QUANTITATIVE STABILITY ANALYSIS OF STOCHASTIC GENERALIZED EQUATIONS

YONGCHAO LIU, WERNER RÖMISCH, AND HUIFU XU

Abstract. We consider the solution of a system of stochastic generalized equations (SGE) where the underlying functions are mathematical expectation of random set-valued mappings. SGE has many applications such as characterizing optimality conditions of a nonsmooth stochastic optimization problem or equilibrium conditions of a stochastic equilibrium problem. We derive quantitative continuity of expected value of the set-valued mapping with respect to the variation of the underlying probability measure in a metric space. This leads to the subsequent qualitative and quantitative stability analysis of solution set mappings of the SGE. Under some metric regularity conditions, we derive Aubin’s property of the solution set mapping with respect to the change of probability measure. The established results are applied to stability analysis of stochastic variational inequality, stationary points of classical one-stage and two-stage stochastic minimization problems, two-stage stochastic mathematical programs with equilibrium constraints, and stochastic programs with second order dominance constraints.

Key words. stochastic generalized equations, stability analysis, equicontinuity, one-stage stochastic programs, two-stage stochastic programs, two-stage SMPECs, stochastic semi-infinite programming

AMS subject classifications. 90C15, 90C30, 90C33

DOI. 10.1137/120880434

1. Introduction. In this paper, we consider the following stochastic generalized equation (SGE):

\[ 0 \in E_P[\Gamma(x,\xi)] + G(x), \]

where \( \Gamma : \mathcal{X} \times \Xi \to 2^\mathcal{Y} \) and \( G : \mathcal{X} \to 2^\mathcal{Y} \) are closed set-valued mappings, \( \mathcal{X} \) and \( \mathcal{Y} \) are subsets of Banach spaces \( X \) and \( Y \), respectively, \( \xi : \Omega \to \Xi \) is a random vector defined on a probability space \((\Omega, \mathcal{F}, P)\) with support set \( \Xi \subseteq \mathbb{R}^d \) and probability distribution \( P \), and \( E_P[\cdot] \) denotes the expected value with respect to \( P \), that is,

\[ E_P[\Gamma(x,\xi)]:=\int_{\Xi} \Gamma(x,\xi)dP(\xi) = \left\{ \int_{\Xi} \psi(\xi)P(d\xi) : \psi \text{ is a Bochner integrable selection of } \Gamma(x,\cdot) \right\} . \]

The expected value of \( \Gamma \) is widely known as Aumann’s integral of the set-valued mapping; see [2, 3, 17].
The SGE formulation extends deterministic generalized equations [33] and underlines first order optimality/equilibrium conditions of nonsmooth stochastic optimization problems and stochastic equilibrium problems and stochastic games; see [30, 31] and references therein. In a particular case when $\Gamma$ is single valued and $\mathcal{G}(x)$ is a normal cone of a set, (1.1) is also known as a stochastic variational inequality for which a lot of research has been carried out over the past few years; see, for instance, [8, 46].

Our concern here is the stability of solutions of (1.1) as the underlying probability measure $P$ varies in some metric space. Apart from theoretical interest, the research is also numerically motivated: in practice, the probability measure $P$ may be unknown or numerically intractable but it can be estimated from historical data or approximated by numerically tractable measures. Consequently there is a need to establish a relationship between the set of solutions of the true problem and that of the approximated problem.

Let $Q$ denote a perturbation of the probability measure $P$. We consider the following perturbed SGE:

$$0 \in E_Q[\Gamma(x, \xi)] + \mathcal{G}(x).$$

Let $S(Q)$ and $S(P)$ denote the set of solutions to (1.1) and (1.2), respectively. We study the relationship between $S(Q)$ and $S(P)$ as $Q$ approximates $P$ under some appropriate metric.

There are two issues that we need to look into: (a) When $Q$ is “close” to $P$, does (1.2) have a solution? (b) Can we obtain a bound for the distance between the solutions to (1.1) and (1.2) in terms of certain distance between $Q$ and $P$? The first issue was investigated by Kummer [21] for a general class of deterministic parametric generalized equations in terms of solvability and further discussed by King and Rockafellar [20] under subinvertibility of a set-valued mapping. The second issue was considered in [45] under the context of perturbation of deterministic generalized equations.

In this paper, we derive quantitative continuity of $E_P[\Gamma(\cdot, \xi)]$ with respect to the variation of the probability measure $P$ in some metric spaces. This leads to the subsequent qualitative and quantitative stability analysis of the solution mappings of the SGE. Under some metric regularity conditions, we derive Aubin’s property of the solution set mapping with respect to the change of probability measure. The results are applied to study the stability of stationary points of a number of stochastic optimization problems. This effectively extends the stability analysis in the literature of stochastic optimization (see, e.g., Rachev and Römisch [29] and Römisch [36]) which relates optimal values and optimal solutions to stationary points. Moreover, the general framework of probability measure approximation extends recent work by Ralph and Xu [30] on asymptotic convergence of sample average approximation of SGEs where the true probability measure is approximated through a sequence of empirical probability measures and has a potential to be exploited to convergence analysis of stationary points when quasi–Monte Carlo methods are applied to nonsmooth stochastic optimization problems and nonsmooth stochastic games/equilibrium problems.

The rest of the paper is organized as follows. We start in section 2 by recalling some basic notions, concepts, and results on generalized equations, set-valued analysis, and Aumann’s integral of a set-valued mapping. In section 3, we present the main stability results concerning SGEs with respect to the perturbation of the probability measure. Applications of the established results to classical one-stage and two-stage linear stochastic programs and two-stage stochastic mathematical programs.
with complementarity constraints are in section 4, and finally we apply the results to stochastic programs with second order dominance constraints in section 5.

Throughout the paper, we use the following notation. \( Z \) denotes a Banach space with norm \( \| \cdot \|_Z \) and \( \mathbb{R}^n \) denotes \( n \) dimensional Euclidean space. By convention, we write \( \langle u, z \rangle \) for dual pairing of \( z \in Z \) which is bilinear, where \( u \) is from the dual space of \( Z \). In the case when \( Z \) is finite dimensional, the dual pairing reduces to a scalar product. Given a point \( x \in X \), we write \( d(x, S) := \inf_{z' \in S} \| z - z' \|_Z \) for the distance from \( x \) to \( S \). For two closed sets \( C \) and \( D \),

\[
\mathbb{D}(C, D) := \sup_{z \in C} d(z, D)
\]

stands for the deviation of set \( C \) from set \( D \), while \( \mathbb{H}(C, D) \) represents the Hausdorff distance between the two sets, that is,

\[
\mathbb{H}(C, D) := \max (\mathbb{D}(C, D), \mathbb{D}(D, C))
\]

In the case when \( C = \{0\} \), \( \mathbb{H}(0, D) = \mathbb{D}(D, 0) \) and we use \( \|D\| \) to denote the quantity. We use \( B(z, \delta) \) to denote the closed ball with radius \( \delta \) and center \( z \), that is, \( B(z, \delta) := \{ z' : \|z' - z\|_Z \leq \delta \} \), and \( B \) to denote the unit ball \( \{ z : \|z\| \leq 1 \} \) in a space. Finally, for a sequence of subsets \( \{S_k\} \) in a metric space, we follow the standard notation [2] by using \( \liminf_{k \to \infty} S_k \) to denote its upper limit, that is,

\[
\liminf_{k \to \infty} S_k = \{ x : \liminf_{k \to \infty} d(x, S_k) = 0 \}.
\]

2. Preliminary results. Let \( \Psi : X \to 2^Y \) be a set-valued mapping. Recall that \( \Psi \) is said to be closed at \( \bar{x} \) if \( x_k \in X, x_k \to \bar{x}, y_k \in \Psi(x_k) \), and \( y_k \to \bar{y} \) implies \( \bar{y} \in \Psi(\bar{x}) \). \( \Psi \) is said to be upper semicontinuous at \( \bar{x} \in X \) if and only if for any neighborhood \( U \) of \( \Psi(\bar{x}) \), there exists a positive number \( \delta > 0 \) such that for any \( x' \in B(x, \delta) \cap X \), \( \Psi(x') \subset U \). When \( \Psi(\bar{x}) \) is compact, \( \Psi \) is upper semicontinuous at \( \bar{x} \) if and only if for every \( \epsilon > 0 \), there exists a constant \( \delta > 0 \) such that

\[
\Psi(\bar{x} + \delta B) \subset \Psi(\bar{x}) + \epsilon B.
\]

\( \Psi \) is said to be lower semicontinuous at \( \bar{x} \in X \) if and only if for any \( \bar{y} \in \Psi(\bar{x}) \) and any sequence \( \{x_k\} \subset X \) converging to \( \bar{x} \), there exists a sequence \( \{y_k\} \), where \( y_k \in \Psi(x_k) \), converging to \( \bar{y} \). The lower semicontinuity holds if and only if for any open set \( U \) with \( U \cap \Psi(\bar{x}) \neq \emptyset \), the set \( \{ x \in X : U \cap \Psi(x) \neq \emptyset \} \) is a neighborhood of \( \bar{x} \). \( \Psi \) is said to be continuous at \( \bar{x} \) if it is both upper and lower semicontinuous at the point; see [2] for details.

2.1. Existence of a solution. We start by presenting a result that states existence of a solution to the perturbed generalized equations (1.2). The issue has been well investigated in the literature of deterministic generalized equations. For instance, Kummer [21] derived a number of sufficient conditions which ensure solvability (existence of a solution) of perturbed generalized equations. Similar conditions were further investigated by King and Rockafellar [20]. Here we present a stochastic analogue of one of the Kummer’s results.
Assumption 2.1. Let $Q$ be a perturbation of probability measure $P$ such that
(a) $\mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x)$ is nonempty and convex;
(b) for any $\epsilon > 0$, there exists a $\delta > 0$ such that
\begin{equation}
\mathbb{E}_{\epsilon}[\Gamma(x, \xi)] \subset \mathbb{E}_Q[\Gamma(x, \xi)] + \epsilon \mathcal{B}
\end{equation}
for all $x \in \mathcal{X}$ and $Q$ with $Q$ being sufficiently close to $P$ under some metric;
(c) for $\alpha \in \mathbb{R}_+$, the set
\[
\left\{ x \in \mathcal{X} : \inf_{\zeta \in \mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x)} \langle \zeta, u \rangle > \alpha \right\}
\]
is open for each $u$ in the unit ball of the dual space of $Y$.

The following result is a direct application of [21, Proposition 3].

Proposition 2.2. Let Assumption 2.1 hold. The perturbed generalized equations (1.2) have a solution for all $Q$ sufficiently close to $P$ if
\begin{equation}
\Delta(P) := \sup_{\|u\| = 1} \inf_{x \in \mathcal{X}} \inf_{\zeta \in \mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x)} \langle u, \zeta \rangle < 0.
\end{equation}

Proof. Let $\epsilon \in (0, \Delta(P))$ and $Q$ satisfy (2.1). Then for each $u$ in the unit ball of the dual space of $Y$
\[
\inf_{\zeta \in \mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x)} \langle u, \zeta \rangle \geq \inf_{x \in \mathcal{X}} \inf_{\zeta \in \mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x)} \langle u, \zeta \rangle - \epsilon.
\]
Therefore
\[
\inf_{x \in \mathcal{X}} \inf_{\zeta \in \mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x)} \langle u, \zeta \rangle \leq \inf_{x \in \mathcal{X}} \inf_{\zeta \in \mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x)} \langle u, \zeta \rangle + \epsilon \leq \Delta(P) + \epsilon < 0.
\]

By [21, Proposition 2], (1.2) has a solution. ☐

Assumption 2.1(a) is satisfied when $\Gamma(x, \xi)$ is convex set-valued mapping for almost every $\xi$ and $\mathcal{G}(x)$ is a convex set-valued mappings. In the case when $\Gamma$ is the Clarke subdifferential of a random function and $\mathcal{G}(x)$ is a normal cone to a convex set, the assumption is obviously satisfied. We will come back to this in sections 4 and 5. Assumption 2.1(b) means uniform Hausdorff continuity of set-valued mapping $\mathbb{E}_Q[\Gamma(x, \xi)]$ w.r.t. $Q$ at $Q = P$ in the case when the set-valued mapping is upper semicontinuous w.r.t. $Q$. Under a pseudometric to be defined in section 3, the continuity is guaranteed when $\Gamma(x, \xi)$ is bounded and continuous w.r.t. $\xi$ independent of $x$. Assumption 2.1(c) means that the set
\[
\left\{ x \in \mathcal{X} : \inf_{\zeta \in \mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x)} \langle u, \zeta \rangle \leq \alpha \right\}
\]
is closed and hence $\inf_{x \in \mathcal{X}} \inf_{\zeta \in \mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x)} \langle u, \zeta \rangle$ is well defined provided the quantity is lower bounded. Condition $\Delta(P) < 0$ implies that for any $u \in \mathcal{B}$, there exists $x \in \mathcal{X}$ such that $\inf_{\zeta \in \mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x)} \langle u, \zeta \rangle < 0$. By [21, Proposition 2] or the separation theorem, the latter means $0 \in \mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x)$.

2.2. Metric regularity.

Definition 2.3. Let $\Psi : \mathcal{X} \to 2^Y$ be a closed set-valued mapping. For $\bar{x} \in \mathcal{X}$ and $\bar{y} \in \Psi(\bar{x})$, $\Psi$ is said to be metrically regular at $\bar{x}$ for $\bar{y}$ if there exist a constant $\alpha > 0$, neighborhoods of $U$ of $\bar{x}$ and $V$ of $\bar{y}$ such that
\[
d(x, \Psi^{-1}(y)) \leq \alpha d(y, \Psi(x)) \quad \forall x \in U, y \in V.
\]
Here the inverse mapping \( \Psi^{-1} \) is defined as \( \Psi^{-1}(y) = \{ x \in X : y \in \Psi(x) \} \) and the minimal constant \( \alpha < \infty \) which makes the above inequality hold is called *regularity modulus* and is denoted by \( \text{reg} \Psi(\bar{x}|\bar{y}) \). \( \Psi(x) \) is said to be *strongly metrically regular* at \( \bar{x} \) for \( \bar{y} \) if it is metrically regular and there exist neighborhoods \( U_{\bar{x}} \) and \( U_{\bar{y}} \) such that for \( y \in U_{\bar{y}} \) there is only one \( x \in U_{\bar{x}} \cap \Psi^{-1}(y) \).

Metric regularity is a generalization of Jacobian nonsingularity of a vector-valued function to a set-valued mapping [32]. The property is equivalent to nonsingularity of the coderivative of \( \Psi \) at \( \bar{x} \) for \( \bar{y} \) and to Aubin’s property of \( \Psi^{-1} \). For a comprehensive discussion of the history and recent development of the notion, see [14, 35] and references therein.

Using the notion of metric regularity, one can analyze the stability of generalized equations. The following result is well known; see, for example, [46, Lemma 2.2].

**Proposition 2.4.** Let \( \Psi, \tilde{\Psi} : X \to 2^Y \) be two set-valued mappings. Let \( \bar{x} \in X \) and \( 0 \in \Psi(\bar{x}) \). Suppose that \( \Psi \) is metrically regular at \( \bar{x} \) for 0 with the neighborhoods of \( U_{\bar{x}} \) of \( \bar{x} \) and \( V_{\bar{y}} \) of \( \bar{y} \). If \( 0 \in \Psi(x) \) with \( x \in U_{\bar{x}} \), then

\[
\begin{align*}
d(x, \Psi^{-1}(0)) &\leq \alpha \mathcal{D}(\tilde{\Psi}(x), \Psi(x)),
\end{align*}
\]

where \( \alpha \) is the regularity modulus of \( \Psi \) at \( \bar{x} \) for 0. If \( \Psi(x) \) is strongly metrically regular at \( \bar{x} \) for 0, that is, there exist neighborhoods \( U_{\bar{x}} \) and \( U_{\bar{y}} \) such that for \( y \in U_{\bar{y}} \) there is only one \( x \in U_{\bar{x}} \cap \Psi^{-1}(y) \), then

\[
\|x - \bar{x}\| \leq \alpha \mathcal{D}(\tilde{\Psi}(x), \Psi(x)).
\]

**2.3. Fubini’s theorem for Aumann’s integral.** Let \( E \) be a Hausdorff locally convex vector space and \( E' \) the dual space. Let \( S \) be a nonempty subset of \( E \). The *support function* of \( S \) is the function defined on \( E' \) by

\[
u \to \sigma(S, u) = \sup_{a \in S} \langle u, a \rangle.
\]

The following result, which is widely known as the Hörmander theorem, establishes a relationship between the distance of two sets in \( E \) and the distance of their support functions over a unit ball in \( E' \).

**Lemma 2.5** (see [7, Theorem II-18]). Let \( \mathcal{C}, \mathcal{D} \) be nonempty, compact, and convex subsets of \( E \) with support functions \( \sigma(u, \mathcal{C}) \) and \( \sigma(u, \mathcal{D}) \). Then

\[
\begin{align*}
\mathcal{D}(\mathcal{C}, \mathcal{D}) &= \max_{\|u\| \leq 1} (\sigma(\mathcal{C}, u) - \sigma(\mathcal{D}, u))
\end{align*}
\]

and

\[
\begin{align*}
\mathcal{H}(\mathcal{C}, \mathcal{D}) &= \max_{\|u\| \leq 1} |\sigma(\mathcal{C}, u) - \sigma(\mathcal{D}, u)|.
\end{align*}
\]

Let \( X \) and \( Y \) be Banach spaces and \( Z \) a Hausdorff locally convex vector space. (Here we are slightly abusing the notation as \( X \) and \( Y \) have already been used in the definition of generalized equations (1.1).) Let \( \mu, \mu_x, \text{ and } \mu_y \) denote the bounded Borel measures in \( X \times Y \), \( X \), and \( Y \), respectively. Consider a nonempty compact and convex set-valued mapping \( \Psi : X \times Y \to 2^Z \) and its Aumann’s integrals \( \int X \times Y \Psi(x, y) \mu(dx dy) \), \( \int X \int Y \Psi(x, y) \mu_y(dy) \mu_x(dx) \), and \( \int Y \int X \Psi(x, y) \mu_x(dx) \mu_y(dy) \), where \( X \) and \( Y \) are nonempty compact subsets of \( X \) and \( Y \). The following proposition states that under some appropriate conditions, the three integrals are equal.

**Proposition 2.6.** Let \( X \) and \( Y \) be separable Banach space. Assume that \( \Psi \) is upper semicontinuous with respect to \( x \) and \( y \). Then the following assertions hold:
(i) $\sigma(\Psi(x,y), u)$ is upper semicontinuous in $x$ and $y$ uniformly w.r.t. $u$.

If, in addition, $\Psi$ is $\mu$-integrably bounded, that is, there exists a nonnegative $\mu$-integrable function $\kappa(x,y)$ with $\int_{X\times Y} \kappa(x,y) \mu(dx\,dy) < \infty$ such that
\[
\|\Phi(x,y)\| \leq \kappa(x,y),
\]
then
(ii) $\Psi(\cdot, y)$ and $\Psi(x, \cdot)$ are $\mu_x$ and $\mu_y$ integrably bounded for each $y$ and $x$, respectively, and
\[
\int_{X\times Y} \Psi(x,y) \mu(dx\,dy) = \int_X \int_Y \Psi(x,y) \mu_y(dy) \mu_x(dx) = \int_Y \int_X \Psi(x,y) \mu_x(dx) \mu_y(dy);
\]
(iii) for any $x', x \in \mathcal{X}$,
\[
\mathbb{H} \left( \int_X \Psi(x',y) \mu_y(dy), \int_X \Psi(x,y) \mu_y(dy) \right) \leq \int_X \mathbb{H}(\Psi(x',y), \Psi(x,y)) \mu_y(dy).
\]

Proof. The results are well known; see, for instance, [17, 47]. We give a proof for completeness.

Part (i). Since $\Psi$ is upper semicontinuous w.r.t. $x$ and $y$, it follows by Hörmander’s theorem that
\[
\sigma(\Psi(x',y'), u) - \sigma(\Psi(x,y), u) \leq \mathbb{D}(\Psi(x',y'), \Psi(x,y))
\]
which indicates that $\sigma(\Psi(x,y), u)$ is upper semicontinuous in $x$ and $y$ uniformly w.r.t. $u$.

Part (ii). By assumption, $\Psi$ is nonempty, compact, convex, and integrably bounded. It follows by [17, Theorem 5.4] that $\int_X \int_Y \Psi(x,y) \mu_y(dy) \mu_x(dx)$ and $\int_Y \int_X \Psi(x,y) \mu_y(dy) \mu_x(dx)$ are nonempty, compact, and convex. By Hörmander’s theorem (Lemma 2.5)
\[
\mathbb{D} \left( \int_X \int_Y \Psi(x,y) \mu_y(dy) \mu_x(dx), \int_Y \int_X \Psi(x,y) \mu_x(dx) \mu_y(dy) \right)
= \sup_{\|u\| \leq 1} \left[ \sigma \left( \int_X \int_Y \Psi(x,y) \mu_y(dy) \mu_x(dx), u \right) - \sigma \left( \int_Y \int_X \Psi(x,y) \mu_y(dy) \mu_x(dx), u \right) \right].
\]
Applying [27, Proposition 3.4] to the support function above, we have
\[
\sigma \left( \int_X \int_Y \Psi(x,y) \mu_y(dy) \mu_x(dx), u \right) = \int_X \int_Y \sigma(\Psi(x,y), u) \mu_y(dy) \mu_x(dx)
\]
and
\[
\sigma \left( \int_Y \int_X \Psi(x,y) \mu_x(dx) \mu_y(dy), u \right) = \int_Y \int_X \sigma(\Psi(x,y), u) \mu_x(dx) \mu_y(dy).
\]
It follows from part (i) that $\sigma(\Psi(x,y), u)$ is upper semicontinuous in $x$ and $y$. Since $\mathcal{X}$ and $\mathcal{Y}$ are compact sets $\Psi(x,y)$ is bounded, which implies the boundedness of $\sigma(\Psi(x,y), u)$. By Fubini’s theorem
\[
\int_X \int_Y \sigma(\Psi(x,y), u) \mu_y(dy) \mu_x(dx) = \int_Y \int_X \sigma(\Psi(x,y), u) \mu_x(dx) \mu_y(dy).
\]
The discussions above yield

\[
D \left( \int_X \int_Y \Psi(x,y) \mu_y(dy) \mu_x(dx), \int_Y \int_X \Psi(x,y) \mu_x(dx) \mu_y(dy) \right) \\
= \sup_{\|u\| \leq 1} \left[ \int_X \int_Y \sigma(\Psi(x,y), u) \mu_y(dy) \mu_x(dx) - \int_Y \int_X \sigma(\Psi(x,y), u) \mu_x(dx) \mu_y(dy) \right] \\
= 0.
\]

**Part (iii).** Indeed, following similar arguments as in the proof of part (ii), we have

\[
H \left( \int_Y \Psi(x',y) \mu_y(dy), \int_Y \Psi(x,y) \mu_y(dy) \right) \leq \int_Y \sup_{\|u\| \leq 1} \left| \sigma(\Psi(x,y), u) \right| \mu_y(dy) \\
= \int_Y H(\Psi(x',y), \Psi(x,y)) \mu_y(dy).
\]

The proof is complete. \(\square\)

### 3. Stability of SGEs

Let \(\mathcal{P}(\Xi)\) denote the set of all Borel probability measures on \(\Xi\). Assuming \(Q\) is close to \(P\) under some metric to be defined shortly, we investigate the relationship between the solution set of SGE (1.2) and that of (1.1).

Let \(\Gamma(x, \xi)\) be defined as in (1.1) and \(\sigma(\Gamma(x, \cdot), u)\) be its support function. Let \(\mathcal{X}\) be a compact subset of \(X\). Throughout this section, we assume that \(\Gamma(x, \xi)\) is nonempty, compact, and convex for every \(x \in X\) and \(\xi \in \Xi\). Define

\[
(3.1) \quad \mathcal{F} := \{ g(\cdot) : g(\xi) := \sigma(\Gamma(x, \xi), u) \text{ for } x \in \mathcal{X}, \|u\| \leq 1 \}.
\]

Then \(\mathcal{F}\) consists of all functions generated by the support function \(\sigma(\Gamma(x, \cdot), u)\) over the set \(\mathcal{X} \times \{ u : \|u\| \leq 1 \}\). Let

\[
\mathcal{D}(Q, P) := \sup_{g(\xi) \in \mathcal{F}} \left( \mathbb{E}_Q[g(\xi)] - \mathbb{E}_P[g(\xi)] \right)
\]

and

\[
\mathcal{H}(Q, P) := \max(\mathcal{D}(Q, P), \mathcal{D}(P, Q)).
\]

It is easy to verify that

\[
\mathcal{D}(Q, P) \geq \sup_{\|u\| \leq 1} \mathbb{E}_Q[\sigma(\Gamma(x, \xi), u)] - \mathbb{E}_P[\sigma(\Gamma(x, \xi), u)] \geq 0 \quad \forall x \in \mathcal{X}.
\]

We will use this relationship later on. Note that by [27, Proposition 3.4],

\[
\mathbb{E}_Q[\sigma(\Gamma(x, \xi), u)] - \mathbb{E}_P[\sigma(\Gamma(x, \xi), u)] = \sigma(\mathbb{E}_Q[\Gamma(x, \xi)], u) - \sigma(\mathbb{E}_P[\sigma(\Gamma(x, \xi)], u).
\]

By Lemma 2.5, the inequality above implies

\[
\mathcal{D}(Q, P) \geq \mathbb{D}(\mathbb{E}_Q[\Gamma(x, \xi)], \mathbb{E}_P[\Gamma(x, \xi)]) \geq 0 \quad \forall x \in \mathcal{X}
\]

and hence

\[
\mathcal{D}(Q, P) = 0 \implies \mathbb{E}_Q[\Gamma(x, \xi)] \subseteq \mathbb{E}_P[\Gamma(x, \xi)] \quad \forall x \in \mathcal{X}.
\]
Likewise
\[ H(Q, P) = 0 \implies \mathbb{E}_Q[\Gamma(x, \xi)] = \mathbb{E}_P[\Gamma(x, \xi)] \quad \forall x \in X. \]
Neither \( H \) nor \( D \) is a metric but one may enlarge the set \( \mathcal{F} \) so that \( H(Q, P) = 0 \) implies \( Q = P \). We call \( H(Q, P) \) a pseudometric. It is also known as a distance of \( \zeta \)-structure; see [48].

Recall that for a sequence of probability measures \( \{P_N\} \) in \( \mathcal{P}(\Xi) \), \( P_N \) is said to converge weakly to \( P \) if
\[
\lim_{N \to \infty} \mathbb{E}_{P_N}[g(\xi)] = \mathbb{E}_P[g(\xi)]
\]
for every bounded continuous real-valued function \( g \) on \( \Xi \).

Let \( \mathcal{F} \) be defined by (3.1) and \( \{P_N\} \subset \mathcal{P}(\Xi) \). We say \( \mathcal{F} \) defines an upper \( P \)-uniformity class of functions if
\[
\lim_{N \to \infty} \mathcal{F}(P_N, P) = 0
\]
for every sequence \( \{P_N\} \) which converges weakly to \( P \) and a \( P \)-uniformity class if
\[
\lim_{N \to \infty} H(P_N, P) = 0.
\]
A family of functions \( \mathcal{K} \) is said to be equicontinuous at a point \( x_0 \) if for every \( \epsilon > 0 \), there exists a \( \delta > 0 \) such that \( \|f(x_0) - f(x)\| < \epsilon \) for all \( f \in \mathcal{K} \) and all \( x, x_0 \) such that \( \|x_0 - x\| < \delta \). A sufficient condition for \( \mathcal{K} \) to be a \( P \)-uniformity class is that \( \mathcal{K} \) is uniformly bounded and
\[
P(\{\xi \in \Xi : \mathcal{K} \text{ is not equicontinuous at } \xi\}) = 0;
\]
see [41]. In our context, the latter is implied by
\[
\lim_{\xi \to \xi', x \in X} \mathbb{E}(\Gamma(x, \xi'), \Gamma(x, \xi)) = 0
\]
for almost every \( \xi \) w.r.t. probability measure \( P \). At this point, we refer readers to the work by Artstein and Wets [1] on the approximation of Aumann’s integral of multifunctions, where the authors showed \( \mathbb{E}_{P_N}[\Gamma(x, \xi)] \) converges to \( \mathbb{E}_P[\Gamma(x, \xi)] \) when \( P_N \) converges weakly to \( P \) and \( \Gamma \) takes convex and compact values and is continuous in \( \xi \); see [1, Theorem 3.1].

**Theorem 3.1.** Consider the SGE (1.1) and its perturbation (1.2). Let \( X \) be a compact subset of \( X \), and let \( S(P) \) and \( S(Q) \) denote the sets of solutions of (1.1) and (1.2) restricted to \( X \), respectively, with \( \text{cl} S(P) \neq X \), where \( \text{cl} \) denotes the closure of a set. Assume (a) \( Y \) is a Euclidean space and \( \Gamma \) is a set-valued mapping taking convex and compact sets in \( Y \); (b) \( \Gamma \) is upper semicontinuous with respect to to \( x \) for every \( \xi \in \Xi \) and bounded by a \( P \)-integrable function \( \kappa(\xi) \) for \( x \in X \); (c) \( \mathcal{G} \) is upper semicontinuous; (d) \( S(Q) \) is nonempty for \( Q \in \mathcal{P}(\Omega) \) and \( \mathcal{D}(Q, P) \) sufficiently small. Then the following assertions hold:

(i) For any small positive number \( \epsilon \), let
\[
R(\epsilon) := \inf_{x \in X, d(x, S(P)) \geq \epsilon} d(0, \mathbb{E}_P[\Gamma(x, \xi)] + \mathcal{G}(x)).
\]
Then
\[
\mathbb{D}(S(Q), S(P)) \leq R^{-1}(2\mathcal{D}(Q, P)),
\]
where \( R^{-1}(\epsilon) := \min\{t \in \mathbb{R}_+ : R(t) = \epsilon\} \), and \( R^{-1}(\epsilon) \to 0 \) as \( \epsilon \downarrow 0 \).
(ii) For any small positive number $\epsilon$, there exists a $\delta > 0$ such that if $\mathcal{D}(Q,P) \leq \delta$, then $\mathbb{D}(S(Q),S(P)) \leq \epsilon$.

(iii) If $x^* \in S(P)$ and $\Phi(x) := \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x)$ is metrically regular at $x^*$ for 0 with regularity modulus $\alpha$, then there exists neighborhood $U_{x^*}$ of $x^*$ such that

$$
(3.8) 
\quad \quad d(x,S(P)) \leq \alpha \mathcal{D}(Q,P)
$$

for $x \in S(Q) \cap U_{x^*}$: if $\Phi$ is strongly metrically regular at $x^*$ for 0 with the same regularity modulus and neighborhood, then

$$
(3.5) 
\quad \quad \|x - x^*\| \leq \alpha \mathcal{D}(Q,P)
$$

for $x \in S(Q)$ close to $\Phi^{-1}(0)$.

Proof. Let $\{x_N\} \subset \mathcal{X}$ be a sequence such that $x_N \rightarrow x$ as $N \rightarrow \infty$. Under conditions (a) and (b), $\Gamma(x,\xi)$ is upper semicontinuous and integrably bounded, and the space $Y$ is finite dimensional (separable and reflexive). By [18, Theorem 2.8] (see also [24, Theorem 1.43]),

$$
(3.6) 
\quad \quad \limsup_{x_N \rightarrow x, x \in \mathcal{X}} \mathbb{E}_P[\Gamma(x_N,\xi)] \subset \mathbb{E}_P \left[ \limsup_{x_N \rightarrow x, x \in \mathcal{X}} \Gamma(x_N,\xi) \right] \subset \mathbb{E}_P [\Gamma(x,\xi)].
$$

Parts (i) and (ii). Let $R(\epsilon)$ be defined by (3.3). It is easy to observe that $R(0) = 0$ and $R(\epsilon)$ is nondecreasing on $[0, \infty)$. In what follows, we show that $R(\epsilon) > 0$ for $\epsilon > 0$. Assume for a contradiction that $R(\epsilon) = 0$. Then there exists a sequence $\{x_N\} \subset \mathcal{X}$ with $d(x_N,S(P)) \geq \epsilon$ such that

$$
lim_{N \rightarrow \infty} d(0,\mathbb{E}_P[\Gamma(x_N,\xi)] + \mathcal{G}(x_N)) = 0,
$$

which is equivalent to

$$
(3.7) 
\quad \quad 0 \in \limsup_{x_N \rightarrow x, x \in \mathcal{X}} (\mathbb{E}_P[\Gamma(x_N,\xi)] + \mathcal{G}(x_N)).
$$

Since $\mathcal{X}$ is a compact set, we may assume without loss of generality that $x_N \rightarrow x^*$ for some $x^* \in \mathcal{X}$. Using the upper semicontinuity of $\mathcal{G}(x)$ and (3.6), we derive from (3.7) that

$$
0 \in \limsup_{N \rightarrow \infty} (\mathbb{E}_P[\Gamma(x_N,\xi)] + \mathcal{G}(x_N)) \subseteq \mathbb{E}_P \left[ \limsup_{N \rightarrow \infty} \Gamma(x_N,\xi) \right] + \mathcal{G}(x^*) \subset \mathbb{E}[\Gamma(x^*,\xi)] + \mathcal{G}(x^*).
$$

The formula above shows $x^* \in S(P)$, which contradicts the fact that $d(x^*,S(P)) \geq \epsilon$. This implies that $R^{-1}(\epsilon) \rightarrow 0$ as $\epsilon \downarrow 0$.

Let $\delta := R(\epsilon)/2$ and $\mathcal{D}(Q,P) \leq \delta$. Let $\rho' := \min_{x \in \mathcal{X}} d(0,\mathcal{G}(x))$. Under the closedness and upper semicontinuity of $\mathcal{G}(-)$, it is easy to verify that $\rho' < \infty$. Let

$$
\rho := \rho' + \sup_{x \in \mathcal{X}} \max(\|\mathbb{E}_P[\Gamma(x,\xi)]\|,\|\mathbb{E}_Q[\Gamma(x,\xi)]\|).
$$

Under condition (b) and compactness of $\mathcal{X}$, it is easy to show that $\rho < \infty$. Let $t$ be any fixed positive number such that $t > \rho$. Then for any point $x \in \mathcal{X}$ with $d(x,S(P)) > \epsilon$,

$$
\begin{align*}
\mathbb{D}(0,\mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x)) &= d(0,\mathbb{E}_Q[\Gamma(x,\xi)] + \mathcal{G}(x) \cap t\mathcal{B}) \\
&\geq d(0,\mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x) \cap t\mathcal{B}) \\
&\quad - \mathbb{D}(\mathbb{E}_Q[\Gamma(x,\xi)] + \mathcal{G}(x) \cap t\mathcal{B},\mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x) \cap t\mathcal{B}) \\
\end{align*}
$$

(3.8)
where $\mathcal{B}$ denotes the unit ball in space $\mathcal{Y}$. Using the definition of $\mathcal{D}$, it is easy to show that

$$
\mathcal{D}(\mathbb{E}_Q[\Gamma(x,\xi)] + \mathcal{G}(x) \cap t\mathcal{B}, \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x) \cap t\mathcal{B}) \leq \mathcal{D}(\mathbb{E}_Q[\Gamma(x,\xi)], \mathbb{E}_P[\Gamma(x,\xi)]);
$$

see, for instance, the proof of [45, Lemma 4.2]. By invoking Hörmander’s theorem and [27, Proposition 3.4], we have

$$
\mathcal{D}(\mathbb{E}_Q[\Gamma(x,\xi)], \mathbb{E}_P[\Gamma(x,\xi)]) = \sup_{\|u\| \leq 1} (\sigma(\mathbb{E}_Q[\Gamma(x,\xi)], u) - \sigma(\mathbb{E}_P[\Gamma(x,\xi)], u))
= \sup_{\|u\| \leq 1} (\mathbb{E}_Q[\sigma(\Gamma(x,\xi), u)] - \mathbb{E}_P[\sigma(\Gamma(x,\xi), u)]).
$$

By the definition of $\mathcal{D}(Q,P)$,

$$
\sup_{\|u\| \leq 1} (\mathbb{E}_Q[\sigma(\Gamma(x,\xi), u)] - \mathbb{E}_P[\sigma(\Gamma(x,\xi), u)]) \leq \mathcal{D}(Q,P).
$$

Combining (3.8)–(3.11), we have

$$
d(0, \mathbb{E}_Q[\Gamma(x,\xi) + \mathcal{G}(x)] \cap t\mathcal{B}) \geq d(0, \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x)) - \mathcal{D}(Q,P)
\geq R(\epsilon) - \delta
\geq \delta > 0.
$$

This shows $x \notin S(Q)$ for any $x \in \mathcal{X}$ with $d(x, S(P)) > \epsilon$, which implies

$$
\mathcal{D}(S(Q), S(P)) \leq \epsilon.
$$

Let $\epsilon$ be the minimal value such that $\frac{1}{2}R(\epsilon) = \mathcal{D}(Q,P) = \delta$. Then (3.12) implies

$$
\mathcal{D}(S(Q), S(P)) \leq \epsilon = R^{-1}(2\mathcal{D}(Q,P)).
$$

**Part (iii).** Let $\mathcal{B}$ denote the unit ball of $\mathcal{Y}$ and $t$ be a constant such that

$$
t > \text{max}\{\|\mathbb{E}_Q[\Gamma(x,\xi)]\|, \|\mathbb{E}_P[\Gamma(x,\xi)]\|\}.
$$

Then for any $x \in \Phi^{-1}(0) \cap \mathcal{X}$

$$
0 \in \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x) \cap t\mathcal{B}.
$$

Likewise, for $x \in S(Q)$,

$$
0 \in \mathbb{E}_Q[\Gamma(x,\xi)] + \mathcal{G}(x) \cap t\mathcal{B}.
$$

On the other hand, the metric regularity of $\Phi(x)$ at $x^*$ for 0 with regularity modulus $\alpha$ implies that there exists neighborhood $U_{x^*}$ of $x^*$ such that

$$
d(x, S(P)) \leq \alpha d(0, \Phi(x))
$$

for all $x \in S(Q) \cap U_{x^*}$. Since

$$
\Phi(x) = \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x) \supset \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x) \cap t\mathcal{B},
$$
then
\[ d(0, \Phi(x)) \leq d(0, E_P[\Gamma(x, \xi)] + G(x) \cap tB) \]
and hence
\[ d(x, S(P)) \leq ad(0, E_P[\Gamma(x, \xi)] + G(x) \cap tB) \]
\[ \leq aD(E_Q[\Gamma(x, \xi)] + G(x) \cap tB, E_P[\Gamma(x, \xi)] + G(x) \cap tB) \]
(3.15)
for all \( x \in S(Q) \cap U_{x^*} \). The second inequality is due to (3.13) and the definition of \( D \). Note that for any bounded sets \( \mathcal{C}, \mathcal{C}', \mathcal{D}, \mathcal{D}' \), it is easy to verify that
\[ D(\mathcal{C} + \mathcal{C}', \mathcal{D} + \mathcal{D}') \leq D(\mathcal{C}, \mathcal{D}) + D(\mathcal{C}', \mathcal{D}'). \]

Using this relationship and (3.9)–(3.11), we obtain
\[ D(E_Q[\Gamma(x, \xi)] + G(x) \cap tB, E_P[\Gamma(x, \xi)] + G(x) \cap tB) \leq D(E_Q[\Gamma(x, \xi)], E_P[\Gamma(x, \xi)]) \]
(3.16)
\[ \leq \mathcal{D}(Q, P). \]
Combining (3.14), (3.15), and (3.16), we obtain (3.4). Inequality (3.5) follows straightforwardly from (3.4) and strong metric regularity.

In general, it is difficult to derive the rate function \( R^{-1}(\epsilon) \). Here we consider two particular cases that we may derive an estimate of \( R^{-1}(\epsilon) \).

**Corollary 3.2.** Let \( \Phi(x) := E_P[\Gamma(x, \xi)] + G(x) \) and \( V := \{ v : v \in \Phi(x) \text{ for all } x \in S(P) \} \). Let \( \epsilon \) be a small positive number. Assume that for any \( x \) with \( d(x, S(P)) \geq \epsilon \), there exist positive constants \( C \) and \( \tau \) (depending on \( \epsilon \)) such that
\[ \|v' - v\| \geq Cd(x, S(P))^\tau \quad \forall v' \in \Phi(x), v \in V. \]
Then there exists a positive constant \( \alpha \) such that
\[ R^{-1}(\epsilon) \leq \alpha \epsilon^{\frac{1}{\tau}}. \]

**Proof.** By definition,
\[ R(\epsilon) = \inf_{x \in X, d(x, S(P)) \geq \epsilon} d(0, E_P[\Gamma(x, \xi)] + G(x)) \]
\[ \geq \inf_{x \in X, d(x, S(P)) \geq \epsilon} \inf_{x^* \in S(P)} \inf_{v \in \Phi(x^*)} d(v, E_P[\Gamma(x, \xi)] + G(x)) \]
\[ = \inf_{x \in X, d(x, S(P)) \geq \epsilon} \inf_{x^* \in S(P)} \inf_{v, v' \in \Phi(x)} \|v - v'|\| \]
\[ \geq \inf_{x \in X, d(x, S(P)) \geq \epsilon} Cd(x, S(P))^\tau \]
\[ \geq C\epsilon^\tau, \]
where the second to last inequality follows from (3.17). The conclusion follows by setting \( \alpha := C^{-\frac{1}{\tau}} \).

Condition (3.17) is a kind of growth condition for the set-valued mapping \( \Phi(x) \). To see this, consider a simple example with \( \Phi(x) = x^2 \), where \( x \in \mathbb{R} \). In this case, \( S(P) = \{0\} \) and \( V = \{0\} \). For any fixed \( \epsilon \),
\[ \|v' - v\| = |x^2 - 0| \geq \epsilon d(x, 0) \quad \forall v' \in \Phi(x), v \in V. \]
Moreover, since \( F \) which is known as Robinson’s normal map. Let
\[
\|v' - v, x' - x\| \geq C^*\|x' - x\|^2 \quad \forall \ v' \in \Phi(x'), \ v \in \Phi(x);
\]
see [5, 6] and [35, Definition 12.53] for the finite dimensional case. Under the strong
monotonicity
\[
\|v' - v\|\|x' - x\| \geq (v' - v, x' - x) \geq C^*\|x' - x\|^2 \quad \forall \ v' \in \Phi(x'), \ v \in \Phi(x),
\]
which implies \( \|x' - x^*\| \geq d(x, S(P)) \) and hence (3.17) with \( C = C^* \) and \( \tau = 1 \). A well
known example for strong monotonicity is the subdifferential mapping of a strongly
convex function; see [35] for the latter.

Let us now consider the case when \( \Gamma(\cdot, \xi) \) is single valued for almost every \( \xi \) and it is Lipschitz continuous over \( \mathcal{X} \subseteq \mathbb{R}^n \) with integrable Lipschitz modulus \( \kappa(\xi) \). Moreover \( \mathcal{G}(x) = \mathcal{N}_K(x) \), where \( K \) is a polyhedral in \( \mathbb{R}^n \) and \( \mathcal{N}_K(x) \) denotes the
normal cone to \( K \) at point \( x \). Under these circumstances, SGE (1.1) can be written
as a stochastic variational inequality problem (SVIP)
\[
(3.19) \quad 0 \in \mathbb{E}_P[\Gamma(x, \xi)] + \mathcal{N}_K(x).
\]
Observe that
\[
d(0, \mathbb{E}_P[\Gamma(x, \xi)] + \mathcal{N}_K(x)) = d(-\mathbb{E}_P[\Gamma(x, \xi)], \mathcal{N}_K(x)).
\]
By [16, Proposition 1.5.14],
\[
d(-\mathbb{E}_P[\Gamma(x, \xi)], \mathcal{N}_K(x)) = \inf\{\|F_K^{\text{nor}}(z)\| : z \in \Pi_K^{-1}(x)\},
\]
where \( \Pi_K \) denotes the Euclidean projection onto \( K \) and \( \Pi_K^{-1} \) its inverse,
\[
F_K^{\text{nor}}(z) := \mathbb{E}_P[\Gamma(\Pi_K(z), \xi)] + z - \Pi_K(z),
\]
which is known as Robinson’s normal map. Let \( Z = \mathcal{X} \). It is easy to verify that
\( F_K^{\text{nor}}(z) \) is Lipschitz continuous on \( Z \) and with modulus being bounded by \( \mathbb{E}[\kappa(\xi)] + 2 \).
Moreover, since \( K \) is polyhedral, it follows by [26, Theorem 2.7] that \( \mathcal{N}_K \) is a polyhe-
dral multifunction and through [26, Theorem 2.4] is locally upper Lipschitz continuous.
Using the relationship
\[
\Pi_K^{-1}(x) = (\mathcal{N}_K + I)(x),
\]
where \( I \) denotes the identity mapping, we conclude that the set-valued mapping \( \Pi_K^{-1} \)
is locally upper Lipschitz continuous.

**Corollary 3.3.** Consider (3.19). Let \( S(P) \) denote its solution set, \( x^* \in S(P) \),
and \( Z^* = \Pi_K^{-1}(S(P)) \). Assume that there exist positive constants \( C \) and \( \tau \) such that
\[
(3.20) \quad \|F_K^{\text{nor}}(z)\| \geq Cd(z, Z^*)^\tau \quad \forall z \in \Pi_K^{-1}(\mathcal{X}),
\]
that is, \( \|F_K^{\text{nor}}(z)\| \) satisfies some growth condition as \( z \) deviating from \( Z^* \). Then
there exists a positive constant \( \alpha \) such that (3.18) holds. If, in addition, \( F_K^{\text{nor}}(z) \) is

\[\text{Note that there are many types of SVIP models. For instances, Chen, Wets, and Zhang [9]}
\[\text{recently proposed an SVIP model where } K \text{ depends on every realization of random variable } \xi \text{ and}
\text{a deterministic solution } x \text{ is sought for solving } 0 \in \Gamma(x, \xi) + \mathcal{N}_K(\xi)(x) \text{ for all } \xi \in \Xi. \text{ It is unclear}
\text{whether results to be presented here can be established for the new SVIP models in the same manner.}\]
a locally Lipschitz homeomorphism near \( z^* \), that is, there exist neighborhoods of \( z^* \) and \( F_K^{\text{nor}}(z^*) \) such that the map \( F_K^{\text{nor}}(\cdot) \) restricted to the neighborhood is bijective and its inverse is also Lipschitz, then (3.18) holds with \( \tau = 1 \).

Proof. Let \( x \in \mathcal{X} \). Note that \( \Pi_K^{-1}(x) \) may be set valued. Under condition (3.20),

\[
\inf_{z \in \Pi_K^{-1}(x)} \| F_K^{\text{nor}}(z) \| = \inf_{z \in \Pi_K^{-1}(x), z^* \in Z^*} (\| F_K^{\text{nor}}(z) \| - \| F_K^{\text{nor}}(z^*) \|)
\] \[
\geq \inf_{z \in \Pi_K^{-1}(x)} C d(z, Z^*)^\tau.
\]

With this, we can estimate \( R(\epsilon) \). By definition,

\[
R(\epsilon) = \inf_{x \in \mathcal{X}, \|x - x^*\| \geq \epsilon} d(0, \mathbb{E}_P[\Gamma(x, \xi)] + \mathcal{N}_K(x)) = \inf_{z \in \Pi_K^{-1}(x), x \in \mathcal{X}, \|d(x, S(P))\| \geq \epsilon} \| F_K^{\text{nor}}(z) \|
\] \[
\geq \inf_{z \in \Pi_K^{-1}(x), x \in \mathcal{X}, \|d(x, S(P))\| \geq \epsilon} C \|x - x^*\| \tau \quad \text{(since \( \|z - z^*\| \geq \|x - x^*\| \))}
\] \[
= C \epsilon^\tau.
\]

If, in addition, \( F_K^{\text{nor}}(z) \) is a locally Lipschitz homeomorphism near \( z^* \), then \( Z^* \) reduces to a singleton, denoted by \( \{z^*\} \), and \( S(P) \) to \( \{x^*\} \). Following an argument similar to the first part of the proof, we have

\[
\inf_{z \in \Pi_K^{-1}(x)} \| F_K^{\text{nor}}(z) \| = \inf_{z \in \Pi_K^{-1}(x)} (\| F_K^{\text{nor}}(z) \| - \| F_K^{\text{nor}}(z^*) \|)
\] \[
\geq \inf_{z \in \Pi_K^{-1}(x)} C' d(z, z^*),
\]

where \( C' \) is a positive constant. Consequently

\[
R(\epsilon) = \inf_{x \in \mathcal{X}, \|x - x^*\| \geq \epsilon} d(0, \mathbb{E}_P[\Gamma(x, \xi)] + \mathcal{N}_K(x)) = \inf_{z \in \Pi_K^{-1}(x), x \in \mathcal{X}, \|d(x, S(P))\| \geq \epsilon} \| F_K^{\text{nor}}(z) \|
\] \[
\geq \inf_{z \in \Pi_K^{-1}(x), x \in \mathcal{X}, \|d(x, S(P))\| \geq \epsilon} C' \|z - z^*\| \quad \text{(since \( \|z - z^*\| \geq \|x - x^*\| \))}
\] \[
= C' \epsilon.
\]

The conclusions follow.

Note that part (iii) of Theorem 3.1 is derived under metric regularity. It is difficult to verify the condition in general. However, when either \( \mathbb{E}_P[\Gamma(x, \xi)] \) or \( G(x) \) reduces to a singleton, then we may characterize the metric regularity of \( \Psi(x) = \mathbb{E}_P[\Gamma(x, \xi)] + G(x) \) through the Mordukhovich coderivative; see [14]. For example, when \( \mathbb{E}_P[\Gamma(x, \xi)] \) is single valued and \( G(x) \) is a normal cone, [14, Theorem 5.1] gives details on this. In the case when the SGE represents the KKT conditions of a one-stage smooth stochastic equality and inequality constrained problems, the metric regularity conditions are equivalent to some stability/error bound conditions and the latter are implied by the second order sufficient conditions; see this sort of discussions in [19] for deterministic minimization problems.
Note also that when $\Gamma(\cdot, \xi)$ is continuously differentiable for every $\xi$ and $G(x)$ is independent of $x$, e.g., $G(x) = G$, where $G$ is a closed convex set in $Y$, the SGE recovers a stochastic cone constraint. In that case, the metric regularity of $\mathbb{E}_P[\Gamma(x, \xi)] + G$ is equivalent to Robinson’s constraint qualification, that is, there exists $x_0 \in X$ such that

$$0 \in \text{int}\{\mathbb{E}_P[\Gamma(x, \xi)] + \langle \mathbb{E}_P[\nabla_x \Gamma(x, \xi)], X \rangle + G,$$

where “int” denotes interior of a set; see [4, Proposition 2.89].

Finally, it is possible to obtain a linear lower bound for $R(\epsilon)$ without metric regularity condition. Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be such that $F(x) = (x_1, 0)$. Consider equation

$$F(x) = 0$$

and its solution restricted to set $X = \{(x_1, x_2) : ||x||_{\infty} \leq 1\}$. This is a very special generalized equation: it is deterministic and linear. The solution set $X^* = \{(x_1, x_2) \in X : x_1 = 0\}$. The function $F(x)$ is not metrically regular because its Jacobian is singular. However, for small positive number $\epsilon$,

$$R(\epsilon) = \inf_{x \in X, d(x, X^*) \geq \epsilon} d(0, F(x)) = \inf_{x \in X, |x_1| \geq \epsilon} |x_1| = \epsilon.$$

Remark 3.4. The assumption of $Y$ to be a Euclidean space (finite dimensional) is only required in (3.6). In some applications, $\Gamma$ may consist of components which are single valued. It is easy to observe that so long as the set-valued components are finite dimensional, the conclusion holds even when the single-valued components are infinite dimensional. We need this argument in section 5.

4. Stochastic minimization problems. In this section, we use the stability results on the SGEs derived in the preceding section to study stability of stationary points of stochastic optimization problems. This is motivated to complement the existing research on stability analysis of optimal values and optimal solutions in stochastic programming [36].

4.1. One-stage stochastic programs with deterministic constraints. Let us start with one-stage problems. To simplify notation, we consider the following nonsmooth stochastic minimization problem:

$$\min_{x} \mathbb{E}_P[f(x, \xi)]$$

subject to $x \in X$,

(4.1)

where $X$ is a closed subset of $\mathbb{R}^n$, $f : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous, and for every fixed $\xi \in \Xi$, the function $f(\cdot, \xi)$ is locally Lipschitz continuous on its domain but not necessarily continuously differentiable or convex, and $P$ is the probability distribution of random vector $\xi : \Omega \to \Xi \subset \mathbb{R}^k$ defined on some probability space $(\Omega, \mathcal{F}, P)$. Note that by allowing $f$ to be nonsmooth, the model subsumes a number of stochastic optimization problems with stochastic constraints and two-stage stochastic optimization problems.

To simplify the discussion, we assume that $\mathbb{E}_P[f(\cdot, \xi)]$ is well defined for some $x_0 \in X$ and the Lipschitz modulus of $f(\cdot, \xi)$ is integrably bounded with respect to the probability measure $P$. It is easy to observe that the assumption implies $\mathbb{E}_P[f(x, \xi)]$ is well defined for every $x \in X$ and that $\mathbb{E}_P[f(\cdot, \xi)]$ is locally Lipschitz continuous.
Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a locally Lipschitz continuous function. Recall that the Clarke subdifferential of $\psi$ at $x$, denoted by $\partial \psi(x)$, is defined as follows:

$$\partial \psi(x) := \text{conv} \left\{ \lim_{x' \to x} \nabla \psi(x') \right\},$$

where $D$ denotes the set of points near $x$ at which $\psi$ is Fréchet differentiable, $\nabla \psi(x)$ denotes the gradient of $\psi$ at $x$, and conv denotes the convex hull of a set; see [10] for details.

Using Clarke’s subdifferential, we may consider the first order optimality conditions of problem (4.1). Under some appropriate constraint qualifications, a local optimal solution $x^* \in X$ to problem (4.1) necessarily satisfies the following:

$$0 \in \partial_E P[f(x,\xi)] + N_X(x).$$

The condition is also sufficient if $f(\cdot,\xi)$ is convex for almost every $\xi$. In general, a point $x \in X$ satisfying (4.2) is called a stationary point. A slightly weaker first optimality condition which is widely discussed in the literature is

$$0 \in E_P[\partial_x f(x,\xi)] + N_X(x).$$

The condition is weaker in that $\partial E_P[f(x,\xi)] \subseteq E_P[\partial_x f(x,\xi)]$ and equality holds only under some regularity conditions. A point $x \in X$ satisfying (4.3) is called a weak stationary point of problem (4.1). For a detailed discussion on the well-definedness of (4.2) and (4.3) and the relationship between stationary point and weak stationary point, see [45] and references therein.

Let us now consider a perturbation of the stochastic minimization problem:

$$\min_x E_Q[f(x,\xi)]$$

s.t. $x \in X$,

where $Q$ is a perturbation of the probability measure $P$ such that $E_Q[f(x,\xi)]$ is well defined for some $x_0 \in X$ and the Lipschitz modulus of $f$ is integrably bounded with respect to $Q$. In the literature of stochastic programming, quantitative stability analysis concerning optimal values and optimal solutions in relation to the variation of the underlying probability measure is well known; see, for instance, [36, 29]. Our focus here is on stationary points. Let $X(P)$ and $X(Q)$ denote the sets of stationary points of problems (4.1) and (4.4) and $\tilde{X}(P)$ and $\tilde{X}(Q)$ the sets of weak stationary points, respectively. We use Theorem 3.1 to investigate stability of the stationary points.

**Theorem 4.1.** Let $f^u(x,\xi;u)$ denote the Clarke generalized directional derivative for a given nonzero vector $u$ and

$$\mathcal{F} := \{ g : g(\cdot) := f^u(x,\cdot;u) \text{ for } x \in X, \|u\| \leq 1 \}.$$

(i) Assume: (a’) $f(\cdot,\xi)$ is locally Lipschitz continuous for every $\xi$ with $P$-integrable modulus; (b’) $Q \in \mathcal{P}(\Xi)$; (c’) $X$ is a compact set; (d’) $X(P)$ and $X(Q)$ are nonempty. Then we obtain the following estimate for the sets of weak stationary points:

$$\mathbb{D}(\tilde{X}(Q), \tilde{X}(P)) \leq \tilde{R}^{-1}(2\mathcal{D}(Q,P)), $$

where $\mathcal{D}$ denotes the distance between two sets and $\mathbb{D}$ denotes the Hausdorff distance.
where $\tilde{R}$ is the growth function

$$\tilde{R}(\epsilon) := \inf_{x \in X, d(x, \tilde{X}(P)) \geq \epsilon} d(0, E_P(\partial_x f(x, \xi)) + N_X(x))$$

and

$$\mathcal{D}(Q, P) := \sup_{g \in \mathcal{G}} (E_Q[g(\xi)] - E_P[g(\xi)]).$$

(ii) Assume that there exists a nondecreasing continuous function $h$ on $[0, +\infty)$ such that $h(0) = 0$, $\sup\{h(2t)/h(t) : t > 0\} < +\infty$ and

\begin{equation}
\sup_{x \in X} \sup_{\tau \in (0, \delta]} \|u\| \leq 1 \left| \frac{1}{\tau} \left( f(x + \tau u, \xi) - f(x, \xi) \right) - \frac{1}{\tau} \left( f(x + \tau u, \tilde{\xi}) - f(x, \tilde{\xi}) \right) \right| \leq h(||\xi - \tilde{\xi}||)
\end{equation}

holds for all $\xi, \tilde{\xi} \in \Xi$ and for $\delta > 0$ sufficiently small. Then the estimate

$$\mathbb{D}(X(Q), X(P)) \leq R^{-1}(2\sup_{x \in X} \mathcal{D}(\partial F_Q(x), \partial F_P(x)))$$

is valid for the sets of stationary points, where $R$ is the growth function

$$R(\epsilon) := \inf_{x \in X, d(x, X(P)) \geq \epsilon} d(0, E_P[\partial_x f(x, \xi)] + N_X(x))$$

and $\zeta_h$ the Kantorovich–Rubinstein functional

\begin{equation}
\zeta_h(P, Q) = \inf_{\Xi \times \Xi} \int h(||\xi - \tilde{\xi}||) d\eta(\xi, \tilde{\xi}),
\end{equation}

where the infimum is over all finite measures $\eta$ on $\Xi \times \Xi$ with $P_1\eta - P_2\eta = P - Q$ and $P_1\eta$ denoting the $i$th projection of $\eta$.

Proof. Part (i). For the proof we use Theorem 3.1. Therefore it suffices to verify the conditions of the theorem for $\Gamma(x, \xi) = \partial_x f(x, \xi)$ and $G(x) = N_X(x)$. Conditions (a) and (c) of Theorem 3.1 are satisfied under the assumption that $f$ is locally Lipschitz continuous w.r.t. $x$ with $P$-integrably Lipschitz constant and the fact that the Clarke subdifferential $\partial_x f(x, \xi)$ is convex and compact and upper semicontinuous w.r.t. $x$ for every fixed $\xi$. Conditions (d) follows from condition (d') and the fact that $\partial E_P[f(x, \xi)] \subseteq E_P[\partial_x f(x, \xi)]$.

Part (ii). Analogous to the proofs of Theorem 3.1, we can derive

$$\mathbb{D}(X(Q), X(P)) \leq R^{-1}\left(2 \sup_{x \in X} \mathbb{D}(\partial E_Q[f(x, \xi)], \partial E_P[f(x, \xi)])\right).$$

In what follows, we use the notation $F_P(x) := E_P[f(x, \xi)]$ and $F_Q(x) := E_Q[f(x, \xi)]$ and estimate $D_* := \sup_{x \in X} \mathbb{D}(\partial F_Q(x), \partial F_P(x))$. By Hörmander’s theorem and the definition of the Clarke subdifferential,
Here, we used for the first estimate the fact that the inequality

\[
\limsup_{\tau \to 0} \sup_{x' \in X, \tau \leq 0} \frac{1}{\tau} \left( F_P(x' - \tau u) - F_Q(x') \right) \leq \sup_{x \in X, \tau \leq 0} \left| \frac{1}{\tau} \left( F_P(x' + \tau u) - F_P(x') \right) - \frac{1}{\tau} \left( F_P(x' + \tau u) - F_P(x') \right) \right|
\]

holds for any bounded sequences \( \{a_k\} \) and \( \{b_k\} \). For the final estimate we used the duality theorem [28, Theorem 5.3.2], implying

\[
\zeta_h(P, Q) = \sup_{g \in G_h} \left| \int_{\Xi} g(\xi) d(P - Q)(\xi) \right|
\]

where the set \( G_h \) is defined by

\[
G_h = \{ g : \Xi \to \mathbb{R} : |g(\xi) - g(\tilde{\xi})| \leq h(\|\xi - \tilde{\xi}\|) \forall \xi, \tilde{\xi} \in \Xi \}
\]

and the conditions imposed for \( h \) are needed for the validity of the duality theorem. The proof is complete. □

**Remark 4.2.** If the integrand \( f(\cdot, \xi) \) is Clarke regular on \( \mathbb{R}^n \) for every \( \xi \), i.e., in particular, if the integrand is convex, the functions \( g = f^a(x, \cdot, u) \) belong to the class \( G_h \) and, hence, we also obtain the estimate

\[
\mathbb{D}(\tilde{X}(Q), \tilde{X}(P)) \leq R^{-1}(2\zeta_h(Q, P))
\]

as a conclusion of part (i) of the previous theorem.

The Kantorovich–Rubinstein functional \( \zeta_h(P, Q) \) is finite if the probability measures \( P \) and \( Q \) belong to the set

\[
\mathcal{P}_h(\Xi) = \left\{ Q \in \mathcal{P}(\Xi) : \int_{\Xi} h(\|\xi\|) dQ(\xi) < +\infty \right\}
\]

Note that \( \zeta_h \) is a (so-called) simple distance on \( \mathcal{P}_h(\Xi) \) (see [28, section 3.2]), which means that (i) \( P = Q \) if and only if \( \zeta_h(P, Q) = 0 \), (ii) \( \zeta_h(P, Q) = \zeta_h(Q, P) \), and (iii) \( \zeta_h(P, Q) \leq K_h(\zeta_h(P, Q) + \zeta_h(Q, Q)) \) for all \( P, Q, \tilde{Q} \in \mathcal{P}_h(\Xi) \), where \( K_h \) is a positive constant depending on function \( h \). An important special case is \( h(t) = t^p \) with \( p \geq 1 \).

In that case, one may deduce the Wasserstein metric of order \( p \) or \( L_p \)-minimal metric \( \ell_p \) by setting \( \ell_p(P, Q) = (\zeta_h(P, Q))^{\frac{1}{p}} \) with \( \mathcal{P}_h(\Xi) \) being the set of all probability measures having finite \( p \)th order moments.

Alternatively, one might require in (4.5) that the term \( h(\|\xi - \tilde{\xi}\|) \) is replaced by

\[
\max\{1, \|\xi\|^{p-1}, \|\tilde{\xi}\|^{p-1}\}\|\xi - \tilde{\xi}\|
\]
for some \( p \geq 1 \). In that case the distance \( \zeta_h \) is replaced by the \( p \)th order Fortet–Mourier metric \( \zeta_p \) (see [28, section 5.1]) and \( \mathcal{P}_h(\Xi) \) by the set of all probability measures having finite \( p \)th order moments.

In the case when \( f \) is convex w.r.t. \( x \) for almost every \( \xi \), one can show that \( \mathbb{E}_Q[f(x, \xi)] \) converges to \( \mathbb{E}_P[f(x, \xi)] \) uniformly over any compact of \( \mathbb{R}^n \) as \( \mathcal{D}(Q, P) \to 0 \). Attouch’s theorem [35, Theorem 12.35] implies that \( \mathbb{E}_Q[f(\cdot, \xi)] \) converges graphically to \( \partial \mathbb{E}_P[f(\cdot, \xi)] \). However, the graphical convergence does not quantify the rate of convergence, while Theorem 4.1 does.

Note that Liu, Xu, and Lin [23] also investigated the stability of problem (4.1) by looking into the impact on stationary points when \( P \) is approximated through a sequence of probability measures. Theorem 4.1 strengthens [23, Theorem 5.3] by quantifying the rate of the approximation/convergence of the stationary points.

Note also that in the case when \( f(x, \xi) \) is continuously differentiable w.r.t. \( x \) for every \( \xi \), the first order optimality condition (4.2) coincides with (4.3). In that case, it is possible to explore metric regularity of \( \mathbb{E}_P[\nabla_x f(x, \xi)] + \mathcal{N}_X(x) \); see Corollary 3.3 and the following remarks. Subsequently, we may obtain a linear bound for the inverse of the growth functions and hence establish a linear bound for \( D(X(Q), X(P)) \).

### 4.2. Two-stage linear recourse problems

In what follows, we consider a linear two-stage recourse minimization problem:

\[
\begin{align*}
\min_{x \in \mathbb{R}^n} & \quad c^\top x + \mathbb{E}_P[v(x, \xi)] \\
\text{s.t.} & \quad Ax = b, x \geq 0,
\end{align*}
\]

(4.7)

where \( v(x, \xi) \) is the optimal value function of the second-stage problem

\[
\begin{align*}
\min_{y \in \mathbb{R}^m} & \quad q(\xi)^\top y \\
\text{s.t.} & \quad T(\xi)x + Wy = h(\xi), \ y \geq 0,
\end{align*}
\]

(4.8)

where \( W \in \mathbb{R}^{r \times m} \) is a fixed recourse matrix, \( T(\xi) \in \mathbb{R}^{r \times n} \) is a random matrix, and \( h(\xi) \in \mathbb{R}^r \) and \( q(\xi) \in \mathbb{R}^m \) are random vectors. We assume that \( T(\cdot), h(\cdot), \) and \( q(\cdot) \) are affine functions of \( \xi \) and that \( \Xi \) is a polyhedral subset of \( \mathbb{R}^n \) (for example, \( \Xi = \mathbb{R}^n \)). If we consider the set \( X = \{ x \in \mathbb{R}^n : Ax = b, x \geq 0 \} \) and define the integrand \( f \) by

\[
f(x, \xi) = c^\top x + v(x, \xi)
\]

the linear two-stage model (4.7) is of the form of problem (4.1). Let

\[
\phi_P(x) = \mathbb{E}_P[v(x, \xi)].
\]

By [43, Theorem 4.7], the domain of \( \phi_P \) is a convex polyhedral subset of \( \mathbb{R}^n \) and it holds that

\[
\text{dom } \phi_P = \{ x \in \mathbb{R}^n : h(\xi) - T(\xi)x \in \text{pos } W \ \forall \xi \in \Xi \},
\]

where “\( \text{pos } W \)” denotes the positive hull of the matrix \( W \), that is, \( \text{pos } W := \{ Wy : y \geq 0 \} \). Next, we recall some properties of \( v \).

**Lemma 4.3.** Let \( \mathcal{M}(q(\xi)) := \{ \pi \in \mathbb{R}^r : W^\top \pi \leq q(\xi) \} \) be nonempty for every \( \xi \in \Xi \). Then there exists a constant \( L > 0 \) such that \( v \) satisfies the local Lipschitz continuity property

\[
|v(x, \xi) - v(\hat{x}, \hat{\xi})| \leq L(\max\{1, \|\xi\|, \|\hat{\xi}\|\}^2\|\hat{x} - x\| + \max\{1, \|x\|, \|\hat{x}\|\} \max\{1, \|\xi\|, \|\hat{\xi}\|\} \|\hat{\xi} - \xi\|)
\]

(4.9)
for all pairs \((x, \xi), (\hat{x}, \hat{\xi}) \in (X \cap \text{dom } \phi_P) \times \Xi\) and some constant \(\hat{L}\). Moreover, \(v(\cdot, \xi)\) is convex for every \(\xi \in \Xi\).

Proof. \(v(x, \xi)\) is the optimal value of the linear program

\[
(4.10) \quad \min \{b^T y : W y = a, y \geq 0\},
\]

where \(a = a(x, \xi) = h(\xi) - T(\xi)x\) and \(b = b(\xi) = q(\xi)\). Let \(\text{val}(a, b)\) denote the optimal value of (4.10). It is known from \([42, 25]\) that the domain of \(\text{val}\) is a polyhedral cone in \(\mathbb{R}^m \times \mathbb{R}^r\) and there exist finitely many matrices \(C_j\) and polyhedral cones \(K_j, j = 1, \ldots, \ell\), such that \(v\) and its domain allow the representation

\[
\text{dom}(\text{val}) = \bigcup_{j=1}^{\ell} K_j \quad \text{and} \quad \text{val}(a, b) = (C_j a)^T b \quad \text{if } (a, b) \in K_j.
\]

Furthermore, it holds that \(\text{int } K_j \neq \emptyset\) and \(K_j \cap K_i = \emptyset, \ i \neq j, i, j = 1, \ldots, \ell\). Hence, \(\text{val}\) satisfies the following continuity property on its domain:

\[
|\text{val}(a, b) - \text{val}(\bar{a}, \bar{b})| \leq L (\max\{1, \|b\|, \|\bar{b}\|\} \|a - \bar{a}\| + \max\{1, \|a\|, \|\bar{a}\|\} \|b - \bar{b}\|)
\]

with some constant \(L > 0\). Moreover, \(\text{val}(\cdot, b)\) is convex for each \(b\). Hence, the mapping \(x \rightarrow v(x, \xi) = \text{val}(h(\xi) - T(\xi)x, q(\xi))\) is convex for each \(\xi \in \Xi\). Furthermore, we obtain

\[
|v(x, \xi) - v(\hat{x}, \hat{\xi})| \leq |v(x, \xi) - v(\hat{x}, \xi)| + |v(\hat{x}, \xi) - v(\hat{x}, \hat{\xi})| \\
\leq |\text{val}(h(\xi) - T(\xi)x, q(\xi)) - \text{val}(h(\xi) - T(\xi)\hat{x}, q(\xi))| \\
+ |\text{val}(h(\hat{\xi}) - T(\hat{\xi})\hat{x}, q(\xi)) - \text{val}(h(\hat{\xi}) - T(\hat{\xi})\hat{x}, q(\hat{\xi}))| \\
\leq L (\max\{1, \|q(\xi)\|, \|q(\hat{\xi})\|\} \|T(\xi)(x - \hat{x})\| \\
+ \max\{1, \|h(\xi) - T(\xi)x\|, \|h(\hat{\xi}) - T(\hat{\xi})\hat{x}\|\} \|q(\xi) - q(\hat{\xi})\|).
\]

Using that \(h, q,\) and \(T\) are affine functions of \(\xi\) then leads to the desired estimate (4.9). \(\square\)

For each \(x \in \text{dom } \phi_P\) it follows from [38, Proposition 2.8] that

\[
(4.11) \quad \partial \phi_P(x) = -E_P[T(\xi)^T D(x, \xi)] + N_{\text{dom } \phi_P}(x),
\]

where \(\partial\) denotes the usual convex subdifferential [34] and \(D(x, \xi)\) the solution set of the dual to (4.8), that is,

\[
D(x, \xi) := \arg \max_{\xi \in \mathcal{M}(q(\xi))} \xi^T (h(\xi) - T(\xi)x).
\]

The proposition below states an existence result and the first order optimality condition for the two-stage minimization problem (4.7).

**Proposition 4.4.** Assume that \(X \cap \text{dom } \phi_P\) is nonempty and bounded, \(\mathcal{M}(q(\xi))\) is nonempty for each \(\xi \in \Xi\) and \(P\) has finite second order moments, i.e., \(E_P[\|\xi\|^2] < \infty\). Then there exists a minimizer \(x^* \in X \cap \text{dom } \phi_P\) of (4.7). Furthermore, \(x^* \in X\) is a minimizer of (4.7) if and only if it satisfies the generalized equations

\[
(4.12) \quad 0 \in E_P[c - T(\xi)^T D(x, \xi)] + N_{X \cap \text{dom } \phi_P}(x).
\]

Here, \(N_{X \cap \text{dom } \phi_P}(x)\) denotes the normal cone to the polyhedral set \(X \cap \text{dom } \phi_P\).
Proof. Lemma 4.3 implies that $E_P[v(x, \xi)]$ is finite for every $x \in X \cap \text{dom } \phi_P$. Hence, the existence follows from the Weierstrass theorem and the first order optimality condition from [35, Theorem 8.15].

The polyhedral set $\text{dom } \phi_P$ may contain some induced constraints. If one assumes relatively complete recourse, i.e., $X \subset \text{dom } \phi_P$, the optimality condition (4.12) coincides with the one in [38, Theorem 2.11]. Our interest here is to apply the stability results of SGEs in section 3 to (4.12) when the probability measure $P$ is perturbed. To this end, we look at properties of the set-valued mapping $\Gamma$ given by

$$\Gamma(x, \xi) := c - T(\xi)^T D(x, \xi) = c - T(\xi)^T \arg \max_{W^T \xi \leq q(\xi)} \zeta^T (h(\xi) - T(\xi)x).$$

**Proposition 4.5.** Let $D(x, \xi)$ be defined as above and assume that $M(q(\xi))$ is nonempty and bounded for every $\xi \in \Xi$. Then $\Gamma$ is locally upper Lipschitz continuous at any $(x, \xi)$ in $(X \cap \text{dom } \phi_P) \times \Xi$ and there exists $L > 0$ such that

$$\Gamma(\tilde{x}, \tilde{\xi}) \subseteq \Gamma(x, \xi) + \hat{L}(\max\{1, \|\xi\|, \|\tilde{\xi}\|\})^2 \|\tilde{x} - x\|
+ \max\{1, \|x\|, \|\tilde{x}\|\} \max\{1, \|\xi\|, \|\tilde{\xi}\|\} \|\tilde{\xi} - \xi\||B$$

for all pairs $(x, \xi), (\tilde{x}, \tilde{\xi}) \in (X \cap \text{dom } \phi_P) \times \Xi$. Here, $B$ denotes the unit ball in $\mathbb{R}^n$.

Proof. Let $S(a, b)$ denote the dual solution set of (4.10). Since the objective function of the dual has linear growth, the upper semicontinuity behavior of the solution set $S$ is very similar to that of $v$ (see (4.9)), namely,

$$S(\tilde{a}, \tilde{b}) \subseteq S(a, b) + L_1(\max\{1, \|b\|, \|\tilde{b}\|\}) \|a - \tilde{a}\| + \max\{1, \|a\|, \|\tilde{a}\|\} \|b - \tilde{b}\||B$$

for some constant $L_1 > 0$ and all pairs $(a, b), (\tilde{a}, \tilde{b}) \in \text{dom}(v)$. Since it holds that $D(x, \xi) = S(h(\xi) - T(\xi)x, q(\xi))$ and $h$, $q$, and $T$ are affine functions of $\xi$, $D$ is locally upper Lipschitz continuous at any pair $(x, \xi) \in X \cap \text{dom } \phi_P \times \Xi$ and it holds that

$$D(\tilde{x}, \tilde{\xi}) \subseteq D(x, \xi) + \hat{L}(\max\{1, \|\xi\|, \|\tilde{\xi}\|\})^2 \|\tilde{x} - x\|
+ \max\{1, \|x\|, \|\tilde{x}\|\} \max\{1, \|\xi\|, \|\tilde{\xi}\|\} \|\tilde{\xi} - \xi\||B.$$

The result follows in a straightforward way from the local upper Lipschitz property of $D$.

We are ready to state our quantitative stability result for the solution set $S(P)$ of (4.7) if the probability distribution $P$ is perturbed by another probability distribution $Q$.

**Theorem 4.6.** Assume that (a) relatively complete recourse is satisfied, (b) $M(q(\xi)) = \{\pi : W^T \pi \leq q(\xi)\}$ is nonempty and bounded for every $\xi \in \Xi$, (c) $P$ has finite second order moments, i.e., $E_P[\|\xi\|^2] < +\infty$, and (d) $X$ is nonempty and bounded. Then it holds for any probability measure $Q$ such that $\mathcal{D}(Q, P)$ is sufficiently small

$$\mathcal{D}(S(Q), S(P)) \leq R^{-1}(2\mathcal{D}(Q, P)),$$

where the function $R$ is defined by

$$R(\epsilon) := \inf_{x \in X, d(x, S(P)) \geq \epsilon} d(0, E_P[\Gamma(x, \xi)] + N_X(x)),$$

and the distance $\mathcal{D}$ is defined in section 3.
Proof. We intend to apply Theorem 3.1 to the SGE

\[ 0 \in E[\Gamma(x, \xi)] + N_X(x) \]

and check the corresponding assumptions. The set-valued mapping \( \Gamma \) takes convex polyhedral and compact values according to condition (b) and is upper semi-continuous with respect to \( x \) for every fixed \( \xi \in \Xi \) according to Proposition 4.6. The set \( D(x, \xi) \) is contained in \( \mathcal{M}(q(\xi)) \); thus, it suffices to show that

\[
\kappa(\xi) = ||c|| + ||T(\xi)\top||\mu(\xi), \quad \text{where} \quad ||\pi|| \leq \mu(\xi) \quad \forall \pi \in \mathcal{M}(q(\xi)),
\]

is \( P \)-integrable. The set-valued mapping \( \mathcal{M} \) assigning to each \( q \in \mathbb{R}^m \) the set \( \mathcal{M}(q) \) has a closed polyhedral graph and hence is Hausdorff Lipschitz continuous on its domain (say, with modulus \( L_M \)). Let \( \xi \) be fixed in \( \Xi \) and \( \xi \in \Xi \) be arbitrary. Then we have for any \( \pi \in \mathcal{M}(q(\xi)) \)

\[
d(\pi, \mathcal{M}(q(\hat{\xi}))) \leq L_M||\xi - \hat{\xi}||.
\]

Hence, there exists \( \bar{\pi} \in \mathcal{M}(q(\hat{\xi})) \) such that \( ||\pi|| \leq ||\bar{\pi}|| + L_M||\xi - \hat{\xi}||. \) Since \( \mathcal{M}(q(\xi)) \) is bounded, there exists a constant \( C \) such that we may choose the function \( \mu \) as \( \mu(\xi) = C + L_M(||\xi|| + ||\xi||) \). We conclude that the function \( \kappa \) given by (4.13) depends on \( ||\xi|| \) at most quadratically. Hence, \( \kappa \) is \( P \)-integrable according to assumption (c). Finally, we note that the normal cone mapping \( N_X \) is upper semi-continuous and \( S(Q) \) is always nonempty due to the compactness of \( X \) and the fact the \( S(Q) \) is the solution set of the minimization problem (4.8) with continuous objective function.

In order to compare the previous novel stability result for two-stage models with earlier ones, it is of interest to characterize the distance \( \mathcal{D} \) and the function \( R_P \) in this particular case. While the function \( R_P \) depends intrinsically of the probability measure \( P \), we may provide more insight of the distance \( \mathcal{D} \).

**Proposition 4.7.** Let the assumptions of the previous theorem be satisfied. Then the function class \( \mathcal{F} \) defined by (3.1) is contained in the function class

\[
\mathcal{F} = \{ g : \Xi \to \mathbb{R} : g(x) - g(\hat{\xi}) \leq C \max\{1, ||\xi||, ||\hat{\xi}||\}^2||\xi - \hat{\xi}|| \quad \forall \xi, \hat{\xi} \in \Xi \}
\]

for some constant \( C > 0 \). Consequently, the estimate

\[
\mathcal{D}(P, Q) \leq C\zeta_3(P, Q)
\]

holds, where \( \zeta_3 \) denotes the third order Fortet–Mourier metric (see Remark 4.2).

**Proof.** Let \( u \in \mathbb{R}^n \) with \( ||u|| \leq 1, x \in X \) and \( \xi, \hat{\xi} \in \Xi \). We consider \( g(\xi) = \sigma(\Gamma(x, \xi), u) \) and know from Proposition 4.5 that

\[
g(\xi) - g(\hat{\xi}) = \sigma(\Gamma(x, \xi), u) - \sigma(\Gamma(x, \hat{\xi}), u) \leq D(\Gamma(x, \xi), \Gamma(x, \hat{\xi})) \leq L \max\{1, ||x||, ||\hat{x}||\} \max\{1, ||\xi||, ||\hat{\xi}||\}^2||\xi - \hat{\xi}||.
\]

Since \( X \) is bounded, we may choose the constant \( C \) such that \( L \max\{1, ||x||, ||\hat{x}||\} \leq C \) for all \( x \in X \). Since \( \zeta_3 \) is slightly stronger than the second order Fortet–Mourier metric \( \zeta_2 \), which appears in the stability analysis for two-stage models in [36], Theorem 4.6 is slightly weaker than earlier ones.
4.3. Two-stage stochastic mathematical problem with equilibrium constraints. In this subsection, we consider application of the stability analysis established in section 3 to a two-stage stochastic mathematical program with complementarity constraints (MPCC) defined as follows:

\[
\begin{align*}
\min_{x, y(\cdot) \in \mathcal{Y}} & \quad \mathbb{E}_P[f(x, y(\omega), \xi(\omega))] \\
\text{s.t.} & \quad x \in X \text{ and for almost every } \omega \in \Omega:
\begin{align*}
g(x, y(\omega), \xi(\omega)) & \leq 0, \\
h(x, y(\omega), \xi(\omega)) & = 0, \\
0 \leq G(x, y(\omega), \xi(\omega)) & \perp H(x, y(\omega), \xi(\omega)) \geq 0,
\end{align*}
\end{align*}
\]

(4.14)

where \(X\) is a nonempty closed convex subset of \(\mathbb{R}^n\), \(f, g, h, G, H\) are continuously differentiable functions from \(\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q\) to \(\mathbb{R}, \mathbb{R}^r, \mathbb{R}^s, \mathbb{R}^m, \mathbb{R}^m\), respectively, \(\xi : \Omega \to \Xi\) is a vector of random variables defined on probability \((\Omega, \mathcal{F}, P)\) with compact support set \(\Xi \subset \mathbb{R}^q\), \(\mathbb{E}_P[\cdot]\) denotes the expected value with respect to probability measure \(P\), \(\perp\) denotes the perpendicularity of two vectors, and \(\mathcal{Y}\) is a space of functions \(y(\cdot) : \Omega \to \mathbb{R}^m\) such that \(\mathbb{E}_P[f(x, y(\omega), \xi(\omega))]\) is well defined. Stability analysis of problem (4.14) has been discussed in [23] through nonlinear programming regularization. Our interest here is in a direct stability analysis on the stationary point of the problem using the SGE scheme discussed in section 3.

Observe first that problem (4.14) can be written as

\[
\begin{align*}
P_\vartheta : \quad & \min_x \vartheta(x) = \mathbb{E}_P[v(x, \xi(\omega))] \\
\text{s.t.} & \quad x \in X,
\end{align*}
\]

as long as \(\mathbb{E}_P[(v(x, \xi))_+] < \infty\) and \(\mathbb{E}_P[(-v(x, \xi))_+] < \infty\), where \((a)_+ = \max(0, a)\) and \(v(x, \xi)\) denotes the optimal value function of the following second-stage problem:

\[
\text{MPCC}(x, \xi) : \quad \min_y \quad f(x, y, \xi) \\
\text{s.t.} \quad g(x, y, \xi) \leq 0, \\
h(x, y, \xi) = 0, \\
0 \leq G(x, y, \xi) \perp H(x, y, \xi) \geq 0.
\]

The reformulation is well known in stochastic programming; see, for example, [37, Proposition 5, Chapter 1] and a discussion in [39, section 2] in the context of two-stage stochastic mathematical problems with equilibrium constraints (MPECs).

Define the Lagrangian function of the second-stage problem MPCC\((x, \xi)\):

\[
\mathcal{L}(x, y, \xi; \alpha, \beta, u, v) := f(x, y, \xi) + g(x, y, \xi)^\top \alpha + h(x, y, \xi)^\top \beta \\
- G(x, y, \xi)^\top u - H(x, y, \xi)^\top v.
\]

We consider the following KKT conditions of MPCC\((x, \xi)\):

\[
\begin{align*}
0 & = \nabla_y \mathcal{L}(x, y, \xi; \alpha, \beta, u, v), \\
y & \in \mathcal{F}(x, \xi), \\
0 & \leq \alpha \perp -g(x, y, \xi) \geq 0, \\
0 & = u_i, \quad i \notin \mathcal{I}_G(x, y, \xi), \\
0 & = v_i, \quad i \notin \mathcal{I}_H(x, y, \xi), \\
0 & \leq u_i v_i, \quad i \in \mathcal{I}_G(x, y, \xi) \cap \mathcal{I}_H(x, y, \xi),
\end{align*}
\]
where $\mathcal{F}(x, \xi)$ denotes the feasible set of MPCC($x, \xi$) and
\[
\mathcal{I}_G(x, y, \xi) := \{ i \mid G_i(x, y, \xi) = 0, \ i = 1, \ldots, m \},
\]
\[
\mathcal{I}_H(x, y, \xi) := \{ i \mid H_i(x, y, \xi) = 0, \ i = 1, \ldots, m \}.
\]
Let $W(x, \xi)$ denote the set of KKT pairs $(y; \alpha, \beta, u, v)$ satisfying the above conditions for given $(x, \xi)$ and $S(x, \xi)$ the corresponding set of stationary points, that is, $S(x, \xi) = \Pi_y W(x, \xi)$. For each $(y; \alpha, \beta, u, v)$, $y$ is a $C$-stationary point of problem MPCC($x, \xi$) and $(\alpha, \beta, u, v)$ are the corresponding Lagrange multipliers. When the stationary points are restricted to global minimizers, we denote the set of KKT pairs by $W^*(x, \xi)$, i.e., $W^*(x, \xi) = \{(y; \alpha, \beta, u, v) \in W(x, \xi), y \in Y_{\text{sol}}(x, \xi)\}$, where $Y_{\text{sol}}(x, \xi)$ denotes the set of optimal solutions of MPCC($x, \xi$).

Let $(x^*, \xi)$ be fixed. Recall that MPCC($x^*, \xi$) is said to satisfy the MPEC-Mangasarian–Fromovitz constraint qualification (MPEC-MFCQ) at a feasible point $y^*$ if the gradient vectors
\[
\{\nabla g_i(x^*, y^*, \xi)\}_{i=1,\ldots,r} \cup \{\nabla H_i(x^*, y^*, \xi)\}_{i=\mathcal{I}_G(x^*, \xi)} \cup \{\nabla H_i(x^*, y^*, \xi)\}_{i=\mathcal{I}_H(x^*, \xi)}
\]
are linearly independent and there exists a vector $d \in \mathbb{R}^n$ perpendicular to the vectors such that
\[
\nabla g_i(x^*, y^*, \xi)^\top d < 0 \quad \forall i \in \mathcal{I}_g(x^*, y^*, \xi),
\]
where
\[
\mathcal{I}_g(x^*, y^*, \xi) := \{ i \mid g_i(x^*, y^*, \xi) = 0, \ i = 1, \ldots, s \}.
\]

The following results are derived in [23].

**Proposition 4.8.** Let $x^* \in X$. Suppose that there exist constants $\delta$, $t^* > 0$, a compact set $Y \subset \mathbb{R}^m$, and a neighborhood $U$ of $x^*$ such that
\[
0 \neq \{ y : f(x, y, \xi) \leq \delta \ \text{and} \ y \in \mathcal{F}(x, \xi) \} \subset Y
\]
for all $(x, \xi) \in U \times \Xi$. Suppose also that problem MPCC($x^*, \xi$) satisfies MPEC-MFCQ at every point $y$ in the solution set of MPCC($x^*, \xi$), denoted by $Y_{\text{sol}}(x^*, \xi)$. Then there exists a neighborhood $U$ of $x^*$ such that

(i) $v(\cdot, \xi)$ is locally Lipschitz continuous on $U$;

(ii) for any $x \in U$ and $\xi \in \Xi,
\[
\partial_x v(x, \xi) \subseteq \Phi(x, \xi),
\]
and $\Phi(\cdot, \cdot)$ is upper semicontinuous on $U \times \Xi$, where
\[
\Phi(x, \xi) := \text{conv} \left\{ \bigcup_{(y; \alpha, \beta, u, v) \in W^*(x, \xi)} \nabla_x \mathcal{L}(x, y, \xi; \alpha, \beta, u, v) \right\}.
\]

Using $\partial_x v(x, \xi)$ and $\Phi(x, \xi)$, we can define the weak KKT conditions of problem (4.14)
\[
(4.15) \quad 0 \in \mathbb{E}_P[\partial_x v(x, \xi)] + \mathcal{N}_X(x)
\]
and its relaxation
\[
(4.16) \quad 0 \in \mathbb{E}_P[\Phi(x, \xi)] + \mathcal{N}_X(x).
\]
Both systems are SGEs. If the probability measure $P$ is perturbed by another probability measure $Q$, the weak KKT conditions of problem (4.14) and its relaxation should be

$$0 \in \mathbb{E}_Q[\partial_x v(x,\xi)] + \mathcal{N}_X(x)$$

and

$$0 \in \mathbb{E}_Q[\Phi(x,\xi)] + \mathcal{N}_X(x),$$

respectively.

**Theorem 4.9.** Let $v^o(x,\xi;u)$ denote the Clarke generalized directional derivative of $v(x,\xi)$ and for a given nonzero vector $u$

$$\mathcal{F} := \{g : g(\cdot) := v^o(x,\cdot;u) \text{ for } x \in X, \|u\| \leq 1\}.$$  

Let $\hat{X}(Q)$ and $\hat{X}(P)$ denote the set of solutions of (4.17) and (4.15), respectively. Then

$$D(\hat{X}(Q), \hat{X}(P)) \leq \hat{R}^{-1}(2\mathcal{D}(Q,P)),$$

where $\hat{R}$ is the growth function

$$\hat{R}(\epsilon) := \inf_{x \in X, d(x,\hat{X}(P)) \geq \epsilon} d(0, \mathbb{E}_P[\partial_x v(x,\xi)] + \mathcal{N}_X(x))$$

and

$$\mathcal{D}(Q,P) := \sup_{g \in \mathcal{F}} (\mathbb{E}_Q[g(\xi)] - \mathbb{E}_P[g(\xi)]).$$

**Remark 4.10.** The key condition in the conclusion (i) of Theorem 4.1 is the Lipschitz continuity of $v(x,\xi)$ which follows from Proposition 4.8. It is possible to derive a conclusion similar to Theorem 4.1(ii). To see this, it suffices to verify the existence of a nondecreasing continuous function $h$. To ease the technical details, let us consider a special case of MPCC$(x,\xi)$:

$$\text{MPCC}'(x,\xi) : \min_y f(x,y,\xi)$$

$$\text{subject to } y \perp H(x,y,\xi) \geq 0,$$

where $H$ is uniformly strongly monotone w.r.t. $y$, that is, there exists a positive constant $C_1 > 0$ such that $\|\nabla_y H(x,y,\xi)\| \leq C_1$ for all $x, y, \xi$. By [44, Theorem 2.3], the complementarity inequality constraint defines a unique feasible solution $y(x,\xi)$ which is piecewise smooth provided $H$ is smooth w.r.t. $x$ and $\xi$. Moreover, if $H$ is uniformly globally Lipschitz continuous w.r.t. $\xi$, then $y(x,\xi)$ is also uniformly globally Lipschitz continuous w.r.t. $\xi$. Assuming that $f(x,y,\xi)$ is continuously differentiable and uniformly globally Lipschitz continuous w.r.t. $y$ and $\xi$, then

$$v(x,\xi) = f(x,y(x,\xi),\xi)$$

is also piecewise continuously differentiable and uniformly globally Lipschitz continuous w.r.t. $\xi$. Denote the Lipschitz modulus of $y(x,\cdot)$ and $f(x,\cdot,\cdot)$ by $L_1$ and $L_2$, respectively. Then

$$|v(x,\xi) - v(x,\xi')| = |f(x,y(x,\xi),\xi) - f(x,y(x,\xi'),\xi')|$$

$$\leq L_2(\|y(x,\xi) - y(x,\xi')\| + \|\xi - \xi'\|)$$

$$\leq L_2(L_1 + 1)\|\xi - \xi'\|. $$
Let \( L := 2L_2(L_1 + 1) \) and \( h(t) := Lt \). Then

\[
\sup_x \sup_{\tau \in (0,\delta), \|u\| \leq 1} \frac{1}{\tau} (v(x + \tau u, \xi) - v(x, \xi)) - \frac{1}{\tau} (v(x + \tau u, \hat{\xi}) - v(x, \hat{\xi})) \leq h(\|\xi - \hat{\xi}\|),
\]

which means (4.5).

5. **Stochastic semi-infinite programming.** In this section, we discuss the application of our perturbation theory developed in section 3 to a class of nonsmooth stochastic semi-infinite programming problems defined as follows:

\[
\begin{align*}
\min_x & \quad \mathbb{E}_P[f(x, \xi)] \\
\text{s.t.} & \quad \mathbb{E}_P[(\eta - G(x, \xi))_+] \leq \mathbb{E}_P[(\eta - Y(\xi))_+] \quad \forall \eta \in [a, b], \\
& \quad x \in X,
\end{align*}
\]

where \( X \) is a closed convex subset in \( \mathbb{R}^n \), \((z)_+\) denotes \( \max\{z, 0\}\) for a real number \( z \), \( f, G : \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R} \) are continuously differentiable functions, \( Y : \mathbb{R}^q \to \mathbb{R} \) is a continuous function, \( \xi : \Omega \to \Xi \) is a vector of random variables defined on probability space \((\Omega, \mathcal{F}, P)\) with support set \( \Xi \subset \mathbb{R}^q \), \( \mathbb{E}_P[\cdot] \) denoting the expected value with respect to probability measure \( P \), and \([a, b]\) is a closed interval in \( \mathbb{R} \).

Problem (5.1) is a key intermediate formulation in the subject of stochastic programs with second order dominance constraints. For the detailed discussions of the latter, see [11, 12, 13] and the references therein. Liu and Xu [22] studied stability of optimal value and optimal solutions of (5.1) through exact penalization. They also investigated approximation of stationary points of the penalized problem when the latter is approximated by an empirical probability measure (Monte Carlo sampling). However, there is a gap between the stationary point of (5.1) and its penalized problem: a stationary point of the latter is not necessarily that of the former.

Our focus here is to carry out stability analysis of the stationary point of (5.1) directly rather than through its penalized problem. Moreover, we consider a general probability measure approximation to \( P \) rather than restricted to an empirical probability measure approximation. Specifically if the probability measure \( Q \) is a perturbation of \( P \), we would like to analyze the approximation of the stationary points of the perturbed problem

\[
\begin{align*}
\min_x & \quad \mathbb{E}_Q[f(x, \xi)] \\
\text{s.t.} & \quad \mathbb{E}_Q[(\eta - G(x, \xi))_+] \leq \mathbb{E}_Q[(\eta - Y(\xi))_+] \quad \forall \eta \in [a, b], \\
& \quad x \in X,
\end{align*}
\]

as \( Q \) tends to \( P \). To this end, we need to consider the first order optimality conditions of the problems.

For simplicity of notation, let

\[
H(x, \eta, \xi) := (\eta - G(x, \xi))_+ - (\eta - Y(\xi))_+.
\]

It is easy to observe that (a) \( H(x, \eta, \xi) \) is globally Lipschitz continuous in \( \eta \) uniformly w.r.t. \( x \) and \( \xi \), and (b) \( H(x, \eta, \xi) \) is Lipschitz continuous w.r.t. \( x \) if \( G(x, \xi) \) is so and they have the same Lipschitz modulus.

Recall that the Bouligand tangent cone to a set \( X \subset \mathbb{R}^n \) at a point \( x \in X \) is defined as follows:

\[
\mathcal{T}_X(x) := \{ h \in \mathbb{R}^n : d(x + th, X) = o(t), t \geq 0 \}.
\]
The normal cone to $X$ at $x$, denoted by $\mathcal{N}_X(x)$, is defined as the polar of the tangent cone:

$$\mathcal{N}_X(x) := \{ h \in \mathbb{R}^n : \zeta^T h \leq 0, \forall \zeta \in T_X(x) \}$$

and $\mathcal{N}_X(x) = \emptyset$ if $x \notin X$.

**Definition 5.1.** Problem (5.2) is said to satisfy the differential constraint qualification at a point $x_0 \in X$ if there exist a feasible point $x_*$ and a constant $\delta > 0$ such that

$$\sum_{\zeta \in \partial_z \mathbb{E}_P[H(x, \eta, \xi)]} \zeta^T(x_* - x_0) \leq -\delta$$

for all $\eta \in \mathcal{I}(x_0)$, where $\mathcal{I}(x_0) := \{ \eta : \mathbb{E}_P[H(x, \eta, \xi)] = 0, \ \eta \in [a, b] \}$ and $\partial_z \mathbb{E}_P[H(x, \eta, \xi)]$ denotes the Clarke subdifferential of $\mathbb{E}_P[H(x, \eta, \xi)]$ w.r.t. $x$.

The constraint qualification was introduced by Dentcheva and Ruszczyński in [13]. Under the condition, they derived the following first order optimality conditions of (5.1) in terms of Clarke subdifferentials.

Let $x^* \in X$ be a local optimal solution of the true problem (5.1) and assume that the differential constraint qualification is satisfied at $x^*$. Then there exists $\mu^* \in \mathcal{M}_+(\mathbb{R}_+ \times [a, b])$ such that $(x^*, \mu^*)$ satisfies the following:

$$\begin{cases}
0 \in \nabla \mathbb{E}_P[f(x, \xi)] + \int_a^b \mathbb{E}_P[\partial_z H(x, \eta, \xi)] d\eta + \mathcal{N}_X(x), \\
\mathbb{E}_P[H(x, \eta, \xi)] \leq 0 \ \forall \eta \in [a, b], \\
\int_a^b \mathbb{E}_P[H(x, \eta, \xi)] d\eta = 0,
\end{cases}$$

(5.3)

where $\mathcal{M}_+(\mathbb{R}_+ \times [a, b])$ is the set of positive measures in the space of regular countably additive measures on $[a, b]$ having finite variation (see [4, Example 2.63] and the references in [11]); $\partial_z H(x, \eta, \xi)$ denotes the Clarke subdifferential of $H$ w.r.t. $x$, that is,

$$\partial_z H(x, \eta, \xi) := \begin{cases}
\{0\} & \text{if } G(x, \xi) > \eta, \\
\text{conv}\{0, -\nabla_z G(x, \xi)\} & \text{if } G(x, \xi) = \eta, \\
\{-\nabla_z G(x, \xi)\} & \text{if } G(x, \xi) < \eta,
\end{cases}$$

and the mathematical expectation/integeral of the subdifferential (w.r.t. the distribution of $\xi$ and $\mu(\cdot)$) is taken in the sense of Aumann. We call a tuple $(x^*, \mu^*)$ a KKT pair of problem (5.1), $x^*$ a Clarke stationary point, and $\mu^*$ the corresponding Lagrange multiplier.

Under similar conditions, we can derive the first order optimality conditions of the perturbed problem (5.2) as follows:

$$\begin{cases}
0 \in \nabla \mathbb{E}_Q[f(x, \xi)] + \int_a^b \mathbb{E}_Q[\partial_z H(x, \eta, \xi)] d\eta + \mathcal{N}_X(x), \\
\mathbb{E}_Q[H(x, \eta, \xi)] \leq 0 \ \forall \eta \in [a, b], \\
\int_a^b \mathbb{E}_Q[H(x, \eta, \xi)] d\eta = 0,
\end{cases}$$

(5.4)

Our aim in this section is to investigate the approximation of the stationary points defined by (5.4) to those of (5.3) as $Q$ approximates $P$. To this end, we reformulate the optimality conditions as a system of SGEs so that we can apply Theorem 3.1. Since $G(x, \xi)$ is Lipschitz continuous in $(x, \xi)$ and the modulus in $x$ is bounded by a positive constant $L_1$, $H(x, \eta, \xi)$ is Lipschitz continuous in $(x, \eta, \xi)$. Then by [45, Proposition 2.1], $\partial_z H(x, \eta, \xi)$ is measurable with respect to $\eta, \xi$. Moreover $\partial_z H(x, \eta, \xi)$ is bounded by $L_1$. By invoking Proposition 2.6, we have
\[
\int_a^b E_P[\partial_x H(x, \eta, \xi)] \mu(d\eta) = E_P \left[ \int_a^b \partial_x H(x, \eta, \xi) \mu(d\eta) \right].
\]

Let
\[
\Gamma(x, \mu, \xi) := \begin{pmatrix}
\nabla_x f(x, \xi) + \int_a^b \partial_x H(x, \eta, \xi) \mu(d\eta) \\
\int_a^b H(x, \eta, \xi) \mu(d\eta)
\end{pmatrix}
\]
and
\begin{align}
G(x, \mu) := \begin{pmatrix}
\mathcal{N}_X(x)
\\
\mathcal{C}_+(\mathcal{I}, [a, b])
\\
0
\end{pmatrix}.
\end{align}

To simplify the notation, let \( z := (x, \mu) \). Then we can reformulate the KKT conditions (5.3) as the following SGE:
\begin{align}
(5.6) \quad 0 &\in E_P[\Gamma(z, \xi)] + G(z),
\end{align}
where the norm in space \( \mathcal{C}([a, b]) \) is \( \| \cdot \|_\infty \). Obviously (5.6) falls into the framework of the SGE (1.1). Likewise, we can reformulate the KKT conditions (5.4) as the SGE
\begin{align}
(5.7) \quad 0 &\in E_Q[\Gamma(z, \xi)] + G(z).
\end{align}

In what follows, we investigate the approximation of the set of solutions of (5.7) to that of (5.6) as \( Q \rightarrow P \).

We need to introduce some new notation. Let \( Z \) denote a compact subset of \( X \times \mathcal{M}_+(\mathcal{I}, [a, b]) \),
\[
\mathcal{F} := \{ g(\xi) : g(\xi) := \sigma(\Gamma(z, \xi), u) \text{ for } z \in Z, \| u \| \leq 1 \}.
\]

Let
\[
\mathcal{D}_S(Q, P) := \sup_{g(\xi) \in \mathcal{F}} (E_Q[g(\xi)] - E_P[g(\xi)])
\]
and
\[
\mathcal{H}_S(Q, P) := \max \{ \mathcal{D}_S(Q, P), \mathcal{D}_S(P, Q) \}.
\]

Let \( \mathcal{S}(P) \) and \( \mathcal{S}(Q) \) denote respectively the set of stationary points of problems (5.1) and (5.2), or equivalently the set of solutions of generalized equations (5.6) and (5.7). Let \( S(P) := \mathcal{S}(P) \cap Z \) and \( S(Q) := \mathcal{S}(Q) \cap Z \).

We are now ready to study the relationship between \( S(Q) \) and \( S(P) \), that is, the stability of stationary points.

**Theorem 5.2.** Consider the SGE (5.6) and its perturbation (5.7). Assume (a') \( G(x, \xi) \) is Lipschitz continuous in \( x \) for every \( \xi \) with modulus \( L_1 \) (independent of \( x \) and \( \xi \)), (b') \( |G(x, \xi)| \) is bounded by a positive constant \( L_2 \) (independent of \( x \) and \( \xi \)), (c') \( f(x, \xi) \) is Lipschitz continuous in \( x \) for every \( \xi \) and the Lipschitz modulus is bounded by an integrable function \( \kappa(\xi) \), and (d') \( S(P) \) and \( S(Q) \) are nonempty. Then the conclusions (i)-(iii) of Theorem 3.1 hold for \( S(P) \) and \( S(Q) \).
The thrust of the proof is to apply Theorem 3.1 to generalized equation (5.6) and its perturbation (5.7), taking into account Remark 3.4 as the single-valued component of \( \Gamma \) is infinite dimensional. To this end, we verify hypotheses of Theorem 3.1. Note that hypothesis (c) is satisfied as \( \mathcal{G}(\cdot) \) (defined by (5.5)) is upper semicontinuous, while (d) coincides with (d')'. Therefore we are left to verify (a) and (b).

Observe first that \( \partial_s H(x, \eta, \xi) \) is convex and compact valued (bounded by \( L_1 \)) and by [3, Theorems 1 and 4] of Aumann's integral, \( \int_a^b \partial_s H(x, \eta, \xi) \mu(d\eta) \) is also compact and convex set valued. Since the other components of \( \Gamma(x, \mu, \xi) \) are single valued, this shows \( \Gamma \) is convex and compact valued and hence verifies (a).

In what follows, we verify (b), that is, upper semicontinuity of \( \Gamma(x, \mu, \xi) \) with respect to \( (x, \mu) \) and its integrable boundedness. Let us look into the third component \( \int_a^b \partial_s H(x, \eta, \xi) \mu(d\eta) \). Under condition (b'), i.e., the boundedness of \( G(x, \xi) \), it is easy to see that \( H(x, \eta, \xi) \) is also bounded (by \( L_2 \)). Moreover, since the Lebesgue measure \( \mu(\cdot) \) is bounded, then \( \int_a^b H(x, \eta, \xi) \mu(d\eta) \) is continuous w.r.t. \((x, \mu)\).

Let us now consider the second component of \( \Gamma(x, \mu, \xi) \), that is, the functional \( H(x, \cdot, \xi) \) defined on interval \([a, b]\) w.r.t. \( x \). By the definition

\[
\|H(x, \cdot, \xi) - H(x', \cdot, \xi)\|_\infty = \sup_{\eta \in [a, b]} \|\eta - G(x, \xi)\_+ - (\eta - G(x', \xi))_+\|
\]

\[
\leq |G(x, \xi) - G(x', \xi)| \leq \kappa(\xi) \|x - x'\|
\]

which implies the continuity of \( H(x, \cdot, \xi) \) w.r.t. \( x \).

Finally, we consider the first component of \( \Gamma(x, \mu, \xi) \), that is, \( \nabla_x f(x, \xi) + \int_a^b \partial_s H(x, \eta, \xi) \mu(d\eta) \). Since \( f \) is assumed to be continuously differentiable, it suffices to verify the upper semicontinuity of \( \int_a^b \partial_s H(x, \eta, \xi) \mu(d\eta) \) w.r.t. \((x, \mu)\). Using the property of \( \mathbb{D} \), we have

\[
\mathbb{D} \left( \int_a^b \partial_s H(x', \eta, \xi) \mu'(d\eta), \int_a^b \partial_s H(x, \eta, \xi) \mu(d\eta) \right)
\]

\[
\leq \mathbb{D} \left( \int_a^b \partial_s H(x', \eta, \xi) \mu'(d\eta), \int_a^b \partial_s H(x, \eta, \xi) \mu'(d\eta) \right)
\]

\[
+ \mathbb{D} \left( \int_a^b \partial_s H(x, \eta, \xi) \mu'(d\eta), \int_a^b \partial_s H(x, \eta, \xi) \mu(d\eta) \right).
\]

Since \( \partial_s H(x, \eta, \xi) \) is convex and compact set valued, by Hörmander’s theorem and [27, Proposition 3.4]

\[
\mathbb{D} \left( \int_a^b \partial_s H(x', \eta, \xi) \mu'(d\eta), \int_a^b \partial_s H(x, \eta, \xi) \mu'(d\eta) \right)
\]

\[
= \sup_{\|u\| \leq 1} \left( \int_a^b \sigma(\partial_s H(x', \eta, \xi), u) - \sigma(\partial_s H(x, \eta, \xi), u) \mu'(d\eta) \right).
\]

It is easy to verify that \( \partial_s H(X', \eta, \xi) \) is upper semicontinuous in \( x \) for every fixed \( \eta \) and \( \xi \) and it is bounded by \( \|\nabla_x G(x, \xi)\| \) which is integrably bounded by assumption.

By [3, Corollary 5.2],

\[
\lim_{x' \to x} \int_a^b \partial_s H(x', \eta, \xi) \mu'(d\eta) \subseteq \int_a^b \partial_s H(x, \eta, \xi) \mu'(d\eta),
\]
which implies
\[
\lim_{x' \to x} \sigma \left( \int_a^b \partial_x H(x', \eta, \xi) \mu'(d\eta), u \right) \leq \sigma \left( \int_a^b \partial_x H(x, \eta, \xi) \mu'(d\eta), u \right)
\]
for any \( u \) with \( \|u\| \leq 1 \). Through [27, Proposition 3.4], the latter inequality can be written as
\[
(5.8) \quad \lim_{x' \to x} \int_a^b \sigma(\partial_x H(x', \eta, \xi), u) \mu'(d\eta) \leq \int_a^b \sigma(\partial_x H(x, \eta, \xi), u) \mu'(d\eta).
\]
Let \( x_k \to x \) and \( u_k \) be such that \( \|u_k\| \leq 1 \) and
\[
\sup_{\|u\| \leq 1} \left( \int_a^b [\sigma(\partial_x H(x_k, \eta, \xi), u_k) - \sigma(\partial_x H(x, \eta, \xi), u_k)] \mu'(d\eta) \right)
\]
\[
= \int_a^b [\sigma(\partial_x H(x_k, \eta, \xi), u_k) - \sigma(\partial_x H(x, \eta, \xi), u_k)] \mu'(d\eta).
\]
Assume by taking a subsequence if necessary that \( u_k \to u \). Using the continuity of the support function w.r.t. \( u \) and the inequality (5.8), we obtain
\[
\lim_{k \to \infty} \int_a^b [\sigma(\partial_x H(x_k, \eta, \xi), u_k) - \sigma(\partial_x H(x, \eta, \xi), u_k)] \mu'(d\eta) \leq 0.
\]
Since \( x_k \) is arbitrary, this implies
\[
\lim_{x' \to x} \sup_{\|u\| \leq 1} \left( \int_a^b [\sigma(\partial_x H(x', \eta, \xi), u) - \sigma(\partial_x H(x, \eta, \xi), u)] \mu'(d\eta) \right) \leq 0.
\]
On the other hand, it follows by [23, Lemma 5.1] that
\[
\mathbb{D} \left( \int_a^b \partial_x H(x, \eta, \xi) \mu'(d\eta), \int_a^b \partial_x H(x, \eta, \xi) \mu(d\eta) \right) \to 0
\]
as \( \mu' \to \mu \). The discussions above show that \( \int_a^b \partial_x H(x, \eta, \xi) \mu(d\eta) \) is upper semicontinuous w.r.t. \( (x, \mu) \).

To complete the verification of (b), we need to show the integrable boundedness of \( \Gamma(x, \mu, \xi) \). It is easy to observe that \( \partial_x H(x, \eta, \xi) \) is bounded by \( L_1 \) and hence \( \int_a^b \partial_x H(x, \eta, \xi) \mu(d\eta) \) is bounded by \( L_1 \mu([a, b]) \). The boundedness of \( G(x, \xi) \) by \( L_2 \) implies the same boundedness of \( \|H(x, \cdot, \xi)\|_\infty \) and \( \int_a^b \partial_x H(x, \eta, \xi) \mu(d\eta) \). Together with the boundedness of \( \nabla_x f(x, \xi) \) (by an integrable \( \kappa(\xi) \)), we have shown that \( \Gamma(x, \mu, \xi) \) is integrably bounded. The proof is complete.

Note that condition (d') implicitly assumes that the Lagrange multipliers of problems (5.1) and (5.2) are bounded at some stationary points. A sufficient condition for this is that the problems satisfy certain constraint qualifications. The issue has been investigated by Sun and Xu in [40, section 3]; we refer interested readers to [40, Proposition 3.1].
Acknowledgment. We would like to thank two referees for insightful comments which significantly help us improve the quality of the paper.

REFERENCES


