# Quantitative Stability of Two Stage Linear Second-order Conic Stochastic Programs with Full Random Recourse 

Qingsong Duan, Mengwei $\mathrm{Xu}^{\dagger}$, Shaoyan Guo, ${ }^{\ddagger}$ and Liwei Zhang ${ }^{\S}$

October 7, 2017


#### Abstract

In this paper, we consider quantitative stability for full random two stage linear stochastic program with second-order conic constraints when the underlying probability distribution is subjected to perturbation. We first investigate locally Lipschitz continuity of feasible set mappings of the primal and dual problems in the sense of Hausdorff distance which derives the Lipschitz continuity of the objective function, and then establish the quantitative stability results of the optimal value function and the optimal solution mapping for the perturbation problem. Finally, the obtained results are applied to the convergence analysis of optimal values and solution sets for empirical approximations of the stochastic problems.


Key words: second order conic optimization, optimal value function, solution mapping, quantitative stability.

2010 MSC: 90C15, 90C31.

## 1 Introduction

Consider the following stochastic programming:

$$
\begin{equation*}
\min \left\{\int_{\Xi} f_{0}(x, \xi) d P(\xi): x \in X, \int_{\Xi} f_{j}(x, \xi) d P(\xi) \leq 0, j=1, \cdots, d\right\} \tag{1.1}
\end{equation*}
$$

[^0]where $X \subseteq \Re^{n}$ and $\Xi \subseteq \Re^{s}$ are closed set, $f_{j}$ mapping from $\Re^{n} \times \Xi$ to $\bar{\Re}$ are normal integrands for $j=0, \cdots, d, \xi: \Omega \rightarrow \Xi$ is a random variable and $P$ is Borel probability measure on $\Xi$.

Qualitative and quantitative stability properties for the optimal value function and optimal solution mapping of stochastic programs play an important role in both theoretical and numerical points of view. For problem (1.1), the authors in [9] studied the upper semi-continuity of local optimal solution mapping and Lipschitz continuity of optimal value function which were also applied to specific models: chance-constrained, linear two stage and mixed-integer two stage models. With weak convergence of probability measures, Schultz [16] discussed stability of chance-constrained model with linear equality constraint and integer variables. First-order and second-order directional differentiability properties of the optimal value function and optimal solution mapping were implied in [2]. Römisch and Schultz investigated the quantitative stability of the optimal value function under different probability distances [12] and established Hölder continuity of optimal solution mappings under the Hausdorff distance in [11]. For two stage stochastic programs, the authors [13] proved the Lipschitz continuity to the optimal solution mapping by introducing the subgradient distance. Shapiro [17] showed that the upper bound for the rate of convergence implies the upper Lipschitz continuity of the optimal solution mapping for two stage program respect to the Kolmogorov-Smirnov distance.

For a full random stochastic program, the quantitative stability of the optimal value function and the optimal solution mapping was investigated under epi-convergence framework in [15]. The authors obtained the Lipschitz continuity of the optimal value function and the $\varepsilon$ approximated solution mapping under the Hausdorff distance with respect to Fortet-Mourier metric of probability distributions. They also applied these results to linear two stage programs. The authors [7] derived stability results for full random linear two stage stochastic programs with recourse.

In this paper, we consider a two stage second-order conic stochastic optimization problem as follows:

$$
\begin{equation*}
\min \mathbb{E}_{P}\left[f_{0}(x, \xi)\right] \tag{1.2}
\end{equation*}
$$

where $f_{0}(x, \xi):=d^{T} x+\theta(x, \xi), d \in \Re^{n}$ and $\theta(x, \xi)$ is the optimal value function of the second stage problem:

$$
\begin{align*}
\mathrm{P}(x, \xi) \quad \min _{y \in \Re^{m}} & c^{T} y  \tag{1.3}\\
\text { s.t. } & a_{i}^{T} y+q_{i}^{T} x-b_{i} \geq\left\|B^{i} y\right\|_{2}, \quad i=1, \ldots, l,
\end{align*}
$$

where $\xi=(c ; A ; Q ; B ; b)$ is a random vector with $c: \Omega \rightarrow \Re^{m}, A=\left(a_{1}, \cdots, a_{l}\right)^{T}: \Omega \rightarrow$ $\Re^{l \times m}, Q=\left(q_{1}, \cdots, q_{l}\right)^{T}: \Omega \rightarrow \Re^{l \times n}, b: \Omega \rightarrow \Re^{l}, B=\left(B^{1} ; \ldots ; B^{l}\right)$ with $B^{i}: \Omega \rightarrow \Re^{J \times m}, i=$
$1, \ldots, l$. For convenience, we define the norm of $\xi$ by $\|\xi\|=\|c\|+\|A\|+\sum_{j=1}^{J}\left\|B_{j}\right\|+\|Q\|+\|b\|$, where $\|\cdot\|$ denotes the 1-norm for a vector, that is $\|c\|_{1}:=\sum_{i}\left|c_{i}\right|$ or infinity-norm of a matrix, that is $\|A\|_{\infty}:=\max \left\{\left\|a_{i}\right\|_{1}, i=1, \cdots, l\right\}$.

Let $g^{i}(x, y ; \xi):=\left(B^{i} y, a_{i}^{T} y+q_{i}^{T} x-b_{i}\right), i=1, \ldots, l$ and $\mathcal{Q}_{J+1} \subset \Re^{J+1}$ be the second-order cone defined by

$$
\mathcal{Q}_{J+1}=\left\{(s, t) \in \Re^{J} \times \Re: t \geq\|s\|_{2}\right\} .
$$

Then Problem (1.3) can be reformulated to

$$
\begin{array}{ll}
\min _{y} & c^{T} y  \tag{1.4}\\
\text { s.t. } & g^{i}(x, y ; \xi) \in \mathcal{Q}_{J+1}, \quad i=1, \ldots, l
\end{array}
$$

In Duan et. al. [3], the authors discussed the stability properties of Problem (1.3) when $\xi=(c ; A ; Q ; B ; b)$ is perturbed to $\tilde{\xi}=(\tilde{c} ; \tilde{A} ; \tilde{Q} ; \tilde{B} ; \tilde{b})$, especially the differentiability property of $\theta(\cdot, \cdot)$. They obtained the upper semi-continuity of solution mappings for both the original problem and its Lagrange dual problem. Furthermore, the locally Lipschitz continuity of $\theta$ and its Hadamard directional differentiability at a given point were established.

In this paper, we extend the study about the quantitative stability of two stage linear stochastic programs with linear constraints to that with second-order conic constraints. First, we introduce the dual problem of the second stage problem (1.3) and investigate the locally Lipschitz continuity in the sense of Hausdorff distance of the primal and dual feasible set mappings. The results can be extended to global Lipschitz continuity when a mild condition holds. Then we establish locally Lipschitz continuity of the objective function and the quantitative stability of the optimal value function and optimal solution mapping for the perturbation problem. We give the convergence property of the optimal values and solution sets for empirical approximations of two stage stochastic programs with second-order conic constraints.

The remaining parts of this paper are organized as follows. In section 2, we give some preliminaries which will be used in this paper. In Section 3, we prove the locally Lipschitz continuity property of the feasible sets for primal and dual problem in sense of Hausdorff distance. We also give the quantitative stability results for the optimal value function and the optimal solution mapping of perturbed two stage linear second-order conic stochastic programs. Asymptotic behavior of an empirical approximation problem is given in Section 4. We conclude our paper in Section 5.

## 2 Preliminary

We present kinds of distance and some background materials on perturbation analysis, which will be used in the following section. Detailed discussions on these subjects can be found in [1, 10].

Various kinds of distance are used for Lipschitz continuous property of the optimal solution mapping.

Definition 2.1. The distance from a point $x$ to the set $C$ can be written as

$$
d(x, C):=\inf \{\|x-\tilde{x}\|: \forall \tilde{x} \in C\}
$$

Definition 2.2. The epi-distance of the objective functions can be bounded by some probability semimetric of the form

$$
\begin{equation*}
d_{\mathcal{F}}(P, Q):=\sup _{f \in \mathcal{F}}\left|\int_{\Xi} f(\xi) P(d \xi)-\int_{\Xi} f(\xi) Q(d \xi)\right| \tag{2.1}
\end{equation*}
$$

where $\mathcal{F}$ denotes a class of measurable functions from $\Xi$ to $\bar{\Re}$ and $P, Q$ belongs to $\mathcal{P}_{\mathcal{F}}$ denoted by

$$
\mathcal{P}_{\mathcal{F}}:=\left\{Q \in \mathcal{P}(\Xi): \int_{\Xi} \inf _{x} f_{0}(x, \xi) Q(d \xi)>-\infty, \sup _{x} \int_{\Xi} f_{0}(x, \xi) Q(d \xi)<\infty\right\} .
$$

Actually, the distance in (2.1) is also called Zolotarev's pseudometric or $\zeta$ - structure. An important probability metric for stochastic programs with locally Lipschitz continuous integrands is the $p$-th order Fortet-Mourier metric.

Definition 2.3. [5] The $p$-th order Fortet-Mourier metric $\zeta_{p}(p \geq 1)$ is defined on $\mathcal{P}_{p}(\Xi)$ by

$$
\zeta_{p}(P, Q):=\sup _{f \in \mathcal{F}_{p}(\Xi)}\left|\int_{\Xi} f(\xi)(P-Q)(d \xi)\right|
$$

for $P, Q \in \mathcal{P}_{p}(\Xi):=\left\{Q \in \mathcal{P}(\Xi): \int_{\Xi}\|\xi\|^{p} Q(d \xi)<\infty\right\}$. Here,

$$
\begin{aligned}
\mathcal{F}_{p}(\Xi):= & \{f: \Xi \mapsto \mathbb{R}:|f(\xi)-f(\tilde{\xi})| \\
& \left.\leq \max \{1,\|\xi\|,\|\tilde{\xi}\|\}^{p-1}\|\xi-\tilde{\xi}\|, \forall \xi, \tilde{\xi} \in \Xi\right\}
\end{aligned}
$$

In order to establish quantitative stability of stochastic program, we need a suitable probability measure.

Definition 2.4. For any $\rho>0$ and probability measures $P, Q \in \mathcal{P}_{\mathcal{F}}, d_{\mathcal{F}, \rho^{-}}$distance is defined by

$$
d_{\mathcal{F}, \rho}(P, Q):=\sup _{x \in \rho \mathbb{B}}\left|\mathbb{E}_{P}\left[f_{0}(x)\right]-\mathbb{E}_{Q}\left[f_{0}(x)\right]\right|
$$

where $\mathbb{E}_{P}\left[f_{0}(x)\right]:=\int_{\Xi} f_{0}(\xi, x) P(d \xi)$.

Definition 2.5. (Hausdorff distance) The Hausdorff distance of closed sets $C$ and $D$ can be written as follows

$$
d_{H}(C, D)=\max \left\{d^{*}(C, D), d^{*}(D, C)\right\}
$$

with $d^{*}(C, D)=\sup \{d(c, D), \forall c \in C\}$.
Definition 2.6. [1] We say that the set-valued mapping $\Psi: X \rightarrow 2^{Y}$ is metric regular at a point $\left(x_{0}, y_{0}\right) \in \operatorname{gph}(\Psi)$, at a rate $c$, if for all $(x, y)$ in a neighborhood of $\left(x_{0}, y_{0}\right)$

$$
\begin{equation*}
d\left(x, \Psi^{-1}(y)\right) \leq c d(y, \Psi(x)) \tag{2.2}
\end{equation*}
$$

Lemma 2.1. [1](Robinson-Ursescu stability theorem) Let $\Psi: X \rightarrow 2^{Y}$ be a closed convex setvalued mapping. Then $\Psi$ is metric regular at $\left(x_{0}, y_{0}\right) \in \operatorname{gph}(\Psi)$ if and only if the regularity condition

$$
y_{0} \in \operatorname{int}(\operatorname{range} \Psi)
$$

holds. More precisely, let $(x, y)$ be such that

$$
\left\|x-x_{0}\right\|<\frac{1}{2} \nu, \quad\left\|y-y_{0}\right\|<\frac{1}{8} \eta .
$$

Then (2.2) holds with constant $c=4 \nu / \eta$.
Lemma 2.2. [1]Let $x_{0} \in \Phi\left(u_{0}\right)$ be such that Robinson's constraint qualification holds. Then for all $(x, u)$ in a neighborhood of $\left(x_{0}, u_{0}\right)$, one has

$$
d(x, \Phi(u))=O(d(G(x, u), K))
$$

where $\Phi(u):=\{x \in X: G(x, u) \in K\}, K$ is a closed convex subset of $Y$ and $G: X \times U \rightarrow Y$ is a continuous mapping.

## 3 Quantitative stability for full random two stage linear secondorder conic stochastic programs

In this section, we derive the Lagrange dual problem to study the quantitative stability for the two stage second-order conic stochastic program. Let $\lambda=\left(\lambda^{1}, \cdots, \lambda^{l}\right) \in \mathcal{Q}:=\mathcal{Q}_{J+1}^{l}$,

$$
\begin{aligned}
L(x, y, \xi, \lambda) & :=c^{T} y-\sum_{i=1}^{l}\left\langle\lambda^{i}, g^{i}(x, y ; \xi)\right\rangle \\
& =c^{T} y-\langle\lambda, \mathcal{A} y\rangle-\sum_{i=1}^{l} \lambda_{J+1}^{i}\left(q_{i}^{T} x-b_{i}\right),
\end{aligned}
$$

where $\mathcal{A}: \Re^{m} \rightarrow \Re^{(J+1) \times l}$ is a linear operator defined by

$$
\mathcal{A} y=\left\{\binom{B^{1} y}{a_{1}^{T} y}, \cdots,\binom{B^{l} y}{a_{l}^{T} y}\right\} .
$$

Then the Lagrange dual of Problem (1.3) becomes

$$
\begin{array}{ll}
\max & \sum_{i=1}^{l} \lambda_{J+1}^{i}\left(b_{i}-q_{i}^{T} x\right) \\
\text { s.t. } & c-\mathcal{A}^{*} \lambda=0,  \tag{3.1}\\
& \lambda \in \mathcal{Q},
\end{array}
$$

where $\mathcal{A}^{*}$ is the adjoint of $\mathcal{A}$ and $\mathcal{A}^{*} \lambda$ is calculated by

$$
\mathcal{A}^{*} \lambda=\sum_{i=1}^{l}\left[\left(B^{i}\right)^{T}, a_{i}\right] \lambda^{i} .
$$

We denote the feasible set of problem (3.1) and problem (1.3) by

$$
\begin{aligned}
& \Lambda(\xi):=\left\{\lambda \in \mathcal{Q}: c-\mathcal{A}^{*} \lambda=0\right\} \\
& Y(x, \xi):=\left\{y \in \Re^{m}: g^{i}(x, y ; \xi) \in \mathcal{Q}_{J+1}, i=1, \ldots, l\right\} .
\end{aligned}
$$

for $\xi \in \Xi$.
To analyze the stability properties of the problem (1.3) when a given parameter $\xi_{0}=$ $\left(c_{0}, A_{0}, Q_{0}, B_{0}, b_{0}\right)$ is perturbed by $\xi=(c, A, Q, B, b)$, we make the following assumptions throughout the paper.

Assumption 3.1. The set $X \subset \Re^{n}$ is a non-empty compact and convex set.
Assumption 3.2. For each $x \in X$ and a given $\xi_{0}$, the optimal value of problem (1.3) is finite and the solution set for problem (1.3) is compact.

Assumption 3.3. The slater condition of problem (1.3) holds for $\xi_{0}$ and each $x \in X$, namely for each $x \in \Re^{n}$, there exists $y_{x}$ such that

$$
g^{i}\left(x, y_{x} ; \xi_{0}\right) \in \operatorname{int} \mathcal{Q}_{J+1}, i=1, \ldots, l .
$$

If Assumption 3.3 is satisfied, it's well known that the dual problem (3.1) has a nonempty compact solution set and the duality gap between (1.3) and its dual problem is zero.

From Assumption 3.1, $X$ is compact and thus bounded. Assume there exists $\gamma>0$ such that for all $x \in X,\|x\| \leq \gamma$. The following theorem reveals that the primal and dual feasible set-valued mappings $\xi \mapsto \Lambda(\xi)$ and $\xi \mapsto Y(x, \xi)$ are locally Lipschitz continuous with respect to the Hausdorff distance.

Theorem 3.1. Suppose that Assumptions 3.1-3.3 hold. For a given $\xi_{0} \in \Xi$ and $\delta>0$, for any $\xi_{1}, \xi_{2} \in \mathbb{B}_{\delta}\left(\xi_{0}\right)$, there exist $L_{\lambda}\left(\xi_{0}\right)$ and $L_{y}\left(\xi_{0}\right)$ such that

$$
\begin{align*}
& d_{H}\left(\Lambda\left(\xi_{1}\right), \Lambda\left(\xi_{2}\right)\right) \leq L_{\lambda}\left(\xi_{0}\right) \max \left\{1,\left\|\xi_{1}\right\|,\left\|\xi_{2}\right\|\right\}\left\|\xi_{1}-\xi_{2}\right\|,  \tag{3.2}\\
& d_{H}\left(Y\left(x, \xi_{1}\right), Y\left(x, \xi_{2}\right)\right) \leq L_{y}\left(\xi_{0}\right) \max \left\{1,\left\|\xi_{1}\right\|,\left\|\xi_{2}\right\|\right\}\left\|\xi_{1}-\xi_{2}\right\| \tag{3.3}
\end{align*}
$$

Proof From Lemma 2.1 and Lemma 3.1 in [3], the slater condition of Problem (1.3) and (3.1) hold around $\xi_{0}$ under Assumptions 3.1-3.3. We first prove the locally Lipschitz continuity of set-valued mapping $\xi \mapsto \Lambda(\xi)$ with respect to the Hausdorff distance around $\xi_{0}$. Define

$$
F(\lambda, \xi)=\left[\begin{array}{c}
\lambda \\
c-\mathcal{A}^{*} \lambda
\end{array}\right] \quad \text { and } \quad K=\mathcal{Q} \times\{0\} .
$$

Then $\Lambda(\xi)$ can be rewritten as

$$
\Lambda(\xi)=\{\lambda: F(\lambda, \xi) \in K\}
$$

From Lemma 3.2 in [3] for fixed $\xi_{0}$, we can obtain that

$$
\begin{equation*}
\lambda_{0} \in \operatorname{int}\left(\operatorname{range} \Lambda\left(\xi_{0}\right)\right) \tag{3.4}
\end{equation*}
$$

Because of Lemma 2.1 and (3.4), we have that $\Lambda\left(\xi_{0}\right)$ is metrically regular at $\left(\xi_{0}, \lambda_{0}\right) \in$ $\operatorname{gph}\left(\Lambda\left(\xi_{0}\right)\right)$ and thus the Robinson's constraint qualification of $\lambda_{0} \in \Lambda\left(\xi_{0}\right)$ holds. For $\delta>0$, $\xi_{1}, \xi_{2} \in \mathbb{B}_{\delta}\left(\xi_{0}\right)$, let $\lambda_{1} \in \Lambda\left(\xi_{1}\right)$ and $\lambda_{2} \in \Lambda\left(\xi_{2}\right)$ such that

$$
\begin{aligned}
& d^{*}\left(\Lambda\left(\xi_{1}\right), \Lambda\left(\xi_{2}\right)\right)=d\left(\lambda_{1}, \Lambda\left(\xi_{2}\right)\right) \\
& d^{*}\left(\Lambda\left(\xi_{2}\right), \Lambda\left(\xi_{1}\right)\right)=d\left(\lambda_{2}, \Lambda\left(\xi_{1}\right)\right)
\end{aligned}
$$

Because of Lemma 3.2 in [3], we have the continuous differentiability of $\Lambda(\xi)$ and there exists $\varepsilon>0$ such that $\lambda_{1}, \lambda_{2} \in \mathbb{B}_{\varepsilon}\left(\lambda_{0}\right)$. Now we choose the couple ( $\xi_{2}, \lambda_{1}$ ), it follows from Lemma 2.2 that

$$
\begin{aligned}
d\left(\lambda_{1}, \Lambda\left(\xi_{2}\right)\right) & \leq \kappa_{1} d\left(F\left(\lambda_{1}, \xi_{2}\right), K\right) \leq \kappa_{1}\left(\left\|F\left(\lambda_{1}, \xi_{2}\right)-F\left(\lambda_{1}, \xi_{1}\right)\right\|+d\left(F\left(\lambda_{1}, \xi_{1}\right), K\right)\right) \\
& \leq \kappa_{1}\left\|c_{1}-c_{2}+\left(\mathcal{A}_{2}^{*}-\mathcal{A}_{1}^{*}\right) \lambda_{1}\right\| \leq \kappa_{1} \max \left\{1,\left\|\lambda_{1}\right\|\right\}\left\|\xi_{1}-\xi_{2}\right\|
\end{aligned}
$$

with $\kappa_{1}>0$. For the couple $\left(\xi_{1}, \lambda_{2}\right)$, we similarly obtain that

$$
d\left(\lambda_{2}, \Lambda\left(\xi_{1}\right)\right) \leq \kappa_{2} \max \left\{1,\left\|\lambda_{2}\right\|\right\}\left\|\xi_{1}-\xi_{2}\right\|
$$

with $\kappa_{2}>0$.

We now prove the operator $\mathcal{A}_{0}^{*}$ is onto when Assumption 3.1 holds. Suppose that there exists $d_{y} \in \Re^{m}$ such that $\mathcal{A}_{0} d_{y}=0$, then we obtain that $c_{0}^{T} d_{y}=0$ by

$$
\left\langle c_{0}-\mathcal{A}_{0}^{*} \lambda_{0}, d_{y}\right\rangle=0 \Leftrightarrow c_{0}^{T} d_{y}-\left\langle\mathcal{A}_{0}^{*} \lambda_{0}, d_{y}\right\rangle=0 \Leftrightarrow c_{0}^{T} d_{y}-\left\langle\lambda_{0}, \mathcal{A}_{0} d_{y}\right\rangle=0
$$

For $\beta>0$ and $y^{*}$ is an optimal solution of problem $P\left(x, \xi_{0}\right)$, then $\bar{y}=y^{*}+\beta d_{y}$ is also an optimal solution. $\bar{y}$ is unbounded when $\beta \rightarrow \infty$, which is a contradiction with Assumption 3.2 and thus $d_{y}=0$. Then we have that $\operatorname{ker} \mathcal{A}_{0}=\{0\}$ and operator $\mathcal{A}_{0}^{*}$ is onto.

Define $M(\xi):=\left[\left(\left(B^{1}\right)^{T}, a_{1}\right), \cdots,\left(\left(B^{l}\right)^{T}, a_{l}\right)\right] \in \Re^{m \times l(J+1)}$ and $M\left(\xi_{i}\right)=M_{i}$, in view of $\mathcal{A}_{0}^{*}$, we have that matrix $M_{0}$ is of row full rank. Then for $d_{y} \neq 0$, we have

$$
\left\|M_{0}^{T} d_{y}\right\|_{2}^{2}>0 \Rightarrow\left(M_{0}^{T} d_{y}\right)^{T} M_{0}^{T} d_{y}>0 \Rightarrow d_{y}^{T}\left(M_{0} M_{0}^{T}\right) d_{y}>0
$$

and thus $M_{0} M_{0}^{T}$ is positive definite. Let $\Delta \mathcal{N}_{i j}=\Delta M_{i j} M_{i}^{T}+M_{i} \Delta M_{i j}^{T}+\Delta M_{i j} \Delta M_{i j}^{T}$, where $\Delta M_{i j}=M_{i}-M_{j}, 0 \leq j \leq i \leq 2$. When $\Delta M_{i 0}$ is small enough, $M_{i} M_{i}^{T}=\left(M_{0}+\Delta M_{i 0}\right)\left(M_{0}+\right.$ $\left.\Delta M_{i 0}\right)^{T}=M_{0} M_{0}^{T}+\Delta \mathcal{N}_{i 0}, i=1,2$ are nonsingular. We assume that $\delta_{i 0}>0$ such that $\left\|\Delta M_{i 0}\right\| \leq$ $\delta_{i 0}$ satisfying that $M_{i} M_{i}^{T}$ is nonsingular. Then we obtain from Sherman-Morrison-Woodbury formula that

$$
\begin{aligned}
M_{i}^{\dagger} & :=M_{i}^{T}\left(M_{i} M_{i}^{T}\right)^{-1} \\
& =\left(M_{j}+\Delta M_{i j}^{T}\right)\left(M_{j} M_{j}^{T}+\Delta \mathcal{N}_{i j}\right)^{-1} \\
& =\left(M_{j}+\Delta M_{i j}^{T}\right)\left[\left(M_{j} M_{j}^{T}\right)^{-1}-\left(M_{j} M_{j}^{T}\right)^{-1} \Delta \mathcal{N}_{i j}\left[I_{m}+\left(M_{j} M_{j}^{T}\right)^{-1} \Delta \mathcal{N}_{i j}\right]^{-1}\left(M_{j} M_{j}^{T}\right)^{-1}\right] \\
& =M_{j}^{\dagger}+\Delta \Sigma_{i j} .
\end{aligned}
$$

From Theorem 3.8 in [19], we know that there exists a constant $\mu>0$ such that the following estimation holds:

$$
\left\|\Delta \Sigma_{i j}\right\|=\left\|M_{i}^{\dagger}-M_{j}^{\dagger}\right\| \leq \mu \max \left\{\left\|M_{i}^{\dagger}\right\|,\left\|M_{j}^{\dagger}\right\|\right\}\left\|\Delta M_{i j}\right\|
$$

for $0 \leq j \leq i \leq 2$. Since $\Delta M_{i 0}$ is small enough and $\xi_{0}$ is fixed, as well as $\left\|M_{0}^{\dagger}\right\|$, then $\left\|M_{i}^{\dagger}\right\|, i=$ 1,2 are bounded and constrained by some $L\left(\xi_{0}\right) \geq 1$ without loss of generality. On the other hand, from the constraints of the problem (3.1), we have that

$$
\begin{aligned}
& c_{1}-\mathcal{A}_{1}^{*} \lambda_{1}=0 \Rightarrow c_{1}=M_{1} \lambda_{1} \Rightarrow \lambda_{1}=M_{1}^{T}\left(M_{1} M_{1}^{T}\right)^{-1} c_{1} \\
\Rightarrow & \lambda_{1}=M_{1}^{\dagger} c_{1} \Rightarrow\left\|\lambda_{1}\right\| \leq\left\|M_{1}^{\dagger}\right\|\left\|c_{1}\right\| \leq L\left(\xi_{0}\right)\left\|\xi_{1}\right\| .
\end{aligned}
$$

Similarly we have that for any $\lambda \in \Lambda(\xi), \xi \in \mathbb{B}_{\delta}\left(\xi_{0}\right)$,

$$
\begin{equation*}
\|\lambda\| \leq L\left(\xi_{0}\right)\|\xi\| \tag{3.5}
\end{equation*}
$$

Taking $L_{\lambda}\left(\xi_{0}\right)=L\left(\xi_{0}\right) \max \left\{\kappa_{1}, \kappa_{2}\right\}$, from the above discussion,

$$
\begin{aligned}
& d_{H}\left(\Lambda\left(\xi_{1}\right), \Lambda\left(\xi_{2}\right)\right)=\max \left\{d\left(\lambda_{1}, \Lambda\left(\xi_{2}\right)\right), d\left(\lambda_{2}, \Lambda\left(\xi_{1}\right)\right)\right\} \\
\leq & \max \left\{\kappa_{1}, \kappa_{2}\right\} \max \left\{1,\left\|\lambda_{1}\right\|,\left\|\lambda_{2}\right\|\right\}\left\|\xi_{1}-\xi_{2}\right\| \\
\leq & L_{\lambda}\left(\xi_{0}\right) \max \left\{1,\left\|\xi_{1}\right\|,\left\|\xi_{2}\right\|\right\}\left\|\xi_{1}-\xi_{2}\right\|,
\end{aligned}
$$

which means that $\xi \mapsto \Lambda(\xi)$ is locally Lipschitz continuous in the sense of Hausdorff distance.
Now we consider the locally Lipshitz continuous property of $Y(x, \xi)$ in the sense of Hausdorff distance. Because of Assumption 3.3, the Robinson's constraint qualification of $y_{0} \in Y\left(x, \xi_{0}\right)$ holds. For $\delta>0, \tilde{\varepsilon}>0, \xi_{1}, \xi_{2} \in \mathbb{B}_{\delta}\left(\xi_{0}\right)$, there exists $y_{1} \in Y\left(x, \xi_{1}\right)$ and $y_{2} \in Y\left(x, \xi_{2}\right)$ such that

$$
\begin{aligned}
& d^{*}\left(Y\left(x, \xi_{1}\right), Y\left(x, \xi_{2}\right)\right)=d\left(y_{1}, Y\left(x, \xi_{2}\right)\right), \\
& d^{*}\left(Y\left(x, \xi_{2}\right), Y\left(x, \xi_{1}\right)\right)=d\left(y_{2}, Y\left(x, \xi_{1}\right)\right) .
\end{aligned}
$$

From the continuous differentiability of $Y(x, \xi)$, there exists $\tilde{\varepsilon}>0$ such that $y_{1}, y_{2} \in \mathbb{B}_{\tilde{\varepsilon}}\left(y_{0}\right)$ from Lemma 2.2 in [3]. For $\left(y_{1}, \xi_{2}\right),\left(y_{2}, \xi_{1}\right)$ in the neighborhood of $\left(y_{0}, \xi_{0}\right)$ and fixed $x$, we have that

$$
\begin{aligned}
\operatorname{dist}\left(y_{1}, Y\left(x, \xi_{2}\right)\right) & \leq \kappa_{3} \operatorname{dist}\left(g\left(x, y_{1}, \xi_{2}\right), \mathcal{Q}\right) \\
& \leq \kappa_{3}\left(\left\|g\left(x, y_{1}, \xi_{2}\right)-g\left(x, y_{1}, \xi_{1}\right)\right\|+\operatorname{dist}\left(g\left(x, y_{1}, \xi_{1}\right), \mathcal{Q}\right)\right) \\
& \leq \kappa_{3}\left(\sum_{j=1}^{l}\left\|B_{1}^{j}-B_{2}^{j}\right\|\left\|y_{1}\right\|+\left\|A_{1}-A_{2}\right\|\left\|y_{1}\right\|+\left\|Q_{1}-Q_{2}\right\|\|x\|+\left\|b_{1}-b_{2}\right\|\right) \\
& \leq \kappa_{3} \max \left\{1, \gamma,\left\|y_{1}\right\|\right\}\left\|\xi_{1}-\xi_{2}\right\|
\end{aligned}
$$

where $\kappa_{3}>0, g(x, y, \xi)=(B y, A y+Q x-b)$ and are in $\mathcal{Q}$ for couples $\left(y_{1}, \xi_{1}\right),\left(y_{2}, \xi_{2}\right)$. Then for $\kappa_{4}>0$,

$$
d\left(y_{2}, Y\left(x, \xi_{1}\right)\right) \leq \kappa_{4} \max \left\{1, \gamma,\left\|y_{2}\right\|\right\}\left\|\xi_{1}-\xi_{2}\right\|
$$

can be obtained.
Let

$$
Y(x, \xi, \alpha):=Y(x, \xi) \cap l e v_{\leq \alpha} f(y, \xi), \quad \forall \alpha,
$$

where $\operatorname{lev}_{\leq \alpha} f(y, \xi)=\left\{y \in \Re^{m}: c^{T} y \leq \alpha\right\}$, we now prove that the set $Y(x, \xi, \alpha)$ is bounded. We only need to prove that for $\alpha \geq c_{0}^{T} y^{*}, y^{*}$ is an optimal solution of (1.3) and there exists $\delta>0$ and a bounded set $\mathcal{D} \subset \Re^{m}$ such that $Y(x, \xi, \alpha) \subset \mathcal{D}, \forall \xi \in \mathbb{B}_{\delta}\left(\xi_{0}\right)$. We prove the result by contradiction. Suppose that there exist a sequence $\left\{\xi^{k}\right\}$ such that $\xi^{k} \rightarrow \xi_{0}$ and $y^{k} \in Y\left(x, \xi^{k}, \alpha\right)$ with $\left\|y^{k}\right\| \rightarrow \infty$. Let $d_{y}^{k}=y^{k} /\left\|y^{k}\right\|$, we can find a subsequence $k_{j}$ such that $d_{y}^{k_{j}} \rightarrow d_{y}$ for $\left\|d_{y}\right\|=1$. In view of $y^{k_{j}} \in Y\left(x, \xi^{k_{j}}, \alpha\right)$, one has

$$
\begin{aligned}
c_{0}^{k_{j} T} y^{k_{j}} & \leq \alpha \\
a_{0, i}^{k_{j} T} y^{k_{j}} & +q_{0, i}^{k_{j} T} x-b_{0, i}^{k_{j}} \geq\left\|\left[B_{0}^{k_{j}}\right]^{i} y^{k_{j}}\right\|_{2}, i=1, \cdots, l .
\end{aligned}
$$

Dividing both sides of the above inequalities by $\left\|y^{k_{j}}\right\|$, we obtain

$$
\begin{aligned}
& c_{0}^{k_{j} T} d_{y}^{k_{j}} \leq \alpha /\left\|y^{k_{j}}\right\| \\
& a_{0, i}^{k_{j} T} d_{y}^{k_{j}}+q_{0, i}^{k_{j} T} x /\left\|y^{k_{j}}\right\|-b_{0, i}^{k_{j}} /\left\|y^{k_{j}}\right\| \geq\left\|\left[B_{0}^{k_{j}}\right]^{i} d_{y}^{k_{j}}\right\|_{2}, i=1, \cdots, l .
\end{aligned}
$$

Taking the limits by $j \rightarrow \infty$, we have

$$
c_{0}^{T} d_{y} \leq 0, \quad a_{0, i}^{T} d_{y} \geq\left\|B_{0}^{i} d_{y}\right\|_{2} \geq 0, i=1, \cdots, l .
$$

Let $\beta>0$. We now show that $\bar{y}=y^{*}+\beta d_{y}$ is a feasible point of (1.3). For any $i=1, \cdots, l$,

$$
a_{0, i}^{T} \bar{y}+q_{0, i}^{T} x-b_{0, i}=a_{0, i}^{T} y^{*}+q_{0, i}^{T} x-b_{0, i}+\beta a_{0, i}^{T} d_{y} \geq\left\|B_{0}^{i} y^{*}\right\|+\beta\left\|B_{0}^{i} d_{y}\right\|_{2} \geq\left\|B_{0}^{i} \bar{y}\right\|_{2}
$$

Since $c_{0}^{T} d_{y} \leq 0$, we have that

$$
c_{0}^{T} \bar{y}=c_{0}^{T}\left(y^{*}+\beta d_{y}\right)=c_{0}^{T} y^{*}+\beta c_{0}^{T} d_{y} \leq c_{0}^{T} y^{*} .
$$

There $\bar{y}$ is also an optimal solution, which implies $d_{y}=0$ and then we get the contradiction with $\left\|d_{y}\right\|=1$. Thus we obtain that $Y\left(x, \xi_{0}, \alpha\right)$ is bounded. Since $Y(x, \xi, \alpha)$ is continuous around $\xi_{0}$ by Lemma 2.2 in [3], $Y(x, \xi, \alpha)$ is also bounded.

Let $\bar{B}=\left(\bar{b}_{1}, \cdots, \bar{b}_{l}\right)^{T} \in \Re^{l \times m}$ such that $\bar{b}_{i}^{T} y=\left\|B^{i} y\right\|_{2}, i=1, \cdots, l$. Then

$$
Q x-b \geq(\bar{B}-A) y
$$

Since $Y(x, \xi, \alpha)$ is bounded, it is known from Theorem 9.3 in [6] that $0 \in \operatorname{int}(H)$, where $H=$ $\operatorname{con}\left\{h_{i} \mid h_{i}=\bar{b}_{i}-a_{i}, i=1, \cdots, l\right\}$, so we have $q_{k}^{T} x-b_{k}>0$ for some $k: 1 \leq k \leq l$. For $0 \neq y \in Y(x, \xi, \alpha)$, we consider $\zeta=y /\|y\|$, then there must exists some $k \in 1, \cdots, l$ to make $h_{k} \zeta>0$. Then we have that

$$
\begin{aligned}
& \|y\| \times \frac{h_{k} y}{\|y\|} \leq q_{k}^{T} x-b_{k} \\
\Rightarrow & \|y\| \leq \min \left\{\left.\frac{q_{k}^{T} x-b_{k}}{h_{k} \zeta} \right\rvert\, h_{k} \zeta>0,1 \leq k \leq l\right\} \\
\Rightarrow & \|y\| \leq \frac{\max \left\{q_{k}^{T} x-b_{k}: 1 \leq k \leq l\right\}}{\max \left\{h_{k} \zeta \mid h_{k} \zeta>0,1 \leq k \leq l\right\}} \\
\Rightarrow & \|y\| \leq L_{\zeta}\left(\xi_{0}\right) \max \{1, \gamma\}\|\xi\|,
\end{aligned}
$$

where $L_{\zeta}\left(\xi_{0}\right)=1 / \max \left\{h_{k} \zeta \mid h_{k} \zeta>0,1 \leq k \leq l\right\}$. Furthermore, we can obtain that

$$
d_{H}\left(Y\left(x, \xi_{1}\right), Y\left(x, \xi_{2}\right)\right) \leq L_{y}\left(\xi_{0}\right)\left\{1,\left\|\xi_{1}\right\|,\left\|\xi_{2}\right\|\right\}\left\|\xi_{1}-\xi_{2}\right\|
$$

with $L_{y}\left(\xi_{0}\right)=\max \left\{\kappa_{3}, \kappa_{4}\right\} \max \{1, \gamma\} \max \left\{1, L_{\zeta}\left(\xi_{0}\right)\right\}$. Then $\xi \mapsto Y(\xi)$ is locally Lischitz continuous in the sense of Hausdorff distance. We complete the proof.

We now verify that locally Lipschitz property of the primal feasible set-valued mapping $\xi \mapsto \Lambda(\xi)$ imply locally Lipschitz continuity of $f_{0}(\cdot, x)$.

Theorem 3.2. Assume that Assumptions 3.1- 3.3 hold. Then, for $\xi_{0}=\left(c_{0}, A_{0}, Q_{0}, B_{0}, b_{0}\right)$ and $\delta>0$, there exist constants $\hat{L}\left(\xi_{0}\right)>0, \bar{L}\left(\xi_{0}\right)>0$ and $\tilde{L}\left(\xi_{0}\right)>0$ such that

$$
\begin{aligned}
& f_{0}\left(x, \xi_{1}\right)-f_{0}\left(x, \xi_{2}\right) \leq \bar{L}\left(\xi_{0}\right) \max \left\{1,\left\|\xi_{1}\right\|,\left\|\xi_{2}\right\|\right\}^{2}\left\|\xi_{1}-\xi_{2}\right\|, \\
& f_{0}\left(x_{1}, \xi\right)-f_{0}\left(x_{2}, \xi\right) \leq \hat{L}\left(\xi_{0}\right) \max \{1,\|\xi\|\}^{2}\left\|x_{1}-x_{2}\right\|, \\
& \left\|f_{0}(x, \xi)\right\| \leq \tilde{L}\left(\xi_{0}\right) \max \{1,\|\xi\|\}^{2},
\end{aligned}
$$

for all $\xi_{1}, \xi_{2} \in \mathbb{B}_{\delta}\left(\xi_{0}\right), x_{1}, x_{2} \in X$.
Proof. Since the duality gap between problem (1.3) and its dual problem is zero, we set $\mu:=\lambda_{J+1}$, for $x_{1}, x_{2} \in X$, and $\xi_{1}, \xi_{2} \in \mathbb{B}_{\delta}\left(\xi_{0}\right)$,

$$
\begin{aligned}
f_{0}\left(x_{1}, \xi_{1}\right)-f_{0}\left(x_{2}, \xi_{2}\right) & =d^{T} x_{1}+\theta\left(x_{1}, \xi_{1}\right)-\left(d^{T} x_{2}+\theta\left(x_{2}, \xi_{2}\right)\right) \\
& =d^{T}\left(x_{1}-x_{2}\right)+\left\langle\mu_{1}^{*}, b_{1}-Q_{1} x_{1}\right\rangle-\left\langle\mu_{2}^{*}, b_{2}-Q_{2} x_{2}\right\rangle \\
& \leq d^{T}\left(x_{1}-x_{2}\right)+\left\langle\mu_{1}^{*}, b_{1}-Q_{1} x_{1}\right\rangle-\left\langle\bar{\mu}_{2}, b_{2}-Q_{2} x_{2}\right\rangle
\end{aligned}
$$

where $\mu_{1}^{*} \in \Lambda\left(\xi_{1}\right), \mu_{2}^{*} \in \Lambda\left(\xi_{2}\right)$ are the dual optimal solution for $\left(x_{1}, \xi_{1}\right),\left(x_{2}, \xi_{2}\right)$ respectively and we denote by $\bar{\mu}_{2}$ the projection of $\mu_{1}^{*}$ onto $\Lambda\left(\xi_{2}\right)$. Thus we can have that by equations (3.2) and (3.5),

$$
\begin{align*}
& \left\|\mu_{1}^{*}-\bar{\mu}_{2}\right\| \leq L_{\lambda}\left(\xi_{0}\right) \max \left\{1,\left\|\xi_{1}\right\|,\left\|\xi_{2}\right\|\right\}\left\|\xi_{1}-\xi_{2}\right\|  \tag{3.6}\\
& \left\|\bar{\mu}_{2}\right\| \leq L\left(\xi_{0}\right)\left\|\xi_{2}\right\| .
\end{align*}
$$

For any fixed $x \in X$, we have

$$
\begin{aligned}
& f_{0}\left(x, \xi_{1}\right)-f_{0}\left(x, \xi_{2}\right) \\
\leq & \left\langle\mu_{1}^{*}, b_{1}-Q_{1} x\right\rangle-\left\langle\bar{\mu}_{2}, b_{2}-Q_{2} x\right\rangle \\
= & \left\langle\mu_{1}^{*}-\bar{\mu}_{2}, b_{1}-Q_{1} x\right\rangle+\left\langle\bar{\mu}_{2},\left(b_{1}-b_{2}\right)-\left(Q_{1} x-Q_{2} x\right)\right\rangle \\
\leq & \left\|\mu_{1}^{*}-\bar{\mu}_{2}\right\|\left\|b_{1}-Q_{1} x\right\|+\left\|\bar{\mu}_{2}\right\|\left\|\left(b_{1}-b_{2}\right)-\left(Q_{1} x-Q_{2} x\right)\right\| \\
\leq & L_{\lambda}\left(\xi_{0}\right) \max \{1, \gamma\} \max \left\{1,\left\|\xi_{1}\right\|,\left\|\xi_{2}\right\|\right\}\left\|\xi_{1}-\xi_{2}\right\|\left\|\xi_{1}\right\|+L\left(\xi_{0}\right) \max \{1, \gamma\}\left\|\xi_{2}\right\|\left\|\xi_{1}-\xi_{2}\right\| \\
\leq & \bar{L}\left(\xi_{0}\right) \max \left\{1,\left\|\xi_{1}\right\|,\left\|\xi_{2}\right\|\right\}^{2}\left\|\xi_{1}-\xi_{2}\right\|,
\end{aligned}
$$

where $\bar{L}\left(\xi_{0}\right)=\max \left\{L_{\lambda}\left(\xi_{0}\right), L\left(\xi_{0}\right)\right\} \max \{1, \gamma\}$. The last inequality holds since

$$
\left\|\xi_{1}\right\|,\left\|\xi_{2}\right\| \leq \max \left\{1,\left\|\xi_{1}\right\|,\left\|\xi_{2}\right\|\right\} \leq \max \left\{1,\left\|\xi_{1}\right\|,\left\|\xi_{2}\right\|\right\}^{2}
$$

Let $\xi_{1}=\xi_{2}=\xi, \Lambda\left(\xi_{1}\right)=\Lambda\left(\xi_{2}\right)$, we have

$$
\begin{aligned}
& f_{0}\left(x_{1}, \xi\right)-f_{0}\left(x_{2}, \xi\right) \\
\leq & d^{T}\left(x_{1}-x_{2}\right)+\left\langle\mu_{1}^{*}, b-Q x_{1}\right\rangle-\left\langle\mu_{1}^{*}, b-Q x_{2}\right\rangle \\
= & d^{T}\left(x_{1}-x_{2}\right)+\left\langle\mu_{1}^{*}, Q x_{2}-Q x_{1}\right\rangle \\
\leq & \|d\|\left\|x_{1}-x_{2}\right\|+\left\|\mu_{1}^{*}\right\|\| \| x_{2}-Q x_{1} \| \\
\leq & \|d\|\left\|x_{1}-x_{2}\right\|+L\left(\xi_{0}\right) \max \{1,\|\xi\|\}\|Q\|\left\|x_{1}-x_{2}\right\| \\
\leq & \hat{L}\left(\xi_{0}\right) \max \{1,\|\xi\|\}^{2}\left\|x_{1}-x_{2}\right\|,
\end{aligned}
$$

where $\hat{L}\left(\xi_{0}\right)=\max \left\{\|d\|, L\left(\xi_{0}\right)\right\}$. For any $\xi \in \mathbb{B}_{\delta}\left(\xi_{0}\right), x \in X, \mu^{*}$ is the dual problem solution for $(x, \xi)$,

$$
\begin{align*}
& \left|f_{0}(x, \xi)\right| \\
\leq & \|d\|\|x\|+\left\|\mu^{*}\right\|\|b-Q x\| \\
\leq & \|d\| \gamma+L\left(\xi_{0}\right) \max \{1, \gamma\} \max \{1,\|\xi\|\}\|\xi\| \\
\leq & \tilde{L}\left(\xi_{0}\right) \max \{1,\|\xi\|\}^{2}, \tag{3.7}
\end{align*}
$$

where $\tilde{L}\left(\xi_{0}\right)=\max \left\{\|d\|, L\left(\xi_{0}\right)\right\} \max \{1, \gamma\}$.
Since the probability measure $P$ would be unknown in most reality examples, we consider the case that $P$ is approximated by $Q$ in the rest of this section.

Note that problem (1.2) can be reformulated as

$$
\begin{equation*}
\min \left\{\int_{\Xi} f_{0}(x, \xi) d P(\xi): x \in X\right\} . \tag{3.8}
\end{equation*}
$$

We denote the optimal value function and the optimal solution mapping of problem (3.8) by

$$
\begin{aligned}
& v(P):=\inf \left\{\mathbb{E}_{p}\left[f_{0}(x, \xi)\right]: x \in X\right\}, \\
& S(P):=\operatorname{argmin}\left\{\mathbb{E}_{p}\left[f_{0}(x, \xi)\right]: x \in X\right\} .
\end{aligned}
$$

In order to investigate the stability analysis of (3.8) when $P$ is perturbed, we give the following assumptions.

Assumption 3.4. Let $\Xi_{0}$ be a countable dense set of $\Xi$. For $\xi \in \Xi_{0}$, Assumptions 3.1- 3.3 hold.
Assumption 3.5. Let $\Xi_{0}$ be a countable dense set of $\Xi$, the optimal solution mapping of dual problem is uniformly bounded by $C$ over $\Xi_{0}$.

Under Assumptions 3.4, 3.5, we show that Theorems 3.1 and 3.2 hold for every $\xi \in \Xi$.

Corollary 3.1. Let Assumption 3.1- 3.4 hold, then for any $\xi \in \Xi, \delta>0$, for any $\tilde{\xi}, \hat{\xi} \in \mathbb{B}_{\delta}(\xi)$, there exists $L_{\xi}$ such that

$$
\begin{equation*}
\max \left\{d_{H}(\Lambda(\tilde{\xi}), \Lambda(\hat{\xi})), d_{H}(Y(x, \tilde{\xi}), Y(x, \hat{\xi}))\right\} \leq L_{\xi}\|\tilde{\xi}-\hat{\xi}\| \tag{3.9}
\end{equation*}
$$

Proof. For any $\xi \in \Xi, \delta>0$, there exist $\xi^{\prime} \in \Xi_{0}$ such that $\xi \in \mathbb{B}_{\frac{\delta}{2}}\left(\xi^{\prime}\right)$ and Assumptions 3.1-3.3 hold at $\xi^{\prime}$. From Theorem 3.1, for any $\tilde{\xi}, \hat{\xi} \in \mathbb{B}_{\frac{\delta}{2}}(\xi) \subseteq \mathbb{B}_{\delta}\left(\xi^{\prime}\right)$, there exists $L_{\xi}=$ $\max \left\{L_{y}\left(\xi^{\prime}\right) \max \{1,\|\tilde{\xi}\|,\|\hat{\xi}\|\}, L_{\lambda}\left(\xi^{\prime}\right) \max \{1,\|\tilde{\xi}\|,\|\hat{\xi}\|\}\right\}$ such that (3.9) holds.

Therefore the Lipschitz property hold for the Pompeiu-Hausdorff distance about the mapping $\xi \rightarrow \Lambda(\xi)$ and $\xi \rightarrow Y(\xi), \forall \xi \in \Xi$.

Corollary 3.2. Assume that Assumptions 3.1-3.4 hold. Then for any $\xi \in \Xi$ and $\delta>0$, $\xi_{1}, \xi_{2} \in \mathbb{B}_{\delta}(\xi), x_{1}, x_{2} \in X$, there exist constants $\hat{L}(\xi)>0, \bar{L}(\xi)>0$ and $\tilde{L}(\xi)>0$ such that

$$
\begin{align*}
& f_{0}\left(x, \xi_{1}\right)-f_{0}\left(x, \xi_{2}\right) \leq \bar{L}(\xi) \max \left\{1,\left\|\xi_{1}\right\|,\left\|\xi_{2}\right\|\right\}^{2}\left\|\xi_{1}-\xi_{2}\right\|,  \tag{3.10}\\
& f_{0}\left(x_{1}, \xi\right)-f_{0}\left(x_{2}, \xi\right) \leq \hat{L}(\xi) \max \{1,\|\xi\|\}^{2}\left\|x_{1}-x_{2}\right\|,  \tag{3.11}\\
& \left\|f_{0}(x, \xi)\right\| \leq \tilde{L}(\xi) \max \{1,\|\xi\|\}^{2} . \tag{3.12}
\end{align*}
$$

Furthermore, if Assumption 3.5 holds, for any $\xi \in \Xi$ and $x, x_{1}, x_{2} \in X$, there exist constants $\hat{L}, \tilde{L}$ such that

$$
\begin{align*}
& f_{0}\left(x_{1}, \xi\right)-f_{0}\left(x_{2}, \xi\right) \leq \hat{L} \max \{1,\|\xi\|\}\left\|x_{1}-x_{2}\right\|,  \tag{3.13}\\
& \left\|f_{0}(x, \xi)\right\| \leq \tilde{L} \max \{1,\|\xi\|\} . \tag{3.14}
\end{align*}
$$

Proof. Similarly with Corollary 3.1, equations (3.10) - (3.12) derive from the Theorem 3.2. For any $x \in X, \xi \in \Xi$, there exists $\xi^{\prime} \in \Xi_{0}$ such that $\xi \in \mathbb{B}_{\frac{\delta}{2}}\left(\xi^{\prime}\right)$, for dual problem solutions $\mu \in \Lambda(\xi)$ and $\mu^{\prime} \in \Lambda\left(\xi^{\prime}\right)$, we have $\|\mu\| \leq\left\|\mu^{\prime}\right\|+1 \leq C+1$ from Assumption 3.5. Let $\mu_{1}$ be the dual problem solution for $\left(x_{1}, \xi\right),\|\mu\| \leq C+1$,

$$
\begin{aligned}
& f_{0}\left(x_{1}, \xi\right)-f_{0}\left(x_{2}, \xi\right) \\
\leq & d^{T}\left(x_{1}-x_{2}\right)+\left\langle\mu_{1}, Q x_{2}-Q x_{1}\right\rangle \\
\leq & \|d\|\left\|x_{1}-x_{2}\right\|+\left\|\mu_{1}\right\|\left\|Q x_{2}-Q x_{1}\right\| \\
\leq & \|d\|\left\|x_{1}-x_{2}\right\|+(C+1)\|Q\|\left\|x_{1}-x_{2}\right\| \\
\leq & \hat{L} \max \{1,\|\xi\|\}\left\|x_{1}-x_{2}\right\|,
\end{aligned}
$$

where $\hat{L}=\max \{\|d\|, C+1\}$ and

$$
\begin{aligned}
& \left|f_{0}(x, \xi)\right|=d^{T} x+\langle\mu, b-Q x\rangle \\
\leq & \|d\|\|x\|+\|\mu\|\|b-Q x\| \\
\leq & \|d\| \gamma+(C+1) \max \{1, \gamma\}\|\xi\| \\
\leq & \tilde{L} \max \{1,\|\xi\|\},
\end{aligned}
$$

where $\tilde{L}=\max \{\|d\|, C+1\} \max \{1, \gamma\}$.
Inspired by Theorem 2.1 [15], the following proposition holds.
Proposition 3.1. Let $P \in \mathcal{P}_{\mathcal{F}}$. Suppose $S(P)$ is nonempty and bounded. Then there exist constants $\rho>0$ and $\delta>0$ such that

$$
\begin{aligned}
& |v(P)-v(Q)| \leq d_{\mathcal{F}, \rho}(P, Q) \\
& \emptyset \neq S(Q) \subset S(P)+\Psi_{P}\left(d_{\mathcal{F}, \rho}(P, Q)\right) \mathbb{B}
\end{aligned}
$$

hold for all $Q \in \mathcal{P}_{\mathcal{F}}$ with $d_{\mathcal{F}, \rho}(P, Q)<\delta$, where $\Psi_{P}$ is a conditioning function associated with our given problem; more precisely,

$$
\begin{equation*}
\Psi_{P}(\eta):=\eta+\psi_{P}^{-1}(2 \eta), \quad \eta \geq 0 \tag{3.15}
\end{equation*}
$$

with

$$
\psi_{P}(\tau):=\min \left\{\mathbb{E}_{P}\left[f_{0}(x)\right]-v(P): d(x, S(P)) \geq \tau, x \in X\right\}, \quad \tau \geq 0
$$

and

$$
\psi_{P}^{-1}(t):=\sup \left\{\tau \in \mathbb{R}_{+}: \psi_{P}(\tau) \leq t\right\}
$$

Proof. We only need to verify that $f_{0}(x, \xi)$ is a convex random lower semi-continuious function.

From Corollary 8.14 in [14], we consider that for any $\xi \in \Xi, f_{0}(\cdot, \xi)$ is convex. For $\tau \in[0,1]$ and $x_{1}, x_{2} \in X$, we have

$$
\begin{aligned}
& f_{0}\left(\tau x_{1}+(1-\tau) x_{2}, \xi\right) \\
= & d^{T}\left(\tau x_{1}+(1-\tau) x_{2}\right)+\max _{\lambda}\left\{\left\langle\lambda_{J+1}, b-Q\left(\tau x_{1}+(1-\tau) x_{2}\right)\right\rangle: c-\mathcal{A}^{*} \lambda=0, \lambda \in \mathcal{Q}\right\} \\
\leq & \tau d^{T} x_{1}+\tau \max _{\lambda}\left\{\left\langle\lambda_{J+1}, b-Q x_{1}\right\rangle: c-\mathcal{A}^{*} \lambda=0, \lambda \in \mathcal{Q}\right\} \\
& +(1-\tau) d^{T} x_{2}+(1-\tau) \max _{\lambda}\left\{\left\langle\lambda_{J+1}, b-Q x_{2}\right\rangle: c-\mathcal{A}^{*} \lambda=0, \lambda \in \mathcal{Q}\right\} \\
= & \tau f_{0}\left(x_{1}, \xi\right)+(1-\tau) f_{0}\left(x_{2}, \xi\right),
\end{aligned}
$$

which implies $f_{0}(\cdot, \xi)$ is a convex and thus is locally Lipschitz continuous for $x \in X$.

From Corollary 3.2, we know $f_{0}(x, \cdot)$ is locally Lipschitz continuous over $\Xi$. Therefore $f(\cdot, \cdot)$ is continuous on $X \times \Xi$. It follows that $f_{0}(x, \xi)$ is a convex random lower semi-continuious function. The proof is completed by Theorem 2.1 [15].

Both functions $\psi_{P}$ and $\Psi_{P}$ depend on $P$ and they are lower semicontinuous on $\mathcal{R}_{+}, \psi_{P}$ is nondecreasing and $\Psi_{P}$ is increasing. The previous results derive when $p=3$, we can rewrite the $p$ th order Fortet-Mourier metric $\zeta_{p}$ defined on $\mathcal{P}_{p}(\Xi)$ by

$$
\zeta_{3}(P, Q):=\sup _{f \in \mathcal{F}_{3}(\Xi)}\left|\int_{\Xi} f(\xi)(P-Q)(d \xi)\right|
$$

for $P, Q \in \mathcal{P}_{3}(\Xi):=\left\{Q \in \mathcal{P}(\Xi): \int_{\Xi}\|\xi\|^{3} Q(d \xi)<\infty\right\}$, where

$$
\begin{equation*}
\mathcal{F}_{3}(\Xi):=\left\{f: \Xi \mapsto \mathbb{R}:\left|f\left(\xi_{1}\right)-f\left(\xi_{2}\right)\right| \leq \max \left\{1,\left\|\xi_{1}\right\|,\left\|\xi_{2}\right\|\right\}^{2}\left\|\xi_{1}-\xi_{2}\right\|, \forall \xi_{1}, \xi_{2} \in \Xi\right\} \tag{3.16}
\end{equation*}
$$

Theorem 3.3. Assume the Assumptions 3.1-3.5 hold and $\Xi$ is a convex set. Let $S(P)$ be nonempty and bounded for $P \in \mathcal{P}_{3}(\Xi)$, then there exist constants $L>0$ and $\delta>0$ such that

$$
\begin{align*}
& |v(P)-v(Q)| \leq L \zeta_{3}(P, Q)  \tag{3.17}\\
& \emptyset \neq S(Q) \subset S(P)+\Psi_{P}\left(L \zeta_{3}(P, Q)\right) \mathbb{B} \tag{3.18}
\end{align*}
$$

whenever $Q \in \mathcal{P}_{3}(\Xi)$ and $\zeta_{3}(P, Q)<\delta$.
Proof. Let $\Xi:=\Xi_{\leq}+\Xi_{>}$, where $\Xi_{\leq}:=\{\xi \in \Xi:\|\xi\| \leq R\}$ and $\Xi_{>}:=\{\xi \in \Xi:\|\xi\|>R\}$ and set $R \geq \max \left\{1, \zeta_{3}^{-1}(P, Q)\right\}$.
(i)Firstly we want to show that there exists a constant $L_{1}>0$ such that

$$
\frac{1}{L_{1}} f_{0}(x, \xi) \mathcal{X}_{\Xi_{\leq}}(\xi) \subseteq \mathcal{F}_{3}(\Xi),
$$

where $\mathcal{X}_{\Xi_{\leq}}$is the indicator function denoted over $\Xi_{\leq}$. From the Heine-Borel Theorem, there exist a finite number of points $\xi_{1}, \xi_{2}, \cdots, \xi_{m} \in \Xi_{\leq}$and positive constant $\delta\left(\xi_{i}\right), i=1, \cdots, m$ such that

$$
\Xi_{\leq} \subseteq \bigcup_{i=1}^{m} \Xi_{i}
$$

where $\Xi_{i}:=\mathbb{B}_{\delta\left(\xi_{i}\right)}\left(\xi_{i}\right)$ denotes an open ball for $i=1, \cdots, m$. From Corollary 3.2, for each $\xi_{i}$, there exists $\bar{L}\left(\xi_{i}\right)$ such that for any $\tilde{\xi}_{i}, \hat{\xi}_{i} \in \Xi_{i}$,

$$
f_{0}\left(x, \tilde{\xi}_{i}\right)-f_{0}\left(x, \hat{\xi}_{i}\right) \leq \bar{L}\left(\xi_{i}\right) \max \left\{1,\left\|\tilde{\xi}_{i}\right\|,\left\|\hat{\xi}_{i}\right\|\right\}^{2}\left\|\tilde{\xi}_{i}-\hat{\xi}_{i}\right\|
$$

We consider the following cases.
Case (a). Assume that $\xi_{i} \in \Xi_{i}, \xi_{k} \in \Xi_{k}$ for some $i, k \in\{1, \cdots, m\}$. Since $\Xi$ is convex, we consider the lines segment $\left[\xi_{i}, \xi_{k}\right]=\left\{\xi(\eta)=(1-\eta) \xi_{i}+\eta \xi_{k}: \eta \in[0,1]\right\} \subseteq \Xi_{\leq}$, there exist indices
$i_{j}, j=1, \cdots, l$ such that $i_{1}=i, i_{l}=k,\left[\xi_{i}, \xi_{k}\right] \cap \Xi_{i_{j}} \neq \emptyset$ for each $j=1, \cdots, l$ and $\left[\xi_{i}, \xi_{k}\right] \subseteq \bigcup_{j=1}^{l} \Xi_{i_{j}}$. Then there exist increasing numbers $\eta_{i_{j}} \in[0,1]$ for each $j=1, \cdots, l$ such that $\xi\left(\eta_{i_{1}}\right)=\xi(0)=\xi_{i}$, $\xi\left(\eta_{i_{l}}\right)=\xi(1)=\xi_{k}, \xi\left(\eta_{i_{j}}\right) \in \Xi_{i_{j-1}} \cap \Xi_{i_{j}}, j=2, \cdots, l$ and $\xi\left(\eta_{i_{k}}\right) \notin \Xi_{i_{j}}$ if $1 \leq i_{k}<i_{j}$. Then we can ensure that there are only two spots $\xi\left(\eta_{i_{j}}\right)$ and $\xi\left(\eta_{i_{j+1}}\right)$ in $\Xi_{i_{j}}$. For fixed $x$ and $i \neq k$, $\xi_{i} \in \Xi_{i}, \xi_{k} \in \Xi_{k}$,

$$
\begin{aligned}
& \left|f_{0}\left(x, \xi_{i}\right) \mathcal{X}_{\Xi_{\leq}}\left(\xi_{i}\right)-f_{0}\left(x, \xi_{k}\right) \mathcal{X}_{\Xi_{\leq}}\left(\xi_{k}\right)\right| \\
= & \left|f_{0}\left(x, \xi_{i}\right)-f_{0}\left(x, \xi_{k}\right)\right| \\
\leq & \sum_{j=1}^{l-1}\left|f_{0}\left(x, \xi\left(\eta_{i_{j}}\right)\right)-f_{0}\left(x, \xi\left(\eta_{i_{j+1}}\right)\right)\right| \\
\leq & \sum_{j=1}^{l-1} \bar{L}\left(\xi_{i_{j}}\right) \max \left\{1,\left\|\xi\left(\eta_{i_{j}}\right)\right\|,\left\|\xi\left(\eta_{i_{j+1}}\right)\right\|\right\}^{2}\left\|\xi\left(\eta_{i_{j}}\right)-\xi\left(\eta_{i_{j+1}}\right)\right\| \\
\leq & \max _{j=1, \cdots, l-1} \bar{L}\left(\xi_{i_{j}}\right) \max \left\{1,\left\|\xi_{i}\right\|,\left\|\xi_{k}\right\|\right\}^{2} \sum_{j=1}^{l-1}\left\|\xi\left(\eta_{i_{j}}\right)-\xi\left(\eta_{i_{j+1}}\right)\right\| \\
\leq & \bar{L}_{i} \max \left\{1,\left\|\xi_{i}\right\|,\left\|\xi_{k}\right\|\right\}^{2} \sum_{j=1}^{l-1}\left\|\left(\eta_{i_{j}}-\eta_{i_{j+1}}\right)\left(\xi_{i}-\xi_{k}\right)\right\| \\
= & \bar{L}_{i} \max \left\{1,\left\|\xi_{1}\right\|,\left\|\xi_{k}\right\|\right\}^{2}\left\|\xi\left(\eta_{i_{1}}\right)-\xi\left(\eta_{i_{l}}\right)\right\| \\
= & \bar{L}_{i} \max \left\{1,\left\|\xi_{i}\right\|,\left\|\xi_{k}\right\|\right\}^{2}\left\|\xi_{i}-\xi_{k}\right\|,
\end{aligned}
$$

where $\bar{L}_{i}=\max _{j=1, \cdots, l-1} \bar{L}\left(\xi_{i_{j}}\right)$.
Case (b). For $\tilde{\xi}_{i}, \hat{\xi}_{i} \in \Xi_{i}, i=1, \cdots, m$,

$$
\begin{align*}
& \left|f_{0}\left(x, \tilde{\xi}_{i}\right) \mathcal{X}_{\Xi_{\leq}}\left(\tilde{\xi}_{i}\right)-f_{0}\left(x, \hat{\xi}_{i}\right) \mathcal{X}_{\Xi_{\leq}}\left(\hat{\xi}_{i}\right)\right| \\
= & \left|f_{0}\left(x, \tilde{\xi}_{i}\right)-f_{0}\left(x, \hat{\xi}_{i}\right)\right| \\
\leq & \tilde{L}\left(\xi_{i}\right) \max \left\{1,\left\|\tilde{\xi}_{i}\right\|,\left\|\hat{\xi}_{i}\right\|\right\}^{2}\left\|\tilde{\xi}_{i}-\hat{\xi}_{i}\right\| . \tag{3.19}
\end{align*}
$$

From Case (a) and Case (b), let $\bar{L}:=\max _{i=1, \cdots, m}\left\{\bar{L}_{i}, \tilde{L}\left(\xi_{i}\right)\right\}$, then for any $\tilde{\xi}, \hat{\xi} \in \Xi_{\leq}$, we have that

$$
\begin{equation*}
\left|f_{0}(x, \tilde{\xi}) \mathcal{X}_{\Xi_{\leq}}(\tilde{\xi})-f_{0}(x, \hat{\xi}) \mathcal{X}_{\Xi_{\leq}}(\hat{\xi})\right| \leq \bar{L} \max \{1,\|\tilde{\xi}\|,\|\hat{\xi}\|\}^{2}\|\tilde{\xi}-\hat{\xi}\| . \tag{3.20}
\end{equation*}
$$

Case (c). We now consider that $\hat{\xi} \in \Xi_{\leq}$and the $\tilde{\xi} \in \Xi_{>}$and the ligature $[\hat{\xi}, \tilde{\xi}]$ and the boundary of $\Xi_{\leq}$intersect at $\xi^{*}$. From the equation (3.14), $\left\|f_{0}(x, \xi)\right\| \leq \tilde{L} \max \{1,\|\xi\|\}$, we can
obtain that

$$
\begin{aligned}
& \left\|f_{0}(x, \hat{\xi}) \mathcal{X}_{\Xi_{\leq}}(\hat{\xi})-f_{0}(x, \tilde{\xi}) \mathcal{X}_{\Xi_{\leq}}(\tilde{\xi})\right\| \\
\leq & \left\|f_{0}(x, \hat{\xi})-f_{0}\left(x, \xi^{*}\right)\right\|+\left\|f_{0}\left(x, \xi^{*}\right)-0\right\| \\
\leq & \bar{L} \max \left\{1,\|\hat{\xi}\|,\left\|\xi^{*}\right\|\right\}^{2}\left\|\hat{\xi}-\xi^{*}\right\|+\tilde{L} \max \left\{1,\left\|\xi^{*}\right\|\right\} \\
\leq & \max \{\bar{L}, \tilde{L}\} \max \{1,\|\hat{\xi}\|,\|\tilde{\xi}\|\}^{2}\left\{\left\|\hat{\xi}-\xi^{*}\right\|+1\right\}
\end{aligned}
$$

Assume that $\left\|\xi^{*}-\tilde{\xi}\right\|>1$, then $\left\|\hat{\xi}-\xi^{*}\right\|+1 \leq\left\|\hat{\xi}-\xi^{*}\right\|+\left\|\xi^{*}-\tilde{\xi}\right\|=\|\hat{\xi}-\tilde{\xi}\|$ and thus

$$
\begin{equation*}
\left\|f_{0}(x, \hat{\xi}) \mathcal{X}_{\Xi_{\leq}}(\hat{\xi})-f_{0}(x, \tilde{\xi}) \mathcal{X}_{\Xi_{\leq}}(\tilde{\xi})\right\| \leq \max \left\{\bar{L}_{i}, \tilde{L}\right\} \max \{1,\|\hat{\xi}\|,\|\tilde{\xi}\|\}^{2}\|\hat{\xi}-\tilde{\xi}\| \tag{3.21}
\end{equation*}
$$

Otherwise if $d\left(\tilde{\xi}, \Xi_{\leq}\right) \leq\left\|\xi^{*}-\tilde{\xi}\right\| \leq 1$, then we denote $\Xi_{\leq}^{\prime}:=\left\{\xi \in \Xi:\|\xi\| \leq R^{\prime}\right\}, R^{\prime}:=R+1$ and $\hat{\xi}, \tilde{\xi} \in \Xi_{\leq}^{\prime}$. From the Heine-Borel Theorem, similarly with (3.20), there exists $\hat{L}$ such that

$$
\begin{equation*}
\left\|f_{0}(x, \hat{\xi}) \mathcal{X}_{\Xi_{\leq}}(\hat{\xi})-f_{0}(x, \tilde{\xi}) \mathcal{X}_{\Xi_{\leq}}(\tilde{\xi})\right\| \leq \hat{L} \max \{1,\|\hat{\xi}\|,\|\tilde{\xi}\|\}^{2}\|\hat{\xi}-\tilde{\xi}\| \tag{3.22}
\end{equation*}
$$

Combing with (3.20)-(3.22), we denote $L_{1}:=\max \{\bar{L}, \tilde{L}, \hat{L}\}$, for any $\hat{\xi}, \tilde{\xi} \in \Xi$,

$$
\left\|f_{0}(x, \hat{\xi}) \mathcal{X}_{\Xi_{\leq}}(\hat{\xi})-f_{0}(x, \tilde{\xi}) \mathcal{X}_{\Xi_{\leq}}(\tilde{\xi})\right\| \leq L_{1} \max \{1,\|\hat{\xi}\|,\|\tilde{\xi}\|\}^{2}\|\hat{\xi}-\tilde{\xi}\|
$$

and thus

$$
\frac{1}{L_{1}} f_{0}(x, \xi) \mathcal{X}_{\Xi_{\leq}}(\xi) \subseteq \mathcal{F}_{3}(\Xi)
$$

Therefore, we can obtain the following estimation on $\Xi$

$$
\begin{align*}
& \int_{\Xi_{\leq}} f_{0}(x, \xi)(P-Q)(d \xi) \\
= & L_{1} \int_{\Xi} \frac{1}{L_{1}} f_{0}(x, \xi) \mathcal{X}_{\Xi_{\leq}}(\xi)(P-Q)(d \xi) \\
\leq & L_{1} \zeta_{3}(P, Q) \tag{3.23}
\end{align*}
$$

(ii) Denote the upper bound of $\left\{\int_{\Xi}\|\xi\|^{3} Q(d \xi): Q \in \mathcal{P}_{3}(\Xi)\right\}$ by $\Pi>0$. From the equation (3.14), $\left\|f_{0}(x, \xi)\right\| \leq \tilde{L} \max \{1,\|\xi\|\}$, we can obtain that

$$
\begin{aligned}
& \int_{\Xi_{>}} f_{0}(x, \xi)(P-Q)(d \xi) \\
\leq & \tilde{L} \int_{\left\{\xi \in \Xi_{>}:\|\xi\|>R\right\}}\|\xi\|(P+Q)(d \xi) \\
\leq & \tilde{L} \int_{\left\{\xi \in \Xi_{>}:\|\xi\|>R\right\}} \frac{\|\xi\|^{2}}{R}(P+Q)(d \xi) \\
\leq & \frac{\tilde{L}}{R}\left(\int_{\Xi_{>}}\|\xi\|^{2} d P(\xi)+\int_{\Xi_{>}}\|\xi\|^{2} d Q(\xi)\right) \\
\leq & \frac{\tilde{L}}{R}\left(\int_{\Xi_{>}}\|\xi\|^{3} d P(\xi)+\int_{\Xi_{>}}\|\xi\|^{3} d Q(\xi)\right) \\
\leq & \frac{2 \tilde{L} \Pi}{R} \leq \min \left\{1, \zeta_{3}(P, Q)\right\} 2 \tilde{L} \Pi .
\end{aligned}
$$

Combine the above discussions, we have that

$$
\begin{aligned}
& \int_{\Xi} f_{0}(x, \xi)(P-Q)(d \xi) \\
= & \int_{\Xi_{\leq}} f_{0}(x, \xi)(P-Q)(d \xi)+\int_{\Xi_{>}} f_{0}(x, \xi)(P-Q)(d \xi) \\
\leq & L_{1} \zeta_{3}(P, Q)+2 \tilde{L} \Pi \zeta_{3}(P, Q) .
\end{aligned}
$$

Thus

$$
\sup _{x \in X}\left|\int_{\Xi} f_{0}(x, \xi)(P-Q)(d \xi)\right| \leq L \zeta_{3}(P, Q),
$$

where $L:=L_{1}+2 \tilde{L} \Pi$. Then we have $d_{\mathcal{F}, \rho}(P, Q) \leq L \zeta_{3}(P, Q)$.
From the second inequality of Proposition 3.1 and the increasing property of $\Psi_{P}$, there exists $\tilde{x} \in S(Q)$, we can obtain that

$$
d(\tilde{x}, S(P)) \leq \Psi_{P}\left(d_{\mathcal{F}, \rho}(P, Q)\right) \leq \Psi_{P}\left(L \zeta_{3}(P, Q)\right)
$$

Then $\emptyset \neq S(Q) \subset S(P)+\Psi_{P}\left(L \zeta_{3}(P, Q)\right) \mathbb{B}$. We complete the proof.

## 4 Empirical approximations of two-stage stochastic programs

In this section, we consider problem (3.8) when the probability distribution $P \in \mathcal{P}(\Xi)$ is estimated by empirical measures and investigate the asymptotic behavior of the approximate problems.

Let $P \in \mathcal{P}(\Xi)$ and $\xi_{1}, \xi_{2}, \ldots, \xi_{n}, \ldots$ be independent identically distributer variables on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ having the joint distribution $P$, i.e., $P=\mathbb{P} \xi_{1}^{-1}$ with support set $\Xi$. We consider the empirical measures

$$
P_{n}(\omega):=\frac{1}{n} \sum_{i=1}^{n} \delta_{\xi_{i}(\omega)}, \quad(\omega \in \Omega ; n \in \mathbb{N})
$$

where $\delta_{\xi}$ denotes the unit mass at $\xi \in \Xi$, and the empirical approximations of the stochastic program (1.2) with sample size $n$, i.e.,

$$
\begin{equation*}
\min \left\{\frac{1}{n} \sum_{i=1}^{n} f_{0}\left(x, \xi_{i}(\cdot)\right): x \in X\right\} . \tag{4.1}
\end{equation*}
$$

Since the objective function of (4.1) is a random lsc function from $\mathcal{R}^{n} \times \Omega$ to $\overline{\mathcal{R}}$, the optimal value $v\left(P_{n}(\cdot)\right)$ of (4.1) is measurable from $\Omega$ to $\mathcal{R}$ and the optimal solution mapping $S\left(P_{n}(\cdot)\right)$ is a closed-valued measurable set-valued mapping from $\Omega$ to $\mathcal{R}^{n}$ from [10].

Another measurability question arises when studying uniform convergence properties of the empirical process

$$
\left\{n^{\frac{1}{2}}\left(P_{n}(\cdot)-P\right) f=n^{-\frac{1}{2}} \sum_{i=1}^{n}\left(f\left(\xi_{i}(\cdot)\right)-P f\right)\right\}_{f \in \mathcal{F}}
$$

indexed by some class $\mathcal{F}=\left\{f_{0}(x, \cdot): x \in X\right\}$. Here we set $P f:=\int_{\Xi} f_{0}(x, \xi) P(d \xi)$ for $P \in \mathcal{P}(\Xi)$ and $f \in \mathcal{F}$. Uniform convergence properties refer to the convergence or to the convergence rate of

$$
\begin{equation*}
d_{\mathcal{F}}\left(P_{n}(\cdot), P\right)=\sup _{f \in \mathcal{F}}\left|P_{n}(\cdot) f-P f\right| \tag{4.2}
\end{equation*}
$$

to 0 in terms of some stochastic convergence. Since the supremum in (4.2) is nonmeasurable in general, we introduce a common condition on $\mathcal{F}$ that can be satisfied in most stochastic programming models.

Definition 4.1. [7] A class of measurable functions from $\Xi$ to $\mathcal{R}$ is called P-permissible if for $P \in \mathcal{P}(\Xi)$, there exists a countable subset $\mathcal{F}_{0}$ of $\mathcal{F}$ such that for each function $f \in \mathcal{F}$, there exists a sequence $f_{n}$ in $\mathcal{F}_{0}$ of $\mathcal{F}$ converging pointwise to $f$ such that the sequence $\left\{P f_{n}\right\}$ also converges to $P f$.

If $\mathcal{F}$ is $P$-permissible, then

$$
d_{\mathcal{F}}\left(P_{n}(\omega), P\right)=d_{\mathcal{F}_{0}}\left(P_{n}(\omega), P\right)
$$

for $\omega \in \Omega$, when $n \rightarrow \infty$. Then the convergence analysis reduces to a countable class.
Definition 4.2. [15] A P-permissible class is called a P-Glivenko-Cantelli class if the sequence $d_{\mathcal{F}}\left(P_{n}(\cdot), P\right)$ of random variables converges to $0 \mathbb{P}$-almost surely.

In order to judge whether a given class $\mathcal{F}$ is a P-Glivenko-Cantelli class, we introduce the following concepts.

Definition 4.3. Let $\mathcal{F}$ be a subset of the normed space $L_{r}(\Xi, P)$, i.e., for some $r \geq 1$, equipped with the usual norm $\|F\|_{P, r}=\left(P|F|^{r}\right)^{\frac{1}{r}}$. The covering number $N(\varepsilon, \mathcal{F},\|\cdot\|)$ is the minimal number of balls $\left\{g \in L_{r}(\Xi, P):\|g-f\|_{P, r}<\varepsilon\right\}$ of radius $\varepsilon$ needed to cover the set $\mathcal{F}$.

The centers of the balls need not belong to $\mathcal{F}$, but they should have finite norms. The entropy (with bracketing) is the logarithm of the coving number.

Given two functions $f_{1}$ and $f_{2}$ from $L_{r}(\Xi, P)$, the bracket $\left[f_{1}, f_{2}\right.$ ] is the set of all functions $f \in L_{r}(\Xi, P)$ with $f_{1}(\xi) \leq f(\xi) \leq f_{2}(\xi)$. An $\varepsilon$-bracket is a bracket $\left[f_{1}, f_{2}\right]$ with $\left\|f_{1}-f_{2}\right\| \leq \varepsilon$.

The bracketing number $N_{[]}(\varepsilon, \mathcal{F},\|\cdot\|)$ is the minimum number of $\varepsilon$-brackets needed to cover $\mathcal{F}$. The entropy with bracketing is the logarithm of the bracketing number. In the definition of the bracketing number, the upper and lower bounds $f_{1}$ and $f_{2}$ of the brackets need not belong to $\mathcal{F}$ themselves but are assumed to have finite norms.

Definition 4.4. If $F$ is called uniformly bounded in probability with tail $C_{\mathcal{F}}(\cdot)$ if the function $C_{\mathcal{F}}(\cdot)$ is defined on $(0,+\infty)$ and decreasing to 0 with the estimate

$$
\begin{equation*}
\mathbb{P}\left\{\omega: n^{\frac{1}{2}} d_{\mathcal{F}}\left(P_{n}(\omega), P\right) \geq \varepsilon\right\} \leq C_{\mathcal{F}}(\varepsilon) \tag{4.3}
\end{equation*}
$$

holds for each $\varepsilon \geq 0$ and $n \in \mathbb{N}$.
To state the following results, we denote the set of all real-valued random variables on $(\Omega, \mathcal{A}, \mathbb{P})$ by $\mathcal{K}(\Re)$. Then the Ky Fan distance [4, Theorem 9.21$] \kappa$ of two real random variable $\mathcal{X}, \mathcal{Y} \in \mathcal{K}(R)$ is denoted by

$$
\begin{equation*}
\kappa(\mathcal{X}, \mathcal{Y}):=\inf \{\eta>0: P(|\mathcal{X}-\mathcal{Y}|>\eta) \leq \eta\}, \tag{4.4}
\end{equation*}
$$

where the equality is understood in sense of $\mathbb{P}$-almost surely.
Lemma 4.1. [20, Theorem 2.7.11] Suppose that

$$
\left|f_{s}(x)-f_{t}(x)\right| \leq d(s, t) F(x)
$$

for some metric $d$ on the index set, function $F$ on the sample space, and for every $x$. Let $\mathcal{F}=\left\{f_{t}: t \in T\right\}$ be a class of functions satisfying the preceding display for every $s$ and $t$ and some fixed function $F$. Then for any form $\|\cdot\|$,

$$
N_{[]}(2 \varepsilon\|F\|, \mathcal{F},\|\cdot\|) \leq N(\varepsilon, T, d) .
$$

Lemma 4.2. [18, 20]Let $\mathcal{F}$ be P-permissible with envelope $F_{\mathcal{F}}$. If $P F_{\mathcal{F}}<\infty$ and

$$
\begin{equation*}
\sup _{P} N_{[]}\left(\varepsilon\left\|F_{\mathcal{F}}\right\|_{P, 1}, \mathcal{F}, L_{1}(\Xi, P)\right)<\infty, \tag{4.5}
\end{equation*}
$$

then $\mathcal{F}$ is a P-Glivenko-Cantelli class. If $\mathcal{F}$ is uniformly bounded and there exist constants $r \geq 1$ and $R \geq 1$ such that

$$
\begin{equation*}
\sup _{P} N_{[\jmath]}\left(\varepsilon\left\|F_{\mathcal{F}}\right\|_{P, 2}, \mathcal{F}, L_{2}(\Xi, P)\right)<\left(\frac{R}{\varepsilon}\right)^{r} \tag{4.6}
\end{equation*}
$$

holds for all $\varepsilon>0$, then the empirical process indexed by $\mathcal{F}$ is uniformly bounded in probability with exponential tail $C_{\mathcal{F}}(\varepsilon)=\left(K(R) \varepsilon r^{-\frac{1}{2}}\right)^{r} \exp \left(-2 \varepsilon^{2}\right)$ with some constant $K(R)$ depending on $R$.

Theorem 4.1. Let the assumptions in Theorem 3.3 hold, then for sufficiently large $n \in N$,

$$
\begin{aligned}
& \kappa\left(v\left(P_{n}(\cdot)\right), v(P)\right) \leq L \kappa\left(\zeta_{3}\left(P_{n}(\cdot), P\right), 0\right), \\
& \kappa\left(\sup _{x \in S\left(P_{n}(\cdot)\right)} d(x, S(P)), 0\right) \leq \Psi_{P}\left(\kappa\left(\zeta_{3}\left(P_{n}(\cdot), P\right), 0\right)\right),
\end{aligned}
$$

where $L>0$ is the constant in Theorem 3.3 and $\Psi_{P}$ is the associated function (3.15).
Proof Since $X$ is a compact set, there exists a countable and dense subset $X_{0}$. Let $\mathcal{F}_{0}=$ $\left\{f_{0}(x, \xi), x \in X_{0}\right\}$. Then $\mathcal{F}_{0}$ is a countable subset of $\mathcal{F}$, which implies that $\mathcal{F}$ is permissible with respect to $P$. Then we prove $\mathcal{F}$ is a $P$-Glivenko-Cantelli class. From Corollary 3.2,

$$
f_{0}\left(x_{1}, \xi\right)-f_{0}\left(x_{2}, \xi\right) \leq \hat{L} \max \{1,\|\xi\|\}\left\|x_{1}-x_{2}\right\|,
$$

for all $x_{1}, x_{2} \in \gamma \mathbb{B}$, and $\xi \in \Xi$.
Setting $f_{s}(x):=f_{0}\left(x_{1}, \cdot\right), f_{t}(x):=f_{0}\left(x_{2}, \cdot\right), F(\xi):=\hat{L} \max \{1,\|\xi\|\}, x_{1}, x_{2} \in \rho \mathbb{B}, d(s, t)=$ $\left\|x_{1}-x_{2}\right\|$. By applying Lemma 4.1 we can obtain that

$$
\begin{equation*}
N_{[]}\left(2 \varepsilon\|F\|_{P, 1}, \mathcal{F}, L_{1}(\Xi, P)\right) \leq N\left(\varepsilon, X, \mathcal{R}^{n}\right) \leq K \varepsilon^{-n} \tag{4.7}
\end{equation*}
$$

for $0<\varepsilon<1$ and some constant $K>0$ depending only on $n$ and the diameter of $X$. Since $\|F\|_{P, 1}$ is finite, we may replace $\varepsilon$ by $\varepsilon / 2\|F\|_{P, 1}$ in (4.7) and obtain that $N_{[]}\left(\varepsilon, \mathcal{F}, L_{1}(\Xi, P)\right)$ is finite for all $\varepsilon>0 . \mathcal{F}$ is a Glivenko-Cantelli class because of Lemma 4.2.

Set $\varepsilon_{n}:=\kappa\left(\zeta_{3}\left(P_{n}(\cdot), P\right), 0\right)$, from Theorem $3.3\left|v(P)-v\left(P_{n}(\cdot)\right)\right| \leq L \zeta_{3}\left(P, P_{n}(\cdot)\right)$, thus

$$
\mathbb{P}\left(\left|v(P), v\left(P_{n}(\cdot)\right)\right|>L \varepsilon_{n}\right) \leq \mathbb{P}\left(\zeta_{3}\left(P_{n}(\cdot), P\right)>\varepsilon_{n}\right) \leq \varepsilon_{n},
$$

for sufficiently large $n \in \mathbb{N}$. Since $\mathcal{F}$ is a Glivenko-Cantelli class, thus the sequence $\varepsilon_{n}$ tends to 0 . From the definition of $\kappa$, we obtain that

$$
\kappa\left(v\left(P_{n}(\cdot)\right), v(P)\right) \leq \max \left\{\varepsilon_{n}, L \varepsilon_{n}\right\}=\max \{1, L\} \kappa\left(\zeta_{3}\left(P_{n}(\cdot), P\right), 0\right) .
$$

Then we conclude from Theorem 3.3 and the nature of increasing about $\Psi_{P}(\eta):=\eta+\psi_{P}^{-1}(2 \eta)$ that

$$
\begin{aligned}
\mathbb{P}\left(\sup _{x \in S\left(P_{n}(\cdot)\right)} d(x, S(P))>\Psi_{P}\left(\varepsilon_{n}\right)\right) & \leq \mathbb{P}\left(\Psi_{P}\left(L \zeta_{3}\left(P_{n}(\cdot), P\right)>\Psi_{P}\left(\varepsilon_{n}\right)\right)\right. \\
& =\mathbb{P}\left(\zeta_{3}\left(P_{n}(\cdot), P\right)>\frac{\varepsilon_{n}}{L}\right) \leq \frac{\varepsilon_{n}}{L} \leq \frac{1}{L} \Psi_{P}\left(\varepsilon_{n}\right)
\end{aligned}
$$

for sufficiently large $n \in \mathbb{N}$. Then we obtain that

$$
\kappa\left(\sup _{x \in S\left(P_{n}(\cdot)\right)} d(x, S(P)), 0\right) \leq \max \left\{1, \frac{1}{L}\right\} \Psi_{P}\left(\varepsilon_{n}\right) .
$$

Note that $L=L_{1}+2 \tilde{L} \Pi>\max \{\bar{L}, \tilde{L}, \hat{L}\}>1$, where $\bar{L}, \tilde{L}, \hat{L}$ are constants denoted in Theorem 3.3 and then we complete the proof.

Theorem 4.2. Assume the conditions of Theorem 3.2 are satisfied and $\Xi$ is bounded, then

$$
\begin{aligned}
& \kappa\left(v(P), v\left(P_{n}\right)\right)=O\left((\log n)^{1 / 2} n^{-1 / 2}\right), \\
& \kappa\left(\sup _{x \in S\left(P_{n}(\cdot)\right)} d(x, S(P)), 0\right)=O\left(\Psi_{P}(\log n)^{1 / 2} n^{-1 / 2}\right) .
\end{aligned}
$$

Proof Since $\Xi$ is bounded, the class $\mathcal{F}$ is uniformly bounded. When $p=3$, due to Lemma 4.1 we have that $N_{[]}\left(\varepsilon, \mathcal{F}, L_{2}(\Xi, P)\right) \leq C \varepsilon^{-n}$ and directly obtain the result by adopting the similar method used in Proposition 4.2 in [9].

Actually, we could show that whatever p is, the results in Theorem 4.1 and Theorem 4.2 hold. For $p>q \geq 1$ and $\varepsilon_{p}=\kappa\left(\zeta_{p}\left(P_{n}(\cdot), P\right), 0\right), \varepsilon_{q}=\kappa\left(\zeta_{q}\left(P_{n}(\cdot), P\right), 0\right)$, since $\mathcal{F}_{p}(\Xi) \supseteq \mathcal{F}_{q}(\Xi)$, we can obtain that

$$
\zeta_{p}\left(P_{n}(\cdot), P\right) \geq \zeta_{q}\left(P_{n}(\cdot), P\right)
$$

To the contrary we assume that $\varepsilon_{q}>\varepsilon_{p}$, then

$$
\varepsilon_{q}>\varepsilon_{p} \geq \mathbb{P}\left(\zeta_{p}\left(P_{n}(\cdot), P\right)>\varepsilon_{p}\right) \geq \mathbb{P}\left(\zeta_{q}\left(P_{n}(\cdot), P\right)>\varepsilon_{p}\right) .
$$

We can obtain that $\varepsilon_{p} \geq \mathbb{P}\left(\zeta_{q}\left(P_{n}(\cdot), P\right)>\varepsilon_{p}\right)$, thus $\varepsilon_{q}$ is not the minimum value of $\kappa\left(\zeta_{q}\left(P_{n}(\cdot), P\right)\right)$ by $\varepsilon_{q}>\varepsilon_{p}$. By contradiction, then $\varepsilon_{p} \geq \varepsilon_{q}$. We also can obtain that

$$
\begin{aligned}
\zeta_{p}\left(P_{n}(\cdot), P\right) & =\sup _{f \in \mathcal{F}_{p}(\Xi)}\left|\int_{\Xi} f(\xi)\left(P_{n}-P\right) d \xi\right| \\
& \leq \sup _{f \in \mathcal{F}(\Xi)}\left|\int_{\Xi} f(\xi)\left(P_{n}-P\right) d \xi\right|=d_{\mathcal{F}}\left(P_{n}(\cdot), P\right) .
\end{aligned}
$$

Because $C_{\mathcal{F}}(\cdot)$ is decreasing to 0 and equation (4.3), we have that

$$
\begin{aligned}
& \mathbb{P}\left(\zeta_{p}\left(P_{n}(\cdot), P\right) \geq \varepsilon_{p}\right) \leq \mathbb{P}\left(d_{\mathcal{F}}\left(P_{n}(\cdot), P\right) \geq \varepsilon_{p}\right) \\
\leq & C_{\mathcal{F}}\left(\sqrt{n} \varepsilon_{p}\right) \leq C_{\mathcal{F}}\left(\sqrt{n} \varepsilon_{q}\right) \leq C_{\mathcal{F}}\left(\sqrt{n} \varepsilon_{1}\right)=\left(K(R) \varepsilon_{1} \sqrt{\frac{n}{r}}\right)^{r} \exp \left(-2 \sqrt{n} \varepsilon_{1}^{2}\right)
\end{aligned}
$$

holds for some constant $K(R)$ depending on $R, \varepsilon_{1}>0$ and nature number $n$. Replacing $\varepsilon_{1}$ by $(\log n)^{1 / 2} n^{-1 / 2}$ leads to the estimate

$$
\mathbb{P}\left(\zeta_{p}\left(P_{n}(\cdot), P\right)>(\log n)^{1 / 2} n^{-1 / 2}\right)=O\left((\log n)^{r / 2} n^{-2}\right)
$$

and hence, to $\kappa\left(\zeta_{p}\left(P_{n}(\cdot), P\right), 0\right)=O\left((\log n)^{1 / 2} n^{-1 / 2}\right)$. Then we complete the proof.

## 5 Conclusions

Based on the research of the stability of a second-order conic optimization problem where all parameters are perturbed by random vectors [3], we investigate the quantitative stability property of the problem with probability distribution being perturbed and approximated in this paper. We derive quantitative continuity properties of the optimal value function and solution mapping by using the locally Lipschitz continuity properties of the feasible set-valued mapping with respect to Hausdorff distance. We also discuss the stability property for optimal values and solution sets for empirical approximations of two stage stochastic programs with second-order conic constraints. In the future research, we intend to study two stage problem with general conic constraints or bilevel programming and carry out quantitative stability under different probability metrics.

## References

[1] J. F. Bonnans and A. Shapiro, Perturbation Analysis of Optimization Problems, Springer Series in Operations Research, New York, 2000.
[2] D. Dentcheva and W. Römisch, Differential stability of two-stage stochastic programs, SIAM J. Optim., 11(1)(2000), 87-112.
[3] Qingsong Duan, Liwei Zhang and Sainan Zhang, Hadamard directional differentiability of the optimal value of a linear second-order conic programming problem, submitted.
[4] R. M. Dudley, Real analysis and prbability. Wadsworth \& Brooks/Cole, Pacific Grove, 1989.
[5] R. Fortet and F. Mourier, Convergence de la répartition empirique vers la répartition théorique, Ann. Sci. Ecole Norm. Sup., 70(1953), 266-285.
[6] M. A. Goberna and M. A. López, Linear Semi-infinite Optimization, Wiley, Chichester, 1998.
[7] Youpan Han and Zhiping Chen, Quantitative stability of full random two-stage stochastic programs with recourse, Optim. Lett., 9(2015), 1075-1090.
[8] D. Klatte, On quantitative stability for non-isolated minima, Contral Cybernet., 23 (1994), 183-200.
[9] S.T. Rachev and W. Römisch, Quantiitative stability in stochastic programming: The method of probability metric, Math. Oper. Res., 27 (2002), 792-818.
[10] R.T. Rockafellar and R.J.-B. Wets, Variational Analysis, Springer, Berlin, 1998.
[11] W. Römisch and R. Schultz, Stability analysis for stochastic programs, Ann. Oper. Res., 30(1991), 241-266.
[12] W. Römisch and R. Schultz, Distribution sensitivity in stochastic programming, Math. Program., 50(1991), 197-226.
[13] W. Römisch and R. Schultz, Lipschitz stability for stochastic programs with complete recourse, SIAM J. Optim., 6 (1996), 531-547.
[14] W. Römisch, Stability of Stochastic Programming, in Stochastic Programming, Handbooks in Operations Research and Management Science, Elsevier, Amsterdam, 2003.
[15] W. Römisch and R. J.-B. Wets, Stability of $\varepsilon$-approximate solutions to convex stochastic programs, SIAM J. Optim., 18(3)(2007), 961-979.
[16] R. Schultz, Some aspects of stability in stochastic programming, Ann. Oper. Res., 100(2000), 55-84.
[17] A. Shapiro, Quantitative stability in stochastic programming, Math. Program., 67 (1994), 99-108.
[18] M. Talagrand, Sharper bounds for Gaussian and empirical process, Ann. Pro., 22(1994), 28-76.
[19] G. Stewart and J. Sun, Matrix Perturbation Theory, Academic Press, New York, 1990.
[20] A.W. Van Der Vaart and J.A. Wellner, Weak Convergence and Empirical Processes, Springer, New York, 1996.


[^0]:    *School of Mathematical Sciences, Dalian University of Technology, China. Email: duanqs@mail.dlut.edu.cn
    ${ }^{\dagger}$ School of Mathematics, Tianjin University, Tianjin, 300072, China. E-mail: xumengw@hotmail.com. The research of this author was supported by the National Natural Science Foundation of China under project No. 11601376
    ${ }^{\ddagger}$ School of Mathematical Sciences, Dalian University of Technology, China. Email: syguo@dlut.edu.cn
    ${ }^{\text {§ S School of Mathematical Sciences, Dalian University of Technology, China. E-mail: lwzhang@dlut.edu.cn. }}$ The research of this author was supported by the National Natural Science Foundation of China under projects No. 11571059 and No. 91330206.

