ASYMPTOTIC CONVERGENCE OF STATIONARY POINTS OF STOCHASTIC MULTIOBJECTIVE PROGRAMS WITH PARAMETRIC VARIATIONAL INEQUALITY CONSTRAINT VIA SAA APPROACH

Liping Pang
School of Mathematical Sciences
Dalian University of Technology
Dalian 116024, China

Fanyun Meng* and Jinhe Wang
School of Information and Control Engineering
Qingdao University of Technology
Qingdao 266033, China

(Communicated by Oleg Prokopyev)

Abstract. We consider the sample average approximation method for a stochastic multiobjective programming problem constrained by parametric variational inequalities. The first order necessary conditions for the original problem and its sample average approximation (SAA) problem are established under constraint qualifications. By graphical convergence of set-valued mappings, the stationary points of the SAA problem converge to the stationary points of the true problem when the sample size tends to infinity. Under proper assumptions, the convergence of optimal solutions of SAA problems is also obtained. The numerical experiments on stochastic multiobjective optimization problems with variational inequalities are given to illustrate the efficiency of SAA estimators.

1. Introduction. We consider the following stochastic multiobjective mathematical program constrained by parametric variational inequality

\[
\begin{align*}
\min & \quad \mathbb{E}[\phi(x, y, \xi(w))] \\
\text{s. t.} & \quad y \in \Gamma, (\mathbb{E}[F(x, y, \xi(w))), z - y) \leq 0, \forall z \in \Gamma, \\
& \quad (x, y) \in C,
\end{align*}
\]

where \(\phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^l\) is locally Lipschitz continuous and \(F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m\) is continuously differentiable. \(\Gamma = \{y \in \mathbb{R}^m | \mathbb{E}[g_i(y, \xi(w))] \leq 0, i = 1, 2, \cdots, q; \mathbb{E}[g_i(y, \xi(w))] = 0, i = q + 1, \cdots, p\}\). The function \(g_i : \mathbb{R}^m \to \mathbb{R}\) is convex, twice continuously differentiable, \(i = 1, 2, \cdots, q\); and \(g_i : \mathbb{R}^m \to \mathbb{R}\) is affine, \(i = q + 1, \cdots, p\). \(C\) is a nonempty closed convex subset of \(\mathbb{R}^n \times \mathbb{R}^m\). \(\xi : \Omega \to \Theta \subset \mathbb{R}^d\) is a vector of random variables defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\). \(\mathbb{E}[\cdot]\) denotes

2010 Mathematics Subject Classification. Primary: 90C30, 90C29; Secondary: 65K10.

Key words and phrases. Stochastic programming, multiobjective programming, parametric variational inequalities, sample average approximation.

The research is partially supported by Huzhou science and technology plan on No.2016GY03, the Natural Science Foundation of China, Grant 11601061.

* Corresponding author: Fan-Yun Meng.
the expected value with respect to the distribution of $\xi$. To simplify the notation, we write $\xi(w)$ as $\xi$ in this context.

Many practical problems can be translated into the above problem (P). For example, consider a firm that use $n$ different resources to produce $m$ products. The firm wishes to maximize the expected profit and minimize the expected excess of employees subjected to its own resources constraints and the maintenance of expected market share constraints. Let $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ be the firm’s production and marketing level. $(u, v) \in \mathbb{R}^n \times \mathbb{R}^m$ is the competitor’s production and marketing level. The variable $\xi$ denotes the market demand which is unknown and is usually determined by some random factors such as prices and consumers. Let the objective be the vector-valued function $\mathbb{E}[\phi(x, y, \xi)] := (-\mathbb{E}[\phi_1(x, y, \xi)], \mathbb{E}[\phi_2(x, y, \xi)])$, where $\mathbb{E}[\phi_1(x, y, \xi)]$ is the expected profit and $\mathbb{E}[\phi_2(x, y, \xi)]$ is the expected excess of employees. The resource utilization of the firm is denoted by $s_j(x, y)$ for product $j$ and the firm’s market share is denoted by $f_i(x, y, u, v, \xi)$ for product $i$. For a given $(x, y)$, the firm’s minimum expected market share for the $i$-th production is given by

$$h_i(x, y) = \min_{(u,v) \in \Gamma} \{\mathbb{E}[f_i(x, y, u, v, \xi)]\},$$

where $\Gamma = \{(u, v) | \mathbb{E}[g(u, v, \xi)] \leq 0\}$ is the feasible set of competitors. Hence the optimal production and market level of the firm can be obtained by the following multiobjective optimization problem

$$\min \ \mathbb{E}[\phi(x, y, \xi)]$$

s. t. $s_j(x, y) \leq a_j, j = 1, 2, \cdots, n,$

$$h_i(x, y) \geq b_i, i = 1, 2, \cdots, m,$$

where $a_j$ denotes the total amount of resource $j$ and $b_i$ denotes the minimum threshold of market share of product $i$. Suppose that the functions $f_i$ and $g$ are differentiable in $(u, v)$ and $\Gamma$ is a convex set, then the above problem is equivalent to the following optimization problem with parametric variational inequality constraint

$$\min \ \mathbb{E}[\phi(x, y, \xi)]$$

s. t. $s_j(x, y) \leq a_j, j = 1, 2, \cdots, n,$

$$h_i(x, y) \geq b_i, i = 1, 2, \cdots, m,$$

$$(u, v) \in \Gamma, (\nabla_{x,y} \mathbb{E}[f_i(x, y, u, v, \xi)], (u, v) - (u', v')) \leq 0,$$

$$\forall (u', v') \in \Gamma, i = 1, 2, \cdots, m.$$
Analogous to ordinary deterministic models, Patriksson [25] firstly studied the stochastic mathematical program with equilibrium constraints. It is a generalization of deterministic MPECs and has been further investigated in [28], [32]. For stochastic multiobjective problems with equilibrium constraints, there is little result on this aspect. A popular method to deal with stochastic programming problems is the sample average approximation (SAA) method where the expected valued is approximated by its sample average. In this paper, we exploit this method to approximate the problem (P). Let $\xi^1, \ldots, \xi^N$ be an independent and identically distribution (iid) sampling of $\xi$, then we consider the following sample approximation

\[
\begin{align*}
\min \quad & \frac{1}{N} \sum_{k=1}^{N} \phi(x, y, \xi^k) \\
\text{s. t.} \quad & y \in \Gamma_N, \left(\frac{1}{N} \sum_{k=1}^{N} F(x, y, \xi^k), z - y\right) \leq 0, \forall z \in \Gamma_N, \\
& (x, y) \in C,
\end{align*}
\]

where $\Gamma_N = \{ y \in \mathbb{R}^m | \frac{1}{N} \sum_{k=1}^{N} g_i(y, \xi^k) \leq 0, i = 1, 2, \ldots, q; \frac{1}{N} \sum_{k=1}^{N} g_i(y, \xi^k) = 0, i = q + 1, \ldots, p \}.$

Sample average approximation methods have been employed in solving stochastic multiobjective programming problems. [9] used the SAA method to solve the one-stage stochastic multiobjective optimization problem. It proposed a scalarization approach to settle the SAA problem and analyzed the convergence of SAA problems when the sample size was increased to infinity. [29] seemed to be first to deal with the two-stage stochastic MPEC by SAA methods and obtained the asymptotic consistency of optimal values and optimal solutions of the problem. Other comprehensive discussion on theory, numerical methods about solving stochastic multiobjective programming problems by SAA refers to [2], [15], [31]. Following the techniques of sample average approximation methods, we use the sample average to approximate the expectations of the objectives and constraints, then analyze the convergence of SAA problems.

For the problems (P) and (SAA-P), we change the multiobjective programming problems into their equivalent scalar optimization problems by evaluation functions. Due to the inclusion relation of subdifferentials for composite functions, the first order necessary condition of the problem is constructed. Under linear independent constraint qualification and strict complementary condition, we obtain the stability of composite set-valued mappings generated by parametric variational inequalities. Utilizing graph convergence results of variational analysis, we prove that the stationary points of (SAA-P) converge to these which are also the stationary points of (P). Furthermore, we establish the convergence of optimal solutions of (SAA-P) by posing some proper conditions.

The rest of this paper is organized as follows. Section 2 gives some basic definitions and results on variational analysis. In section 3, firstly, the first order necessary conditions for the original problem and its SAA problem are provided. Secondly, the asymptotic convergence of stationary points of SAA problems is proved. Furthermore, we obtain the convergence of optimal solutions of SAA problems. Some numerical results on stochastic multiobjective optimization problems with variational inequalities are given in section 4.

2. Preliminaries. This section gives the notations and reviews some results in variational analysis. The detailed discussion on these subjects can be found in [7],
\[ \|M\| := \max_{M \in \mathcal{M}} \|M\| \]

if \( \mathcal{M} \) is a compact set of vectors. Moreover, \( d(x, D) := \inf_{x' \in D} \|x - x'\| \) denotes the distance from a point \( x \) to a set \( D \). For two compact sets \( \mathcal{C} \) and \( \mathcal{D} \), denote by
\[
\mathcal{D}(\mathcal{C}, \mathcal{D}) := \sup_{x \in \mathcal{C}} d(x, \mathcal{D})
\]

the deviation from \( \mathcal{C} \) to \( \mathcal{D} \). \( B(x, \delta) \) denotes the open ball with radius \( \delta \) and center \( x \), which is \( B(x, \delta) := \{ x' : \|x' - x\| < \delta \} \). \( B \) is used to denote a closed unit ball in a finite dimensional space. For a continuously differentiable mapping \( F : \mathbb{R}^n \to \mathbb{R}^m \), \( \mathcal{J}F(x) \) denotes the Jacobian of \( F \) at \( x \).

Let \( S : \mathbb{R}^m \Rightarrow \mathbb{R}^m \) be a set-valued mapping and \( \limsup_{x \to \bar{x}} S(x) \) denotes the Painlevé-Kuratowski outer-limit
\[
\limsup_{x \to \bar{x}} S(x) := \{ v \in \mathbb{R}^m \mid \exists \text{sequence} x_k \to \bar{x}, v_k \to v \text{ with } v_k \in S(x_k), k = 1, 2, \ldots \}. 
\]

Given a nonempty set \( \Omega \subset \mathbb{R}^n \), fix \( \bar{x} \in \Omega \) and define
\[
\hat{N}_\Omega(\bar{x}) = \{ x^* \in \mathbb{R}^n \mid \limsup_{x \to \bar{x}} \frac{\langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0 \}
\]

the regular normal cone (known also as the prenormal cone or Fréchet normal cone) to \( \Omega \) at \( \bar{x} \). The basic normal cone to \( \Omega \) at \( \bar{x} \) is defined by
\[
N_\Omega(\bar{x}) = \limsup_{x \to \bar{x}} \hat{N}_\Omega(x),
\]

which is also well known as the limiting or Mordukhovich normal cone. This two cones agree with the normal cone of convex analysis when \( \Omega \) is convex. The following proposition provides the normal cone to sets with constraint structure.

**Proposition 1.** [26] Let \( \Gamma = \{ x \in \Gamma \mid F(x) \in D \} \) for closed set \( \Gamma \subset \mathbb{R}^n, D \subset \mathbb{R}^m \), a \( C^1 \) mapping \( F : \mathbb{R}^n \to \mathbb{R}^m \). Written componentwise as \( F(x) = (f_1(x), \ldots, f_m(x)) \).

At any \( \bar{x} \in C \), one has
\[
\hat{N}_C(\bar{x}) \supset \left\{ \sum_{i=1}^m y_i \nabla f_i(\bar{x}) \mathbf{1} + z \mid y \in \hat{N}_D(F(\bar{x})), z \in \hat{N}_\Gamma(\bar{x}) \right\},
\]

where \( y = (y_1, \ldots, y_m) \). On the other hand, one has
\[
N_C(\bar{x}) \subset \left\{ \sum_{i=1}^m y_i \nabla f_i(\bar{x}) \mathbf{1} + z \mid y \in N_D(F(\bar{x})), z \in N_\Gamma(\bar{x}) \right\}
\]

at any \( \bar{x} \in C \) satisfying the constraint qualification that
\[
\left\{ \begin{array}{l}
\text{the only vector } y \in N_D(F(\bar{x})) \text{ for which } \sum_{i=1}^m y_i \nabla f_i(\bar{x}) \in N_\Gamma(\bar{x}) \text{ is } y = (0, \ldots, 0). \\
\end{array} \right. 
\]

(1)

If in addition to this constraint qualification the set \( \Gamma \) is regular at \( \bar{x} \) and \( D \) is regular at \( F(\bar{x}) \), then \( C \) is regular at \( \bar{x} \) and
\[
N_C(\bar{x}) = \left\{ \sum_{i=1}^m y_i \nabla f_i(\bar{x}) \mathbf{1} + z \mid y \in N_D(F(\bar{x})), z \in N_\Gamma(\bar{x}) \right\}.
\]
Given a set-valued mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a point $(\bar{x}, \bar{y})$ with its graph $\text{gph} S := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid y \in S(x)\}$, the coderivative $D^* S(x, y) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is defined by
\[
D^* S(x, y)(y^*) = \{x^* \in \mathbb{R}^n \mid (x^*, -y^*) \in \mathcal{N}_{\text{gph} S}(\bar{x}, \bar{y})\}.
\]
Recall that $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ has the (Aubin) Lipschitz-like property around $(\bar{x}, \bar{y}) \in \text{gph} S$ with modulus $l \geq 0$ if there are neighborhoods $U$ of $\bar{x}$ and $V$ of $\bar{y}$ such that
\[
S(x) \cap V \subset S(u) + l\|x - u\|_B, \forall x, u \in U.
\]
A weaker property than Lipschitz-like property is the calmness, which is that $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is said to be calm at $(\bar{x}, \bar{y}) \in \text{gph} S$ with $l \geq 0$ if there are neighborhoods $U$ of $\bar{x}$ and $V$ of $\bar{y}$ such that
\[
S(x) \cap V \subset S(\bar{x}) + l\|x - \bar{x}\|_B, \text{for all } x \in U.
\]
An important condition to fully characterize the Lipschitz-like property and calmness property is the coderivative criterion, which is also called the Mordukhovich criterion, that is
\[
D^* S(\bar{x}, \bar{y})(0) = \{0\}.
\]
Recall that the multiobjective optimization problem (MOP) $\min_{(x, y) \in \Lambda} f(x, y)$, where $f = (f_1, f_2, \cdots, f_l)$ is defined from $\mathbb{R}^n \times \mathbb{R}^m$ to $\mathbb{R}^l$ and $\Lambda$ is the feasible set of (MOP). A point $(\bar{x}, \bar{y}) \in \Lambda$ is said to be

- a Pareto solution if and only if there is no element $(x, y) \in \Lambda$ satisfying $\forall j \in \{1, 2, \cdots, l\}, f_j(x, y) \leq f_j(\bar{x}, \bar{y})$ and $\exists j_0 \in \{1, 2, \cdots, l\}, f_{j_0}(x, y) < f_{j_0}(\bar{x}, \bar{y})$;
- a weakly Pareto solution if and only if there is no element $(x, y) \in \Lambda$ satisfying $\forall j \in \{1, 2, \cdots, l\}, f_j(x, y) < f_j(\bar{x}, \bar{y})$.

The following lemma gives the relations of the solutions between the multiobjective optimization problem and its scalar problem.

**Lemma 2.1.** The function $\psi : \mathbb{R}^l \rightarrow \mathbb{R}$ is defined on the function value space $f(\Lambda)$. If $\psi$ is a strictly monotonically increasing (monotonically increasing) function of $f(x, y)$, then the optimal solution of single-objective optimization problem (SOP) $\min_{(x, y) \in \Lambda} (\psi \circ f)(x, y)$ is equivalent to the Pareto solution (weakly Pareto solution) of (MOP).

**Proof.** It is easy to prove that the Pareto solution (weakly Pareto solution) of (MOP) is the optimal solution of (SOP) and we omit the proof of this part. Suppose that $(\bar{x}, \bar{y})$ is an optimal solution of (SOP). By a contradiction that $(\bar{x}, \bar{y})$ is not the Pareto solution of (MOP), then there exists $(x^*, y^*) \in \Lambda$ such that
\[
\begin{align*}
f_j(x^*, y^*) &\leq f_j(\bar{x}, \bar{y}), \forall j \in \{1, 2, \cdots, l\}, \\
f_{j_0}(x^*, y^*) &< f_{j_0}(\bar{x}, \bar{y}), \exists j_0 \in \{1, 2, \cdots, l\}.
\end{align*}
\]
Because that $\psi$ is strictly monotonically increasing, it follows from (2) that
\[
(\psi \circ f)(x^*, y^*) < (\psi \circ f)(\bar{x}, \bar{y}),
\]
which is a contradiction, and the conclusion holds.

The subdifferentials of Lipschitz composite functions have the following property.
Lemma 2.2. [21] Let \( f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^l \) be strictly Lipschitz at \((\bar{x}, \bar{y})\) and \( \psi : \mathbb{R}^l \to \mathbb{R} \) is Lipschitz continuous around \( f(\bar{x}, \bar{y}) \), then

\[
\partial (\psi \circ f)(\bar{x}, \bar{y}) \subset \bigcup_{u^* \in \partial \psi(f(\bar{x}, \bar{y}))} \partial (u^*, f)(\bar{x}, \bar{y}).
\]

At last we review the outer semicontinuity of set valued mappings. A set-valued mapping \( S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m \) is said to be closed at \( \bar{x} \) if for \( x_k \to \bar{x} \), \( y_k \in S(x_k) \) and \( y_k \to \bar{y} \), it implies \( \bar{y} \in S(\bar{x}) \).

Definition 2.3. A set-valued mapping \( S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m \) is outer semicontinuous (osc) at \( \bar{x} \) if

\[
\limsup_{x \to \bar{x}} S(x) \subset S(\bar{x})
\]

or equivalently \( \limsup_{x \to \bar{x}} S(x) = S(\bar{x}) \).

The outer continuity convergence of set-valued mappings can be described equivalently from several aspects. It helps to prove the outer semicontinuity of composite set-valued mappings.

Proposition 2. Let \( S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m \) be a set-valued mapping, then the following assertions are satisfied:

(i) \( S \) is osc (everywhere) if and only if \( \text{gph} S \) is closed in \( \mathbb{R}^n \times \mathbb{R}^m \);
(ii) \( \limsup_{x \to \bar{x}} S(x^v) \subset S(\bar{x}) \iff \text{D}(S(x^v), S(\bar{x})) \to 0 \) as \( v \to \infty \).

Proof. (i) This equivalence is the result of [26, Theorem 5.7 (a)].

(ii) The equivalence is obtained directly from [11, p.59]. \( \square \)

Proposition 3. Let \( S : \mathbb{R}^n \to \mathbb{R}^m \) be a single-valued continuous mapping and \( T : \mathbb{R}^m \rightrightarrows \mathbb{R}^p \) is a set-valued mapping with a closed graph, then the following statements hold.

(i) The coderivative mapping \( D^*T(\cdot, \cdot)(\cdot) \) is osc on \( \mathbb{R}^n \).
(ii) The composite set-valued mapping \( S \circ T := S(\cdot)D^*T(\cdot, \cdot)(\cdot) \) is osc on \( \mathbb{R}^n \).

Proof, (i) From [31, Proposition 2.7], the graph of \( D^*T(\cdot, \cdot)(\cdot) \) is closed, then the assertion holds from Proposition 2 (i).

(ii) Because of Proposition 2, it only needs to verify the closedness of the graph of \( S \circ T \). Take \((z_k, \vartheta_k) \in \text{gph}(S \circ T)\) and \( \vartheta_k = u_k \zeta_k \), where \( u_k \in S(z_k) \), \( \zeta_k \in D^*T(z_k,v_k)(\eta_k) \), \( z_k \in T(v_k) \). Suppose that \((z_k, v_k) \to (z, v)\), \( u_k \to u \), \( (\zeta_k, \eta_k) \to (\zeta, \eta) \), \( \vartheta_k \to \vartheta \). Since

\[
\zeta_k \in D^*T(z_k,v_k)(\eta_k) \iff (\zeta_k, -\eta_k) \in N_{\text{gph}T}(z_k,v_k)
\]

and the normal cone mapping is osc, we have \((\zeta, -\eta) \in N_{\text{gph}T}(z,v)\) and \( \zeta \in D^*T(z,v)(\eta) \). The mapping \( S \) is continuous, then \( \vartheta = u\zeta \in S(z)D^*T(z,v)(\eta) \), which means \((z, \vartheta) \in \text{gph}(S \circ T)\). This proves the closeness of \( S \circ T \). \( \square \)

3. Asymptotic convergence of stationary points. In this section we intend to investigate the asymptotic convergence of stationary points of (SAA-P) when the sample size tends to infinity. Firstly, the first order necessary conditions for the original problem (P) and its SAA problem (SAA-P) are established. Secondly, we prove that the stationary points of (SAA-P) converge to these which are also the stationary points of (P). At last, we give the convergence of optimal solutions of (SAA-P).
3.1. First order necessary condition for the original problem. Before the main discussion, some assumptions are made on the objective functions and constraint functions. They are popularly used to analyze stochastic programming problems by SAA methods.

Assumption 3.1. Let \( \phi(x, y, \xi), f(x, y, \xi), g_i(y, \xi), i = 1, 2, \ldots, p \) be defined as in (P).

(A1) \( \mathbb{E}[\phi(x, y, \xi)], \mathbb{E}[f(x, y, \xi)] \) and \( \mathbb{E}[g_i(y, \xi)], i = 1, 2 \ldots, p \), are well defined on \( \mathbb{R}^n \times \mathbb{R}^m \).

(A2) Let \( \vartheta(x, y, \xi) \) and \( \sigma(y, \xi) \) denote any element in the collection of functions \( \{\phi(x, y, \xi), \nabla f(x, y, \xi)\} \) and \( \{\nabla g_i(y, \xi), \nabla^2 g_i(y, \xi), i = 1, 2, \ldots, p\} \). There exist positive measurable functions \( \kappa(\xi) \) and \( \kappa'(\xi) \) such that

\[
\|\vartheta(x', y', \xi) - \vartheta(x, y, \xi)\| \leq \kappa(\xi)(\|x' - x\| + \|y' - y\|), \quad \forall x', x \in \mathbb{R}^n, y', y \in \mathbb{R}^m,
\]

\[
\|\sigma(y', \xi) - \sigma(y, \xi)\| \leq \kappa'(\xi)\|y' - y\|, \quad \forall y', y \in \mathbb{R}^m
\]

for almost every \( \xi \), where \( [\kappa(\xi)] < \infty \) and \( [\kappa'(\xi)] < \infty \).

(A3) There exist nonnegative measurable functions \( \kappa(\xi) < \infty \) and \( \kappa'(\xi) < \infty \) such that

\[
\sup_{x, y} \max\{\|\phi(x, y, \xi)\|\} \leq \kappa(\xi),
\]

\[
\sup_y \max\{\|\rho(y, \xi)\|\} \leq \kappa'(\xi)
\]

for \( \xi \), \( \psi(x, y, \xi) \) and \( \rho(y, \xi) \) are any element in the collection of functions \( \{\phi(x, y, \xi), f(x, y, \xi), \nabla \phi(x, y, \xi), \nabla f(x, y, \xi)\}, \{g_i(y, \xi), \nabla g_i(y, \xi), \nabla^2 g_i(x, y, \xi), i = 1, 2, \ldots, p\} \).

To ease the notation, we rewrite the set \( \Gamma \) in parametric variational inequalities as the following representation

\[ \Gamma = \{y \in \mathbb{R}^n \mid \mathbb{E}[g(y, \xi)] \in \mathcal{K}\}, \]

where \( g(y, \xi) := (g_1(y, \xi), \ldots, g_p(y, \xi)) \) and \( \mathcal{K} := \mathbb{R}^q \times \{0\}_{p-q} \). Replace the parametric variational inequality in (P) by

\[ 0 \in \Psi(x, y) + \mathcal{N}_\Gamma(y), \quad (3) \]

where \( \Psi(x, y) := \mathbb{E}[F(x, y, \xi)] \). Under Assumption 3.1, \( \mathbb{E}[g(y, \xi)] \) is continuously differentiable and \( \mathbb{E}\mathbb{E}[g(y, \xi)] = \mathbb{E}[\mathcal{J}g(y, \xi)] \). When \( \mathbb{E}[\mathcal{J}g(y, \xi)] \) is linearly independent for \( y \in \mathbb{R}^m \), by Proposition 1, we have

\[ \mathcal{N}_\Gamma(y) = p(y)^T \mathcal{N}_\mathcal{K}(r(y)), \quad (4) \]

where

\[ p(y) = \mathbb{E}[\mathcal{J}g(y, \xi)], \quad r(y) = \mathbb{E}[g(y, \xi)]. \]

Denote by \( \Xi := \{(x, y) \in C \mid 0 \in \Psi(x, y) + \mathcal{N}_\Gamma(y)\} \) the feasible set of (P).

Let \( \mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \rightarrow \mathbb{R}^m \) be an auxiliary function defined by

\[ \mathcal{L}(x, y, d) = \Psi(x, y) + p(y)^T d. \quad (5) \]

Define the set-valued mapping \( \mathcal{M} : \mathbb{R}^2p \rightrightarrows \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \) by

\[ \mathcal{M}(\vartheta) = \{(x, y, d) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \mid (r(y), d)^T + \vartheta \in \text{gph}\mathcal{N}_\mathcal{K}\}. \quad (6) \]

It follows from [22] that if \( (x, y) \) is a feasible solution of (P) and \( p(y) \) has full row rank, then \( \mathcal{L}(x, y, d) = 0 \) has a unique solution \( d \in \mathcal{N}_\mathcal{K}(r(y)) \). Denote by \( \Lambda(x, y) = \{d \in \mathbb{R}^q \mid \mathcal{L}(x, y, d) = 0, d \in S(y)\} \) where \( S(y) = \mathcal{N}_\mathcal{K}(r(y)) \). For \( \bar{y} \in \mathbb{R}^m \),
denote \( I(\bar{y}) := \{i | \mathbb{E}[g_i(\bar{y}, \xi)] = 0, \ 1 \leq i \leq q\} \), \( J := \{1, 2, \cdots, q\}\setminus I \) and \( L := \{q + 1, q + 2, \cdots, p\}. \)

To simplify the main results, the coderivative of a normal cone mapping is given under assumption of strict complementarity conditions.

**Assumption 3.2.** The strict complementarity condition holds at \( \bar{y} \), that is
\[
\bar{d}_i > 0, \ i \in I(\bar{y})
\]
for all \( \bar{d} \in \mathcal{N}_K(r(\bar{y})). \)

**Proposition 4.** Suppose that Assumption 3.1 and Assumption 3.2 are satisfied. Let \( Q(\bar{y}) := \mathcal{N}_F(\bar{y}) \) and \( \bar{v} \in Q(\bar{y}). \) Assume that \( p_{1\cup J}(\bar{y}) \) has full row rank, then for \( \bar{d} \in \Lambda(\bar{x}, \bar{y}) \) and \( \bar{u} \in \mathbb{R}^m \), it holds that
\[
D^*Q(\bar{y}, \bar{v})(\bar{u}) = \bigcup_{\bar{d} \in \Lambda(\bar{x}, \bar{y})} \left\{ (\mathcal{J}p(\bar{y})^T \bar{d})^T \bar{u} + \mathcal{J}r(\bar{y})^T D^*p_{1\cup J}(\bar{y}, \bar{d})(p(\bar{y}) \bar{u}) \right\}.
\] (7)

**Proof.** Since \( p_{1\cup J}(\bar{y}) \) has full row rank at \( \bar{y} \), from [24, Theorem 4.1], the mapping \( \mathcal{M} \) is clams at \((0, \bar{y}, \bar{d})\), then we have
\[
D^*Q(\bar{y}, \bar{v})(\bar{u}) \subset \bigcup_{\bar{d} \in \Lambda(\bar{x}, \bar{y})} \left\{ (\mathcal{J}p(\bar{y})^T \bar{d})^T \bar{u} + \mathcal{J}r(\bar{y})^T D^*p_{1\cup J}(\bar{y}, \bar{d})(p(\bar{y}) \bar{u}) \right\}.
\]

Under Assumption 3.2, the equality is obtained from [35, Lemma 3.3]. \( \Box \)

Next we put forward the first order necessary condition of the true problem under the linear independent constraint qualification.

**Assumption 3.3.** The linear independent constraint qualification is satisfied, namely the matrix
\[
\begin{bmatrix}
\mathcal{J}_{x,y}, \mathcal{L}(\bar{x}, \bar{y}, \bar{d}) \\
p_{1\cup J}(\bar{y})
\end{bmatrix}
\]
is of full row rank.

**Theorem 3.1.** Suppose that Assumption 3.1-3.3 are satisfied. The function \( \phi \) is Lipschitz continuous for almost every \( \xi \in \Theta \) and \( \mathcal{J}_{x,y}F(x, y, \xi) \) has full row rank at \((\bar{x}, \bar{y})\) for almost every \( \xi \in \Theta \). The probability space of \( \xi \) is nonatomic. Let \((\bar{x}, \bar{y})\) be a local weakly Pareto solution of \((P)\) and \( \bar{d} \) is a solution of \( \mathcal{L}(\bar{x}, \bar{y}, \bar{d}) = 0. \) Then there exist \( \mu^* \neq 0 \) and \( \bar{u} \in \mathbb{R}^m \) such that
\[
0 \in \mathbb{E}[\partial(\mu^*, \phi(\cdot, \cdot, \xi))(\bar{x}, \bar{y})] + \mathbb{E}[\nabla F(\bar{x}, \bar{y})]^T \bar{u} + \mathcal{N}_C(\bar{x}, \bar{y}) + \mathbb{E}[\nabla^2 g(\bar{y}, \xi)] \bar{d} \bar{u} + \mathbb{E}[\nabla g(\bar{y}, \xi)] D^* \mathcal{N}_E(\mathbb{E}[\nabla g(\bar{y}, \xi)], \bar{d})(\mathbb{E}[\nabla g(\bar{y}, \xi)]) \bar{u}.
\] (8)

**Proof.** Since \((\bar{x}, \bar{y})\) is the weakly Pareto solution of \((P)\), it follows from Lemma 2.1 that \((\bar{x}, \bar{y})\) is the optimal solution of \( \{\min(\psi \circ \mathbb{E}[\phi(\cdot, \cdot, \xi)])(x, y) \mid (x, y) \in \Xi\} \). By Lemma 2.2, there is \( \mu^* \in \partial \psi(\mathbb{E}[\phi(\bar{x}, \bar{y}, \xi)]) \) such that
\[
0 \in \partial(\mu^*, \mathbb{E}[\phi(\cdot, \cdot, \xi)])(\bar{x}, \bar{y}) + \mathcal{N}_E(\bar{x}, \bar{y}).
\] (9)

Since \( \phi \) is Lipschitz continuous with respect to \( x, y \) for almost every \( \xi \in \Theta \), from Assumption 3.1, we obtain that \( \mathbb{E}[\phi(\cdot, \cdot, \xi)] \) is Lipschitz and therefore \( \langle \mu^*, \mathbb{E}[\phi(\cdot, \cdot, \xi)] \rangle \) is also Lipschitz. It follows from [33, Theorem 2.1] that
\[
\partial(\mu^*, \mathbb{E}[\phi(\cdot, \cdot, \xi)])(\bar{x}, \bar{y}) \subset \mathbb{E}[\partial(\mu^*, \phi(\cdot, \cdot, \xi))(\bar{x}, \bar{y})].
\]

Hence, (9) becomes
\[
0 \in \mathbb{E}[\partial(\mu^*, \phi(\cdot, \cdot, \xi))(\bar{x}, \bar{y})] + \mathcal{N}_E(\bar{x}, \bar{y}).
\] (10)
Under Assumption 3.3, the mapping $\mathcal{P}$ is calm at $(0_{2m}, 0_{2p}, \bar{x}, \bar{y}, \bar{d})$, where $\mathcal{P} : \mathbb{R}^m \times \mathbb{R}^p \rightrightarrows \mathbb{R}^m \times \mathbb{R}^p$ is given by

$$\mathcal{P}(x, \vartheta) = \{(x, y, d) \in \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^p \mid \mathcal{L}(x, y, d) + z = 0\} \cap \mathcal{M}(\vartheta).$$  \hfill (11)

From [24, Theorem 3.2] and $\mathcal{J}_{x,y}F(x, y, \xi)$ having full row rank at $(\bar{x}, \bar{y})$, we have

$$\mathcal{N}_{\Xi}(\bar{x}, \bar{y}) = \bigcup_{\vartheta \in \Lambda(\xi, y)} [\mathcal{N}_C(\bar{x}, \bar{y}) + \nabla \Psi(\bar{x}, \bar{y})^T \bar{u} + D^* Q(\bar{y}, -\Psi(\bar{x}, \bar{y}))(\bar{u})].$$  \hfill (12)

Combining (10), (12) and Proposition 4, we obtain that (8) is satisfied. \hfill \square

**Remark.** We discuss the subdifferential calculus of random functions. When $f(x, y, \xi)$ is locally Lipschitz continuous with respect to $x, y$ for almost every $\xi$, we have

$$\partial f(x, y, \xi) \subset \partial^C f(x, y, \xi) = \text{clconv} \partial f(x, y, \xi),$$

where $\partial f, \partial^C f$ denote the limiting subdifferential and Clarke generalized subdifferential of $f$. Moreover, $\mathbb{E}[f(x, y, \xi)]$ is also Lipschitz continuous for $x, y$ and it holds that

$$\partial \mathbb{E}[f(x, y, \xi)] \subset \mathbb{E}[\partial f(x, y, \xi)].$$

When $f$ is Clarke regular at $x, y$ for almost every $\xi$, the equality holds. While the probability space $(\Omega, \mathcal{F}, P)$ is non-atomic or the atoms are convex, one has

$$\mathbb{E}[\partial f(x, y, \xi)] = \mathbb{E}[\partial^C f(x, y, \xi)] = \mathbb{E}[\text{conv} \partial f(x, y, \xi)] = \mathbb{E}[\partial f(x, y, \xi)].$$

If $f$ is convex with respect to $x, y$ for almost every $\xi$, the above subdifferentials are same. In general, the Clarke generalized subdifferential is larger than the limiting subdifferential and has better properties when the probability space $(\Omega, \mathcal{F}, P)$ is non-atomic. Hence, using Clarke generalized subdifferentials to characterize the first order necessary conditions is quite favorable when $\phi(x, y, \xi)$ is nonconvex and nonsmooth for fixed $\xi$.

### 3.2. First order necessary condition for the SAA problem.

Analogously, we establish the first order necessary condition of corresponding SAA problems. Rewrite the sample average approximation of parametric variational inequalities admitting the following representation

$$0 \in \hat{F}^N(x, y) + \mathcal{N}_{\Gamma_N}(y)$$  \hfill (13)

with $\Gamma_N = \{y \in \mathbb{R}^m \mid \hat{g}^N(y) \in \mathcal{K}\}$, where

$$\hat{F}^N(x, y) := \frac{1}{N} \sum_{k=1}^N F(x, y, \xi^k), \quad \hat{g}^N(y) := \frac{1}{N} \sum_{k=1}^N g(y, \xi^k).$$  \hfill (14)

Similarly, denote $\hat{p}^N(y) := \mathcal{J} \hat{g}^N(y)$ and $\Xi_N := \{(x, y) \in C \mid 0 \in \hat{F}^N(x, y) + \mathcal{N}_{\Gamma_N}(y)\}$ which is the approximation feasible set of (SAA-P). The SAA auxiliary function of $\mathcal{L}(x, y, d)$ is given by

$$\hat{\mathcal{L}}^N(x, y, d) := \hat{F}^N(x, y) + \hat{p}^N(y)^T d.$$  

The SAA multifunctions $\hat{\mathcal{M}}^N(\vartheta)$ and $\hat{\mathcal{P}}^N(z, \vartheta)$ are provided by

$$\hat{\mathcal{M}}^N(\vartheta) := \{(x, y, d) \in \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^p \mid (\hat{g}^N(y), d)^T + \vartheta \in \text{gph} \mathcal{N}_{\mathcal{K}}\}$$

and

$$\hat{\mathcal{P}}^N(z, \vartheta) := \{(x, y, d) \in \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^p \mid \hat{\mathcal{L}}^N(x, y, d) + z = 0\} \cap \hat{\mathcal{M}}^N(\vartheta).$$

Moreover, denote by $\Lambda_N(x, y) = \{d \in \mathbb{R}^p \mid \hat{\mathcal{L}}^N(x, y, d) = 0, d \in S_N(y)\}$ the SAA mapping of $\Lambda(x, y)$ where $S_N(y) = \mathcal{N}_{\mathcal{K}}(\hat{g}^N(y))$. 
The following proposition estimates the coderivative of $\hat{Q}^N(y) := \mathcal{N}_\mathcal{F}(y)$ in (13).

**Proposition 5.** Let $(\bar{x}^N, \bar{y}^N) \to (\bar{x}, \bar{y})$ as $N \to \infty$ with $(\bar{x}^N, \bar{y}^N) \in \Xi_N$ for each $N$. Taking $\bar{w}^N \in \hat{Q}^N(y)$, suppose that $p_{I\cup J}(y)$ has full row rank at $\bar{y}$. For each $N$, there is $\bar{d}^N \in \Lambda_N(\bar{x}^N, \bar{y}^N)$. Then for every $\bar{w}^N \in \mathbb{R}^p$, it holds that

$$D^*\hat{Q}^N(\bar{y}^N, \bar{w}^N)(\bar{u}^N) = \bigcup_{\bar{d}^N \in \Lambda_N(\bar{x}^N, \bar{y}^N)} \left\{ \hat{p}^N(\bar{y}^N)^T D^*\mathcal{N}_K(\hat{g}^N(\bar{y}^N), \bar{d}^N)(\bar{w}^N) + \nabla^2 \hat{g}^N(\bar{y}^N) \bar{d}^N \bar{u}^N \right\},$$

where $\bar{w}^N := \hat{p}^N(\bar{y}^N)^T \bar{u}^N$.

**Proof.** When $p_{I\cup J}(y)$ has full row rank at $\bar{y}$, it follows from [24, Theorem 4.1] that $\mathcal{M}_N$ is clasm at $(0_{2p}, \bar{y}^N)$ almost surely with $\bar{d}^N \in \Lambda_N(\bar{x}^N, \bar{y}^N)$ for each $N$, then

$$D^*\hat{Q}^N(\bar{y}^N, \bar{w}^N)(\bar{u}^N) = \nabla^2 \hat{g}^N(\bar{y}^N) \bar{d}^N \bar{u}^N + D^*(\mathcal{N}_K \circ \hat{g}^N)((\bar{y}^N), \bar{w}^N)(\hat{p}^N(\bar{y}^N)^T \bar{u}^N).$$

Consider the set-valued mapping $\mathcal{T} := \mathcal{N}_K \circ \hat{g}^N$ and observe that

$$\text{gph} \mathcal{T} = \{(y, d) \in \mathbb{R}^m \times \mathbb{R}^p \mid (\hat{g}^N(y), d)^T \in \text{gph} \mathcal{N}_K\}.$$  

Because gph $\mathcal{T} = M(0)$ in [13, Theorem 4.4] and $p_{I\cup J}(y)$ has full row rank at $\bar{y}$, we obtain

$$\mathcal{N}_{\text{gph} \mathcal{T}}(\bar{y}^N, \bar{d}^N) \subset \begin{pmatrix} \hat{p}^N(\bar{y}^N) \\ 0 \end{pmatrix} \mathcal{N}_{\text{gph} \mathcal{N}_K}(\hat{g}^N(\bar{y}^N), \bar{d}^N).$$

It follows from Proposition 1 that

$$\mathcal{N}^{*}_{\text{gph} \mathcal{T}}(\bar{y}^N, \bar{d}^N) \supset \begin{pmatrix} \hat{p}^N(\bar{y}^N) \\ 0 \end{pmatrix} \mathcal{N}^{*}_{\text{gph} \mathcal{N}_K}(\hat{g}^N(\bar{y}^N), \bar{d}^N).$$

Since the strict complementarity condition is satisfied at $(\bar{y}, \bar{d})$, gph $\mathcal{N}_K$ is normally regular at $(r(\bar{y}), \bar{d})$. Combining (17) and (18), we conclude that the equality (17) holds. Hence,

$$D^*(\mathcal{N}_K \circ \hat{g}^N)((\bar{y}^N), \bar{d}^N)(\hat{p}^N(\bar{y}^N)^T \bar{u}^N) = \hat{p}^N(\bar{y}^N)^T D^*\mathcal{N}_K(\hat{g}^N(\bar{y}^N), \bar{d}^N)(\hat{p}^N(\bar{y}^N)^T \bar{u}^N).$$

From (16) and (19), we have that (15) is satisfied.  

The following lemma provides the convergence properties of approximation functions.

**Lemma 3.2.** Suppose that Assumption 3.1 are satisfied. Let $\{(\bar{x}^N, \bar{y}^N)\}$ be a sequence converging to $(\bar{x}, \bar{y})$ as $N$ tends to infinity with $(\bar{x}^N, \bar{y}^N) \in \Xi_N$ for each $N$. Then

$$\hat{\phi}^N(\bar{x}^N, \bar{y}^N) \to \mathbb{E}[\phi(\bar{x}, \bar{y}, \bar{\xi})],$$

$$\hat{F}^N(\bar{x}^N, \bar{y}^N) \to \mathbb{E}[F(\bar{x}, \bar{y}, \bar{\xi})],$$

$$\hat{g}^N(\bar{y}^N) \to \mathbb{E}[g(\bar{y}, \bar{\xi})],$$

$$\nabla \hat{F}^N(\bar{x}^N, \bar{y}^N) \to \mathbb{E}[\nabla F(\bar{x}, \bar{y}, \bar{\xi})] = \nabla \mathbb{E}[F(\bar{x}, \bar{y}, \bar{\xi})],$$

$$\nabla \hat{g}^N(\bar{y}^N) \to \mathbb{E}[\nabla g(\bar{y}, \bar{\xi})] = \nabla \mathbb{E}[g(\bar{y}, \bar{\xi})],$$

$$\nabla^2 \hat{g}^N(\bar{y}^N) \to \mathbb{E}[\nabla^2 g(\bar{y}, \bar{\xi})] = \nabla^2 \mathbb{E}[g(\bar{y}, \bar{\xi})].$$

**Proof.** The proof is similar to [35, Lemma 3.1].
In what follows, we derive the first order necessary condition for the SAA problem.

**Theorem 3.3.** Suppose that Assumption 3.1-3.3 are satisfied. \( \phi \) is Lipschitz continuous for almost every \( \xi \in \Theta \) and \( J_{x,y}F(x,y,\xi) \) has full row rank at \((\bar{x}, \bar{y})\) for almost every \( \xi \in \Theta \). Let \((\bar{x}^N, \bar{y}^N)\) be a local weakly Pareto solution of (SAA-P) and \((\bar{x}, \bar{y})\) is a cluster point of \((\bar{x}^N, \bar{y}^N)\). Let \( \bar{d}^N \) be a solution of \( \tilde{F}(\bar{x}^N, \bar{y}^N, \bar{d}) = 0 \). Then there exists a bounded multiplier \( \bar{\mu}^N \in \mathbb{R}^m \) such that

\[
0 \in \mathcal{F}_N(\bar{x}^N, \bar{y}^N) + \nabla \hat{F}(\bar{x}^N, \bar{y}^N)^T \bar{d}^N + \mathcal{N}_C(\bar{x}^N, \bar{y}^N) + \nabla^2 \hat{g}(\bar{y}^N) \bar{d}^N \bar{u}^N,
\]

where \( \mathcal{F}_N(\bar{x}^N, \bar{y}^N) = \frac{1}{N} \sum_{k=1}^{N} \partial(\mu^*, \phi^*)(\bar{x}^N, \bar{y}^N) \). We call the pair \((\bar{x}, \bar{y}, \bar{u})\) which satisfies (20) the stationary pair of the problem (SAA-P).

**Proof.** When Assumption 3.2 and 3.3 hold, the multifunctions \( \mathcal{M}^N \) and \( \hat{P}^N \) are clamped at points \((0_{2p}, \bar{y}^N, \bar{d}^N)\) and \((0_{m}, 0_{2p}, \bar{x}^N, \bar{y}^N, \bar{d}^N)\) almost surely with \( \bar{d}^N \in \Lambda_N(\bar{x}^N, \bar{y}^N) \) for each \( N \). From Lemma 2.1, there is \( \mu^* \in \partial \psi(\phi^N(\bar{x}^N, \bar{y}^N)) \) such that

\[
0 \in \partial(\mu^*, \phi^*)(\bar{x}^N, \bar{y}^N) + \mathcal{N}_E(\bar{x}^N, \bar{y}^N).
\]

Because

\[
\partial(\mu^*, \phi^*)(\bar{x}^N, \bar{y}^N) = \partial \left[ \frac{1}{N} \sum_{k=1}^{N} (\mu^*, \phi(\cdot, \cdot, \xi^k)) \right] \bar{x}^N, \bar{y}^N \]

\[
\subset \frac{1}{N} \sum_{k=1}^{N} \partial(\mu^*, \phi(\cdot, \cdot, \xi^k)) \bar{x}^N, \bar{y}^N,
\]

combining (21), (22) and Proposition 5, we obtain that (20) holds.

Next we prove the boundedness of \( \bar{u}^N \). Assume a contradiction that \( \|\bar{u}^N\| \to \infty \) as \( N \to \infty \), then there exists \( \alpha^N \to 0 \) as \( N \to \infty \) such that almost surely \( \alpha^N \bar{u}^N \to \bar{u}, \|\bar{u}\| \neq 0 \). Multiplying (20) by \( \alpha^N \), we obtain

\[
0 \in \partial(\mu^*, \phi^*)(\bar{x}^N, \bar{y}^N)\alpha^N + \nabla \hat{F}(\bar{x}^N, \bar{y}^N)^T \alpha^N \bar{u}^N + \alpha^N \mathcal{N}_C(\bar{x}^N, \bar{y}^N) + \nabla^2 \hat{g}(\bar{y}^N) \bar{d}^N \alpha^N \bar{u}^N + \alpha^N \mathcal{N}_C(\bar{x}^N, \bar{y}^N) + \nabla^2 \hat{g}(\bar{y}^N) \bar{d}^N \alpha^N \bar{u}^N .
\]

Since \( \phi(x,y,\xi) \) is Lipschitz, \( \langle \mu^*, \phi^* \rangle \) is also Lipschitz. From [21, Corollary 1.81], \( \bar{w}^N \) is bounded. By virtue of [26, Proposition 6.6], \( \mathcal{N}_C(\cdot, \cdot, \cdot) \) is outer semicontinuous, that is

\[
\limsup_{N \to \infty} \mathcal{N}_C(\bar{x}^N, \bar{y}^N) \subset \mathcal{N}_C(\bar{x}, \bar{y}).
\]

From Proposition 3 (i), we obtain that \( D^+ \mathcal{N}_K \) is outer semicontinuous. It follows from Lemma 3.2 that

\[
\limsup_{N \to \infty} D^+ \mathcal{N}_K(\bar{g}(\bar{y}, \xi), \bar{d})(\nabla \hat{g}(\bar{y}^N) \bar{u}^N)
\subset D^+ \mathcal{N}_K(\bar{g}(\bar{y}, \xi), \bar{d})(\nabla \hat{g}(\bar{y}, \xi) \bar{u}).
\]
Combining (24) and (25), let $N \to \infty$, it exists $\nu \in D^*\mathcal{N}_K(\mathbb{E}[g(\bar{y}, \xi)], \bar{d})(\mathbb{E}[\nabla g(\bar{y}, \xi)]\bar{u})$ and $\rho \in \mathcal{N}_C(\bar{x}, \bar{y})$ almost surely such that

$$0 + \mathbb{E}[\nabla F(\bar{x}, \bar{y}, \xi)]^T\bar{u} + 0 \cdot \rho + \mathbb{E}[\nabla^2 g(\bar{y}, \xi)]^T\bar{u} + 0 \cdot \mathbb{E}[\nabla g(\bar{y}, \xi)]^Tv = 0.$$ 

From Assumption 3.3, we obtain $\bar{u} = 0$, which is a contradiction, then $\bar{u}^N$ is bounded. 

3.3. Convergence of stationary points of the SAA problem. This subsection gives the convergence of stationary points of (SAA-P). The stationary sequence of (SAA-P) converges to the points which are also the stationary points of (P) when the sample size tends to infinity. To begin with, the following proposition provides the convergence of SAA coderivatives.

**Proposition 6.** Suppose that Assumption 3.1 and Assumption 3.2 are satisfied. Let $\{(\bar{y}^N, \bar{v}^N)\} \subset \mathbb{R}^m \times \mathbb{R}^m$ be a sequence converging to $(\bar{y}, \bar{v})$ as $N \to \infty$ with $(\bar{y}^N, \bar{v}^N) \in \text{gph} \tilde{Q}^N$. If $\text{p}(\bar{y})$ has full row rank at $\bar{y}$, then it holds that

$$\mathbb{D}(D^*\tilde{Q}^N(\bar{y}^N, \bar{v}^N)(\bar{u}^N), D^*Q(\bar{y}, \bar{v})(\bar{u})) \to 0 \text{ as } N \to \infty.$$ 

**Proof.** By Proposition 2, it suffices to prove $\limsup_{N \to \infty} D^*\tilde{Q}^N(\bar{y}^N, \bar{v}^N)(\bar{u}^N) \subset D^*Q(\bar{y}, \bar{v})(\bar{u})$. Let $\tilde{\zeta}^N \to \zeta$, $N \to \infty$ and $\tilde{\zeta}^N \in D^*\tilde{Q}^N(\bar{y}^N, \bar{v}^N)(\bar{u}^N)$. Since $\tilde{\zeta}^N \in D^*\tilde{Q}^N(\bar{y}^N, \bar{v}^N)(\bar{u}^N)$, from (15) in Proposition 5, there exists $d^N \in \Lambda_N(\bar{x}^N, \bar{y}^N)$ satisfying

$$\tilde{\zeta}^N = \nabla^2 \bar{g}^N(\bar{y}^N)d^N\bar{u}^N + \tilde{p}^N(\bar{y}^N)^TD^*\mathcal{N}_K(\bar{g}^N(\bar{y}^N), d^N) (\tilde{p}^N(\bar{y}^N)^T\bar{u}^N).$$

Hence there is $\eta^N \in D^*\mathcal{N}_K(\bar{g}^N(\bar{y}^N), d^N) (\tilde{p}^N(\bar{y}^N)^T\bar{u}^N)$ such that

$$\tilde{\zeta}^N = \nabla^2 \bar{g}^N(\bar{y}^N)d^N\bar{u}^N + \tilde{p}^N(\bar{y}^N)^T\eta^N. \quad (26)$$

From [31, Proposition 2.7], $D^*\mathcal{N}_K$ is osc and there is $\eta \in \mathbb{R}^p$ such that $\tilde{\eta}^N \to \eta$ as $N \to \infty$, $\tilde{\eta} \in D^*\mathcal{N}_K(r(\bar{y}), \bar{d}) (p(\bar{y})^Tu^N)$. Let $N \to \infty$ in (26), we obtain

$$\zeta = (\mathbb{E}[\nabla^2 g(\bar{y}, \xi)]\bar{d})^T\bar{u} + p(\bar{y})^T\bar{\eta},$$

which means

$$\zeta \in (Jp(\bar{y})^T\bar{d})^T\bar{u} + Jr(\bar{y})^T D^*\mathcal{N}_K(r(\bar{y}), \bar{d})(p(\bar{y})^Tu)$$

with $\bar{d} \in \Lambda(\bar{x}, \bar{y})$. From (7) in Proposition 4, we obtain $\zeta \in D^*Q(\bar{y}, \bar{v})(\bar{u})$ and the proof is completed. 

Now we give the convergence of stationary points. For simplicity, let

$$\mathcal{A}(\bar{x}, \bar{y}, \bar{u}) := \mathbb{E}[\theta(\mu^*, \phi(\cdot, \cdot, \xi))(\bar{x}, \bar{y})] + \mathbb{E}[\nabla F(\bar{x}, \bar{y}, \xi)]^T\bar{u} + \mathcal{N}_C(\bar{x}, \bar{y}) + \mathbb{E}[\nabla^2 g(\bar{y}, \xi)\bar{d}]\bar{u} + \mathbb{E}[\nabla g(\bar{y}, \xi)]D^*\mathcal{N}_K(\mathbb{E}[g(\bar{y}, \xi)], \bar{d})(\mathbb{E}[\nabla g(\bar{y}, \xi)]\bar{u})$$

and

$$\mathcal{F}^N(\bar{x}^N, \bar{y}^N, \bar{u}^N) := \mathbb{E}[\nabla F(\bar{x}^N, \bar{y}^N, \bar{u}^N) + \nabla^2 g(\bar{y}^N)\bar{d}^N\bar{u}^N] + \mathcal{N}_C(\bar{x}^N, \bar{y}^N) + \mathbb{E}[\nabla g(\bar{y}^N)D^*\mathcal{N}_K(\mathbb{E}[g(\bar{y}^N), \bar{d}^N])](\mathbb{E}[\nabla g(\bar{y}^N)]\bar{u}^N).$$

The first order necessary conditions in (8) and (20) can be written as

$$0 \in \mathcal{A}(\bar{x}, \bar{y}, \bar{u}) \quad (29)$$

and

$$0 \in \mathcal{F}^N(\bar{x}^N, \bar{y}^N, \bar{u}^N) \quad (30)$$
correspondingly. From [26, Theorem 5.37], if $A^N$ converges to $A$ graphically at $(\bar{x}, \bar{y}, \bar{u})$ and (30) is satisfied, then (29) holds, which means that $(\bar{x}, \bar{y}, \bar{u})$ is the stationary pair of (P). Firstly, recall some basic definitions related to the graphical convergence.

**Definition 3.4.** [26] For a sequence of set-valued mappings $S^v : R^n \rightrightarrows R^m$, the graphical outer limit, denoted by $(g - \limsup_v S^v)(x)$, is the mapping having as its graph of the set $\limsup_v (\text{gph } S^v)$:

$$\text{gph}(g - \limsup_v S^v)(x) = \limsup_v (\text{gph } S^v).$$

The graphical inner limit, denoted by $(g - \liminf_v S^v)(x)$, is the mapping having as its graph of the set $\liminf_v (\text{gph } S^v)$:

$$\text{gph}(g - \liminf_v S^v)(x) = \liminf_v (\text{gph } S^v).$$

When the graphical outer and inner limits agree, the graphical limit $(g - \lim_v S^v)(x)$ is said to exist.

**Proposition 7.** [26] For any sequence $S^v : R^n \rightrightarrows R^m$, one has

$$(g - \limsup_v S^v)(x) = \bigcup_{\{x^v \to x\}} \limsup_{v \to \infty} S^v(x^v) = \lim_{\delta \to 0} \left[ \limsup_{v \to \infty} S^v(x + \delta B) \right]$$

and

$$(g - \liminf_v S^v)(x) = \bigcup_{\{x^v \to x\}} \liminf_{v \to \infty} S^v(x^v) = \lim_{\delta \to 0} \left[ \liminf_{v \to \infty} S^v(x + \delta B) \right].$$

Furthermore, $S^v$ converges graphically to $S$ if and if only, at each point $\bar{x} \in R^n$, one has

$$\bigcup_{\{x^v \to x\}} \limsup_{v \to \infty} S^v(x^v) \subset S(\bar{x}) \subset \bigcup_{\{x^v \to x\}} \liminf_{v \to \infty} S^v(x^v).$$

Now we give the main convergent results.

**Theorem 3.5.** Let $A : R^n \times R^m \rightrightarrows R^{m+n}$ and $A^N : R^n \times R^m \rightrightarrows R^{m+n}$ be defined as in (27) and (28) respectively. Suppose that the assumptions in Theorem 3.1 are satisfied. Let $\{(\bar{x}^N, \bar{y}^N)\}$ be a sequence of stationary points satisfying (20) and $(\bar{x}, \bar{y})$ is a cluster point of $(\bar{x}^N, \bar{y}^N)$, then

$$(g - \limsup_N A^N)(\bar{x}, \bar{y}, \bar{u}) \subset A(\bar{x}, \bar{y}, \bar{u}). \quad (31)$$

Moreover, $(\bar{x}, \bar{y})$ satisfies the first order necessary condition (8).

**Proof.** Firstly, we prove that (31) is satisfied. From Proposition 7, it needs to prove

$$\bigcup_{\{(\bar{x}^N, \bar{y}^N, \bar{u}^N) \to (\bar{x}, \bar{y}, \bar{u})\}} \limsup_{N \to \infty} A^N(\bar{x}^N, \bar{y}^N, \bar{u}^N) \subset A(\bar{x}, \bar{y}, \bar{u}).$$

Without loss of generality, assume that $(\bar{x}^N, \bar{y}^N, \bar{u}^N) \to (\bar{x}, \bar{y}, \bar{u})$ as $N \to \infty$. From [1], one has

$$\Upsilon^N(\bar{x}^N, \bar{y}^N) \to E[\text{conv}\partial(\mu^*, \phi(\cdot, \cdot, \xi))(\bar{x}, \bar{y})]$$

as $N \to \infty$. Since the probability space of $\xi$ is nonatomic, it holds that $E[\text{conv}\partial(\mu^*, \phi(\cdot, \cdot, \xi))(\bar{x}, \bar{y})] = E[\partial(\mu^*, \phi(\cdot, \cdot, \xi))(\bar{x}, \bar{y})]$, and we obtain

$$\Upsilon^N(\bar{x}^N, \bar{y}^N) \to E[\partial(\mu^*, \phi(\cdot, \cdot, \xi))(\bar{x}, \bar{y})]$$

as $N \to \infty$. \quad (32)
It follows from Lemma 3.2 that
\[ \lim_{N \to \infty} F^N(\bar{x}^N, \bar{y}^N)^T \bar{u}^N = \mathbb{E}[\nabla F(\bar{x}, \bar{y}, \xi)]^T \bar{u}. \] (33)

Similarly, we get
\[ \lim_{N \to \infty} \nabla^2 \hat{g}^N(\bar{y}^N) \hat{d}^N \bar{u}^N = \mathbb{E}[\nabla^2 g(\bar{y}, \xi) \hat{d}] \bar{u}. \] (34)

Put \( S := \nabla g \) and \( T := N_K \) in Proposition 3 (ii) and we obtain that \( \nabla g D^* N_K \) is outer semicontinuous, hence,
\[ \limsup_{N \to \infty} \nabla \hat{g}^N(\bar{y}^N) D^* N_K(\hat{g}^N(\bar{y}^N), \hat{d}^N)(\nabla \hat{g}^N(\bar{y}^N) \bar{u}^N) \subset \mathbb{E}[\nabla g(\bar{y}, \xi)] D^* N_K(\mathbb{E}[g(\bar{y}, \xi)], \hat{d})(\mathbb{E}[\nabla g(\bar{y}, \xi)] \bar{u}). \] (35)

Combining Lemma 3.2 and (32-35), we have
\[ \limsup_{N \to \infty} A^N(\bar{x}^N, \bar{y}^N, \bar{u}^N) \subset A(\bar{x}, \bar{y}, \bar{u}), \]
and therefore (31) holds. Since \( (\bar{x}^N, \bar{y}^N) \) satisfies the first order necessary condition (20) which is \( 0 \in A^N(\bar{x}^N, \bar{y}^N, \bar{u}^N) \), according to the above proof, \( A^N \) graphically converges to \( A \). Then \( (\bar{x}, \bar{y}) \) is the stationary point of (8) by applying [26, Theorem 5.37] to \( A^N \). \( \square \)

3.4. Convergence of optimal solutions. Another important focus on SAA methods is the convergence of optimal solutions. As we know, when a problem is convex, the stationary point is equivalent to the optimal solution. Exploiting well known scalar results for (MOP), referring to [6, Theorem 2.5], [14] and [19], the convergence of optimal solutions of SAA problems is obtained.

The assumptions on the functions are given as follows.

\( (H_1) \) \( \phi(x, y, \xi) \) is strictly convex on \( R^n \times R^m \) for almost every \( \xi \in \Theta \).

\( (H_2) \) \( F(x, y, \xi) \) is monotone on \( R^n \times R^m \) for almost every \( \xi \in \Theta \).

**Theorem 3.6.** Suppose that the assumptions \((H_1), (H_2)\) are satisfied, then we have
\[ \bigcup_{\mu \in R_+^m \setminus \{0\}} \min_{(x, y) \in \Xi} \langle \mu, \mathbb{E}[\phi(x, y, \xi)] \rangle = E_W. \]

Moreover, for each \( N \) and almost every \( \xi \in \Theta \), we have
\[ \bigcup_{\mu \in R_+^m \setminus \{0\}} \min_{(x, y) \in \Xi^N} \langle \mu, \phi_N(x, y) \rangle = E_W^N. \]

**Proof.** From \( (H_1) \), \( (H_2) \) and [10], \( \Xi \) is convex for almost every \( \xi \in \Theta \). By [3, Theorem 2], the first equality holds. In the same way, the second equality holds from [6, Theorem 2.5]. \( \square \)

The following proposition provides the properties of weakly Pareto solution sets.

**Proposition 8.** Under \((H_1)\) and \((H_2)\), the following assertions hold:

(i) The weakly Pareto solution set \( E_W \) is compact, and for each \( N \), the SAA weakly Pareto solution set \( E_W^N \) is also compact for almost every \( \xi \in \Theta \).

(ii) The weakly Pareto solution set \( E_W \neq \emptyset \), and for each \( N \), the SAA weakly Pareto solution set \( E_W^N \neq \emptyset \) for almost every \( \xi \in \Theta \).
Theorem 3.7. Suppose that assumptions \((H_1, H_2)\) are satisfied and \(p_{I \cup J}(y)\) has full row rank at \(\bar{y}\) for almost every \(\xi \in \Theta\). The sets \(E_W\) and \(E_W^N\) are given in Theorem 3.6. Assume that, for some \((\bar{x}, \bar{y}) \in E_W\), there exists a sequence \((\bar{x}^N, \bar{y}^N) \in \Xi_N\) such that \((\bar{x}^N, \bar{y}^N) \to (\bar{x}, \bar{y})\) as \(N \to \infty\). Then we have

\[
\mathbb{D}(E_W^N, E_W) \to 0 \quad \text{as} \quad N \to \infty.
\]

Proof. It is sufficient to prove \(\lim \sup_{N \to \infty} E_W^N \subseteq E_W\). Take \((\bar{x}^N, \bar{y}^N) \in \Xi_N\) and let \((\bar{x}^N, \bar{y}^N) \to (\bar{x}, \bar{y})\) as \(N \to \infty\). Since \((\bar{x}^N, \bar{y}^N) \in \Xi_N\), by Theorem 3.6, there is \(\mu \in \Xi\) such that \(\mu \in \Xi_N\). Then, we have

\[
\mathbb{D}(E_W^N, E_W) \to 0 \quad \text{as} \quad N \to \infty.
\]

Before proving the convergence of optimal solutions, it needs to establish the convergence of feasible sets of (SAA-P), which is equivalent to prove the convergent property of solution mappings of SAA parametric variational inequalities.

**Proposition 9.** Suppose Assumption 3.1 are satisfied. Let \(S(x)\) and \(S_N(x)\) be the solution mappings to (3) and (12). \(\{\bar{x}^N\} \subset \mathbb{R}^n\) is a sequence converging to \(\bar{x}\) as \(N \to \infty\). Moreover, \(p_{I \cup J}(y)\) has full row rank at \(\bar{y}\), then we have

\[
\mathbb{D}(S_N(\bar{x}^N), S(\bar{x})) \to 0 \quad \text{as} \quad N \to \infty.
\]

Proof. From Proposition 9 (ii), it suffices to prove \(\lim \sup_{N \to \infty} S_N(\bar{x}^N) \subseteq S(\bar{x})\). Let \(\bar{y}\) be a cluster point of \(\bar{y}^N\) and \(\bar{y}^N \in S_N(\bar{x}^N)\). Without loss of generality, assume \(\bar{y}^N \to \bar{y}\) as \(N \to \infty\). Since \(\bar{y}^N \in S_N(\bar{x}^N)\), one has

\[
0 \in F_N(\bar{x}^N, \bar{y}^N) + N_{\Gamma_N}(\bar{y}^N).
\]

The matrix \(p_{I \cup J}(y)\) has full row rank at \(\bar{y}\), then

\[
N_{\Gamma_N}(\bar{y}^N) = \nabla \bar{g}^N(\bar{y}^N) N_K(\bar{y}^N).
\]

Combining (36) and (37), there exists \(\bar{d}^N \in N_K(\bar{y}^N)\) such that

\[
F_N(\bar{x}^N, \bar{y}^N) + \nabla \bar{g}^N(\bar{y}^N) \bar{d}^N = 0.
\]

Now we prove that \(\bar{d}^N\) is bounded. By contradiction, there exists \(\alpha^N \to 0\) such that \(\alpha^N \bar{d}^N \to \bar{d}, \bar{d} \neq 0\) as \(N \to \infty\). Multiplying both sides of (38) by \(\alpha^N\) and let \(N \to \infty\), we obtain \(p(\bar{y}) \bar{d} = 0\). Since \(p_{I \cup J}(\bar{y})\) has full row rank, it holds \(\bar{d}_{i,I \cup J} = 0\) when \(i \in I \cup J\). By \(E[\bar{y}^N(\bar{y}, \xi)] < 0\) and the definitions of normal cones, it has \(\bar{d}_i = 0, i \notin I \cup J\). Hence \(\bar{d} = 0\), which is a contradiction. Without loss of generality, there exists \(\bar{d} \in R^n\) such that \(\bar{d}^N \to \bar{d}, \bar{d} \in N_K(\bar{y})\) as \(N \to \infty\). From (38), we have

\[
0 \in \Psi(\bar{x}, \bar{y}) + p(\bar{y})^T N_K(\bar{y}),
\]

which means \(\bar{y} \in S(\bar{x})\) under the basic constraint qualification. The proof of this theorem is completed. \(\square\)

Now we give the almost surely convergence of weakly Pareto solution sets of (SAA-P) under Theorem 3.6.
For the problem becomes the following minimization of functions as follows:

\[ \psi = \min \{0 + C \xi \} \]

**Example 1.** We consider the following stochastic bi-objective optimization problem:

\[
\begin{align*}
\min \quad & \mathbb{E}[\phi_1(x, y, \xi), \phi_2(x, y, \xi)] \\
\text{s.t.} \quad & 0 \leq y \perp F(x, y, \xi) \geq 0, \\
& (x, y) \in C,
\end{align*}
\]

where

\[
\begin{pmatrix}
\phi_1(x, y, \xi) \\
\phi_2(x, y, \xi)
\end{pmatrix} = \begin{pmatrix}
x^2 - 2x - y + \xi \\
\frac{1}{4}x^2 - x + \frac{5}{8} + \frac{1}{2}y + \frac{3}{4} \xi
\end{pmatrix}, \quad F(x, y, \xi) = y - x - \xi.
\]

The random variable \( \xi \) follows a uniform distribution on \([1, 2]\) and the feasible set \( C = [0, 2] \times [0, 2] \). From the complementarity constraint problem, there is a unique solution \( y \) related to \( x \) and \( \xi \), which is

\[
y(x, \xi) = \begin{cases} x - \xi, & \text{if } \xi \leq x, \\ 0, & \text{otherwise.}\end{cases}
\]

Consequently, the expectations of bi-objective functions are as follows:

\[
\begin{align*}
\mathbb{E}[\phi_1(x, y, \xi)] &= x^2 - 2x - \int_1^x (x - t)dt + \frac{3}{2} = \frac{1}{2}x^2 - x + 1, \\
\mathbb{E}[\phi_2(x, y, \xi)] &= \frac{1}{4}x^2 - x + \frac{5}{8} + \frac{1}{2} \int_1^x (x - t)dt + \frac{3}{4} = \frac{1}{2}x^2 - \frac{3}{2}x + \frac{13}{8}.
\end{align*}
\]

From the graph of bifunctions, it is easy to obtain \( E_w = [1, \frac{3}{2}] \times [0, \frac{1}{2}] \). Take the evaluation function \( \psi(z_1, z_2) = \frac{1}{2}z_1 + 2z_2 \), then the above multiobjective optimization problem becomes the following minimization of functions as follows:

\[
\min \psi(\mathbb{E}[\phi_1(x, y, \xi), \phi_2(x, y, \xi)]) = \min_{x \in [0, 2]} \left[ \frac{5}{4}x^2 - \frac{7}{2}x + \frac{15}{4} \right].
\]

It is easily obtained that the optimal value of above problem is \( \psi_{\min} = 1.3 \) and the optimal solution is \( x_{\min} = \frac{7}{2} \).

Now, let \( \xi = (\xi^1, \xi^2, \cdots, \xi^N) \) be given, for each sample \( \xi^k, k = 1, 2, \cdots, N \), and the corresponding solution of complementarity constraint problems is provided by

\[
y^k(x, \xi^k) = \begin{cases} x - \xi^k, & \text{if } \xi^k \leq x, \\ 0, & \text{otherwise,}\end{cases}
\]
which means \( y^k(x, \xi^k) = \max\{x - \xi^k, 0\} \). The SAA bi-objective functions become

\[
\hat{\phi}^N_1(x, y) = x^2 - 2x - \frac{1}{N} \sum_{k=1}^{N} \max\{x - \xi^k, 0\} + \frac{1}{N} \sum_{k=1}^{N} \xi^k,
\]

\[
\hat{\phi}^N_2(x, y) = \frac{1}{4} x^2 - x + \frac{5}{8} + \frac{1}{2N} \sum_{k=1}^{N} \max\{x - \xi^k, 0\} + \frac{1}{2N} \sum_{k=1}^{N} \xi^k,
\]

and the SAA scalarization optimization problem becomes

\[
\min \psi(\hat{\phi}^N_1, \hat{\phi}^N_2) = \min \left[ x^2 - 3x + \frac{5}{4} + \frac{1}{2N} \sum_{k=1}^{N} \max\{x - \xi^k, 0\} + \frac{3}{2N} \sum_{k=1}^{N} \xi^k \right],
\]

where \( \hat{\phi}^N_i := \hat{\phi}^N_i(x, y), i = 1, 2 \).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Convergence of SAA optimal values.}
\end{figure}

For the sample size \( N = 1, 2, \ldots, 300 \), we give the optimal values and optimality solutions of SAA scalarization problems and numerical results are provided in Figure 1, Figure 2. From Figure 1, the optimal value of SAA scalarization problems tends to the true optimal value as the sample size \( N \) increases and the errors between two values are small. It illustrates that the SAA scalarization method dealing with this problem is reasonable. The component \( x \) in \( E_w \) is \([1, \frac{3}{2}] \). By virtue of Theorem 3.6, the weakly (SAA) Pareto solution set is equivalent to the optimal solution set of (SAA) scalarization problems. In Figure 2, it provides the SAA optimal solutions of SAA problems. It is shown that most solutions belong to \([1, \frac{3}{2}]\), which means that the deviation between the weakly SAA Pareto solution set and its true Pareto solution set is small.

**Example 2.** In this example we consider a stochastic bi-objective optimization problem with complementarity constraints in two dimensions
Figure 2. Convergence of SAA optimal solutions.

$$\begin{align*}
\min \quad & \mathbb{E}[\phi_1(x, y, \xi), \phi_2(x, y, \xi)] \\
\text{s. t.} \quad & 0 \leq y \perp F(x, y, \xi) \geq 0, \\
& (x, y) \in C,
\end{align*}$$

where $x = (x_1, x_2), y = (y_1, y_2)$ and $\xi = (\xi_1, \xi_2)$. The bi-objective functions is given by

$$\begin{pmatrix}
\phi_1(x, y, \xi) \\
\phi_2(x, y, \xi)
\end{pmatrix} = \begin{pmatrix}
\frac{x_1^2}{2} + \frac{x_2^2}{2} + x_1 + x_2 - y_1 - y_2 + 2\xi_1 + \xi_2 \\
-2x_1x_2 + 3(x_1 - x_2) + 2y_1 + 2y_2 + \frac{2}{3}\xi_1 + 4\xi_2
\end{pmatrix}.$$  

The mapping $F$ and the feasible set $C$ are given by

$$F(x, y, \xi) = \begin{pmatrix}
y_1 - x_1 + \xi_1 \\
y_2 - x_2 + \xi_2
\end{pmatrix}, \quad C = [0, 1] \times [0, 1] \times [0, 1] \times [0, 1].$$

The random variable $\xi$ is uniformly distributed on $[0, 1] \times [0, 1]$. From the complementarity constraint problem, the unique solution $y$ related to $x$ and $\xi$ is

$$y_i(x_i, \xi_i) = \begin{cases} x_i - \xi_i, & \text{if } \xi_i \leq x_i, \\
0, & \text{otherwise.}
\end{cases}$$

where $i = 1, 2$. Similarly, the expectations of bi-objective functions are

$$\mathbb{E}[\phi_1(x, y, \xi)] = \frac{x_1^2}{2} + \frac{x_2^2}{2} + x_1 + x_2 - \int_0^{x_1} (x_1 - t) dt - \int_0^{x_2} (x_2 - t) dt$$
$$+ 2\int_0^1 t dt + \int_0^1 t dt$$
$$= x_1 + x_2 + \frac{3}{2},$$
\[ E[\phi_2(x,y,\xi)] = -2x_1x_2 + 3(x_1 - x_2) + 2 \int_0^{x_1} (x_1 - t) \, dt + 2 \int_0^{x_2} (x_2 - t) \, dt + \frac{2}{3} \int_0^1 t \, dt + 4 \int_0^1 t \, dt = x_1^2 + x_2^2 - 2x_1x_2 + 3x_1 - 3x_2 + \frac{7}{3}. \]

Figure 3. The boundary of set \( E[\phi(C,\cdot)]. \)

The Pareto solution set image \( E[\phi(C,\cdot)] \) is the quadrilateral curvilinear domain \( \overline{A_1 A_2 A_3 A_4} \) with \( A_1(\frac{3}{2},\frac{7}{3}), A_2(\frac{5}{2},\frac{1}{3}), A_3(\frac{7}{2},\frac{19}{3}), A_4(\frac{7}{2},\frac{7}{3}). \) The parabolic arcs \( \overline{A_1 A_2}, \overline{A_1 A_3}, \overline{A_2 A_4}, \overline{A_3 A_4} \) are given by following representations: for \( s \in [0,1], \)

\[ \overline{A_1 A_2} : s \rightarrow (s + \frac{3}{2}, s^2 - 3s + \frac{7}{3}), \quad \overline{A_1 A_3} : s \rightarrow (s + \frac{3}{2}, s^2 + 3s + \frac{7}{3}), \]
\[ \overline{A_2 A_4} : s \rightarrow (s + \frac{5}{2}, s^2 + s + \frac{1}{3}), \quad \overline{A_3 A_4} : s \rightarrow (s + \frac{5}{2}, s^2 - 5s + \frac{19}{3}). \]

The true Pareto solution set image is the arc \( \overline{A_1 A_2} \) (in bold in Figure 3). By the definitions of Pareto solution sets and complementarity constraints, the true Pareto solution set is \( E = \{0\} \times [0,1] \times \{0\} \times [0,1]. \) Choose \( \psi(z_1,z_2) = z_1 + z_2, \) then the above multiobjective optimization problem becomes

\[ \min \psi(E[\phi_1(x,y,\xi), \phi_2(x,y,\xi)]) = \min \left[ (x_1 - x_2)^2 + 4x_1 - 2x_2 + \frac{23}{6} \right]. \]

By computation, we obtain \( \psi_{\min} = 2.8333 \) and \( x_{\min} = (0.0000, 0.0990). \)

Analogous to Example 1, the SAA bi-objective functions are provided by

\[ \hat{\phi}_1^N(x,y) = x_1^2 + x_2^2 + x_1 + x_2 - \frac{1}{2} \sum_{k=1}^N \max\{x_1 - \xi_1^k, 0\} + \frac{2}{N} \sum_{k=1}^N \xi_1^k \]
\[ - \frac{1}{N} \sum_{k=1}^N \max\{x_2 - \xi_2^k, 0\} + \frac{1}{N} \sum_{k=1}^N \xi_2^k, \]
\[ \hat{\phi}^N_2(x, y) = -2x_1x_2 + 3(x_1 - x_2) + \frac{2}{N} \sum_{k=1}^{N} \max\{x_1 - \xi^k_1, 0\} + \frac{2}{3N} \sum_{k=1}^{N} \xi^k_1 \]
\[ + \frac{2}{N} \sum_{k=1}^{N} \max\{x_2 - \xi^k_2, 0\} + \frac{4}{N} \sum_{k=1}^{N} \xi^k_2. \]

Hence the SAA scalarization problem becomes
\[
\min \left[ \frac{x_1^2}{2} + \frac{x_2^2}{2} - 2x_1x_2 + 4x_1 - 2x_2 + \frac{1}{N} \sum_{k=1}^{N} \max\{x_1 - \xi^k_1, 0\} + \frac{8}{3N} \sum_{k=1}^{N} \xi^k_1 \right. \\
\left. + \frac{1}{N} \sum_{k=1}^{N} \max\{x_2 - \xi^k_2, 0\} + \frac{5}{N} \sum_{k=1}^{N} \xi^k_2 \right].
\] (39)

**Figure 4.** Convergence of SAA optimal values.

In this example, for \( N = 1, 2, \ldots, 300 \), we give the convergence of optimal values of SAA scalarization problems when the sample size \( N \) increases and the numerical results are presented in Figure 4. From Figure 4, the optimal value of SAA scalarization problems converges to the true optimal value as the sample size \( N \) grows and the convergence tends to stable gradually. It demonstrates that this example is well approximated by the sample average and the performance of SAA methods is carried out quite well.

**Example 3.** We consider the following stochastic bi-objective optimization problem with variational inequality constraint
\[
\min \quad \mathbb{E}[\phi_1(x, y, \xi), \phi_2(x, y, \xi)] \\
\text{s. t.} \quad \langle \mathbb{E}[F(x, y, \xi)], z - y \rangle \leq 0, \ z \in \Gamma, \ (x, y) \in C,
\]
where
\[
\begin{pmatrix}
\phi_1(x, y, \xi) \\
\phi_2(x, y, \xi)
\end{pmatrix} = \begin{pmatrix}
5x^2 - 8x - y + \xi \\
2x^2 - 16x + 2y - \xi
\end{pmatrix}, \quad F(x, y, \xi) = 2(y - x - \xi).
The set $\Gamma = \{ y \mid E[g(y, \xi)] \leq 0 \}$ and $g(y, \xi) = \xi + 2 - 2y$. The random variable $\xi$ follows a uniform distribution on $[0, 2]$ and the feasible set $C = [0, 3] \times [0, 3]$. From the variational inequality problem, there is a unique solution $y$ related to $x$ and $\xi$, which is

$$y(x, \xi) = \begin{cases} 1 + \frac{1}{2} \xi, & \text{if } \xi \leq 2 - 2x, \\ x + \xi, & \text{otherwise.} \end{cases}$$

Consequently, the expectations of bi-objective functions are as follows:

$$E[\phi_1(x, y, \xi)] = 5x^2 - 8x - \int_0^2 (1 + \frac{1}{2} t - t)dt - \int_2^{2x} (x + t - t)dt$$

$$= 4x^2 - 8x - 1,$$

$$E[\phi_2(x, y, \xi)] = 2x^2 - 16x + \int_0^2 [2(1 + \frac{1}{2} t) - t]dt + \int_2^{2x} [2(x + t - t)]dt$$

$$= 4x^2 - 16x + 4.$$

From the graph of bifunctions, it is easy to obtain $E_w = [1, 2] \times [2, 3]$. Take the evaluation function $\psi(z_1, z_2) = z_1 + z_2$, then the above multiobjective optimization problem becomes the following minimization of functions as follows:

$$\min \psi(E[\phi_1(x, y, \xi), \phi_2(x, y, \xi)]) = \min_{x \in [0, 3]} \left[ 8x^2 - 24x + 3 \right].$$

It is easily obtained that the optimal optimal solution is $x_{\min} = \frac{3}{2}$.

Let $\xi = (\xi^1, \xi^2, \cdots, \xi^N)$ be given, for each sample $\xi^k, k = 1, 2, \cdots, N$, and the corresponding solution of complementarity constraint problems is provided by

$$y^k(x, \xi^k) = \begin{cases} 1 + \frac{1}{2} \xi^k, & \text{if } \xi^k \leq 2 - 2x, \\ x + \xi^k, & \text{otherwise.} \end{cases}$$

The SAA bi-objective functions become

$$\hat{\phi}_1(x, y) = 5x^2 - 8x - \frac{1}{N} \sum_{k=1}^{N} y^k(x, \xi^k) + \frac{1}{N} \sum_{k=1}^{N} \xi^k,$$

$$\hat{\phi}_2(x, y) = 2x^2 - 16x + \frac{2}{N} \sum_{k=1}^{N} y^k(x, \xi^k) - \frac{1}{N} \sum_{k=1}^{N} \xi^k,$$

and the SAA scalarization optimization problem becomes

$$\min \psi(\hat{\phi}_1(x, y), \hat{\phi}_2(x, y)) = \min \left[ 7x^2 - 24x + \frac{1}{N} \sum_{k=1}^{N} y^k(x, \xi^k) \right].$$

For the sample size $N = 1, 2, \cdots, 300$, the optimal solutions of SAA scalarization problems and numerical results are provided in Figure 5. The component $x$ in $E_w$ is $[1, 2]$. By virtue of Theorem 3.6, the weakly (SAA) Pareto solution set is equivalent to the optimal solution set of (SAA) scalarization problems. It is shown that all the solutions belong to $[1, 2]$, which means that the deviation between the weakly SAA Pareto solution set and its true Pareto solution set is zero.

**Acknowledgments.** The authors wish to express their sincere thanks to two anonymous referees for their careful reading, helpful suggestions and comments, which improved the manuscript greatly.
Convergence of SAA optimal solutions.

REFERENCES

[1] Z. Artstein and R. A. Vitale, A strong law of large numbers for random compact sets, Ann. Probab., 3 (1975), 879-882.

[2] H. Bonnel and J. Collonge, Stochastic optimization over a Pareto set associated with a stochastic multi-objective optimization problem, J. Optim. Theory Appl., 162 (2014), 405–427.

[3] R. Caballero, E. Cerdá, M. M. Muñoz, L. Rey and I. M. Stancu-Minasian, Efficient solution concepts and their relations in stochastic multiobjective programming, J. Optim. Theory Appl., 110 (2001), 53–74.

[4] R. Caballero, E. Cerdá, M. M. Muñoz and L. Rey, Stochastic approach versus multiobjective approach for obtaining efficient solutions in stochastic multiobjective programming problems, European Journal of Operational Research, 158 (2004), 633–648.

[5] A. Chen, J. Kim, S. Lee and, Y. C. Kim, Stochastic multi-objective models for network design problem, Expert Systems with Applications, 37 (2010), 1608–1619.

[6] G. Y. Chen, X. X. Huang and X. Q. Yang, Vector Optimization Set-valued and Variational Analysis, Springer, Berlin, 2005.

[7] F. H. Clarke, Optimization and Nonsmooth Analysis, John Wiley & Sons, Inc., New York, 1983.

[8] K. Deb and A. Sinha, An efficient and accurate solution methodology for bilevel multi-objective programming problems using a hybrid evolutionary-local-search algorithm, Evolutionary Computation, 18 (2010), 403–449.

[9] J. Fliege and H. F. Xu, Stochastic multiobjective optimization: Sample average approximation and applications, J. Optim. Theory Appl., 151 (2011), 135–162.

[10] M. Fukushima, Fundamentals of Nonlinear Optimization, Asakura Shoten, Tokyo. 2001.

[11] A. Göpfert, C. Tammer, H. Riahi and C. Zălinescu, Variational Methods in Partially Ordered Spaces, Springer, New York, 2003.

[12] W. J. Gutjahr and A. Pichler, Stochastic multi-objective optimization: A survey on non-scalarizing methods, Ann. Oper. Res., 236 (2016), 475–499.

[13] R. Henrion, A. Jourani and J. Outrata, On the calmness of a class of multifunction, SIAM J. Optimization, 13 (2002), 603–618.

[14] J. Jahn, Vector Optimization, Springer, Berlin, 2004.

[15] S. Kim and J. H. Ryu, The sample average approximation method for multi-objective stochastic optimization, Proceeding of the 2011 Winter Simulation Conference, (2011), 4026–4037.
[16] G. H. Lin, D. L. Zhang and Y. C. Liang, Stochastic multiobjective problems with complementarity constraints and applications in healthcare management, European Journal of Operational Research, 226 (2013), 461–470.

[17] G. H. Lin, X. J. Chen and M. Fukushima, Solving stochastic mathematical programs with equilibrium constraints via approximation and smoothing implicit programming with penalization, Math. Program. Ser. B, 116 (2009), 343–368.

[18] K. Massimiliano and R. Daris, Multi-Objective stochastic optimization programs for a non-life insurance company under solvency constraints, Risk, 3 (2015), 390–419.

[19] K. Miettinen, Nonlinear Multiobjective Optimization, Kluwer Academic Publishers, Boston, MA, 1999.

[20] B. S. Mordukhovich, Equilibrium problems with equilibrium constraints via multiobjective optimization, Optimization Methods and Software, 19 (2004), 479–492.

[21] B. Mordukhovich, Variational Analysis and Generalized Differentiation I, Springer, Berlin, 2006.

[22] B. S. Mordukhovich and J. V. Outrata, Coderivative analysis of quasi-variational inequalities with applications to stability and optimization, SIAM J. Optim., 18 (2007), 752–777.

[23] B. S. Mordukhovich, Multiobjective optimization problems with equilibrium constraints, Mathematical Programming, 117 (2009), 331–354.

[24] L. P. Pang, F. Y. Meng, S. Chen and D. Li, Optimality condition for multi-objective optimization problem constrained by parameterized variational inequalities, Set-Valued and Variational Analysis, 22 (2014), 285–298.

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

[26] R. T. Rockafellar and R. J. B. Wets, Variational Analysis, Springer, Berlin, 1998.

[27] E. Roghanian, S. J. Sadjadi and M. B. Aryanezhad, A probabilistic bi-level linear multiobjective programming problem to supply chain planning, Applied Mathematics and Computation, 188 (2007), 786–800.

[28] A. Shapiro, Stochastic programming with equilibrium constraints, J. Optim. Theory Appl., 128 (2006), 223–243.

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

[30] R. Slowinski and J. Teghem, Stochastic Versus Fuzzy Approaches to Multiobjective Mathematical Programming Under Uncertainty, Kluwer Academic, Dordrecht, 1990.

[31] H. F. Xu and J. J. Ye, Approximating stationary points of stochastic mathematical programs with equilibrium constraints via sample averaging, Set-Valued Analysis, 19 (2011), 283–309.

[32] H. F. Xu and J. J. Ye, Necessary conditions for two-stage stochastic programs with equilibrium constraints, SIAM J. Optim, 20 (2010), 1685–1715.

[33] H. F. Xu, Uniform exponential convergence of sample average random functions under general sampling with applications in stochastic programming, Journal of Mathematical Analysis and Applications, 368 (2010), 692–710.

[34] J. J. Ye and Q. J. Zhu, Multiobjective optimization problems with variational inequality constraints, Mathematical Program, 96 (2003), 139–160.

[35] J. Zhang, L. W. Zhang and L. P. Pang, On the convergence of coderivative of SAA solution mapping for a parametric stochastic variational inequality, Set-Valued Ana, 20 (2012), 75–109.