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.

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

$$0 \in \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x), \tag{1.1}$$

where $\Gamma : \mathcal{X} \times \Xi \to 2^{\mathcal{Y}}$ and $\mathcal{G} : \mathcal{X} \to 2^{\mathcal{Y}}$ are closed set-valued mappings, $\mathcal{X}$ and $\mathcal{Y}$ are subsets of Banach spaces X and Y (with norm $\|\cdot\|_X$ and $\|\cdot\|_Y$) 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 $\mathbb{E}_P[\cdot]$ denotes the expected value with respect to P, that is,

$$\begin{aligned}
\mathbb{E}_P[\Gamma(x,\xi)] &:= \int_\Xi \Gamma(x,\xi) dP(\xi) \\
&= \left\{ \int_\Xi \psi(\xi) P(d\xi) : \psi \text{ is a Bochner integrable selection of } \Gamma(x,\cdot) \right\}.
\end{aligned}$$

The expected value of Γ 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 Γ is single valued and $\mathcal{G}(x)$ is a normal cone of a set, (1.1) is also known as 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 on 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 true problem and that of the approximated problem.

---

[*]Department of Mathematics, Dalian Maritime University, Dalian 116026, China. Part of this author's work is carried out while he was visiting the third author in the School of Mathematics, University of Southampton sponsored by China Scholarship Council. The work of this author is supported in part by NSFC Grant #11201044 and FRFCU Grant #3132013099.

[†]Institute of Mathematics, Humboldt University Berlin, D-10099 Berlin, Germany. The work of this author is supported by the DFG Research Center MATHEON at Berlin.

[‡] School of Engineering and Mathematical Sciences, City University London, EC1V 0HB, London, UK. Part of this author's work was carried out during his visit to the second author at Berlin in 2011 and second author's visit to Southampton in 2012.

Let Q denote a perturbation of the probability measure P. We consider the following perturbed stochastic generalized equations:

$$0 \in \mathbb{E}_Q[\Gamma(x,\xi)] + \mathcal{G}(x). \tag{1.2}$$

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 equations (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 $\mathbb{E}_P[\Gamma(\cdot,\xi)]$ 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 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 stochastic generalized equations where the true probability measure is approximated through 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 stochastic generalized equations with respect to the perturbation of the probability measure. Applications of the established results to classical one stage and two stages linear stochastic programs and two stage stochastic mathematical programs with complementarity constraints 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. $\mathcal{Z}$ denotes a Banach space with norm $\|\cdot\|_{\mathcal{Z}}$ and $\mathbb{R}^n$ denotes n dimensional Euclidean space. By convention, we write $\langle u, z \rangle$ for dual pairing of $z \in \mathcal{Z}$ which is bilinear, where u is from the dual space of $\mathcal{Z}$. In the case when $\mathcal{Z}$ is finite dimensional, the dual pairing reduces to scalar product. Given a point $z \in \mathcal{Z}$ and a set $\mathcal{D}$, we write $d(z,\mathcal{D}) := \inf_{z' \in \mathcal{D}} \|z - z'\|_{\mathcal{Z}}$ for the distance from z to $\mathcal{D}$. For two closed sets $\mathcal{C}$ and $\mathcal{D}$,

$$\mathbb{D}(\mathcal{C},\mathcal{D}) := \sup_{z \in \mathcal{C}} d(z,\mathcal{D})$$

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

$$\mathbb{H}(\mathcal{C},\mathcal{D}) := \max\left(\mathbb{D}(\mathcal{C},\mathcal{D}), \mathbb{D}(\mathcal{D},\mathcal{C})\right).$$

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

$$\overline{\lim_{k \to \infty}} S_k = \{x : \liminf_{k \to \infty} d(x, S_k) = 0\}.$$

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

$$\Psi(\bar{x} + \delta\mathcal{B}) \subset \Psi(\bar{x}) + \epsilon\mathcal{B}.$$

Ψ is said to be *lower semi-continuous* (lsc for short) 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 $\mathcal{U}$ with $\mathcal{U} \cap \Psi(\bar{x}) \neq \emptyset$, the set $\{x \in X : \mathcal{U} \cap \Psi(x) \neq \emptyset\}$ is a neighborhood of $\bar{x}$. Ψ is said to be *continuous* at $\bar{x}$ if it is both usc and lsc 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 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

$$\mathbb{E}_P[\Gamma(x,\xi)] \subset \mathbb{E}_Q[\Gamma(x,\xi)] + \epsilon\mathcal{B} \tag{2.1}$$

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*

$$\Delta(P) := \sup_{\|u\|=1} \inf_{x \in \mathcal{X}} \inf_{\zeta \in \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x)} \langle u, \zeta \rangle < 0. \tag{2.2}$$

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

$$\inf_{\zeta \in \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x)} \langle u, \zeta \rangle \geq \inf_{\zeta \in \mathbb{E}_Q[\Gamma(x,\xi)] + \mathcal{G}(x)} \langle u, \zeta \rangle - \epsilon.$$

Therefore

$$\inf_{x \in \mathcal{X}} \inf_{\zeta \in \mathbb{E}_Q[\Gamma(x,\xi)] + \mathcal{G}(x)} \langle u, \zeta \rangle \leq \inf_{x \in \mathcal{X}} \inf_{\zeta \in \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x)} \langle u, \zeta \rangle + \epsilon \leq \Delta(P) + \epsilon < 0.$$

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 ξ and $\mathcal{G}(x)$ is a convex set-valued mappings. In the case when Γ 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 usc 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. ξ independent of x. Assumption 2.1 (c) means that the set

$$\left\{ x \in \mathcal{X} : \inf_{\zeta \in \mathbb{E}_Q[\Gamma(x,\xi)] + \mathcal{G}(x)} \langle u, \zeta \rangle \le \alpha \right\}$$

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 has a 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* $\Psi : X \to 2^Y$ *be a closed set valued mapping. For* $\bar{x} \in X$ *and* $\bar{y} \in \Psi(\bar{x})$, Ψ *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*

$$d(x, \Psi^{-1}(y)) \le \alpha d(y, \Psi(x)), \quad \forall x \in U, \ y \in V.$$

Here the inverse mapping Ψ^{-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 $\operatorname{reg} \Psi(\bar{x}|\bar{y})$. $\Psi(x)$ 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 Ψ at $\bar{x}$ for $\bar{y}$ and to Aubin's property of Ψ^{-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* Ψ *is metrically regular at* $\bar{x}$ *for* 0 *with the neighborhoods of* $U_{\bar{x}}$ *of* $\bar{x}$ *and* V_0 *of* 0. *If* $0 \in \tilde{\Psi}(x)$ *with* $x \in U_{\bar{x}}$, *then*

$$d\left(x, \Psi^{-1}(0)\right) \le \alpha \mathbb{D}(\tilde{\Psi}(x), \Psi(x)),$$

where α *is the regularity modulus of* Ψ *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*

$$\|x - \bar{x}\| \le \alpha \mathbb{D}(\tilde{\Psi}(x), \Psi(x)).$$

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 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. *([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*

$$\mathbb{D}(\mathcal{C}, \mathcal{D}) = \max_{\|u\| \le 1} (\sigma(\mathcal{C}, u) - \sigma(\mathcal{D}, u))$$

and

$$\mathbb{H}(\mathcal{C}, \mathcal{D}) = \max_{\|u\| \leq 1} |\sigma(\mathcal{C}, u) - \sigma(\mathcal{D}, u)|.$$

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 μ, μ_x and μ_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 $\int_{\mathcal{X} \times \mathcal{Y}} \Psi(x, y)\mu(dxdy)$, $\int_{\mathcal{X}} \int_{\mathcal{Y}} \Psi(x, y)\mu_y(dy)\mu_x(dx)$ and $\int_{\mathcal{Y}} \int_{\mathcal{X}} \Psi(x, y)\mu_x(dx)\mu_y(dy)$, where $\mathcal{X}$ and $\mathcal{Y}$ are nonempty compact subset 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 Ψ is upper semi-continuous with respect to x and y. Then the following assertions hold.*

(i) $\sigma(\Psi(x, y), u)$ is upper semi-continuous in x and y uniformly w.r.t. u;

if, in addition, Ψ is μ-integrably bounded, that is, there exists a nonnegative μ-integrable function $\kappa(x, y)$ with $\int_{\mathcal{X} \times \mathcal{Y}} \kappa(x, y)\mu(dxdy) < \infty$ such that

$$\|\Phi(x, y)\| \leq \kappa(x, y),$$

then

(ii) $\Psi(\cdot, y)$ and $\Psi(x, \cdot)$ are μ_x and μ_y integrably bounded for each y and x respectively, and

$$\int_{\mathcal{X} \times \mathcal{Y}} \Psi(x, y)\mu(dxdy) = \int_{\mathcal{X}} \int_{\mathcal{Y}} \Psi(x, y)\mu_y(dy)\mu_x(dx) = \int_{\mathcal{Y}} \int_{\mathcal{X}} \Psi(x, y)\mu_x(dx)\mu_y(dy);$$

(iii) for any $x', x \in \mathcal{X}$,

$$\mathbb{H}\left(\int_{\mathcal{Y}} \Psi(x', y)\mu_y(dy), \int_{\mathcal{Y}} \Psi(x, y)\mu_y(dy)\right) \leq \int_{\mathcal{Y}} \mathbb{H}(\Psi(x', y), \Psi(x, y))\mu_y(dy).$$

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

Part (i). Since Ψ is upper semi-continuous 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 semi-continuous in x and y uniformly w.r.t. u.

Part (ii) is well known, see [47, Theorem 2.1]. Here we include a proof for completeness. By assumption, Ψ is nonempty, compact, convex and integrably bounded. It follows by [17, Theorem 5.4], that $\int_{\mathcal{X}} \int_{\mathcal{Y}} \Psi(x, y)\mu_y(dy)\mu_x(dx)$ and $\int_{\mathcal{Y}} \int_{\mathcal{X}} \Psi(x, y)\mu_x(dx)\mu_y(dy)$ are nonempty, compact and convex. By Hörmander's theorem (Lemma 2.5)

$$\mathbb{D}\left(\int_{\mathcal{X}} \int_{\mathcal{Y}} \Psi(x, y)\mu_y(dy)\mu_x(dx), \int_{\mathcal{Y}} \int_{\mathcal{X}} \Psi(x, y)\mu_x(dx)\mu_y(dy)\right)$$
$$= \sup_{\|u\| \leq 1} \left[\sigma\left(\int_{\mathcal{X}} \int_{\mathcal{Y}} \Psi(x, y)\mu_y(dy)\mu_x(dx), u\right) - \sigma\left(\int_{\mathcal{Y}} \int_{\mathcal{X}} \Psi(x, y)\mu_x(dx)\mu_y(dy), u\right)\right].$$

Applying [27, Proposition 3.4] to the support function above, we have

$$\sigma\left(\int_{\mathcal{X}} \int_{\mathcal{Y}} \Psi(x, y)\mu_y(dy)\mu_x(dx), u\right) = \int_{\mathcal{X}} \int_{\mathcal{Y}} \sigma(\Psi(x, y), u)\mu_y(dy)\mu_x(dx)$$

and

$$\sigma\left(\int_{\mathcal{Y}} \int_{\mathcal{X}} \Psi(x, y)\mu_x(dx)\mu_y(dy), u\right) = \int_{\mathcal{Y}} \int_{\mathcal{X}} \sigma(\Psi(x, y), u)\mu_x(dx)\mu_y(dy).$$

It follows from part (i) that $\sigma(\Psi(x,y),u)$ is upper semi-continuous 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

$$\int_{\mathcal{X}}\int_{\mathcal{Y}}\sigma(\Psi(x,y),u)\mu_y(dy)\mu_x(dx) = \int_{\mathcal{Y}}\int_{\mathcal{X}}\sigma(\Psi(x,y),u)\mu_x(dx)\mu_y(dy).$$

The discussions above yield

$$\mathbb{D}\left(\int_{\mathcal{X}}\int_{\mathcal{Y}}\Psi(x,y)\mu_y(dy)\mu_x(dx), \int_{\mathcal{Y}}\int_{\mathcal{X}}\Psi(x,y)\mu_x(dx)\mu_y(dy)\right)$$
$$= \sup_{\|u\|\leq 1}\left[\int_{\mathcal{X}}\int_{\mathcal{Y}}\sigma(\Psi(x,y),u)\mu_y(dy)\mu_x(dx) - \int_{\mathcal{Y}}\int_{\mathcal{X}}\sigma(\Psi(x,y),u)\mu_x(dx)\mu_y(dy)\right]$$
$$= 0.$$

Part (iii). The result is well known. See [17, Theorem 5.4]. Indeed, following similar arguments as in the proof of Part (ii), we have

$$\mathbb{H}\left(\int_{\mathcal{Y}}\Psi(x',y)\mu_y(dy), \int_{\mathcal{Y}}\Psi(x,y)\mu_y(dy)\right) \leq \int_{\mathcal{Y}}\sup_{\|u\|\leq 1}|\sigma(\Psi(x,y),u) - \sigma(\Psi(x,y),u)|\,\mu_y(dy)$$
$$= \int_{\mathcal{Y}}\mathbb{H}(\Psi(x',y),\Psi(x,y))\mu_y(dy).$$

The proof is complete. ∎

3. Stability of stochastic generalized equations. Let $\mathscr{P}(\Xi)$ denote the set of all Borel probability measures on Ξ. Assuming Q is close to P under some metric to be defined shortly, we investigate the relationship between the solution set of stochastic generalized equations (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

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

Then $\mathscr{F}$ 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

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

and

$$\mathscr{H}(Q,P) := \max\left(\mathscr{D}(Q,P), \mathscr{D}(P,Q)\right).$$

It is easy to verify that

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

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

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

and hence

$$\mathscr{D}(Q,P) = 0 \implies \mathbb{E}_Q[\Gamma(x,\xi)] \subseteq \mathbb{E}_P[\Gamma(x,\xi)], \quad \forall x \in \mathcal{X}.$$

Likewise

$$\mathscr{H}(Q,P) = 0 \implies \mathbb{E}_Q[\Gamma(x,\xi)] = \mathbb{E}_P[\Gamma(x,\xi)], \quad \forall x \in \mathcal{X}.$$

Neither $\mathscr{H}$ nor $\mathscr{D}$ is a metric but one may enlarge the set $\mathscr{F}$ so that $\mathscr{H}(Q,P) = 0$ implies $Q = P$. We call $\mathscr{H}(Q,P)$ a *pseudometric*. It is also known as a distance of probability measures having ζ-structure, see [48].

Recall that for a sequence of probability measures $\{P_N\}$ in $\mathscr{P}(\Xi)$, 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 Ξ.

Let $\mathscr{F}$ be defined by (3.1) and $\{P_N\} \subset \mathscr{P}(\Xi)$. We say $\mathscr{F}$ defines an *upper P-uniformity class* of functions if

$$\lim_{N \to \infty} \mathscr{D}(P_N, P) = 0$$

for every sequence $\{P_N\}$ which converges weakly to P, and a *P-uniformity class* if

$$\lim_{N \to \infty} \mathscr{H}(P_N, P) = 0.$$

A family of functions $\mathscr{K}$ 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 $f \in \mathscr{K}$ and all x, x_0 such that $\|x_0 - x\| \le \delta$. A sufficient condition for $\mathscr{K}$ to be a P-uniformity class is that $\mathscr{K}$ is uniformly bounded and

$$P(\{\xi \in \Xi : \mathscr{K} \text{ is not equicontinuous at } \xi\}) = 0,$$

see [41]. In our context, the latter is implied by

$$\lim_{\xi' \to \xi} \sup_{x \in \mathcal{X}} \mathbb{H}(\Gamma(x,\xi'), \Gamma(x,\xi)) = 0 \tag{3.2}$$

for almost every ξ w.r.t. probability measure P. At this point, we may refer readers to the work by Artstein and Wets [1] on 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 Γ takes convex and compact values and is continuous in ξ, see [1, Theorem 3.1].

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

(i) *For any small positive number ϵ, let*

$$R(\epsilon) := \inf_{x \in \mathcal{X}, \, d(x, S(P)) \ge \epsilon} d(0, \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x)). \tag{3.3}$$

Then

$$\mathbb{D}(S(Q), S(P)) \le R^{-1}(2\mathscr{D}(Q,P)),$$

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 ϵ, there exists a $\delta > 0$ such that if $\mathscr{D}(Q,P) \le \delta$, then $\mathbb{D}(S(Q), S(P)) \le \epsilon$.*

(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 α, then there exists neighborhood U_{x^*} of x^* such that*

$$d(x, S(P)) \leq \alpha \mathscr{D}(Q, P) \tag{3.4}$$

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

$$\|x - x^*\|_X \leq \alpha \mathscr{D}(Q, P) \tag{3.5}$$

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

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

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

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 $\{x_N\} \subseteq \mathcal{X}$ with $d(x_N, S(P)) \geq \epsilon$ such that

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

which is equivalent to

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

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

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

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

Let $\delta := R(\epsilon)/2$ and $\mathscr{D}(Q, P) \leq \delta$. Let $\rho' := \min_{x \in \mathcal{X}} d(0, \mathcal{G}(x))$. Under the closedness and upper semi-continuity of $\mathcal{G}(\cdot)$, 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 $d(x, S(P)) > \epsilon$,

$$
\begin{aligned}
d(0, \mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x)) &= d(0, \mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x) \cap t\mathcal{B}) \\
&\geq d(0, \mathbb{E}_P[\Gamma(x, \xi)] + \mathcal{G}(x) \cap t\mathcal{B}) \\
&\quad - \mathbb{D}(\mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x) \cap t\mathcal{B}, \mathbb{E}_P[\Gamma(x, \xi)] + \mathcal{G}(x) \cap t\mathcal{B}),
\end{aligned} \tag{3.8}
$$

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

$$\mathbb{D}(\mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x) \cap t\mathcal{B}, \mathbb{E}_P[\Gamma(x, \xi)] + \mathcal{G}(x) \cap t\mathcal{B}) \leq \mathbb{D}(\mathbb{E}_Q[\Gamma(x, \xi)], \mathbb{E}_P[\Gamma(x, \xi)]), \tag{3.9}$$

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

$$
\begin{aligned}
\mathbb{D}(\mathbb{E}_Q[\Gamma(x, \xi)], \mathbb{E}_P[\Gamma(x, \xi)]) &= \sup_{\|u\| \leq 1} (\sigma(\mathbb{E}_Q[\Gamma(x, \xi)], u) - \sigma(\mathbb{E}_P[\Gamma(x, \xi)], u)) \\
&= \sup_{\|u\| \leq 1} (\mathbb{E}_Q[\sigma(\Gamma(x, \xi), u)] - \mathbb{E}_P[\sigma(\Gamma(x, \xi), u)]).
\end{aligned} \tag{3.10}
$$

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

$$\sup_{\|u\| \leq 1} (\mathbb{E}_Q[\sigma(\Gamma(x, \xi), u)] - \mathbb{E}_P[\sigma(\Gamma(x, \xi), u)]) \leq \mathscr{D}(Q, P). \tag{3.11}$$

Combining (3.8)–(3.11), we have

$$\begin{aligned} d(0, \mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x)) &\geq d(0, \mathbb{E}_P[\Gamma(x, \xi)] + \mathcal{G}(x)) - \mathscr{D}(Q, P) \\ &\geq R(\epsilon) - \delta \\ &= \delta > 0. \end{aligned} \tag{3.12}$$

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

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

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

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

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

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

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 $x \in S(Q)$,

$$0 \in \mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x) \cap t\mathcal{B}. \tag{3.13}$$

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

$$d(x, S(P)) \leq \alpha d(0, \Phi(x)) \tag{3.14}$$

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

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

and hence

$$\begin{aligned} d(x, S(P)) &\leq \alpha d(0, \mathbb{E}_P[\Gamma(x, \xi)] + \mathcal{G}(x) \cap t\mathcal{B}) \\ &\leq \alpha \mathbb{D}(\mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x) \cap t\mathcal{B}, \mathbb{E}_P[\Gamma(x, \xi)] + \mathcal{G}(x) \cap t\mathcal{B}) \end{aligned} \tag{3.15}$$

for all $x \in S(Q) \cap U_{x^*}$. The second inequality is due to (3.13) and the definition of $\mathbb{D}$. Note that for any bounded sets $\mathcal{C}, \mathcal{C}', \mathcal{D}, \mathcal{D}'$, it is easy to verify that

$$\mathbb{D}(\mathcal{C} + \mathcal{C}', \mathcal{D} + \mathcal{D}') \leq \mathbb{D}(\mathcal{C}, \mathcal{D}) + \mathbb{D}(\mathcal{C}', \mathcal{D}').$$

Using this relationship and (3.9)–(3.11), we obtain

$$\begin{aligned} \mathbb{D}(\mathbb{E}_Q[\Gamma(x, \xi)] + \mathcal{G}(x) \cap t\mathcal{B}, \mathbb{E}_P[\Gamma(x, \xi)] + \mathcal{G}(x) \cap t\mathcal{B}) &\leq \mathbb{D}(\mathbb{E}_Q[\Gamma(x, \xi)], \mathbb{E}_P[\Gamma(x, \xi)]) \\ &\leq \mathscr{D}(Q, P). \end{aligned} \tag{3.16}$$

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 $\Phi(x) := \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x)$ and $V := \{v : v \in \Phi(x), \forall x \in S(P)\}$. Let ϵ be a small positive number. Assume that, for any x with $d(x, S(P)) \geq \epsilon$, there exists positive constants C and τ (depending on ϵ) such that*

$$\|v' - v\| \geq Cd(x, S(P))^\tau, \quad \forall\, v' \in \Phi(x),\ v \in V. \tag{3.17}$$

Then there exists a positive constant α such that

$$R^{-1}(\epsilon) \leq \alpha\epsilon^{\frac{1}{\tau}}. \tag{3.18}$$

Proof. By definition,

$$
\begin{aligned}
R(\epsilon) &= \inf_{x \in \mathcal{X},\, d(x, S(P)) \geq \epsilon} d(0, \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x)) \\
&\geq \inf_{x \in \mathcal{X},\, d(x, S(P)) \geq \epsilon} \inf_{x^* \in S(P)} \inf_{v \in \Phi(x^*)} d(v, \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x)) \\
&= \inf_{x \in \mathcal{X},\, d(x, S(P)) \geq \epsilon} \inf_{x^* \in S(P)} \inf_{v \in \Phi(x^*), v' \in \Phi(x)} \|v - v'\| \\
&\geq \inf_{x \in \mathcal{X},\, d(x, S(P)) \geq \epsilon} Cd(x, S(P))^\tau \\
&\geq C\epsilon^\tau,
\end{aligned}
$$

where the second 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 ϵ,

$$\|v' - v\| = \|x^2 - 0\| \geq \epsilon d(x, 0), \quad \forall\, v' \in \Phi(x),\ v \in V.$$

Note also that (3.17) is implied by strong monotonicity of $\Phi(\cdot)$, that is, for all $x, x' \in \mathcal{X}$

$$\langle v' - v, x' - x \rangle \geq C^*\|x' - x\|^2, \quad \forall\, v' \in \Phi(x'),\ v \in \Phi(x),$$

see [5, 6] and [35, Definition 12.53] for finite dimensional case. Under the strong monotonicity

$$\|v' - v\|\|x' - x\| \geq \langle v' - v, x' - x \rangle \geq C^*\|x' - x\|^2, \quad \forall\, v' \in \Phi(x'),\ v \in \Phi(x),$$

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 ξ 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)[1]

$$0 \in \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{N}_K(x). \tag{3.19}$$

Observe that

$$d(0, \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{N}_K(x)) = d(-\mathbb{E}_P[\Gamma(x,\xi)], \mathcal{N}_K(x)).$$

---

[1]Note that there are many types of SVIP models. For instances, Chen et al [9] recently proposed a stochastic VIP model where K depends on every realization of random variable ξ and a deterministic solution x is sought for solving $0 \in \Gamma(x,\xi) + \mathcal{N}_{K(\xi)}(x), \forall \xi \in \Xi$. It is unclear whether results to be presented here can be established for the new SVIP models in the same manner.

By [16, Proposition 1.5.14],

$$d(-\mathbb{E}_P[\Gamma(x,\xi)], \mathcal{N}_K(x)) = \inf\{\|F_K^{\mathrm{nor}}(z)\| : z \in \Pi_K^{-1}(x)\},$$

where Π_K denotes the Euclidean projection onto K and Π_K^{-1} its inverse,

$$F_K^{\mathrm{nor}}(z) := \mathbb{E}_P[\Gamma(\Pi_K(z), \xi)] + z - \Pi_K(z),$$

which is known as Robinson's normal map. Let $\mathcal{Z} = \mathcal{X}$. It is easy to verify that $F_K^{\mathrm{nor}}(z)$ is Lipschitz continuous on $\mathcal{Z}$ 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 polyhedral multifunction and through [26, Theorem 2.4], 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 Π_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 exists positive constants C and τ such that*

$$\|F_K^{\mathrm{nor}}(z)\| \geq Cd(z, Z^*)^\tau, \forall z \in \Pi_K^{-1}(\mathcal{X}), \tag{3.20}$$

that is, $\|F_K^{\mathrm{nor}}(z)\|$ satisfies some growth condition as z deviating from Z^*. Then there exists a positive constant α such that (3.18) holds. If, in addition, $F_K^{\mathrm{nor}}(z)$ is locally Lipschitz homeomorphism near z^*, that is, there exist neighborhoods of z^* and $F_K^{\mathrm{nor}}(z^*)$ such that the map $F_K^{\mathrm{nor}}(\cdot)$ 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),

$$\inf_{z \in \Pi_K^{-1}(x)} \|F_K^{\mathrm{nor}}(z)\| = \inf_{z \in \Pi_K^{-1}(x), z^* \in Z^*} (\|F_K^{\mathrm{nor}}(z)\| - \|F_K^{\mathrm{nor}}(z^*)\|)$$
$$\geq \inf_{z \in \Pi_K^{-1}(x)} Cd(z, Z^*)^\tau.$$

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

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

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

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

where C' is a positive constant. Consequently

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

The conclusions follow.   ∎

Note that part (iii) of Theorem 3.1 is derived under metric regularity. It is difficulty to verify the condition in general. However, when either $\mathbb{E}_P[\Gamma(x,\xi)]$ or $\mathcal{G}(x)$ reduces to a singleton, then we may characterize the metric regularity of $\Psi(x) = \mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}(x)$ through Mordukhovich coderivative, see [14]. For example, when $\mathbb{E}_P[\Gamma(x,\xi)]$ is single valued and $\mathcal{G}(x)$ is a normal cone, [14, Theorem 5.1] gives details on this. In the case when 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 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 ξ and $\mathcal{G}(x)$ is independent of x, e.g., $\mathcal{G}(x) = \mathcal{G}$, where $\mathcal{G}$ is a closed convex set in Y, the SGE recovers a stochastic cone constraint. In that case, the metric regularity of $\mathbb{E}_P[\Gamma(x,\xi)] + \mathcal{G}$ is equivalent to Robinson's constraint qualification, that is, there exists $x_0 \in \mathcal{X}$ such that

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

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

Finally, note that 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 equations

$$F(x) = 0$$

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

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

REMARK 3.4. *The assumption of Y to be a Euclidean space (finite dimensional) is only required in (3.6). In some applications, Γ 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 stochastic generalized equations 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

$$\begin{aligned} \min_x \quad & \mathbb{E}_P[f(x,\xi)] \\ \text{s.t.} \quad & x \in X, \end{aligned} \tag{4.1}$$

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 semi-continuous 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, 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 models 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 Clarke subdifferential of ψ at x, denoted by $\partial\psi(x)$, is defined as follows:

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

where D denotes the set of points near x at which ψ is Fréchet differentiable, $\nabla\psi(x)$ denotes the gradient of ψ 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:

$$0 \in \partial\mathbb{E}_P[f(x,\xi)] + \mathcal{N}_X(x). \tag{4.2}$$

The condition is also sufficient if $f(\cdot,\xi)$ is convex for almost every ξ. 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

$$0 \in \mathbb{E}_P[\partial_x f(x,\xi)] + \mathcal{N}_X(x). \tag{4.3}$$

The condition is weaker in that $\partial\mathbb{E}_P[f(x,\xi)] \subseteq \mathbb{E}_P[\partial_x f(x,\xi)]$ 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:

$$\begin{aligned} \min_{x} \quad & \mathbb{E}_Q[f(x,\xi)] \\ \text{s.t.} \quad & x \in X, \end{aligned} \tag{4.4}$$

where Q is a perturbation of the probability measure P such that $\mathbb{E}_Q[f(x,\xi)]$ 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 $f^o(x,\xi;u)$ denote the Clarke generalized directional derivative for a given nonzero vector u and*

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

(i) *Assume: (a') $f(\cdot,\xi)$ is locally Lipschitz continuous for every ξ with P-integrable modulus; (b') $Q \in \mathscr{P}(\Xi)$; (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:*

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

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

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

and

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

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

$$\sup_{x \in X} \sup_{\tau \in (0,\delta)} \sup_{\|u\| \leq 1} \left| \frac{1}{\tau}(f(x + \tau u, \xi) - f(x, \xi)) - \frac{1}{\tau}(f(x + \tau u, \tilde{\xi}) - f(x, \tilde{\xi})) \right| \leq h(\|\xi - \tilde{\xi}\|) \quad (4.5)$$

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

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

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

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

and ζ_h the Kantorovich-Rubinstein functional

$$\zeta_h(P, Q) = \inf \int_{\Xi \times \Xi} h(\|\xi - \tilde{\xi}\|) d\eta(\xi, \tilde{\xi}), \quad (4.6)$$

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

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 $\mathcal{G}(x) = \mathcal{N}_X(x)$. 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 semi-continuous w.r.t. x for every fixed ξ. Conditions (d) follows from condition (d') and the fact that $\partial \mathbb{E}_P[f(x, \xi)] \subseteq \mathbb{E}_P[\partial_x f(x, \xi)]$.

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

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

In what follows, we use the notation $F_P(x) := \mathbb{E}_P[f(x, \xi)]$ and $F_Q(x) := \mathbb{E}_Q[f(x, \xi)]$, 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,

$$\begin{aligned}
D_* &= \sup_{x \in X, \|u\| \leq 1} (\sigma(\partial F_Q(x), u) - \sigma(\partial F_P(x), u) \\
&= \sup_{x \in X, \|u\| \leq 1} \left(\limsup_{x' \to x, \tau \downarrow 0} \frac{1}{\tau}(F_Q(x' + \tau u) - F_Q(x')) - \limsup_{x' \to x, \tau \downarrow 0} \frac{1}{\tau}(F_P(x' + \tau u) - F_P(x')) \right) \\
&\leq \sup_{x \in X, \|u\| \leq 1} \limsup_{x' \to x, \tau \downarrow 0} \left| \frac{1}{\tau}(F_Q(x' + \tau u) - F_Q(x')) - \frac{1}{\tau}(F_P(x' + \tau u) - F_P(x')) \right| \\
&= \sup_{x \in X, \|u\| \leq 1} \limsup_{x' \to x, \tau \downarrow 0} \left| \int_\Xi \frac{1}{\tau}(f(x' + \tau u, \xi) - f(x', \xi)) d(Q - P)(\xi) \right| \\
&\leq \sup_{x \in X, \|u\| \leq 1, \tau \in (0,\delta)} \left| \int_\Xi \frac{1}{\tau}(f(x + \tau u, \xi) - f(x, \xi)) d(Q - P)(\xi) \right| \\
&\leq \zeta_h(P, Q).
\end{aligned}$$

Here, we used for the first estimate the fact that the inequality

$$\left| \limsup_{k \to \infty} a_k - \limsup_{k \to \infty} b_k \right| \leq \limsup_{k \to \infty} |a_k - b_k|$$

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

$$\zeta_h(P,Q) = \sup_{g \in \mathscr{G}_h} \left| \int_\Xi g(\xi) d(P-Q)(\xi) \right|,$$

where the set $\mathscr{G}_h$ is defined by

$$\mathscr{G}_h = \{g : \Xi \to \mathbb{R} : |g(\xi) - g(\tilde{\xi})| \le h(\|\xi - \tilde{\xi}\|), \forall \xi, \tilde{\xi} \in \Xi\}$$

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 ξ, i.e., in particular, if the integrand is convex, the functions $g = f^o(x, \cdot; u)$ belong to the class $\mathscr{G}_h$ and, hence, we also obtain the estimate

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

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

$$\mathscr{P}_h(\Xi) = \{Q \in \mathscr{P}(\Xi) : \int_\Xi h(\|\xi\|) dQ(\xi) < +\infty\}.$$

Note that ζ_h is a (so-called) *simple distance* on $\mathscr{P}_h(\Xi)$ (see [28, Section 3.2]) which means that (i) $P = Q$ iff $\zeta_h(P,Q) = 0$, (ii) $\zeta_h(P,Q) = \zeta_h(Q,P)$, and (iii) $\zeta_h(P,Q) \le K_h(\zeta_h(P,\tilde{Q}) + \zeta_h(\tilde{Q},Q))$ for all $P, Q, \tilde{Q} \in \mathscr{P}_h(\Xi)$, where K_h is a positive constant depending on function h. An important special case is $h(t) = t^p$ with $p \ge 1$. In that case, one may deduce the *Wasserstein metric of order p* or *L_p-minimal metric ℓ_p* by setting $\ell_p(P,Q) = (\zeta_h(P,Q))^{\frac{1}{p}}$ with $\mathscr{P}_h(\Xi)$ being the set of all probability measures having finite pth 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 $p \ge 1$. In that case the distance ζ_h is replaced by the pth order Fortet-Mourier metric ζ_p (see [28, Section 5.1])) and $\mathscr{P}_h(\Xi)$ by the set of all probability measures having finite pth order moments.

In the case when f is convex w.r.t. x for almost every ξ, 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 $\mathscr{D}(Q,P) \to 0$. By Attouch's theorem ([35, Theorem 12.35]), which implies $\partial \mathbb{E}_Q[f(\cdot,\xi)]$ 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 et al [23] also investigated 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 ξ, 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 linear bound for $\mathbb{D}(X(Q), X(P))$.

4.2. Two-stage linear recourse problems. In what follows, we consider a linear two stage recourse minimization problem:

$$\begin{aligned}
\min_{x \in \mathbb{R}^n} \quad & c^\top x + \mathbb{E}_P[v(x,\xi)] \\
\text{s.t.} \quad & Ax = b, x \ge 0,
\end{aligned} \tag{4.7}$$

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

$$
\begin{aligned}
\min_{y \in \mathbb{R}^m} \quad & q(\xi)^\top y \\
\text{s.t.} \quad & T(\xi)x + Wy = h(\xi), \ y \ge 0,
\end{aligned}
\tag{4.8}
$$

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 ξ and that Ξ is a polyhedral subset of $\mathbb{R}^s$ (for example, $\Xi = \mathbb{R}^s$). If we consider the set $X = \{x \in \mathbb{R}^n : Ax = b, x \ge 0\}$ 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 ϕ_P is a convex polyhedral subset of $\mathbb{R}^n$ and it holds

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

where "pos W" denotes the positive hull of the matrix W, that is, $\mathrm{pos}\, W := \{Wy : y \ge 0\}$. Next, we recall some properties of v.

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

$$
|v(x, \xi) - v(\tilde{x}, \tilde{\xi})| \le \hat{L}(\max\{1, \|\xi\|, \|\tilde{\xi}\|\}^2 \|\tilde{x} - x\| + \max\{1, \|x\|, \|\tilde{x}\|\} \max\{1, \|\xi\|, \|\tilde{\xi}\|\} \|\tilde{\xi} - \xi\|)
\tag{4.9}
$$

for all pairs $(x, \xi), (\tilde{x}, \tilde{\xi}) \in (X \cap \mathrm{dom}\, \phi_P) \times \Xi$ *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

$$
\min\{b^\top y : Wy = a, y \ge 0\},
\tag{4.10}
$$

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

$$
\mathrm{dom}(\mathrm{val}) = \bigcup_{j=1}^{\ell} \mathcal{K}_j \quad \text{and} \quad \mathrm{val}(a, b) = (C_j a)^\top b \quad \text{if } (a, b) \in \mathcal{K}_j.
$$

Furthermore, it holds $\mathrm{int}\, \mathcal{K}_j \ne \emptyset$ and $\mathcal{K}_j \cap \mathcal{K}_i = \emptyset$, $i \ne j$, $i, j = 1, \dots, \ell$. Hence, val satisfies the following continuity property on its domain

$$
|\mathrm{val}(a, b) - \mathrm{val}(\tilde{a}, \tilde{b})| \le L(\max\{1, \|b\|, \|\tilde{b}\|\} \|a - \tilde{a}\| + \max\{1, \|a\|, \|\tilde{a}\|\} \|b - \tilde{b}\|)
$$

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

$$
\begin{aligned}
|v(x, \xi) - v(\tilde{x}, \tilde{\xi})| \le{} & |v(x, \xi) - v(\tilde{x}, \xi)| + |v(\tilde{x}, \xi) - v(\tilde{x}, \tilde{\xi})| \\
\le{} & |\mathrm{val}(h(\xi) - T(\xi)x, q(\xi)) - \mathrm{val}(h(\xi) - T(\xi)\tilde{x}, q(\xi))| \\
& + |\mathrm{val}(h(\xi) - T(\xi)\tilde{x}, q(\xi)) - \mathrm{val}(h(\tilde{\xi}) - T(\tilde{\xi})\tilde{x}, q(\tilde{\xi}))| \\
\le{} & L(\max\{1, \|q(\xi)\|, \|q(\tilde{\xi})\|\} \|T(\xi)(x - \tilde{x})\| \\
& + \max\{1, \|h(\xi) - T(\xi)\tilde{x}\|, \|h(\tilde{\xi}) - T(\tilde{\xi})\tilde{x}\|\} \|q(\xi) - q(\tilde{\xi})\|)
\end{aligned}
$$

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

For each $x \in \operatorname{dom} \phi_P$ it follows from [38, Proposition 2.8] that

$$\partial \phi_P(x) = -\mathbb{E}_P[T(\xi)^\top D(x,\xi)] + \mathcal{N}_{\operatorname{dom} \phi_P}(x), \tag{4.11}$$

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

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

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

$$0 \in \mathbb{E}_P[c - T(\xi)^\top D(x,\xi)] + \mathcal{N}_{\mathcal{X} \cap \operatorname{dom} \phi_P}(x). \tag{4.12}$$

Here, $\mathcal{N}_{X \cap \operatorname{dom} \phi_P}(x)$ denotes the normal cone to the polyhedral set $X \cap \operatorname{dom} \phi_P$.

Proof. Lemma 4.3 implies that $\mathbb{E}_P[v(x,\xi)]$ is finite for every $x \in X \cap \operatorname{dom} \phi_P$. Hence, the existence follows from Weierstrass theorem and the first order optimality condition from [35, Theorem 8.15]. ∎

The polyhedral set $\operatorname{dom} \phi_P$ may contain some induced constraints. If one assumes *relatively complete recourse*, i.e., $X \subset \operatorname{dom} \phi_P$, the optimality condition (4.12) coincides with the one in [38, Theorem 2.11]. Our interest here is to apply the stability results of stochastic generalized equations 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 Γ given by

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

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

$$\Gamma(\tilde{x},\tilde{\xi}) \subseteq \Gamma(x,\xi) + \hat{L}(\max\{1, \|\xi\|, \|\tilde{\xi}\|\}^3 \|\tilde{x} - x\| + \max\{1, \|x\|, \|\tilde{x}\|\} \max\{1, \|\xi\|, \|\tilde{\xi}\|\}^2 \|\tilde{\xi} - \xi\|)\mathcal{B}$$

for all pairs $(x,\xi), (\tilde{x},\tilde{\xi}) \in (X \cap \operatorname{dom} \phi_P) \times \Xi$. Here, $\mathcal{B}$ 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 semi-continuity behavior of the solution set S is very similar to that of v (see (4.9)), namely,

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

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

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

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) $\mathcal{M}(q(\xi)) = \{\pi : W^\top \pi \leq q(\xi)\}$ is nonempty and bounded for every $\xi \in \Xi$, (c) P has finite second order moments, i.e., $\mathbb{E}_P[\|\xi\|^2] < +\infty$ and (d) X is nonempty and bounded. Then it holds for any probability measure Q such that $\mathscr{D}(Q, P)$ is sufficiently small*

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

where the function R is defined by

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

and the distance $\mathscr{D}$ is defined in Section 3.

Proof: We intend to apply Theorem 3.1 to the stochastic generalized equations

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

and check the corresponding assumptions. The set-valued mapping Γ 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

$$\kappa(\xi) = \|c\| + \|T(\xi)^\top\|\mu(\xi) \quad \text{where} \quad \|\pi\| \leq \mu(\xi) \ \forall \pi \in \mathcal{M}(q(\xi)) \tag{4.13}$$

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

$$d(\pi, \mathcal{M}(q(\bar{\xi}))) \leq L_M \|\xi - \bar{\xi}\|.$$

Hence, there exists $\bar{\pi} \in \mathcal{M}(q(\bar{\xi}))$ such that $\|\pi\| \leq \|\bar{\pi}\| + L_M \|\xi - \bar{\xi}\|$. Since $\mathcal{M}(q(\bar{\xi}))$ is bounded, there exists a constant $\bar{C}$ such that, we may choose the function μ as $\mu(\xi) = \bar{C} + L_M(\|\xi\| + \|\bar{\xi}\|)$. We conclude that the function κ given by (4.13) depends on $\|\xi\|$ at most quadratically. Hence, κ is P-integrable according to assumption (c). Finally, we note that the normal cone mapping $\mathcal{N}_X$ 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 $\mathscr{D}$ 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 $\mathscr{D}$.

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

$$\mathscr{F} = \{g : \Xi \to \mathbb{R} : g(x) - g(\tilde{\xi}) \leq C \max\{1, \|\xi\|, \|\tilde{\xi}\|\}^2 \|\xi - \tilde{\xi}\|, \forall \xi, \tilde{\xi} \in \Xi\}$$

for some constant $C > 0$. Consequently, the estimate

$$\mathscr{D}(P, Q) \leq C\zeta_3(P, Q)$$

holds, where ζ_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 $\xi, \tilde{\xi} \in \Xi$. We consider $g(\xi) = \sigma(\Gamma(x, \xi), u)$ and know from Proposition 4.5 that

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

Since X is bounded, we may choose the constant C such that $\hat{L} \max\{1, \|x\|, \|\tilde{x}\|\} \leq C$ for all $x \in X$. ∎

Since the distance ζ_3 is slightly stronger than the second order Fortet-Mourier metric ζ_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 SMPEC. In this subsection, we consider application of the stability analysis established in Section 3 to a two stage stochastic mathematical program with complementarity constraints (SMPCC) defined as follows:

$$
\begin{aligned}
\min_{x,\,y(\cdot)\in\mathscr{Y}} \quad & \mathbb{E}_P[f(x,y(\omega),\xi(\omega))] \\
\text{s.t.} \quad & x \in X \text{ and for almost every } \omega \in \Omega: \\
& g(x,y(\omega),\xi(\omega)) \le 0, \\
& h(x,y(\omega),\xi(\omega)) = 0, \\
& 0 \le G(x,y(\omega),\xi(\omega)) \perp H(x,y(\omega),\xi(\omega)) \ge 0,
\end{aligned}
\tag{4.14}
$$

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 $\mathbb{R}, \mathbb{R}^s, \mathbb{R}^r, \mathbb{R}^m, \mathbb{R}^m$, respectively, $\xi : \Omega \to \Xi$ is a vector of random variables defined on probability $(\Omega, \mathcal{F}, P)$ with compact support set $\Xi \subset \mathbb{R}^q$, and $\mathbb{E}_P[\cdot]$ denotes the expected value with respect to probability measure P, and '$\perp$' denotes the perpendicularity of two vectors, $\mathscr{Y}$ 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 NLP regularization. Our interest here is in a *direct* stability analysis on the stationary point of the problem using the stochastic generalized equations scheme discussed in section 3.

Observe first that problem (4.14) can be written as

$$
\mathrm{P}_\vartheta : \qquad \min_x \quad \vartheta(x) = \mathbb{E}_P[v(x,\xi(\omega))] \\
\text{s.t.} \quad x \in X,
$$

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:

$$
\mathrm{MPCC}(x,\xi) : \qquad \min_y \quad f(x,y,\xi) \\
\begin{aligned}
\text{s.t.} \quad & g(x,y,\xi) \le 0, \\
& h(x,y,\xi) = 0, \\
& 0 \le G(x,y,\xi) \perp H(x,y,\xi) \ge 0.
\end{aligned}
$$

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 SMPECs.

Define the Lagrangian function of the second stage problem $\mathrm{MPCC}(x,\xi)$:

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

We consider the following KKT conditions of $\mathrm{MPCC}(x,\xi)$:

$$
\begin{cases}
0 = \nabla_y \mathcal{L}(x,y,\xi;\alpha,\beta,u,v), \\
y \in \mathcal{F}(x,\xi), \\
0 \le \alpha \perp -g(x,y,\xi) \ge 0, \\
0 = u_i, & i \notin \mathcal{I}_G(x,y,\xi), \\
0 = v_i, & i \notin \mathcal{I}_H(x,y,\xi), \\
0 \le u_i v_i, & i \in \mathcal{I}_G(x,y,\xi) \cap \mathcal{I}_H(x,y,\xi),
\end{cases}
$$

where $\mathcal{F}(x,\xi)$ denotes the feasible set of $\mathrm{MPCC}(x,\xi)$ and

$$
\begin{aligned}
\mathcal{I}_G(x,y,\xi) &:= \{i \,|\, G_i(x,y,\xi) = 0, \, i = 1, \cdots, m\}, \\
\mathcal{I}_H(x,y,\xi) &:= \{i \,|\, H_i(x,y,\xi) = 0, \, i = 1, \cdots, m\}.
\end{aligned}
$$

Let $\mathcal{W}(x,\xi)$ denote the set of KKT pairs $(y;\alpha,\beta,u,v)$ satisfying the above conditions for given (x,ξ) and $S(x,\xi)$ the corresponding set of stationary points, that is, $S(x,\xi) = \Pi_y \mathcal{W}(x,\xi)$. For each $(y;\alpha,\beta,u,v)$, y is a *C-stationary point* of problem $\mathrm{MPCC}(x,\xi)$ and (α,β,u,v) the corresponding Lagrange multipliers. When the stationary points are restricted to global minimizers, we denote the set of KKT pairs by $\mathcal{W}^*(x,\xi)$, i.e.,

$\mathcal{W}^*(x,\xi) = \{(y; \alpha, \beta, u, v) \in \mathcal{W}(x,\xi), y \in Y_{sol}(x,\xi)\}$, where $Y_{sol}(x,\xi)$ denotes the set of optimal solutions of MPCC(x,ξ).

Let (x^*,ξ) be fixed. Recall that MPCC(x^*,ξ) is said to satisfy *MPEC-Mangasarian-Fromowitz Constraint Qualification* (MPEC-MFCQ for short) at a feasible point y^* if the gradient vectors

$$\{\nabla_y h_i(x^*, y^*, \xi)\}_{i=1,\cdots,r}; \{\nabla G_i(x^*, y^*, \xi)\}_{i \in \mathcal{I}_G(x^*,y^*,\xi)}; \{\nabla H_i(x^*, y^*, \xi)\}_{i \in \mathcal{I}_H(x^*,y^*,\xi)}$$

are linearly independent and there exists a vector $d \in R^n$ perpendicular to the vectors such that

$$\nabla g_i(x^*, y^*, \xi)^\top d < 0, \qquad \forall i \in \mathcal{I}_g(x^*, y^*, \xi),$$

where

$$\mathcal{I}_g(x^*, y^*, \xi) := \{i \,|\, g_i(x^*, y^*, \xi) = 0, \, i = 1, \cdots, s\}.$$

The following results are derived in [23].

PROPOSITION 4.8. *Let $x^* \in X$. Suppose that there exist constants δ, $t^* > 0$, a compact set $Y \subset \mathbb{R}^m$ and a neighborhood U of x^* such that*

$$\emptyset \neq \{y : f(x,y,\xi) \leq \delta \text{ and } y \in \mathcal{F}(x,\xi)\} \subset Y,$$

for all $(x,\xi) \in U \times \Xi$. Suppose also that problem MPCC(x^*,ξ) satisfies MPEC-MFCQ at every point y in solution set of MPCC(x^*,ξ), denoted by $Y_{sol}(x^*,\xi)$. 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 semi-continuous on $U \times \Xi$, where

$$\Phi(x,\xi) := \text{conv}\left\{\bigcup_{(y;\alpha,\beta,u,v)\in\mathcal{W}^*(x,\xi)} \nabla_x \mathcal{L}(x,y,\xi;\alpha,\beta,u,v)\right\}.$$

Using $\partial_x v(x,\xi)$ and $\Phi(x,\xi)$, we can define the weak KKT conditions of problem (4.14)

$$0 \in \mathbb{E}_P[\partial_x v(x,\xi)] + \mathcal{N}_X(x) \tag{4.15}$$

and its relaxation

$$0 \in \mathbb{E}_P[\Phi(x,\xi)] + \mathcal{N}_X(x). \tag{4.16}$$

Both of the systems are stochastic generalized equations. 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:

$$0 \in \mathbb{E}_Q[\partial_x v(x,\xi)] + \mathcal{N}_X(x) \tag{4.17}$$

and

$$0 \in \mathbb{E}_Q[\Phi(x,\xi)] + \mathcal{N}_X(x), \tag{4.18}$$

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*

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

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

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

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

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

REMARK 4.10. The key condition in the conclusion (i) of Theorem 4.1 is the Lipschiz 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 non-decreasing continuous function h. To ease the technical details, let us consider a special case of MPCC(x,ξ)

$$\text{MPCC}'(x,\xi): \quad \min_{y} \quad f(x,y,\xi)$$
$$\text{s.t.} \quad 0 \leq y \perp H(x,y,\xi) \geq 0,$$

where H is uniformly strongly monotone w.r.t. y, that is, there exists a positive constant $C_1 > 0$ such that $\|\nabla_y H(x,y,\xi)^{-1}\| \leq C_1$ for all x, y, ξ. 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 ξ. Moreover, if H is uniformly globally Lipschitz continuous w.r.t. ξ, then $y(x,\xi)$ is also uniformly globally Lipschitz continuous w.r.t. ξ. Assuming that $f(x,y,\xi)$ is continuously differentiable and uniformly globally Lipschitz continuous w.r.t. y and ξ, then

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

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

$$\begin{aligned}
|v(x,\xi) - v(x,\xi')| &= |f(x, y(x,\xi), \xi) - f(x, y(x,\xi'), \xi')| \\
&\leq L_2(\|y(x,\xi) - y(x,\xi')\| + \|\xi - \xi'\|) \\
&\leq L_2(L_1 + 1)\|\xi - \xi'\|.
\end{aligned}$$

Let $L := 2L_2(L_1 + 1)$ and $h(t) := Lt$. Then

$$\sup_{x \in X} \sup_{\tau \in (0,\delta)} \sup_{\|u\| \leq 1} \left| \frac{1}{\tau}(v(x+\tau u, \xi) - v(x,\xi)) - \frac{1}{\tau}(v(x+\tau u, \hat{\xi}) - v(x,\hat{\xi})) \right| \leq h(\|\xi - \hat{\xi}\|)$$

which means (4.5).

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

$$\begin{aligned}
\min_{x} \quad & \mathbb{E}_P[f(x,\xi)] \\
\text{s.t.} \quad & \mathbb{E}_P[(\eta - G(x,\xi))_+] \leq \mathbb{E}_P[(\eta - Y(\xi))_+], \ \forall \eta \in [a,b], \\
& x \in X,
\end{aligned} \quad (5.1)$$

where X is a closed convex subset in $\mathbb{R}^n$, $f, G : \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}$ 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 $(\Omega, \mathcal{F}, P)$ with support set $\Xi \subset \mathbb{R}^q$, $\mathbb{E}_P[\cdot]$ denotes 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 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 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 following perturbed problem

$$\begin{aligned} \min_{x} \quad & \mathbb{E}_Q[f(x,\xi)] \\ \text{s.t.} \quad & \mathbb{E}_Q[(\eta - G(x,\xi))_+] \leq \mathbb{E}_Q[(\eta - Y(\xi))_+], \ \forall \eta \in [a,b], \\ & x \in X, \end{aligned} \tag{5.2}$$

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

For the simplicity of notation, let

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

It is easy to observe: (a) $H(x,\eta,\xi)$ is globally Lipschitz continuous in η uniformly w.r.t. x and ξ, (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 Bouligrand 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:

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

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

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

$$\sum_{\zeta \in \partial_x \mathbb{E}_P[H(x,\eta,\xi)]} \zeta^\top (x_s - x_0) \leq -\delta$$

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 $\partial_x \mathbb{E}_P[H(x,\eta,\xi)]$ 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 $\mu^* \in \mathcal{M}_+([a,b])$ such that (x^*, μ^*) satisfies the following:

$$\begin{cases} 0 \in \nabla \mathbb{E}_P[f(x,\xi)] + \int_a^b \mathbb{E}_P[\partial_x H(x,\eta,\xi)]\mu(d\eta) + \mathcal{N}_X(x), \\ \mathbb{E}_P[H(x,\eta,\xi)] \leq 0, \forall \eta \in [a,b], \\ \int_a^b \mathbb{E}_P[H(x,\eta,\xi)]\mu(d\eta) = 0, \end{cases} \tag{5.3}$$

where $\mathcal{M}_+([a,b])$ is the set of positive measures in the the space of regular countably additive measures on $[a,b]$ having finite variation, see [4, Example 2.63], [11] the references therein; $\partial_x H(x,\eta,\xi)$ denotes the

Clarke subdifferential of H w.r.t. x, that is,

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

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

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

$$\begin{cases} 0 \in \nabla \mathbb{E}_Q[f(x,\xi)] + \int_a^b \mathbb{E}_Q[\partial_x H(x,\eta,\xi)]\mu(d\eta) + \mathcal{N}_X(x), \\ \mathbb{E}_Q[H(x,\eta,\xi)] \leq 0, \forall \eta \in [a,b], \\ \int_a^b \mathbb{E}_Q[H(x,\eta,\xi)]\mu(d\eta) = 0. \end{cases} \tag{5.4}$$

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 stochastic generalized equations so that we can apply Theorem 3.1. Since $G(x,\xi)$ is Lipschitz continuous in (x,ξ) and the modulus in x is bounded by a positive constant L_1, $H(x,\eta,\xi)$ is Lipschitz continuous in (x,η,ξ). Then by [45, Proposition 2.1], $\partial_x H(x,\eta,\xi)$ is measurable with respect to η, ξ. Moreover $\partial_x H(x,\eta,\xi)$ is bounded by L_1. By invoking Proposition 2.6, we have

$$\int_a^b \mathbb{E}_P[\partial_x H(x,\eta,\xi)]\mu(d\eta) = \mathbb{E}_P\left[\int_a^b \partial_x H(x,\eta,\xi)\mu(d\eta)\right].$$

Let

$$\Gamma(x,\mu,\xi) := \begin{pmatrix} \nabla_x f(x,\xi) + \int_a^b \partial_x H(x,\eta,\xi)\mu(d\eta) \\ H(x,\eta,\xi) : \eta \in [a,b] \\ \int_a^b H(x,\eta,\xi)\mu(d\eta) \end{pmatrix}$$

and

$$\mathcal{G}(x,\mu) := \begin{pmatrix} \mathcal{N}_X(x) \\ \mathscr{C}_+([a,b]) \\ 0 \end{pmatrix}. \tag{5.5}$$

To simplify the notation, let $z := (x,\mu)$. Then we can reformulate the KKT conditions (5.3) as the following stochastic generalized equations

$$0 \in \mathbb{E}_P[\Gamma(z,\xi)] + \mathcal{G}(z), \tag{5.6}$$

where the norm in space $\mathscr{C}([a,b])$ is $\|\cdot\|_\infty$. Obviously (5.6) falls into the framework of the stochastic generalized equations (1.1). Likewise, we can reformulate the KKT conditions (5.4) as the stochastic generalized equations

$$0 \in \mathbb{E}_Q[\Gamma(z,\xi)] + \mathcal{G}(z). \tag{5.7}$$

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

We need to introduce some new notation. Let Z denote a compact subset of $X \times \mathscr{M}_+([a,b])$,

$$\mathscr{F} := \{g(\xi) : g(\xi) := \sigma(\Gamma(z,\xi), u), \text{ for } z \in Z, \|u\| \leq 1\}.$$

Let

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

and

$$\mathscr{H}_S(Q,P) := \max\big(D_S(Q,P), \mathscr{D}_S(P,Q)\big).$$

Let $\tilde{S}(P)$ and $\tilde{S}(Q)$ 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 $S(P) := \tilde{S}(P) \cap Z$ and $S(Q) := \tilde{S}(Q) \cap Z$.

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 stochastic generalized equations (5.6) and its perturbation (5.7). Assume: (a') $G(x,\xi)$ is Lipschitz continuous in x for every ξ with modulus L_1 (independent of x and ξ), (b') $|G(x,\xi)|$ is bounded by a positive constant L_2 (independent of x and ξ), (c') $f(x,\xi)$ is Lipschitz continuous in x for every ξ and the Lipschitz modulus is bounded by an integrable function $\kappa(\xi)$, (d') $S(P)$ and $S(Q)$ are nonempty. Then the conclusions (i)-(iii) of Theorem 3.1 hold for $S(P)$ and $S(Q)$.*

Proof. The thrust of the proof is to apply Theorem 3.1 to generalized equations (5.6) and its perturbation (5.7), taking into account Remark 3.4 as the single valued components of Γ 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 semi-continuous, while (d) coincides with (d'). Therefore we are left to verify (a) and (b).

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

In what follows, we verify (b), that is, upper semi-continuity of $\Gamma(x,\mu,\xi)$ with respect to (x,μ) and its integrable boundedness. Let us look into the third component $\int_a^b H(x,\eta,\xi)\mu(d\eta)$. 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,μ).

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, $\nabla_x f(x,\xi) + \int_a^b \partial_x H(x,\eta,\xi)\mu(d\eta)$. Since f is assumed to be continuously differentiable, it suffices to verify the upper semi-continuity of $\int_a^b \partial_x H(x,\eta,\xi)\mu(d\eta)$ w.r.t. (x,μ). Using property of $\mathbb{D}$, we have

$$\mathbb{D}\left(\int_a^b \partial_x H(x',\eta,\xi)\mu'(d\eta), \int_a^b \partial_x H(x,\eta,\xi)\mu(d\eta)\right)$$
$$\leq \mathbb{D}\left(\int_a^b \partial_x H(x',\eta,\xi)\mu'(d\eta), \int_a^b \partial_x H(x,\eta,\xi)\mu'(d\eta)\right)$$
$$+ \mathbb{D}\left(\int_a^b \partial_x H(x,\eta,\xi)\mu'(d\eta), \int_a^b \partial_x H(x,\eta,\xi)\mu(d\eta)\right).$$

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

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

It is easy to verify that $\partial_x H(x', \eta, \xi)$ is upper semi-continuous in x for every fixed η and ξ and it is bounded by $\|\nabla_x G(x, \xi)\|$ which is integrably bounded by assumption.

By [3, Corollary 5.2],

$$\overline{\lim_{x' \to x}} \int_a^b \partial_x H(x', \eta, \xi) \mu'(d\eta) \subseteq \int_a^b \partial_x H(x, \eta, \xi) \mu'(d\eta)$$

which implies

$$\overline{\lim_{x' \to x}} \, \sigma\left(\int_a^b \partial_x H(x', \eta, \xi) \mu'(d\eta), u \right) \leq \sigma\left(\int_a^b \partial_x H(x, \eta, \xi) \mu'(d\eta), u \right)$$

for any u with $\|u\| \leq 1$. Through [27, Proposition 3.4], the latter inequality can be written as

$$\overline{\lim_{x' \to x}} \int_a^b \sigma(\partial_x H(x', \eta, \xi), u) \mu'(d\eta) \leq \int_a^b \sigma(\partial_x H(x, \eta, \xi), u) \mu'(d\eta). \tag{5.8}$$

Let $x_k \to x$ and u_k be such that $\|u_k\| \leq 1$ and

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

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 from

$$\overline{\lim_{k \to \infty}} \int_a^b [\sigma(\partial_x H(x_k, \eta, \xi), u_k) - \sigma(\partial_x H(x, \eta, \xi), u_k)] \mu'(d\eta) \leq 0.$$

Since x_k is arbitrary, this implies

$$\overline{\lim_{x' \to x}} \sup_{\|u\| \leq 1} \left(\int_a^b [\sigma(\partial_x H(x', \eta, \xi), u) - \sigma(\partial_x H(x, \eta, \xi), u)] \mu'(d\eta) \right) \leq 0.$$

On the other hand, it follows by [23, Lemma 5.1]

$$\mathbb{D}\left(\int_a^b \partial_x H(x, \eta, \xi) \mu'(d\eta), \int_a^b \partial_x H(x, \eta, \xi) \mu(d\eta) \right) \to 0$$

as $\mu' \to \mu$. The discussions above show that $\int_a^b \partial_x H(x, \eta, \xi) \mu(d\eta)$ is upper semi-continuous w.r.t. (x, μ).

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 $\int_a^b H(x, \eta, \xi) \mu(d\eta)$. 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].

REFERENCES

[1] Z. ARTSTEIN AND R.-J. B. WETS, *Approximating the integral of a multifunction*, J. Multivariate Anal., 24 (1988), pp. 285-308.

[2] J.-P. AUBIN AND H. FRANKOWSKA, *Set-Valued Analysis*, Birkhauser, Boston, 1990.

[3] R. J. AUMANN, *Integrals of set-valued functions*, J. Math. Anal. Appl., 12 (1965), pp. 1-12.

[4] J. F. BONNANS AND A. SHAPIRO, *Perturbation Analysis of Optimization Problems*, Springer Series in Operations Research, Springer-Verlag, 2000.

[5] F. E. BROWDER, *Multivalued monotone nonlinear mappings and duality mappings in Banach spaces*, Trans. Amer. Math. Soc., 118 (1965), pp. 338-351.

[6] J. M. BORWEIN AND Q. J. ZHU, *Techniques of Variational Analysis*, Springer, New York, 2005.

[7] C. CASTAING AND M. VALADIER, *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Mathematics, Vol 580, Springer, Berlin, 1977.

[8] X. CHEN AND M. FUKUSHIMA, *Expected residual minimization method for stochastic linear complementarity Problems*, Math. Oper. Res., 30 (2005), pp. 1022-1038.

[9] X. CHEN, R. B-J WETS AND Y. ZHANG, *Stochastic variational inequalities: residual minimization smoothing/sample average approximations*, SIAM J. Optim., 22 (2012), pp. 649-673.

[10] F. H. CLARKE, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[11] D. DENTCHEVA AND A. RUSZCZYŃSKI, *Optimization with stochastic dominance constraints*, SIAM J. Optim., 14 (2003), pp. 548-566.

[12] D. DENTCHEVA AND A. RUSZCZYŃSKI, *Optimality and duality theory for stochastic optimization with nonlinear dominance constraints*, Math. Program., 99 (2004), pp. 329-350.

[13] D. DENTCHEVA AND A. RUSZCZYŃSKI, *Composite semi-infinite optimization*, Control Cybernet., 36 (2007), pp. 1-14.

[14] A. L. DONTCHEV, A. S. LEWIS AND R. T. ROCKAFELLAR, *The radius of metric regularity*, Trans. Amer. Math. Soc., 355 (2003), pp. 493-517.

[15] A. L. DONTCHEV AND R. T. ROCKAFELLAR *Implicit Functions and Solution Mappings: A View from Variational Analysis*, Springer, 2009.

[16] F. FACCHINEI AND J-S. PANG, *Finite-dimensional Variational Inequalities and Complementarity Problems*, Springer, 2003.

[17] C. HESS, *Set-valued integration and set-valued probability theory: an overview*, Handbook of measure theory, Vol. I, II, 617-673, North-Holland, Amsterdam, 2002.

[18] F. HIAI, *Convergence of conditional expectations and strong laws of large numbers for multivalued random variables*, Trans. Amer. Math. Soc., 291 (1985), pp. 603-627.

[19] A. F. IZMAILOV, A. S. KURENNOY AND M. V. SOLODOV, *A note on upper Lipschitz stability, error bounds, and critical multipliers for Lipschitz-continuous KKT systems*, to appear in Mathematical Programming Series A.

[20] A. J. KING AND R. T. ROCKAFELLAR, *Sensitivity analysis for nonsmooth generalized equations*, Math. Program., 55 (1992), pp. 191-212.

[21] B. KUMMER, *Generalized equations: solvability and regularity*, Math. Program. Stud., 21 (1984), pp. 199-212.

[22] Y. LIU AND H. XU, *Stability and sensitivity analysis of stochastic programs with second order dominance constraints*, to appear in Mathematical Programming Series A.

[23] Y. LIU, H. XU AND G.H. LIN, *Stability analysis of two stage stochastic mathematical programs with complementarity constraints via NLP-regularization*, SIAM J. Optim., 21 (2011), pp. 609-705.

[24] I. MOLCHANOV, *Theory of Random Sets*, in Probability and Its Applications, J. Gani et al eds. Springer, 2005.

[25] F. NOŽIČKA, J. GUDDAT, H. HOLLATZ AND B. BANK, *Theorie der Linearen Parametrischen Optimierung* (in German), Akademie-Verlag, Berlin, 1974.

[26] J. OUTRATA, M. KOCVARA AND J. ZOWE, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results*, Kluwer Academic Publishers, Boston, 1998.

[27] N. PAPAGEORGIOU, *On the theory of Banach space valued multifunctions 1. Integration and conditional expectation*, J. Multivariate Anal., 17 (1985), pp. 185-206.

[28] S. T. RACHEV, *Probability Metrics and the Stability of Stochastic Models*, Wiley, Chichester, 1991.

[29] S. T. RACHEV AND W. RÖMISCH, *Quantitative stability in stochastic programming: The method of probability metrics*, Math. Oper. Res., 27 (2002), pp. 792-818.

[30] D. RALPH AND H. XU, *Asympototic analysis of stationary points of sample average two stage stochastic programs: A generalized equations approach*, Math. Oper. Res., 36 (2011), pp. 568-592.

[31] U. RAVAT AND U.V. SHANBHAG, *On the characterization of solution sets of smooth and nonsmooth convex stochastic Nash games*, SIAM J. Optim., 21 (2011), pp. 1168-1199.

[32] S. M. ROBINSON, *Regularity and stability for convex multivalued functions*, Math. Oper. Res., 1 (1976), pp. 131-143.

[33] S. M. ROBINSON, *Generalized Equations*. In A. Bachem, M. Grötschel, and B. Korte, editors, Mathematical Programming: The State of the Art, Speinger-Verlag, Berlin, pp. 346-367, 1983.

[34] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[35] R. T. ROCKAFELLAR AND R. J.-B. WETS, *Variational Analysis*, Springer-Verlag, Berlin, 1998.

[36] W. RÖMISCH, *Stability of Stochastic Programming Problems, in: Stochastic Programming* (A. Ruszczynski and A. Shapiro eds.), Handbooks in Operations Research and Management Science Vol. 10, Elsevier, Amsterdam, pp. 483-554, 2003.

[37] A. SHAPIRO, *Monte Carlo Sampling Methods, in: Stochastic Programming* (A. Ruszczynski and A. Shapiro eds.), Handbooks in Operations Research and Management Science Vol. 10, Elsevier, Amsterdam, pp. 353-425, 2003.

[38] A. SHAPIRO, D. DENTCHEVA AND A. RUSCZYŃSKI, *Lecture Notes on Stochastic Programming: Modelling and Theory*, MPS-SIAM Series on Optimization, Philadelphia, 2009.

[39] A. Shapiro and H. Xu, *Stochastic mathematical programs with equilibrium constraints, Modeling and sample average approximation*, Optimization, 57 (2008), pp. 395-418.

[40] H. Sun and H. Xu, *Convergence analysis of stationary points in sample average approximation of stochastic programs with second order stochastic dominance constraints*, Math. Prog. Series A, 2013.

[41] F. Topsoe, *On the connection between P-continuity and P-uniformity in weak convergence*, Theory Probab. Appl., 12 (1967), pp. 281-290.

[42] D. Walkup and R. J-B. Wets, *Lifting projections of convex polyhedra*, Pac. J. Math., 28 (1969), pp. 465-475.

[43] R. J-B. Wets, *Stochastic programs with fixed recourse: The equivalent deterministic program*, SIAM Rev., 16 (1974), pp. 309-339.

[44] H. Xu, *An implicit programming approach for a class of stochastic mathematical programs with complementarity constraints*, SIAM J. Optim., 16 (2006), pp. 670-696.

[45] H. Xu, *Uniform exponential convergence of sample average random functions under general sampling with applications in stochastic programming*, J. Math. Anal. Appl., 368 (2010), pp. 692-710.

[46] H. Xu, *Sample average approximation methods for a class of stochastic variational inequality problems*, Asian-Pac. J. Oper. Res., 27 (2010), pp. 103-119.

[47] D. Zhang and C. Guo, *Fubini theorem for F-valued integrals*, Fuzzy Sets and Systems, 62 (1994), pp. 355-358.

[48] V. M. Zolotarev, *Probability metrics*, Theory Probab. Appl., 28 (1983), pp. 278-302.