^* \in \Re^{s_2}, \gamma^* \in \Re^m$, and a Lagrangian multiplier function $\delta^* : \Omega \to \Re^m$ satisfying (2.6)–(2.11). Let

$$
\begin{aligned}
\bar{\alpha} &:= \alpha^*, && (2.20)\\
\bar{\beta} &:= \beta^*, && (2.21)\\
\bar{\gamma} &:= \gamma^* - \mathbb{E}[\mathrm{diag}(F_1(x^*, y^*, \omega), \cdots, F_m(x^*, y^*, \omega))\, \delta^*(\omega)], && (2.22)\\
\bar{\lambda}(\omega) &:= 2\sigma z^*(\omega) - \mathrm{diag}(y_1^*, \cdots, y_m^*)\, \delta^*(\omega), && (2.23)\\
\bar{\mu}(\omega) &:= \mathrm{diag}(y_1^*, \cdots, y_m^*)\, \delta^*(\omega). && (2.24)
\end{aligned}
$$

We then have (2.14)–(2.16) immediately. On the other hand, since

$$\begin{aligned}
\nabla_x(y \circ F(x,y,\omega)) &= \nabla_x F(x,y,\omega)\mathrm{diag}(y_1,\cdots,y_m), && (2.25)\\
\nabla_y(y \circ F(x,y,\omega)) &= \nabla_y F(x,y,\omega)\mathrm{diag}(y_1,\cdots,y_m) + \mathrm{diag}(F_1(x,y,\omega),\cdots,F_m(x,y,\omega)) && (2.26)
\end{aligned}$$

for any $(x,y) \in \Re^{n+m}$ and $\omega \in \Omega$, conditions (2.12) and (2.13) follow from (2.6)–(2.7) and (2.20)–(2.24).

We next show (2.17) and (2.18). Note that, from (2.11),

$$y_i^* \delta_i^*(\omega) F_i(x^*,y^*,\omega) = 0, \qquad \omega \in \Omega \ \text{ a.s.} \tag{2.27}$$

holds for each $i = 1,\cdots,m$. Moreover, it is not difficult to show that

$$\mathcal{I}_W^* = \Big\{ i \,\Big|\, \mathbb{E}[F_i(x^*,y^*,\omega) + z_i^*(\omega)] = 0 \Big\}.$$

Suppose that $i \notin \mathcal{I}_Y^*$, which means $y_i^* > 0$. It then follows from (2.10) that $\gamma_i^* = 0$. Moreover, by (2.27), there holds $\delta_i^*(\omega)F_i(x^*,y^*,\omega) = 0$ for almost every $\omega \in \Omega$ and hence $\mathbb{E}[\delta_i^*(\omega)F_i(x^*,y^*,\omega)] = 0$. Therefore, we have

$$\bar{\gamma}_i = \gamma_i^* - \mathbb{E}[\delta_i^*(\omega)F_i(x^*,y^*,\omega)] = 0.$$

Suppose that $i \in \mathcal{I}_Y^* \cap \mathcal{I}_W^*$. From the feasibility of $(x^*,y^*,z^*(\cdot))$ in (1.2), we have $F_i(x^*,y^*,\omega) + z_i^*(\omega) = 0$ for almost every $\omega \in \Omega$. This indicates that $F_i(x^*,y^*,\omega) \le 0$ for almost all $\omega \in \Omega$ and hence $\mathbb{E}[\delta_i^*(\omega)F_i(x^*,y^*,\omega)] \le 0$. Since $\gamma_i^* \ge 0$ by (2.10), we have

$$\bar{\gamma}_i = \gamma_i^* - \mathbb{E}[\delta_i^*(\omega)F_i(x^*,y^*,\omega)] \ge 0.$$

This shows (2.17). In a similar way, we can show (2.18).

Finally we show (2.19). It is obvious that $\bar{\mu}(\omega) \ge 0$ and $z^*(\omega) \ge 0$ for almost all $\omega \in \Omega$. Noticing that $\bar{\mu}_i(\omega)F_i(x^*,y^*,\omega) = y_i^*\delta_i^*(\omega)F_i(x^*,y^*,\omega) = 0$ by (2.27) and $z_i^*(\omega) = \max\{-F_i(x^*,y^*,\omega), 0\}$ by the definition, we have $\bar{\mu}_i(\omega)z_i^*(\omega) = 0$ for each i and almost every $\omega \in \Omega$. As a result, there must hold (2.19).

In consequence, the multipliers defined by (2.20)–(2.24) satisfy conditions (2.12)–(2.19). Namely, $(x^*,y^*,z^*(\cdot))$ is a strongly stationary point of problem (1.2) ∎

3 Monte Carlo Sampling and Penalty Approximations

Let Ω be a compact set and $\phi : \Omega \to \Re$ be a function. The Monte Carlo sampling estimate for $\mathbb{E}[\phi(\omega)]$ is obtained by taking independently and identically distributed random samples

$\{\omega_1, \cdots, \omega_k\}$ from Ω and letting $\mathbb{E}[\phi(\omega)] \approx \frac{1}{k}\sum_{\ell=1}^{k}\phi(\omega_\ell)$. The strong law of large numbers guarantees that this procedure converges with probability one (abbreviated by "w.p.1" below), i.e.,

$$\lim_{k\to\infty}\frac{1}{k}\sum_{\ell=1}^{k}\phi(\omega_\ell) = \mathbb{E}[\phi(\omega)] := \int_\Omega \phi(\omega)dp \qquad \text{w.p.1.} \tag{3.1}$$

See [19, 24] for more details about the Monte Carlo sampling methods.

Applying the above method and using a penalty technique, we obtain the problem

$$\begin{array}{ll} \min & \displaystyle\frac{1}{k}\sum_{\ell=1}^{k}\Big(f(x,y,\omega_\ell) + \sigma\|u(x,y,\omega_\ell)\|^2 + \rho_k\,\|y\circ v(x,y,\omega_\ell)\|^2\Big) \\ \text{s.t.} & g(x)\le 0,\ \ h(x)=0,\ \ y\ge 0, \end{array} \tag{3.2}$$

which is a smooth approximation of problem (1.3). Here, $\rho_k > 0$ is a penalty parameter tending to ∞ as $k \to \infty$, $u : \Re^{n+m} \times \Omega \to \Re^m$ is defined by (1.4), and $v : \Re^{n+m} \times \Omega \to \Re^m$ is given by

$$v(x,y,\omega) := \max\{F(x,y,\omega),\ 0\}. \tag{3.3}$$

Note that, by (1.4) and (3.3), we have

$$v(x,y,\omega) = F(x,y,\omega) + u(x,y,\omega), \quad (x,y,\omega) \in \Re^{n+m}\times\Omega. \tag{3.4}$$

Problem (3.2) is neither a semi-infinite program nor an MPEC and it is generally much easier to deal with than those problems. In the rest of the paper, we denote the feasible region of problem (3.2) by $\mathcal{F}$. Note that $\mathcal{F}$ does not depend on k.

We next discuss the existence conditions of solutions of problem (3.2). Let F be affine with respect to (x, y) and given by

$$F(x,y,\omega) := N(\omega)x + M(\omega)y + q(\omega), \tag{3.5}$$

where $N : \Omega \to \Re^{m\times n}$, $M : \Omega \to \Re^{m\times m}$, and $q : \Omega \to \Re^m$ are all continuous.

Definition 3.1 *We call $\bar{M} \in \Re^{m\times m}$ an R_0-matrix if*

$$y \ge 0,\ \bar{M}y \ge 0,\ y^T\bar{M}y = 0 \quad\Longrightarrow\quad y = 0.$$

It is well-known that any P-matrix is an R_0-matrix [6]. We have the following result.

Lemma 3.1 *Let $\{M_k\} \subset \Re^{m\times m}$ be convergent to $\bar{M} \in \Re^{m\times m}$ and $\bar{M}$ be an R_0-matrix. Then, there exists an integer $k_0 > 0$ such that M_k is an R_0-matrix for every $k \ge k_0$.*

Theorem 3.1 *Suppose that the set $\mathcal{X} := \{x \in \Re^n \mid g(x) \le 0, h(x) = 0\}$ is nonempty and bounded, the function f is bounded below on $\mathcal{F} \times \Omega$, and $\lim_{k\to\infty} \rho_k = +\infty$. Let F be defined by (3.5) and $\bar{M} := \int_\Omega M(\omega)dp$ be an R_0-matrix. We then have the following statements almost surely.*

(i) *Problem (3.2) has at least one optimal solution when k is sufficiently large.*

(ii) *Let (x^k, y^k) be an optimal solution of (3.2) for each k sufficiently large. Then the sequence $\{(x^k, y^k)\}$ is bounded.*

Proof. (i) For each k, let $M_k := \frac{1}{k}\sum_{\ell=1}^k M(\omega_\ell)$. It then follows from (3.1) that $\bar{M} = \lim_{k\to\infty} M_k$ with probability one. Since $\bar{M}$ is an R_0-matrix, by Lemma 3.1, there almost surely exists an integer $k_0 > 0$ such that M_k is an R_0-matrix for every $k \ge k_0$.

Let $k \ge k_0$ be fixed and suppose M_k is an R_0-matrix. It is easy to see that $\mathcal{F}$ is a nonempty closed set and the objective function of problem (3.2) is bounded below on $\mathcal{F}$. Then, there exists a sequence $\{(x^j, y^j)\} \subseteq \mathcal{F}$ such that

$$\begin{aligned} &\lim_{j\to\infty} \frac{1}{k}\sum_{\ell=1}^k \Big(f(x^j, y^j, \omega_\ell) + \sigma\|u(x^j, y^j, \omega_\ell)\|^2 + \rho_k\|y^j \circ v(x^j, y^j, \omega_\ell)\|^2 \Big) \\ = &\inf_{(x,y)\in\mathcal{F}} \frac{1}{k}\sum_{\ell=1}^k \Big(f(x, y, \omega_\ell) + \sigma\|u(x, y, \omega_\ell)\|^2 + \rho_k\|y \circ v(x, y, \omega_\ell)\|^2 \Big). \end{aligned} \tag{3.6}$$

Since f is bounded below and ρ_k is a positive constant, it follows from (3.6) that the sequences

$$\Big\{ \frac{1}{k}\sum_{\ell=1}^k \|u(x^j, y^j, \omega_\ell)\|^2 \Big\}_{j=0,1,\cdots} \quad \text{and} \quad \Big\{ \frac{1}{k}\sum_{\ell=1}^k \|y^j \circ v(x^j, y^j, \omega_\ell)\|^2 \Big\}_{j=0,1,\cdots}$$

are bounded. This along with (3.4) implies that

$$\Big\{ \frac{1}{k}\sum_{\ell=1}^k u(x^j, y^j, \omega_\ell) \Big\}_{j=0,1,\cdots} \text{and} \ \Big\{ \frac{1}{k}\sum_{\ell=1}^k (y^j)^T \Big(N(\omega_\ell)x^j + M(\omega_\ell)y^j + q(\omega_\ell) + u(x^j, y^j, \omega_\ell) \Big) \Big\}_{j=0,1,\cdots}$$

are also bounded. Note that the latter sequence can be rewritten as

$$\Big\{ (y^j)^T \Big(\frac{1}{k}\sum_{\ell=1}^k N(\omega_\ell)x^j + M_k y^j + \frac{1}{k}\sum_{\ell=1}^k q(\omega_\ell) + \frac{1}{k}\sum_{\ell=1}^k u(x^j, y^j, \omega_\ell) \Big) \Big\}_{j=0,1,\cdots}. \tag{3.7}$$

Moreover, by the boundedness of the set $\mathcal{X}$, the sequence $\{x^j\}$ is bounded. On the other hand, it is obvious from the feasibility of (x^j, y^j) in (3.2) and the definition of u that, for each j,

$$y^j \ge 0, \qquad \frac{1}{k}\sum_{\ell=1}^k N(\omega_\ell)x^j + M_k y^j + \frac{1}{k}\sum_{\ell=1}^k q(\omega_\ell) + \frac{1}{k}\sum_{\ell=1}^k u(x^j, y^j, \omega_\ell) \ge 0. \tag{3.8}$$

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

$$\lim_{j\to\infty} \|y^j\| = +\infty, \quad \lim_{j\to\infty} \frac{y^j}{\|y^j\|} = \bar{y}, \quad \|\bar{y}\| = 1. \tag{3.9}$$

Then, dividing (3.7) and (3.8) by $\|y^j\|^2$ and $\|y^j\|$, respectively, and letting $j \to +\infty$, we obtain

$$0 \leq \bar{y} \perp M_k \bar{y} \geq 0.$$

Since M_k is an R_0-matrix, we have $\bar{y} = 0$. This contradicts (3.9) and hence $\{y^j\}$ is bounded.

Therefore, $\{(x^j, y^j)\}$ is bounded. Since $\mathcal{F}$ is closed, we see from (3.6) that any accumulation point of $\{(x^j, y^j)\}$ must be an optimal solution of (3.2). This completes the proof of (i).

(ii) Let (x^k, y^k) be an optimal solution of (3.2) for each sufficiently large k. The boundedness of $\{x^k\}$ follows from the boundedness of the set $\mathcal{X}$ immediately. We next prove that $\{y^k\}$ is almost surely bounded. To this end, we choose a vector $\bar{x} \in \mathcal{X}$ arbitrarily. Then, $(\bar{x}, 0)$ is feasible to problem (3.2). Since (x^k, y^k) is an optimal solution of (3.2), we have

$$\begin{aligned} & \frac{1}{k}\sum_{\ell=1}^{k}\Big(f(x^k, y^k, \omega_\ell) + \sigma\|u(x^k, y^k, \omega_\ell)\|^2 + \rho_k\|y^k \circ v(x^k, y^k, \omega_\ell)\|^2\Big) \\ \leq\ & \frac{1}{k}\sum_{\ell=1}^{k}\Big(f(\bar{x}, 0, \omega_\ell) + \sigma\|u(\bar{x}, 0, \omega_\ell)\|^2\Big) \end{aligned} \tag{3.10}$$

and, by the definitions (1.4) and (3.5),

$$\frac{1}{k}\sum_{\ell=1}^{k} N(\omega_\ell)x^k + \frac{1}{k}\sum_{\ell=1}^{k} M(\omega_\ell)y^k + \frac{1}{k}\sum_{\ell=1}^{k} q(\omega_\ell) + \frac{1}{k}\sum_{\ell=1}^{k} u(x^k, y^k, \omega_\ell) \geq 0, \quad y^k \geq 0. \tag{3.11}$$

It follows from (3.10) that

$$\begin{aligned} 0 \ &\leq\ \frac{\sigma}{k}\sum_{\ell=1}^{k}\|u(x^k, y^k, \omega_\ell)\|^2 + \frac{\rho_k}{k}\sum_{\ell=1}^{k}\|y^k \circ v(x^k, y^k, \omega_\ell)\|^2 \\ &\leq\ \frac{1}{k}\sum_{\ell=1}^{k}\Big(f(\bar{x}, 0, \omega_\ell) - f(x^k, y^k, \omega_\ell)\Big) + \frac{\sigma}{k}\sum_{\ell=1}^{k}\|u(\bar{x}, 0, \omega_\ell)\|^2. \end{aligned}$$

Since f is bounded below, we have from (3.1) that

$$\left\{\frac{1}{k}\sum_{\ell=1}^{k}\Big(f(\bar{x}, 0, \omega_\ell) - f(x^k, y^k, \omega_\ell)\Big)\right\} \quad \text{and} \quad \left\{\frac{\sigma}{k}\sum_{\ell=1}^{k}\|u(\bar{x}, 0, \omega_\ell)\|^2\right\}$$

are almost surely bounded. In consequence, the sequences

$$\left\{\frac{1}{k}\sum_{\ell=1}^{k}\|u(x^k, y^k, \omega_\ell)\|^2\right\} \quad \text{and} \quad \left\{\frac{\rho_k}{k}\sum_{\ell=1}^{k}\|y^k \circ v(x^k, y^k, \omega_\ell)\|^2\right\}$$

are almost surely bounded. By Cauchy-Schwartz inequality, we have

$$\Big(\sum_{\ell=1}^{k} u_i(x^k, y^k, \omega_\ell)\Big)^2 \leq k\sum_{\ell=1}^{k}\Big(u_i(x^k, y^k, \omega_\ell)\Big)^2, \qquad i = 1, \cdots, m$$

for each k and hence

$$\begin{aligned} \Big\|\frac{1}{k}\sum_{\ell=1}^{k} u(x^k, y^k, \omega_\ell)\Big\|^2 \ &=\ \frac{1}{k^2}\sum_{i=1}^{m}\Big(\sum_{\ell=1}^{k} u_i(x^k, y^k, \omega_\ell)\Big)^2 \\ &\leq\ \frac{1}{k}\sum_{i=1}^{m}\sum_{\ell=1}^{k}\Big(u_i(x^k, y^k, \omega_\ell)\Big)^2 \ =\ \frac{1}{k}\sum_{\ell=1}^{k}\|u(x^k, y^k, \omega_\ell)\|^2. \end{aligned} \tag{3.12}$$

Similarly, we have

$$\begin{aligned}
&\Big|\frac{1}{k}\sum_{\ell=1}^{k}(y^k)^T\Big(N(\omega_\ell)x^k+M(\omega_\ell)y^k+q(\omega_\ell)+u(x^k,y^k,\omega_\ell)\Big)\Big|^2\\
=\;& \frac{1}{k^2}\Big|\sum_{i=1}^{m}\sum_{\ell=1}^{k}y_i^k v_i(x^k,y^k,\omega_\ell)\Big|^2 \;\le\; \frac{m}{k^2}\sum_{i=1}^{m}\Big(\sum_{\ell=1}^{k}y_i^k v_i(x^k,y^k,\omega_\ell)\Big)^2\\
\le\;& \frac{m}{k}\sum_{i=1}^{m}\sum_{\ell=1}^{k}\Big(y_i^k v_i(x^k,y^k,\omega_\ell)\Big)^2 \;=\; \frac{m}{k}\sum_{\ell=1}^{k}\|y^k\circ v(x^k,y^k,\omega_\ell)\|^2.
\end{aligned} \tag{3.13}$$

It follows from (3.12) and (3.13) that both $\left\{\frac{1}{k}\sum_{\ell=1}^{k}u(x^k,y^k,\omega_\ell)\right\}$ and

$$\left\{\frac{1}{k}\sum_{\ell=1}^{k}(y^k)^T\Big(N(\omega_\ell)x^k+M(\omega_\ell)y^k+q(\omega_\ell)+u(x^k,y^k,\omega_\ell)\Big)\right\} \tag{3.14}$$

are almost surely bounded. Suppose that the sequence $\{y^k\}$ is unbounded with probability one. Taking a subsequence if necessary, we assume that

$$\lim_{k\to\infty}\|y^k\|=+\infty,\quad \lim_{k\to\infty}\frac{y^k}{\|y^k\|}=\bar{y},\quad \|\bar{y}\|=1. \tag{3.15}$$

Note that the sequences $\{x^k\}$ and $\left\{\frac{1}{k}\sum_{\ell=1}^{k}u(x^k,y^k,\omega_\ell)\right\}$ are bounded and, by (3.1),

$$\lim_{k\to\infty}\frac{1}{k}\sum_{\ell=1}^{k}M(\omega_\ell)=\bar{M},\quad \lim_{k\to\infty}\frac{1}{k}\sum_{\ell=1}^{k}N(\omega_\ell)=\int_\Omega N(\omega)dp,\quad \lim_{k\to\infty}\frac{1}{k}\sum_{\ell=1}^{k}q(\omega_\ell)=\int_\Omega q(\omega)dp.$$

Dividing (3.11) and (3.14) by $\|y^k\|$ and $\|y^k\|^2$, respectively, and letting $k\to+\infty$, we obtain $0\le\bar{y}\perp\bar{M}\,\bar{y}\ge0$. Since $\bar{M}$ is an R_0-matrix, we have $\bar{y}=0$ with probability one. This contradicts (3.15) and hence, the sequence $\{y^k\}$ is almost surely bounded. This completes the proof of (ii). ∎

4 Convergence Analysis

In this section, we investigate convergence properties of the Monte Carlo sampling and penalty method. For each k, we let $\{\omega_1,\cdots,\omega_k\}$ be independently and identically distributed random samples drawn from Ω.

Definition 4.1 [20] Let $\tau>0$ and $\kappa\ge0$ be constants. We say $G:\Re^s\to\Re^t$ is *Hölder continuous* on $K\subseteq\Re^s$ with order τ and Hölder constant κ if

$$\|G(u)-G(v)\|\le\kappa\|u-v\|^\tau$$

holds for all u and v in K.

This concept is a generalization of the Lipschitz continuity, which is, by definition, Hölder continuity with order $\tau = 1$. Note that, for two different positive numbers τ and τ', Hölder continuous functions with order τ and those with order τ' constitute different subclasses. For example, the function $G(u) := \sqrt{\|u\|}$ is Hölder continuous with order $\tau = \frac{1}{2}$ but not Lipschitz continuous.

4.1 Limiting behavior of optimal solutions

We first study the convergence of optimal solutions of problems (3.2). Recall that $\mathcal{F}$ denotes the feasible region of problem (3.2).

Theorem 4.1 *Suppose that both f and F are Hölder continuous in (x, y) on $\mathcal{F}$ with order $\tau > 0$ and Hölder constant $\kappa(\omega)$ satisfying $\mathbb{E}[\kappa(\omega)] < +\infty$. Assume that $\lim_{k\to\infty} \rho_k = +\infty$, (x^k, y^k) solves problem (3.2) for each k, and the sequence $\{(x^k, y^k)\}$ is bounded. Let (x^*, y^*) be an accumulation point of $\{(x^k, y^k)\}$. Then (x^*, y^*) is an optimal solution of problem (1.3) with probability one.*

Proof. Without loss of generality, we suppose $\lim_{k\to\infty}(x^k, y^k) = (x^*, y^*)$.

(a) We first prove that (x^*, y^*) is almost surely feasible to (1.3). It is obvious that (x^*, y^*) satisfies the constraints of problem (3.2). Therefore, it is sufficient to show that there holds

$$y^* \circ F(x^*, y^*, \omega) \le 0, \quad \omega \in \Omega \ \text{ a.s.} \tag{4.1}$$

In fact, since (x^k, y^k) is an optimal solution of problem (3.2) and $(x^*, 0)$ is a feasible point of (3.2), we have

$$\begin{aligned} & \frac{1}{k}\sum_{\ell=1}^{k}\Big(f(x^k, y^k, \omega_\ell) + \sigma\|u(x^k, y^k, \omega_\ell)\|^2 + \rho_k\|y^k \circ v(x^k, y^k, \omega_\ell)\|^2\Big) \\ \le\ & \frac{1}{k}\sum_{\ell=1}^{k}\Big(f(x^*, 0, \omega_\ell) + \sigma\|u(x^*, 0, \omega_\ell)\|^2\Big), \end{aligned}$$

which yields

$$\frac{\rho_k}{k}\sum_{\ell=1}^{k}\|y^k \circ v(x^k, y^k, \omega_\ell)\|^2 \le \frac{1}{k}\sum_{\ell=1}^{k}\Big(f(x^*, 0, \omega_\ell) + \sigma\|u(x^*, 0, \omega_\ell)\|^2\Big) - \frac{1}{k}\sum_{\ell=1}^{k} f(x^k, y^k, \omega_\ell). \tag{4.2}$$

By the assumptions of the theorem, there hold

$$\|f(x, y, \omega) - f(x', y', \omega)\| \le \kappa(\omega)\|(x, y) - (x', y')\|^\tau$$

and

$$\|F(x, y, \omega) - F(x', y', \omega)\| \le \kappa(\omega)\|(x, y) - (x', y')\|^\tau \tag{4.3}$$

for any $(x,y) \in \mathcal{F}, (x',y') \in \mathcal{F}$, and $\omega \in \Omega$. Therefore, we have from (3.1) that

$$\lim_{k\to\infty} \Big| \frac{1}{k} \sum_{\ell=1}^{k} \Big(f(x^k, y^k, \omega_\ell) - f(x^*, y^*, \omega_\ell) \Big) \Big| \le \lim_{k\to\infty} \|(x^k, y^k) - (x^*, y^*)\|^\tau \, \frac{1}{k} \sum_{\ell=1}^{k} \kappa(\omega_\ell) = 0$$

and hence

$$\lim_{k\to\infty} \frac{1}{k} \sum_{\ell=1}^{k} f(x^k, y^k, \omega_\ell) = \lim_{k\to\infty} \frac{1}{k} \sum_{\ell=1}^{k} f(x^*, y^*, \omega_\ell) = \int_\Omega f(x^*, y^*, \omega) dp \qquad \text{w.p.1.}$$

This indicates that $\left\{ \frac{1}{k} \sum_{\ell=1}^{k} f(x^k, y^k, \omega_\ell) \right\}$ is bounded with probability one. Moreover, since Ω is compact, both $f(x^*, 0, \cdot)$ and $u(x^*, 0, \cdot)$ are bounded on Ω. Thus, we have from (4.2) that the sequence $\left\{ \frac{\rho_k}{k} \sum_{\ell=1}^{k} \|y^k \circ v(x^k, y^k, \omega_\ell)\|^2 \right\}$ is bounded with probability one. As a result, the sequence $\left\{ \frac{\rho_k}{k} \sum_{\ell=1}^{k} (y_i^k)^2 \left(F_i(x^k, y^k, \omega_\ell) + u_i(x^k, y^k, \omega_\ell) \right)^2 \right\}$ is almost surely bounded for each i and, since

$$y^k \ge 0, \qquad F(x^k, y^k, \omega_\ell) + u(x^k, y^k, \omega_\ell) = v(x^k, y^k, \omega_\ell) \ge 0$$

for every k and ℓ, $\left\{ \frac{\rho_k}{k} \sum_{\ell=1}^{k} (y^k)^T (F(x^k, y^k, \omega_\ell) + u(x^k, y^k, \omega_\ell)) \right\}$ is almost surely bounded. Noting that $\lim_{k\to\infty} \rho_k = +\infty$, we have

$$\lim_{k\to\infty} \frac{1}{k} \sum_{\ell=1}^{k} (y^k)^T \Big(F(x^k, y^k, \omega_\ell) + u(x^k, y^k, \omega_\ell) \Big) = 0 \qquad \text{w.p.1.} \tag{4.4}$$

On the other hand, we have from (4.3) that, for any k and ℓ,

$$\begin{aligned} & \|(F(x^k, y^k, \omega_\ell) + u(x^k, y^k, \omega_\ell)) - (F(x^*, y^*, \omega_\ell) + u(x^*, y^*, \omega_\ell))\| \\ \le \; & 2\|F(x^k, y^k, \omega_\ell) - F(x^*, y^*, \omega_\ell)\| \\ \le \; & 2\kappa(\omega_\ell)\|(x^k, y^k) - (x^*, y^*)\|^\tau \end{aligned}$$

and then

$$\begin{aligned} & \lim_{k\to\infty} \Big| \frac{1}{k} \sum_{\ell=1}^{k} (y^k)^T \Big((F(x^k, y^k, \omega_\ell) + u(x^k, y^k, \omega_\ell)) - (F(x^*, y^*, \omega_\ell) + u(x^*, y^*, \omega_\ell)) \Big) \Big| \\ \le \; & \lim_{k\to\infty} 2\,\|y^k\|\, \|(x^k, y^k) - (x^*, y^*)\|^\tau \, \frac{1}{k} \sum_{\ell=1}^{k} \kappa(\omega_\ell) \\ = \; & 0 \qquad \text{w.p.1.} \end{aligned} \tag{4.5}$$

It follows from (4.4) and (4.5) that

$$\begin{aligned} 0 &= \lim_{k\to\infty} \frac{1}{k} \sum_{\ell=1}^{k} (y^k)^T \Big(F(x^k, y^k, \omega_\ell) + u(x^k, y^k, \omega_\ell) \Big) \\ &= \lim_{k\to\infty} \frac{1}{k} \sum_{\ell=1}^{k} (y^k)^T \Big(F(x^*, y^*, \omega_\ell) + u(x^*, y^*, \omega_\ell) \Big) \\ &= \int_\Omega (y^*)^T (F(x^*, y^*, \omega) + u(x^*, y^*, \omega_\ell)) dp \qquad \text{w.p.1,} \end{aligned} \tag{4.6}$$

where the last equality follows from (3.1). Noting that both $(y^*)^T(F(x^*,y^*,\cdot)+u(x^*,y^*,\cdot))$ and $(y^*)^T u(x^*,y^*,\cdot)$ are nonnegative on Ω, we obtain (4.1) from (4.6) immediately.

(b) Let (x,y) be an arbitrary feasible solution of problem (1.3). It is obvious that (x,y) is feasible to problem (3.2). Moreover, if $y_i > 0$ for some i, there must hold $F_i(x,y,\omega) \le 0$ for almost all $\omega \in \Omega$ and so $u_i(x,y,\omega) = -F_i(x,y,\omega)$ for almost all $\omega \in \Omega$. This means

$$y \circ v(x,y,\omega) = y \circ (F(x,y,\omega) + u(x,y,\omega)) = 0, \quad \omega \in \Omega \text{ a.s.} \tag{4.7}$$

Since (x^k, y^k) is an optimal solution of problem (3.2), we have almost surely that

$$\begin{aligned}
& \frac{1}{k}\sum_{\ell=1}^{k}\Big(f(x,y,\omega_\ell) + \sigma\|u(x,y,\omega_\ell)\|^2\Big) \\
= \ & \frac{1}{k}\sum_{\ell=1}^{k}\Big(f(x,y,\omega_\ell) + \sigma\|u(x,y,\omega_\ell)\|^2 + \rho_k\|y \circ v(x,y,\omega_\ell)\|^2\Big) \\
\ge \ & \frac{1}{k}\sum_{\ell=1}^{k}\Big(f(x^k,y^k,\omega_\ell) + \sigma\|u(x^k,y^k,\omega_\ell)\|^2 + \rho_k\|y^k \circ v(x^k,y^k,\omega_\ell)\|^2\Big) \\
\ge \ & \frac{1}{k}\sum_{\ell=1}^{k}\Big(f(x^k,y^k,\omega_\ell) + \sigma\|u(x^k,y^k,\omega_\ell)\|^2\Big).
\end{aligned}$$

As a result, we have

$$\begin{aligned}
& \frac{1}{k}\sum_{\ell=1}^{k}\Big(f(x^*,y^*,\omega_\ell) + \sigma\|u(x^*,y^*,\omega_\ell)\|^2\Big) - \frac{1}{k}\sum_{\ell=1}^{k}\Big(f(x,y,\omega_\ell) + \sigma\|u(x,y,\omega_\ell)\|^2\Big) \\
\le \ & \frac{1}{k}\sum_{\ell=1}^{k}\Big(f(x^*,y^*,\omega_\ell) + \sigma\|u(x^*,y^*,\omega_\ell)\|^2\Big) - \frac{1}{k}\sum_{\ell=1}^{k}\Big(f(x^k,y^k,\omega_\ell) + \sigma\|u(x^k,y^k,\omega_\ell)\|^2\Big) \\
\le \ & \frac{1}{k}\sum_{\ell=1}^{k}\Big(|f(x^*,y^*,\omega_\ell) - f(x^k,y^k,\omega_\ell)| \\
& + \sigma\|u(x^*,y^*,\omega_\ell) - u(x^k,y^k,\omega_\ell)\|\,(\|u(x^*,y^*,\omega_\ell)\| + \|u(x^k,y^k,\omega_\ell)\|)\Big) \qquad \text{w.p.1.}
\end{aligned} \tag{4.8}$$

Note that the Hölder continuity of f yields

$$\lim_{k\to\infty}\frac{1}{k}\sum_{\ell=1}^{k}|f(x^*,y^*,\omega_\ell) - f(x^k,y^k,\omega_\ell)| = 0. \tag{4.9}$$

On the other hand, it follows from (1.4) and (4.3) that

$$\begin{aligned}
\|u(x^*,y^*,\omega_\ell) - u(x^k,y^k,\omega_\ell)\| &\le \|F(x^*,y^*,\omega_\ell) - F(x^k,y^k,\omega_\ell)\| \\
&\le \kappa(\omega_\ell)\|(x^k,y^k) - (x^*,y^*)\|^\tau, \qquad \ell = 1,\cdots,k.
\end{aligned}$$

By the boundedness of the sequence $\Big\{\frac{1}{k}\sum_{\ell=1}^{k}(\|u(x^*,y^*,\omega_\ell)\| + \|u(x^k,y^k,\omega_\ell)\|)\Big\}$, we have

$$\lim_{k\to\infty}\frac{\sigma}{k}\sum_{\ell=1}^{k}\|u(x^*,y^*,\omega_\ell) - u(x^k,y^k,\omega_\ell)\|\Big(\|u(x^*,y^*,\omega_\ell)\| + \|u(x^k,y^k,\omega_\ell)\|\Big) = 0. \tag{4.10}$$

Letting $k \to +\infty$ in (4.8) and taking (4.9), (4.10) and (3.1) into account, we obtain

$$\mathbb{E}[f(x^*, y^*, \omega) + \sigma\|u(x^*, y^*, \omega)\|^2] \leq \mathbb{E}[f(x, y, \omega) + \sigma\|u(x, y, \omega)\|^2] \qquad \text{w.p.1},$$

which indicates that (x^*, y^*) is an optimal solution of (1.3) with probability one. ■

Remark 4.1 In Theorem 4.1, to avoid notational complication, we simply assume that the functions f and F are Hölder continuous with same order. However, it is easy to see from the proof that the assumption can be relaxed to allow those functions to have different orders. This remark also applies to Theorem 4.2 in the next subsection.

4.2 Limiting behavior of stationary points

In general, it is difficult to obtain an optimal solution, whereas computation of stationary points is relatively easy. Therefore, it is important to study the limiting behavior of stationary points of problems (3.2). In what follows, we let $\mathcal{E}$ denote the feasible region of problem (1.3). Note that, for any $(x, y) \in \mathcal{E}$, the standard Mangasarian-Fromovitz constraint qualification does not hold at (x, y) if there holds $y_i = 0$ for some index i. In this paper, we define a generalized constraint qualification for problem (1.3) as follows.

Definition 4.2 *Let $(x^*, y^*) \in \mathcal{E}$. We say the generalized Mangasarian-Fromovitz constraint qualification (GMFCQ) holds at (x^*, y^*) if*

- *the gradients $\nabla h_i(x^*), i = 1, \cdots, s_2$, are linearly independent;*
- *there exists a vector $\binom{dx}{dy} \in \Re^{n+m}$ such that*

$$\begin{aligned} (dx)^T \nabla h_i(x^*) &= 0, & & i = 1, \cdots, s_2; \\ (dx)^T \nabla g_i(x^*) &< 0, & & i \in \mathcal{I}_g(x^*); \\ (dy)^T e_i &> 0, & & i \in \mathcal{I}_Y^*; \\ \binom{dx}{dy}^T \nabla_{(x,y)}[y_i^* F_i(x^*, y^*, \omega)] &< 0, & & \omega \in \hat{\Omega}_i \text{ a.s., } i = 1, \cdots, m, \end{aligned}$$

where

$$\begin{aligned} \hat{\Omega}_i &:= \left\{\omega \in \Omega_i' \,\middle|\, (y_i^*)^2 + F_i^2(x^*, y^*, \omega) \neq 0\right\}, \\ \Omega_i' &:= \left\{\omega \in \Omega \,\middle|\, y_i^* F_i(x^*, y^*, \omega) = 0\right\} \end{aligned}$$

for each i.

Note that, for $i \in \mathcal{I}_Y^*$, Ω_i' is equal to Ω, and hence $\hat{\Omega}_i$ equals the set $\{\omega \in \Omega \mid F_i(x^*, y^*, \omega) \neq 0\}$. The main convergence result can be stated as follows.

Theorem 4.2 *Suppose $\nabla_{(x,y)} f, F, \nabla_{(x,y)} F$ are all Hölder continuous in (x, y) on $\mathcal{F}$ with order $\tau > 0$ and Hölder constant $\kappa(\omega)$ satisfying $\mathbb{E}[\kappa(\omega)] < +\infty$ and $\lim_{k\to\infty} \rho_k = +\infty$. Let (x^k, y^k) be a Karush-Kuhn-Tucker point of (3.2) for each k and $(x^*, y^*) \in \mathcal{E}$ be an accumulation point of $\{(x^k, y^k)\}$. Suppose that the GMFCQ holds at (x^*, y^*), and for each $i \in \mathcal{I}_Y^*$, either $p(\omega \in \Omega : F_i(x^*, y^*, \omega) = 0) = 0$ or $p(\omega \in \Omega : F_i(x^*, y^*, \omega) > 0) > 0$ holds. Then (x^*, y^*) is a stationary point of problem (1.3) with probability one.*

Proof. Without loss of generality, we suppose that $\lim_{k\to\infty}(x^k, y^k) = (x^*, y^*)$. Since (x^k, y^k) is a Karush-Kuhn-Tucker point of problem (3.2), there must exist Lagrangian multiplier vectors $\alpha^k \in \Re^{s_1}, \beta^k \in \Re^{s_2}$, and $\gamma^k \in \Re^m$ such that

$$\begin{aligned} 0 &= \frac{1}{k}\sum_{\ell=1}^k \Big(\nabla_x f(x^k, y^k, \omega_\ell) - 2\sigma \nabla_x F(x^k, y^k, \omega_\ell) u(x^k, y^k, \omega_\ell) \\ &\quad + 2\rho_k \nabla_x F(x^k, y^k, \omega_\ell) \mathrm{diag}(y_1^k, \cdots, y_m^k)(y^k \circ v(x^k, y^k, \omega_\ell))\Big) + \nabla g(x^k)\alpha^k + \nabla h(x^k)\beta^k, \end{aligned} \tag{4.11}$$

$$\begin{aligned} 0 &= \frac{1}{k}\sum_{\ell=1}^k \Big(\nabla_y f(x^k, y^k, \omega_\ell) - 2\sigma \nabla_y F(x^k, y^k, \omega_\ell) u(x^k, y^k, \omega_\ell) \\ &\quad + 2\rho_k \Big(\nabla_y F(x^k, y^k, \omega_\ell) \mathrm{diag}(y_1^k, \cdots, y_m^k) \\ &\quad + \mathrm{diag}(v_1(x^k, y^k, \omega_\ell), \cdots, v_m(x^k, y^k, \omega_\ell))\Big)(y^k \circ v(x^k, y^k, \omega_\ell))\Big) - \gamma^k, \end{aligned} \tag{4.12}$$

$$0 \;\leq\; \alpha^k \;\perp\; -g(x^k) \;\geq\; 0, \tag{4.13}$$

$$\beta^k \;:\; \text{free}, \qquad h(x^k) = 0, \tag{4.14}$$

$$0 \;\leq\; \gamma^k \;\perp\; y^k \;\geq\; 0. \tag{4.15}$$

(i) We first show that the sequences $\{\alpha^k\}$, $\{\beta^k\}$, $\{\gamma^k\}$, and $\Big\{\frac{1}{k}\sum_{\ell=1}^k \rho_k y_i^k v_i(x^k, y^k, \omega_\ell)\Big\}, i = 1, \cdots, m$, are all bounded with probability one. To this end, we let

$$\tau_k := \sum_{i=1}^m \frac{1}{k}\sum_{\ell=1}^k 2\rho_k y_i^k v_i(x^k, y^k, \omega_\ell) + \sum_{i=1}^{s_1} \alpha_i^k + \sum_{i=1}^{s_2} |\beta_i^k| + \sum_{i=1}^m \gamma_i^k. \tag{4.16}$$

Suppose that there is an unbounded sequence among the above sequences. Then, taking a subsequence if necessary, we may assume that $\lim_{k\to\infty} \tau_k = +\infty$. Note that, by the definition (3.3), there holds $(v_i(x, y, \omega))^2 = v_i(x, y, \omega)F_i(x, y, \omega)$ for any (x, y, ω) and any i. Then, we can rewrite (4.11) and (4.12) as follows:

$$\begin{aligned} & -\frac{1}{k}\sum_{\ell=1}^k \Big(\nabla_{(x,y)} f(x^k, y^k, \omega_\ell) - 2\sigma \nabla_{(x,y)} F(x^k, y^k, \omega_\ell) u(x^k, y^k, \omega_\ell)\Big) \\ = \;& \frac{1}{k}\sum_{\ell=1}^k \sum_{i=1}^m 2\rho_k y_i^k v_i(x^k, y^k, \omega_\ell) \begin{pmatrix} y_i^k \nabla_x F_i(x^k, y^k, \omega_\ell) \\ y_i^k \nabla_y F_i(x^k, y^k, \omega_\ell) + F_i(x^k, y^k, \omega_\ell) e_i \end{pmatrix} \\ & + \sum_{i=1}^{s_1} \alpha_i^k \begin{pmatrix} \nabla g_i(x^k) \\ 0 \end{pmatrix} + \sum_{i=1}^{s_2} \beta_i^k \begin{pmatrix} \nabla h_i(x^k) \\ 0 \end{pmatrix} - \sum_{i=1}^m \gamma_i^k \begin{pmatrix} 0 \\ e_i \end{pmatrix}. \end{aligned} \tag{4.17}$$

Dividing both sides of (4.17) and (4.16) by τ_k, we get

$$\begin{pmatrix} -\frac{1}{k\tau_k}\sum_{\ell=1}^k(\nabla_{(x,y)}f(x^k,y^k,\omega_\ell)-2\sigma\nabla_{(x,y)}F(x^k,y^k,\omega_\ell)u(x^k,y^k,\omega_\ell)) \\ 1 \end{pmatrix}$$

$$\begin{aligned} = \quad & \sum_{i=1}^m \frac{1}{k}\sum_{\ell=1}^k \frac{2\rho_k y_i^k v_i(x^k,y^k,\omega_\ell)}{\tau_k}\begin{pmatrix} y_i^k\nabla_x F_i(x^k,y^k,\omega_\ell) \\ y_i^k\nabla_y F_i(x^k,y^k,\omega_\ell)+F_i(x^k,y^k,\omega_\ell)e_i \\ 1 \end{pmatrix} \\ & +\sum_{i=1}^{s_1}\frac{\alpha_i^k}{\tau_k}\begin{pmatrix}\nabla g_i(x^k) \\ 0 \\ 1\end{pmatrix}+\sum_{i=1}^{s_2}\frac{\beta_i^k}{\tau_k}\begin{pmatrix}\nabla h_i(x^k) \\ 0 \\ \mathrm{sign}(\beta_i^k)\end{pmatrix}-\sum_{i=1}^m\frac{\gamma_i^k}{\tau_k}\begin{pmatrix}0 \\ e_i \\ 1\end{pmatrix}. \end{aligned}$$

Let B be an arbitrary subset of Ω with $p(B)=0$. We then have

$$\begin{aligned} & \begin{pmatrix} -\frac{1}{k\tau_k}\sum_{\ell=1}^k(\nabla_{(x,y)}f(x^k,y^k,\omega_\ell)-2\sigma\nabla_{(x,y)}F(x^k,y^k,\omega_\ell)u(x^k,y^k,\omega_\ell)) \\ 1 \end{pmatrix} \\ & -\sum_{i=1}^m \frac{1}{k}\sum_{\substack{1\le\ell\le k \\ \omega_\ell\in B}} \frac{2\rho_k y_i^k v_i(x^k,y^k,\omega_\ell)}{\tau_k}\begin{pmatrix} y_i^k\nabla_x F_i(x^k,y^k,\omega_\ell) \\ y_i^k\nabla_y F_i(x^k,y^k,\omega_\ell)+F_i(x^k,y^k,\omega_\ell)e_i \\ 1 \end{pmatrix} \\ = \quad & \sum_{i=1}^m \frac{1}{k}\sum_{\substack{1\le\ell\le k \\ \omega_\ell\notin B}} \frac{2\rho_k y_i^k v_i(x^k,y^k,\omega_\ell)}{\tau_k}\begin{pmatrix} y_i^k\nabla_x F_i(x^k,y^k,\omega_\ell) \\ y_i^k\nabla_y F_i(x^k,y^k,\omega_\ell)+F_i(x^k,y^k,\omega_\ell)e_i \\ 1 \end{pmatrix} \\ & +\sum_{i=1}^{s_1}\frac{\alpha_i^k}{\tau_k}\begin{pmatrix}\nabla g_i(x^k) \\ 0 \\ 1\end{pmatrix}+\sum_{i=1}^{s_2}\frac{\beta_i^k}{\tau_k}\begin{pmatrix}\nabla h_i(x^k) \\ 0 \\ \mathrm{sign}(\beta_i^k)\end{pmatrix}-\sum_{i=1}^m\frac{\gamma_i^k}{\tau_k}\begin{pmatrix}0 \\ e_i \\ 1\end{pmatrix}. \end{aligned} \tag{4.18}$$

Note that, by the assumptions, there hold

$$\begin{aligned} \|\nabla_{(x,y)}f(x,y,\omega)-\nabla_{(x,y)}f(x',y',\omega)\| &\le \kappa(\omega)\|(x,y)-(x',y')\|^\tau, \\ \|F(x,y,\omega)-F(x',y',\omega)\| &\le \kappa(\omega)\|(x,y)-(x',y')\|^\tau, \\ \|\nabla_{(x,y)}F(x,y,\omega)-\nabla_{(x,y)}F(x',y',\omega)\| &\le \kappa(\omega)\|(x,y)-(x',y')\|^\tau \end{aligned}$$

for any $(x,y)\in\mathcal{F},(x',y')\in\mathcal{F}$, and any $\omega\in\Omega$. Therefore, for each k, we have

$$\begin{aligned} & \lim_{k\to\infty}\left\|\frac{1}{k}\sum_{\ell=1}^k\Big(\nabla_{(x,y)}f(x^k,y^k,\omega_\ell)-\nabla_{(x,y)}f(x^*,y^*,\omega_\ell)\Big)\right\| \\ \le\quad & \lim_{k\to\infty}\frac{1}{k}\sum_{\ell=1}^k\Big\|\nabla_{(x,y)}f(x^k,y^k,\omega_\ell)-\nabla_{(x,y)}f(x^*,y^*,\omega_\ell)\Big\| \\ \le\quad & \lim_{k\to\infty}\|(x^k,y^k)-(x^*,y^*)\|^\tau\,\frac{1}{k}\sum_{\ell=1}^k\kappa(\omega_\ell) \\ =\quad & 0 \qquad \text{w.p.1.} \end{aligned}$$

It then follows that

$$\begin{aligned} \lim_{k\to\infty}\frac{1}{k}\sum_{\ell=1}^k\nabla_{(x,y)}f(x^k,y^k,\omega_\ell) &= \lim_{k\to\infty}\frac{1}{k}\sum_{\ell=1}^k\nabla_{(x,y)}f(x^*,y^*,\omega_\ell) \\ &= \int_\Omega\nabla_{(x,y)}f(x^*,y^*,\omega)\,dp \qquad \text{w.p.1,} \end{aligned} \tag{4.19}$$

where the last equality follows from (3.1). Moreover, since

$$
\begin{aligned}
& \lim_{k\to\infty} \Big\| \frac{1}{k}\sum_{\ell=1}^{k}\Big(\nabla_{(x,y)}F(x^k,y^k,\omega_\ell)u(x^k,y^k,\omega_\ell) - \nabla_{(x,y)}F(x^*,y^*,\omega_\ell)u(x^*,y^*,\omega_\ell)\Big)\Big\| \\
\le\ & \lim_{k\to\infty} \frac{1}{k}\sum_{\ell=1}^{k}\Big\|\nabla_{(x,y)}F(x^k,y^k,\omega_\ell)\Big(u(x^k,y^k,\omega_\ell)-u(x^*,y^*,\omega_\ell)\Big) \\
& \qquad + \Big(\nabla_{(x,y)}F(x^k,y^k,\omega_\ell)-\nabla_{(x,y)}F(x^*,y^*,\omega_\ell)\Big)u(x^*,y^*,\omega_\ell)\Big\| \\
\le\ & \lim_{k\to\infty} \frac{1}{k}\sum_{\ell=1}^{k}\Big(\|\nabla_{(x,y)}F(x^k,y^k,\omega_\ell)\|\,\|F(x^k,y^k,\omega_\ell)-F(x^*,y^*,\omega_\ell)\| \\
& \qquad + \|\nabla_{(x,y)}F(x^k,y^k,\omega_\ell)-\nabla_{(x,y)}F(x^*,y^*,\omega_\ell)\|\,\|u(x^*,y^*,\omega_\ell)\|\Big) \\
\le\ & \lim_{k\to\infty} 2\,C\,\|(x^k,y^k)-(x^*,y^*)\|^\tau\,\frac{1}{k}\sum_{\ell=1}^{k}\kappa(\omega_\ell) \\
=\ & 0 \qquad \text{w.p.1,}
\end{aligned}
$$

where $C>0$ is an upper bound of $\{\nabla_{(x,y)}F(x^k,y^k,\omega_\ell)\}$ and $\{u(x^*,y^*,\omega_\ell)\}$, we have

$$
\begin{aligned}
\lim_{k\to\infty}\frac{1}{k}\sum_{\ell=1}^{k}\nabla_{(x,y)}F(x^k,y^k,\omega_\ell)u(x^k,y^k,\omega_\ell) &= \lim_{k\to\infty}\frac{1}{k}\sum_{\ell=1}^{k}\nabla_{(x,y)}F(x^*,y^*,\omega_\ell)u(x^*,y^*,\omega_\ell) \\
&= \int_\Omega \nabla_{(x,y)}F(x^*,y^*,\omega)u(x^*,y^*,\omega)\,dp \quad \text{w.p.1.} \qquad (4.20)
\end{aligned}
$$

It follows from (4.19) and (4.20) that

$$
\lim_{k\to\infty}\frac{1}{k\tau_k}\sum_{\ell=1}^{k}\Big(\nabla_{(x,y)}f(x^k,y^k,\omega_\ell)-2\sigma\nabla_{(x,y)}F(x^k,y^k,\omega_\ell)u(x^k,y^k,\omega_\ell)\Big)=0 \qquad \text{w.p.1.} \quad (4.21)
$$

Now let l_k^B be the cardinality of the sample subset $\{\omega_\ell\in B\mid \ell=1,\cdots,k\}$. Since $p(B)=0$, there almost surely holds $\lim\limits_{k\to\infty}\frac{l_k^B}{k}=0$. Therefore, taking into account the fact that the sequence

$$
\left\{\sum_{i=1}^{m}\frac{2\rho_k y_i^k v_i(x^k,y^k,\omega_\ell)}{\tau_k}\begin{pmatrix} y_i^k\nabla_x F_i(x^k,y^k,\omega_\ell) \\ y_i^k\nabla_y F_i(x^k,y^k,\omega_\ell)+F_i(x^k,y^k,\omega_\ell)e_i \\ 1\end{pmatrix}\right\}
$$

is bounded, we have

$$
\lim_{k\to\infty}\frac{1}{k}\sum_{\substack{1\le\ell\le k\\ \omega_\ell\in B}}\sum_{i=1}^{m}\frac{2\rho_k y_i^k v_i(x^k,y^k,\omega_\ell)}{\tau_k}\begin{pmatrix} y_i^k\nabla_x F_i(x^k,y^k,\omega_\ell) \\ y_i^k\nabla_y F_i(x^k,y^k,\omega_\ell)+F_i(x^k,y^k,\omega_\ell)e_i \\ 1\end{pmatrix}=0 \quad \text{w.p.1.} \quad (4.22)
$$

Furthermore, taking a subsequence if necessary, we may assume that the limits

$$
\bar\alpha := \lim_{k\to\infty}\frac{\alpha^k}{\tau_k}, \qquad \bar\beta := \lim_{k\to\infty}\frac{\beta^k}{\tau_k}, \qquad \bar\gamma := \lim_{k\to\infty}\frac{\gamma^k}{\tau_k} \qquad (4.23)
$$

exist. Thus, letting $k \to +\infty$ in (4.18), we have

$$
\begin{aligned}
&\lim_{k\to\infty}\sum_{i=1}^{m}\frac{1}{k}\sum_{\substack{1\le \ell\le k\\ \omega_\ell\notin B}}\frac{2\rho_k y_i^k v_i(x^k,y^k,\omega_\ell)}{\tau_k}\begin{pmatrix} y_i^k\nabla_x F_i(x^k,y^k,\omega_\ell)\\ y_i^k\nabla_y F_i(x^k,y^k,\omega_\ell)+F_i(x^k,y^k,\omega_\ell)e_i\\ 1\end{pmatrix}\\
=\;&\begin{pmatrix}0\\0\\1\end{pmatrix}-\sum_{i=1}^{s_1}\bar{\alpha}_i\begin{pmatrix}\nabla g_i(x^*)\\0\\1\end{pmatrix}-\sum_{i=1}^{s_2}\bar{\beta}_i\begin{pmatrix}\nabla h_i(x^*)\\0\\ \mathrm{sign}(\beta_i^k)\end{pmatrix}+\sum_{i=1}^{m}\bar{\gamma}_i\begin{pmatrix}0\\e_i\\1\end{pmatrix}\qquad \text{w.p.1.}
\end{aligned}\tag{4.24}
$$

In addition, in a similar way to (4.19) and (4.20), we can show that, for each i, there holds

$$
\begin{aligned}
&\lim_{k\to\infty}\frac{1}{k}\sum_{\substack{1\le \ell\le k\\ \omega_\ell\notin B}}\frac{2\rho_k y_i^k v_i(x^k,y^k,\omega_\ell)}{\tau_k}\begin{pmatrix} y_i^k\nabla_x F_i(x^k,y^k,\omega_\ell)\\ y_i^k\nabla_y F_i(x^k,y^k,\omega_\ell)+F_i(x^k,y^k,\omega_\ell)e_i\\ 1\end{pmatrix}\\
=\;&\lim_{k\to\infty}\frac{1}{k}\sum_{\substack{1\le \ell\le k\\ \omega_\ell\notin B}}\frac{2\rho_k y_i^k v_i(x^k,y^k,\omega_\ell)}{\tau_k}\begin{pmatrix} y_i^*\nabla_x F_i(x^*,y^*,\omega_\ell)\\ y_i^*\nabla_y F_i(x^*,y^*,\omega_\ell)+F_i(x^*,y^*,\omega_\ell)e_i\\ 1\end{pmatrix}
\end{aligned}\tag{4.25}
$$

with probability one. Note that $y_i^* F_i(x^*,y^*,\omega)\le 0, \omega\in\Omega$ a.s., for each i, since $(x^*,y^*)\in\mathcal{E}$ by the given assumption. For each i and any $\epsilon>0$, let

$$
\Omega_i^\epsilon := \left\{\omega\in\Omega \;\middle|\; y_i^* F_i(x^*,y^*,\omega)\le -\epsilon\right\}.
$$

If Ω_i^ϵ is nonempty for some i, we have $F_i(x^*,y^*,\omega)\le -\epsilon/y_i^* < 0$ for every $\omega\in\Omega_i^\epsilon$. Since Ω_i^ϵ is a closed subset of the compact set Ω, there must exist a neighborhood U^* of (x^*,y^*) such that the function F_i is uniformly continuous on $U^*\times\Omega_i^\epsilon$. It then follows that, when k is sufficiently large,

$$
v_i(x^k,y^k,\omega)=\max\left\{F_i(x^k,y^k,\omega),0\right\}=0 \tag{4.26}
$$

holds for every $\omega\in\Omega_i^\epsilon$. Therefore, we have from (4.24)–(4.26) that

$$
\begin{aligned}
&\begin{pmatrix}0\\0\\1\end{pmatrix}-\sum_{i=1}^{s_1}\bar{\alpha}_i\begin{pmatrix}\nabla g_i(x^*)\\0\\1\end{pmatrix}-\sum_{i=1}^{s_2}\bar{\beta}_i\begin{pmatrix}\nabla h_i(x^*)\\0\\ \mathrm{sign}(\beta_i^k)\end{pmatrix}+\sum_{i=1}^{m}\bar{\gamma}_i\begin{pmatrix}0\\e_i\\1\end{pmatrix}\\
=\;&\sum_{i=1}^{m}\lim_{k\to\infty}\frac{1}{k}\sum_{\substack{1\le \ell\le k\\ \omega_\ell\notin B}}\frac{2\rho_k y_i^k v_i(x^k,y^k,\omega_\ell)}{\tau_k}\begin{pmatrix} y_i^*\nabla_x F_i(x^*,y^*,\omega_\ell)\\ y_i^*\nabla_y F_i(x^*,y^*,\omega_\ell)+F_i(x^*,y^*,\omega_\ell)e_i\\ 1\end{pmatrix}\\
=\;&\sum_{i=1}^{m}\lim_{k\to\infty}\frac{1}{k}\sum_{\substack{1\le \ell\le k\\ \omega_\ell\notin B,\omega_\ell\notin\Omega_i^\epsilon}}\frac{2\rho_k y_i^k v_i(x^k,y^k,\omega_\ell)}{\tau_k}\begin{pmatrix} y_i^*\nabla_x F_i(x^*,y^*,\omega_\ell)\\ y_i^*\nabla_y F_i(x^*,y^*,\omega_\ell)+F_i(x^*,y^*,\omega_\ell)e_i\\ 1\end{pmatrix}\\
\in\;&\sum_{i=1}^{m}\overline{\mathrm{cone}}\left\{\begin{pmatrix} y_i^*\nabla_x F_i(x^*,y^*,\omega)\\ y_i^*\nabla_y F_i(x^*,y^*,\omega)+F_i(x^*,y^*,\omega)e_i\\ 1\end{pmatrix}\;\middle|\;\omega\in\Omega\setminus(B\cup\Omega_i^\epsilon)\right\}\qquad\text{w.p.1.}
\end{aligned}
$$

The arbitrariness of B and ϵ yields

$$
\begin{aligned}
&\begin{pmatrix}0\\0\\1\end{pmatrix}-\sum_{i=1}^{s_1}\bar{\alpha}_i\begin{pmatrix}\nabla g_i(x^*)\\0\\1\end{pmatrix}-\sum_{i=1}^{s_2}\bar{\beta}_i\begin{pmatrix}\nabla h_i(x^*)\\0\\\mathrm{sign}(\beta_i^k)\end{pmatrix}+\sum_{i=1}^{m}\bar{\gamma}_i\begin{pmatrix}0\\e_i\\1\end{pmatrix}\\
\in\ &\sum_{i=1}^{m}\bigcap_{\substack{B\subset\Omega\\p(B)=0}}\overline{\mathrm{cone}}\left\{\begin{pmatrix}y_i^*\nabla_x F_i(x^*,y^*,\omega)\\y_i^*\nabla_y F_i(x^*,y^*,\omega)+F_i(x^*,y^*,\omega)e_i\\1\end{pmatrix}\ \middle|\ \omega\in\Omega_i'\setminus B\right\}\quad \text{w.p.1.}
\end{aligned}
$$

In consequence, as shown in Section 2, there almost surely exist some finite-valued nonnegative measurable functions $\bar{\delta}_i, i=1,\cdots,m$, defined on Ω_i' such that

$$
\begin{aligned}
&\begin{pmatrix}0\\0\\1\end{pmatrix}-\sum_{i=1}^{s_1}\bar{\alpha}_i\begin{pmatrix}\nabla g_i(x^*)\\0\\1\end{pmatrix}-\sum_{i=1}^{s_2}\bar{\beta}_i\begin{pmatrix}\nabla h_i(x^*)\\0\\\mathrm{sign}(\beta_i^k)\end{pmatrix}+\sum_{i=1}^{m}\bar{\gamma}_i\begin{pmatrix}0\\e_i\\1\end{pmatrix}\\
=\ &\sum_{i=1}^{m}\int_{\Omega_i'}\begin{pmatrix}y_i^*\nabla_x F_i(x^*,y^*,\omega)\\y_i^*\nabla_y F_i(x^*,y^*,\omega)+F_i(x^*,y^*,\omega)e_i\\1\end{pmatrix}\bar{\delta}_i(\omega)dp,
\end{aligned}
$$

which is equivalent to

$$
\begin{aligned}
0\ =\ &\sum_{i=1}^{m}\int_{\Omega_i'}\begin{pmatrix}y_i^*\nabla_x F_i(x^*,y^*,\omega)\\y_i^*\nabla_y F_i(x^*,y^*,\omega)+F_i(x^*,y^*,\omega)e_i\end{pmatrix}\bar{\delta}_i(\omega)dp\\
&+\sum_{i=1}^{s_1}\bar{\alpha}_i\begin{pmatrix}\nabla g_i(x^*)\\0\end{pmatrix}+\sum_{i=1}^{s_2}\bar{\beta}_i\begin{pmatrix}\nabla h_i(x^*)\\0\end{pmatrix}-\sum_{i=1}^{m}\bar{\gamma}_i\begin{pmatrix}0\\e_i\end{pmatrix}, \qquad (4.27)\\
1\ =\ &\sum_{i=1}^{m}\int_{\Omega_i'}\bar{\delta}_i(\omega)dp+\sum_{i=1}^{s_1}\bar{\alpha}_i+\sum_{i=1}^{s_2}|\bar{\beta}_i|+\sum_{i=1}^{m}\bar{\gamma}_i. \qquad (4.28)
\end{aligned}
$$

Note that, by (4.13) and (4.15), there hold

$$\bar{\alpha}_i=0,\qquad i\notin\mathcal{I}_g(x^*), \tag{4.29}$$

$$\bar{\gamma}_i=0,\qquad i\notin\mathcal{I}_Y^*. \tag{4.30}$$

By taking into account the definition of $\hat{\Omega}_i$ along with (4.29) and (4.30), we may rewrite (4.27) as

$$
\begin{aligned}
0\ =\ &\sum_{i=1}^{m}\int_{\hat{\Omega}_i}\begin{pmatrix}y_i^*\nabla_x F_i(x^*,y^*,\omega)\\y_i^*\nabla_y F_i(x^*,y^*,\omega)+F_i(x^*,y^*,\omega)e_i\end{pmatrix}\bar{\delta}_i(\omega)dp\\
&+\sum_{i\in\mathcal{I}_g(x^*)}\bar{\alpha}_i\begin{pmatrix}\nabla g_i(x^*)\\0\end{pmatrix}+\sum_{i=1}^{s_2}\bar{\beta}_i\begin{pmatrix}\nabla h_i(x^*)\\0\end{pmatrix}-\sum_{i\in\mathcal{I}_Y^*}\bar{\gamma}_i\begin{pmatrix}0\\e_i\end{pmatrix}.
\end{aligned}
$$

Since the GMFCQ holds at (x^*,y^*), it follows that

$$
\begin{aligned}
\bar{\alpha}_i&=0, && i\in\mathcal{I}_g(x^*),\\
\bar{\beta}_i&=0, && i=1,\cdots,s_2,\\
\bar{\gamma}_i&=0, && i\in\mathcal{I}_Y^*,
\end{aligned}
$$

and

$$\bar{\delta}_i(\omega) = 0, \qquad \forall \omega \in \hat{\Omega}_i \tag{4.31}$$

for each i. This together with (4.28)–(4.30) yields

$$1 = \sum_{i=1}^{m} \int_{\Omega_i' \setminus \hat{\Omega}_i} \bar{\delta}_i(\omega) dp = \sum_{i \in \mathcal{I}_Y^*} \int_{\Omega_i' \setminus \hat{\Omega}_i} \bar{\delta}_i(\omega) dp,$$

where the second equality follows from the fact that $\Omega_i' \setminus \hat{\Omega}_i$ must be empty when $i \notin \mathcal{I}_Y^*$. Therefore, there exists an index $i_0 \in \mathcal{I}_Y^*$ such that

$$\int_{\Omega_{i_0}' \setminus \hat{\Omega}_{i_0}} \bar{\delta}_{i_0}(\omega) dp > 0. \tag{4.32}$$

This indicates that $p(\Omega_{i_0}' \setminus \hat{\Omega}_{i_0}) = p(\omega \in \Omega : F_{i_0}(x^*, y^*, \omega) = 0) > 0$. Then, from the assumptions of the theorem, $p(\omega \in \hat{\Omega}_{i_0} : F_{i_0}(x^*, y^*, \omega) > 0)$ is positive. We further can choose a number $\bar{\epsilon} > 0$ such that the probability measure of the set $\bar{\Omega}_{i_0} := \{\omega \in \hat{\Omega}_{i_0} \mid F_{i_0}(x^*, y^*, \omega) \geq \bar{\epsilon}\}$ is also positive. Note that both $\Omega_{i_0}' \setminus \hat{\Omega}_{i_0}$ and $\bar{\Omega}_{i_0}$ are compact. It follows from the definition (3.3) that, when k is sufficiently large,

$$v_{i_0}(x^k, y^k, \omega') \leq \bar{\epsilon} \leq v_{i_0}(x^k, y^k, \omega'')$$

holds for any $\omega' \in \Omega_{i_0}' \setminus \hat{\Omega}_{i_0}$ and any $\omega'' \in \bar{\Omega}_{i_0}$. Let k be sufficiently large. Noting that both $p(\Omega_{i_0}' \setminus \hat{\Omega}_{i_0})$ and $p(\bar{\Omega}_{i_0})$ are positive, we may choose independently and identically distributed random samples, denoted by $\{\omega_1', \cdots, \omega_k'\}$ and $\{\omega_1'' \cdots, \omega_k''\}$, respectively, from $\Omega_{i_0}' \setminus \hat{\Omega}_{i_0}$ and $\bar{\Omega}_{i_0}$. Since $y_{i_0}^k \geq 0$, it then follows that

$$\frac{1}{k} \sum_{\ell=1}^{k} \frac{2\rho_k y_{i_0}^k v_{i_0}(x^k, y^k, \omega_\ell')}{\tau_k} \leq \frac{1}{k} \sum_{\ell=1}^{k} \frac{2\rho_k y_{i_0}^k v_{i_0}(x^k, y^k, \omega_\ell'')}{\tau_k}.$$

Letting $k \to +\infty$ and taking into (3.1) account, we have

$$\int_{\Omega_{i_0}' \setminus \hat{\Omega}_{i_0}} \bar{\delta}_{i_0}(\omega) dp \leq \int_{\bar{\Omega}_{i_0}} \bar{\delta}_{i_0}(\omega) dp = 0$$

with probability one, where the equality follows from (4.31). This contradicts (4.32). Hence, the sequences $\{\alpha^k\}$, $\{\beta^k\}$, $\{\gamma^k\}$, and $\left\{\frac{1}{k} \sum_{\ell=1}^{k} \rho_k y_i^k v_i(x^k, y^k, \omega_\ell)\right\}, i = 1, \cdots, m$, are all bounded with probability one.

(ii) We next show that there exist multiplier vectors $\alpha^* \in \Re^{s_1}, \beta^* \in \Re^{s_2}, \gamma^* \in \Re^m$, and a multiplier function $\delta^* : \Omega \to \Re^m$ such that there hold (2.6)–(2.11) with probability one. First of all, without loss of generality, we may assume that the following limits exist:

$$\alpha^* := \lim_{k \to \infty} \alpha^k, \quad \beta^* := \lim_{k \to \infty} \beta^k, \quad \gamma^* := \lim_{k \to \infty} \gamma^k.$$

Recall that the sequences $\left\{\frac{1}{k}\sum_{\ell=1}^{k}\rho_k y_i^k v_i(x^k, y^k, \omega_\ell)\right\}, i = 1, \cdots, m$, are all bounded with probability one. In a similar way to (4.27) and taking (4.19)–(4.20) into account, we can get from (4.17) that there almost surely exist some finite-valued nonnegative measurable functions $\delta_i^*, i = 1, \cdots, m$, defined on Ω_i' such that

$$\begin{aligned}
& -\int_\Omega \Big(\nabla_{(x,y)} f(x^*, y^*, \omega) - 2\sigma\nabla_{(x,y)} F(x^*, y^*, \omega) u(x^*, y^*, \omega)\Big)\, dp \\
= & \sum_{i=1}^{m}\int_{\Omega_i'}\begin{pmatrix} y_i^*\nabla_x F_i(x^*, y^*, \omega) \\ y_i^*\nabla_y F_i(x^*, y^*, \omega) + F_i(x^*, y^*, \omega)e_i\end{pmatrix}\delta_i^*(\omega)dp \\
& + \sum_{i=1}^{s_1}\alpha_i^*\begin{pmatrix}\nabla g_i(x^*) \\ 0\end{pmatrix} + \sum_{i=1}^{s_2}\beta_i^*\begin{pmatrix}\nabla h_i(x^*) \\ 0\end{pmatrix} - \sum_{i=1}^{m}\gamma_i^*\begin{pmatrix}0 \\ e_i\end{pmatrix}.
\end{aligned}$$

For each i, we define $\delta_i^*(\omega) := 0$ for $\omega \in \Omega \setminus \Omega_i'$. It then follows that

$$\begin{aligned}
& -\int_\Omega \Big(\nabla_{(x,y)} f(x^*, y^*, \omega) - 2\sigma\nabla_{(x,y)} F(x^*, y^*, \omega) u(x^*, y^*, \omega)\Big)\, dp \\
= & \sum_{i=1}^{m}\int_{\Omega}\begin{pmatrix} y_i^*\nabla_x F_i(x^*, y^*, \omega) \\ y_i^*\nabla_y F_i(x^*, y^*, \omega) + F_i(x^*, y^*, \omega)e_i\end{pmatrix}\delta_i^*(\omega)dp \\
& + \sum_{i=1}^{s_1}\alpha_i^*\begin{pmatrix}\nabla g_i(x^*) \\ 0\end{pmatrix} + \sum_{i=1}^{s_2}\beta_i^*\begin{pmatrix}\nabla h_i(x^*) \\ 0\end{pmatrix} - \sum_{i=1}^{m}\gamma_i^*\begin{pmatrix}0 \\ e_i\end{pmatrix} \qquad \text{w.p.1.}
\end{aligned} \tag{4.33}$$

Taking (2.25) and (2.26) into account, we obtain (2.6) and (2.7) from (4.33) with probability one. Moreover, (2.8)–(2.10) follow from (4.13)–(4.15) immediately. In addition, it is obvious that $\delta_i^*(\omega) \geq 0$ for any $\omega \in \Omega$. Since $y_i^* F_i(x^*, y^*, \omega) < 0$ is equivalent to $\omega \in \Omega \setminus \Omega_i'$ for each i, (2.11) is also valid.

Therefore, $(\alpha^*, \beta^*, \gamma^*, \delta^*(\cdot))$ satisfies (2.6)–(2.11) with probability one and hence (x^*, y^*) is almost surely a stationary point of problem (1.3). This completes the proof of the theorem. ■

5 Conclusions

We have presented a new formulation (1.2) of the SMPECs with recourse and shown that the new formulation is actually equivalent to a smooth semi-infinite programming problem. We have deduced the optimality conditions for the problems and investigated the connections among the conditions. Then, we have employed a Monte Carlo sampling method and a penalty technique to get some approximations to the problem. Under appropriate assumptions, we have established convergence of the proposed method. Recall that the sample space Ω is assumed to have infinitely many elements. Actually, if Ω has only a finite number of elements, we may present a similar method without resort to a Monte Carlo sampling approximation technique.

Acknowledgement. The authors are grateful to the anonymous referees whose helpful comments and suggestions have led to much improvement of the paper.

References

[1] S.I. Birbil, G. Gürkan, and O. Listes, *Simulation-based solution of stochastic mathematical programs with complementarity constraints: Sample-path analysis*, Econometric Institute Report EI 2004-03, Erasmus University Rotterdam, 2004.

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

[3] C.W. Burrill, Measure, Integration, and Probability, McGraw-Hill, 1972.

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

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

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

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

[8] M.A. Goberna and M.A. Lopez, Linear Semi-Infinite Optimization, John Wiley & Sons, Chichester, 1998.

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

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

[11] R. Hettich and K.O. Kortanek, *Semi-infinite programming: Theory, methods, and applications*, SIAM Review, 35 (1993), 380–429.

[12] V. Jeyakumar, *Characterizing set containments involving infinite convex constraints and reverse-convex constraints*, SIAM Journal on Optimization, 13 (2003), 947–959.

[13] G.H. Lin, X. Chen and M. Fukushima, *Smoothing implicit programming approaches for stochastic mathematical programs with linear complementarity constraints*, Technical Report 2003-006, Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto, Japan, 2003.

[14] G.H. Lin, X. Chen and M. Fukushima, *Solving stochastic mathematical programs with equilibrium constraints via approximation and smoothing implicit programming with penalization*, Mathematical Programming, to appear.

[15] G.H. Lin and M. Fukushima, *A class of stochastic mathematical programs with complementarity constraints: Reformulations and algorithms*, Journal of Industrial and Management Optimization, 1 (2005), 99–122.

[16] G.H. Lin and M. Fukushima, *New reformulations for stochastic complementarity problems*, Optimization Methods and Software, to appear.

[17] G.H. Lin and M. Fukushima, *Regularization method for stochastic mathematical programs with complementarity constraints*, European Series of Applied and Industrial Mathematics (ESAIM): Control, Optimisation and Calculus of Variations, 11 (2005), 252–265.

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

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

[20] J.M. Ortega and W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

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

[22] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, NJ, 1970.

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

[24] A. Shapiro, *Monte Carlo sampling approach to stochastic programming*, European Series of Applied and Industrial Mathematics (ESAIM): Proceedings, 13 (2003), 65–73.

[25] A. Shapiro, *Stochastic mathematical programs with equilibrium constraints*, Journal of Optimization Theory and Applications, to appear.

[26] A. Shapiro and H. Xu, *Stochastic mathematical programs with equilibrium constraints, modeling and sample average approximation*, Preprint, School of Industrial and System Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA, 2005.

[27] H. Xu, *An implicit programming approach for a class of stochastic mathematical programs with linear complementarity constraints*, SIAM Journal on Optimization, to appear.

Appendix: Equivalence between problems (1.2) and (1.3). *If (x^*, y^*) solves problem (1.3), then $(x^*, y^*, u(x^*, y^*, \cdot))$ is an optimal solution of problem (1.2). Conversely, if $(x^*, y^*, z^*(\cdot))$ is an optimal solution of problem (1.2), then (x^*, y^*) solves problem (1.3).*

Proof. (i) Suppose that (x^*, y^*) is an optimal solution of (1.3). We then have from (1.4) that

$$F(x^*, y^*, \omega) + u(x^*, y^*, \omega) \geq 0, \qquad \forall \omega \in \Omega.$$

Note that, if $y_i^* > 0$ for some i, there must hold $F_i(x^*, y^*, \omega) \leq 0$ for almost all $\omega \in \Omega$ and so $u_i(x^*, y^*, \omega) = -F_i(x^*, y^*, \omega)$ for almost all $\omega \in \Omega$. Therefore, we have

$$(y^*)^T (F(x^*, y^*, \omega) + u(x^*, y^*, \omega)) = 0, \quad \omega \in \Omega \ \text{ a.s.}$$

This indicates that $(x^*, y^*, u(x^*, y^*, \cdot))$ is feasible to problem (1.2). Let $(x, y, z(\cdot))$ be an arbitrary feasible point of problem (1.2). It then follows that, for almost every $\omega \in \Omega$,

$$z(\omega) - u(x, y, \omega) = \min\{F(x, y, \omega) + z(\omega), \ z(\omega)\} \geq 0$$

and hence $z(\omega) \geq u(x, y, \omega) \geq 0$. This implies that $\mathbb{E}[\, \|z(\omega)\|^2 - \|u(x, y, \omega)\|^2 \,] \geq 0$. On the other hand, it follows from the feasibility of $(x, y, z(\cdot))$ in problem (1.2) that

$$y \circ F(x, y, \omega) = -y \circ z(\omega) \leq 0, \quad \omega \in \Omega \ \text{ a.s.},$$

and so the point (x, y) is a feasible point of problem (1.3). Thus, we have from the optimality of (x^*, y^*) in (1.3) that

$$\mathbb{E}[f(x, y, \omega) + \sigma \|u(x, y, \omega)\|^2] \ \geq \ \mathbb{E}[f(x^*, y^*, \omega) + \sigma \|u(x^*, y^*, \omega)\|^2].$$

Therefore, there holds

$$\begin{aligned}
& \mathbb{E}[f(x, y, \omega) + \sigma \|z(\omega)\|^2] - \mathbb{E}[f(x^*, y^*, \omega) + \sigma \|u(x^*, y^*, \omega)\|^2] \\
= \ & \mathbb{E}[f(x, y, \omega) + \sigma \|u(x, y, \omega)\|^2] - \mathbb{E}[f(x^*, y^*, \omega) + \sigma \|u(x^*, y^*, \omega)\|^2] + \sigma \mathbb{E}[\|z(\omega)\|^2 - \|u(x, y, \omega)\|^2] \\
\geq \ & 0.
\end{aligned}$$

This indicates that $(x^*, y^*, u(x^*, y^*, \cdot))$ is an optimal solution of problem (1.2).

(ii) Suppose that $(x^*, y^*, z^*(\cdot))$ is an optimal solution of (1.2). It is not difficult to see that $z^*(\omega) = u(x^*, y^*, \omega)$ for almost all $\omega \in \Omega$. Let (x, y) be an arbitrary feasible point of (1.3). In a similar way to (i), we can show that $(x, y, u(x, y, \cdot))$ is feasible to problem (1.2). Since $(x^*, y^*, z^*(\cdot))$ solves (1.2), there holds

$$\begin{aligned} \mathbb{E}[f(x, y, \omega) + \sigma\|u(x, y, \omega)\|^2] &\geq \mathbb{E}[f(x^*, y^*, \omega) + \sigma\|z^*(\omega)\|^2] \\ &= \mathbb{E}[f(x^*, y^*, \omega) + \sigma\|u(x^*, y^*, \omega)\|^2]. \end{aligned}$$

This implies that (x^*, y^*) is an optimal solution of problem (1.3). ■